# On the algebra generated by three commuting matrices: combinatorial cases

Ron Cherny Matthew Satriano Yohan Song

Submitted: Mar 25, 2024; Accepted: Sep 30, 2024; Published: Nov 15, 2024



**Abstract**

Gerstenhaber proved in 1961 that the unital algebra generated by a pair of commuting $d \times d$ matrices over a field has dimension at most $d$. It is an open problem whether the analogous statement is true for triples of matrices which pairwise commute. We answer this question for special classes of triples of matrices arising from combinatorial data.

**Mathematics Subject Classifications:** 05E40, 15A27

## 1 Introduction

Given a field $k$, the well-known Cayley–Hamilton theorem asserts that every matrix $A \in \mathrm{M}_d(k)$ is a root of its characteristic polynomial. In particular, the (unital) algebra generated by $A$ is a $k$-vector space of dimension at most $d$. It follows from a result of Motzkin and Taussky [MT55] (shown independently by Gerstenhaber [Ger61]) that if $A, B \in \mathrm{M}_d(k)$ are commuting $d \times d$ matrices, then the algebra they generate has dimension at most $d$. In contrast, it is known that for all $n \geqslant 4$, there exists $d$ and pairwise commuting matrices $A_1, \ldots, A_n \in \mathrm{M}_d(k)$ such that the algebra generated by the $A_i$ has dimension strictly larger than $d$, see Example 1. It has been a longstanding open question (referred to, e.g. in [O'M20], as the *Gerstenhaber problem*) to determine whether pairwise commuting matrices $A, B, C \in \mathrm{M}_d(k)$ generate an algebra of dimension at most $d$.

Let us begin by discussing some of the known cases of the Gerstenhaber problem. To begin, the results in [MT55, Ger61] are algebro-geometric, showing irreducibility of the algebraic variety $C(2, d)$ parameterizing pairs $(A, B)$ of commuting $d \times d$ matrices. From this, one reduce to the case of *generic* pairs $(A, B)$ of commuting matrices, which are simultaneously diagonalizable; hence, the result follows from Cayley–Hamilton. In fact, this technique of reducing to simultaneously diagonalizable matrices works whenever we have irreducibility of the variety $C(n, d)$, parameterizing pairwise commuting $d \times d$

Department of Pure Mathematics, University of Waterloo, Waterloo, ON, Canada (msatriano@uwaterloo.ca, rcherny@uwaterloo.ca, y343song@uwaterloo.ca).

matrices $(A_1, \ldots, A_n)$. Unfortunately, for $n \geqslant 3$, $C(n,d)$ is notoriously complicated. For $n \geqslant 4$ and $d \geqslant 4$, $C(n,d)$ has multiple irreducible components, see [Ger61, Gur92]. For $n = 3$, much less is known: the variety $C(n,d)$ is irreducible for $d \leqslant 10$ [Š12] and reducible for $d \geqslant 29$ [HO01, NŠ14]. See also [JŠ22] for further results on the structure of components of $C(n,d)$. In general, the Gerstenhaber problem is reduced to checking at the generic points of every irreducible component of $C(n,d)$; however such an approach is essentially intractable.

The aforementioned results are geometric and concern the structure of $C(n,d)$. Other commutative algebraic and linear algebraic proofs that commuting pairs of $d \times d$ matrices generate an algebra of dimension at most $d$ were later discovered [BH90, LL91, NS99, Ber13]. In addition to the case $d \leqslant 10$ mentioned above, several other cases of the Gerstenhaber problem are known when one imposes linear algebraic constraints. For example, if one of the three $d \times d$ matrices $A_1, A_2, A_3$ has nullity at most 3, then it was shown in [GS00, Š12] that the algebra the matrices generate has dimension at most $d$. The Gerstenhaber problem is also known if one of the matrices has index at most 2, i.e., some $A_i^2 = 0$, see [HO01]. We refer to [Set11, HO15] for a survey of further results.

In this paper, the viewpoint we take is to break up the Gerstenhaber problem based on the minimal number of *generating vectors* required. Given pairwise commuting $d \times d$ matrices $(A_1, \ldots, A_n)$, let $\mathcal{A}$ be the algebra they generate. We say $v_1, \ldots, v_r \in k^d$ are generating vectors if

$$\operatorname{Span}\{Av_j \mid A \in \mathcal{A},\, j \leqslant r\} = k^d.$$

When $r = 1$, it is straightforward to check that $\dim_k \mathcal{A} \leqslant d$. On the other hand, for $r = 2$, the Gerstenhaber problem is still open and highly non-trivial. Indeed, the simplest subcase when $r = 2$ is when $A_i v_2 \in \operatorname{Span}\{Av_1 \mid A \in \mathcal{A},\, j \leqslant r\}$ for all $i$; this case was only recently resolved by Rajchgot and the second author [RS18, Theorem 1.5], showing $\dim_k \mathcal{A} \leqslant d$.

Following [RSS20], our current paper considers a broad class of combinatorially motivated examples when $r = 2$. Before giving the formal definition, we begin with an example.

**Example 1.** As mentioned above, there exist choices of pairwise commuting matrices $A_1, \ldots, A_4 \in \mathrm{M}_d(k)$ such that the algebra $\mathcal{A}$ generated by the $A_i$ has $\dim_k \mathcal{A} > d$. The standard such example is given by taking $d = 4$ and letting the four pairwise commuting matrices be $E_{13}$, $E_{23}$, $E_{14}$, and $E_{24}$. Then $\mathcal{A}$ has a basis given by these matrices as well as the identity matrix $I$, hence $\dim_k \mathcal{A} = 5 > 4 = d$.

This example can be understood combinatorially as follows. Let $S = k[x_1, \ldots, x_4]$ and consider the monomial ideals

$$I = (x_1, x_2)^2 + (x_3, x_4) \quad \text{and} \quad J = (x_3, x_4)^2 + (x_1, x_2).$$

We consider the $S$-module $M$ obtained from $S/I \oplus S/J$ by gluing $(x_1, 0)$ to $(0, x_3)$, and gluing $(x_2, 0)$ to $(0, x_4)$, i.e.,

$$M = (S/I \oplus S/J)/\langle (x_1, 0) - (0, x_3), (x_2, 0) - (0, x_4) \rangle.$$

Then $M$ is a vector space of dimension $d = 4$ with basis $(1, 0)$, $(0, 1)$, $(x_1, 0)$, and $(x_2, 0)$. Multiplication by $x_i$ on $M$ yields $n = 4$ commuting matrices. These matrices are precisely the same as the standard example given in the previous paragraph. $\diamond$

With this as motivation, we now define our combinatorial matrices, cf. [RSS20, §4.2].

**Definition 2.** Let $S = k[x_1, \dots, x_n]$. Let $I \subset K \subset S$ and $J \subset L \subset S$ be monomial ideals with $\dim_k S/I < \infty$ and $\dim_k S/J < \infty$. Given an isomorphism $\phi\colon K/I \xrightarrow{\simeq} L/J$ of $S$-modules sending monomials to monomials, we obtain an $S$-module

$$M = (S/I \oplus S/J)/\langle\, (f, -\phi(f)) \mid f \in K/I\,\rangle$$

equipped with the natural monomial basis. Letting $d = \dim M$ and $A_i$ be the $d \times d$ matrix given by the linear map $M \xrightarrow{\cdot x_i} M$, we say $(A_1, \dots, A_n)$ are *associated to* $(I, J, K, L, \phi)$.

*Remark* 3. With notation as in Definition 2, by construction $(A_1, \dots, A_n)$ pairwise commute and require at most two generating vectors, namely $(1, 0), (0, 1) \in M$. In Section 2, we review how to think of such $(A_1, \dots, A_n)$ as coming from $n$-dimensional partition shapes.

*Remark* 4. The $n$-tuples of commuting $d \times d$ matrices $(A_1, \dots, A_n)$ arising from Definition 2 have the following equivalent formulation. They are those tuples for which there are generating vectors $v_1$ and $v_2$ such that for each $j \in \{1, 2\}$, we have: (i) the annihilator $\mathrm{Ann}_{\mathcal{A}}(v_j) := \{A \in \mathcal{A} \mid Av_j = 0\}$ is generated by monomials $\prod_i A_i^{k_i}$, and (ii) if $\prod_i A_i^{\ell_i} v_j \in \mathcal{A}v_{3-j}$, then there are $m_i$ with $\prod_i A_i^{\ell_i} v_j = \prod_i A_i^{m_i} v_{3-j}$.

**Example 5.** In Example 1, we have $K = (x_1, x_2)$ and $L = (x_3, x_4)$. The isomorphism $\phi\colon K/I \xrightarrow{\simeq} L/J$ is given by $\phi(ax_1 + bx_2) = ax_3 + bx_4$ for $a, b \in k$.

