Pacific Journal of Mathematics

Volume 344 No. 1 September 2026

PACIFIC JOURNAL OF MATHEMATICS

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

msp.org/pjm

EDITORS

Don Blasius (Managing Editor)
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
blasius@math.ucla.edu

Matthias Aschenbrenner
Fakultät für Mathematik
Universität Wien
Vienna, Austria
matthias.aschenbrenner@univie.ac.at

Vyjayanthi Chari
Department of Mathematics
University of California
Riverside, CA 92521-0135
chari@math.ucr.edu

Atsushi Ichino
Department of Mathematics
Kyoto University
Kyoto 606-8502, Japan
atsushi.ichino@gmail.com

Kefeng Liu
School of Sciences
Chongqing University of Technology
Chongqing 400054, China
liu@math.ucla.edu

Sucharit Sarkar
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
sucharit@math.ucla.edu

Dimitri Shlyakhtenko
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
shlyakht@ipam.ucla.edu

Ruixiang Zhang
Department of Mathematics
University of California
Berkeley, CA 94720-3840
ruixiang@berkeley.edu

PRODUCTION

Silvio Levy, Scientific Editor, production@msp.org

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

The subscription price for 2026 is US $710/year for the electronic version, and $965/year for print and electronic.
Subscriptions, requests for back issues and changes of subscriber address should be sent to Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004, 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 Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004, is published twelve times a year. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004.

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

© 2026 Mathematical Sciences Publishers

PACIFIC JOURNAL OF MATHEMATICS
Vol. 344, No. 1, 2026
https://doi.org/10.2140/pjm.2026.344.1

MINIMAL RATIONAL GRAPHS ADMITTING A ℚHD SMOOTHING

MÁRTON BEKE

Using the picture deformation technique of de Jong and van Straten we show that no singularity whose resolution graph has 3 or 4 large nodes, i.e., nodes satisfying $d(v)+e(v)\leq -2$, has a ℚHD smoothing. This is achieved by providing a general reduction algorithm for graphs with ℚHD smoothings, and enumeration. New examples and families are presented, which admit a combinatorial ℚHD smoothing, i.e., the incidence relations for a sandwich presentation can be satisfied. We also give a new proof of the Bhupal–Stipsicz theorem on the classification of weighted homogeneous singularities admitting ℚHD smoothings with this method by using cusp singularities.

1. Introduction

The study of complex surface singularities is a central area of interest in algebraic geometry. A key problem is understanding the various ways a singularity can be deformed into a smooth surface through a process called smoothing. This paper investigates a specific type of deformation known as a rational homology disk (ℚHD) smoothing. These are the topologically simplest kinds of deformations, giving us hope of a complete classification. The conjecture is attributed to Wahl, who had an unpublished list of singularities admitting a ℚHD smoothing, stating that this list is complete. This paper presents further evidence for Wahl's conjecture by showing the following:

Theorem 1.1. *No complex surface singularity whose resolution graph has at most three large nodes of any degree, or at most four nodes of degree* 3*, and no further nodes can admit a* ℚHD *smoothing.*

We also give a new proof of the classification theorem for weighted homogeneous graphs.

Theorem 1.2 [Bhupal and Stipsicz 2011]. *The star-shaped graphs admitting a* ℚHD *smoothing are precisely the graphs depicted in Figure 2 in Section 2 and Figures 10–16 in Section 6.*

MSC2020: 14B07, 14J17, 32S25, 32S30, 51E30.

Keywords: rational homology disk, surface singularity, rational blowdown, Wahl's conjecture, smoothing, plane curve deformations.

© 2026 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

Besides the algebro-geometric perspective, these spaces are also important in low-dimensional topology motivated by the rational blowdown operation of [Fintushel and Stern 1998], which uses $\mathbb{Q}$HD fillings of linear graphs and a cut-and-paste operation to generate topologically smaller exotic 4-manifolds. In a way this problem also tries to classify all possible rational blowdowns of 4-manifolds.

Our result can also be interpreted as a generalization of an observation in [Bhupal and Stipsicz 2013, Theorem 2.5, Remark 2.6]: a star-shaped graph, which admits a $\mathbb{Q}$HD smoothing, cannot have a too negative node, which is the same as the large nodes discussed here ($e(v) + d(v) \leq -2$).

In order to achieve these results, we use the deformation theory of sandwiched singularities developed in [de Jong and van Straten 1998]. This reduces the question of finding deformations of surface singularities to deformations of curves with some intersections prescribed by the surface singularity. We will be mostly dealing with the combinatorial properties of these intersection patterns. Using Donaldson's diagonalizability theorem, strong necessary conditions were imposed on graphs admitting a $\mathbb{Q}$HD smoothing by Stipsicz, Szabó and Wahl [Stipsicz et al. 2008], in particular the study of graphs with $\mathbb{Q}$HD smoothings is essentially reduced to the families $\mathcal{G}$, $\mathcal{W}$, $\mathcal{N}$, $\mathcal{M}$, $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, a fact we will rely on in what follows. The proof of Theorem 1.1 is as follows: Under the assumption of a graph having large nodes and a $\mathbb{Q}$HD smoothing, we find a reduction algorithm (Section 3.2), so if a graph admits (the combinatorics of a) $\mathbb{Q}$HD smoothing, then some smaller graph does so too. This, and information about the general shape of our graphs, reduces the problem to a finite number of graphs (Lemma 3.23). We then show (Theorem 4.1) that any graph that admits a combinatorial solution to its $\mathbb{Q}$HD smoothing problem satisfies $Z_K^2 + n = 0$, where n is the number of vertices of the graph, and Z_K is the anticanonical cycle. Using this statement, we can search through all possible graphs to finish the proof of the theorem.

The proof of the Bhupal–Stipsicz theorem is a case analysis of possible solutions to the $\mathbb{Q}$HD smoothing problem of graphs in $\mathcal{A}, \mathcal{B}, \mathcal{C}$ (since the other cases have been settled already; see [Stipsicz et al. 2008, Section 8]). This argument is notably simpler and shorter than the original proof, which utilizes the combinatorics of symplectic curves and McDuff's theorem. Accordingly the result is weaker, since we only rule out the $\mathcal{A}, \mathcal{B}, \mathcal{C}$ graphs with one star from having a $\mathbb{Q}$HD smoothing, whereas Bhupal and Stipsicz [2011] ruled them out from having a symplectic filling. This gap is filled by new results of Plamenevskaya and Starkston. In [Plamenevskaya and Starkston 2023] they expand the de Jong–van Straten theory into the symplectic case, if the corresponding curve singularity has only smooth branches, and in [Plamenevskaya and Starkston 2025] to the general case. This means, that if the intersection pattern for a $\mathbb{Q}$HD smoothing is impossible then the graph cannot admit a symplectic $\mathbb{Q}$HD filling.

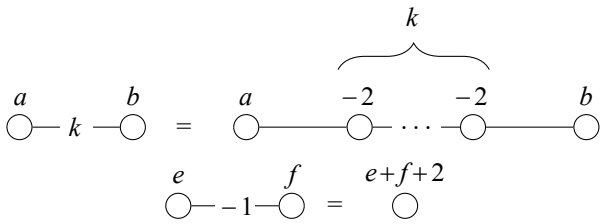


Figure 1. An illustration of the notation.

The paper is organized as follows: In Section 2, we review details on the $\mathcal{A}, \mathcal{B}, \mathcal{C}$ families, and the deformation theory for sandwiched singularities. Section 3 gives a specific introduction to minimal rational graphs, and describes and proves the reduction algorithm. Section 4 shows some necessary conditions for a graph to admit a combinatorial ℚHD smoothing. Section 5 is dedicated to examples of graphs that admit combinatorial ℚHD smoothings, and Section 6 consists of an introduction to the case when the curves can have cusp singularities, as well as the proof of the Bhupal–Stipsicz theorem.

1.1. ***Notation.*** A number k on the edge of a graph Γ means a path consisting of k vertices framed -2. In some cases, k is allowed to be -1; in this case, we identify the two endpoints of the edge with k and frame it $e+f+2$, where e, f are the framings of the two endpoints of the edge with k over it. See Figure 1.

2. Preliminaries

2.1. ***Graphs admitting a ℚHD smoothing.*** Stipsicz, Szabó and Wahl [2008] derived strong restrictions on the resolution graph of a singularity if it admits a ℚHD smoothing.

Definition 2.1. A resolution graph Γ on n vertices is called a *symplectic plumbing tree* if its intersection lattice Q_Γ admits an embedding into the negative definite trivial lattice $(\mathbb{Z}\langle e_1, \dots, e_n\rangle, n\langle -1\rangle)$ such that $K = \sum e_i$ represents the anticanonical cycle of Q_Γ.

Theorem 2.2 [Stipsicz et al. 2008, Corollary 2.5]. *If a complex surface singularity admits a ℚHD smoothing, its resolution graph is a symplectic plumbing tree.*

The minimal resolution graphs with this property are completely classified; they belong to one of seven classes. The first, $\mathcal{G}$, consists of linear graphs, with framings given by the negatives of the Hirzebruch–Jung continued fraction coefficients of $p^2/(pq-1)$ for $p > q > 0$ coprime integers. The next three $\mathcal{W}, \mathcal{N}, \mathcal{M}$ are depicted in Figure 2.

The remaining $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are defined by repeated blowups of an edge next to the unique -1 vertex or the -1 vertex itself, beginning from the graph of Figure 3.

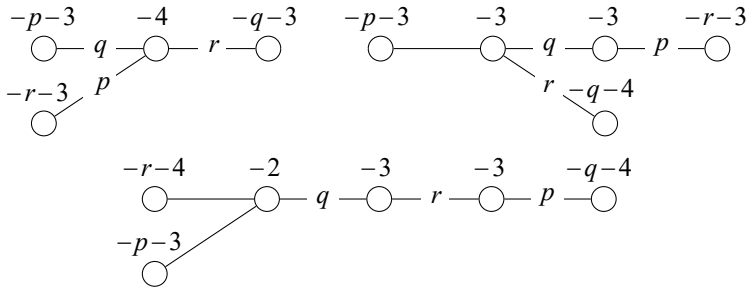


Figure 2. The graphs of $\mathcal{W}, \mathcal{N}$ and $\mathcal{M}$ where $p, q, r \geq -1$.

After the blowups, the framing of the -1 vertex (which is always unique) is changed to -4, -3, -2 to obtain an element of the $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ classes, respectively.

Theorem 2.3 [Stipsicz et al. 2008, Theorem 1.8]. *The set of minimal symplectic plumbing trees $\mathcal{S}$ equals $\mathcal{G} \cup \mathcal{W} \cup \mathcal{N} \cup \mathcal{M} \cup \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$.*

This is only a necessary condition; for the first four classes, a converse of Theorem 2.2 also holds.

Theorem 2.4 [Stipsicz et al. 2008, Theorem 1.10]. *For every $\Gamma \in \mathcal{G} \cup \mathcal{W} \cup \mathcal{N} \cup \mathcal{M}$, there is a complex surface singularity with resolution graph Γ admitting a $\mathbb{Q}$HD smoothing.*

The case of the $\mathcal{A}, \mathcal{B}, \mathcal{C}$ classes is only partially resolved, but there is a complete classification for weighted homogeneous singularities in [Bhupal and Stipsicz 2011, Theorems 1.4, 1.6], which will be proven in Section 6.

2.2. *Deformations of sandwiched singularities.*

Definition 2.5. A resolution graph Γ is called a *sandwich graph* if it arises as a subgraph of a graph that can be blown down to the empty graph. A singularity is called *sandwiched* if its resolution graph is a sandwich graph.

Note that the deformation theory of sandwiched singularities essentially reduces to deformations of plane curves, as will be explained momentarily. We rely on the picture deformation method developed in [de Jong and van Straten 1998].

Let $\rho : (\tilde{A}, E) \to (\mathbb{C}^2, 0)$ be obtained by point blowups starting from $(\mathbb{C}^2, 0)$ and the singularity X be obtained by contracting the non-(-1) curves from $\tilde{A}$. Choose

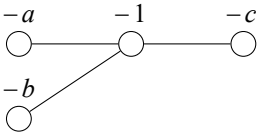


Figure 3. The starting graph for the construction of the $\mathcal{A}, \mathcal{B}, \mathcal{C}$ families, where $(a, b, c) = (3, 3, 3), (2, 4, 4), (2, 3, 6)$, respectively.

one transverse *curvetta* $\widetilde{C}_i$ on each -1 curve of $\tilde{A}$. Now $\tilde{A}$ can be seen as a good embedded resolution of $\rho(\widetilde{C})$.

Reversing the blowup process, de Jong and van Straten define:

Definition 2.6 [de Jong and van Straten 1998, Definitions 1.3, 4.1]. A *decorated curve* is a pair (C, l), where $C = \bigcup_I C_i \subset (\mathbb{C}^2, 0)$ is a curve with $|I|$ irreducible components and a function $l : I \to \mathbb{N}$ assigning to each irreducible component a number. Let $m(i)$ denote the sum of the multiplicities of the i-th branch in the minimal resolution of C. Then

$$m(i) \leq l(i) \tag{1}$$

for each $i \in I$. For discussing deformations, it is useful to encode the decoration l as a divisor on the normalization $\widetilde{C}$ of C.

From a decorated curve, we obtain a surface singularity as follows:

Definition 2.7 [de Jong and van Straten 1998, Definition 1.4]. Let (C, l) be a decorated curve. The modification $\tilde{A}(C, l) \to (\mathbb{C}^2, 0)$ determined by (C, l) is obtained from the minimal resolution of C by $l(i) - m(i)$ further blowups at the i-th branch of C. The analytic set $X(C, l)$ is obtained by contracting the exceptional curves not intersecting any branch of C in $\tilde{A}(C, l)$.

Remark 2.8. If during the blowup process, we don't reach the minimal *good* resolution of C, the analytic set can have multiple singular points.

Note that the same surface singularity may be presented in multiple ways from different decorated curves.

Next, de Jong and van Straten define the corresponding notion of deformation of sandwich singularities from the decorated curve data.

Definition 2.9 [de Jong and van Straten 1998, Definition 4.2]. A one-parameter deformation of a decorated curve (C, l) is a δ-constant deformation $C_\Delta \to \Delta$ of C (where $\Delta := (\mathbb{C}, 0)$) and a flat deformation $l_\Delta \subset \widetilde{C} \times \Delta$ (where $\widetilde{C}$ is the normalization of C) under the multiplicity condition of (1).

The multiplicity condition means a divisor on each fiber $l_t = l_\Delta \cap C_t = \sum_{p \in C_t} m_p p$ subject to $\sum_{p \in (C_t)_i} m_p = \sum_{p \in (C_0)_i} m_p = l(i)$, where $(C)_i$ denotes the i-th irreducible component of the curve C.

The important theorem for us is the following:

Theorem 2.10 [de Jong and van Straten 1998, Theorem 4.4]. *For every one-parameter deformation of* (C, l), *there is a one-parameter deformation of* $X(C, l)$, *and every one-parameter deformation of* $X(C, l)$ *appears in this form*!

Mostly we are interested in smoothing components. It is folklore that for rational singularities (which sandwich singularities are; see [Spivakovsky 1990]) every

component in the base space of the semiuniversal deformation space of X (denoted by $\mathcal{B}(X)$) is a smoothing component. That is, smoothing occurs over a generic point of this component. A one-parameter deformation is always induced from a well-defined component of $\mathcal{B}(X)$, and the generic fiber can be described by a particularly simple special case of the previous theorem.

Definition 2.11 [de Jong and van Straten 1998, Definition 4.6]. A one parameter deformation (C_Δ, l_Δ) is called a *picture deformation*, if for generic $t \neq 0$ the divisor l_t is reduced.

Remark 2.12. It follows that the generic fiber of the δ-constant deformation C_t only has transverse m-tuple points as singularities. Following [de Jong and van Straten 1998] we call 1-tuple points "*free*."

The general fiber of a picture deformation in general consists of a collection of curves, with the m-tuple points marked for $m > 1$, and possibly some further marked points on each component, so that in total, each component C_i contains $l(i)$ marked points. Since we will only be working with the intersection patterns of these curves, we will identify curves with their marked points, so that they are finite (multi)sets.

We mention one particularly useful construction, which produces a smoothing over the *Artin component* of $\mathcal{B}(X)$, called the *Scott deformation* of X.

Theorem 2.13 (see [de Jong and van Straten 1998, Proposition 1.10; Plamenevskaya and Starkston 2023, Proposition 4.1]). *Let (C, o) be an isolated curve singularity of multiplicity m. There is a one-parameter deformation, the general fiber of which only has singularities of the following two types*:

- *singularities of C after blowing up o,*
- *an m-tuple point.*

Given a sandwiched singularity using Theorem 2.10 and iterating Theorem 2.13, we obtain a picture deformation of X, the Milnor fiber of which is isomorphic to the resolution of X. This will be useful when we compare $\mathbb{Q}$HD smoothings to the resolution of the singularity in Section 4.

One can gather all of the incidence information of this combinatorial setup into a $|\{\text{points}\}| \times |\{\text{curves}\}|$ matrix $I := (\iota_{mn})$ where $\iota_{mn} = 1$ if point m is on curve n, and 0 otherwise.

Theorem 2.14 [de Jong and van Straten 1998, Theorem 5.2]. *There is an exact sequence*

$$0 \to H_2(F) \to \mathbb{Z}\langle p \in C_t : m_p \neq 0\rangle \xrightarrow{I} \mathbb{Z}\langle (C_t)_i\rangle \to H_1(F) \to 0, \tag{2}$$

where F is the Milnor fiber of the picture deformation.

Using [Greuel and Steenbrink 1983, Theorem 2], which states that for any smoothing of a normal surface singularity, the Milnor fiber F satisfies $b_1(F) = 0$, we get:

Corollary 2.15. $\operatorname{rk} H_2(F) =: \mu(F) = \#\{\text{points}\} - \#\{\text{curves}\}$.

Finally, we mention a method to recover the intersection lattice of the Milnor fiber of the smoothing from a picture deformation.

Theorem 2.16 [de Jong and van Straten 1998, Theorems 5.5, 5.11]. *The intersection form on $H_2(F) \subset \mathbb{Z}\langle p \in C_t : m_p \neq 0\rangle$ is the restriction of the Euclidean negative definite inner product on the latter space. Moreover, the canonical class is represented by* $(1, 1, \ldots, 1)$ *in this basis.*

3. Combinatorial smoothings

3.1. ***The minimal rational case.*** Throughout this section, we assume Γ to be a minimal rational graph with a given sandwich presentation where each component of the corresponding decorated curve is smooth.

Definition 3.1. In a tree graph Γ we denote the unique path between two vertices $v, w \in \Gamma$ by $p(v, w) := (v = v_1, \ldots, v_n = w) \subset V(\Gamma)$.

Definition 3.2. Given a minimal rational graph Γ, choose a vertex $v \in \Gamma$. Add $-(\deg(v) + e(v)) - 1 \geq 0$ many -1 leaves to it, and $-(\deg(w) + e(w))$ many -1 leaves to every other vertex (deg denotes the degree of the vertex in the graph Γ, and e is the Euler number of the vertex). This larger graph $\widetilde{\Gamma}$ is a *sandwich presentation* of Γ *with smooth branches*, where *the blowdown ends at* v.

We will be only interested in the incidence structure of the smoothings. To this end, we identify the components of the deformed curve $(C)_i$ with the support of their decoration divisor $(l)_i$. We make this identification concrete with the following definition.

Definition 3.3. Given a sandwich presentation of Γ with smooth branches define a *combinatorial smoothing* to be a tuple $(V, \mathcal{C})$, where V is a set of points, and $\mathcal{C} \subset \mathcal{P}(V)$ represents the curves corresponding to each -1 leaf of the presentation. We denote a curve on a vertex $w \in \Gamma$ by C_w, and v denotes the last vertex of the blowdown. The curves have to abide by the following two rules:

(1) $|C_w| = |p(w, v)| + 1$.

(2) $|C_w \cap C_z| = |p(w, v) \cap p(z, v)|$.

Remark 3.4. Using the Noether multiplicity formula, it is clear that given a sandwiched singularity with smooth branches, this is the incidence structure of a picture deformation.

Again, curves are identified with the support of their decoration divisor, and so are finite sets.

There are many different ways to give a sandwich presentation of Γ with smooth curves. This corresponds to the choice of v in Definition 3.2, the vertex where we put one curve less than would be required to blow it down. De Jong and van Straten [1998, Theorem 4.16.] explained how to relate combinatorial smoothings under a change of this last vertex v.

Theorem 3.5 [loc. cit.]. *Given a sandwich presentation of Γ with smooth branches, ending at a vertex v, a combinatorial smoothing $(V, \mathcal{C})$ and another vertex w with $\deg(w) + e(w) < 0$ one gets another sandwich presentation of Γ with smooth branches, ending at w by adding an extra -1 leaf to v, and removing one from w. There is a combinatorial smoothing of this graph $(\widetilde{V}, \tilde{\mathcal{C}})$ obtained as follows*:

(1) *Define $\widetilde{V} := V$.*

(2) *Choose an arbitrary C_w curve to remove.*

(3) *Define the new curve $\widetilde{C}_v := C_w$.*

(4) *For $C \neq C_w$ and $p \in V$ we have $\widetilde{C} = C_w \triangle C$.*

Remark 3.6. Switching from C_v to C_w, and then back to C_v, recovers the original smoothing.

Definition 3.7. We put a partial ordering on the graph: $x > y$ if $y \in p(x, v)$ where v is the last vertex to be blown down.

Remark 3.8. From the definition it is clear that if $x \leq y$, then $|C_x \cap C_y| = |C_x| - 1$, or stated differently $\exists! p \in C_x : C_x \setminus \{p\} \subset C_y$. On the other hand, if $x \not\geqslant y$ and $z \geq y$, then $|C_x \cap C_y| = |C_x \cap C_z|$. These basic properties will be used many times in the following.

Definition 3.9. A *reducing triple* consists of three (not necessarily distinct) vertices $v \leq w \leq z$, subject to the following conditions:

(1) $\{\deg(v), \deg(w), \deg(z)\} \subset \{1, 2\}$ or $v = w = z$.

(2) There is one curve on w and at least one on v and z.

(3) No vertex in $p(v, z) \setminus \{v, w, z\}$ has any curves.

(4) $\Gamma \setminus p(v, z)$ and $p(v, z)$ only connect with edges in Γ at v and z.

Item (2) is to be understood additively, i.e., if say $v = w$, then it should have at least two curves (one for w and at least one more for v).

Definition 3.10. For a reducing triple $v \leq w \leq z$, we let the blowdown end at v and pick a distinguished curve on z and denote it by C_z. The curve on w is denoted by C_w. We name $C_w \setminus C_z =: \{Q\}$, $C_w \cap C_z =: P$, $C_z \setminus P =: Q'$. Similarly, we will denote curves on a vertex x by C_x^1, C_x^2, and so on.

3.2. ***The reduction algorithm.*** Next, we prove a few simple statements in the context of a given reducing triple. From now on, we also suppose that every node satisfies $\deg(v) + e(v) \leq -2$.

Lemma 3.11. *Let $w \neq x \in \Gamma$ be a vertex. If $Q \in C_x$, then $Q \in C_a$ for all $x \leq a$.*

Proof. Since $Q \in C_x$, either $Q' \subset C_x$ if $x \geq z$, or $\exists! q' \in Q' : q' \in C_x$ if $x \not\geq z$. Now supposing $Q \notin C_a$ implies $C_x \setminus \{Q\} \subset C_a$ since $x \leq a$. In the first case this implies that $P \subset C_a$ by intersection count with C_w and since $Q' \subset C_x \setminus \{Q\}$ we get that $C_z = P \cup Q' \subset C_a$, a contradiction. Similarly in the second case $q' \in C_a$ implies that $Q \in C_a$, which is a contradiction. □

Corollary 3.12. *If $Q \in C_x$, $Q \notin C_y$, then $x \neq y$ unless $w = z = x = y$.* □

Remark 3.13. By this corollary, we can say that a *vertex* $v \in \Gamma$ *contains* Q or not, dependent on if its curves do.

Lemma 3.14. *If $a, b \in \Gamma$ are in different components of $\Gamma \setminus \{v\}$ and $Q \in C_a$, then $Q \notin C_b$.*

Proof. The $\deg(v) = 1$ case is vacuous. Secondly, if $\deg(v) = 2$ and $Q \in C_a \cap C_b$, suppose $z \leq a$. Now b being in a different component means that C_b has to intersect C_a once, but from $Q \in C_b$, we get that there is also a $q' \in Q'$ which is in C_b since it has to intersect C_z once. The relation $z \leq a$ implies that $Q' \subset C_a$, and thus $|C_a \cap C_b| \geq 2$, which is a contradiction.

In the third case, when the reducing triple is v, v, v, Q' consists of a single point, and the same argument shows that $Q \in C_a \cap C_b$ implies $Q' \in C_a \cap C_b$, but $|C_a \cap C_b| = 1$ by assumption. □

Lemma 3.15. *Let $(w \neq) x < a, b$ with $a \not\geq b$ in two different components of $\Gamma \setminus \{x\}$. If $Q \notin C_x^1, C_x^2$, then Q cannot be in both C_a and C_b.*

Proof. Assume $x > z$ and let $n + 1 = |C_x^i|$. If $Q \in C_a$, then there is a point $q' \in Q'$ which is in C_a and a point $p \in P$ with $p \notin C_a$. Call $C_x^i \setminus C_x^j = \{p_i\}$. Since C_a does not contain a point of P and contains Q, it also contains both p_i's. This means that if $Q \in C_b$ were true, then C_a and C_b would intersect in at least $n - 2$ points of $C_x^1 \cap C_x^2$, in Q and p_1, p_2 for a total of at least $n + 1$ points, which is a contradiction since they have to intersect n times by assumption.

If x is in another component of $\Gamma \setminus \{v\}$, we argue similarly. The C_x^i intersect C_w, C_z in a single point, which can be in $C_x^1 \cap C_x^2$, or the p_i's. In either case, C_a, C_b have to intersect in the remaining at least n points of $C_x^1 \cup C_x^2$, and in Q, which is a contradiction. □

We can extend the notion of reducing triple to the case where we replace Definition 3.9(1) with $v = w$, $\deg(v) = \deg(w) < 3$ and $\deg(z) > 2$.

Lemma 3.16. *With the modified assumption on the reducing triple, we have that* $C_z^1 \cap C_z^2 \cap C_v = \{P\}$.

Proof. Let the blowdown end at v. We choose one of the curves on z to be $C_z := C_z^1$, and define $C_z \cap C_v =: \{P\}$, $C_v \setminus \{P\} =: \{Q\}$ and $C_z \setminus \{Q\} =: Q'$ as usual. Furthermore, assume that $Q \in C_z^2$. Pick $x > z$ with a curve. Without loss of generality, assume that $P \in C_x$. Then $Q' \subset C_x$, since it missed $Q \in C_z^2$, and it cannot intersect C_z^2 enough times otherwise. This implies $C_z \subset C_x$, which is a contradiction. □

This tells us which point of C_v to label Q. Now we only need to check that only one component of $\Gamma \setminus \{z\}$ and $\Gamma \setminus \{v\}$ can contain curves with Q. Afterwards Lemmas 3.11 and 3.15 are applicable.

Lemma 3.17. *If* $x \not\gtrless y$ *are in two different components of* $\Gamma \setminus \{z\}$, *then* x, y *cannot both contain* Q.

Proof. The assumption $Q \in C_x \cap C_y$ implies that C_x and C_y intersect $n+1$ times, since they have to contain $(C_z^1 \cup C_z^2) \setminus \{P\}$, which is a contradiction. □

Lemma 3.18. *Consider* $x > z$ *and* y *in another component of* $\Gamma \setminus \{v\}$. *Then* x, y *cannot both contain* Q.

Proof. By $Q \in C_x$, we get that $Q' \subset C_x$, and by the assumption on y, we get that $C_y \cap Q' = \varnothing$. So $C_y \cap C_z = \varnothing$, which is a contradiction. □

Consider a graph Γ and a $\mu = 0$ combinatorial smoothing of it $(V, \mathcal{C})$. In this case, the map I of Theorem 2.14 is injective, in particular invertible, so by [de Jong and van Straten 1998, Corollary 5.12 and Theorem 13], such a combinatorial smoothing is a qG-smoothing (see [Kollár 1991]) and, in particular, has no free points. This guarantees us that there will be a curve besides C_w containing Q. The preceding lemmas tell us that there is an edge of the graph that separates the vertices containing Q from the vertices not containing it. We call such an edge a *separating edge*.

Now, if there is a reducing triple in Γ, we can construct a smaller graph and a combinatorial smoothing for it by contracting a separating edge, deleting Q from V and C_w from $\mathcal{C}$. This can be iterated until the graph has no further reducing triples. A graph with no reducing triples consists of nodes with two curves and other vertices supporting at most one curve, such that every component of $\Gamma \setminus \{\text{nodes}\}$ has at most two curves total.

Definition 3.19. We call such a graph *reduced*.

Remark 3.20. During reduction, the number of nodes, leaves, and components of $\Gamma \setminus \{\text{nodes}\}$ can only decrease.

Definition 3.21. Given a sandwich graph Γ we define

$$\delta(\Gamma) := \left(\sum_{v \in \Gamma} -e(v) - d(v) \right) - 1 - |V(\Gamma)|.$$

Remark 3.22. Given $(\mathcal{C}, V)$, a combinatorial smoothing of Γ with smooth branches, we see $\delta = |\mathcal{C}| - |V(\Gamma)|$. Notice also that $\delta(\Gamma) = 1 - \sum_\Gamma (e(v) + 3)$, and compare with Wahl's formula [2011, Theorem 8.1], which states that the dimension of a ℚHD smoothing component is given by $h^1(S) - \sum_\Gamma (e(v) + 3)$, where S is the sheaf of *logarithmic vector fields*, the kernel of the map $T_X \to \oplus \nu_v \to 0$.

For graphs in $\mathcal{A}, \mathcal{B}, \mathcal{C}$, this quantity is always $3 - s$ if the graph only has degree-3 nodes, and $2 - s$ if it has a degree-4 node, where $s := \#\{v \in \Gamma : \deg(v) > 2\}$.

The above proposed algorithm leaves this quantity unchanged.

Lemma 3.23. *After reducing a graph $\Gamma \in \mathcal{A}^3, \mathcal{B}^3, \mathcal{C}^3$ with s many nodes, the reduced graph γ satisfies $|\gamma| \leq 7s - 2$. Similarly for a graph $\Gamma \in \mathcal{A}^4, \mathcal{B}^4, \mathcal{C}^4$ we have $|\gamma| \leq 7s + 1$.*

Proof. In the first case $\Gamma \setminus \{\text{nodes}\}$ has at most $2s + 1$ components, so γ has at most this many.

This implies that γ has at most $2s + 2(2s + 1) - 1 = 6s + 1$ curves. Writing $3 - s \leq 6s + 1 - |\gamma|$ and rearranging gives the result.

For the case where Γ has a degree-4 node, we compute in a similar way that $2s + 2(2s + 2) - 1 = 6s + 3$ is the upper bound on the curves, $2 - s \leq 6s + 3 - |\gamma|$ and the statement follows. □

Using Lemma 3.23 and computer checking we obtain special cases of [Bhupal and Stipsicz 2011, Theorems 1.4, 1.6; Wahl 2011, Theorem 8.6]:

Corollary 3.24. *Let $\Gamma \in \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ be a graph with one node of degree* 3 *(resp.* 4*) with framing at most* -5 *(resp.* -6*). Then Γ cannot admit a* ℚHD *smoothing.*

Proof. Suppose the graph has a ℚHD smoothing; apply the reduction algorithm to its incidence structure. The reduced graph still has $\delta = 2$ (1), with $e(v) = -2$ if $\deg(v) = 1$, $e(v) \in \{-2, -3\}$ if $\deg(v) = 2$, $e(v) = -\deg(v) - 2$ otherwise. A linear graph under these framing restrictions has $\delta \leq -1$, which is an immediate contradiction. Thus, the graphs have a node and at least 3 leaves. This means that they have $\delta \leq 0$, which is a contradiction. □

Corollary 3.25. *Let $\Gamma \in \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ be a graph with two nodes of degree* 3 *and framing at most* -5*. Then Γ cannot admit a* ℚHD *smoothing.*

Proof. Suppose Γ admits a ℚHD smoothing. Choosing a sandwich presentation for Γ with smooth curvettas (since it is assumed to be minimal rational), we get a combinatorial ℚHD smoothing, to which we apply the reduction algorithm, after which we obtain a reduced graph γ.

By Lemma 3.16, we have the same framing possibilities as before. Since δ stays constant, the reduced graph has $\delta = 1$, since this is true for the starting graphs. By the previous corollary, graphs with a single node have $\delta \leq 0$, so all possible graphs have two nodes. Checking all trees γ with $|\gamma| \leq 12$ under the constraints derived thus far, we see that there is no reduced graph with nonsquare determinant. □

Further, we can extend this to graphs with a degree-4 and a degree-3 node, and to the cases with 3 and 4 nodes as well. In these cases, the determinant can be a square, so we need to rely on Theorem 4.1, proved below, which states that for a graph admitting a combinatorial $\mathbb{Q}$HD smoothing, one has $Z_K^2 + |\Gamma| = 0$. Unfortunately the computation of the d-invariant[1] for general reduced graphs remains elusive.

Theorem 3.26. *Let $\Gamma \in \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ be a graph with one node of valency 4 and framing at most -6, and at most 2 nodes of valency 3 and framing at most -5. Then Γ cannot admit a $\mathbb{Q}$HD smoothing.* □

Theorem 3.27. *Let $\Gamma \in \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ be a graph with at most four nodes of valency 3, and framing at most -5. Then Γ cannot admit a $\mathbb{Q}$HD smoothing.* □

4. Lattice embeddings

In this brief section, we wish to prove the following:

Theorem 4.1. *If a sandwiched graph Γ admits a combinatorial rational homology disk smoothing, then the anticanonical cycle of the graph Z_K satisfies $Z_K^2 + |\Gamma| = 0$.*

Proof. Consider a sandwiched graph Γ with a fixed sandwich presentation $\widetilde{\Gamma}$, which admits a combinatorial $\mathbb{Q}$HD smoothing. Following [de Jong and van Straten 1998], we let $P' = \mathbb{Z}^l$ denote the free abelian group generated by the points of the combinatorial $\mathbb{Q}$HD smoothing and endow it with the trivial negative definite intersection form. Similarly, let $L = \mathbb{Z}^l$ denote the free abelian group generated by the set of curvettas of the sandwich presentation $\widetilde{\Gamma}$. Consider the map $J : P' \to L$ where $p \mapsto \sum_{p \in L_i} L_i$, i.e., the incidence matrix of the combinatorial smoothing. Since $\operatorname{rk} P' = \operatorname{rk} L$, by [de Jong and van Straten 1998, Corollary 5.10], the matrix of this map in the natural bases is invertible over $\mathbb{Q}$.

Consider also the incidence map of the Scott deformation $I : P = \mathbb{Z}^{l+n} \to L$. As mentioned in [de Jong and van Straten 1998, Corollary 1.11; Plamenevskaya and Starkston 2023, Subsection 4.2], the Milnor fiber of this deformation is diffeomorphic to the resolution of the singularity; thus, by Theorem 2.16, we have $Q_\Gamma \hookrightarrow P$.

Next note that $II^* = JJ^* = Q_{\text{artin}}$ in the language of [de Jong and van Straten 1998]; thus $J^{-1}II^*J^{*-1} = (J^{-1}I)(J^{-1}I)^* = \mathrm{id}$, which means that the rows of the matrix $J^{-1}I$ form an orthonormal basis for $\ker(J^{-1}I)^\perp = \ker(I)^\perp \leq P \otimes \mathbb{Q}$.

[1] That is, $\frac{1}{4}(Z_K^2 + |\Gamma|)$; see, e.g., [Némethi 2022].

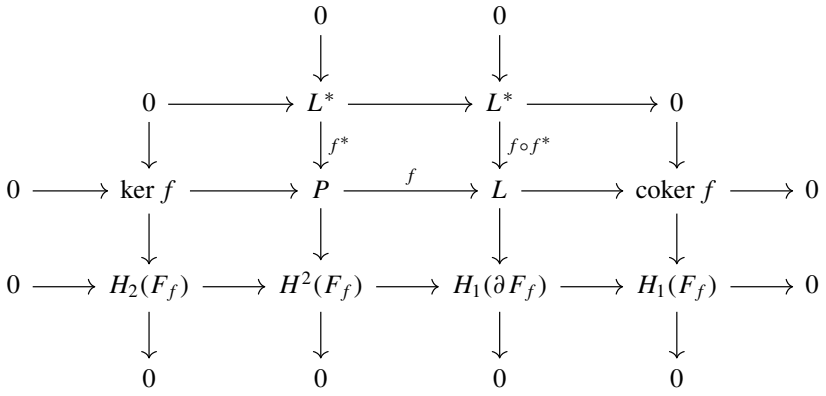


Figure 4. f denotes the incidence matrix of a combinatorial smoothing, and F_f the Milnor fiber [de Jong and van Straten 1998, Diagram 5.9].

Let $1_\alpha = (1)_1^\alpha$. The product $I\,1_{l+n}$ equals $(|L_i|)_1^l$. Since the number of points on each curvetta is determined by the sandwich presentation $\widetilde{\Gamma}$, the same is true for $J\,1_l$. Thus $J^{-1}I\,1_{l+n} = 1_l$. This means that denoting the rows of $J^{-1}I$ by f_i we obtain a basis where taking $K = \sum_{p\in P} p$ we have

$$(J^{-1}I)^*(J^{-1}I)K = \sum f_i. \tag{3}$$

Until now, we were concerned with $\ker(I)^\perp$. We have $P\otimes\mathbb{Q} = (\mathbb{Q}^{l+n}, \langle -1\rangle^{l+n})$, and an embedded $(\mathbb{Q}^l, \langle -1\rangle^l)$. By Witt's cancellation theorem (see Lemma 4.4 in [Milnor and Husemoller 1973]) this means that the orthogonal complement generated by $\ker(J^{-1}I)$ is isomorphic to $(\mathbb{Q}^n, \langle -1\rangle^n)$.

By (3) the orthogonal projection of K onto $\ker(I)^\perp$ is equal to $\sum f_i$, and so has square $-l$, which implies that the orthogonal projection to $\ker(J^{-1}I)$ will have square $-n$. From Theorem 2.16, this projection represents the anticanonical cycle of Γ; thus we get that $Z_K^2 = -n$. □

Proposition 4.2. *If a sandwiched graph Γ admits a combinatorial rational homology disk smoothing, then $Q_\Gamma \subset U$ for some unimodular lattice U, with* $\operatorname{rk} U = |\Gamma|$.

Proof. Let $I : P \to L$ denote the incidence matrix of the Scott deformation, and $J : P' \to L$ denote the incidence matrix of a combinatorial ℚHD deformation, as previously. By Figure 4 [de Jong and van Straten 1998, Diagram 5.9] we get that I is surjective and J is injective. In particular $H^2(F_J) \le H_1(\partial F_J)$ is a subgroup of the discriminant group of Q_Γ, which is isomorphic to $H_1(\partial F)$ of size $\sqrt{|H_1(\partial F)|}$. By commutativity of the diagram and the fact that P is endowed with the trivial negative definite intersection form, this subgroup is isotropic, which by [Looijenga and Wahl 1986, 1.7.1] gives an integral unimodular overlattice U for Q_Γ in Q_Γ^*. □

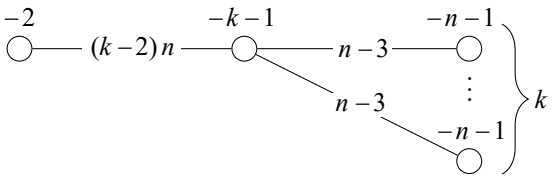


Figure 5. The graph $\mathrm{Cl}(k,n)$. The number x on an edge represents a path of x many -2 vertices.

5. Some graphs admitting a combinatorial ℚHD smoothing

Example 5.1. Let us denote by $\mathrm{fpp}(n)$ the star-shaped graph with n^2+n+1 arms, each of length $n-1$, with the node having framing $-n^2-n-2$, and all other vertices -2. A combinatorial ℚHD smoothing of this graph is the finite projective plane of order n. Each graph $\mathrm{fpp}(n)$ satisfies $\det(\mathrm{fpp}(n)) = n^{(n+1)n}(n+1)^2$, and they all have $Z_K^2 + n = 0$.

By modifying the finite projective plane configuration, we obtain further graphs with combinatorial ℚHD smoothings. Pick a point x and all lines through it. Add further points $a_1,\dots,a_l$, and curves $\{x,a_i\}$ to the configuration, and for all lines that do not contain x, add all of the a_i. This is a ℚHD smoothing for the graph we denote by $\mathrm{fpp}(n)_l$.

Example 5.2. Consider k "*clusters*" consisting of n points each: $\{A_i\}_1^k$ with $|A_i| = n$. Define sets of curves as $C^i_j = \bigcup_{l\neq i} A_l \cup \{a_j\}$ where $a_j \in A_i$. This is a combinatorial ℚHD smoothing of a graph, which we denote by $\mathrm{Cl}(k,n)$. See Figure 5.

Example 5.3. We define the cluster extension of the graphs $\mathrm{Cl}(k,n)$. Here, consider another set B, with $|B| = b$, and add its points to all existing curves. Furthermore add new curves $D_i = \bigcup A_l \cup \{b_i\}$. The construction corresponds to a combinatorial smoothing of a graph if $b \geq n \geq 2$. See Figure 6.

Example 5.4. Another construction one can make is the star extension of $\mathrm{Cl}(k,n)$. Add a new point x and a new set B of size b. Add the elements of B to the existing curves, and define new curves $S_i = \{x\} \cup \{b_i\}$ and $F = \bigcup_1^k A_i \cup \{x\}$. See Figure 7.

Example 5.5. Using the complement idea, we generalize the family of Wahl [1981, Section 5.9.2]. In this case we consider three sets A, B, C of sizes $|a|$, $|b|$, $|c|$,

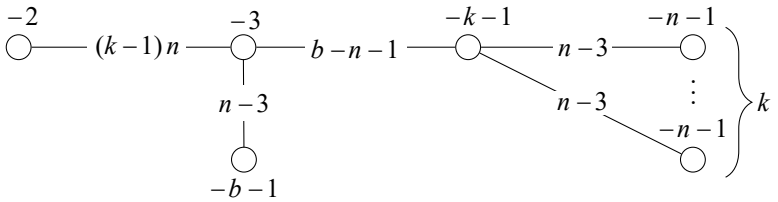


Figure 6. The graph we obtain after cluster extending $\mathrm{Cl}(k,n)$.

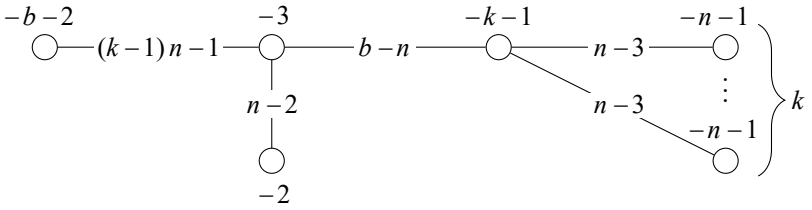


Figure 7. The graph we obtain after star extending $\mathrm{Cl}(k, n)$.

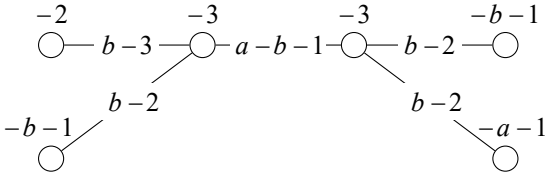


Figure 8. The graph $t(-a, -b, -b)$.

the curves are $A \cup \{b_i\}$, $B \cup \{c_i\}$, $C \cup \{a_i\}$ with $a_i \in A$, $b_i \in B$, $c_i \in C$, or one can take the complements of the singletons inside their containing sets. We denote this choice by flipping the sign of the set that the corresponding curves contain, i.e., $a, -b, c$ uses the curves $A \cup \{b_i\}$, $B \cup (C \setminus \{c_i\})$, $C \cup \{a_i\}$. This gives combinatorial smoothings for the graphs we denote by $t(a, b, c)$. If all parameters are positive, we get Wahl's family $\mathcal{W}$ (Figure 2, left), if one is negative, we get the $\mathcal{N}$ family (Figure 2, right). In the case $t(a, -b, -c)$ for the curves to be the smoothing of a graph we need either $a = 2$ or $c = 2$ and get the subfamily of N with $p = 0$ and the subfamily of W with $q = 0$. If all three are negative with $|a| > |b| = |c|$ we get the novel graph of Figure 8.

6. Weighted homogeneous singularities — stars

This section presents a new proof of the main theorem of [Bhupal and Stipsicz 2011], namely the classification of weighted homogeneous singularities admitting a ℚHD smoothing. From Theorem 2.4, we only need to check the star-shaped elements of $\mathcal{A}, \mathcal{B}, \mathcal{C}$. In contrast to Section 3, not all graphs will be minimal rational; most curvettas will be ordinary cusps.

6.1. ***The degree-3 case.*** Let us begin with some generalities, which will be used multiple times in the following. We will be dealing with star-shaped graphs with three arms, where two of the arms have length 1 and decorations $-a$ and $-b$, respectively. We will always put $\max\{a-2, 0\}$ many -1 framed vertices on the $-a$ vertex and $\max\{b-3, 0\}$ many -1 framed vertices to the $-b$ vertex. Denote curvettas on these vertices by S_i and L_i, respectively. The third "long" arm will be blown down starting from the end and decorated as necessary. Take an arbitrary

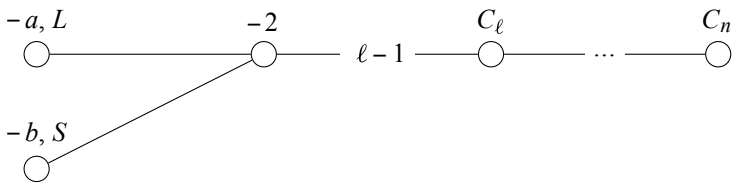


Figure 9. The general case graph with the curves indicated.

order on the -1 leaves on every vertex, and denote the j-th curvetta of the i-th vertex by C_i^j. See Figure 9 for an illustration.

Since the node has framing at most -2, we know that the C_i^j will be star-shaped and the L_i, S_i will be smooth. We provide their intersections in Table 1, right.

The multiplicity sequence of these cusps are easily computed to be $(2, 1, \dots, 1)$, so C_i^j must all have one double (2-tuple) point and further simple points (i.e., at all but one point the curve is locally irreducible). For a cusp curve C, we define $X(C)$ to be the point in P which it contains with multiplicity 2. From this, it is simple to compute how many points each curve has to contain and their intersection multiplicities, this is collected in Table 1, left. We denote the points where these curves intersect by $\{p_i : 1 \leq i \leq 4+n\}$.

In this section, n will always denote the number of vertices on the "long" arm (on which the C_i^j are located). From the construction of the $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ classes, it is easy to see by induction that there will always be $4+n$ many curves in total in the sandwich presentation described above. We will mostly be interested in the smooth curves, the curves at the very end C_n^α, and the curve closest to the node. We let ℓ denote the index where there exists C_ℓ^1, and for every $k < \ell$, there is no C_k^1.

Remark 6.1. As before, we work with the intersection pattern of the given sandwich presentation of the graph, and by a small abuse of notation, we identify curves with the intersection points they contain.

We note some simple facts about these curves:

curve	L_i	S_i	C_i^j
# of points	2	3	5+i

$(\cdot,\cdot)$	L_a	S_a	C_a^b
L_i	1	1	2
S_i		2	3
C_i^j			$6+\min\{i,a\}$

Table 1. The combinatorial information in our setup. Left: Point data. Right: Intersection data.

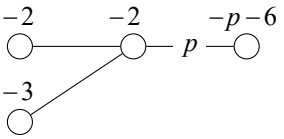


Figure 10. Graphs in $\mathcal{C}_6$ admitting a ℚHD smoothing.

Lemma 6.2. *Consider C_i^j and C_a^b for $i \le a$. If $X(C_i^j) \neq X(C_a^b)$, then $C_i^j \subset C_a^b$ (as sets, without multiplicities). If $X(C_i^j) = X(C_a^b)$, then there is a single $p \in C_i^j$ with $p \notin C_a^b$.*

Proof. If their double points differ, then we have $p_d := X(C_a^b) \in C_i^j$, otherwise, $(C_a^b, C_i^j) \le 5+i < 6+i$, and if $p_d \in C_i^j$, then it has to go through every point for the intersection to reach $6+i$.

On the other hand, if their double points coincide, then there are a further $3+i$ simple points to make $2+i$ intersections. □

Corollary 6.3. *The C_n^α curves all have distinct double points, and all other C_i^j curves must pass through these double points with multiplicity* 1. *The L_i and S_i cannot go through these points.*

Proof. Since the C_n^α curves contain $5+n$ points with multiplicity, and there are $4+n$ points in total, they go through all points, i.e., it is impossible for the second case of Lemma 6.2 to occur.

Both L_i and S_i would intersect the C_n^α curves too many times if they went through the double points. □

We state the following for emphasis:

Corollary 6.4. *All C_i^j must go through every other C_a^b's double point.*

6.1.1. *Graphs in $\mathcal{C}$.* We begin the case analysis with the class $\mathcal{C}$.

Theorem 6.5. *In $\mathcal{C}_6$, only the family depicted in Figure 10 can admit a* ℚHD *smoothing.*

Proof. In this case, there are no L_i, S_i, only the cusps. The known case is $\ell = n$, and we want to rule out any other possibility. It is easy to see by induction that there will be $6+\ell$ many C_n^α curves if $\ell < n$. By Corollary 6.4, C_ℓ^1 would have to contain $2+6+\ell > 5+\ell$ many points, which is a contradiction. □

Theorem 6.6. *In $\mathcal{C}_3$, only the graphs of Figure 11 can admit* ℚHD *smoothings.*

Proof. In this case, we have three L_i's in addition to the cusps. If $\ell = n$, we are done, so let $\ell < n$. Simple induction shows that there will be $3+\ell$ many C_n^α's, so C_ℓ^1 contains their double points, its double point and no further point. This means that the L_i have to go through C_ℓ^1's double point to intersect it twice by Corollary 6.3.

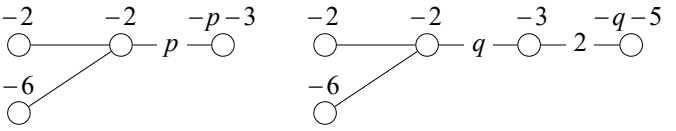


Figure 11. The graphs for the C_3 case, $p \geq 0$, $q \geq -1$.

If there would be another C_k^α with $\ell \leq k < n$, then it intersects C_ℓ^1 least $7+\ell$ times by Corollary 6.4, a contradiction. So there is C_ℓ^1, and the C_n^α curves. This means that $3+\ell = 4+n-3-1$, ergo $\ell = n-3$, and this is the other known case. □

Theorem 6.7. *In $\mathcal{C}_2$, only the graphs of Figure 12 can admit a $\mathbb{Q}$HD smoothing.*

Proof. In this case, there will be four S_i curves and the cusps. Let $\ell < n$. There will be $2+\ell$ many cusps on the last vertex.

Counting the number of points of C_ℓ^1 it has (besides the C_n^α double points) $p_d := X(C_\ell^1)$ and a further simple point denoted by p_s. All of the S_i have to go through p_d and p_s since $(S_i, C_\ell^1) = 3$. This also means that they have a further point each p_{S_i}. If there is a second C_i^j for some $\ell \leq i < n$, then $X(C_i^j) = p_d$ or p_s.

If $X(C_i^j) = p_d$, then p_s will be the point C_i^j misses by Lemma 6.2, and $p_{S_i} \in C_i^j$. If there were a third C_a^b curve with $i \leq a < n$, we see that $X(C_a^b) = p_d$ by Corollary 6.4. By the same argument we get that $p_{S_i} \in C_a^b$ for $i = 1, \ldots, 4$.

Every subsequent curve meets these same requirements. Let $K := \mathcal{C} \setminus \{S_i, C_n^\alpha, C_\ell^1\}$, with $|K| = k$. Let $\widetilde{K} := \bigcup K \setminus \{X(C_n^\alpha) : 1 \leq \alpha \leq \ell+2\} \subset V$ with $|\widetilde{K}| = \tilde{k}$. Counting the curves and points of the configuration, we get

$$2+\ell+1+4+k = 2+\ell+2+4+\tilde{k} = 4+n, \tag{4}$$

implying $k = 1+\tilde{k}$.

Claim. $C_i^j = C_{\ell+3}^1$.

Proof. If C_i^j would contain any further points, the curves in K restricted to the points of $\widetilde{K}$ give a solution to the combinatorial smoothing problem for a linear graph, since $X(C) = p_d$ for all $C \in K$. This implies $k \leq \tilde{k}$, which is a contradiction. □

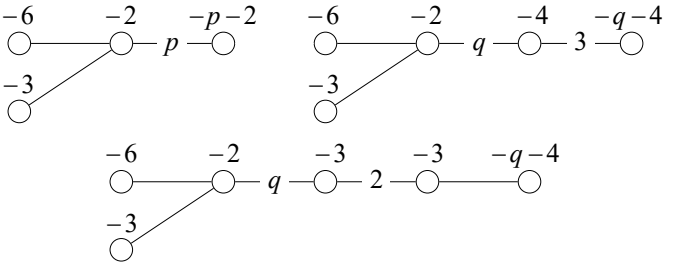


Figure 12. The graphs for the C_2 case, $p \geq 0$, $q \geq -1$.

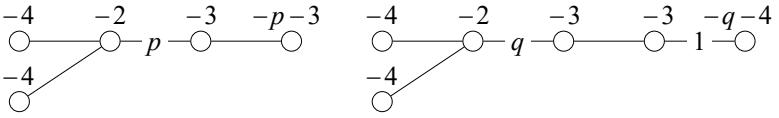


Figure 13. The graphs for the B_2 case, $p \geq 0$, $q \geq -1$.

Any further curve would also go through p_d and p_{S_i}, which means it would intersect $C^1_{\ell+3}$ $10+\ell > 9+\ell$ times, which is a contradiction. Substituting $k=1$ into (4) we see that $\ell = n-4$; this is the third graph of Figure 12.

The other case plays out similarly. If $X(C_i^j) = p_s$, then every other curve has to have its double point at p_s and go through p_d to intersect S_i and C^1_ℓ enough times. The same argument shows that $C_i^j = C^2_\ell$, there are no further curves and $\ell = n-4$. This is the second graph of Figure 12.

If there is only one cusp curve that is not on the last vertex, then we have $4+n = 2+\ell+1+4$, ergo $\ell = n-3$. The C^α_n's require $2+n-3 = n-1$ points, C^1_ℓ an additional 2. We have only 3 points left, which is a contradiction since the S_i need 4 additional points besides the previously mentioned ones.

Finally, if $\ell = n$, we get the first graph of Figure 12. □

6.1.2. *Graphs in $\mathcal{B}$.* We continue with the $\mathcal{B}$ class.

Theorem 6.8. *In $\mathcal{B}_2$, only the graphs of Figure 13 can admit a* ℚHD *smoothing.*

Proof. In this case, there are L_1, S_1, S_2 and the cusps. Suppose that not every blowup happened on the edge next to the leaf. In this case, the setup of the cusps is identical to the $\mathcal{C}_2$ case, the curve C^1_ℓ has (besides the $X(C^\alpha_n)$ points) a double point p_d and one further simple point p_s. The points p_s, p_d have to be contained in the S_i in order to intersect C^1_ℓ three times (and both have one further point p_{S_i}). The curve L_1 has to go through p_d to intersect C^1_ℓ twice, and thus it contains one further point p_{L_1}.

We continue similarly to the $\mathcal{C}_2$ case. Any further curve can have its double point at either p_d or p_s, and all of them have to be at the same point. Denote the curve with the next smallest index after ℓ by C_i^j. From here, the same argument provides that C_i^j is $C^1_{\ell+1}$ in both cases (if its double point is p_d, it must contain p_{S_1} and p_{S_2}; if its double point is p_s, it must contain p_d and p_{L_1}). In both cases, there are no further curves because they would intersect $C^1_{\ell+1}$ too many times. This gives the second graph of Figure 13.

If every blowup happens next to the leaf, we get the first graph of Figure 13. □

Theorem 6.9. *In $\mathcal{B}_4$, only the graph of Figure 14 can admit a* ℚHD *smoothing.*

Proof. In this case, we have L_1 and the cusps. Suppose not every blowup happened next to the leaf. Then a simple induction shows that there are $4+\ell$ many C^α_n curves. This means C^1_ℓ goes through $6+\ell$ points with multiplicity, which is an immediate contradiction. □

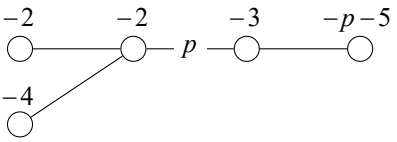


Figure 14. The graph for the B_4 case, $p \geq 0$.

6.1.3. *Graphs in $\mathcal{A}$.*

Theorem 6.10. *In $\mathcal{A}_3$, only the graph of Figure 15 can admit a $\mathbb{Q}$HD smoothing.*

Proof. In this case, there are S_1 and the cusps. Suppose not every blowup happened next to the leaf. There will be $3+\ell$ curves at the last vertex by induction. The curve C_ℓ^1 will go through the double points of the C_n^α curves, and it has its double point and no further points. By Corollary 6.3, S_1 can only intersect C_ℓ^1 at p_d, so at most 2 times instead of the required 3, which is an immediate contradiction. Thus, we are left with the case when every blowup happens next to the leaf, depicted in Figure 15. □

6.2. ***The degree-4 case.*** Most of the setup stays unchanged in this case; there will still be $4+n$ many curves, the end curves have to go through all points, etc. In addition to the L_i, S_i, there will also be some *exceptional cusps* Γ_i. They intersect the L_i, S_i in the same manner as the C_i^j curves do, and intersect the C_i^j curves 6 times.

Lemma 6.11. *$X(C_n^\alpha) \notin \Gamma_i$ for any i.*

Proof. If this were not the case, they would intersect 7 times. □

In the following, we will assume that not all blowups happen next to the leaf. As in the previous cases, a simple calculation shows, that C_ℓ^1 has three points (besides the double points of the C_n^α curves of course): its double point denoted by p_d, and two simple points p_{s_1} and p_{s_2}. We fix this notation in this section.

Lemma 6.12. *Every curve after C_ℓ^1 can have its double point at a maximum of two of the three points p_d, p_{s_1}, p_{s_2}, i.e.,*

$$\bigcup_{j>1} X(C_\ell^j) \cup \bigcup_{\substack{n>i>\ell \\ j>0}} X(C_i^j) \subsetneq \{p_d, p_{s_1}, p_{s_2}\}.$$

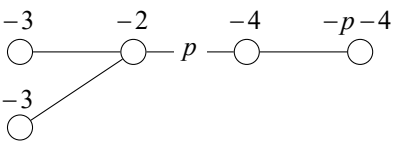


Figure 15. The graphs for the A_3 case, $p \geq 0$.

Proof. There are two cases to consider. The first curves have their double points at p_d, p_{s_1} or p_{s_1}, p_{s_2} and both are similar. If there is a curve with double point at p_d besides C^1_ℓ, it has to miss one of the p_{s_i}, say p_{s_2}. This means no curve can have its double point at p_{s_2} by Corollary 6.4.

Symmetrically, if every subsequent curve has its double point at the p_{s_i}, no curve can have its double point at p_d since it would have to miss one of the p_{s_i} by Lemma 6.2. □

Corollary 6.13. *If $X(C^j_i) = p_d$, then $X(C^b_a) = p_d$ for any $a < i$.*

Proof. Such curves have to miss one of the p_{s_i}'s, and so a point of any curve with double point at any of these two points, which is a contradiction. □

Lemma 6.14. *If every curve B after C^1_ℓ until C^j_i has $X(B) = p_d$ and the next curve C satisfies $X(C) = p_{s_1}$, then all subsequent curves have their double points at p_{s_1}. The same statement is true if we change the order of p_d, p_{s_1} or replace p_d, p_{s_1} with p_{s_1}, p_{s_2}.*

Proof. Let us denote a curve after C by C' with $X(C) \neq X(C')$. If $X(C^j_i) = p_d$, then by assumption we know that $p_{s_1} \in C^j_i$ and $p_{s_2} \notin C^j_i$; thus we only have to rule out $X(C') = p_d$. This is automatic by Lemma 6.2: since $p_d, p_{s_1}, p_{s_2} \in C$, the curve C' will also contain these points, and thus all points of C^1_ℓ. This is a contradiction, since $X(C^1_\ell) = X(C')$.

The case where $X(C^j_i) = p_{s_1}$ is simpler; if $X(C) = p_d$, then $p_{s_2} \notin C$, which is a contradiction since $p_{s_2} \in C^j_i$.

If $X(C) = p_{s_2}$, then $X(C') \neq p_d$, since it would miss a point of C^j_i, C. Since $X(C) \neq X(C') = X(C^j_i)$ we get that $C^j_i \subset C \subset C'$, which is a contradiction. □

The strategy of the proof follows the $\mathcal{C}_2$, $\mathcal{B}_2$ cases: we analyze the configuration of the existing L_i, S_i, Γ_i curves in relation to C^1_ℓ, then we take every other cusp, which is not a C^α_n, and use them construct a solution to the combinatorial smoothing problem of a linear graph, i.e., a sequence of sets A_i satisfying Definition 3.3.

This is done by observing that if some cusps C_i have $X(C_i) = X(C_{i'})$, then by "forgetting" the multiplicity of the double point the definition is already satisfied. Furthermore if there are larger (in the sense, that they correspond to vertices of the graph farther from the node) curves D_i with $X(D_i) = X(D_{i'}) \neq X(C_i)$ if we remove $X(C_i)$ from D_i the definition is satisfied. Furthermore we delete points, which are contained in every curve, and identify points which we know must be either contained or missed together.

6.2.1. *Graphs in $\mathcal{C}$.*

Theorem 6.15. *In $\mathcal{C}^4$, only the graph of Figure 16 can admit a ℚHD smoothing.*

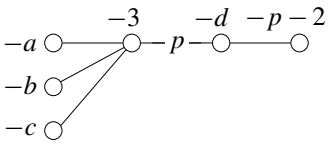


Figure 16. The valency-4 graphs, which admit a $\mathbb{Q}$HD smoothing, where $[a, b, c; d] \in \{[3, 3, 3; 4], [2, 4, 4; 3], [2, 3, 6; 2]\}$ and $p \geq 0$.

Proof. We have five Γ_i exceptional cusps and the C_i^j on the fourth arm. If there is no C_ℓ^1, we are done; otherwise, there are $1+\ell$ many C_n^α curves. Counting points, we see that C_ℓ^1 has three points (besides the double points of the C_n^α): its double point p_d, and two simple points p_{s_i}. The case $(\Gamma_i, C_\ell^1) = 6$ can only happen if every Γ_i has its double point at p_d and contains both p_{s_i} by Lemma 6.11.

A simple checking of cases shows that all of the Γ_i have to go through a further shared point (denoted by p_Γ), and each must have a final point not on any other Γ_j (denoted by p_i). Every C_i^j curve which is not C_ℓ^1 or C_n^α will have to contain either p_Γ or the p_i to intersect the Γ_i 6 times (but not both). Note, that a "switch" lemma, similar to Lemma 6.14 applies to these two possibilities as well, if some curves contain, say p_Γ, and the next curve contains the p_i, then every subsequent curve has to contain the p_i as well, since the curves C_i^j have to contain the previous ones' points except for at most one.

Every further curve C_i^j with $i < n$ has to have its double point at p_d or one of the p_{s_i}. By Lemmata 6.12–6.14 we have to consider the cases where some curves after C_ℓ^1 have double points at p_d or p_{s_1} and possibly some later curves have their double points at p_{s_1} in the former, or p_{s_2} in the latter case.

We construct a new system of curves: Take the points of $X := C_\ell^1 \setminus \{A\}$, where A denotes the double point of the first curve after C_ℓ^1. Take every curve with index $(i, j) \neq (\ell, 1)$ or (n, α). Our new curves will be $C_i^j \setminus X$ if the double point of this curve was at A, or $C_i^j \setminus \{X \cup \{A\}\}$ if its double point was not at A, with the modification that the p_i are combined into a single point P, and finally with an added new curve $S_0 = \{p_\Gamma, P\}$. These curves give a combinatorial solution to the smoothing problem of a linear graph, and thus in particular it has at least as many points as curves. From this we deduce that $k+1 \leq \tilde{k}+3$, where k denotes the number of curves in the original configuration besides C_ℓ^1, C_n^α, Γ_i, and $\tilde{k}$ denotes the number of points in the original configuration besides the points of C_ℓ^1, p_Γ, $p_1, \ldots, p_5$.

Suppose there existed a $\mathbb{Q}$HD smoothing. Then we would have

$$1 + \ell + 5 + 1 + k = 1 + \ell + 3 + 1 + 5 + \tilde{k},$$

which is equivalent to

$$k = \tilde{k} + 3.$$

Thus $\tilde{k} + 4 \leq \tilde{k} + 3$, which is a contradiction. □

6.2.2. *Graphs in $\mathcal{B}$.*

Theorem 6.16. *In $\mathcal{B}^4$, only the known case admits a* ℚHD *smoothing.*

Proof. We have L_1, S_1, S_2, Γ_1 besides the C_i^j. If every blowup happens next to the leaf of the fourth arm, we are in the known case. Otherwise there are $1+\ell$ many C_n^α curves, and C_ℓ^1 has three points (which are not double points of the C_n^α curves), denoted by p_d, p_{s_1}, p_{s_2} as before. Besides these three points, Γ_1 has two further points p_Γ, p_1. The curve L_1 contains two points. These can be either p_d and a new point p_L, or p_{s_1}, p_{s_2}. In this latter case, there can be no further C_i^j, since either $X(C_i^j) = p_d$ which forces $(C_i^j, L_1) = 1$, or $X(C_i^j) = p_{s_1}$, which implies $((C_i^j), L_1) = 3$. This means that this case cannot occur since we would have fewer curves than points. The S_i have to go through p_d to intersect C_ℓ^1 three times. From here, either they both go through p_{s_1}, and each has one extra point (p_2 and p_3, respectively), or they split the simple points ($p_{s_1} \in S_1$ and $p_{s_2} \in S_2$) and meet at a further point p_S.

In the latter case, no further C_i^j can exist. This is because $X(C_i^j) = p_d$ would be forced, meaning $p_{s_1} \in C_i^j$ and $p_{s_2} \notin C_i^j$; thus $(C_i^j, S_2) = 2$, which is a contradiction. Thus this case cannot happen, since there are at least two curves on the long arm, which are not on the end by the construction of the $\mathcal{B}$ class.

If both S_i go through p_{s_1} then by Lemma 6.14 the sequence of double points for the subsequent C_i^j curves must follow one of the following patterns:

(1) p_{s_1} then p_{s_2},

(2) p_d then p_{s_1},

(3) p_d then p_{s_2},

(4) entirely at p_{s_2}.

In the first three patterns the "switch" may or may not occur.

For all of these cases we derive a combinatorial smoothing for a linear graph using the same method as previously. Let A denote $X(C_i^j)$ where C_i^j is the next curve after C_ℓ^1. Let $X := (C_\ell^1 \cup L_1 \cup S_1 \cup S_2) \setminus \{A\}$. If $X(C) = A$, then consider $C \setminus X$, otherwise we take a curve to be $C \setminus (X \cup \{A\})$. Finally add a curve $\{p_\Gamma, p_1\}$. The validity of this configuration shows that $k+1 \leq \tilde{k}+3$. On the other hand we can calculate

$$1+\ell+5+k = 1+\ell+8+\tilde{k} \iff k = \tilde{k}+3,$$

which is a contradiction. □

6.2.3. *Graphs in $\mathcal{A}$.*

Theorem 6.17. *In $\mathcal{A}^4$, only the known case admits a* ℚHD *smoothing.*

Proof. We have Γ_1, Γ_2, S_1 and the cusps. The known case is when every blowup happens towards the leaf. Otherwise, $1+\ell$ is the number of C_n^α curves. Besides the double points of the C_n^α curves, C_ℓ^1 has its own double point p_d, and two simple points p_{s_1} and p_{s_2}. The curves Γ_1 and Γ_2 have their double points at p_d to intersect C_ℓ^1 enough times. They intersect at one additional point p_Γ, and each has a unique simple point respectively denoted by p_1, p_2. The curve S_1 contains p_d, p_{s_1} and a further point p_S.

The proof is essentially the same as the previous two cases. The double points of subsequent curves are restricted to at most two of the three possible points p_d, p_{s_1}, p_{s_2}. Let A denote the double point of the first curve. We create a solution to the combinatorial smoothing problem of a linear graph by identifying the points p_1, p_2 into a single point P, taking $X = C_\ell^1 \setminus \{A\}$ and considering the curves $C_i^j \setminus X$ if $X(C_i^j) = A$ and $C_i^j \setminus (X \cup \{A\})$ if $X(C_i^j) \neq A$, and finally adding $S_0 = \{p_\Gamma, P\}$ to the beginning of the configuration. This means that $k+1 \leq \tilde{k}+3$. Assuming the existence of a $\mathbb{Q}$HD smoothing, we compute

$$1+\ell+4+k = 1+\ell+7+\tilde{k} \iff k = \tilde{k}+3. \qquad \square$$

Remark 6.18. The known families already ruled out in, for instance, [Gay and Stipsicz 2012, Proposition 4.2] can also be ruled out from having a $\mathbb{Q}$HD smoothing using this method. Similarly to Section 5 one can find many graphs with a fixed sandwiched presentation admitting a combinatorial $\mathbb{Q}$HD smoothing. Unlike the minimal rational case, this is even possible inside the $\mathcal{A}, \mathcal{B}, \mathcal{C}$ families.

More generally, there are graphs which are rational, and have more than one -2 node in the $\mathcal{A}, \mathcal{B}, \mathcal{C}$ families, that don't have a sandwich presentation using only smooth and cusp curves. In fact there are rational graphs in $\mathcal{A}, \mathcal{B}, \mathcal{C}$ that are not sandwiched at all.

Acknowledgements

The author is thankful to András Némethi and András Stipsicz for continued help and support with this project. The author is thankful for enlightening conversations to Marco Golla, László Koltai and Olga Plamenevskaya.

The project supported by the Doctoral Excellence Fellowship Programme (DCEP) is funded by the National Research Development and Innovation Fund of the Ministry of Culture and Innovation and the Budapest University of Technology and Economics and by ERC Advanced grant KnotSurf4d.

References

[Bhupal and Stipsicz 2011] M. Bhupal and A. I. Stipsicz, "Weighted homogeneous singularities and rational homology disk smoothings", *Amer. J. Math.* **133**:5 (2011), 1259–1297. MR

[Bhupal and Stipsicz 2013] M. Bhupal and A. I. Stipsicz, "Smoothings of singularities and symplectic topology", pp. 57–97 in *Deformations of surface singularities*, edited by A. Némethi and A. Szilárd, Bolyai Soc. Math. Stud. **23**, János Bolyai Math. Soc., Budapest, 2013. MR

[Fintushel and Stern 1998] R. Fintushel and R. J. Stern, "Knots, links, and 4-manifolds", *Invent. Math.* **134**:2 (1998), 363–400. MR

[Gay and Stipsicz 2012] D. T. Gay and A. I. Stipsicz, "On symplectic caps", pp. 199–212 in *Perspectives in analysis*, *geometry*, *and topology*, edited by I. Itenberg et al., Progr. Math. **296**, Springer, 2012. MR

[Greuel and Steenbrink 1983] G.-M. Greuel and J. Steenbrink, "On the topology of smoothable singularities", pp. 535–545 in *Singularities*, *I* (Arcata, CA, 1981), edited by P. Orlik, Proc. Sympos. Pure Math. **40**, Amer. Math. Soc., Providence, RI, 1983. MR

[de Jong and van Straten 1998] T. de Jong and D. van Straten, "Deformation theory of sandwiched singularities", *Duke Math. J.* **95**:3 (1998), 451–522. MR

[Kollár 1991] J. Kollár, "Flips, flops, minimal models, etc.", pp. 113–199 in *Surveys in differential geometry* (Cambridge, MA, 1990), Lehigh Univ., Bethlehem, PA, 1991. MR

[Looijenga and Wahl 1986] E. Looijenga and J. Wahl, "Quadratic functions and smoothing surface singularities", *Topology* **25**:3 (1986), 261–291. MR

[Milnor and Husemoller 1973] J. Milnor and D. Husemoller, *Symmetric bilinear forms*, Ergebnisse der Math. **73**, Springer, 1973. MR

[Némethi 2022] A. Némethi, *Normal surface singularities*, Ergebnisse der Math. (3) **74**, Springer, 2022. MR

[Plamenevskaya and Starkston 2023] O. Plamenevskaya and L. Starkston, "Unexpected Stein fillings, rational surface singularities and plane curve arrangements", *Geom. Topol.* **27**:3 (2023), 1083–1202. MR

[Plamenevskaya and Starkston 2025] O. Plamenevskaya and L. Starkston, "Sandwiched singularities and nearly Lefschetz fibrations", 2025. arXiv 2507.21293

[Spivakovsky 1990] M. Spivakovsky, "Sandwiched singularities and desingularization of surfaces by normalized Nash transformations", *Ann. of Math.* (2) **131**:3 (1990), 411–491. MR

[Stipsicz et al. 2008] A. I. Stipsicz, Z. Szabó, and J. Wahl, "Rational blowdowns and smoothings of surface singularities", *J. Topol.* **1**:2 (2008), 477–517. MR

[Wahl 1981] J. Wahl, "Smoothings of normal surface singularities", *Topology* **20**:3 (1981), 219–246. MR

[Wahl 2011] J. Wahl, "On rational homology disk smoothings of valency 4 surface singularities", *Geom. Topol.* **15**:2 (2011), 1125–1156. MR

Received April 9, 2025.

MÁRTON BEKE
ALFRÉD RÉNYI INSTITUTE OF MATHEMATICS
BUDAPEST
HUNGARY

and

UNIVERSITY OF TECHNOLOGY AND ECONOMICS
BUDAPEST
HUNGARY
bekem@renyi.hu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 344, No. 1, 2026
https://doi.org/10.2140/pjm.2026.344.27

IWAHORI–HECKE MODEL FOR THE UNIVERSAL SUPERSINGULAR REPRESENTATION

ANAND CHITRAO AND ASFAK SONEJI

Let F be a nonarchimedean local field with residue field $\mathbb{F}_q$. Every supersingular representation of $\mathrm{GL}_2(F)$ is a quotient of a universal supersingular representation π_r for some integer $0 \le r \le q-1$. When F is a finite extension of $\mathbb{Q}_p$, Anandavardhanan–Borisagar and Anandavardhanan–Jana introduced an Iwahori–Hecke model for π_r in the *regular* case $0 < r < q-1$. When $F = \mathbb{Q}_p$, Chitrao introduced an Iwahori–Hecke model for π_r when $r = 0$, $p-1$. This model was then used by Chitrao–Ghate to compute the reduction mod p of two-dimensional irreducible semistable representations of $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$. We extend this model to an arbitrary local field F for $r = 0$, $q-1$.

We also write down an explicit nonsplit self-extension of π_r which has a four-dimensional space of $I(1)$-invariants when $F = \mathbb{Q}_p$ and $r = 0$, $p-1$.

1. Notation and conventions

For a prime p, let $\mathbb{F}_q = \mathbb{F}_{p^f}$ be a degree-f extension of the finite field $\mathbb{F}_p$. Let $\overline{\mathbb{F}}_p$ denote an algebraic closure of $\mathbb{F}_p$. We fix an embedding $\mathbb{F}_q \hookrightarrow \overline{\mathbb{F}}_p$. Let F be a local field with ring of integers $\mathcal{O}$, uniformizer ϖ and residue field $\mathbb{F}_q$. Let G be the group $\mathrm{GL}_2(F)$ and Z be its center. For $\mu \in \mathbb{F}_q$, the element $[\mu] \in \mathcal{O}$ denotes the Teichmüller lift of μ in the mixed characteristic case and denotes the image of μ under the natural embedding in the equal characteristic case. Let $K = \mathrm{GL}_2(\mathcal{O})$ be the maximal compact subgroup. Let I denote the Iwahori subgroup of upper triangular matrices mod ϖ, and $I(1)$ be the pro-p-Iwahori subgroup of K consisting of matrices that are upper triangular unipotent mod ϖ. Let

$$w := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \alpha := \begin{pmatrix} 1 & 0 \\ 0 & \varpi \end{pmatrix}, \quad \beta := \begin{pmatrix} 0 & 1 \\ \varpi & 0 \end{pmatrix} = \alpha w.$$

Let $\underline{r} = (r_0, \dots, r_{f-1})$ be an f-tuple of integers where $r_j \ge 0$ for $0 \le j \le f-1$. Let V_{r_j} denote the symmetric r_j-th power of the standard two-dimensional representation of $\mathrm{GL}_2(\mathbb{F}_q)$. We identify it with the homogeneous polynomials of degree r_j in variables X_j and Y_j for all $0 \le j \le f-1$. By $V_{r_j}^{\mathrm{Fr}^j}$ we mean that the action of $\mathrm{GL}_2(\mathbb{F}_q)$

MSC2020: 11F70, 20G05, 22E50.
Keywords: modular representations, Iwahori–Hecke model, supersingular representations.

© 2026 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

is twisted by the j-th power of Frobenius. We view these as representations of K, via the reduction mod ϖ map and then as representations of KZ by making the scalar matrix ϖ act trivially. For such an f-tuple $\underline{r}$, let $V_{\underline{r}}$ be the representation $V_{r_0} \otimes_{\mathbb{F}_q} V_{r_1}^{\mathrm{Fr}} \otimes_{\mathbb{F}_q} \cdots \otimes_{\mathbb{F}_q} V_{r_{f-1}}^{\mathrm{Fr}^{f-1}}$ of KZ. For $0 \le r < q-1$, let d^r be the character $I \to \mathbb{F}_q^\times$ sending $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ to $d^r \bmod \varpi$ and extend it to IZ by mapping the scalar matrix $\varpi \in Z$ to 1. If $0 < r < q-1$, we say that d^r is *regular* and if $r = 0$, we say that d^r is *nonregular*.

Denote by $\mathrm{ind}_{KZ}^G \sigma$ the compact induction of a representation σ of KZ. Following notation of [Barthel and Livné 1994], for $g \in G$ and $v \in \sigma$, we use the symbol $[g, v] \in \mathrm{ind}_{KZ}^G \sigma$ to denote the function from G to σ which is supported on the coset KZg^{-1} and which sends g^{-1} to v. To distinguish elements of different induced representations, we will denote elements of compact induction $\mathrm{ind}_{IZ}^G \chi$ by $[[g, v]]$ and those of $\mathrm{ind}_{IZ}^{KZ} \sigma$ by $[[[g, v]]]$. For an irreducible representation χ, we define the Iwahori–Hecke algebra as the endomorphism algebra $\mathrm{End}_G(\mathrm{ind}_{IZ}^G \chi)$. Similarly, for an irreducible representation σ of KZ, the spherical Hecke algebra is defined as $\mathrm{End}_G(\mathrm{ind}_{KZ}^G \sigma)$.

2. Introduction

For a local field F, the classification of irreducible smooth representations over $\overline{\mathbb{F}}_p$ admitting a central character of G began with the notable work of Barthel and Livné [1994]. In addition to many interesting results, they proved the existence of irreducible smooth representations, known as supersingular representations, which cannot be obtained as a subquotient of a parabolically induced representation.

It is proved in [Barthel and Livné 1994, Theorem 34] that up to a twist by a character, a supersingular representation can be realized as an irreducible quotient (after a possible extension of scalars) of a universal representation

$$\pi_r = \frac{\mathrm{ind}_{KZ}^G V_{\underline{r}}}{(T)}$$

for $\underline{r} = (r_0, \ldots, r_{f-1})$ with $0 \le r_j \le p-1$ for all $0 \le j \le f-1$. Here, we have $r = r_0 + r_1 p + \cdots + r_{f-1} p^{f-1}$ and T is the standard spherical Hecke operator defined in [Barthel and Livné 1994, Section 3.1]. Since π_r is defined over $\mathbb{F}_q$, we assume that all representations that we work with are defined over $\mathbb{F}_q$. Computing $I(1)$-invariants, Breuil [2003, Corollaries 4.1.1, 4.1.4] showed that π_r is irreducible when $F = \mathbb{Q}_p$. However, if $[F : \mathbb{Q}_p] > 1$, the structure of π_r is mysterious. For instance, they are of infinite length. Schein [2011, Theorem 1.3] established a criterion for irreducibility of quotients of π_r. In [Schein 2014, Proposition 1.3], he also studied the K-structure of the G-representations π_r when $F/\mathbb{Q}_p$ is totally ramified and $[F : \mathbb{Q}_p] > 1$. It can be seen in [Breuil and Paškūnas 2012] that explicit construction of a supersingular representation is a difficult task when

$F \supsetneq \mathbb{Q}_p$ is unramified. Schein [2023, Theorem 1.1] constructed a family of supersingular representations with good socle for a ramified extension F of $\mathbb{Q}_p$, with the ramification index $e \leq \frac{p-1}{2}$ using diagrams. For a finite extension F of $\mathbb{Q}_p$, Hendel [2019, Theorem 1.2] computed the space of $I(1)$-invariants of π_r, and when F is unramified over $\mathbb{Q}_p$, constructed a universal supersingular representation having all supersingular representations with a good K-socle as quotients. Jana [2023, Theorem 1.4] has extended Hendel's work concerning the $I(1)$-invariants of π_r. Using the results of this paper, we have recently computed the space of $I(1)$-invariants of π_r for $r = 0, q-1$ with Jana [Chitrao et al. 2026, Theorem 2.1].

A universal supersingular representation can also be constructed using the Iwahori–Hecke operators instead of the spherical Hecke operator T for a finite extension of $\mathbb{Q}_p$ (see [Anandavardhanan and Borisagar 2015, Theorem 1.1; Anandavardhanan and Jana 2021, Theorem 1.1]). For $0 < r < q-1$, consider the *regular* character d^r of IZ. It is shown in [Barthel and Livné 1994, Proposition 13] that the corresponding Iwahori–Hecke algebra is commutative and is given by

$$\mathrm{End}_G(\mathrm{ind}_{IZ}^G d^r) \cong \frac{\mathbb{F}_q[T_{-1,0}, T_{1,2}]}{(T_{-1,0}T_{1,2}, T_{1,2}T_{-1,0})}.$$

Here, $T_{-1,0}$ and $T_{1,2}$ are Iwahori–Hecke operators as defined in [Barthel and Livné 1994, (17) and (16)]. The precise formulas for these operators are given in [Anandavardhanan and Borisagar 2015, (9) and (10)] for $F = \mathbb{Q}_p$ and in [Anandavardhanan and Jana 2021, (3-2) and (3-3)] for an arbitrary $F/\mathbb{Q}_p$. When F is a totally ramified extension of $\mathbb{Q}_p$, by [Anandavardhanan and Borisagar 2015, Remark 1], we have

$$\text{(2-1)} \qquad \mathrm{Im}\, T_{-1,0} = \mathrm{Ker}\, T_{1,2} \quad \text{and} \quad \mathrm{Im}\, T_{1,2} = \mathrm{Ker}\, T_{-1,0}.$$

When F is not totally ramified, there are strict inclusions

$$\text{(2-2)} \qquad \mathrm{Im}\, T_{-1,0} \subsetneq \mathrm{Ker}\, T_{1,2} \quad \text{and} \quad \mathrm{Im}\, T_{1,2} \subsetneq \mathrm{Ker}\, T_{-1,0}$$

as shown in [Anandavardhanan and Jana 2021, Remark 5]. When F is a finite extension of $\mathbb{Q}_p$, the comparison between Iwahori induction and spherical induction is given in [Anandavardhanan and Jana 2021, Theorem 1.1].

The Iwahori model for the universal supersingular representations is also useful when one wants to compute the invariants in π_r under the action of principal congruence subgroups. This is because using (2-1) one can write π_r as a quotient of an induced representation by kernels of certain Iwahori–Hecke operators. This idea is worked out in [Anandavardhanan and Borisagar 2013, Theorem 1.2] to give an easier proof of Morra's result [2013, Theorem 1.4] on the K-structure of the invariants of π_r under the principal congruence subgroups when $r \neq 0, p-1$.

For the *nonregular* case $r \in \{0, q-1\}$, the Iwahori–Hecke algebra is noncommutative and is generated by Iwahori–Hecke operators $T_{1,0}$ and $T_{1,2}$ satisfying

$$T_{1,0}^2 = \mathrm{Id} \quad \text{and} \quad T_{1,2}T_{1,0}T_{1,2} = -T_{1,2} \tag{2-3}$$

as given in [Barthel and Livné 1994, Proposition 11]. The precise formulas for these operators are given in Section 3. As far as the remaining cases are concerned, Chitrao [2025, Theorem 3.13] provides an explicit Iwahori–Hecke model for π_r in the *nonregular* case, though exclusively for $\mathbb{Q}_p$, i.e., for $r = 0$ and $p-1$. We extend this model to include arbitrary local fields F; see Theorem 4.4. We apply the Iwahori–Hecke model for $\mathrm{GL}_2(\mathbb{Q}_p)$ to construct self-extensions of supersingular representations, which have four-dimensional $I(1)$-invariants in the *nonregular* case, as is outlined in Theorems 5.5 and 5.6. This explicit construction partially generalizes the results of [Anandavardhanan and Borisagar 2015, Theorem 1.2] for the *regular* case. Furthermore, we establish the correct analogues of the inclusions in (2-1) and (2-2) for the *nonregular* case over an arbitrary local field F. Specifically, in Lemmas 5.7 and 5.8, we show that these inclusions get replaced by

$$\begin{aligned} &\mathrm{Im}(T_{1,2}+T_{1,0}) = \mathrm{Ker}(T_{-1,0}), \quad \mathrm{Im}(T_{-1,0}) = \mathrm{Ker}(T_{1,2}+T_{1,0}),\\ &\mathrm{Im}(T_{1,2}) = \mathrm{Ker}(T_{1,0}+T_{-1,0}), \quad \mathrm{Im}(T_{1,0}+T_{-1,0}) = \mathrm{Ker}(T_{1,2}). \end{aligned}$$

This highlights a striking structural dissimilarity between the *regular* and *nonregular* cases.

One of the motivations behind this paper was to write down an Iwahori-theoretic modular local Langlands correspondence between $\mathrm{Gal}(\bar{F}/F)$ and $\mathrm{GL}_2(F)$ as was done in [Chitrao 2025] for $F = \mathbb{Q}_p$. In [Chitrao 2025, Theorem 2.2], Breuil's correspondence was rewritten entirely in the language of Iwahori induction. For a parameter $r \in \{0, \dots, p-1\}$ and a smooth character $\eta : \mathbb{Q}_p^\times \to \bar{\mathbb{F}}_p^\times$, the correspondence in the supersingular case (where the Hecke eigenvalue is $\lambda = 0$) takes the form

$$(\mathrm{ind}\,\omega_2^{r+1}) \otimes \eta \iff \frac{\mathrm{ind}_{IZ}^G d^r}{(T_{-1,0} + \delta_{r,p-1}T_{1,0}) + (T_{1,2} + \delta_{r,0}T_{1,0})} \otimes \eta.$$

Here, ω_2 is the fundamental character of level 2, and δ denotes the Kronecker delta. The correspondence in the language of induction from the Iwahori subgroup (times the center) is especially useful while computing the reduction of semi-stable representations of $\mathrm{Gal}(\bar{F}/F)$ when $F = \mathbb{Q}_p$. It turns out that the automorphic representation associated by the p-adic local Langlands correspondence with a two-dimensional, irreducible, semi-stable representation of the absolute Galois group of $\mathbb{Q}_p$ has a dense subrepresentation isomorphic to a quotient of a representation induced from the Iwahori subgroup to $\mathrm{GL}_2(\mathbb{Q}_p)$. Therefore, when computing the reduction mod p of the automorphic representation, one must work with Iwahori

induction. Using this idea, it is possible to compute the reduction mod p of two-dimensional, irreducible, semi-stable representations of the absolute Galois group of $\mathbb{Q}_p$ with *arbitrarily large* Hodge–Tate weights. This has been done in [Chitrao and Ghate 2023, Theorem 1.1] for weights at most $p+1$, including the previously untreated weights p and $p+1$, and is further explored in [Chitrao and Ghate 2026] for weights at most p^2. Similarly, the results here might be useful in computing the reduction mod p of representations of $\mathrm{Gal}(\overline{F}/F)$ for arbitrary $F/\mathbb{Q}_p$. The progress in this direction, however, is heavily dependent on a (yet to be discovered) mod p local Langlands correspondence for representations of $\mathrm{Gal}(\overline{F}/F)$.

3. Preliminaries

We state Lucas' classical result in modular combinatorics, which computes the residue of the binomial coefficient $\binom{n}{i}$ modulo p.

Theorem 3.1 (Lucas). *Let $n, i \in \mathbb{N}$ be such that $n = \sum_{j=0}^{k} n_j p^j$ and $i = \sum_{j=0}^{k} i_j p^j$, where $0 \le n_j \le p-1$ and $0 \le i_j \le p-1$. Then*

$$\binom{n}{i} \equiv \prod_{j=0}^{k} \binom{n_j}{i_j} \mod p.$$

Corollary 3.2. *Let $n, i \in \mathbb{N}$. Then $p \mid \binom{n}{i}$ if and only if $n_j < i_j$ for some $0 \le j \le k$.*

As mentioned in the introduction, the Iwahori–Hecke algebra is noncommutative in the *nonregular* case. It is generated by the Iwahori–Hecke operators $T_{1,0}$ and $T_{1,2}$, which satisfy (2-3). We write down the explicit formulas for these operators as given in [Barthel and Livné 1995, Lemma 6]:

$$T_{1,0}[[g,1]] = [[g\beta,1]] \text{ and } T_{1,2}[[g,1]] = \sum_{\lambda\in\mathbb{F}_q} \left[\left[g \begin{pmatrix} 1 & 0 \\ \varpi[\lambda] & \varpi \end{pmatrix}, 1 \right]\right].$$

According to [Barthel and Livné 1995, Lemma 8(4)], the operator $T_{-1,0}$ satisfies

$$T_{-1,0} = T_{1,0}T_{1,2}T_{1,0}. \tag{3-1}$$

Its formula is

$$T_{-1,0}[[g,1]] = \sum_{\lambda\in\mathbb{F}_q} \left[\left[g \begin{pmatrix} \varpi & [\lambda] \\ 0 & 1 \end{pmatrix}, 1 \right]\right].$$

The spherical Hecke operator T is determined by

$$T[g,1] = \sum_{\lambda\in\mathbb{F}_q} \left[g\begin{pmatrix} \varpi & [\lambda] \\ 0 & 1 \end{pmatrix}, 1\right] + \left[g\begin{pmatrix} 1 & 0 \\ 0 & \varpi \end{pmatrix}, 1\right] \quad \text{if } r = 0,$$

$$T[g, X_0^{p-1}\cdots X_{f-1}^{p-1}] = \sum_{\lambda\in\mathbb{F}_q} \left[g\begin{pmatrix} \varpi & [\lambda] \\ 0 & 1 \end{pmatrix}, X_0^{p-1}\cdots X_{f-1}^{p-1}\right] \qquad \text{if } r = q-1.$$

4. Comparison between Iwahori induction and spherical induction

In this section, we prove the main comparison theorem between Iwahori induction and spherical induction.

The following lemma extends a case of [Ghate and Vangala 2022, Lemma 2.15]. This lemma also appears in [Ghate and Pandey 2026].

Lemma 4.1. *For $0 \le j \le f-1$, fix integers $r_j \ge 0$ such that*

$$0 < r = r_0 + r_1 p + \cdots + r_{f-1} p^{f-1} \quad \text{is divisible by } q-1.$$

Then for any integer $i = i_0 + i_1 p + \cdots + i_{f-1} p^{f-1} \in [0, q-1)$, let K_i be the set of f-tuples $(k_0, \ldots, k_{f-1})$ of integers such that $0 \le k_j \le r_j$ for all $0 \le j \le f-1$ and $k_0 + k_1 p + \cdots + k_{f-1} p^{f-1} \equiv i \bmod q-1$. Then we have

$$\sum_{(k_0,\ldots,k_{f-1}) \in K_i} \binom{r_0}{k_0} \cdots \binom{r_{f-1}}{k_{f-1}} = \begin{cases} (-1)^i \bmod p & \text{if } i \in (0, q-1), \\ 2 \bmod p & \text{if } i = 0. \end{cases}$$

Proof. Consider the polynomial

$$f(x) = (1+x)^{r_0} (1+x^p)^{r_1} \cdots (1+x^{p^{f-1}})^{r_{f-1}}$$

in the ring $R = \mathbb{F}_p[x]/(x^{q-1}-1)$. For $i \in [0, q-1)$, the coefficient of x^i in $f(x)$ is given by the sum, over the set K_i,

$$\sum_{(k_0,\ldots,k_{f-1}) \in K_i} \binom{r_0}{k_0} \cdots \binom{r_{f-1}}{k_{f-1}}.$$

However, $f(x) = (1+x)^r$. Let $r = s_0 + s_1 p^f + \cdots + s_n p^{nf}$ be the expansion of r in base p^f. Using the identity $(1+x)^q = (1+x)$ in R, we see that $f(x) = (1+x)^s$, where $s = s_0 + s_1 + \cdots + s_n$. If $r \ge q$, then $s < r$ and $s \equiv r \bmod q-1$. Iterating this process, we may assume that $s \in (0, q-1]$. Since $s \equiv r \equiv 0 \bmod q-1$, we see that $s = q-1$. Therefore $f(x) = (1+x)^{q-1}$ in R.

If $i \in (0, q-1)$, the coefficient of x^i in $f(x)$ is

$$\binom{(p-1)p^{f-1} + (p-1)p^{f-2} + \cdots + (p-1)}{i_{f-1} p^{f-1} + i_{f-2} p^{f-2} + \cdots + i_0} = \prod_{j=0}^{f-1} \binom{p-1}{i_j} = (-1)^i$$

in $\mathbb{F}_p$. The last equality follows from Theorem 3.1 and the fact $\binom{p-1}{l} \equiv (-1)^l \bmod p$ for $0 \le l \le p-1$. If $i = 0$, however, the relation $x^{q-1} - 1$ makes the leading term constant. So, the coefficient of x^i receives contributions from the constant term and the leading term, which proves the lemma. □

For $j \in \mathbb{Z}/f\mathbb{Z}$, let $\theta_j = X_j Y_{j-1}^p - Y_j X_{j-1}^p$ denote the twisted Dickson polynomials introduced in [Ghate and Jana 2025]. Using this notation, we state the following theorem, which is the $m = 0$ case of [Ghate and Jana 2025, Theorem 2.18].

Theorem 4.2. *Let $r = r_0 + r_1 p + \cdots + r_{f-1} p^{f-1}$ be such that $r_j \geq q = p^f$ for all $0 \leq j \leq f-1$, and r is divisible by $q-1$. The map*

$$\psi : \bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j} \to \mathrm{ind}_{B(\mathbb{F}_q)}^{\mathrm{GL}_2(\mathbb{F}_q)} \mathbb{1}$$

that sends a polynomial $P = P_0(X_0, Y_0) \otimes \cdots \otimes P_{f-1}(X_{f-1}, Y_{f-1}) \in \bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}$ to the function $\psi_P : \mathrm{GL}_2(\mathbb{F}_q) \to \mathbb{F}_q$ defined by

$$\psi_P\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = P_0(c, d) \otimes \cdots \otimes P_{f-1}(c^{p^{f-1}}, d^{p^{f-1}})$$

induces an isomorphism of $\mathrm{GL}_2(\mathbb{F}_q)$-representations

$$\frac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \ldots, \theta_{f-1} \rangle} \xrightarrow{\sim} \mathrm{ind}_{B(\mathbb{F}_q)}^{\mathrm{GL}_2(\mathbb{F}_q)} \mathbb{1}.$$

For brevity, let us define some notation. For $\underline{r} = (r_0, \ldots, r_{f-1})$ with $r_j \geq 0$ for all $0 \leq j \leq f-1$, we write $\underline{X}^{\underline{r}}$ and $\underline{Y}^{\underline{r}}$ to represent $X_0^{r_0} \cdots X_{f-1}^{r_{f-1}}$ and $Y_0^{r_0} \cdots Y_{f-1}^{r_{f-1}}$, respectively. We also define $\underline{X}^{\underline{p-1}} := X_0^{p-1} \cdots X_{f-1}^{p-1}$, $\underline{Y}^{\underline{p-1}} := Y_0^{p-1} \cdots Y_{f-1}^{p-1}$ and $\underline{X}^{\underline{r-p+1}} \underline{Y}^{\underline{p-1}} := X_0^{r_0-(p-1)} Y_0^{p-1} \cdots X_{f-1}^{r_{f-1}-(p-1)} Y_{f-1}^{p-1}$.

Proposition 4.3. *Let $f \geq 1$. Let $r = r_0 + r_1 p + \cdots + r_{f-1} p^{f-1}$ be such that $r_j \geq q = p^f$ for all $0 \leq j \leq f-1$. Assume r is divisible by $q-1$. Then, the maps*

$$\iota_0 : \bigotimes_{j=0}^{f-1} V_0^{\mathrm{Fr}^j} \to \frac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \ldots, \theta_{f-1} \rangle}, \quad 1 \mapsto \underline{X}^{\underline{r}} - \underline{X}^{\underline{r-p+1}} \underline{Y}^{\underline{p-1}} + \underline{Y}^{\underline{r}},$$

and

$$\iota_{p-1} : \bigotimes_{j=0}^{f-1} V_{p-1}^{\mathrm{Fr}^j} \to \frac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \ldots, \theta_{f-1} \rangle},$$

$$\underline{X}^{\underline{p-1-i}} \underline{Y}^{\underline{i}} \mapsto \underline{X}^{\underline{r-i}} \underline{Y}^{\underline{i}} \quad \text{if } (i_0, \ldots, i_{f-1}) \neq (p-1, \ldots, p-1),$$

$$\underline{Y}^{\underline{p-1}} \mapsto \underline{Y}^{\underline{r}} \quad \text{if } (i_0, \ldots, i_{f-1}) = (p-1, \ldots, p-1),$$

induce an isomorphism

$$\bigotimes_{j=0}^{f-1} V_0^{\mathrm{Fr}^j} \oplus \bigotimes_{j=0}^{f-1} V_{p-1}^{\mathrm{Fr}^j} \simeq \frac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \ldots, \theta_{f-1} \rangle}.$$

Proof. By Theorem 4.2, we have that

$$\frac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \ldots, \theta_{f-1} \rangle} \simeq \mathrm{ind}_{B(\mathbb{F}_q)}^{\mathrm{GL}_2(\mathbb{F}_q)} \mathbb{1}.$$

There is a copy of $\bigotimes_{j=0}^{f-1} V_0^{\mathrm{Fr}^j}$ in $\mathrm{ind}_{B(\mathbb{F}_q)}^{\mathrm{GL}_2(\mathbb{F}_q)} \mathbb{1}$ given by the constant functions. A brief calculation shows that under the isomorphism presented above, the polynomial $\underline{X}^{\underline{r}} - \underline{X}^{\underline{r-p+1}}\underline{Y}^{\underline{p-1}} + \underline{Y}^{\underline{r}}$ maps to the constant function 1.

Next, we prove that the map

$$\bigotimes_{j=0}^{f-1} V_{p-1}^{\mathrm{Fr}^j} \to \frac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \dots, \theta_{f-1} \rangle},$$

$$\underline{X}^{\underline{p-1-i}}\underline{Y}^{\underline{i}} \mapsto \underline{X}^{\underline{r-i}}\underline{Y}^{\underline{i}} \quad \text{if } (i_0, \dots, i_{f-1}) \neq (p-1, \dots, p-1),$$

$$\underline{Y}^{\underline{p-1}} \mapsto \underline{Y}^{\underline{r}} \quad \text{if } (i_0, \dots, i_{f-1}) = (p-1, \dots, p-1),$$

is $\mathrm{GL}_2(\mathbb{F}_q)$-equivariant.

Since $\bigotimes_{j=0}^{f-1} V_{p-1}^{\mathrm{Fr}^j}$ is irreducible, it is generated by any nonzero vector. In particular, it is generated by $\underline{X}^{\underline{p-1}}$. Therefore to show ι_{p-1} is $\mathrm{GL}_2(\mathbb{F}_q)$-equivariant, it is enough to show

$$\iota_{p-1}(g\underline{X}^{\underline{p-1}}) = g\iota_{p-1}(\underline{X}^{\underline{p-1}}) \quad \text{for all } g \in \mathrm{GL}_2(\mathbb{F}_q).$$

Noting the Bruhat decomposition $\mathrm{GL}_2(\mathbb{F}_q) = B \sqcup BwB$, it is enough to show

$$\iota_{p-1}(M\underline{X}^{\underline{p-1}}) = M\iota_{p-1}(\underline{X}^{\underline{p-1}}), \quad \iota_{p-1}(M\underline{Y}^{\underline{p-1}}) = M\iota_{p-1}(\underline{Y}^{\underline{p-1}})$$

for all $M \in B$. The first identity is easy to check since B stabilizes $\underline{X}^{\underline{p-1}}$ and $\underline{X}^{\underline{r}}$. We now verify the second identity. Let $M = \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$. Then the left-hand side of the second equation above is

$$\iota_{p-1}[(bX_0 + dY_0)^{p-1} \cdots (b^{p^{f-1}} X_{f-1} + d^{p^{f-1}} Y_{f-1})^{p-1}]$$
$$= \iota_{p-1}\left[\sum_{(i_0,\dots,i_{f-1})} (-1)^i b^{q-1-i} d^i X_0^{p-1-i_0} Y_0^{i_0} \cdots X_{f-1}^{p-1-i_{f-1}} Y_{f-1}^{i_{f-1}} \right],$$

where the sum runs over all tuples $(i_0, \dots, i_{f-1})$ for which $0 \le i_j \le p-1$ for all $0 \le j \le f-1$. Separating the $(i_0, \dots, i_{f-1}) = (p-1, \dots, p-1)$ term, the last expression becomes

$$\left[\sum_{(i_0,\dots,i_{f-1}) \neq (p-1,\dots,p-1)} (-1)^i b^{q-1-i} d^i X_0^{r_0-i_0} Y_0^{i_0} \cdots X_{f-1}^{r_{f-1}-i_{f-1}} Y_{f-1}^{i_{f-1}} \right] + Y_0^{r_0} \cdots Y_{f-1}^{r_{f-1}}.$$

Next, we compute the right-hand side as

$$M\iota_{p-1}(\underline{Y}^{\underline{p-1}}) = M(\underline{Y}^{\underline{r}})$$
$$= (bX_0 + dY_0)^{r_0} \cdots (b^{p^{f-1}} X_{f-1} + d^{p^{f-1}} Y_{f-1})^{r_{f-1}}$$
$$= \sum_{(k_0,\dots,k_{f-1})} \binom{r_0}{k_0} \cdots \binom{r_{f-1}}{k_{f-1}} b^{r-k} d^k X_0^{r_0-k_0} Y_0^{k_0} \cdots X_{f-1}^{r_{f-1}-k_{f-1}} Y_{f-1}^{k_{f-1}},$$

where the sum runs over all tuples $(k_0, \dots, k_{f-1})$ for which $0 \le k_j \le r_j$ for all $0 \le j \le f-1$. Note that the coefficients of $\underline{X}^r$ at the end of the two displays above match, and so do the coefficients of $\underline{Y}^r$. To show that the remaining expressions are equal, it is enough to check

$$\sum_{(k_0,\dots,k_{f-1})\in K_i} \binom{r_0}{k_0} \cdots \binom{r_{f-1}}{k_{f-1}} \equiv (-1)^i \mod p \qquad \text{for } i \in (0,\ p^{f-1}-1),$$

$$\sum_{\substack{(k_0,\dots,k_{f-1})\in K_i \\ (k_0,\dots,k_{f-1})\neq(0,\dots,0) \\ (k_0,\dots,k_{f-1})\neq(r_0,\dots,r_{f-1})}} \binom{r_0}{k_0} \cdots \binom{r_{f-1}}{k_{f-1}} \equiv 0 \mod p \qquad \text{for } i = 0.$$

Here, the indexing set K_i is defined as in Lemma 4.1, which also provides the proof for these congruences. Finally, a dimension count shows

$$\bigotimes_{j=0}^{f-1} V_0^{\mathrm{Fr}^j} \oplus \bigotimes_{j=0}^{f-1} V_{p-1}^{\mathrm{Fr}^j} \simeq \frac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \dots, \theta_{f-1} \rangle}. \qquad \square$$

The following theorem generalizes [Chitrao 2025, Theorem 3.13], which was the special case $F = \mathbb{Q}_p$.

Theorem 4.4. *Let $f \ge 1$. Let $r = r_0 + r_1 p + \cdots + r_{f-1} p^{f-1}$ be such that $r_j \ge q = p^f$ for all $0 \le j \le f-1$. Assume r is divisible by $q-1$. The map*

$$\Theta : \mathrm{ind}_{IZ}^G \mathbb{1} \to \mathrm{ind}_{KZ}^G \frac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \dots, \theta_{f-1} \rangle}, \quad [[\mathrm{id}, 1]] \mapsto [\mathrm{id}, \underline{Y}^r - \underline{X}^{\underline{r-p+1}} \underline{Y}^{\underline{p-1}}],$$

induces isomorphisms

$$\frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{1,2}T_{1,0})} \simeq \mathrm{ind}_{KZ}^G \bigotimes_{j=0}^{f-1} V_0^{\mathrm{Fr}^j} \quad \textit{and} \quad \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(\mathrm{Id} + T_{1,2}T_{1,0})} \simeq \mathrm{ind}_{KZ}^G \bigotimes_{j=0}^{f-1} V_{p-1}^{\mathrm{Fr}^j}.$$

Moreover, the operators $T_{-1,0} + T_{1,0}$, $T_{-1,0}$ on the left correspond to the operator T on the right under the isomorphism above. Note that $T \in \mathrm{End}_G\big(\mathrm{ind}_{IZ}^G \bigotimes_{j=0}^{f-1} V_0^{\mathrm{Fr}^j}\big)$ in the first case is different from $T \in \mathrm{End}_G\big(\mathrm{ind}_{IZ}^G \bigotimes_{j=0}^{f-1} V_{p-1}^{\mathrm{Fr}^j}\big)$ in the second case.

Proof. We first prove that the map

$$\Theta : \mathrm{ind}_{IZ}^G \mathbb{1} \to \mathrm{ind}_{KZ}^G \frac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \dots, \theta_{f-1} \rangle}, \quad [[\mathrm{id}, 1]] \mapsto [\mathrm{id}, \underline{Y}^r - \underline{X}^{\underline{r-p+1}} \underline{Y}^{\underline{p-1}}],$$

is an isomorphism. Since $IZ \subseteq KZ \subseteq G$ with $IZ \subseteq KZ$ of finite index, there is an isomorphism from $\mathrm{ind}_{IZ}^G \mathbb{1}$ to $\mathrm{ind}_{KZ}^G \mathrm{ind}_{IZ}^{KZ} \mathbb{1}$ defined by

$$[[\mathrm{id}, 1]] \mapsto [\mathrm{id}, [[[\mathrm{id}, 1]]]].$$

A brief computation using Theorem 4.2 shows the map sending $\underline{Y}^r - \underline{X}^{\underline{r-p+1}}\underline{Y}^{\underline{p-1}}$ to [[[id, 1]]] is an isomorphism between

$$\frac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \dots, \theta_{f-1}\rangle} \quad \text{and} \quad \mathrm{ind}_{IZ}^{KZ}\, \mathbb{1}.$$

Since $-T_{1,2}T_{1,0}$ is an idempotent by (2-3) in the noncommutative Hecke algebra, we get a splitting

$$\mathrm{ind}_{IZ}^{G}\, \mathbb{1} = \mathrm{Im}(-T_{1,2}T_{1,0}) \oplus \mathrm{Im}(\mathrm{Id} + T_{1,2}T_{1,0}).$$

We also know from Proposition 4.3 that

$$\mathrm{ind}_{KZ}^{G} \frac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \dots, \theta_{f-1}\rangle} = \mathrm{ind}_{KZ}^{G} \bigotimes_{j=0}^{f-1} V_0^{\mathrm{Fr}^j} \oplus \mathrm{ind}_{KZ}^{G} \bigotimes_{j=0}^{f-1} V_{p-1}^{\mathrm{Fr}^j},$$

where the space $\mathrm{ind}_{KZ}^{G}\big(\bigotimes_{j=0}^{f-1} V_{p-1}^{\mathrm{Fr}^j}\big)$ is identified with the subspace generated by the element $[\mathrm{id}, \underline{X}^r]$ and the space $\mathrm{ind}_{KZ}^{G}\big(\bigotimes_{j=0}^{f-1} V_0^{\mathrm{Fr}^j}\big)$ is identified with subspace generated by the element $[\mathrm{id}, \underline{X}^r - \underline{X}^{\underline{r-p+1}}\underline{Y}^{\underline{p-1}} + \underline{Y}^r]$. Therefore, to prove the first part of this theorem, we need to show that

- $\Theta(\mathrm{Im}(-T_{1,2}T_{1,0})) = \mathrm{ind}_{KZ}^{G}\big(\bigotimes_{j=0}^{f-1} V_{p-1}^{\mathrm{Fr}^j}\big)$,
- $\Theta(\mathrm{Im}(\mathrm{Id} + T_{1,2}T_{1,0})) = \mathrm{ind}_{KZ}^{G}\big(\bigotimes_{j=0}^{f-1} V_0^{\mathrm{Fr}^j}\big)$.

Indeed, under the map

$$\mathrm{ind}_{IZ}^{G}\, \mathbb{1} \to \mathrm{ind}_{KZ}^{G} \frac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \dots, \theta_{f-1}\rangle},$$

the element $T_{1,2}[[\mathrm{id}, 1]]$ maps to

$$\sum_{\lambda\in\mathbb{F}_q}\left[\begin{pmatrix} 1 & 0 \\ \varpi[\lambda] & \varpi \end{pmatrix}, \underline{Y}^r - \underline{X}^{\underline{r-p+1}}\underline{Y}^{\underline{p-1}}\right] = \sum_{\lambda\in\mathbb{F}_q}\left[\beta \begin{pmatrix} 1 & [\lambda] \\ 0 & 1 \end{pmatrix} w, \underline{Y}^r - \underline{X}^{\underline{r-p+1}}\underline{Y}^{\underline{p-1}}\right].$$

Since $\left(\begin{smallmatrix} 1 & [\lambda] \\ 0 & 1 \end{smallmatrix}\right) w \in KZ$, we can move it to the other side to get

$$\sum_{\lambda\in\mathbb{F}_q}\big[\beta, X_0^{r_0}\cdots X_{f-1}^{r_{f-1}} - (\lambda X_0 + Y_0)^{r_0-(p-1)}X_0^{p-1} \cdots (\lambda^{p^{f-1}}X_{f-1} + Y_{f-1})^{r_{f-1}-(p-1)}X_{f-1}^{p-1}\big].$$

Expanding $(\lambda^{p^j}X_j + Y_j)^{r_j-(p-1)}$ for all $0 \le j \le f-1$, the above expression is

$$-\left[\beta, \sum_{(k_0,\dots,k_{f-1})} \binom{r_0-(p-1)}{k_0} \cdots \binom{r_{f-1}-(p-1)}{k_{f-1}} \sum_{\lambda\in\mathbb{F}_q^\times} \lambda^{r-(q-1)-k}\, \underline{X}^{\underline{r-k}}\underline{Y}^{\underline{k}}\right] - [\beta, \underline{Y}^{\underline{r-p+1}}\underline{X}^{\underline{p-1}}],$$

where the sum runs over all tuples $(k_0, \dots, k_{f-1})$ for which $0 \le k_j \le r_j - (p-1)$ for all $0 \le j \le f-1$ and where $k = k_0 + k_1 p + \cdots + k_{f-1} p^{f-1}$.

Using the identity

$$\sum_{\mu \in \mathbb{F}_q} \mu^l = \begin{cases} 0 & \text{if } q-1 \nmid l, \\ q-1 & \text{if } q-1 \mid l \end{cases} \qquad \text{for } l \ge 1, \tag{4-1}$$

we see that the inner sum over $\lambda \in \mathbb{F}_q^\times$ is nonzero modulo p if and only if the exponent satisfies $k \equiv r \pmod{q-1}$. Consequently, only these terms contribute to the sum. In the quotient

$$\frac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \dots, \theta_{f-1} \rangle},$$

we use the relations defined by the polynomials θ_j to rewrite the sum of the remaining terms as a linear combination of the monomials $\underline{X}^{\underline{r-k}} \underline{Y}^{\underline{k}}$ for $\underline{k}$ satisfying $0 \le k_j \le p-1$. Finally, applying Lemma 4.1 with r_j replaced by $r_j - (p-1)$ allows us to evaluate the resulting sum of binomial coefficients. Thus,

$$T_{1,2}[[\mathrm{id}, 1]] \mapsto [\beta, \underline{X}^{\underline{r}}]. \tag{4-2}$$

- Using (4-2), we see that

$$T_{1,2} T_{1,0}[[\mathrm{id}, 1]] = T_{1,2}[[\beta, 1]] \mapsto [\mathrm{id}, \underline{X}^{\underline{r}}].$$

  Moreover, $[\mathrm{id}, \underline{X}^{\underline{r}}]$ generates $\mathrm{ind}_{KZ}^G\big(\bigotimes_{j=0}^{f-1} V_{p-1}^{\mathrm{Fr}^j}\big)$ in $\mathrm{ind}_{KZ}^G \dfrac{\bigotimes_{j=0}^{f-1} V_{r_j}^{\mathrm{Fr}^j}}{\langle \theta_0, \dots, \theta_{f-1} \rangle}$.

- Again using (4-2), we see that

$$\begin{aligned} &(\mathrm{Id} + T_{1,2} T_{1,0})[[\mathrm{id}, 1]] = [[\mathrm{id}, 1]] + T_{1,2}[[\beta, 1]], \\ &[[\mathrm{id}, 1]] + T_{1,2}[[\beta, 1]] \mapsto [\mathrm{id}, \underline{X}^{\underline{r}} - \underline{X}^{\underline{r-p+1}} \underline{Y}^{\underline{p-1}} + \underline{Y}^{\underline{r}}]. \end{aligned}$$

  Moreover $[\mathrm{id}, \underline{X}^{\underline{r}} - \underline{X}^{\underline{r-p+1}} \underline{Y}^{\underline{p-1}} + \underline{Y}^{\underline{r}}]$ generates $\mathrm{ind}_{KZ}^G\big(\bigotimes_{j=0}^{f-1} V_0^{\mathrm{Fr}^j}\big)$.

We have proved that the natural map induces isomorphisms

$$\begin{aligned} \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{1,2} T_{1,0})} &\xrightarrow{\sim} \mathrm{ind}_{KZ}^G\Bigg(\bigotimes_{j=0}^{f-1} V_0^{\mathrm{Fr}^j} \Bigg), \\ \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(\mathrm{Id} + T_{1,2} T_{1,0})} &\xrightarrow{\sim} \mathrm{ind}_{KZ}^G\Bigg(\bigotimes_{j=0}^{f-1} V_{p-1}^{\mathrm{Fr}^j} \Bigg). \end{aligned} \tag{4-3}$$

These maps are respectively given by

$$[[\mathrm{id}, 1]] \mapsto [\mathrm{id}, 1] \quad \text{and} \quad [[\mathrm{id}, 1]] \mapsto [\mathrm{id}, -(\underline{X}^{\underline{p-1}})].$$

Indeed, the first map sends [[id, 1]] to

$$[\mathrm{id}, \underline{Y}^r - \underline{X}^{r-p+1}\underline{Y}^{p-1}] \equiv [\mathrm{id}, \underline{X}^r - \underline{X}^{r-p+1}\underline{Y}^{p-1} + \underline{Y}^r] \mod \mathrm{ind}_{KZ}^G\left(\bigotimes_{j=0}^{f-1} V_{p-1}^{\mathrm{Fr}^j}\right)$$

and the second map sends [[id, 1]] to

$$[\mathrm{id}, \underline{Y}^r - \underline{X}^{r-p+1}\underline{Y}^{p-1}] \equiv [\mathrm{id}, -(\underline{X}^r)] \mod \mathrm{ind}_{KZ}^G\left(\bigotimes_{j=0}^{f-1} V_0^{\mathrm{Fr}^j}\right).$$

Now using (2-3) and (3-1), we see that

$$\mathrm{Im}\, T_{1,2}T_{1,0} \subseteq \mathrm{Ker}(T_{-1,0} + T_{1,0}) \quad \text{and} \quad \mathrm{Im}(\mathrm{Id} + T_{1,2}T_{1,0}) \subseteq \mathrm{Ker}\, T_{-1,0}.$$

Therefore, they are operators on the quotients

$$\frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{1,2}T_{1,0})} \quad \text{and} \quad \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(\mathrm{Id} + T_{1,2}T_{1,0})},$$

respectively. Then, under the first map in (4-3), we see that

$$(T_{-1,0} + T_{1,0})[[\mathrm{id}, 1]] = \sum_{\lambda \in \mathbb{F}_q} \left[\left[\begin{pmatrix} \varpi & [\lambda] \\ 0 & 1 \end{pmatrix}, 1\right]\right] + [[\beta, 1]]$$

$$\mapsto \sum_{\lambda \in \mathbb{F}_q} \left[\begin{pmatrix} \varpi & [\lambda] \\ 0 & 1 \end{pmatrix}, 1\right] + [\alpha, 1] = T[\mathrm{id}, 1]$$

since $\beta = \alpha w$ and $w \in KZ$. Similarly, under the second map in (4-3), we get

$$T_{-1,0}[[\mathrm{id}, 1]] = \sum_{\lambda \in \mathbb{F}_q} \left[\left[\begin{pmatrix} \varpi & [\lambda] \\ 0 & 1 \end{pmatrix}, 1\right]\right]$$

$$\mapsto \sum_{\lambda \in \mathbb{F}_q} \left[\begin{pmatrix} \varpi & [\lambda] \\ 0 & 1 \end{pmatrix}, -(\underline{X}^{p-1})\right] = T[\mathrm{id}, -(\underline{X}^{p-1})]. \quad \square$$

5. Self extensions of a supersingular representation of $\mathbf{GL_2(\mathbb{Q}_p)}$

Throughout this section, let p be an odd prime. Let $G = \mathrm{GL}_2(\mathbb{Q}_p)$ and $K = \mathrm{GL}_2(\mathbb{Z}_p)$ be its standard maximal compact subgroup. Let I be the Iwahori subgroup of G, and $I(1)$ be the pro-p-Iwahori subgroup of G. Let $\pi_r = \mathrm{ind}_{KZ}^G V_r/(T)$, where $r \in \{0, p-1\}$, be a supersingular representation of G.

Breuil [2003, Corollary 4.1.3] proved that the map $[\mathrm{id}, 1] \mapsto [\alpha, Y^{p-1}]$ induces an isomorphism

$$\frac{\mathrm{ind}_{KZ}^G V_0}{(T)} \cong \frac{\mathrm{ind}_{KZ}^G V_{p-1}}{(T)}.$$

Before we come to results, we need to consider the implications of the above isomorphism for finding the self-extensions of π_r. A key property of the Hecke operator T is its injectivity which implies that we have an exact sequence

$$0 \to \pi_r \to \frac{\operatorname{ind}_{KZ}^{G} V_r}{(T^2)} \to \pi_r \to 0.$$

The following lemma is well-known. We could not find its proof in the literature. Hence, we provide a short proof. One can compare this lemma on the automorphic side with [Bhattacharya et al. 2018, Theorem 7.53] on the Galois side using the nonsemisimple mod p local Langlands correspondence.

Lemma 5.1. *For $r \in \{0, p-1\}$, the exact sequence*

$$0 \to \pi_r \xrightarrow{T} \frac{\operatorname{ind}_{KZ}^{G} V_0}{(T^2)} \to \pi_r \to 0$$

is nonsplit.

Proof. Since $\pi_0 \cong \pi_{p-1}$, it suffices to prove that the sequence is nonsplit for $r = 0$. Suppose that the given sequence is split. So, there exists a retraction map

$$S : \frac{\operatorname{ind}_{KZ}^{G} V_0}{(T^2)} \to \pi_0$$

for the map T. Using the fact that [id, 1] is fixed under $I(1)$, alongside Breuil's determination of the $I(1)$-invariants of π_0 [2003, Theorem 3.2.4], we deduce that $S([\mathrm{id}, 1]) = [\mathrm{id}, a] + [\alpha, b]$ for some $a, b \in \mathbb{F}_p$. Also, since the element [id, 1] is fixed by w in $\frac{\operatorname{ind}_{KZ}^{G} V_0}{(T^2)}$, its image is fixed by w. Then, a quick computation using [Barthel and Livné 1994, Lemma 20] shows $b = 0$. Writing $T[\mathrm{id}, 1]$ as a sum of $\mathrm{GL}_2(\mathbb{Q}_p)$-translates of [id, 1], we see that S commutes with T. Consequently, $S(T[\mathrm{id}, 1]) = 0$ in π_0. However, since S is a retraction map for T, we get a contradiction. □

It is known using [Paškūnas 2010, Proposition 9.1, Theorem 10.13, Lemma 9.3; Breuil and Paškūnas 2012, Corollaries 6.6, 6.7] that the one-dimensional subspace of the three-dimensional $\operatorname{Ext}^1_G(\pi_r, \pi_r)$ spanned by the class of this self-extension consists of all the self extensions of π_r which have a four-dimensional space of $I(1)$-invariants, where Ext is computed in the category of smooth representations with a given central character ζ. This is in contrast with the *regular* case, and hence we write the argument in the form of the following proposition.

Proposition 5.2. *Let $r \in \{0, 1, \ldots, p-1\}$. Let $\operatorname{Ext}^1_G(\pi_r, \pi_r)$ be the space of self-extensions of π_r with a fixed central character ζ. Then, the subspace of $\operatorname{Ext}^1_G(\pi_r, \pi_r)$ consisting of extensions having a four-dimensional space of $I(1)$-invariants has dimension 1 if $r \in \{0, p-1\}$ or 2 if $r \notin \{0, p-1\}$.*

Proof. We claim that the subspace of $\mathrm{Ext}^1_G(\pi_r, \pi_r)$ consisting of the extensions with a four-dimensional space of $I(1)$-invariants is the image of $\mathrm{Ext}^1_{\mathcal{H}}(\pi_r^{I(1)}, \pi_r^{I(1)})$ under the map obtained using the functor $-\otimes_{\mathcal{H}} \mathrm{ind}_{I(1)Z}^G \zeta$. It is shown in [Paškūnas 2010, Proposition 9.1] that $\phi : \mathrm{Ext}^1_{\mathcal{H}}(\pi_r^{I(1)}, \pi_r^{I(1)}) \to \mathrm{Ext}^1_G(\pi_r, \pi_r)$ is injective. Moreover, [Paškūnas 2010, Lemma 9.3] says that

$$\dim \mathrm{Ext}^1_{\mathcal{H}}(\pi_r^{I(1)}, \pi_r^{I(1)}) = \begin{cases} 1 & \text{if } r \in \{0, p-1\}, \\ 2 & \text{otherwise.} \end{cases}$$

This is where the difference between the *regular* and the *nonregular* cases lies.

Next, we justify why every self-extension with a four-dimensional space of $I(1)$-invariants lies in the image of ϕ. Suppose E is one such self-extension of π_r. Then, taking $I(1)$-invariants, we get

$$0 \to \pi_r^{I(1)} \to E^{I(1)} \to \pi_r^{I(1)} \to 0.$$

This sequence is exact because of the assumption that $\dim E^{I(1)} = 4$ and the fact that $\dim \pi_r^{I(1)} = 2$. Finally, we note that the two copies of π_r appearing as the submodule and the quotient module of E are generated as G-modules by their $I(1)$-invariants. Therefore E is also generated as a G-module by $E^{I(1)}$. This means

$$E \simeq E^{I(1)} \otimes_{\mathcal{H}} \mathrm{ind}_{I(1)Z}^G \zeta,$$

and thus the element in $\mathrm{Ext}^1_G(\pi_r, \pi_r)$ corresponding to E is in the image of ϕ. □

In this section, we focus on giving a realization of the self-extensions of π_r, belonging to the above-mentioned one-dimensional space, using the Iwahori–Hecke model for the *nonregular* case, i.e., $r \in \{0, p-1\}$, given in [Chitrao 2025]. For the *regular* case $0 < r < p-1$, see [Anandavardhanan and Borisagar 2015, Theorem 1.2].

The following theorem gives the Iwahori–Hecke model for a supersingular representation of $\mathrm{GL}_2(\mathbb{Q}_p)$ in the *nonregular* case.

Theorem 5.3. *We have*

$$\frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{1,2}T_{1,0})} \cong \mathrm{ind}_{KZ}^G V_0 \quad \textit{and} \quad \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(\mathrm{Id} + T_{1,2}T_{1,0})} \cong \mathrm{ind}_{KZ}^G V_{p-1}.$$

Moreover, the operators $T_{-1,0} + T_{1,0}$ and $T_{-1,0}$ on the left correspond to the operator T on the right.

Proof. See Theorem 3.13 in [Chitrao 2025]. □

Explicitly, Breuil's isomorphism in the *nonregular* case is given as follows.

Theorem 5.4 (Breuil's isomorphism). *We have*

$$\frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{-1,0}) + (T_{1,2} + T_{1,0})} \cong \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{1,0} + T_{-1,0}) + (T_{1,2})}.$$

Proof. We construct the isomorphism using the action of $T_{1,0}$ on $\mathrm{ind}_{IZ}^{G}\mathbb{1}$. Since $T_{1,0}$ is invertible, it induces an automorphism of the entire space. Using (2-3) and (3-1), we see that $T_{1,0}$ interchanges the defining operators of the quotients via $T_{1,0}T_{-1,0}T_{1,0} = T_{1,2}$ and $T_{1,0}(T_{1,2}+T_{1,0})T_{1,0} = T_{-1,0}+T_{1,0}$. Consequently, $T_{1,0}$ maps the subspace $(T_{-1,0})+(T_{1,2}+T_{1,0})$ bijectively onto $(T_{1,0}+T_{-1,0})+(T_{1,2})$, ensuring the induced map on the quotients is a well-defined isomorphism. □

Theorem 5.5. *Consider the irreducible supersingular representation* $\pi_0 = \frac{\mathrm{ind}_{KZ}^{G} V_0}{(T)}$ *of* $\mathrm{GL}_2(\mathbb{Q}_\mathrm{p})$. *Then, a basis for the subspace of* $\mathrm{Ext}_G^1(\pi_0,\pi_0)$ *consisting of isomorphism class of short exact sequence*

$$0 \to \pi_0 \to \tau \to \pi_0 \to 0,$$

where the representation τ *has a four-dimensional space of* $I(1)$*-invariants* (*as explicitly computed in Remark 5.9*), *is given by isomorphism class of*

$$0 \to \pi_0 \to \frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{-1,0})+(T_{1,2}+T_{1,0})^2} \cong \frac{\mathrm{ind}_{KZ}^{G} V_0}{(T^2)} \to \pi_0 \to 0.$$

Theorem 5.6. *Consider the irreducible supersingular representation* $\pi_{p-1} = \frac{\mathrm{ind}_{KZ}^{G} V_{p-1}}{(T)}$ *of* $\mathrm{GL}_2(\mathbb{Q}_\mathrm{p})$. *Then, a basis for the subspace of* $\mathrm{Ext}_G^1(\pi_{p-1},\pi_{p-1})$ *consisting of isomorphism class of short exact sequence*

$$0 \to \pi_{p-1} \to \tau' \to \pi_{p-1} \to 0,$$

where the representation τ' *has a four-dimensional space of* $I(1)$*-invariants* (*see Remark 5.9*), *is given by isomorphism class of*

$$0 \to \pi_{p-1} \to \frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{1,0}+T_{-1,0})+(T_{1,2}^2)} \cong \frac{\mathrm{ind}_{KZ}^{G} V_{p-1}}{(T^2)} \to \pi_{p-1} \to 0.$$

Before we prove the theorems we note that

$$\tau = \frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{-1,0})+(T_{1,2}+T_{1,0})^2} \quad \text{and} \quad \tau' = \frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{1,0}+T_{-1,0})+(T_{1,2}^2)}$$

are isomorphic by defining a map

$$\Phi : \frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{X=[(T_{1,0}+T_{-1,0})+(T_{1,2}^2)]} \to \frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{[(T_{-1,0})+(T_{1,2}+T_{1,0})^2]=Y}$$

as

$$\Phi(f+X) = (-T_{1,2})(f)+Y.$$

It can be checked that Φ is a well-defined G-linear map. Indeed, we must show that Φ maps $(T_{1,0}+T_{-1,0})$ and $(T_{1,2}^2)$ to 0. The former is a consequence of

$\mathrm{Im}(T_{1,0}+T_{-1,0}) = \mathrm{Ker}\, T_{1,2}$. For the latter, note that

$$\begin{aligned}(T_{1,2}+T_{1,0})^2(-T_{1,2}) &= (T_{1,2}^2 + T_{1,2}T_{1,0} + T_{1,0}T_{1,2} + \mathrm{Id})(-T_{1,2})\\ &= -T_{1,2}^3 - (-T_{1,2}) - T_{1,0}T_{1,2}^2 - T_{1,2}\\ &= -T_{1,2}^3 - T_{1,0}T_{1,2}^2\\ &= -T_{1,2}^3 - T_{-1,0}T_{1,0}T_{1,2}.\end{aligned}$$

Therefore, $(T_{1,2}^3)$ belongs to $(T_{-1,0}) + (T_{1,2}+T_{1,0})^2$. Also, Φ makes the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{1,0}+T_{-1,0})+(T_{1,2})} & \xrightarrow{T_{1,2}} & \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{X} & \longrightarrow & \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{1,0}+T_{-1,0})+(T_{1,2})} & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle T_{1,0}} & & \downarrow{\scriptstyle -T_{1,2}=\Phi} & & \downarrow{\scriptstyle T_{1,0}} & & \\ 0 & \longrightarrow & \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{-1,0})+(T_{1,2}+T_{1,0})} & \xrightarrow{T_{1,2}+T_{1,0}} & \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{Y} & \longrightarrow & \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{-1,0})+(T_{1,2}+T_{1,0})} & \longrightarrow & 0 \end{array} \tag{5-1}$$

commute. Indeed, it is easy to see that the second square commutes. For the first square, we must show that $-T_{1,2}^2 f = (T_{1,2}T_{1,0}+\mathrm{Id})f$ in $\frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{Y}$ for any $f \in \mathrm{ind}_{IZ}^G \mathbb{1}$. Note that

$$(T_{1,2}+T_{1,0})^2 - T_{-1,0}T_{1,0} = T_{1,2}^2 + T_{1,2}T_{1,0} + \mathrm{Id}.$$

Hence, Φ is an isomorphism.

5.1. *Proof of Theorem 5.5.*

Proof. We prove that

$$\tau = \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{-1,0}) + (T_{1,2}+T_{1,0})^2} \in \mathrm{Ext}_G^1(\pi_0, \pi_0).$$

Let $\pi_0 \cong \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{-1,0})+(T_{1,2}+T_{1,0})}$ be an irreducible supersingular representation of $\mathrm{GL}_2(\mathbb{Q}_p)$. Consider the natural surjection

$$\begin{aligned}\Psi : \tau &\to \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{-1,0}) + (T_{1,2}+T_{1,0})},\\ f + (T_{-1,0}) + (T_{1,2}+T_{1,0})^2 &\mapsto f + (T_{-1,0}) + (T_{1,2}+T_{1,0}).\end{aligned}$$

We easily see $\mathrm{Ker}\,\Psi = \frac{(T_{-1,0})+(T_{1,2}+T_{1,0})}{Y=[(T_{-1,0})+(T_{1,2}+T_{1,0})^2]}$. We claim $\mathrm{Ker}\,\Psi \cong \frac{\mathrm{ind}_{IZ}^G \mathbb{1}}{(T_{-1,0})+(T_{1,2}+T_{1,0})}$. Consider a map

$$\begin{aligned}\phi : \mathrm{ind}_{IZ}^G \mathbb{1} &\to \frac{(T_{-1,0}) + (T_{1,2}+T_{1,0})}{Y = [(T_{-1,0}) + (T_{1,2}+T_{1,0})^2]},\\ f &\mapsto (T_{1,2}+T_{1,0})(f) + Y.\end{aligned}$$

Note that ϕ is a G-linear onto map. We show that $\mathrm{Ker}\,\phi = (T_{-1,0}) + (T_{1,2}+T_{1,0})$. For that, we prove the following result.

Lemma 5.7. *For an arbitrary local field F, we have* $\mathrm{Ker}(T_{-1,0}) = \mathrm{Im}(T_{1,2} + T_{1,0})$ *and* $\mathrm{Ker}(T_{1,2} + T_{1,0}) = \mathrm{Im}(T_{-1,0})$.

Proof. Since $T_{1,0}$ is invertible, $\mathrm{Ker}(T_{1,2}T_{1,0}) = \mathrm{Ker}(T_{1,0}T_{1,2}T_{1,0}) = \mathrm{Ker}(T_{-1,0})$ and $\mathrm{Im}(T_{1,2} + T_{1,0}) = \mathrm{Im}(T_{1,2}T_{1,0} + \mathrm{Id})$. Using the fact $-T_{1,2}T_{1,0}$ is an idempotent, we see $\mathrm{Ker}(T_{1,2}T_{1,0}) = \mathrm{Im}(T_{1,2}T_{1,0} + \mathrm{Id})$. Therefore, we get the first equality. For the second equality, a similar proof applies using $\mathrm{Im}(T_{1,0}T_{1,2}) = \mathrm{Ker}(T_{1,0}T_{1,2} + \mathrm{Id})$. □

We can observe from Lemma 5.8 that $\mathrm{Im}(T_{1,2}+T_{1,0})+\mathrm{Im}(T_{-1,0}) \subseteq \mathrm{Ker}\,\phi$. Since $\frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{-1,0})+(T_{1,2}+T_{1,0})}$ is irreducible and ϕ is nonzero, $\mathrm{Ker}\,\phi = (T_{-1,0}) + (T_{1,2} + T_{1,0})$. Hence,

$$\frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{-1,0}) + (T_{1,2} + T_{1,0})} \cong \frac{(T_{-1,0}) + (T_{1,2} + T_{1,0})}{X = [(T_{-1,0}) + (T_{1,2} + T_{1,0})^2]} = \mathrm{Ker}\,\Psi,$$

which proves our claim $\tau = \frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{-1,0})+(T_{1,2}+T_{1,0})^2} \in \mathrm{Ext}_G^1(\pi_0, \pi_0)$. □

5.2. *Proof of Theorem 5.6.*

Proof. We prove that

$$\tau' = \frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{1,0} + T_{-1,0}) + (T_{1,2}^2)} \in \mathrm{Ext}_G^1(\pi_{p-1}, \pi_{p-1}).$$

Let $\pi_{p-1} \cong \frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{1,0}+T_{-1,0})+(T_{1,2})}$ be an irreducible supersingular representation of the group $\mathrm{GL}_2(\mathbb{Q}_p)$. Consider the natural surjection

$$\Psi' : \tau' \to \frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{1,0} + T_{-1,0}) + (T_{1,2})},$$
$$f + (T_{1,0} + T_{-1,0}) + (T_{1,2}^2) \mapsto f + (T_{1,0} + T_{-1,0}) + (T_{1,2}).$$

We easily see $\mathrm{Ker}\,\Psi' = \frac{(T_{1,0}+T_{-1,0})+(T_{1,2})}{X=[(T_{1,0}+T_{-1,0})+(T_{1,2}^2)]}$. We claim $\mathrm{Ker}\,\Psi' \cong \frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{1,0}+T_{-1,0})+(T_{1,2})}$. Consider a map

$$\phi' : \mathrm{ind}_{IZ}^{G}\mathbb{1} \to \frac{(T_{1,0} + T_{-1,0}) + (T_{1,2})}{X = [(T_{1,0} + T_{-1,0}) + (T_{1,2}^2)]},$$
$$f \mapsto (T_{1,2})(f) + X.$$

Note that ϕ' is a G-linear onto map. We show that $\mathrm{Ker}\,\phi' = (T_{1,0} + T_{-1,0}) + (T_{1,2})$. For that, we prove the following result.

Lemma 5.8. *For an arbitrary local field F, we have* $\mathrm{Ker}(T_{1,0} + T_{-1,0}) = \mathrm{Im}(T_{1,2})$ *and* $\mathrm{Ker}(T_{1,2}) = \mathrm{Im}(T_{1,0} + T_{-1,0})$.

Proof. Since $T_{1,0}$ is invertible,

$$\mathrm{Ker}(T_{1,0}+T_{-1,0}) = \mathrm{Ker}(\mathrm{Id}+T_{1,2}T_{1,0}) = \mathrm{Im}(T_{1,2}T_{1,0}) = \mathrm{Im}(T_{1,2}).$$

For the second equality, use that $\mathrm{Ker}(T_{1,2}T_{1,0}) = \mathrm{Im}(T_{1,2}T_{1,0}+\mathrm{Id})$, $T_{1,0}$ induces an isomorphism $\mathrm{Ker}(T_{1,2}) \cong \mathrm{Ker}(T_{1,2}T_{1,0})$ and $\mathrm{Im}(T_{1,0}+T_{-1,0}) \cong \mathrm{Im}(T_{1,2}T_{1,0}+\mathrm{Id})$. □

We can observe from Lemma 5.7 that $\mathrm{Im}(T_{-1,0}+T_{1,0})+\mathrm{Im}(T_{1,2}) \subseteq \mathrm{Ker}\,\phi'$. Since $\frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{1,0}+T_{-1,0})+(T_{1,2})}$ is irreducible and ϕ' is nonzero, $\mathrm{Ker}\,\phi' = (T_{1,0}+T_{-1,0})+(T_{1,2})$. Hence,

$$\frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{1,0}+T_{-1,0})+(T_{1,2})} \cong \frac{(T_{1,0}+T_{-1,0})+(T_{1,2})}{X=[(T_{1,0}+T_{-1,0})+(T_{1,2}^2)]} = \mathrm{Ker}\,\Psi',$$

which proves our claim $\tau' = \frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{1,0}+T_{-1,0})+(T_{1,2}^2)} \in \mathrm{Ext}_G^1(\pi_{p-1}, \pi_{p-1})$. □

Remark 5.9. We first recall the vector space isomorphisms for the two-dimensional spaces of $I(1)$-invariants associated to the base representations:

$$\left(\frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{-1,0})+(T_{1,2}+T_{1,0})}\right)^{I(1)} \cong \left(\frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{1,0}+T_{-1,0})+(T_{1,2})}\right)^{I(1)}.$$

Next, the space of $I(1)$-invariants in τ is given by

$$\left(\frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{-1,0})+(T_{1,2}+T_{1,0})^2}\right)^{I(1)} = \left\langle \overline{[[\mathrm{id},1]]}, \overline{[[\beta,1]]}, \overline{(T_{1,2}+T_{1,0})([[\mathrm{id},1]])}, \overline{(T_{1,2}+T_{1,0})([[\beta,1]])} \right\rangle.$$

Furthermore, τ and τ' are isomorphic (see (5-1)). Consequently, τ' also possesses a four-dimensional space of $I(1)$-invariants given by

$$\left(\frac{\mathrm{ind}_{IZ}^{G}\mathbb{1}}{(T_{1,0}+T_{-1,0})+(T_{1,2}^2)}\right)^{I(1)} = \left\langle \overline{[[\mathrm{id},1]]}, \overline{[[\beta,1]]}, \overline{T_{1,2}([[\mathrm{id},1]])}, \overline{T_{1,2}([[\beta,1]])} \right\rangle.$$

Acknowledgements

We thank Prof. Ghate, Prof. Anandavardhanan, Prof. Borisagar, and Jana for fruitful discussions. We also thank the referee whose careful reading and insightful comments helped improve the quality of this paper. Chitrao gratefully acknowledges support from the HRI postdoctoral fellowship. Soneji thanks Prof. Ghate for the invitation and warm hospitality during his visit to TIFR, Mumbai. Part of the work was done during Soneji's visit to TIFR. Additionally, Soneji acknowledges the financial assistance as a Junior Research Fellowship (NET) from UGC, India.

References

[Anandavardhanan and Borisagar 2013] U. K. Anandavardhanan and G. H. Borisagar, "On the $K(n)$-invariants of a supersingular representation of $GL_2(\mathbb{Q}_p)$", pp. 55–75 in *The legacy of Srinivasa Ramanujan*, edited by B. C. Berndt and D. Prasad, Ramanujan Math. Soc. Lect. Notes Ser. **20**, Ramanujan Math. Soc., Mysore, 2013. MR

[Anandavardhanan and Borisagar 2015] U. K. Anandavardhanan and G. H. Borisagar, "Iwahori–Hecke model for supersingular representations of $\mathrm{GL}_2(\mathbb{Q}_p)$", *J. Algebra* **423** (2015), 1–27. MR

[Anandavardhanan and Jana 2021] U. K. Anandavardhanan and A. Jana, "Iwahori–Hecke model for mod p representations of GL(2, F)", *Pacific J. Math.* **315**:2 (2021), 255–283. MR

[Barthel and Livné 1994] L. Barthel and R. Livné, "Irreducible modular representations of GL_2 of a local field", *Duke Math. J.* **75**:2 (1994), 261–292. MR

[Barthel and Livné 1995] L. Barthel and R. Livné, "Modular representations of GL_2 of a local field: the ordinary, unramified case", *J. Number Theory* **55**:1 (1995), 1–27. MR

[Bhattacharya et al. 2018] S. Bhattacharya, E. Ghate, and S. Rozensztajn, "Reductions of Galois representations of slope 1", *J. Algebra* **508** (2018), 98–156. MR

[Breuil 2003] C. Breuil, "Sur quelques représentations modulaires et p-adiques de $\mathrm{GL}_2(\mathbb{Q}_p)$, I", *Compositio Math.* **138**:2 (2003), 165–188. MR

[Breuil and Paškūnas 2012] C. Breuil and V. Paškūnas, *Towards a modulo p Langlands correspondence for* GL_2, Mem. Amer. Math. Soc. **1016**, Amer. Math. Soc., Providence, RI, 2012. MR

[Chitrao 2025] A. Chitrao, "An Iwahori theoretic mod p local Langlands correspondence", *Canad. Math. Bull.* **68**:3 (2025), 805–817. MR

[Chitrao and Ghate 2023] A. Chitrao and E. Ghate, "Reductions of semi-stable representations using the Iwahori mod p local Langlands correspondence", 2023. arXiv 2311.03740

[Chitrao and Ghate 2026] A. Chitrao and E. Ghate, "Reduction mod p of semi-stable representations of some super-Breuil weights", 2026. arXiv 2604.16867

[Chitrao et al. 2026] A. Chitrao, A. Jana, and A. Soneji, "On the $I(1)$-invariants: non-abelian Hecke algebra case", 2026. arXiv 2604.20218

[Ghate and Jana 2025] E. Ghate and A. Jana, "Modular representations of $\mathrm{GL}_2(\mathbb{F}_q)$ using calculus", *Forum Math.* **37**:5 (2025), 1503–1543. MR

[Ghate and Pandey 2026] E. Ghate and S. Pandey, "Reductions of $\mathrm{GL}_2(\mathbb{Q}_{p^f})$-Banach spaces of slopes in (0, 1)", 2026. arXiv 2605.02654

[Ghate and Vangala 2022] E. Ghate and R. Vangala, "The monomial lattice in modular symmetric power representations", *Algebr. Represent. Theory* **25**:1 (2022), 121–185. MR

[Hendel 2019] Y. I. Hendel, "On the universal mod p supersingular quotients for $\mathrm{GL}_2(F)$ over $\overline{\mathbb{F}}_p$ for a general $F/\mathbb{Q}_p$", *J. Algebra* **519** (2019), 1–38. MR

[Jana 2023] A. Jana, "The pro-p-Iwahori invariants of the universal quotient representation for $\mathrm{GL}_2(F)$", *Int. J. Number Theory* **19**:8 (2023), 1853–1879. MR

[Morra 2013] S. Morra, "Invariant elements for p-modular representations of $\mathrm{GL}_2(\mathbb{Q}_p)$", *Trans. Amer. Math. Soc.* **365**:12 (2013), 6625–6667. MR

[Paškūnas 2010] V. Paškūnas, "Extensions for supersingular representations of $\mathrm{GL}_2(\mathbb{Q}_p)$", pp. 317–353 in *Représentations p-adiques de groupes p-adiques*, *III*: *Méthodes globales et géométriques*, edited by L. Berger et al., Astérisque **331**, Soc. Math. France, Paris, 2010. MR

[Schein 2011] M. M. Schein, "An irreducibility criterion for supersingular mod p representations of $\mathrm{GL}_2(F)$ for totally ramified extensions F of $\mathbb{Q}_p$", *Trans. Amer. Math. Soc.* **363**:12 (2011), 6269–6289. MR

[Schein 2014] M. M. Schein, "On the universal supersingular mod p representations of $\mathrm{GL}_2(F)$", *J. Number Theory* **141** (2014), 242–277. MR

[Schein 2023] M. M. Schein, "A family of irreducible supersingular representations of $\mathrm{GL}_2(F)$ for some ramified p-adic fields", *Israel J. Math.* **255**:2 (2023), 911–930. MR

Received August 26, 2025. Revised May 11, 2026.

ANAND CHITRAO
DEPARTMENT OF MATHEMATICAL SCIENCES
ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY
ULSAN
SOUTH KOREA
anandchitrao@gmail.com

ASFAK SONEJI
DEPARTMENT OF BASIC SCIENCES
INSTITUTE OF INFRASTRUCTURE, TECHNOLOGY, RESEARCH AND MANAGEMENT
AHMEDABAD
INDIA
sonejiasfak96@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 344, No. 1, 2026
https://doi.org/10.2140/pjm.2026.344.47

THE CANONICAL EXACT SEQUENCE OF DIFFERENTIAL MODULES FOR 0-DIMENSIONAL SCHEMES

TRAN N. K. LINH AND LE NGOC LONG

Given a 0-dimensional scheme $\mathbb{X}$ in $\mathbb{P}^n_K$ over a perfect field K, we examine the second differential power of its homogeneous vanishing ideal. This enables us to establish the canonical exact sequence for the associated Kähler differential module. We also provide a formula for the Hilbert polynomial of Kähler differential modules when $\mathbb{X}$ is either a fat point scheme or a 0-dimensional locally monomial Gorenstein scheme.

1. Introduction

Let K be a perfect field, and let R be a finitely generated K-algebra of the form $R = P/I$, where $P = K[X_0, \ldots, X_n]$ is a polynomial ring over K and I is an ideal of P. One of the powerful tools for investigating the K-algebra R is its associated Kähler differential module. This module is precisely defined as $\Omega^1_R = \Delta_R/\Delta_R^2$, where Δ_R is the kernel of the multiplication map $\mu : R \otimes_K R \to R$ with $f \otimes g \mapsto fg$. The module Ω^1_R is finitely generated and it reflects many structural properties of the K-algebra R, including criteria for smoothness, regularity, and ramification (see, for instance, [EGA IV$_4$ 1967; Kunz 1986; Scheja and Storch 1974]).

It is well known that Ω^1_R admits the canonical exact sequence of R-modules

$$I/I^2 \xrightarrow{\delta} \Omega^1_P/I\Omega^1_P \xrightarrow{\gamma} \Omega^1_R \to 0,$$

where $\delta(f + I^2) = df + I\Omega^1_P$ for each $f \in I$ and the map γ is induced from the functorial map $\Omega^1_P \to \Omega^1_R$. A significant result by Mohan Kumar [1978] demonstrates that when R is reduced, I is a complete intersection if and only if the exact canonical sequence forms a free resolution of Ω^1_R, i.e., if and only if

$$0 \to I/I^2 \xrightarrow{\delta} R^{n+1} \xrightarrow{\gamma} \Omega^1_R \to 0$$

is exact and I/I^2 is a free R-module. This exactness condition has been further explored in subsequent works. De Dominicis and Kreuzer [1999] proved that if

MSC2020: primary 13N05; secondary 13D40, 13N10, 14N05.

Keywords: Kähler differential modules, differential powers, Hilbert functions, zero-dimensional schemes, locally monomial Gorenstein schemes.

© 2026 The Authors, under license to MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

$I = I_{\mathbb{X}}$ is the homogeneous vanishing ideal of a finite set of points $\mathbb{X} = \{p_1, \dots, p_s\}$ in $\mathbb{P}^n_K$ over a field K of characteristic zero, then the sequence of graded R-modules

$$0 \to I_{\mathbb{X}}/I_{\mathbb{Y}} \xrightarrow{\delta} R^{n+1}(-1) \xrightarrow{\gamma} \Omega^1_R \to 0 \tag{1}$$

is exact, where $I_{\mathbb{Y}}$ is the homogeneous vanishing ideal of the double point scheme $\mathbb{Y}$ in $\mathbb{P}^n_K$. Recall that, for K-rational points $p_1, \dots, p_s$ and positive integers $m_1, \dots, m_s$, the saturated homogeneous ideal $I_{\mathbb{X}} = I_{p_1}^{m_1} \cap \cdots \cap I_{p_s}^{m_s}$ of P defines a fat point scheme $\mathbb{X} = m_1 p_1 + \cdots + m_s p_s$ in $\mathbb{P}^n_K$. Recently, when $\operatorname{char}(K) = 0$, the canonical exact sequence (1) was generalized in [Kreuzer et al. 2015, Theorem 1.7] to such a fat point scheme $\mathbb{X} = m_1 p_1 + \cdots + m_s p_s$ using its fattening $\mathbb{Y} = (m_1+1)p_1 + \cdots + (m_s+1)p_s$ in $\mathbb{P}^n_K$. Inspired by these advancements, it is natural to ask the following question:

Question 1.1. *Given an arbitrary* 0*-dimensional scheme* $\mathbb{X}$ *in* $\mathbb{P}^n_K$ *over a perfect field* K *of characteristic* $p \geq 0$*, does a subscheme* $\mathbb{Y}$ *of* $\mathbb{P}^n_K$ *exist for which the canonical exact sequence* (1) *for the Kähler differential module for* $\mathbb{X}$ *remains valid*?

To answer this question, we will first deal with the differential powers of the homogeneous vanishing ideal $I_{\mathbb{X}}$ of the scheme $\mathbb{X}$. Throughout, we assume that the support of $\mathbb{X}$ is disjoint from the hyperplane $\mathcal{Z}(X_0)$. This can be achieved without loss of generality by a suitable linear change of coordinates, possibly after a small base field extension if K is finite. The set $\{X_0^{\alpha_0} \cdots X_n^{\alpha_n} \mid (\alpha_0, \dots, \alpha_n) \in \mathbb{N}^{n+1}\}$ of all terms in P is denoted by $\mathbb{T}^{n+1}$. For $m \geq 0$, let D_P^m be the P-module of K-linear differential operators on P of order at most m. On account of [Dao et al. 2018, Definition 2.2], the m-th differential power of the ideal $I_{\mathbb{X}}$ is given by $I_{\mathbb{X}}^{\langle m \rangle} = \{f \in P \mid \delta(f) \in I_{\mathbb{X}} \text{ for all } \delta \in D_P^{m-1}\}$. The interplay between differential and symbolic powers of an ideal has been studied extensively (see, for instance, [Dao et al. 2018; De Stefani et al. 2020; Hochster and Huneke 2002; 2007]). For our purpose, we focus on investigating the second differential power $I_{\mathbb{X}}^{\text{diff}} := I_{\mathbb{X}}^{\langle 2 \rangle}$. Our first result is the following theorem that shows how to compute a finite generating system of $I_{\mathbb{X}}^{\text{diff}}$, when $\mathbb{X}$ contains only a monomial point.

Theorem 1.2 (Theorem 3.2). *Let* K *be a perfect field of characteristic* $p \geq 0$*. Assume that the homogeneous vanishing ideal* $I_{\mathbb{X}}$ *of the scheme* $\mathbb{X}$ *is a proper monomial ideal of* P*, minimally generated by terms* $t_1, \dots, t_r \in \mathbb{T}^{n+1}$*.*

(a) *The ideal* $I_{\mathbb{X}}^{\text{diff}}$ *is a monomial subideal of* $I_{\mathbb{X}}$*.*

(b) *For* $1 \leq j \leq r$*, we write* $t_j = X_0^{\alpha_{0j}} \cdots X_n^{\alpha_{nj}}$*, and for* $2 \leq k \leq n+1$ *and* $1 \leq j_1 \leq \cdots \leq j_k \leq r$*, we define* $t_{j_1,\dots,j_k} := X_0^{\beta_0} \cdots X_n^{\beta_n}$*, where, for* $i = 0, \dots, n$*,*

$$\beta_i := \begin{cases} 0 & \text{if } \alpha_{ij_1} = \cdots = \alpha_{ij_k} = 0, \\ \alpha_{ij_1} & \text{if } \alpha_{ij_1} = \cdots = \alpha_{ij_k} > 0 \text{ and } p \mid \alpha_{ij_1}, \\ \alpha_{ij_1} + 1 & \text{if } \alpha_{ij_1} = \cdots = \alpha_{ij_k} > 0 \text{ and } p \nmid \alpha_{ij_1}, \\ \max\{\alpha_{ij_1}, \dots, \alpha_{ij_k}\} & \text{otherwise.} \end{cases} \tag{2}$$

Then

$$I_{\mathbb{X}}^{\mathrm{diff}} = \langle t_{j_1,\dots,j_k} \mid 1 \le j_1 \le \cdots \le j_k \le r;\ 2 \le k \le n+1 \rangle.$$

In the general case of a 0-dimensional scheme $\mathbb{X}$, we show that $I_{\mathbb{X}}^{\mathrm{diff}}$ is indeed a saturated homogeneous ideal in P, and so it defines a 0-dimensional scheme $\mathbb{Y}$ in $\mathbb{P}_K^n$ which is called the *differential fattening* of $\mathbb{X}$. This notion can be viewed as a natural extension of the classical notion of fattening for fat point schemes. Specifically, when the characteristic of the base field does not divide the multiplicities of a fat point scheme, its first fattening is precisely the scheme defined by the second differential power of its vanishing ideal (see Remark 3.6). By x_i we denote the image of X_i in $R = P/I_{\mathbb{X}}$ for $i = 0, \dots, n$. Using the above terminology, we give an affirmative answer for Question 1.1.

Theorem 1.3 (Theorem 3.8). *Let $\mathbb{X}$ be a 0-dimensional scheme in $\mathbb{P}_K^n$ with homogeneous vanishing ideal $I_{\mathbb{X}}$, and let $\mathbb{Y}$ be the differential fattening of $\mathbb{X}$ with its homogeneous vanishing ideal $I_{\mathbb{Y}}$. The sequence of graded R-modules*

$$0 \to I_{\mathbb{X}}/I_{\mathbb{Y}} \xrightarrow{\delta} R^{n+1}(-1) \xrightarrow{\gamma} \Omega_R^1 \to 0$$

is exact, where $\delta(f + I_{\mathbb{Y}}) = \sum_{i=0}^{n} (\partial f/\partial x_i) e_i$ for every homogeneous polynomial $f \in I_{\mathbb{X}}$ and $\gamma(e_i) = dx_i$ for $i = 0, \dots, n$.

Clearly, the theorem enables us to compute the Hilbert function of the Kähler differential module Ω_R^1 and to bound its regularity index via those of $\mathbb{X}$ and of $\mathbb{Y}$. When $\mathbb{X} = m_1 p_1 + \cdots + m_s p_s$ is a fat point scheme in $\mathbb{P}_K^n$, Kreuzer, Linh, and Long [Kreuzer et al. 2025, Remark 7.1] posed the question: *does the Hilbert polynomial of the m-form Kähler differential module Ω_R^m depend only on m, n and the multiplicities $m_1, \dots, m_s$?* A positive answer for this question is given in [Kreuzer et al. 2025, Theorem 7.12] when either $\mathrm{char}(K) = 0$ or $\mathrm{char}(K) > \max\{m_j \mid j = 1, \dots, s\}$. The proof therein combines localization arguments together with either the exactness of the Koszul–Euler complex for Kähler differential modules or Gröbner basis techniques to control the Hilbert function. However, these methods rely essentially on the condition that the characteristic does not divide the multiplicities, and hence they do not extend to the remaining case $0 < \mathrm{char}(K) \le \max\{m_j \mid j = 1, \dots, s\}$. In this range, the subtle behavior of differential operators causes the standard connection between differential and symbolic powers to fail (see [De Stefani et al. 2020, Examples 3.7–3.8] or Example 3.3). Building on Theorems 1.2 and 1.3, we give a formula for the Hilbert polynomial of Ω_R^1 and this improves [Kreuzer et al. 2025, Theorem 7.12] for $m = 1$ and $0 < \mathrm{char}(K) \le \max\{m_j \mid j = 1, \dots, s\}$ (see Corollary 3.10). Unfortunately, we do not know an answer to this question when $m \ge 2$ and $0 < \mathrm{char}(K) \le \max\{m_j \mid j = 1, \dots, s\}$.

On the other hand, we are also interested in finding a formula for the Hilbert polynomial of the m-form Kähler differential module for 0-dimensional locally

monomial Gorenstein schemes in $\mathbb{P}^n_K$. Here, a 0-dimensional scheme $\mathbb{X}$ with support $\mathrm{Supp}(\mathbb{X}) = \{p_1, \ldots, p_s\}$ in $\mathbb{P}^n_K$ is a locally monomial Gorenstein scheme if, for each $j \in \{1, \ldots, s\}$, the local ring $\mathcal{O}_j$ of $\mathbb{X}$ at p_j satisfies $\mathcal{O}_j \cong K[X_1, \ldots, X_n]/\mathfrak{Q}_j$, where $\mathfrak{Q}_j$ is a monomial Gorenstein ideal generated by pure powers of the indeterminates. The class of these schemes includes the class of 0-dimensional curvilinear schemes which have been shown to play an important role in the proof of the so-called Alexander–Hirschowitz theorem (see, for instance, [Chandler 2001]).

Corollary 1.4 (Corollary 4.8). *Let $\mathbb{X}$ be a 0-dimensional locally monomial Gorenstein scheme in $\mathbb{P}^n_K$ with $\mathrm{Supp}(\mathbb{X}) = \{p_1, \ldots, p_s\}$, and let $p = \mathrm{char}(K)$. For $j = 1, \ldots s$, we write $\mathcal{O}_j \cong K[X_1, \ldots, X_n]/\mathfrak{Q}_j$, where $\mathfrak{Q}_j = \langle X_1^{k_{1j}}, \ldots, X_n^{k_{nj}} \rangle$ with $k_{ij} \geq 1$. For $1 \leq i_1 < \cdots < i_m \leq n$ and $1 \leq j \leq s$, let*

$$\Gamma_{(i_1,\ldots,i_m),j} = \{i \in \{i_1, \ldots, i_m\} : p \nmid k_{ij}\}.$$

For $m \geq 1$, we have

$$\mathrm{HP}(\Omega^m_R) = \sum_{j=1}^{s} \Bigg[\sum_{1 \leq i_1 < \cdots < i_{m-1} \leq n} \prod_{i \notin \Gamma_{(i_1,\ldots,i_{m-1}),j}} k_{ij} \prod_{i \in \Gamma_{(i_1,\ldots,i_{m-1}),j}} (k_{ij} - 1) + \sum_{1 \leq l_1 < \cdots < l_m \leq n} \prod_{i \notin \Gamma_{(l_1,\ldots,l_m),j}} k_{ij} \prod_{i \in \Gamma_{(l_1,\ldots,l_m),j}} (k_{ij} - 1) \Bigg].$$

This paper is outlined as follows. In Section 2, we first recall the needed facts about the Kähler differential modules, the ring of K-linear differential operators, and the K-linear differential powers. Then we consider these invariants for a 0-dimensional scheme in $\mathbb{P}^n_K$. In Section 3, we prove Theorems 1.2 and 1.3 and derive their consequences (see Corollaries 3.5, 3.10, and 3.11). In the final section, we establish a formula for the Hilbert polynomial of the m-form Kähler differential module associated to a 0-dimensional locally monomial Gorenstein scheme (see Proposition 4.4 and Corollary 1.4).

Unless explicitly stated otherwise, we adhere to the definitions and notation introduced in [Kreuzer and Robbiano 2000; 2005; Kunz 1986]. All examples in this paper were calculated by using the computer algebraic system `ApCoCoA` [$\geq$ 2026].

2. Differential modules and differential powers of ideals

In the following, let K be a perfect field and let $P = K[X_0, \ldots, X_n]$ be a polynomial ring in indeterminates $X_0, \ldots, X_n$ over K. Let R be a finitely generated K-algebra of the form $R = P/I$, where I is an ideal in P. The image of X_i in R is denoted by x_i for $i = 0, \ldots, n$. We denote by μ the multiplication map $\mu : R \otimes_K R \to R$ given by $\mu(f \otimes g) = fg$ for $f, g \in R$. This is a K-algebra homomorphism and its kernel $\Delta_R := \mathrm{Ker}(\mu)$ is an ideal of the algebra $R \otimes_K R$. In particular, the ideal Δ_R is generated by $\{1 \otimes x_i - x_i \otimes 1 \mid i \in \{0, \ldots, n\}\}$.

Definition 2.1. Let $R = P/I$ be a finitely generated K-algebra.

(a) We call the finitely generated R-module $\Omega^1_{R/K} = \Delta_R/\Delta_R^2$ the *Kähler differential module* (or the *module of Kähler differentials*) of R/K.

(b) For every $m \geq 0$, we call the m-th exterior power $\Omega^m_{R/K} := \Lambda^m_R \Omega^1_{R/K}$ the *m-form Kähler differential module* (or the *module of Kähler differential m-forms*) of R/K.

(c) We call the exterior algebra $\Omega^\bullet_{R/K} := \bigoplus_{m\geq 0} \Omega^m_{R/K}$ the *Kähler differential algebra* of R/K.

Unless explicitly stated otherwise, all algebras will be K-algebras. So, we usually simplify the notation and write Ω^m_R instead of $\Omega^m_{R/K}$ for all $m \geq 0$, as well as $\Omega^\bullet_R$ instead of $\Omega^\bullet_{R/K}$. Notice that $\Omega^0_R = R$ and if I is a homogeneous ideal of P, then Ω^m_R is a graded R-module for every $m \geq 0$. The Kähler differential module of R/K comes together with the *universal derivation* $d_R : R \to \Omega^1_R$ which is defined by $d_R(f) = f \otimes 1 - 1 \otimes f + \Delta^2_R$ for $f \in R$. We write d for d_R, if no confusion can arise.

As in [Kunz 1986, Section 1], the set of all K-linear derivations $\partial : R \to R$ is an R-module and is denoted by $\mathrm{Der}_K(R)$. By [Kunz 1986, Proposition 1.23], there is an isomorphism

$$\mathrm{Hom}_R(\Omega^1_R, R) \to \mathrm{Der}_K(R), \quad \varphi \mapsto \varphi \circ d.$$

Also, both $\mathrm{Der}_K(R)$ and $\mathrm{Hom}_R(R, R)$ are submodules of $\mathrm{Hom}_K(R, R)$. The latter, $\mathrm{Hom}_K(R, R)$, is the ring of endomorphisms of R/K, with operations of sum and composition of endomorphisms.

If $\delta, \theta \in \mathrm{Hom}_K(R, R)$, then their commutator is defined as

$$[\delta, \theta] := \delta \circ \theta - \theta \circ \delta.$$

Obviously, $[\delta, \theta] \in \mathrm{Hom}_K(R, R)$. Every element in $\mathrm{Hom}_R(R, R)$ is simply a multiplication map $\mu_f : R \to R$ for some $f \in R$, where $\mu_f(g) = fg$ for all $g \in R$. We write $[\delta, f]$ for $[\delta, \mu_f]$, so that

$$[\delta, f] = \delta \circ \mu_f - \mu_f \circ \delta \quad \text{and} \quad [f, g] = 0$$

for all $f, g \in R$ and $\delta \in \mathrm{Hom}_K(R, R)$.

Based on [Heyneman and Sweedler 1969, Section 2; Coutinho 1995, Chapter 3, Section 1], we have the following notion.

Definition 2.2. Let $R = P/I$ be a finitely generated K-algebra, and let $m \in \mathbb{N}$. The *K-linear differential operators of R of order at most m*, $D^m_R \subseteq \mathrm{Hom}_K(R, R)$, are defined inductively as

(i) $D^0_R = \mathrm{Hom}_R(R, R)$;

(ii) $D^m_R = \{\delta \in \mathrm{Hom}_K(R, R) \mid [\delta, f] \in D^{m-1}_R \text{ for every } f \in R\}$.

The *ring of K-linear differential operators* of R is defined by $D_R = \bigcup_{m\in\mathbb{N}} D^m_R$.

Note that the ring structure on D_R is also given by composition and satisfies $D_R^m \subseteq D_R^{m+1}$ for $m \geq 0$ and $D_R^m \circ D_R^k \subseteq D_R^{m+k}$ (see, e.g., [Coutinho 1995, Chapter 3, Proposition 1.2]). The ring D_R is a noncommutative ring, and it is in general not Noetherian (see [Bernstein et al. 1972]).

When $R = P = K[X_0, \ldots, X_n]$ and K is of characteristic zero, the ring D_P is known as the Weyl algebra of P and it has a concrete description below (see, for instance, [Björk 1979, Chapter I, Proposition 1.2; Hart 1983, Theorem 2]). We refer to [Smith 1986] for further information on D_P in positive characteristic. In the following, $\partial_0, \ldots, \partial_n$ are the partial derivatives defined by $\partial_i(f) = \partial f/\partial X_i$ for $f \in P$. For $\alpha = (\alpha_0, \ldots, \alpha_n) \in \mathbb{N}^{n+1}$, by X^α we mean the term $X_0^{\alpha_0} \cdots X_n^{\alpha_n}$ and its degree is $|\alpha| = \alpha_0 + \cdots + \alpha_n$, and similarly, for $\beta = (\beta_0, \ldots, \beta_n) \in \mathbb{N}^{n+1}$, by ∂^β we denote a ∂-term $\partial_0^{\beta_0} \cdots \partial_n^{\beta_n}$.

Proposition 2.3. *If K is a field of characteristic zero, then*

$$D_P = P[\mathrm{Der}_K(P)] = K[X_0, \ldots, X_n, \partial_0, \ldots, \partial_n]$$

and the set $\{X^\alpha \partial^\beta \mid \alpha, \beta \in \mathbb{N}^{n+1}\}$ is a K-basis of D_P. In particular, any $\delta \in D_P^m$ can be uniquely written as

$$\delta = \sum_{|\alpha| \leq m} f_\alpha \partial^\alpha \quad (f_\alpha \in P).$$

For the finitely generated K-algebra $R = P/I$, there is an $(R \otimes_K R)$-module structure on $\mathrm{Hom}_K(R, R)$ given by the rule

$$((f \otimes g) \cdot \theta)(h) = f\theta(gh)$$

for $f, g, h \in R$ and $\theta \in \mathrm{Hom}_K(R, R)$. Via this module structure, we have the following characterization of differential operators (see [Heyneman and Sweedler 1969, Lemma 2.2.1 and Proposition 2.2.3]).

Proposition 2.4. *The set D_R^m of K-linear differential operators of order $\leq m$ is an $(R \otimes_K R)$-submodule (and so an R-submodule) of $\mathrm{Hom}_K(R, R)$. Moreover, $\delta \in D_R^m$ if and only if $\Delta_R \cdot \delta \subseteq D_R^{m-1}$ if and only if $\Delta_R^{m+1} \cdot \delta = 0$.*

The following notion of differential powers of an ideal in R was introduced in [Dao et al. 2018, Definition 2.2].

Definition 2.5. Let J be an ideal of R and let m be a positive integer. The *m-th K-linear differential power* of J is defined by

$$J^{\langle m \rangle} := \{f \in R \mid \delta(f) \in J \text{ for all } \delta \in D_R^{m-1}\}.$$

It is straightforward to verify that $J^{\langle m \rangle}$ is a subideal of J in R and if $J_1 \subseteq J_2$ then $J_1^{\langle m \rangle} \subseteq J_2^{\langle m \rangle}$. The next proposition gathers some additional properties of differential powers from [Dao et al. 2018, Section 2].

Proposition 2.6. *Let J, J_1, J_2 be ideals of the finitely generated K-algebra R, and let m be a positive integer. Then*

(a) $J^m \subseteq J^{\langle m\rangle} \subseteq J$;

(b) $(J_1 \cap J_2)^{\langle m\rangle} = J_1^{\langle m\rangle} \cap J_2^{\langle m\rangle}$;

(c) *If $R = P$ and J is a maximal ideal of P, then $J^{\langle m\rangle} = J^m$.*

In what follows, let $\mathbb{X}$ be a 0-dimensional scheme in the n-dimensional projective space $\mathbb{P}^n_K$ over the perfect field K, and let $I_{\mathbb{X}}$ be the saturated homogeneous vanishing ideal of $\mathbb{X}$ in $P = K[X_0, \ldots, X_n]$. Then the homogeneous coordinate ring of $\mathbb{X}$ is $R = P/I_{\mathbb{X}}$. The ring R is a 1-dimensional, standard graded, Cohen–Macaulay K-algebra. The set of closed points of $\mathbb{X}$ is known as the support of $\mathbb{X}$ and is denoted by $\mathrm{Supp}(\mathbb{X}) = \{p_1, \ldots, p_s\}$. To each point $p_j \in \mathrm{Supp}(\mathbb{X})$ we associate the homogeneous prime ideal $\mathfrak{P}_j \subseteq P$, whose image $\mathfrak{p}_j$ in R is a minimal homogeneous prime ideal. In particular, the set of minimal homogeneous prime ideals of R is $\{\mathfrak{p}_1, \ldots, \mathfrak{p}_s\}$. The local ring of $\mathbb{X}$ at p_j is then the homogeneous localization $R_{(\mathfrak{p}_j)}$ of R at $\mathfrak{p}_j$.

Assumption 2.7. No point of the support of $\mathbb{X}$ lies on the hyperplane $\mathcal{Z}(X_0)$.

When this assumption fails, it suffices to perform a homogeneous linear change of coordinates (potentially requiring a small base field extension if K is finite) without affecting any of the results in this paper. Under this assumption, the elements x_0 and $x_0 - 1$ are nonzerodivisors of R, the ring $\bar{R} = R/\langle x_0\rangle$ is a 0-dimensional local graded K-algebra, especially, $\bar{R}$ is a finite-dimensional K-vector space. Moreover, $\mathbb{X}$ is contained in the affine space $D_+(X_0) \cong \mathbb{A}^n_K$. The affine coordinate ring of $\mathbb{X}$, viewed as a subscheme of $\mathbb{A}^n_K$, is given by $S = R/\langle x_0 - 1\rangle \cong A/J_{\mathbb{X}}$, where $A = K[X_1, \ldots, X_n]$ and $J_{\mathbb{X}} = I_{\mathbb{X}}^{\mathrm{deh}}$ is the dehomogenization of $I_{\mathbb{X}}$ with respect to X_0. The ring S is a 0-dimensional affine K-algebra and hence a finite-dimensional K-vector space.

For $j \in \{1, \ldots, s\}$, let I_j denote the $\mathfrak{P}_j$-primary component of $I_{\mathbb{X}}$, and define $\mathfrak{Q}_j := I_j^{\mathrm{deh}} \subseteq A$ and $\mathcal{O}_j := A/\mathfrak{Q}_j$. According to [Kreuzer and Robbiano 2005, Propositions 4.3.10 and 4.3.12], the homogeneous primary decomposition $I_{\mathbb{X}} = I_1 \cap \cdots \cap I_s$ induces the primary decomposition $J_{\mathbb{X}} = \mathfrak{Q}_1 \cap \cdots \cap \mathfrak{Q}_s$. Consequently, the Chinese remainder theorem yields

$$S \cong A/\mathfrak{Q}_1 \times \cdots \times A/\mathfrak{Q}_s = \mathcal{O}_1 \times \cdots \times \mathcal{O}_s.$$

Furthermore, the local ring of $\mathbb{X}$ at p_i satisfies

$$R_{(\mathfrak{p}_j)} \cong \mathcal{O}_j \quad \text{for each } j \in \{1, \ldots, s\},$$

as addressed in [Hartshorne 1977, Chapter I, Section 3; Kunz 2013, Chapter III, Section 2].

Corollary 2.8. *Let $\mathbb{X}$ be a 0-dimensional scheme in $\mathbb{P}^n_K$ with $\operatorname{Supp}(\mathbb{X}) = \{p_1, \ldots, p_s\}$, and let I_j be the $\mathfrak{P}_j$-primary component of $I_{\mathbb{X}}$ corresponding to the point p_j for $j \in \{1, \ldots, s\}$. For $m \geq 1$, we have*

$$I_{\mathbb{X}}^{\langle m \rangle} = I_1^{\langle m \rangle} \cap \cdots \cap I_s^{\langle m \rangle}$$

and $\sqrt{I_j^{\langle m \rangle}} = \sqrt{I_j} = \mathfrak{P}_j$ for $j \in \{1, \ldots, s\}$.

Proof. This follows from the primary decomposition $I_{\mathbb{X}} = I_1 \cap \cdots \cap I_s$ and by Proposition 2.6(a)–(b). □

The m-form Kähler differential module Ω_R^m for the scheme $\mathbb{X}$ is a finitely generated graded R-module and the Kähler differential algebra for the scheme $\mathbb{X}$ is the exterior algebra $\Omega_R^\bullet = \bigoplus_{m \geq 0} \Omega_R^m$.

Proposition 2.9 [Kreuzer et al. 2025, Proposition 4.2]. *Let $\mathbb{X}$ be a 0-dimensional scheme in $\mathbb{P}^n_K$ with homogeneous coordinate ring $R = P/I_{\mathbb{X}}$.*

(a) *If $m \geq n+2$, then $\Omega_R^m = \langle 0 \rangle$.*

(b) *For $m = 0, \ldots, n+1$, the P-module Ω_P^m is a free P-module of rank $\binom{n+1}{m}$ with basis*

$$\{dX_{i_1} \wedge \cdots \wedge dX_{i_m} \mid 0 \leq i_1 < \cdots < i_m \leq n\}.$$

(c) *For $m \geq 1$, the module of Kähler differential m-forms satisfies*

$$\Omega_R^m \cong \Omega_P^m / (I_{\mathbb{X}} \Omega_P^m + dI_{\mathbb{X}} \wedge \Omega_P^{m-1}),$$

where $dI_{\mathbb{X}}$ is the P-submodule of Ω_P^1 generated by $\{df \mid f \in I_{\mathbb{X}}\}$.

(d) *Let $I_{\mathbb{X}} = \langle f_1, \ldots, f_k \rangle$ for some $f_1, \ldots, f_k \in P$. Then*

$$dI_{\mathbb{X}} = \langle df_1, \ldots, df_k \rangle + \langle f_i dX_j \mid i \in \{1, \ldots, k\},\ j \in \{0, \ldots, n\} \rangle.$$

Definition 2.10. Let $M = \bigoplus_{i \in \mathbb{Z}} M_i$ be a finitely generated graded R-module.

(a) The map $\mathrm{HF}_M : \mathbb{Z} \to \mathbb{Z}$ given by $\mathrm{HF}_M(i) = \dim_K(M_i)$ for all $i \in \mathbb{Z}$ is called the *Hilbert function* of M. The *Hilbert polynomial* of M, denoted by HP_M, always exists. If HP_M is a constant polynomial, it is also written as $\mathrm{HP}(M)$.

(b) The number $\mathrm{ri}(M) = \min\{i \in \mathbb{Z} \mid \mathrm{HF}_M(j) = \mathrm{HP}_M(j) \text{ for all } j \geq i\}$ is called the *regularity index* of M.

Remark 2.11. The function $\mathrm{HF}_{\mathbb{X}}(i) := \mathrm{HF}_R(i)$ for all $i \in \mathbb{Z}$ is called the *Hilbert function of* $\mathbb{X}$ and the number $r_{\mathbb{X}} := \mathrm{ri}(R)$ is called the *regularity index of* $\mathbb{X}$. We have

$$1 = \mathrm{HF}_{\mathbb{X}}(0) < \mathrm{HF}_{\mathbb{X}}(1) < \cdots < \mathrm{HF}_{\mathbb{X}}(r_{\mathbb{X}}) = \mathrm{HF}_{\mathbb{X}}(r_{\mathbb{X}} + 1) = \cdots = \deg(\mathbb{X}),$$

where $\deg(\mathbb{X})$ is the degree of $\mathbb{X}$, i.e., the constant Hilbert polynomial of R.

It is worth noting that $\deg(\mathbb{X}) = \dim_K(S)$. For $i \geq r_{\mathbb{X}}$, the map $\varepsilon_i : R_i \to S$ defined by $f \mapsto f^{\text{deh}}$ is an isomorphism of K-vector spaces (see [Kreuzer et al. 2018]). This structure map induces a valuable connection between the modules Ω^m_R and Ω^m_S (see [Kreuzer et al. 2025]), ultimately leading to the proof of the following result.

Proposition 2.12 [Kreuzer et al. 2025, Proposition 5.4]. *Let $\mathbb{X}$ be a 0-dimensional scheme in $\mathbb{P}^n$ with homogeneous coordinate ring R and affine coordinate ring $S = R/\langle x_0 - 1\rangle$, and let $m \geq 1$. Then*

$$\mathrm{HP}(\Omega^m_R) = \dim_K(\Omega^m_S) + \dim_K(\Omega^{m-1}_S) \quad \textit{and} \quad \mathrm{ri}(\Omega^m_R) \leq 2r_{\mathbb{X}} + m.$$

In addition, if $\mathrm{char}(K) = 0$ or if $\mathrm{char}(K) = p > 0$ and $p \nmid (2r_{\mathbb{X}} + n + 1)$, then $\mathrm{ri}(\Omega^{n+1}_R) \leq 2r_{\mathbb{X}} + n$.

3. The canonical exact sequence for Kähler differentials

As in the previous section, let K be a perfect field, let $\mathbb{X}$ be a 0-dimensional scheme in $\mathbb{P}^n_K$ with homogeneous vanishing ideal $I_{\mathbb{X}}$ and homogeneous coordinate ring $R = P/I_{\mathbb{X}}$. Further, let $\mathrm{Supp}(\mathbb{X}) = \{p_1, \ldots, p_s\}$, and let $\mathfrak{P}_j$ be the homogeneous prime ideal of the point p_j for every $j \in \{1, \ldots, s\}$. In the following, suppose that Assumption 2.7 holds for $\mathbb{X}$. Our goal in this section is to give an answer to Question 1.1 posted in Section 1, regarding the existence of the canonical exact sequence for the Kähler differential module Ω^1_R of $\mathbb{X}$. To achieve this goal, we will first examine the second differential power of the homogeneous vanishing ideal $I_{\mathbb{X}}$ of the scheme $\mathbb{X}$.

Let $\mathfrak{M} = \langle X_0, \ldots, X_n\rangle \subseteq P$ and let D_P denote the ring of K-linear differential operators on P. The second differential power $I^{\langle 2\rangle}_{\mathbb{X}}$ of $I_{\mathbb{X}}$ is given by

$$I^{\langle 2\rangle}_{\mathbb{X}} = \{f \in P \mid \delta(f) \in I_{\mathbb{X}} \text{ for all } \delta \in D^1_P\}.$$

This ideal is also written as $I^{\mathrm{diff}}_{\mathbb{X}}$. For $i = 0, \ldots, n$, we will use ∂_i to represent $\partial/\partial X_i$.

Proposition 3.1. *Suppose that $\mathbb{X}$ is a 0-dimensional scheme in $\mathbb{P}^n_K$ with support $\mathrm{Supp}(\mathbb{X}) = \{p_1, \ldots, p_s\}$. For $j \in \{1, \ldots, s\}$, let $\mathfrak{P}_j$ denote the homogeneous prime ideal of p_j and let I_j be the corresponding $\mathfrak{P}_j$-primary component of $I_{\mathbb{X}}$.*

(a) *The ideal $I^{\mathrm{diff}}_{\mathbb{X}}$ is a saturated homogeneous ideal of P and*

$$I^{\mathrm{diff}}_{\mathbb{X}} = \{f \in I_{\mathbb{X}} \mid \partial_i f \in I_{\mathbb{X}} \textit{ for all } i \in \{0, \ldots, n\}\} \subseteq P. \tag{3}$$

(b) *We have $I^{\mathrm{diff}}_{\mathbb{X}} = I^{\mathrm{diff}}_1 \cap \cdots \cap I^{\mathrm{diff}}_s$, and $\sqrt{I^{\mathrm{diff}}_j} = \sqrt{I_j} = \mathfrak{P}_j$ for $j = 1, \ldots, s$.*

Proof. (a) According to [Coutinho 1995, Chapter 3, Proposition 1.1], we first observe that $D^1_P = P + \mathrm{Der}_K(P)$, where the elements of P are identified with their multiplication operators of order zero. By [Kunz 1986, Example 1.22(c)],

$\mathrm{Der}_K(P)$ is generated as a P-module by the partial derivatives $\partial_0, \dots, \partial_n$. They form in fact a basis of $\mathrm{Der}_K(P)$ since if $\sum_{i=0}^n f_i \partial_i = 0$ for $f_0, \dots, f_n \in P$ then $f_k = \sum_{i=0}^n f_i \partial_i(X_k) = 0$ for $k = 0, \dots, n$. Thus we have $D_P^1 = P \oplus P\partial_0 \oplus \cdots \oplus P\partial_n$. Consequently, (3) is true for $I_{\mathbb{X}}^{\mathrm{diff}}$.

Now let $g = g_0 + \cdots + g_d \in I_{\mathbb{X}}^{\mathrm{diff}}$ be a nonzero polynomial of degree d, where $g_k \in P_k$ is the k-th homogeneous component of g. For each $i \in \{0, \dots, n\}$, we have $\partial_i(g) = \partial_i(g_0) + \cdots + \partial_i(g_d) \in I_{\mathbb{X}}$. Since $I_{\mathbb{X}}$ is homogeneous, $\partial_i(g_j) \in I_{\mathbb{X}}$ for $j \in \{0, \dots, d\}$. Thus $g_j \in I_{\mathbb{X}}^{\mathrm{diff}}$ for $j \in \{0, \dots, d\}$, and hence $I_{\mathbb{X}}^{\mathrm{diff}}$ is homogeneous.

Next, we will verify that $I_{\mathbb{X}}^{\mathrm{diff}}$ is saturated. It suffices to prove $I_{\mathbb{X}}^{\mathrm{diff}} = I_{\mathbb{X}}^{\mathrm{diff}} :_P \mathfrak{M}$. Let $f \in I_{\mathbb{X}}^{\mathrm{diff}} :_P \mathfrak{M}$ and $i \in \{0, \dots, n\}$. Then $X_k f \in I_{\mathbb{X}}^{\mathrm{diff}}$ for all $k = 0, \dots, n$, and so $\partial_i(X_k f) \in I_{\mathbb{X}}$. Observe that $\partial_i(X_0 f) \in I_{\mathbb{X}}$ and $X_0 f \in I_{\mathbb{X}}$. In R, we have $x_0 \bar{f} = 0$. Under Assumption 2.7, x_0 is a nonzerodivisor of R, and this yields $\bar{f} = 0$, in other words, $f \in I_{\mathbb{X}}$. Consequently, we have

$$X_0 \partial_i(f) = \partial_i(X_0 f) - f\partial_i(X_0) \in I_{\mathbb{X}}$$

for all $i = 0, \dots, n$. This also implies $\partial_i(f) \in I_{\mathbb{X}}$ for all $i = 0, \dots, n$. Hence $f \in I_{\mathbb{X}}^{\mathrm{diff}}$.

(b) This follows from Corollary 2.8. □

When the scheme $\mathbb{X}$ consists of a single monomial point, the homogeneous ideal $I_{\mathbb{X}}^{\mathrm{diff}}$ can be explicitly described by the following theorem. Here the set of all terms in P is given by $\mathbb{T}^{n+1} = \{X_0^{\alpha_0} \cdots X_n^{\alpha_n} \mid (\alpha_0, \dots, \alpha_n) \in \mathbb{N}^{n+1}\}$.

Theorem 3.2. *Let K be a perfect field of characteristic $p \geq 0$. Assume that the homogeneous vanishing ideal $I_{\mathbb{X}}$ of the scheme $\mathbb{X}$ is a proper monomial ideal of P, minimally generated by terms $t_1, \dots, t_r \in \mathbb{T}^{n+1}$.*

(a) *The ideal $I_{\mathbb{X}}^{\mathrm{diff}}$ is a monomial subideal of $I_{\mathbb{X}}$.*

(b) *For $1 \leq j \leq r$, we write $t_j = X_0^{\alpha_{0j}} \cdots X_n^{\alpha_{nj}}$, and for $2 \leq k \leq n+1$ and $1 \leq j_1 \leq \cdots \leq j_k \leq r$, we define $t_{j_1,\dots,j_k} := X_0^{\beta_0} \cdots X_n^{\beta_n}$, where, for $i = 0, \dots, n$,*

$$\beta_i := \begin{cases} 0 & \text{if } \alpha_{ij_1} = \cdots = \alpha_{ij_k} = 0, \\ \alpha_{ij_1} & \text{if } \alpha_{ij_1} = \cdots = \alpha_{ij_k} > 0 \text{ and } p \mid \alpha_{ij_1}, \\ \alpha_{ij_1} + 1 & \text{if } \alpha_{ij_1} = \cdots = \alpha_{ij_k} > 0 \text{ and } p \nmid \alpha_{ij_1}, \\ \max\{\alpha_{ij_1}, \dots, \alpha_{ij_k}\} & \text{otherwise.} \end{cases} \tag{4}$$

Then

$$I_{\mathbb{X}}^{\mathrm{diff}} = \langle t_{j_1,\dots,j_k} \mid 1 \leq j_1 \leq \cdots \leq j_k \leq r;\ 2 \leq k \leq n+1 \rangle.$$

Proof. Note that $I_{\mathbb{X}}^{\mathrm{diff}} \subseteq I_{\mathbb{X}}$ by Proposition 3.1(a), and that a nonconstant term $t = X_0^{\alpha_0} \cdots X_n^{\alpha_n} \in I_{\mathbb{X}}$ belongs to $I_{\mathbb{X}}^{\mathrm{diff}}$ if and only if $\alpha_i t / X_i \in I_{\mathbb{X}}$ for $i = 0, \dots, n$.

(a) Let $f \in I_{\mathbb{X}}^{\mathrm{diff}}$ be a nonzero homogeneous polynomial of degree d and write $f = c_1 t_1' + \cdots + c_m t_m'$ with $c_1, \ldots, c_m \in K \setminus \{0\}$ and $t_1', \ldots, t_m' \in \mathbb{T}^{n+1}$. It suffices to show that $t_j' \in I_{\mathbb{X}}^{\mathrm{diff}}$ for $j = 1, \ldots, m$. Let $i \in \{0, \ldots, n\}$ and write $t_j' = X_0^{\gamma_{0j}} \cdots X_n^{\gamma_{nj}}$ with $(\gamma_{0j}, \ldots, \gamma_{nj}) \in \mathbb{N}^{n+1}$ and $\gamma_{0j} + \cdots + \gamma_{nj} = d$ for $j = 1, \ldots, m$. When $\gamma_{ij} = 0$ in K, it directly follows that $\partial_i(t_j') = 0 \in I_{\mathbb{X}}$. Now, consider the case, where $\gamma_{ij} \neq 0$ in K. In this case, for $1 \leq j < k \leq m$, we have $t_j' \neq t_k'$ and

$$\partial_i(t_j') = \frac{\gamma_{ij} t_j'}{X_i} \neq \frac{\gamma_{ik} t_k'}{X_i} = \partial_i(t_k').$$

So, the element t_j'/X_i appears in the expansion of $\partial_i(f) = c_1 \partial_i(t_1') + \cdots + c_m \partial_i(t_m')$ with a nonzero coefficient. Since $I_{\mathbb{X}}$ is a monomial ideal and $\partial_i(f) \in I_{\mathbb{X}}$, this implies

$$\frac{\partial_i(t_j')}{\gamma_{ij}} = \frac{t_j'}{X_i} \in I_{\mathbb{X}}.$$

Consequently, $\partial_i(t_j') \in I_{\mathbb{X}}$ for every $j = 1, \ldots, m$ and for every $i = 0, \ldots, n$. Hence, we find $t_1', \ldots, t_m' \in I_{\mathbb{X}}^{\mathrm{diff}}$.

(b) Let

$$\Gamma = \{t_{j_1,\ldots,j_k} \mid 1 \leq j_1 \leq \cdots \leq j_k \leq r;\ 2 \leq k \leq n+1\},$$

where $t_{j_1,\ldots,j_k} = X_0^{\beta_0} \cdots X_n^{\beta_n} \in \Gamma$ has the exponents β_i as given in (4). For such a term $t_{j_1,\ldots,j_k} \in \Gamma$, it is easily seen that

$$\beta_i \geq \max\{\alpha_{ij_1}, \ldots, \alpha_{ij_k}\}$$

for all $i = 0, \ldots, n$, and so $t_{j_1,\ldots,j_k} \in I_{\mathbb{X}}$. In particular, $\Gamma \subseteq I_{\mathbb{X}}$. Moreover, either $p \,|\, \beta_i$ or both $p \nmid \beta_i$ and $\beta_i \geq \min\{\alpha_{ij_1}, \ldots, \alpha_{ij_k}\} + 1$. If $p \,|\, \beta_i$ then $\partial_i(t_{j_1,\ldots,j_k}) = 0 \in I_{\mathbb{X}}$. When $p \nmid \beta_i$, we may assume without loss of generality that $\alpha_{ij_1} = \min\{\alpha_{ij_1}, \ldots, \alpha_{ij_k}\}$, and we get $\beta_i \geq \alpha_{ij_1} + 1$. Then

$$\partial_i(t_{j_1,\ldots,j_k}) = \frac{\beta_i t_{j_1,\ldots,j_k}}{X_i} = \beta_i X_0^{\beta_0} \cdots X_{i-1}^{\beta_{i-1}} X_i^{\beta_i - 1} X_{i+1}^{\beta_{i+1}} \cdots X_n^{\beta_n}$$

is a multiple of t_{j_1}, and hence it belongs to $I_{\mathbb{X}}$. Thus, $\partial_i(t_{j_1,\ldots,j_k}) \in I_{\mathbb{X}}$ for all $i \in \{0, \ldots, n\}$, which means $t_{j_1,\ldots,j_k} \in I_{\mathbb{X}}^{\mathrm{diff}}$. Therefore, $\langle \Gamma \rangle \subseteq I_{\mathbb{X}}^{\mathrm{diff}}$.

For the "$\supseteq$" inclusion, let $u = X_0^{\gamma_0} \cdots X_n^{\gamma_n}$ with $(\gamma_0, \ldots, \gamma_n) \in \mathbb{N}^{n+1}$ be a term in $I_{\mathbb{X}}^{\mathrm{diff}}$. By (3), we have $u \subseteq I_{\mathbb{X}}$ and $\partial_i(u) \in I_{\mathbb{X}}$ for $i = 0, \ldots, n$. Since $\{t_1, \ldots, t_r\}$ is the minimal monomial system of generators of $I_{\mathbb{X}}$, there exists $j \in \{1, \ldots, r\}$ such that $t_j \,|\, u$. Let T_u be the set of all terms in $\{t_1, \ldots, t_r\}$ which divide u. After renumbering the indices of t_j, we may assume that $T_u = \{t_1, \ldots, t_s\}$ with $s \leq r$. Obviously, we have

$$\gamma_i \geq \max\{\alpha_{ij} \mid j = 1, \ldots, s\} \quad (i = 0, \ldots, n). \tag{5}$$

Next, we construct a specific term $t_{j_1,\dots,j_k} \in \Gamma$ that divides u. Set $k = n+1$. For $i \in \{0,\dots,n\}$, if $\partial_i(u) = 0$ then we choose $j_{i+1} = 1$. If $\partial_i(u) \neq 0$, then $\partial_i(u) \in I_{\mathbb{X}}$ and $\partial_i(u)$ divides u, and so we can choose an index $j_{i+1} \in \{1,\dots,s\}$ such that $t_{j_{i+1}} \,|\, \partial_i(u)$. Now we consider the term $t_{j_1,\dots,j_{n+1}} = X_0^{\beta_0} \cdots X_n^{\beta_n}$ formed by these chosen indices $j_1,\dots,j_{n+1}$ according to (4). In order to show that $t_{j_1,\dots,j_{n+1}} \,|\, u$, we need demonstrate that $\gamma_i \geq \beta_i$ for all $i = 0,\dots,n$.

From (4) and (5), $\gamma_i \geq \beta_i$ is true, except in the case where $\alpha_{ij_1} = \cdots = \alpha_{ij_{n+1}} > 0$ and $p \nmid \alpha_{ij_1}$. In this exceptional case, we also have $\beta_i = \alpha_{ij_1} + 1$. If $\partial_i(u) = 0$, then the characteristic of the field K must be $p > 0$ with $p \,|\, \gamma_i$. We know $\gamma_i \geq \alpha_{ij_1} > 0$ and $p \nmid \alpha_{ij_1}$. It follows that $\gamma_i \geq \alpha_{ij_1} + 1 = \beta_i$. If $\partial_i(u) \neq 0$, by the choice of j_i, we find $t_{j_{i+1}} \,|\, \partial_i(u)$. This implies $\gamma_i - 1 \geq \alpha_{ij_{i+1}}$, and so $\gamma_i \geq \alpha_{ij_{i+1}} + 1 = \alpha_{ij_1} + 1 = \beta_i$.

Altogether, we have $\gamma_i \geq \beta_i$ for every $i \in \{0,\dots,n\}$, and hence the constructed term $t_{j_1,\dots,j_{n+1}} \in \Gamma$ divides u. It follows that $u \in \langle \Gamma \rangle$. Therefore, $\langle \Gamma \rangle \supseteq I_{\mathbb{X}}^{\mathrm{diff}}$. □

In view of (4), the generating system for $I_{\mathbb{X}}^{\mathrm{diff}}$ can be restricted to terms $t_{j_1,\dots,j_k}$ whose indices satisfy $j_1 \leq j_2$ for $k = 2$ and $1 \leq j_1 < \cdots < j_k \leq r$ for $k \geq 3$.

Example 3.3. In $\mathbb{P}^2_K$, let $\mathbb{X}$ be the 0-dimensional scheme defined by the monomial ideal $I_{\mathbb{X}} = \langle X_1, X_2 \rangle^5 \subseteq P = K[X_0, X_1, X_2]$. The minimal generators of $I_{\mathbb{X}}$ are $t_1 = X_2^5$, $t_2 = X_1X_2^4$, $t_3 = X_1^2X_2^3$, $t_4 = X_1^3X_2^2$, $t_5 = X_1^4X_2$, and $t_6 = X_1^5$.

(a) *Suppose* $\mathrm{char}(K) = 0$: Following (4), we compute various terms $t_{1,1} = X_2^6$, $t_{2,2} = X_1^2X_2^5$, $t_{1,5} = X_1^5X_2$, $t_{1,2,3} = X_1^5X_2^4$, $t_{3,4,5} = X_1^4X_2^3$, etc. Consider the element $u = X_1^3X_2^4 \in I_{\mathbb{X}}^{\mathrm{diff}}$. As in the proof of Theorem 3.2(b), we have $T_u = \{t_2, t_3, t_4\}$. By choosing the indices $j_1 = 2$, $j_2 = 2$, $j_3 = 3$, we find $t_{2,2,3} = t_{2,3} = X_1^2X_2^4$, which divides u. A computation using `ApCoCoA` gives a generating system of $I_{\mathbb{X}}^{\mathrm{diff}}$ consisting of 56 terms. In particular, we get $I_{\mathbb{X}}^{\mathrm{diff}} = \langle X_1, X_2 \rangle^6$.

(b) *Suppose* $\mathrm{char}(K) = 5$: Observe that $X_1^5, X_2^5 \in I_{\mathbb{X}}^{\mathrm{diff}}$, since their partial derivatives are zero in K. A minimal system of generators of $I_{\mathbb{X}}^{\mathrm{diff}}$ is calculated by $\{X_1^5, X_1^2X_2^4, X_1^3X_2^3, X_1^4X_2^2, X_2^5\}$. In this case, we have $I_{\mathbb{X}}^{\mathrm{diff}} \neq \langle X_1, X_2 \rangle^6$. It is worth noting that if $\mathrm{char}(K) \neq 5$ then one can also verify that $I_{\mathbb{X}}^{\mathrm{diff}} = \langle X_1, X_2 \rangle^6$.

Remark 3.4. The proof of Theorem 3.2 remains valid for any arbitrary nonzero monomial ideal I in P. Specifically, a finite monomial system of generators for I^{diff} can be explicitly constructed from the minimal generators of I using the formulas established above. If I is minimally generated by r terms in $\mathbb{T}^{n+1}$, then the complexity of computing I^{diff} using Theorem 3.2 is given by $O(n^2r^{n+1})$. In addition, applying the above method to other special classes of ideals in P is an intriguing direction for the authors. Nevertheless, extending these results beyond the monomial case will likely involve further constraints or the theory of Gröbner bases.

As a direct consequence of Theorem 3.2, we obtain the following corollary. The proof of part (a) for the case $\operatorname{char}(K) = 0$ is discussed in [Dao et al. 2018, Exercise 2.8], and the straightforward but tedious proof of the remaining cases is left to the interested reader.

Corollary 3.5. *Let K be a perfect field of characteristic $p \geq 0$. Suppose that $\mathbb{X}$ contains exactly one point $p_1 = (1 : 0 : \cdots : 0)$.*

(a) *If $I_{\mathbb{X}} = \langle X_1, \ldots, X_n \rangle^m$ for some $m \geq 1$, then*

$$I_{\mathbb{X}}^{\text{diff}} = \begin{cases} \langle X_1, \ldots, X_n \rangle^{m+1} & \text{if } p \nmid m, \\ \langle X_1, \ldots, X_n \rangle^{m+1} + \left\langle X_1^{p\alpha_1} \cdots X_n^{p\alpha_n} \mid \sum_{i=1}^n \alpha_i = \frac{m}{p} \right\rangle & \text{if } p \mid m. \end{cases}$$

(b) *If $I_{\mathbb{X}} = \langle X_1^{k_1}, \ldots, X_n^{k_n} \rangle$ for some integers $k_1, \ldots, k_n \geq 1$, then*

$$I_{\mathbb{X}}^{\text{diff}} = \begin{cases} \langle X_1^{k_1+1}, \ldots, X_n^{k_n+1} \rangle + \langle X_i^{k_i} X_j^{k_j} \mid 1 \leq i < j \leq n \rangle & \text{if } p = 0, \\ \langle X_i^{k_i} \mid i = 1, \ldots, n;\ p \mid k_i \rangle + \langle X_i^{k_i+1} \mid i = 1, \ldots, n;\ p \nmid k_i \rangle & \\ \qquad + \langle X_i^{k_i} X_j^{k_j} \mid 1 \leq i < j \leq n \rangle & \text{if } p > 0. \end{cases}$$

Remark 3.6. Let $p_1, \ldots, p_s$ be distinct K-rational points in $\mathbb{P}_K^n$, and let $I_{p_j} \subseteq P$ be the homogeneous vanishing ideal of p_j for $j = 1, \ldots, s$. Given positive integers $m_1, \ldots, m_s$, the saturated homogeneous ideal $I_{\mathbb{X}} := I_{p_1}^{m_1} \cap \cdots \cap I_{p_s}^{m_s}$ of P defines a fat point scheme $\mathbb{X} = m_1 p_1 + \cdots + m_s p_s$ in $\mathbb{P}_K^n$. The number m_j is the multiplicity of the point p_j for $j = 1, \ldots, s$. As in [Kreuzer et al. 2018, Definition 4.1], the first fattening of this scheme is the fat point scheme $\mathbb{Y} = (m_1 + 1) p_1 + \cdots + (m_s + 1) p_s$ in $\mathbb{P}_K^n$. When $\mathbb{X} = m_1 p_1 + \cdots + m_s p_s$ and $\operatorname{char}(K) \nmid m_j$ for every $j = 1, \ldots, s$, Proposition 3.1 and Corollary 3.5(a) imply that $I_{\mathbb{X}}^{\text{diff}}$ is the homogeneous vanishing ideal of the first fattening $\mathbb{Y}$ of $\mathbb{X}$ in $\mathbb{P}_K^n$.

The above observation indicates that for a fat point scheme whose multiplicities are not divisible by the characteristic of the base field, its first fattening coincides with the scheme defined by the second differential power of its vanishing ideal. This motivates the following general definition for arbitrary 0-dimensional schemes.

Definition 3.7. Let $\mathbb{X}$ be a 0-dimensional scheme in $\mathbb{P}_K^n$ with homogeneous vanishing ideal $I_{\mathbb{X}}$. Then the saturated homogeneous ideal $I_{\mathbb{X}}^{\text{diff}}$ defines a 0-dimensional scheme $\mathbb{Y}$ in $\mathbb{P}_K^n$. The scheme $\mathbb{Y}$ is called the *differential fattening* of $\mathbb{X}$.

With this terminology in hand, we can now provide an affirmative answer to Question 1.1, regarding the existence of the canonical exact sequence of the Kähler differential module Ω_R^1 for an arbitrary 0-dimensional scheme $\mathbb{X}$.

Theorem 3.8. *Let $\mathbb{X}$ be a 0-dimensional scheme in $\mathbb{P}_K^n$ with homogeneous vanishing ideal $I_{\mathbb{X}}$, and let $\mathbb{Y}$ be the differential fattening of $\mathbb{X}$ with its homogeneous vanishing ideal $I_{\mathbb{Y}}$. The sequence of graded R-modules*

$$0 \to I_{\mathbb{X}}/I_{\mathbb{Y}} \xrightarrow{\delta} R^{n+1}(-1) \xrightarrow{\gamma} \Omega_R^1 \to 0$$

is exact, where $\delta(f+I_{\mathbb{Y}}) = (\partial_0(f),\dots,\partial_n(f)) = \sum_{i=0}^n \partial_i(f)e_i$ *for every homogeneous polynomial* $f \in I_{\mathbb{X}}$ *and* $\gamma(e_i) = dx_i$ *for* $i = 0,\dots,n$.

Proof. Note that $\{e_0,\dots,e_n\}$ is the standard basis of the graded free R-module R^{n+1} with $\deg(e_i) = 1$ for every $i = 0,\dots,n$. We also have $R^{n+1}(-1) \cong \Omega^1_P/I_{\mathbb{X}}\Omega^1_P$ and $I^2_{\mathbb{X}} \subseteq I_{\mathbb{Y}} = I^{\text{diff}}_{\mathbb{X}}$. By [Kunz 1986, Proposition 4.17], we have the canonical exact sequence of graded R-modules

$$I_{\mathbb{X}}/I^2_{\mathbb{X}} \xrightarrow{\delta^*} \Omega^1_P/I_{\mathbb{X}}\Omega^1_P \xrightarrow{\gamma^*} \Omega^1_R \to 0,$$

where $\delta^*(f+I^2_{\mathbb{X}}) = df + I_{\mathbb{X}}\Omega^1_P$ for each homogeneous polynomial $f \in I_{\mathbb{X}}$ and $\gamma^*(dX_i + I_{\mathbb{X}}\Omega^1_P) = dx_i$ for $i = 0,\dots,n$. So, it suffices to show that

$$\operatorname{Ker}(\delta^*) = I^{\text{diff}}_{\mathbb{X}}/I^2_{\mathbb{X}}.$$

Indeed, for any polynomial $f \in I_{\mathbb{X}}$, we have $f + I^2_{\mathbb{X}} \in \operatorname{Ker}(\delta^*)$ if and only if $df = \sum_{i=0}^n \partial_i(f)dX_i \in I_{\mathbb{X}}\Omega^1_P$. This is equivalent to the condition that $\partial_i(f) \in I_{\mathbb{X}}$ for all $i = 0,\dots,n$, which in turn is equivalent to $f \in I^{\text{diff}}_{\mathbb{X}}$. □

More generally, the preceding proof also implies that for any ideal $I \subseteq P$ and $R = P/I$, there is a canonical short exact sequence

$$0 \to I/I^{\langle 2\rangle} \to \Omega^1_P/I\Omega^1_P \to \Omega^1_R \to 0.$$

Remark 3.9. Theorem 3.8 and Remark 3.6 together generalize the result of [Kreuzer et al. 2015, Theorem 1.7] for fat point schemes $\mathbb{X} = m_1p_1 + \cdots + m_sp_s$ in $\mathbb{P}^n_K$. This holds even when the base field K is not necessary of characteristic zero, provided that $\operatorname{char}(K) \nmid m_j$ for all $j = 1,\dots,s$.

Corollary 3.10. *Let* $\mathbb{X}$ *be a* 0*-dimensional scheme in* $\mathbb{P}^n_K$*, and let* $\mathbb{Y}$ *be the differential fattening of* $\mathbb{X}$.

(a) *For* $i \in \mathbb{N}$, *we have*

$$\operatorname{HF}_{\Omega^1_R}(i) = (n+1)\operatorname{HF}_{\mathbb{X}}(i-1) + \operatorname{HF}_{\mathbb{X}}(i) - \operatorname{HF}_{\mathbb{Y}}(i).$$

In particular, $\operatorname{HP}(\Omega^1_R) = (n+2)\deg(\mathbb{X}) - \deg(\mathbb{Y})$ *and the regularity index of* Ω^1_R *is bounded by* $\operatorname{ri}(\Omega^1_R) \le \max\{r_{\mathbb{X}}+1, r_{\mathbb{Y}}\}$.

(b) *Suppose* $\mathbb{X} = m_1p_1 + \cdots + m_sp_s$ *is a fat point scheme in* $\mathbb{P}^n_K$ *and let* $\operatorname{char}(K) = p$.

(i) *We have*

$$\operatorname{HP}(\Omega^1_R) = \sum_{j=1}^s \left[(n+2)\binom{m_j+n-1}{n} - \binom{m_j+n}{n}\right] + \sum_{p\mid m_j}\binom{m_j/p+n-1}{n-1}.$$

(ii) *We have* $m_1 = \cdots = m_s = 1$ *if and only if* $\operatorname{HP}(\Omega^1_R) = s$.

Proof. Claim (a) directly follows from Theorem 3.8. Now we will prove the two parts in claim (b). For (i), we have $\mathrm{HP}(\Omega^1_R) = (n+2)\deg(\mathbb{X}) - \deg(\mathbb{Y})$ by (a). For $j \in \{1, \dots, s\}$, let $\mathbb{X}_j = m_j p_j$ and let $\mathbb{Y}_j$ be the differential fattening of $\mathbb{X}_j$. Then $\deg(\mathbb{X}) = \sum_{j=1}^s \deg(\mathbb{X}_j) = \sum_{j=1}^s \binom{m_j+n-1}{n}$ and $\deg(\mathbb{Y}) = \sum_{j=1}^s \deg(\mathbb{Y}_j)$. So, it is sufficient to locally examine $\deg(\mathbb{Y}_j)$. After a homogeneous linear change of coordinates, we may assume $p_j = (1 : 0 : \cdots : 0)$. Then $I_{\mathbb{X}_j} = \langle X_1, \dots, X_n \rangle^{m_j}$. If $p \nmid m_j$, then, by Corollary 3.5(a), $I_{\mathbb{Y}_j} = \langle X_1, \dots, X_n \rangle^{m_j+1}$ and $\mathbb{Y}_j = (m_j+1)p_j$, and consequently $\deg(\mathbb{Y}_j) = \binom{m_j+n}{n}$. When $p \mid m_j$, Corollary 3.5(a) yields

$$I_{\mathbb{Y}_j} = \langle X_1, \dots, X_n \rangle^{m_j+1} + \left\langle X_1^{p\alpha_1} \cdots X_n^{p\alpha_n} \;\middle|\; \sum_{i=1}^n \alpha_i = \frac{m_j}{p} \right\rangle.$$

Letting $A = K[X_1, \dots, X_n]$, it follows that $I^{\mathrm{deh}}_{\mathbb{Y}_j} = I_{\mathbb{Y}_j} \cap A$ and

$$\begin{aligned} \deg(\mathbb{Y}_j) &= \dim_K (A/(I_{\mathbb{Y}_j} \cap A)) \\ &= \binom{m_j+n}{n} - \#\left\{ (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n \;\middle|\; \sum_{i=1}^n \alpha_i = \frac{m_j}{p} \right\} \\ &= \binom{m_j+n}{n} - \sum_{p \mid m_j} \binom{m_j/p+n-1}{n-1}. \end{aligned}$$

Hence we get the desired formula for $\mathrm{HP}(\Omega^1_R)$ as in (i).

Finally, from (i), we have

$$\begin{aligned} \mathrm{HP}(\Omega^1_R) &= \sum_{j=1}^s \left[(n+2)\binom{m_j+n-1}{n} - \binom{m_j+n}{n} \right] + \sum_{p \mid m_j} \binom{m_j/p+n-1}{n-1} \\ &= \sum_{j=1}^s \left[\binom{m_j+n-1}{n} + \binom{m_j+n-1}{n-1}(m_j-1) \right] + \sum_{p \mid m_j} \binom{m_j/p+n-1}{n-1}. \end{aligned}$$

Therefore, $\mathrm{HP}(\Omega^1_R) = s$ if and only if $m_1 = \cdots = m_s = 1$, and (ii) follows. □

Next, let us consider a 0-dimensional scheme $\mathbb{X} \subseteq \mathbb{P}^n_K$ with support $\mathrm{Supp}(\mathbb{X}) = \{p_1, \dots, p_s\} \subseteq D_+(X_0)$. For $j \in \{1, \dots, s\}$, let $\mathfrak{P}_j$ be the homogeneous prime ideal of p_j, and let I_j be the $\mathfrak{P}_j$-primary component of $I_{\mathbb{X}}$. In the sense of [Kreuzer and Robbiano 2016, Definition 4.3.1], the scheme $\mathbb{X}$ is termed a curvilinear scheme if, for each $j \in \{1, \dots, s\}$, there exists a homogeneous linear change of coordinates such that $I_j = \langle X_1, \dots, X_{n-1}, X_n^{k_j} \rangle$ for some positive integer k_j. For such a curvilinear scheme $\mathbb{X}$, the ideal $\mathfrak{Q}_j = I_j^{\mathrm{deh}}$ in $A = K[X_1, \dots, X_n]$ is also generated by $\{X_1, \dots, X_{n-1}, X_n^{k_j}\}$ after an affine linear change of coordinates. The affine ideal of $\mathbb{X}$ is $J_{\mathbb{X}} = \mathfrak{Q}_1 \cap \cdots \cap \mathfrak{Q}_s$, and its affine coordinate ring satisfies $S = A/J_{\mathbb{X}} \cong \mathcal{O}_1 \times \cdots \times \mathcal{O}_s$, where $\mathcal{O}_j \cong A/\mathfrak{Q}_j$.

The next corollary describes the Hilbert polynomial of the Kähler differential module Ω^1_R for a curvilinear scheme (see also [Kreuzer et al. 2025, Corollary 6.7] for an alternative proof).

Corollary 3.11. *Let $\mathbb{X}$ be a 0-dimensional curvilinear scheme in $\mathbb{P}^n_K$, and let $k_j = \dim_K(\mathcal{O}_j) = \dim_K(A/\mathfrak{Q}_j)$ for each $j \in \{1, \ldots, s\}$. Then*

$$\mathrm{HP}(\Omega^1_R) = 2\sum_{j=1}^{s} k_j - \sum_{\mathrm{char}(K)\nmid k_j} 1.$$

Proof. Let $\mathbb{Y}$ be the differential fattening of $\mathbb{X}$. By Corollary 3.10, we have

$$\begin{aligned}
\mathrm{HP}(\Omega^1_R) &= (n+2)\deg(\mathbb{X}) - \deg(\mathbb{Y}) \\
&= (n+2)\dim_K(S) - \dim_K(A/J_{\mathbb{X}}^{\mathrm{diff}}) \\
&= (n+2)\sum_{j=1}^{s}\dim_K(\mathcal{O}_j) - \sum_{j=1}^{s}\dim_K(A/\mathfrak{Q}_j^{\mathrm{diff}}) \\
&= \sum_{j=1}^{s}[(n+2)k_j - \dim_K(A/\mathfrak{Q}_j^{\mathrm{diff}})].
\end{aligned}$$

Fix $j \in \{1, \ldots, s\}$. Without loss of generality, we may assume that

$$\mathfrak{Q}_j = I_j^{\mathrm{deh}} = \langle X_1, \ldots, X_{n-1}, X_n^{k_j}\rangle.$$

If $\mathrm{char}(K)\nmid k_j$ then Corollary 3.5(b) implies that

$$\mathfrak{Q}_j^{\mathrm{diff}} = \langle X_1, \ldots, X_{n-1}\rangle^2 + \langle X_i X_n^{k_j} \mid i = 1, \ldots, n\rangle,$$

and hence $\dim_K(A/\mathfrak{Q}_j^{\mathrm{diff}}) = nk_j + 1$. If $\mathrm{char}(K)\,|\,k_j$ then Corollary 3.5(b) yields

$$\mathfrak{Q}_j^{\mathrm{diff}} = \langle X_1, \ldots, X_{n-1}\rangle^2 + \langle X_n^{k_j}\rangle,$$

and consequently $\dim_K(A/\mathfrak{Q}_j^{\mathrm{diff}}) = nk_j$. Thus we obtain the above formula for the constant Hilbert polynomial $\mathrm{HP}(\Omega^1_R)$. □

4. Differential modules for 0-dimensional LMG-schemes

We continue with the established notation. More precisely, let K be a perfect field, and let $\mathbb{X}$ be a 0-dimensional scheme in $\mathbb{P}^n_K$ with $\mathrm{Supp}(\mathbb{X}) = \{p_1, \ldots, p_s\} \subseteq D_+(X_0)$. Its homogeneous vanishing ideal in $P = K[X_0, \ldots, X_n]$ is $I_{\mathbb{X}}$ and its homogeneous coordinate ring is $R = P/I_{\mathbb{X}}$. Furthermore, let $A = K[X_1, \ldots, X_n]$, let $J_{\mathbb{X}} \subseteq A$ be the affine vanishing ideal of $\mathbb{X}$, and let $J_{\mathbb{X}} = \mathfrak{Q}_1 \cap \cdots \cap \mathfrak{Q}_s$ be the primary decomposition of $J_{\mathbb{X}}$. Then the affine coordinate ring of $\mathbb{X}$ is $S = R/\langle x_0 - 1\rangle = A/J_{\mathbb{X}}$ and it can be decomposed as $S = \mathcal{O}_1 \times \cdots \times \mathcal{O}_s$, where $\mathcal{O}_j = A/\mathfrak{Q}_j$ is the 0-dimensional local ring of $\mathbb{X}$ at the point p_j.

In this section, we turn our focus to the following special class of 0-dimensional schemes that generalizes curvilinear schemes.

Definition 4.1. (a) The scheme $\mathbb{X}$ is called a *locally monomial scheme* if, for each $j \in \{1, \ldots, s\}$, there is a linear change of coordinates $\varphi_j : A \to A$ such that $\mathfrak{Q}'_j = \varphi_j(\mathfrak{Q}_j)$ is a monomial ideal in A.

(b) The scheme $\mathbb{X}$ is called a *locally monomial Gorenstein scheme* (or, for short, an *LMG-scheme*), if $\mathbb{X}$ is a locally monomial scheme and its local rings $\mathcal{O}_j \cong A/\mathfrak{Q}_j$ are Gorenstein for $j = 1, \ldots, s$.

It is well known that every locally complete intersection is locally Gorenstein; however, the converse does not hold in general. The following characterization of an LMG-scheme at a given point can be found in [Herzog and Hibi 2011, Corollary 1.3.2 and Proposition A.6.5].

Proposition 4.2. *Let $\mathfrak{Q}$ be a 0-dimensional monomial ideal in $A = K[X_1, \ldots, X_n]$. Then the following conditions are equivalent*:

(a) *The ring $A/\mathfrak{Q}$ is a Gorenstein ring.*

(b) *The ring $A/\mathfrak{Q}$ is a complete intersection.*

(c) *The ideal $\mathfrak{Q}$ is irreducible.*

(d) $\mathfrak{Q} = \langle X_1^{k_1}, \ldots, X_n^{k_n} \rangle$ *for some positive integers $k_1, \ldots, k_n$.*

Corollary 4.3. *The scheme $\mathbb{X}$ is an LMG-scheme if and only if, for $j \in \{1, \ldots, s\}$, there are positive integers $k_1, \ldots, k_n$ such that $\mathcal{O}_j \cong A/\langle X_1^{k_1}, \ldots, X_n^{k_n} \rangle$ after a suitable linear change of coordinates.*

Now we aim to establish a formula for the Hilbert polynomial of Ω_R^m, the m-form Kähler differential module associated to LMG-schemes $\mathbb{X}$. Since $\Omega_R^m = 0$ for all $m > n + 1$, we only need to consider the range $1 \le m \le n + 1$. When $m = 1$, we can apply Corollary 3.5 to derive the following formula for $\mathrm{HP}(\Omega_R^1)$.

Proposition 4.4. *Let $\mathbb{X}$ be a 0-dimensional LMG-scheme in $\mathbb{P}_K^n$ with support* $\mathrm{Supp}(\mathbb{X}) = \{p_1, \ldots, p_s\}$, *and let p denote the characteristic of K. For $j = 1, \ldots s$, we write $\mathcal{O}_j \cong A/\mathfrak{Q}_j$, where $\mathfrak{Q}_j = \langle X_1^{k_{1j}}, \ldots, X_n^{k_{nj}} \rangle$ with $k_{ij} \ge 1$. Then*

$$\mathrm{HP}(\Omega_R^1) = \sum_{j=1}^{s} \Big[(n+1)k_{1j} \cdots k_{nj} - \sum_{p \nmid k_{ij}} k_{1j} \cdots \widehat{k_{ij}} \cdots k_{nj} \Big].$$

Proof. Let $\mathbb{Y}$ be the differential fattening of $\mathbb{X}$. By Corollary 3.10, we have

$$\begin{aligned} \mathrm{HP}(\Omega_R^1) &= (n+2)\deg(\mathbb{X}) - \deg(\mathbb{Y}) \\ &= (n+2)\sum_{j=1}^{s} \dim_K(\mathcal{O}_j) - \sum_{j=1}^{s} \dim_K(A/\mathfrak{Q}_j^{\mathrm{diff}}) \\ &= \sum_{j=1}^{s} [(n+2)k_{1j} \cdots k_{nj} - \dim_K(A/\mathfrak{Q}_j^{\mathrm{diff}})]. \end{aligned}$$

Letting $j \in \{1, \dots, s\}$, we will compute $\dim_K(A/\mathfrak{Q}_j^{\text{diff}})$. Without loss of generality, we may assume that $p \mid k_{ij}$ for $i \le e$ and $p \nmid k_{ij}$ for $i > e$. An application of Corollary 3.5(b) gives

$$\mathfrak{Q}_j^{\text{diff}} = \langle X_1^{k_{1j}}, \dots, X_e^{k_{ej}}, X_{e+1}^{k_{e+1j}+1}, \dots, X_n^{k_{nj}+1} \rangle + \langle X_{i_1}^{k_{i_1 j}} X_j^{k_{i_2 j}} \mid e < i_1 < i_2 \le n \rangle.$$

Consider a term $t = X_1^{\alpha_1} \cdots X_n^{\alpha_n} \in \mathbb{T}^n$, where $\mathbb{T}^n$ is the set of all terms in A. Note that $t \notin \mathfrak{Q}_j^{\text{diff}}$ yields $\alpha_i \ne k_{ij}$ for $i = 1, \dots, e$. Thus, $t \in \mathbb{T}^n \setminus \mathfrak{Q}_j^{\text{diff}}$ if and only if one of the following holds:

- $0 \le \alpha_i < k_{ij}$ for all $i = 1, \dots, n$.
- if $\alpha_i = k_{ij}$ for some $i \in \{e+1, \dots, n\}$, then $0 \le \alpha_l < k_{lj}$ for all $l \ne i$.

It follows that

$$\begin{aligned} \dim_K(A/\mathfrak{Q}_j^{\text{diff}}) = \#(\mathbb{T}^n \setminus \mathfrak{Q}_j^{\text{diff}}) &= k_{1j} \cdots k_{nj} + \sum_{i=e+1}^{n} k_{1j} \cdots \widehat{k_{ij}} \cdots k_{nj} \\ &= k_{1j} \cdots k_{nj} + \sum_{p \nmid k_{ij}} k_{1j} \cdots \widehat{k_{ij}} \cdots k_{nj}, \end{aligned}$$

where $\widehat{k_{ij}}$ means that the term k_{ij} is omitted from the product. Consequently, we get the desired formula for $\mathrm{HP}(\Omega_R^1)$. □

Example 4.5. Let $K = \mathbb{F}_3$ and let $\mathbb{X}$ be a 0-dimensional scheme in $\mathbb{P}_K^3$ defined by $I_{\mathbb{X}} = \langle f_1, \dots, f_6 \rangle \subseteq P = K[X_0, X_1, X_2, X_3]$, where

$$\begin{aligned} f_1 &= X_1^3 + X_2^3, \\ f_2 &= X_0^2 X_1 + X_0 X_1^2 + X_0^2 X_2 + X_0 X_1 X_2 + X_1^2 X_2 - X_2^3, \\ f_3 &= X_0^2 X_2 X_3 + X_0 X_1 X_2 X_3 + X_1^2 X_2 X_3 - X_3^4, \\ f_4 &= X_0^3 X_3 + X_2^3 X_3 + X_3^4, \\ f_5 &= X_0^3 X_2 + X_0^2 X_2^2 + X_0 X_1 X_2^2 + X_1^2 X_2^2 + X_2^4, \\ f_6 &= X_0^2 X_2^3 + X_0 X_1 X_2^3 + X_1^2 X_2^3 - X_0^2 X_3^3 - X_0 X_1 X_3^3 - X_1^2 X_3^3. \end{aligned}$$

Then $\mathrm{Supp}(\mathbb{X}) = \{p_1, p_2, p_3\}$. In $A = K[X_1, X_2, X_3]$, the affine vanishing ideal $J_{\mathbb{X}}$ decomposes as $J_{\mathbb{X}} = \mathfrak{Q}_1' \cap \mathfrak{Q}_2' \cap \mathfrak{Q}_3'$, where

$$\begin{aligned} \mathfrak{Q}_1' &= \langle X_1, X_2, X_3 \rangle, \\ \mathfrak{Q}_2' &= \langle X_1 + 1, X_2 - 1, (X_3 - 1)^3 \rangle, \\ \mathfrak{Q}_3' &= \langle X_1^2 + X_1 + 1, X_2^3 + 1, X_3^4 \rangle. \end{aligned}$$

Letting $\mathfrak{Q}_1 = \langle X_1, X_2, X_3 \rangle$, $\mathfrak{Q}_2 = \langle X_1, X_2, X_3^3 \rangle$, and $\mathfrak{Q}_3 = \langle X_1^2, X_2^3, X_3^4 \rangle$, we find $\mathcal{O}_j \cong A/\mathfrak{Q}_j$ for $j = 1, 2, 3$. Thus $\mathbb{X}$ is a 0-dimensional LMG-scheme. An application of Proposition 4.4 yields

$$\mathrm{HP}(\Omega_R^1) = 1 + 6 + 78 = 85.$$

To treat the case that $2 \le m \le n+1$, we apply the following observation. By [Kunz 1986, Proposition 4.12], the module Ω^m_S decomposes as $\Omega^m_S \cong \Omega^m_{\mathcal{O}_1} \times \cdots \times \Omega^m_{\mathcal{O}_s}$. Furthermore, according to Proposition 2.12, the Hilbert polynomial of Ω^m_R satisfies

$$\begin{aligned}\mathrm{HP}(\Omega^m_R) &= \dim_K(\Omega^m_S) + \dim_K(\Omega^{m-1}_S) \\ &= \sum_{j=1}^{s} [\dim_K(\Omega^m_{\mathcal{O}_j}) + \dim_K(\Omega^{m-1}_{\mathcal{O}_j})]\end{aligned}$$

and the regularity index of Ω^m_R is bounded by $\mathrm{ri}(\Omega^m_R) \le 2r_{\mathbb{X}} + m$. This reduces the problem to computing $\dim_K(\Omega^m_S)$ in the local setting where $\mathbb{X}$ contains exactly one point.