We wish to emphasize that, despite the simplicity of Definition 2, this case is already non-trivial. Indeed, letting $\mathfrak{m} := (x_1, \dots, x_n)$, note that in Example 1, we have $K/I \simeq (S/\mathfrak{m})^{\oplus 2}$. Thus, even if one restricts attention to examples where $K/I$ is as simple as possible, namely $(S/\mathfrak{m})^{\oplus c}$, we already obtain a broad enough class to encompass the standard example of four pairwise commuting $d \times d$ matrices $(A_1, \dots, A_4)$ with $\dim_k \mathcal{A} > d$.

**Definition 6.** Let $S = k[x_1, \dots, x_n]$ and $N$ be an $S$-module. We say the *combinatorial Gerstenhaber problem holds when gluing along* $N$ if for all $(I, J, K, L, \phi)$ as in Definition 2 with $K/I \simeq N$, we have

$$\dim_k \mathcal{A} \leqslant \dim M,$$

where $(A_1, \dots, A_n)$ is associated to $(I, J, K, L, \phi)$ and $\mathcal{A}$ is the algebra generated by $(A_1, \dots, A_n)$.

In this paper, we prove:

**Theorem 7.** *If $S = k[x_1, x_2, x_3]$ and $N = \bigoplus_i S/(x_1, x_2, x_3^{n_i})$, then the combinatorial Gerstenhaber problem holds when gluing along $N$.*

*Remark* 8. Theorem 7 has the following matrix-theoretic description. It proves $\dim_k \mathcal{A} \leqslant d$ for all triples of commuting $d \times d$ matrices $(A_1, A_2, A_3)$ with generating vectors $v_1$ and $v_2$ satisfying conditions (i)–(ii) of Remark 4 and the additional constraint that whenever $w \in \mathcal{A}v_1 \cap \mathcal{A}v_2$, we have $A_1 w = A_2 w = 0$.

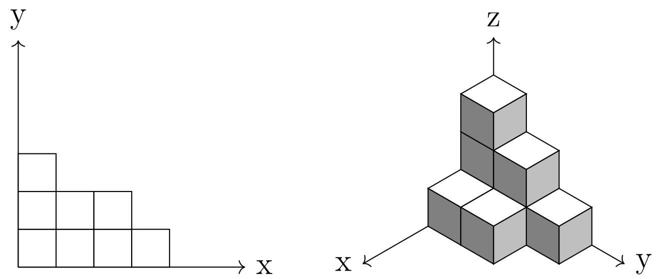
Theorem 7 generalizes [RSS20, Theorem 4], where the result was shown when all $n_i = 1$; in terms of triples of matrices, [RSS20, Theorem 4] handled those $(A_1, A_2, A_3)$ as in Remark 8 but where the condition $A_1 w = A_2 w = 0$ was replaced with $A_1 w = A_2 w = A_3 w = 0$; note Theorem 7 is a marked improvement. The proof given in [RSS20] does not lend itself to generalization and a new idea was required. Our proof of Theorem 7 uses a series of reductions to turn this three-dimensional problem into a two-dimensional one using objects that we call floor plans. We ultimately prove the main theorem by constraining the shape that such floor plans can assume.

*Conventions.* We let $\mathbb{N}$ be the set of non-negative integers.

## 2 Combinatorial Gerstenhaber problem and Young diagrams

In this section, we reduce Theorem 7 to a problem in combinatorics. Much of the material in this section is based on [RSS20, §4.1–4.2].

To begin, the monomials in $S := k[x_1, \ldots, x_n]$ are in bijection with elements of $\mathbb{N}^n$ by identifying $a = (a_1, \ldots, a_n) \in \mathbb{N}^n$ with $x^a := x_1^{a_1} \cdots x_n^{a_n}$. Recall that an *n-dimensional Young diagram* (also known as a *standard set* or *staircase diagram*) is a finite subset $\lambda \subset \mathbb{N}^n$ such that for all $v, w \in \mathbb{N}^n$ with $v \leqslant w$ (in the standard partial order on $\mathbb{N}^n$), if $w \in \lambda$ then $v \in \lambda$. Throughout the paper we fix a convention for drawing Young diagrams. On the left are the axes for 2-dimensional young diagrams, and on the right are the axes for 3-dimensional young diagrams




Given a monomial ideal $I \subset S$ with $\dim_k S/I < \infty$, we obtain an $n$-dimensional Young diagram $\lambda \subset \mathbb{N}^n$ given by the set of $a \in \mathbb{N}^n$ with $x^a \notin I$; moreover, this yields an inclusion-reversing bijection between such monomial ideals and $n$-dimensional Young diagrams, see, e.g., [MS05, Chapter 3] for further details.

Next, if $I$ is as above and $I \subset K$ with $K$ a monomial ideal, then let $\nu \subset \mathbb{N}^n$ be the set of $a \in \mathbb{N}^n$ with $x^a \in K/I$. We see $\nu = \lambda \smallsetminus \lambda'$ where $\lambda$ (resp. $\lambda'$) is the $n$-dimensional Young diagram associated to $I$ (resp. $K$). Sets obtained as the difference of two $n$-dimensional Young diagrams are referred to as *skew shapes*. Let $e_m := (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{N}^n$ with

1 in the $m$-th entry. We define an equivalence relation $\sim_\epsilon$ as follows: given $a, b \in \epsilon$, we write $a \sim_\epsilon b$ if there exist sequences $a_0, \ldots, a_\ell \in \epsilon$ and $j_0, \ldots, j_{\ell-1} \in \{1, \ldots, n\}$ such that $a_0 = a$, $a_\ell = b$, and $a_{i+1} = a_i \pm e_{j_i}$. The equivalence classes are referred to as the *connected components* of $\epsilon$; we say $\epsilon$ is *connected* if it is a single equivalence class.

**Lemma 9.** *Let $S = k[x_1, \ldots, x_n]$. Let $I \subset K \subset S$ and $J \subset L \subset S$ be monomial ideals with $\dim_k S/I < \infty$ and $\dim_k S/J < \infty$. Let $\nu$ (resp. $\epsilon$) be the skew shape associated with $K/I$ (resp. $L/J$). Let $\nu = \nu_1 \amalg \cdots \amalg \nu_m$ and $\epsilon = \epsilon_1 \amalg \cdots \amalg \epsilon_\ell$ be the decompositions into connected components.*

*Every isomorphism $\phi\colon K/I \xrightarrow{\simeq} L/J$ of $S$-modules sending monomials to monomials is given as follows. We have a bijection $\sigma\colon \{1, \ldots, m\} \to \{1, \ldots, \ell\}$ and an element $c_i \in \mathbb{Z}^n$ such that $\nu_i + c_i = \epsilon_{\sigma(i)}$. The map $\phi$ is then given by $\phi(x^a) = x^{a+c_i}$ for all $a \in \nu_i$.*

*Proof.* Since $\phi$ is an $S$-module isomorphism and induces a bijection from monomials of $K/I$ to monomials of $L/J$, we see $\phi(x^a) = x^{\psi(a)}$, where $\psi\colon \nu \to \epsilon$ is an isomorphism of posets.

Thus, we need only to prove that if $a$ and $b$ are in the same connected component of $\nu$, then $\psi(a)$ and $\psi(b)$ are in the same connected component of $\epsilon$. To see this, suppose $a \sim_\nu b$; we must show $\psi(a) \sim_\epsilon \psi(b)$. By definition of the equivalence relation $\sim_\nu$, it suffices to assume $b = a + e_m$. Then

$$x^{\psi(b)} = \phi(x^b) = \phi(x_m x^a) = x_m \phi(x^a) = x_m x^{\psi(a)} = x^{\psi(a)+e_m}$$

in $L/J$. Since $\phi$ is an isomorphism and $b \in \nu$, we see $\phi(x^b) \neq 0$, i.e., both sides of the equation are non-zero. Hence, $\psi(b) = \psi(a) + e_m$, showing $\psi(a) \sim_\epsilon \psi(b)$. □

In light of Lemma 9, we say two skew shapes $\nu, \epsilon \subset \mathbb{N}^n$ are *translationally equivalent* if there exists $w \in \mathbb{Z}^n$ such that $\nu + w = \epsilon$. Note that this defines an equivalence relation on skew shapes. Note further that each skew shape $\nu$ has a unique lex-smallest point $\ell_\nu$. We may therefore normalize our skew shapes by considering $\nu - \ell_\nu$; we refer to this translated skew shape as an *abstract skew shape*. Thus, every skew shape is of the form $\alpha + w$ where $\alpha$ is an abstract skew shape, $w \in \mathbb{Z}^n$, and $\alpha + w \subset \mathbb{N}^n$. We have natural notions of connectedness and connected components for abstract skew shapes.