Proposition 4.6. *Let $m, k_1, \ldots, k_n$ be positive integers and $p = \mathrm{char}(K)$. Suppose that $S = A/\mathfrak{Q}$, where $A = K[X_1, \ldots, X_n]$ and $\mathfrak{Q} = \langle X_1^{k_1}, \ldots, X_n^{k_n} \rangle \subseteq A$. For $1 \le i_1 < \cdots < i_m \le n$, let*

$$\Gamma_{(i_1,\ldots,i_m)} = \{i \in \{i_1, \ldots, i_m\} : p \nmid k_i\}$$

and define

$$J_{(i_1,\ldots,i_m)} = \mathfrak{Q} + \langle X_i^{k_i - 1} \mid i \in \Gamma_{(i_1,\ldots,i_m)} \rangle.$$

Then

$$\Omega^m_S \cong \bigoplus_{1 \le i_1 < \cdots < i_m \le n} (A/J_{(i_1,\ldots,i_m)})(-m)$$

and

$$\dim_K(\Omega^m_S) = \sum_{1 \le i_1 < \cdots < i_m \le n} \prod_{i \notin \Gamma_{(i_1,\ldots,i_m)}} k_i \prod_{i \in \Gamma_{(i_1,\ldots,i_m)}} (k_i - 1).$$

Proof. We equip A with the standard grading, where $\deg(X_i) = 1$ and $\deg(dX_i) = 1$ for every $i = 1, \ldots, n$. Then Ω^m_A is a graded free A-module with basis

$$\{dX_{i_1} \wedge \cdots \wedge dX_{i_m} \mid 1 \le i_1 < \cdots < i_m \le n\}$$

by Proposition 2.9(b). Let us set $e_{i_1,\ldots,i_m} = dX_{i_1} \wedge \cdots \wedge dX_{i_m}$ and write

$$\Omega^m_A = \bigoplus_{1 \le i_1 < \cdots < i_m \le n} A e_{i_1,\ldots,i_m}.$$

By [Kunz 1986, Proposition 4.12], we have the exact sequence of graded A-modules

$$0 \to \mathfrak{Q}\Omega^m_A + d\mathfrak{Q} \wedge \Omega^{m-1}_A \to \Omega^m_A \to \Omega^m_S \to 0.$$

Since $\mathfrak{Q}$ is a monomial ideal, the graded A-module $\mathfrak{Q}\Omega^m_A + d\mathfrak{Q} \wedge \Omega^{m-1}_A$ is a monomial module in the sense of [Kreuzer and Robbiano 2000, Definition 1.3.7]. Of course, for $i = 1, \ldots, n$, we have $dX_i^{k_i} = \sum_{i=1}^n \partial_i(X_i^{k_i}) = k_i X_i^{k_i - 1} dX_i$ if $p \nmid k_i$ and $dX_i^{k_i} = 0$

if $p|k_i$. By [Kreuzer and Robbiano 2000, Theorem 1.3.9] and Proposition 2.9(d), a calculation gives

$$\mathfrak{Q}\Omega_A^m + d\mathfrak{Q} \wedge \Omega_A^{m-1} = \bigoplus_{1\le i_1<\cdots<i_m\le n} J_{(i_1,\ldots,i_m)} e_{i_1,\ldots,i_m}.$$

It follows that

$$\Omega_S^m = \bigoplus_{1\le i_1<\cdots<i_m\le n} (A/J_{(i_1,\ldots,i_m)})(-m).$$

Moreover, we have

$$\dim_K(A/J_{(i_1,\ldots,i_m)}) = \prod_{i\notin\Gamma_{(i_1,\ldots,i_m)}} k_i \prod_{i\in\Gamma_{(i_1,\ldots,i_m)}} (k_i-1).$$

Therefore, we obtain the above formula for $\dim_K(\Omega_S^m)$. □

Example 4.7. Let us continue looking at the LMG-scheme $\mathbb{X}$ given in Example 4.5. We know that $\mathrm{Supp}(\mathbb{X}) = \{p_1, p_2, p_3\}$ and $\mathcal{O}_j \cong A/\mathfrak{Q}_j$ for $j = 1, 2, 3$, where $\mathfrak{Q}_1 = \langle X_1, X_2, X_3\rangle$, $\mathfrak{Q}_2 = \langle X_1, X_2, X_3^3\rangle$, and $\mathfrak{Q}_3 = \langle X_1^2, X_2^3, X_3^4\rangle$. Since K is a perfect field and $\mathcal{O}_1 \cong K$, we have $\Omega^m_{\mathcal{O}_1} = 0$ for $m \ge 1$. Notice that $\Omega^m_{\mathcal{O}_j} = 0$ for $m \ge 4$ and $j = 2, 3$. By Proposition 4.6, we find

$$\begin{aligned}
&\dim_K(\Omega^1_{\mathcal{O}_2}) = 3, \quad \dim_K(\Omega^1_{\mathcal{O}_3}) = 54,\\
&\dim_K(\Omega^2_{\mathcal{O}_2}) = 0, \quad \dim_K(\Omega^2_{\mathcal{O}_3}) = 39,\\
&\dim_K(\Omega^3_{\mathcal{O}_2}) = 0, \quad \dim_K(\Omega^3_{\mathcal{O}_3}) = 9.
\end{aligned}$$

Consequently, by Propositions 2.12, we get

$$\mathrm{HP}(\Omega^1_R) = \deg(\mathbb{X}) + \sum_{j=1}^{3} \dim_K \Omega^1_{\mathcal{O}_j} = 28+0+3+54 = 85,$$

$$\mathrm{HP}(\Omega^2_R) = \sum_{j=1}^{3} (\dim_K \Omega^1_{\mathcal{O}_j} + \dim_K \Omega^2_{\mathcal{O}_j}) = 0+3+93 = 96,$$

$$\mathrm{HP}(\Omega^3_R) = \sum_{j=1}^{3} (\dim_K \Omega^2_{\mathcal{O}_j} + \dim_K \Omega^3_{\mathcal{O}_j}) = 0+0+48 = 48.$$

When $\mathbb{X}$ is a 0-dimensional LMG-scheme, an application of Propositions 2.12 and 4.6 yields the following formula for the Hilbert polynomial of Ω^m_R.

Corollary 4.8. *Let $\mathbb{X}$ be a 0-dimensional LMG-scheme in $\mathbb{P}^n_K$, and let $p = \mathrm{char}(K)$. For $j = 1, \ldots s$, we write $\mathcal{O}_j \cong A/\mathfrak{Q}_j$, where $\mathfrak{Q}_j = \langle X_1^{k_{1j}}, \ldots, X_n^{k_{nj}}\rangle$ with $k_{ij} \ge 1$. For $1 \le i_1 < \cdots < i_m \le n$ and $1 \le j \le s$, let*

$$\Gamma_{(i_1,\ldots,i_m),j} = \{i \in \{i_1, \ldots, i_m\} : p \nmid k_{ij}\}.$$

For $m \geq 1$, we have

$$\mathrm{HP}(\Omega^m_R) = \sum_{j=1}^{s} \Bigg[\sum_{1 \leq i_1 < \cdots < i_{m-1} \leq n \; i \notin \Gamma_{(i_1,\ldots,i_{m-1}),j}} \prod k_{ij} \prod_{i \in \Gamma_{(i_1,\ldots,i_{m-1}),j}} (k_{ij} - 1) + \sum_{1 \leq l_1 < \cdots < l_m \leq n \; i \notin \Gamma_{(l_1,\ldots,l_m),j}} \prod k_{ij} \prod_{i \in \Gamma_{(l_1,\ldots,l_m),j}} (k_{ij} - 1) \Bigg].$$

Acknowledgements

The authors would like to thank Martin Kreuzer for his generous support and encouragement. This work is supported by the *Vietnam Ministry of Education and Training* under the grant B2025-DHH-02. The authors thank the referees for a careful reading of the paper and many detailed and helpful remarks and suggestions.

References

[ApCoCoA ≥ 2026] ApCoCoA, "ApCoCoA: applied computations in computer algebra", available at apcocoa.uni-passau.de.

[Bernstein et al. 1972] I. N. Bernstein, I. M. Gelfand, and S. I. Gelfand, "Differential operators on a cubic cone", *Uspekhi Mat. Nauk* **27**:1(163) (1972), 185–190. In Russian; translated in *Russ. Math. Surv.* **27**:1 (1972), 169–174. MR

[Björk 1979] J.-E. Björk, *Rings of differential operators*, North-Holland Mathematical Library **21**, North-Holland, Amsterdam, 1979. MR

[Chandler 2001] K. A. Chandler, "A brief proof of a maximal rank theorem for generic double points in projective space", *Trans. Amer. Math. Soc.* **353**:5 (2001), 1907–1920. MR

[Coutinho 1995] S. C. Coutinho, *A primer of algebraic D-modules*, London Mathematical Society Student Texts **33**, Cambridge Univ. Press, 1995. MR

[Dao et al. 2018] H. Dao, A. De Stefani, E. Grifo, C. Huneke, and L. Núñez Betancourt, "Symbolic powers of ideals", pp. 387–432 in *Singularities and foliations*: *geometry*, *topology and applications*, edited by R. N. Araújo dos Santos et al., Springer Proc. Math. Stat. **222**, Springer, 2018. MR

[De Stefani et al. 2020] A. De Stefani, E. Grifo, and J. Jeffries, "A Zariski–Nagata theorem for smooth $\mathbb{Z}$-algebras", *J. Reine Angew. Math.* **761** (2020), 123–140. MR

[de Dominicis and Kreuzer 1999] G. de Dominicis and M. Kreuzer, "Kähler differentials for points in $\mathbb{P}^n$", *J. Pure Appl. Algebra* **141**:2 (1999), 153–173. MR

[EGA IV$_4$ 1967] A. Grothendieck, "Éléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, IV", *Inst. Hautes Études Sci. Publ. Math.* **32** (1967), 5–361. MR

[Hart 1983] R. Hart, "Differential operators on affine algebras", *J. London Math. Soc.* (2) **28**:3 (1983), 470–476. MR

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, 1977. MR

[Herzog and Hibi 2011] J. Herzog and T. Hibi, *Monomial ideals*, Graduate Texts in Mathematics **260**, Springer, 2011. MR

[Heyneman and Sweedler 1969] R. G. Heyneman and M. E. Sweedler, "Affine Hopf algebras, I", *J. Algebra* **13** (1969), 192–241. MR

[Hochster and Huneke 2002] M. Hochster and C. Huneke, "Comparison of symbolic and ordinary powers of ideals", *Invent. Math.* **147**:2 (2002), 349–369. MR

[Hochster and Huneke 2007] M. Hochster and C. Huneke, "Fine behavior of symbolic powers of ideals", *Illinois J. Math.* **51**:1 (2007), 171–183. MR

[Kreuzer and Robbiano 2000] M. Kreuzer and L. Robbiano, *Computational commutative algebra, I*, Springer, 2000. MR

[Kreuzer and Robbiano 2005] M. Kreuzer and L. Robbiano, *Computational commutative algebra, II*, Springer, 2005. MR

[Kreuzer and Robbiano 2016] M. Kreuzer and L. Robbiano, *Computational linear and commutative algebra*, Springer, 2016. MR

[Kreuzer et al. 2015] M. Kreuzer, N. K. L. Tran, and L. N. Long, "Kähler differentials and Kähler differents for fat point schemes", *J. Pure Appl. Algebra* **219**:10 (2015), 4479–4509. MR

[Kreuzer et al. 2018] M. Kreuzer, T. N. K. Linh, and L. N. Long, "Kähler differential algebras for 0-dimensional schemes", *J. Algebra* **501** (2018), 255–284. MR

[Kreuzer et al. 2025] M. Kreuzer, T. N. K. Linh, and L. N. Long, "Differential theory of zero-dimensional schemes", *J. Pure Appl. Algebra* **229**:1 (2025), art. id. 107815. MR

[Kumar 1978] N. M. Kumar, "On two conjectures about polynomial rings", *Invent. Math.* **46**:3 (1978), 225–236. MR

[Kunz 1986] E. Kunz, *Kähler differentials*, Friedr. Vieweg & Sohn, Braunschweig, 1986. MR

[Kunz 2013] E. Kunz, *Introduction to commutative algebra and algebraic geometry*, Springer, 2013. MR

[Scheja and Storch 1974] G. Scheja and U. Storch, *Lokale Verzweigungstheorie*, Schriftenreihe des Mathematischen Institutes der Universität Freiburg **5**, Université de Fribourg, Institut des Mathématiques, Fribourg, 1974. MR

[Smith 1986] S. P. Smith, "Differential operators on the affine and projective lines in characteristic $p > 0$", pp. 157–177 in *Séminaire d'algèbre Paul Dubreil et Marie-Paule Malliavin*, 37*ème année* (Paris, 1985), edited by M.-P. Malliavin, Lecture Notes in Math. **1220**, Springer, 1986. MR

Received June 25, 2025. Revised April 16, 2026.

TRAN N. K. LINH
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF EDUCATION, HUE UNIVERSITY
HUE
VIETNAM
tnkhanhlinh@hueuni.edu.vn

LE NGOC LONG
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF EDUCATION, HUE UNIVERSITY
HUE
VIETNAM
lengoclong@dhsphue.edu.vn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 344, No. 1, 2026
https://doi.org/10.2140/pjm.2026.344.69

ON THE MEASURE OF PRIMITIVE REPRESENTATIONS OF AN INTEGER BY A QUADRATIC FORM

MANABU MURATA

We explicitly compare a measure defined by Siegel with a mass defined by Shimura, both in the context of primitive representations of an integer by an indefinite integral quadratic form. The main result gives an explicit formula for the Siegel's measure under the assumption that the lattices are maximal. The result is applied to the computation of these measures in some families, such as that of hyperbolic forms.

Introduction

Our study concerns the integral representations of an integer by an *indefinite* quadratic form with integer coefficients. Namely, we take an indefinite symmetric matrix S in $\mathrm{GL}_n(\mathbf{Q})$ with entries in $\mathbf{Z}$ and consider the solutions over $\mathbf{Z}$ of the quadratic Diophantine equation

$$ {}^{\mathsf{T}}xSx = t \tag{0-1}$$

with an integer t. Unlike the definite case, the set of solutions of (0-1) has infinitely many elements in general. Siegel considered a measure on such an infinite set by defining the volumes associated with orbits of the solutions under the unit group of S and called it *the measure of representations of t by S*. In [Si3, Theorem 1] he then proved the product formula, which expresses that measure as an infinite product of local representation densities of t by S. There are many researches concerning his product formula, and recently the formula was generalized to the case where S is defined over an arbitrary algebraic number field by Kojima and Sakata in [KS]. However, the explicit results for computation seem to be not so much known. The reader is referred to Ratcliffe and Tschantz [RT, Theorem 12], in which they gave an explicit formula for the representation measures of integers by Lorentzian forms. In this paper we give a way of computing Siegel's representation measure with the primitive condition explained below.

MSC2020: primary 11D09, 11E12; secondary 11E20.
Keywords: quadratic Diophantine equation, quadratic form, primitive representation.

© 2026 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

With S and t above, we set

$$B(S,t) = \left\{x \in \mathbf{Q}^n \mid {}^{\mathsf{T}}xSx = t,\ \mathbf{Z}_n^1 x = \mathbf{Z}\right\},$$

where $\mathbf{Z}_n^1$ is the $\mathbf{Z}$-module of all $(1 \times n)$-matrices with entries in $\mathbf{Z}$. The condition $\mathbf{Z}_n^1 x = \mathbf{Z}$ means that x is *primitive* in the traditional sense. We also set

$$\Omega(S) = \{\gamma \in \mathrm{GL}_n(\mathbf{R}) \mid {}^{\mathsf{T}}\gamma S\gamma = S\}, \quad \Omega(S,C) = \{\gamma \in \Omega(S) \mid \gamma C = C\},$$
$$\Delta(S) = \Omega(S) \cap \mathrm{GL}_n(\mathbf{Z}), \qquad \Delta(S,C) = \Omega(S,C) \cap \mathrm{GL}_n(\mathbf{Z}),$$

with $C \in B(S,t)$. Since $B(S,t)$ is stable under the map $x \mapsto \gamma x$ for $\gamma \in \Delta(S)$, we can consider the set of the $\Delta(S)$-orbits of $B(S,t)$, which we denote by $\Delta(S)\backslash B(S,t)$. For each representative C for $\Delta(S)\backslash B(S,t)$ we consider $(C\ \ Y) \in \mathrm{GL}_n(\mathbf{R})$ with a variable matrix $Y \in \mathbf{R}_{n-1}^n$; then

$${}^{\mathsf{T}}(C\ \ Y)S(C\ \ Y) = \begin{bmatrix} t & Q \\ {}^{\mathsf{T}}Q & R \end{bmatrix}$$

has the same signature as S, where $Q \in \mathbf{R}_{n-1}^1$ and ${}^{\mathsf{T}}R = R \in \mathbf{R}_{n-1}^{n-1}$. Fixing such a matrix T_1, we have the set $\mathcal{P}(S,C,T_1)$ of all solutions $(C\ \ Y)$ of the equation ${}^{\mathsf{T}}(C\ \ Y)S(C\ \ Y) = T_1$. This can be viewed as a differentiable manifold on which $\Omega(S,C)$ acts on the left. In [Si3] Siegel defined the volume $\rho(S,C)$ of $\Delta(S,C)\backslash\mathcal{P}(S,C,T_1)$ using an $\Omega(S,C)$-invariant volume element on $\mathcal{P}(S,C,T_1)$; similarly, the volume $\rho(S)$ was defined for the equation ${}^{\mathsf{T}}XSX = T$; see Section 1.3 for details. Then he considered the quantity

$$\nu(S,t) = \sum_{\Delta(S)C \in \Delta(S)\backslash B(S,t)} \rho(S,C)/\rho(S). \tag{0-2}$$

Following Siegel, we call $\nu(S,t)$ *the measure of primitive representations of t by S*. It is noted that $\Delta(S)\backslash B(S,t)$ is a finite set.

Siegel proved the product formula

$$\nu(S,t) = \prod_p \beta_p(S,t), \tag{0-3}$$

where p runs over all prime numbers, $\beta_p(S,t)$ is the primitive representation density of t by S at p, and $n \geq 4$. In fact, this is the specialization of [Si3, Theorem 2] to the scalar-valued Diophantine equation ${}^{\mathsf{T}}xSx = t$. (As for the case $n = 4$, see a remark in [Si3, Section 12].) By virtue of (0-3), $\nu(S,t)$ can be computed by determining all the $\beta_p(S,t)$ in the product. Conversely, the determination of the measure in question provides useful information on such local densities. Several explicit results are known for local representation densities; examples are due to Siegel [Si1, Hilfssatz 16], Yang [Y, Theorem 4.1], and more recently Ikeda and Katsurada [IK, Lemma 5.4] for hyperbolic forms.

On the other hand, Shimura considered the solutions of (0-1) in a different perspective from Siegel's and introduced the mass of the set of such solutions in [S2, (13.14)]. Since our main result is obtained from his mass formula in [S2, Theorem 13.12], let us briefly recall Shimura's treatment on the Diophantine equation.

Let φ be an indefinite symmetric matrix in $\mathrm{GL}_n(\mathbf{Q})$ and assume that φ is semi-integral; namely, the diagonal entries are in $\mathbf{Z}$ and the others in $2^{-1}\mathbf{Z}$. We use the same symbol φ for the bilinear symmetric form defined by $(x, y) \mapsto x\varphi \cdot {}^{\mathrm{T}}y$. We set $V = \mathbf{Q}_n^1$, $L = \mathbf{Z}_n^1$, $G = \{\gamma \in SL_n(\mathbf{Q}) \mid \gamma\varphi \cdot {}^{\mathrm{T}}\gamma = \varphi\}$, and $\Gamma(L) = \{\gamma \in G \mid L\gamma = L\}$. We consider the set of solutions $L[q, c\mathbf{Z}]$ defined by

$$L[q, c\mathbf{Z}] = \{x \in V \mid \varphi(x, x) = q,\ \varphi(x, L) = c\mathbf{Z}\} \tag{0-4}$$

with $q, c \in \mathbf{Q}^\times$. Then $L[q, c\mathbf{Z}]$ is stable under the map $x \mapsto x\gamma$ for $\gamma \in \Gamma(L)$. Pick $h \in L[q, c\mathbf{Z}]$ and consider the complement $W = \{y \in V \mid \varphi(h, y) = 0\}$, the restriction ψ of φ to W, and $H = \{\gamma \in G \mid h\gamma = h\}$, which is regarded as the special orthogonal group of ψ. Then the mass of $L[q, c\mathbf{Z}]$ is defined by

$$\mathfrak{m}(L[q, c\mathbf{Z}]) = \sum_{\beta \in B} \nu(H \cap \beta\Gamma(L)\beta^{-1})/\nu(\Gamma(L)) \tag{0-5}$$

with a finite subset B of G such that $L[q, c\mathbf{Z}] = \bigsqcup_{\beta \in B} h\beta\Gamma(L)$. Here $\nu(\Gamma(L))$ is the constant $\#(\Gamma(L) \cap \{\pm 1\})^{-1}$ times the measure of the quotient $\Gamma(L)\backslash\mathcal{Z}^\varphi$ with respect to an invariant measure on the symmetric space $\mathcal{Z}^\varphi = \mathcal{Z}(r, \theta)$ on which G acts with $0 \le r \in \mathbf{Z}$ and $0 < {}^{\mathrm{T}}\theta = \theta \in \mathrm{GL}_{n-2r}(\mathbf{R})$ (see Section 1.2 for details). The quantity $\nu(H \cap \beta\Gamma(L)\beta^{-1})$ is defined similarly with H and $\mathcal{Z}^\psi = \mathcal{Z}(r_1, \theta_1)$; in particular if ψ is definite, $r_1 = 0$ and $\nu(H \cap \beta\Gamma(L)\beta^{-1}) = [H \cap \beta\Gamma(L)\beta^{-1} : 1]^{-1}$. We note that $\mathfrak{m}(L[q, c\mathbf{Z}]) = \#\{L[q, c\mathbf{Z}]\}$ when φ is definite.

The arithmetic nature of (0-4) can be clarified when L is $\mathbf{Z}$-*maximal* in the sense that L is a $\mathbf{Z}$-lattice in V such that $\varphi(x, x) \in \mathbf{Z}$ for every $x \in L$ and L is maximal among such lattices. To be precise, let $G_{\boldsymbol{A}}$ be the adelization of G and put $C(L) = \{\alpha \in G_{\boldsymbol{A}} \mid L\alpha = L\}$. For $x \in V$ and $\alpha \in G_{\boldsymbol{A}}$, $x\alpha$ is meaningful in $V_{\boldsymbol{A}}$ $(= V \otimes_{\mathbf{Q}} \mathbf{Q}_{\boldsymbol{A}})$ as V is embedded in $V_{\boldsymbol{A}}$. Shimura called the set $V \cap hC(L)$ *the genus of* h and in [S4, Theorem 2.2(iv)] he proved that if L is $\mathbf{Z}$-maximal and $n > 2$, then

$$V \cap hC(L) = L[q, c\mathbf{Z}]. \tag{0-6}$$

This means that every element of $L[q, c\mathbf{Z}]$ locally belongs to $hC(L)_p$ for every prime p if L is maximal, where the subscript p means the p-component. Shimura found a principle governing such genera $V \cap hC(L)\alpha^{-1}$ for $\alpha \in G_{\boldsymbol{A}}$ and gave the mass formula based on that principle in [S4, Theorems 2.2 and 3.2]. In our maximal case, that mass formula can be stated as follows:

$$\mathfrak{m}(L[q, c\mathbf{Z}])\nu(\Gamma(L)) = \mathfrak{m}(H, H_{\boldsymbol{A}} \cap C(L)), \tag{0-7}$$

where $H_{\boldsymbol{A}}$ is the adelization of H and $\mathfrak{m}(H, H_{\boldsymbol{A}} \cap C(L))$ is the mass of H relative to $H_{\boldsymbol{A}} \cap C(L)$ (see (4-5) for the definition). By virtue of the classical mass formula for orthogonal groups, $\mathfrak{m}(H, H_{\boldsymbol{A}} \cap C(L))$ is essentially the inverse of an infinite product of local representation densities of 2ψ by 2ψ itself. Thus (0-7) relates the mass of the genus of a solution of the equation $x\varphi \cdot {}^{\mathrm{t}}x = q$ to the local representation densities of 2ψ by 2ψ. Note that (0-7) is not applicable to $L[q, c\mathbf{Z}]$ with an arbitrary lattice L in general, because it may not be maximal; see [S2, Theorem 13.12] for the precise statement. In [S2, §13.15] Shimura asked how his mass formula can be connected with Siegel's product formulas.

The purpose of this paper is to connect Shimura's mass with Siegel's representation measure and by using the relationship, to give an explicit formula for $\nu(S, t)$ under the assumption that the lattices L and $L \cap W$ are maximal. With φ and $L = \mathbf{Z}_n^1$ as above, we have:

Theorem. *Assume that φ is indefinite and semi-integral with $n > 2$. Pick $h \in L[q, c\mathbf{Z}]$ with $q, c \in \mathbf{Q}^\times$ and consider (W, ψ) and H in the above manner. Assume also that ψ is anisotropic if $n = 3$. Let $\mathfrak{m}(L[q, c\mathbf{Z}])$ be the mass of $L[q, c\mathbf{Z}]$ defined with $\mathcal{Z}^\varphi = \mathcal{Z}(r, \theta)$ and $\mathcal{Z}^\psi = \mathcal{Z}(r_1, \theta_1)$. Set $S = \ell_\varphi(2\varphi)^{-1}$ and $t = 2^{-1}\ell_\varphi c^{-2} q$, where ℓ_φ denotes the level of φ. Then $\mathfrak{m}(L[q, c\mathbf{Z}])$ is related to the measure $\nu(S, t)$ by*

$$\text{(0-8)}\quad \mathfrak{m}(L[q, c\mathbf{Z}]) = 2^{(2+r_1(n-1)-rn)/2} |c^{-2}q|^{(n-2)/2} |\det(\varphi)|^{1/2} \ell_\varphi^{-1} \cdot \frac{\det(\theta_1)^{r_1/2} \rho_{n-r}\rho_r}{\det(\theta)^{r/2}\rho_{n-1-r_1}\rho_{r_1}}\, \nu(S, t).$$

Here $\rho_m = \prod_{k=1}^m \pi^{k/2}\Gamma(k/2)^{-1}$ if $m > 0$ and $\rho_0 = 1$; as usual, $\Gamma(s)$ is the gamma function.

It is noted that S in the Theorem defines a quadratic form with *even-integer* coefficients, because $2^{-1}S$ is semi-integral by the definition of ℓ_φ. The assumption on ψ when $n = 3$ is needed to avoid the case where the volume $\rho(\psi)$ becomes infinite for a binary isotropic form ψ.

By the Theorem combined with (0-3) the local primitive representation densities of t by S are related to $\mathfrak{m}(L[q, c\mathbf{Z}])$. Further if L is maximal, the latter is related to the local representation densities of 2ψ by 2ψ via (0-7).

One virtue of our result is that it bypasses the classical product formula of Siegel, which is not readily computable in practice. In his formal expression, the local densities are difficult to compute explicitly, especially under the primitive condition for indefinite forms. Our comparison approach trades the difficulty of local representations for a more tractable (though still delicate) problem of evaluating masses of arithmetic quotients of type $\Gamma\backslash\mathcal{Z}$, which can at least be computed in the maximal case through Shimura's mass formula as will be seen from Theorem 4.1.

The Theorem will be proved in Sections 2 and 3 after some preliminaries in Section 1. We shall explicitly compare $\nu(H \cap \beta\Gamma(L)\beta^{-1})$ with $\rho(S, C_\beta)$ for some C_β by relating both to the same volume $\rho(a\psi_\beta)$, where a is a nonzero constant and ψ_β is the matrix representing ψ with a basis of the lattice $L\beta^{-1} \cap W$. In particular, Section 3 is devoted to connecting $\rho(S, C_\beta)$ with $\rho(a\psi_\beta)$ by means of some technique of integration such as Siegel did in [Si3, §10].

The main result of this paper is an explicit formula for $\nu(S, t)$ that follows from the Theorem combined with (0-7). To state it simply, suppose n is even and put $M = L \cap W$ with $h \in L[q, c\mathbf{Z}]$. Set $S = \ell_\varphi(2\varphi)^{-1}$ and $t = 2^{-1}\ell_\varphi c^{-2}q$ as in the Theorem. Then for $\mathbf{Z}$-maximal lattices L and M we have (see Theorem 4.1)

$$\nu(S, t) = 2\ell_\varphi[\tilde{L} : L]^{-1}[C(M) : H_A \cap C(L)]L(n/2, \chi)^{-1} \prod_{p \mid \mathfrak{e}} (\lambda_p^\varphi)^{-1} \prod_{p \mid \mathfrak{e}_1} \lambda_p^\psi.$$

Here $\tilde{L} = \{x \in V \mid 2\varphi(x, L) \subset \mathbf{Z}\}$, $\mathfrak{e}$ (resp. $\mathfrak{e}_1$) is the product of all the prime numbers for which $\tilde{L}_p \neq L_p$ (resp. $\tilde{M}_p \neq M_p$), $L(s, \chi)$ is the L-function of the primitive Dirichlet character χ corresponding to $\mathbf{Q}(\sqrt{\delta})$ with $\delta = (-1)^{n/2}\det(\varphi)$, and the constants λ_p^φ and λ_p^ψ are explicitly given in Theorem 4.1. A similar formula will be proven for odd n. Section 4.1 explains the actual computations on our formula.

As an application of Theorem 4.1, in Section 5 we shall consider the hyperbolic form $\eta_r = \left[\begin{smallmatrix} 0 & 1_r \\ 1_r & 0 \end{smallmatrix}\right]$ and obtain an explicit formula for $\nu(\eta_r, 2q)$ for every squarefree integer q. We shall also deduce the local density $\beta_p(\eta_r, 2q)$ at $p = 2$ by comparing our result with Siegel's formula (0-3). A similar application will be discussed for the Lorentzian form $S_n = \left[\begin{smallmatrix} 1_{n-1} & 0 \\ 0 & -1 \end{smallmatrix}\right]$ with $n = 5$ after the remark that the lattice $\mathbf{Z}_n^1$ is not $\mathbf{Z}$-maximal with respect to S_n and so Theorem 4.1 is not applicable to $\varphi = S_n$ and $L = \mathbf{Z}_n^1$ for any $n > 2$.

Another consequence of the Theorem is Corollary 5.9, on the invariance of the mass of $L[q, c\mathbf{Z}]$ for indefinite φ with $n > 3$: *Integral* $\mathbf{Z}$*-lattices* L_1 *in the same G-genus have the same masses* $\mathfrak{m}(L[q, c\mathbf{Z}])$, *provided* $L_1[q, c\mathbf{Z}] \neq \varnothing$. This invariance does not hold when φ is definite; see, e.g., the table in [M2, Introduction].

Notation. Throughout the paper we follow the notation and terminology in [S2]. We denote by $\mathbf{Z}_p$ the ring of p-adic integers and by $\mathbf{Q}_p$ the field of p-adic numbers. We put $F^{\times 2} = \{x^2 \mid x \in F^\times\}$ for $F = \mathbf{Q}$ or $\mathbf{Q}_p$.

If A is a commutative associative ring with identity element and if X is an A-module, then we write $A^\times$ for the group of all invertible elements of A and X_n^m for the A-module of all $(m \times n)$-matrices with entries in X. We put $\hat{x} = {}^\top x^{-1}$ if x is an invertible matrix. We denote by 1_n the identity element of A_n^n. We put $\mathrm{GL}_n(A) = (A_n^n)^\times$ and $SL_n(A) = \{x \in \mathrm{GL}_n(A) \mid \det(x) = 1\}$. For a symmetric matrix S in A_n^n we put $S[x] = {}^\top xSx$ for $x \in A_m^n$. We write $\mathrm{diag}[a_1, \ldots, a_s]$ for the matrix with square matrices $a_1, \ldots, a_s$ in the diagonal blocks and 0 in all other blocks. If a union $\bigcup_{i \in I} X_i$ is disjoint, we indicate this by writing $\bigsqcup_{i \in I} X_i$.

1. Preliminaries

1.1. ***Orbits of $L[q, \mathfrak{b}]$ and $B(S, t)$.*** Let V be a vector space of dimension $n > 2$ over $\mathbf{Q}$ and φ a nondegenerate $\mathbf{Q}$-bilinear symmetric form on $V \times V$. By a $\mathbf{Z}$-lattice in V we mean a finitely generated $\mathbf{Z}$-submodule L of V that spans V over $\mathbf{Q}$. We call L *integral* with respect to φ if $\varphi(x, x) \in \mathbf{Z}$ for every $x \in L$. If L is maximal among such $\mathbf{Z}$-lattices, we call it $\mathbf{Z}$-*maximal* with respect to φ. We put

$$\widetilde{L} = \{x \in V \mid 2\varphi(x, L) \subset \mathbf{Z}\}.$$

If L is integral with respect to φ, then L is contained in $\widetilde{L}$. We let the automorphism group $\mathrm{GL}(V)$ act on V on the right and define the orthogonal group of φ by $O(\varphi) = \{\gamma \in \mathrm{GL}(V) \mid \varphi(x\gamma, y\gamma) = \varphi(x, y) \text{ for every } x, y \in V\}$. Put

$$\Gamma^*(L) = \{\gamma \in O(\varphi) \mid L\gamma = L\},$$
$$\mathrm{SO}(\varphi) = \{\gamma \in O(\varphi) \mid \det(\gamma) = 1\},$$

and $\Gamma(L) = \mathrm{SO}(\varphi) \cap \Gamma^*(L)$. We have $[\Gamma^*(L) : \Gamma(L)] \leq 2$ since $[O(\varphi) : \mathrm{SO}(\varphi)] = 2$.

Let h be an element of V such that $\varphi(h, h) \neq 0$, and set $q = \varphi(h, h)$ and $\mathfrak{b} = \varphi(h, L)$. We consider the orthogonal complement

$$W = (\mathbf{Q}h)^{\perp} = \{x \in V \mid \varphi(x, h) = 0\}$$

and the restriction ψ of φ to W. We then regard $O(\psi)$ as the subgroup

$$H^* = \{\gamma \in O(\varphi) \mid h\gamma = h\}$$

of $O(\varphi)$ and $\mathrm{SO}(\psi)$ as $H = \mathrm{SO}(\varphi) \cap H^*$. Hence $\Gamma(M)$ means the subgroup $\{\gamma \in H \mid M\gamma = M\}$ for a $\mathbf{Z}$-lattice M in W.

With q and $\mathfrak{b}$ as above and $L[q, \mathfrak{b}]$ as in (0-4), we have $h \in L[q, \mathfrak{b}]$. If $k \in L[q, \mathfrak{b}]$, there exists $\alpha \in \mathrm{SO}(\varphi)$ such that $k = h\alpha$, by [S2, Lemma 1.5]. Then, since $L[q, \mathfrak{b}]$ is stable under the map $x \mapsto x\gamma$ for $\gamma \in \Gamma(L)$, we have

$$L[q, \mathfrak{b}] = \bigsqcup_{\beta \in B} h\beta\Gamma(L) \tag{1-1}$$

with a subset B of $\mathrm{SO}(\varphi)$. It can be seen from [S2, Theorem 11.1] that B is a finite set. We may assume that B contains the identity element.

If $[\Gamma^*(L) : \Gamma(L)] = 1$, we write $B = B_1 \sqcup B_2$, with

$$B_1 = \{\beta \in B \mid \Gamma^*(L\beta^{-1} \cap W) = \Gamma(L\beta^{-1} \cap W)\},$$
$$B_2 = \{\beta \in B \mid \Gamma^*(L\beta^{-1} \cap W) \neq \Gamma(L\beta^{-1} \cap W)\}.$$

If instead $[\Gamma^*(L) : \Gamma(L)] = 2$, fix $\delta \in \Gamma^*(L)$ such that $\Gamma^*(L) = \Gamma(L) \sqcup \delta\Gamma(L)$, and put

$$B_1 = \{\beta \in B \mid h\beta\delta \in h\beta\Gamma(L)\}.$$

If $\beta_1, \beta_2 \in B$ are such that $\beta_1 \notin B_1$ and $h\beta_1\delta \in h\beta_2\Gamma(L)$, associate to the *unordered* pair β_1, β_2 either of β_1 or β_2. Let B_2 be the subset of B consisting of elements so chosen. Then put

$$B_{21} = \{\beta \in B_2 \mid \Gamma^\bullet(L\beta^{-1} \cap W) \neq \Gamma(L\beta^{-1} \cap W)\},$$
$$B_{22} = \{\beta \in B_2 \mid \Gamma^\bullet(L\beta^{-1} \cap W) = \Gamma(L\beta^{-1} \cap W)\},$$
$$B_{2i}^\delta = \{\beta' \in B \mid h\beta\delta \in h\beta'\Gamma(L) \text{ for } \beta \in B_{2i}\} \quad \text{for } i = 1, 2,$$
$$B_2^\delta = B_{21}^\delta \sqcup B_{22}^\delta.$$

Then we see that $B_2 = B_{21} \sqcup B_{22}$ and $B = B_1 \sqcup B_2 \sqcup B_2^\delta$.

Lemma 1.1. *With the notation above, let $\beta \in B$. Then*:

(1) $[H^\bullet \cap \Gamma^\bullet(L\beta^{-1}) : H \cap \Gamma(L\beta^{-1})] = 1$ *if* $[\Gamma^\bullet(L\beta^{-1}) : \Gamma(L\beta^{-1})] = 1$.

(2) $[H^\bullet \cap \Gamma^\bullet(L\beta^{-1}) : H \cap \Gamma(L\beta^{-1})] = 1$ *if* $[\Gamma^\bullet(L\beta^{-1} \cap W) : \Gamma(L\beta^{-1} \cap W)] = 1$.

(3) *Suppose that* $[\Gamma^\bullet(L) : \Gamma(L)] = 2$. *Let* $\Gamma^\bullet(L) = \Gamma(L) \sqcup \delta\Gamma(L)$. *Then*:

(i) $\beta \in B_1 \iff [H^\bullet \cap \Gamma^\bullet(L\beta^{-1}) : H \cap \Gamma(L\beta^{-1})] = 2$.

(ii) $\beta \in B_1 \sqcup B_{21} \sqcup B_{21}^\delta \iff [\Gamma^\bullet(L\beta^{-1} \cap W) : \Gamma(L\beta^{-1} \cap W)] = 2$.

Consequently, $\beta \in B_1 \sqcup B_{22} \sqcup B_{22}^\delta$ if and only if

$$[\Gamma^\bullet(L\beta^{-1} \cap W) : H^\bullet \cap \Gamma^\bullet(L\beta^{-1})] = [\Gamma(L\beta^{-1} \cap W) : H \cap \Gamma(L\beta^{-1})].$$

(4) *If n is even, then* $[\Gamma^\bullet(L\beta^{-1} \cap W) : \Gamma(L\beta^{-1} \cap W)] = 2$ *for every $\beta \in B$.*

(5) *If L is $\mathbf{Z}$-maximal with respect to φ, then $L\beta^{-1} \cap W$ belongs to the genus of $L \cap W$ for every $\beta \in B$.*

Proof. (1) If $\gamma \in H^\bullet \cap \Gamma^\bullet(L\beta^{-1})$, then $\gamma \in \Gamma(L\beta^{-1})$ by the assumption. Hence γ belongs to $H \cap \Gamma(L\beta^{-1})$.

(2) Take $\gamma \in H^\bullet \cap \Gamma^\bullet(L\beta^{-1})$ so that $\gamma \notin H$. Clearly $\gamma \in \Gamma^\bullet(L\beta^{-1} \cap W)$. If it belongs to $\Gamma(L\beta^{-1} \cap W)$, then $\gamma \in H$, which is a contradiction. This proves that $[H^\bullet \cap \Gamma^\bullet(L\beta^{-1}) : H \cap \Gamma(L\beta^{-1})] = 2$ implies $[\Gamma^\bullet(L\beta^{-1} \cap W) : \Gamma(L\beta^{-1} \cap W)] = 2$.

(3) We see that $h\beta\delta \in h\beta\Gamma(L)$ if and only if $\beta\gamma\delta^{-1}\beta^{-1} \in H^\bullet \cap \Gamma^\bullet(L\beta^{-1})$, which is so if and only if $[H^\bullet \cap \Gamma^\bullet(L\beta^{-1}) : H \cap \Gamma(L\beta^{-1})] = 2$. This proves the first assertion. To prove (ii) for $\beta \in B_1$, suppose that $[\Gamma^\bullet(L\beta^{-1} \cap W) : \Gamma(L\beta^{-1} \cap W)] = 1$. Then $\Gamma^\bullet(L\beta^{-1} \cap W)$ must coincide with $H^\bullet \cap \Gamma^\bullet(L\beta^{-1})$, and so $\Gamma(L\beta^{-1} \cap W) = H^\bullet \cap \Gamma^\bullet(L\beta^{-1})$, in contradiction with (i). Hence (ii) is valid for $\beta \in B_1$. To prove (ii) in the remaining case, let $\beta \in B_2$ and $\beta' \in B_2^\delta$ be such that $h\beta\delta \in h\beta'\Gamma(L)$. We easily see that $L(\beta')^{-1} \cap W = (L\beta^{-1} \cap W)\eta$, with some $\eta \in H^\bullet$. In view of this, we only have to prove (ii) for $\beta \in B_2$. Then the assertion immediately follows from the definition of B_{21}. Since $H^\bullet \cap \Gamma^\bullet(L\beta^{-1}) = H \cap \Gamma(L\beta^{-1})$ by (i), we have the last assertion.

(4) Fixing a basis of V, we employ the matrix representation of φ, where we identify V with $\mathbf{Q}_n^1$. Let $L\beta^{-1} \cap W = \mathbf{Z}_{n-1}^1 k_\beta$ with $k_\beta \in \mathbf{Q}_n^{n-1}$ and put $\sigma = \left[\begin{smallmatrix} h \\ k_\beta \end{smallmatrix}\right]$; clearly $\sigma \in \mathrm{GL}_n(\mathbf{Q})$ since $L\beta^{-1} \cap W$ has rank $n-1$. Then the mapping $x\sigma \mapsto x$ gives an isomorphism $(V, \varphi) \cong (\mathbf{Q}_n^1, \sigma\varphi \cdot {}^{\mathsf{T}}\sigma)$ that maps $L\beta^{-1} \cap W$ to $\sum_{i=2}^n \mathbf{Z}e_i$, where $\{e_i\}_{i=1}^n$ is the standard basis of $\mathbf{Q}_n^1$. Observe that $\sigma\varphi \cdot {}^{\mathsf{T}}\sigma = \mathrm{diag}[q, \psi_\beta]$ with $\psi_\beta = k_\beta \varphi \cdot {}^{\mathsf{T}}k_\beta$, $\sigma H^\bullet \sigma^{-1} = \{\gamma \in O(\sigma\varphi \cdot {}^{\mathsf{T}}\sigma) \mid e_1\gamma = e_1\}$, and

$$\sigma\Gamma^\bullet(L\beta^{-1} \cap W)\sigma^{-1} = \left\{\mathrm{diag}[1, \gamma_0] \;\middle|\; \gamma_0 \in O(\psi_\beta) \cap \mathrm{GL}_{n-1}(\mathbf{Z})\right\}.$$

Now we find that $\mathrm{diag}[1, -1_{n-1}] \in \sigma\Gamma^\bullet(L\beta^{-1} \cap W)\sigma^{-1}$. Since $n-1$ is odd, this proves (4).

(5) By [M2, Lemma 4.1] and since L is assumed maximal, the mapping $h\beta\Gamma(L) \mapsto (L \cap (\mathbf{Q}h\beta)^\perp)\beta^{-1}H$ defines a well-defined map of $L[q, \mathfrak{b}]/\Gamma(L)$ into the set of all H-classes in the H-genus of $L \cap W$. Then $(L \cap (\mathbf{Q}h\beta)^\perp)\beta^{-1} = L\beta^{-1} \cap W$, which proves (5). □

Let φ_0 be the matrix that represents φ with respect to a $\mathbf{Z}$-basis of L. With this basis, we identify φ with φ_0, V with $\mathbf{Q}_n^1$, and L with $\mathbf{Z}_n^1$, respectively; we may put $\varphi(x, y) = x\varphi \cdot {}^{\mathsf{T}}y$ for $x, y \in V$. Then

$$L[q, \mathfrak{b}] = \{x \in \mathbf{Q}_n^1 \mid x\varphi \cdot {}^{\mathsf{T}}x = q,\ x\varphi\mathbf{Z}_1^n = \mathfrak{b}\}.$$

The groups $O(\varphi)$ and $\Gamma^\bullet(L)$ are defined by

$$O(\varphi) = \{\gamma \in \mathrm{GL}_n(\mathbf{Q}) \mid \gamma\varphi \cdot {}^{\mathsf{T}}\gamma = \varphi\}, \quad \Gamma^\bullet(L) = O(\varphi) \cap \mathrm{GL}_n(\mathbf{Z}).$$

Also $\mathrm{SO}(\varphi) = O(\varphi) \cap SL_n(\mathbf{Q})$ and $\Gamma(L) = \mathrm{SO}(\varphi) \cap \mathrm{GL}_n(\mathbf{Z})$, which is often denoted by $\Gamma(\varphi)$. A matrix is called *integral* if all the entries are in $\mathbf{Z}$. A symmetric matrix is called *semi-integral* if the diagonal entries are in $\mathbf{Z}$ and the others in $2^{-1}\mathbf{Z}$. We note that φ_0 is semi-integral if L is integral with respect to φ. Supposing φ is semi-integral under the identification above, we denote by $\ell(\varphi)$ the *level* of φ; namely, $\ell(\varphi)$ is the smallest positive integer N such that $N(2\varphi)^{-1}$ is an integral matrix whose diagonal entries are all even.

For a symmetric matrix $S \in \mathrm{GL}_n(\mathbf{Q})$ we put

$$\Omega(S) = \{\gamma \in \mathrm{GL}_n(\mathbf{Q}) \mid {}^{\mathsf{T}}\gamma S\gamma = S\}, \quad \Delta(S) = \Omega(S) \cap \mathrm{GL}_n(\mathbf{Z}). \tag{1-2}$$

We easily see that $\Omega(\hat{S}) = O(S)$ and the map $\gamma \mapsto \gamma^{-1}$ (resp. $\gamma \mapsto {}^{\mathsf{T}}\gamma$) gives an isomorphism $O(S) \cong \Omega(\hat{S})$ (resp. $O(S) \cong \Omega(S)$). Supposing that S is integral, we set

$$B(S, t) = \{x \in \mathbf{Q}_1^n \mid {}^{\mathsf{T}}xSx = t,\ \mathbf{Z}_n^1 x = \mathbf{Z}\}$$

with $t \in \mathbf{Z}$. This is stable under the map $x \mapsto \gamma x$ for $\gamma \in \Delta(S)$.

Lemma 1.2. *Put* $\mathfrak{b} = c\mathbf{Z}$ *and* $\ell = 2^{-1}\ell(\varphi)$ *in the above setting. Also put* $S = \ell\hat{\varphi}$, $t = c^{-2}\ell q$, *and* $C_\beta = {}^{\mathsf{T}}(c^{-1}h\beta\varphi)$ *for* $\beta \in B$. *Then the mapping* $x \mapsto {}^{\mathsf{T}}(c^{-1}x\varphi)$ *defines a bijection of* $L[q, c\mathbf{Z}]$ *onto* $B(S,t)$. *This also gives a well-defined bijection of* $L[q, c\mathbf{Z}]/\,\Gamma^*(L)$ *onto* $\Delta(S)\backslash B(S,t)$. *Consequently, we have*

$$L[q, c\mathbf{Z}] = \bigsqcup_{\beta \in B_1 \sqcup B_2} h\beta\Gamma^*(L), \quad B(S,t) = \bigsqcup_{\beta \in B_1 \sqcup B_2} \Delta(S)C_\beta.$$

Here B_1 *and* B_2 *are as in Lemma 1.1.*

Proof. We first observe that $\hat{\varphi}[{}^{\mathsf{T}}(c^{-1}x\varphi)] = c^{-2}q$ and $\mathbf{Z}_n^1 \cdot {}^{\mathsf{T}}(c^{-1}x\varphi) = \mathbf{Z}$ for $x \in L[q, c\mathbf{Z}]$. It follows that ${}^{\mathsf{T}}(c^{-1}x\varphi) \in \mathbf{Z}_1^n$. Since $\ell\hat{\varphi}$ is integral by the definition of $\ell(\varphi)$, we have $c^{-2}\ell q \in \mathbf{Z}$. Hence our map is well defined and bijective. Furthermore, we have ${}^{\mathsf{T}}(c^{-1}x\varepsilon\varphi) = \varepsilon^{-1} \cdot {}^{\mathsf{T}}(c^{-1}x\varphi)$ and $\varepsilon^{-1} \in \Delta(\ell\hat{\varphi})$ for $\varepsilon \in \Gamma^*(L)$ since $\Gamma^*(L) = O(\varphi) \cap \mathrm{GL}_n(\mathbf{Z})$ in the present setting. Thus the map $x\Gamma^*(L) \mapsto \Delta(S) \cdot {}^{\mathsf{T}}(c^{-1}x\varphi)$ is also well defined and bijective, which proves the second assertion. Now, the $\Gamma(L)$-orbits of $L[q, \mathfrak{b}]$ can be found after (1-1). From this and the definitions of B_{2i} and B_{2i}^δ if $[\Gamma^*(L) : \Gamma(L)] = 2$ we can see the $\Gamma^*(L)$-orbits of $L[q, \mathfrak{b}]$ in the statement. For example, if $[\Gamma^*(L) : \Gamma(L)] = 2$, then

$$\begin{aligned} L[q, \mathfrak{b}] &= \bigsqcup_{\beta \in B_1 \sqcup B_2 \sqcup B_2^\delta} h\beta\Gamma(L) \\ &= \bigsqcup_{\beta \in B_1} h\beta\Gamma^*(L) \sqcup \bigsqcup_{\beta \in B_2} h\beta\Gamma(L) \sqcup \bigsqcup_{\beta' \in B_2^\delta} h\beta'\Gamma(L) \\ &= \bigsqcup_{\beta \in B_1 \sqcup B_2} h\beta\Gamma^*(L). \end{aligned}$$

The $\Delta(S)$-orbits of $B(S,t)$ follow from the bijection of $L[q, c\mathbf{Z}]/\,\Gamma^*(L)$ above. □

1.2. ***The symmetric space*** $\mathcal{Z}$**.** Keeping the notation of the previous section, assume that φ is indefinite. We put $\mathrm{SO}(\varphi)_{\boldsymbol{a}} = \{\alpha \in SL_n(\mathbf{R}) \mid \alpha\varphi \cdot {}^{\mathsf{T}}\alpha = \varphi\}$. By a Witt decomposition, φ can be represented by

$$\text{(1-3)} \qquad \begin{bmatrix} 0 & 0 & -1_r \\ 0 & \theta & 0 \\ -1_r & 0 & 0 \end{bmatrix}, \quad {}^{\mathsf{T}}\theta = \theta \in \mathrm{GL}_t(\mathbf{R}),$$

with $r > 0$, where θ is positive or negative definite. If $t = 0$, we ignore θ. We take $\kappa \in \mathbf{Q}^\times$ so that $\kappa\varphi$ has signature $(r + t, r)$. Also take $\sigma \in \mathrm{GL}_n(\mathbf{R})$ so that

$$\text{(1-4)} \qquad \kappa\sigma\varphi \cdot {}^{\mathsf{T}}\sigma \text{ is of the form (1-3) with } 0 < \theta \in \mathrm{GL}_t(\mathbf{R}).$$

We put $\varphi' = \kappa\sigma\varphi \cdot {}^{\mathsf{T}}\sigma$. Then following [S2, (16.3b)], we define a set $\mathcal{Z}^\varphi$ by

$$\mathcal{Z}^\varphi = \mathcal{Z}(r, \theta) = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}_r^{r+t} \;\middle|\; x \in \mathbf{R}_r^r,\ y \in \mathbf{R}_r^t,\ {}^{\mathsf{T}}x + x > {}^{\mathsf{T}}y\theta^{-1}y \right\}.$$

Notice that $\mathcal{Z}^\varphi$ depends on the choice of (1-4) above.

For $\alpha \in O(\varphi')_{\boldsymbol{a}}$ and $z \in \mathcal{Z}^{\varphi}$ we can define $\alpha(z) \in \mathcal{Z}^{\varphi}$ in the manner of [S2, §16.3]. Then by [S2, Proposition 16.6(i), (iii), and (v)], $\mathrm{SO}(\varphi')_{\boldsymbol{a}}$ acts transitively on $\mathcal{Z}^{\varphi}$, $C' = \{\gamma \in \mathrm{SO}(\varphi')_{\boldsymbol{a}} \mid \gamma(\mathbb{1}) = \mathbb{1}\}$ is a maximal compact subgroup of $\mathrm{SO}(\varphi')_{\boldsymbol{a}}$, and $\{\alpha \in \mathrm{SO}(\varphi')_{\boldsymbol{a}} \mid \alpha(z) = z \text{ for every } z \in \mathcal{Z}^{\varphi}\} = \mathrm{SO}(\varphi')_{\boldsymbol{a}} \cap \{\pm 1_n\}$, where $\mathbb{1} = \left[\begin{smallmatrix} 1_r \\ 0 \end{smallmatrix}\right]$. Since $\mathrm{SO}(\kappa\varphi) = \mathrm{SO}(\varphi)$ and $\sigma\,\mathrm{SO}(\varphi)\sigma^{-1} = \mathrm{SO}(\varphi')$, we can define the action of $\mathrm{SO}(\varphi)_{\boldsymbol{a}}$ on $\mathcal{Z}^{\varphi}$ by $\alpha(z) = (\sigma\alpha\sigma^{-1})(z)$ for $\alpha \in \mathrm{SO}(\varphi)_{\boldsymbol{a}}$ and $z \in \mathcal{Z}^{\varphi}$. Hence $\mathrm{SO}(\varphi)$ acts on $\mathcal{Z}^{\varphi}$ via the projection of $\mathrm{SO}(\varphi)$ into $\mathrm{SO}(\varphi)_{\boldsymbol{a}}$.

The space $\mathcal{Z}^{\varphi}$ can be viewed as a Riemannian manifold with the $\mathrm{SO}(\varphi)_{\boldsymbol{a}}$-invariant metric ds^2 defined by [S2, (16.8c)]. Furthermore, the $\mathrm{SO}(\varphi)_{\boldsymbol{a}}$-action is transitive. Hence $\mathcal{Z}^{\varphi}$ is a symmetric space on which $\mathrm{SO}(\varphi)$ acts; see [M3, Lemma 2.1] for details with Corrections at the end of this paper. We define the $\mathrm{SO}(\varphi)_{\boldsymbol{a}}$-invariant measure on $\mathcal{Z}^{\varphi}$ by $\boldsymbol{d}z = \delta(z)^{-n/2} \prod_{i=1}^{r+t} \prod_{j=1}^{r} dz_{ij}$, where $z = [z_{ij}] = \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in \mathcal{Z}^{\varphi}$, $\delta(z) = \det(2^{-1}({}^{\mathrm{t}}x + x - {}^{\mathrm{t}}y\theta^{-1}y))$, and $\prod_{i=1}^{r+t} \prod_{j=1}^{r} dz_{ij}$ denotes the standard measure on $\mathbf{R}^{r(r+t)}$ $(= \mathbf{R}^1_{r(r+t)})$.

Let D be an open subgroup of $\mathrm{SO}(\varphi)_{\boldsymbol{A}}$ such that D contains $\mathrm{SO}(\varphi)_{\boldsymbol{a}}$ and $\mathrm{SO}(\varphi)_{\boldsymbol{h}} \cap D$ is compact, where $\mathrm{SO}(\varphi)_{\boldsymbol{A}}$ denotes the adelization of $\mathrm{SO}(\varphi)$ and $\mathrm{SO}(\varphi)_{\boldsymbol{h}}$ its nonarchimedean part. Put $\Gamma = \mathrm{SO}(\varphi) \cap D$. Since Γ is discrete in $\mathrm{SO}(\varphi)_{\boldsymbol{a}}$, Γ acts properly discontinuously on $\mathcal{Z}^{\varphi}$. Denote by $\mathrm{vol}(\Gamma\backslash\mathcal{Z}^{\varphi})$ the measure of $\Gamma\backslash\mathcal{Z}^{\varphi}$ with respect to $\boldsymbol{d}z$ and assume that $\mathrm{vol}(\Gamma\backslash\mathcal{Z}^{\varphi})$ is finite. Then put

$$\nu(\Gamma) = [\Gamma \cap \{\pm 1_n\} : 1]^{-1} \mathrm{vol}(\Gamma\backslash\mathcal{Z}^{\varphi}). \tag{1-5}$$

Let us take D by $C(L) = \{\alpha \in \mathrm{SO}(\varphi)_{\boldsymbol{A}} \mid L\alpha = L\}$ with $L = \mathbf{Z}^1_n$. Clearly $\Gamma(L) = \mathrm{SO}(\varphi) \cap C(L) = \Gamma(\varphi)$. Then we may assume that $\nu(\Gamma(c\varphi)) = \nu(\Gamma(\varphi))$ for any $0 < c \in \mathbf{R}$. In fact, $c\varphi$ has the form in (1-4) with $c^{-1/2}\sigma$ in place of σ; hence $\mathcal{Z}^{c\varphi} = \mathcal{Z}^{\varphi}$. Noticing that the $O(\varphi)$-action does not depend on this change of σ, we have the desired fact.

The mass of $L[q, \mathfrak{b}]$ is defined by (0-5) with the $\Gamma(L)$-orbits of $L[q, \mathfrak{b}]$ as in (1-1). As for $\nu(H \cap \beta\Gamma(L)\beta^{-1})$, if ψ is definite, we put $\nu(H \cap \beta\Gamma(L)\beta^{-1}) = [H \cap \beta\Gamma(L)\beta^{-1} : 1]^{-1}$. In this case we understand that the space $\mathcal{Z}^{\psi}$ consists of a single point, which can be denoted by $\mathcal{Z}(0, \kappa_1\psi)$ with $\kappa_1 \in \{\pm 1\}$. If ψ is indefinite, we have to realize $\mathcal{Z}^{\psi}$ in the following procedure: We first represent ψ by a matrix ψ_0 with a fixed basis of W; we then define $\mathcal{Z}^{\psi}$ by $\mathcal{Z}^{\psi_0}$ with a fixed matrix equivalent to ψ_0 in the same manner as in (1-4). Since $H \cap \beta\Gamma(L)\beta^{-1}$ is identified with the subgroup of $\mathrm{SO}(\psi_0)$ with respect to the fixed basis, we define $\nu(H \cap \beta\Gamma(L)\beta^{-1})$ in the manner of (1-5). It can be verified that the action of H on $\mathcal{Z}^{\psi}$ does not depend on the choice of matrix representation ψ_0.

1.3. ***The surfaces $\mathcal{P}(S, T)$ and $\mathcal{P}(S, C, T_1)$.*** We recall some notation from [Si3]. For a symmetric matrix S in $\mathrm{GL}_n(\mathbf{Q})$, the orthogonal group $\Omega(S)_{\boldsymbol{a}} = \{\gamma \in \mathbf{R}^n_n \mid {}^{\mathrm{t}}\gamma S\gamma = S\}$ acts on $\mathrm{GL}_n(\mathbf{R})$ via $X \mapsto \gamma X$ for $\gamma \in \Omega(S)_{\boldsymbol{a}}$. Let $\mathcal{F}_S$ be the mapping

$X \mapsto S[X]$ of $\mathrm{GL}_n(\mathbf{R})$ into the set of symmetric elements T of $\mathrm{GL}_n(\mathbf{R})$. We consider the n^2 entries of X as differentiable functions of the $n(n+1)/2$ independent variables in T and of $n(n-1)/2$ new independent variables $x_1, \ldots, x_{n(n-1)/2}$. This is valid on a suitable neighborhood of X_0 for each $X_0 \in \mathrm{GL}_n(\mathbf{R})$ by virtue of the inverse function theorem. Let $J_S(T, x)$ be the Jacobian of this transformation of variables.

For a symmetric matrix T in $\mathrm{GL}_n(\mathbf{R})$ with the same signature as S we have the $n(n-1)/2$-dimensional surface $\mathcal{P}(S, T) = \mathcal{F}_S^{-1}(T)$. The group $\Omega(S)_{\boldsymbol{a}}$ acts transitively on this surface. Notice that $J_S(T, x)$ is a function of the coordinates $x = (x_1, \ldots, x_{n(n-1)/2})$. We then define a volume element on $\mathcal{P}(S, T)$ by

$$dv = |\det(TS^{-1})|^{1/2} |J_S(T, x)| \prod_{i=1}^{n(n-1)/2} dx_1 \cdots dx_{n(n-1)/2}. \tag{1-6}$$

This is invariant under the action of $\Omega(S)_{\boldsymbol{a}}$. Also dv does not depend on the choice of T. Let $\Delta(S)$ be as in (1-2). This is a discrete subgroup in $\Omega(S)_{\boldsymbol{a}}$ and acts discontinuously on $\mathcal{P}(S, T)$. We denote by $\rho(S)$ the volume of $\Delta(S)\backslash\mathcal{P}(S, T)$ computed with dv.

Let us fix a nonzero vector $C \in \mathbf{Q}_1^n$ and put $S[C] = t \in \mathbf{Q}$. Set $X = (C \;\; Y)$ in $\mathrm{GL}_n(\mathbf{R})$ with a variable matrix $Y \in \mathbf{R}_m^n$, where $m = n - 1$. The subgroup

$$\Omega(S, C)_{\boldsymbol{a}} = \{\gamma \in \Omega(S)_{\boldsymbol{a}} \mid \gamma C = C\}$$

of $\Omega(S)_{\boldsymbol{a}}$ acts on the set of such elements X. Also $\mathcal{F}_S$ maps X to a symmetric element $T_1 \in \mathrm{GL}_n(\mathbf{R})$ with the same signature as S, which is denoted in the form

$$T_1 = \begin{bmatrix} t & Q \\ {}^{\mathrm{T}}Q & R \end{bmatrix} \quad \text{with } Q \in \mathbf{R}_m^1 \text{ and } {}^{\mathrm{T}}R = R \in \mathbf{R}_m^m. \tag{1-7}$$

Noticing that Y is related to (Q, R) by the equations ${}^{\mathrm{T}}CSY = Q$ and $S[Y] = R$, in a similar way to the case of $\mathrm{GL}_n(\mathbf{R})$ above we consider the nm entries of Y as differentiable functions of the $m(n+2)/2$ independent variables in T_1 and of $m(m-1)/2$ new independent variables $y_1, \ldots, y_{m(m-1)/2}$. Let $J_S^C(T_1, y)$ be the Jacobian of this transformation.

Let T_1 be as above. We have the $m(m-1)/2$-dimensional surface $\mathcal{P}(S, C, T_1) = \mathcal{F}_S^{-1}(T_1) \cap \{(C \;\; Y) \mid Y \in \mathbf{R}_m^n\}$ on which $\Omega(S, C)_{\boldsymbol{a}}$ acts transitively; namely,

$$\mathcal{P}(S, C, T_1) = \left\{(C \;\; Y) \in \mathrm{GL}_n(\mathbf{R}) \;\middle|\; Y \in \mathbf{R}_{n-1}^n,\ S[(C \;\; Y)] = T_1\right\}.$$

With the function $J_S^C(T_1, y)$ of the coordinates $y = (y_1, \ldots, y_{m(m-1)/2})$, we define the $\Omega(S, C)_{\boldsymbol{a}}$-invariant volume element on $\mathcal{P}(S, C, T_1)$ by

$$dv_1 = |\det(T_1 S^{-1})|^{1/2} |J_S^C(T_1, y)| \prod_{i=1}^{m(m-1)/2} dy_1 \cdots dy_{m(m-1)/2}. \tag{1-8}$$

This does not depend on the choice of T_1. Now, we put

$$\Omega(S, C) = \Omega(S, C)_{\boldsymbol{a}} \cap \mathrm{GL}_n(\mathbf{Q}), \quad \Delta(S, C) = \Omega(S, C) \cap \mathrm{GL}_n(\mathbf{Z}).$$

Then we denote by $\rho(S, C)$ the volume of $\Delta(S, C)\backslash\mathcal{P}(S, C, T_1)$ computed with dv_1.

It is known that $\rho(S)$ is finite if $n > 2$ or if $n = 2$ and the quadratic form represented by S is anisotropic. Also if S is integral and C is primitive, then $\rho(S, C)$ is related to $\rho(S')$ with a suitable symmetric element S' of $\mathrm{GL}_{n-1}(\mathbf{Q})$. These facts can be found in [Si3, (110) and (11)], for example. We note that $\rho(S)$ and $\rho(S, C)$ are independent of the choice of T and T_1, respectively. Also $\rho(S, C)$ is independent of the choice of a representative in the $\Delta(S)$-orbit of C.

2. Proof of the Theorem

2.1. ***The volumes $\rho(S)$ and $\rho(S, C_\beta)$.*** Throughout this and the next sections, let the notation be as in Section 1.1. Assume that the symmetric matrix $\varphi \in \mathrm{GL}_n(\mathbf{Q})$ is indefinite and semi-integral, and that

(2-1) $\qquad \psi$ is anisotropic if $n = 3$.

We note that (2-1) holds if and only if $-q\det(\varphi) \notin \mathbf{Q}^{\times 2}$. Put $m = n - 1 \geq 2$, and recall that $L = \mathbf{Z}_n^1$. With the notation of (1-1) we set

$$L\beta^{-1} \cap W = \mathbf{Z}_m^1 k_\beta, \quad \psi_\beta = k_\beta \varphi \cdot {}^{\mathrm{T}}k_\beta,$$

with $k_\beta \in \mathbf{Q}_n^m$.

We start with the $\Delta(\ell\hat{\varphi})$-orbits of $B(\ell\hat{\varphi}, c^{-2}\ell q)$ given in Lemma 1.2. Set

$$S = \ell\hat{\varphi}, \quad t = c^{-2}\ell q, \quad C_\beta = {}^{\mathrm{T}}(c^{-1}h\beta\varphi) \ \text{ for } \beta \in B.$$

Then the measure of primitive representations of t by S is defined by

$$\nu(S, t) = \sum_{\beta \in B_1 \sqcup B_2} \rho(S, C_\beta)/\rho(S).$$

To study the relationship between $\rho(S)$ and $\nu(\Gamma(L))$, we define $\nu(\Gamma(L))$ with $\mathcal{Z}^\varphi = \mathcal{Z}(r, \theta)$ and $\boldsymbol{d}z$ in the setting of (1-4). Since $\hat{\varphi} = \varphi^{-1}\varphi\hat{\varphi}$, a matrix equivalent to $\hat{\varphi}$ over $\mathbf{R}$ is given by (1-3). Then $\mathcal{Z}^{\hat{\varphi}} = \mathcal{Z}^\varphi = \mathcal{Z}(r, \theta)$ and the $O(\varphi)$-action on $\mathcal{Z}^\varphi$ is the same as the $O(\hat{\varphi})$-action on $\mathcal{Z}^{\hat{\varphi}}$ under the isomorphism $O(\varphi) \cong O(\hat{\varphi})$ via $\gamma \mapsto \varphi^{-1}\gamma\varphi$.

Consider the isomorphism $x \mapsto x\varphi$ of $(\mathbf{Q}_n^1, \varphi)$ onto $(\mathbf{Q}_n^1, \hat{\varphi})$ over $\mathbf{Q}$. This maps L to $\mathbf{Z}_n^1\varphi$ and also $\widetilde{L}$ to $2^{-1}\mathbf{Z}_n^1$ because $\widetilde{L} = \mathbf{Z}_n^1(2\varphi)^{-1}$. Since $\Gamma(\widetilde{L}) = \Gamma(L)$ in $\mathrm{SO}(\varphi)$, it follows that $\Gamma(\mathbf{Z}_n^1) = \Gamma(\mathbf{Z}_n^1\varphi)$ in $\mathrm{SO}(\hat{\varphi})$. Consequently we have $\Gamma(\varphi) \cong \Gamma(\hat{\varphi})$, and hence $\nu(\Gamma(\varphi)) = \nu(\Gamma(\hat{\varphi}))$ in the above setting. Further, in view of the fact mentioned

after (1-5), we have $\nu(\Gamma(\hat{\varphi})) = \nu(\Gamma(2^{-1}\ell\hat{\varphi}))$. Applying [M3, Proposition 3.3] to the form $2^{-1}\ell\hat{\varphi}$ with $\kappa \in \{\pm 1\}$, we have thus

Lemma 2.1. *Put* $c_\varphi = 2^{rn/2}\det(\theta)^{r/2}\rho_{r+t}^{-1}\rho_r^{-1}$ *with the notation of* (1-4). *Then*

$$\nu(\Gamma(L)) = c_\varphi |\det(\ell\hat{\varphi})|^{(n+1)/2}\,[\Gamma^\bullet(L):\Gamma(L)]\,\rho(S).$$

As for $\rho(S, C_\beta)$, it is related to the volume $\rho(c^{-2}\ell\psi_\beta)$ as follows.

Lemma 2.2. *Assume that* φ *is indefinite and* (2-1). *For each* $h\beta \in L[q, c\mathbf{Z}]$ *with* $\beta \in B$ *we have*

$$\rho(S, C_\beta) = |c^{-2}\ell q|^{1-n}\,|\det(c^{-1}\varphi)|^n\,|\det({}^{\mathsf{T}}h\,{}^{\mathsf{T}}k_\beta)|^n \cdot \left[\Gamma^\bullet(L\beta^{-1}\cap W) : H^\bullet \cap \beta\Gamma^\bullet(L)\beta^{-1}\right]\rho(c^{-2}\ell\psi_\beta).$$

This will be proven in Section 3.

2.2. *The case of indefinite* ψ. We shall prove a relationship between $\nu(H \cap \beta\Gamma(L)\beta^{-1})$ and $\rho(c^{-2}\ell\psi_\beta)$ for $\beta \in B$ under (2-1).

We define $\mathcal{Z}^\psi$ by $\mathcal{Z}^{\psi_\beta}$ for each β. To be precise, let us first fix the matrix ψ_0 representing ψ with respect to a basis of W over $\mathbf{Q}$ and also fix a matrix equivalent to ψ_0 over $\mathbf{R}$ as follows:

$$\text{(2-2)} \qquad \kappa_1\sigma_0\psi_0\cdot{}^{\mathsf{T}}\sigma_0 = \begin{bmatrix} 0 & 0 & -1_{r_1} \\ 0 & \theta_1 & 0 \\ -1_{r_1} & 0 & 0 \end{bmatrix}, \quad 0 < {}^{\mathsf{T}}\theta_1 = \theta_1 \in \mathrm{GL}_{t_1}(\mathbf{R}),$$

where $r_1 > 0$, $\kappa_1 \in \{\pm 1\}$, and $\sigma_0 \in \mathrm{GL}_m(\mathbf{R})$; If $t_1 = 0$, we ignore θ_1. We define $\mathcal{Z}^\psi = \mathcal{Z}(r_1, \theta_1)$. Noticing $\mathbf{Q}^1_m k_\beta = W$, we next take $a_\beta \in \mathrm{GL}_m(\mathbf{Q})$ so that $\psi_0 = a_\beta\psi_\beta\cdot{}^{\mathsf{T}}a_\beta$. Since $\kappa_1\sigma_0 a_\beta\psi_\beta\cdot{}^{\mathsf{T}}(\sigma_0 a_\beta) = \kappa_1\sigma_0\psi_0\cdot{}^{\mathsf{T}}\sigma_0$, we have then $\mathcal{Z}^{\psi_\beta} = \mathcal{Z}(r_1, \theta_1) = \mathcal{Z}^\psi$. As mentioned in Section 1.2, the H-action on $\mathcal{Z}^\psi$ does not depend on $\beta \in B$.

Now, $H \cap \beta\Gamma(L)\beta^{-1}$ is contained in $\Gamma(L\beta^{-1}\cap W)$. Under the identification with respect to k_β, we can regard $L\beta^{-1}\cap W$ as $\mathbf{Z}^1_m$, H as $\mathrm{SO}(\psi_\beta)$, and $\Gamma(L\beta^{-1}\cap W)$ as $\Gamma(\mathbf{Z}^1_m) = \mathrm{SO}(\psi_\beta)\cap\mathrm{GL}_m(\mathbf{Z}) = \Gamma(\psi_\beta)$, respectively. Then by applying [S5, Theorem 10(iii)] we have

$$\nu(H\cap\beta\Gamma(L)\beta^{-1}) = \left[\Gamma(L\beta^{-1}\cap W) : H\cap\beta\Gamma(L)\beta^{-1}\right]\nu(\Gamma(\psi_\beta)).$$

Lemma 2.3. *Let the notation be as above. For each* $\beta \in B$ *we have*

$$\nu(H\cap\beta\Gamma(L)\beta^{-1}) = c_\psi|\det(c^{-2}\ell\psi_\beta)|^{n/2}\left[\Gamma^\bullet(L\beta^{-1}\cap W) : H\cap\beta\Gamma(L)\beta^{-1}\right]\cdot\rho(c^{-2}\ell\psi_\beta),$$

where $c_\psi = 2^{r_1 m/2}\det(\theta_1)^{r_1/2}\rho_{r_1+t_1}^{-1}\rho_{r_1}^{-1}$ *with the matrix fixed in* (2-2).

Proof. Noting that $\nu(\Gamma(\psi_\beta)) = \nu(\Gamma(2^{-1}c^{-2}\ell\psi_\beta))$, we apply Proposition 3.3 of [M3] to the matrix $2^{-1}c^{-2}\ell\psi_\beta$ to obtain our result.

In fact that proposition, with the notation used there, applies to the case $n > 2$, but actually it is valid even when $n = 2$ and φ is anisotropic. Indeed $\mathrm{vol}(\mathfrak{H}/\Delta)$ (see [M3, §3.2]) is finite for $n > 1$ except when $n = 2$ and S is isotropic, and then equality (3.5) of [M3] holds (see [Si3, (110)] for the proof). Consequently, the proposition is applicable to $2^{-1}c^{-2}\ell\psi_\beta$ even if $m = 2$ (with the present notation) under assumption (2-1). □

Proposition 2.4. *The quantity $\nu(H \cap \beta\Gamma(L)\beta^{-1})$ is related to the volume $\rho(S, C_\beta)$ for $\beta \in B$ as follows*:

$$\nu(H \cap \beta\Gamma(L)\beta^{-1}) = c_\psi |c^{-2}q|^{(n-2)/2}\ell^{(n-1)(n+2)/2}|\det(\varphi)|^{-n/2} \cdot \left[H^\bullet \cap \beta\Gamma^\bullet(L)\beta^{-1} : H \cap \beta\Gamma(L)\beta^{-1}\right]\rho(S, C_\beta).$$

Here c_ψ is as in Lemma 2.3.

Proof. Combining Lemmas 2.2 and 2.3, we obtain

$$\nu(H \cap \beta\Gamma(L)\beta^{-1}) = c_\psi |c|^{2-n}|\det(\ell\psi_\beta)|^{n/2}|\ell q|^{n-1}|\det(\varphi)|^{-n}|\det({}^{\mathrm{T}}h\,{}^{\mathrm{T}}k_\beta)|^{-n} \cdot \left[H^\bullet \cap \beta\Gamma^\bullet(L)\beta^{-1} : H \cap \beta\Gamma(L)\beta^{-1}\right]\rho(S, C_\beta).$$

Since $\varphi[({}^{\mathrm{T}}h\,{}^{\mathrm{T}}k_\beta)] = \mathrm{diag}[q, \psi_\beta]$, we have

$$|\det({}^{\mathrm{T}}h\,{}^{\mathrm{T}}k_\beta)| = \left|q\,\det(\psi_\beta)\,\det(\varphi)^{-1}\right|^{1/2}. \tag{2-3}$$

From these the desired relation follows. □

Now, it can be seen from Section 1.1 that

$$L[q, c\mathbf{Z}] = \begin{cases} \bigsqcup_{\beta \in B_1 \sqcup B_2} h\beta\Gamma(L) & \text{if } [\Gamma^\bullet(L) : \Gamma(L)] = 1, \\ \bigsqcup_{\beta \in B_1 \sqcup B_2} h\beta\Gamma(L) \sqcup \bigsqcup_{\beta' \in B_2^\delta} h\beta'\Gamma(L) & \text{if } [\Gamma^\bullet(L) : \Gamma(L)] = 2. \end{cases}$$

By employing Lemma 1.1(1) and (3) we find that

$$[H^\bullet \cap \Gamma^\bullet(L\beta^{-1}) : H \cap \Gamma(L\beta^{-1})] = \begin{cases} 2 & \text{if } [\Gamma^\bullet(L) : \Gamma(L)] = 2 \text{ and } \beta \in B_1, \\ 1 & \text{if } [\Gamma^\bullet(L) : \Gamma(L)] = 1 \text{ or if } [\Gamma^\bullet(L) : \Gamma(L)] = 2 \text{ and } \beta \in B_2 \cup B_2^\delta. \end{cases}$$

The $\Delta(S)$-orbits of $B(S, t)$ are given by C_β for $\beta \in B_1 \sqcup B_2$ by Lemma 1.2. We also note that $\rho(S, C_\beta) = \rho(S, C_{\beta'})$ if $\beta \in B_2$ and $\beta' \in B_2^\delta$ satisfy $h\beta\delta \in h\beta'\Gamma(L)$.

Then by Proposition 2.4 the quantity $\mathfrak{m}(L[q, c\mathbf{Z}])\nu(\Gamma(L))$ is given by

$$\text{(2-4)}\quad \sum_{\beta\in B}\nu(H\cap\beta\Gamma(L)\beta^{-1})$$
$$=\begin{cases}\sum_{\beta\in B_1\sqcup B_2}\nu(H\cap\beta\Gamma(L)\beta^{-1}) & \text{if } [\Gamma^{\bullet}(L):\Gamma(L)]=1,\\ \sum_{\beta\in B_1\sqcup B_2\sqcup B_2^{\delta}}\nu(H\cap\beta\Gamma(L)\beta^{-1}) & \text{if } [\Gamma^{\bullet}(L):\Gamma(L)]=2,\end{cases}$$
$$=c_{\psi}|c^{-2}q|^{(n-2)/2}\ell^{(n-1)(n+2)/2}|\det(\varphi)|^{-n/2}\,[\Gamma^{\bullet}(L):\Gamma(L)]\sum_{\beta\in B_1\sqcup B_2}\rho(S,C_{\beta}).$$

Therefore, from (2-4) and Lemma 2.1 we obtain (0-8). This completes the proof of the Theorem when ψ is indefinite.

2.3. ***The case of definite ψ.*** Since $\nu(H\cap\beta\Gamma(L)\beta^{-1})=[H\cap\beta\Gamma(L)\beta^{-1}:1]^{-1}$ by definition, the mass $\mathfrak{m}(L[q,c\mathbf{Z}])$ is then given by

$$\mathfrak{m}(L[q,c\mathbf{Z}])=\nu(\Gamma(L))^{-1}\sum_{\beta\in B}\left[H\cap\beta\Gamma(L)\beta^{-1}:1\right]^{-1}.$$

On the other hand, $\Omega(c^{-2}\ell\psi_{\beta})_{\boldsymbol{a}}$ is compact since ψ_{β} is definite. Then by using [Si3, Lemma 26] we see that

$$\int_{\Omega(c^{-2}\ell\psi_{\beta})_{\boldsymbol{a}}}du=\rho_{n-1}\left|\det(c^{-2}\ell\psi_{\beta})\right|^{-n/2},$$

where du is the volume element on $\Omega(c^{-2}\ell\psi_{\beta})_{\boldsymbol{a}}$ as in (1-6) with m and $c^{-2}\ell\psi_{\beta}$ in place of n and S there. Noticing that $\Delta(c^{-2}\ell\psi_{\beta})$ is a finite subgroup in $\Omega(c^{-2}\ell\psi_{\beta})_{\boldsymbol{a}}$, we thus have

$$\rho(c^{-2}\ell\psi_{\beta})=\rho_{n-1}\left|\det(c^{-2}\ell\psi_{\beta})\right|^{-n/2}[\Delta(c^{-2}\ell\psi_{\beta}):1]^{-1}.$$

Since $[\Delta(c^{-2}\ell\psi_{\beta}):1]=[\Gamma^{\bullet}(c^{-2}\ell\psi_{\beta}):1]=[\Gamma^{\bullet}(L\beta^{-1}\cap W):1]$, Lemma 2.2 combined with (2-3) shows that

$$\rho(S,C_{\beta})=\rho_{n-1}\left|c^{n-2}q^{(2-n)/2}\right|\ell^{1-n(n+1)/2}\,|\det(\varphi)|^{n/2}\left[H^{\bullet}\cap\Gamma^{\bullet}(L\beta^{-1}):1\right]^{-1}.$$

By the same way as in the indefinite case we can verify that the quantity $\mathfrak{m}(L[q,c\mathbf{Z}])\nu(\Gamma(L))$ is given by (2-4). From this together with Lemma 2.1 we obtain (0-8) with $r_1=0$, which completes the proof of the Theorem.

3. Proof of Lemma 2.2

We preserve the notation of Section 2.1, and we introduce a bit more, before proceeding to the proof.

If $X=[x_{ij}]$ is a variable matrix in $\mathbf{R}_q^p$, we put $\{dX\}=\prod_{i=1}^{p}\prod_{j=1}^{q}dx_{ij}$ on $\mathbf{R}^{pq}$. Similarly for symmetric X, we put $\{dX\}=\prod_{i\le j}dx_{ij}$.

3.1. Put $p = n - r$ with the matrix in (1-4). Then S has the signature $(p, n-p)$. By [Si3, Lemma 27] there is $\varepsilon \in \mathrm{GL}_n(\mathbf{Z})$ such that ${}^{\mathrm{T}}\varepsilon S\varepsilon = \left[\begin{smallmatrix} S_1 & * \\ * & * \end{smallmatrix}\right]$ with $0 < -S_1 \in \mathrm{GL}_{n-p}(\mathbf{Q})$. Since ${}^{\mathrm{T}}\varepsilon S\varepsilon[\varepsilon^{-1}X] = S[X]$ under the map $X \mapsto \varepsilon^{-1}X$ and $\Omega({}^{\mathrm{T}}\varepsilon S\varepsilon) = \varepsilon^{-1}\Omega(S)\varepsilon$, we see that $\rho(S) = \rho({}^{\mathrm{T}}\varepsilon S\varepsilon)$ and $\rho(S, C_\beta) = \rho({}^{\mathrm{T}}\varepsilon S\varepsilon, \varepsilon^{-1}C_\beta)$. Hence we may assume that $S = \left[\begin{smallmatrix} S_1 & * \\ * & * \end{smallmatrix}\right]$ with S_1 as above.

Put

$$\mathfrak{X}_\beta = \{(C_\beta\ X_1\ X_2) \in \mathrm{GL}_n(\mathbf{R}) \mid (X_1, X_2) \in \mathbf{R}^n_p \times \mathbf{R}^n_{m-p},\ S[X_1] > 0\},$$
$$\mathfrak{W} = \left\{(Q, R) \in \mathbf{R}^1_m \times \mathbf{R}^m_m \;\middle|\; \left[\begin{smallmatrix} t & Q \\ {}^{\mathrm{T}}Q & R \end{smallmatrix}\right] \text{ has the same signature as } S\right\},$$
$$\mathfrak{H} = \{0 < {}^{\mathrm{T}}P = P \in \mathbf{R}^n_n \mid PS^{-1}P = S\}.$$

Clearly, $\Omega(S, C_\beta)$ acts on $\mathfrak{X}_\beta$. Let $\Omega(S)$ act on $\mathfrak{H}$ by $P \mapsto P[\gamma]$ for $\gamma \in \Omega(S)$. It can be seen that $\mathfrak{H}$ is a symmetric space on which $\Omega(S)_a$ acts transitively. We consider the mapping $\mathcal{F}$ of $\mathfrak{X}_\beta$ into $\mathfrak{H} \times \mathfrak{W}$ defined by

$$\mathcal{F}:\ (C_\beta\ X_1\ X_2) \mapsto (P, Q, R),$$

where

$$S[(C_\beta\ X_1\ X_2)] = \left[\begin{smallmatrix} t & Q \\ {}^{\mathrm{T}}Q & R \end{smallmatrix}\right], \quad P = 2(S[X_1])^{-1}[{}^{\mathrm{T}}X_1 S] - S,$$

with

$$X_1 = \left[\begin{smallmatrix} Y \\ 1_p \end{smallmatrix}\right] X_0,\ Y \in \mathbf{R}^{n-p}_p,\ \text{and } X_0 \in \mathrm{GL}_p(\mathbf{R}).$$

By virtue of [Si3, Lemma 4] together with the assumption on S and S_1, $\mathcal{F}$ is well defined and Y bijectively corresponds to a point P on $\mathfrak{H}$ in the above way. Moreover, $U^{-1}(C_\beta\ X)$ corresponds to $P[U]$ for $U \in \Omega(S, C_\beta)$ under the correspondence $Y \mapsto P$, where $X = (X_1\ X_2)$. Let H_0 be a fundamental domain for $\Delta(S)$ in $\mathfrak{H}$. Then it can be verified that

$$H(C_\beta) = \bigsqcup_{\gamma \in \Delta(S)/\Delta(S, C_\beta)\{\pm 1_n\}} H_0[\gamma]$$

is a fundamental domain for $\Delta(S, C_\beta)$ in $\mathfrak{H}$.

Let D be a domain in $\mathfrak{W}$ and D_0 a domain in $\mathfrak{X}_\beta$ of all $(C_\beta\ X)$ such that $\mathcal{F}((C_\beta\ X)) \in H(C_\beta) \times D$. For $(Q, R) \in D$ let T_1 be as in (1-7) and consider the surface $\mathcal{P}(S, C_\beta, T_1)$. Since the $\Omega(S, C_\beta)$-action on $\mathcal{P}(S, C_\beta, T_1)$ corresponds to the $\Omega(S, C_\beta)$-action on $\mathfrak{H}$, the intersection $J(C_\beta, T_1)$ of D_0 with $\mathcal{P}(S, C_\beta, T_1)$ is a fundamental domain for $\Delta(S, C_\beta)\backslash\mathcal{P}(S, C_\beta, T_1)$. Then in the same way as in [Si3, (116)] we have

$$\int_{D_0} \{dX\} = |\det(S)|^{1/2} \int_D |\det(T_1)|^{-1/2} \rho(S, C_\beta)\{dQ\}\{dR\}, \tag{3-1}$$

where $\rho(S, C_\beta)$ is the volume of $J(C_\beta, T_1)$ computed with dv_1 of (1-8) with $\ell\hat{\varphi}$ and C_β in place of S and C.

3.2. We put $\delta = (C_\beta\ {}^\top(c^{-1}k_\beta\beta\varphi)) \in \mathrm{GL}_n(\mathbf{Q})$, $\Psi_\beta = \hat{\varphi}[\delta]$, $e_1 = \left[\begin{smallmatrix}1\\0_1^m\end{smallmatrix}\right]$, and

$$\mathfrak{Y}_\beta = \{(e_1\ Y_1\ Y_2) \in \mathrm{GL}_n(\mathbf{R}) \mid (Y_1, Y_2) \in \mathbf{R}_p^n \times \mathbf{R}_{m-p}^n, \ell\Psi_\beta[Y_1] > 0\}.$$

Since $W\beta = (\mathbf{Q}h\beta)^\perp$ in (V, φ) and $L \cap W\beta = \mathbf{Z}_m^1 k_\beta\beta$, we have

$$\Psi_\beta = c^{-2}\begin{bmatrix} q & 0 \\ 0 & \psi_\beta \end{bmatrix}.$$

Clearly $\Omega(\ell\Psi_\beta, e_1)$ acts on $\mathfrak{Y}_\beta$ and $\Omega(\ell\Psi_\beta, e_1) = \delta^{-1}\Omega(S, C_\beta)\delta$ since $\delta e_1 = C_\beta$. Observe that $\hat{\varphi}[\delta Y_1] > 0$ for $(e_1\ Y_1\ Y_2) \in \mathfrak{Y}_\beta$. Then under the transformation of variables

$$\mathfrak{Y}_\beta \ni (e_1\ Y) \mapsto \delta(e_1\ Y) \in \mathfrak{X}_\beta \quad \text{with } Y = (Y_1\ \ Y_2)$$

we have

$$\int_{D_0}\{dX\} = \int_{\delta^{-1}D_0} |\det(\delta)|^m\{dY\} = |\det(c^{-1}\varphi)|^m[L/N_\beta]^m \int_{\delta^{-1}D_0}\{dY\}, \tag{3-2}$$

where $N_\beta = \mathbf{Z}h + L\beta^{-1} \cap W$ and $[L/N_\beta] = |\det({}^\top h\ {}^\top k_\beta)|$.

To proceed in the integration over $\mathfrak{Y}_\beta$, we put

$$\mathfrak{H}_1 = \{0 < {}^\top P_1 = P_1 \in \mathbf{R}_n^n \mid P_1(\ell\Psi_\beta)^{-1}P_1 = \ell\Psi_\beta\}.$$

Then the mapping $P \mapsto P[\delta]$ gives a bijection of $\mathfrak{H}$ onto $\mathfrak{H}_1$ and $P[\gamma][\delta] = P[\delta][\delta^{-1}\gamma\delta]$ for $\gamma \in \Omega(S, C_\beta)$. Hence $H(C_\beta)[\delta]$ is a fundamental domain for $\delta^{-1}\Delta(S, C_\beta)\delta$ in $\mathfrak{H}_1$. Thus we can consider the mapping $\mathcal{F}_1$, similar to $\mathcal{F}$, of $\mathfrak{Y}_\beta$ into $\mathfrak{H}_1 \times \mathfrak{W}$ via $(e_1\ Y) \mapsto (P_1, Q, R)$. Notice that $P[\delta] = P_1$ if $\mathcal{F}((C_\beta\ X)) = (P, Q, R)$ and $\mathcal{F}_1(\delta^{-1}(C_\beta\ X)) = (P_1, Q, R)$. From this it follows that $\delta^{-1}D_0$ is a domain in $\mathfrak{Y}_\beta$ of all $(e_1\ Y)$ such that $\mathcal{F}_1((e_1\ Y)) \in H(C_\beta)[\delta] \times D$. Since $U_1^{-1}(e_1\ Y) \in \mathfrak{Y}_\beta$ corresponds to $P_1[U_1] \in \mathfrak{H}_1$ for $U_1 \in \Omega(\ell\Psi_\beta, e_1)$, the intersection $J(e_1, T_1)$ of $\delta^{-1}D_0$ with $\mathcal{P}(\ell\Psi_\beta, e_1, T_1)$ is a fundamental domain for $\delta^{-1}\Delta(S, C_\beta)\delta$ in this surface for a fixed T_1 with $(Q, R) \in D$. Thus we get

$$\int_{\delta^{-1}D_0}\{dY\} = |\det(\ell\Psi_\beta)|^{1/2}\int_D |\det(T_1)|^{-1/2}\rho(\tilde{\Delta})\{dQ\}\{dR\}. \tag{3-3}$$

Here $\tilde{\Delta} = \delta^{-1}\Delta(S, C_\beta)\delta$ and $\rho(\tilde{\Delta})$ is the volume of $J(e_1, T_1)$ computed with the volume element on $\mathcal{P}(\ell\Psi_\beta, e_1, T_1)$, which is defined in the manner of (1-8) with $\ell\Psi_\beta$ and e_1 in place of S and C. Combining (3-3) with (3-2) and (3-1), we obtain

$$\rho(S, C_\beta) = |\det(c^{-1}\varphi)|^n[L/N_\beta]^n\rho(\tilde{\Delta}). \tag{3-4}$$

3.3. To compute $\rho(\tilde{\Delta})$, let D^* be the projection of $\mathcal{F}_1(\delta^{-1}D_0)$ into $\mathfrak{W}$. Put

$$D' = \{Q \in \mathbf{R}_m^1 \mid (Q, R) \in D^* \text{ with } R \in \mathbf{R}_m^m\},$$
$$D(Q) = \{{}^{\mathrm{T}}R = R \in \mathbf{R}_m^m \mid (Q, R) \in D^*\} \quad \text{for } Q \in D'.$$

Then the integral in the right-hand side of (3-3) is given by

$$\int_{D'} \int_{D(Q)} |\det(T_1)|^{-1/2} \rho(\tilde{\Delta})\{dR\}\{dQ\}. \tag{3-5}$$

Furthermore, for $(e_1\ Y) \in \delta^{-1}D_0$ we put $Y = \left[\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right]$ with $y_1 \in \mathbf{R}_m^1$ and $y_2 \in \mathrm{GL}_m(\mathbf{R})$. Let D_1 be the set of all such y_1 and put

$$D(y_1) = \left\{y_2 \in \mathrm{GL}_m(\mathbf{R}) \;\middle|\; \left[\begin{smallmatrix} 1 & y_1 \\ 0 & y_2 \end{smallmatrix}\right] \in \delta^{-1}D_0\right\}$$

for $y_1 \in D_1$. Also put $R_1 = c^{-2}\ell\psi_\beta[y_2]$ for a given y_2. Then we observe that

$$\begin{bmatrix} t & Q \\ {}^{\mathrm{T}}Q & R \end{bmatrix} = \ell\Psi_\beta[(e_1\ Y)] = \begin{bmatrix} t & ty_1 \\ {}^{\mathrm{T}}(ty_1) & t[y_1] + c^{-2}\ell\psi_\beta[y_2] \end{bmatrix},$$
$$\{dQ\} = |t|^m\{dy_1\}, \quad \{dR\} = \{dR_1\} \quad \text{for fixed } Q \in D'.$$

It can be verified that $D' = tD_1$ and the map $R_1 \mapsto t[y_1] + R_1$ is a bijection of $E(y_1)$ onto $D(Q)$ for each $Q \in D'$, where $y_1 = t^{-1}Q$ and $E(y_1) = c^{-2}\ell\psi_\beta[D(y_1)]$. Moreover, $\det(T_1) = t\det(R_1)$ because

$$\begin{bmatrix} t & Q \\ {}^{\mathrm{T}}Q & R \end{bmatrix}\begin{bmatrix} 1 & -t^{-1}Q \\ 0 & 1_m \end{bmatrix} = \begin{bmatrix} t & 0 \\ {}^{\mathrm{T}}Q & R - t^{-1}[Q] \end{bmatrix}.$$

Hence under the transformations $y_1 \mapsto Q$ and $R_1 \mapsto R$, (3-5) is given by

$$\int_{D_1} \int_{E(y_1)} |t|^{-1/2} |\det(R_1)|^{-1/2} \rho(\tilde{\Delta})\{dR_1\}|t|^m\{dy_1\}.$$

Combining this with (3-3), we obtain

$$\int_{\delta^{-1}D_0} \{dY\} = |t|^m |\det(c^{-2}\ell\psi_\beta)|^{1/2} \int_{D_1} \int_{E(y_1)} |\det(R_1)|^{-1/2} \rho(\tilde{\Delta})\{dR_1\}\{dy_1\}. \tag{3-6}$$

On the other hand, we have

$$\int_{\delta^{-1}D_0} \{dY\} = \int_{D_1} \int_{D(y_1)} \{dy_2\}\{dy_1\}.$$

Since $c^{-2}\ell\psi_\beta$ maps y_2 to a symmetric element R_1 of $\mathrm{GL}_m(\mathbf{R})$, we consider the m^2 entries of y_2 as differentiable functions of the $m(m+1)/2$ variables in R_1 and of

new variables $u_1, \dots, u_{m(m-1)/2}$. We define the volume element du by (1-6) with m, $c^{-2}\ell\psi_\beta$, and R_1 in place of n, S, and T. Then we have

$$\int_{D(y_1)} \{dy_2\} = |\det(c^{-2}\ell\psi_\beta)|^{1/2} \int_{E(y_1)} |\det(R_1)|^{-1/2} \int_{J(Q,R_1)} du\{dR_1\},$$

where $J(Q, R_1)$ is the intersection of $D(y_1)$ with $\mathcal{P}(c^{-2}\ell\psi_\beta, R_1)$; here $Q = ty_1$. Combining these with (3-6), we get

$$\rho(\tilde{\Delta}) = |t|^{-m} \int_{J(Q,R_1)} du \tag{3-7}$$

for each R_1, that is, each $(Q, R) \in D^*$.

Now, we may choose $T_1 = \ell\Psi_\beta$ with $(Q, R) = (0^1_m, c^{-2}\ell\psi_\beta) \in D^*$, which corresponds to $R_1 = R = c^{-2}\ell\psi_\beta$. It can be seen that

$$\mathcal{P}(\ell\Psi_\beta, e_1, \ell\Psi_\beta) = \Omega(\ell\Psi_\beta, e_1)_{\boldsymbol{a}} = \left\{ \left[\begin{smallmatrix} 1 & 0 \\ 0 & y \end{smallmatrix}\right] \;\middle|\; y \in \Omega(c^{-2}\ell\psi_\beta)_{\boldsymbol{a}} \right\}.$$

Hence the mapping $\mathfrak{Y}_\beta \ni (e_1\ Y) \mapsto (0^m_1\ 1_m)(e_1\ Y) \cdot {}^{\mathsf{T}}(0^m_1\ 1_m) \in \mathrm{GL}_m(\mathbf{R})$ gives the bijective projection

$$\Omega(\ell\Psi_\beta, e_1)_{\boldsymbol{a}} \ni \left[\begin{smallmatrix} 1 & 0 \\ 0 & y \end{smallmatrix}\right] \mapsto y \in \Omega(c^{-2}\ell\psi_\beta)_{\boldsymbol{a}}. \tag{3-8}$$

Note that $J(0^1_m, c^{-2}\ell\psi_\beta) = D(0^1_m) \cap \Omega(c^{-2}\ell\psi_\beta)_{\boldsymbol{a}} = (0^m_1\ 1_m)\delta^{-1}D_0 \cdot {}^{\mathsf{T}}(0^m_1\ 1_m)$. Since the $\Omega(\ell\Psi_\beta, e_1)$-action on $\Omega(\ell\Psi_\beta, e_1)$ corresponds to the $\Omega(c^{-2}\ell\psi_\beta)$-action on $\Omega(c^{-2}\ell\psi_\beta)$ under the projection, $J(0^1_m, c^{-2}\ell\psi_\beta)$ is a fundamental domain for $\Delta\backslash\Omega(c^{-2}\ell\psi_\beta)_{\boldsymbol{a}}$, where Δ is the image of $\tilde{\Delta}$ under (3-8). Let $\rho(\Delta)$ be the volume of $J(0^1_m, c^{-2}\ell\psi_\beta)$ computed with du. Then by (3-7) we have $\rho(\tilde{\Delta}) = |t|^{-m}\rho(\Delta)$. Together with (3-4), this shows that

$$\rho(S, C_\beta) = |t|^{-m}|\det(c^{-1}\varphi)|^n [L/N_\beta]^n \rho(\Delta). \tag{3-9}$$

3.4. We observe that $\Omega(\hat{\varphi}, C_\beta) = O(\varphi, h\beta) = \beta^{-1}H^\cdot\beta$. Then

$$\Delta(S, C_\beta) = \beta^{-1}H^\cdot\beta \cap \Gamma^\cdot(L),$$

which is contained in $\Gamma^\cdot(L \cap W\beta)$. Since $\delta^{-1}\Omega(\hat{\varphi}, C_\beta)\delta = \Omega(\Psi_\beta, e_1)$, we have

$$\beta^{-1}H^\cdot\beta \cap \Gamma^\cdot(L) \cong \tilde{\Delta} \subset \Gamma^\cdot((L \cap W\beta)\delta).$$

Now, $(L \cap W\beta)\delta = \mathbf{Z}^1_m k_\beta\beta \cdot c^{-1}\varphi({}^{\mathsf{T}}(h\beta)\ {}^{\mathsf{T}}(k_\beta\beta)) = \mathbf{Z}^1_m(0^m_1\ c^{-1}\psi_\beta)$. From this it follows that

$$\Gamma^\cdot((L \cap W\beta)\delta) = \left\{ \left[\begin{smallmatrix} 1 & 0 \\ 0 & \gamma_0 \end{smallmatrix}\right] \;\middle|\; \gamma_0 \in \Omega(c^{-2}\ell\psi_\beta) \cap \mathrm{GL}_m(\mathbf{Z}) \right\}.$$

In fact, if $\gamma = \mathrm{diag}[1, \gamma_0] \in \Gamma^\cdot((L \cap W\beta)\delta)$ with $\gamma_0 \in \Omega(c^{-2}\ell\psi_\beta)$, then $(L \cap W\beta)\delta = (L \cap W\beta)\delta\gamma = \mathbf{Z}^1_m\widehat{\gamma_0}(0^m_1\ c^{-1}\psi_\beta)$. This shows that $\mathbf{Z}^1_m\widehat{\gamma_0} = \mathbf{Z}^1_m$, which proves

the desired fact. Hence the image of $\Gamma^\bullet((L\cap W\beta)\delta)$ under (3-8) is $\Delta(c^{-2}\ell\psi_\beta)$. Summing up these, we see that

$$[\Delta(c^{-2}\ell\psi_\beta):\Delta]=[\Gamma^\bullet(L\cap W\beta):\beta^{-1}H^\bullet\beta\cap\Gamma^\bullet(L)],$$
$$\rho(\Delta)=[\Gamma^\bullet(L\beta^{-1}\cap W):H^\bullet\cap\Gamma^\bullet(L\beta^{-1})]\rho(c^{-2}\ell\psi_\beta),$$

where $\rho(c^{-2}\ell\psi_\beta)$ is the volume of $\Delta(c^{-2}\ell\psi_\beta)\backslash\Omega(c^{-2}\ell\psi_\beta)_{\boldsymbol{a}}$ computed with du. This combined with (3-9) gives the desired relation, which completes the proof of Lemma 2.2.

4. A formula for $\nu(S,t)$

4.1. ***Statement.*** Given a $\mathbf{Z}$-lattice L in a vector space V over $\mathbf{Q}$, we denote by L_p the $\mathbf{Z}_p$-lattice in $V\otimes_{\mathbf{Q}}\mathbf{Q}_p$ generated by L over $\mathbf{Z}_p$. For the special orthogonal group G of a quadratic form defined on V, we denote by $G_{\boldsymbol{A}}$ the adelization of G and put $C(L)=\{\alpha\in G_{\boldsymbol{A}}\mid L\alpha=L\}$.

Theorem 4.1. *Let φ be an indefinite symmetric matrix in $\mathrm{GL}_n(\mathbf{Q})$ and set $L=\mathbf{Z}_n^1$ with $n>2$. Pick $h\in L[q,2^{-1}b\mathbf{Z}]$ for $q,b\in\mathbf{Q}^\times$. Set $W=(\mathbf{Q}h)^\perp$, $M=L\cap W$, and $H=\{\gamma\in\mathrm{SO}(\varphi)\mid h\gamma=h\}$; let ψ be the restriction of the quadratic form represented by φ to W. Assume that L is $\mathbf{Z}$-maximal with respect to φ, M is $\mathbf{Z}$-maximal with respect to ψ, and ψ satisfies* (2-1). *Set $S=\ell(\varphi)(2\varphi)^{-1}$ and $t=2b^{-2}\ell(\varphi)q$, where $\ell(\varphi)$ is the level of φ. Then the measure of primitive representations of t by S is given by*

$$\nu(S,t)=2\ell(\varphi)[\widetilde{L}:L]^{-1}[C(M):H_{\boldsymbol{A}}\cap C(L)]\prod_{p\mid\mathfrak{e}}(\lambda_p^\varphi)^{-1}\prod_{p\mid\mathfrak{e}_1}\lambda_p^\psi$$
$$\cdot\begin{cases}L((n-1)/2,\chi_1)\zeta(n-1)^{-1} & \text{if } n \text{ is odd},\\ L(n/2,\chi)^{-1} & \text{if } n \text{ is even}.\end{cases}$$

Here $\mathfrak{e}$ (resp. $\mathfrak{e}_1$) is the product of all the prime numbers for which $\widetilde{L}_p\neq L_p$ (resp. $\widetilde{M}_p\neq M_p$), ζ is the Riemann zeta function, $L(s,\chi)$ (resp. $L(s,\chi_1)$) is the L-function of the primitive Dirichlet character χ (resp. χ_1) corresponding to $\mathbf{Q}(\sqrt{(-1)^{n/2}\det(\varphi)})$ (resp. $\mathbf{Q}(\sqrt{(-1)^{(n-1)/2}\det(\psi)})$), and the constants λ_p^φ and λ_p^ψ are given as follows:

(1) *The case of odd n*:

$$\lambda_p^\varphi=\begin{cases}1 & \text{if } \det(\varphi)\in\mathbf{Z}_p^\times\mathbf{Q}_p^{\times2},\ t_p=1, \text{ and } p\mid 2,\\ 2^{-1}(1+p)^{-1}(1-p^{1-n}) & \text{if } \det(\varphi)\in\mathbf{Z}_p^\times\mathbf{Q}_p^{\times2} \text{ and } t_p=3,\\ 2^{-1}(1+p^{(1-n)/2}) & \text{if } \det(\varphi)\notin\mathbf{Z}_p^\times\mathbf{Q}_p^{\times2} \text{ and } t_p=1,\\ 2^{-1}(1-p^{(1-n)/2}) & \text{if } \det(\varphi)\notin\mathbf{Z}_p^\times\mathbf{Q}_p^{\times2} \text{ and } t_p=3,\end{cases}$$

$$\lambda_p^\psi=\begin{cases}2^{-1}(1+p)^{-1}(1-\chi_1(p)p^{(3-n)/2})(1-\chi_1(p)p^{(1-n)/2}) & \text{if } \chi_1(p)\neq0,\\ 2^{-1} & \text{if } \chi_1(p)=0.\end{cases}$$

(2) *The case of even n*:

$$\lambda_p^{\varphi} = \begin{cases} 2^{-1}(1+p)^{-1}(1-\chi(p)p^{(2-n)/2})(1-\chi(p)p^{-n/2}) & \text{if } \chi(p) \neq 0, \\ 2^{-1} & \text{if } \chi(p) = 0, \end{cases}$$

$$\lambda_p^{\psi} = \begin{cases} 1 & \text{if } \det(\psi) \in \mathbf{Z}_p^{\times}\mathbf{Q}_p^{\times 2},\ s_p = 1, \text{ and } p \mid 2, \\ 2^{-1}(1+p)^{-1}(1-p^{2-n}) & \text{if } \det(\psi) \in \mathbf{Z}_p^{\times}\mathbf{Q}_p^{\times 2} \text{ and } s_p = 3, \\ 2^{-1}(1+p^{(2-n)/2}) & \text{if } \det(\psi) \notin \mathbf{Z}_p^{\times}\mathbf{Q}_p^{\times 2} \text{ and } s_p = 1, \\ 2^{-1}(1-p^{(2-n)/2}) & \text{if } \det(\psi) \notin \mathbf{Z}_p^{\times}\mathbf{Q}_p^{\times 2} \text{ and } s_p = 3. \end{cases}$$

Here t_p (resp. s_p) is the core dimension of φ (resp. ψ) at p.

The ideal $2^{-1}b\mathbf{Z}$ in Theorem 4.1 satisfies

$$b^2\mathbf{Z} = 2q[\tilde{L}/L][\tilde{M}/M]^{-1}, \tag{4-1}$$

where $[\tilde{L}/L]$ (resp. $[\tilde{M}/M]$) is the *discriminant ideal* of φ (resp. ψ) defined in [S3, §6.1]. It can be seen from [M1, Theorem 4.2] that if $h \in L[q, 2^{-1}b\mathbf{Z}]$ with b above, then $L \cap W$ is $\mathbf{Z}$-maximal with respect to ψ, and vice versa. Hence the assumption on the maximality of $L \cap W$ in Theorem 4.1 can be replaced by the condition that $h \in L[q, 2^{-1}b\mathbf{Z}]$ with b of (4-1).

As for the computation of the discriminant ideal of φ, [S3, Theorem 6.2] describes $[\tilde{L}/L]$ and also $\ell(\varphi)$ in terms of the *invariants* of φ in the sense of [S3] (see Section 4.2 below). Similarly for the discriminant ideal of ψ, we need the invariants of ψ, which can be determined by [M1, Theorem 1.1]. We note simple facts: (i) $[\tilde{L}/L] = [\tilde{L} : L]\mathbf{Z}$. (ii) If φ is semi-integral and if $\det(2\varphi)\mathbf{Z}$ coincides with the discriminant ideal of φ, then $L(= \mathbf{Z}_n^1)$ is $\mathbf{Z}$-maximal with respect to φ.

As for the constants λ_p^{φ} and λ_p^{ψ}, we need the core dimensions of φ and ψ at each prime $p \mid \mathfrak{e}$ and $p \mid \mathfrak{e}_1$, respectively. These dimensions can be seen from the invariants of φ and ψ. The index $[C(M) : H_{\boldsymbol{A}} \cap C(L)]$ is computable by [MY, Theorem 3.8] together with the invariants of φ and ψ, since $M = L \cap W$ is maximal.

As can be seen from the formula preceding Corollary 4.3, the maximality of L is indispensable for our proof of Theorem 4.1. Indeed, (0-6) is valid only for maximal lattices, and thus for maximal L we can connect $\mathfrak{m}(L[q, 2^{-1}b\mathbf{Z}])$ with the mass $\mathfrak{m}(H, C(M))$ via the mass formula in (0-7).

4.2. ***Invariants and discriminant ideal of a quadratic form.*** The invariants of φ consist of four data; the *dimension* n of V, the *discriminant field* $K = \mathbf{Q}(\sqrt{\delta})$ with $\delta = (-1)^{n(n-1)/2}\det(\varphi)$, the *characteristic quaternion algebra*, denoted by $Q(\varphi)$, and the *index*, denoted by $s_\infty(\varphi)$. We denote a set of these invariants by

$$\{n, K, Q(\varphi), s_\infty(\varphi)\}.$$

By [S3, Theorem 4.2] the invariants of φ determine the isomorphism class of the quadratic space (V, φ). The algebra $Q(\varphi)$ is the quaternion algebra over $\mathbf{Q}$

determined by an isomorphism $A(\varphi) \cong Q(\varphi)_m^m$ or $A^+(\varphi) \cong Q(\varphi)_m^m$ according to whether n is even or odd with some $0 < m \in \mathbf{Z}$, where $A(\varphi)$ (resp. $A^+(\varphi)$) is the Clifford algebra (resp. the even Clifford algebra) of φ; we put $Q(\varphi) = \mathbf{Q}_2^2$ if $n = 1$. The index of φ is defined by $s_\infty(\varphi) = p - q$ if φ is equivalent to $\mathrm{diag}[1_p, -1_q]$ over $\mathbf{R}$.

Let t_p be the core dimension of φ at a prime p. From [S3, §3.2] we have

$$t_p = \begin{cases} 0 & \text{if } \chi(p) = 1 \text{ and } p \text{ is unramified in } Q(\varphi), \\ 1 & \text{if } p \text{ is unramified in } Q(\varphi), \\ 2 & \text{if } \chi(p) \neq 1, \\ 3 & \text{if } p \text{ is ramified in } Q(\varphi), \\ 4 & \text{if } \chi(p) = 1 \text{ and } p \text{ is ramified in } Q(\varphi), \end{cases} \tag{4-2}$$

where χ is the primitive Dirichlet character corresponding to K.

The discriminant ideal of φ is defined by $[\widetilde{L}/L]$ with a $\mathbf{Z}$-maximal lattice L in V with respect to φ. Here, in general, for two lattices M, N in V we denote by $[M/N]$ the fractional ideal generated over $\mathbf{Z}$ by $\det(\alpha)$ of all $\alpha \in \mathrm{GL}_n(\mathbf{Q})$ such that $M\alpha \subset N$. The ideal $[\widetilde{L}/L]$ is independent of the choice of L and in fact, it can be given by [S3, Theorem 6.2] as follows:

$$[\widetilde{L}/L] = \begin{cases} D_K (D_1)^2 \mathbf{Z} & \text{if } n \text{ is even}, \\ 2\delta_0^{-1} (D_{Q(\varphi)})^2 \mathbf{Z} \cap 2\delta_0 \mathbf{Z} & \text{if } n \text{ is odd}, \end{cases} \tag{4-3}$$

where D_K is the discriminant of K, δ_0 is the squarefree factor of δ, $D_{Q(\varphi)}$ is the product of all prime numbers ramified in $Q(\varphi)$, and D_1 is the product of all the prime numbers for which $p \mid D_{Q(\varphi)}$ and $p \nmid D_K$. The same theorem gives the level of φ_0, where φ_0 is the matrix that represents φ with respect to a basis of L:

$$\ell(\varphi_0)\mathbf{Z} = \begin{cases} D_K D_1 \mathbf{Z} & \text{if } n \text{ is even}, \\ 4\delta_0 \mathbf{Z} \cap D_{Q(\varphi)} \mathbf{Z} & \text{if } n \text{ is odd}. \end{cases} \tag{4-4}$$

4.3. ***Proof of the formula.*** We need a mass formula for indefinite quadratic forms. We start by recalling the definition of the mass in the sense of [S1, (5.5.1)].

Given a symmetric element φ of $\mathrm{GL}_n(\mathbf{Q})$ and a $\mathbf{Z}$-lattice L in $\mathbf{Q}_n^1$, we put $G = \mathrm{SO}(\varphi)$, $C(L) = \{\alpha \in G_{\boldsymbol{A}} \mid L\alpha = L\}$, and $\Gamma(L) = G \cap C(L)$, where $G_{\boldsymbol{A}}$ denotes the adelization of G. Let $G_{\boldsymbol{A}} = \bigsqcup_{\alpha \in \mathcal{A}} G\alpha C(L)$ with a finite subset $\mathcal{A}$ of $G_{\boldsymbol{A}}$. Then *the mass of G relative to $C(L)$* is defined by

$$\mathfrak{m}(G, C(L)) = \sum_{\alpha \in \mathcal{A}} \nu(G \cap \alpha C(L) \alpha^{-1}).$$

This is independent of the choice of $\mathcal{A}$. Also when φ is indefinite, $\mathfrak{m}(G, C(L))$ depends on the choice of (1-4). Similarly, with the notation of (0-7), the mass

$\mathfrak{m}(H,\ H_A \cap C(L))$ is defined by

$$\mathfrak{m}(H,\ H_A \cap C(L)) = \sum_{\beta \in \mathcal{B}} \nu(H \cap \beta C(L)\beta^{-1}), \tag{4-5}$$

where $\mathcal{B}$ is a complete set of representatives for $H\backslash H_A/(H_A \cap C(L))$.

Proposition 4.2. *Let φ be an indefinite symmetric matrix in $\mathrm{GL}_n(\mathbf{Q})$ with $n \geq 3$ and L a $\mathbf{Z}$-maximal lattice in $\mathbf{Q}_n^1$ with respect to φ. Then the genus of L consists of a single class with respect to $\mathrm{SO}(\varphi)$ and the mass of $\mathrm{SO}(\varphi)$ relative to $C(L)$ is given by*

$$\mathfrak{m}(\mathrm{SO}(\varphi), C(L)) = \nu(\Gamma(L)) = 2^{rn/2+2-n} \det(\theta)^{r/2} \rho_{r+t}^{-1} \rho_r^{-1} [\tilde{L} : L]^{(n-1)/2} \cdot \prod_{p|\mathfrak{e}} \lambda_p^{\varphi} \cdot \begin{cases} \prod_{k=1}^{(n-1)/2} \zeta(2k) & \text{if } n \text{ is odd,} \\ \prod_{k=1}^{(n-2)/2} \zeta(2k) L(n/2, \chi) & \text{if } n \text{ is even,} \end{cases}$$

where $\mathfrak{m}(\mathrm{SO}(\varphi), C(L))$ is defined with a fixed matrix as in (1-4), *and $\mathfrak{e}$, λ_p^{φ}, and χ are as in Theorem 4.1.*

Proof. The first assertion follows from [S2, Theorem 9.26] (cf. [S4, Remark 2.4(5)]) since L is $\mathbf{Z}$-maximal with respect to indefinite φ. Hence by definition we have $\mathfrak{m}(\mathrm{SO}(\varphi), C(L)) = \nu(\Gamma(L))$. With a $\mathbf{Z}$-basis of L, we may identify φ with φ_0 and L with $\mathbf{Z}_n^1$, where φ_0 is the matrix representing φ with respect to the basis of L. Then φ is semi-integral and by the first assertion the genus of φ consists of a single class with respect to SO. Thus by [M3, Corollary 3.4] we obtain

$$\nu(\Gamma(L)) = 2^{2+rn/2} \det(\theta)^{r/2} \rho_{r+t}^{-1} \rho_r^{-1} |\det(2\varphi)|^{(n+1)/2} \left(\prod_p 2^{-1} e_p(2\varphi) \right)^{-1}, \tag{4-6}$$

where the last product is taken over all prime numbers p and $e_p(2\varphi)$ is the representation density of 2φ at p. By [S1, Theorem 8.6(2)] each $e_p(2\varphi)$ can be given by $e_p(2\varphi) = p^{n\mathrm{ord}_p(2)-n(n-1)/2+\mathrm{ord}_p(\det(2\varphi))}[C_p^{\varphi} : (D_p)_1^{\varphi}]$, where ord_p is the order function of $\mathbf{Q}_p$, $C_p^{\varphi} = \{\gamma \in \mathrm{GL}_n(\mathbf{Q}_p) \mid \gamma\varphi \cdot {}^{\mathrm{T}}\gamma = \varphi, L_p\gamma = L_p\}$, and $(D_p)_1^{\varphi} = \{\gamma \in C_p^{\varphi} \mid \tilde{L}_p(\gamma - 1_n) \subset pL_p\}$. Since L_p is maximal for every prime p, we can compute such an index as follows.

If n is even, then by [S1, Proposition 3.9] we see that

$$[C_p^{\varphi} : (D_p)_1^{\varphi}] = 2p^{n(n-1)/2} \prod_{i=1}^{(n-2)/2} (1 - p^{-2i}) \cdot \begin{cases} 1 - \chi(p) p^{-n/2} & \text{if } \chi(p) \neq 0 \text{ and } p \nmid Q(\varphi), \\ 2(1+p)(1 - \chi(p) p^{1-n/2})^{-1} & \text{if } \chi(p) \neq 0 \text{ and } p \mid Q(\varphi), \\ 2 & \text{if } \chi(p) = 0. \end{cases}$$

If n is odd, by [S1, Proposition 3.9] (if $\det(\varphi) \in \mathbf{Z}_p^{\times}\mathbf{Q}_p^{\times 2}$) or [MY, Lemma 2.5]

(otherwise), we see that

$$[C_p^{\varphi} : (D_p)_1^{\varphi}] = 2p^{n(n-1)/2} \prod_{i=1}^{(n-1)/2} (1 - p^{-2i}) \cdot \begin{cases} 1 & \text{if } \det(\varphi) \in \mathbf{Z}_p^{\times}\mathbf{Q}_p^{\times 2} \text{ and } t_p = 1, \\ 2(1+p)(1-p^{1-n})^{-1} & \text{if } \det(\varphi) \in \mathbf{Z}_p^{\times}\mathbf{Q}_p^{\times 2} \text{ and } t_p = 3, \\ 2(1+p^{(1-n)/2})^{-1} & \text{if } \det(\varphi) \notin \mathbf{Z}_p^{\times}\mathbf{Q}_p^{\times 2} \text{ and } t_p = 1, \\ 2(1-p^{(1-n)/2})^{-1} & \text{if } \det(\varphi) \notin \mathbf{Z}_p^{\times}\mathbf{Q}_p^{\times 2} \text{ and } t_p = 3. \end{cases}$$

Now, we observe that $p \mid \mathfrak{e}$ if and only if p divides $[\widetilde{L}/L]$. In view of this, combining all results on $e_p(2\varphi)$ and $[C_p^{\varphi} : (D_p)_1^{\varphi}]$, we can compute the product $\prod_p 2^{-1} e_p(2\varphi)$. Since $[\widetilde{L}/L] = \det(2\varphi)\mathbf{Z}$ in our setting, the desired formula follows from (4-6). □

The notation being as in Theorem 4.1, let $M = \mathbf{Z}_{n-1}^1 k$ and set $\psi_1 = k\varphi \cdot {}^{\mathrm{t}}k$. We fix two matrices equivalent to φ and ψ_1 over $\mathbf{R}$ as follows:

$$\kappa\sigma\varphi \cdot {}^{\mathrm{t}}\sigma = \begin{bmatrix} 0 & 0 & -1_r \\ 0 & \theta & 0 \\ -1_r & 0 & 0 \end{bmatrix}, \quad \kappa_1\sigma_1\psi_1 \cdot {}^{\mathrm{t}}\sigma_1 = \begin{bmatrix} 0 & 0 & -1_{r_1} \\ 0 & \theta_1 & 0 \\ -1_{r_1} & 0 & 0 \end{bmatrix}, \tag{4-7}$$

with $0 < {}^{\mathrm{t}}\theta = \theta \in \mathrm{GL}_t(\mathbf{R})$ and $0 < {}^{\mathrm{t}}\theta_1 = \theta_1 \in \mathrm{GL}_{t_1}(\mathbf{R})$, where $\kappa, \kappa_1 \in \{\pm 1\}$, $\sigma \in \mathrm{GL}_n(\mathbf{R})$, and $\sigma_1 \in \mathrm{GL}_{n-1}(\mathbf{R})$. In particular, $r_1 = 0$ if ψ is definite. We then define $\nu(\Gamma(L))$, and also $\nu(\Gamma)$ for a subgroup Γ of H if ψ is indefinite, as in (1-5).

Since L is $\mathbf{Z}$-maximal with respect to indefinite φ, by Proposition 4.2 the genus of L consists of a single class with respect to $\mathrm{SO}(\varphi)$. By virtue of [S4, Theorem 3.2] (or by (0-7)) we obtain

$$\mathfrak{m}(L[q, 2^{-1}b\mathbf{Z}])\nu(\Gamma(L)) = \mathfrak{m}(H, H_A \cap C(L)) = [C(M) : H_A \cap C(L)]\mathfrak{m}(H, C(M)).$$

Corollary 4.3. *In the notation of Theorem 4.1, the mass of $L[q, 2^{-1}b\mathbf{Z}]$ is given by*

$$\mathfrak{m}(L[q, 2^{-1}b\mathbf{Z}]) = 2^{(2+r_1(n-1)-rn)/2}[C(M) : H_A \cap C(L)] \prod_{p|\mathfrak{e}} (\lambda_p^{\varphi})^{-1} \prod_{p|\mathfrak{e}_1} \lambda_p^{\psi} \cdot \frac{\det(\theta_1)^{r_1/2}\rho_{r+t}\rho_r[\widetilde{M} : M]^{(n-2)/2}}{\det(\theta)^{r/2}\rho_{r_1+t_1}\rho_{r_1}[\widetilde{L} : L]^{(n-1)/2}} \cdot \begin{cases} L((n-1)/2, \chi_1)\zeta(n-1)^{-1} & \textit{if } n \textit{ is odd}, \\ L(n/2, \chi)^{-1} & \textit{if } n \textit{ is even}. \end{cases}$$

Proof. Suppose that ψ is indefinite and $n > 3$. Since M is $\mathbf{Z}$-maximal with respect to ψ, we have $\mathfrak{m}(H, C(M)) = \nu(\Gamma(M))$ by Proposition 4.2. Therefore

$$\mathfrak{m}(L[q, 2^{-1}b\mathbf{Z}]) = \big[C(M) : H_A \cap C(L)\big]\nu(\Gamma(M))/\nu(\Gamma(L)).$$

Applying the formula in Proposition 4.2 to these $\nu(\Gamma(L))$ and $\nu(\Gamma(M))$, we obtain the desired result.

If $n = 3$, then ψ is anisotropic by (2-1). In this case [S1, Theorem 5.8] is applicable to $\mathfrak{m}(H, C(M))$ and Proposition 4.2 is applicable to $\nu(\Gamma(L))$. Combining these results, we have the formula in the statement. This settles the case of indefinite ψ.

Suppose ψ is definite. Then $\mathfrak{m}(H, C(M))$ is given by [MY, Corollary 2.7]. Notice that the expression

$$2^{n-1}\rho_{n-1}\prod_{k=1}^{[(n-2)/2]}(2k-1)!\,(2\pi)^{-2k}\cdot\begin{cases}((n-3)/2)!\,(2\pi)^{(1-n)/2} & \text{if } n \text{ is odd},\\ 2^{(2-n)/2} & \text{if } n \text{ is even}\end{cases}$$

equals 2 for $n > 2$. Combining these facts with Proposition 4.2, we have the desired formula with $r_1 = 0$ ($t_1 = n - 1 > 1$). □

Now we put $2^{-1}b\mathbf{Z} = c\mathbf{Z}$ and observe that

$$|c|^{2-n} = |2q^{-1}|^{(n-2)/2}[\tilde{L} : L]^{(2-n)/2}[\tilde{M} : M]^{(n-2)/2}.$$

Then our Theorem (page 72) shows that in the present case

$$\mathfrak{m}(L[q, 2^{-1}b\mathbf{Z}]) = 2^{(r_1(n-1)-rn)/2}\ell(\varphi)^{-1}[\tilde{L} : L]^{(3-n)/2}[\tilde{M} : M]^{(n-2)/2}\cdot\frac{\det(\theta_1)^{r_1/2}\rho_{r+t}\rho_r}{\det(\theta)^{r/2}\rho_{r_1+t_1}\rho_{r_1}}\,\nu(S, t). \tag{4-8}$$

Here $\mathfrak{m}(L[q, 2^{-1}b\mathbf{Z}])$ is the same quantity as in Corollary 4.3. This is because if ψ is indefinite, the spaces $\mathcal{Z}^{\varphi}$ and $\mathcal{Z}^{\psi}$ in Theorem are respectively defined by $\mathcal{Z}(r, \theta)$ and $\mathcal{Z}(r_1, \theta_1)$ and we may assume that these are defined with the setting of (4-7). That is also true for definite ψ in a similar reason. Thus (4-8) combined with Corollary 4.3 leads to the formula for $\nu(S, t)$. This proves Theorem 4.1.

5. Applications

With S and t as in (0-2), we denote by $\mu(S, t)$ the measure of representations of t by S mentioned in the Introduction. This representation measure is related to the primitive representation measures in the formula (see [Si3, Lemma 33])

$$\mu(S, t) = \sum_c c^{2-n}\,\nu(S, c^{-2}t) \tag{5-1}$$

for $0 \neq t \in \mathbf{Z}$, where c runs over all positive integers satisfying $c^{-2}t \in \mathbf{Z}$. By [Si3, Theorem 1], $\mu(S, t)$ can be expressed as an infinite product of the representation densities $\alpha_p(S, t)$ of t by S for all primes p. We here recall the two kinds of densities, $\alpha_p(S, t)$ and $\beta_p(S, t)$, for our application. By [Si1, Hilfssatz 13] their

values are

$$\begin{aligned} d &= 2\mathrm{ord}_p(2t)+1,\\ \alpha_p(S,t) &= p^{-3d}\cdot \#A_{p^d}(S,t), \quad \beta_p(S,t) = p^{-3d}\cdot \#B_{p^d}(S,t),\\ A_{p^d}(S,t) &= \left\{z \in (\mathbf{Z}_p/p^d\mathbf{Z}_p)_1^n \mid S[z]-t \in p^d\mathbf{Z}_p\right\},\\ B_{p^d}(S,t) &= \left\{z \in (\mathbf{Z}_p/p^d\mathbf{Z}_p)_1^n \mid S[z]-t \in p^d\mathbf{Z}_p, (\mathbf{Z}_p)_n^1 z = \mathbf{Z}_p\right\}. \end{aligned}$$

We note that if p does not divide the square factor of t, then $\beta_p(S,t) = \alpha_p(S,t)$.

5.1. ***The quadratic form η_r.*** We consider the quadratic form in $2r$ variables defined by $\varphi = 2^{-1}\eta_r$, where

$$\eta_r = \begin{bmatrix} 0 & 1_r \\ 1_r & 0 \end{bmatrix}$$

for $2 \le r \in \mathbf{Z}$. Clearly φ is indefinite and semi-integral. The invariants of φ are $\{n, \mathbf{Q}, M_2(\mathbf{Q}), 0\}$ with $n = 2r$, where $M_m(\mathbf{Q}) = \mathbf{Q}_m^m$ as usual. In fact, $\det(\varphi) = (-1)^r 2^{-n}$ and

$$\begin{bmatrix} 1_r & 1_r \\ 1_r & -1_r \end{bmatrix}\begin{bmatrix} 0 & 1_r \\ 1_r & 0 \end{bmatrix}^{\mathsf{T}}\begin{bmatrix} 1_r & 1_r \\ 1_r & -1_r \end{bmatrix} = 2\begin{bmatrix} 1_r & 0 \\ 0 & -1_r \end{bmatrix},$$

which reveals the discriminant field and the index of φ. As for the characteristic algebra, since $A(\varphi) \cong M_{2^r}(\mathbf{Q})$ by [S2, Theorem 2.6], we have $Q(\varphi) = M_2(\mathbf{Q})$.

We set $L = \mathbf{Z}_n^1$. This is integral with respect to φ as φ is semi-integral. We then observe that $[\tilde{L}/L] = \det(2\varphi)\mathbf{Z} = \mathbf{Z}$, which coincides with the discriminant ideal of φ by (4-3). Hence L is $\mathbf{Z}$-maximal with respect to φ.

Given $q \in \mathbf{Z}$, $\neq 0$, set $q = ab^2$ with squarefree $a \in \mathbf{Z}$ and $b \in \mathbf{Z}$. Let us take $h = abe_r + be_n$, where $\{e_i\}_{i=1}^n$ is the standard basis of $\mathbf{Q}_n^1$. We easily see that $h \in L[q, 2^{-1}b\mathbf{Z}]$. Set $W = (\mathbf{Q}h)^\perp$, $\psi = \varphi|_W$, and $H = \{\gamma \in \mathrm{SO}(\varphi) \mid h\gamma = h\}$. Then the invariants of ψ are $\{n-1, \mathbf{Q}(\sqrt{-q}), M_2(\mathbf{Q}), s_\infty\}$ with $s_\infty = -1$ or 1 according as $q > 0$ or < 0; in particular, ψ is indefinite.

To prove this, we first observe that

$$\begin{aligned} W &= \mathbf{Q}(abe_r - be_n) + \sum_{i=1}^{r-1}(\mathbf{Q}e_i + \mathbf{Q}e_{r+i}),\\ L \cap W &= \mathbf{Z}(abe_r - be_n) + \sum_{i=1}^{r-1}(\mathbf{Z}e_i + \mathbf{Z}e_{r+i}). \end{aligned}$$

In the basis $\{abe_r - be_n, e_i, e_{r+i} \mid i = 1, \dots, r-1\}$, ψ is represented by the matrix $\psi_0 = \mathrm{diag}[-q, 2^{-1}\eta_{r-1}]$. Hence we have a Witt decomposition

$$\sigma\psi_0 \cdot {}^{\mathsf{T}}\sigma = \begin{bmatrix} 0 & 0 & 2^{-1}1_{r-1} \\ 0 & -q & 0 \\ 2^{-1}1_{r-1} & 0 & 0 \end{bmatrix},$$

with $\sigma \in \mathrm{GL}_{n-1}(\mathbf{Q})$. Then $A^+(\psi) \cong M_{2^{r-1}}(\mathbf{Q})$ by [S2, Theorem 2.6], which shows $Q(\psi) = M_2(\mathbf{Q})$. The other invariants can be seen with ψ_0 above.

Proposition 5.1. *For every squarefree integer a and $2 \le r \in \mathbf{Z}$ we have*

$$\nu(\eta_r, 2a) = 2\zeta(r)^{-1} \prod_{p|a} (1 + p^{1-r}). \tag{5-2}$$

Proof. The discriminant ideal of ψ is given by $2a\mathbf{Z}$ by (4-3). Since ψ_0 is semi-integral and $|\det(2\psi_0)| = 2|a|$, $L \cap W$ is $\mathbf{Z}$-maximal with respect to ψ. Therefore Theorem 4.1 is applicable to φ and $h \in L[q, 2^{-1}b\mathbf{Z}]$. In view of the invariants of φ, we see that $\ell(\varphi) = 1$ by (4-4), $\chi = 1$, and $\mathfrak{e} = 1$ as $\tilde{L} = L$ with the notation of Theorem 4.1. Then $2^{-1}\hat{\varphi} = \eta_r$. Also, $p \mid \mathfrak{e}_1$ if and only if $p \mid 2a$. Hence we have

$$\nu(\eta_r, 2a) = 2[C(L \cap W) : H_{\boldsymbol{A}} \cap C(L)]\zeta(r)^{-1} \prod_{p|2a} \lambda_p^{\psi}.$$

Observe that the core dimension of ψ at p equals 1 for any prime p, because of $Q(\psi) = M_2(\mathbf{Q})$. Noticing that $\det(\psi) \notin \mathbf{Z}_p^{\times}\mathbf{Q}_p^{\times 2}$ if and only if $p \mid a$, by [MY, Theorem 3.8] we have $[C(L \cap W) : H_{\boldsymbol{A}} \cap C(L)] = 2^{\lambda}$, where λ is the number of prime factors of a. Further, λ_p^{ψ} equals 1 or $2^{-1}(1 + p^{1-r})$ according as $p \mid 2$ or $p \mid a$. Combining these, we obtain the desired formula. □

On the other hand, Siegel's product formula (0-3) specialized to $(S, t) = (\eta_r, 2a)$ shows that

$$\nu(\eta_r, 2a) = \prod_p \beta_p(\eta_r, 2a) = \beta_2(\eta_r, 2)(1 - 2^{-r})^{-1}\zeta(r)^{-1} \prod_{2 \ne p|a} (1 + p^{1-r}).$$

Indeed, it can be seen from [Si1, Hilfssatz 16] that $\beta_p(\eta_r, 2a) = 1 - p^{-r}$ or $(1 - p^{-r})(1 + p^{1-r})$ according as $p \nmid a$ or $p \mid a$ for every odd prime p. Comparing this with (5-2), we have

Corollary 5.2. *The representation density $\beta_2(\eta_r, 2a)$ is given by*

$$\beta_2(\eta_r, 2a) = 2(1 - 2^{-r}) \cdot \begin{cases} 1 & \text{if } a \text{ is odd,} \\ 1 + 2^{1-r} & \text{if } a \text{ is even.} \end{cases} \tag{5-3}$$

We mention that Siegel in [Si2, (90)] took up an integral matrix S equivalent to η_2 over $\mathbf{Z}$ and computed $\mu(S, 2\ell)$ for every positive integer ℓ, which hence coincides with (5-2) with $r = 2$ if ℓ is odd squarefree. We note that the present $\mu(S, 2\ell)$ is denoted by $\mu(S, 2\ell)/\mu(S)$ with the notation of [Si2]. We also mention that (5-3) in the case of even a can be verified by [IK, Lemma 5.4] together with the second assertion of [Y, Theorem 4.1].

Now, we shall discuss a calculation of the representation measure $\mu(\eta_r, 2q)$ by means of (5-1). To make our argument short, let $q = ab^2$ with a squarefree integer a

and a prime number b not dividing a. We consider the set

$$L[q] = \{x \in L \mid x\varphi \cdot {}^{\mathrm{t}}x = q\}.$$

By [M1, (4.1)] a fractional ideal $\mathfrak{b}(q)$ of $\mathbf{Q}$ can be defined by the equality $\mathfrak{b}(q)^2 = 2q[\widetilde{L} : L][\widetilde{M} : M]^{-1}$, where $[\widetilde{L} : L]$ is the discriminant ideal of φ and $[\widetilde{M} : M]$ is that of ψ. Since $[\widetilde{L} : L] = \mathbf{Z}$ and $[\widetilde{M} : M] = 2a\mathbf{Z}$, we have $\mathfrak{b}(q) = b\mathbf{Z}$. Then from [M1, (4.4)] we see that

$$L[q] = L[q, 2^{-1}b\mathbf{Z}] \cup L[q, 2^{-1}\mathbf{Z}].$$

Observe that $\mathfrak{m}(L[q, 2^{-1}b\mathbf{Z}])$ is related to $\nu(\eta_r, 2a)$ and $\mathfrak{m}(L[q, 2^{-1}\mathbf{Z}])$ to $\nu(\eta_r, 2q)$ by the Theorem. Then by (5-1), $\mu(\eta_r, 2q)$ is given by

$$\mu(\eta_r, 2q) = b^{2-2r}\nu(\eta_r, 2a) + \nu(\eta_r, 2q). \tag{5-4}$$

Here $\nu(\eta_r, q/2) = \nu(\eta_r, a/2) = 0$ whenever a is even, because of $B(\eta_r, t) = \varnothing$ for any odd integer t. Thus formula (5-2) computes $\nu(\eta_r, 2a)$ in (5-4).

Another measure $\nu(\eta_r, 2q)$, however, cannot be computed by Theorem 4.1. This is because $L \cap W_1$ is not $\mathbf{Z}$-maximal with respect to ψ_1, where $W_1 = (\mathbf{Q}k)^{\perp}$, $\psi_1 = \varphi|_{W_1}$, and $H_1 = \{\gamma \in \mathrm{SO}(\varphi) \mid k\gamma = k\}$ with $k \in L[q, 2^{-1}\mathbf{Z}]$; see (4-1). Since still the present L is maximal, Theorem combined with (0-7) connects $\nu(\eta_r, 2q)$ with $\mathfrak{m}(H_1, C(L \cap W_1))$. In view of this, it is useful in application to remove the condition of the maximality of $L \cap W$ from Corollary 4.3.

Let us show an easy example of $\nu(\eta_r, 2q)$ by employing Siegel's formula (0-3). Assume $r = 2$ and put $\eta = \eta_2$ for simplicity. Let $q = 2 \cdot 3^2$ with $a = 2, b = 3$, for example. Then $d = 7, 5$, or 1 according as $p = 2, 3$, or $p \nmid 6$. If $p \nmid 6$, by [Si1, Hilfssatz 16] we have $\beta_p(\eta, 36) = 1 - p^{-2}$. If $p \mid 6$, direct calculations (by computer) show that $\#B_{2^7}(\eta, 36) = 4718592$ and $\#B_{3^5}(\eta, 36) = 17006112$; hence we have $\beta_2(\eta, 36) = \frac{9}{4}$ and $\beta_3(\eta, 36) = \frac{32}{27}$. Thus by (0-3),

$$\nu(\eta, 36) = \beta_2(\eta, 36)\beta_3(\eta, 36)\prod_{p \nmid 6}(1 - p^{-2}) = 4\zeta(2)^{-1}.$$

Therefore by (5-4) together with (5-2) we obtain

$$\mu(\eta, 36) = 3^{-2}\nu(\eta, 4) + \nu(\eta, 36) = 13/3\zeta(2).$$

We can check this result by using [Si3, Theorem 1], or simply, by [Si2, (90)] mentioned above.

5.2. ***The quadratic form S_n.*** In this subsection we assume $n > 2$ and set

$$S_n = \begin{bmatrix} 1_{n-1} & 0 \\ 0 & -1 \end{bmatrix}, \quad L = \mathbf{Z}_n^1.$$

Clearly S_n is indefinite and semi-integral, $S_n^{-1} = S_n$ and $\det(S_n) = -1$. By definition the level of S_n is 4 for every n. We shall first observe the relation between $\mathfrak{m}(L[\varepsilon, \mathbf{Z}])$ and $\nu(2S_n, 2\varepsilon)$ by applying Theorem to S_n and L, where $\varepsilon = 1$ or -1.

Lemma 5.3. *The invariants of S_n are*

$$\begin{aligned}
\{n, \mathbf{Q}(\sqrt{-1}), B_{2,\infty}, n-2\} &\quad \text{if } n \equiv 0, 5 \pmod 8,\\
\{n, \mathbf{Q}(\sqrt{-1}), M_2(\mathbf{Q}), n-2\} &\quad \text{if } n \equiv 1, 4 \pmod 8,\\
\{n, \mathbf{Q}, M_2(\mathbf{Q}), n-2\} &\quad \text{if } n \equiv 2, 3 \pmod 8,\\
\{n, \mathbf{Q}, B_{2,\infty}, n-2\} &\quad \text{if } n \equiv 6, 7 \pmod 8,
\end{aligned}$$

where $B_{2,\infty}$ is a definite quaternion algebra over $\mathbf{Q}$ ramified exactly at 2.

Proof. We first recall the characteristic algebra $Q(1_m)$ for $m > 0$:

$$Q(1_m) = \begin{cases} M_2(\mathbf{Q}) & \text{if } m \equiv 0, \pm 1, 2 \pmod 8,\\ B_{2,\infty} & \text{if } m \equiv 4, \pm 3, -2 \pmod 8. \end{cases} \tag{5-5}$$

Suppose $n \equiv 0 \pmod 8$. Put $n = 8k$. Then $(-1)^{n/2}\det(S_n) = (-1)^{4k+1}$, and hence by definition the discriminant field is $\mathbf{Q}(\sqrt{-1})$. Clearly the index of S_n equals $n-2$. By applying [S3, Lemma 2.8(i)] to 1_{n-2} and diag$[1, -1]$, we see that

$$A(S_n) \cong A(1_{n-2}) \otimes_{\mathbf{Q}} A((-1)^{(n-2)/2}\mathrm{diag}[1, -1]) \cong M_s(B_{2,\infty}) \otimes_{\mathbf{Q}} M_2(\mathbf{Q})$$

with $0 < s \in \mathbf{Z}$. From this it follows that $Q(S_n) = B_{2,\infty}$. Hence we have $\{n, \mathbf{Q}(\sqrt{-1}), B_{2,\infty}, n-2\}$ if $n \equiv 0 \pmod 8$. The other cases can be handled in a similar way. □

Let $\{e_i\}_{i=1}^n$ be the standard basis of $\mathbf{Q}_n^1$. For $h \in L[\varepsilon, \mathbf{Z}]$ put $W = (\mathbf{Q}h)^\perp$ with respect to S_n and $H = \{\gamma \in \mathrm{SO}(S_n) \mid h\gamma = h\}$. Let ψ be the restriction of the quadratic form represented by S_n to W.

Lemma 5.4. *The quantity $\nu(2S_n, 2\varepsilon)/\mathfrak{m}(L[\varepsilon, \mathbf{Z}])$ can be given by*

$$\frac{\nu(2S_n, 2\varepsilon)}{\mathfrak{m}(L[\varepsilon, \mathbf{Z}])} = \begin{cases} 4\pi^{(1-n)/2}\Gamma((n-1)/2) & \text{if } n > 3 \text{ and } \varepsilon = 1,\\ 2^n & \text{if } n > 2 \text{ and } \varepsilon = -1. \end{cases}$$

Here $\mathfrak{m}(L[\varepsilon, \mathbf{Z}])$ is defined as in (0-5) *with $\mathcal{Z}^{S_n} = \mathcal{Z}(1, 2\cdot 1_{n-2})$, $\mathcal{Z}^{\psi} = \mathcal{Z}(1, 2\cdot 1_{n-3})$ if $\varepsilon = 1$, and $\mathcal{Z}^{\psi} = \mathcal{Z}(0, 1_{n-1})$ if $\varepsilon = -1$.*

Proof. With $\mathcal{Z}^{S_n} = \mathcal{Z}(r, \theta)$ and $\mathcal{Z}^{\psi} = \mathcal{Z}(r_1, \theta_1)$, our theorem shows that

$$\frac{\nu(2S_n, 2\varepsilon)}{\mathfrak{m}(L[\varepsilon, \mathbf{Z}])} = 2^{(2+rn-r_1(n-1))/2} \cdot \frac{\det(\theta)^{r/2}\rho_{n-1-r_1}\rho_{r_1}}{\det(\theta_1)^{r_1/2}\rho_{n-r}\rho_r}. \tag{5-6}$$

Since $S_n = \mathrm{diag}[1, 1_{n-2}, -1]$ is equivalent to

$$\begin{bmatrix} 0 & 0 & -1/2 \\ 0 & 1_{n-2} & 0 \\ -1/2 & 0 & 0 \end{bmatrix}$$

over $\mathbf{Q}$, we may set $\mathcal{Z}^{S_n} = \mathcal{Z}(1, 2\cdot 1_{n-2})$. We pick e_1 and fix it. Clearly $e_1 \in L[1, \mathbf{Z}]$. Take W and ψ as above with e_1 in place of h. Then we have $W = \sum_{i=2}^{n} \mathbf{Q}e_i$. With this basis ψ is represented by $\psi_0 = \mathrm{diag}[1_{n-2}, -1]$, which is isotropic. Hence we have to assume that $n > 3$ in the application of Theorem. Since ψ_0 is equivalent to

$$\begin{bmatrix} 0 & 0 & -1 \\ 0 & 2\cdot 1_{n-3} & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

over $\mathbf{R}$, we set $\mathcal{Z}^{\psi} = \mathcal{Z}(1, 2\cdot 1_{n-3})$. Clearly $\rho_{n-2}/\rho_{n-1} = \pi^{(1-n)/2}\Gamma((n-1)/2)$. Combining these facts with (5-6), we have the relation in the case of $\varepsilon = 1$. For the case of $\varepsilon = -1$, we pick $e_n \in L[-1, \mathbf{Z}]$. Then ψ is positive definite. The desired relation follows from (5-6) with $r_1 = 0$. □

Now we cannot apply Theorem 4.1 — more precisely, (0-7) — to $\mathfrak{m}(L[\varepsilon, \mathbf{Z}])$ because the present L is *not* $\mathbf{Z}$-maximal with respect to S_n for any $n > 2$ as will be shown below. On the other hand, (0-3) is applicable to $\nu(2S_n, 2\varepsilon)$ if $n \geq 4$ and so we can calculate $\mathfrak{m}(L[\varepsilon, \mathbf{Z}])$ by Lemma 5.4. To see this, let us take $n = 4$ and $\varepsilon = 1$, for example. By [Si1, Hilfssatz 16] we have $\alpha_p(2S_4, 2) = 1 - \xi(p)p^{-2}$ for every $p \nmid 2$, where ξ is the primitive Dirichlet character corresponding to $\mathbf{Q}(\sqrt{-1})$. As for $p = 2$, a direct calculation shows that $\#B_{2^5}(2S_4, 2) = 98304$. Hence we have $\beta_2(2S_4, 2) = 3$ and thus $\nu(2S_4, 2) = 3L(2, \xi)^{-1}$ by (0-3). This equals $\mu(2S_4, 2)$ since $t = 2$ is squarefree. Therefore, by Lemma 5.4,

$$\mathfrak{m}(L[1, \mathbf{Z}]) = \frac{\pi^{3/2}}{4\Gamma\left(\frac{3}{2}\right)}\nu(2S_4, 2) = \frac{3\pi}{2}L(2, \xi)^{-1}.$$

As mentioned in the Introduction, our theorem treats the quadratic form represented by S (in the statement), whose coefficients are all even integers. For this reason, $\nu(S_n, k)$ cannot be derived from the theorem. There is an explicit formula for the representation measure $\mu(S_n, k)$ valid for $0 \neq k \in \mathbf{Z}$ if $n > 3$. According to that formula, $\mu(S_4, 1)$ equals $\frac{3}{2}L(2, \xi)^{-1}$; see [RT, Theorem 12]. For the equality $\alpha_2(S_4, 1) = \frac{3}{2}$ the reader might refer to [Y, Theorem 4.1].

Returning to our example, take a $\mathbf{Z}$-maximal lattice L_0 in $\mathbf{Q}_n^1$ with respect to S_n. Let S_0 be the matrix that represents S_n with respect to a $\mathbf{Z}$-basis of L_0. Then by (4-4) and (4-3) together with Lemma 5.3 we can determine the level of S_0 and $[\widetilde{L}_0 : L_0]$:

Lemma 5.5. *With the notation above, we have*

$$\ell(S_0) = \begin{cases} 1 & \text{if } n \equiv 2 \pmod 8, \\ 2 & \text{if } n \equiv 6 \pmod 8, \\ 4 & \text{otherwise}; \end{cases} \qquad [\widetilde{L}_0 : L_0] = \begin{cases} 1 & \text{if } n \equiv 2 \pmod 8, \\ 2 & \text{if } n \equiv 1, 3 \pmod 8, \\ 4 & \text{if } n \equiv 0, 4, 6 \pmod 8, \\ 8 & \text{if } n \equiv 5, 7 \pmod 8. \end{cases}$$

On the other hand, we have

$$[\widetilde{L} : L] = |\det(2S_n)| = 2^n.$$

Thus L is not $\mathbf{Z}$-maximal with respect to S_n for any $n > 2$ as mentioned.

If $L_0[q, 2^{-1}b\mathbf{Z}] \neq \varnothing$ for a given $q \in \mathbf{Q}^\times$ with b satisfying (4-1), one can compute $\nu(S, t)$ by means of Theorem 4.1, where $S = \ell(S_0)(2S_0)^{-1}$ and $t = 2b^{-2}\ell(S_0)q$. We illustrate this in the case of $n = 5$:

Example. The lattice

$$L_0 = \sum_{i=1}^{3} \mathbf{Z}e_i + \mathbf{Z}f_4 + \mathbf{Z}e_5, \quad f_4 = 2^{-1}(e_1 + e_2 + e_3 + e_4). \tag{5-7}$$

is $\mathbf{Z}$-maximal with respect to S_5, because representing S_5 (in the basis of L_0) by

$$S_0 = \begin{bmatrix} 1 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}, \quad \text{whose inverse is} \quad \begin{bmatrix} 2 & 1 & 1 & -2 & 0 \\ 1 & 2 & 1 & -2 & 0 \\ 1 & 1 & 2 & -2 & 0 \\ -2 & -2 & -2 & 4 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix},$$

we easily see that S_0 is semi-integral and $[\widetilde{L}_0 : L_0] = |\det(2S_0)| = 8$. Together with Lemma 5.5, this proves the desired fact.

We consider the set $L[q, \mathbf{Z}]$ with S_0. Notice that $L = \mathbf{Z}_5^1$ is $\mathbf{Z}$-maximal with respect to S_0.

Suppose $h \in L[q, \mathbf{Z}]$ with a squarefree odd positive integer q. With this h we set W, ψ, and H as before. The invariants of ψ are $\{4, \mathbf{Q}(\sqrt{-q}), Q(\psi), 2\}$, where $Q(\psi)$ is the indefinite quaternion algebra over $\mathbf{Q}$ of discriminant P or $2P$ according as $q \equiv 1 \pmod 4$ or $q \equiv 3 \pmod 4$, and P is the product of all the prime factors of q such that $p \equiv 3 \pmod 4$. The discriminant ideal of ψ is $4q\mathbf{Z}$. Together with $[\widetilde{L}_0 : L_0] = 8$, this shows that (4-1) is satisfied in the present case. Hence Theorem 4.1 is applicable to $h \in L[q, \mathbf{Z}]$. Clearly $\ell(S_0) = 4$ by Lemma 5.5.

Proposition 5.6. *Assume that $L[q, \mathbf{Z}] \neq \varnothing$ with a squarefree odd positive integer q in the above setting. Then $\nu(2S_0^{-1}, 2q)$ is given by*

$$\nu(2S_0^{-1}, 2q) = \frac{2^4 L(2, \chi)}{5\zeta(4)} \cdot \begin{cases} 1 & \text{if } q \equiv 1 \pmod 4, \\ 3^{-1}(1-\chi(2)2^{-1})(1-\chi(2)2^{-2}) & \text{if } q \equiv 3 \pmod 4, \end{cases}$$

where χ is the primitive Dirichlet character corresponding to $\mathbf{Q}(\sqrt{-q})$.

Proof. Put $K = \mathbf{Q}(\sqrt{-q})$ and $M = L \cap W$. We see that $\widetilde{M}_p \neq M_p$ if and only if $p \mid 2q$. Then we have

$$\nu(2S_0^{-1}, 2q) = [C(M) : H_A \cap C(L)]L(2, \chi)\zeta(4)^{-1} \prod_{p \mid 2} (\lambda_p^{\varphi})^{-1} \prod_{p \mid 2q} \lambda_p^{\psi}.$$

Since $Q(\varphi) = B_{2,\infty}$, the core dimension of φ at $p = 2$ equals 3 by (4-2). Hence $\lambda_2^{\varphi} = 2^{-1} \cdot 3^{-1}(1 - 2^{-4})$. Also since $D_K = -4q$ or $-q$ according to whether $q \equiv 1 \pmod 4$ or $q \equiv 3 \pmod 4$, we have $\lambda_p^{\psi} = 2^{-1}$ for any $p \mid q$ in both cases. As for $p = 2$, if $q \equiv 1 \pmod 4$, then $\lambda_2^{\psi} = 2^{-1}$; if $q \equiv 3 \pmod 4$, then $\lambda_2^{\psi} = 2^{-1} \cdot 3^{-1}(1 - \chi(2)2^{-1})(1 - \chi(2)2^{-2})$. It can be seen from [MY, Theorem 3.8] that $[C(M) : H_A \cap C(L)] = 2^{\lambda}$ with the number λ of all prime factors of q. Combining all this, we have the desired formula. □

Corollary 5.7. *The representation density $\beta_2(2S_0^{-1}, 2q)$ is given by*

$$\beta_2(2S_0^{-1}, 2q) = \begin{cases} 3 & \text{if } q \equiv 1 \pmod 4, \\ 1 - \chi(2)2^{-1} & \text{if } q \equiv 3 \pmod 4. \end{cases}$$

Proof. For each $p \nmid 2$, we have $\alpha_p(2S_0^{-1}, 2q) = (1 - p^{-4})(1 - \chi(p)p^{-2})^{-1}$ by [Si1, Hilfssatz 16]. Thus (0-3) in the present case shows that

$$\nu(2S_0^{-1}, 2q) = \frac{\beta_2(2S_0^{-1}, 2q)}{1 - 2^{-4}} \frac{L(2, \chi)}{\zeta(4)} \cdot \begin{cases} 1 & \text{if } q \equiv 1 \pmod 4, \\ 1 - \chi(2)2^{-2} & \text{if } q \equiv 3 \pmod 4. \end{cases}$$

Comparing this with Proposition 5.6, we obtain $\beta_2(2S_0^{-1}, 2q)$. □

We end this subsection by showing, under some conditions, the existence of the element $h \in L[q, \mathbf{Z}]$ needed in Proposition 5.6. Clearly, $L[q, \mathbf{Z}] \neq \varnothing$ if and only if $L_0[q, \mathbf{Z}] \neq \varnothing$ with respect to S_5, where L_0 is the maximal lattice given in (5-7).

Lemma 5.8. *For a positive integer q the following assertions hold.*

(1) *If q is even and squarefree, then $L_0[q, \mathbf{Z}] \neq \varnothing$.*

(2) *If q is odd and squarefree, then $L_0[q, 2^{-1}\mathbf{Z}] \neq \varnothing$.*

(3) *If q is odd and $q + 1$ is squarefree, then $L_0[q, \mathbf{Z}] \neq \varnothing$.*

Proof. We shall employ the results from the 4-dimensional quadratic space $(\mathbf{Q}_4^1, 1_4)$. Clearly $\mathbf{Q}_4^1$ can be identified with the subspace $\sum_{i=1}^4 \mathbf{Q}e_i$ of $\mathbf{Q}_5^1$ under the map $x \mapsto [x\ \ 0]$; notice that $[x\ \ 0]S_5 \cdot {}^{\mathrm{T}}[x\ 0] = x \cdot {}^{\mathrm{T}}x$ for $x \in \mathbf{Q}_4^1$. Then by (4-3) and (5-5) we see that $M = \sum_{i=1}^3 \mathbf{Z}e_i + \mathbf{Z}f_4$ is a $\mathbf{Z}$-maximal lattice in $(\mathbf{Q}_4^1, 1_4)$. By a result of Yoshinaga [Yo, Lemma 7.4(1)],

$$\begin{aligned} M[q, \mathbf{Z}] &\neq \varnothing \quad \text{if } q \text{ is even and squarefree,} \\ M[q, 2^{-1}\mathbf{Z}] &\neq \varnothing \quad \text{if } q \text{ is odd and squarefree.} \end{aligned} \tag{5-8}$$

Assertions (1) and (2) follow from these results since $L_0 = M + \mathbf{Z}e_5$ and $(\mathbf{Q}e_5)^{\perp} = \mathbf{Q}_4^1$ with respect to $S_5 = \mathrm{diag}[1_4, -1]$. To prove (3), set $q + 1 = 2a$ with a squarefree odd integer $a > 0$. Then by (5-8) we can find $g \in M[2a, \mathbf{Z}]$. Putting $h = g + e_5$, we easily see that $h \in L_0[q, \mathbf{Z}]$. □

5.3. ***Invariance of the mass*** $\mathfrak{m}(L[q, \mathfrak{b}])$ ***for indefinite*** φ**.** Let (V, φ) be a quadratic space over $\mathbf{Q}$, $\mathrm{SO}(\varphi)_A$ the adelization of $\mathrm{SO}(\varphi)$, and L a $\mathbf{Z}$-lattice in V. We recall that the genus of L is the set of all $\mathbf{Z}$-lattices in V of the form $L\alpha$ with $\alpha \in \mathrm{SO}(\varphi)_A$. There is a correspondence between the genus of a lattice and the genus of a matrix; see [S2, §9.27] for details.

Corollary 5.9. *Let the notation and assumptions be as in Theorem with* $n > 3$*. Let* L_1 *be a* $\mathbf{Z}$*-lattice in the genus of* L *for which* $L_1[q, c\mathbf{Z}] \neq \varnothing$*. Pick* $h_1 \in L_1[q, c\mathbf{Z}]$ *and let* ψ_1 *be the restriction of* φ *to the complement* $(\mathbf{Q}h_1)^{\perp}$ *in* (V, φ)*. Define the mass* $\mathfrak{m}(L_1[q, c\mathbf{Z}])$ *as in* (0-5) *with* $\mathcal{Z}^{\varphi} = \mathcal{Z}(r, \theta)$ *and* $\mathcal{Z}^{\psi_1} = \mathcal{Z}(r_1, \theta_1)$*. Then*

$$\mathfrak{m}(L_1[q, c\mathbf{Z}]) = \mathfrak{m}(L[q, c\mathbf{Z}]).$$

Proof. Since $\varphi[h_1] = \varphi[h]$, we have $h_1 = h\gamma$ with $\gamma \in \mathrm{SO}(\varphi)$. Hence we can take a matrix representing ψ_1 by the matrix representing ψ with respect to a basis of W. We then define $\mathfrak{m}(L_1[q, c\mathbf{Z}])$ as in the statement.

Let φ_1 be the matrix that represents φ with respect to a $\mathbf{Z}$-basis of L_1. Then φ_1 belongs to the genus of φ; namely, $\varphi_1 = \alpha_v \varphi \cdot {}^{\mathrm{T}}\alpha_v$ with $\alpha_v \in \mathrm{GL}_n(\mathbf{Z}_v)$ for every prime spot v of $\mathbf{Q}$, where $\mathbf{Z}_v = \mathbf{R}$ if v is the archimedean prime. Since $\det(\varphi)$ and $\ell(\varphi)$ are genus invariants, these coincide with $\det(\varphi_1)$ and $\ell(\varphi_1)$, respectively. Put $\ell = 2^{-1}\ell(\varphi)$. It can be seen from the definition that $\ell\hat{\varphi}_1$ belongs to the genus of $S(= \ell\hat{\varphi})$. By virtue of [Si3, Theorem 2] the measure $\nu(S, t)$ is a genus invariant, and thus $\nu(\ell\hat{\varphi}_1, t) = \nu(S, t)$ with $t = c^{-2}\ell q$. This is valid even when $n = 4$; see a remark in [Si3, Section 12]. Then we can find the desired equality by applying our theorem to $\mathfrak{m}(L_1[q, c\mathbf{Z}])$. □

Acknowledgments

The author expresses his gratitude to the reviewer for helpful comments on the earlier version of this paper. He also thanks the referee very much for carefully reading the manuscript and for many remarks on the revision. The author would like to thank Professor Don Blasius for giving him the chance to improve the Introduction with encouraging suggestions.

Corrections to [M3]

- Page 91, line 3 from the bottom: "a $\mathbf{Z}$-lattice L" should be "the lattice $L = \mathbf{Z}_n^1$".
- Page 99, Lemma 2.1: "C_v" should be "$C_v \cap (G_v)_0$". We also mention that $(G_v)_0/(C_v \cap (G_v)_0)$ is analytically diffeomorphic to G_v/C_v with the notation there.

References

[IK] T. Ikeda and H. Katsurada, "An explicit formula for the Siegel series of a quadratic form over a non-archimedean local field", *J. Reine Angew. Math.* **783** (2022), 1–47. MR

[KS] H. Kojima and H. Sakata, "On Siegel's product formulas for quadratic forms over algebraic number fields in higher dimensional cases", *J. Anal.* **33**:2 (2025), 619–646. MR

[M1] M. Murata, "Invariants and discriminant ideals of orthogonal complements in a quadratic space", preprint, 2011. arXiv 1112.2454

[M2] M. Murata, "On quadratic Diophantine equations in four variables and orders associated with lattices", *Doc. Math.* **19** (2014), 247–284. MR

[M3] M. Murata, "The orbits of primitive representations of a quadratic form by a quadratic form", *Osaka J. Math.* **63**:1 (2026), 91–110. MR

[MY] M. Murata and T. Yoshinaga, "On the solutions of quadratic Diophantine equations, II", *J. Math. Soc. Japan* **70**:3 (2018), 895–919. MR

[RT] J. G. Ratcliffe and S. T. Tschantz, "On the representation of integers by the Lorentzian quadratic form", *J. Funct. Anal.* **150**:2 (1997), 498–525. MR

[S1] G. Shimura, "An exact mass formula for orthogonal groups", *Duke Math. J.* **97**:1 (1999), 1–66. MR

[S2] G. Shimura, *Arithmetic and analytic theories of quadratic forms and Clifford groups*, Mathematical Surveys and Monographs **109**, Amer. Math. Soc., 2004. MR

[S3] G. Shimura, "Classification, construction, and similitudes of quadratic forms", *Amer. J. Math.* **128**:6 (2006), 1521–1552. MR

[S4] G. Shimura, "Integer-valued quadratic forms and quadratic Diophantine equations", *Doc. Math.* **11** (2006), 333–367. MR

[S5] G. Shimura, "Quadratic Diophantine equations, the class number, and the mass formula", *Bull. Amer. Math. Soc.* (*N.S.*) **43**:3 (2006), 285–304. MR

[Si1] C. L. Siegel, "Über die analytische Theorie der quadratischen Formen", *Ann. of Math.* (2) **36**:3 (1935), 527–606. MR

[Si2] C. L. Siegel, "Über die analytische Theorie der quadratischen Formen. II", *Ann. of Math.* (2) **37**:1 (1936), 230–263. MR

[Si3] C. L. Siegel, "On the theory of indefinite quadratic forms", *Ann. of Math.* (2) **45** (1944), 577–622. MR

[Y] T. Yang, "An explicit formula for local densities of quadratic forms", *J. Number Theory* **72**:2 (1998), 309–356. MR

[Yo] T. Yoshinaga, "On the solutions of quadratic Diophantine equations", *Doc. Math.* **15** (2010), 347–385. MR

Received February 8, 2025. Revised April 14, 2026.

MANABU MURATA
DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
RITSUMEIKAN UNIVERSITY
KUSATSU
JAPAN
murata31@pl.ritsumei.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 344, No. 1, 2026
https://doi.org/10.2140/pjm.2026.344.103

GROUP TOPOLOGIES ON GROUPS OF ABSOLUTELY BICONTINUOUS HOMEOMORPHISMS

J. DE LA NUEZ GONZÁLEZ

The group of homeomorphisms of the closed interval that are absolutely continuous and have an absolutely continuous inverse was shown by Solecki to admit a natural Polish group topology τ_{ac}. We show that under mild conditions on a compact space endowed with a finite Borel measure such a topology can be defined on the subgroup of the homeomorphism group consisting of those elements g such that g and g^{-1} preserve the class of null sets.

We use a probabilistic argument to show that in the case of a compact topological manifold equipped with an Oxtoby–Ulam measure, as well as in that of the Cantor space endowed with some natural Borel measures, there is no group topology between τ_{ac} and the restriction τ_{co} of the compact-open topology. In fact, we show that any separable group topology strictly finer than τ_{co} must be also finer than τ_{ac}. For one-dimensional manifolds we also show that τ_{co} and τ_{ac} are the only Hausdorff group topologies coarser than τ_{ac}, and one can read our result as evidence for the nonexistence of a good notion of regularity between continuity and absolute continuity.

We also show that while Solecki's example is not Roelcke precompact, the group of absolutely bicontinuous homeomorphisms of the Cantor space endowed with the measure given by the Fraïssé limit of the class of measured boolean algebras with rational probability measures is Roelcke precompact. This can be interpreted as showing the nonexistence of a good notion of regularity between continuity and absolute continuity.

Introduction

Recall the notion of absolute continuity from classical real analysis.

Definition 0.1. We say that $\phi : [0,1] \to \mathbb{R}$ is *absolutely continuous* if there exists some $\psi \in L^1([0,1])$ such that for all $t \in [0,1]$ we have $\phi(t) - \phi(0) = \int_{[0,t]} \psi \, d\lambda$, where λ is the Lebesgue measure on the interval.

Work supported by Samsung Science and Technology Foundation under Project Number SSTF-BA1301-51, by KIAS individual grant MG084001 and by the National Research Foundation of Korea under Mid-Career Research Program RS-2023-00278510.

MSC2020: primary 57S05, 57Sxx; secondary 28-XX.

Keywords: group topology, absolutely continuous, homeomorphism group.

© 2026 The Author, under license to MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

If f is absolutely continuous, then the derivative of f exists and coincides with ψ almost everywhere. In [21, Lemma 2.4] Solecki considers the group

$$\mathcal{H}^+_{bac}([0,1]) := \big\{g \in \mathcal{H}^+([0,1]) \mid g \text{ and } g^{-1} \text{ are absolutely continuous}\big\},$$

where $\mathcal{H}([0,1])$ denotes the homeomorphism group of a topological space X and $\mathcal{H}^+([0,1])$ the subgroup of those which are orientation preserving in case $[0,1]$ is a manifold, and he shows that it is a Polish group with the topology τ_{ac} induced by the metric $\rho(f,g) := \|f' - g'\|_1$. One can define the group $\mathcal{H}_{bac}(S^1)$ analogously (see [4; 9]) and prove Polishness in essentially the same way, although in this case an additional term controlling the displacement of a reference point is needed in the definition of ρ.

Even though absolute continuity is particularly interesting in the 1-manifold setting, in so far as it can be regarded as the first of an infinite family of intermediate regularity conditions defining some Polish subgroup of the $\mathcal{H}([0,1])$ (see [4]), we believe worth pointing out that Solecki's Polish group generalizes to a wide variety of other contexts, some of which will be explored here.

Recall that given two nonnegative measures μ, ν on a measurable space $(X, \mathcal{A})$ we say that ν is absolutely continuous with respect to μ, written $\nu \ll \mu$, if and only if $\nu(A) = 0$ for any $A \in \mathcal{A}$ with $\mu(A) = 0$. We are interested in the case where μ is a Borel measure, i.e., a measure on the Borel σ-algebra $\mathcal{B}(X)$ of X. The group $\mathcal{H}(X)$ acts on the space of Borel probability measures in a natural way by letting $g \cdot \mu$ be the push-forward measure: $(g \cdot \mu)(A) = \mu(g^{-1}(A))$ for $A \in \mathcal{B}(X)$.

Definition 0.2. Let X be a topological space and μ a Borel measure on X. We let $\mathcal{H}_{bac}(X, \mu)$ be the subgroup of $\mathcal{H}(X)$ consisting of all those $g \in \mathcal{H}(X)$ such that $g \cdot \mu \ll \mu$ and $\mu \ll g \cdot \mu$ (equivalently, $g^{-1} \cdot \mu \ll \mu$), that is, the subgroup of all $g \in \mathcal{H}(X)$ such that for $A \in \mathcal{B}(X)$ we have $\mu(A) = 0$ if and only if $\mu(g(A)) = 0$.

Using a different terminology, $\mathcal{H}_{bac}(X, \mu)$ is the group of all homeomorphisms of X that preserve the measure class of μ. In the case in which $X = [0,1]$ and μ is Lebesgue measure this definition agrees with the one in classical analysis given in Definition 0.1. The condition $g \cdot \mu \ll \mu$ for a homeomorphism in g is satisfied precisely when g is absolutely continuous in the sense of Banach, a terminology frequently found in older literature. We could thus refer to $\mathcal{H}_{bac}(X, \mu)$ as the set of absolutely bicontinuous homeomorphisms in the sense of Banach.

Fact 0.3 (Radon–Nikodym; see [13], Corollary 7.34). *If μ, ν are two σ-finite measures on a measurable space $(X, \mathcal{A})$ such that $\nu \ll \mu$, then there exists a unique $\phi \in L^1(X, \mathcal{A}, \mu)$ (that is, well-defined up to a μ-null set) such that $\nu(A) = \int_A \phi \, d\mu$ for all $A \in \mathcal{A}$. We write such $\phi \geq 0$ as $d\nu/d\mu$. It follows that for all $\mathcal{A}$-measurable $h : X \to \mathbb{R}$ we have $h \in L^1(X, \mathcal{A}, \nu)$ if and only if $h\phi \in L^1(X, \mathcal{A}, \mu)$, with equality $\int h \, d\nu = \int h\phi \, d\mu$ in that case.*

This allows one to define the following generalization of the absolute value of the Jacobi determinant, central to the entire subsequent discussion.

Definition 0.4. If $g \in \mathcal{H}_{bac}(X, \mu)$, we write $\mathcal{D}(g) := \frac{dg^*\mu}{d\mu}$.

We establish the following generalization of Solecki's result.

Proposition A. *Let X be a metrizable compact topological space and μ a finite Borel measure on X such that X has a basis of open sets with μ-null frontier. Then the topology τ_{ac} on $\mathcal{H}_{bac}(X, \mu)$ generated by the restriction of the compact-open topology on $\mathcal{H}(X)$ and the topology induced by the distance $\rho_{ac}(g, h) := \|\mathcal{D}(g) - \mathcal{D}(h)\|_1$ it is a Polish group topology.*

When $X \in \{[0, 1], S^1\}$ and μ is Lebesgue measure we recover Solecki's result [21]. Even in the original setting, our proof that τ_{ac} is a group topology diverges somewhat from that in [21]. Our result applies, in particular, to compact manifolds of arbitrary dimension equipped with an Oxtoby–Ulam measure, as defined below.

Definition 0.5. An Oxtoby–Ulam measure on a compact connected manifold X is a finite Borel measure (necessarily regular) satisfying the following properties:

- $\mu(U) > 0$ for all nonempty open $U \subseteq X$,
- $\mu(\{x\}) = 0$ for all $x \in X$
- and $\mu(\partial X) = 0$.

We will refer to (X, μ) for X and μ a probability measure as above as an OU pair.

It is easy to establish the existence of Oxtoby–Ulam measures on an arbitrary compact manifold M. If the manifold is orientable and admits a differentiable structure, one may choose a volume form and define an Oxtoby–Ulam measure, up to renormalization, by letting the measure of an open set be the absolute value of the integral of the volume form over that set. In the general case, we may start with a finite cover of M by embedded topological closed balls of the same dimension and then obtain an Oxtoby–Ulam measure by adding an Oxtoby–Ulam measure on each of the balls.

Another natural way of producing pairs (X, μ) as in the statement of the proposition is the following. Suppose given an inverse system $\pi_i : X_{i+1} \to X_i$ of continuous epimorphisms between compact topological metric spaces, together with finite Borel measures μ_i on X_i satisfying $(\phi_i)_*(\mu_{i+1}) = \mu_i$. Then by Theorem 2.2 in [2] the inverse limit $X := \varprojlim X_i$ can be endowed with a limit Borel measure $\mu := \varprojlim \mu_i$ such that for each of the natural projections $\pi_i : X \to X_i$ we have $(\pi_i)_*(\mu) = \mu_i$. If X_i is generated by sets of μ_i-null frontier for all i, it can be easily checked that the same property holds for (X, μ).

Two natural Borel measures on the Cantor space $\mathcal{C}$ are worth considering. Given some finite measure μ_0 on $\{1, \dots r\}$ one can take the power μ_0^ω measure on

$\{1, \dots r\}^\omega \cong \mathcal{C}$ positive on singletons. We will refer to this as a generalized Bernoulli measure.

Other than this, given any countable field $K \subseteq \mathbb{R}$, one can consider the Fraïssé limit $\mathbb{B}_K$ of the class of all the finite Boolean algebras (see, for instance, [11] for $K = \mathbb{Q}$) equipped with probability measures with values on $K \cap [0, 1]$. By Stone duality, one can identify $\mathbb{B}_\mathbb{Q}$ with the family of clopen sets of $\mathcal{C}$, with the measure on $\mathbb{B}_K$ extending to a unique (up to this identification) Borel measure on $\mathcal{C}$, which we will denote by μ_K.

A gap between group topologies. We will use the term τ_{co} to refer to the restriction of the compact-open topology to the group at hand. The main result of the paper is Theorem B below, which establishes the existence of a gap between τ_{co} and τ_{ac} in a range of situations.

Theorem B. *Let (X, μ) be either*

(I) *an OU pair, or*

(II) *the Cantor space $\mathcal{C} = X$ equipped with either μ_K for some dense subfield $K \subseteq \mathbb{R}$ or a generalized Bernoulli measure.*

Then there is no group topology τ on $\mathcal{H}_{bac}(X, \mu)$ with $\tau_{co} \subsetneq \tau \subsetneq \tau_{ac}$. In fact, more is true: for any separable group topology τ on $\mathcal{H}_{bac}(X, \mu)$ if $\tau_{co} \subsetneq \tau$, then $\tau_{ac} \subseteq \tau$.

In the case of a one-dimensional manifold X, one can read the result above as stating that there is no notion of regularity (for homeomorphisms) sitting between continuity and absolute continuity in the same way absolute continuity sits between C^1-regularity and continuity. A more general set of sufficient conditions for the first claim in the statement is given in Proposition 7.1 and the reader interested only in this claim for the examples covered by Theorem B may skip the discussion in Section 6. In the one-dimensional case we actually know that τ_{co} is the coarsest (not just minimal) Hausdorff group topology on $\mathcal{H}_{bac}(X, \mu)$ by the arguments in [8]. Therefore, we can conclude:

Corollary C. *Let (X, μ) be an OU pair with $X \in \{[0, 1], S^1\}$ and μ be as above. Then τ_{ac} and τ_{co} are the only Hausdorff group topologies on $\mathcal{H}_{bac}(X, \mu)$ coarser than τ_{co}.*

Strategy of the proof. Write $\mathfrak{G} := \mathcal{H}_{bac}(X, \mu)$. The proof of the first statement of Theorem B goes roughly as follows. Assume that we are given a group topology τ with $\tau_{co} \subseteq \tau \subsetneq \tau_{ac}$. By the continuity of multiplication, it suffices to find a fixed integer $K > 0$ such that for every $\mathbb{U} \in \mathrm{N}_{\mathrm{Id}}(\tau)$ there is $\mathbb{U}' \in \mathrm{N}_{\mathrm{Id}}(\tau_{co})$ inside $\mathbb{U}^K$. The second inclusion implies that $\mathbb{U} \in \mathrm{N}_{\mathrm{Id}}(\tau)$ contains some $\mathcal{V} \in \mathrm{N}_{\mathrm{Id}}(\tau_{ac})$. It is not hard to show that for any $g \in \mathfrak{G}$ supported on some set $U \subseteq X$ we have $\|g\|_{ac} \leq 2\mu(U)$ ksee Lemma 3.5), so any such $\mathcal{V}$ will contain the subgroup $\mathfrak{G}_U$ of all elements

supported on U for any open $U \subseteq X$ with $\mu(U) \leq \kappa_{\mathbb{U}}$, for some constant κ. The local fragmentation property for homeomorphism groups tells us that a finite product of the form $\mathfrak{G}_{U_1} \cdots \mathfrak{G}_{U_k}$ is a neighborhood of the identity in τ_{co} if $X = \bigcup_{i=1}^{k} U_i$. The argument above yields the existence of many similar products inside $\mathbb{U}^k$, but only under constraints on $\mu(U_i)$ which might rule out the covering condition.

Note that we can assume $\mathbb{U} = \mathbb{U}^{-1}$, so that for any $U \subseteq X$ with $\mu(U) \leq \kappa_{\mathbb{U}}$ and $g \in \mathbb{U}$ we have $\mathfrak{G}_{g(U)} = g\mathfrak{G}_U g^{-1} \subseteq \mathbb{U}^3$. If $\mu(X \setminus g(U)) < \kappa_{\mathbb{U}}$, then surely we can find an open $U' \subseteq X$ with $\mu(U') \leq \kappa_{\mathbb{U}}$ and $U \cup U'$ and conclude using fragmentation, so that the problem reduces to the existence of $g \in \mathbb{U}$ for which a suitable U exists. This is done in Corollary 5.9, and the main difficulty is how to get there.

It is not difficult to prove that if $\|g\|_{ac}$ is sufficiently close to 2, the supremum of all possible values, then such a U will exist, even though it might be very far from being connected. In order to proceed, we turn to the assumption $\tau \subsetneq \tau_{ac}$, which amounts to the existence of some $\epsilon_0 > 0$ (depending only on τ) such that in every $\mathcal{W} \in \mathrm{N}_{\mathrm{Id}}(\tau)$ we can find $\|g\|_{ac} > \epsilon_0$. Our goal is now to get from "a definite amount" to "almost as large as possible".

Continuity of multiplication is far from enough for that purpose, since, while it is true that any $\mathcal{V} \in \mathrm{N}_{\mathrm{Id}}(\tau)$ will contain arbitrary finite products of the form $g_1 \cdots g_n$ with $\|g_i\|_{ac} > \epsilon_0$, a careless choice of g_i might result in $\ln(\mathcal{D}(g_i))$ for different g_i canceling out under composition, canceling their good properties instead of combining them. On the other hand, we have very little control over the choice of the g_i, and there is no obvious way to ensure that the sign of $\ln(\mathcal{D}(g_i))$ for different values of i agrees in any suitable set. Here, probability comes to our rescue. Given a measurable real function ψ (i.e., a random variable) on a probability space (X, μ), as well as a constant $C > 0$, Chebyshev's inequality allows us to infer a particular behavior of the function on a suitably large set from the properties of an aggregate function of ψ, namely

$$\mu(\{x \in X \mid |\psi(x)| > C\}) \leq \frac{\|\psi\|_2^2}{C^2}, \quad C > 0. \tag{1}$$

And it involves precisely aggregate values derived from $\mathcal{D}(g)$, the kind of thing we have access to.

For the sake of additivity (in the cocycle sense) it is natural and standard to replace $\mathcal{D}(g)$ with $\phi := \overline{\mathcal{D}}(g) := \ln(\mathcal{D}(g))$ and we are interested in our candidate element g being "almost optimal" in the sense above, so that ϕ has very large positive values on the points of some set $X_+ \subseteq X$ of very small measure and negative values of very large absolute value on the points of some set $X_- \subseteq X$ of almost total measure. Note that in any case $\int \mathcal{D}(g)\, d\mu = 1$ (see Observation 1.1), which implies that in passing from $\mathcal{D}(g) - 1$ to $\phi = \bar{D}(g)$ we expect the contribution to the expectation $\int \phi^2\, d\mu$ of the locus of positivity of ϕ to be uniformly bounded, while the contribution of the locus of negativity is expected to blow up. Hence, it makes sense to try to find g

in our arbitrary neighborhood of the identity such that for a given large $C > 0$ we have $\phi \leq -2C + \psi$ so as to conclude $\mu(\{x \in X \mid \phi(x) \leq -C\}) \leq \frac{1}{C}$.

Recall that one can assume that $g = g_1 \cdots g_n$, where $\|g_i\|_{ac} > \epsilon_0$ for $\epsilon_0(\tau)$, but the latter is far from enough for our purposes. The key will be to show that for constants $K, c > 0$ the g_i can be chosen satisfying the following properties:

(a) $\overline{\mathcal{D}}(g)$ is well-approximated by $\sum_{i=1}^{n} \bar{D}(g_i)$.

(b) Each $\overline{\mathcal{D}}(g_i)$ can be approximately bounded from above by a measurable function of the form $-c + \psi_i$, where $\int \psi_i \, d\mu = 0$ and $\left\| \sum_{l=1}^{n} \psi_l \right\|_2^2 \leq K$.

(c) $(\psi_i)_i$ is the sequence of differences of a martingale sequence, a weak form of probabilistic independence (see Definition 3.1).

Writing $\psi = \sum_{j=1}^{n} \psi_j$, (c) implies that $\|\psi\|_2^2 \leq Kn$, so that in (1) we can set $C = nc$ and still have the right-hand side converge to zero at a controlled speed.

Showing this involves two steps. Section 4 is dedicated to the proof of Corollary 4.7, which establishes that the elements g_i in our product can be taken with $\|g_i\|_{ac}$ not only uniformly nonzero, but also sufficiently evenly spread over X in a technical sense. The subsequent Section 5 is almost entirely concerned with the existence of $(g_i)_{i=1}^{n}$ admitting a sequence (ψ_i) as above, stated precisely in Proposition 5.3, from which Corollary 5.9 follows. The property established in Corollary 4.7 is crucial in ensuring that the bound in (b) (uniform on an almost full set) can feature the constant $c > 0$, which stands for the negative excess resulting from applying the logarithm to negative values.

Key to the proof of Proposition 5.3 is that, given $g_0 \in \mathfrak{G}$, if $g_1 \in \mathfrak{G}$ is chosen sufficiently close to the identity in τ_{co} (as a function of g_1), then $\overline{\mathcal{D}}(g_0 g_1)$ is well approximated by the sum $\overline{\mathcal{D}}(g_0) + \overline{\mathcal{D}}(g_1)$, and the two functions behave independently in a certain sense. It is in order to exploit this phenomenon that the assumption $\tau_{co} \subseteq \tau$ is used.

Roelcke precompactness. We say that a topological group (G, τ) is Roelcke precompact if it is totally bounded with respect to the Roelcke uniformity or, in more concrete terms, if for any neighborhood $\mathcal{V}$ of the identity in G there are finitely many elements $\{g_1, \ldots g_r\}$ such that $G = \bigcup_{l=1}^{r} \mathcal{V} g_l \mathcal{V}$. Roelcke precompactness is the topological correlate of the model theoretical notion of ω-categoricity, which explains some of its importance.

In [19] Rosendal develops a theory that endows every topological group in some large family that contains all Polish groups with a canonical metrizable coarse structure. Roelcke precompactness implies boundedness in the sense of said coarse structure. It is shown in [9] that the group $(\mathcal{H}_{bac}(X, \mu), \tau_{ac})$, where $X \in \{[0,1], S^1\}$ and μ is the standard measure, is bounded. One might wonder whether this could just be a consequence of $\mathcal{H}_{bac}(X, \mu)$ being Roelcke precompact, but that is not so:

Theorem D. *If (X, μ) is a 1-dimensional OU pair, then $\mathcal{H}_{bac}(X, \mu)$ is not Roelcke precompact: indeed, if $(g_n)_{n\in\mathbb{N}} \subset \mathcal{H}_{bac}(X, \mu)$ is a sequence such that*

- *$(g_n)_{n\in\mathbb{N}}$ converges to the identity in τ_{co}, and*
- *$(g_n)_{n\in\mathbb{N}}$ does not converge to* Id *in τ_{ac},*

then $(g_n)_{n\in\mathbb{N}}$ is not totally bounded for the Roelcke uniformity.

Remark. When (X, μ) is an OU pair with $\dim(X) \geq 2$, the non-Roelcke precompactness of $\mathcal{H}_{bac}(X, \mu)$ follows from that of $(\mathcal{H}(X), \tau_{co})$ (see [18]) together with Corollary 2.9.

One might wonder whether it might be possible to abstract the proof of Theorem D so as to show that the conclusion holds for any sufficiently nontrivial instance of Proposition A. However, some degree of uniform control of the topology of X over the measure seems to be needed, as evidenced by our last result, which plays out in the opposite direction.

Theorem E. *Let K a countable subfield of $\mathbb{R}$. Then $(\mathcal{H}_{bac}(\mathcal{C}, \mu_K), \tau_{ac})$ is Roelcke precompact.*

Structure of the paper. Section 1 is dedicated to the proof of Proposition A, while introducing useful notation and results used later on. Section 2 focuses on the properties of $\mathcal{H}_{bac}(X, \mu)$ in the case in which (X, μ) is an OU pair. It collects technical statements generalizing properties of the group of homeomorphisms of a compact manifold, such as a fragmentation property relative to τ_{co}. Section 3 collects a handful of miscellaneous results and definitions of a probabilistic nature to be used in later sections. Sections 4 and 5 are directly concerned with the proof of Proposition 5.3, key to Theorem B, as described in more detailed in the summary of the proof of Theorem B provided above.

Section 6 considers a group topology τ in $\mathcal{H}_{bac}(X, \mu)$ of which it is only assumed to be separable in the case in which (X, μ) is an OU pair. Adapting well known arguments found in the literature on automatic continuity of homeomorphism groups yields conclusions which fall short of automatic continuity for $\mathcal{H}_{bac}(X, \mu)$, but which can be combined with Corollary 5.9 to recover the stronger claim in Theorem B. The theorem itself is proved in Section 7, while Section 8 is devoted to the proofs of Theorems D and E.

1. A family of Polish groups

Let X be a compact metric space and μ a finite Borel measure on X. We will assume throughout the paper that in such situation some compatible metric d on X has been fixed. For the sake of simplicity, we write $\mathfrak{G} := \mathcal{H}_{bac}(X, \mu)$.

For readability, in the presence of integrals we will often use $\circ$ to denote the composition of elements in $\mathcal{H}(X)$ and write Id for the identity map. We write $\mathbb{1}_A$ for the indicator function of a set $A \subseteq X$, and simply 1 in place of $\mathbb{1}_X$.

The following two observations are well-known and easy to show (see [21]). We include a proof of the next one for completeness.

Observation 1.1. *For any $g \in \mathfrak{G}$ and any $\phi \in L^1(X, \mu)$ we have*

$$\int (\phi \circ g)\mathcal{D}(g)\, d\mu = \int \phi\, d\mu$$

Proof. $\int (\phi \circ g)\mathcal{D}(g)\, d\mu = \int \phi \circ g\, d(g^{-1} \cdot \mu) = \int \phi \circ g \circ g^{-1}\, d\mu = \int \phi\, d\mu$. □

Using this and the fact that $\mathcal{D}(g)$ is well-defined up to a set of measure zero, one also checks:

Observation 1.2. *For any $g_1, g_2 \in \mathfrak{G}$ we have*

$$\mathcal{D}(g_1 \circ g_2) = (\mathcal{D}(g_1) \circ g_2)\mathcal{D}(g_2), \quad \mathcal{D}(g_1^{-1}) = \frac{1}{\mathcal{D}(g_1) \circ g_1^{-1}}, \quad \text{a.e.}$$

Recall that given $g_1, g_2 \in \mathfrak{G}$ we write

$$\rho_{ac}(g_1, g_2) := \|\mathcal{D}(g_1) - \mathcal{D}(g_2)\|_1, \quad \|g_1\|_{ac} := \rho_{ac}(g_1, \mathrm{Id}).$$

Clearly ρ_{ac} is a pseudometric on $\mathfrak{G}$. For $A \in \mathcal{B}(X)$, $f \in L^1(X, \mu)$ and $g \in \mathfrak{G}$ we also write

$$\|f\|_{1,A} := \|\mathbb{1}_A f\|_1, \quad \|g\|_{ac,A} := \|\mathcal{D}(g) - 1\|_{1,A}.$$

Write $\tau_{ac} := \tau_{ac}(X, \mu)$ for the topology generated by the restriction to $\mathfrak{G}$ of the compact-open (uniform convergence) topology on $\mathcal{H}(X)$, which we shall denote henceforth simply as τ_{co}, and the topology induced by the pseudometric ρ_{ac} above.

Fix a compatible metric d on X (necessarily complete).

Definition 1.3. Given $\delta > 0$ we write

$$\mathcal{V}_\delta := \{g \in \mathfrak{G} \mid \sup_{x \in X} d(g(x), x) \leq \delta\}, \quad \mathcal{V}'_\delta := \{g \in \mathfrak{G} \mid \|g\|_{ac} \leq \delta\}, \quad \mathcal{W}_\delta := \mathcal{V}_\delta \cap \mathcal{V}'_\delta,$$

so that a system of neighborhoods of Id in τ_{ac} is given by $\{\mathcal{W}_\delta\}_{\delta > 0}$.

The following observation is well-known (the first statement is a formal generalization of a result in [21]), but we include a proof for completeness.

Observation 1.4. *For any $h_1, h_2, g \in \mathfrak{G}$ we have $\rho_{ac}(h_1, h_2) = \rho_{ac}(h_1 \circ g, h_2 \circ g)$. Therefore, the following equalities hold*:

$$\|h_1\|_{ac} = \|h_1^{-1}\|_{ac}, \quad \|h_1 \circ h_2\|_{ac} \leq \|h_1\|_{ac} + \|h_2\|_{ac}.$$

Proof. The first equality is an easy consequence of Observations 1.1 and 1.2, while the second equality follows by taking $h_2 = \mathrm{Id}$, $g = h_1^{-1}$. For the final inequality:

$$\|h_1 \circ h_2\|_{ac} = \rho_{ac}(h_1, h_2^{-1}) \leq \rho_{ac}(h_1, 1) + \rho_{ac}(h_2^{-1}, 1) = \|h_1\|_{ac} + \|h_2\|_{ac},$$

where the last step follows from the equality $\|h_2^{-1}\|_{ac} = \|h_2\|_{ac}$. □

Definition 1.5. Given $f \in C^0(X)$ we write

$$\omega_f : \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}, \quad \omega_f(\delta) = \sup\big\{|f(x) - f(y)| \,|\, d(x, y) \leq \delta\big\}.$$

The following fact will be used throughout the paper:

Observation 1.6. *Let $g \in \mathfrak{G}$ and $A \in \mathcal{B}(X)$. If $\phi \in C^0(X)$ satisfies $\|\phi - \mathcal{D}(g)\|_{1,A} \leq \epsilon$ and $\omega_\phi(\delta) \leq \epsilon'$ then for any $h \in \mathcal{V}_\delta$ we have*

$$\|\mathcal{D}(g \circ h) - \phi\mathcal{D}(h)\|_{1,A} \leq \epsilon + \epsilon' \mu(h(A)).$$

If $\omega_\phi(2\delta) \leq \epsilon'$ and $\psi := \sum_{i=1}^m \lambda_i \mathbb{1}_{A_i}$ for some partition $\{A_i\}_{i=1}^m$ of X into Borel sets of diameter at most δ and $\lambda_i \in [\inf_{A_i} \phi, \sup_{A_i} \phi]$, then the inequality above holds with ψ in place of ϕ.

Proof. Using the composition rule and the triangular inequality for the first inequality and Observation 1.1 in the equality after that we obtain

$$\begin{aligned}\|\mathcal{D}(g\circ h) - \phi\mathcal{D}(h)\|_{1,A} &\leq \|((\mathcal{D}(g) - \phi)\circ h)\mathcal{D}(h)\|_{1,A} + \|(\phi \circ h - \phi)\mathcal{D}(h)\|_{1,A} \\ &= \|\mathcal{D}(g) - \phi\|_{1,A} + \|(\phi\circ h - \phi)\mathcal{D}(h)\|_{1,A} \\ &\leq \epsilon + \|\phi\circ h - \phi\|_\infty \|\mathcal{D}(h)\|_{1,A} \leq \epsilon + \epsilon' \mu(h(A)).\end{aligned}$$

The second claim follows similarly, since the assumption that $\omega_\phi(2\delta) < \epsilon'$ implies that $\|\phi \circ h - \psi\|_\infty \leq \epsilon'$. □

Lemma 1.7. *For any $g \in \mathfrak{G}$ and $\epsilon > 0$ there is $\delta = \delta(\epsilon, g)$ such that for any $h \in \mathcal{W}_\delta$ we have $\|h^{g^{-1}}\|_{ac} = \rho_{ac}(g, g \circ h) \leq \epsilon$.*

Proof. Choose some $\phi \in C^0(X)$ such that $\|\phi - \mathcal{D}(g)\|_1 < \frac{\epsilon}{4}$ and $\delta > 0$ such that

$$\delta \leq \min\Big\{\omega_\phi\Big(\frac{\epsilon}{4}\Big),\ \frac{\epsilon}{4\|\phi\|_\infty}\Big\}.$$

Then

$$\begin{aligned}\|\mathcal{D}(g) - (\mathcal{D}(g)\circ h)\mathcal{D}(h)\|_1 &\leq \|\phi - \phi\mathcal{D}(h)\|_1 + \|\phi - \mathcal{D}(g)\| + \|((\phi - \mathcal{D}(g))\circ h)\mathcal{D}(h)\|_1 \\ &\leq \|\phi - \phi\mathcal{D}(h)\|_1 + \tfrac{\epsilon}{4} + \big(\tfrac{\epsilon}{4} + \tfrac{\epsilon}{4}\big) \\ &\leq \|\phi\|_\infty \|\mathcal{D}(h) - 1\|_1 + \tfrac{3}{4}\epsilon \leq \epsilon,\end{aligned}$$

where we used Observation 1.6 in the second inequality. □

Corollary 1.8. *Under the assumptions on X and μ made at the beginning of the section, τ_{ac} is a group topology on $\mathfrak{G}$.*

Proof. Lemma 1.7, together with the fact that τ_{co} is a group topology, implies that the unique left-invariant topology on $\mathfrak{G}$ for which $\{\mathcal{W}_\delta\}_{\delta>0}$ is a system of neighborhoods of the identity is in fact a group topology. It has to coincide with τ_{ac} by the same lemma. □

It remains to show that $(\mathfrak{G}, \tau_{ac})$ is Polish. Consider the metric $\rho : \mathfrak{G} \times \mathfrak{G} \to \mathbb{R}_{\geq 0}$ on $\mathfrak{G}$ given by

$$\rho(g_1, g_2) = \sup_{x\in X} d(g_1(x), g_2(x)) + \sup_{x\in X} d(g_1^{-1}(x), g_2^{-1}(x)) + \rho_{ac}(g_1, g_2) + \rho_{ac}(g_1^{-1}, g_2^{-1}).$$

This induces the topology τ_{ac} on $\mathfrak{G}$ by Observation 1.4, together with the fact that the metric given by the sum of the first two terms above is a complete metric for the compact-open topology.

Fact 1.9. *Suppose we are given two finite measures μ, ν on a measurable space $(X, \mathcal{A})$ for which $\nu \ll \mu$. Then for any $\epsilon > 0$ there exists some $\delta > 0$ such that $\nu(A) < \epsilon$ for all $A \in \mathcal{A}$ with $\mu(A) < \delta$.*

Given $g \in \mathfrak{G}$ and $\delta > 0$ we write

$$\Omega_g(\delta) := \sup\left\{\mu(g(A)) \mid A \in \mathcal{B}(X),\ \mu(A) \leq \delta\right\}.$$

Lemma 1.10. *Let $\underline{g} = (g_n)_{n\geq 0}$ be a sequence in $\mathfrak{G}$ that is Cauchy with respect to ρ. Then the function $\Omega_{\underline{g}} = \sup_{n\geq 0} \Omega_{g_n}$ satisfies $\lim_{\epsilon\to 0} \Omega_{\underline{g}}(\epsilon) = 0$.*

Proof. For a fixed $\epsilon > 0$, take n_0 such that for all $n \geq n_0$ we have $\|\rho_{ac}(g_n, g_{n_0})\| \leq \frac{\epsilon}{2}$. Then $\sup_{n\geq 0} \mu(g_n(A)) \leq \epsilon$ for any $A \in \mathcal{B}(X)$ such that $\Omega_{g_k}(\mu(A)) \leq \frac{\epsilon}{2}$ for all $0 \leq k \leq n_0$. □

Lemma 1.11. *ρ is a complete metric on $\mathfrak{G}$.*

Proof. Let $\underline{g} = (g_n)_{n\geq 0}$ be a Cauchy sequence in $\mathfrak{G}$ with respect to ρ. Then $(g_n)_{n\geq 0}$ must converge in τ_{co} to some $g_\infty \in \mathcal{H}(X)$, by the completeness of the metric $\rho_\infty(g, h) = \sup_{x\in X} d(g_1(x), g_2(x)) + \sup_{x\in X} d(g_1^{-1}(x), g_2^{-1}(x))$ on $\mathcal{H}(X)$. We claim that $g_\infty \in \mathcal{H}_{bac}(X, \mu)$. Note that $(\mathcal{D}(g_n))_{n\geq 0}$ must be Cauchy in $L^1(X, \mu)$ and thus must converge in $L^1(X, \mu)$ to some function $\Psi \geq 0$, which implies that, for any $A \in \mathcal{B}(X)$,

$$\mu(g_n(A)) = \int_A \mathcal{D}(g_n)\, d\mu \xrightarrow{n\to\infty} \int_A \Psi\, d\mu. \tag{2}$$

Since $g_n \to g_\infty$ in τ_{co}, we have

$$g_\infty(U) \subseteq \bigcup_{n\geq 0} \bigcap_{k\geq n} g_k(U)$$

for any open set $U \subseteq X$. The continuity of μ then allows us to conclude that

$$\mu(g_\infty(U)) \leq \lim_{n\to\infty} \mu(g_n(U)) \tag{3}$$

and so, in particular, $\mu(g_\infty(U)) \le \Omega_{\underline{g}}(\mu(U))$ for any such U, which implies that $\mu(g_\infty(A)) = 0$ for any $A \in \mathcal{B}(X)$ with $\mu(A) = 0$, by Lemma 1.10 and the regularity of μ. A similar observation can be made for the inverse g_∞^{-1} and the (Cauchy) sequence $(g_n^{-1})_{n\ge 0}$ and so we conclude that $g_\infty \in \mathcal{H}_{bac}(X, \mu)$.

Now if $U \subseteq X$ is an open set with $\mu(\mathrm{Fr}(U)) = 0$ then $\mu(\mathrm{Fr}(g_\infty(U))) = 0$, and thus by (3) we have

$$\mu(X) = \mu(g_\infty(U)) + \mu(g_\infty(X \setminus \bar{U})) \le \lim_{n\to\infty} [\mu(g_n(U)) + \mu(g_n(X \setminus \bar{U}))] = \mu(X),$$

which implies

$$\mu(g_\infty(U)) = \lim_{n\to\infty} \mu(g_n(U)) = \int_U \Psi \, d\mu$$

for all such U. Any finite union of open sets with μ-null frontier also has μ-null frontier, and open sets with μ-null frontier generate the topology of X by assumption. This implies that every closed (i.e., compact) set $F \subseteq X$ is the intersection of a descending sequence of open sets $U_0 \supset U_1 \supset \cdots$ with μ-null frontier, as follows: for each n the set F is covered by finitely many open sets of diameter at most $\frac{1}{n}$ with μ-null frontier, so we can take U_n as their union (intersected with U_{n-1} to ensure nestedness).

Using the dominated convergence theorem, we now have

$$\mu(g_\infty(F)) = \lim_n \mu(g_\infty(U_n)) = \lim_n \int_{U_n} \Psi \, d\mu = \int_F \Psi \, d\mu.$$

The regularity of the measures $g_\infty^*\mu$ and $A \mapsto \int_A \Psi \, d\mu$ now easily shows that $\mu(g_\infty(A)) = \int_A \Psi \, d\mu$ for all $A \in \mathcal{B}(X)$ and thus that $\mathcal{D}(g_\infty) = \Psi$ and $\mathcal{D}(g_n)$ converge to $\mathcal{D}(g_\infty)$ in L_1. Arguing similarly for g_∞^{-1} and $(g_n^{-1})_{n\ge 0}$, we conclude that $\lim_{n\to\infty} \rho(g_n, g_\infty) = 0$. □

To conclude the proof of Proposition A we note that the separability of X implies that of $\mathcal{H}_{bac}(X, \mu)$.

Observation 1.12. *The topology τ_{ac} is separable.*

Proof. Every compact metric space is second countable. It follows that $\mathfrak{h}(X)$ is itself separable. Let $\{\phi_n\}_{n\ge 1}$ be some dense subset of $L^1(X, \mu)$ and $\{h_n\}_{n\ge 1}$ a countable dense subset of $\mathcal{H}(X)$ for the compact-open topology. For every pair of integers n, m and $N \ge 1$, let $g_{n,m,N} \in \mathcal{H}_{bac}(X, \mu)$ be either an element such that $\|\mathcal{D}(g_{n,m,N}) - \phi_n\|_1 \le \frac{1}{N}$ and $\sup_{x\in X} d(g_{n,m,N}(x), h_n(x)) \le \frac{1}{N}$ in the case that such an element exists and arbitrary otherwise. It is easy to check that $\{g_{n,m,N}\}_{n,m,N\ge 1}$ is dense in τ_{ac}. □

2. Some properties of $\mathcal{H}_{bac}(X, \mu)$ for an Oxtoby–Ulam pair (X, μ)

Fix an OU pair (X, μ). The goal of this section is to generalize some basic properties of homeomorphism groups and the compact-open topology to our group

$\mathfrak{G} = \mathcal{H}_{bac}(X, \mu)$ of absolutely bicontinuous homeomorphisms with the restriction topology. We introduce ball transitivity (Corollary 2.4), local fragmentation (Lemma 2.13), and uniform perfectness of the group (Lemma 2.15). Of these, only fragmentation will be required in the proof of the first statement of Theorem B, whereas the rest will be used only to generalize some arguments from existing automatic continuity proofs in Section 6.

For any subset A of a topological space we will use the notation $\mathring{A}$, $\mathrm{Fr}(A)$, $\bar{A}$ to refer to the interior, frontier and closure of A in X, so that $\mathrm{Fr}(A) = \bar{A} \cap \overline{X \setminus A}$. In the context of our topological manifold X we will use the notation ∂X to refer to the boundary of X. By a *nicely embedded ball* in X we mean the homeomorphic image in $\mathrm{Int}(X)$ of the standard closed m-ball $\bar{B}^m(1)$, which can be extended to a homeomorphic embedding of the open ball $B^m(2)$. By a *boundary half-ball* we mean the image of $\bar{B}^m(1) \cap \mathbb{R}^{m-1} \times \mathbb{R}_{\geq 0}$ by a homeomorphic embedding $h : \bar{B}^m(1) \cap \mathbb{R}^{m-1} \times \mathbb{R}_{\geq 0}$, which extends to a homeomorphic embedding $\tilde{h} : B^m(2) \cap \mathbb{R}^{m-1} \times \mathbb{R}_{\geq 0} \to X$ with the property that $\tilde{h}^{-1}(\partial X) = B^m(2) \cap \{x_m = 0\}$. By a *$\mu$-adapted ball* (*resp. boundary half-ball*) we mean a nicely embedded ball (resp. boundary half-ball) whose frontier is μ-null.

We denote by $\mathcal{H}(X, \partial X)$ the group of homeomorphisms of X fixing ∂X pointwise. Note that this is still a Polish group when endowed with the compact-open topology.

A useful tool will be the following classical result due to Oxtoby and Ulam.

Fact 2.1 [16, Theorem 2]. *Let X be a compact manifold and μ, μ' two Oxtoby–Ulam measures on X with $\mu(X) = \mu'(X)$. Then there exists some $h \in \mathcal{H}(X, \partial X)$ such that $h_*\mu = \mu'$.*

Given $g \in \mathcal{H}(X)$, write $\mathrm{supp}(g) := \{x \in X \mid g(x) \neq x\}$ and denote its closure by $\overline{\mathrm{supp}}(g)$. Given an open set $U \subseteq X$ and $\mathfrak{h} \leq \mathcal{H}(X)$ we write

$$\mathfrak{h}_U := \{g \in \mathfrak{h} \mid \mathrm{supp}(g) \subseteq U\}, \qquad \mathfrak{h}_U^c := \{g \in \mathfrak{h} \mid \overline{\mathrm{supp}}(g) \subseteq U\}.$$

Given $A \in \mathcal{B}(X)$ we say that $h \in \mathcal{H}(X)$ is measure-class preserving on A if for all $C \in \mathcal{B}(X)$, $C \subseteq A$ we have that $\mu(C) = 0$ if and only if $\mu(h(C)) = 0$. The following is obvious from σ-additivity:

Observation 2.2. *Let $F_k \in \mathcal{B}(X)$, $k \geq 0$. Then $g \in \mathcal{H}(X)$ is measure-class preserving on $\bigcup_{k\geq 0} F_k$ if it is on F_k for all $k \geq 0$.*

Observation 2.3. *Let $h \in \mathcal{H}(X)$ and assume that Y is a closed subspace of X homeomorphic to a compact m-manifold such that $\mu(\partial Y) = 0$. Then there exists some $g \in \mathcal{H}(X)$ supported on Y such that hg is measure-class preserving on $\bigcup_{j\in\mathcal{F}} Y_j$. We may choose g so that*

$$\|hg\|_{ac,Y} = \big|\mu(Y) - \mu(h(Y))\big|.$$

Proof. Apply Fact 2.1 to the two OU measures

$$\nu_1 := \frac{1}{\mu(h(Y))} h^*(\mu_{\restriction h(Y)}) \quad \text{and} \quad \nu_2 := \frac{1}{\mu(Y)} \mu_{\restriction Y}$$

on the compact manifold Y. This yields an element $\gamma \in \mathcal{H}(Y, \partial Y)$ such that $\gamma^*(\nu_1) = \nu_2$. We then have

$$(h \circ \gamma)^*(\mu_{\restriction h(Y)}) = \mu(h(Y))\gamma^*(\nu_1) = \mu(h(Y))\nu_2 = \frac{\mu(h(Y))}{\mu(Y)} \mu_{\restriction Y}.$$

Therefore, if we extend γ by the identity outside Y to an element of $\mathcal{H}(X)$, then hg is measure-class preserving on Y. Thus, $\mathcal{D}(hg)$ is well-defined on Y and equal to the constant function $\mu(h(Y))/\mu(Y)$, so that $\|hg\|_{ac,Y}$ is also well-defined and equal to $\big|\mu(Y) - \mu(h(Y))\big|$. □

The following is most likely known, but we have failed to find an appropriate reference.

Corollary 2.4. *Let $\{B_i\}_{i=1}^r$, $\{B_i'\}_{i=1}^r$ be two collections of disjoint μ-adapted balls and boundary half-balls X with the property that B_i is a boundary half-ball if and only if B_i' is, in which case they both intersect the same boundary component. If $m = 1$, we assume that the two collections are ordered in the same way up to orientation. If $m = 2$, then we also assume that the intervals in which B_i and B_i' touch the boundary components of X are ordered identically for the two collections. Then there exists $h \in \mathfrak{G}$ such that $h(B_i) = B_i'$ for all $1 \le i \le r$.*

Proof. A standard application of the annulus theorem of Kirby [12] and Quinn [17] shows the existence of $h_0 \in \mathcal{H}(X)$ with $h_0(B_i) = B_i'$ for all $1 \le i \le r$. Note that for this to work it is crucial that B_i and B_i' are nicely embedded, as there are exotically embedded balls to which the theorem does not apply.

The result then follows by applying Observation 2.3 to each of the members of the covering of X given by $\{B_i\}_{i=1}^r \cup \{\overline{X \setminus \bigcup_{i=1}^r B_i}\}$. For each member A of this covering, we thus obtain an element g_A supported on A with the property that $h_0 \circ g_A$ is measure-class preserving on A. The product of all the (pairwise commuting) g_A is an element g with the property that $h := h_0 g \in \mathfrak{G}$ and acts on the balls as needed.[1] □

[1]The annulus theorem (originally an important conjecture) states that for any two m-balls B_0, B_1 with B_1 nicely embedded in B_0, the complement $B_0 \setminus \mathring{B}_1$ is homeomorphic to $S^{m-1} \times [0,1]$. For the statement above in the boundaryless case one first finds disjoint nicely embedded balls $\{E_i\}_{i=1}^r$ such that $B_i, B_i' \subseteq \mathring{E}_i$. The annulus theorem provides homeomorphisms $\phi_i : E_i \setminus \mathring{B}_i \cong E_i \setminus \mathring{B}_i'$, which can be pasted with suitable homeomorphisms $\psi_i : B_i \cong B_i'$ and the identity outside of $\bigcup_{i=1}^r E_i$ to yield the desired global homeomorphism h_0. In the presence of boundary half-balls one can argue in a similar way. All one needs to show is that given boundary half balls B_0, B_1 with $B_1 \subseteq \mathring{B}_0$ the homeomorphism type of the pair $(B_0 \setminus \mathring{B}_1, B_0 \setminus \mathring{B}_1 \cap \partial M)$ does not depend on B_0 and B_1. Note that the complement $B_2 := B_0 \setminus \mathring{B}_1$ is homeomorphic to an m-ball by the generalized Schoenflies

Fact 2.5 [14, Lemma 3.10]. *Given a compact manifold X, there is some $N = N(X) > 0$ such that for any $\epsilon > 0$ we can write $X = \bigcup_{l=1}^{N} V_l$ with $V_l = \coprod_{j \in J_l} \mathring{F}_{j,l}$, where $\{F_{j,l}\}_{l \in J_l}$ is a collection of disjoint balls and boundary balls of diameter less than ϵ.*

Lemma 2.6. *In the conclusion of Fact 2.5 we may assume that all balls and boundary half-balls $F_{j,l}$ involved are μ-adapted.*

Proof. It is enough to show that given a finite collection $\{D_i\}_{i \in I}$ of balls and boundary half balls whose interiors cover our compact manifold X one can always find μ-adapted (boundary half) balls $D_i' \subseteq D_i$, $i \in I$ of the same type that are μ-adapted and whose interiors cover X.

Claim 2.7. *There is $\delta > 0$ such that the union of the interiors of some collection $\{D_i'\}_{i \in I}$ as above (not necessarily μ-adapted) will cover X provided D_i is contained in $D_i' \cup \mathrm{N}_\delta(\mathrm{Fr}(D_i))$.*

In view of this, all we need to show is that for every D_i there is another ball or boundary half ball of the same type, $D_i' \subseteq D_i$, such that

(i) D_i' is μ-adapted and

(ii) $D_i \subseteq \mathrm{N}_\delta(\mathrm{Fr}(D_i)) \cup D_i'$.

Such D_i' can be found as follows. Consider the Borel measure $\mu_{\mathring{D}_i}$ on D_i obtained by removing all μ-mass from $\mathrm{Fr}(D_i)$, i.e., $\mu_{\mathring{D}_i}(C) = \mu(C \cap \mathring{D}_i)$ for all $C \in \mathcal{B}(D_i)$. It is trivial to check that this is an Oxtoby–Ulam measure on D_i. Thus Fact 2.1 applies, and so for some $\theta > 0$ we have that $\theta \mu_{\mathring{D}_i}$ is the push-forward of the Lebesgue measure λ on the euclidean unit ball $\overline{B}^m$ by some homeomorphism $\phi : \overline{B}^m \to D_i$. The rescaled balls $\alpha \overline{B}^m \subseteq B^m$, $\alpha \in (0,1)$, have λ-null frontier, so that

$$\mu(\mathrm{Fr}(\phi(\alpha \overline{B}^m))) = \mu_{\mathring{D}_i}(\phi(\mathrm{Fr}(\alpha \overline{B}^m))) = \theta^{-1} \lambda(\mathrm{Fr}(\alpha \overline{B}^m)) = 0,$$

that is, $\phi(\alpha \overline{B}^m)$ satisfies (i). By continuity, it will satisfy (ii) for α sufficiently close to 1, and so we are done.

As for the claim, the open cover $\{\mathring{D}_i\}_{i \in I}$ must admit a positive Lebesgue number $\delta > 0$, meaning that for any point $x \in X$ we have $B(x, \delta) \subseteq \mathring{D}_i$ for some $i \in I$. We show that we can take this to be the δ in the statement of the claim. For the sake of contradiction, assume that there is $x \in X \setminus \bigcup_{i \in I} \mathring{D}_i'$. If $x \in \mathring{D}_i$, then $x \in \mathring{D}_i \setminus \mathring{D}_i' \subseteq \mathrm{N}_\delta(\mathrm{Fr}(D_i))$ so that it cannot be the case that $B(x, \delta) \subseteq D_i$. It follows that $B(x, \delta) \nsubseteq \mathring{D}_i$ for all i, contradicting the choice of δ. □

For $\delta > 0$ write $\tilde{\mathcal{V}}_\delta := \{g \in \mathcal{H}(X) \mid \forall x \in X \, d(x, g(x)) \leq \delta\}$, so that $\mathcal{V}_\delta = \tilde{\mathcal{V}}_\delta \cap \mathfrak{G}$.

theorem [15], so all that is left to show is that the homeomorphism type of the pair $(\partial B_2, \partial B_2 \cap \partial M)$ does not depend on B_0 and B_1, which is a consequence of the annulus theorem in dimension $m - 1$.

Lemma 2.8. *Let $g \in \mathcal{H}(X)$ and let U be an open subset of X, $F \subseteq U$ a compact set, and $\epsilon > 0$. Then there exists some $h \in \mathcal{H}(X, \partial X)$ such that*

- $F \subseteq \mathrm{supp}(h) \subseteq \overline{\mathrm{supp}}(h) \subseteq U$,
- $h \in \tilde{\mathcal{V}}_\epsilon$, *and*
- gh *is measure-class preserving on* $\overline{\mathrm{supp}}(h)$.

Proof. Let $N(X)$ be as in Fact 2.5. Using Lemma 2.6 construct N finite families $\{B_i^l\}_{i \in \mathcal{J}_l}$ of μ-adapted balls and boundary half balls such that $\mathrm{diam}(B_i^l) < \epsilon/N$ for all $1 \leq l \leq N$ and $i \in \mathcal{J}_l$ and the union of the interiors of all the B_i^l covers X.

Let $C_1, \ldots C_m$ be an enumeration of all sets of the form B_i^l that intersect F. Clearly, $F \subseteq \bigcup_{j=1}^m \mathring{C}_j$ and, by taking ϵ to be small enough, we may assume that $\bigcup_{j=1}^m C_j \subset U$.

Using Observation 2.3 we can find $h_1 \in \mathcal{H}(X, \partial X)$ with $\mathrm{supp}(h_1) \subseteq C_1$ such that gh_1 is measure-class preserving on C_1. Applying 2.3 once more (to the homeomorphism $g' = gh_1$) yields $h_2 \in \mathcal{H}(X, \partial X)$ such that gh_1h_2 is measure-class preserving on C_2. Since $(gh_1)_{\restriction C_1 \setminus C_2} = (gh_1h_2)_{\restriction C_1 \setminus C_2}$, gh_1h_2 is also measure-class preserving on $C_1 \setminus C_2$. By Observation 2.2 gh_1h_2 is thus measure-class preserving on $C_1 \cup C_2$. Continuing in this fashion, we obtain $h_1, \ldots h_m \in \mathcal{H}(X, \partial X)$ with h_j supported on C_j and such that if $h := h_1h_2 \cdots h_m$, the composition gh is measure-class preserving on $\bigcup_{j=1}^m C_j$, which is a neighborhood of F contained in U.

Since $\mathrm{diam}(C_j) < \epsilon/N$ and one can still separate the family $\{C_j\}_{j=1}^m$ into N distinct families of disjoint sets, it is not hard to see that $h \in \tilde{\mathcal{V}}_\epsilon$ and so we are done. □

Corollary 2.9. *For any OU pair (X, μ) the group $\mathcal{H}_{bac}(X, \mu)$ is dense in $\mathcal{H}(X)$ with respect to τ_{co}.*

Local fragmentation for the restriction of the compact-open topology. We write $\mathrm{N}_\epsilon(A)$ for the metric neighborhood $\mathrm{N}_\epsilon(A) := \{x \in X \mid d(x, A) < \epsilon\}$.

Definition 2.10. We say that a subgroup $\mathfrak{h} \leq \mathcal{H}(X)$ satisfies the weak local fragmentation property for τ_{co} if for any finite collection $U_1, \ldots U_r, V$ of open subsets of X with $\overline{V} \subseteq \bigcup_{j=1}^r U_j$ there is $\delta > 0$ such that

$$\mathfrak{h}_V \cap \tilde{\mathcal{V}}_\delta \subseteq \mathfrak{h}_{U_1}^c \cdots \mathfrak{h}_{U_r}^c.$$

If in the situation above for any $\epsilon > 0$ there is $\delta > 0$ such that

$$\mathfrak{h}_V \cap \tilde{\mathcal{V}}_\delta \subseteq (\mathfrak{h}_{U_1}^c \cap \tilde{\mathcal{V}}_\epsilon) \cdots (\mathfrak{h}_{U_r}^c \cap \tilde{\mathcal{V}}_\epsilon),$$

then we say that $\mathfrak{h}$ satisfies the local fragmentation property for τ_{co}.

Lemma 2.11. *For any compact manifold X, if $\mathfrak{h} \leq \mathfrak{h}(X)$ $\mathfrak{h}$ satisfies the weak local fragmentation property with respect to τ_{co}, then it satisfies the local fragmentation property with respect to τ_{co}.*

Proof. Let $V, U_1, \dots U_r \subseteq X$ be open subsets with $\overline{V} \subseteq \bigcup_{i=1}^r U_i$ and $\epsilon > 0$. Let $N = N(X)$ be as in Fact 2.5. Then $\overline{V} \subseteq \bigcup_{i=1}^r \bigcup_{j=1}^N V_{i,j}$, where each $V_{i,j}$ is the union of the interiors of some disjoint collection of compact subsets of V_i of diameter at most ϵ/N. The desired conclusion now follows from the weak local fragmentation property applied to V and $\{V_{i,j} \mid 1 \le i \le r,\ 1 \le j \le N\}$ and the observation that $\mathfrak{h}_{V_{i,1}} \cdots \mathfrak{h}_{V_{i,N}} \subseteq (\tilde{\mathcal{V}}_{\epsilon/N})^N \subseteq \tilde{\mathcal{V}}_\epsilon$. □

We recall the following fact, due to Edwards and Kirby [6].

Fact 2.12. *Let M be a compact manifold. Then $\mathcal{H}(X)$ and $\mathcal{H}(X, \partial X)$ have the weak local fragmentation property for τ_{co}.*

Lemma 2.13. *If (X, μ) is an OU pair, then $\mathfrak{G} := \mathcal{H}_{bac}(X, \mu)$ satisfies the weak local fragmentation property for τ_{co}.*

Proof. Write $\mathfrak{h}$ for $\mathcal{H}(X)$ and let $U_1, \dots U_r, V$ be open sets in X such that $\overline{V} \subseteq \bigcup_{j=1}^r U_j$. We will show by induction on r the existence of some $\delta > 0$ such that

$$\mathfrak{G}_V \cap \mathcal{V}_\delta \subseteq \mathfrak{G}^c_{U_1} \cdots \mathfrak{G}^c_{U_r}. \tag{4}$$

The weak local fragmentation property for τ_{co} will then follow. The base case $r = 1$ is trivial, so assume $r \ge 2$. It is easy to find an open set U' such that $\overline{U}' \subseteq \bigcup_{l \ne r} U_l$ and $\overline{V} \subseteq U' \cup U_r$. By the induction hypothesis, there exists $\delta_0 > 0$ such that

$$\mathfrak{G}_{U'} \cap \mathcal{V}_{\delta_0} \subseteq \mathfrak{G}^c_{U_1} \cdots \mathfrak{G}^c_{U_{r-1}}. \tag{5}$$

Next, let $\epsilon := \delta_0/2$. By the local fragmentation property for the ambient group $\mathfrak{h}$ there exists $\delta > 0$ such that

$$\mathfrak{h}_V \cap \tilde{\mathcal{V}}_\delta \subseteq (\mathfrak{h}^c_{U'} \cap \tilde{\mathcal{V}}_\epsilon)(\mathfrak{h}^c_{U_r} \cap \tilde{\mathcal{V}}_\epsilon). \tag{6}$$

We claim that condition (4) holds for the δ given above. To see this, choose an arbitrary $g \in \mathfrak{G}_V \cap \mathcal{V}_\delta$. By (6) we can write $g = \hat{g}_1 \hat{g}_2$, where $\hat{g}_1 \in \mathfrak{h}^c_{U'} \cap \tilde{\mathcal{V}}_\epsilon$ and $\hat{g}_2 \in \mathfrak{h}^c_{U_r} \cap \tilde{\mathcal{V}}_\epsilon$.

By applying Lemma 2.8 with $F := \overline{\mathrm{supp}}(\hat{g}_1) \cap \overline{\mathrm{supp}}(\hat{g}_2)$ and $U := U' \cap U_r$, there exists some $h \in \mathcal{H}(X, \partial X)$ such that

- $F \subseteq \mathrm{supp}(h) \subseteq \overline{\mathrm{supp}}(h) \subseteq U' \cap U_r$,
- $h \in \tilde{\mathcal{V}}_\epsilon$, and
- $\tilde{g} := \hat{g}_1 h$ is measure-class preserving on $\mathrm{supp}(h)$.

Observe that

$$\overline{\mathrm{supp}}(\tilde{g}) \subseteq \mathrm{supp}(\hat{g}_1) \cup \mathrm{supp}(h) \subseteq \overline{\mathrm{supp}}(\hat{g}_1) \cup \overline{\mathrm{supp}}(h) \subseteq U'.$$

Clearly, $\tilde{g}$ agrees with $\hat{g}_1$ on $\mathrm{supp}(\tilde{g}) \setminus \mathrm{supp}(h)$. We also have

$$\mathrm{supp}(\tilde{g}) \setminus \mathrm{supp}(h) \subseteq \mathrm{supp}(\hat{g}_1) \setminus \mathrm{supp}(h) \subseteq \mathrm{supp}(\hat{g}_1) \setminus F. \tag{7}$$

But $\hat{g}_1$ agrees with g on the set $\mathrm{supp}(\hat{g}_1) \setminus F$. Since g is measure-class preserving everywhere, this implies that $\tilde{g}$ is also measure-class preserving on $X \setminus \mathrm{supp}(h)$. From Observation 2.2, we conclude that $\tilde{g} \in \mathfrak{G}$. The choice of ϵ also guaranties that $\tilde{g} \in \mathcal{V}_{2\epsilon} \subseteq \mathcal{V}_{\delta_0}$. Therefore, by (5) there are $g_i \in \mathfrak{G}^c_{U_i}$, $1 \leq i \leq r-1$ such that $\tilde{g} = g_1 \cdots g_{r-1}$.

Finally, write $g_r := h^{-1}\hat{g}_2$. Clearly g_2 is in $\mathfrak{G}$, since $\tilde{g}$ and g are, while also

$$\overline{\mathrm{supp}}(g_2) \subseteq \overline{\mathrm{supp}}(\hat{g}_2) \cup \overline{\mathrm{supp}}(h) \subseteq U_r. \qquad \square$$

The next result follows from an approximation argument like that of [10, Lemma 5].

Corollary 2.14. *Let $X = [0,1]$ and suppose that $U_1, \ldots U_r$ are open sets with $(0,1) = \bigcup_{j=1}^r U_j$. Then there exists some continuous function $\epsilon : X \to \mathbb{R}_{\geq 0}$ with $\epsilon^{-1}(0) = \{0,1\}$ such that*

$$\left\{ g \in \mathfrak{G} \mid \forall x \in (0,1)\ d(g(x), x) \leq \epsilon(x) \right\} \subseteq \mathfrak{G}^c_{U_1} \cdots \mathfrak{G}^c_{U_r}.$$

Uniform perfectness.

Lemma 2.15. *Assume that $D, D' \subseteq X$ are either μ-adapted balls or boundary half-balls in X satisfying $D \subseteq \mathring{D}'$. In case they are boundary half-balls, assume in the latter case that $\dim(X) \geq 2$. Then there is $h \in \mathfrak{G}_{D'}$ such that $\{h^k(D)\}_{k\in\mathbb{Z}}$ are disjoint.*

Proof. Assume without loss of generality that D and D' are embedded balls and pick some μ-adapted ball D_0 with $D \subseteq \mathring{D}_0 \subseteq D_0 \subseteq \mathring{D}'$ and some point $p_0 \in \mathring{D}' \setminus D_0$. Using Fact 2.1, one can easily see that D_0 is the first entry of a sequence of μ-adapted balls $\{D_k\}_{k\geq 0}$ such that, for all $k \geq 0$,

- $D_{l+1} \subset \mathring{D}_l$ for all $0 \leq l < k$,
- there is some constant $\lambda \in (0,1)$ such that $\mu(D_{k+1}) = \lambda\mu(D_k)$ for all $k \geq 0$, and
- D_k converges to $\{p_0\}$ in the Hausdorff topology as k goes to infinity.

Using Observation 2.3 it is easy to construct a sequence of elements $(h_k)_{k\geq 0} \subset \mathcal{H}(X)$ supported on D' such that

- h_k agrees with h_{k+1} on $D' \setminus \mathring{D}_k$, and
- h_k maps D_l homeomorphically onto D_{l+1} for $0 \leq l \leq k$.

In view of the statement of 2.3, we may also assume

$$\begin{aligned}
&\|h_k\|_{bac, D'\setminus D_k} \\
&= \big(\mu(D' \setminus D_1) - \mu(D' \setminus D_0)\big) + \sum_{l=0}^{k-1} \big(\mu(D_l) - \mu(D_{l+1}) - (\mu(D_{l+1}) - \mu(D_{l+2}))\big) \\
&\leq \mu(D_0 \setminus D_1) + \sum_{l=0}^{k-1} (1-\lambda)^2 \lambda^l \mu(D_0) \leq \mu(D_0 \setminus D_1) + (1-\lambda)\mu(D_0).
\end{aligned}$$

The sequence $(h_k)_{k\geq 0}$ is seen to converge in τ_{ac} to some $h_\infty \in \mathfrak{G}$ with the desired property. □

The following follows easily from a standard argument due to Anderson [1] and is likely known.

Corollary 2.16. *Let D, $D' \subseteq X$ be two embedded (boundary) balls such that $D \subseteq \mathring{D}'$. Then for any $g \in \mathfrak{G}_{\mathring{D}}$ there exists $h_1, h_2 \in \mathfrak{G}_{\mathring{D}'}$ such that $g = [h_1, h_2]$.*

Proof. Take h_2 as given by Lemma 2.15 and let

$$h_1 \in \mathfrak{G}_{\bigcup_{k\geq 0} h^k(\mathring{D})}$$

be the element that acts on $h^k(\mathring{D})$ as $g^{h^{-k}}$ for all $k \geq 0$ (using Observation 2.2). □

3. Some general preliminaries for Theorem B

Here we collect some material needed in the next two sections. Most of it applies in the general setting of Proposition A. We first recall the very basic probability theory needed in the proof, namely Chebyshev's inequality, martingales, and the behavior of their variance. The second subsection relates the behavior of $\mathcal{D}(g)$ to the measure distortion of the element $g \in \mathfrak{G}$. The third one introduces the technical notion of filling blocks and good covering sequences, needed in the following two subsections to precisely state and proof the existence of elements g such that $\mathcal{D}(g)$ deviates from $\mathcal{D}(\mathrm{Id})$ in a evenly spread manner, and we prove existence in the case of an OU pair.