We can now reinterpret the data in Definition 2 purely combinatorially.

**Corollary 10.** *Let $S = k[x_1, \ldots, x_n]$. Giving $(I, J, K, L, \phi)$ as in Definition 2 is equivalent to giving*

1. *$n$-dimensional Young diagrams $\lambda$ and $\mu$,*

2. *connected abstract skew shapes $\nu_1, \ldots, \nu_r$, and*

3. *lattice points $b_1, \ldots, b_r, c_1, \ldots, c_r \in \mathbb{Z}^n$*

*such that*

(a) $\nu_1 + b_1, \ldots, \nu_r + b_r$ *are disjoint and contained in* $\lambda$,

(b) *for all* $w \in \mathbb{N}^n$ *and all* $v_i \in \nu_i$ *such that* $v_i + b_i + w \in \lambda$, *we have* $v_i + w \in \nu_i$,

(c) $\nu_1 + c_1, \ldots, \nu_r + c_r$ *are disjoint and contained in* $\mu$, *and*

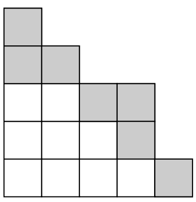
(d) *for all* $w \in \mathbb{N}^n$ *and all* $v_i \in \nu_i$ *such that* $v_i + c_i + w \in \mu$, *we have* $v_i + w \in \nu_i$.

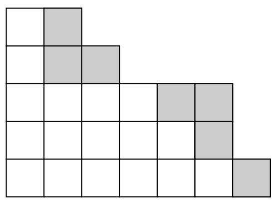
*Moreover, under this equivalence,* $\lambda$ *(resp.* $\mu$*) is the* $n$*-dimensional Young diagram corresponding to* $I$ *(resp.* $J$*), and* $\nu_i + b_i$ *(resp.* $\nu_i + c_i$*) are the connected components of the skew shape associated to* $K/I$ *(resp.* $L/J$*).*

*Proof.* If $(I, J, K, L, \phi)$ are as in Definition 2, let $\lambda$ (resp. $\mu$) be the $n$-dimensional Young diagram corresponding to $I$ (resp. $J$). Let $\beta$ (resp. $\gamma$) be the skew shape corresponding to $K/I$ (resp. $L/J$). By Lemma 9, we may order the connected components $\beta_1, \ldots, \beta_r$ of $\beta$ and $\gamma_1, \ldots, \gamma_r$ of $\gamma$ such that $\beta_i$ and $\gamma_i$ are translationally equivalent. Letting $\nu_i$ be the abstract skew shape associated to $\beta_i$, we then see there are $b_1, \ldots, b_r, c_1, \ldots, c_r \in \mathbb{Z}^n$ such that $\beta_i = \nu_i + b_i$ and $\gamma_i = \nu_i + c_i$. Lastly, note that if $v_i \in \nu_i$, $w \in \mathbb{N}^n$, and $v_i + b_i + w \in \lambda$, then $x^{v_i+b_i+w}$ is not in $I$ hence also not in $K$, i.e., $v_i + b_i + w \notin \nu_i + b_i$, so $v_i + w \notin \nu_i$.

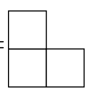
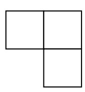
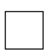
Conversely, given data in (1)–(3) satisfying properties (a)–(d), let $I$ (resp. $J$) be the monomial ideal corresponding to $\lambda$ (resp. $\mu$). Let $K'$ be the set of monomials $x^a$ such that $a \in \nu_i + b_i$ for some $i$; let $K''$ be linear combinations of the monomials in $K'$. Then $K := K'' + I$ is a monomial ideal containing $I$, and property (b) shows that $K'$ forms a monomial basis for $K/I$. Thus, $\coprod_i(\nu_i + b_i)$ is the skew shape corresponding to $K/I$ with connected components $\nu_i + b_i$. Similarly, we obtain a monomial ideal $L \supset J$ whose such that the skew shape corresponding to $L/J$ has connected components $\nu_i + c_i$. The lattice points $c_i - b_i \in \mathbb{Z}^n$ then yield an $S$-module isomorphism $\phi\colon K/I \xrightarrow{\simeq} L/J$ as in Lemma 9. □

**Example 11.** Consider the following example (a variant on [RSS20, Example 4]). We have 2-dimensional Young diagrams (i.e., partitions) $\lambda$ and $\mu$ corresponding respectively to the ideals $I = (x^5, x^4y, x^2y^3, xy^4, y^5)$, on the left, and $J = (x^7, x^6y, x^3y^3, x^2y^4, y^5)$, on the right.

$\lambda =$ 

$\mu =$ 

The grey regions represent the connected skew shapes $\nu_1, \nu_2, \nu_3$. These abstract skew shapes are depicted as follows.

$\nu_1, \nu_2, \nu_3 =$  ,  , 

Explicitly,

$$\begin{aligned} &\nu_1 = \{(0,0),(1,0),(0,1)\}, && \nu_2 = \{(1,0),(0,1),(1,1)\}, && \nu_3 = \{(0,0)\}, \\ &b_1 = (0,3), && b_2 = (2,1), && b_3 = (4,0), \\ &c_1 = (1,3), && c_2 = (4,1), && c_3 = (6,0). \end{aligned}$$

Here $K = (y^3, x^2y^2, x^3y, x^4)$ and $L = (xy^3, x^4y^2, x^5y, x^6)$. $\diamond$

Having now understood Definition 2 combinatorially, the next result gives a combinatorial reinterpretation of the Gerstenhaber problem, cf. [RSS20, Equation (6)].

**Proposition 12.** *Keep the notation of Definition 2 and let $\mathcal{A}$ be the algebra generated by $(A_1, \ldots, A_n)$. Let $\lambda$ (resp. $\mu$) be the $n$-dimensional Young diagram corresponding to $I$ (resp. $J$). Let $\nu$ be the skew shape associated to $K/I$. Then*

$$\dim_k \mathcal{A} \leqslant \dim_k M \iff |\nu| \leqslant |\lambda \cap \mu|.$$

*Proof.* By [RSS20, Proposition 1], $\dim_k \mathcal{A} \leqslant \dim_k M$ if and only if $\dim_k S/\mathrm{Ann}(M) \leqslant \dim_k M$. Lemma 8 and the paragraph afterwards in (loc. cit) shows that

$$\dim_k S/\mathrm{Ann}(M) = |\lambda \cup \mu|.$$

Thus,

$$\begin{aligned} \dim_k M - \dim_k S/\mathrm{Ann}(M) &= \dim_k(M) - |\lambda \cup \mu| \\ &= |\lambda| + |\mu| - |\nu| - |\lambda \cup \mu| = |\lambda \cap \mu| - |\nu|. \end{aligned}$$

This proves the equivalence of $|\nu| \leqslant |\lambda \cap \mu|$ and $\dim_k S/\mathrm{Ann}(M) \leqslant \dim_k M$, hence also the equivalence of these inequalities with $\dim_k \mathcal{A} \leqslant \dim_k M$. $\square$

*Remark* 13. Whether or not the inequality $\dim_k \mathcal{A} \leqslant \dim_k M$ holds is independent of the choice of $\phi$, as Proposition 12 shows. Correspondingly, via Corollary 10, whether or not the inequality holds is independent of the choice of the $b_i$ and $c_i$.

We end this section by specializing to our case of interest.

**Definition 14.** We say $(\lambda, \nu, b)$ is a *tower* if $\lambda$ is a 3-dimensional Young diagram, $\nu = (\nu_1, \ldots, \nu_r)$ where each $\nu_i$ is an abstract skew shape of the form $\{(0,0,z) \mid 0 \leqslant z < n_i, z \in \mathbb{N}\}$, and $b_i \in \mathbb{Z}^n$ satisfies properties (a) and (b) of Corollary 10. We say $(\lambda, \mu, \nu, b, c)$ is a *compatible tower* if $(\lambda, \nu, b)$ and $(\mu, \nu, c)$ are towers. We define $|\nu| := \sum_{i=1}^r |\nu_i|$ and frequently refer to $\nu_i$ as a $1 \times 1 \times n_i$ shape.