Martingales and Chebyshev's inequality.

Definition 3.1. Let $(X, \mathcal{A}, \mu)$ be a probability space. We say that a sequence of functions $\psi_1, \dots \psi_n : X \to \mathbb{R}$ in $L^1(X, \mu)$ is a sequence of martingale increments if as random variables the expectation of ψ_k conditional on $(\psi_l)_{l<k}$ is 0. In other words, if $\int \psi_1 \, d\mu = 0$ and for every $1 \leq k \leq n-1$ and every finite tuple of real intervals $\underline{J} := (J_1, \dots J_k)$ if we let $A_{\underline{J}} := \{x \mid \psi_i(x) \in J_i \text{ for } 1 \leq i \leq k\}$ then we have $\int_{A_{\underline{J}}} \psi_{k+1} \, d\mu = 0$.[2]

For the following, see, for instance, [13], Example 10.2 and Theorem 10.4.

Fact 3.2. *For any sequence of martingale increments $\psi_1, \dots \psi_n$ we have*

$$\left\| \sum_{l=1}^{n} \psi_l \right\|_2^2 = \sum_{l=0}^{n} \| \psi_l \|_2^2.$$

[2]Equivalently, if the sequence $(\sum_{j=1}^{k} \psi_j)_{k=1}^{n}$ is a martingale that starts at zero.

Fact 3.3 (Chebyshev's inequality). *For any probability space* $(X, \mathcal{A}, \mu)$, *any* $f \in L^2(X, \mu)$ *and any* $\epsilon > 0$ *we have*

$$\mu\big(\{x \in X \mid |f(x)| > \epsilon\}\big) \leq \frac{\|f\|_2^2}{\epsilon^2}.$$

The following is a standard fact and an immediate corollary of Facts 3.3 and 3.2.

Corollary 3.4. *For any* $\epsilon, C > 0$ *there exists some* $N = N(C, \epsilon) \in \mathbb{N}$ *such that for any* $n \geq N$ *and any sequence of martingale increments* $\psi_0, \psi_1, \ldots$ *with* $\|\psi_l\|_2^2 \leq C$ *for all* $1 \leq l \leq n$ *we have*

$$\mu\Big(\Big\{x \,\Big|\, \Big|\frac{1}{n}\sum_{l=0}^{n} \psi_l(x)\Big| > \epsilon\Big\}\Big) < \epsilon.$$

Two observations on derivatives and integrals. Let X and μ be as in the assumptions of Proposition A and write $\mathfrak{G} := \mathcal{H}_{bac}(X, \mu)$. Without loss of generality, let μ be a probability measure. Given a real function f on X, we write $\phi_+ := \phi \mathbb{1}_{\{\phi \geq 0\}}$.

Lemma 3.5. *Let* $g \in \mathfrak{G}$ *and let* $A \in \mathcal{B}(X)$. *Then*

$$\|g\|_{ac,A} = 2\int_A (1 - \mathcal{D}(g))_+ \, d\mu + (\mu(g(A)) - \mu(A)) \leq \mu(A) + \mu(g(A)).$$

Proof. Note that

$$\mu(g(A)) - \mu(A) = \int_A (\mathcal{D}(g) - 1)\, d\mu = \int_A (\mathcal{D}(g) - 1)_+ \, d\mu - \int_A (1 - \mathcal{D}(g))_+ \, d\mu.$$

Since $f_+ + (-f)_+ = |f|$ for all $f \in L^1(X, \mu)$, it follows that

$$2\int_A (1 - \mathcal{D}(g))_+ \, d\mu + \mu(g(A)) - \mu(A) = \int_A |1 - \mathcal{D}(g)| \, d\mu = \|g\|_{ac,A}.$$

On the other hand, since $\mathcal{D}(g) > 0$ almost everywhere, we have $\int_A (1 - \mathcal{D}(g))_+ \, d\mu \leq \mu(A)$, from which the inequality in the statement follows. □

Corollary 3.6. *Suppose that* $A \in \mathcal{B}(X)$ *and* $g \in \mathfrak{G}$ *satisfy*

(1) $\|g\|_{ac,A} \geq \epsilon \mu(A)$ *for some* $\epsilon > 0$, *and*

(2) $\mu(g(A)) \leq (1 + \delta)\mu(A)$ *for some* $\delta > 0$.

Then

$$\mu(A_0) \geq \tfrac{1}{4}((\epsilon - \delta)\mu(A)),$$

where $A_0 := \big\{x \in A \mid \mathcal{D}(g)(x) \leq 1 - \tfrac{1}{4}(\epsilon - \delta)\big\}$.

Proof. By Lemma 3.5 we have

$$\int_A (1 - \mathcal{D}(g))_+ \, d\mu = \tfrac{1}{2}(\|g\|_{ac,A} - (\mu(g(A)) - \mu(A))) \geq \tfrac{1}{2}(\epsilon - \delta)\mu(A).$$

On the other hand, we have

$$\int_A (1-\mathcal{D}(g))_+ \, d\mu \leq \mu(A_0) + \tfrac{1}{4}(\epsilon-\delta)\mu(A\setminus A_0) \leq \mu(A_0) + \tfrac{1}{4}(\epsilon-\delta)\mu(A),$$

since $\mathcal{D}(g) > 0$ a.e. The result follows. □

Filling blocks and good covering sequences.

Definition 3.7. Given $\epsilon > 0$, we say that $A \subseteq X$ is an ϵ-filling block if $\mu(\mathrm{Fr}(A)) = 0$, $\mu(A) > 0$ and there exist $n > 1$ and a finite set $\{g_l\}_{l=1}^n \subseteq \mathfrak{G}$ such that $\mathcal{D}(g_l)$ is constant almost everywhere in some neighborhood of $\bar{A}$ and if we let $A_l := g_l(A)$ we have

- $\mu(A_l \cap A_{l'}) = 0$ for $1 \leq l < l' \leq n$, and
- $\mu\big(X \setminus \bigcup_{l=1}^n A_l\big) < \epsilon$.

The following result is due to Fathi [7], based on the work of Brown for the unmeasured case; see also [3, Chapter 17].

Fact 3.8. *For each OU pair (X, μ) there exists a continuous surjection*

$$f : I^n \to X$$

such that the following hold:

(a) *The restriction of f to $I^n \setminus \partial I^n$ is an embedding onto an open dense subset of X, so that in particular, $f(I^n \setminus \partial I^n) \subseteq X \setminus \partial X$ and thus $\partial X \subseteq f(\partial I^n)$.*

(b) *We have $f(I^n \setminus \partial I^n) \cap f(\partial I^n) = \varnothing$.*

(c) *The measure μ is the push-forward of Lebesgue measure on I^n by f.*

Here I^n is only a particular description of the standard n-dimensional ball as a topological manifold.

Definition 3.9. By a *covering sequence* (for (X, μ)) we mean a sequence

$$\mathscr{S} = (\mathcal{P}_m, \mathscr{L}_m, \mathscr{L}'_m)_{m\geq 0},$$

where $\mathcal{P}_m = \{P(m,l)\}_{l\in\mathscr{L}_m}$ is a collection of finitely many measurable sets and $\mathscr{L}'_m \subseteq \mathscr{L}_m$ is such that the following conditions are satisfied.

(1) $\bigcup_{l\in\mathscr{L}_m} P(m,l) = X$.

(2) For distinct $l, l' \in \mathscr{L}_m$ we have $P(m,l)\cap P(m,l') \subseteq \mathrm{Fr}(P(m,l))\cap\mathrm{Fr}(P(m,l'))$.

(3) $\mu(\mathrm{Fr}(P(m,l))) = 0$ and $\mu(P(m,l)) > 0$ for all $m \geq 0$ and $l \in \mathscr{L}_m$.

(4) For any $m \geq 0$ and $l \in \mathscr{L}_{m+1}$ there is $l' \in \mathscr{L}_m$ with $P(m+1,l) \subseteq P(m,l')$.

(5) $\sup_{l\in\mathscr{L}_m} \mathrm{diam}(P(m,l))$ converges to 0 as m tends to infinity.

(6) For $m' > m \geq 0$ and every $l \in \mathscr{L}_m$ we have

$$\tfrac{1}{2}\mu(P(m,l)) \leq \mu\Big(P(m,l) \cap \bigcup_{l\in\mathscr{L}'_{m'}} P(m',l')\Big).$$

If, in addition, for some ϵ we have

(7) $P(m, l)$ is an ϵ-filling block for all $m \geq 0$, $l \in \mathscr{L}'_m$,

then we say $\mathcal{P}$ is *ϵ-filling*.

We say that (X, μ) is *well-covered* if it admits an ϵ-filling good covering sequence for all $\epsilon > 0$.

Corollary 3.10. *If (X, μ) is an OU pair, then any μ-adapted ball $D \subseteq X$ is an ϵ-filling block for all $\epsilon > 0$ and (X, μ) is well-covered.*

Proof. Fix a continuous surjection $f : I^n \to X$ as given by Fact 3.8. By Corollary 2.4, the property of being an ϵ-filling blocks holds for some ball D if and only if it holds for any μ-adapted ball of the same measure. We can thus assume that D is disjoint from $f(\partial I^n)$. The preimage $f^{-1}(D)$, which is disjoint from ∂I^n, is easily shown to be an ϵ-filling block in I^n for every $\epsilon > 0$ with respect to the Lebesgue measure, as witnessed by elements of $\mathcal{H}_{bac}(X, \mu)$ which fix the boundary of I^n pointwise. The properties of f ensure that those elements push forward by f to homeomorphisms of X which in turn witness the fact that D is an ϵ-filling block in (X, μ) for all ϵ.

To construct a covering sequence $\mathscr{S}$, for each $m \geq 0$ it suffices to push onto X the cover of I^n by the collection of all cubes of the form $\left[0, \frac{1}{2^m}\right]^m + \bar{\alpha}$, where $\bar{\alpha}$ is an arbitrary n-tuple with values in $\frac{1}{2^m}(\mathbb{N} \cap [0, 2^m - 1])$. The subset of indices $\mathscr{L}'_m$ will correspond to the images of the small cubes that do not touch ∂I^n, which are all ϵ-filling for any ϵ, by the previous paragraph. □

4. Density of elements with well-distributed norm around the identity

Throughout this section and the next, we assume that (X, μ) satisfies the hypotheses of Proposition A, with $\mu(X) = 1$, and d is a compatible metric on X.

The main goal of this section is to prove Corollary 4.7, whose content can be rephrased as saying that for any group topology τ on $\mathfrak{G}$ finer than τ_{co} but strictly coarser than τ_{ac} there is some constant ϵ such that for δ-small enough and any finite collection $\{A_i\}_{i=1}^r$ of δ-filling blocks with pairwise μ-null intersection the collection of $g \in \mathfrak{G}$ such that $\mathcal{D}(g)$ deviates at least ϵ_0 on average from 1 on each A_i (i.e., $\|g\|_{ac,A_i} > \mu(A_i)$ for all i) accumulates at the identity with respect to τ.

For a group topology $\tau \supseteq \tau_{co}$, being strictly coarser than τ_{ac} is equivalent to the existence of $\epsilon_0 > 0$ such that the collection of elements g with $\|g\|_{ac} > \epsilon_0$ accumulates at the identity and the proof of Corollary 4.7 involves the following two ingredients.

• Lemma 4.1, which shows, using an argument involving the continuity of conjugation of a flavor often seen in minimality proofs, that in the situation above for some

fixed $\epsilon, \epsilon' > 0$ and every ϵ-filling block $A \in \mathcal{B}(X)$ the collection of $g \in \mathfrak{G}$ such that $\|g\|_{ac,A_i} > \epsilon'\mu(A)$ accumulates on the identity with respect to τ.

• Continuity of multiplication with respect to τ now easily implies that given $\{A_i\}_{i=1}^r$ as in the statement of Corollary 4.7 we can find in every neighborhood of identity products of the form $g_1 \circ \cdots \circ g_r$, where $\|g_l\|_{ac,A_l} > \epsilon'\mu(A_l)$ for each $1 \le l \le r$. To prevent composition from mostly canceling out the good properties of the individual g_l rather than aggregating them we use Lemma 4.4, which states that the composition $g_0 \circ g_1$ of two elements from $\mathfrak{G}$ preserves enough of the maximum of $\|g_0\|_{ac,A}$ and $\|g_1\|_{ac,A}$ on any Borel set A, provided g_1 is chosen sufficiently close to the identity in τ_{co} in terms of τ_{co}.

We start with some notation and facts that will be useful in the proof of Lemma 4.4, and also later on.

Lemma 4.1. *Suppose that τ is a group topology on $\mathfrak{G}$ finer than τ_{co} and such that there is $\epsilon_0 > 0$ with the property that for all $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ there is $g \in \mathcal{V}$ with $\|g\|_{ac} > \epsilon_0$.*

Then for every $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ any $\frac{1}{4}\epsilon_0$-filling block A there is some $g \in \mathcal{V}$ such that $\|g\|_{ac,A} > \frac{1}{4}\epsilon_0\mu(A)$.

Proof. Assume for the sake of contradiction the existence of some $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ and an $\frac{1}{4}\epsilon_0$-filling block A such that for all $g \in \mathcal{V}$ we have $\|g\|_{ac,A} \le \frac{1}{4}\epsilon_0\mu(A)$.

According to Definition 3.7 there exists $n > 1$ and a finite set $\{g_l\}_{l=1}^n \subseteq \mathfrak{G}$ such that $\mathcal{D}(g_l)$ is constant on some neighborhood U_l of A and, if we let $A_l := g_l(A)$, we have

- $\mu(A_l \cap A_{l'}) = 0$ for $1 \le l < l' \le n$ and
- $\mu\big(X \setminus \bigcup_{l=1}^n A_l\big) < \epsilon_0$.

If we let $\delta_l > 0$ be small enough that $\mathrm{N}_{\delta_l}(A_l) \subseteq g_l(U)$, then for any element $g \in \mathcal{V}_{\delta_l}$ a simple application of the chain rules implies that for almost all $x \in A$ we have $\mathcal{D}(g^{g_l})(x) = \mathcal{D}(g)(g_l(x))$, from which the equality $\mu(A_l)\|g\|_{ac,A} = \mu(A)\|g\|_{ac,A_l}$ follows.

Let $\delta^* = \min_{1\le l\le n} \delta_l$ and consider $\mathcal{V}' := \mathcal{V}_{\delta^*} \cap \bigcap_{l=1}^n \mathcal{V}^{g_l^{-1}} \in \mathcal{N}_\tau(\mathrm{Id})$ (where we use $\tau_{co} \subseteq \tau$). By assumption, we can find $g \in \mathcal{V}'$ with the property that $\|g\|_{ac} > \epsilon_0$. In particular $\|g^{g_l}\|_{ac,A} \le \frac{1}{4}\epsilon_0\mu(A)$ for all $1 \le l \le n$. By the discussion just above, $\|g\|_{ac,A_l} \le \frac{1}{4}\epsilon_0\mu(A_l)$ for all $1 \le l \le n$, and if we write $C := \bigcup_{l=1}^n A_l$, then clearly also

$$\|g\|_{ac,C} = \frac{\mu(C)}{\mu(A)}\|g\|_{ac,A} \le \tfrac{1}{4}\epsilon_0\mu(C) \le \tfrac{1}{4}\epsilon_0.$$

Therefore

$$\mu(g(X \setminus C)) - \mu(X \setminus C) = \mu(C) - \mu(g(C)) \le \tfrac{1}{4}\epsilon_0,$$

and thus $\|g\|_{ac,X\setminus C} \leq 2\mu(X \setminus C) + \frac{1}{4}\epsilon_0 \leq \frac{3}{4}\epsilon_0$ by Lemma 3.5. This leads to a contradiction:

$$\|g\|_{ac} = \|g\|_{ac,C} + \|g\|_{ac,X\setminus C} \leq \tfrac{1}{4}\epsilon_0 + \tfrac{3}{4}\epsilon_0 = \epsilon_0. \qquad \square$$

Definition 4.2. Given $A \in \mathcal{B}(X)$ with $\mu(\mathrm{Fr}(A)) = 0$, $\mu(A) > 0$ and $\epsilon > 0$ we write

$$\mathfrak{p}_A(\epsilon) = \sup\{\delta > 0 \,|\, \mu(g(A)\triangle A) < \epsilon\mu(A)\},$$

so that, in particular, $|\mu(A) - \mu(g(A))| < \epsilon$ if $g \in \mathcal{V}_\delta$ for $\delta < \mathfrak{p}_A(\epsilon)$.

For example, if $X = I$, μ the Lebesgue measure, d the standard distance, and $A \subseteq \mathring{I}$ an interval, then $\mathfrak{p}_A(\epsilon) = \epsilon/(2\mu(A))$ for ϵ small enough.

Observation 4.3. *Given A as in 4.2, we have $\mathfrak{p}_A(\epsilon) > 0$ for every $\epsilon > 0$.*

Proof. It suffices to show that for any sequence $(g_n)_n \in \mathcal{H}(X)$ converging to Id in τ_{co} the measure of $\mu(g_n(A) \setminus A)$ and that of $\mu(A \setminus g_n(A))$ both converge to 0. As in the proof of Lemma 1.11, we must have $\mathring{A} \subseteq \bigcup_{n\in\mathbb{N}} \bigcap_{m\geq n} g_m(A)$ and thus $\bigcap_{n\in\mathbb{N}} \bigcup_{m\geq n}(A \setminus g_m(A)) \subseteq \mathrm{Fr}(A)$, which implies that $\lim_n A \setminus \mu(g_n(A)) = 0$. Arguing with $X \setminus A$ instead of A one can likewise deduce $\lim_n \mu(g_n(A) \setminus A) = 0$. $\square$

We call to the reader's attention that the proof of the following lemma contains a nested chain of sublemmas, with 4.6 subordinate to 4.5, in turn subordinate to 4.4. We use the width of the end-of-proof symbol to indicate their degree of nesting.

Lemma 4.4. *Given any $g_0 \in \mathfrak{G}$, any compact $A \subseteq X$ and any $\eta > 0$ there is some $\delta(g_0, \eta) > 0$ such that for any $g_1 \in \mathcal{V}_\delta$ we have*

$$\|g_0 \circ g_1\|_{ac,A} \geq \max\{\|g_0\|_{ac,A}, \tfrac{1}{8}\|g_1\|_{ac,A}\} - \eta.$$

Proof. We start by choosing $\phi \in C^0(X, \mathbb{R}_{>0})$ such that $\|\mathcal{D}(g_0) - \phi\| \leq \frac{1}{4}\eta$ and $\lambda > 0$ such that $\omega_\phi(2\lambda) < \frac{1}{8}\eta$, where ω_ϕ is the function introduced in Definition 1.5. Consider now some partition $A = \coprod_{i=1}^n A_i$ into finitely many Borel sets of diameter less than λ and with a μ-null frontier. For $1 \leq i \leq n$ choose an arbitrary $x_i \in A_i$. Let $\mathcal{I}_+$ be the set of indices i such that $\phi(x_i) > 1$ and $\mathcal{I}_- := \{1, \ldots n\} \setminus \mathcal{I}_+$. We will now check that any δ with

$$0 < \delta < \min\Big\{\lambda, \mathfrak{p}_A(1), \mathfrak{p}_{A_i}(2), \mathfrak{p}_{A_i}\Big(\frac{\eta}{4n\phi(x_i)\mu(A_i)}\Big)\Big\}_{i=1}^n \tag{8}$$

satisfies the conditions of the statement. Fix some arbitrary $g_1 \in \mathcal{V}_\delta$ and let $\psi := \sum_{i=1}^n \phi(x_i)\mathbb{1}_{A_i}$.

Lemma 4.5. $\|\psi\mathcal{D}(g_1) - 1\|_{1,A} \geq \max\big\{\|\psi - 1\|_{1,A}, \tfrac{1}{8}\|g_1\|_{ac,A}\big\} - \frac{1}{4}\eta.$

Proof. We will show that for any $i \in \{1, 2, \ldots n\}$ we have

$$\int_{A_i} |\phi(x_i)\mathcal{D}(g_1) - 1|\, d\mu \geq \max\big\{|\phi(x_i) - 1|\mu(A_i), \tfrac{1}{8}\|g_1\|_{ac,A_i}\big\} - \frac{\eta}{4n}. \tag{9}$$

Summing this over $1 \le i \le n$ yields the lemma. To prove (9), first take $i \in \mathcal{I}^+$. Since $g_1 \in \mathcal{V}_\delta$ and $\delta < \eta/\big(4n\phi(x_i)\mu(A_i)\big)$ we get

$$\begin{aligned}(10)\qquad \int_{A_i} |\phi(x_i)\mathcal{D}(g_1) - 1|\, d\mu &\ge \int_{A_i} (\phi(x_i)\mathcal{D}(g_1) - 1)\, d\mu \\ &= \phi(x_i) \int_{A_i} \mathcal{D}(g_1)\, d\mu - \mu(A_i) \\ &= \phi(x_i)\mu(g_1(A_i)) - \mu(A_i) \\ &\ge \phi(x_i)\Big(\mu(A_i) - \frac{\eta}{4n\phi(x_i)}\Big) - \mu(A_i) \\ &= |\phi(x_i) - 1|\mu(A_i) - \frac{\eta}{4n}.\end{aligned}$$

For $i \in \mathcal{I}_-$ we can argue similarly:

$$\begin{aligned}(11)\qquad \int_{A_i} |\phi(x_i)\mathcal{D}(g_1) - 1|\, d\mu &\ge \int_{A_i} (1 - \phi(x_i)\mathcal{D}(g_1))\, d\mu \\ &= \mu(A_i) - \phi(x_i)\mu(g_1(A_i)) \\ &\ge \mu(A_i) - \phi(x_i)\Big(\mu(A_i) + \frac{\eta}{4n\phi(x_i)}\Big) \\ &= (1 - \phi(x_i))\mu(A_i) - \frac{\eta}{4n} \\ &= |\phi(x_i) - 1|\mu(A_i) - \frac{\eta}{4n}.\end{aligned}$$

This establishes the first half of (9) for an arbitrary value $i \in \{1, \dots n\}$. It remains to show that

$$\int_{A_i} |\phi(x_i)\mathcal{D}(g_1) - 1|\, d\mu \ge \tfrac{1}{8}\|g_1\|_{ac,A_i} - \frac{\eta}{4n}.$$

and, clearly, we may assume

$$(12)\qquad |\phi(x_i) - 1|\,\mu(A_i) < \tfrac{1}{8}\|g_1\|_{ac,A_i},$$

so that then

$$|\phi(x_i) - 1| < \frac{1}{8\mu(A_i)}\|g_1\|_{ac,A_i} \le \frac{1}{8\mu(A_i)}\big(\mu(A_i) + \mu(g_1(A_i))\big) \le \frac{4\mu(A_i)}{8\mu(A_i)} \le \tfrac{1}{2},$$

where we have used Lemma 3.5 in the second inequality and the fact that $g_1 \in \mathcal{V}_\delta$ and $\delta < \mathfrak{p}_{A_i}(2)$ in the last one. It follows that $\phi(x_i) \ge \frac{1}{2}$.

We will also need the following:

Lemma 4.6. *Let $A \in \mathcal{B}(X)$ and $\theta \in L^1(X, \mu)$. Then for all $\beta > 0$ we have*

$$\|\beta\theta - 1\|_{1,A} \ge \beta\|\theta - 1\|_{1,A} - |\beta - 1|\,\mu(A).$$

Proof. Let $A_+ := \{x \in A \mid \theta(x) \geq 1\}$ and $A_- := \{x \in A \mid \theta(x) < 1\}$. Then

$$\begin{aligned}\int_{A_-} |\beta\theta - 1|\, d\mu &\geq \int_{A_-} \beta(1-\theta)\, d\mu + \int_{A_-} (1-\beta)\, d\mu \\ &= \beta\|\theta - 1\|_{1,A_-}\, d\mu + (1-\beta)\mu(A_-).\end{aligned}$$

Arguing analogously, one obtains

$$\begin{aligned}\int_{A_+} |\beta\theta - 1|\, d\mu &\geq \int_{A_+} \beta(\theta-1)\, d\mu + \int_{A_+} (\beta-1)\, d\mu \\ &= \beta\|\theta - 1\|_{ac,A_+} + (\beta-1)\mu(A_+).\end{aligned}$$

Adding the inequalities in the last two displays yields the desired result, since

$$(1-\beta)(\mu(A_-) - \mu(A_+)) \leq |1-\beta|\mu(A). \qquad \square$$

To conclude the proof of Lemma 4.5 we write

$$\begin{aligned}\|\phi(x_i)\mathcal{D}(g_1) - 1\|_{ac,A_i} &\geq \phi(x_i)\|g_1\|_{ac,A_i} - |\phi(x_i) - 1|\,\mu(A_i) \\ &\geq \tfrac{1}{2}\|g_1\|_{ac,A_i} - \tfrac{1}{8}\|g_1\|_{ac,A_i} \geq \tfrac{1}{8}\|g_1\|_{ac,A_i},\end{aligned}$$

where we first used Lemma 4.6 with $\beta = \phi(x_i)$, then applied the inequalities $\phi(x_i) \geq \frac{1}{2}$ and (12). $\square$

To conclude the proof of Lemma 4.4 we apply Observation 1.6. We use the fact that $\|\mathcal{D}(g_0) - \phi\|_{1,A} \leq \frac{1}{4}\eta$, as well as $\delta < \lambda$, $g_1 \in \mathcal{V}_\delta$, and $\omega_\phi(2\lambda) < \frac{1}{8}\eta$. We obtain

$$\|g_0 \circ g_1\|_{ac,A} \geq \|\psi\mathcal{D}(g_1) - 1\|_{1,A} - \tfrac{1}{4}\eta - \tfrac{1}{8}\mu(g_1(A))\eta \geq \|\psi\mathcal{D}(g_1) - 1\|_{1,A} - \tfrac{3}{8}\eta. \tag{13}$$

But by the choice of λ and ϕ we have

$$\|\psi - 1\|_{1,A} \geq \|\phi - 1\|_{1,A} - \tfrac{1}{8}\eta \geq \|g_0\|_{ac,A} - \tfrac{3}{8}\eta,$$

which together with Lemma 4.5 yields

$$\begin{aligned}\|\psi\mathcal{D}(g_1)\|_{1,A} - \tfrac{3}{8}\eta &\geq \max\big\{\|\psi - 1\|_{1,A},\, \tfrac{1}{8}\|\mathcal{D}(g_1)\|_{ac,A}\big\} - \tfrac{5}{8}\eta \\ &\geq \max\big\{\|g_0\|_{ac,A},\, \tfrac{1}{8}\|\mathcal{D}(g_1)\|_{ac,A}\big\} - \eta.\end{aligned}$$

Combining this with (13), the conclusion of Lemma 4.4 follows. $\square$

Corollary 4.7. *For every $\epsilon > 0$ there are $\epsilon', \epsilon'' > 0$ such that the following happens. Let τ be a group topology finer than τ_{co} such that for all $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ there is $g \in \mathcal{V}$ with $\|g\|_{ac} > \epsilon$. Let $A_1, \ldots A_n$ be a finite collection of ϵ'-filling blocks in X whose pairwise intersection is μ-null. Then for any $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ there is $g \in \mathcal{V}$ such that $\|g\|_{ac,A_i} \geq \epsilon''\mu(A_i)$ for all $1 \leq i \leq n$.*

Proof. We show that we can take $\epsilon' := \frac{1}{4}\epsilon$, and $\epsilon'' := \frac{1}{64}\epsilon$. Pick $\mathbb{U} \in \mathcal{N}_\tau(\mathrm{Id})$ such that $\mathbb{U}^n \subseteq \mathcal{V}$. We choose $g_1, \dots g_n \in \mathfrak{G}$ iteratively in the following manner, where we write $\bar{g}_0 = \mathrm{Id}$ and $\bar{g}_i = g_1 g_2 \dots g_i$ for $1 \leq i \leq n$. Having found $g_1, \dots g_{i-1}$, we choose g_i as follows:

(1) Let $\mathcal{V}_1$ stand for $\mathfrak{G}$ if $i = 1$, and otherwise for $\mathcal{V}_\delta$, where $\delta = \delta\big(\bar{g}_{i-1}, \epsilon_0/2^{n+7}\big)$ is given by Lemma 4.4.

(2) Using the fact that $\tau_{co} \subseteq \tau$ and Lemma 4.1 one can pick $g_i \in \mathbb{U} \cap \mathcal{V}_1$ with $\|g_i\|_{ac,A_i} \geq \frac{1}{4}\epsilon\mu(A_i)$.

An iterated application of Lemma 4.4 easily shows that $g := \bar{g}_n \in \mathcal{V}$ satisfies $\|g\|_{ac,A_i} \geq \frac{1}{64}\epsilon\mu(A_i)$ for all $1 \leq i \leq n$. □

5. Covering sequences and the density of elements with very large norm

The goal of this section is to establish Proposition 5.3, which, given a sequence $(g_i)_{i\geq 1} \subseteq \mathfrak{G}$ that meets certain conditions, allows us to approximate the logarithm of $\mathcal{D}(g_1 g_2 \cdots g_k)$ by $\theta_k - \epsilon' k$, where $\epsilon' > 0$ and $(\theta_i)_{i>1}$ is a sequence in $L^1(X, \mu)$ that as a sequence of random variables is a martingale process. This result is quite technical, and its proof involves some structural subtleties, which we outline at the outset. Its goal is to show Corollary 5.9, which states that under certain conditions on a group topology τ on $\mathfrak{G}$ every neighborhood of the identity must contain elements that contract some Borel sets of measure arbitrarily close to $\mu(X) = 1$ to a set of measure arbitrarily close to 0.

The following trivial upper bound Λ on the logarithm will be useful. Suppose given an approximation $\phi > 0$ of $\mathcal{D}(g_i)$ for one of the g_i. At some point, we will need to bound $\ln(\phi)$ from above by the sum of a negative constant and some function of bounded variation. On any point $x \in X$ at which $\phi(x) \leq 1$, we have $\ln(\phi(x)) \leq \phi(x) - 1$ and the negative difference $\ln(\phi(x)) - \phi(x) + 1$ can be offloaded on either the constant or the inequality. Keeping track of how this "gain" aggregates over different sets of points will be a key ingredient in the proof of Proposition 5.3.

Definition 5.1. Let $\Lambda : (0, \infty) \to \mathbb{R}$ be the continuous function defined by

$$\Lambda(t) := \begin{cases} \ln(t) & \text{if } t \geq 1, \\ t - 1 & \text{if } t < 1. \end{cases}$$

Observation 5.2. *Let $\phi : X \to \mathbb{R}_{>0}$ and $\rho : X \to [-1, 1]$ be measurable functions, $A \in \mathcal{B}(X)$. Then*

$$\|\Lambda(\phi) + \rho\|_{2,A}^2 \leq 2\|\phi\|_{1,A} + 6\mu(A).$$

Proof. Consider an arbitrary $\gamma \in [-1, 1]$. For $y \geq 1$ and we have

$$(\Lambda(y) + \gamma)^2 = (\ln(y) + \gamma)^2 \leq (\ln(y) + 1)^2 \leq 2\big(1 + \ln(y) + \tfrac{1}{2}\ln(y)^2\big) \leq 2e^{\ln(y)} = 2y,$$

while for any $0 \le y \le 1$ we have

$$(\Lambda(y)+\gamma)^2 = (y+\gamma-1)^2 \le 4.$$

One can use this to conclude:

$$\begin{aligned}\|\Lambda(\phi)+\rho\|_{2,A}^2 &\le 2\|(\phi-1)_+ + 1\|_{1,A} + 4\mu(\{x\in X \mid \phi(x)\le 1\})\\ &\le 2\|\phi\|_1 + 4\mu(A).\end{aligned}$$

□

Proposition 5.3. *For $\epsilon > 0$ small enough the following holds. Let $(\mathcal{P}_m, \mathscr{L}_m, \mathscr{L}'_m)_{m\ge 0}$ be a good covering sequence as defined in Definition 3.9. For any $\eta > 0$, $k \in \mathbb{N}^+$ there are functions*

$$\boldsymbol{\delta}_{\eta,k} : \mathfrak{G}^{k-1} \to \mathbb{R}_{>0} \quad \textit{and} \quad \boldsymbol{m}_{\eta,k} : \mathfrak{G}^k \to \mathbb{N}$$

with the properties described below.[3]

For any sequence $g_1, \dots g_n$ of elements of $\mathfrak{G}$, writing $m_0 = 0$ and

$$\delta_k := \boldsymbol{\delta}_{\eta,k-1}(g_1,\dots,g_{k-1}), \quad m_k := \boldsymbol{m}_{\eta,k}(g_1,\dots g_k)$$

for $k \ge 1$, if

(I) $g_k \in \mathcal{V}_{\delta_k}$ *and*

(II) $\|g_k\|_{ac,P(m_{k-1},l)} > \epsilon\mu(P(m_{k-1},l))$ *for all $l \in \mathscr{L}'_{m_k}$,*

then there exist

- *a sequence $(\psi_k)_{k=1}^n$ of martingale increments on (X,μ) such that $\|\psi_k\|_2^2 \le 8$,*
- $\epsilon' := \frac{1}{64}\epsilon\big(-\frac{1}{32}\epsilon - \ln\big(1-\frac{1}{32}\epsilon\big)\big) > 0$, *and*
- $\Theta \in L^1(X,\mu)$ *with* $\|\Theta\|_1 \le \eta$,

such that the inequality

$$\mathcal{D}(g_1 \circ g_2 \cdots g_n) \le \Theta + e^{-n\epsilon' + \sum_{k=1}^n \psi_l} \tag{14}$$

holds almost everywhere.

Overview of the proof. We work with a fixed $\eta > 0$. The construction is by induction on k. We assume that $\boldsymbol{\delta}_{j,\eta}$ and $\boldsymbol{m}_{j,\eta}$ are already defined for $1 \le j < k$ and we are given a tuple $\bar{g}_{k-1} := (g_1, \dots g_{k-1}) \in \mathfrak{G}^{k-1}$, which satisfies the conditions of the statement of Proposition 5.3. We also assume that the construction has provided us with step functions $\psi_1, \dots \psi_{k-1}$, together with some additional auxiliary data, to be described in detail below.

We begin the k-th step of the construction using the previous information to find $\delta_k := \boldsymbol{\delta}_{\eta,k}(\bar{g}_{k-1})$, after which, given $g_k \in \mathfrak{G}$ satisfying (I) and (II), we describe how to find the aforementioned auxiliary data, consisting of

- a continuous $\theta_k : X \to \mathbb{R}_{>0}$ which approximates $\mathcal{D}(g)$ sufficiently well in L^1;

[3] In particular, $\boldsymbol{m}_{1,\eta}$, $\boldsymbol{\delta}_{1,\eta}$ are constants dependent on η.

- a positive constant $\lambda_k > 0$ depending on the modulus of continuity of θ_k;
- an integer $m_k := \boldsymbol{m}_{\eta,k}(\bar{g}_{k-1}, g_k) > m_{k-1}$ large enough to ensure that the sets in $\mathcal{P}_{m_k}$ have uniformly small diameter in relation to λ_k; and
- a step function $\phi_k : X \to \mathbb{R}_{>0}$ defined in terms θ_k and the covering $\mathcal{P}_{m_k}$, which is meant to approximate $\mathcal{D}(g_k)$ still reasonably well by the choice of λ_k and m_k.

We next need to find the next martingale increment ψ_k in the sequence. In Lemma 5.4 we describe how to find a step function ψ_k that bounds $\ln(\phi_k)$ in a uniformly strict manner, integrates to zero on each of the sets in the previous partition $\mathcal{P}(m_{k-1})$ (i.e., its expectation conditional to $P(m_{k-1}, l)$ is null) and has a uniformly bounded L^2-norm (variance). The proof is technical and contains the two nested sublemmas 5.5 and 5.6.

To conclude, we need to show that for any tuple $(g_1, \dots g_n) \in \mathfrak{G}^n$ as in the hypothesis, with a resulting sequence of step functions $\psi_1, \dots \psi_n$, the control of ψ_j over ϕ_j for all $1 \leq j \leq n$, as well $\mathcal{D}(g_j)$ being well-approximated by ϕ_j, translates into the existence of a suitable Θ such that inequality (14) is satisfied, which comes down to combining together all the information coming from the previous steps in a telescopic sum (which explains the powers of two in some of the definitions). □

Proof of Proposition 5.3. Recall that $\eta > 0$, which is given to us, is fixed and most of the functions defined above are dependent on it (even if the dependence is hidden from the notation from this point on).

Notation. Recall that $m_0 = 0$.

- $\hat{\epsilon} := \big(-\frac{1}{32}\epsilon - \ln\big(1 - \frac{1}{32}\epsilon\big)\big)$ so that $\epsilon' = \frac{1}{64}\epsilon\hat{\epsilon}$,
- $\Delta_j := e^{-j\epsilon' + \sum_{l=1}^{j} \psi_l}$ (in particular, $\Delta_0 = 1$),
- $\mathfrak{p}_j := \min_{l \in \mathscr{L}_{m_j}} \mathfrak{p}_{P(m_j,l)}$ (see Definition 4.2),
- $\nu_{j,l} := \mu(P(m_j, l))$.

Choice of δ_k, m_k and the auxiliary data. We adopt the convention $\lambda_0 := 1$ and initially set

$$\delta_k := \tfrac{1}{2} \min\big\{ \lambda_{k-1}, \mathfrak{p}_{k-1}\big(\tfrac{1}{2}\epsilon'\big), \mathfrak{p}_{k-1}\big(\tfrac{1}{4}\epsilon\big), \delta_{k-1} \big\}. \tag{15}$$

Given $g_k \in \mathfrak{G}$ satisfying conditions (I) and (II) we find m_k and ψ_k through the following sequence of choices:

(i) $\theta_k \in C^0(X)$, $\theta_k > 0$ such that
  (a) $\|\theta_k - \mathcal{D}(g_k)\|_1 < \dfrac{\eta}{2^{k+2}\|\Delta_{k-1}\|_\infty}$ and
  (b) $\|\theta_k - \mathcal{D}(g_k)\|_{1,P(m_{k-1},l)} < \min\big\{\frac{\epsilon'}{2}, \big(\frac{\epsilon}{32}\big)^2\big\}\nu_{k-1,l}$ for $l \in \mathscr{L}_{m_{k-1}}$,

(ii) $\lambda_k \in (0, \delta_k)$ such that $\omega_{\theta_k}(2\lambda_k) < \dfrac{\eta}{2^{k+2}\|\Delta_{k-1}\|_\infty}$,

(iii) and finally $m_k > m_{k-1}$ such that $\sup_{l\in\mathscr{L}_{m_k}} \operatorname{diam}(P(m_k,l)) < \lambda_k$.

Define ϕ_k as the step function

$$\phi_k := \sum_{l\in\mathscr{L}_{m_k}} \Big(\inf_{x\in P(m_k,l)} \theta_k(x)\Big)\mathbb{1}_{P(m_k,l)}.$$

We start the proof of the proposition with the following sublemma.

Lemma 5.4. *There is some function ψ_k constant on $P(m_k,l)$ for every $l\in\mathscr{L}_{m_k}$ and satisfying*

(a) $\ln(\phi_k) \le -\epsilon' + \psi_k$ *almost everywhere*,

(b) $\int_{P(m_{k-1},l)} \psi_k \, d\mu = 0$ *for each* $l\in\mathscr{L}_{m_{k-1}}$, *and*

(c) $\|\psi_k\|_2^2 \le 8$.

Proof. Let $l\in\mathscr{L}_{m_{k-1}}$. By (15), the assumption $g_k\in\mathcal{V}_{\delta_k}$, and (I) we have

$$\mu(g_k(P(m_{k-1},l))) \le \big(1+\min\{\tfrac{1}{2}\epsilon', \tfrac{1}{4}\epsilon\}\big)\nu_{k-1,l}. \tag{16}$$

On the other hand, since at least μ-half of $P(m_{k-1},l)$ is covered by sets of the form $P(m_k,l')$ with $l'\in\mathscr{L}'_m$ (item (6) of 3.9), it follows from (II) in the choice of g_k that

$$\|g_k\|_{ac,P(m_{k-1},l)} \ge \tfrac{1}{2}\epsilon\nu_{k-1,l}. \tag{17}$$

Together with (16), this allows us to apply Corollary 3.6 with $\frac{1}{2}\epsilon$ in place of the ϵ in the statement of the Corollary and $\frac{1}{4}\epsilon$ in place of δ, and obtain the inequality

$$\mu\big(\{x\in P(m_{k-1},l)\mid \mathcal{D}(g_k)(x)\le 1-\tfrac{1}{16}\epsilon\}\big) \ge \tfrac{1}{16}\epsilon\nu_{k-1,l}.$$

From the bound $\|\mathcal{D}(g_k)-\theta_k\|_{1,P(m_{k-},l)} \le \frac{1}{2}\epsilon'\nu_{k-1,l}$ in (1b) it follows that

$$\mu\big(\{x\in P(m_{k-1},l)\;\big|\;|\mathcal{D}(g_k)(x)-\theta_k(x)|\ge\tfrac{1}{32}\epsilon\}\big) \le \tfrac{1}{32}\epsilon\nu_{k-1,l}. \tag{18}$$

Write $A_l := \{x\in P(m_{k-1},l)\mid \phi_k \le 1-\frac{1}{32}\epsilon\}$. This is a union of sets of the form $P(m_k,l')$, $u\in\mathscr{L}_{m_k}$, $P(m_k,l')\subseteq P(m_{k-1},l)$. We can use inequalities (17) and (18) to conclude that

$$\mu(A_l) \ge \mu\big(\{x\in P(m_{k-1},l)\mid \theta_k \le 1-\tfrac{1}{32}\epsilon\}\big) \ge \tfrac{1}{32}\epsilon\nu_{k-1,l}, \tag{19}$$

since clearly $\mathcal{D}(x)\le 1-\frac{1}{16}\epsilon$ and $|\mathcal{D}(g_k)(x)-\theta(x)|<\frac{1}{32}\epsilon$ imply $\theta(x)<\frac{1}{32}\epsilon$ for $x\in P(m_{k-1},l)$.

Set $K_l := \int_{P(m_{k-1},l)} \Lambda(\phi_k)\, d\mu$.

Lemma 5.5. *The inequalities $-\nu_{k-1,l} \le K_l \le \epsilon'\nu_{k-1,l}$ hold.*

Proof. We start by bounding from above a natural approximation of K_l using (16):

$$\int_{P(m_{k-1},l)} \Lambda(\mathcal{D}(g_k))\, d\mu \le \int_{P(m_{k-1},l)} (\mathcal{D}(g_k)-1)\, d\mu = \tfrac{1}{2}\epsilon'\nu_{k-1,l}. \tag{20}$$

Using the monotonicity of Λ for the first inequality below, the fact that Λ is 1-Lipschitz for the third one, and (20) together with item (1b) for the last one we obtain

$$
\begin{aligned}
(21) \qquad K_l &\le \int_{P(m_{k-1},l)} \Lambda(\theta_k)\, d\mu \\
&\le \int_{P(m_{k-1},l)} \Lambda(\mathcal{D}(g_k))\, d\mu + \|\Lambda(\mathcal{D}(g_k)) - \Lambda(\theta_k)\|_{1,P(m_{k-1},l)} \\
&\le \int_{P(m_{k-1},l)} \Lambda(\mathcal{D}(g_k))\, d\mu + \|\mathcal{D}(g_k) - \theta_k\|_{1,P(m_{k-1},l)} \\
&\le \tfrac{1}{2}\epsilon' \nu_{k-1,l} + \tfrac{1}{2}\epsilon' \nu_{k-1,l} = \epsilon' \nu_{k-1,l}.
\end{aligned}
$$

The bound $K_l \ge -\nu_{k-1,l}$ follows trivially since $\Lambda \circ \phi > -1$ everywhere. □

Next, write

$$
\gamma_l := \max\Big\{0, \frac{K_l + \epsilon' \nu_{k-1,l}}{\mu(A_l)}\Big\}, \quad \sigma_l := \frac{\gamma_l \mu(A_l) - K_l}{\nu_{k-1,l}}.
$$

Lemma 5.6. *For all $l \in \mathscr{L}_{m_{k-1}}$ we have*

$$
\sigma_l \ge \epsilon' \quad \textit{and} \quad \{\sigma_l, \sigma_l - \gamma_l\} \subseteq [-1, 1].
$$

Proof. Let $l \in \mathscr{L}_{m_{k-1}}$. Then either

- $\gamma_l = \frac{K_l + \epsilon' \nu_{k-1,l}}{\mu(A_l)} > 0$ and $\sigma_l = \epsilon'$, or
- $\gamma_l = 0$, the quantity $\frac{K_l + \epsilon' \nu_{k-1,l}}{\mu(A_l)}$ is negative, and $\sigma_l = -\frac{K_l}{\nu_{k-1,l}}$.

Using the upper bound in Lemma 5.5 in the first inequality below and inequality (19) in the second we get

$$
\gamma_l \le \frac{2\epsilon' \nu_{k-1,l}}{\mu(A_l)} \le \frac{64\epsilon'}{\epsilon} = \hat{\epsilon}.
$$

If $\gamma_l > 0$, then $\sigma_l = \epsilon'$ and the inclusion $\{\sigma_l, \sigma_l - \gamma_l\} \subseteq [-1, 1]$ is clear (for ϵ small enough). In the second case we have

$$
\epsilon' \le \sigma_l = \sigma_l - \gamma_l = -\frac{K_l}{\nu_{k-1,l}} \le 1,
$$

where the first inequality follows from $-K_l + \epsilon' \nu_{k-1,l} < 0$ and the second from Lemma 5.5. □

To conclude the proof of Lemma 5.4, define

$$
\psi_k := \Lambda(\phi_k) + \sum_{l \in \mathscr{L}_{m_{k-1}}} (\sigma_l \mathbb{1}_{P(m_{k-1},l)} - \gamma_l \mathbb{1}_{A_l}).
$$

We proceed to check that ψ_k satisfies the properties in the conclusion of Lemma 5.4.

(a) The inequality

$$\ln(\phi_k) \leq -\epsilon' + \psi_k$$

holds almost everywhere. Indeed, by definition, $\hat{\epsilon} = -\frac{\epsilon}{32} - \ln(1 - \frac{32}{\epsilon})$, that is $\hat{\epsilon} = \Gamma(1 - \frac{\epsilon}{32})$, where

$$\Gamma(y) = (y-1) - \ln(y).$$

It is not hard to see that Γ is strictly decreasing on $(0, 1]$. For all for all $x \in A_l$ we have $\phi(x) \leq \theta(x) \leq 1 - \frac{\epsilon}{32}$ and $\Lambda(\phi(x)) = x - 1$. Therefore for all $x \in A_l$ outside a set of measure zero we have

$$\ln(\phi_k(x)) \leq \Lambda(\phi(x)) - \hat{\epsilon} \leq \Lambda(\phi(x)) - \gamma_l = -\sigma_l + \psi(x) \leq -\epsilon' + \psi(x).$$

The inequality between first and last terms can be proven for almost all $x \in P(m_{k-1}, l) \setminus A_l$ in a similar way. Recall that $\epsilon' \leq \sigma_l$ is from Lemma 5.6.

(b) For all $l \in \mathscr{L}_{m_{k-1}}$ the choice of σ_l implies that

$$\int_{P(m_{k-1},l)} \psi_k \, d\mu = K_l + \sigma_l \nu_{k-1,l} - \gamma_l \mu(A_l) = 0.$$

(c) Finally, the bound $\|\psi_k\|_2 \leq 8$ holds, as a consequence of Observation 5.2, which applies with $\rho := \sum_{l \in \mathscr{L}_{m_{k-1}}, l} (\sigma_l \mathbb{1}_{P(m_{k-1},l)} - \gamma_l \mathbb{1}_{A_l})$, using the conclusion of Lemma 5.6 and the inequality $\|\phi_k\|_{1,P} \leq 2$. □

End of the proof of Proposition 5.3: L^1-approximate bound of the derivative. We are now in a position to finish the proof of Proposition 5.3. Assume that the process has run for n steps, producing elements $g_1, \ldots g_n \in \mathfrak{G}$, together with all the other data. Write $\hat{g}_k := g_{k+1} \circ \ldots g_n$ for all $0 \leq k \leq n$. Set

$$\Theta_k := \Delta_{k-1}(\mathcal{D}(g_k \circ \hat{g}_k) - \phi_k \mathcal{D}(\hat{g}_k)) \in L^1(X, \mu), \quad 1 \leq k \leq n,$$
$$\Xi_k := \Delta_k \mathcal{D}(\hat{g}_k), \qquad 0 \leq k \leq n.$$

For all $1 \leq k \leq n$ noting that $g_k \circ \hat{g}_k = \hat{g}_{k-1}$ we see that, almost everywhere,

$$\Xi_{k-1} - \Theta_k = \Delta_{k-1}\phi_k \mathcal{D}(\hat{g}_k) \leq \Delta_{k-1} e^{-\epsilon' + \psi_k} \mathcal{D}(\hat{g}_k) = \Delta_k \mathcal{D}(\hat{g}_k) = \Xi_k, \tag{22}$$

where the first equality follows from simple algebraic manipulations, the inequality follows from the conclusion of Lemma 5.4 and the final equality from the definition $\Delta_j := e^{-j\epsilon' + \sum_{l=1}^{j} \psi_l}$.

Adding inequality (22) for all values $1 \leq k \leq n$ we get the validity almost everywhere of

$$\mathcal{D}(g_1 \circ \ldots g_n) = \Xi_0 \leq \Theta + \Xi_n = \Theta + e^{-n\epsilon' + \sum_{k=1}^{n} \psi_l},$$

where $\Theta := \sum_{k=1}^{n} \Theta_k$. To conclude the proof of Proposition 5.3 it suffices to show that $\|\Theta\|_1 \leq \eta$, which boils down to the following:

Lemma 5.7. *The inequality* $\|\mathcal{D}(g_k \circ \hat{g}_k) - \phi_k \mathcal{D}(\hat{g}_k)\|_1 \leq \dfrac{\eta}{2^k \|\Delta_{k-1}\|_\infty}$ *holds for* $1 \leq k \leq n$.

Proof. We will verify that we may apply Observation 1.6 with $\eta/(2^{k+1}\|\Delta_{k-1}\|_\infty)$ in place of the values of ϵ and ϵ' in the statement of that observation. Recall that by construction $\sup_{l \in \mathscr{L}_{m_k}} \operatorname{diam}(P(m_k, l)) < \lambda_k$, and ϕ_k agrees with $\inf_{x \in P(m_k,l)} \theta_k(x)$ on each of those sets.

The L^1 bound in the assumptions of Observation 1.6 follows from (1a), combined from the L^∞ bound resulting from the choice of λ_k in (ii), as well as the properties of ϕ_k mentioned above:

$$\|\mathcal{D}(g_k) - \phi_k\|_1 \leq \|\mathcal{D}(g_k) - \theta_k\|_1 + \|\theta_k - \phi_k\|_\infty \leq 2\frac{\eta}{2^{k+2}\|\Delta_{k-1}\|_\infty} = \frac{\eta}{2^{k+1}\|\Delta_{k-1}\|_\infty}.$$

In order to verify the conditions pertaining to the ϵ' in Observation 1.6, observe first that for all $1 \leq s \leq n$ we have $\delta_{s+1} \leq \frac{1}{2}\lambda_s < \frac{1}{2}\delta_s$, as well as $g_s \in \mathcal{V}_{\delta_s}$. It follows that $\hat{g}_k \in \mathcal{V}_{2\delta_{k+1}} \subseteq \mathcal{V}_{\lambda_k}$. □

From Lemma 5.7 we get

$$\|\Theta_k\|_1 \leq \|\Delta_{k-1}\|_\infty \|(\mathcal{D}(g_k) \circ \hat{g}_k - \phi_k)\mathcal{D}(\hat{g}_k)\|_1 \leq \|\Delta_{k-1}\|_\infty \frac{2\eta}{2^{k+1}\|\Delta_{k-1}\|_\infty} = \frac{\eta}{2^k},$$

from which we easily get $\|\Theta\|_1 < \eta$, concluding the proof of Proposition 5.3. □

Lemma 5.8. *Assume that* (X, μ) *is well-covered and let* τ *be a group topology on* $\mathcal{H}_{bac}(X, \mu)$ *satisfying* $\tau_{co} \subsetneq \tau$ *and* $\tau_{ac} \not\subseteq \tau$. *Then for any* $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ *and any* $\delta > 0$ *there exists some* $g \in \mathcal{V}$ *such that* $\mu(\{x \in X \mid \mathcal{D}(g)(x) \geq \delta\}) < \delta$.

Proof. Pick $\epsilon_0 > 0$ such that for every $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ there is some $g \in \mathcal{V}$ with $\|g\|_{ac} > \epsilon_0$ and write $\epsilon = \frac{1}{64}\epsilon_0$ and $\epsilon' := \frac{1}{64}\epsilon\big(\frac{1}{32}\epsilon - \ln\big(1 - \frac{1}{32}\epsilon\big) - 1\big)$. Let also $(\mathcal{P}_m, \mathscr{L}_m, \mathscr{L}'_m)_{m \geq 0}$ be an $\frac{1}{4}\epsilon_0$-filling good covering sequence for (X, μ). Let

$$\nu := \min\Big\{\frac{2}{\epsilon'}, \frac{\delta}{2}\Big\}, \quad N_0 = \max\Big\{\Big\lceil \frac{2\ln(2/\delta)}{\epsilon'} \Big\rceil,\ N(8, \nu)\Big\},$$

where $N : \mathbb{R}^2_{>0} \to \mathbb{N}$ is the function given by Corollary 3.4. Pick $\mathbb{U} = \mathbb{U}^{-1} \in \mathcal{N}_\tau(\mathrm{Id})$ with $\mathbb{U}^{N_0} \subseteq \mathcal{V}$. Using Corollary 4.7 and the assumption that $\tau_{co} \subseteq \tau$ we can inductively choose $g_1, \ldots g_{N_0} \in \mathbb{U}$ such that for some $0 < \eta < \big(\frac{1}{2}\delta\big)^2$ and $\delta_k := \boldsymbol{\delta}_{\eta,k}(g_1, \ldots g_{k-1})$, $m_k = \boldsymbol{m}_{\eta,k}(g_1, \ldots g_k)$, where $\boldsymbol{\delta}_{\eta,k}, \boldsymbol{m}_{\eta,k}$ are given by Proposition 5.3, then the following conditions are satisfied:

- $g_k \in \mathcal{V}_{\delta_k}$,
- $\|g_k\|_{ac, P(m_k,l)} > \epsilon\mu(P(m_k, l))$ for $l \in \mathscr{L}'_{m_k}$.

If we write $g := g_1 \circ \cdots g_n$, then 5.3 yields $\Theta \in L^1(X.\mu)$ with $\|\Theta\|_1 < (\frac{1}{2}\delta)^2$ and a sequence of martingale increments $\psi_1, \ldots \psi_n$ such that $\|\psi_l\|_2^2 \leq 8$ and

$\mathcal{D}(g) \le \Theta + e^{-N_0\epsilon' + \sum_{l=1}^{N_0} \psi_l}$ almost everywhere. Using one of the bounds on N_0 in each of the last two inequalities below, we get

$$\begin{aligned}\mu\big(\{x \in X \mid (\mathcal{D}(g) - \Theta)(x) \ge \tfrac{1}{2}\delta\}\big) &\le \mu\big(\{x \in X \mid e^{-N_0\epsilon' + \sum_{l=1}^{N_0} \psi_l} \ge \tfrac{1}{2}\delta\}\big) \\ &\le \mu\Big(\Big\{x \in X \;\Big|\; -\epsilon' + \frac{1}{N_0}\sum_{l=1}^{N_0} \psi_l(x) \ge \frac{\ln(\delta/2)}{N_0}\Big\}\Big) \\ &\le \mu\Big(\Big\{x \in X \;\Big|\; \Big|\frac{1}{N_0}\sum_{l=1}^{N_0} \psi_l(x)\Big| \ge \tfrac{1}{2}\epsilon'\Big\}\Big) \le \tfrac{1}{2}\delta.\end{aligned}$$

On the other hand, $\|\Theta\|_1 < \big(\frac{1}{2}\delta\big)^2$ implies that $\mu\big(\{x \in X \mid \Theta(x) > \frac{1}{2}\delta\}\big) \le \frac{1}{2}\delta$ and thus

$$\begin{aligned}\mu\big(\{x \in X \mid \mathcal{D}(g) \ge \delta\}\big) &\le \mu\big(\{x \in X \mid (\mathcal{D}(g) - \Theta)(x) \ge \tfrac{1}{2}\delta\}\big) + \mu\big(\{x \in X \mid \Theta(x) \ge \tfrac{1}{2}\delta\}\big) \\ &\le \delta. \qquad \square\end{aligned}$$

Using the regularity of the measure μ together with Observation 1.1 one easily concludes:

Corollary 5.9. *Under the assumptions of Lemma 5.8 for any $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ and any $\delta > 0$ there is some open set $V \subseteq X$ and some $g \in \mathcal{V}$ such that $\mu(V) \ge 1 - \delta$ and $\mu(g(V)) \le \delta$.*

6. Separable topologies on $\mathcal{H}_{bac}(X, \mu)$ in the Oxtoby–Ulam case

Let X be a compact manifold and μ an OU measure on X, and set $\mathfrak{G} := \mathcal{H}_{bac}(X, \mu)$. Without loss of generality, we assume that μ is a probability measure. In this section we show Lemma 6.2 below, as well as a variation, Lemma 6.7, with a weaker conclusion in the case $X = [0, 1]$. Both lemmas can be interpreted as saying that any separable group topology on $\mathfrak{G}$ is reasonably close to being finer than τ_{ac}, in the sense that any neighborhood $\mathcal{V}$ of the identity in τ will contain any set of the form $\mathfrak{G}_U$ for an open set U satisfying certain conditions, the most important being that $\mu(U) < \kappa$ for some $\kappa > 0$ depending on $\mathcal{V}$. We do this by adapting, in a rather straightforward manner, arguments from the literature on automatic continuity of homeomorphism groups. Unfortunately, we fall short of proving automatic continuity for $(\mathfrak{G}, \tau_{ac})$, as the final steps in those proofs fail in this context and we do not know how to fill the gap.

Definition 6.1. Assume that $X \ne [0, 1]$. Given $\epsilon > 0$ by an ϵ-fine collection we intend a finite disjoint collection $\{D_l\}_{l \in \Lambda}$ of embedded balls and boundary half-balls such that

- $\mathrm{diam}(D_l) \le \epsilon$ for all $l \in \Lambda$, and
- $\mu\big(\bigcup_{l \in \Lambda} D_l\big) \le \epsilon$.

By an ϵ-fine open subset of X we intend one of the form $\bigcup_{l\in\Lambda} \mathring{D}_l$ for some ϵ-fine collection $\{D_l\}_{l\in\Lambda}$.

Lemma 6.2. *If $X \neq [0,1]$, then there is some number N such that for any separable group topology τ on $\mathfrak{G} := \mathcal{H}_{bac}(X,\mu)$ and every $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ there exists some $\epsilon > 0$ such that for any ϵ-fine open subset V of X we have $\mathfrak{G}_V \subseteq \mathcal{V}$.*

Proof. By continuity of multiplication, it suffices to show, after fixing ϵ, that for any arbitrary $\mathbb{U} = \mathbb{U}^{-1} \in \mathcal{N}_\tau(\mathrm{Id})$ the result is true with $\mathcal{V}$ replaced by $\mathbb{U}^{14}$ in the conclusion. Fix such $\mathbb{U}$. For the following see Lemma 3.2 in [14], which uses only the Baire property.

Lemma 6.3. *The closure of $\mathbb{U}$ contains the standard neighborhood of the identity $\mathcal{W}_\delta$ of τ_{ac} (Definition 1.3) for some $\delta > 0$.*

Take $\epsilon := \frac{1}{4}\delta$ for the above δ and fix an ϵ-fine collection $\{D_l\}_{l\in\Lambda}$. The following follows from Observations 2.2 by the same argument as in Lemma 3.8 in [14].

Lemma 6.4. *Assume that for all $l \in \Lambda$ we are given a countable disjoint collection $\{D_{l,k}\}_{k\geq 0}$ of sets of the same type as D_l. Then there is some choice function $\rho : \Lambda \to \mathbb{N}$ such that for all $l \in \Lambda$ and any choice of elements $g_l \in \mathfrak{G}_{D_{l,\rho(l)}}$ there is some $g \in \mathfrak{G}_{\bigcup_{l\in\Lambda,k\geq 0} D_{l,k}} \cap \mathbb{U}^2$ such that $g_{\restriction D_{l,\rho(l)}}$ for all l.*

Using Corollary 2.16 and Lemma 2.15 together with the argument in the proof of Lemma 3.7 in [14] and the inclusion $\tau_{co} \subseteq \tau_{ac}$ we obtain the following refinement of Lemma 6.3:

Lemma 6.5. *Let $D_{l,k} \subseteq X$, $\lambda \in \Lambda$, $k \geq 0$ be as in Lemma 6.4 and $\rho : \Lambda \to \mathbb{N}$ the choice function provided therein. Assume that for all $l \in \Lambda$ a set $D'_l \subseteq \mathring{D}_{l,\rho(l)}$ of the same type as D_l is given. Then for any choice of elements $g_l \in \mathfrak{G}_{\mathring{D}_{l,\rho(l)}}$ there is $g \in \mathfrak{G}_{\bigcup_{l\in\Lambda} D'_l} \cap \mathbb{U}^{12}$ such that $g_{\restriction D'_{l,\rho(l)}} = g_l$ for all $1 \leq l \leq r$.*

To conclude the proof of 6.2 let $\{D'_l\}_{l\in\Lambda}$ be as given by Lemma 6.5. We can use Observation 1.6 to find $h_0 \in \mathcal{W}_\delta$ such that $h_0(D_l) \subseteq \mathring{D}'_l$ for all $l \in \Lambda$ (recall that Λ is finite) and by Lemma 6.3 (in fact, we only need τ_{co}-density) we may assume that $h_0 \in \mathbb{U}$. It follows that $\mathfrak{G}_{\bigcup_{l\in\Lambda} D_l} \subseteq \mathbb{U}^{h_0} \subseteq \mathbb{U}^{14}$. □

Definition 6.6. Let μ be the Lebesgue measure on $X = [0,1]$, $\epsilon > 0$ and $\omega : [0,1] \to \mathbb{R}_{\geq 0}$ some continuous function with $\omega^{-1}(0) = \{0,1\}$. By an (ω,ϵ)-fine collection we intend a countable disjoint collection $\{D_l\}_{l\in\Lambda}$ of closed subintervals of $(0,1)$ such that

- $\mathrm{diam}(D_l) \leq \sup_{x\in D_l} \omega(x)$ for all $l \in \Lambda$, and
- $\mu\big(\bigcup_{l\in\Lambda} D_l\big) \leq \epsilon$.

An (ω,ϵ)-*fine set* is defined as a set $\bigcup_{l\in\Lambda} \mathring{D}_l$ for some collection $\{D_l\}_{l\in\Lambda_l}$ as above.

Lemma 6.7. *Let $X = [0,1]$ and μ be the Lebesgue measure. Then there is some number N such that for any separable group topology τ on $\mathfrak{G} := \mathcal{H}_{bac}(X,\mu)$ and every $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ there are ω, ϵ as in Definition 6.6 such that for any (ω,ϵ)-fine set $V \subseteq X$ we have $\mathfrak{G}_V \subseteq \mathcal{V}$.*

Proof. Let $\tilde{d}$ be some compatible proper metric on $(0,1)$. One can define a modified metric $\tilde{\rho}$ on $\mathcal{H}_{bac}(0,1)$ by

$$\tilde{\rho}(g,h) := \min\Big\{1,\ \sup_{x\in(0,1)} \tilde{d}(g(x),h(x)) + \sup_{x\in(0,1)} \tilde{d}(g^{-1}(x),h^{-1}(x)) + \|\mathcal{D}(g)-\mathcal{D}(h)\|_1 + \|\mathcal{D}(g^{-1})-\mathcal{D}(h^{-1})\|_1\Big\}.$$

The same argument as in Proposition A shows that $\tilde{\rho}$ is complete and induces a group topology. It is not separable, but the resulting topological space will have the Baire property nevertheless, so that Lemma 6.3 will hold with the topology induced by $\tilde{\rho}$ in place of τ_{ac}. From this point on, the argument continues in the same fashion as in Lemma 6.2. Alternatively, one may use the existence of generic conjugacy classes [10] and the argument in the proof of automatic continuity for $\mathrm{Homeo}^+(\mathbb{R})$ found in [20]. □

7. Proof of Theorem B

We restate the result toward whose proof we have been working in the previous three sections.

Theorem B. *Let (X,μ) be either*

(I) *an OU pair,*

(II) *or the Cantor space $\mathcal{C} = X$ equipped with either μ_K for some dense subfield $K \subseteq \mathbb{R}$ or a generalized Bernoulli measure.*

Then there is no group topology τ on $\mathcal{H}_{bac}(X,\mu)$ with $\tau_{co} \subsetneq \tau \subsetneq \tau_{ac}$. In fact, more is true: for any separable group topology τ on $\mathcal{H}_{bac}(X,\mu)$, if $\tau_{co} \subsetneq \tau$, then $\tau_{ac} \subseteq \tau$.

Given an open set $V \subseteq X$ we let $\mathrm{cdiam}(V)$ stand for the supremum of the diameter of all connected components of V.

Proposition 7.1. *Let (X,μ) be as in Proposition A and let $\mathfrak{G} := \mathcal{H}_{bac}(X,\mu)$. Assume also that*

(a) *(X,μ) is well-covered;*

(b) *$\mathfrak{G}$ has the local fragmentation property for τ_{co};*

(c) *we are given some collection $\mathscr{C}$ of open sets in X and $N > 0$ with the following properties:*

  (i) *For any $U \in \mathscr{C}$ and $\mathcal{F} \subseteq \pi_0(U)$ we have $\bigcup \mathcal{F} \in \mathscr{C}$.*

  (ii) *For any $\epsilon > 0$ there is a covering of X by sets $V_1, \dots V_N \in \mathscr{C}$ with $\mathrm{cdiam}(V_j) \le \epsilon$.*

Then τ_{co} is the only group topology τ on $\mathfrak{G}$ satisfying the following conditions:

(1) $\tau_{ac} \nsubseteq \tau$, $\tau_{co} \subseteq \tau$.

(2) *For any $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ there is $\epsilon > 0$ such that for any open $V \in \mathscr{C}$ with $\mu(\bar{V}) \leq \epsilon$ and* $\mathrm{cdiam}(V) \leq \epsilon$ *we have $\mathfrak{G}_V \subseteq \mathcal{V}$.*

In particular, there is no group topology τ on $\mathcal{H}_{bac}(X, \mu)$ with $\tau_{co} \subsetneq \tau \subsetneq \tau_{ac}$.

Proof. Assume $\tau_{ac} \nsubseteq \tau$ and $\tau_{co} \subseteq \tau$ and pick arbitrary $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ and $\mathbb{U} = \mathbb{U}^{-1} \in \mathcal{N}_\tau(\mathrm{Id})$ with $\mathbb{U}^{2N+2} \subseteq \mathcal{V}$. Let $\epsilon > 0$ be the constant associated with $\mathbb{U}$ by (2). By Corollary 5.9 there is some open set $V \subseteq X$ and some $g \in \mathbb{U}$ such that $\mu(V) \geq 1 - \frac{1}{2}\epsilon$ and $\mu(g(V)) \leq \frac{1}{2}\epsilon$. Using the regularity and continuity of the measure, one can easily find open sets V', V'' with $\bar{V}' \subseteq V$, $V' \cup V'' = X$, and $\mu(\bar{V}') \leq \epsilon$.

Now, fix some $\delta > 0$ such that $\delta \leq \epsilon$, $\omega_g(\delta) \leq \epsilon$, and δ is smaller than the Lebesgue number of the open cover $\{V', V''\}$. Choose $V_1, \dots V_n \in \mathscr{C}$ with $X = \bigcup_{j=1}^N$ and $\mathrm{cdiam}(V_j) < \delta$. For each $1 \leq j \leq N$ define $V_j' \in \mathscr{C}$ as the union of the connected components of V_j contained in V' and define $V_j'' \in \mathscr{C}$ analogously. Our choice of δ ensures that $V_j = V_j' \cup V_j''$ and that $\mathrm{cdiam}(g(V_j')) \leq \epsilon$. The choice of ϵ implies that $\mathfrak{G}^c_{g(V_j')} \cup \mathfrak{G}^c_{V_j''} \subseteq \mathbb{U}$.

The fragmentation property implies the existence of some $\eta > 0$ such that

$$\mathcal{V}_\eta \subseteq \mathfrak{G}^c_{V_1''} \cdots \mathfrak{G}^c_{V_N''} \mathfrak{G}^c_{V_1'} \cdots \mathfrak{G}^c_{V_N'} \subseteq \mathbb{U}^N (\mathbb{U}^g)^N \subset \mathbb{U}^{2N+2} \subseteq \mathcal{V},$$

and so we are done. □

Proof of Theorem B. Assume first that (X, μ) is an OU pair with $X \neq [0,1]$, $\mathfrak{G} := \mathcal{H}_{bac}(X, \mu)$. We will apply Proposition A. We need to check that conditions (a), (b) of the proposition are satisfied by (X, μ) and that the separability hypothesis on τ implies (c) on τ, where $\mathscr{C}$ is the collection of all finite unions of balls and boundary half-balls. (Note that (c) is trivial when $\tau \subseteq \tau_{ac}$, by Lemma 3.5.) This follows from Corollary 3.10 (for filling partitions), Lemma 2.13 (for local fragmentation) and Lemma 2.6 in combination with Lemma 6.2 (to ensure property (2)).

In case (II) ($X = \mathcal{C}$) of the statement, conditions (a) and (b) of Proposition A are trivial to check. For (c), note that there clopen sets can be made to play the role of embedded balls, so that the relevant notion of an ϵ-fine collection is just that of a finite collection $\{C_l\}_{l \in \Lambda}$ of disjoint clopen sets with $\mathrm{diam}(C_l) < \epsilon$, $\mu\big(\bigcup C_l\big) < \epsilon$. The analog of Fact 2.5 in this context is immediate, since any finite covering by open sets is always refined by a covering by disjoint clopen sets. The analog of Lemma 6.2 can be shown using an simplified version of the same argument, along the lines of [20] (using the commutator trick instead of the existence of generic conjugacy classes), establishing property (2).

The case $X = [0,1]$ is somewhat more involved. We cannot apply Proposition A directly, but we can argue along similar lines. For any $\epsilon > 0$ and any continuous

function $\omega : [0,1] \to \mathbb{R}_{\geq 0}$ with $\omega^{-1}(0) = \{0,1\}$, let $\mathscr{C}_{\omega,\epsilon}$ be the collection of all unions of the interior of an (ω,ϵ)-fine collections in the sense of Definition 6.6. Fix a separable group topology τ on $\mathfrak{G}$ satisfying $\tau_{ac} \not\subseteq \tau$ and $\tau_{co} \subseteq \tau$. We begin by establishing an analog of property (2) in the statement of Proposition A, namely, that for any such τ and any $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ there are ϵ, ω as above such that $\mathfrak{G}_U \subseteq \mathcal{V}$ for any $U \in \mathscr{C}_{\omega,\epsilon}$. This is the content of Lemma 6.7.

We need to show the existence of some constant $M > 0$ such that for any symmetric $\mathcal{V} \in \mathcal{N}_\tau(\mathrm{Id})$ the set $\mathcal{V}^M$ contains some neighborhood of the identity in τ_{co}. Fix $\mathcal{V}$ from here on and let (ω,ϵ) the pair associated to $\mathcal{V}$ by the above discussion.

As in the proof of the proposition, we note that by Corollary 5.9 there is some open set $V \subseteq [0,1]$ and some $g \in \mathbb{U}$ such that $\mu(V) \geq 1 - \frac{1}{2}\epsilon$ and $\mu(g(V)) \leq \frac{1}{2}\epsilon$. Using the regularity and continuity of the measure one can easily find open sets V_0, V_1, V_2 such that $\overline{V}_0 \subseteq V$, $\bigcup_i V_i = (0,1)$ and all of $g(V_0)$, V_1, and V_2 are (ω,ϵ)-fine open sets. Arguing as in the proof of Proposition A, we conclude that $\mathfrak{G}_{V_0}\mathfrak{G}_{V_1}\mathfrak{G}_{V_2} \subseteq \mathcal{V}^5$. This does not imply that $\mathcal{V}^5 \in \mathcal{N}_\tau(1)$ right away, since the applicable variant of fragmentation is Corollary 2.14, which has a weaker conclusion than that of Lemma 2.13. Fortunately, this issue is not new and can be handled by a straightforward adaptation of arguments in the literature.

For $\delta > 0$ write $U_\delta := [0,\delta) \cup (1-\delta, 1]$, and note that by Lemma 2.13 it suffices to show that for some integer constant M' we have $\mathfrak{G}_{U_\delta} \subseteq \mathcal{V}^{M'}$ for some $\delta > 0$. We also know that some neighborhood of the identity in τ_{ac}, and thus $\mathfrak{G}_{U_\delta} \subseteq \mathcal{V}^{M'}$ for some δ small enough, must be contained in the closure of $\mathcal{V}$ in τ_{ac} and thus also contained in the closure of $\mathcal{V}$ in τ_{co}. To find M' with the required properties, one can now apply a set-theoretical argument originally due to Truss [22], as is done in [20]. We refer the reader especially to the discussion starting in the last paragraph of p. 366. □

8. Roelcke precompactness: proofs of Theorems D and E

Theorem D. *If (X,μ) is a 1-dimensional OU pair, then $\mathcal{H}_{bac}(X,\mu)$ is not Roelcke precompact: indeed, if $(g_n)_{n\in\mathbb{N}} \subset \mathcal{H}_{bac}(X,\mu)$ is a sequence such that*

- *$(g_n)_{n\in\mathbb{N}}$ converges to the identity in τ_{co}, and*
- *$(g_n)_{n\in\mathbb{N}}$ does not converge to* Id *in τ_{ac},*

then $(g_n)_{n\in\mathbb{N}}$ is not totally bounded for the Roelcke uniformity.

Proof. For simplicity, assume that μ is the Lebesgue measure and d the standard distance.

Observation 8.1. *Suppose given $u, \tilde{u}, g, g' \in \mathfrak{G}$ with $g = \tilde{u}g'u$ and $g' \in \mathcal{V}_\delta$, as well as some finite collection of disjoint open intervals $\{J_i\}_{1\leq i\leq k}$ in X. Then*

$$\|u\|_{ac} + \|\tilde{u}\|_{ac} \geq \sum_{i=1}^{k} |\mu(J_i) - \mu(g(J_i))| - 2k\delta.$$

Proof. For $1 \leq i \leq k$ let $\Delta_i := |\mu(J_i) - \mu(u(J_i))|$. Let also $J_i' := g'u(J_i)$ and $\Delta_i' := |\mu(J_i') - \mu(\tilde{u}(J_i'))|$. Since $g' \in \mathcal{V}_\delta$, it is clear that $|\mu(J_i) - \mu(g(J_i))| \leq 2\delta + \Delta_i + \Delta_i'$ for all i. But then

$$\sum_{i=1}^{k} |\mu(J_i) - \mu(g(J_i))| \leq 2k\delta + \sum_{i=1}^{k} \Delta_i + \sum_{i=1}^{k} \Delta_i' \leq 2k\delta + \|u\|_{ac} + \|\tilde{u}\|_{ac},$$

as needed. □

Let $(g_n)_{n\in\mathbb{N}} \subseteq \mathfrak{G}$ be a sequence such that $g_n \in \mathcal{V}_{\delta_n}$ for some sequence $(\delta_n)_n \subseteq \mathbb{R}_{>0}$ converging to zero, but such that $(g_n)_n$ does not converge to Id in τ_{ac}. Up to passing to a subsequence, we may assume $\|g_n\| \geq \epsilon$ for some $\epsilon > 0$. If the conclusion of the theorem does not hold, then up to replacing $(g_n)_n$ by a subsequence we may assume that for any $n > 0$ there are $u_n, \tilde{u}_n \in \mathcal{V}_{\epsilon/8}$ such that $g_n = u_n g_0 \tilde{u}_n$. An approximation argument such as in Observation 1.6 yields the existence of some finite disjoint collection of intervals $J_1, \dots J_k \subseteq X$ with disjoint interiors such that $\sum_{i=1}^{k} |\mu(g_0(J_i)) - \mu(J_i)| \geq \frac{1}{2}\epsilon$, which together with Observation 8.1 leads to a contradiction:

$$\tfrac{1}{4}\epsilon \geq \|u_n\|_{ac} + \|\tilde{u}_n\|_{ac} \geq \tfrac{1}{2}\epsilon - 2k\delta_n \xrightarrow{n\to\infty} \tfrac{1}{2}\epsilon. \qquad \square$$

Key to making this proof work is that the condition $g \in \mathcal{V}_\delta$ implies some uniformly bounded distortion under the action of g for all intervals. The absence of a similar property for the Cantor space with any reasonably homogeneous Borel probability measure is what makes Theorem E possible. Before proving it, we state two simple preliminary results. The first follows easily from the density of continuous functions in $L1(X, \mu)$ for any finite Borel measure μ on a separable topological space X; see [5, p. 92].

Fact 8.2. *Let μ be a Borel probability measure on the Cantor space $\mathcal{C}$, $M > 0$ a positive integer and $\phi \in L^1(\mathcal{C}, \mu)$, $\phi > 0$ such that $\|\phi - \phi \mathbb{1}_{\phi^{-1}(0,1]}\|_1 \leq 1/(2M)$. Then there is a step function $\psi := \sum_{j=1}^{3M} \frac{k_j}{3M} \mathbb{1}_{A_j}$, where the A_j are clopen and $\|\phi - \psi\|_1 \leq 1/M$.*

Observation 8.3. *Let K be a countable subfield of $\mathbb{R}$ and $\mathfrak{G} = \mathcal{H}_{bac}(\mathcal{C}, \mu_K)$. For any two partitions $\{A_i\}_{i=1}^r$ and $\{B_i\}_{i=1}^r$ of $\mathcal{C}$ into clopen sets, there exists some $g \in \mathfrak{G}$ such that $g(A_i) = B_i$ for all $1 \leq i \leq r$ and $\mathcal{D}(g)_{\restriction A_i}$ is constantly equal to $\mu_k(B_i)/\mu_K(A_i)$.*

Proof. The assumption that K is a field, together with the uniqueness of Fraïssé limits, implies that for any clopen set $A \subseteq \mathcal{C}$ there is a homeomorphism $h : \mathcal{C} \simeq A$ such that $h_*\mu_K = \mu_K(A)(\mu_K)_{\restriction A}$, from which the result easily follows (compare with Observation 2.3). □

Theorem E. *Let K a countable subfield of $\mathbb{R}$. Then $(\mathcal{H}_{bac}(\mathcal{C}, \mu_K), \tau_{ac})$ is Roelcke precompact.*

Proof. Let $\mathbb{B}$ denote the boolean algebra of clopen subsets of $\mathcal{C}$, $\mathfrak{G} := \mathcal{H}_{bac}(\mathcal{C}, \mu)$. To prove $(\mathfrak{G}, \tau_{ac})$ is Roelcke precompact, it suffices to show that for all $\mathcal{W} \in \mathcal{N}_{\tau_{ac}}(\mathrm{Id})$ and any infinite sequence $(g_n)_{n\geq 0} \subseteq \mathfrak{G}$ there is some infinite subsequence $(g_{n_k})_{k\geq 0}$ such that $g_{n_l} \in \mathcal{W} g_{n_k} \mathcal{W}$ for all $k \neq l$. Without loss of generality there is some integer constant $N > 0$ and some partition $\{A_j\}_{j=1}^r$ of $\mathcal{C}$ into clopen sets such that writing $\mathfrak{h}$ for the subgroup consisting of all the elements of $\mathfrak{G}$ preserving each A_j setwise, we have that $\mathcal{W}$ consists precisely of those elements $g \in \mathfrak{h}$ satisfying $\|g\|_{ac} \leq 1/N$. For $m \geq 0$, $i, j \in \{1, \dots r\}$ let

$$\begin{aligned} A_{i,j}^m &:= A_i \cap g_m^{-1}(A_j), \\ A_{i,j}^{m,+} &:= \{x \in A_{i,j}^m \mid \mathcal{D}(g_m)(x) \geq 1\}, \\ A_{i,j}^{m,-} &:= \{x \in A_{i,j}^m \mid \mathcal{D}(g_m)(x) < 1\}. \end{aligned}$$

Write $K := 6Nr^2$. For $m \geq 0$ and $(i, j) \in \{1, \dots r\}^2$ we can choose a partition $\{\hat{A}_{i,j}^{m,+}, \hat{A}_{i,j}^{m,-}\} \subseteq \mathbb{B}$ of $A_{i,j}^m$ such that if we write $D_{i,j}^{m,+} := A_{i,j}^{m,+} \triangle \hat{A}_{i,j}^{+} = A_{i,j}^{m,-} \triangle \hat{A}_{i,j}^{-}$, then

$$\|g_m\|_{ac, D_{i,j}^m} \leq \frac{1}{2K}, \quad \|g_m^{-1}\|_{ac, g_m(D_{i,j}^m)} \leq \frac{1}{2K}. \tag{23}$$

For convenience, we will use the abbreviations

$$\begin{gathered} \hat{B}_{i,j}^{m,\pm} := g_m(\hat{A}_{i,j}^{m,\pm}), \quad C_{i,j}^{m,-} := \hat{A}_{i,j}^{m,-}, \quad C_{i,j}^{m,+} := \hat{B}_{i,j}^{m,+}, \\ h_{i,j}^{m,-} := (g_m)_{\restriction \hat{A}_{i,j}^{m,-}}, \quad h_{i,j}^{m,+} := (g_m)^{-1}_{\restriction \hat{B}_{i,j}^{m,+}}. \end{gathered}$$

For $n \geq 0$ and $(i, j) \in \mathcal{P}$ we can find by Fact 8.2 and (23) a partition

$$\{C_{i,j}^{m,\pm}(1), \dots, C_{i,j}^{m,\pm}(3K)\} \subseteq \mathbb{B}$$

of $C_{i,j}^{m,\pm}$ such that

$$\|\mathcal{D}(h_{i,j}^{m,\pm}) - \sum_{k=1}^{3} K \frac{k}{3K} \mathbb{1}_{C_{i,j}^{m,\pm}(k)}\|_{1, C_{i,j}^{m,\pm}} \leq \frac{1}{K}. \tag{24}$$

Up to passing to a further subsequence we may assume that for all $1 \leq k \leq 3K$ we have

$$\left|\mu(\hat{A}_{i,j}^{m,\pm}(k)) - \mu(\hat{A}_{i,j}^{n,\pm}(k))\right| \leq \frac{1}{3K^2}, \quad \left|\mu(\hat{B}_{i,j}^{m,\pm}(k)) - \mu(\hat{B}_{i,j}^{n,\pm}(k))\right| \leq \frac{1}{3K^2}. \tag{25}$$

By Observation 8.3 we can find $v, w \in \mathfrak{h}$ such that for all $(i, j) \in \mathcal{P}$ and $1 \leq k \leq 3K$ the following properties are satisfied:

(i) $v(\hat{A}_{i,j}^{0,\pm}(k)) = \hat{A}_{i,j}^{1,\pm}(k)$.
(ii) $\mathcal{D}(v)$ has constant value $\dfrac{\mu(\hat{A}_{i,j}^{1,\pm}(k))}{\mu(\hat{A}_{i,j}^{0,\pm}(k))}$ on $\hat{A}_{i,j}^{0,\pm}(k)$.
(iii) $w(\hat{B}_{i,j}^{1,\pm}(k)) = \hat{B}_{i,j}^{0,\pm}(k)$.
(iv) $\mathcal{D}(w)$ has constant value $\dfrac{\mu(\hat{B}_{i,j}^{0,\pm}(k))}{\mu(\hat{B}_{i,j}^{1,\pm}(k))}$ on $\hat{B}_{i,j}^{0,\pm}(k)$.

Define A^+, A^- by $A^\pm := \bigcup_{i,j=1}^{r} \bigcup_{k=1}^{3} K\hat{A}_{i,j}^{0,\pm}(k)$ and similarly for $B^\pm$. From (25) and (ii) we get

$$\|v\|_{ac,\hat{A}_{i,j}^{0,\pm}} = \sum_{k=1}^{3} K\left|\mu(\hat{A}_{i,j}^{0,\pm}(k)) - \mu(\hat{A}_{i,j}^{1,\pm}(k))\right| \le \frac{1}{K}, \quad \|v\|_{ac} \le \frac{2r^2}{K} < \frac{1}{2N}. \tag{26}$$

Clearly, similar inequalities hold for w and $\hat{B}_{i,j}^{l,\pm}$ in place of v and $\hat{A}_{i,j}^{l,\pm}$.

Let $\tilde{g}_1 := wg_1v$ and note that $\tilde{g}_1(\hat{A}_{i,j}^{0,\pm}(k)) = g_0(\hat{A}_{i,j}^{0,\pm}(k)) = \hat{B}_{i,j}^{0,\pm}(k)$ for all $1 \le i, j \le r$ and $1 \le k \le 3K$. It is easy to see that there are elements $u, \tilde{u} \in \mathfrak{h}$ with

$$u_{\restriction A^+} = \mathrm{Id}_{A^+}, \quad \tilde{u}_{\restriction B^-} = \mathrm{Id}_{B^-}, \quad g_0 = \tilde{u}\tilde{g}_1 u,$$

and with u preserving each $\hat{A}_{i,j}^{0,\pm}(k)$ and $\tilde{u}$ preserving each $\hat{B}_{i,j}^{0,\pm}(k)$.

Lemma 8.4. *The inequalities* $\|u\|_{ac}, \|\tilde{u}\|_{ac} < \dfrac{1}{2N}$ *hold.*

Proof. We get the estimate

$$\begin{aligned}
&\left\|\mathcal{D}(\tilde{g}_1) - \sum_{k=1}^{3} K\frac{k}{3K}\mathbb{1}_{\hat{A}_{i,j}^{0,-}(k)}\right\|_{1,\hat{A}_{i,j}^{0,-}} \\
&\quad= \left\|\mathcal{D}(wg_1) - \sum_{k=1}^{3} K\frac{k\mu(\hat{A}_{i,j}^{0,-}(k))}{3K\mu(\hat{A}_{i,j}^{1,-}(k))}\mathbb{1}_{\hat{A}_{i,j}^{1,-}(k)}\right\|_{1,\hat{A}_{i,j}^{1,-}} \\
&\quad\le \left\|\mathcal{D}(g_1) - \sum_{k=1}^{3} K\frac{k}{3K}\mathbb{1}_{\hat{A}_{i,j}^{1,-}(k)}\right\|_{1,\hat{A}_{i,j}^{1,-}} + \sum_{k=1}^{3} K\frac{k}{3K}\left|\mu(\hat{A}_{i,j}^{0,-}(k)) - \mu(\hat{A}_{i,j}^{1,-}(k))\right| \\
&\qquad\qquad + \|\mathcal{D}(wg_1) - \mathcal{D}(g_1)\|_{1,\hat{A}_{i,j}^{1,-}} \\
&\quad\le \frac{1}{K} + \frac{1}{K} + \|w\|_{ac,\hat{B}_{i,j}^{1}} \le \frac{3}{K} \le \frac{1}{2N}.
\end{aligned}$$

In this display the equality uses Observation 1.1 (after an application of the chain rule to rewrite $D(\tilde{g}_1)$ as $D(wg_1 \circ v)D(v)$), plus the identities (i) and (ii); while in the last inequality we used (24) to bound the first term, Observation 1.4 once again and (25) to bound the second term. Adding over all $1 \le i \le j \le r$ and combining the result with (24) for $m = 0$ yields $\|\mathcal{D}(\tilde{g}_1) - \mathcal{D}(g_0)\|_{1,A^-} \le 4(|\mathcal{P}^+| + |\mathcal{P}^-|)/K = 1/(2N)$. Using Observation 1.4 again and then the inequality above we obtain

$$\|\tilde{u}\|_{ac} = \|\tilde{u}\|_{ac,B^-} = \|\mathcal{D}(\tilde{u}\tilde{g}_1) - \mathcal{D}(\tilde{g}_1)\|_{1,B^-} = \|\mathcal{D}(g_0) - \mathcal{D}(\tilde{g}_1)\|_{1,B^-} \le \frac{1}{2N}.$$

By estimating $\|\mathcal{D}(\tilde{g}_1^{-1}) - \mathcal{D}(g_0^{-1})\|_{1,B^+}$ symmetrically we also get $\|u\|_{ac} \le \dfrac{1}{2N}$. □

Now $g_0 = \tilde{u}wg_1vu$ with $\tilde{u}w, vu \in \mathfrak{h}$. On the other hand,

$$\|\tilde{u}w\|_{ac} \le \|\tilde{u}\|_{ac} + \|w\|_{ac} \le \frac{1}{N}$$

by Observation 1.4 and similarly $\|vu\| \le \frac{1}{N}$, so $\tilde{u}w, vu \in \mathcal{W}$, contradicting the choice of $(g_n)_{n\ge 0}$. This concludes the proof. □

References

[1] R. D. Anderson, "The algebraic simplicity of certain groups of homeomorphisms", *Amer. J. Math.* **80** (1958), 955–963. MR

[2] J. R. Choksi, "Inverse limits of measure spaces", *Proc. London Math. Soc.* (3) **8** (1958), 321–342. MR

[3] C. O. Christenson and W. L. Voxman, *Aspects of topology*, Pure and Applied Mathematics **39**, Dekker, New York, 1977. MR

[4] M. P. Cohen, "Polishability of some groups of interval and circle diffeomorphisms", *Fund. Math.* **248**:1 (2020), 91–109. MR

[5] J. L. Doob, *Measure theory*, Graduate Texts in Mathematics **143**, Springer, 1994. MR

[6] R. D. Edwards and R. C. Kirby, "Deformations of spaces of imbeddings", *Ann. of Math.* (2) **93** (1971), 63–88. MR

[7] A. Fathi, "Structure of the group of homeomorphisms preserving a good measure on a compact manifold", *Ann. Sci. École Norm. Sup.* (4) **13**:1 (1980), 45–93. MR

[8] P. Gartside and A. Glyn, "Autohomeomorphism groups", *Topology Appl.* **129**:2 (2003), 103–110. MR

[9] J. Herndon, "Absolute continuity and large-scale geometry of Polish groups", preprint, 2018. arXiv 1802.10239

[10] D. T. Ihli, "Generation and genericity of the group of absolutely continuous homeomorphisms of the interval", preprint, 2022. arXiv 2203.02854

[11] A. S. Kechris and C. Rosendal, "Turbulence, amalgamation, and generic automorphisms of homogeneous structures", *Proc. Lond. Math. Soc.* (3) **94**:2 (2007), 302–350. MR

[12] R. Kirby, "Stable homeomorphisms and the annulus conjecture", *Ann. Math.* **89**:3 (1969).

[13] A. Klenke, *Probability theory: a comprehensive course*, 2nd ed., Springer, 2014. MR

[14] K. Mann, "Automatic continuity for homeomorphism groups and applications", *Geom. Topol.* **20**:5 (2016), 3033–3056. MR

[15] B. C. Mazur, "On embeddings of spheres", *Acta Math.* **105** (1961), 1–17. MR

[16] J. C. Oxtoby and S. M. Ulam, "Measure-preserving homeomorphisms and metrical transitivity", *Ann. of Math.* (2) **42** (1941), 874–920. MR

[17] F. Quinn, "Ends of maps, III: Dimensions 4 and 5", *J. Differential Geometry* **17**:3 (1982), 503–521. MR

[18] C. Rosendal, "Global and local boundedness of Polish groups", *Indiana Univ. Math. J.* **62**:5 (2013), 1621–1678. MR

[19] C. Rosendal, *Coarse geometry of topological groups*, Cambridge Tracts in Mathematics **223**, Cambridge University Press, 2022. MR

[20] C. Rosendal and S. Solecki, "Automatic continuity of homomorphisms and fixed points on metric compacta", *Israel J. Math.* **162** (2007), 349–371. MR

[21] S. Solecki, "Polish group topologies", pp. 339–364 in *Sets and proofs* (Leeds, 1997), London Math. Soc. Lecture Note Ser. **258**, Cambridge Univ. Press, 1999. MR

[22] J. K. Truss, "Generic automorphisms of homogeneous structures", *Proc. London Math. Soc.* (3) **65**:1 (1992), 121–141. MR

Received October 6, 2025. Revised February 26, 2026.

J. DE LA NUEZ GONZÁLEZ
KOREA INSTITUTE FOR ADVANCED STUDY (KIAS)
SEOUL
SOUTH KOREA
jnuezgonzalez@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 344, No. 1, 2026
https://doi.org/10.2140/pjm.2026.344.145

ADJUNCTION INEQUALITY FOR SPATIALLY REFINED s-INVARIANTS

QIUYU REN

We note an adjunction inequality in the space $k\overline{\mathbb{CP}^2}$ for the s-version of the Sq^1-refinement of Rasmussen's s-invariant. This does not hold for general spatial refinements of s-invariants.

Ren [7, Corollary 1.4] proved that for an oriented link cobordism $\Sigma : L_0 \to L_1$ in some $k\overline{\mathbb{CP}^2} \setminus (\mathrm{int}(B^4) \sqcup \mathrm{int}(B^4))$ between two oriented links $L_0, L_1 \subset S^3$ with $\pi_0(L_1) \to \pi_0(\Sigma)$ surjective, we have an adjunction-type inequality for Rasmussen's s-invariant [2; 6] over any coefficient field $\mathbb{F}$ of the form

$$s(L_1) \le s(L_0) - \chi(\Sigma) - [\Sigma]^2 - |[\Sigma]|. \tag{1}$$

However, Ciprian Manolescu informed us that the adjunction inequality (for knots) may fail for spatially refined s-invariants defined by Lipshitz and Sarkar [5], Sarkar, Scaduto and Stoffregen [8] and Schütz [10]. The counterexample he found is the knot 9_{42} which bounds a disk in the punctured $\overline{\mathbb{CP}^2}$ with homology class $[\overline{\mathbb{CP}^1}]$ whose geometric intersection number with $\overline{\mathbb{CP}^1}$ is 3 (see Figure 1), but for which some versions of the odd Sq^1-, Sq^2- and even Sq^2-refined s-invariants are equal to 2, violating (1).

Let us see what goes wrong with the proof in [7] if it were applied to these refined s-invariants. The proof relies on the following three facts about s-invariants:

(a) s is defined for all oriented links in S^3. If $\Sigma : L_0 \to L_1$ is an oriented cobordism in $S^3 \times I$ with $\pi_0(L_0) \to \pi_0(\Sigma)$ surjective, then $s(L_1) - s(L_0) \ge \chi(\Sigma)$.

(b) If $T(n,n)_{p,q}$ is the torus link $T(n,n)$, equipped with an orientation where p strands are oriented against the other q strands, $p \ge q$, $p+q=n$, then $s(T(n,n)_{p,q}) = (p-q)^2 - 2p + 1$.

(c) $s(L \sqcup T(n,n)_{p,q}) = s(L) + s(T(n,n)_{p,q}) - 1$ for any oriented link L.

Lipshitz and Sarkar's refined s-invariants are only defined for knots, but it is not hard to extend the definition to links and address item (a), which we will do in a while. For this natural generalization, however, one can verify using KnotJob [9]

MSC2020: primary 57K18; secondary 57K10, 57K40.
Keywords: Rasmussen s-invariant, adjunction inequality, spatially refinement, Khovanov homology.

© 2026 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

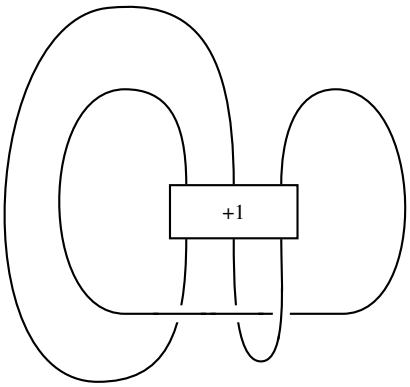


Figure 1. A diagram of the knot 9_{42}. Untwisting the twist region results in the unknot.

that $s_+^{\mathrm{Sq}^2}(T(3,3)_{2,1}) = s(T(3,3)_{2,1}) + 2$, violating item (b). The same violation for $T(3,3)_{2,1}$ appears for the odd Sq^1-refinement and the s-version of one of Schütz's two odd Sq^2-refinements (once appropriately defined). These contribute to the failure of the corresponding refined s-invariants to satisfy (1) for the example of 9_{42}.

On the positive side, we show the following.

Theorem 1. *The adjunction inequality* (1) *holds for* $s_+^{\mathrm{Sq}^1}$*, the* s*-version of the even* Sq^1*-refinement of* s (*defined for knots in* [5], *and extended to all oriented links as in Section 0.1*).

For the proof of (1) in [7] to work, it suffices to justify (a)–(c). We address each of them in the following three sections, respectively.

0.1. ***Spatially refined s-invariants for links.*** We recall the construction and some properties of even spatial refinements of Rasmussen's s-invariant from Lipshitz and Sarkar's work [5], with the relevant definitions extended from knots to oriented links in the natural way. Compare with [5, Definitions 1.1–1.2, Theorem 1].

Let θ be a stable cohomology operation of degree m over a field $\mathbb{F}$ and let L be a nonempty oriented link in the sphere S^3 with orientation $\mathfrak{o}$. Also, let $W(L) = \mathrm{span}\{[\mathfrak{s}_\mathfrak{o}], [\mathfrak{s}_{\bar{\mathfrak{o}}}]\} \subset \mathrm{Kh}^0_{\mathrm{Lee}}(L;\mathbb{F})$ (use Bar-Natan homology $\mathrm{Kh}_{\mathrm{BN}}$ [1; 11] instead of Lee homology $\mathrm{Kh}_{\mathrm{Lee}}$ [4] if $\mathrm{char}(\mathbb{F}) = 2$, and do the same below) denote the 2-dimensional subspace spanned by the canonical generators corresponding to the orientation of L and its reverse. For an integer $q \equiv \#\pi_0(L) \pmod 2$, consider the maps

$$\mathrm{Kh}^{-m,q}(L;\mathbb{F}) \xrightarrow{\theta} \mathrm{Kh}^{0,q}(L;\mathbb{F}) \xleftarrow{p} H^0(\mathrm{CKh}^{\geq q}_{\mathrm{Lee}}(L;\mathbb{F})) \xrightarrow{j} \mathrm{Kh}^0_{\mathrm{Lee}}(L;\mathbb{F}), \tag{2}$$

where Kh is the Khovanov homology [3], $\mathrm{CKh}_{\mathrm{Lee}}$ is the Lee chain complex, and $\mathrm{CKh}^{\geq q}_{\mathrm{Lee}}$ is its subcomplex consisting of elements with filtered degree at least q.

With respect to the link L, the integer q is *θ-half-full* (resp. *θ-full*) if the subspace $V^q(L) := j\big(p^{-1}(\mathrm{im}(\theta))\big) \cap W(L)$ of $W(L)$ is at least 1-dimensional (resp. 2-dimensional); it is *half-full* (resp. *full*) if the subspace $\mathrm{im}(j) \cap W(L)$ of $W(L)$ is at least 1-dimensional (resp. 2-dimensional). Hence

- if q is θ-half-full (resp. θ-full), then q is half-full (resp. full);
- if q is half-full (resp. full), then $q-2$ is θ-half-full (resp. θ-full).

Note that in this language, the classical s-invariant of L over $\mathbb{F}$ is

$$s^{\mathbb{F}}(L) = \max\{q : q \text{ is half-full}\} - 1 = \max\{q : q \text{ is full}\} + 1.$$

Define the θ-refined s-invariants of L by

$$r_+^\theta(L) := \max\{q : q \text{ is } \theta\text{-half-full}\} + 1, \quad s_+^\theta(L) := \max\{q : q \text{ is } \theta\text{-full}\} + 3.$$

Thus $r_+^\theta(L), s_+^\theta(L) \in \{s^{\mathbb{F}}(L), s^{\mathbb{F}}(L)+2\}$. If 0 denotes the zero stable cohomology operation over $\mathbb{F}$, then $s^{\mathbb{F}} = r_+^0 = s_+^0$.

For computational purposes, we note that $s_+^\theta(L) = s^{\mathbb{F}}(L) + 2$ if and only if there exists an $x \in H^0(\mathrm{CKh}_{\mathrm{Lee}}^{\geq s^{\mathbb{F}}(L)-1}(L;\mathbb{F}))$ with $j(x) = [\mathfrak{s}_o]$ and $p(x) \in \mathrm{im}(\theta)$. Similarly, $r_+^\theta(L) = s^{\mathbb{F}}(L) + 2$ if and only if there exists an $x \in H^0(\mathrm{CKh}_{\mathrm{Lee}}^{\geq s^{\mathbb{F}}(L)+1}(L;\mathbb{F}))$ with $j(x) = [\mathfrak{s}_o] \pm [\mathfrak{s}_{\bar{o}}]$ and $p(x) \in \mathrm{im}(\theta)$ for some sign $\pm$.

If $\Sigma : L_0 \to L_1$ is an oriented link cobordism in $S^3 \times I$ between nonempty links with $\pi_0(L_0) \to \pi_1(\Sigma)$ surjective, then $\mathrm{Kh}_{\mathrm{Lee}}(\Sigma)$ maps $W(L_0)$ isomorphically onto $W(L_1)$, which restricts to monomorphisms $V^q(L_0) \to V^{q+\chi(\Sigma)}(L_1)$ (this uses the naturality of stable cohomology operations on Kh; see [5, Theorem 4]). It follows that

$$r_+^\theta(L_1) - r_+^\theta(L_0) \geq \chi(\Sigma), \quad s_+^\theta(L_1) - s_+^\theta(L_0) \geq \chi(\Sigma). \tag{3}$$

Finally, we define $r_+^\theta(\varnothing) = s_+^\theta(\varnothing) = 1$. It is straightforward to check that (3) still holds when L_0, L_1 are allowed to be the empty link.

The minus versions of the spatial refinements of s are defined as $r_-^\theta(L) := -r_+^\theta(\bar{L})$, $s_-^\theta(L) := -s_+^\theta(\bar{L})$, where $\bar{L}$ denotes the mirror image of L. We will not use them.

0.2. Sq^1-*refined s-invariants for torus links.*

Proposition 2. *For $p \geq q$, $p+q = n$, we have*

$$s_+^{\mathrm{Sq}^1}(T(n,n)_{p,q}) = s^{\mathbb{F}_2}(T(n,n)_{p,q}) = (p-q)^2 - 2p + 1.$$

The same holds for $r_+^{\mathrm{Sq}^1}$ when $p = q$.

Proof. Write $T = T(n,n)_{p,q}$ and $s(\cdot) = s^{\mathbb{F}_2}(\cdot)$. By Theorems 1.1 and 2.1 of [7], we know $\mathrm{Kh}^{0,s(T)-1}(T) = \mathbb{Z}$ and $\mathrm{Kh}^{1,s(T)-1}(T) = 0$. Since Sq^1 is the Bockstein homomorphism, it follows that $\mathrm{Sq}^1 : \mathrm{Kh}^{-1,s(T)-1}(T;\mathbb{F}_2) \to \mathrm{Kh}^{0,s(T)-1}(T;\mathbb{F}_2)$ is

the zero map. Hence $s_+^{\mathrm{Sq}^1}(T) = s(T)$. When $p = q$, the induction scheme in [7, Section 5] implies that $\mathrm{Kh}^{-1,*}(T) = 0$ for all $*$. Hence $r_+^{\mathrm{Sq}^1}(T) = s(T)$. The last equality is [7, Theorem 1.1]. □

0.3. $s_+^{\mathrm{Sq}^1}$ *under disjoint union with nice links.*

Proposition 3. *Let T be an oriented link in S^3 such that*

$$\mathrm{Sq}^1 : \mathrm{Kh}^{i-1,s^{\mathbb{F}_2}(T)-1}(T;\mathbb{F}_2) \to \mathrm{Kh}^{i,s^{\mathbb{F}_2}(T)-1}(T;\mathbb{F}_2)$$

is zero for $i = 0, 1$. Then, for any oriented link L in S^3, we have

$$s_+^{\mathrm{Sq}^1}(L \sqcup T) = s_+^{\mathrm{Sq}^1}(L) + s_+^{\mathrm{Sq}^1}(T) - 1.$$

By arguments in the proof of Proposition 2 above, we know all torus links $T = T(n,n)_{p,q}$ satisfy the conditions in Proposition 3.

Proof. Write for short $s(\,\cdot\,) = s^{\mathbb{F}_2}(\,\cdot\,)$. When $L = \varnothing$ or $T = \varnothing$, the statement is trivial. Assume T, L are nonempty. Since $\mathrm{Sq}^1 : \mathrm{Kh}^{-1,s(T)-1}(T;\mathbb{F}_2) \to \mathrm{Kh}^{0,s(T)-1}(T;\mathbb{F}_2)$ is zero, we know $s_+^{\mathrm{Sq}^1}(T) = s(T)$.

Since $s(L \sqcup T) = s(L) + s(T) - 1$, it suffices to show that $s_+^{\mathrm{Sq}^1}(L \sqcup T) = s(L \sqcup T) + 2$ if and only if $s_+^{\mathrm{Sq}^1}(L) = s(L) + 2$. Although the "if" direction is simpler, we prove the two directions together.

We first observe that any filtered cochain complex C^* over a field $\mathbb{F}$ is filtered isomorphic to a filtered complex of the form

$$C^* \cong H^*(C) \oplus \left(\bigoplus_\alpha E_\alpha^*\right) \oplus \left(\bigoplus_\beta A_\beta^*\right),$$

where each E_α (resp. A_β) is a 2-term filtered cochain complex $\mathbb{F} \xrightarrow{=} \mathbb{F}$ in some homological degree, which is not (resp. is) filtration-preserving. Thus, C^* is filtered chain homotopy equivalent to $H^*(C) \oplus \left(\bigoplus_\alpha E_\alpha^*\right)$.

In particular, for each $X \in \{L, T\}$, we can choose such a filtered chain homotopy equivalence (which we simply denote as $=$) $\mathrm{CKh}^*_{\mathrm{BN}}(X) = \mathrm{Kh}^*_{\mathrm{BN}}(X) \oplus \left(\bigoplus_\alpha E^*_{X,\alpha}\right)$. Then

$$\mathrm{Kh}^{*,*}(X;\mathbb{F}_2) = \mathrm{gr}^*(\mathrm{Kh}^*_{\mathrm{BN}}(X)) \oplus \left(\bigoplus_\alpha \mathrm{gr}^*(E^*_{X,\alpha})\right),$$

and $p_X(j_X^{-1}([\mathfrak{s}_{o_X}]) \subset \mathrm{Kh}^{0,s(X)-1}(X;\mathbb{F}_2)$ (see (2) for notation; here $q = s(X) - 1$ and subscripts indicate dependence on the link) is the affine space over

$$V_X := p_X(\ker(j_X)) = \bigoplus_\alpha \mathrm{im}(H^0(E^{*,\geq s(X)-1}_{X,\alpha}) \to \mathrm{gr}^{s(X)-1}(E^0_{X,\alpha}))$$

containing $x_X := \mathrm{gr}^{s(X)-1}([\mathfrak{s}_{o_X}])$.

For the disjoint union, we have

$$\mathrm{CKh}^*_{\mathrm{BN}}(L \sqcup T) = \mathrm{Kh}^*_{\mathrm{BN}}(L) \otimes \mathrm{Kh}^*_{\mathrm{BN}}(T) \oplus \left(\bigoplus_{\alpha} E^*_{L,\alpha} \otimes \mathrm{Kh}^*_{\mathrm{BN}}(T) \right) \oplus \left(\bigoplus_{\beta} \mathrm{Kh}^*_{\mathrm{BN}}(L) \otimes E^*_{T,\beta} \right) \oplus \left(\bigoplus_{\alpha,\beta} E^*_{L,\alpha} \otimes E^*_{T,\beta} \right),$$

and we will use the decomposition

$$\begin{aligned} \mathrm{Kh}^{0,s(L\sqcup T)-1}(L \sqcup T; \mathbb{F}_2) &= \bigoplus_{h,q \in \mathbb{Z}} (\mathrm{Kh}^{h,s(L)-1+q}(L; \mathbb{F}_2) \otimes \mathrm{Kh}^{-h,s(T)-1-q}(T; \mathbb{F}_2)) \\ &= \mathrm{Kh}^{0,s(L)-1}(L; \mathbb{F}_2) \otimes \mathrm{gr}^{s(T)-1}(\mathrm{Kh}^0_{\mathrm{BN}}(T)) \oplus H_{\mathrm{skew}}, \end{aligned}$$

where

$$H_{\mathrm{skew}} = \mathrm{Kh}^{0,s(L)-1}(L; \mathbb{F}_2) \otimes \left(\bigoplus_{\beta} \mathrm{gr}^{s(T)-1}(E^0_{T,\beta}) \right) \oplus \left(\bigoplus_{(h,q)\neq(0,0)} \mathrm{Kh}^{h,s(L)-1+q}(L; \mathbb{F}_2) \otimes \mathrm{Kh}^{-h,s(T)-1-q}(T; \mathbb{F}_2) \right).$$

We see that $p_{L\sqcup T}(j^{-1}_{L\sqcup T}([\mathfrak{s}_{\mathfrak{o}_L}] \otimes [\mathfrak{s}_{\mathfrak{o}_T}])) \subset \mathrm{Kh}^{0,s(L\sqcup T)-1}(L \sqcup T; \mathbb{F}_2)$ is the space containing $x_L \otimes x_T$ affine over

$$V_{L\sqcup T} := p_{L\sqcup T}(\ker(j_{L\sqcup T})) = V_L \otimes \mathrm{gr}^{s(T)-1}(\mathrm{Kh}^0_{\mathrm{BN}}(T)) \oplus W$$

for some subspace $W \subset H_{\mathrm{skew}}$.

On the other hand, since

$$\mathrm{Sq}^1_{L\sqcup T} = \mathrm{Sq}^1_L \otimes 1 + 1 \otimes \mathrm{Sq}^1_T,$$

by our assumption on T, the image $\mathrm{im}(\mathrm{Sq}^1_{L\sqcup T}) \subset \mathrm{Kh}^{0,s(L\sqcup T)-1}(L \sqcup T; \mathbb{F}_2)$ is equal to

$$\mathrm{im}(\mathrm{Sq}^1_L : \mathrm{Kh}^{-1,s(L)-1}(L; \mathbb{F}_2) \to \mathrm{Kh}^{0,s(L)-1}(L; \mathbb{F}_2)) \otimes \mathrm{gr}^{s(T)-1}(\mathrm{Kh}^0_{\mathrm{BN}}(T)) \oplus W'$$

for some subspace $W' \subset H_{\mathrm{skew}}$. It follows that

$$\begin{aligned} s^{\mathrm{Sq}^1}_+(L \sqcup T) = s(L \sqcup T) + 2 &\iff x_L \otimes x_T \in V_{L\sqcup T} + \mathrm{im}(\mathrm{Sq}^1_{L\sqcup T}) \\ &\iff x_L \in V_L + \mathrm{im}(\mathrm{Sq}^1_L) \\ &\iff s^{\mathrm{Sq}^1}_+(L) = s(L) + 2. \end{aligned}$$

□

Acknowledgment

We thank Ciprian Manolescu and Robert Lipshitz for helpful discussions and correspondence. We thank the referee for helpful comments.

References

[1] D. Bar-Natan, "Khovanov's homology for tangles and cobordisms", *Geom. Topol.* **9** (2005), 1443–1499. MR

[2] A. Beliakova and S. Wehrli, "Categorification of the colored Jones polynomial and Rasmussen invariant of links", *Canad. J. Math.* **60**:6 (2008), 1240–1266. MR

[3] M. Khovanov, "A categorification of the Jones polynomial", *Duke Math. J.* **101**:3 (2000), 359–426. MR

[4] E. S. Lee, "An endomorphism of the Khovanov invariant", *Adv. Math.* **197**:2 (2005), 554–586. MR

[5] R. Lipshitz and S. Sarkar, "A refinement of Rasmussen's S-invariant", *Duke Math. J.* **163**:5 (2014), 923–952. MR

[6] J. Rasmussen, "Khovanov homology and the slice genus", *Invent. Math.* **182**:2 (2010), 419–447. MR

[7] Q. Ren, "Lee filtration structure of torus links", *Geom. Topol.* **28**:8 (2024), 3935–3960. MR

[8] S. Sarkar, C. Scaduto, and M. Stoffregen, "An odd Khovanov homotopy type", *Adv. Math.* **367** (2020), art. id. 107112. MR

[9] D. Schütz, "KnotJob", 2025, https://www.maths.dur.ac.uk/users/dirk.schuetz/knotjob.html.