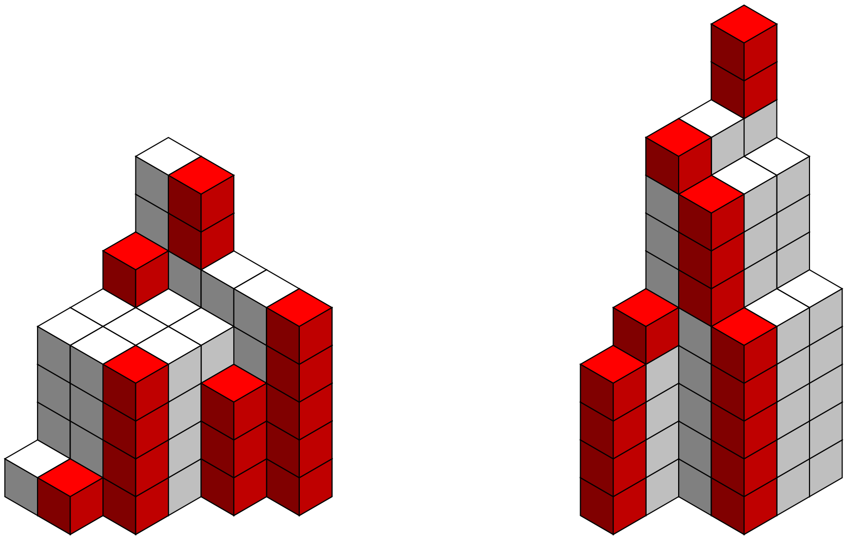
**Example 15.** Throughout this paper, we will consider the following running example. Below, we have a compatible tower $(\lambda, \mu, \nu, b, c)$ where

$$\begin{aligned} b &= ((4,1,0),(1,0,4),(0,1,5),(1,3,0),(2,3,0),(0,4,0)) \\ c &= ((3,0,4),(2,0,8),(0,0,9),(2,1,5),(4,0,0),(2,2,0)) \end{aligned}$$

and $\nu_i$ is $1 \times 1 \times n_i$ with

$$(n_1, \ldots, n_6) = (1, 1, 2, 3, 4, 5)\,;$$

$\lambda$ is on the left and $\mu$ is on the right, with $\nu_i$'s highlighted in red:



By Corollary 10 and Proposition 12, we see Theorem 7 is equivalent to

**Theorem 16.** *If* $(\lambda, \mu, \nu, b, c)$ *is a compatible tower, then*

$$|\nu| \leqslant |\lambda \cap \mu|.$$

The remainder of this paper is devoted to proving Theorem 16.

*Remark* 17. It is often useful to think of a three-dimensional Young diagram $\lambda$ in terms of two-dimensional data. Let $\pi\colon \mathbb{N}^3 \to \mathbb{N}^2$ be the projection onto the first two coordinates. Then specifying $\lambda$ is equivalent giving a function $H_\lambda\colon \mathbb{N}^2 \to \mathbb{N}$ with the property that $H_\lambda(v) \geqslant H_\lambda(w)$ whenever $v \leqslant w$ in the poset partial order. Specifically, this equivalence is given by letting $H_\lambda(v) = |\lambda \cap \pi^{-1}(v)|$. We sometimes refer to $H_\lambda(v)$ as the *height of* $\lambda$ *over* $v$.

## 3 Restricting the shape of counter-examples: scaffolded towers

We introduce the following partial order which will play a crucial role in our paper.

**Definition 18.** We define a partial order on compatible towers as follows: $(\lambda, \mu, \nu, b, c) \leqslant (\lambda', \mu', \nu', b', c')$ if the following hold:

1. $\lambda \subseteq \lambda'$ and $\mu \subseteq \mu'$,

2. letting $\nu = (\nu_i \mid i \in I)$ and $\nu' = (\nu'_i \mid i \in J)$ there exists an injection $\iota\colon I \hookrightarrow J$ such that $\nu_i \subseteq \nu'_{\iota(i)}$,

3. $|\lambda \cap \mu| - |\nu| \leqslant |\lambda' \cap \mu'| - |\nu'|$.

*Remark* 19. By (3) and Proposition 12, if $(\lambda, \mu, \nu, b, c) \leqslant (\lambda', \mu', \nu', b', c')$ and $(\lambda', \mu', \nu', b', c')$ is a counter-example to the Gerstenhaber problem, then $(\lambda, \mu, \nu, b, c)$ is as well.

Using Remark 19, we will reduce the study of potential counter-examples to certain compatible towers, which we call *scaffolded pairs*. We introduce this definition after first recalling the notion of order ideals.

**Definition 20.** For $T \subseteq \mathbb{N}^d$, we define the *order ideal* generated by $T$ to be

$$\langle T \rangle = \{w \in \mathbb{N}^d : w \leqslant u \text{ for some } u \in T\},$$

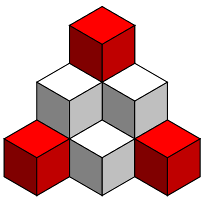
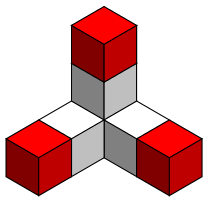
where $\leqslant$ is the entrywise partial order: $w = (w_1, \ldots, w_d) \leqslant u = (u_1, \ldots, u_d)$ if and only if $w_i \leqslant u_i$ for all $i$.

**Definition 21.** We say a tower $(\lambda, \nu, b)$ is *scaffolded* if

$$\lambda = \left\langle \bigcup_{i=1}^{r} (\nu_i + b_i) \right\rangle$$

where $\nu = (\nu_1, \ldots, \nu_r)$. We say $(\lambda, \mu, \nu, b, c)$ is a compatible tower that is *scaffolded* (or a *compatible scaffolded tower*) if both $(\lambda, \nu, b)$ and $(\mu, \nu, c)$ are scaffolded.

**Example 22.** Let $e_i$ denote the $i$-th standard basis vector of $\mathbb{Z}^n$. Let $\nu_1 = \nu_2 = \nu_3 = \{(0, 0, 0)\}$ and $b_i = 2e_i$. Then the tower $(\lambda, \nu, b)$ depicted below on the left is not scaffolded, whereas the tower $(\lambda', \nu, b)$ on the right is scaffolded.

$\lambda =$     $\lambda' =$ 

**Example 23.** Our running example $(\lambda, \mu, \nu, b, c)$ introduced in Example 15 is a compatible scaffolded tower. $\diamond$

The following lemma shows that it suffices to consider compatible scaffolded towers.

**Lemma 24.** *Let $(\lambda, \mu, \nu, b, c)$ be a compatible tower with $\nu = (\nu_1, \ldots, \nu_r)$. Letting*

$$\lambda' = \left\langle \bigcup_{i=1}^{r} (\nu_i + b_i) \right\rangle \quad \text{and} \quad \mu' = \left\langle \bigcup_{i=1}^{r} (\nu_i + c_i) \right\rangle,$$

*we see $(\lambda', \mu', \nu, b, c)$ is scaffolded, and $(\lambda', \mu', \nu, b, c) \leqslant (\lambda, \mu, \nu, b, c)$.*

*Proof.* We see $(\lambda', \mu', \nu, b, c)$ is a scaffolded by definition. Since $\lambda$ is a three-dimensional Young diagram, if $p, q \in \mathbb{N}^3$ and $p \in \lambda$, then $q \in \lambda$. Since $\nu_i + b_i \subset \lambda$, it follows that the order ideal generated by $\nu_i + b_i$ is contained in $\lambda$, and similarly for $\mu$. Thus, $(\lambda', \mu', \nu, b, c) \leqslant (\lambda, \mu, \nu, b, c)$. $\square$

# 4 Reducing to two-dimensional data: floor plans

Due to Remark 19 and Lemma 24, we have reduced Theorem 16 to the case of compatible scaffolded towers. In this section, we further reduce to two-dimensional data. For this, we introduce the following definition.

**Definition 25.** A *floor plan* is a pair $(P, h)$ of sequences $P = (p_1, \dots, p_r)$, $h = (h_1, \dots, h_r)$ with $p_i \in \mathbb{N}^2$ and $h_i \in \mathbb{Z}^+$. A *compatible floor plan* is a triple $(P, Q, h)$ where $(P, h)$ and $(Q, h)$ are floor plans.