[10] D. Schütz, "Two second Steenrod squares for odd Khovanov homology", *Fund. Math.* **271**:1 (2025), 29–69. MR

[11] P. R. Turner, "Calculating Bar-Natan's characteristic two Khovanov homology", *J. Knot Theory Ramifications* **15**:10 (2006), 1335–1356. MR

Received September 8, 2025. Revised April 3, 2026.

QIUYU REN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA, BERKELEY
BERKELEY, CA
UNITED STATES
qiuyu_ren@berkeley.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 344, No. 1, 2026
https://doi.org/10.2140/pjm.2026.344.151

ON THE INVARIANTS OF FINITE GROUPS ARISING IN A TOPOLOGICAL QUANTUM FIELD THEORY

CHRISTOPHER A. SCHROEDER AND HUNG P. TONG-VIET

We investigate structural properties of finite groups that are detected by certain group invariants arising from Dijkgraaf–Witten theory, a topological quantum field theory, in one space and one time dimension. In this setting, each finite group G determines a family of numerical invariants associated with closed orientable surfaces, expressed in terms of the degrees of the complex irreducible characters of G. These invariants can be viewed as natural extensions of the commuting probability $d(G)$, which measures the probability that two randomly chosen elements of G commute and has been extensively studied in the literature. By analyzing these higher-genus analogues, we establish new quantitative criteria relating the values of these invariants to key structural features of finite groups, such as commutativity, nilpotency, supersolvability and solvability. Our results generalize several classical theorems concerning the commuting probability, thereby linking ideas from finite group theory and topological quantum field theory.

1. Introduction

There is a long history of using invariants to deduce structural properties of finite groups. These invariants are often constructed by counting objects associated with the group, such as elements, conjugacy classes or irreducible characters. In this paper, we take a different approach. Given the remarkable compatibility between mathematics and physics, it is reasonable to expect that group invariants arising naturally in physics are useful for characterizing group structure. Here, we consider Dijkgraaf–Witten theory, a topological quantum field theory (TQFT), in one space dimension and one time dimension. In short, such a TQFT determines a commutative Frobenius algebra that gives rise to invariants of surfaces. If we assume that this algebra is the center $\boldsymbol{Z}(\mathbb{C}G)$ of the complex group algebra of a finite group G, then the invariant of a genus-h surface takes the form

$$\mathcal{Q}_h(G) = \sum_{\chi \in \mathrm{Irr}(G)} \left(\frac{|G|}{\chi(1)}\right)^{2h-2},$$

MSC2020: primary 20C15, 20E45; secondary 20C20, 20D10, 20D15, 20D20.
Keywords: character degree, conjugacy class, invariant, commuting probability, TQFT.

© 2026 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

where $\mathrm{Irr}(G)$ is the collection of complex irreducible characters of G. Note that $\mathcal{Q}_0(G) = 1/|G|$ and $\mathcal{Q}_1(G) = k(G)$, where $k(G)$ is the number of conjugacy classes of G (equivalently, the number $|\mathrm{Irr}(G)|$ of irreducible complex characters). If we consider the invariants $\mathcal{Q}_h(G)$ as arising from a sequence of increasingly complicated experiments, we might imagine that we first measure $\mathcal{Q}_0(G)$, then $\mathcal{Q}_1(G)$, and so on. With this in mind, it is natural to consider the following scaled invariants, which, as we will show, are very well-behaved:

$$q_h(G) := \frac{1}{|G|} \sum_{\chi \in \mathrm{Irr}(G)} \left(\frac{1}{\chi(1)}\right)^{2h-2}.$$

Now note that $q_0(G) = 1$ and $q_1(G) = k(G)/|G|$, where the latter is the so-called commuting probability $d(G)$, which was first considered by Joseph [20; 21], then by Gustafson [15] and many others. Viewed in this way, the "quantum invariants" $q_h(G)$ are generalizations of the commuting probability, for which the following structure criteria are known:

(a) If $d(G) > d(\mathrm{D}_8) = \frac{5}{8}$, then G is abelian (Gustafson [15]).

(b) If $d(G) > d(\mathrm{S}_3) = \frac{1}{2}$, then G is nilpotent (Lescot [23]).

(c) If $d(G) > d(\mathrm{A}_4) = \frac{1}{3}$, then G is supersolvable (Barry, MacHale, Ní Shé [3]).

(d) If $d(G) > d(\mathrm{A}_5) = \frac{1}{12}$, then G is solvable (Dixon [10]).

In this paper, we investigate which structural properties of finite groups are witnessed by the invariants $q_h(G)$. Our main results establish broad generalizations of the classical structure criteria above to arbitrary genus $h \geq 1$. The bounds in our first theorem are clearly best possible.

Theorem 1.1. *Let G be a finite group, and let h be a positive integer.*

(a) *If $q_h(G) > q_h(\mathrm{D}_8)$, then G is abelian.*

(b) *If $q_h(G) > q_h(\mathrm{S}_3)$, then G is nilpotent.*

(c) *If $q_h(G) > q_h(\mathrm{A}_4)$, then G is supersolvable.*

(d) *If $q_h(G) > q_h(\mathrm{A}_5)$, then G is solvable.*

Guralnick and Robinson showed in [14, Lemma 2(x)] that if p is a prime and $d(G) > 1/p$, then G has a normal Sylow p-subgroup. Our next result generalizes this p-closed criterion to all genera.

Theorem 1.2. *Let G be a finite group, let p be a prime, and let h be a positive integer. If $q_h(G) > \gamma(h, p)$, where $\gamma(h, p) = (1 + 1/p^{2h-1})/(p+1)$, then G has a normal Sylow p-subgroup.*

This criterion is the best possible. For example, if $p = 2^f - 1$ is a Mersenne prime, then the Frobenius group $G = (C_2)^f \rtimes C_p$ satisfies $q_h(G) = \gamma(h, p)$ and does not have a normal Sylow p-subgroup. Although Theorem 1.2 implies Theorem 1.1(b) and much of part (c), we will in fact use the latter to prove this p-closed criterion.

Given a prime p, it is also natural to consider a p-local version of the invariants $q_h(G)$ by summing over irreducible p-Brauer characters instead of irreducible complex characters. We define

$$q_{h,p'}(G) = \frac{1}{|G|_{p'}} \sum_{\varphi \in \mathrm{IBr}(G)} \frac{1}{\varphi(1)^{2h-2}}.$$

Note that $q_{1,p'}(G) = k_{p'}(G)/|G|_{p'}$, where $k_{p'}(G)$ is the number of irreducible p-Brauer characters (equivalently, the number of p-regular conjugacy classes). This invariant is denoted by $d_{p'}(G)$ in the literature, and it was shown in [33, Theorem 1.1] that if p is an odd prime and $d_{p'}(G) > 1/(p-1)$, then G is p-solvable. This result is best possible for $p > 3$ since $d_{p'}(\mathrm{PSL}_2(p)) = 1/(p-1)$. Our next theorem extends this result to all values of the genus.

Theorem 1.3. *Let G be a finite group, h a positive integer, and p an odd prime. If $q_{h,p'}(G) > \alpha(h, p)/(p-1)$, where $\alpha(h, p) = (2^{2h-2} + \sqrt{p-1})/(2^{2h-2}\sqrt{p-1})$, then G is p-solvable.*

Note that $\alpha(h, p) < 1$ when both $h > 1$ and $p > 2$. Theorem 1.3 is not best possible for $h = 1$, in which case, again, the best bound is $1/(p-1)$ from [33]. We suspect, but have not proved, that the result is also not best possible for $h > 1$. Our proof relies on a lower bound for $k_{p'}(G)$ when G is not p-solvable. At the time of writing, the best known bound is $k_{p'}(G) > \sqrt{p-1}$, which was proved by Hung and Maróti in [30]. An improvement in this bound would lead directly to an improvement in Theorem 1.3.

Finally, it is natural to consider an invariant $\tilde{q}_h(G)$ dual to $q_h(G)$ by summing over the sizes of the conjugacy classes $C \in \mathrm{Cl}(G)$ of a finite group G, where $\mathrm{Cl}(G)$ denotes the set of all conjugacy classes of G. We define the invariant

$$\tilde{q}_h(G) = \frac{1}{|G|} \sum_{C \in \mathrm{Cl}(G)} \frac{1}{|C|^{h-1}}.$$

Note that $\tilde{q}_0(G) = 1$ and $\tilde{q}_1(G) = d(G)$. From the definition, $\tilde{q}_h(G) \geq \tilde{q}_{h+1}(G)$ for all integers $h \geq 0$. We again prove generalizations of the structure criteria witnessed by the commuting probability to all values of the genus h.

Theorem 1.4. *Let G be a finite group, and let h be a positive integer.*

(a) *If $\tilde{q}_h(G) > \tilde{q}_h(\mathrm{D}_8)$, then G is abelian.*

(b) *If $\tilde{q}_h(G) > \tilde{q}_h(\mathrm{S}_3)$, then G is nilpotent.*

(c) *If* $\tilde{q}_h(G) > \tilde{q}_h(\mathrm{A}_4)$, *then* G *is supersolvable.*

(d) *If* $\tilde{q}_h(G) > \tilde{q}_h(\mathrm{A}_5)$, *then* G *is solvable.*

We also prove a dual criterion for the existence of a normal Sylow p-subgroup.

Theorem 1.5. *Let* G *be a finite group, let* p *be a prime and let* h *be a positive integer. If* $\tilde{q}_h(G) > \tilde{\gamma}(h, p)$, *where* $\tilde{\gamma}(h, p) = (1+1/p^{h-1}+(p-1)/(p+1)^{h-1})/(p(p+1))$, *then* G *has a normal Sylow* p*-subgroup.*

This criterion is best possible, as again the Frobenius group $G = (C_2)^f \rtimes C_p$ with $p = 2^f - 1$ a Mersenne prime satisfies $\tilde{q}_h(G) = \tilde{\gamma}(h, p)$.

We conclude this Introduction with some remarks.

Remark 1.6. The content of Theorems 1.1(a)–(d) and 1.4(a)–(d) is that the group-theoretic properties of commutativity, nilpotency, supersolvability and solvability are witnessed by the value of $q_h(G)$ or $\tilde{q}_h(G)$ for a particular "minimal" group G for all $h \geq 1$. For example, Theorem 1.1(d) says that solvability is witnessed by A_5 for all values of the genus h. Of course, our theorems would have trivial content if the invariants preserved the ordering of finite groups for all h; that is, if for all finite groups G and H it were true that $q_h(G) > q_h(H)$ for some $h > 1$ implied that $d(G) > d(H)$, then all our theorems would follow trivially from the corresponding result for $d(G)$. But this is not the case. For example, if $G = \mathrm{S}_6$ and $H = \mathrm{PGL}_2(9)$, then $q_h(G) > q_h(H)$ and $\tilde{q}_h(G) > \tilde{q}_h(H)$ for all $h > 1$, but $d(G) = d(H)$.

Our next example illustrates a general phenomenon: Notice that for a finite group G, the invariant $q_h(G)$ approaches $1/|G'|$ as h approaches infinity, where G' denotes the derived subgroup of G. Therefore, if $d(G) > d(H)$ for some finite groups G and H, but $1/|G'| < 1/|H'|$, then there must exist some positive integer N such that $q_h(G) < q_h(H)$ for all $h > N$. For example, let $G = \mathrm{A}_5$, and let H be an extraspecial p-group of order p^3 for the prime $p = 59$. Then $d(G) = \frac{1}{12}$ and $1/|G'| = \frac{1}{60}$, while $d(H) = (p^2+p-1)/p^3$ and $1/|H'| = 1/p$. Then $d(G) > d(H)$, while $1/|G'| < 1/|H'|$, and the preceding observation applies. In fact, a calculation shows that $q_h(G) > q_h(H)$ for $h = 1, 2, 3$, while $q_h(G) < q_h(H)$ for $h \geq 4$.

For the dual invariant, take $G = \mathrm{A}_5$ and let H be an extraspecial p-group of order 11^3. Then $d(G) < d(H)$ yet $\tilde{q}_h(G) > \tilde{q}_h(H)$ for all integers $h \geq 2$.

Remark 1.7. By examining our proofs, it is easy to see that our results hold not only when the exponent in $q_h(G)$ is the negative of the Euler characteristic $2-2h$ of a genus-h surface, but for any real number $s \geq 1$. In this way, our results are related to the so-called zeta function $\zeta_G(s) = \sum_{\chi \in \mathrm{Irr}(G)} \chi(1)^{-s}$ of a finite group G, which was used by Liebeck and Shalev to prove various results on Fuchsian groups, coverings of Riemann surfaces, subgroup growth, random walks and representation varieties [24; 25; 26]. We also note that the zeta function for conjugacy class sizes was studied in [25] for finite groups of Lie type, with applications to base

sizes of primitive permutation groups. Since our motivation stems from topological quantum field theories, we have stated and proved our results in this special case.

Remark 1.8. It is known that $d(G) \leq d(N)d(G/N)$ when $N \trianglelefteq G$ [14, Lemma 2(ii)]. Our proofs would be simplified if $q_h(G) \leq q_h(N)q_h(G/N)$ for all h. We have not been able to prove this statement, however. If it were true, then it must of course be true for $N = \boldsymbol{Z}(G)$. But this already seems to be a hard question about projective character degrees: letting $Z = \boldsymbol{Z}(G)$,

$$q_h(G) = \frac{1}{|G/Z||Z|} \sum_{\lambda \in \mathrm{Irr}(Z)} \sum_{\chi \in \mathrm{Irr}(G|\lambda)} \frac{1}{\chi(1)^{2h-2}}.$$

So to prove that $q_h(G) \leq q_h(G/Z)$, it would suffice to show, for example, that

$$\sum_{\chi \in \mathrm{Irr}(G|\lambda)} \frac{1}{\chi(1)^{2h-2}} \leq \sum_{\chi \in \mathrm{Irr}(G/Z)} \frac{1}{\chi(1)^{2h-2}};$$

that is, we want to show the sum over projective character degrees of G/Z is less than the sum over ordinary character degrees. We know that they are fewer in number: $|\mathrm{Irr}(G|\lambda)| \leq k(G/Z)$ as a consequence of Gallagher's result on "θ-good" conjugacy classes (see Theorem 1.19 and Corollary 1.23 in [19]).

Remark 1.9. The ratios $|G|/\chi(1)$ for $\chi \in \mathrm{Irr}(G)$ are the character degree quotients originally considered by Chillag and Herzog in [5]. In order to be compatible with group quotients, many authors have considered instead the character codegrees $[G : \ker(\chi)]/\chi(1)$. Note that if G is simple, then the character codegrees and character degree quotients of nontrivial characters coincide.

Remark 1.10. Our "dual invariant" $\tilde{q}_h(G)$ also has a physical interpretation in the context of lattice Hamiltonians for Yang–Mills gauge theories with finite gauge group [27]. These theories are motivated by the potential for quantum simulation of gauge theories.

Outline. Our paper is structured as follows. As our methods are purely group-theoretic, we defer an introduction to Dijkgraaf–Witten theory to Section 6. In Section 2, we make some remarks and collect some known facts that will be needed in the rest of the paper. In Section 3, we prove Theorems 1.1(a)–(d) and Theorem 1.3. In Section 4, we prove Theorem 1.2. In Section 5 we prove Theorems 1.4(a)–(d) and Theorem 1.5. Only the proof of Theorem 1.3 requires us to cite results that depend on the classification of finite simple groups. We do use Brauer's $k(GV)$-theorem in the proof of Theorem 1.1(c), but there the group G is solvable, and the proof in that case does not rely on the classification. All groups considered in this paper are assumed to be finite.

Notation. Our notation is standard. We follow [18] for finite group theory and [17] for the character theory of finite groups. Let n be a positive integer. We denote by A_n and S_n the alternating and symmetric groups of degree n, respectively. The dihedral group of order $2n$ is denoted by D_{2n}, and a cyclic group of order n by C_n. The special linear, projective general linear, and projective special linear groups of degree n over the finite field of size q are denoted by $\mathrm{SL}_n(q)$, $\mathrm{PGL}_n(q)$, and $\mathrm{PSL}_n(q)$, respectively. For a finite group G and a subset $X \subseteq G$, we write $\boldsymbol{C}_G(X)$ and $\boldsymbol{N}_G(X)$ for the centralizer and normalizer of X in G, respectively. The derived subgroup of G is denoted by G', and $\boldsymbol{Z}(G)$ denotes the center of G. The set of complex irreducible characters of G is denoted by $\mathrm{Irr}(G)$, and we let $k(G)$ be the number of conjugacy classes of G. We will repeatedly use that $k(G) = |\mathrm{Irr}(G)|$. We denote by $\mathrm{Cl}(G)$ the set of all conjugacy classes of G. If $N \leq G$ and $\lambda \in \mathrm{Irr}(N)$, then $\mathrm{Irr}(G|\lambda)$ denotes the set of all irreducible characters of G lying above λ. If p is a prime, the set of irreducible p-Brauer characters of G is denoted by $\mathrm{IBr}(G)$. All other notation is standard.

2. Preliminaries

As noted in the introduction, the invariants $q_h(G)$ can be viewed as generalizations of the commuting probability $d(G) = k(G)/|G|$ first considered by Joseph [20; 21], then by Gustafson [15]. Our proofs are based on the extensive literature on this invariant. Our first lemma collects some of the facts that are already known.

Lemma 2.1. *Let G be a finite group.*

(a) *We have $q_1(G) = d(G)$ and $q_h(G) \geq q_{h+1}(G)$ for all $h \geq 0$.*

(b) *If G is perfect and $d(G) > \frac{1}{20}$, then $G \cong \mathrm{A}_5$ or $G \cong \mathrm{SL}_2(5)$.*

(c) *If p is a prime such that a Sylow p-subgroup of G is not normal, then $d(G) \leq 1/p$.*

(d) *If G is a nonabelian p-group for some prime p, then $d(G) \leq 1/p + 1/p^2$.*

(e) *If G is nilpotent with a nonabelian Sylow p-subgroup for some prime p, then $d(G) < 1/p$ or $|G'| = p$.*

(f) *Whenever $N \trianglelefteq G$, we have $d(G) \leq d(G/N)d(N)$. In particular, we always have $d(G) \leq d(G/N)$.*

Proof. Part (a) follows from the definition using $|\mathrm{Irr}(G)| = k(G)$ and $\chi(1) \geq 1$ for all $\chi \in \mathrm{Irr}(G)$. Part (b) is [6, Proposition 1.6]. Part (c) is [14, Lemma 2(x)]. Part (d) is [14, Lemma 2(viii)]. Part (e) is [14, Lemma 2(ix)] and part (f) is [14, Lemma 2(ii)]. □

Our next lemma shows that $q_h(G)$ is monotone with respect to subgroups $H \leq G$ for all $h \geq 1$.

Lemma 2.2. *Let G be a finite group, and let h be any positive integer. If $H \leq G$, then $q_h(H) \geq q_h(G)$.*

Proof. We have

$$q_h(G) = \frac{1}{|G|} \sum_{\chi \in \mathrm{Irr}(G)} \frac{1}{\chi(1)^{2h-2}} \leq \frac{1}{|G|} \sum_{\psi \in \mathrm{Irr}(H)} \sum_{\chi \in \mathrm{Irr}(G|\psi)} \frac{1}{\chi(1)^{2h-2}}$$
$$\leq \frac{1}{|G|} \sum_{\psi \in \mathrm{Irr}(H)} \frac{[G:H]}{\psi(1)^{2h-2}} = \frac{1}{|H|} \sum_{\psi \in \mathrm{Irr}(H)} \frac{1}{\psi(1)^{2h-2}} = q_h(H).$$

The first inequality follows because we have enumerated all the $\chi \in \mathrm{Irr}(G)$ (possibly multiple times). The characters in $\mathrm{Irr}(G|\psi)$ are the constituents of ψ^G by Frobenius reciprocity, and $\psi^G(1) = [G:H]\psi(1)$. The sum $\sum_{\chi \in \mathrm{Irr}(G|\psi)} 1/\chi(1)^{2h-2}$ is as large as possible when all the degrees are as small as possible and $|\mathrm{Irr}(G|\psi)|$ is as large as possible. Since ψ lies under χ, $\chi(1) \geq \psi(1)$. If equality holds for all $\chi \in \mathrm{Irr}(G|\psi)$, then $\chi(1)$ is as small as possible and $|\mathrm{Irr}(G|\psi)| = [G:H]$ is as large as possible, which yields the second inequality. □

The next result slightly improves [14, Lemma 2(x)] for nonabelian simple groups.

Lemma 2.3. *Let G be a finite nonabelian simple group, and let p be a prime dividing $|G|$. Then $d(G) \leq 1/(p+1)$.*

Proof. Suppose for a contradiction that $d(G) > 1/(p+1)$. Since G is a finite nonabelian simple group, by [10] we know that $d(G) \leq \frac{1}{12}$. Hence $1/(p+1) < \frac{1}{12}$, which implies that $p \geq 13$ as p is a prime. Let P be a Sylow p-subgroup of G. Suppose that $[G : \boldsymbol{N}_G(P)] \leq p$. Since $[G : \boldsymbol{N}_G(P)]$ is coprime to p, we deduce that $[G : \boldsymbol{N}_G(P)] \leq p-1$. Since G is nonabelian simple, the core of $\boldsymbol{N}_G(P)$ in G is trivial, so G embeds into S_{p-1}, the symmetric group of degree $p-1$. But then $|G|$ is not divisible by p, which is a contradiction. Thus $[G : \boldsymbol{N}_G(P)] \geq p+1$. By Burnside's normal p-complement theorem [16, Theorem IV.2.6], we have $\boldsymbol{N}_G(P) \neq \boldsymbol{C}_G(P)$ and hence $|\boldsymbol{N}_G(P)| \geq 2p$. Combining these two inequalities, we obtain

$$|G| = [G : \boldsymbol{N}_G(P)] \cdot |\boldsymbol{N}_G(P)| \geq 2p(p+1).$$

Let d be the minimal degree of a nontrivial complex irreducible character of G. By [12, Theorem 1], we have $d \geq (p-1)/2$. By [14, Lemma 2(vi)], noting that $G' = G$, we have

$$d(G) < \frac{1}{d^2} + \Big(1 - \frac{1}{d^2}\Big)\frac{1}{|G'|} \leq \frac{4}{(p-1)^2} + \frac{1}{2p(p+1)}.$$

Since $d(G) > 1/(p+1)$, it follows that

$$\frac{1}{p+1} < \frac{4}{(p-1)^2} + \frac{1}{2p(p+1)}.$$

Simplifying this inequality gives $2p^3 - 13p^2 - 4p < 1$, i.e., $2p^3 - 13p^2 - 4p \leq 0$. This can be rewritten as $(2p-13)p \leq 4$ which is impossible for $p \geq 13$. Hence, the assumption $d(G) > 1/(p+1)$ leads to a contradiction, and the lemma follows. □

We also need the following easy lemma due to Lescot [23], whose proof we include for completeness.

Lemma 2.4. *If $d(G) > \frac{1}{4}$, then*

$$|G'| \leq \frac{3}{4d(G)-1}.$$

Proof. Since $|G| = \sum_{\chi \in \mathrm{Irr}(G)} \chi(1)^2$ and the number of linear characters is $[G : G']$, we have

$$|G| - [G : G'] = \sum_{\chi(1)>1} \chi(1)^2 \geq 4(k(G) - [G : G']).$$

Dividing both sides by $|G|$ and rearranging yields the claim. □

Remark 2.5. We will repeatedly use the fact that if the smallest degree of a nonlinear irreducible character of G is n, then $d(G)$ and $q_h(G)$ can be bounded in terms of n and $|G'|$, with equality if all nonlinear characters have degree n:

$$q_h(G) \leq \frac{1}{|G|}\Big([G : G'] + \frac{k(G) - [G : G']}{n^{2h-2}}\Big) = \frac{1}{|G'|}\Big(1 - \frac{1}{n^{2h-2}}\Big) + \frac{d(G)}{n^{2h-2}}.$$

We will derive similar inequalities for $\tilde{q}_h(G)$ in the course of our proofs.

We now turn to the properties of a finite group G that are witnessed by the dual invariant $\tilde{q}_h(G) = 1/|G| \sum_{C \in \mathrm{Cl}(G)} 1/|C|^{h-1}$, where $\mathrm{Cl}(G)$ is the collection of conjugacy classes of G. Whereas we showed that $q_h(G)$ is monotone with respect to subgroups in Lemma 2.2, we now prove, in a nicely dual result, that $\tilde{q}_h(G)$ is monotone with respect to quotients.

Lemma 2.6. *Let G be a finite group, let $N \trianglelefteq G$, and let h be any positive integer. Then $\tilde{q}_h(G) \leq \tilde{q}_h(G/N)$.*

Proof. Let $\overline{G} = G/N$ and use the bar notation. Let $\pi : G \to G/N$ denote the canonical homomorphism. Given a conjugacy class $D \in \mathrm{Cl}(\overline{G})$, the preimage $\pi^{-1}(D) \subseteq G$ is closed under conjugation, so it is a union of conjugacy classes C_i of G, say

$$\pi^{-1}(D) = C_1 \cup C_2 \cup \cdots \cup C_k.$$

Let $S(D) = \{C_1, \ldots, C_k\}$. Choosing some $x_i \in C_i$ for $i = 1, \ldots, k$, we have $|C_i| = [G : \boldsymbol{C}_G(x_i)]$. Similarly, $|D| = [\overline{G} : \boldsymbol{C}_{\overline{G}}(\bar{x}_i)]$ for all $i = 1, \ldots, k$ since $\pi(x_i) \in D$. Writing $\boldsymbol{C}_{\overline{G}}(\bar{x}_i) = H/N$ for some subgroup $H \leq G$, we have $|D| = [G : H]$. Note

that $\boldsymbol{C}_G(x_i) \leq H$. Therefore,

$$|C_i| = [G : \boldsymbol{C}_G(x_i)] = [G : H][H : \boldsymbol{C}_G(x_i)] \geq |D|.$$

Furthermore, $|\pi^{-1}(D)| = |D||N|$ since π is an $|N|$-to-1 map. Since $|C_i| \geq |D|$, we have

$$|N||D| = \sum_{i=1}^{k} |C_i| \geq k|D|,$$

and so $k \leq |N|$.

Each conjugacy class of G lies in the preimage of exactly one conjugacy class of $\overline{G}$. Now we calculate

$$\tilde{q}_h(G) = \frac{1}{|G|} \sum_{D \in \mathrm{Cl}(\overline{G})} \sum_{C \in S(D)} \frac{1}{|C|^{h-1}} \leq \frac{1}{|G|} \sum_{D \in \mathrm{Cl}(\overline{G})} \frac{|N|}{|D|^{h-1}} = \tilde{q}_h(\overline{G}),$$

which is what we wanted to show. □

The following easy lemma is our main tool for proving the structure criteria in Theorem 1.4.

Lemma 2.7. *Let G be a finite group, and let n be the smallest size of the noncentral conjugacy classes of G. Let $h \geq 2$ be an integer, and assume that G is not abelian. Let N be an integer such that $d(G)/\tilde{q}_h(G) \leq N$. Then either*

(1) $n^{h-1} \leq N$, *or*

(2) $[G : \mathbf{Z}(G)] \leq \dfrac{N}{\tilde{q}_h(G)(N+1) - d(G)} \leq \dfrac{d(G)}{\tilde{q}_h(G)^2}$.

Proof. Since G is not abelian, $\tilde{q}_h(G) < d(G)$. If $n^{h-1} \leq N$, then part (1) is satisfied. Assume that $n^{h-1} > N$. So $n^{h-1} > N \geq d(G)/\tilde{q}_h(G)$. It follows that $\tilde{q}_h(G)n^{h-1} - d(G) > 0$ and

$$n^{h-1} \geq N + 1 \geq d(G)/\tilde{q}_h(G) + 1.$$

We have

$$\tilde{q}_h(G) \leq \frac{1}{|G|}\Big(|\mathbf{Z}(G)| + \frac{k(G) - |\mathbf{Z}(G)|}{n^{h-1}}\Big).$$

Using that $d(G) = k(G)/|G|$ and rearranging, we find that

$$[G : \mathbf{Z}(G)] \leq \frac{n^{h-1} - 1}{\tilde{q}_h(G)n^{h-1} - d(G)}.$$

Since the right-hand side is a decreasing function of n^{h-1}, we have

$$[G : \mathbf{Z}(G)] \leq \frac{N}{\tilde{q}_h(G)(N+1) - d(G)} \leq \frac{d(G)}{\tilde{q}_h(G)^2},$$

as claimed. □

3. Commutativity, nilpotency, supersolvability and solvability criteria

In this section, we prove the four claims of Theorem 1.1.

Proof of the commutativity criterion, Theorem 1.1(a). Suppose for a contradiction that G is a counterexample to the claim with minimal order. The degrees of the complex irreducible characters of D_8 are 1, 1, 1, 1 and 2. By a result due to Gustafson [15], $d(G) \leq \frac{5}{8}$ since G is nonabelian. Altogether, we have

$$\tfrac{1}{2} < \tfrac{1}{2}(1 + 2^{-2h}) = q_h(\mathrm{D}_8) < q_h(G) \leq d(G) \leq \tfrac{5}{8}.$$

We first claim that G is nilpotent; that is, all Sylow subgroups of G are normal in G. By Lemma 2.1(c), if a Sylow p-subgroup of G is not normal for some prime p, then $d(G) \leq 1/p$. But then $\frac{1}{2} < \frac{1}{p}$, which is impossible.

Next, we claim that all Sylow subgroups of odd order are abelian. Note that if $G = A \times B$, then $q_h(G) = q_h(A)q_h(B)$. Since G is nilpotent by the last paragraph, this means that $q_h(G) = \prod q_h(P)$, where the product runs over the set of Sylow subgroups of G. In particular, $q_h(P) > \frac{1}{2}$ for every Sylow subgroup P. However, by Lemma 2.1(d), if P is nonabelian then $d(P) < 1/p + 1/p^2$. Thus, $1/p + 1/p^2 > \frac{1}{2}$. This inequality only holds if $p = 2$, so all Sylow subgroups of odd order are abelian.

We may now assume that G is a nonabelian 2-group such that $q_h(G) > q_h(\mathrm{D}_8)$. Now, by Lemma 2.1(e), either $d(G) < \frac{1}{2}$ or $|G'| = 2$. We have $d(G) > \frac{1}{2}$ by the first paragraph, so $|G'| = 2$. Therefore,

$$q_h(\mathrm{D}_8) < q_h(G) \leq \tfrac{1}{2}(1 - 2^{-(2h-2)}) + d(G) \cdot 2^{-(2h-2)},$$

where we used that nonlinear irreducible characters of G have degree at least 2 and the observation in Remark 2.5. Thus,

$$d(G) > \tfrac{1}{2} + 2^{2h-2}\big(q_h(\mathrm{D}_8) - \tfrac{1}{2}\big) = \tfrac{1}{2} + 2^{2h-2} \cdot 2^{-(2h+1)} = \tfrac{1}{2} + \tfrac{1}{8} = \tfrac{5}{8}.$$

This contradicts the bound $d(G) \leq \frac{5}{8}$ from the first paragraph, proving the result. □

Next, we prove Theorem 1.1(b). We will show that a minimal counterexample is a *minimal nonnilpotent group*; that is, a nonnilpotent group whose proper subgroups are nilpotent. The structure of such groups, which we summarize in the following lemma, is well understood:

Lemma 3.1 [16, Theorem III.5.2]. *Suppose G is a minimal nonnilpotent group. Then $G \cong Q \rtimes P$, where Q is a Sylow q-subgroup for some prime q, while P is a cyclic Sylow p-subgroup for some prime p and $\Phi(P) \leq \mathbf{Z}(G)$.*

Proof of the nilpotency criterion, Theorem 1.1(b). Suppose for a contradiction that G is a counterexample to the claim with minimal order. The irreducible character degrees of S_3 are 1, 1 and 2. Since G is not nilpotent, it contains a nonnormal Sylow p-subgroup for some prime p dividing $|G|$. This means $d(G) \leq 1/p \leq \frac{1}{2}$ by

Lemma 2.1(c). Altogether, we have

$$\tfrac{1}{3} < \tfrac{1}{3}(1+2^{-(2h-1)}) = q_h(\mathrm{S}_3) < q_h(G) \le d(G) \le \tfrac{1}{p} \le \tfrac{1}{2}.$$

This means that the only nonnormal Sylow subgroup is the Sylow 2-subgroup.

Now, by the monotonicity of Lemma 2.2, every proper subgroup of G is nilpotent, so G is a minimal nonnilpotent group by induction and has the structure described in Lemma 3.1. We know that $p = 2$ by the last paragraph, and $q \neq 2$ since G is not nilpotent. As P is cyclic, $[P : \Phi(P)] = 2$.

Consider the structure of G': As G is not nilpotent, of course $G' > 1$. As $d(G) > \frac{1}{3}$, we have $|G'| < 9$ by Lemma 2.4. Since $G/Q \cong P$ is abelian, $G' \le Q$. So $q = 3$, 5 or 7 and G' is cyclic of prime order q. Since $G' \trianglelefteq Q$, we have $G' \cap \boldsymbol{Z}(Q) > 1$, so $G' \le \boldsymbol{Z}(Q)$ as G' has prime order. So Q centralizes G'. If $G' < Q$, then $G'P < G$ is nilpotent by induction on the group order, so P centralizes G', as well. But then $G' \le \boldsymbol{Z}(G)$ and G is nilpotent, a contradiction. So we have that $G' = Q$ is cyclic of order 3, 5 or 7.

Since $\Phi(P) \le \boldsymbol{Z}(G)$, the product $Q\Phi(P)$ is abelian of index 2 in G. So every nonlinear irreducible character of G has degree 2 since it is a constituent of a character induced from a (necessarily linear) character of $Q\Phi(P)$. As G has $[G : G'] = |P|$ linear characters, the remaining nonlinear characters must be $\frac{1}{4}(|Q|-1)|P|$ in number since the resulting sum of squared degrees is $|P| + \frac{1}{4}(|Q|-1)|P| \cdot 2^2 = |G|$. Therefore,

$$q_h(G) = \frac{1}{|G|}\left(|P| + \frac{(|Q|-1)|P|}{4}\frac{1}{2^{2h-2}}\right) = \frac{1}{q}\left(1 + \frac{q-1}{2^{2h}}\right) > \frac{1}{3}\left(1+\frac{2}{2^{2h}}\right) = q_h(\mathrm{S}_3).$$

But the inequality implies that $q < 3$. We have reached our final contradiction and proved the claim. □

To prove the supersolvability criterion of Theorem 1.1(c), we will show that a minimal counterexample is a *minimal nonsupersolvable group*; that is, a nonsupersolvable group whose proper subgroups are supersolvable. The structure of such groups was determined by Doerk [11], and can be read off from [2, Theorem 12] and [16, Chapter VI, Exercise 16]. We summarize the results we will initially need:

Lemma 3.2 [11]. *Suppose G is a minimal nonsupersolvable group. Then G has exactly one normal Sylow p-subgroup P and a complement Q such that $|Q|$ is divisible by at most two distinct primes. Furthermore, $P/\Phi(P)$ is a minimal normal subgroup of $G/\Phi(P)$ and is not cyclic, and $Q \cap \boldsymbol{C}_G(P/\Phi(P)) = Q \cap \Phi(G) = \Phi(Q) \cap \Phi(G)$.*

Proof of the supersolvability criterion, Theorem 1.1(c). Let G be a counterexample to the theorem with minimal order. The irreducible character degrees of A_4 are

1, 1, 1 and 3. Since G is not supersolvable, $d(G) \leq \frac{1}{3}$ by [3]. Altogether, we have

$$\tfrac{1}{4} < \tfrac{1}{4}(1 + 3^{-(2h-1)}) = q_h(\mathrm{A}_4) < q_h(G) \leq d(G) \leq \tfrac{1}{3}.$$

Then by Lemma 2.2, for every proper subgroup H of G, we have $q_h(H) \geq q_h(G) > q_h(\mathrm{A}_4)$ and hence by the minimality of $|G|$, H is supersolvable. As G is not supersolvable, G is a minimal nonsupersolvable group and has the structure described in Lemma 3.2. Now we work in $\overline{G} = G/\Phi(G)$, which is again minimal nonsupersolvable, since if $\overline{G}$ is supersolvable then G is supersolvable by [16, Theorem VI.8.5]. Recall that $\Phi(P) \leq \Phi(G)$ by [16, Lemma III.3.3]; since $P/\Phi(P)$ is minimal normal in $G/\Phi(P)$ by Lemma 3.2, $\Phi(P) = P \cap \Phi(G)$. Therefore, $\overline{P}$ is a minimal normal, elementary abelian, noncyclic subgroup of $\overline{G}$, where $\overline{G} = \overline{P}\,\overline{Q}$ and $\overline{Q}$ acts coprimely, faithfully and irreducibly on $\overline{P}$. By Brauer's $k(GV)$-theorem [13, Theorem], we know that $k(\overline{G}) = k(\overline{P} \rtimes \overline{Q}) \leq |\overline{P}|$. Using this and Lemma 2.1(f), we have

$$\frac{1}{4} < d(G) \leq d(\overline{G}) \leq \frac{1}{|\overline{Q}|},$$

and so $|\overline{Q}| \leq 3$. Since $\overline{Q}$ has prime order and $\overline{P}$ is an irreducible $\overline{Q}$-module, this means $\boldsymbol{C}_{\overline{G}}(\bar{x}) \leq \overline{P}$ for all nontrivial $\bar{x} \in \overline{P}$, so $\overline{G}$ is a Frobenius group. Now, if $\overline{Q} = \langle \bar{t} \rangle$ has order 2, then $\bar{t}$ acts as inversion on $\overline{P}$ (see [17, Lemma 7.21] and its proof). So every subgroup $\langle \bar{x} \rangle \leq \overline{P}$ is fixed by $\overline{Q}$, which means $\overline{P} = \langle \bar{x} \rangle$ is cyclic as it is an irreducible $\overline{Q}$-module. This contradiction implies that $|\overline{Q}| = 3$.

Now we count the characters of the Frobenius group $\overline{G}$: Since $\overline{G}' = \overline{P}$, there are $[\overline{G} : \overline{G}'] = |\overline{Q}|$ linear characters, and each of the nonlinear irreducible characters of $\overline{G}$ is induced from an irreducible character of $\overline{P}$ lying in one of the $(|\overline{P}| - 1)/|\overline{Q}|$ orbits under $\overline{Q}$. So

$$\frac{1}{4} < d(\overline{G}) = \frac{1}{|\overline{P}|\,|\overline{Q}|}\left(|\overline{Q}| + \frac{|\overline{P}| - 1}{|\overline{Q}|}\right).$$

Rearranging and using that $|\overline{Q}| = 3$, we find that $p^2 \leq |\overline{P}| < \frac{32}{5} < 7$. Therefore, $|\overline{P}| = 4$ and so $\overline{G} \cong \mathrm{A}_4$. In particular, P is a 2-group.

Let r be a prime divisor of $|G|$, and let R be a Sylow r-subgroup of G. If R is not normal in G, then $d(G) \leq \frac{1}{r}$. Since $d(G) \geq q_h(G) > q_h(\mathrm{A}_4) > \frac{1}{4}$ we deduce that $\frac{1}{r} > \frac{1}{4}$ which forces $r < 4$ and hence $r \in \{2, 3\}$. In particular, Q is a π-group, where $\pi \subseteq \{2, 3\}$. Since we showed that the Sylow p-subgroup P is a 2-group, Q must then be a 3-group. Now, since $\Phi(Q) \geq \Phi(Q) \cap \Phi(G) = Q \cap \boldsymbol{C}_G(P/\Phi(P))$ and $\overline{Q} \cong Q/(\Phi(Q) \cap \Phi(G))$ is cyclic of order 3, we have $\Phi(Q) = \Phi(Q) \cap \Phi(G)$. So Q is itself cyclic by Burnside's basis theorem [16, Theorem III.3.15]. We have shown that $G = PQ$, where P is a 2-group and Q is a cyclic 3-group, and $\overline{G} = G/\Phi(G) \cong \mathrm{A}_4$.

Now we turn to the more precise classification of minimal nonsupersolvable groups in [2, Theorem 10]. Since Q is cyclic of 3-power order and $p = 2$, inspecting the list of possibilities in [2, Theorem 9] shows that only Types 2 and 3 are possible. We inspect each of these in turn and derive contradictions:

Type 2: *$Q = \langle z \rangle$ is cyclic, P is an irreducible Q-module over the field of p elements with kernel $\langle z^q \rangle$ in Q.*

Since $P \trianglelefteq G$ is an irreducible Q-module, P has no Q-invariant nontrivial proper subgroup. In particular, $P \cap \Phi(G) = 1$ and so $\bar{P} \cong P$ is elementary abelian of order 4 by what we showed above. The kernel $K = \langle z^3 \rangle$ is central in G and has index 3 in Q, so $|Q| = 3|K|$ and $|G| = 12|K|$. We have $[G : G'] = [G : P] = |Q| = 3|K|$, and the nonlinear irreducible characters of G are induced from characters $\psi_1 \times \psi_2 \in \mathrm{Irr}(P \times K)$ with $\psi_1 \neq 1$, which split into $|K|$ orbits under G. Thus, each of these $|K|$ nonlinear characters of G has degree 3. So

$$q_h(G) = \frac{1}{12|K|}\left(3|K| + \frac{|K|}{3^{2h-2}}\right) = \frac{1}{4}\left(1 + \frac{1}{3^{2h-1}}\right) = q_h(\mathrm{A}_4),$$

which is a contradiction. Note that if $K = 1$, then $G \cong \mathrm{A}_4$.

Type 3: *P is a nonabelian special p-group of rank $2m$, the order of p modulo q being $2m$, $Q = \langle z \rangle$ is cyclic, z induces an automorphism on P such that $P/\Phi(P)$ is a faithful and irreducible Q-module, and z centralizes $\Phi(P)$. Furthermore, $|P/\Phi(P)| = p^{2m}$ and $|P'| \leq p^m$.*

Since $|\bar{P}| = 4$, we have $m = 1$. So $|P'| \leq 2$, which means P is an extraspecial 2-group of order 2^3. Since $P/\Phi(P)$ is a faithful Q-module, Q is cyclic of order 3. If P is an extraspecial 2-group of order 2^3, then P has 4 linear characters and 1 nonlinear irreducible character of degree 2. The nontrivial linear characters form one Q-orbit and induce irreducibly to the same nonlinear character of G of degree 3. The unique nonlinear character of P is necessarily G-invariant, and so extends to G since $[G : P] = 3$ is prime (see [17, Corollary 6.19]). This gives rise to $|\mathrm{Irr}(G/P)| = 3$ distinct irreducible characters of G of degree 2 by [17, Corollary 6.17]. There are $[G : G'] = [G : P] = 3$ linear characters. So

$$q_h(G) = \frac{1}{24}\left(3 + \frac{3}{2^{2h-2}} + \frac{1}{3^{2h-2}}\right) \leq \frac{1}{24}\left(6 + \frac{1}{3^{2h-2}}\right) = \frac{1}{4}\left(1 + \frac{1}{2 \cdot 3^{2h-1}}\right) < q_h(\mathrm{A}_4),$$

again a contradiction. The result is proved. □

Proof of the solvability criterion, Theorem 1.1(d). If $h = 1$, then $q_1(G) = d(G) > d(\mathrm{A}_5) = \frac{1}{12}$, and G is already known to be solvable by a result of Dixon [10]. So, now assume that $h \geq 2$, and suppose that G is a counterexample to the claim with minimal order. By Lemma 2.1(a), we have $d(G) \geq q_h(G) > q_h(\mathrm{A}_5) > 1/|\mathrm{A}_5| = \frac{1}{60}$.

We first note that $q_h(G') \geq q_h(G)$ by Lemma 2.2. So if $G' < G$, then G' is solvable by induction on the order of the group. But then G is also solvable, a

contradiction. So $G = G'$ is perfect and nontrivial irreducible characters have degree at least 2. Thus,

$$q_h(G) \le \frac{1}{|G|}\left(1 + \frac{k(G)-1}{2^{2h-2}}\right) = \frac{1}{|G|}\left(\frac{k(G)+2^{2h-2}-1}{2^{2h-2}}\right).$$

Next, we have the following inequalities when $k(G) \ge 5$:

$$\frac{k(G)+2^{2h-2}-1}{2^{2h-2}} \le \frac{2k(G)}{5} \qquad \text{for } h = 2,$$
$$\frac{k(G)+2^{2h-2}-1}{2^{2h-2}} \le \frac{k(G)}{3} \qquad \text{for } h > 2.$$

It is not difficult to see that if $k(G) < 5$, then $|G| \le 12$, and so G is solvable (see [4, Note A] or [34, Table 1]). Since G is nonsolvable, we must have $k(G) \ge 5$. Therefore, we have

$$d(G) = k(G)/|G| \ge \tfrac{5}{2} q_2(G) > \tfrac{5}{2} q_2(\mathrm{A}_5) = \tfrac{5}{2}\,\tfrac{4769}{216000} > \tfrac{1}{20},$$
$$d(G) = k(G)/|G| \ge 3q_h(G) > 3q_h(\mathrm{A}_5) > \tfrac{3}{60} = \tfrac{1}{20} \qquad \text{for } h > 2.$$

Therefore, G is perfect with $d(G) > \frac{1}{20}$ for all $h \ge 2$. Then by Lemma 2.1(b), $G \cong \mathrm{A}_5$ or $G \cong \mathrm{SL}_2(5)$. But

$$q_h(\mathrm{SL}_2(5)) = \frac{1}{120}\left(1 + \frac{2}{2^{2h-2}} + \frac{2}{3^{2h-2}} + \frac{2}{4^{2h-2}} + \frac{1}{5^{2h-2}} + \frac{1}{6^{2h-2}}\right)$$
$$< \frac{1}{120}\left(2 + \frac{4}{3^{2h-2}} + \frac{2}{4^{2h-2}} + \frac{2}{5^{2h-2}}\right) = q_h(\mathrm{A}_5).$$

This is our final contradiction, and Theorem 1.1 is proved. □

Next, the proof of Theorem 1.3 runs along the lines of that of Theorem 1.1(d). Although it may be true that $q_{h,p'}(G)$ is monotone with respect to subgroups, the proof of Lemma 2.2 will not work when the prime p divides $|G|$, because Frobenius reciprocity does not hold in this case. However, we can again show that a minimal counterexample to the theorem must be perfect by using some basic results from the theory of Brauer characters. Before we begin, recall that we defined $\alpha(h, p) = (2^{2h-2} + \sqrt{p-1})/(2^{2h-2}\sqrt{p-1})$.

Proof of Theorem 1.3. Suppose for a contradiction that G is a counterexample to the claim with minimal order. We first show that if $N \le G$ is a normal subgroup with index equal to any prime r, then $q_{h,p'}(G) \le q_{h,p'}(N)$. This will then imply that G is perfect, since if not, then there exists such an $N \trianglelefteq G$ with $\alpha(h, p)/(p-1) < q_{h,p'}(G) \le q_{h,p'}(N)$, so N is p-solvable by induction on the order of the group. But then G is also p-solvable since the quotient G/N is cyclic, which is a contradiction.

Now, each character $\varphi \in \mathrm{IBr}(G)$ lies above at least one character $\theta \in \mathrm{IBr}(N)$. Denoting by $\mathrm{IBr}(G|\theta)$ the collection of irreducible p-Brauer characters of G lying

over $\theta \in \mathrm{IBr}(N)$, we have

$$q_{h,p'}(G) \leq \frac{1}{|G|_{p'}} \sum_{\theta \in \mathrm{IBr}(N)} \sum_{\varphi \in \mathrm{IBr}(G|\theta)} \frac{1}{\varphi(1)^{2h-2}}.$$

Since $N \trianglelefteq G$ has prime index r, there are two possibilities for each $\theta \in \mathrm{IBr}(N)$:

(1) θ is G-invariant, in which case θ extends to an irreducible Brauer character $\hat{\theta}$ of G by [29, Theorem 8.12] since G/N is cyclic. In addition, the characters $\beta\hat{\theta}$ for $\beta \in \mathrm{IBr}(G/N)$ are all the irreducible constituents of θ^G by [29, Corollary 8.20]. Note that $\beta(1)\hat{\theta}(1) = \theta(1)$ since β is linear. Finally, since $N \trianglelefteq G$, $\varphi \in \mathrm{IBr}(G)$ is an irreducible constituent of θ^G if and only if θ is an irreducible constituent of φ_N by [29, Corollary 8.7]. Therefore, $|\mathrm{IBr}(G|\theta)| = |\mathrm{IBr}(G/N)| = [G : N]_{p'}$.

(2) The stabilizer of θ in G is N, in which case the induced character $\varphi = \theta^G$ is irreducible by Clifford's theorem [29, Theorem 8.9]. So $|\mathrm{IBr}(G|\theta)| = 1$, again by [29, Corollary 8.7], and $\varphi(1) > \theta(1)$.

Therefore, for each $\theta \in \mathrm{IBr}(N)$ we have

$$\sum_{\varphi \in \mathrm{IBr}(G|\theta)} \frac{1}{\varphi(1)^{2h-2}} \leq \frac{[G : N]_{p'}}{\theta(1)^{2h-2}},$$

and so

$$q_{h,p'}(G) \leq \frac{[G : N]_{p'}}{|G|_{p'}} \sum_{\theta \in \mathrm{IBr}(N)} \frac{1}{\theta(1)^{2h-2}} = q_{h,p'}(N),$$

which is what we wanted to show.

Since we now know that the counterexample G is perfect, all nontrivial irreducible Brauer characters are nonlinear, and we have

$$q_{h,p'}(G) \leq \frac{1}{|G|_{p'}} \left(1 + \frac{k_{p'}(G) - 1}{2^{2h-2}}\right) = \frac{1}{|G|_{p'}} \left(\frac{k_{p'}(G) + 2^{2h-2} - 1}{2^{2h-2}}\right),$$

where the inequality follows since $\varphi(1) \geq 2$ for all nonlinear $\varphi \in \mathrm{IBr}(G)$.

Now, since G is non-p-solvable, it is known that $k_{p'}(G) > \sqrt{p-1}$ by [30, Theorem 1.2]. Note that the following inequality is satisfied for $k_{p'}(G) > \sqrt{p-1}$ and each $h \geq 1$:

$$\frac{k_{p'}(G) + 2^{2h-2} - 1}{2^{2h-2}} < \alpha(h, p) k_{p'}(G).$$

To see this, multiply through by 2^{2h-2}, move all terms involving $k_{p'}(G)$ to the right-hand side, and use that $k_{p'}(G) > \sqrt{p-1}$. This implies that

$$d_{p'}(G) = \frac{k_{p'}(G)}{|G|_{p'}} > \frac{1}{\alpha(h, p)} q_{h,p'}(G) > \frac{1}{p-1}.$$

But this contradicts [33, Theorem 1.1], and the theorem is proved. □

4. Criterion for the existence of normal Sylow subgroups

In this section, we prove Theorem 1.2, which is a criterion for the existence of a normal Sylow subgroup. Let p be a prime, and let h be a positive integer. Recall that for a finite group G and a prime p, G is said to be p-closed if G has a normal Sylow p-subgroup. Define

$$\gamma(h, p) = \frac{1}{p+1}\Big(1 + \frac{1}{p^{2h-1}}\Big).$$

Note that $\gamma(h, p) > 1/(p+1)$. We now prove Theorem 1.2.

Proof of Theorem 1.2. Let G be a counterexample to the theorem with minimal order. Then $q_h(G) > \gamma(h, p)$ but G does not have a normal Sylow p-subgroup. By Lemma 2.1(c), we have $d(G) \leq 1/p$. Consequently,

$$\frac{1}{p} \geq d(G) \geq q_h(G) > \gamma(h, p) > \frac{1}{p+1}.$$

By Lemma 2.2, for each proper subgroup H of G, we have $q_h(H) \geq q_h(G) > \gamma(h, p)$ and thus by the minimality of $|G|$, H is p-closed. In particular, every proper subgroup of G is p-closed. Also, if N is a normal subgroup of G, then every proper subgroup of the quotient group G/N is p-closed.

Assume that $p = 2$. Then $\gamma(h, p) = q_h(\mathrm{S}_3)$ and since $q_h(G) > q_h(\mathrm{S}_3)$, G is nilpotent by Theorem 1.1(b). But then G has a normal Sylow 2-subgroup, a contradiction. Therefore, we may assume that $p \geq 3$.

Let P be a Sylow p-subgroup of G.

Claim 1: *$\langle P^G \rangle = G$. Hence, G has no proper normal subgroup of p'-index.*

Let $L = \langle P^G \rangle$. Then $P \leq L \trianglelefteq G$. Suppose that $L < G$. By the minimality of $|G|$, L has a normal Sylow p-subgroup, which implies that $P \trianglelefteq L$. Since P is characteristic in L, we deduce that $P \trianglelefteq G$, which is a contradiction. Therefore, $G = L$ as desired.

Claim 2: *G is not a nonabelian simple group.*

Since $d(G) > 1/(p+1)$, this follows from Lemma 2.3.

Claim 3: *Let N be a maximal normal subgroup of G. Then $G = \langle x \rangle N$ with $[G : N] = p$, where N has a normal Sylow p-subgroup, which is $O_p(G)$, and $x \in P$ is a p-element and $P = \langle x \rangle O_p(G)$. In particular, G is p-solvable with a Hall p'-subgroup R and $G = PR$ with $[P : O_p(G)] = p$.*

Since G is not a simple group, G has a maximal normal subgroup, say N. It follows that N has a normal Sylow p-subgroup, say P_1. As $P_1 \trianglelefteq N \trianglelefteq G$, we deduce that $P_1 \trianglelefteq G$ and thus $P_1 \leq O_p(G)$. Since G/N is a simple group and $d(G/N) \geq d(G) > 1/(p+1)$ and p divides $|G/N|$ by Claim 1, Lemma 2.3 yields

that G/N is abelian. Hence, G/N is cyclic of order p. It follows that $G = PN$ and $P \cap N = P_1 \trianglelefteq G$. As $[G : N] = [P : P_1] = p$ and P is not normal in G, we deduce that $P_1 = O_p(G)$. Let $x \in P \setminus P_1$. Then $P = \langle x \rangle P_1 = \langle x \rangle O_p(G)$. Since $N/O_p(G)$ is a p'-group and both $O_p(G)$ and G/N are p-groups, G is p-solvable. So G has a Hall p'-subgroup R and thus $N = RO_p(G)$ and $G = PN = PR$ with $[G : N] = [P : O_p(G)] = p$.

Claim 4: *R is a Sylow r-group of G for some prime r.*

Let $\overline{G} = G/O_p(G)$. Then $\overline{P} \cong C_p$ acts coprimely and nontrivially on $\overline{N} = \overline{R} \cong R$. Suppose for a contradiction that $|\overline{R}| = |R|$ is divisible by at least two distinct primes. Let q be a prime divisor of $|R|$. By the coprime action theorem [18, Theorem 3.23], $\overline{P}$ stabilizes a Sylow q-subgroup $\overline{Q}$ of $\overline{R}$. But then $\overline{P}\,\overline{Q}$ is a proper subgroup of $\overline{G}$, and hence it has a normal Sylow p-subgroup. Thus, $\overline{P}$ centralizes $\overline{Q}$. It follows that $|\overline{G} : \boldsymbol{C}_{\overline{G}}(\overline{P})|$ is not divisible by q. Clearly, this index is also not divisible by p. As this is true for all prime divisors of $|R|$, we must have $\overline{G} = \boldsymbol{C}_{\overline{G}}(\overline{P})$, which is a contradiction. Thus R is an r-group for some prime r, and so it is a Sylow r-subgroup of G.

Claim 5: *Let n be the smallest nontrivial character degree of G. Then $n \geq (p-1)/2$.*

Note that G is nonabelian and so such a minimal nontrivial degree n exists. Let $\chi \in \mathrm{Irr}(G)$ such that $\chi(1) = n$ and let K be the kernel of χ. Then K is a proper normal subgroup of G. We consider the following cases.

Case 1: G/K has a normal Sylow p-subgroup. Then $PK/K \trianglelefteq G/K$ as PK/K is a Sylow p-subgroup of G/K. By Claim 1, we have $G = PK$ and thus $G/K \cong P/P\cap K$, which is a p-group. Since $\chi \in \mathrm{Irr}(G/K)$ is nonlinear, G/K is a nonabelian p-group and thus $n = \chi(1) \geq p > (p-1)/2$.

Case 2: G/K does not have a normal Sylow p-subgroup. It follows that G/K has a faithful irreducible character of degree n and does not have a normal Sylow p-subgroup. By [12, Theorem 1], $p \leq 2n+1$ or $n \geq (p-1)/2$.

Claim 6: *$G' = RU$, where $U = G' \cap O_p(G) \trianglelefteq G$.*

Note that G is a $\{p, r\}$-group by Claim 4, so it is solvable and $G' \leq N \trianglelefteq G$. Since $\langle P^G \rangle = G$, we deduce that G/G' is an abelian p-group. Thus, G' contains a Sylow r-group of G. It follows that $R \leq G'$. By Dedekind's modular law, $G' = R(G' \cap O_p(G))$, where $G' \cap O_p(G)$ is a normal Sylow p-subgroup of G'.

Claim 7: *The quotient $G/R'O_p(G)$ is a Frobenius group with a cyclic complement of order p and Frobenius kernel an abelian 2-group of order 2^m for some $m \geq 2$, and $p = 2^m - 1$.*

Note that $N/O_p(G) = RO_p(G)/O_p(G) \trianglelefteq G/O_p(G)$ and thus $(N/O_p(G))' = R'O_p(G)/O_p(G)$. Hence, $R'O_p(G) \trianglelefteq G$. Let $\overline{G} = G/R'O_p(G)$. Then $O_p(\overline{G}) = 1$,

$\bar{R} \trianglelefteq \bar{G}$, $\bar{G}' = \bar{R}$ and $\bar{G} = \langle \bar{x} \rangle \bar{R}$ with $|\bar{x}| = p$. Moreover, $\bar{R}$ is an abelian r-group and $|\bar{G}'| = |\bar{R}| = r^m$ for some positive integer m. It follows that p is the only nontrivial character degree of G. Hence,

$$d(\bar{G}) = \frac{1}{p^2} + \left(1 - \frac{1}{p^2}\right)\frac{1}{|\bar{G}'|},$$

by [14, Lemma 2(vi)]. Since $d(\bar{G}) > 1/(p+1)$, we deduce that

$$r^m < \frac{(p+1)(p^2-1)}{p^2-p-1}.$$

Since $p \geq 3$, we can check that

$$\frac{(p+1)(p^2-1)}{p^2-p-1} \leq 2p+1.$$

Hence $r^m \leq 2p$. As $r \neq p$, we have $r^m \leq 2p-1$. Moreover, as

$$\frac{1}{|\bar{G}'|} \leq d(\bar{G}) \leq \frac{1}{p},$$

we deduce that $r^m \geq p$. Thus

$$(*) \qquad p \leq r^m \leq 2p-1.$$

By Fitting's theorem [18, Theorem 4.34], $\bar{R} = [\bar{R}, \bar{P}] \times \boldsymbol{C}_{\bar{R}}(\bar{P})$. If $\boldsymbol{C}_{\bar{R}}(\bar{P})$ is nontrivial, then $[\bar{R}, \bar{P}]\bar{P}$ is a proper subgroup of $\bar{G}$, which implies that it is p-closed and hence $\bar{P}$ centralizes $[\bar{R}, \bar{P}]$. This is impossible as it would imply that $\bar{P}$ centralizes $\bar{R}$ and thus $\bar{G}$ has a normal Sylow p-subgroup. Therefore, $\boldsymbol{C}_{\bar{R}}(\bar{P}) = 1$. It follows that $\bar{G}$ is a Frobenius group with kernel $\bar{R}$ and a cyclic complement $\bar{P}$. This implies that $r^m - 1 = kp$ for some positive integer k. Combining with $(*)$, we get $r^m - 1 = p$. As $p \geq 3$, $p+1 = r^m$ must be even. Hence $r = 2$ and

$$2^m - 1 = p.$$

In particular, $m \geq 2$ is a prime and p is a Mersenne prime.

For brevity, we define $\Gamma(p, m) := C_2^m \rtimes C_p$ to be a Frobenius group with kernel an elementary abelian 2-group of order 2^m and cyclic complement of order p with $2^m - 1 = p$.

Claim 8: *We have $R' = 1$, so $G/O_p(G) = G/R'O_p(G) \cong \Gamma(p, m)$ with $p = 2^m - 1 \geq 7$ and $m \geq 3$ an odd prime.*

Assume that $m = 2$. Then $p = 3$ and

$$\gamma(h, 3) = \tfrac{1}{4}(1 + 3^{-(2h-1)}) = q_h(\mathrm{A}_4).$$

Hence, $q_h(G) > q_h(\mathrm{A}_4)$ and so by Theorem 1.1(c), G is supersolvable. By [16, Theorem VI.9.1(c)], as G is a $\{2,3\}$-group, G has a normal Sylow 3-subgroup, which is a contradiction. Therefore, we may assume from now on that $m \geq 3$ and $p = 2^m - 1 \geq 7$.

Working in the quotient group $G/O_p(G)$, we may assume $O_p(G) = 1$. By Claim 6, $G' = R$ and so $G'' = R'$ with $[R : R'] = 2^m$ by Claim 7. Recall that n is the smallest nontrivial character degree of G. Assume first that $n \geq p$. By [14, Lemma 2(vi)], we have

$$\frac{1}{p+1} < d(G) \leq \frac{1}{n^2} + \left(1 - \frac{1}{n^2}\right)\frac{1}{|G'|} \leq \frac{1}{p^2} + \left(1 - \frac{1}{p^2}\right)\frac{1}{|G'|}.$$

As above, we deduce that $p \leq |G'| \leq 2p - 1$. Write $|G'| = |R| = 2^{m+c}$ for some integer $c \geq 0$. Recall that $2^m = 1 + p$ so

$$p \leq |G'| = (1+p)2^c = 2^c p + 2^c \leq 2p - 1.$$

This implies that $2^c = 1$, and hence $G' = R$ is abelian.

Now assume that $n < p$. As $|G| = 2^{m+c}p$ and n divides $|G|$ and is coprime to p, we deduce that $n = 2^a$ for some positive integer $a \geq 1$, and hence that

$$\tfrac{1}{2}(p-1) \leq 2^a < p = 2^m - 1,$$

the first inequality coming from Claim 5. It follows that

$$\tfrac{1}{2}(2^m - 2) \leq 2^a \leq p - 1 = 2^m - 2.$$

Therefore,

$$2^{m-1} - 1 = \tfrac{1}{2}(2^m - 2) \leq 2^a \leq 2^m - 2.$$

We deduce that $n = 2^a = 2^{m-1}$. By [14, Lemma 2 (vi)], we have

$$\frac{1}{p+1} = \frac{1}{2^m} < d(G) \leq \frac{1}{2^{2(m-1)}} + \left(1 - \frac{1}{2^{2(m-1)}}\right)\frac{1}{|G'|}.$$

Since $m \geq 3$, simplifying the previous inequality, we obtain

$$|G'| < 2^m + \frac{2^m - 1}{2^{m-2} - 1}.$$

Now, the second summand on the right hand side of the inequality is at most 7 and thus $2^{m+c} = |G'| < 2^m + 7$. However, this is impossible if $c \geq 1$ as $m \geq 3$.

We showed in both cases that $G' = R$ is abelian. If R is not elementary abelian, then the subgroup S generated by all elements of R of order 2 is a proper subgroup of R stabilized by P, so $SP < G$ has a normal Sylow p-subgroup. Then $P \trianglelefteq SP$ fixes every element of order 2 in R, and so P acts trivially on R by [18, Corollary 4.35].

But then $P \trianglelefteq G$, a contradiction. Thus, R is elementary abelian and $G \cong \Gamma(p, m)$, as claimed.

Claim 9: *The final contradiction.*

Assume $O_p(G) = 1$. Then $G \cong \Gamma(p, m)$, and so $[G : G'] = p$ and p is the unique nontrivial character degree of G. Hence, $q_h(G) = \gamma(h, p)$, which is a contradiction.

Assume that $O_p(G) > 1$. Note that $G' = RU$ and $U = G' \cap O_p(G)$ is a normal Sylow p-subgroup of G'.

(i) Assume that $p \nmid |G'|$. So $G' = R$ is a normal Sylow r-subgroup of G. Now $P = \langle x \rangle O_p(G)$ is abelian. In this case, $G'\langle x \rangle \trianglelefteq G$. If $G'\langle x \rangle < G$, then x centralizes G', which is impossible. Thus $G = G'\langle x \rangle$. In particular, $P = \langle x \rangle$ and $y = x^p$ centralizes $G' = R$.

By Claim 8, $G/O_p(G) \cong \Gamma(p, m)$ is a Frobenius group with $G' = R$ an elementary abelian subgroup of order $2^m = p + 1$. Thus $\boldsymbol{F}(G) = G' \times \langle y \rangle = G' \times O_p(G)$ is abelian and in fact $O_p(G) = \langle y \rangle$ is central in G. It follows that p is the only nontrivial character degree of G. Since $[G : G'] = |P|$, we have

$$q_h(G) = \frac{1}{|G|}\left(|P| + \frac{k(G) - |P|}{p^{2h-2}}\right) \text{ and } |G| = [G : G'] + p^2(k(G) - |P|).$$

Solving for $k(G) - |P|$ from the second equation and plugging it in the former, we get $q_h(G) = \gamma(h, p)$, a contradiction.

(ii) Assume $p \mid |G'|$. So $|G'| = |R|p^e$ for some positive integer e. Let n be the minimal nontrivial degree of G. By [14, Lemma 2 (vi)], we have

$$\frac{1}{p+1} < \frac{1}{n^2} + \left(1 - \frac{1}{n^2}\right)\frac{1}{|G'|}.$$

Assume first that $n \geq p$. The preceding inequality then implies that

$$|G'| < \frac{(p+1)(p^2-1)}{p^2 - p - 1}.$$

As in the proof of Claim 8, we deduce that $|G'| \leq 2p - 1$. However, since $|G'| = |R|p^e$ with $e \geq 1$ and $|R| = 2^m = 1 + p \geq 8$, we see that $|G'| > 2p - 1$, which is a contradiction.

Assume that $n < p$. By Claim 5, we have $n \geq (p-1)/2$. Thus $(p-1)/2 \leq n \leq p-1$. Let $\chi \in \mathrm{Irr}(G)$ with $\chi(1) = n$. Recall that $2^m = p+1$ and $G/O_p(G) \cong \Gamma(p, m)$. It follows that $\chi(1)$ divides $|R|$ where $|R| = 2^m$ since $p \nmid \chi(1)$ and $\chi(1) \mid |G|$. So $n = 2^a$ for some positive integer a. Since $(p-1)/2 = (2^m - 2)/2 = 2^{m-1} - 1 < 2^a \leq p - 1 = 2^m - 2$, we deduce that $n = 2^{m-1}$.

Employing the same argument as in the proof of Claim 8, we obtain a contradiction as $m \geq 3$.

The proof of the theorem is now complete. □

5. The dual invariant

In this section, we prove Theorems 1.4 and 1.5.

Proof of the abelian criterion, Theorem 1.4(a). Assume for a contradiction that G is a counterexample to the claim with minimal order. Note that the sizes of the conjugacy classes of D_8 are $1, 1, 2, 2, 2$. Since $\tilde{q}_1(G) = d(G)$, we must have $h \geq 2$. As G is nonabelian, we have $d(G) \leq \frac{5}{8}$ by [15]. Every noncentral conjugacy class of G has size at least 2, so we have

$$\begin{aligned}\frac{1}{4}\Big(1+\frac{3}{2^h}\Big) = \tilde{q}_h(\mathrm{D}_8) < \tilde{q}_h(G) &\leq \frac{1}{|G|}\left(|\mathbf{Z}(G)| + \frac{k(G)-|\mathbf{Z}(G)|}{2^{h-1}}\right)\\ &= \frac{1}{[G:\mathbf{Z}(G)]}\Big(1-\frac{1}{2^{h-1}}\Big)+\frac{d(G)}{2^{h-1}}\\ &\leq \frac{1}{[G:\mathbf{Z}(G)]}\Big(1-\frac{1}{2^{h-1}}\Big)+\frac{1}{4}\cdot\frac{5}{2^h}.\end{aligned}$$

We find immediately from these inequalities that $[G:\mathbf{Z}(G)] < 4$. But then $G/\mathbf{Z}(G)$ is cyclic and G is abelian, a contradiction. □

Proof of the nilpotency criterion, Theorem 1.4(b). Assume for a contradiction that G is a counterexample to the claim with minimal order. Since $\tilde{q}_1(G) = d(G)$, we must have $h \geq 2$. We have

$$\frac{1}{6} < \frac{1}{6}\Big(1+\frac{1}{2^{h-1}}+\frac{1}{3^{h-1}}\Big) = \tilde{q}_h(\mathrm{S}_3) < \tilde{q}_h(G) \leq d(G) \leq \frac{1}{2}.$$

As $\tilde{q}_h(G/\mathbf{Z}(G)) \geq \tilde{q}_h(G)$ by Lemma 2.6, the quotient $G/\mathbf{Z}(G)$ is nilpotent if $\mathbf{Z}(G) > 1$ by the minimality of $|G|$. But then G is nilpotent, a contradiction. Therefore, we may assume that $|\mathbf{Z}(G)| = 1$.

Since $d(G)/\tilde{q}_h(G) \leq \frac{1}{2}/\frac{1}{6} = 3$, we may take $N = 3$ in Lemma 2.7. Letting $n \geq 2$ be the minimal nontrivial class size, if $n^{h-1} \leq N = 3$, then we must have $h = 2$. In this case,

$$\frac{11}{36} = \tilde{q}_2(\mathrm{S}_3) < \tilde{q}_2(G) \leq \frac{1}{|G|}\Big(1+\frac{k(G)-1}{2}\Big) = \frac{1}{2|G|}+\frac{d(G)}{2} \leq \frac{1}{2|G|}+\frac{1}{4},$$

which means $|G| < 9$ and so $G \cong \mathrm{S}_3$, a contradiction. Therefore, it only remains to consider case (2) of Lemma 2.7; that is, $|G| \leq d(G)/\tilde{q}_h(G)^2 < 18$. The only nonnilpotent groups with trivial center of order less than 18 are dihedral groups D_{2k} for $k = 3, 5, 7$ and A_4. Since

$$\tilde{q}_h(\mathrm{D}_{2k}) = \frac{1}{2k}\Big(1+\frac{(k-1)/2}{2^{h-1}}+\frac{1}{k^{h-1}}\Big)$$

when k is odd, $\tilde{q}_h(\mathrm{D}_{2k})$ is a decreasing function of k, and so $\tilde{q}_h(G) \le \tilde{q}_h(\mathrm{S}_3)$ for all h in these cases. The class sizes of A_4 are $1, 3, 4, 4$, so

$$\begin{aligned}\tilde{q}_h(\mathrm{A}_4) &= \frac{1}{12}\Big(1+\frac{1}{3^{h-1}}+\frac{2}{4^{h-1}}\Big)\\ &< \frac{1}{12}\Big(2+\frac{2}{3^{h-1}}+\frac{2}{4^{h-1}}\Big) = \frac{1}{6}\Big(1+\frac{1}{3^{h-1}}+\frac{1}{4^{h-1}}\Big) < \tilde{q}_h(\mathrm{S}_3),\end{aligned}$$

a contradiction. The result is proved. □

Proof of the supersolvability criterion, Theorem 1.4(c). Assume for a contradiction that G is a counterexample to the claim with minimal order. Since $\tilde{q}_1(G) = d(G)$, we must have $h \ge 2$. We have

$$\frac{1}{12} < \frac{1}{12}\Big(1+\frac{1}{3^{h-1}}+\frac{2}{4^{h-1}}\Big) = \tilde{q}_h(\mathrm{A}_4) < \tilde{q}_h(G) \le d(G) \le \frac{1}{3}.$$

Since $d(G) > \frac{1}{12}$, G is solvable by [10]. Since G is supersolvable if $G/\Phi(G)$ is supersolvable by [16, Theorem VI.8.6(a)] and G is supersolvable if $G/\boldsymbol{Z}(G)$ is supersolvable by the definition of supersolvability [16, Definition VI.8.5(b)], it follows by Lemma 2.6 and induction that both the Frattini subgroup and the center of G are trivial.

Suppose that G has two distinct minimal normal subgroups, say M_1 and M_2. By Lemma 2.6 and the minimality of $|G|$, both G/M_1 and G/M_2 are supersolvable, and so G is supersolvable as G embeds into a supersolvable group $G/M_1 \times G/M_2$. This is a contradiction. Therefore, G has a unique minimal normal subgroup, say M, and since G is solvable, M is an elementary abelian p-group of order p^m for some prime p and positive integer $m \ge 1$. If $m = 1$, then G is supersolvable, so we may assume that $m > 1$. As $\Phi(G) = 1$, we have $\boldsymbol{F}(G) = M$ by Gaschütz' theorem [16, Theorem III.4.5], and hence $\boldsymbol{C}_G(M) = M$ by [16, Theorem III.4.2 b)].

First assume that $n = 2$, where n is the minimal size of a noncentral class of G. Let $x \in G$ be such that $|x^G| = [G : \boldsymbol{C}_G(x)] = 2$. Then $\boldsymbol{C}_G(x) \trianglelefteq G$, and so $M \le \boldsymbol{C}_G(x)$ as M is the unique minimal normal subgroup of G. Hence, x lies in $\boldsymbol{C}_G(M) = M$. Since $\langle x^G\rangle$ is a normal subgroup of G and $x \in M$, we must have $\langle x^G\rangle = M$. As $|x^G| = 2$, $M = \langle x, x^t\rangle$ for some $t \in G \setminus \boldsymbol{C}_G(x)$. In particular, $|M| = p^2$ as M is noncyclic elementary abelian and generated by two elements. Now, $M = \boldsymbol{C}_G(M) = \boldsymbol{C}_G(x) \cap \boldsymbol{C}_G(x^t) = \boldsymbol{C}_G(x) \cap \boldsymbol{C}_G(x)^t = \boldsymbol{C}_G(x)$ as $\boldsymbol{C}_G(x)$ is normal in G. Thus, $[G : M] = [G : \boldsymbol{C}_G(x)] = 2$, and so $|G| = 2p^2$. If $p = 2$, then G is a finite 2-group, so it is nilpotent and hence it is supersolvable. Therefore, we may assume that $p \ge 3$. Since $|G| = 2p^2$, G has a cyclic Sylow 2-subgroup $Q = \langle g\rangle$, where g is an involution, acting on the unique minimal normal subgroup $M \cong C_p \times C_p$ of G by conjugation. But this is impossible. To see this, let $1 \ne a \in M$. We have $(aa^g)^g = a^g a = aa^g$, so $\langle aa^g\rangle$ is a normal subgroup of G. If $aa^g \ne 1$, then $\langle aa^g\rangle$ is a nontrivial cyclic normal subgroup of G properly contained in the

noncyclic group M, contradicting the fact that M is minimal normal. On the other hand, if $aa^g = 1$, then $a^g = a^{-1}$ and $\langle a \rangle$ is again a nontrivial normal subgroup of G properly contained in M.

Therefore, may assume that $n \geq 3$. Since $d(G)/\tilde{q}_h(G) \leq \frac{1}{3}/\frac{1}{12} = 4$, we may choose to set $N = 8$ in Lemma 2.7. Since $3^{h-1} \leq n^{h-1} \leq N = 8$, we must have $h = 2$. We have

$$\frac{11}{72} = \tilde{q}_2(\mathrm{A}_4) < \tilde{q}_2(G) \leq \frac{1}{|G|} + \frac{d(G)}{n} \leq \frac{1}{|G|} + \frac{1}{9}.$$

Simplifying this inequality, we obtain $|G| < 24$. Since G is nonsupersolvable with trivial center, the only possibility for G is A_4. This is a contradiction.

There remains to consider case (2) of Lemma 2.7. Since $\boldsymbol{Z}(G) = 1$ by the first paragraph and we chose $N = 8$, we have

$$|G| \leq \frac{8}{\frac{1}{12} \cdot 9 - \frac{1}{3}} = \frac{96}{5} < 20.$$

The only nonsupersolvable group of order less than 20 with trivial center is again A_4. This contradiction completes the proof of this item. □

Proof of the solvability criterion, Theorem 1.4(d). Assume for a contradiction that G is a counterexample to the claim with minimal order. Since $\tilde{q}_1(G) = d(G)$, we must have $h \geq 2$. We have

$$\tfrac{1}{60} < \tilde{q}_h(\mathrm{A}_5) < \tilde{q}_h(G) \leq d(G) \leq \tfrac{1}{12}.$$

As $\tilde{q}_h(G/\mathrm{Sol}(G)) \geq \tilde{q}_h(G)$ by Lemma 2.6, the quotient $G/\mathrm{Sol}(G)$ is solvable if the solvable radical $\mathrm{Sol}(G)$ is nontrivial by the minimality of $|G|$. But then G is solvable, a contradiction. Therefore, we may assume that $|\mathrm{Sol}(G)| = 1$.

Let n be the minimal nontrivial class size of G. If n is a prime power, then the subgroup generated by that class is a solvable normal subgroup of G by [22]. As G has no nontrivial solvable normal subgroups, we must have $n \geq 6$. Since $d(G)/\tilde{q}_h(G) \leq (\frac{1}{12})/(\frac{1}{60}) = 5$, we may take $N = 5$ in Lemma 2.7. This means that $n^{h-1} \leq N$ never occurs, and we need only consider case (2) of Lemma 2.7; that is,

$$|G| \leq d(G)/\tilde{q}_h(G)^2 < 300.$$

The only nonsolvable groups of order less than 300 with trivial solvable radical are A_5, S_5, and $\mathrm{PSL}_3(2)$. The conjugacy class sizes of A_5 are 1, 12, 12, 15, 20; those of S_5 are 1, 10, 15, 20, 20, 24, 30; and those of $\mathrm{PSL}_3(2)$ are 1, 21, 24, 24, 42, 56. Then one can check by direct calculation that $\tilde{q}_h(\mathrm{PSL}_3(2)) < \tilde{q}_h(\mathrm{S}_5) < \tilde{q}_h(\mathrm{A}_5)$. This means G is not isomorphic to A_5, S_5, or $\mathrm{PSL}_3(2)$, and we are done. □

We now prove Theorem 1.5. For a prime p and a positive integer h, define

$$\tilde{\gamma}(h, p) = \frac{1}{p(p+1)}\left(1 + \frac{1}{p^{h-1}} + \frac{p-1}{(p+1)^{h-1}}\right).$$

Proof of Theorem 1.5. Since $\tilde{\gamma}(h, 2) = \tilde{q}_h(\mathrm{S}_3)$, if $p = 2$, then G is nilpotent by Theorem 1.4(b) and thus has a normal Sylow 2-subgroup. Therefore, we may assume $p \geq 3$. As $\tilde{q}_1(G) = d(G)$ and $\tilde{\gamma}(1, p) = 1/p$, the theorem follows from Lemma 2.1(c) when $h = 1$. So we may assume that $h \geq 2$. Let G be a counterexample to the theorem with minimal order. Then

$$\frac{1}{p} \geq d(G) \geq \tilde{q}_h(G) > \tilde{\gamma}(h, p) > \frac{1}{p(p+1)}$$

but G is not p-closed. Let P be a Sylow p-subgroup of G. We may assume that $O_p(G) = 1$ as otherwise, by Lemma 2.6 and the minimality of $|G|$, we deduce that $G/O_p(G)$ has a normal Sylow p-subgroup and so does G, a contradiction. Next, assume that $\boldsymbol{Z}(G) > 1$. Then $\boldsymbol{Z}(G)$ is a p'-group, and $G/\boldsymbol{Z}(G)$ has a normal Sylow p-subgroup $P\boldsymbol{Z}(G)/\boldsymbol{Z}(G)$ again by Lemma 2.6. It follows that $P \trianglelefteq P\boldsymbol{Z}(G) \trianglelefteq G$ which implies that $P \trianglelefteq G$, a contradiction. Therefore, $\boldsymbol{Z}(G) = 1$.

Next, we claim that G has a unique minimal normal subgroup. Suppose that G has two distinct minimal normal subgroups, say N_1 and N_2. By Lemma 2.6 and the minimality of $|G|$, both G/N_1 and G/N_2 are p-closed. Since G embeds into $G/N_1 \times G/N_2$, the group G is p-closed as subgroups and direct products of p-closed groups are p-closed. It follows that G has a unique minimal normal subgroup N, and $N \cong S^m$ for some simple group S and some integer $m \geq 1$.

Let $H = \boldsymbol{N}_G(P)$. Since $PN \trianglelefteq G$ by Lemma 2.6 and the minimality of $|G|$, we have $G = HN$ by Frattini's argument.

Claim 1: *If C is a noncentral conjugacy class of G, then $|C| \geq p$.*

If $C_0 = x^G$ is a noncentral conjugacy class of G having minimal size, we claim that $|C_0| \geq p$. For if $|C_0| = [G : \boldsymbol{C}_G(x)] < p$ instead, and $U = \bigcap_{g \in G} \boldsymbol{C}_G(x)^g$ is the core of $\boldsymbol{C}_G(x)$ in G, then $U \trianglelefteq G$ and $[G : U] = |C_0| \cdot [\boldsymbol{C}_G(x) : U]$. As G/U embeds into the symmetric group S_t, where $t = |C_0| = [G : \boldsymbol{C}_G(x)] < p$, $|G/U|$ is not divisible by p. Since $U \trianglelefteq G$ and Sylow subgroups are conjugate by Sylow's theorems, $P \leq U$. Hence U is nontrivial and thus $N \leq U \leq \boldsymbol{C}_G(x)$. In particular, $x \in \boldsymbol{C}_G(N)$.

If N is nonabelian, then $\boldsymbol{C}_G(N) = 1$, which leads to a contradiction. Now assume that N is abelian. Then $S \cong C_r$ for some prime r and so N is an elementary abelian r-group of order r^m. Since $G = HN$ and N is the unique minimal normal subgroup of G, we must have $\Phi(G) = 1$ and that $\boldsymbol{F}(G) = O_r(G)$ is a product of minimal normal subgroups of G by Gaschütz' theorem [16, Theorem III.4.5]. Thus $\boldsymbol{F}(G) = N$ and so $\boldsymbol{C}_G(N) = N$ by [16, Theorem III.4.2 b)]. It follows that $x \in N$. Furthermore, since $P \leq \boldsymbol{C}_G(x)$, we have $x \in \boldsymbol{C}_N(P)$. Observe that P

acts coprimely and nontrivially on N. By Fitting's Theorem [18, Theorem 4.34], $N = [N, P] \times \boldsymbol{C}_N(P)$. Note that $\boldsymbol{N}_N(P) = N \cap \boldsymbol{N}_G(P) = N \cap H$. As $N \trianglelefteq G$, we have $H \cap N \trianglelefteq H$ and since N is abelian, $H \cap N \trianglelefteq N$. It follows that $H \cap N \trianglelefteq G = NH$, and by the minimality of N, we have $N \cap H = 1$. Now $\boldsymbol{C}_N(P) \leq \boldsymbol{N}_N(P) = 1$, which implies that $x \in \boldsymbol{C}_N(P) = 1$, a contradiction.

Therefore, we have shown that $|C_0| \geq p$, and thus since $\boldsymbol{Z}(G) = 1$, for each nontrivial conjugacy class C of G, we have $|C| \geq p$.

Claim 2: *$|G| < 2p(p+1)$ and $|P| = p$. Let $n_p = [G : \boldsymbol{N}_G(P)]$. Then $n_p = 1 + p$ or $1 + 2p$. In particular, $H = P$ and $G = PN$.*

Assume for a contradiction that $|G| \geq 2p(p+1)$. The hypothesis yields

$$\frac{1}{p(p+1)} < \tilde{q}_h(G) \leq \frac{1}{|G|}\left(1 + \sum_{\substack{C \in \mathrm{Cl}(G) \\ C \neq 1}} \frac{1}{|C|^{h-1}}\right) \leq \frac{1}{|G|}\left(1 + \frac{k(G)-1}{p^{h-1}}\right)$$

$$= \frac{1}{|G|}\left(1 - \frac{1}{p^{h-1}}\right) + \frac{d(G)}{p^{h-1}}.$$

Then

$$\frac{1}{p(p+1)} < \frac{1}{2p(p+1)} + \frac{d(G)}{p^{h-1}}.$$

Since $d(G) \leq 1/p$, we obtain $p^h < 2p(p+1)$, which implies $h = 2$ as $p \geq 3$.

For $h = 2$, the hypothesis yields (noting that $d(G) \leq 1/p$):

$$\tilde{\gamma}(2, p) = \frac{1}{p(p+1)}\left(1 + \frac{1}{p} + \frac{p-1}{p+1}\right) < \tilde{q}_h(G) \leq \frac{1}{|G|}\left(1 - \frac{1}{p}\right) + \frac{d(G)}{p}$$

$$\leq \frac{1}{|G|}\left(1 - \frac{1}{p}\right) + \frac{1}{p^2}.$$

Simplifying, we obtain $|G| < (p+1)^2 < 2p(p+1)$, a contradiction.

Assume that $|P| \geq p^2$. If N is nonabelian, then $|N|$ is divisible by 4 and $|G| \geq 4p^2 > 2p(p+1)$, a contradiction. Assume that N is abelian. Since $\boldsymbol{C}_G(N) = N$, P acts nontrivially and coprimely on N. Hence $|N| > p$ as P has at least one nontrivial orbit on $N \setminus \{1\}$. Then $H \cap N = 1$ by the argument in Claim 1, and thus $|G| = |H||N| \geq p^2|N| > p^3 > 2p(p+1)$. In both cases, we obtain a contradiction. Thus, $|P| = p$ as claimed.

By Sylow's theorem, we know that $n_p \equiv 1 \bmod p$, so that $n_p = 1 + kp$ for some positive integer k. Since $n_p = [G : H] \leq [G : P] < 2(p+1)$, we have $n_p < 2(p+1)$ and so $n_p \in \{1 + p, 1 + 2p\}$.

Now since $|G| = |H| n_p \in \{(1+p)|H|, (1+2p)|H|\}$ and $|G| < 2p(p+1)$, we must have $|H| < 2p$, forcing $|H| = p$ and thus $H = \boldsymbol{C}_G(P) = P$. In particular, $G = PN$.

Claim 3: *N is abelian.*

Suppose otherwise. Then $N \cong S^m$ for some nonabelian simple group S. By the uniqueness of N, we deduce that G has a trivial solvable radical.

If $|C| = p$ for some class C, then by [22], $\langle C \rangle$ is a normal solvable subgroup of G, which is impossible. Thus $|C| \geq p+1$ for every nontrivial class C of G.

If p divides $|S|$, then $p = |G|_p \geq (|S|_p)^m$ means that $m = 1$ and $G = PS = S$ is simple. However, in this case, $|G| \geq 2p(p+1)$ by the proof of Lemma 2.3. Thus p does not divide $|S|$ and so N is a p'-group. So $|G| = p|S|^m$ with $m = 1$ or $m = p$. (If N is not simple, then $m = p$ since P acts transitively on the simple factors of N as N is a minimal normal subgroup of G.)

Now $|N| = [G : H] = n_p \in \{1+p, 1+2p\}$, so we have $|N| = |S|^m = 1+p$ or $1+2p$. Since P has order p and $\boldsymbol{C}_G(P) = P$ by the last step, the only point of N fixed by P under conjugation is the identity and every other orbit has size p. This implies that P has at most 3 orbits in its conjugation action on N, which is impossible since $|N|$ is divisible by at least 3 distinct primes by Burnside's $p^a q^b$-theorem, so N must have at least 4 orbits under the action of P.

Claim 4: *The final contradiction.*

We now have that $G = PN$ with $|P| = p$, and N is an elementary abelian r-group of order r^m for some prime r and $r^m = |N| = 1+p$ or $1+2p$. Since $\boldsymbol{Z}(G) = 1$, we see that G is a Frobenius group with Frobenius kernel N and complement P. Therefore, G has the following class structure: the $r^m - 1$ elements of order r split into classes of size p, and the $r^m(p-1)$ elements of order p split into classes of size r^m. Since $r^m = 1+kp$ for $k = 1$ or $k = 2$, we have

$$\tilde{q}_h(G) = \frac{1}{|G|}\left(1 + \frac{(r^m-1)/p}{p^{h-1}} + \frac{p-1}{(r^m)^{h-1}}\right) = \frac{1}{p(1+kp)}\left(1 + \frac{k}{p^{h-1}} + \frac{p-1}{(1+kp)^{h-1}}\right).$$

If $|N| = 1+p$, then substituting $k = 1$ into the equation above yields $\tilde{q}_h(G) = \tilde{\gamma}(h, p)$, a contradiction. So $r^m = |N| = 1+2p$, and we can calculate $\tilde{q}_h(G)$ by substituting $k = 2$ into the equation above.

Since

$$\frac{1}{p(p+1)} < \tilde{q}_h(G) \leq \frac{1}{|G|} + \frac{d(G)}{p^{h-1}} \leq \frac{1}{p(1+2p)} + \frac{1}{p^h},$$

we deduce that $p^h < (1+p)(1+2p)$, forcing $h = 3$ and $p = 3$, or $h = 2$. In the former case, we calculate directly that $\tilde{\gamma}(3,3) > \tilde{q}_3(G)$, a contradiction. Similarly, if $h = 2$, then we calculate directly that we must have

$$\frac{1}{p(p+1)}\left(1 + \frac{1}{p} + \frac{p-1}{p+1}\right) = \tilde{\gamma}(2, p) < \tilde{q}_2(G) = \frac{1}{p(1+2p)}\left(1 + \frac{2}{p} + \frac{p-1}{1+2p}\right).$$

This can be simplified to $5p^4 + 2p^3 - 3p^2 - 3p < 1$, which is impossible since $p \geq 3$. The proof is now complete. □

6. Introduction to Dijkgraaf–Witten theory

For mathematicians, a topological quantum field theory (TQFT) is a machine for generating topological invariants of manifolds. For example, Witten [35] showed that the Jones polynomial is a topological invariant of manifolds in a certain TQFT. Atiyah [1] axiomatized topological quantum field theories in the language of category theory as a functor from a particular category of bordisms to the category of vector spaces. The invariants that we consider arise in the context of Dijkgraaf–Witten theory [7; 9] in one space dimension and one time dimension. Such a TQFT determines a commutative Frobenius algebra that gives rise to topological invariants of surfaces. The new contribution of this paper is that, upon assuming this Frobenius algebra is the center of a complex group algebra, we instead consider these invariants as invariants of groups. It should be noted, however, that Dijkgraaf–Witten theory has already been used to study finite groups; for example, row and column sums in their character tables [28; 31; 32]. In this section, we derive these topological invariants in an elementary, explicit way, without assuming any previous knowledge of quantum field theory. For a similarly gentle introduction, we recommend Dijkgraaf's 2015 lecture at the Institute for Advanced Study [8].

Let X, Y be two closed $(n-1)$-dimensional manifolds. A bordism from X to Y is a quadruple (Z, p, f_X, f_Y) where Z is an n-manifold with boundary δZ, $p : \delta Z \to \{0, 1\}$ is a continuous map (partition of the boundary into two components), $f_X : X \to p^{-1}(0)$, $f_Y : Y \to p^{-1}(1)$ are diffeomorphisms. (The empty set is considered an $(n-1)$-manifold for all n.) Two such bordisms Z, Z' from X to Y are considered equivalent if there is a diffeomorphism from Z to Z' respecting the maps p and f in the obvious way. Declaring Mor (X,Y) to be the set of all equivalence classes of bordisms from X to Y, we get the structure of a category $B(n-1)$ on closed $(n-1)$-manifolds. The composition of two bordisms is obtained by gluing them. This category $B(m)$ is a symmetric monoidal category with disjoint union as the monoidal product. Any covariant functor of symmetric monoidal categories from $B(m)$ to vector spaces Vect_k over a field k, with tensor product as the monoidal product, is an example of a topological quantum field theory.

When $n = 2$, closed 1-manifolds are just disjoint unions of a finite number of circles S^1. So a TQFT T of the above type on objects of $B(1)$ is determined by specifying $T(S^1) = V$. The images of morphisms via T are linear maps between appropriate tensor powers of V. For example, a cylinder is a bordism between two circles, and it corresponds to the identity map $\mathrm{id} : V \to V$. Perhaps most importantly, the "pair of pants" corresponds to a multiplication of quantum states $V \otimes V \to V$, which gives V the structure of an algebra. This multiplication is commutative, which we can see from topological considerations: Interchanging the legs in a pair of pants (which corresponds, on the vector space side, to switching the order

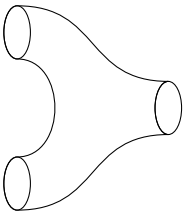
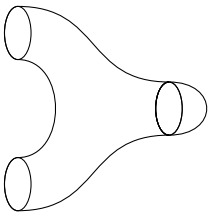
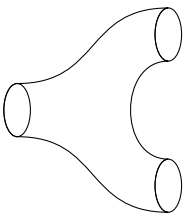
bordism	map	$V = \mathbf{Z}(\mathbb{C}G)$
	$V \otimes V \to V$	$e_\chi \cdot e_{\chi'} = \delta_{\chi,\chi'} e_\chi$
	$\langle\ \rangle : V \otimes V \to \mathbb{C}$	$\langle e_\chi, e_{\chi'} \rangle = \delta_{\chi,\chi'} \left(\frac{\chi(1)}{\lvert G \rvert}\right)^2$
	$\mathbb{C} \to V, \quad 1 \mapsto e$	$e = \sum_{\chi \in \mathrm{Irr}(G)} e_\chi$
	$V \to \mathbb{C}$	$e_\chi \mapsto \left(\frac{\lvert G \rvert}{\chi(1)}\right)^2$
	$V \to V \otimes V$	$e_\chi \mapsto \left(\frac{\lvert G \rvert}{\chi(1)}\right)^2 e_\chi \otimes e_\chi$

Table 1. Bordisms giving rise to linear maps of vector spaces.

of the tensor product) is a diffeomorphism, so two such bordisms are equivalent. The bordisms giving rise to a multiplicative identity $e \in V$, a comultiplication $V \to V \otimes V$ and a bilinear form $\langle\ \rangle : V \otimes V \to \mathbb{C}$ are pictured in Table 1. By considering appropriate bordisms, it is not difficult to show that the multiplication is associative and the bilinear form satisfies the identity

$$\langle a, b \cdot c \rangle = \langle a \cdot b, c \rangle$$

of a Frobenius algebra. A genus-h surface can be decomposed into a composition of the bordisms in Table 1, giving rise to a map $\mathbb{C} \to \mathbb{C}$. Since the map is linear, it is defined by the image of 1, and this complex number is a topological invariant of the surface.

Now we assume that $V = \boldsymbol{Z}(\mathbb{C}G)$ is the center of the complex group algebra of a finite group G. We compute the maps in terms of a basis for $\boldsymbol{Z}(\mathbb{C}G)$ that is indexed by the irreducible complex characters $\chi \in \mathrm{Irr}(G)$ of G; namely, the primitive central idempotents

$$e_\chi = \frac{\chi(1)}{|G|} \sum_{g \in G} \overline{\chi(g)} g.$$

In this basis, the multiplicative identity decomposes as $e = \sum_{\chi \in \mathrm{Irr}(G)} e_\chi$, and the multiplication takes the form $e_\chi \cdot e_{\chi'} = \delta_{\chi,\chi'} e_\chi$. We define the bilinear form to be

$$\langle e_\chi, e_{\chi'} \rangle = \delta_{\chi,\chi'} \left(\frac{\chi(1)}{|G|} \right)^2,$$

which is the coefficient of the identity in the product $e_\chi \cdot e_{\chi'}$ times $1/|G|$. This is the so-called special Frobenius algebra structure on $\boldsymbol{Z}(\mathbb{C}G)$.

To show that the "cup" in the third row of Table 1 gives rise to a multiplicative identity of V, note that the bordism in Figure 1, top left, is homeomorphic to the cylinder, which is the identity map. To show that the "cap" in the fourth row of Table 1 sends e_χ to $(\chi(1)/|G|)^2$, note that the cap is homeomorphic to adding a cup to the bordism for the bilinear form, as in Figure 1, top middle. Finally, to show that the comultiplication is $e_\chi \mapsto (|G|/\chi(1))^2 e_\chi \otimes e_\chi$ in the last row of Table 1, note that capping off the comultiplication is homeomorphic to the cylinder, which represents the identity map, as in Figure 1, top right.

Now, we compute the map $\mathbb{C} \to \mathbb{C}$ for a genus-h surface by composing the maps

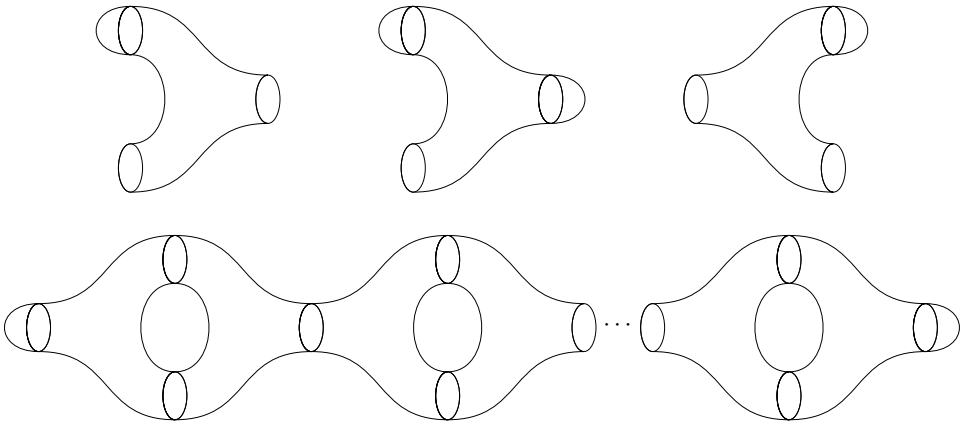


Figure 1. Bordisms for calculating (top row from the left) the cup, the cap and the comultiplication. Decomposing the genus-h surface on the bottom into bordisms calculated in Table 1 yields an invariant $\mathcal{Q}_h(G)$ of the surface.

in Table 1:

$$1 \overset{\bigcirc}{\longmapsto} \sum_{\chi\in\mathrm{Irr}(G)} e_\chi \overset{\cap}{\longmapsto} \sum_{\chi\in\mathrm{Irr}(G)} \left(\frac{|G|}{\chi(1)}\right)^2 e_\chi \otimes e_\chi \overset{\cup}{\longmapsto} \sum_{\chi\in\mathrm{Irr}(G)} \left(\frac{|G|}{\chi(1)}\right)^2 e_\chi$$
$$\overset{\cap}{\longmapsto} \cdots \overset{\cup}{\longmapsto} \sum_{\chi\in\mathrm{Irr}(G)} \left(\frac{|G|}{\chi(1)}\right)^{2h} e_\chi \overset{\bigcirc}{\longmapsto} \sum_{\chi\in\mathrm{Irr}(G)} \left(\frac{|G|}{\chi(1)}\right)^{2h-2} = \mathcal{Q}_h(G).$$

The number $\mathcal{Q}_h(G)$ is the topological invariant of a genus-h surface in this TQFT. In this paper, we considered a more well-behaved invariant of groups by scaling $\mathcal{Q}_h(G)$ by an appropriate factor of $|G|$, namely

$$q_h(G) = \frac{1}{|G|} \sum_{\chi\in\mathrm{Irr}(G)} \left(\frac{1}{\chi(1)}\right)^{2h-2}.$$

Acknowledgment

The authors would like to thank Dr. Alessandro Mariani and Prof. Marcin Mazur for comments on a previous version of the paper, as well as the anonymous referee for their close reading of the paper and helpful comments.

References

[1] M. Atiyah, "Topological quantum field theories", *Inst. Hautes Études Sci. Publ. Math.* **68** (1988), 175–186. MR

[2] A. Ballester-Bolinches and R. Esteban-Romero, "On minimal non-supersoluble groups", *Rev. Mat. Iberoam.* **23**:1 (2007), 127–142. MR

[3] F. Barry, D. MacHale, and A. Ní Shé, "Some supersolvability conditions for finite groups", *Math. Proc. R. Ir. Acad.* **106A**:2 (2006), 163–177. MR

[4] W. Burnside, *Theory of groups of finite order*, 2nd ed., Cambridge University Press, 1911. Reprinted Dover, New York, 1955; reissued in TEX by Project Gutenberg. MR

[5] D. Chillag and M. Herzog, "On character degrees quotients", *Arch. Math.* (*Basel*) **55**:1 (1990), 25–29. MR

[6] Y. Choi, "Small values and forbidden values for the Fourier anti-diagonal constant of a finite group", *J. Aust. Math. Soc.* **118**:3 (2025), 297–316. MR

[7] R. Dijkgraaf, "Fields, strings and duality", pp. 3–147 in *Symétries quantiques* (Les Houches, 1995), North-Holland, Amsterdam, 1998. MR

[8] R. Dijkgraaf, "PiTP 2015: Introduction to topological and conformal field theory", video lecture part 1, available at http://youtu.be/jEEQO-tcyHc.

[9] R. Dijkgraaf and E. Witten, "Topological gauge theories and group cohomology", *Comm. Math. Phys.* **129**:2 (1990), 393–429. MR

[10] J. D. Dixon, "Solution to Problem 176", *Canadian Math. Bull.* **16** (1973), 302.

[11] K. Doerk, "Minimal nicht überauflösbare, endliche Gruppen", *Math. Z.* **91** (1966), 198–205. MR

[12] W. Feit and J. G. Thompson, "Groups which have a faithful representation of degree less than $(p-1/2)$" [sic; for $(p-1)/2$], *Pacific J. Math.* **11** (1961), 1257–1262. MR

[13] D. Gluck, K. Magaard, U. Riese, and P. Schmid, "The solution of the $k(GV)$-problem", *J. Algebra* **279**:2 (2004), 694–719. MR

[14] R. M. Guralnick and G. R. Robinson, "On the commuting probability in finite groups", *J. Algebra* **300**:2 (2006), 509–528. MR

[15] W. H. Gustafson, "What is the probability that two group elements commute?", *Amer. Math. Monthly* **80** (1973), 1031–1034. MR

[16] B. Huppert, *Finite groups, I*, Grundl. Math. Wissen. **364**, Springer, 2025. MR

[17] I. M. Isaacs, *Character theory of finite groups*, Academic Press, New York, 1976. Corrected reprint, Dover, New York, 1994. MR

[18] I. M. Isaacs, *Finite group theory*, Graduate Studies in Mathematics **92**, Amer. Math. Soc., 2008. MR

[19] I. M. Isaacs, *Characters of solvable groups*, Graduate Studies in Mathematics **189**, Amer. Math. Soc., 2018. MR

[20] K. S. Joseph, *Commutativity in non-abelian groups*, Ph.D. thesis, University of California, Los Angeles, 1969, available at https://www.proquest.com/docview/302401138. MR

[21] K. S. Joseph, "Research problems: several conjectures on commutativity in algebraic structures", *Amer. Math. Monthly* **84**:7 (1977), 550–551. MR

[22] L. S. Kazarin, "Burnside's p^α-lemma", *Mat. Zametki* **48**:2 (1990), 45–48, 158. In Russian; translated in *Math. Notes* **48**:1–2 (1990), 749–751. MR

[23] P. Lescot, "Sur certains grupes finis", *Rev. Math. Spéciales* **8** (1987), 267–277.

[24] M. W. Liebeck and A. Shalev, "Fuchsian groups, coverings of Riemann surfaces, subgroup growth, random quotients and random walks", *J. Algebra* **276**:2 (2004), 552–601. MR

[25] M. W. Liebeck and A. Shalev, "Character degrees and random walks in finite groups of Lie type", *Proc. London Math. Soc.* (3) **90**:1 (2005), 61–86. MR

[26] M. W. Liebeck and A. Shalev, "Fuchsian groups, finite simple groups and representation varieties", *Invent. Math.* **159**:2 (2005), 317–367. MR

[27] A. Mariani, S. Pradhan, and E. Ercolessi, "Hamiltonians and gauge-invariant Hilbert space for lattice Yang–Mills-like theories with finite gauge group", *Phys. Rev. D* **107**:11 (2023), art. id. 114513, 15 pp. MR

[28] R. de Mello Koch, Y.-H. He, G. Kemp, and S. Ramgoolam, "Integrality, duality and finiteness in combinatoric topological strings", *J. High Energy Phys.* **2022**:1 (2022), art. id. 071, 56 pp. MR

[29] G. Navarro, *Characters and blocks of finite groups*, London Mathematical Society Lecture Note Series **250**, Cambridge University Press, 1998. MR

[30] H. N. Nguyen and A. Maróti, "p-regular conjugacy classes and p-rational irreducible characters", *J. Algebra* **607** (2022), 387–425. MR

[31] A. Padellaro, R. Radhakrishnan, and S. Ramgoolam, "Row-column duality and combinatorial topological strings", *J. Phys. A* **57**:6 (2024), art. id. 065202, 68 pp. MR

[32] S. Ramgoolam and E. Sharpe, "Combinatoric topological string theories and group theory algorithms", *J. High Energy Phys.* **2022**:10 (2022), art. id. 147, 56 pp. MR

[33] C. A. Schroeder, "Finite groups with many p-regular conjugacy classes", *J. Algebra* **641** (2024), 716–734. MR

[34] A. Vera López and J. Vera López, "Classification of finite groups according to the number of conjugacy classes", *Israel J. Math.* **51**:4 (1985), 305–338. MR

[35] E. Witten, "Quantum field theory and the Jones polynomial", pp. 239–329 in *Braid group*, *knot theory and statistical mechanics*, Advanced Series in Mathematical Physics **9**, World Sci. Publ., Teaneck, NJ, 1989. MR

Received December 3, 2025. Revised April 26, 2026.

CHRISTOPHER A. SCHROEDER
DEPARTMENT OF MATHEMATICS AND STATISTICS
BINGHAMTON UNIVERSITY
BINGHAMTON, NY 13902-6000
UNITED STATES

cschroe2@binghamton.edu

HUNG P. TONG-VIET
DEPARTMENT OF MATHEMATICS AND STATISTICS
BINGHAMTON UNIVERSITY
BINGHAMTON, NY 13902-6000
UNITED STATES

htongvie@binghamton.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 344, No. 1, 2026
https://doi.org/10.2140/pjm.2026.344.183

FROM BRICK MANIFOLDS TO GRASSMANNIANS OF BIMODULES

EVGENY FEIGIN AND MARKUS REINEKE

We study a class of Grassmannians of subbimodules over the path algebras of quivers. Our quiver Grassmannians include Escobar's brick manifolds as well as Labelle's generalizations. We give an explicit construction of the varieties in question, provide examples and clarify the connection with the quiver representation spaces. We also prove smoothness of our Grassmannians, construct cellular decompositions and derive a realization as framed moduli spaces. The framed moduli realization leads to a recursive formula for the motives.

1. Introduction

The brick manifolds introduced in [Escobar 2016] are certain fibers of the Bott–Samelson varieties which are not birationally isomorphic to the flag varieties. They turn out to be useful in various situations; see, e.g., [Dotsenko et al. 2019; Casals et al. 2021; Speyer 2023]. In the type-A case under certain conditions the brick manifolds are smooth toric varieties admitting an explicit realization as collections of subspaces in an ambient vector space satisfying a number of incidence conditions.

A generalization of the type-A brick manifolds was suggested in [Labelle 2025]. The generalized brick manifolds are no longer toric, but still smooth and admit an explicit realization in terms of collections of subspaces. Conjecturally, their Poincaré polynomials are equal to a specialization of Toda eigenfunctions [Braverman and Finkelberg 2005; Givental and Lee 2003]. The conjecture is equivalent to a certain fermionic-type recursion for the Poincaré polynomials [Braverman and Finkelberg 2014; Feigin et al. 2009; Labelle 2025].

All of these constructions have to do with the equioriented type-A quiver. More precisely, the generalized brick varieties can be realized as Grassmannians of bimodules of the corresponding path algebras. Our goal is to extend the setup to the case of arbitrary quivers with no parallel paths and to study the varieties thus obtained. Below we provide some details of the construction and formulate the main results.

MSC2020: 14M15, 16G20.
Keywords: quivers, representations, Grassmannians.

© 2026 The Authors, under license to MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

Let Q be an acyclic quiver without parallel paths; let Q_0 be the set of vertices and Q_1 be the set of edges. We fix a collection of vector spaces $V_* = (V_i)_{i\in Q_0}$ and for each pair $(i, j) \in Q_0^2$ of vertices of Q we define the vector space

$$M(V_*)_{i,j} = \bigoplus_{\omega,\omega'} V_p \quad \text{such that } j \overset{\omega'}{\rightsquigarrow} p \overset{\omega}{\rightsquigarrow} i$$

(i.e., ω is a path in Q from a vertex p to i and ω' is a path in Q from j to p). The collection $M(V_*) = (M(V_*)_{i,j})_{i,j}$ has a natural structure of a bimodule over the path algebra A of the quiver Q; in other words, $M(V_*)$ carries a natural structure of a module over $A \otimes A^{\mathrm{op}}$. Our main objects of study are the varieties

$$X(V_*) = \mathrm{Gr}^{\boldsymbol{f}(V_*)}_{A\otimes A^{\mathrm{op}}}(M(V_*))$$

of subbimodules of $M(V_*)$ of certain codimension $\boldsymbol{f}(V_*) = (\boldsymbol{f}_{i,j})_{i,j}$. The $\boldsymbol{f}_{i,j}$ vanish unless there exists a path from j to i in Q; if such a path does exist, then $\boldsymbol{f}_{i,j} = \dim V_i$. These varieties can be also realized as quiver Grassmannians for certain quivers (see also [Cerulli Irelli et al. 2012; Caldero and Reineke 2008; Cerulli Irelli 2020]).

One observes that if Q is an equioriented type-A quiver, then the Grassmannians above are isomorphic to the generalized brick manifolds from [Labelle 2025]. We prove the following theorem, where $\boldsymbol{d} = (\dim V_i)_{i\in Q_0}$ and $G_{\boldsymbol{d}} = \prod_{i\in Q_0} \mathrm{GL}_{d_i}$.

Theorem A. *Assume that Q is acyclic and has no parallel paths. Then $X(V_*)$ is a smooth $G_{\boldsymbol{d}}$-equivariant compactification of the representation space $R_{\boldsymbol{d}}(Q)$.*

Our next task is to study the motives $[X(V_*)]$ of the varieties $X(V_*)$. More precisely, we construct a cellular decomposition of our Grassmannians of bimodules and derive a recursive formula for their motives $[X(V_*)]$ which involves more general Grassmannians. We note that our formula in the equioriented type-A case is different from the one conjectured in [Labelle 2025].

Theorem B. *Let Q be an acyclic quiver with no parallel paths. Then $X(V_*)$ admits a cellular decomposition with the cells labeled by the torus fixed points and the motives $[X(V_*)]$ satisfy the recursive formula from Theorem 7.4.*

Our construction works for quivers with parallel paths as well. However, the varieties thus constructed are more complicated; in particular, they are no longer smooth in general. We provide the definitions and examples at the end of our paper.

The paper is organized as follows. In Section 2 we introduce basic notation on quivers and bimodules over path algebras. In Section 3 we introduce our main objects of study — certain quiver Grassmannians of subbimodules. In Section 4 we provide several examples of the general construction. In Section 5 we show that our Grassmannians compactify representation spaces and in Section 6 we give a proof

of smoothness. In Section 7 we provide a framed moduli realization and derive a recursive formula for the motives of the framed moduli spaces. Finally, in Section 8 we generalize our construction to the case of quivers with parallel paths.

2. Quiver representations and bimodules

We start with introducing basic definitions from the theory of quivers (see, e.g., [Assem et al. 2006; Schiffler 2014]). Let k be an algebraically closed field, and let Q be an acyclic quiver with set of vertices Q_0 and arrows written as $\alpha : i \to j$. We denote by $A = kQ$ the (finite-dimensional) path algebra of Q over k. The unit of the path algebra decomposes into orthogonal idempotents $1 = \sum_{i \in Q_0} e_i$. Then $e_i A e_j$ is linearly spanned by the paths from j to i.

Left modules over A correspond to representations of Q over k; namely, to a representation $V = ((V_i)_i, (V_\alpha)_\alpha)$ we associate the module with underlying k-vector space $V = \bigoplus_i V_i$, the generator e_i of A acting as the projection to the V_i-component of V, and the generator α acting via projecting from V to the V_i-component, mapping with V_α to V_j, and embedding V_j into V. Conversely, given a left module V over A, we define $V_i = e_i V$ for $i \in Q_0$ and V_α as the action of α on V, which gives a well-defined map from V_i to V_j for all $\alpha : i \to j$.