We have a function

$$\mathcal{F}\colon \{\text{scaffolded towers}\} \to \{\text{floor plans}\}$$

defined as follows. If $(\lambda, \nu, b)$ is a scaffolded tower with $\nu = (\nu_1, \dots, \nu_r)$ and $\pi\colon \mathbb{N}^3 \to \mathbb{N}^2$ is given by $\pi(x, y, z) = (x, y)$, then $\mathcal{F}(\lambda, \nu, b) = ((p_i)_i, (h_i)_i)$, where $h_i = |\nu_i|$ and $p_i = \pi(b_i)$. Similarly, we have an induced function

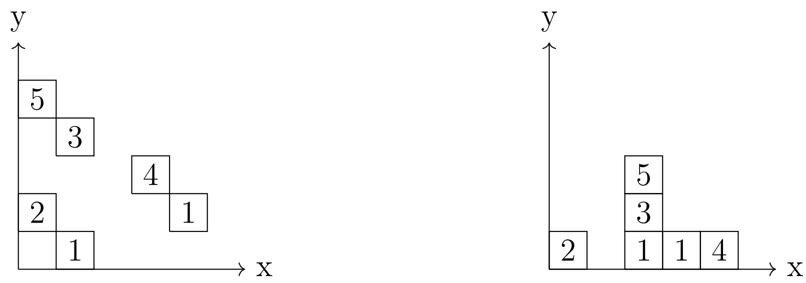
$$\mathcal{F}\colon \{\text{compatible scaffolded towers}\} \to \{\text{compatible floor plans}\}$$

which we also denote by $\mathcal{F}$. The notation $(P, h)$ is chosen since one should think of $P$ as the positions of the $\nu_i$ and $h$ as the associated heights.

**Example 26.** Letting $(\lambda, \mu, \nu, b, c)$ be as in Example 15, we have $(P, Q, h) = \mathcal{F}(\lambda, \mu, \nu, b, c)$ with

$$\begin{aligned} P &= ((4,1), (1,0), (0,1), (1,3), (3,2), (0,4)) \\ Q &= ((3,0), (2,0), (0,0), (2,1), (4,0), (2,2)) \\ h &= (1, 1, 2, 3, 4, 5). \end{aligned}$$

We depict such floor plans below with $(P, h)$ on the left and $(Q, h)$ on the right, where each $p_i \in P$ and $q_i \in Q$ is represented with a box at its place with the numbers in each box being $h_i$




Our next goal is to construct suitable sections of the maps denoted by $\mathcal{F}$. For this, we need some preliminary definitions. A *North-East path* is a sequence $\gamma := (q_1, q_2, \dots, q_m)$ with $q_i \in \mathbb{N}^2$ and $q_{i+1} - q_i \in \{(1,0), (0,1)\}$. Note that we do not require $q_1 = (0,0)$. We say $\gamma$ *originates at* $q_1$.

**Definition 27.** Given a floor plan $(P,h)$ with $P=(p_1,\dots,p_r)$ and $h=(h_1,\dots,h_r)$, the *score* of a North-East path $\gamma=(q_1,\dots,q_m)$ is

$$\mathrm{score}_{(P,h)}(\gamma) := \sum_{p_i\in\{q_1,\dots,q_m\}} h_i.$$

The *max score* of a lattice point $q\in\mathbb{N}^2$ is

$$\max\mathrm{score}_{(P,h)}(q) := \max\{\mathrm{score}_{(P,h)}(\gamma) \mid \gamma \text{ originates at } q\}.$$

Any North-East path $\gamma$ originating at $q$ for which $\max\mathrm{score}_{(P,h)}(q)=\mathrm{score}(\gamma)$ is referred to as a *winning path* for $q$. When $(P,h)$ is understood from context, we suppress it in the notation for score and max score.

**Definition 28.** Given a floor plan $(P,h)$, its *minimal realization* $\mathcal{T}(P,h)$ is the tower $(\lambda,\nu,b)$ defined as follows. Let $\nu_i$ be a $1\times1\times h_i$ shape and let $\lambda$ be the three-dimensional Young diagram whose corresponding function $H_\lambda\colon\mathbb{N}^2\to\mathbb{N}$ is given by $H_\lambda(q):=\max\mathrm{score}_{(P,h)}(q)$, see Remark 17. Letting $P=(p_i)_i$, we let $b_i$ be the lattice point whose $(x,y)$-coordinates are given by $p_i$ and whose $z$-coordinate is $H_\lambda(p_i)-h_i$, i.e., $b_i=p_i+(0,0,H_\lambda(p_i)-h_i)$.

Similarly, given a compatible floor plan $(P,Q,h)$, its *minimal realization* $\mathcal{T}(P,Q,h)$ is $(\lambda,\mu,\nu,b,c)$ where $(\lambda,\nu,b)=\mathcal{T}(P,h)$ and $(\mu,\nu,c)=\mathcal{T}(Q,h)$.

**Example 29.** Let $(P,Q,h)$ be the compatible floor plan given in Example 26. Its minimal realization $\mathcal{T}(P,Q,h)$ is precisely given by our running example $(\lambda,\mu,\nu,b,c)$ from Example 15. $\diamond$

**Proposition 30.** *The minimal realizations $\mathcal{T}$ yield sections of the two maps denoted by $\mathcal{F}$. Furthermore, if $(\lambda,\mu,\nu,b,c)$ is scaffolded, then*

$$\mathcal{T}(\mathcal{F}(\lambda,\mu,\nu,b,c)) \leqslant (\lambda,\mu,\nu,b,c)$$

*with respect to the partial order in Definition 18.*

*Remark* 31. In fact, the proof of Proposition 30 below shows the stronger statement that if $(\lambda,\nu,b)$ is a scaffolded tower, then $(\lambda',\nu,b')=\mathcal{T}(\mathcal{F}(\lambda,\nu,b))$ satisfies $\lambda'\subseteq\lambda$.

*Proof of Proposition 30.* Let $(P,h)$ be a floor plan with $P=(p_i)_i$. We first show that $(\lambda,\nu,b):=\mathcal{T}(P,h)$ is scaffolded. Let $q\in\mathbb{N}^2$ with $m:=\max\mathrm{score}_{(P,h)}(q)>0$. We must show $q+(0,0,m)$ is in the order ideal generated by the $(0,0,h_i)+b_i$.

For this, let $\gamma=(q_1,q_2,\dots,q_s)$ be a winning path originating at $q_1:=q$. Since $m>0$, there exists $q_\ell\in P$. Without loss of generality, $q_\ell=p_1$ and $q_i\notin P$ for $i<\ell$. Then $(q_\ell,q_{\ell+1},\dots,q_s)$ is a winning path originating at $q_\ell$; indeed, if $\gamma'$ is a path originating at $q_\ell$ with a strictly larger score, then $(q_1,\dots,q_{\ell-1})$ concatenated with $\gamma'$ would yield a strictly larger score than $\gamma$. Thus, $q+(0,0,m)$ is in the order ideal generated by $q_\ell+(0,0,m)$. Now $q_\ell=p_1$ and $m=H_\lambda(p_1)$, so $q_\ell+(0,0,m)=b_1+(0,0,h_1)$. We have therefore shown $\mathcal{T}(P,h)$ is scaffolded.

Next, let $(\lambda, \nu, b)$ be a scaffolded tower, $(P, h) := \mathcal{F}(\lambda, \nu, b)$ the associated floor plan, and $(\lambda', \nu, b') = \mathcal{T}(P, h)$. Note that $\nu$ remains unchanged under the operation $\mathcal{T} \circ \mathcal{F}$. Thus, to finish the proof, it suffices to show $\lambda' \subseteq \lambda$, i.e., for all $q \in \mathbb{N}^2$ and any North-East path $\gamma$ originating at $q$, $H_\lambda(q) \geqslant \text{score}_{(P,h)}(\gamma)$. Let $\gamma = (q_1, \ldots, q_s)$ and assume without loss of generality that $q_{i_j} = p_j$ for $i_1 < \cdots < i_\ell$, and for $t \neq i_j$ we have $q_t \notin P$. We know $\nu_j + b_j \subset \lambda$ and by property (b) of Corollary 10, $w + b_j \notin \lambda$ for all $w \in \mathbb{N}^3$ with $w \notin \nu_j$. Therefore,

$$H_\lambda(q_{i_j}) \geqslant |\nu_j| + H_\lambda(q_{i_j+1}) = h_j + H_\lambda(q_{1+i_j}) \geqslant h_j + H_\lambda(q_{i_{j+1}}).$$

Thus,

$$H_\lambda(q) \geqslant H_\lambda(q_{i_1}) \geqslant h_1 + H_\lambda(q_{i_2}) \geqslant h_1 + h_2 + H_\lambda(q_{i_3}) \geqslant \ldots \geqslant \sum_j h_j = \text{score}_{(P,h)}(\gamma),$$

as desired. □

## 5 Constraints on the border of a floor plan

In this section, we further reduce Theorem 16 to the study of floor plans whose borders are highly constrained. Throughout this section, we use the following notation. Let $e_1, e_2, e_3$ be the standard basis vectors of $\mathbb{Z}^3$. For any $p \in \mathbb{N}^3$, we let $x(p)$, $y(p)$, and $z(p)$ denote the $x$, $y$, and $z$ coordinates of $p$. If $(P, h)$ is a floor plan with $P = (p_i)_i$, we will sometimes write $p \in P$ to mean $p = p_i$ for some $i$.

**Definition 32.** Let $(P, h)$ and $(P', h)$ be floor plans[1] and let $(\lambda, \nu, b) = \mathcal{T}(P, h)$ and $(\lambda', \nu, b') = \mathcal{T}(P', h)$. We write

$$(P', h) \leqslant (P, h) \quad \text{if} \quad \lambda' \subset \lambda.^2$$

Similarly, given compatible floor plans $(P, Q, h)$ and $(P', Q', h)$, we write

$$(P, Q, h) \leqslant (P', Q', h) \quad \text{if} \quad \mathcal{T}(P, Q, h) \leqslant \mathcal{T}(P', Q', h)$$

with respect to the partial ordering in Definition 18. We say $(P, h)$, respectively $(P, Q, h)$, is *minimal* if it is minimal with respect to these partial orders.

**Definition 33.** Let $(P, h)$ be a floor plan. We define the *support* of $(P, h)$ to be

$$\text{supp}(P, h) := \{q \in \mathbb{N}^2 \mid \max \text{score}_{(P,h)}(q) > 0\}.$$

The *border* $B(P, h)$ is then defined as all $q \in \text{supp}(P, h)$ such that $q + e_1$ or $q + e_2$ is not in $\text{supp}(P, h)$.

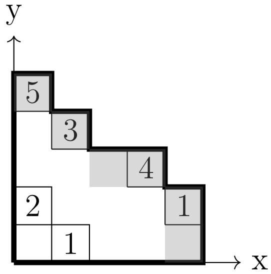
[1] Note that the second coordinates of these floor plans are the same.

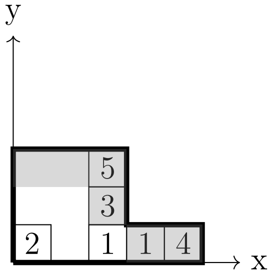
[2] This is equivalent to $\mathcal{T}(P', h) \leqslant \mathcal{T}(P, h)$ with respect to the partial ordering in Definition 18.

*Remark* 34. Note that that $\mathrm{supp}(P,h) = \langle P \rangle$, which we will make use of in Proposition 41.

*Remark* 35. Note that $\mathrm{supp}(P,h)$ is the projection of $\mathcal{T}(P,h)$ onto the $xy$-plane.

**Example 36.** The support and border in Example 26 are illustrated below for each floor plan. The support is the region enclosed by a bold line and the border is indicated by shaded grey boxes.




Our first goal is to prove:

**Proposition 37.** *Let $(P,h)$ be a minimal floor plan. Then*

$$B(P,h) \subseteq P.$$

*In particular, if $(P,Q,h)$ is a minimal compatible floor plan, then $B(P,h) \subseteq P$ and $B(Q,h) \subseteq Q$.*

We prove this proposition after a preliminary result.

**Lemma 38.** *Let $(P,h)$ be a floor plan. Suppose there exists $i$ such that*

1. *$x(p_i) > 0$ and*
2. *for all $j$ with $x(p_j) = x(p_i) - 1$, we have $y(p_j) < y(p_i)$.*

*Then letting $p'_i = p_i - e_1$ and $p'_k = p_k$ for all $k \neq i$, we have $(P',h) < (P,h)$ where $P' = (p'_j)_j$.*

*Proof.* Note that $\max\mathrm{score}_{(P',h)}(p_i) = \max\mathrm{score}_{(P,h)}(p_i) - h_i < \max\mathrm{score}_{(P,h)}(p_i)$ and that for all $p$ with $p \not\leqslant p_i$, we have $\max\mathrm{score}_{(P',h)}(p) = \max\mathrm{score}_{(P,h)}(p)$. Further note that $\max\mathrm{score}_{(P',h)}(p) \leqslant \max\mathrm{score}_{(P,h)}(p)$ if $p < p_i$ with $x(p) = x(p_i)$; indeed, letting $\gamma$ be a North-East path originating at $p$, if $\gamma$ contains $p_i$ then $\mathrm{score}_{(P',h)}(\gamma) = \mathrm{score}_{(P,h)}(\gamma) - h_i$, and if $\gamma$ does not contain $p_i$ then $\mathrm{score}_{(P',h)}(\gamma) = \mathrm{score}_{(P,h)}(\gamma)$.

Thus, to prove $(P',h) < (P,h)$, it suffices to show $\max\mathrm{score}_{(P',h)}(v) \leqslant \max\mathrm{score}_{(P,h)}(v)$ for all $v \leqslant p_i - e_1$. For this, consider a North-East path $\gamma$ which contains $p'_i$. Let $\gamma = (q_1, \ldots, q_s)$ with $q_\ell = p'_i$. Say $x(q_j) = x(p'_i)$ for $\ell \leqslant j \leqslant m$ and that $x(q_{m+1}) > x(p'_i)$; this implies $q_{m+1} = q_m + e_1$. Let

$$\gamma' = (q_1, \ldots, q_\ell, q_\ell + e_1, \ldots, q_m + e_1, q_{m+2}, \ldots, q_s).$$

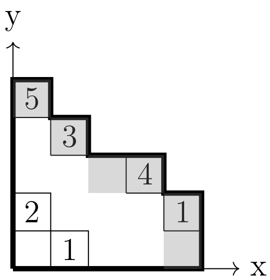
By hypothesis, $q_j \notin P$ for $\ell \leqslant j \leqslant m$, so

$$\mathrm{score}_{(P',h)}(\gamma) \leqslant \mathrm{score}_{(P',h)}(\gamma') = \mathrm{score}_{(P,h)}(\gamma')$$

which proves the result. $\square$

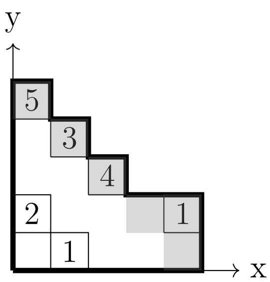
*Proof of Proposition 37.* Let $v \in B(P,h) \setminus P$. Note that we cannot have $v+e_1 \notin \mathrm{supp}(P)$ and $v+e_2 \notin \mathrm{supp}(P)$ since this implies then $\max \mathrm{score}_{(P,h)}(v) = \mathrm{score}_{(P,h)}(\gamma) = 0$, where $\gamma$ is the singleton path $(v)$. Thus, without loss of generality, $v + e_1 \in \mathrm{supp}(P)$ and $v + e_2 \notin \mathrm{supp}(P)$. As a result, $\max \mathrm{score}_{(P,h)}(v)$ is the sum of the $h_i$ for all $i$ with $y(p_i) = y(v)$. Since this quantity is non-zero, we may assume without loss of generality that $y(v) = y(p_1)$, $x(v) < x(p_1)$, and there are no $p \in P$ such that $y(v) = y(p)$ and $x(v) < x(p) < x(p_1)$. Let $P' = (p'_i)_i$ where $p'_1 = p_1 - e_1$ and $p'_i = p_i$ for $i \neq 1$. Then by Lemma 38, $(P', h) < (P, h)$, showing that $(P, h)$ is not minimal. $\square$

**Example 39.** Here is a demonstration of the proof of Proposition 37. We use $(P, h)$ from Example 36, which is



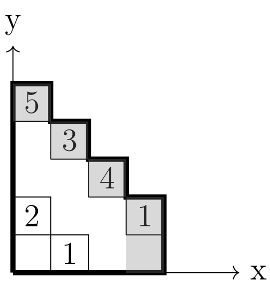

$$P = ((4,1),(1,0),(0,1),(1,3),(3,2),(0,4)),$$
$$h = (1, \quad 1, \quad 2, \quad 3, \quad 4, \quad 5).$$

We observe that $p_4 = (3, 2)$ satisfies the premise of Lemma 38, as $x(p_4) = 3 > 0$ and there is no $j$ such that $x(p_j) = x(p_4) - 1$. Therefore, we let $p'_4 = p_4 - e_1 = (2, 2)$ to update our floor plan to be



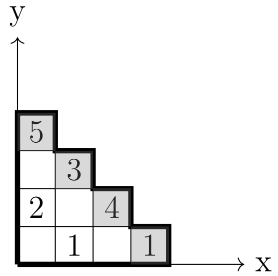

$$P = ((4,1),(1,0),(0,1),(1,3),(2,2),(0,4)),$$
$$h = (1, \quad 1, \quad 2, \quad 3, \quad 4, \quad 5).$$

For simplicity, we will call the *updated* floor plan $(P, h)$ as before. Note that the border $B(P, h)$ is still not contained in $P$ and that $p_1 = (4, 1)$ is such that there is no $x(p_j) = x(p_1) - 1$. Hence we let $p'_1 = p_1 - e_1 = (3, 1)$.