In particular, the indecomposable projective representation P_i corresponds to Ae_i and the indecomposable injective representation I_i corresponds to $A^* e_i$ (where $A^* = \mathrm{Hom}_k(A, k)$ denotes the linear dual). Every left A-module V then admits a standard projective resolution

$$0 \to \bigoplus_{\alpha : i \to j} Ae_j \otimes_k e_i V \xrightarrow{\alpha} \bigoplus_i Ae_i \otimes_k e_i V \xrightarrow{\beta} V \to 0,$$

where the maps are given as follows: for a path ω in Q starting in i (which is an element of Ae_i) and an element v of $e_i V$, we define $\beta(\omega \otimes v) = V_\omega(v)$. Here V_ω is defined as

$$V_{\alpha_s} \circ \cdots \circ V_{\alpha_1} \quad \text{if } \omega = \alpha_s \cdots \alpha_1.$$

For a path ω in Q starting in j and an element v of $e_i V$, we define

$$\alpha(\omega \otimes v) = \omega\alpha \otimes v - \omega \otimes V_\alpha(v).$$

Now we consider A-A-bimodules. By definition, these are equivalent to left $A \otimes A^{\mathrm{op}}$-modules. We have $(kQ)^{\mathrm{op}} = k(Q^{\mathrm{op}})$, where Q^{op} denotes the opposite quiver, with an arrow $\alpha^* : j \to i$ for every arrow $\alpha : i \to j$ in Q. It is then easy to see

$$A \otimes A^{\mathrm{op}} \simeq k(Q \times Q^{\mathrm{op}})/I, \tag{2-1}$$

where $Q \times Q^{\mathrm{op}}$ is the quiver that has vertices (i, j) for $i, j \in Q_0$ and arrows $(\alpha, j) : (i, j) \to (i', j)$ for all $\alpha : i \to i'$ in Q and $j \in Q_0$, as well as arrows

$(i, \alpha^*) : (i, j') \to (i, j)$ for all $i \in Q_0$ and $\alpha : j \to j'$ in Q. The ideal I is generated by all commutativity relations, that is, by all

$$(\alpha, j)(i, \beta^*) - (i', \beta^*)(\alpha, j')$$

for $\alpha : i \to i'$ and $\beta : j \to j'$ in Q. We can thus view A-A-bimodules M as representations of $Q \times Q^{\mathrm{op}}$, given by k-vector spaces $M_{i,j}$ for $i, j \in Q_0$ and linear maps $M_{\alpha,j} : M_{i,j} \to M_{i',j}$ for all $\alpha : i \to i'$ in Q and $j \in Q_0$, as well as linear maps $M_{i,\alpha^*} : M_{i,j'} \to M_{i,j}$ for all $i \in Q_0$ and $\alpha : j \to j'$ in Q, such that

$$M_{\alpha,j} \circ M_{i,\beta^*} = M_{i',\beta^*} \circ M_{\alpha,j'}$$

for $\alpha : i \to i'$ and $\beta : j \to j'$ in Q.

The A-A-bimodule $A \otimes_k A^*$ decomposes into indecomposable projective bimodules $P_{ij} = Ae_i \otimes e_j A$ for $i, j \in Q_0$. We have

$$\mathrm{Hom}_{A \otimes A^{\mathrm{op}}}(P_{ij}, P_{ql}) \simeq e_i A e_q \otimes e_l A e_j$$

for all $i, j, q, l \in Q_0$.

3. A class of Grassmannians of subbimodules

From now on and till the last section we assume that Q is a tree, that is, there are no parallel paths in Q. For every $i \in Q_0$ we fix a k-vector space V_i, and consider the bimodule $Ae_i \otimes_k V_i \otimes_k e_i A$ (which is isomorphic to $(\dim V_i)$-many copies of the indecomposable projective bimodule P_{ii}). Finally, we define

$$M(V_*) = \bigoplus_{i \in Q_0} Ae_i \otimes V_i \otimes e_i A. \tag{3-1}$$

To describe this as a representation of $Q \times Q^{\mathrm{op}}$ fulfilling the commutativity relations, we calculate

$$M(V_*)_{i,j} = e_i M(V_*) e_j = \bigoplus_p e_i A e_p \otimes V_p \otimes e_p A e_j = \bigoplus_{(\omega,\omega')} V_p,$$

where the final direct sum ranges over all pairs of paths (ω, ω') in Q such that

$$j \overset{\omega'}{\rightsquigarrow} p \overset{\omega}{\rightsquigarrow} i.$$

All maps are given by obvious inclusions; we only treat the case of the map $M(V_*)_{\alpha,j}$ for an arrow $\alpha : i \to i'$ in Q: given a component V_p of $M(V_*)_{i,j}$ corresponding to a pair of paths $j \overset{\omega'}{\rightsquigarrow} p \overset{\omega}{\rightsquigarrow} i$, we have a pair of paths $j \overset{\omega'}{\rightsquigarrow} p \overset{\alpha\omega}{\rightsquigarrow} i'$, yielding a corresponding component V_p in $M(V_*)_{i',j}$.

Let $\overline{Q}_0 \subset Q_0 \times Q_0$ be the set of pairs of vertices (i, j) such that there exists a path from j to i in Q. Then the target of any arrow in $Q \times Q^{\mathrm{op}}$ whose source belongs to $\overline{Q}_0$ is also in $\overline{Q}_0$. We denote by $\overline{Q}$ the full subquiver of $Q \times Q^{\mathrm{op}}$ with

the set of vertices $\overline{Q}_0$. It follows from the definition that $M(V_*)$ is completely determined by its restriction to $\overline{Q}$.

There is a natural structure of a poset on the set of vertices $\overline{Q}_0$. Namely, given two pairs $(i_1, j_1), (i_2, j_2) \in Q_0^2$ connected by paths $j_1 \overset{\omega_1}{\rightsquigarrow} i_1$ and $j_2 \overset{\omega_2}{\rightsquigarrow} i_2$ we write $(i_1, j_1) \leq (i_2, j_2)$ if ω_1 is a subpath of ω_2.

Recall that an upper ideal is a subset C such that $c_1 \in C$ implies that $c_2 \in C$ for any $c_2 \geq c_1$. We call an upper ideal indecomposable if it cannot be written as a disjoint union of two upper ideals.

Lemma 3.1. *There is one-to-one correspondence between the indecomposable subrepresentations of $M(V_*)$ and the indecomposable upper ideals of the poset $\overline{Q}_0$.*

Proof. Let $C \subset \overline{Q}_0$ be an indecomposable upper ideal. We define the $\overline{Q}$ module $M(C)$ with $\dim M(C)_c = 1$ for $c \in C$ and $M(C)_c = 0$ for $c \notin C$ in an obvious way, i.e., with all arrows between nonzero spaces being isomorphisms. Then $M(C)$ is obviously an indecomposable subrepresentation of $M(V_*)$.

Now assume we are given an indecomposable $\overline{Q}$ module $M \subset M(V_*)$. Let $C \subset \overline{Q}_0$ be the set of elements $c \in Q_0$ such that $M_c \neq 0$. We claim that C is an indecomposable upper ideal and $M \simeq M(C)$. First, C is an upper ideal because all the maps in $M(V_*)$ are injective. Second, it is indecomposable, since a decomposition $C = C_1 \sqcup C_2$ into a disjoint union of two upper ideals would lead to the decomposition of the module M. Finally, let us show that all the nontrivial components of M are one-dimensional. We choose a minimal element c of the poset $\overline{Q}_0$ such that $M_c \neq 0$. Choosing a vector $v \in M_c$ we construct a subrepresentation $N \subset M$ as follows. First, $v \in N$. Second, N contains $N(v)$ — a subrepresentation of M generated by v. Third, we add to N all vectors of M which are mapped to $N(v)$ by all possible compositions of arrows. Since all the maps of $M(V_*)$ are embeddings, all the nontrivial components of N are one-dimensional. We are left to show that N is a direct summand of M. This follows from the injectivity of the arrows and from the assumption that Q has no parallel paths. □

We define a dimension vector $\boldsymbol{f} = \boldsymbol{f}(V_*)$ for $Q \times Q^{\mathrm{op}}$ (or, equivalently, for $\overline{Q}$) by

$$\boldsymbol{f}_{i,j} = \begin{cases} \dim V_i & \text{if there exists a path from } j \text{ to } i \text{ in } Q, \\ 0 & \text{otherwise.} \end{cases}$$

Our central object of interest is the Grassmannian of quotient bimodules

$$X = \mathrm{Gr}^{\boldsymbol{f}(V_*)}_{A \otimes A^{\mathrm{op}}}(M(V_*)).$$

Lemma 3.2. *Let $\boldsymbol{e} = \dim M(V_*) - \boldsymbol{f}(V_*)$. Then the Grassmannian X is isomorphic to the Grassmannian of A subbimodules ($A \otimes A^{\mathrm{op}}$ modules) $\mathrm{Gr}_{\boldsymbol{e}}(M(V_*))$.*

Proof. The desired isomorphism sends an element of $\mathrm{Gr}^{\boldsymbol{f}(V_*)}_{A \otimes A^{\mathrm{op}}}(M(V_*))$ to the quotient representation. □

4. Examples

4.1. ***Linearly oriented type-A.*** Let Q be the linearly oriented type-A_n quiver given by $1 \to 2 \to \cdots \to n$. Then

$$M(V_*)_{ij} = V_j \oplus \cdots \oplus V_i \quad \text{for } j \le i,$$

and zero otherwise. All maps in this representation are given by the natural inclusions of direct summands. We have $\boldsymbol{f}_{i,j} = \dim V_i$ for $j \le i$, and zero otherwise. We can interpret $X = \mathrm{Gr}^{\boldsymbol{f}}_{A\times A^{\mathrm{op}}}(M(V_*))$ as the Grassmannian of subbimodules

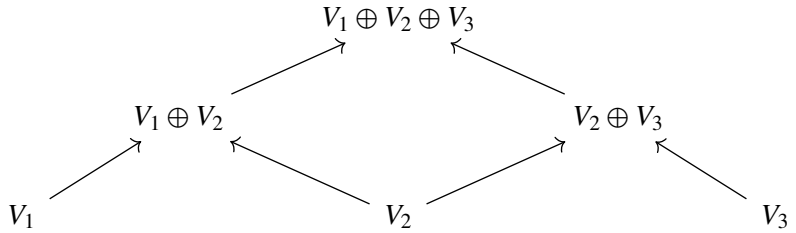
$$\mathrm{Gr}^{A\otimes A^{\mathrm{op}}}_{\boldsymbol{e}}(M(V_*)),$$

where $\boldsymbol{e}$ is given by

$$\boldsymbol{e}_{i,j} = \dim V_j + \cdots + \dim V_{i-1}$$

for $j < i$, and zero otherwise (we note that $\boldsymbol{e} = \dim N(V_*)$ for certain natural projective bimodule; see (5-1)). We thus arrive at the variety $X(d_1, \ldots, d_n)$ defined in [Labelle 2025], where $d_i = \dim V_i$ (if all $d_i = 1$, one gets the Escobar brick manifolds [2016]).

Here is a picture for the representation $M(V_*)$ for $Q = A_3$:



The dimension vector $\boldsymbol{e}$ is given by

$$\begin{array}{ccccc} & & d_1+d_2 & & \\ & d_1 & & d_2 & \\ 0 & & 0 & & 0 \end{array}$$

4.2. ***Alternating orientation quivers.*** Let Q be a quiver with alternating orientation, that is, every vertex is a source or a sink. We then find

$$M(V_*)_{i,j} = \begin{cases} V_i & \text{if } i = j, \\ V_i \oplus V_j & \text{if there is an arrow } \alpha : j \to i, \\ 0 & \text{otherwise.} \end{cases}$$

The only nonzero maps in $M(V_*)$ are the inclusions of direct summands

$$M(V_*)_{\alpha,i} : V_i \to V_i \oplus V_j, \quad M(V_*)_{j,\alpha^*} : V_j \to V_i \oplus V_j$$

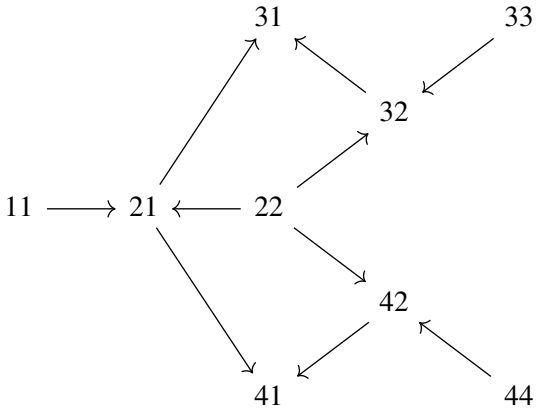
for all $\alpha : i \to j$. In particular, there are no proper commutativity relations fulfilled by this representations. In fact, we can view it as a representation of the quiver Γ with vertices $i \in Q_0$ and v_α for α an arrow in Q, and arrows $i \to v_\alpha \leftarrow j$ for all $\alpha : i \to j$ in Q. This is again a quiver with alternating orientation, and $M(V_*)$ is the projective representation

$$\bigoplus_{i \in Q_0} P_i \otimes V_i.$$

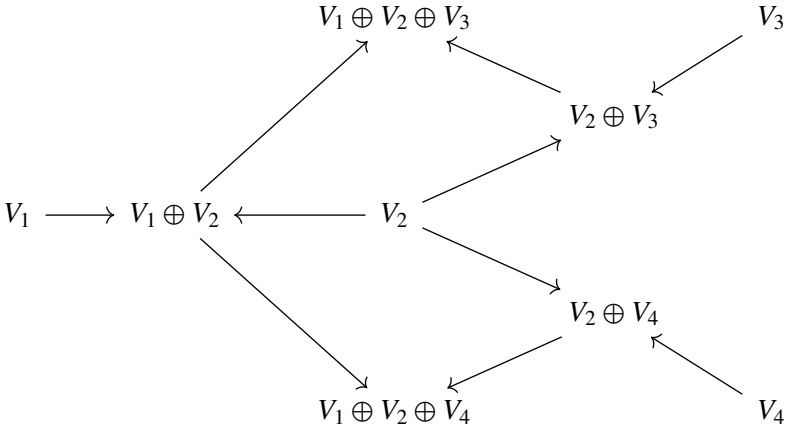
The nonzero components of the dimension vector $\boldsymbol{e} = \boldsymbol{e}(V_*)$ are given by $\boldsymbol{e}_{i,j} = \dim V_j$ for all arrows $j \to i$ in Q. Hence one gets

$$\mathrm{Gr}_{\boldsymbol{e}}(M(V_*)) \simeq \prod_{\alpha \in Q_1} \mathrm{Gr}_{\dim V_{s(\alpha)}}(V_{s(\alpha)} \oplus V_{t(\alpha)}).$$

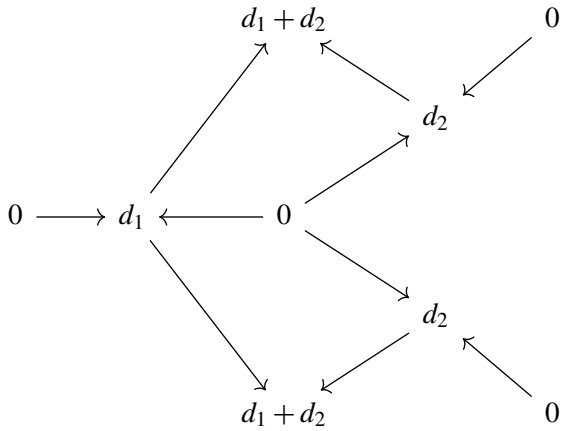
4.3. ***Type D quiver.*** Let Q be the D_4 quiver with vertices 1, 2, 3, 4 and arrows $1 \to 2$, $2 \to 3$ and $2 \to 4$. Then the quiver $\overline{Q}$ is of the form



The representation $M(V_*)$ is given by



and the dimension vector $\boldsymbol{e}$ is given by



Hence the Grassmannian of subbimodules $\mathrm{Gr}_{\boldsymbol{e}}(M(V_*))$ is isomorphic to the quiver Grassmannian for the alternating A_5 quiver with representation

$$V_2 \oplus V_3 \to V_1 \oplus V_2 \oplus V_3 \leftarrow V_1 \oplus V_2 \to V_1 \oplus V_2 \oplus V_4 \leftarrow V_2 \oplus V_4$$

and the dimension vector $(d_2, d_1+d_2, d_1, d_1+d_2, d_2)$.

5. Compactification of representation spaces

The goal of this section is to prove our Grassmannians of subbimodules equivariantly compactify the representation spaces.

Let $\boldsymbol{d} = (\dim V_i)_{i\in Q_0}$ and let

$$R_{\boldsymbol{d}}(Q) = \bigoplus_{\alpha: i\to j} \mathrm{Hom}_k(V_i, V_j)$$

be the representation space. The group $G_{\boldsymbol{d}} = \prod_{i\in Q_0} \mathrm{GL}(V_i)$ acts on $R_{\boldsymbol{d}}(Q)$ by the base change; the orbits of this action correspond to isomorphism classes of k-representations of Q of dimension vector $\boldsymbol{d}$.

Recall the notation $X = \mathrm{Gr}^{\boldsymbol{f}(V_*)}_{A\otimes A^{\mathrm{op}}}(M(V_*))$. The variety X admits a natural $G_{\boldsymbol{d}}$ action induced by the obvious action on the spaces V_i.

Theorem 5.1. *There is a $G_{\boldsymbol{d}}$ equivariant embedding $R_{\boldsymbol{d}}(Q) \subset X$ such that the image is dense in X.*

Proof. We consider the projective bimodule

$$N(V_*) = \bigoplus_{\alpha: i\to j} Ae_j \otimes V_i \otimes e_i A \tag{5-1}$$

and consider bimodule maps from $N(V_*)$ to $M(V_*)$ (inspired by the form of the standard projective resolution above). We then have

$$\begin{aligned}
&\operatorname{Hom}_{A\otimes A^{\mathrm{op}}}(N(V_*), M(V_*)) \\
&\qquad \simeq \bigoplus_{\alpha: i\to j} \bigoplus_{p\in Q_0} \operatorname{Hom}_{A\otimes A^{\mathrm{op}}}(Ae_j \otimes V_i \otimes e_i A,\, Ae_p \otimes V_p \otimes e_p A) \\
&\qquad \simeq \bigoplus_{\alpha: i\to j} \bigoplus_{p\in Q_0} \operatorname{Hom}_k(V_i, \operatorname{Hom}_{A\otimes A^{\mathrm{op}}}(Ae_j \otimes e_i A,\, Ae_p \otimes e_p A) \otimes V_p) \\
&\qquad \simeq \bigoplus_{\alpha: i\to j} \operatorname{Hom}_k(V_i,\, V_i \oplus V_j).
\end{aligned}$$

Here we use the fact that

$$\operatorname{Hom}_{A\otimes A^{\mathrm{op}}}(Ae_j \otimes e_i A,\, Ae_p \otimes e_p A) \simeq e_j Ae_p \otimes e_p Ae_i$$

is spanned by $\alpha \otimes e_i$ and $e_j \otimes \alpha$, making use of the assumption that there are no parallel paths in Q.

Similarly, we can compute

$$\operatorname{End}_{A\otimes A^{\mathrm{op}}}(M(V_*)) \simeq \bigoplus_{i\in Q_0} \operatorname{End}_k(V_i)$$

(even without the assumption on Q) and

$$\operatorname{End}_{A\otimes A^{\mathrm{op}}}(N(V_*)) \simeq \bigoplus_{\alpha: i\to j} \operatorname{End}_k(V_i)$$

(again using the assumption on Q).

Moreover, let us note that

$$\mathbf{dim}\, M(V_*) - \mathbf{dim}\, N(V_*) = \boldsymbol{f}.$$

Namely, we have already proved that

$$M(V_*)_{i,j} = \bigoplus_{j \overset{\omega'}{\rightsquigarrow} p \overset{\omega}{\rightsquigarrow} i} V_p,$$

and similarly, we find

$$N(V_*)_{i,j} = \bigoplus_{j \overset{\omega'}{\rightsquigarrow} p \overset{\alpha}{\to} q \overset{\omega}{\rightsquigarrow} i} V_p.$$

The only summand of $M(V_*)_{i,j}$ not appearing as a summand of $N(V_*)_{i,j}$ is the summand V_i corresponding to $j \overset{\omega'}{\rightsquigarrow} i \overset{e_i}{\rightsquigarrow} i$ in case there exists a path ω' from j to i (again using the assumption of Q admitting no parallel paths). This proves the identity of dimension vectors.

Inside $X = \mathrm{Gr}^{f(V_*)}_{A\otimes A^{\mathrm{op}}}(M(V_*))$, we thus have the locally closed subset consisting of quotients whose kernel is isomorphic to $N(V_*)$. This subset is isomorphic to the quotient of the set

$$\mathrm{Hom}^{\mathrm{inj}}_{A\otimes A^{\mathrm{op}}}(N(V_*), M(V_*))$$

of injective maps by the group

$$\mathrm{Aut}_{A\otimes A^{\mathrm{op}}}(N(V_*)).$$

By the above calculation, this is nothing else than the quotient of

$$Z = \bigoplus_{\alpha: i\to j} \mathrm{Hom}_k(V_i, V_i \oplus V_j)$$

by the action of

$$H = \prod_{\alpha: i\to j} \mathrm{GL}(V_i).$$

On the open subset Z^0 of Z consisting of tuples

$$\left(F_\alpha = \begin{bmatrix} F^1_\alpha \\ F^2_\alpha \end{bmatrix} : V_i \to V_i \oplus V_j\right)_{\alpha: i\to j},$$

where all F^1_α are invertible, we can use the action of H to normalize all F^1_α to identity maps. This shows that

$$Z^0/H \simeq R_{\boldsymbol{d}}(Q).$$

The action of $\mathrm{Aut}_{A\otimes A^{\mathrm{op}}}(M(V_*)) \simeq G_{\boldsymbol{d}}$ on X then restricts to the standard change of base action of $G_{\boldsymbol{d}}$ on

$$R_{\boldsymbol{d}} \simeq Z^0/H \subset Z/H \subset X,$$

yielding the claimed equivariant embedding in to X, which is a projective variety. □

6. Smoothness

This section's goal is to prove smoothness of the Grassmannians of subbimodules, generalizing the corresponding results from [Labelle 2025]. As above, we assume that the quiver Q has no parallel paths.

Theorem 6.1. *The varieties* $\mathrm{Gr}^{f(V_*)}_{A\otimes A^{\mathrm{op}}}(M(V_*))$ *are smooth.*

Proof. We use a slight generalization of [Cerulli Irelli et al. 2013, Proposition 7.1] (see also [Cerulli Irelli et al. 2014]): if B is an algebra of global dimension at most two and M is a representation of B such that $\mathrm{Ext}^1_B(M, M) = 0$ and

$$\mathrm{Ext}^1(N, M/N) = 0 = \mathrm{Ext}^2(N, M/N)$$

for all B-subrepresentations N of M of dimension vector $\boldsymbol{e}$, then $\mathrm{Gr}_{\boldsymbol{e}}^{B}(M)$ is smooth and irreducible of dimension $\langle \boldsymbol{e}, \mathbf{dim}\, M - \boldsymbol{e}\rangle$. A direct inspection of the proof of [Cerulli Irelli et al. 2013, Proposition 7.1] immediately shows that the proof carries over to the present assumptions.

The algebra $A \otimes A^{\mathrm{op}}$ is of global dimension at most two as it is the product of two hereditary algebras. Alternatively, we can easily construct projective resolutions of the simple representations $S_i \otimes S_j$ as

$$0 \to \bigoplus_{\substack{i\to p\\ q\to j}} P_p \otimes P_q \to \bigoplus_{i\to p} P_p \otimes P_j \oplus \bigoplus_{q\to j} P_i \otimes P_q \to P_i \otimes P_j \to S_i \otimes S_j \to 0.$$

Obviously, $P_i \otimes P_j$ has vanishing Ext^1 since it is a projective bimodule.

Now assume we have an exact sequence

$$0 \to N \to M = \bigoplus_i P_i \otimes V_i \otimes P_i \to M/N \to 0.$$

Using the long exact sequence induced by $\mathrm{Hom}_B(\,\cdot\,, M/N)$ and the projectivity of M, we find

$$\begin{aligned} 0 = \mathrm{Ext}^1_B(M, M/N) \to \mathrm{Ext}^1_B(N, M/N) \to \mathrm{Ext}^2_B(M/N, M/N)\\ \to 0 = \mathrm{Ext}^2_B(M, M/N) \to \mathrm{Ext}^2_B(N, M/N) \to \mathrm{Ext}^3_B(M/N, M/N) = 0; \end{aligned}$$

thus $\mathrm{Ext}^2_B(N, M/N) = 0$ and

$$\mathrm{Ext}^1_B(N, M/N) \simeq \mathrm{Ext}^2_B(M/N, M/N).$$

The right-hand side is a factor of $\mathrm{Ext}^2_B(M/N, M)$; thus we are finished if we can prove that the latter vanishes.

We can compute $\mathrm{Ext}^2_B(S_i \otimes S_j, M)$ using the above projective resolution of $S_i \otimes S_j$. Namely, we have a right exact sequence

$$\bigoplus_{i\to p} M_{pj} \oplus \bigoplus_{q\to j} M_{iq} \to \bigoplus_{\substack{i\to p\\ q\to j}} M_{pq} \to \mathrm{Ext}^2_B(S_i \otimes S_j, M) \to 0.$$

It is then easy to see that

$$\mathrm{Ext}^2_B(S_i \otimes S_j, P_l \otimes P_l) = 0$$

for all vertices i, j, l of Q such that there exists a path $j \rightsquigarrow i$, again using the assumption that there are no parallel paths. Now the proof can be finished as above: we only have to prove that $\mathrm{Ext}^2_B(M/N, M) = 0$. But M/N, being of dimension vector $\boldsymbol{f}$, has a composition series with subquotients $S_i \otimes S_j$ for $j \rightsquigarrow i$, whereas M is a direct sum of the $P_l \otimes P_l$. Thus this Ext-group indeed vanishes. □

7. Framed moduli spaces and motivic generating series

7.1. ***Torus fixed points.*** In this section we describe cellular decompositions of the varieties X. We start with the description of the labeling of the set of cells.

Let us fix bases of all V_i and denote by D_i, $i \in Q_0$, a set of cardinality $d_i = \dim V_i$ labeling the basis elements of V_i (the sets D_i do not intersect). Let $T \subset G_{\boldsymbol{d}}$ be a torus of rank $\sum_i d_i$ acting on $M(V_*)$ (recall that all maps in this representation are just inclusions of direct summands). Define

$$D_{i,j} = \bigcup_{j \overset{\omega'}{\rightsquigarrow} p \overset{\omega}{\rightsquigarrow} i} D_p$$

for all $i, j \in Q_0$. We have natural inclusions maps $\iota_{\alpha,j} : D_{i,j} \subset D_{i',j}$ for all $\alpha : i \to i'$ and $j \in Q_0$, as well as $\iota_{i,\alpha^*} : D_{i,j} \subset D_{i,j'}$ for all $\alpha : j' \to j$ and $i \in Q_0$. Set $\boldsymbol{e} = \dim N(V_*)$ as above. The next lemma follows from the above definitions.

Lemma 7.1. *The T-fixed points in the Grassmannian X are labeled by tuples of subsets $(I_{i,j} \subset D_{i,j})_{i,j \in Q_0}$, $|I_{i,j}| = \boldsymbol{e}_{i,j}$ $(i, j \in Q_0)$, which are compatible with all the inclusion maps $\iota_{\alpha,j}$ and ι_{i,α^*}.*

Proof. Each collection $I_{i,j}$ determines a torus fixed point in the Grassmann variety $\mathrm{Gr}_{\boldsymbol{e}_{i,j}}(M(V_*)_{i,j})$. The product of these points belongs to the Grassmannian of bimodules X if and only if the collections $I_{i,j}$ satisfy all the conditions above. In what follows for a collection $\boldsymbol{I}$ of the admissible sets $I_{i,j}$ we denote the corresponding point in X by $p(\boldsymbol{I})$. □

Recall (see Theorem 6.1) that our Grassmannians of subbimodules are smooth. Choosing a general enough one-dimensional subtorus one applies the Białynicki-Birula theorem [1973; 1974] in order to derive an affine paving of X. The cells of the paving are labeled by the collections $\boldsymbol{I} = (I_{i,j})_{i,j}$ from the above lemma. Recall (see Section 3) the quiver $\overline{Q}$ whose set of vertices consists of pairs $(i, j) \in Q_0 \times Q_0$ such that there is a path from j to i in Q. The vertices $\overline{Q}_0$ form a poset and for each upper ideal C in $\overline{Q}_0$ which cannot be written as a disjoint union of two upper ideals, one gets an indecomposable $\overline{Q}$ module $M(C)$. In particular, all the components of $M(C)$ are at most one-dimensional and the nontrivial components are labeled by the elements of C.

Lemma 7.2. *The point $p(\boldsymbol{I})$ as a $\overline{Q}$ module is isomorphic to the direct sum of modules $M(C)$ over all upper ideals C satisfying that there exists $i = 1, \dots, n$ and $r \in D_i$ such that*

$$r \in I_{a,b} \iff (a, b) \in C.$$

The representation $M(C)$ further decomposes into a direct sum of $M(C_u)$ such that $C = \bigsqcup_u C_u$ and each C_u is an indecomposable upper ideal.

Proof. It is enough to note that the maps $\iota_{\alpha,j}$ and ι_{i,α^*} are embeddings. □

7.2. ***Framed moduli spaces.*** Using the methods of [Reineke 2008], one can prove that X is isomorphic to a framed moduli space. Define vector spaces $M_{i,j}$ as V_i if there exists a path from j to i in Q, and as 0 otherwise. Consider the variety $R_{\boldsymbol{e}}(Q \times Q^{\mathrm{op}}, I)$ of representations of $Q \times Q^{\mathrm{op}}$ on the vector spaces $M_{i,j}$ which satisfy all commutativity relations (making them into a representation M of $kQ \otimes kQ^{\mathrm{op}}$), on which the group

$$G_{\boldsymbol{e}} = \prod_{j \rightsquigarrow i} \mathrm{GL}(M_{i,j})$$

acts. We define the framed representation variety as

$$R_{\boldsymbol{e}}^{\boldsymbol{d}-\mathrm{fr}}(Q \times Q^{\mathrm{op}}, I) = R_{\boldsymbol{e}}(Q \times Q^{\mathrm{op}}, I) \times \bigoplus_{i \in Q_0} \mathrm{Hom}_k(V_i, M_{i,i}),$$

in which we consider the open subset $R_{\boldsymbol{e}}^{\boldsymbol{d}-\mathrm{fr}}(Q \times Q^{\mathrm{op}}, I)^{\mathrm{st}}$ of pairs (M, F) such that the direct sum of the images $F_i(V_i) \subset M_{i,i}$ generates M as a representation of $Q \times Q^{\mathrm{op}}$. The group $G_{\boldsymbol{e}}$ acts freely on the latter, and the quotient

$$M_{\boldsymbol{e}}^{\boldsymbol{d}-\mathrm{fr}}(Q \times Q^{\mathrm{op}}, I) = R_{\boldsymbol{e}}^{\boldsymbol{d}-\mathrm{fr}}(Q \times Q^{\mathrm{op}}, I)^{\mathrm{st}}/G_{\boldsymbol{e}}$$

exists as quasiprojective variety since it is a GIT quotient of a set of stable points.

Proposition 7.3. *The framed moduli space $M_{\boldsymbol{e}}^{\boldsymbol{d}-\mathrm{fr}}(Q \times Q^{\mathrm{op}}, I)$ is isomorphic to the Grassmannian of subbimodules X.*

Proof. It suffices to note that the images of the maps F_i generate M if and only if the representation M is a quotient of the projective representation $M(V_*)$. □

We stratify $R_{\boldsymbol{e}}^{\boldsymbol{d}-\mathrm{fr}}(Q \times Q^{\mathrm{op}}, I)$ by locally closed subsets $R_{\boldsymbol{e}}^{\boldsymbol{d}-\mathrm{fr}}(Q \times Q^{\mathrm{op}}, I)_{\boldsymbol{g}}$ consisting of all pairs (M, F) such that the images $F_i(V_i)$ generate a subbimodule of dimension vector $\boldsymbol{g} \leq \boldsymbol{e}$. In particular,

$$R_{\boldsymbol{e}}^{\boldsymbol{d}-\mathrm{fr}}(Q \times Q^{\mathrm{op}}, I)_{\boldsymbol{e}} = R_{\boldsymbol{e}}^{\boldsymbol{d}-\mathrm{fr}}(Q \times Q^{\mathrm{op}}, I)^{\mathrm{st}}$$

is the open stratum. We have

$$R_{\boldsymbol{e}}^{\boldsymbol{d}-\mathrm{fr}}(Q \times Q^{\mathrm{op}}, I)_{\boldsymbol{g}} \simeq G_{\boldsymbol{e}} \times^{P_{\boldsymbol{g},\boldsymbol{e}-\boldsymbol{g}}} Z_{\boldsymbol{g},\boldsymbol{e}-\boldsymbol{g}},$$

where $Z_{\boldsymbol{g},\boldsymbol{e}-\boldsymbol{g}} \subset R_{\boldsymbol{e}}^{\boldsymbol{d}-\mathrm{fr}}(Q \times Q^{\mathrm{op}}, I)$ consists of pairs

$$\left(\begin{bmatrix} M' & * \\ 0 & M'' \end{bmatrix}, \begin{bmatrix} f' \\ 0 \end{bmatrix} \right),$$

where $(M', f') \in R_{\boldsymbol{g}}^{\boldsymbol{d}-\mathrm{fr}}(Q \times Q^{\mathrm{op}}, I)^{\mathrm{st}}$ and $M'' \in R_{\boldsymbol{f}-\boldsymbol{g}}(Q \times Q^{\mathrm{op}}, I)$. We claim that the natural projection

$$p : Z_{\boldsymbol{g},\boldsymbol{e}-\boldsymbol{g}} \to R_{\boldsymbol{g}}^{\boldsymbol{d}-\mathrm{fr}}(Q \times Q^{\mathrm{op}}, I) \times R_{\boldsymbol{e}-\boldsymbol{g}}(Q \times Q^{\mathrm{op}}, I)$$

is a Zariski-locally trivial vector bundle of rank $\langle \boldsymbol{e}-\boldsymbol{g}, \boldsymbol{g}\rangle_B$, where $B = kQ \otimes kQ^{\mathrm{op}}$, and the homological Euler form $\langle \cdot, \cdot \rangle_B$ can be computed using the projective resolution of simple bimodules as

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \sum_{i,j} x_{ij}y_{ij} - \sum_{\substack{i\to i'\\ j}} x_{i,j}y_{i',j} - \sum_{\substack{i\\ j'\to j}} x_{i,j}y_{i,j'} + \sum_{\substack{i\to i'\\ j'\to j}} x_{ij}y_{i'j'}.$$

Namely, constancy of the fiber dimensions of p is a consequence of the Ext^2_B-vanishing property

$$\mathrm{Ext}^2_B(S_i \otimes S_j, P_l \otimes P_l) = 0, \quad \text{whenever } j \rightsquigarrow i.$$

The above stratification immediately allows us to formulate a recursive formula for motives. We work in the localization $K_0(\mathrm{Var}_k)[\mathbb{L}^{-1}, (1-\mathbb{L}^i)^{-1}, i \geq 1]$ of the Grothendieck ring of k-varieties, for $\mathbb{L} = [\mathbb{A}^1]$ the Lefschetz motive, $[Z]$ denoting the class of a variety Z.

Theorem 7.4. *One has*

$$\mathbb{L}^{\sum_{i\in Q_0} d_i^2} \cdot \frac{[R_{\boldsymbol{e}}(Q\times Q^{\mathrm{op}}, I)]}{[G_{\boldsymbol{e}}]} = \sum_{\boldsymbol{g}\leq \boldsymbol{e}} [M_{\boldsymbol{g}}^{\boldsymbol{d}-\mathrm{fr}}(Q\times Q^{\mathrm{op}}, I)] \cdot \frac{[R_{\boldsymbol{e}-\boldsymbol{g}}(Q\times Q^{\mathrm{op}}, I)]}{[G_{\boldsymbol{e}-\boldsymbol{g}}]} \cdot \mathbb{L}^{\langle \boldsymbol{e}-\boldsymbol{g}, \boldsymbol{g}\rangle_B}.$$

Proof. The proof follows from the stratification described above. □

We note that thanks to the existence of cellular decomposition, the framed moduli spaces are trivially motivated, that is, their motive is a polynomial in the Lefschetz motive, and consequently,

$$[M_{\boldsymbol{g}}^{\boldsymbol{d}-\mathrm{fr}}(Q\times Q^{\mathrm{op}}, I)] = \sum_i \dim H^i(M_{\boldsymbol{g}}^{\boldsymbol{d}-\mathrm{fr}}(Q\times Q^{\mathrm{op}}, I), \mathbb{Q}) \cdot \mathbb{L}^{i/2}.$$

8. Quivers with parallel paths

In this section we drop the condition that Q has no parallel paths and develop a more general construction. We show that in general the Grassmannians of subbimodules are more complicated; in particular, we give a nonsmooth example.

Let Q be a quiver with no cycles (but parallel paths are allowed). Let $\overline{Q}$ be a new quiver whose vertices are paths ω in Q and there are two types of arrows. Namely, let ω be a path in Q with the source $s(\omega) \in Q_0$ and target $t(\omega) \in Q_0$. Let $\alpha, \beta \in Q_1$ be the arrows satisfying $t(\alpha) = s(\omega)$, $s(\beta) = t(\omega)$. Then the arrows in $\overline{Q}$ are

$$\big(s(\omega) \overset{\omega}{\rightsquigarrow} t(\omega)\big) \xrightarrow{\hat{\alpha}(\omega)} \big(s(\alpha) \xrightarrow{\alpha} t(\alpha) = s(\omega) \overset{\omega}{\rightsquigarrow} t(\omega)\big), \tag{8-1}$$

$$\big(s(\omega) \overset{\omega}{\rightsquigarrow} t(\omega)\big) \xrightarrow{\hat{\beta}(\omega)} \big(s(\omega) \overset{\omega}{\rightsquigarrow} t(\omega) = s(\beta) \xrightarrow{\beta} t(\beta)\big). \tag{8-2}$$

We define the ideal I in the path algebra of $\overline{Q}$ generated by the relations $\hat{\alpha}(\omega)\hat{\beta}(\omega) = \hat{\beta}(\omega)\hat{\alpha}(\omega)$ for all $\alpha, \beta \in Q_1$, $\omega \in \overline{Q}_0$, satisfying the conditions as above.

Let us fix a collection of vector spaces V_i, $i \in Q_0$, with $\dim V_i = d_i$. We define a $k\overline{Q}/I$ module $M(V_*)$ as follows. For a path ω (a vertex of $\overline{Q}$) the space $M(V_*)_\omega$ is defined as

$$M(V_*)_\omega = \bigoplus V_p, \quad \omega \text{ passes through } p,$$

where the maps $M(V_*)_{\hat{\alpha}(\omega)}$ and $M(V_*)_{\hat{\beta}(\omega)}$ are the obvious embeddings. Also, given a dimension vector $\dim(V_*) = (d_i)_{i\in Q_0}$ we define the dimension vectors $\boldsymbol{f}(\dim V_*) = \boldsymbol{f} = (f_\omega)_{\omega\in\overline{Q}_0}$, $\boldsymbol{e}(\dim V_*) = \boldsymbol{e} = (e_\omega)_{\omega\in\overline{Q}_0}$ by

$$f_\omega = d_{t(\omega)}, \quad e_\omega = \sum_{\substack{p\in\omega \\ p\neq t(\omega)}} d_p.$$

In particular, $\dim M(V_*) = \boldsymbol{f} + \boldsymbol{e}$. We are interested in the Grassmannian of submodules

$$X = \mathrm{Gr}^{\overline{Q}/I}_{\boldsymbol{e}(\dim V_*)}(M(V_*)). \tag{8-3}$$

Given a Q module V we denote by $\overline{V}$ the $k\overline{Q}/I$ module (of dimension $\boldsymbol{e}(\dim V)$) defined as follows. For a lazy (length zero) path ω in Q the space $\overline{V}_\omega$ is trivial (zero). For a nonlazy path one defines

$$\overline{V}_\omega = \bigoplus_{Q_1\ni\alpha\in\omega} \{x + V_\alpha(x) : x \in s(\alpha)\} \subset \bigoplus_{Q_0\ni p\in\omega} V_i$$

(here the sums are taken over the arrows of Q which show up in the path ω). The maps $\overline{V}_{\hat{\alpha}(\omega)}$ and $\overline{V}_{\hat{\beta}(\omega)}$ are the obvious embeddings.

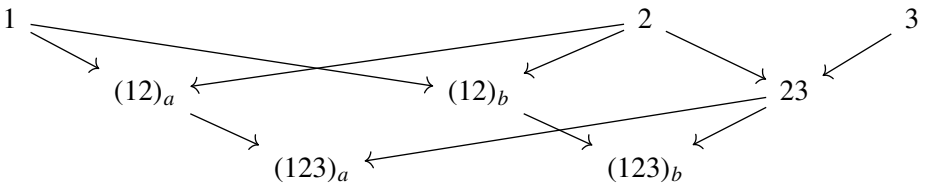
Lemma 8.1. *The $k\overline{Q}/I$ module $\overline{V}$ is a subrepresentation of $\widehat{M}(V_*)$ of dimension $\boldsymbol{e}(\dim V)$.*

Proof. The sum $\bigoplus_{Q_1\ni\alpha\in\omega}\{x + V_\alpha(x) : x \in s(\alpha)\}$ is direct thanks to the acyclicity of Q. □

This lemma allows to embed the representation variety $\mathrm{Rep}_{\boldsymbol{d}}(k\overline{Q}/I)$ into the Grassmannian X (8-3).

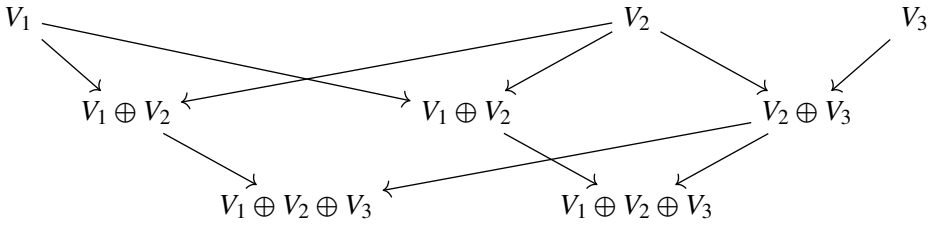
Conjecture 8.2. The closure of the image of the above embedding coincides with X. In particular, X is irreducible.

Let us consider the example $Q = 1 \rightrightarrows 2 \to 3$. Then $\overline{Q}$ is of the form

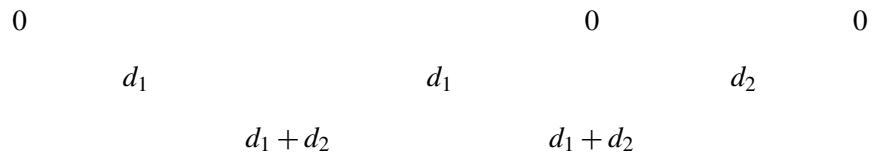


(here the subscripts a and b correspond to the two arrows between the vertices 1 and 2). The ideal I is generated by the two commutative squares in the picture.

Given a collection of vector spaces V_1, V_2, V_3 of dimensions d_1, d_2, d_3 the module $\widehat{M}(V_*)$ is given by



The dimension vector $\boldsymbol{e}(V)$ is given by



The corresponding Grassmannian is irreducible. To see this one notices that we can forget the vertices of $\overline{Q}$ in the upper row (since the values of the dimension vector $\boldsymbol{e}$ are anyway zero); the remaining quiver is of type A_5:

$$(12)_a \to (123)_a \leftarrow 23 \to (123)_b \leftarrow (12)_b.$$

The module $M(V_*)$ takes the form

$$U_{1,2}^{\oplus d_1} \oplus U_{1,5}^{\oplus d_2} \oplus U_{2,4}^{\oplus d_3} \oplus U_{4,5}^{\oplus d_1},$$

where $U_{i,j}$ are indecomposables supported between vertices i and j. The dimension vector $\boldsymbol{e}$ reads as $(d_1, d_1 + d_2, d_2, d_1 + d_2, d_1)$. In particular,

$$\boldsymbol{e} = \dim U_{1,2}^{\oplus d_1} \oplus U_{2,4}^{\oplus d_2} \oplus U_{4,5}^{\oplus d_1}.$$

Now to prove the irreducibility it suffices to observe the Ext vanishing criterion from [Cerulli Irelli et al. 2012].

Finally, let us consider the case $\dim V_* = (1, 1, 1)$. Then one checks that the Poincaré polynomial of the quiver Grassmannian X is equal to

$$1 + 5q + 6q^2 + q^3,$$

and hence X is singular.

Acknowledgement

Feigin was partially supported by the ISF grant 493/24.

References

[Assem et al. 2006] I. Assem, D. Simson, and A. Skowroński, *Elements of the representation theory of associative algebras, I: Techniques of representation theory*, London Mathematical Society Student Texts **65**, Cambridge Univ. Press, 2006. MR

[Białynicki-Birula 1973] A. Białynicki-Birula, "Some theorems on actions of algebraic groups", *Ann. of Math.* (2) **98** (1973), 480–497. MR

[Białynicki-Birula 1974] A. Białynicki-Birula, "On fixed points of torus actions on projective varieties", *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **22** (1974), 1097–1101. MR

[Braverman and Finkelberg 2005] A. Braverman and M. Finkelberg, "Finite difference quantum Toda lattice via equivariant K-theory", *Transform. Groups* **10**:3-4 (2005), 363–386. MR

[Braverman and Finkelberg 2014] A. Braverman and M. Finkelberg, "Semi-infinite Schubert varieties and quantum K-theory of flag manifolds", *J. Amer. Math. Soc.* **27**:4 (2014), 1147–1168. MR

[Caldero and Reineke 2008] P. Caldero and M. Reineke, "On the quiver Grassmannian in the acyclic case", *J. Pure Appl. Algebra* **212**:11 (2008), 2369–2380. MR

[Casals et al. 2021] R. Casals, E. Gorsky, M. Gorsky, and J. Simental, "Positroid links and braid varieties", 2021. arXiv 2105.13948

[Cerulli Irelli 2020] G. Cerulli Irelli, "Three lectures on quiver Grassmannians", pp. 57–89 in *Representation theory and beyond*, edited by J. Šťovíček and J. Trlifaj, Contemp. Math. **758**, Amer. Math. Soc., Providence, RI, 2020. MR

[Cerulli Irelli et al. 2012] G. Cerulli Irelli, E. Feigin, and M. Reineke, "Quiver Grassmannians and degenerate flag varieties", *Algebra Number Theory* **6**:1 (2012), 165–194. MR

[Cerulli Irelli et al. 2013] G. Cerulli Irelli, E. Feigin, and M. Reineke, "Desingularization of quiver Grassmannians for Dynkin quivers", *Adv. Math.* **245** (2013), 182–207. MR

[Cerulli Irelli et al. 2014] G. Cerulli Irelli, E. Feigin, and M. Reineke, "Homological approach to the Hernandez–Leclerc construction and quiver varieties", *Represent. Theory* **18** (2014), 1–14. MR

[Dotsenko et al. 2019] V. Dotsenko, S. Shadrin, and B. Vallette, "Toric varieties of Loday's associahedra and noncommutative cohomological field theories", *J. Topol.* **12**:2 (2019), 463–535. MR

[Escobar 2016] L. Escobar, "Brick manifolds and toric varieties of brick polytopes", *Electron. J. Combin.* **23**:2 (2016), art. id. 2.25. MR

[Feigin et al. 2009] B. Feigin, E. Feigin, M. Jimbo, T. Miwa, and E. Mukhin, "Fermionic formulas for eigenfunctions of the difference Toda Hamiltonian", *Lett. Math. Phys.* **88**:1-3 (2009), 39–77. MR

[Givental and Lee 2003] A. Givental and Y.-P. Lee, "Quantum K-theory on flag manifolds, finite-difference Toda lattices and quantum groups", *Invent. Math.* **151**:1 (2003), 193–219. MR

[Labelle 2025] A. Labelle, "On a specialization of Toda eigenfunctions", 2025. arXiv 2502.10655

[Reineke 2008] M. Reineke, "Framed quiver moduli, cohomology, and quantum groups", *J. Algebra* **320**:1 (2008), 94–115. MR

[Schiffler 2014] R. Schiffler, *Quiver representations*, Springer, 2014. MR

[Speyer 2023] D. E. Speyer, "Richardson varieties, projected Richardson varieties and positroid varieties", 2023. arXiv 2303.04831

Received November 6, 2025. Revised May 23, 2026.

EVGENY FEIGIN
SCHOOL OF MATHEMATICAL SCIENCES
TEL AVIV UNIVERSITY
TEL AVIV
ISRAEL
evgfeig@gmail.com

MARKUS REINEKE
FACULTY OF MATHEMATICS
RUHR-UNIVERSITÄT BOCHUM
BOCHUM
GERMANY
markus.reineke@rub.de

PACIFIC JOURNAL OF MATHEMATICS
Vol. 344, No. 1, 2026
https://doi.org/10.2140/pjm.2026.344.201

ON MIXED CURVATURE FOR HERMITIAN MANIFOLDS

KAI TANG

We consider *mixed curvature* $\mathcal{C}_{\alpha,\beta}$ for Hermitian manifolds, which is a convex combination of the first Chern Ricci curvature and holomorphic sectional curvature introduced by Chu, Lee and Tam (2022). We prove that if a compact Hermitian surface has constant mixed curvature c, then the Hermitian metric must be Kähler unless $c=0$ and $2\alpha+\beta=0$, which extends a previous result by Apostolov, Davidov and Mushkarov (1996). For the higher-dimensional case, we also partially classify compact locally conformal Kähler manifolds with constant mixed curvature. Lastly, we prove that if $\beta \geq 0$ and $\alpha(n+1)+2\beta > 0$, then a compact Hermitian manifold with semipositive but not identically zero mixed curvature has Kodaira dimension $-\infty$.

1. Introduction

1.1. ***Background.*** For a Hermitian manifold (M^n, g, J), the Chern curvature R is the curvature induced by the Chern connection ∇ which is defined as the connection such that $\nabla g = \nabla J = 0$ with no $(1,1)$-components on the torsion. The first Chern Ricci curvature Ric is defined by

$$\mathrm{Ric} = -\sqrt{-1}\partial\bar{\partial}\log\det g,$$

which is a $(1,1)$-form representing the first Chern class $c_1(M)$. The holomorphic sectional curvature H is defined by

$$H(X) = R(X, \overline{X}, X, \overline{X})/|X|^4$$

for a $(1,0)$-tangent vector $X \in T^{1,0}M$ [45]. Given the perplexing relationship between the holomorphic sectional curvature and the first Ricci curvature, one can attempt to interpolate between these curvatures by considering the following curvature constraint: In [11] (see also [5]), Chu, Lee and Tam introduced, for $\alpha, \beta \in \mathbb{R}$, the *mixed curvature*

$$\mathcal{C}_{\alpha,\beta}(X) = \frac{\alpha}{|X|_g^2}\,\mathrm{Ric}(X, \overline{X}) + \beta H(X).$$

The author was supported by the Natural Science Foundation of Zhejiang Province (LMS26A010006).
MSC2020: 32Q15, 53C55.
Keywords: Hermitian manifold, holomorphic sectional curvature, Chern Ricci curvature.

© 2026 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

We always assume that α, β are both not 0, otherwise, the definition is trivial. $\mathcal{C}_{1,0}$ is the Ricci curvature and $\mathcal{C}_{0,1}$ is the holomorphic sectional curvature. When the metric g is Kähler, the mixed curvature encompasses many interesting curvature conditions. $\mathcal{C}_{1,1}$ is the notion $\mathrm{Ric}^+(X, \overline{X})$ introduced by Ni [23]. $\mathcal{C}_{1,-1}$ is the orthogonal Ricci curvature $\mathrm{Ric}^\perp(X, \overline{X})$ introduced by Ni and Zheng [26]. $\mathcal{C}_{k-1,n-k}$ is closely related to the k-Ricci curvature Ric_k introduced by Ni [24].

Recently, there have been many important results on compact Kähler manifolds with $\mathcal{C}_{\alpha,\beta} \geq 0$ or $\mathcal{C}_{\alpha,\beta} \leq 0$. It was proved by Yang [37] that a compact Kähler manifold with $\mathcal{C}_{0,1} > 0$ must be projective and rationally connected, confirming a conjecture of Yau [41]. Ni [23] showed that it is also true if $\mathrm{Ric}_k > 0$ for some $1 \leq k \leq n$. Matsumura [21; 22] established the structure theorems for a projective manifold with $\mathcal{C}_{0,1} \geq 0$. In [5; 11], it was shown that any compact Kähler manifold with $\mathcal{C}_{\alpha,\beta} > 0$, whose constants satisfy $\alpha > 0$ and $3\alpha + 2\beta \geq 0$, must be projective and simply connected. Zhang and Zhang [43] generalized Yang's result to the quasipositive case, confirming a conjecture of Yang affirmatively. Tang [32] showed the projectivity of compact Kähler manifolds with quasipositive $\mathcal{C}_{\alpha,\beta}$. Very recently, Chu, Lee and Zhu [12] established a structure theorem for compact Kähler manifolds with $\mathcal{C}_{\alpha,\beta} \geq 0$. For the nonpositive case, Wu and Yau [34] confirmed a conjecture of Yau that a projective Kähler manifold with $\mathcal{C}_{0,1} < 0$ must have ample canonical line bundle. Tosatti and Yang [33] were able to drop the projectivity assumption in Wu and Yau's theorem. Chu, Lee and Tam [11] proved that a compact Kähler manifold with $\mathcal{C}_{\alpha,\beta} < 0$ and $\alpha \geq 0$ and $\beta \geq 0$ has an ample canonical bundle. For more related works, we refer readers to [6; 13; 16; 18; 25; 27; 31; 35; 39; 42].

Previous results have primarily focused on the study of Kähler cases. In this paper, we mainly investigate Hermitian manifolds.

1.2. ***Constant mixed curvature.*** The holomorphic sectional curvature plays a fundamental role in Hermitian geometry. Complete Kähler manifolds with constant holomorphic sectional curvature H are called complex space forms. They are quotients of complex projective space $\mathbb{CP}^n$, the complex Euclidean space $\mathbb{C}^n$, and the complex hyperbolic space $\mathbb{CH}^n$, equipped with (scaling of) the standard metrics. A long term folklore conjecture in non-Kähler geometry is the following:

Conjecture 1.1. *Let (M^n, g) be a compact Hermitian manifold with $n \geq 2$. Assume that $H = c$ where c is a constant. If $c \neq 0$, then g is Kähler and if $c = 0$, then $R = 0$.*

Note that compact Chern flat manifolds were classified by Boothby [4] in 1958 as the set of all compact quotients of complex Lie groups (equipped with left-invariant metrics compatible with the complex structure). The compactness assumption in the above conjecture is necessary, as there are counterexamples in the noncompact case.

In complex dimension $n = 2$, the conjecture holds. In the $c \leq 0$ case, it was solved by Balas and Gauduchon [3] (see also [2]) in 1985 for Chern connection

and by Sato and Sekigawa [29] in 1990 for Riemannian connection. The general case for $n = 2$ (for both Chern connection and Levi-Civita connection) was solved by Apostolov, Davidov and Mushkarov [1] in 1996, as a corollary of their beautiful classification theorem for compact self-dual Hermitian surfaces.

Theorem 1.2 (Apostolov–Davidov–Mushkarov). *Any compact Hermitian surface with pointwise constant holomorphic sectional curvature with respect to the Levi-Civita connection or the Chern connection must be Kähler.*

For $n \geq 3$, the conjecture is still largely open. Tang [30] confirmed the conjecture under the additional assumption that g is Chern Kähler-like (namely R obeys all Kähler symmetries). Chen, Chen and Nie [10] proved the conjecture under the additional assumption that g is locally conformally Kähler and $c \leq 0$. They also pointed out the necessity of the compactness assumption in the conjecture by explicit examples. Zhou and Zheng [46] proved that any compact Hermitian threefold with vanishing *real bisectional curvature* must be Chern flat. Real bisectional curvature is a curvature notion introduced by Yang and Zheng [40]. It is equivalent to holomorphic sectional curvature H in strength when the metric is Kähler, but is slightly stronger than H when the metric is non-Kähler. Rao and Zheng [28] showed that the conjecture holds if the metric is Bismut Kähler-like, meaning that the curvature of the Bismut connection obeys all Kähler symmetries. Subsequently, Zhao and Zheng [44] proved a very beautiful theorem, which states that all compact Strominger Kähler-like manifolds are pluriclosed. Broder and Tang [7] proved that pluriclosed metrics with vanishing holomorphic sectional curvature on compact Kähler manifolds are Kähler. They also showed that Hermitian metrics with vanishing real bisectional curvature on compact complex manifolds in the Fujiki class $\mathcal{C}$ are Kähler.

On the other hand, if Chern Ricci curvature is nonzero constant λ, namely $\mathrm{Ric} = \lambda g$, then g is Kähler. For these reasons, we naturally consider constant mixed curvature. We assume that α and β are arbitrary fixed real numbers, and $\beta \neq 0$. Mimicking the above folklore conjecture, we propose the following:

Conjecture 1.3. *Let (M^n, g) be a compact Hermitian manifold with $n \geq 2$. Assume that the mixed curvature $\mathcal{C}_{\alpha,\beta}$ equals c where c is a constant. If $c \neq 0$, then g is Kähler.*

Note that if $\mathcal{C}_{\alpha,\beta} = 0$, we may not be able to conclude that $R = 0$. For example, the isosceles Hopf manifold satisfies $\mathcal{C}_{\alpha,\beta} = 0$ with $n\alpha + \beta = 0$, but it does not imply $R = 0$ which we will see in Section 3.

Using similar techniques as in [1], we confirm Conjecture 1.3 in the surface case:

Theorem 1.4. *If (M^2, g) is a compact Hermitian surface with constant mixed curvature $\mathcal{C}_{\alpha,\beta} = c$, then g must be Kähler, unless $c = 0$ and $2\alpha + \beta = 0$. If $c = 0$ and $2\alpha + \beta = 0$, the either M^2 is a Kähler surface, or M^2 is an isosceles Hopf surface.*

In fact, we can prove a stronger result, assuming the mixed curvature $\mathcal{C}_{\alpha,\beta}$ is equal to a pointwise constant f where f is a real-valued function on M^2.

Theorem 1.5. *Assume (M^2, g) is a compact Hermitian surface with pointwise constant mixed curvature $\mathcal{C}_{\alpha,\beta} = f$. If $2\alpha + \beta \neq 0$, then g is Kähler.*

For higher-dimensional cases ($n \geq 3$), we can also restrict ourselves to locally conformally Kähler metrics or Chern Kähler-like metrics. The second main result of this paper is the following theorem, which generalizes Chen, Chen and Nie's result [10] to the mixed curvature case.

Theorem 1.6. *Let (M^n, g) be a compact locally conformal Kähler manifold with constant mixed curvature $\mathcal{C}_{\alpha,\beta} = c$.*

(1) *If $c = 0$ and $n\alpha + \beta \neq 0$, then g is Kähler.*

(2) *If $c < 0$ and $[(n+2)\alpha + 4\beta](n\alpha + \beta)\beta > 0$, then g is Kähler.*

(3) *If $c > 0$ and $[(n+2)\alpha + 4\beta](n\alpha + \beta)\beta < 0$, then g is Kähler.*

If the Hermitian metric is Chern Kähler-like, we can get similar conclusions even without the compactness condition.

Theorem 1.7. *Let (M^n, g) be a Kähler-like manifold with nonzero constant mixed curvature $\mathcal{C}_{\alpha,\beta} = c$. If $(n+2)\alpha + 4\beta \neq 0$, then g is Kähler.*

For other connections, such as the Strominger connection and Gauduchon connection, there has also been some new progress recently. Chen and Zheng [9] proved that any compact Hermitian surface with (pointwise) constant Strominger holomorphic sectional curvature must be either Kähler or an isosceles Hopf surface with an admissible metric. Chen and Nie [8] showed that it is also true for any Hermitian surface with pointwise constant holomorphic sectional curvature with respect to a Gauduchon connection. Additionally, there are interesting discussions regarding the almost Kähler case (see [19; 20] for details).

1.3. ***Semipositive mixed curvature.*** The relationships between the holomorphic sectional curvature, the Ricci curvature, the algebraic positivity of the (anti)canonical line bundles, and some birational invariants of the ambient manifolds are still mysterious. Yang [36] proved that a compact Hermitian manifold with semipositive but not identically zero holomorphic sectional curvature has Kodaira dimension $-\infty$. We extend Yang's result to mixed curvature.

Theorem 1.8. *Let (M^n, g) be a Hermitian manifold with semipositive but not identically zero mixed curvature $\mathcal{C}_{\alpha,\beta}$. If $\beta \geq 0$, $\alpha(n+1) + 2\beta > 0$, then the Kodaira dimension $\kappa(M)$ equals $-\infty$.*

This paper is organized as follows. In Section 2, we give some basic knowledge which will be used in our proofs. In Section 3, we prove Theorems 1.4, 1.5, 1.6 and 1.7. In Section 4, we prove Theorem 1.8.

2. Preliminaries

Let (M, g) be a Hermitian manifold of dimension n and denote by ω the Kähler form associated with g. If ω is closed, that is, if $d\omega = 0$, we call g a Kähler metric. We will directly use g to denote the Kähler form, which should not cause any ambiguity. Denote by ∇ the Chern connection.

Fix any $p \in M$, let $\{e_1, \ldots, e_n\}$ be a frame of (1,0)-tangent vector of M in a neighborhood of p, with $\{\varphi_1, \ldots, \varphi_n\}$ being its dual coframe of (1,0)-form. We will also write $e = {}^t(e_1, \ldots, e_n)$ and $\varphi = {}^t(\varphi_1, \ldots, \varphi_n)$ as column vectors. Let $\langle \cdot, \cdot \rangle$ be the inner product given by Hermitian metric g, and extend it bilinearly over $\mathbb{C}$. Define $|\cdot|^2 = \langle \cdot, \cdot \rangle$. Let θ, τ and Θ be respectively the connection matrix, torsion column vector and curvature matrix under e for the Chern connection. Then the structure equation and Bianchi identities are

$$d\varphi = -{}^t\theta \wedge \varphi + \tau, \quad d\theta = \theta \wedge \theta + \Theta, \quad d\tau = -{}^t\theta \wedge \tau + {}^t\Theta \wedge \varphi, \quad d\Theta = \theta \wedge \Theta - \Theta \wedge \theta.$$

It is well known that the entries of the matrix Θ are all (1,1)-forms, while the entries of the column vector τ are all (2, 0)-forms, under any frame e. Define

$$\Theta_{kl} = R_{i\bar{j}k\bar{l}} \varphi_i \wedge \overline{\varphi}_j.$$

If we choose local holomorphic chart $(z_1, \ldots, z^n)$, we write

$$\omega = \sqrt{-1} \sum_{\substack{i=1 \\ j=1}}^{n} g_{i\bar{j}}\, dz^i \wedge d\bar{z}^j.$$

Then the curvature tensor $R = \{R_{i\bar{j}k\bar{l}}\}$ of the Chern connection is given by

$$R_{i\bar{j}k\bar{l}} = -\frac{\partial^2 g_{k\bar{l}}}{\partial z^i \partial \bar{z}^j} + g^{p\bar{q}} \frac{\partial g_{k\bar{q}}}{\partial z^i} \frac{\partial g_{p\bar{l}}}{\partial \bar{z}^j}.$$

There are four distinct traces of the Chern curvature tensor and, therefore, four distinct Chern Ricci curvatures. The *first* and *second Chern Ricci curvatures* $\mathrm{Ric}^{(1)}_{i\bar{j}} := R_{i\bar{j}} = g^{k\bar{\ell}} R_{i\bar{j}k\bar{\ell}}$ and $\mathrm{Ric}^{(2)}_{k\bar{\ell}} := g^{i\bar{j}} R_{i\bar{j}k\bar{\ell}}$ trace to the same *Chern scalar curvature* u. The *third* and *fourth Chern Ricci curvatures* $\mathrm{Ric}^{(3)}_{i\bar{\ell}} := g^{k\bar{j}} R_{i\bar{j}k\bar{\ell}}$ and $\mathrm{Ric}^{(4)}_{k\bar{j}} := g^{i\bar{\ell}} R_{i\bar{j}k\bar{\ell}}$ trace to what we refer to as the *altered Chern scalar curvature* v, i.e.,

$$u := g^{i\bar{j}}\, \mathrm{Ric}^{(1)}_{i\bar{j}} = g^{k\bar{\ell}}\, \mathrm{Ric}^{(2)}_{k\bar{\ell}}, \quad v := g^{i\bar{\ell}}\, \mathrm{Ric}^{(3)}_{i\bar{\ell}} = g^{k\bar{j}}\, \mathrm{Ric}^{(4)}_{k\bar{j}}.$$

The first Chern Ricci curvature is a (1, 1)-form representing the first Chern class $c_1(M)$. For convenience, we also denote the first, second, third, and fourth Ricci curvatures as $\rho^{(1)}$, $\rho^{(2)}$, $\rho^{(3)}$ and $\rho^{(4)}$ respectively. The components of the Chern torsion T are given by $\mathrm{T}^k_{ij} = g^{k\bar{\ell}}(\partial_i g_{j\bar{\ell}} - \partial_j g_{i\bar{\ell}})$. We write $\eta = \sum_i \eta_i\, dz^i = \sum_{i,k} \mathrm{T}^k_{ik}\, dz^i$ for the

torsion $(1,0)$*-form*. It is well known that

$$\int_M (u-v)dv = \int_M |\eta|^2\, dv.$$

Note that if the g is a Kähler metric, then R obeys all Kähler symmetries, the four Chern Ricci curvatures are the same and $\mathrm{T}=0$.

By the conformal formula for Hermitian metric, the Hermitian $\tilde{g} = e^{2F} g$ satisfies

$$\widetilde{R}_{k\bar{l}i\bar{j}} = e^{2F}(R_{k\bar{l}i\bar{j}} - 2g_{i\bar{j}}F_{k\bar{l}}), \tag{2-1}$$

where $\widetilde{R}$ is the curvature tensor of $\tilde{g}$ and R is the curvature tensor of g.

For Hermitian surfaces (M^2, g), we list a few formulas. By [17, p. 113, (4.2)], we get

$$c_1^2(M) = \frac{1}{8\pi^2}[u^2 - \langle \mathrm{Ric}^{(1)}, \mathrm{Ric}^{(1)}\rangle]g^2. \tag{2-2}$$

From [3] (also see [14]), we have

$$\rho^{(1)} + \rho^{(2)} - 2\mathrm{Re}(\rho^{(3)}) = (u-v)g, \tag{2-3}$$

$$\int_M \langle \rho^{(1)}, \rho^{(2)}\rangle g^2 = \int_M \langle \rho^{(1)}, \mathrm{Re}(\rho^{(3)})\rangle g^2 = \int_M \langle \rho^{(1)}, \rho^{(1)}\rangle g^2 - \int_M u(u-v)g^2. \tag{2-4}$$

We recall the formula of Weyl curvature tensor for a Hermitian surface (M^2, g). One has $W = W^+ + W^-$, and g is said to be *self-dual* if $W^- = 0$. Let $\{e_1, e_2\}$ be a local unitary frame of M. Then

$$\left\{ e_1 \wedge \bar{e}_2,\ e_2 \wedge \bar{e}_1,\ \frac{1}{\sqrt{2}}(e_1 \wedge \bar{e}_1 - e_2 \wedge \bar{e}_2) \right\}$$

forms a basis of the complexification of Λ^2_-, the space on which the Hodge star operator is minus identity. We call this basis the standard basis associated with the unitary frame e. In [1], it was proved that:

Lemma 2.1 [1]. *Let (M^2, g) be a Hermitian surface, and e a local unitary frame. Then under the standard basis associated with e, the components of W^-, the antiself dual part of the Weyl tensor, are given by*

$$\begin{aligned}
W_1^- &= R_{1\bar{2}1\bar{2}},\\
W_2^- &= \frac{1}{\sqrt{2}}(R_{1\bar{2}2\bar{2}} + R_{2\bar{2}1\bar{2}} - R_{1\bar{2}1\bar{1}} - R_{1\bar{1}1\bar{2}}),\\
W_3^- &= \tfrac{1}{6}(R_{1\bar{1}1\bar{1}} + R_{2\bar{2}2\bar{2}} - R_{1\bar{1}2\bar{2}} - R_{2\bar{2}1\bar{1}} - R_{1\bar{2}2\bar{1}} - R_{2\bar{1}1\bar{2}}).
\end{aligned}$$

3. Constant mixed curvature

3.1. ***Hermitian surface.*** We will prove Theorem 1.4 in this section. As in [1], the first step is to show that any Hermitian surface with constant mixed curvature must be self-dual:

Proposition 3.1. *Let (M^2, g) be a Hermitian surface. Assume that the mixed curvature $\mathcal{C}_{\alpha,\beta}$ is pointwise constant, namely $\mathcal{C}_{\alpha,\beta} = f$ where f is a real-valued function on M^2. Then $W_- = 0$, namely, M is self-dual.*

Proof. For any fixed $p \in M$, let $X = e_1$ and $Y = e_2$ be orthonormal $(1,0)$-vectors, and set $\mathcal{C}_{\alpha,\beta} = f(p) = c$. It follows from the average trick that

$$\frac{1}{\mathrm{vol}(\mathbb{S}^3)} \int_{\substack{Z\in T_p^{1,0}M \\ |Z|=1}} \alpha \,\mathrm{Ric}(Z,\bar{Z}) + \beta H(Z)\, d\theta(Z) = (3\alpha+\beta)u + \beta v = 6c. \tag{3-1}$$

Hence

$$(3\alpha+2\beta)(R_{1\bar{1}1\bar{1}}+R_{2\bar{2}2\bar{2}})+(3\alpha+\beta)(R_{1\bar{1}2\bar{2}}+R_{2\bar{2}1\bar{1}})+\beta(R_{1\bar{2}2\bar{1}}+R_{2\bar{1}1\bar{2}})=6c. \tag{3-2}$$

By $\mathcal{C}_{\alpha,\beta}(X) = \mathcal{C}_{\alpha,\beta}(Y) = c$, we have

$$\alpha(R_{1\bar{1}1\bar{1}} + R_{1\bar{1}2\bar{2}}) + \beta R_{1\bar{1}1\bar{1}} = c, \quad \alpha(R_{2\bar{2}1\bar{1}} + R_{2\bar{2}2\bar{2}}) + \beta R_{2\bar{2}2\bar{2}} = c. \tag{3-3}$$

From (3-2), (3-3) and $\beta \neq 0$, we get

$$R_{1\bar{1}1\bar{1}} + R_{2\bar{2}2\bar{2}} - (R_{1\bar{1}2\bar{2}} + R_{2\bar{2}1\bar{1}} + R_{1\bar{2}2\bar{1}} + R_{2\bar{1}1\bar{2}}) = 0. \tag{3-4}$$

So $W_-^3 = 0$. Let $a, b \in \mathbb{C}$ and $|a|^2 + |b|^2 = 1$. Then $aX + bY$ is unit vector. Thus, by carrying out the calculation, we obtain

$$\begin{aligned}
c &= \alpha\,\mathrm{Ric}(aX+bY, \bar{a}\bar{X}+\bar{b}\bar{Y}) + \beta R(aX+bY, \bar{a}\bar{X}+\bar{b}\bar{Y}, aX+bY, \bar{a}\bar{X}+\bar{b}\bar{Y}) \\
&= \alpha\{|a|^2\,\mathrm{Ric}(X,\bar{X}) + a\bar{b}\,\mathrm{Ric}(X,\bar{Y}) + b\bar{a}\,\mathrm{Ric}(Y,\bar{X}) + |b|^2\,\mathrm{Ric}(Y,\bar{Y})\} \\
&\quad + \beta\{|a|^4 R(X,\bar{X},X,\bar{X}) + |b|^4 R(Y,\bar{Y},Y,\bar{Y})) \\
&\quad + |a|^2|b|^2[R(X,\bar{X},Y,\bar{Y}) + R(Y,\bar{Y},X,\bar{X}) + R(X,\bar{Y},Y,\bar{X}) + R(Y,\bar{X},X,\bar{Y})] \\
&\quad + a^2\bar{b}^2 R(X,\bar{Y},X,\bar{Y}) + \bar{a}^2 b^2 R(Y,\bar{X},Y,\bar{X}) \\
&\quad + a^2\bar{a}\bar{b}[R(X,\bar{X},X,\bar{Y}) + R(X,\bar{Y},X,\bar{X})] + \bar{a}^2 ab[R(Y,\bar{X},X,\bar{X}) + R(X,\bar{X},Y,\bar{X})] \\
&\quad + \bar{a}b^2\bar{b}[R(Y,\bar{Y},Y,\bar{X}) + R(Y,\bar{X},Y,\bar{Y})] + a\bar{b}^2 b[R(Y,\bar{Y},X,\bar{Y}) + R(X,\bar{Y},Y,\bar{Y})]\}.
\end{aligned}$$

In fact, by combining (3-4), $\mathcal{C}_{\alpha,\beta}(X) = \mathcal{C}_{\alpha,\beta}(Y) = c$ and $|a|^2 + |b|^2 = 1$, the above can be simplified to

$$\begin{aligned}
0 = {} & \alpha[a\bar{b}\,\mathrm{Ric}(X,\bar{Y}) + b\bar{a}\,\mathrm{Ric}(Y,\bar{X})] \\
& + \beta[a^2\bar{b}^2 R(X,\bar{Y},X,\bar{Y}) + \bar{a}^2 b^2 R(Y,\bar{X},Y,\bar{X})] \\
& + \beta[(a^2\bar{a}\bar{b})A + (\bar{a}^2 ab)B + (\bar{a}b^2\bar{b})C + (a\bar{b}^2 b)D],
\end{aligned} \tag{3-5}$$

where we assume that

$$\begin{aligned} &A = R(X,\overline{X},X,\overline{Y}) + R(X,\overline{Y},X,\overline{X}); \quad B = R(Y,\overline{X},X,\overline{X}) + R(X,\overline{X},Y,\overline{X}); \\ &C = R(Y,\overline{Y},Y,\overline{X}) + R(Y,\overline{X},Y,\overline{Y}); \quad\quad D = R(Y,\overline{Y},X,\overline{Y}) + R(X,\overline{Y},Y,\overline{Y}). \end{aligned} \tag{3-6}$$

Now we choose special a and b. Set $a = \frac{1}{\sqrt{2}}$, $b = \frac{1}{\sqrt{2}}$, and $a = \frac{-1}{\sqrt{2}}$, $b = \frac{1}{\sqrt{2}}$. Then we get

$$R(X,\overline{Y},X,\overline{Y}) + R(Y,\overline{X},Y,\overline{X}) = 0. \tag{3-7}$$

Setting $a = \frac{\sqrt{2}}{2}$, $b = \frac{1-i}{2}$, and $a = \frac{-\sqrt{2}}{2}$, $b = \frac{1-i}{2}$. We have

$$R(X,\overline{Y},X,\overline{Y}) - R(Y,\overline{X},Y,\overline{X}) = 0. \tag{3-8}$$

Hence $R_{1\bar{2}1\bar{2}} = R(X,\overline{Y},X,Y) = 0$. So $W^1_- = 0$. Now (3-5) is further simplified to

$$\begin{aligned} 0 = {} & \alpha[a\bar{b}\,\mathrm{Ric}(X,\overline{Y}) + b\bar{a}\,\mathrm{Ric}(Y,\overline{X})] \\ & + \beta[(a^2\bar{a}\bar{b})A + (\bar{a}^2ab)B + (\bar{a}b^2\bar{b})C + (a\bar{b}^2b)D]. \end{aligned} \tag{3-9}$$

Setting $a = \frac{\sqrt{2}}{2}$, $b = -\frac{\sqrt{2}}{2}i$, and $a = \frac{\sqrt{3}}{2}$, $b = \frac{1}{2}i$, yields

$$A + C = B + D.$$

Setting $a = -\frac{\sqrt{2}}{2}$, $b = \frac{\sqrt{2}}{2}$, and $a = \frac{\sqrt{3}}{2}$, $b = \frac{1}{2}$, yields

$$A + B = C + D.$$

From these, we get

$$R_{1\bar{2}2\bar{2}} + R_{2\bar{2}1\bar{2}} - R_{1\bar{2}1\bar{1}} - R_{1\bar{1}1\bar{2}} = D - A = 0, \tag{3-10}$$

namely, $W^2_- = 0$. We conclude that $W_- = 0$. □

Next, we consider the case under conformal changes. We first establish a lemma.

Lemma 3.2. *Let (M^n, g) be a compact Hermitian manifold. Assume that the mixed curvature $\mathcal{C}_{\alpha,\beta}$ is pointwise constant, namely $\mathcal{C}_{\alpha,\beta} = f$ where f is a real-valued function on M. Then we have*

$$[\alpha(n+2)+\beta]\,\mathrm{Ric}^{(1)} + \beta\,\mathrm{Ric}^{(2)} + 2\beta\,\mathrm{Re}(\mathrm{Ric}^{(3)}) = [2(n+1)f - \alpha u]g. \tag{3-11}$$

Proof. For $p \in M$, let $X = X^i \frac{\partial}{\partial x^i} \in T^{(1,0)}_p M$. Since $\mathcal{C}_{\alpha,\beta}(X) = f(p) = c$ at p,

$$\alpha\,\mathrm{Ric}(X,\overline{X})|X|^2 + \beta R(X,\overline{X},X,\overline{X}) = c|X|^4. \tag{3-12}$$

Specifically, we have

$$\begin{aligned}(3\text{-}13)\quad &\alpha X^i \overline{X}^j X^k \overline{X}^l (R_{i\bar{j}} g_{k\bar{l}} + R_{k\bar{j}} g_{i\bar{l}} + R_{i\bar{l}} g_{k\bar{j}} + R_{k\bar{l}} g_{i\bar{j}}) \\ &\qquad + \beta X^i \overline{X}^j X^k \overline{X}^l (R_{i\bar{j}k\bar{l}} + R_{k\bar{j}i\bar{l}} + R_{i\bar{l}k\bar{j}} + R_{k\bar{l}i\bar{j}}) \\ &= 2c X^i \overline{X}^j X^k \overline{X}^l (g_{i\bar{j}} g_{k\bar{l}} + g_{i\bar{l}} g_{k\bar{j}}) \qquad .\end{aligned}$$

Thus the above equality holds for any $X = X^i \frac{\partial}{\partial x^i}$ if and only if

$$\begin{aligned}(3\text{-}14)\quad &\alpha(R_{i\bar{j}} g_{k\bar{l}} + R_{k\bar{j}} g_{i\bar{l}} + R_{i\bar{l}} g_{k\bar{j}} + R_{k\bar{l}} g_{i\bar{j}}) + \beta(R_{i\bar{j}k\bar{l}} + R_{k\bar{j}i\bar{l}} + R_{i\bar{l}k\bar{j}} + R_{k\bar{l}i\bar{j}}) \\ &\qquad = 2c(g_{i\bar{j}} g_{k\bar{l}} + g_{i\bar{l}} g_{k\bar{j}}).\end{aligned}$$

Taking the trace, we obtain (3-11). □

Proposition 3.3. *Let (M^2, g) be a compact Kähler surface of constant holomorphic sectional curvature $H = c$. Let $\tilde{g} = e^{2F} g$ is a conformal metric with pointwise constant mixed curvature $\mathcal{C}_{\alpha,\beta} = f$, where F is a smooth real-valued function on M.*

(1) *If $2\alpha + \beta \neq 0$, then F is constant, namely, $\tilde{g}$ is Kähler.*

(2) *If f is constant, then F is constant unless $f \equiv 0$ and $2\alpha + \beta = 0$.*

Proof. We denote the first, second and third Ricci curvatures of $\tilde{g}$ as $\tilde{\rho}^{(1)}$, $\tilde{\rho}^{(2)}$, $\tilde{\rho}^{(3)}$ respectively. Let u, v be the scalar and altered scalar curvatures of g, and $\tilde{u}$, $\tilde{v}$ are the scalar and altered scalar curvatures of $\tilde{g}$.

By (2-3) and (3-11), we have

$$\tilde{\rho}^{(1)} + \tilde{\rho}^{(2)} - 2\,\mathrm{Re}(\tilde{\rho}^{(3)}) = (\tilde{u} - \tilde{v})\tilde{g}, \tag{3-15}$$

$$[4\alpha + \beta]\tilde{\rho}^{(1)} + \beta\tilde{\rho}^{(2)} + 2\beta\,\mathrm{Re}(\tilde{\rho}^{(3)}) = [6f - \alpha\tilde{u}]\tilde{g}, \tag{3-16}$$

and since $[3\alpha + \beta]\tilde{u} + \beta\tilde{v} = 6f$, we obtain

$$2\alpha\tilde{\rho}^{(1)} + 2\beta\,\mathrm{Re}(\tilde{\rho}^{(3)}) = (\alpha\tilde{u} + \beta\tilde{v})\tilde{g}. \tag{3-17}$$

By (2-4),

$$\int_M \langle \tilde{\rho}^{(1)}, \mathrm{Re}(\tilde{\rho}^{(3)}) \rangle \tilde{g}^2 = \int_M \langle \tilde{\rho}^{(1)}, \tilde{\rho}^{(1)} \rangle \tilde{g}^2 - \int_M \tilde{u}(\tilde{u} - \tilde{v})\tilde{g}^2. \tag{3-18}$$

Combining (3-17) and (3-18), we get

$$(\alpha + 2\beta) \int_M \tilde{u}^2 \tilde{g}^2 = 2(\alpha + \beta) \int_M \langle \tilde{\rho}^{(1)}, \tilde{\rho}^{(1)} \rangle \tilde{g}^2 + \beta \int_M (\tilde{u}\tilde{v})\tilde{g}^2. \tag{3-19}$$

Setting $c_1^2 = \int_M c_1^2(M)$, according to (2-2), we have

$$c_1^2 = \frac{1}{8\pi^2} \left[\int_M \tilde{u}^2 \tilde{g}^2 - \int_M \langle \tilde{\rho}^{(1)}, \tilde{\rho}^{(1)} \rangle \tilde{g}^2 \right]. \tag{3-20}$$

Hence, we have

$$16\pi^2(\alpha+\beta)c_1^2=\alpha\int_M\tilde{u}^2\tilde{g}^2+\beta\int_M(\tilde{u}\tilde{v})\tilde{g}^2. \tag{3-21}$$

On the other hand, for the Kähler metric g with $H=c$, we have $u=v=3c$. So

$$c_1^2=\frac{1}{16\pi^2}\int_M u^2g^2. \tag{3-22}$$

From (3-21) and (3-22), we have

$$(\alpha+\beta)\int_M u^2g^2=\alpha\int_M\tilde{u}^2\tilde{g}^2+\beta\int_M(\tilde{u}\tilde{v})\tilde{g}^2. \tag{3-23}$$

Since $\tilde{g}=e^{2F}g$, by (2-1), we obtain

$$e^{2F}\tilde{u}=u-4\Delta F;\quad e^{2F}\tilde{v}=v-2\Delta F. \tag{3-24}$$

From these, we get

$$(\alpha+\beta)\int_M u^2g^2=\alpha\int_M(u-4\Delta F)^2g^2+\beta\int_M(u-4\Delta F)(u-2\Delta F)g^2. \tag{3-25}$$

Then

$$(2\alpha+\beta)\int_M(\Delta F)^2g^2=0, \tag{3-26}$$

where we use $\int_M\Delta Fg^2=0$ and $u=3c$. This means that if $2\alpha+\beta\neq 0$, then $\Delta F=0$. So F is constant.

If f is constant, we assume that $f=\tilde{c}$. We only consider the case where $2\alpha+\beta=0$. We have

$$\alpha(\tilde{u}-2\tilde{v})=6\tilde{c}. \tag{3-27}$$

From (3-24) and (3-27), we get

$$2\tilde{c}e^{2F}=-\alpha c. \tag{3-28}$$

If $\tilde{c}\neq 0$, then F is constant. □

Recall that Hopf manifolds are compact complex manifolds whose universal cover is $\mathbb{C}^n\setminus\{0\}$. An isosceles Hopf manifold (see [9]) means $M_\phi^n=\mathbb{C}^n\setminus\{0\}/\langle\phi\rangle$ where

$$\phi:(z_1,z_2,\ldots,z_n)\to(a_1z_1,a_2z_2,\ldots,a_nz_n) \tag{3-29}$$

with $0<|a_1|=|a_2|=\cdots=|a_n|<1$. The standard Hopf metric g with Kähler form $\omega_g=\sqrt{-1}\frac{\partial\partial|z|^2}{|z|^2}$ descends down to M_ϕ^n. Write $h=|z|$, $e_i=h\frac{\partial}{\partial z_i}$, $\varphi_i=\frac{1}{h}dz_i$. Then e

becomes a unitary frame with φ the dual coframe. By the structure equations,

$$\theta_{ij} = \frac{1}{h}(\bar{\partial} - \partial)h\delta_{ij}. \tag{3-30}$$

Hence

$$\Theta_{ij} = -\frac{2}{h^2}[\partial h \wedge (\bar{\partial}h)]\delta_{ij} + \frac{2}{h}(\partial\bar{\partial}h)\delta_{ij}. \tag{3-31}$$

We get

$$R_{i\bar{j}k\bar{l}} = -\frac{\bar{z}_i z_j}{|z|^2}\delta_{kl} + \delta_{ij}\delta_{kl}, \tag{3-32}$$

where $h_i = \frac{\partial h}{\partial z^i}$. From this, the first Ricci curvature is

$$R_{i\bar{j}} = -n\frac{\bar{z}_i z_j}{|z|^2} + n\delta_{ij}. \tag{3-33}$$

So we get scalar curvature $u = n^2 - n$ and altered scalar curvature $v = n - 1$. It implies

$$(R_{i\bar{j}}g_{k\bar{l}} + R_{k\bar{j}}g_{i\bar{l}} + R_{i\bar{l}}g_{k\bar{j}} + R_{k\bar{l}}g_{i\bar{j}}) - n(R_{i\bar{j}k\bar{l}} + R_{k\bar{j}i\bar{l}} + R_{i\bar{l}k\bar{j}} + R_{k\bar{l}i\bar{j}}) = 0. \tag{3-34}$$

So when $n\alpha + \beta = 0$, the Hopf manifold has $\mathcal{C}_{\alpha,\beta} = 0$. Note that the Hopf manifold is not Chern flat. We also obtain the following lemma.

Lemma 3.4. *Let M^2 be a Hopf surface with standard Hopf metric g. Let $\tilde{g} = e^{2F}g$ be a conformal metric with pointwise constant mixed curvature $\mathcal{C}_{\alpha,\beta} = f$, where F is a smooth real-valued function on M. Then the metric $\tilde{g}$ does not exist, unless $f \equiv 0$ and $2\alpha + \beta = 0$.*

Proof. According to (3-21) and (3-33), it is easy to deduce that $c_1^2 = 0$. From (3-21), (3-24) and $u = 2v$, we have $\tilde{u} = 2\tilde{v}$ and

$$\begin{aligned} 0 &= 4\alpha\int_M (v - 2\Delta F)^2 g^2 + 2\beta\int_M (v - 2\Delta F)^2 g^2 \\ &= 2(2\alpha + \beta)\int_M (v - 2\Delta F)^2 g^2. \end{aligned} \tag{3-35}$$

This means that if $2\alpha + \beta \neq 0$, then $\Delta F = \frac{v}{2} = \frac{1}{2}$, namely, F is constant. This of course will contradict with $v \neq 0$.

On the other hand, it is easy to deduce that

$$(2\alpha + \beta)\tilde{v} = 2f. \tag{3-36}$$

When $2\alpha + \beta = 0$, f must be zero. □

Let $\tilde{g} = e^{2F} g$ be a conformal metric with pointwise constant mixed curvature $\mathcal{C}_{\alpha,\beta} = f$ on M^n. We denote curvatures of g as $R_{i\bar{j}k\bar{l}}$. Generally, from (2-1) and (3-14), we can obtain

$$\begin{aligned}(3\text{-}37)\quad &\alpha(R_{i\bar{j}}g_{k\bar{l}} + R_{k\bar{j}}g_{i\bar{l}} + R_{i\bar{l}}g_{k\bar{j}} + R_{k\bar{l}}g_{i\bar{j}}) + \beta(R_{i\bar{j}k\bar{l}} + R_{k\bar{j}i\bar{l}} + R_{i\bar{l}k\bar{j}} + R_{k\bar{l}i\bar{j}})\\ &-2(n\alpha+\beta)[g_{i\bar{j}}F_{k\bar{l}}+g_{k\bar{j}}F_{i\bar{l}}+g_{i\bar{l}}F_{k\bar{j}}+g_{k\bar{l}}F_{i\bar{j}}] = 2fe^{2F}(g_{i\bar{j}}g_{k\bar{l}}+g_{i\bar{l}}g_{k\bar{j}}).\end{aligned}$$

Now we are ready to prove the main result of this article, Theorem 1.4.

Proof of Theorem 1.4. Let (M^2, g) be a compact Hermitian surface with constant mixed curvature $\mathcal{C}_{\alpha,\beta} = c$. By Proposition 3.1, we know that M^2 is self-dual. Hence, by Theorem 1' in [1], g is conformal to a metric h on M^2, which is either

(1) a complex space form, or

(2) the nonflat, conformally flat Kähler metric, or

(3) an isosceles Hopf surface.

Write $g = e^{2F}h$.

For case (1), by Proposition 3.3, if $c \neq 0$ or $2\alpha + \beta \neq 0$, we get that F is constant. Hence g is Kähler.

For case (2), M^2 is a holomorphic bundle over a curve C of genus at least 2, and h is locally a product metric where the factors have constant holomorphic sectional curvature -1 (for the base curve) or 1 (for the fiber). Following [1], let (z_1, z_2) be holomorphic coordinates, and let ω_h be the Kähler form

$$(3\text{-}38)\qquad \omega_h = 2\sqrt{-1}\frac{dz^1 \wedge d\bar{z}_1}{(1-|z_1|^2)^2} + 2\sqrt{-1}\frac{dz^2 \wedge d\bar{z}_2}{(1+|z_2|^2)^2} = \omega_1 + \omega_2.$$

Let e be a unitary frame, such that e_i is parallel to $\frac{\partial}{\partial z_i}$. Then under the frame e,

$$(3\text{-}39)\qquad R_{1\bar{1}1\bar{1}} = -1; \quad R_{2\bar{2}2\bar{2}} = 1;$$

and the other curvatures are zero. So $R_{11} = -1$, $R_{2\bar{2}} = 1$, and the other Ricci curvatures are zero. Define $R_{i\bar{j}} = (-1)^i \delta_{ij}$. Here, let's assume that g has pointwise constant mixed curvature $\mathcal{C}_{\alpha,\beta} = f$, where f is a real-valued function on M^2. According to (3-37), we can deduce that

$$(3\text{-}40)\qquad \begin{aligned}-(\alpha+\beta) - 2(2\alpha+\beta)F_{1\bar{1}} &= fe^{2F};\\ (\alpha+\beta) - 2(2\alpha+\beta)F_{2\bar{2}} &= fe^{2F};\\ (2\alpha+\beta)F_{1\bar{2}} &= 0.\end{aligned}$$

Suppose $2\alpha + \beta \neq 0$, and set $F_{1\bar{1}} = x$, $F_{2\bar{2}} = y$. So $y - x = \frac{\alpha+\beta}{2\alpha+\beta}$, $x_{\bar{2}} = F_{1\bar{1}2} = 0$ and $y_{\bar{1}} = F_{2\bar{2}\bar{1}} = 0$. Then we have

$$(3\text{-}41)\qquad \Delta x = x_{1\bar{1}} = y_{1\bar{1}} = 0; \quad \Delta y = 0.$$

This implies that ΔF is constant. Hence F is constant and g is Kähler. Note that in this case we can only have $f \equiv 0$ and $\alpha + \beta = 0$. If $2\alpha + \beta = 0$, from (3-40), $f = 0$ and $\alpha + \beta$=0, which contradicts with $2\alpha + \beta = 0$.

For case (3), by Lemma 3.4, if $c \neq 0$ or $2\alpha + \beta \neq 0$, the Hermitian metric g does not exist on isosceles Hopf surface. □

Proof of Theorem 1.5. From the above proof of Theorem 1.4, combining with Proposition 3.3 and Lemma 3.4, we can easily draw the conclusion. □

3.2. ***High-dimensional Hermitian manifold.*** We first consider the case of compact locally conformal Kähler manifolds with constant mixed curvature. We also follow similar steps as in [10]. We first present the following lemma.

Lemma 3.5. *Let (M^n, g) be a compact locally conformal Kähler manifold with pointwise constant mixed curvature $\mathcal{C}_{\alpha,\beta} = f$.*

(1) *If $f \equiv 0$ and $n\alpha + \beta \neq 0$, then g is globally Kähler.*

(2) *if $f \leq 0$ and $[(n+2)\alpha + 4\beta](n\alpha + \beta)\beta > 0$, then g is globally Kähler.*

(3) *if $f \geq 0$ and $[(n+2)\alpha + 4\beta](n\alpha + \beta)\beta < 0$, then g is globally Kähler.*

Proof. We denote the first, second and third Ricci curvatures of g as $\rho^{(1)}$, $\rho^{(2)}$, $\rho^{(3)}$ respectively. Let u, v be the scalar and altered scalar curvature of g. Since g is locally conformal Kähler, $\partial\partial^* g = \bar{\partial}\bar{\partial}^* g$ and

$$\rho^{(1)} = \rho^{(3)} + \partial\partial^* g, \tag{3-42}$$

where from (3.12) in [10]. So $\rho^{(3)}$ is a real closed (1,1)-form. According to Proposition 3.2 in [10], we also have

$$\rho^{(1)} - \rho^{(2)} = \frac{1}{n-1}[(v-u)g + n\partial\partial^* g]. \tag{3-43}$$

By Lemma 3.1, we get

$$[\alpha(n+2) + \beta]\rho^{(1)} + \beta\rho^{(2)} + 2\beta\,\mathrm{Re}(\rho^{(3)}) = [2(n+1)f - \alpha u]g; \tag{3-44}$$

$$[\alpha(n+1) + \beta]u + \beta v = n(n+1)f. \tag{3-45}$$

From these, we obtain

$$[\alpha(n+2) + 4\beta]\rho^{(3)} = \frac{1}{n-1}\{[(3n-2)(n+1)f - 2(n\alpha + \beta)u]g + [\beta(2-n) - \alpha(n+2)(n-1)]\partial\partial^* g\}. \tag{3-46}$$

Hence, we differentiate (3-46) to obtain

$$d\{[(3n-2)(n+1)f - 2(n\alpha + \beta)u]g\} = 0. \tag{3-47}$$

Let $\eta = (3n-2)(n+1)f - 2(n\alpha+\beta)u$. Then we will prove that η is either everywhere nonzero or $\eta \equiv 0$.

(1) If η is everywhere nonzero, then g is globally Kähler.

(2) If $\eta = 0$ at some point p on M, then $\eta \equiv 0$ on M (see [10]). Therefore, if $\eta \equiv 0$ we have the following cases:

(a) Suppose $f \equiv 0$ and $n\alpha+\beta \neq 0$. Then $u = v = 0$. This implies that g is balanced. Since g is also locally conformal Kähler, we get g is Kähler.

(b) Suppose $f \leq 0$ and $[(n+2)\alpha+4\beta](n\alpha+\beta)\beta > 0$. We only consider the case where there is at least one point p such that $f(p) < 0$. We have

$$u = \frac{(3n-2)(n+1)f}{2(n\alpha+\beta)}, \quad v = \frac{-(n+1)[(n+2)(n-1)\alpha+(n-2)\beta]f}{2\beta(n\alpha+\beta)}. \tag{3-48}$$

Hence, we get

$$v - u = \frac{-(n+1)(n-1)[(n+2)\alpha+4\beta]f}{2(n\alpha+\beta)\beta}. \tag{3-49}$$

So $\int_M (v-u)g^n > 0$, which is contradiction with a well-known result that $\int_M (v-u)g^n \leq 0$. If $f \geq 0$ and $[(n+2)\alpha+4\beta](n\alpha+\beta)\beta < 0$, this would also lead to a contradiction.

In this way, we have completed the proof. □

Proof of Theorem 1.6. Let g be a compact locally conformal Kähler manifold with constant mixed curvature $\mathcal{C}_{\alpha,\beta} = c$. Under the assumptions of Lemma 3.5, we know that g is globally Kähler. We assume that h is Kähler metric where $dh = 0$, and $g = e^{2F}h$. Set

$$A = [(3n-2)(n+1)c - 2(n\alpha+\beta)u]e^{2F}. \tag{3-50}$$

By (3-47), we get that $d(Ah) = 0$. So A is constant. Since h is Kähler, we get

$$e^{2F}u = e^{2F}v + 2(1-n)\Delta_h F. \tag{3-51}$$

We also have that

$$(\alpha(n+1)+\beta)u + \beta v = n(n+1)c. \tag{3-52}$$

From (3-50), (3-51) and (3-52), we can easily deduce that

$$\begin{aligned} ce^{2F}(n+1)(n-1)[(n+2)\alpha+4\beta] \\ = [\alpha(n+1)+2\beta]A + 4\beta(1-n)(n\alpha+\beta)\Delta_h F. \end{aligned} \tag{3-53}$$

(1) If $c = 0$ and $n\alpha+\beta \neq 0$, by integrating (3-53) with respect to the metric h, we obtain that F is constant. This implies that g is Kähler.

(2) If $c < 0$, $[(n+2)\alpha + 4\beta](n\alpha + \beta)\beta > 0$, we will show that F is constant. Set

$$B = [\alpha(n+1) + 2\beta]A. \tag{3-54}$$

By integrating (3-53) with respect to the metric h, we have

$$c(n+1)(n-1)[(n+2)\alpha + 4\beta] \int_M e^{2F} h^n = B \cdot \mathrm{Vol}(M). \tag{3-55}$$

On the other hand, we also have that

$$\Delta_h F = -\tfrac{1}{2} e^{2F} (\Delta_h e^{-2F}) + 2e^{2F} |\partial F|_g^2. \tag{3-56}$$

Substituting (3-56) into (3-53), and then integrating it, we obtain

$$\begin{aligned} B \int_M e^{-2F} h^n = {} & c(n+1)(n-1)[(n+2)\alpha + 4\beta] \,\mathrm{Vol}(M) \\ & + 8\beta(n-1)(n\alpha + \beta) \int_M |\partial F|_g^2 h^n. \end{aligned} \tag{3-57}$$

By (3-55), (3-56) and the Cauchy–Schwarz inequality, we have

$$(\mathrm{Vol}(M))^2 \le \int_M e^{2F} h^n \int_M e^{-2F} h^n \le (\mathrm{Vol}(M))^2.$$

The equality holds if and only if F is constant. Therefore g is Kähler.

(3) If $c > 0$ and $[(n+2)\alpha + 4\beta](n\alpha + \beta)\beta < 0$, the proof is similar to case (2). □

Now we will prove the Kähler-like case.

Proof of Theorem 1.7. This proof is relatively simple. Let (M^n, g) be a Kähler-like manifold with constant mixed curvature $\mathcal{C}_{\alpha,\beta} = c$. Then $\mathrm{Ric}^{(1)} = \mathrm{Ric}^{(2)} = \mathrm{Ric}^{(3)}$, $u = v$ and by Lemma 3.2, we have

$$\begin{aligned} [(n+2)\alpha + 4\beta] \,\mathrm{Ric}^{(1)} &= [2(n+1)c - \alpha u]g, \\ [(n+1)\alpha + 2\beta]u &= n(n+1)c. \end{aligned}$$

When $(n+2)\alpha + 4\beta \neq 0$, we get that

$$[(n+1)\alpha + 2\beta] \,\mathrm{Ric}^{(1)} = (n+1)cg.$$

If $c \neq 0$, then g is Kähler. □

4. Semipositive mixed curvature

In this section, we provide the proof of Theorem 1.8.

Proof of Theorem 1.8. Let $g_{\mathrm{G}} = f_0^{1/(n-1)} g$ be the Gauduchon metric (i.e., g_{G} satisfies $\partial\bar{\partial}\omega_{\mathrm{G}}^{n-1} = 0$) in the conformal class of g, where $f_0 \in \mathcal{C}^\infty(X, \mathbb{R})$ is a strictly positive function (such a metric always exists [15]). Let $\omega_{\mathrm{G}}(\cdot, \cdot) := g_{\mathrm{G}}(J\cdot, \cdot)$. Let S_{g_G} and $\widetilde{S}_{g_G}$ be the scalar and altered scalar curvatures of g_G, and let S_g and $\widetilde{S}_g$ be the scalar and altered scalar curvatures of g.

Then

$$\begin{aligned}\int_X S_{g_{\mathrm{G}}}\omega_{\mathrm{G}}^n &= n\int_X \mathrm{Ric}^{(1)}_{g_{\mathrm{G}}}\wedge\omega_{\mathrm{G}}^{n-1}\\ &= n\int_X \mathrm{Ric}^{(1)}_{g}\wedge\omega_{\mathrm{G}}^{n-1} = n\int_X f_0\,\mathrm{Ric}^{(1)}_{g}\wedge\omega_{g}^{n-1} = \int_X f_0 S_g\omega_g^n.\end{aligned}$$

A similar computation (see, e.g., [36]) then gives $\int_X \widetilde{S}_{g_{\mathrm{G}}}\omega_{\mathrm{G}}^n = \int_X f_0\widetilde{S}_g\omega_g^n$. Let η_{G} denote the torsion $(1,0)$-form of g_{G}. We then compute

$$\begin{aligned}(\alpha(n+1)+2\beta)&\int_X S_{g_{\mathrm{G}}}\omega_{\mathrm{G}}^n\\ &= \int_X\big((\alpha(n+1)+\beta)S_{g_{\mathrm{G}}}+\beta\widetilde{S}_{g_{\mathrm{G}}}\big)\omega_{\mathrm{G}}^n + \int_X\beta(S_{g_{\mathrm{G}}}-\widetilde{S}_{g_{\mathrm{G}}})\omega_{\mathrm{G}}^n\\ &= \int_X f_0\big((\alpha(n+1)+\beta)S_g+\beta\widetilde{S}_g\big)\omega_g^n + \beta\int_X|\eta_{\mathrm{G}}|^2_{g_{\mathrm{G}}}\omega_{\mathrm{G}}^n.\end{aligned}$$

It follows from the average trick that

$$\frac{1}{\mathrm{vol}(\mathbb{S}^{2n-1})}\int_{\substack{Z\in T^{1,0}_pM\\ |Z|=1}}\alpha\,\mathrm{Ric}_g(Z,\bar{Z})+\beta H_g(Z)d\theta(Z) = \frac{((n+1)\alpha+\beta)S_g+\beta\widetilde{S}_g}{n(n+1)}.$$

If g has semipositive but not identically zero mixed curvature $\mathcal{C}_{\alpha,\beta}$ and $\beta\geq 0$, $\alpha(n+1)+2\beta>0$, then we have

$$\int_X S_{g_{\mathrm{G}}}\omega_{\mathrm{G}}^n > 0.$$

This means that the Gauduchon metric g_{G} has positive total scalar curvature. By Theorem 1.3 in [38], we know that M admits a Hermitian metric with positive scalar curvature, which implies that the Kodaira dimension $\kappa(M)$ equals $-\infty$. □

Acknowledgement

The author is grateful to Professor Fangyang Zheng for constant encouragement and support. We also thank the referees for helpful comments which enhanced the presentation.

References

[1] V. Apostolov, J. Davidov, and O. Mushkarov, "Compact self-dual Hermitian surfaces", *Trans. Amer. Math. Soc.* **348**:8 (1996), 3051–3063. MR

[2] A. Balas, "Compact Hermitian manifolds of constant holomorphic sectional curvature", *Math. Z.* **189**:2 (1985), 193–210. MR

[3] A. Balas and P. Gauduchon, "Any Hermitian metric of constant nonpositive (Hermitian) holomorphic sectional curvature on a compact complex surface is Kähler", *Math. Z.* **190**:1 (1985), 39–43. MR

[4] W. M. Boothby, "Hermitian manifolds with zero curvature", *Michigan Math. J.* **5** (1958), 229–233. MR

[5] K. Broder and K. Tang, "On the weighted orthogonal Ricci curvature", *J. Geom. Phys.* **193** (2023), art. id. 104783. MR

[6] K. Broder and K. Tang, "(ε, δ)-quasi-negative curvature and positivity of the canonical bundle", *J. Geom. Anal.* **34**:6 (2024), art. id. 180. MR

[7] K. Broder and K. Tang, "On Hermitian manifolds with vanishing curvature", *Math. Z.* **309**:1 (2025), art. id. 12. MR

[8] H. Chen and X. Nie, "Compact Hermitian surfaces with pointwise constant Gauduchon holomorphic sectional curvature", *Math. Z.* **302**:3 (2022), 1721–1737. MR

[9] S. Chen and F. Zheng, "On Strominger space forms", *J. Geom. Anal.* **32**:4 (2022), art. id. 141. MR

[10] H. Chen, L. Chen, and X. Nie, "Chern–Ricci curvatures, holomorphic sectional curvature and Hermitian metrics", *Sci. China Math.* **64**:4 (2021), 763–780. MR

[11] J. Chu, M.-C. Lee, and L.-F. Tam, "Kähler manifolds and mixed curvature", *Trans. Amer. Math. Soc.* **375**:11 (2022), 7925–7944. MR

[12] J. Chu, M.-C. Lee, and J. Zhu, "On Kähler manifolds with non-negative mixed curvature", *J. Reine Angew. Math.* **827** (2025), 313–338. MR

[13] S. Diverio and S. Trapani, "Quasi-negative holomorphic sectional curvature and positivity of the canonical bundle", *J. Differential Geom.* **111**:2 (2019), 303–314. MR

[14] P. Gauduchon, "La topologie d'une surface hermitienne d'Einstein", *C. R. Acad. Sci. Paris Sér. A-B* **290**:11 (1980), A509–A512. MR

[15] P. Gauduchon, "La 1-forme de torsion d'une variété hermitienne compacte", *Math. Ann.* **267**:4 (1984), 495–518. MR

[16] G. Heier and B. Wong, "On projective Kähler manifolds of partially positive curvature and rational connectedness", *Doc. Math.* **25** (2020), 219–238. MR

[17] S. Kobayashi, *Differential geometry of complex vector bundles*, Publications of the Mathematical Society of Japan **15**, Princeton Univ. Press, 1987. MR

[18] M.-C. Lee and J. Streets, "Complex manifolds with negative curvature operator", *Int. Math. Res. Not.* **2021**:24 (2021), 18520–18528. MR

[19] M. Lejmi and M. Upmeier, "Closed almost Kähler 4-manifolds of constant non-negative Hermitian holomorphic sectional curvature are Kähler", *Tohoku Math. J.* (2) **72**:4 (2020), 581–594. MR

[20] M. Lejmi and L. Vezzoni, "4-dimensional almost-Kähler Lie algebras of constant Hermitian holomorphic sectional curvature are Kähler", *J. Lie Theory* **29**:1 (2019), 181–190. MR

[21] S.-i. Matsumura, "On the image of MRC fibrations of projective manifolds with semi-positive holomorphic sectional curvature", *Pure Appl. Math. Q.* **16**:5 (2020), 1419–1439. MR

[22] S.-i. Matsumura, "On projective manifolds with semi-positive holomorphic sectional curvature", *Amer. J. Math.* **144**:3 (2022), 747–777. MR

[23] L. Ni, "The fundamental group, rational connectedness and the positivity of Kähler manifolds", *J. Reine Angew. Math.* **774** (2021), 267–299. MR

[24] L. Ni, "Liouville theorems and a Schwarz lemma for holomorphic mappings between Kähler manifolds", *Comm. Pure Appl. Math.* **74**:5 (2021), 1100–1126. MR

[25] L. Ni, "Holonomy and the Ricci curvature of complex Hermitian manifolds", *J. Geom. Anal.* **35**:1 (2025), art. id. 30. MR

[26] L. Ni and F. Zheng, "Comparison and vanishing theorems for Kähler manifolds", *Calc. Var. Partial Differential Equations* **57**:6 (2018), art. id. 151. MR

[27] L. Ni and F. Zheng, "Positivity and the Kodaira embedding theorem", *Geom. Topol.* **26**:6 (2022), 2491–2505. MR

[28] P. P. Rao and F. Y. Zheng, "Pluriclosed manifolds with constant holomorphic sectional curvature", *Acta Math. Sin.* (*Engl. Ser.*) **38**:6 (2022), 1094–1104. MR

[29] T. Sato and K. Sekigawa, "Hermitian surfaces of constant holomorphic sectional curvature", *Math. Z.* **205**:4 (1990), 659–668. MR

[30] K. Tang, "Holomorphic sectional curvature and Kähler-like metric", *Sci. Sin. Math.* **51**:12 (2021), 1–12. In Chinese.

[31] K. Tang, "On almost nonpositive k-Ricci curvature", *J. Geom. Anal.* **32**:12 (2022), art. id. 306. MR

[32] K. Tang, "Quasi-positive mixed curvature, vanishing theorems, and rational connectedness", *Bull. Lond. Math. Soc.* **58**:2 (2026), art. id. e70294. MR

[33] V. Tosatti and X. Yang, "An extension of a theorem of Wu–Yau", *J. Differential Geom.* **107**:3 (2017), 573–579. MR

[34] D. Wu and S.-T. Yau, "Negative holomorphic curvature and positive canonical bundle", *Invent. Math.* **204**:2 (2016), 595–604. MR

[35] D. Wu and S.-T. Yau, "A remark on our paper 'Negative holomorphic curvature and positive canonical bundle"', *Comm. Anal. Geom.* **24**:4 (2016), 901–912. MR

[36] X. Yang, "Hermitian manifolds with semi-positive holomorphic sectional curvature", *Math. Res. Lett.* **23**:3 (2016), 939–952. MR

[37] X. Yang, "RC-positivity, rational connectedness and Yau's conjecture", *Camb. J. Math.* **6**:2 (2018), 183–212. MR

[38] X. Yang, "Scalar curvature on compact complex manifolds", *Trans. Amer. Math. Soc.* **371**:3 (2019), 2073–2087. MR

[39] X. Yang, "RC-positive metrics on rationally connected manifolds", *Forum Math. Sigma* **8** (2020), art. id. e53. MR

[40] X. Yang and F. Zheng, "On real bisectional curvature for Hermitian manifolds", *Trans. Amer. Math. Soc.* **371**:4 (2019), 2703–2718. MR

[41] S. T. Yau, "Problem section", pp. 669–706 in *Seminar on Differential Geometry*, edited by S. T. Yau, Ann. of Math. Stud. **102**, Princeton Univ. Press, 1982. MR

[42] Y. Zhang, "Holomorphic sectional curvature, nefness and Miyaoka–Yau type inequality", *Math. Z.* **298**:3-4 (2021), 953–974. MR

[43] S. Zhang and X. Zhang, "On the structure of compact Kähler manifolds with nonnegative holomorphic sectional curvature", 2023. arXiv 2311.18779

[44] Q. Zhao and F. Zheng, "Strominger connection and pluriclosed metrics", *J. Reine Angew. Math.* **796** (2023), 245–267. MR

[45] F. Zheng, *Complex differential geometry*, AMS/IP Studies in Advanced Mathematics **18**, Amer. Math. Soc., Providence, RI, 2000. MR

[46] W. Zhou and F. Zheng, "Hermitian threefolds with vanishing real bisectional curvature", *Sci. China Math.* **52**:7 (2021), 757–764. In Chinese; translated in arXiv:2103.04296.

Received March 9, 2025. Revised April 1, 2026.

KAI TANG
SCHOOL OF MATHEMATICAL SCIENCES
ZHEJIANG NORMAL UNIVERSITY
JINHUA
CHINA
kaitang001@zjnu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 344, No. 1, 2026
https://doi.org/10.2140/pjm.2026.344.219

THE SIZE OF A SET FROM THE SIZE OF ITS COUNTABLE POWER SET

WILLIAM CHAN

Let $[\omega_1]^\omega$ and $[\omega_1]^{<\omega_1}$ be the set of all increasing ω-length and countable-length sequences of countable ordinals, respectively. Assume $\omega_1 \to_* (\omega_1)^{\omega_1}_2$. For any sequence $\langle Y_n : n \in \omega \rangle$, if $|[\omega_1]^{<\omega_1}| \leq \left|\prod_{n\in\omega} Y_n\right|$, then there exists an $\bar{n} \in \omega$ such that $|[\omega_1]^\omega| \leq |Y_{\bar{n}}|$.

Assume Woodin's extension AD^+ of the axiom of determinacy. For any set X, let $\mathscr{P}(X)$ be the set of subsets of X and let $\mathscr{P}_{\omega_1}(X)$ be the set of countable subsets of X. For any X which is a surjective image of $\mathbb{R}$, if $|\mathscr{P}(\omega_1)| \leq |\mathscr{P}_{\omega_1}(X)|$, then $|[\omega_1]^\omega| < |X|$.