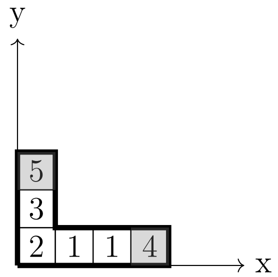

$$P = ((3,1),(1,0),(0,1),(1,3),(2,2),(0,3)),$$
$$h = (1, \quad 1, \quad 2, \quad 3, \quad 4, \quad 5).$$

By symmetry, we apply Lemma 38 in the $y$-direction, which gives us




$$P = ((2,1),(1,0),(0,1),(0,3),(1,2),(0,3)),$$
$$h = (\,1, \quad 1, \quad 2, \quad 3, \quad 4, \quad 5).$$

Applying the same algorithm on $(Q, h)$ gives




$$Q = ((2,0),(1,0),(0,0),(0,1),(3,0),(0,2)),$$
$$h = (\,1, \quad 1, \quad 2, \quad 3, \quad 4, \quad 5).$$

Proposition 37 gives a structure result for minimal floor plans. Our next goal in this section is to prove a further structure result for minimal *compatible* floor plans.

Recall the notation for $H_\epsilon(v)$ for the height of a three-dimensional Young diagram $\epsilon$ over a point $v \in \mathbb{N}^2$ in Remark 17. We observe that for three-dimensional Young diagrams $\lambda$ and $\mu$ we have

$$|\lambda \cap \mu| \; = \; \sum_{v \in \mathbb{N}^2} H_{\lambda\cap\mu}(v).$$

If we also have compatible floor plans $(P, Q, h)$ such that $(\lambda, \mu, \nu, b, c) = \mathcal{T}(P, Q, h)$, then we get

$$H_{\lambda\cap\mu}(v) = \min(\max \mathrm{score}_{(P,h)}(v), \max \mathrm{score}_{(Q,h)}(v)).$$

**Lemma 40.** *If $(P, Q, h)$ is a minimal compatible floor plan, then $P \cap Q = \varnothing$.*

*Proof.* Let $P = (p_i)_i$ and $Q = (q_i)_i$. Let $(\lambda, \mu, \nu, b, c) = \mathcal{T}(P, Q, h)$. We show that if $P \cap Q$ is non-empty, then $(\lambda, \mu, \nu, b, c)$ is not minimal. Without loss of generality, $w := p_i = q_j$ and $\max \mathrm{score}_{(P,h)}(p_i) \geqslant \max \mathrm{score}_{(Q,h)}(q_j)$. Let $h'_j = h_j - 1$ and $h'_m = h_m$ for all $m \neq j$. Let $(\lambda', \mu', \nu', b', c') = \mathcal{T}(P, Q, h')$. Then we see

$$\max \mathrm{score}_{(P,h')}(p) \leqslant \max \mathrm{score}_{(P,h)}(p) \quad \text{and} \quad \max \mathrm{score}_{(Q,h')}(q) \leqslant \max \mathrm{score}_{(Q,h)}(q)$$

for all $p \in P$ and $q \in Q$. Furthermore,

$$\max \mathrm{score}_{(Q,h')}(q_j) = \max \mathrm{score}_{(Q,h)}(q_j) - 1 \text{ and } \max \mathrm{score}_{(P,h')}(p_i) \geqslant \max \mathrm{score}_{(P,h)}(p_i) - 1.$$

Therefore

$$\begin{aligned} H_{\lambda'\cap\mu'}(w) &= \min(\max \mathrm{score}_{(P,h')}(p_i), \max \mathrm{score}_{(Q,h')}(q_j)) \\ &= \max \mathrm{score}_{(Q,h)}(q_j) - 1 = H_{\lambda\cap\mu}(w) - 1 \end{aligned}$$

and so

$$|\lambda' \cap \mu'| = \sum_{v \in \mathbb{N}^2} H_{\lambda' \cap \mu'}(v) \leqslant \sum_{v \in \mathbb{N}^2} H_{\lambda \cap \mu}(v) - 1 = |\lambda \cap \mu| - 1.$$

Since $\lambda' \subset \lambda$ and $\mu' \subsetneq \mu$, we see $(\lambda', \mu', \nu', b', c') < (\lambda, \mu, \nu, b, c)$. □

**Proposition 41.** *If $(P, Q, h)$ is a minimal compatible floor plan, then*

$$\text{supp}(P) \subset \text{supp}(Q) \quad \textit{or} \quad \text{supp}(Q) \subset \text{supp}(P).$$

*Proof.* Let $M(P), M(Q)$ denote the sets of maximal elements of $P, Q$, respectively.

We first note that, given any subset $S \subseteq \mathbb{N}^2$, the order ideal $\langle S \rangle$ generated by $S$ is completely determined by the maximal elements of $S$: that is, if $M(S)$ is the set of maximal elements of $S$, then $\langle S \rangle = \langle M(S) \rangle$. Also note that $\text{supp}(P) = \langle P \rangle, \text{supp}(Q) = \langle Q \rangle$ by definition, so that one has

$$\text{supp}(P) = \langle M(P) \rangle, \qquad \text{supp}(Q) = \langle M(Q) \rangle.$$

By maximality of elements of $M(P)$, every two distinct elements of $M(P)$ are not comparable, so we may order $M(P)$ as

$$M(P) = \{p^k\}_{k=1}^K$$

such that $x(p^k) < x(p^{k'}), y(p^k) > y(p^{k'})$ for each $k < k'$. We use superscript notation so as not to conflict with our running subscript notation $p_i$. We can similarly order $M(Q) = \{q^l\}_{l=1}^L$.

Now observe that the ordering on $M(P), M(Q)$ is chosen in a way to allow a convenient expression of $B(P), B(Q)$. That is, we have

$$B(P) = \bigcup_{l=1}^{L} \left( \{x(p^{l-1}) + 1, \ldots, x(p^l)\} \times \{y(p^l)\} \right) \cup \bigcup_{l=1}^{L} \left( \{x(p^l)\} \times \{y(p^l), \ldots, y(p^{l+1}) + 1\} \right), \tag{5.1}$$

where we simply define $x(p^0) = 0 = y(p^{L+1})$. One has similar formula for $B(Q)$.

We are going to show that $M(P) \subseteq Q$ or $M(Q) \subseteq P$, thereby proving the result. To that end, suppose by symmetry $y(p^1) \leqslant y(q^1)$, and we claim that every $p^k \in Q$. We induct on $k$.

The base case $k = 1$ is given as follows. Let $l_1$ be the maximum index such that $y(p^1) \leqslant y(q^{l_1})$. Such an index exists since $l_1 = 1$ satisfies the inequality. Suppose for contradiction $x(p^1) > x(q^{l_1})$. Then note (5.1) provides that $(x(q^{l_1}), y(p^1)) \in B(P) \cap B(Q)$, which is a contradiction to Proposition 37 and Proposition 40.

The inductive step is essentially the same. Suppose the inductive argument is true for $k - 1$, where $k > 1$. Then note that we can guarantee the existence of maximum index $l_k$ such that $y(p^k) \leqslant y(q^{l_k})$. If we assume $x(p^k) > x(q^{l_k})$, then $(x(q^{l_k}), y(p^k)) \in B(P) \cap B(Q)$, so we run into a contradiction.

Hence every $p^k$ have some $q^{l_k}$ with $p^k \leqslant q^{l_k}$, meaning that

$$\text{supp}(P) = \langle M(P) \rangle \subseteq \langle M(Q) \rangle = \text{supp}(Q).$$

Now we finally invoke Lemma 40 again to conclude that the above inclusion is strict, since otherwise $M(P) = M(Q)$ which implies $P \cap Q \neq \emptyset$. □

## 6 Proof of Theorem 7

We turn now to proving Theorem 7. As shown in §2, this is equivalent to proving Theorem 16.

*Proof of Theorem 16.* Assume there exists a counter-example to the theorem. By Remark 19, Lemma 24, and Proposition 30, we may assume $(\lambda,\mu,\nu,b,c) = \mathcal{T}(P,Q,h)$ with $(P,Q,h)$ a minimal compatible floor plan. By Proposition 41, we may assume without loss of generality that $\operatorname{supp}(P) \subset \operatorname{supp}(Q)$.