1. Introduction

We will show that assuming the partition relation $\omega_1 \to_* (\omega_1)^{\omega_1}_2$ holds, if $[\omega_1]^{<\omega_1}$ injects into an ω-length Cartesian product of the form $\prod_{n\in\omega} Y_n$, then there is an $\bar{n} < \omega$ such that $[\omega_1]^\omega$ injects into $Y_{\bar{n}}$. This is partial evidence towards a conjecture that the ω-length Cartesian product operation $\prod_{n\in\omega} Y_n$ cannot create an injective copy of $[\omega_1]^{<\omega_1}$ unless one of the factors $Y_{\bar{n}}$ already had an injective copy of $[\omega_1]^{<\omega_1}$. The previous result will be used to show that under AD^+, if X is a surjective image of $\mathbb{R}$ and $\mathscr{P}(\omega_1)$ injects into $\mathscr{P}_{\omega_1}(X)$, then $[\omega_1]^\omega$ injects into X. This provides partial evidence towards a conjecture that the countable power set operation $\mathscr{P}_{\omega_1}(X)$ cannot created an injective copy of $\mathscr{P}(\omega_1)$ unless X already had an injective copy of $\mathscr{P}(\omega_1)$.

A cardinality is an equivalence class of a set under the bijection equivalence relation. If X is a set, then the cardinality of X, denoted by $|X|$, is the bijection equivalence class of X. If X and Y are two sets, then one writes $|X| \leq |Y|$ if and only if there is an injection of X into Y. One writes $|X| < |Y|$ if and only if $|X| \leq |Y|$ but $\neg(|Y| \leq |X|)$. A cardinal is an ordinal which does not inject into any smaller ordinal. If X is a well-orderable set, then $|X|$ has a unique cardinal as a member.

The axiom of determinacy, AD, fully classifies the cardinalities below $\mathscr{P}(\omega)$ or $\mathbb{R}$. Regularity properties for $\mathbb{R}$ (including the perfect set property, the property of Baire,

This research was funded in whole or in part by the Austrian Science Fund (FWF) 10.55776/Y1498.
MSC2020: 03E02, 03E60.
Keywords: partition relations, cardinalities.

© 2026 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

and Lebesgue measurability) answer many questions concerning combinatorial properties of the cardinality of $\mathbb{R}$ or $\mathscr{P}(\omega)$.

A major goal of recent research under determinacy is to classify the cardinalities below $\mathscr{P}(\omega_1)$ which means to identify all the cardinalities of subsets of $\mathscr{P}(\omega_1)$ and determine their relationship under the injection comparison relation. A deep understanding of some cardinalities strictly below $\mathscr{P}(\omega_1)$ has been obtained. For example, Woodin [21] provided a complete classification of the cardinalities below $\mathbb{R} \times \omega_1$, $[\omega_1]^\omega$, and $S_1 = \{f \in [\omega_1]^{<\omega_1} : \sup(f) = \omega_1^{L[f]}\}$ under $\mathsf{AD}_\mathbb{R}$ and DC. However, Woodin [21] also showed that the collection of cardinalities below $[\omega_1]^{<\omega_1}$ is very large and exhibits complex behavior under the injection comparison relation. Chan and Jackson [7] studied the structure of the cardinality below $\mathbb{R} \times \omega_1$ and between $[\omega_1]^{<\omega_1}$ and $\mathscr{P}(\omega_1)$ under AD and $V = L(\mathbb{R})$ (where $\mathsf{AD}_\mathbb{R}$ fails). The full classification below $\mathscr{P}(\omega_1)$ and many basic combinatorial properties of $\mathscr{P}(\omega_1)$ are very far from fully understood. Further development toward understanding the cardinality of $\mathscr{P}(\omega_1)$ has been motivated by many conjectures concerning inaccessibility of the cardinality of $\mathscr{P}(\omega_1)$.

One such inaccessibility feature concerns the regularity of $\mathscr{P}(\omega_1)$ and the search for the cofinality of $\mathscr{P}(\omega_1)$. Chan, Jackson, and Trang [10] developed the basic theory of regularity and cofinality in the choiceless determinacy context. The conjecture is that $\mathscr{P}(\omega_1)$ has globally regular cardinality which means that for all sets X which are images of $\mathbb{R}$ such that $\mathscr{P}(\omega_1)$ does not inject into X and all functions $\Phi : \mathscr{P}(\omega_1) \to X$, there is an $x \in X$ such that $|\Phi^{-1}[\{x\}]| = |\mathscr{P}(\omega_1)|$. Substantial evidence for the global regularity of $\mathscr{P}(\omega_1)$ is provided in [10].

We will focus on another inaccessibility conjecture for $\mathscr{P}(\omega_1)$ which was investigated in [7] involving the countable power set operation. If X is a set, then $\mathscr{P}(X)$ is the power set of X consisting of all subsets of X. If κ is a cardinal, then $\mathscr{P}_\kappa(X) = \{Y \subseteq X : |Y| < \kappa\}$ (that is, the collection of all subsets Y of X such that Y injects into κ but κ does not inject into Y). A particularly interesting case is $\mathscr{P}_{\omega_1}(X)$ which is the collection of all countable subsets of X.

The countable power set conjecture states that for any set X, if $\mathscr{P}(\omega_1)$ injects into $\mathscr{P}_{\omega_1}(X)$, then $\mathscr{P}(\omega_1)$ already injected into X. The intuition is that the countable power set operation applied to any set X cannot create a copy of $\mathscr{P}(\omega_1)$ in $\mathscr{P}_{\omega_1}(X)$ unless a copy of $\mathscr{P}(\omega_1)$ was already present in X.

The set $\mathscr{P}(\omega_1)$ contains many important smaller cardinalities which include $\mathbb{R}$, ω_1, $\mathbb{R} \sqcup \omega_1$, $\mathbb{R} \times \omega_1$, $[\omega_1]^\omega$, and $[\omega_1]^{<\omega_1}$. Evidence for the above conjecture can be given by showing that for any set X, if $|\mathscr{P}(\omega_1)| \leq |\mathscr{P}_{\omega_1}(X)|$, then X must have an injective copy of $\mathbb{R}$, ω_1, $\mathbb{R} \sqcup \omega_1$, $\mathbb{R} \times \omega_1$, $[\omega_1]^\omega$, or $[\omega_1]^{<\omega_1}$.

An ∞-Borel code is a pair (S, φ) such that S is a set of ordinals and φ is a formula in the language of set theory. A set $A \subseteq \mathbb{R}$ is ∞-Borel if and only if there is an ∞-Borel code (S, φ) such that $A = \{x \in \mathbb{R} : L[S, x] \models \varphi(S, x)\}$. An

∞-Borel code for A is a highly absolute definition for A in the sense that to query membership of $x \in A$, one only needs to determine the truth of $\varphi(S, x)$ in $L[S, x]$ which is the minimal inner model of ZFC containing x and the code set S. The statement that all subsets of $\mathbb{R}$ have ∞-Borel codes and $\mathsf{DC}_{\mathbb{R}}$ are part of Woodin's extension of determinacy known as AD^+. Chan and Jackson [7, Theorem 6.10] began the investigation of the countable power set conjecture by showing that under AD^+, for any set X which is an image of $\mathbb{R}$, if $|\mathscr{P}(\omega_1)| \leq |\mathscr{P}_{\omega_1}(X)|$, then $|\mathbb{R} \sqcup \omega_1| < |X|$. They also asked if X is an image of $\mathbb{R}$ and $|\mathscr{P}(\omega_1)| \leq |\mathscr{P}_{\omega_1}(X)|$, then does $|\mathbb{R} \times \omega_1| < |X|$ hold. Moreover, since $|\mathbb{R} \times \omega_1| < |[\omega_1]^\omega| < |\mathscr{P}(\omega_1)|$, an even stronger question would be under the same setting, does $|[\omega_1]^\omega| < |X|$ hold. The goal of the paper is to answer the latter question positively which will provide evidence for the countable power set conjecture:

- (Theorem 3.6) Assume AD, $\mathsf{DC}_{\mathbb{R}}$, and all subsets of $\mathbb{R}$ are ∞-Borel. For any set Y which is a surjective image of $\mathbb{R}$, if $|\mathscr{P}(\omega_1)| \leq |\mathscr{P}_{\omega_1}(Y)|$, then $|[\omega_1]^\omega| < |Y|$.

The main theorem requires a sufficiently deep understanding of the cardinalities of the relevant sets to answer a class of related "finite-dimensional combinatorial questions". Vaguely, an n-dimensional combinatorial question on a set X requires finding an $\widetilde{X} \subseteq X$ with $|\widetilde{X}| = |X|$ for which one has complete control over all n-tuples from $\widetilde{X}$. An important one-dimensional combinatorial question concerns regularity and cofinality. For example, for a particular set X and cardinal δ, X is δ-regular if and only if for all functions $\Phi : X \to \delta$, there exists a $\gamma < \delta$ such that $|\Phi^{-1}[\{\gamma\}]| = |X|$. One can ask for which ordinals δ is X δ-regular. There are many important two-dimensional combinatorial questions including primeness and basis for linear orderings. Important finite-dimensional questions (of arbitrary finite dimension) include the Jónssonness property and optimal colorings for partitions on n-tuples through a set X. Primeness is most relevant here. Chan, Jackson, and Trang [9] defined the notion of primeness. A set X is a said to be prime (or 2-prime) if and only if for all sets A and B, if $|X| \leq |A \times B|$, then $|X| \leq |A|$ or $|X| \leq |B|$. Chan, Jackson, and Trang [9] showed that under $\mathsf{AC}^{\mathbb{R}}_{\omega}$ and assuming all subsets of $\mathbb{R}$ have the Baire property, $\mathbb{R}$ and $\mathbb{R}/E_0$ are prime. They also showed that if κ is an uncountable cardinal satisfying $\kappa \to_* (\kappa)^{\omega+\omega}_2$, then $[\kappa]^\omega$ is prime. Moreover, they showed that for all $n < \omega$, $[\omega_n]^\omega$ is prime and even $[\omega_\omega]^\omega$ is prime (even though ω_n does not possess the partition property $\omega_n \to_* (\omega_n)^{\omega+\omega}_2$ since ω_n is singular when $3 \leq n \leq \omega$). More recently, Chan [3] has used a generalized Namba forcing (a form of Cox–Krueger Namba forcing [11] or diagonal Prikry forcing) over HOD-type inner models under AD^+ to determine the exact extent of ordinal regularity for $[\kappa]^\omega$ for all cardinals $\kappa < \Theta$. These Namba forcing methods can be adapted to establish the primeness of $[\kappa]^\omega$ for other $\kappa < \Theta$. These recent developments have provided a deep combinatorial understanding of the cardinality of ω-sequences of ordinals (which corresponds to

the smallest non-well-orderable cardinal exponentiations). In particular, note that the 2-primeness of $[\omega_1]^\omega$ under the partition relation $\omega_1 \to_* (\omega_1)^{\omega+\omega}_2$ can be regarded as an inaccessibility phenomenon for $[\omega_1]^\omega$ with respect to the two-fold Cartesian product operator: For any sets A and B, $A \times B$ has an injective copy of $[\omega_1]^\omega$ if and only if one of the factors A or B already possessed an injective copy of $[\omega_1]^\omega$.

However, one is still missing such understanding for the cardinality of $[\omega_1]^{<\omega_1}$. If κ is an ordinal and X is a set, then we say that X is κ-prime if and only if for any sequence $\langle Y_\alpha : \alpha < \kappa \rangle$, if $|X| \leq \left|\prod_{\alpha<\kappa} Y_\alpha\right|$, then there exists an $\bar{\alpha} < \kappa$ such that $|X| \leq |Y_{\bar{\alpha}}|$. Cardinals of countable cofinality are not ω-prime. It is shown in [9] that if κ is a cardinal and $\lambda < \kappa$, then κ is λ-prime if and only if $\lambda < \mathrm{cof}(\kappa)$ and for all $\delta < \kappa$, $\neg(|\kappa| \leq |^\lambda \delta|)$. Steel [19, Theorem 8.26] and Woodin [20, Theorem 2.16] showed that under AD^+, boldface GCH holds below Θ which is the statement that for all $\kappa < \Theta$, κ^+ does not inject into $\mathscr{P}(\kappa)$. Thus under AD^+, a cardinal κ is λ-prime if and only if $\lambda < \mathrm{cof}(\kappa)$. In particular, any cardinal of uncountable cofinality is ω-prime under AD^+. Note that $[\omega_1]^\omega$ is clearly not ω-prime although it is 2-prime under the partition relation $\omega_1 \to_* (\omega_1)^{\omega+\omega}_2$. In [9], it is shown that under AD^+, a set which is an image of $\mathbb{R}$ and is δ-prime for all $\delta \in \mathrm{ON}$ cannot be linearly orderable. If all subsets of $\mathbb{R}$ have the Baire property, then $\mathbb{R}/E_0$ is not linearly orderable. In [9], it is also shown that $\mathbb{R}/E_0$ is δ-prime for all $\delta \in \mathrm{ON}$ under $\mathsf{AC}^{\mathbb{R}}_\omega$ and all subsets of $\mathbb{R}$ have the Baire property. There are currently no known examples of a linearly orderable non-well-orderable ω-prime set but the natural candidate is $[\omega_1]^{<\omega_1}$. However, it is not presently known if $[\omega_1]^{<\omega_1}$ is even 2-prime. The ω-primeness conjecture for $[\omega_1]^{<\omega_1}$ is the statement that the strong partition property $\omega_1 \to_* (\omega_1)^{\omega_1}_2$ (or the very strong partition property $\omega_1 \to_* (\omega_1)^{\omega_1}_{<\omega_1}$) implies $[\omega_1]^{<\omega_1}$ is ω-prime. Intuitively this conjecture states that $[\omega_1]^{<\omega_1}$ is inaccessible to the countable Cartesian product operation in the sense that for any sequence $\langle Y_n : n \in \omega \rangle$ of sets, $[\omega_1]^{<\omega_1}$ injects into $\prod_{n\in\omega} Y_n$ if and only if there is a factor $Y_{\bar{n}}$ which already has an injective copy of $[\omega_1]^{<\omega_1}$. The key combinatorial fact needed to prove Theorem 3.6 is the following evidence for the ω-primeness conjecture for $[\omega_1]^{<\omega_1}$:

- (Theorem 2.11) Assume $\omega_1 \to_* (\omega_1)^{\omega_1}_2$. Let $\langle Y_n : n \in \omega \rangle$ be a sequence of sets. If $|[\omega_1]^{<\omega_1}| \leq \left|\prod_{n<\omega} Y_n\right|$, then there is an $\bar{n} \in \omega$ such that $|[\omega_1]^\omega| < |Y_{\bar{n}}|$.

The last ingredient to prove the main result of the paper will involve some consequences of Woodin's analysis of nice models of AD^+ by OD ∞-Borel code forcings (a variation of Vopěnka forcing). This forcing will not be used directly here but two consequences will be needed. The first states in nice AD^+ models, if a set $A \subseteq \mathbb{R}$ is ordinal definable via a formula φ, a suitable set z, and some ordinals, then there is a uniform procedure which takes φ, z, and the ordinal parameters to a ∞-Borel code (S, ψ) for A such that S is ordinal definable from z. (See

Fact 3.2 for the precise statement.) Woodin's perfect set dichotomy [1; 2] states that if X is a surjective image of $\mathbb{R}$, then either X is well-orderable or $\mathbb{R}$ injects into X. The second necessary fact is a uniform version of Woodin's perfect set dichotomy which states that if X is a quotient of an equivalence relation E on $\mathbb{R}$ with ∞-Borel code (S, φ) and X is well-orderable, then a well-ordering of X can be found uniformly from the ∞-Borel code (S, φ). (The precise statement is provided in Fact 3.3.)

If Y is a set, let $\mathcal{I}_{<\omega_1}(Y)$ be the set of all injective functions $f : \epsilon \to Y$ where $\epsilon < \omega_1$. The following will be shown:

- (Theorem 3.5) Assume AD, $\mathsf{DC}_{\mathbb{R}}$, and all subsets of $\mathbb{R}$ are ∞-Borel. If Y is a surjective image of $\mathbb{R}$ and $\Phi : \mathscr{P}(\omega_1) \to \mathscr{P}_{\omega_1}(Y)$, then there exists a function $\tilde{\Phi} : \mathscr{P}(\omega_1) \to \mathcal{I}_{<\omega_1}(Y)$ such that $\tilde{\Phi}(A)$ enumerates $\Phi(A)$ for all $A \in \mathscr{P}(\omega_1)$.

What is remarkable about Theorem 3.5 is that there exists generally no map $\Gamma : \mathscr{P}_{\omega_1}(Y) \to \mathcal{I}_{<\omega_1}(Y)$ such that $\Gamma(B)$ is a well-ordered enumeration of B for all $B \in \mathscr{P}_{\omega_1}(Y)$. In fact, there is generally no injection from $\mathscr{P}_{\omega_1}(Y)$ into $\mathcal{I}_{<\omega_1}(Y)$. For example for $Y = \mathbb{R}$, one can show that $|\mathcal{I}_{<\omega_1}(\mathbb{R})| < |\mathscr{P}_{\omega_1}(\mathbb{R})|$.

Given the main result of the paper that for any set X which is an image of $\mathbb{R}$, $|\mathscr{P}(\omega_1)| \leq |\mathscr{P}_{\omega_1}(X)|$ implies $|[\omega_1]^{\omega}| < |X|$, the next natural evidence one would like for the countable power set conjecture would be to show that $|\mathscr{P}(\omega_1)| \leq |\mathscr{P}_{\omega_1}(X)|$ implies $|[\omega_1]^{<\omega_1}| < |X|$. If the ω-primeness conjecture for $[\omega_1]^{<\omega_1}$ can be shown, then the arguments of Section 3 will immediately show that $|\mathscr{P}(\omega_1)| \leq |\mathscr{P}_{\omega_1}(X)|$ implies $|[\omega_1]^{<\omega_1}| < |X|$. One will need a deeper understanding of the cardinality of $[\omega_1]^{<\omega_1}$ to handle ω-primeness and other finite-dimensional combinatorial questions involving $[\omega_1]^{<\omega_1}$.

2. Evidence for primeness

Definition 2.1. Assume $\epsilon \in \mathrm{ON}$ and $f : \epsilon \to \mathrm{ON}$.

- The function f is discontinuous everywhere if and only if for all $\beta < \epsilon$, $\sup(f \restriction \beta) = \sup\{f(\alpha) : \alpha < \beta\} < f(\beta)$ (and note that this implies f is strictly increasing).
- The function f has uniform cofinality ω if and only if there is a function $F : \epsilon \times \omega \to \mathrm{ON}$ such that for all $\alpha < \epsilon$ and $n \in \omega$, $F(\alpha, n) < F(\alpha, n+1)$ and $f(\alpha) = \sup\{F(\alpha, n) : n \in \omega\}$. (A function $F : \epsilon \times \omega \to \mathrm{ON}$ with this property will be called a witness for the uniform cofinality ω of f.)
- The function f has the correct type if and only if f is both discontinuous everywhere and has uniform cofinality ω.

If $X \subseteq \mathrm{ON}$ and $\epsilon \in \mathrm{ON}$, then $[X]^{\epsilon}_{*}$ will denote the collection of $f : \epsilon \to X$ of the correct type.

The following is the correct type club partition relation.

Definition 2.2. Let κ be a cardinal.

- Let $\epsilon \leq \kappa$ and $\gamma < \kappa$. The statement $\kappa \to_* (\kappa)^{\epsilon}_{\gamma}$ asserts that for all $P : [\kappa]^{\epsilon}_* \to \gamma$, there is a $\beta < \gamma$ and a club $C \subseteq \kappa$ such that for all $f \in [C]^{\epsilon}_*$, $P(f) = \beta$.
- Let $\epsilon \leq \kappa$ and $\gamma \leq \kappa$. The statement $\kappa \to_* (\kappa)^{<\epsilon}_{<\gamma}$ asserts that for all $\epsilon' < \epsilon$ and $\gamma' < \gamma$, $\kappa \to_* (\kappa)^{\epsilon'}_{\gamma'}$.
- If κ satisfies $\kappa \to_* (\kappa)^{\kappa}_2$, then κ is called a strong partition cardinal.
- If κ satisfies $\kappa \to_* (\kappa)^{\kappa}_{<\kappa}$, then κ is called a very strong partition cardinal.
- If κ satisfies $\kappa \to_* (\kappa)^{<\kappa}_2$, then κ is called a weak partition cardinal.

See [2; 4; 17] for more information on the combinatorial aspects of partition properties. The ordinary partition relation $\kappa \to (\kappa)^{\epsilon}_{\gamma}$ is the assertion that for all $P : [\kappa]^{\epsilon} \to \kappa$, there is a $\beta < \gamma$ and an $A \subseteq \kappa$ with $|A| = \kappa$ such that for all $f \in [A]^{\epsilon}$, $P(f) = \beta$. The correct type club partition and the ordinary partition relation are closely related after a small shift in exponent: $\kappa \to_* (\kappa)^{\epsilon}_2$ implies $\kappa \to (\kappa)^{\epsilon}_2$ and $\kappa \to (\kappa)^{\omega \cdot \epsilon}_2$ implies $\kappa \to_* (\kappa)^{\epsilon}_2$. The correct type partition relation is much more practical for combinatorial arguments and essentially all applications of partition properties in the determinacy context directly or indirectly use the notion of a correct type partition.

It can be shown that for any $\epsilon < \kappa$, $\kappa \to_* (\kappa)^{\epsilon+\epsilon}_2$ implies that $\kappa \to (\kappa)^{\epsilon}_{<\kappa}$. Thus if κ is a weak partition cardinal, then $\kappa \to_* (\kappa)^{<\kappa}_{<\kappa}$ holds. It is not known if a strong partition cardinal must always be a very strong partition cardinal. The most natural universes to study partition properties are those satisfying AD. Moreover, the only known method to establish the existence of a strong partition cardinal comes from Martin's argument which uses pointclass boundedness principles under determinacy to create good coding systems for functions from certain ordinals to certain ordinals. (See [2; 4; 14; 15] for more information concerning good coding systems.)

If Γ is a pointclass, then let $\breve{\Gamma}$ refer to the dual pointclass. Let $\delta(\Gamma)$ be the supremum of the lengths of the pre-well-orderings in $\Gamma \cap \breve{\Gamma}$. For each $n \in \omega$, $\boldsymbol{\delta}^1_n$ is defined to be $\delta(\Sigma^1_n)$. By using the Kunen–Martin theorem, Kunen and Martin showed that $\boldsymbol{\delta}^1_{2n+2} = (\boldsymbol{\delta}^1_{2n+1})^+$ for all $n \in \omega$. Jackson [12; 13; 14] computed the values of all the projective ordinals $\boldsymbol{\delta}^1_n$ for $1 \leq n < \omega$. In particular, $\boldsymbol{\delta}^1_1 = \omega_1$, and thus $\boldsymbol{\delta}^1_2 = \omega_2$. Martin showed that $\boldsymbol{\delta}^1_3 = \omega_{\omega+1}$, and thus $\boldsymbol{\delta}^1_4 = \omega_{\omega+2}$. The ordinal $\boldsymbol{\delta}^2_1$ is defined to be $\delta(\Sigma^2_1)$. Let $A \in \mathscr{P}(\omega_1)$. The ordinal $\boldsymbol{\delta}_A$ is the least δ such that $L_\delta(A, \mathbb{R}) \prec_1 L(A, \mathbb{R})$ (Σ_1-elementarity in the language of set theory augmented with constant symbols for each real and the collection of all reals and the models are given the intended interpretation). It can be shown that $\boldsymbol{\delta}_\varnothing = (\boldsymbol{\delta}^2_1)^{L(\mathbb{R})}$. The following are some important examples of partition cardinals under the axiom of determinacy for which the results of this paper apply.

Fact 2.3. *Assume* AD.

- (*Martin* [2, Corollary 4.27; 15, Theorem 12.2]) ω_1 *is a very strong partition cardinal.*
- (*Martin–Paris* [2, Theorem 5.19 and Corollary 6.17; 15, Corollary 13.5]) ω_2 *is a weak partition cardinal but not a strong partition cardinal.*
- (*Jackson* [12; 13]) $\boldsymbol{\delta}^1_{2n+1}$ *is a very strong partition cardinal.*
- (*Jackson* [12; 13; 14]) $\boldsymbol{\delta}^1_{2n+2}$ *is a weak partition cardinal and not a strong partition cardinal.*
- (*Kechris–Kleinberg–Moschovakis–Woodin* [16; 18, Theorem 5.8]) *For any* $A \in \mathscr{P}(\kappa)$, $\boldsymbol{\delta}_A$ *is a very strong partition cardinal.*
- (*Kechris–Kleinberg–Moschovakis–Woodin* [16; 18, Theorem 5.5]) ($\mathsf{DC}_{\mathbb{R}}$) $\boldsymbol{\delta}^2_1$ *is a very strong partition cardinal.*

First, one will define a combinatorially useful injection of $[\omega_1]^\omega$ into $[\omega_1]^{<\omega_1}$. The injection is defined in Definition 2.6 and will be used in Theorem 2.11. The definition of the injection will use the linear ordering from Definition 2.4. The main combinatorial properties of the injection is described in Fact 2.7.

Definition 2.4. Let κ be a cardinal and $\Delta : \kappa \times \omega \to \kappa$ be a function. Define $\mathfrak{d}_\Delta : \kappa \to \kappa$ by $\mathfrak{d}_\Delta(\alpha) = \sup\{\Delta(\alpha, n) : n \in \omega\}$. Say that $\Delta : \kappa \times \omega \to \kappa$ is a witness for uniform cofinality ω for some discontinuous function (namely $\mathfrak{d}_\Delta$) if and only if the following hold:

(1) For all $\alpha < \kappa$ and $n \in \omega$, $0 < \Delta(\alpha, n) < \Delta(\alpha, n+1)$.

(2) $\mathfrak{d}_\Delta : \kappa \to \kappa$ is discontinuous everywhere (in the sense of Definition 2.1 which implies $\mathfrak{d}_\Delta$ is strictly increasing).

(Note that Δ witnesses that $\mathfrak{d}_\Delta$ is a function of uniform cofinality ω, and thus $\mathfrak{d}_\Delta \in [\kappa]^\kappa_*$.)

If $\gamma < \kappa$ and $\alpha < \mathfrak{d}_\Delta(\gamma)$, then let n^γ_α be the least n such that $\alpha < \Delta(\gamma, n)$.

Let L_Δ consist of all tuples $(\gamma, \alpha, \beta_0, \ldots, \beta_{n^\gamma_\alpha - 1})$ with the following properties:

(1) $\gamma < \kappa$ and $\alpha < \mathfrak{d}_\Delta(\gamma)$.

(2) If $n^\gamma_\alpha = 0$, then the tuple takes the form (γ, α).

(3) If $n^\gamma_\alpha > 0$, then $\beta_0 < \beta_1 < \cdots < \beta_{n^\gamma_\alpha - 1} < \gamma$.

Let $\prec_\Delta$ be the lexicographic ordering on L_Δ. Let $\mathcal{L}_\Delta$ be the linear ordering $(L_\Delta, \prec_\Delta)$.

Fact 2.5. *Let* κ *be a cardinal. Let* $\Delta : \kappa \times \omega \to \kappa$ *be a witness for uniform cofinality* ω *for some increasing discontinuous function.* $\mathcal{L}_\Delta = (L_\Delta, \prec_\Delta)$ *is a well-ordering of order type* κ.

Definition 2.6. Let κ, Δ, and $\mathfrak{d}_\Delta$ be as in Definition 2.4. A pair (Δ, Σ) is suitable if and only if the following holds:

(1) $\Sigma : L_\Delta \to \kappa$ is an order-preserving function from $\mathcal{L}_\Delta = (L_\Delta, \prec_\Delta)$ into $(\kappa, <)$.

(2) For all $\gamma_0 < \kappa$ and $(\gamma_0, \alpha, \beta_0, \ldots, \beta_{n_\alpha^{\gamma_0}-1}) \in L_\Delta$,

$$\mathfrak{d}_\Delta(\gamma_0) < \Sigma(\gamma_0, \alpha, \beta_0, \ldots, \beta_{n_\alpha^{\gamma_0}-1}) < \mathfrak{d}_\Delta(\gamma_0 + 1).$$

(3) Σ has the correct type in the following natural sense:

(a) Σ is discontinuous: for all $x \in L_\Delta$,

$$\sup(\Sigma \restriction x) = \sup\{\Sigma(y) : y \prec_\Delta x\} < \Sigma(x).$$

(b) Σ has uniform cofinality ω: there is a $\Pi : L_\Delta \times \omega \to \kappa$ such that for all $x \in L_\Delta$ and $n \in \omega$,

$$\Pi(x, n) < \Pi(x, n+1) \quad \text{and} \quad \Sigma(x) = \sup\{\Pi(x, n) : n \in \omega\}.$$

For each $f \in [\kappa]^\omega$, define a function $\Lambda_{(\Delta,\Sigma)}(f) : \mathfrak{d}_\Delta(\sup(f)) \to \kappa$ by

$$\Lambda_{(\Delta,\Sigma)}(f)(\alpha) = \Sigma(\sup(f), \alpha, f(0), \ldots, f(n_\alpha^{\sup(f)} - 1))$$

for each $\alpha < \mathfrak{d}_\Delta(\sup(f))$. Thus $\Lambda_{(\Delta,\Sigma)} : [\kappa]^\omega \to {}^{<\kappa}\kappa$.

Fact 2.7. *Let κ, Δ, $\mathfrak{d}_\Delta$, $\mathcal{L}_\Delta = (L_\Delta, \prec_\Delta)$, Σ, $\Lambda_{(\Delta,\Sigma)}$ be as in Definitions 2.4 and 2.6. The following properties hold*:

(1) *For all $f \in [\kappa]^\omega$ and $\alpha_0 < \alpha_1 < \mathfrak{d}_\Delta(\sup(f))$,*

$$\Lambda_{(\Delta,\Sigma)}(f)(\alpha_0) < \Lambda_{(\Delta,\Sigma)}(f)(\alpha_1).$$

(2) *For all $f \in [\kappa]^\omega$, $\Lambda_{(\Delta,\Sigma)}(f)(0) > \mathfrak{d}_\Delta(\sup(f))$ and $\Lambda_{(\Delta,\Sigma)}(f) \in [\kappa]_*^{\mathfrak{d}_\Delta(\sup(f))}$.*

(3) *For all $f, g \in [\kappa]^\omega$, if $\sup(f) < \sup(g)$, then*

$$\sup(\Lambda_{(\Delta,\Sigma)}(f)) < \mathfrak{d}_\Delta(\sup(g)) < \Lambda_{(\Delta,\Sigma)}(g)(0).$$

(4) *Suppose $f, g \in [\kappa]^\omega$, $\sup(f) = \sup(g)$, $m \in \omega$ is least such that $f(m) \neq g(m)$, and $f(m) < g(m)$. For all $\alpha < \Delta(\sup(f), m)$,*

$$\Lambda_{(\Delta,\Sigma)}(f)(\alpha) = \Lambda_{(\Delta,\Sigma)}(g)(\alpha).$$

For all α such that $\Delta(\sup(f), m) \leq \alpha < \mathfrak{d}_\Delta(\sup(f)) = \mathfrak{d}_\Delta(\sup(g))$,

$$\Lambda_{(\Delta,\Sigma)}(f)(\alpha) < \Lambda_{(\Delta,\Sigma)}(g)(\alpha) < \Lambda_{(\Delta,\Sigma)}(f)(\alpha + 1).$$

(5) *$\Lambda_{(\Delta,\Sigma)} : [\kappa]^\omega \to [\kappa]_*^{<\kappa}$ is an injection. Let $K = \Sigma[L_\Delta]$ and $J = \mathfrak{d}_\Delta[\kappa]$. For all $\epsilon \in J$, $\mathrm{cof}(\epsilon) = \omega$. Thus the range of $\Lambda_{(\Delta,\Sigma)}$ is $\bigcup_{\epsilon \in J} [K \setminus (\epsilon + 1)]_*^\epsilon$.*

Proof. Each of above properties are straightforward observations by comparing the relevant coordinate of a relevant tuple in L_Δ used in the definition of $\Lambda_{(\Delta,\Sigma)}(f)(\alpha)$. The following provides the details:

(1) Let $f \in [\kappa]^\omega$ and $\alpha_0 < \alpha_1 < \mathfrak{d}_\Delta(\sup(f))$. Observe

$$(\sup(f), \alpha_0, f(0), \dots, f(n^{\sup(f)}_{\alpha_0} - 1)) \prec_\Delta (\sup(f), \alpha_1, f(0), \dots, f(n^{\sup(f)}_{\alpha_1} - 1))$$

by comparing the second coordinate. Thus $\Lambda_{(\Delta,\Sigma)}(f)(\alpha_0) < \Lambda_{(\Delta,\Sigma)}(f)(\alpha_1)$ since Σ is order preserving.

(2) By the suitability of the pair (Δ, Σ),

$$\Lambda_{(\Delta,\Sigma)}(f)(0) = \Sigma(\sup(f), 0) > \mathfrak{d}(\sup(f)).$$

Note that (1) shows that $\Lambda_{(\Delta,\Sigma)}(f)$ is an increasing function. For all $\alpha < \mathfrak{d}_\Delta(\sup(f))$,

$$\begin{aligned}\sup(\Lambda_{(\Delta,\Sigma)}(f) \restriction \alpha) &= \sup\big(\Sigma \restriction (\sup(f), \alpha, f(0), \dots, f(n^{\sup(f)}_\alpha - 1))\big)\\ &< \Sigma(\sup(f), \alpha, f(0), \dots, f(n^{\sup(f)}_\alpha - 1))\\ &= \Lambda_{(\Delta,\Sigma)}(f)(\alpha),\end{aligned}$$

using the discontinuity of Σ. Thus $\Lambda_{(\Delta,\Sigma)}(f)$ is discontinuous everywhere. Fix a function $\Pi : L_\Delta \times \omega \to \kappa$ which witnesses that Σ has uniform cofinality ω as in Definition 2.6(3b). Define $F : \mathfrak{d}_\Delta(\sup(f)) \times \omega \to \kappa$ by

$$F(\alpha, n) = \Pi\big((\sup(f), \alpha, f(0), \dots, f(n^{\sup(f)}_\alpha - 1)), n\big).$$

F witnesses that $\Lambda_{(\Delta,\Sigma)}(f)$ has uniform cofinality ω. Thus $\Lambda_{(\Delta,\Sigma)}(f)$ has the correct type, and hence $\Lambda_{(\Delta,\Sigma)}(f) \in [\kappa]^{\mathfrak{d}_\Delta(\sup(f))}_*$.

(3) Suppose $f, g \in [\kappa]^\omega$ and $\sup(f) < \sup(g)$. For all $\alpha < \mathfrak{d}_\Delta(\sup(f))$,

$$(\sup(f), \alpha, f(0), \dots, f(n^{\sup(f)}_\alpha - 1)) \prec_\Delta (\sup(f) + 1, 0)$$

by comparing the first coordinates. Thus

$$\Lambda_{(\Delta,\Sigma)}(f)(\alpha) < \Sigma(\sup(f) + 1, 0) < \mathfrak{d}_\Delta(\sup(g)) < \Lambda_{(\Delta,\Sigma)}(g)(0)$$

for all $\alpha < \mathfrak{d}_\Delta(\sup(f))$ by (2) established above and Definition 2.6(2). Hence $\sup(\Lambda_{(\Delta,\Sigma)}(f)) < \mathfrak{d}_\Delta(\sup(g)) < \Lambda_{(\Delta,\Sigma)}(g)(0)$.

(4) Suppose $f, g \in [\kappa]^\omega$, $\sup(f) = \sup(g)$, m is the least such that $f(m) \neq g(m)$, and $f(m) < g(m)$. Suppose $\alpha < \Delta(\sup(f), m) = \Sigma(\sup(g), m)$. Then

$$n^{\sup(f)}_\alpha = n^{\sup(g)}_\alpha \leq m.$$

Thus

$$\begin{aligned}\Lambda_{(\Delta,\Sigma)}(f)(\alpha) &= \Sigma(\sup(f), \alpha, f(0), \dots, f(n^{\sup(f)}_\alpha - 1))\\ &= \Sigma(\sup(g), \alpha, g(0), \dots, g(n^{\sup(g)}_\alpha - 1))\\ &= \Lambda_{(\Delta,\Sigma)}(g)(\alpha).\end{aligned}$$

Now suppose $\Delta(\sup(f), m) = \Delta(\sup(g), m) \leq \alpha$. Then $n_\alpha^{\sup(f)} = n_\alpha^{\sup(g)} > m$. Thus

$$(\sup(f), \alpha, f(0), \ldots, f(n_\alpha^{\sup(f)} - 1)) \prec_\Delta (\sup(g), \alpha, g(0), \ldots, g(n_\alpha^{\sup(g)} - 1)) \prec_\Delta (\sup(f), \alpha + 1, f(0), \ldots, f(n_{\alpha+1}^{\sup(f)}))$$

by comparing the $(m+2)$-nd entries for the first inequality and comparing the second entry for the second inequality. Thus

$$\Lambda_{(\Delta,\Sigma)}(f)(\alpha) < \Lambda_{(\Delta,\Sigma)}(g)(\alpha) < \Lambda_{(\Delta,\Sigma)}(f)(\alpha + 1).$$

(5) Suppose $f, g \in [\kappa]^\omega$ and $f \neq g$. If $\sup(f) \neq \sup(g)$, then (3) above implies $\Lambda_{(\Delta,\Sigma)}(f) \neq \Lambda_{(\Delta,\Sigma)}(g)$. Suppose $\sup(f) = \sup(g)$. There is a least $m \in \omega$ such that $f(m) \neq g(m)$. Then $\Lambda_{(\Delta,\Sigma)}(f) \neq \Lambda_{(\Delta,\Sigma)}(g)$ by (4). Thus $\Lambda_{(\Delta,\Sigma)}$ is an injection. Let $K = \Sigma[L_\Delta]$ and $J = \mathfrak{d}_\Delta[\kappa]$. By definition of Σ and $\Lambda_{(\Delta,\Sigma)}$ and (2), for all $f \in [\kappa]^\omega$, $\Lambda_{(\Delta,\Sigma)}(f) \in [K]_*^{\mathfrak{d}_\Delta(\sup(f))}$. Since $\Lambda_{(\Delta,\Sigma)}(f)(0) > \mathfrak{d}_\Delta(\sup(f))$ by (2) and $\mathfrak{d}_\Delta(\sup(f)) \in J$, $\Lambda_{(\Delta,\Sigma)}(f) \in \bigcup_{\epsilon \in J} [K \setminus (\epsilon + 1)]_*^\epsilon$. Note that for all $\gamma \in \kappa$, $\mathrm{cof}(\mathfrak{d}_\Delta(\gamma)) = \omega$ as witnessed by the cofinal sequence $\langle \Delta(\gamma, n) : n \in \omega \rangle$. Thus for all $\epsilon \in J$, $\mathrm{cof}(\epsilon) = \omega$. □

Let $X \subseteq \kappa$ with $|X| = \kappa$. Let $\mathrm{enum}_X : \kappa \to X$ be the increasing enumeration of X. For $\alpha, \beta < \kappa$, let $\mathrm{next}_X^\alpha(\beta) = \mathrm{enum}_{X \setminus (\beta+1)}(\alpha)$ which is the α-th-element of X strictly greater than β.

Fact 2.8. *Let κ be a cardinal and $C \subseteq \kappa$ be a club subset of κ. Then there is a suitable (Δ, Σ) such that $\Delta : \kappa \times \omega \to [C]_*^1$ and $\Sigma : L_\Delta \to C$. (Note that $\mathfrak{d}_\Delta : \kappa \to [C]_*^1$ and $\Sigma : L_\Delta \to [C]_*^1$ since Σ has the correct type by suitability.)*

Proof. The function $\Delta : \kappa \times \omega \to [C]_*^1$, $\Sigma : L_\Delta \to C$, and $\Pi : L_\Delta \times \omega \to \kappa$ will be defined by recursion by choosing the minimal element of C which maintains the necessary properties. Let $\gamma < \kappa$. Suppose $\Delta \restriction \gamma \times \omega$ has been defined. This implies one can meaningfully define the objects $\mathfrak{d}_\Delta \restriction \gamma$, n_α^η for all $\eta < \gamma$ and $\alpha < \mathfrak{d}_\Delta(\eta)$, and a linear ordering A_γ using the definition of L_Δ from Definition 2.4 with the restriction that the first coordinate is less than γ. Suppose $\Sigma \restriction A_\gamma$ and $\Pi \restriction A_\gamma \times \omega$ have been defined. Suppose the following properties have been shown:

(1) $\Sigma \restriction A_\gamma$ is increasing.

(2) For all $x \in A_\gamma$ with first coordinate $\eta < \gamma$, $\mathfrak{d}_\Delta(\eta) < \Sigma(x)$.

(3) For all $\gamma_0 < \gamma_1 < \gamma$ and $x \in A_\gamma$ with first coordinate γ_0, $\Sigma(x) < \mathfrak{d}_\Delta(\gamma_1)$.

(4) $\Pi \restriction A_\gamma \times \omega : A_\gamma \times \omega \to \kappa$ witnesses that $\Sigma \restriction A_\gamma$ has uniform cofinality ω.

Let $\delta_\gamma = \sup(\Sigma \restriction A_\gamma)$. Let $\Delta(\gamma, n) = \mathrm{next}_{[C]_*^1}^n(\delta_\gamma)$. Let $\mathfrak{d}_\Delta(\gamma) = \sup\{\Delta(\gamma, n) : n \in \omega\}$. Now one can meaningfully define $A_{\gamma+1}$ using the definition of L_Δ given in Definition 2.4 restricted below $\gamma + 1$. Let $B_\gamma = A_{\gamma+1} \setminus A_\gamma$ which are the elements of $A_{\gamma+1}$ with first coordinate γ. Note that $\mathrm{ot}(A_{\gamma+1}) < \kappa$ and $\mathrm{ot}(B_\gamma) < \kappa$ by the

argument from Fact 2.5. Let $\zeta_\gamma = \mathrm{ot}(B_\gamma)$ and $\pi_\gamma : B_\gamma \to \zeta_\gamma$ be the unique order-preserving bijection. For $x \in B_\gamma$, let $\Sigma(x) = \mathrm{next}_C^{\omega\cdot\pi_\gamma(x)+\omega}(\mathfrak{d}_\Delta(\gamma))$ and for $n \in \omega$, let $\Pi(x,n) = \mathrm{next}_C^{\omega\cdot\pi_\gamma(x)+n}(\mathfrak{d}_\Delta(\gamma))$. One has defined $\Delta \restriction (\gamma+1)\times\omega$, $\Sigma \restriction A_{\gamma+1}$, and $\Pi \restriction A_{\gamma+1}\times\omega$. Using the induction hypothesis that $\Sigma \restriction A_\gamma$ is order preserving, $\Sigma \restriction A_{\gamma+1}$ is order preserving by construction. Note that $\mathfrak{d}_\Delta(\gamma) < \Sigma(x)$ for all $x \in B_\gamma$ by construction. Now suppose $\gamma_0 < \gamma$ and $x \in A_{\gamma+1}$ with first coordinate γ_0. $\Sigma(x) \le \sup(\Sigma \restriction A_\gamma) = \delta_\gamma < \Delta(\gamma,0) < \mathfrak{d}_\Delta(\gamma)$. $\Pi \restriction A_{\gamma+1}\times\omega$ witnesses that $\Sigma \restriction A_{\gamma+1}$ has uniform cofinality ω. Thus the four properties hold at $\gamma+1$ for $\Delta \restriction (\delta+1)\times\omega$, $\Sigma \restriction A_{\gamma+1}$, and $\Pi \restriction A_{\gamma+1}\times\omega$.

Suppose γ is a limit ordinal and for all $\tilde\gamma < \gamma$, $\Delta \restriction \tilde\gamma\times\omega$, $\Sigma \restriction A_{\tilde\gamma}$, and $\Pi \restriction A_{\tilde\gamma}\times\omega$ have been defined with the above four properties. Note $A_\gamma = \bigcup_{\tilde\gamma<\gamma} A_{\tilde\gamma}$. Then $\Delta \restriction \gamma\times\omega$, $\Sigma \restriction A_\gamma$, and $\Pi \restriction A_\gamma\times\omega$ also have these four properties.

This completes the construction of $\Delta : \kappa\times\omega \to [C]^1_*$ and $\Sigma : L_\Delta \to C$. If $\alpha < \beta < \kappa$, then $\mathfrak{d}_\Delta(\alpha) < \sup(\Sigma \restriction A_{\alpha+1}) < \mathfrak{d}_\Delta(\beta)$ using properties (2) and (3) at $\beta+1$. Thus $\mathfrak{d}_\Delta$ is an increasing function. Suppose $\alpha < \kappa$. Then we have $\sup(\mathfrak{d}_\Delta \restriction \alpha) \le \sup(\Sigma \restriction A_\alpha) = \delta_\alpha < \Delta(\alpha,0) < \mathfrak{d}_\Delta(\alpha)$ by property (2) and (3) at $\alpha+1$. Thus $\mathfrak{d}_\Delta$ is discontinuous. The function $\Sigma : L_\Delta \to \kappa$ is increasing by (1). For all $\gamma_0 < \gamma_1 < \kappa$ and $x \in L_\Delta$ with first coordinate γ_0, $\mathfrak{d}_\Delta(\gamma_0) < \Sigma(x) < \mathfrak{d}_\Delta(\gamma_1)$ follows by properties (2) and (3) at γ_1+1. This verifies Definition 2.6(2). Suppose $x \in L_\Delta$ with first coordinate γ, and hence $x \in B_\gamma$. If x is the least element of B_γ, then $\pi_\gamma(x) = 0$ and

$$\begin{aligned}\sup(\Sigma \restriction x) &= \sup(\Sigma \restriction A_\gamma) = \delta_\gamma \\ &< \Delta(\gamma,0) < \mathfrak{d}_\Delta(\gamma) < \mathrm{next}_C^{\omega}(\mathfrak{d}_\Delta(\gamma)) \\ &= \mathrm{next}_C^{\omega\cdot\pi_\gamma(x)+\omega}(\mathfrak{d}_\Delta(\gamma)) = \Sigma(x).\end{aligned}$$

Now suppose x is not the minimal element of B_γ. Then

$$\begin{aligned}\sup(\Sigma \restriction x) &= \sup\{\Sigma(y) : y \prec_\Delta x\} = \sup\{\mathrm{next}_C^{\omega\cdot\pi_\gamma(y)+\omega}(\mathfrak{d}_\Delta(\gamma)) : y \prec_\Delta x\} \\ &= \sup\{\mathrm{next}_C^{\omega\cdot\nu+\omega}(\mathfrak{d}_\Delta(\gamma)) : \nu < \pi_\gamma(x)\} \le \mathrm{next}_C^{\omega\cdot\pi_\gamma(x)}(\mathfrak{d}_\Delta(\gamma)) \\ &< \mathrm{next}_C^{\omega\cdot\pi_\gamma(x)+\omega}(\mathfrak{d}_\Delta(\gamma)) = \Sigma(x).\end{aligned}$$

Thus Σ is discontinuous. Π witnesses that Σ has uniform cofinality ω. Σ has the correct type. Thus (Δ, Σ) is suitable with the desired properties. □

If $\delta < \epsilon$ and $f : \epsilon \to \mathrm{ON}$, then let $\mathrm{drop}(f,\delta) : (\epsilon-\delta) \to \mathrm{ON}$ be defined by $\mathrm{drop}(f,\delta)(\alpha) = f(\delta+\alpha)$. If κ is a cardinal, then we let club_κ denote the set of all club subsets of κ. For any cardinal κ and $\epsilon \le \kappa$, a relation $R \subseteq [\kappa]^\epsilon_* \times \mathsf{club}_\kappa$ is $\subseteq$-downward closed in the club_κ-coordinate if and only if for all $\ell \in [\kappa]^\epsilon_*$ and clubs $C \subseteq D$, if $R(\ell, D)$ holds, then $R(\ell, C)$ holds. Without the axiom of choice, uniform choice of homogeneous sets for many partitions is a very subtle and

challenging problem. "Club uniformization" results are often very important for combinatorial questions involving partition spaces. The club uniformization of Theorem 3.10 in [6] is more than sufficient for the proof of Theorem 2.11 (for just the cardinal ω_1); however, the proof is not purely combinatorial. The following club uniformization results are proved from just the strong partition property and will allow for the selection of clubs homogeneous for various desired properties in the main construction of Theorem 2.11

Fact 2.9. *Assume $\kappa \to_* (\kappa)^\kappa_2$ and $\epsilon < \kappa$. Let $R \subseteq [\kappa]^\epsilon_* \times \mathsf{club}_\kappa$ be $\subseteq$-downward closed in the club_κ-coordinate. Then there is a club $C \subseteq \kappa$ such that for all $\ell \in \mathrm{dom}(R) \cap [C]^\epsilon_*$, $R(\ell, C \setminus (\sup(\ell)+1))$.*

Proof. Fix $\epsilon < \kappa$. If $g \in [\kappa]^\kappa_*$, then let $\mathcal{C}_g$ be the closure of $g[\kappa]$ which is a club. Define $P : [\kappa]^\kappa_* \to 2$ by $P(f) = 0$ if and only if $f \restriction \epsilon \in \mathrm{dom}(R) \Rightarrow R(f \restriction \epsilon, \mathcal{C}_{\mathrm{drop}(f,\epsilon)})$. By $\kappa \to_* (\kappa)^\kappa_2$, there is a club $C \subseteq \kappa$ and an $i \in 2$ such that for all $f \in [C]^\kappa_*$, $P(f) = i$. Pick any $\ell \in [C]^\epsilon_*$. First, suppose $\ell \notin \mathrm{dom}(R)$. Pick any $g \in [C \setminus (\sup(\ell)+1)]^\kappa_*$ and let $f = \ell \frown g$. Then $f \in [C]^\kappa_*$ and $P(f) = 0$. Now suppose $\ell \in \mathrm{dom}(R)$. There is some club $D \subseteq \kappa$ such that $R(\ell, D)$. Let $g \in [C \cap D]^\kappa_*$ with $g(0) > \sup(\ell)$ and observe that $\mathcal{C}_g \subseteq D$. Since R is $\subseteq$-downward closed in the club_κ-coordinate, $R(\ell, \mathcal{C}_g)$. Let $f = \ell \frown g$. Note that $f \in [C]^\kappa_*$ and $g = \mathrm{drop}(f, \epsilon)$. Thus $P(f) = 0$. Since C was assumed to be homogeneous for P, this shows that $i = 0$. Fix an $h \in [C]^\kappa_*$. Let $E = \mathcal{C}_h$ which is a club subset of κ. Suppose $\ell \in \mathrm{dom}(R) \cap [E]^\epsilon_*$. Let $\alpha < \kappa$ be least such that $h(\alpha) > \sup(\ell)$. Let $f = \ell \frown \mathrm{drop}(h, \alpha)$ and note that $f \in [C]^\kappa_*$, $f \restriction \epsilon = \ell$, $\mathrm{drop}(f, \epsilon) = \mathrm{drop}(h, \alpha)$, and $E \setminus (\sup(\ell)+1) = \mathcal{C}_{\mathrm{drop}(h,\alpha)}$. $P(f) = 0$ implies that $R(f \restriction \epsilon, \mathcal{C}_{\mathrm{drop}(f,\epsilon)})$. Thus $R(\ell, E \setminus (\sup(\ell)+1))$. E is the desired club. □

Fact 2.10. *Assume $\kappa \to_* (\kappa)^\kappa_2$. Let $R \subseteq [\kappa]^{<\kappa}_* \times \mathsf{club}_\kappa$ be $\subseteq$-downward closed in the club_κ-coordinate. There is a club $C \subseteq \kappa$ such that $R(\ell, C \setminus (\sup(\ell)+1))$ for all $\ell \in \mathrm{dom}(R) \cap \bigcup_{\epsilon \in [C]^1_*} [C \setminus (\epsilon+1)]^\epsilon_*$.*

Proof. Fix $\epsilon < \kappa$. Define $R^\epsilon \subseteq [\kappa]^\epsilon_* \times \mathsf{club}_\kappa$ by $R^\epsilon(\ell, C)$ if and only if $R(\ell, C)$. R^ϵ is $\subseteq$-downward closed in the club_κ-coordinate. By Fact 2.9, there is a club D such that $R(\ell, D \setminus (\sup(\ell)+1))$ for all $\ell \in \mathrm{dom}(R^\epsilon) \cap [C]^\epsilon_*$. Define $S \subseteq \kappa \times \mathsf{club}_\kappa$ by $S(\epsilon, C)$ if and only if $R^\epsilon(\ell, C \setminus (\sup(\ell)+1))$ for all $\ell \in \mathrm{dom}(R^\epsilon) \cap [C]^\epsilon_*$. By the above discussion, $\mathrm{dom}(S) = \kappa$. By Fact 2.9 applied with $\epsilon = 1$, there is a club $E \subseteq \kappa$ such that $S(\epsilon, E \setminus (\epsilon+1))$ for all $\epsilon \in [E]^1_*$. Let $\ell \in \mathrm{dom}(R) \cap \bigcup_{\epsilon \in [E]^1_*} [E \setminus (\epsilon+1)]^\epsilon_*$. Let $\epsilon = |\ell|$ and note that $\epsilon \in [E]^1_*$ and $\ell \in [E \setminus (\epsilon+1)]^\epsilon_*$. $S(\epsilon, E \setminus (\epsilon+1))$ holds. Since $\ell \in \mathrm{dom}(R^\epsilon) \cap [E \setminus (\epsilon+1)]^\epsilon_*$, $R^\epsilon(\ell, (E \setminus (\epsilon+1)) \setminus (\sup(\ell)+1))$. Since $(E \setminus (\epsilon+1)) \setminus (\sup(\ell)+1) = E \setminus (\sup(\ell)+1)$, $R^\epsilon(\ell, E \setminus (\sup(\ell)+1))$. Thus $R(\ell, E \setminus (\sup(\ell)+1))$. E is the desired club. □

The following will provide evidence for the conjecture that $[\omega_1]^{<\omega_1}$ is ω-prime. In the following argument, there will be a very large family of relevant partitions.

Fact 2.10 will provide a club which is "simultaneously homogeneous" for all the relevant partitions in the formal sense of Fact 2.10. Using this club, one will define an injection of the form presented in Definition 2.6. A number of fine technical details need to be verified to see that this injection interacts meaningfully with all the relevant partitions.

Theorem 2.11. *Assume κ is a cardinal satisfying $\kappa \to_* (\kappa)^\kappa_2$. Let $\lambda < \kappa$ and $\langle Y_\eta : \eta < \lambda\rangle$ be a sequence of sets. If $|[\kappa]^{<\kappa}| \leq \left|\prod_{\eta<\lambda} Y_\eta\right|$, then there is an $\bar\eta < \lambda$ such that $|[\kappa]^\omega| \leq |Y_{\bar\eta}|$.*

Proof. Since $|[\kappa]^{<\kappa}| \leq \left|\prod_{\eta\in\lambda} Y_\eta\right|$, let $\Phi : [\kappa]^{<\kappa} \to \prod_{\eta\in\lambda} Y_\eta$ be an injection. If $f, g \in [\kappa]^{<\kappa}$ and $f \neq g$, then let $\chi(f, g)$ be the least element $\eta \in \lambda$ such that $\Phi(f)(\eta) \neq \Phi(g)(\eta)$. Let $Z \subseteq \kappa$ be the club of indecomposable ordinals below κ and work within Z. Note that $\kappa \to_* (\kappa)^\kappa_2$ implies $\kappa \to_* (\kappa)^{<\kappa}_2$, and hence $\kappa \to_* (\kappa)^{<\kappa}_{<\kappa}$ by the comments in the paragraphs following Definition 2.2.

For any $\epsilon \in Z$ and $\iota \in [Z]^\epsilon_*$, let $\iota^0, \iota^1 \in [Z]^\epsilon_*$ be defined by $\iota^i(\alpha) = \iota(2\cdot\alpha + i)$.

Claim 1. *There is a club $C^0_1 \subseteq \kappa$ and $\bar\eta < \lambda$ such that $\chi(\ell \frown \iota^0, \ell \frown \iota^1) = \bar\eta$ for all $(\delta, \epsilon) \in [C^0_1]^2_*$, all $\ell \in [C^0_1 \setminus (\epsilon + 1)]^\delta_*$, and all $\iota \in [C^0_1 \setminus (\sup(\ell) + 1)]^\epsilon_*$.*

Proof of Claim 1. For any $(\delta, \epsilon) \in [Z]^2_*$ (hence $\delta < \epsilon$) and $\ell \in [Z \setminus (\epsilon + 1)]^\delta_*$, define $P_{\delta,\epsilon,\ell} : [Z \setminus (\sup(\ell) + 1)]^\epsilon_* \to \lambda$ by $P_{\delta,\epsilon,\ell}(\iota) = \chi(\ell \frown \iota^0, \ell \frown \iota^1)$. By $\kappa \to_* (\kappa)^\epsilon_\lambda$, there is an $\eta_{\delta,\epsilon,\ell} \in \lambda$ and a club $C \subseteq Z$ such that $P_{\delta,\epsilon,\ell}(\iota) = \eta_{\delta,\epsilon,\ell}$ for all $\iota \in [C]^\epsilon_*$. For $(\delta, \epsilon) \in [Z]^2_*$, define $P_{\delta,\epsilon} : [Z \setminus (\epsilon + 1)]^\delta_* \to \lambda$ by $P_{\delta,\epsilon}(\ell) = \eta_{\delta,\epsilon,\ell}$. By $\kappa \to_* (\kappa)^\delta_\lambda$, there is a club $C \subseteq Z$ and an $\eta_{\delta,\epsilon} \in \lambda$ such that $P_{\delta,\epsilon}(\ell) = \eta_{\delta,\epsilon,\ell} = \eta_{\delta,\epsilon}$ for all $\ell \in [C]^\delta_*$. Define $P : [Z]^2_* \to \lambda$ by $P(\delta, \epsilon) = \eta_{\delta,\epsilon}$. By $\kappa \to_* (\kappa)^2_\lambda$, fix a club $C^0_0 \subseteq Z$ and an $\bar\eta \in \lambda$ such that $P(\delta, \epsilon) = \eta_{\delta,\epsilon} = \bar\eta$ for all $(\delta, \epsilon) \in [C^0_0]^2_*$. For $(\delta, \epsilon) \in [C^0_0]^2_*$, let $R_{\delta,\epsilon} \subseteq [C^0_0]^\delta_* \times \mathrm{club}_\kappa$ be defined by $R_{\delta,\epsilon}(\ell, C)$ if and only if $P_{\delta,\epsilon,\ell}(\iota) = \bar\eta$ for all $\iota \in [C]^\epsilon_*$. Fix $(\delta, \epsilon) \in [C^0_0]^2_*$. By definition of C^0_0, $\eta_{\delta,\epsilon} = \bar\eta$. By definition of $\eta_{\delta,\epsilon} = \bar\eta$, there is a club $D \subseteq Z$ such that $P_{\delta,\epsilon}(\ell) = \eta_{\delta,\epsilon,\ell} = \eta_{\delta,\epsilon} = \bar\eta$ for all $\ell \in [D]^\delta_*$. Fix an element $\ell \in [D]^\delta_*$. By definition of $\eta_{\delta,\epsilon,\ell} = \bar\eta$, there is a club $E \subseteq Z$ such that $P_{\delta,\epsilon,\ell}(\iota) = \eta_{\delta,\epsilon,\ell} = \eta_{\delta,\epsilon} = \bar\eta$ for all $\iota \in [E]^\epsilon_*$. Therefore $R_{\delta,\epsilon}(\ell, E)$. So $\ell \in \mathrm{dom}(R_{\delta,\epsilon})$. Since $\ell \in [D]^\delta_*$ was arbitrary, $[D]^\delta_* \subseteq \mathrm{dom}(R_{\delta,\epsilon})$. $R_{\delta,\epsilon}$ is $\subseteq$-downward closed in the club_κ-coordinate. By Fact 2.9, there is a club $F \subseteq D$ such that $R_{\delta,\epsilon}(\ell, F \setminus (\sup(\ell) + 1))$ for all $\ell \in \mathrm{dom}(R_{\delta,\epsilon}) \cap [F]^\delta_* = [F]^\delta_*$ (since $[D]^\delta_* \subseteq \mathrm{dom}(R_{\delta,\epsilon})$). Since $(\delta, \epsilon) \in [C^0_0]^2_*$ was arbitrary, it has been shown that for all $(\delta, \epsilon) \in [C^0_0]^2_*$, there is a club C with the property that $R_{\delta,\epsilon}(\ell, C \setminus (\sup(\ell) + 1))$ for all $\ell \in [C]^\delta_*$. Define $R \subseteq [C^0_0]^2_* \times \mathrm{club}_\kappa$ by $R((\delta, \epsilon), C)$ if and only if $R_{\delta,\epsilon}(\ell, C \setminus (\sup(\ell) + 1))$ for all $\ell \in [C]^\delta_*$. Note that $\mathrm{dom}(R) = [C^0_0]^2_*$ by the above discussion. R is $\subseteq$-downward closed in the club_κ-coordinate. By Fact 2.9, there is a club $C^0_1 \subseteq C^0_0$ such that $R((\delta, \epsilon), C^0_1 \setminus (\epsilon + 1))$ for all $(\delta, \epsilon) \in \mathrm{dom}(R) \cap [C^0_1]^2_* = [C^0_1]^2_*$. □

Claim 2. *There is a club $\bar C \subseteq C^0_1$ such that $\Phi(\ell)(\bar\eta) \neq \Phi(\iota)(\bar\eta)$ for all $\delta \in [\bar C]^1_*$, for all $\ell \in [\bar C \setminus (\delta + 1)]^\delta_*$, for all $\epsilon \in [\bar C \setminus (\sup(\ell) + 1)]^1_*$, and for all $\iota \in [\bar C \setminus (\epsilon + 1)]^\epsilon_*$.*

Proof of Claim 2. For $(\delta, \epsilon) \in [Z]^2_*$ and $\ell \in [Z]^\delta_*$, let $Q_{\delta,\ell,\epsilon} : [Z]^\epsilon_* \to 2$ be defined by $Q_{\delta,\ell,\epsilon}(\iota) = 0$ if and only if $\Phi(\ell)(\bar{\eta}) \neq \Phi(\iota)(\bar{\eta})$. By $\kappa \to_* (\kappa)^\epsilon_2$, there is a $j_{\delta,\ell,\epsilon} \in 2$ and a club $C \subseteq Z$ such that $Q_{\delta,\ell,\epsilon}(\iota) = j_{\delta,\ell,\epsilon}$ for all $\iota \in [C]^\epsilon_*$. For all $\delta \in [Z]^1_*$ and $\ell \in [Z]^\delta_*$, define $Q_{\delta,\ell} : [Z]^1_* \to 2$ by $Q_{\delta,\ell}(\epsilon) = j_{\delta,\ell,\epsilon}$. By $\kappa \to_* (\kappa)^1_2$, there is a $j_{\delta,\ell} \in 2$ and a club $C \subseteq Z$ such that $Q_{\delta,\ell}(\epsilon) = j_{\delta,\ell}$ for all $\epsilon \in [C]^1_*$. For all $\delta \in [Z]^1_*$, let $Q_\delta : [Z]^\delta_* \to 2$ be defined by $Q_\delta(\ell) = j_{\delta,\ell}$. By $\kappa \to_* (\kappa)^\delta_2$, there is $j_\delta \in 2$ and a club $C \subseteq Z$ such that $Q_\delta(\ell) = j_{\delta,\ell} = j_\delta$ for all $\ell \in [C]^\delta_*$. Define $Q : [Z]^1_* \to 2$ by $Q(\delta) = j_\delta$. By $\kappa \to_* (\kappa)^1_2$, there is a $\bar{j} \in 2$ and a fixed club $C^1_0 \subseteq Z$ such that $Q(\delta) = j_\delta = \bar{j}$ for all $\delta \in [C^1_0]^1_*$. Let $S \subseteq [Z]^1_* \times \mathrm{club}_\kappa$ be defined by $S(\delta, C)$ if and only if $j_{\delta,\ell} = \bar{j}$ for all $\ell \in [C]^\delta_*$. S is $\subseteq$-downward closed in the club_κ-coordinate. By Fact 2.9, there is a club $C^1_1 \subseteq C^1_0$ such that $S(\delta, C^1_1 \setminus (\delta+1))$ for all $\delta \in \mathrm{dom}(S) \cap [C^1_1]^1_*$. Suppose $\delta \in [C^1_0]^1_*$. By definition of C^1_0, $j_\delta = \bar{j}$. By definition of $j_\delta = \bar{j}$, there is some club $C \subseteq Z$ such that $j_{\delta,\ell} = j_\delta = \bar{j}$ for all $\ell \in [C]^\delta_*$, and hence $S(\delta, C)$. So $\delta \in \mathrm{dom}(S)$. Since $\delta \in [C^1_0]^1_*$ was arbitrary, this shows that $[C^1_0]^1_* \subseteq \mathrm{dom}(S)$, and hence $[C^1_1]^1_* \subseteq [C^1_0]^1_* \subseteq \mathrm{dom}(S)$. Therefore $S(\delta, C^1_1 \setminus (\delta+1))$ for all $\delta \in [C^1_1]^1_*$. Define $T \subseteq [Z]^{<\kappa}_* \times \mathrm{club}_\kappa$ by $T(\ell, C)$ if and only if $Q_{|\ell|,\ell,\epsilon}(\iota) = \bar{j}$ for all $\epsilon \in [C]^1_*$ and for all $\iota \in [C \setminus (\epsilon+1)]^\epsilon_*$. T is $\subseteq$-downward closed in the club_κ-coordinate and therefore by Fact 2.10, there is a club $C^1_2 \subseteq C^1_1$ such that $T(\ell, C^1_2 \setminus (\sup(\ell)+1))$ for all $\ell \in \mathrm{dom}(T) \cap \bigcup_{\delta \in [C^1_2]^1_*} [C^1_2 \setminus (\delta+1)]^\delta_*$. Fix $\ell \in \bigcup_{\delta \in [C^1_1]^1_*} [C^1_1 \setminus (\delta+1)]^\delta_*$. Thus $|\ell| \in [C^1_1]^1_*$. Define $U_\ell \subseteq Z \times \mathrm{club}_\kappa$ by $U_\ell(\epsilon, C)$ if and only if $Q_{|\ell|,\ell,\epsilon}(\iota) = \bar{j}$ for all $\iota \in [C]^\epsilon_*$. Since $|\ell| \in [C^1_1]^1_*$, $\ell \in [C^1_1 \setminus (|\ell|+1)]^{|\ell|}_*$, and $S(|\ell|, C^1_1 \setminus (|\ell|+1))$, one has that $j_{|\ell|,\ell} = \bar{j}$. By definition of $j_{|\ell|,\ell} = \bar{j}$, there is a club $D \subseteq Z$ such that $j_{|\ell|,\ell,\epsilon} = j_{|\ell|,\ell} = \bar{j}$ for all $\epsilon \in [D]^1_*$. Fix $\epsilon \in [D]^1_*$. By definition of $j_{|\ell|,\ell,\epsilon} = \bar{j}$, there is a club $E \subseteq Z$ such that $Q_{|\ell|,\ell,\epsilon}(\iota) = j_{|\ell|,\ell,\epsilon} = \bar{j}$ for all $\iota \in [E]^\epsilon_*$, and thus $U_\ell(\epsilon, E)$. So $\epsilon \in \mathrm{dom}(U_\ell)$. Since $\epsilon \in [D]^1_*$ was arbitrary, this shows that $[D]^1_* \subseteq \mathrm{dom}(U_\ell)$. U_ℓ is $\subseteq$-downward closed in the club_κ-coordinate. Thus by Fact 2.9, there is a club $F \subseteq D$ such that $U_\ell(\epsilon, F \setminus (\epsilon+1))$ for all $\epsilon \in \mathrm{dom}(U_\ell) \cap [F]^1_* = [F]^1_*$. Thus $Q_{|\ell|,\ell,\epsilon}(\iota) = \bar{j}$ for all $\epsilon \in [F]^1_*$ and for all $\iota \in [F \setminus (\epsilon+1)]^\epsilon_*$. Thus $T(\ell, F)$. This shows that ℓ is an element of $\mathrm{dom}(T)$. Since the element $\ell \in \bigcup_{\delta \in [C^1_1]^1_*} [C^1_1 \setminus (\delta+1)]^\delta_*$ was arbitrary, it has been shown that $\bigcup_{\delta \in [C^1_1]^1_*} [C^1_1 \setminus (\delta+1)]^\delta_* \subseteq \mathrm{dom}(T)$, and therefore $\bigcup_{\delta \in [C^1_2]^1_*} [C^1_2 \setminus (\delta+1)]^\delta_* \subseteq \bigcup_{\delta \in [C^1_1]^1_*} [C^1_1 \setminus (\delta+1)]^\delta_* \subseteq \mathrm{dom}(T)$. Thus we have $T(\ell, C^1_2 \setminus (\sup(\ell)+1))$ for all $\ell \in \bigcup_{\delta \in [C^1_2]^1_*} [C^1_2 \setminus (\delta+1)]^\delta_*$. In summary, $Q_{\delta,\ell,\epsilon}(\iota) = \bar{j}$ for all $\delta \in [C^1_2]^1_*$, for all $\ell \in [C^1_2 \setminus (\delta+1)]^\delta_*$, for all $\epsilon \in [C^1_2 \setminus (\sup(\ell)+1)]^1_*$, and all $\iota \in [C^1_2 \setminus (\epsilon+1)]^\epsilon_*$.

Let $\bar{C} = C^0_1 \cap C^1_2$. Let $\delta \in [\bar{C}]^1_*$, $\ell_0 \in [\bar{C} \setminus (\delta+1)]^\delta_*$, $\epsilon \in [\bar{C} \setminus (\sup(\ell_0)+1)]^1_*$, $\ell_1 \in [\bar{C} \setminus (\epsilon+1)]^\delta_*$, and $\iota \in [\bar{C} \setminus (\sup(\ell_1)+1)]^\epsilon_*$. For $i \in 2$, let $\iota^i \in [\bar{C} \setminus (\sup(\ell_1)+1)]^\epsilon_*$ be defined by $\iota^i(\alpha) = \iota(2 \cdot \alpha + i)$ and let $\sigma^i \in [\bar{C} \setminus (\epsilon+1)]^\epsilon_*$ be defined by $\sigma^i = \ell_1 \frown \iota^i$ (and note that $|\sigma^i| = \delta + \epsilon = \epsilon$ since $\epsilon \in Z$ is indecomposable). Since $(\delta, \epsilon) \in [\bar{C}]^2_*$, $\ell_1 \in [\bar{C} \setminus (\epsilon+1)]^\delta_*$, and $\iota \in [\bar{C} \setminus (\sup(\ell_1)+1)]^\epsilon_*$, one has, by Claim 1 and since $\bar{C} \subseteq C^0_1$, that $\chi(\sigma^0, \sigma^1) = \chi(\ell_1 \frown \iota^0, \ell_1 \frown \iota^1) = \bar{\eta}$. Thus $\Phi(\sigma^0)(\bar{\eta}) \neq \Phi(\sigma^1)(\bar{\eta})$. Hence

there must be some $\hat{i} \in 2$ such that $\Phi(\sigma^{\hat{i}})(\bar{\eta}) \neq \Phi(\ell_0)(\bar{\eta})$. Thus $Q_{\delta,\ell_0,\epsilon}(\sigma^{\hat{i}}) = 0$. Since $\delta \in [\bar{C}]^1_*$, $\ell_0 \in [\bar{C} \setminus (\delta+1)]^\delta_*$, $\epsilon \in [\bar{C} \setminus (\sup(\ell_0)+1)]^1_*$, and $\sigma^{\hat{i}} \in [\bar{C} \setminus (\epsilon+1)]^\epsilon_*$, $Q_{\delta,\ell_0,\epsilon}(\sigma^{\hat{i}}) = \bar{j}$. It has been shown that $\bar{j} = 0$. Since $\bar{C} \subseteq C^1_2$, the conclusion of the previous paragraph completes the argument for Claim 2. □

By Fact 2.8, let (Δ, Σ) be a suitable pair such that $\Delta : \kappa \times \omega \to [\bar{C}]^1_*$ and $\Sigma : L_\Delta \to [\bar{C}]^1_*$. Note that $\mathfrak{d}_\Delta : \kappa \to [\bar{C}]^1_*$. Let $\Lambda_{(\Delta,\Sigma)} : [\kappa]^\omega \to \bigcup_{\epsilon \in [\bar{C}]^1_*} [\bar{C} \setminus (\epsilon+1)]^\epsilon_*$ be the injection as defined in Definition 2.6 possessing the properties stated in Fact 2.7. Let $\Gamma : [\kappa]^\omega \to Y_{\bar{\eta}}$ be defined by $\Gamma(f) = \Phi(\Lambda_{(\Delta,\Sigma)}(f))(\bar{\eta})$.

Claim 3. *The function Γ is an injection and therefore $|[\kappa]^\omega| \leq |Y_{\bar{\eta}}|$.*

Proof of Claim 3. Suppose $f, g \in [\kappa]^\omega$ and $f \neq g$. First, suppose $\sup(f) \neq \sup(g)$ and without loss of generality, $\sup(f) < \sup(g)$. Note that $\mathfrak{d}_\Delta(\sup(f)) \in [\bar{C}]^1_*$. By Fact 2.7(2) and (5), we obtain that $\Lambda_{(\Delta,\Sigma)}(f) \in [\bar{C} \setminus (\mathfrak{d}_\Delta(\sup(f))+1)]^{\mathfrak{d}_\Delta(\sup(f))}_*$ and $\Lambda_{(\Delta,\Sigma)}(g) \in [\bar{C} \setminus (\mathfrak{d}_\Delta(\sup(g))+1)]^{\mathfrak{d}_\Delta(\sup(g))}_*$. It follows from Fact 2.7(3) that $\mathfrak{d}_\Delta(\sup(g)) \in [\bar{C} \setminus (\sup(\Lambda_{(\Delta,\Sigma)}(f))+1)]^1_*$. We use Claim 2 to conclude that $\Gamma(f) = \Phi(\Lambda_{(\Delta,\Sigma)}(f))(\bar{\eta}) \neq \Phi(\Lambda_{(\Delta,\Sigma)}(g))(\bar{\eta}) = \Gamma(g)$. Secondly, suppose $\sup(f) = \sup(g)$, $m \in \omega$ is least such that $f(m) \neq g(m)$, and without loss of generality, suppose $f(m) < g(m)$. Let $\delta = \Delta(\sup(f), m) = \Delta(\sup(g), m)$ and note that $\delta \in [\bar{C}]^1_*$ by definition of Δ. Let $\epsilon = \mathfrak{d}_\Delta(\sup(f)) = \mathfrak{d}_\Delta(\sup(g))$ and note that $\epsilon \in [\bar{C}]^1_*$, $\delta < \epsilon$, and hence $(\delta, \epsilon) \in [\bar{C}]^2_*$. By Fact 2.7(4), $\Lambda_{(\Delta,\Sigma)}(f) \restriction \delta = \Lambda_{(\Delta,\Sigma)}(g) \restriction \delta$. Let $\ell = \Lambda_{(\Delta,\Sigma)}(f) \restriction \delta = \Lambda_{(\Delta,\Sigma)}(g) \restriction \delta$. Note that $\epsilon < \ell(0)$ by Fact 2.7(2). The set $A = \{\Lambda_{(\Delta,\Sigma)}(f)(\alpha), \Lambda_{(\Delta,\Sigma)}(g)(\alpha) : \delta = \Delta(\sup(f), m) = \Delta(\sup(g), m) < \alpha < \epsilon\}$ has order type ϵ (since $\epsilon \in \bar{C} \subseteq Z$ is indecomposable). Let $\iota : \epsilon \to A$ be the increasing enumeration of A. Since $\Lambda_{(\Delta,\Sigma)}(f)$ and $\Lambda_{(\Delta,\Sigma)}(g)$ have the correct type, ι has the correct type, and thus $\iota \in [\bar{C} \setminus (\sup(\ell)+1)]^\epsilon_*$. By Fact 2.7(4), $\iota^0 = \mathrm{drop}(\Lambda_{(\Delta,\Sigma)}(f), \delta)$ and $\iota^1 = \mathrm{drop}(\Lambda_{(\Delta,\Sigma)}(g), \delta)$. Hence, $\ell \frown \iota^0 = \Lambda_{(\Delta,\Sigma)}(f)$ and $\ell \frown \iota^1 = \Lambda_{(\Delta,\Sigma)}(g)$. Thus $(\delta, \epsilon) \in [\bar{C}]^1_*$, $\ell \in [\bar{C} \setminus (\epsilon+1)]^\delta_*$, and $\iota \in [\bar{C} \setminus (\sup(\ell)+1)]^\epsilon_*$. By Claim 1 (and since $C^0_1 \supseteq \bar{C}$), $\bar{\eta} = \chi(\ell \frown \iota^0, \ell \frown \iota^1) = \chi(\Lambda_{(\Delta,\Sigma)}(f), \Lambda_{(\Delta,\Sigma)}(g))$. By definition of χ, $\Gamma(f) = \Phi(\Lambda_{(\Sigma,\Delta)}(f))(\bar{\eta}) \neq \Phi(\Lambda_{(\Delta,\Sigma)}(g))(\bar{\eta}) = \Gamma(g)$. □

Fact 2.12 [8]. *Suppose κ is a cardinal such that $\kappa \to_* (\kappa)^{<\kappa}_2$. For all $\chi < \kappa$, there is no injection of $[\kappa]^{<\kappa}$ into ${}^\chi\mathrm{ON}$, the class of χ length sequence of ordinals.*

Under AD^+, Fact 2.12 can be extended to all uncountable cardinals $\kappa < \Theta$. Let $\mathscr{P}_B(\kappa) = \{A \in \mathscr{P}(\kappa) : \sup(A) < \kappa\}$. Note that $|\mathscr{P}_B(\kappa)| \leq |[\kappa]^{<\kappa}|$. If κ is regular, then $|\mathscr{P}_B(\kappa)| = |[\kappa]^{<\kappa}|$. It was shown in [3] that under AD^+, if $\kappa < \Theta$ is singular, then $|\mathscr{P}_B(\kappa)| < |[\kappa]^{<\kappa}|$. The main result of [3] shows that under AD^+, for all $\omega < \kappa < \Theta$ and $\chi < \kappa$, $\mathscr{P}_B(\kappa)$ does not inject into ${}^\chi\mathrm{ON}$, and hence $[\kappa]^{<\kappa}$ does not inject into ${}^\chi\mathrm{ON}$.

Corollary 2.13. *Suppose κ is a cardinal satisfying $\kappa \to_* (\kappa)^\kappa_2$, $\lambda < \kappa$, and Y is a set. If $|[\kappa]^{<\kappa}| \leq |{}^\lambda Y|$, then $|[\kappa]^\omega| < |Y|$.*

Proof. Let $\Phi : [\kappa]^{<\kappa} \to {}^{\lambda}Y$ be an injection. For $\eta < \lambda$, let $Y_\eta = Y$. By applying Theorem 2.11 to $\langle Y_\eta : \eta < \lambda\rangle$, $|[\kappa]^\omega| \leq |Y|$. Suppose $|Y| = |[\kappa]^\omega|$. Let $\Sigma : Y \to [\kappa]^\omega$ be an injection. Define $\Pi : {}^{\lambda}Y \to {}^{\omega\cdot\lambda}\kappa$ by $\Pi(f)(\omega\cdot\delta + n) = \Sigma(f(\delta))(n)$ for all $\delta < \lambda$ and $n \in \omega$. The function Π is an injection. Then $\Pi \circ \Phi : [\kappa]^{<\kappa} \to {}^{\omega\cdot\lambda}\kappa$ is an injection which violates Fact 2.12. This shows that $|[\kappa]^\omega| < |Y|$. □

Remark 2.14. In light of the strict cardinality inequality of Corollary 2.13, one would expect a strict inequality in statement of Theorem 2.11: that is, there is an $\bar\eta$ such that $|[\kappa]^\omega| < |Y_{\bar\eta}|$, in the notation for Theorem 2.11. Unfortunately, the argument for Corollary 2.13 would require the existence of a sequence $\langle \Sigma_\eta : \eta < \lambda\rangle$ such that $\Sigma_\eta : Y_\eta \to [\kappa]^\omega$ is an injection for each $\eta < \lambda$. When $\lambda > \omega$, there is insufficient choice even under determinacy to find such a sequence. Of course, the ideal strategy to produce the strict inequality in Theorem 2.11 would be to answer the question of whether $[\kappa]^{<\kappa}$ is λ-prime for all $\lambda < \kappa$ when κ satisfies suitable partition properties: if $|[\kappa]^{<\kappa}| \leq \left|\prod_{\eta<\lambda} Y_\eta\right|$, then there is an $\bar\eta < \lambda$ such that $|[\kappa]^{<\kappa}| \leq |Y_{\bar\eta}|$.

The next fact says that if κ is a strong partition cardinal, then $\mathscr{P}(\kappa)$ is ON-regular.

Fact 2.15 [4]. *Suppose κ is a cardinal satisfying $\kappa \to_* (\kappa)^\kappa_2$. For all functions $\Psi : \mathscr{P}(\kappa) \to \mathrm{ON}$, there is an $\alpha \in \mathrm{ON}$ such that $|\Psi^{-1}[\{\alpha\}]| = |\mathscr{P}(\kappa)|$.*

Proof. Since $|\mathscr{P}(\kappa)| = |{}^\kappa 2| = |[\kappa]^\kappa_*|$, one will interchangeably use the various presentations of this cardinality. Assume the result fails. By replacement, for every function $\Psi : [\kappa]^\kappa_* \to \mathrm{ON}$, there is a $\delta \in \mathrm{ON}$ such that $\Psi : [\kappa]^\kappa_* \to \delta$. Let δ be the least ordinal such that there is a function $\Psi : [\kappa]^\kappa_* \to \delta$ with the property that for all $\alpha < \delta$, $|\Psi^{-1}[\{\alpha\}]| < |[\kappa]^\kappa_*|$. By minimality, δ must be a cardinal. If $f \in [\kappa]^\kappa_*$, then let f^0, $f^1 \in [\kappa]^\kappa_*$ be defined by $f^i(\alpha) = f(2\cdot\alpha + i)$ for both $i \in 2$. Define $P : [\kappa]^\kappa_* \to 2$ by $P(f) = 0$ if and only if $\Psi(f^0) \leq \Psi(f^1)$. By $\kappa \to_* (\kappa)^\kappa_2$, there is a club $C_0 \subseteq \kappa$ which is homogeneous for P. Suppose C_0 is homogeneous P taking value 1. Fix $g \in [C_0]^\kappa_*$. For each $n \in \omega$, let $g_n \in [C_0]^\kappa_*$ be defined by $g_n(\alpha) = g(\omega\cdot\alpha + n)$. For each $n \in \omega$, there is an $f_n \in [C_0]^\kappa_*$ such that $f^0 = g_n$ and $g^1 = g_{n+1}$. For all $n \in \omega$, since $P(f_n) = 1$, one has $\Psi(g_{n+1}) = \Psi(f^1) < \Psi(f^0) = \Psi(g_n)$. Thus $\langle \Psi(g_n) : n \in \omega\rangle$ is an infinite descending sequence of ordinals which is a contradiction. Thus C_0 is homogeneous for P taking value 0. Let $h \in [C_0]^\kappa_*$. Let $g \in [C_0]^\kappa_*$ be defined by $g(\alpha) = g(3\cdot\alpha + 2)$. Let $\Phi : {}^\kappa 2 \to [C_0]^\kappa_*$ be defined by $\Phi(p)(\alpha) = h(3\cdot\alpha + p(\alpha))$ where $p \in {}^\kappa 2$ implies p takes the form of a function from κ into 2. $\Phi : {}^\kappa 2 \to [C_0]^\kappa_*$ is an injection. For each $p \in {}^\kappa 2$, there is a unique $f_p \in [C_0]^\kappa_*$ such that $f_p^0 = \Phi(p)$ and $f_p^1 = g$. The statement $P(f_p) = 0$ implies that $\Psi(\Phi(p)) = \Psi(f_p^0) \leq \Psi(f_p^1) = \Psi(g)$. Let $\eta = \Psi(g)$ and note that $\eta + 1 < \delta$ since δ is a cardinal. Define $\Sigma : {}^\kappa 2 \to \eta + 1$ by $\Sigma(p) = \Psi(\Phi(p))$. For all $\alpha < \eta + 1$, $|\Sigma^{-1}[\{\alpha\}]| \leq |\Psi^{-1}[\{\alpha\}]| < |\mathscr{P}(\kappa)|$ by the assumption on Ψ and the fact that $\Phi : \Sigma^{-1}[\{\alpha\}] \to \Psi^{-1}[\{\alpha\}]$ is an injection. Thus Σ witnesses that $\eta + 1$ violates the minimality of δ. This contradiction implies that P has no homogeneous club. This contradicts $\kappa \to_* (\kappa)^\kappa_2$. □

Fact 2.16. *If κ is a cardinal satisfying $\kappa \to_* (\kappa)^\kappa_2$, then $|[\kappa]^{<\kappa}| < |\mathscr{P}(\kappa)|$.*

Proof. For the sake of a contradiction, suppose $\Phi : \mathscr{P}(\kappa) \to [\kappa]^{<\kappa}$ is an injection. Let $\Psi : \mathscr{P}(\kappa) \to \kappa$ be defined by $\Psi(X) = \mathrm{dom}(\Phi(f))$. By Fact 2.15, there is an $\epsilon < \kappa$ such that $|\Psi^{-1}[\{\epsilon\}]| = |\mathscr{P}(\kappa)|$. Let $\Sigma : \mathscr{P}(\kappa) \to \Psi^{-1}[\{\epsilon\}]$ be an injection. Then $\Phi \circ \Sigma : \mathscr{P}(\kappa) \to [\kappa]^\epsilon$ is an injection. Since $|[\kappa]^{<\kappa}| \leq |\mathscr{P}(\kappa)|$, this implies there is an injection $\Lambda : [\kappa]^{<\kappa} \to [\kappa]^\epsilon$ which contradicts Fact 2.12. □

Theorem 2.17. *Suppose κ is a cardinal satisfying $\kappa \to_* (\kappa)^\kappa_2$. Let $\langle Y_\alpha : \alpha < \kappa \rangle$ be a sequence of sets. If $|\mathscr{P}(\kappa)| \leq \left|\bigcup_{\lambda<\kappa} \prod_{\alpha<\lambda} Y_\alpha\right|$, then there is $\bar{\lambda} < \kappa$ such that $|[\kappa]^\omega| \leq |Y_{\bar{\lambda}}|$.*

Proof. Let $\Gamma : \mathscr{P}(\kappa) \to \bigcup_{\lambda<\kappa} \prod_{\alpha<\lambda} Y_\alpha$ be an injection. The idea is to consider the assignment of $A \in \mathscr{P}(\kappa)$ to the domain of $\Gamma(A)$ which is an ordinal less than κ. By Fact 2.15, there is a $\lambda < \kappa$ such that a subset of cardinality $\mathscr{P}(\kappa)$ is assigned to λ. This will induce a function from $[\kappa]^{<\kappa}$ into $\prod_{\alpha<\lambda} Y_\alpha$. Then the result will follow by applying Theorem 2.11. The details follow.

Define $\Psi : \mathscr{P}(\kappa) \to \kappa$ by $\Psi(A) = \mathrm{dom}(\Gamma(A))$. (Note that $\Gamma(A) \in \bigcup_{\lambda<\kappa} \prod_{\alpha<\lambda} Y_\alpha$ implies that $\Gamma(A)$ is a function on some $\lambda < \kappa$, and thus $\Psi(A)$ is this λ.) By Fact 2.15, there is a $\lambda < \kappa$ such that $|\Psi^{-1}[\{\lambda\}]| = |\mathscr{P}(\kappa)|$. Let $\Sigma : \mathscr{P}(\kappa) \to \Psi^{-1}[\{\lambda\}]$ be an injection. Since $|[\kappa]^{<\kappa}| < |\mathscr{P}(\kappa)|$, let $\Pi : [\kappa]^{<\kappa} \to \mathscr{P}(\kappa)$ be an injection. Let $\Phi : [\kappa]^{<\kappa} \to \prod_{\alpha<\lambda} Y_\alpha$ be defined by $\Phi = \Gamma \circ \Sigma \circ \Pi$. Since Φ is an injection, there is some $\bar{\lambda} < \kappa$ such that $|[\kappa]^\omega| \leq |Y_{\bar{\lambda}}|$ by Theorem 2.11. □

Corollary 2.18. *Let κ be a cardinal satisfying $\kappa \to_* (\kappa)^\kappa_2$ and let Y be a set. If $|\mathscr{P}(\kappa)| \leq |{}^{<\kappa}Y|$, then $|[\kappa]^\omega| < |Y|$.*

Proof. Let $\Phi : \mathscr{P}(\kappa) \to {}^{<\kappa}Y$ be an injection. For each $\alpha < \kappa$, let $Y_\alpha = Y$. By applying Theorem 2.17 to $\langle Y_\alpha : \alpha < \kappa \rangle$, $|[\kappa]^\omega| \leq |Y|$. Now suppose $|[\kappa]^\omega| = |Y|$. (The key observation is that $|{}^{<\kappa}([\kappa]^\omega)| = |{}^{<\kappa}\kappa|$.) Let $\Sigma : Y \to [\kappa]^\omega$ be an injection. Define $\Pi : {}^{<\kappa}Y \to {}^{<\kappa}\kappa$ as follows: if $f \in {}^\lambda Y$ for some $\lambda < \kappa$, then $\Pi(f) \in {}^{\omega\cdot\lambda}\kappa$ is defined by $\Pi(f)(\omega \cdot \delta + n) = \Sigma(f(\delta))(n)$ for all $\delta < \kappa$ and $n \in \omega$. The function Π is an injection. Then $\Pi \circ \Phi : \mathscr{P}(\kappa) \to {}^{<\kappa}\kappa$ is an injection which violates Fact 2.16. This shows that $|[\kappa]^\omega| < |Y|$. □

3. The well-ordered power set operation

To study the well-ordered power set operation, some results concerning determinacy in the presence of ∞-Borel codes will be necessary. Below the two relevant facts will be summarized. See [2; 18] for more information about ∞-Borel codes.

Definition 3.1. A pair (S, φ) is an ∞-Borel code if and only if S is a set of ordinals and φ is a formula of set theory. A set $A \subseteq \mathbb{R}$ is ∞-Borel if and only if there is an ∞-Borel code (S, φ) such that $A = \{x \in \mathbb{R} : L[S, x] \models \varphi(S, x)\}$.

The following result states that if AD and $\mathsf{DC}_{\mathbb{R}}$ hold and J is a set of ordinals, then there is a uniform procedure which translate an $\mathrm{OD}^{L(J,\mathbb{R})}_{\{J,z\}}$ definition for a set of reals to an $\mathrm{OD}^{L(J,\mathbb{R})}_{\{J,z\}}$ ∞-Borel code for that set of reals. This is done using Woodin's result that any model of the form $L(J,\mathbb{R}) \models \mathsf{AD} + \mathsf{DC}$ is a "symmetric extension" of $\mathrm{HOD}^{L(J,\mathbb{R})}_{\{J\}}$ when J is a set of ordinals. In particular, truth in $L(J,\mathbb{R})$ can be captured by a suitable remainder forcing of the ω-direct limit of the $\mathrm{OD}_{\{J\}}$ ∞-Borel code forcing on the various $\mathbb{R}^n$. See [2; 3] for the details concerning this forcing and these results in the model $L(\mathbb{R})$.

Fact 3.2 (Woodin). *Assume* AD *and* $\mathsf{DC}_{\mathbb{R}}$. *Then there is a function* Ξ *such that* $\Xi(\phi, J, z, \vec{\alpha})$ *is an* ∞*-Borel code for the set* $A = \{x \in \mathbb{R} : L(J,\mathbb{R}) \models \phi(x, J, z, \vec{\alpha})\}$ *for all formula* ϕ *of set theory, sets of ordinals* J, $z \in \mathbb{R}$, *and finite sequences of ordinals* $\vec{\alpha}$. (*Moreover, the definition of* $\Xi(\phi, J, z, \vec{\alpha})$ *is absolute into* $L(J,\mathbb{R})$, *and thus* $\Xi(\phi, J, z, \vec{\alpha}) \in L(J,\mathbb{R})$.)

Proof. This is shown for $L(\mathbb{R})$ in [2, Theorem 7.19]. The same argument works for $L(J,\mathbb{R})$ after suitable relativization. Moreover, the proof provides a uniform manner of obtaining such ∞-Borel code from ϕ, J, z, and $\vec{\alpha}$. Also [3] provides a proof in full generality. □

Fact 3.3 (Woodin perfect set dichotomy; [2, Theorem 8.5]). *Assume* AD *and* $\mathsf{DC}_{\mathbb{R}}$. *Let* E *be an equivalence relation with* ∞*-Borel code* (S, φ). *Then exactly one of the following holds*:

(1) $\mathbb{R}/E \subseteq \mathrm{OD}^{L(S,\mathbb{R})}_{\{S\}}$, *and hence* $\mathbb{R}/E$ *has an* $\mathrm{OD}^{L(S,\mathbb{R})}_{\{S\}}$ *well-ordering* (*which is obtained uniformly from* (S, φ)).

(2) $|\mathbb{R}| \leq |\mathbb{R}/E|$ (*there is an injection of* $\mathbb{R}$ *into* $\mathbb{R}/E$).

Proof. The definability of the well-ordering in the first case can be seen by the inspecting the argument corresponding to the first case. See the proof of this result as presented in [5, Theorem 3.2] which states this more explicitly. □

Definition 3.4. Let X be a set. Let $\mathscr{P}_{\mathrm{WO}}(X)$ be the collection of all sets $A \subseteq X$ such that A is well-orderable. If κ is a cardinal, let $\mathscr{P}_\kappa(X)$ be the collection of all $A \subseteq X$ such that $|A| < \kappa$ (that is, A injects into κ but κ does not inject into A). Note that for all cardinals κ, $\mathscr{P}_\kappa(X) \subseteq \mathscr{P}_{\mathrm{WO}}(X)$.

If $\delta \in \mathrm{ON}$, then let $\mathcal{I}_{<\delta}(X)$ be the set of all injective functions $f : \epsilon \to X$ with $\epsilon < \delta$. Let $\mathcal{I}_{<\mathrm{ON}}(X)$ be the set of all injective functions $f : \epsilon \to X$ for some ordinal ϵ. (By the axiom of replacement, there is a cardinal κ such that $\mathcal{I}_{<\kappa}(X) = \mathcal{I}_{<\mathrm{ON}}(X)$.) Let $\mathfrak{r} : \mathcal{I}_{<\mathrm{ON}}(X) \to \mathscr{P}_{\mathrm{WO}}(X)$ be defined by $\mathfrak{r}(f) = f[\mathrm{dom}(f)]$. The function $\mathfrak{r} : I_{<\mathrm{ON}}(X) \to \mathscr{P}_{\mathrm{WO}}(X)$ is a surjection. For any cardinal κ, the restriction of $\mathfrak{r}$ to $\mathcal{I}_{<\kappa}(X)$ is a surjection of the form $\mathfrak{r} : \mathcal{I}_{<\kappa}(X) \to \mathscr{P}_\kappa(X)$.

Theorem 3.5. *Assume* AD, $\mathsf{DC}_{\mathbb{R}}$, *and all sets of reals have an* ∞*-Borel code. Let* $\kappa < \Theta$ *and* Y *be a surjective image of* $\mathbb{R}$. *Let* $\Phi : \mathscr{P}(\kappa) \to \mathscr{P}_{\mathrm{WO}}(Y)$. *Let*

$\mathfrak{r}: \mathcal{I}_{<\mathrm{ON}}(Y) \to \mathscr{P}_{\mathrm{WO}}(X)$ *be defined by* $\mathfrak{r}(f) = f[\mathrm{dom}(f)]$. *Then there is a function* $\tilde{\Phi}: \mathscr{P}(\kappa) \to \mathcal{I}_{<\mathrm{ON}}(Y)$ *such that* $\Phi = \mathfrak{r} \circ \tilde{\Phi}$.

Proof. For each $A \in \mathscr{P}(\kappa)$, one will uniformly define an equivalence relation E_A on a subset $R_A \subseteq \mathbb{R}$ along with the ∞-Borel code for E_A so that $\Phi(A)$ is in bijection with R_A/E_A. Fact 3.3 will uniformly provide a well-ordering for R_A/E_A which will be used to define $\tilde{\Phi}(A)$. The precise details follow.

Since $\kappa < \Theta$, let $\phi: \mathbb{R} \to \kappa$ be a surjection. Define a pre-well-ordering $\preceq$ on $\mathbb{R}$ by $x \preceq y$ if and only if $\phi(x) \leq \phi(y)$. Let $\mathrm{rk}_{\preceq}(x)$ be the rank of x in the pre-well-ordering $\preceq$ and note that $\mathrm{rk}_{\preceq} = \phi$. Let Γ be any nonselfdual pointclass closed under $\wedge$, $\exists^{\mathbb{R}}$, and containing $\preceq$ (for example, the pointclass $\Sigma^1_1(\preceq)$ frequently used for the Moschovakis coding lemma). Let $U \subseteq \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ be a set in Γ and universal for subsets of $\mathbb{R}^2$ in Γ (which exists by an application of the Wadge lemma). Let $\bar{0} \in \mathbb{R}$ be the constant 0 function. For each $e \in \mathbb{R}$, define the set $C_e = \{\alpha \in \kappa : (\exists x)(\mathrm{rk}_{\preceq}(x) = \alpha \wedge U(e, x, \bar{0}))\}$. For each $A \subseteq \kappa$, let $T_A \subseteq \mathbb{R} \times \mathbb{R}$ be defined by $T_A(x, y)$ if and only if $\mathrm{rk}_{\preceq}(x) \in A \wedge y = \bar{0}$. By the Moschovakis coding lemma, there is a $V \subseteq \mathbb{R}^2$ with $V \in \Gamma$ such that $V \subseteq T_A$ and for all $\alpha < \kappa$, $T_A \cap (\mathrm{rk}_{\preceq}^{-1}[\{\alpha\}] \times \mathbb{R}) \neq \varnothing$ if and only if $V \cap (\mathrm{rk}_{\preceq}^{-1}[\{\alpha\}] \times \mathbb{R}) \neq \varnothing$. Since U is universal for subsets of $\mathbb{R}^2$ in Γ, there is an $e \in \mathbb{R}$ such that $V = U_e = \{(x, y) : U(e, x, y)\}$. Then $C_e = A$. Thus the map $e \mapsto C_e$ is a surjection of $\mathbb{R}$ onto $\mathscr{P}(\kappa)$ via the explicit coding procedure described above using $\preceq$ and the universal set U. Note that this coding is absolute to any inner model M containing all the reals, $\preceq$, and U. This implies that $\mathscr{P}(\kappa) = \mathscr{P}(\kappa) \cap M$.

Since Y is a surjective image of $\mathbb{R}$, without of generality, assume $Y = \mathbb{R}/E$ for some equivalence relation E on $\mathbb{R}$. Define $R \subseteq \mathscr{P}(\kappa) \times \mathbb{R}$ by $R(A, z)$ if and only if $[z]_E \in \Phi(A)$. Note that for all $A \in \mathscr{P}(\kappa)$, $R_A = \{z \in \mathbb{R} : R(A, z)\} = \bigcup \Phi(A)$ and $R_A/E = \Phi(A)$. Define $\widehat{R} \subseteq \mathbb{R} \times \mathbb{R}$ by $\widehat{R}(e, z)$ if and only if $R(C_e, z)$.

Since all sets of reals have ∞-Borel codes, let (S_0, φ_0'), (S_1, φ_1'), (S_2, φ_2'), and (S_3, φ_3') be ∞-Borel codes for $\preceq$, U, E, and $\widehat{R}$, respectively. For notational simplicity, merge all four sets of ordinals S_0, S_1, S_2, and S_3 into a single set of ordinals T in some simple manner. Then one can find four formulas φ_0, φ_1, φ_2, and φ_3 so that (T, φ_0), (T, φ_1), (T, φ_2), and (T, φ_3) are ∞-Borel codes for $\preceq$, U, E, and $\widehat{R}$, respectively.

Thus $\preceq, U, E, \widehat{R} \in L(T, \mathbb{R})$. By the discussion above, $\mathscr{P}(\kappa) \subseteq L(T, \mathbb{R})$. Thus $R \in L(T, \mathbb{R})$ since within $L(T, \mathbb{R})$, one can define $R(A, z)$ if and only if there exists an $e \in \mathbb{R}$ such that $\widehat{R}(e, z)$ and $C_e = A$. Moreover using the ∞-Borel codes (T, φ_0), (T, φ_1), (T, φ_2), and (T, φ_3), the above shows that R is $\mathrm{OD}^{L(T,\mathbb{R})}_{\{T\}}$. Let $A \in \mathscr{P}(\kappa)$. Define $E_A \subseteq \mathbb{R} \times \mathbb{R}$ by $x \; E_A \; y$ if and only if $x \; E \; y \wedge R(A, x) \wedge R(A, y)$ which is an equivalence relation on $R_A = \{x : R(A, x)\}$. Note that E_A and R_A are $\mathrm{OD}^{L(T,\mathbb{R})}_{\{T,A\}}$. Since T and A can be merged into a single set of ordinals and the equivalence relation E_A is $\mathrm{OD}^{L(T,\mathbb{R})}_{\{T,A\}}$, Fact 3.2 implies there is ∞-Borel code

for E_A which is $\mathrm{OD}^{L(T,\mathbb{R})}_{\{T,A\}}$. Let (S_A, φ_A) be the least such ∞-Borel according to the canonical ordering of $\mathrm{OD}^{L(T,\mathbb{R})}_{\{T,A\}}$ (or let it be the ∞-Borel code obtained by the uniform procedure of Fact 3.2). Since $R_A/E_A = R_A/E = \Phi(A)$, R_A/E_A is well-orderable. Therefore by Fact 3.3, R_A/E_A has an $\mathrm{OD}^{L(T,\mathbb{R})}_{\{T,A\}}$ well-ordering. Let $\sqsubset_A$ be the $\mathrm{OD}^{L(T,\mathbb{R})}_{\{T,A\}}$-least such well-ordering (or more explicitly, the well-ordering uniformly defined from (S_A, φ_A) as in the proof of Fact 3.3 found in [2]). Let $\psi_A : \mathrm{ot}(\sqsubset_A) \to R_A/E_A$ be the canonical order-preserving enumeration of $R_A/E_A = \Phi(A)$ via $\sqsubset_A$. Let $\tilde{\Phi}(A) = \psi_A$. Thus $\tilde{\Phi} : \mathscr{P}(\kappa) \to \mathcal{I}_{<\mathrm{ON}}(Y)$. Since $\tilde{\Phi}(A) = \psi_A$ enumerates $\Phi(A) = R_A/E_A$, $\mathfrak{r}(\tilde{\Phi}(A)) = \Phi(A)$. Thus $\mathfrak{r} \circ \tilde{\Phi} = \Phi$. □

In other words, Theorem 3.5 states that for all $\Phi : \mathscr{P}(\kappa) \to \mathscr{P}_{\mathrm{WO}}(Y)$, there is a function $\tilde{\Phi} : \mathscr{P}(\kappa) \to \mathcal{I}_{<\mathrm{ON}}(Y)$ such that $\tilde{\Phi}(A)$ is an well-ordered enumeration of $\Phi(A)$ for any $A \in \mathscr{P}(\kappa)$. Note that there is in general no map $\Gamma : \mathscr{P}_{\mathrm{WO}}(Y) \to \mathcal{I}_{<\mathrm{ON}}(Y)$ which sends a well-orderable $B \subseteq Y$ to a well-ordered enumeration $\Gamma(B)$ of B. In fact, there is in general no injection of $\mathscr{P}_{\mathrm{WO}}(Y)$ into $\mathcal{I}_{<\mathrm{ON}}(Y)$. For example, let $Y = \mathbb{R}$. Let E_0 be the equivalence relation on ${}^{\omega}2$ defined by $x \; E_0 \; y$ if and only if there exists an $m \in \omega$ such that for all $n \geq m$, $x(n) = y(n)$. If $\mathsf{AC}^{\mathbb{R}}_{\omega}$ holds and all subsets of $\mathbb{R}$ have the property of Baire, then $\mathbb{R}/E_0$ is not linearly orderable. $\mathbb{R}/E_0$ injects into $\mathscr{P}_{\omega_1}(\mathbb{R})$ by the inclusion map since E_0 is an equivalence relation with all countable classes. Thus $\mathscr{P}_{\omega_1}(\mathbb{R})$ is not linearly orderable. However $\mathcal{I}_{<\omega_1}(\mathbb{R})$ is linearly orderable by the lexicographic ordering. (Also one can inject $\mathcal{I}_{<\omega_1}(\mathbb{R})$ into $\mathscr{P}(\omega_1)$ and use the lexicographic ordering on $\mathscr{P}(\omega_1)$.) This shows that $\mathscr{P}_{\omega_1}(\mathbb{R})$ does not inject into $\mathcal{I}_{<\omega_1}(\mathbb{R})$. This discussion also shows $|\mathcal{I}_{<\omega_1}(\mathbb{R})| < |\mathscr{P}_{\omega_1}(\mathbb{R})|$.

Theorem 3.6. *Assume* AD, $\mathsf{DC}_{\mathbb{R}}$, *and every subset of* $\mathbb{R}$ *has an* ∞*-Borel code. Let* $\kappa < \Theta$ *satisfy* $\kappa \to_* (\kappa)^{\kappa}_2$. *Let* Y *be a surjective image of* $\mathbb{R}$. *If* $|\mathscr{P}(\kappa)| \leq |\mathscr{P}_{\kappa}(Y)|$, *then* $|[\kappa]^{\omega}| < |Y|$.

Proof. Let $\Phi : \mathscr{P}(\kappa) \to \mathscr{P}_{\kappa}(Y)$ be an injection. By Theorem 3.5, there is a $\tilde{\Phi} : \mathscr{P}(\kappa) \to \mathcal{I}_{<\mathrm{ON}}(Y)$ such that $\Phi = \mathfrak{r} \circ \tilde{\Phi}$. Since Φ is injective, this implies that $\tilde{\Phi}$ must be injective. For all $A \in \mathscr{P}(\kappa)$, $\Phi(A) = \mathfrak{r}(\tilde{\Phi}(A)) = \tilde{\Phi}[\mathrm{dom}(\tilde{\Phi}(A))]$ implies that $\tilde{\Phi}(A)$ is a well-ordered injective enumeration of $\Phi(A)$. Since $\Phi(A) \in \mathscr{P}_{\kappa}(Y)$, $\tilde{\Phi}(A) \in \mathcal{I}_{<\kappa}(Y)$, and thus $\tilde{\Phi} : \mathscr{P}(\kappa) \to \mathcal{I}_{<\kappa}(Y)$. Since $\mathcal{I}_{<\kappa}(Y) \subseteq {}^{<\kappa}Y$, $\tilde{\Phi}$ may be regarded as an injective function $\tilde{\Phi} : \mathscr{P}(\kappa) \to {}^{<\kappa}Y$. Then $|[\kappa]^{\omega}| < |Y|$ by Corollary 2.18. □

References

[1] A. E. Caicedo and R. Ketchersid, "A trichotomy theorem in natural models of AD^{+}", pp. 227–258 in *Set theory and its applications*, edited by L. Babinkostova et al., Contemp. Math. **533**, Amer. Math. Soc., Providence, RI, 2011. MR

[2] W. Chan, "An introduction to combinatorics of determinacy", pp. 21–75 in *Trends in set theory*, edited by S. Coskey and G. Sargsyan, Contemp. Math. **752**, Amer. Math. Soc., Providence, RI, 2020. MR

[3] W. Chan, "Cardinality of the set of bounded subsets of a cardinal", In preparation.

[4] W. Chan, "Definable combinatorics of strong partition cardinals", In preparation.

[5] W. Chan and S. Jackson, "Cardinality of wellordered disjoint unions of quotients of smooth equivalence relations", *Ann. Pure Appl. Logic* **172**:8 (2021), art. id. 102988. MR

[6] W. Chan and S. Jackson, "Definable combinatorics at the first uncountable cardinal", *Trans. Amer. Math. Soc.* **374**:3 (2021), 2035–2056. MR

[7] W. Chan and S. Jackson, "Applications of infinity-Borel codes to definability and definable cardinals", *Fund. Math.* **265**:3 (2024), 215–258. MR

[8] W. Chan, S. Jackson, and N. Trang, "Almost everywhere behavior of functions according to partition measures", *Forum Math. Sigma* **12** (2024), art. id. e16. MR

[9] W. Chan, S. Jackson, and N. Trang, "The primeness property for cardinalities", In preparation.

[10] W. Chan, S. Jackson, and N. Trang, "A search for the cofinality of the power set of the first uncountable cardinal", In preparation.

[11] S. Cox and J. Krueger, "Namba forcing, weak approximation, and guessing", *J. Symb. Log.* **83**:4 (2018), 1539–1565. MR

[12] S. Jackson, "AD and the projective ordinals", pp. 117–220 in *Cabal Seminar* 81–85, edited by A. S. Kechris et al., Lecture Notes in Math. **1333**, Springer, 1988. MR

[13] S. Jackson, *A computation of* δ^1_5, Mem. Amer. Math. Soc. **670**, Amer. Math. Soc., Providence, RI, 1999. MR

[14] S. Jackson, "Structural consequences of AD", pp. 1753–1876 in *Handbook of set theory*, *I*, edited by M. Foreman and A. Kanamori, Springer, 2010. MR

[15] A. S. Kechris, "AD and projective ordinals", pp. 91–132 in *Cabal Seminar* 76–77 (Pasadena, CA, 1976–77), edited by A. S. Kechris and Y. N. Moschovakis, Lecture Notes in Math. **689**, Springer, 1978. MR

[16] A. S. Kechris, E. M. Kleinberg, Y. N. Moschovakis, and W. H. Woodin, "The axiom of determinacy, strong partition properties and nonsingular measures", pp. 75–99 in *Cabal Seminar* 77–79 (Pasadena, CA, 1977–79), edited by A. S. Kechris et al., Lecture Notes in Math. **839**, Springer, 1981. MR

[17] E. M. Kleinberg, *Infinitary combinatorics and the axiom of determinateness*, Lecture Notes in Math. **612**, Springer, 1977. MR

[18] P. B. Larson, *Extensions of the axiom of determinacy*, University Lecture Series **78**, Amer. Math. Soc., Providence, RI, 2023. MR

[19] J. R. Steel, "An outline of inner model theory", pp. 1595–1684 in *Handbook of set theory*, *I*, edited by M. Foreman and A. Kanamori, Springer, 2010. MR

[20] J. R. Steel, "Ordinal definability in models of determinacy: introduction to part V", pp. 3–48 in *Ordinal definability and recursion theory*: *the Cabal seminar*, *III*, edited by A. S. Kechris et al., Lect. Notes Log. **43**, Assoc. Symbol. Logic, Ithaca, NY, 2016. MR

[21] W. H. Woodin, "The cardinals below $|[\omega_1]^{<\omega_1}|$", *Ann. Pure Appl. Logic* **140**:1-3 (2006), 161–232. MR

Received July 13, 2025. Revised March 23, 2026.

WILLIAM CHAN
INSTITUTE FOR DISCRETE MATHEMATICS AND GEOMETRY
VIENNA UNIVERSITY OF TECHNOLOGY
VIENNA
AUSTRIA
william.chan@tuwien.ac.at

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 journals+pjm@msp.org 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.

At submission time a PDF file is required, and it must be derived from a LaTeX source file in which the command `\documentclass[12pt]{amsart}` is active. No changes to horizontal or vertical margins, or line spacing, should be made. Please include in the preamble the command `\usepackage[pagewise]{lineno}\linenumbers` — this will spare referees tedious effort in line counting.

Carefully preserve all related 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 journals+pjm@msp.org.

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 344 No. 1 September 2026

Minimal rational graphs admitting a $\mathbb{Q}$HD smoothing 1
MÁRTON BEKE

Iwahori–Hecke model for the universal supersingular representation 27
ANAND CHITRAO and ASFAK SONEJI

The canonical exact sequence of differential modules for 0-dimensional schemes 47
TRAN N. K. LINH and LE NGOC LONG

On the measure of primitive representations of an integer by a quadratic form 69
MANABU MURATA

Group topologies on groups of absolutely bicontinuous homeomorphisms 103
J. DE LA NUEZ GONZÁLEZ

Adjunction inequality for spatially refined s-invariants 145
QIUYU REN

On the invariants of finite groups arising in a topological quantum field theory 151
CHRISTOPHER A. SCHROEDER and HUNG P. TONG-VIET

From brick manifolds to Grassmannians of bimodules 183
EVGENY FEIGIN and MARKUS REINEKE

On mixed curvature for Hermitian manifolds 201
KAI TANG

The size of a set from the size of its countable power set 219
WILLIAM CHAN