Let $P = (p_1,\dots,p_r)$ and $Q = (q_1,\dots,q_r)$. Reindexing, we may assume there exists $N_1$ such that $h_i = 1$ and $q_i \in \operatorname{supp}(Q)$ is maximal (in the partial ordering on $\mathbb{N}^2$) if and only if $i > N_1$. We may further assume there exists $N_0 \leqslant N_1$ such that $q_i \in \operatorname{supp}(Q)$ is maximal if and only if $i > N_0$. Let

$$P' = (p_i \mid i \leqslant N_1),\ Q' = (q_i \mid i \leqslant N_1), \text{ and } h' = (h_1,\dots,h_{N_0}, h_{N_0+1}-1,\dots,h_{N_1}-1);$$

in other words, $h'$ decreases the value of $h$ at all maximal elements of $\operatorname{supp}(Q)$ and deletes any indices which now have value 0. Let $(\lambda',\mu',\nu',b',c') = \mathcal{T}(P',Q',h')$. We claim $(\lambda',\mu',\nu',b',c') < (\lambda,\mu,\nu,b,c)$, contradicting minimality of $(\lambda,\mu,\nu,b,c)$.

To see this, first note that

$$\max \operatorname{score}_{(P',h')}(p) \leqslant \max \operatorname{score}_{(P,h)}(p)$$

for all $p \in \operatorname{supp}(P)$ and that

$$\max \operatorname{score}_{(P',h')}(p_i) \leqslant \max \operatorname{score}_{(P,h)}(p_i) - 1$$

for all $N_0 < i \leqslant N_1$. Thus, $\lambda' \subseteq \lambda$. Furthermore, $B(Q,h) \subset Q$ by Proposition 37; it follows that for all $q \in \operatorname{supp}(Q)$, any winning path $\gamma$ for $(Q,h)$ originating at $q$ must contain a (necessarily unique) maximal element of $\operatorname{supp}(Q)$. Since the value of $h'$ is one less than the value of $h$ at all maximal elements of $\operatorname{supp}(Q)$, we see

$$\max \operatorname{score}_{(Q',h')}(q) \leqslant \max \operatorname{score}_{(Q,h)}(q) - 1$$

for all $q \in \operatorname{supp}(Q)$. In particular, $\mu' \subsetneq \mu$.

To show $(\lambda',\mu',\nu',b',c') < (\lambda,\mu,\nu,b,c)$, it therefore remains to prove

$$|\lambda' \cap \mu'| - \sum_j |\nu'_j| \leqslant |\lambda \cap \mu| - \sum_i |\nu_i|. \tag{6.2}$$

For $i > N_0$, we see

$$\begin{aligned} H_{\lambda'\cap\mu'}(p_i) &= \min(\max \operatorname{score}_{(P',h')}(p_i), \max \operatorname{score}_{(Q',h')}(p_i)) \\ &\leqslant \min(\max \operatorname{score}_{(P,h)}(p_i), \max \operatorname{score}_{(Q,h)}(p_i)) - 1 = H_{\lambda\cap\mu}(p_i) - 1. \end{aligned}$$

Thus,

$$|\lambda' \cap \mu'| \ = \ \sum_{v\in\mathbb{N}^2} H_{\lambda'\cap\mu'}(v) \ \leqslant \ \sum_{v\in\mathbb{N}^2} H_{\lambda\cap\mu}(v) - (r - N_0) \ \leqslant \ |\lambda\cap\mu| - (r-N_0).$$

Observing that $\sum_i |\nu_i| - \sum_j |\nu'_j| = r - N_0$, we see (6.2) holds. ☐

**Example 42.** Lastly, we compute the basic matrix invariants for the triple $(A_1, A_2, A_3)$ corresponding to our running example (Example 15). Several results to date on Gerstenhaber's problem are proved using matrix-theoretic properties. For example, in [Š12, NS99] the Gerstenhaber problem is proved whenever one of the matrices $A_i$ has nullity at most 3, and in [HO01] it is solved if some $A_i$ has index at most 2; recall the index of a matrix $A$ is the minimum positive integer $k$ such that $A^k = 0$.

Our matrices $(A_1, A_2, A_3)$ from Example 15 have size $132 \times 132$ since $|\lambda| + |\mu| - |\nu| = 71 + 77 - 16 = 132$.

One can easily read off the indices of our matrices from the combinatorial description. Recall that $A_1, A_2, A_3$ correspond respectively to multiplication by $x, y, z$. In terms of our diagram pictured in Example 15, this corresponds to taking a box and shifting in the $x$, $y$, or $z$ directions. Thus, the index of $A_3$ is the maximum of the $z$-height of $\lambda$ and the $z$-height of $\mu$, i.e., the index of $A_3$ is $11 = \max(7, 11)$. Similarly, the index of $A_1$ is the maximum of the $x$-lengths of $\lambda$ and $\mu$, which is $5 = \max(5, 5)$. Similarly, we compute the index of $A_2$ to be $5 = \max(5, 3)$.

One can also easily read off the nullity of our matrices from the combinatorial description. The kernel of $A_3$ corresponds to the boxes at on top layers of $\lambda$ and $\mu$ after taking into that red boxes are glued. Note that there are 17 boxes on the top layer of $\lambda$, 11 on the top layer of $\mu$, and 6 of these are glued, so we see the nullity of $A_3$ is $17+11-6 = 22$. Similarly, we compute that the nullities of $A_1$ and $A_2$ are 37 and 43. Note that the number of boxes on the top layer of $\lambda$ is simply the number of boxes in the projection of $\lambda$ to the $xy$-plane. Thus, the nullity of $A_3$ can be computed as the number of boxes in the projection of $\lambda$ plus the number of boxes in the projection of $\mu$ minus the number of total $\nu$ shapes. The nullity of $A_1$ (resp. $A_2$) is similarly computed as the number of boxes in the projection of $\lambda$ to the $yz$ (resp. $xz$) plane plus the number of boxes in the projection of $\mu$ minus $|\nu|$.

Similarly, it is straightforward to compute the Weyr form of the matrices. Recall that the Weyr form [OCV11] of a matrix $A$ is given by the tuple[3] whose $i$-th entry is the dimension of the annihilator of $A^i$ minus the dimension of the annihilator of $A^{i-1}$. For example, the first entry of the tuple is the nullity of $A$. It is easy to check that if $\lambda_i$ (resp. $\mu_i$) denotes the boxes in $\lambda$ (resp. $\mu$) whose $x$-coordinate is $i$, then the Weyr form of $A_1$ is given by

$$(|\lambda_0| + |\mu_0| - |\nu|, |\lambda_1| + |\mu_1|, |\lambda_2| + |\mu_2|, \dots).$$

In our running example, the Weyr form for $A_1$ is $(37, 38, 34, 17, 6)$.

We see from these combinatorial descriptions that Theorem 7 applies to triples of matrices with arbitrarily large index and nullity, and one has great flexibility in choosing the Weyr form of $A_1$. ⋄

## Acknowledgements

This paper is the outcome of an NSERC-USRA project; we thank NSERC for their support. This project benefited greatly from conversations with Oliver Pechenik and

[3] authors sometimes take the associated ordered partition

Jenna Rajchgot; it is a pleasure to thank them. The second author would especially like to thank his son, Quinn, for letting him use his toy blocks, which helped with three-dimensional visualization. Lastly, we would like to thank the referees for their many suggestions which improved the quality of this paper.

## References

[Ber13] George M. Bergman, *Commuting matrices, and modules over artinian local rings*, arXiv:1309.0053, 2013.

[BH90] José Barría and P. R. Halmos, *Vector bases for two commuting matrices*, Linear and Multilinear Algebra **27** (1990), no. 3, 147–157. MR 1064891

[Ger61] Murray Gerstenhaber, *On dominance and varieties of commuting matrices*, Ann. of Math. (2) **73** (1961), 324–348. MR 0132079

[GS00] Robert M. Guralnick and B. A. Sethuraman, *Commuting pairs and triples of matrices and related varieties*, Linear Algebra Appl. **310** (2000), no. 1-3, 139–148. MR 1753173

[Gur92] Robert M. Guralnick, *A note on commuting pairs of matrices*, Linear and Multilinear Algebra **31** (1992), no. 1-4, 71–75. MR 1199042

[HO01] John Holbrook and Matjaž Omladič, *Approximating commuting operators*, Linear Algebra Appl. **327** (2001), no. 1-3, 131–149. MR 1823346

[HO15] J. Holbrook and K. C. O'Meara, *Some thoughts on Gerstenhaber's theorem*, Linear Algebra Appl. **466** (2015), 267–295. MR 3278252

[JŠ22] Joachim Jelisiejew and Klemen Šivic, *Components and singularities of Quot schemes and varieties of commuting matrices*, J. Reine Angew. Math. **788** (2022), 129–187. MR 4445543

[LL91] Thomas J. Laffey and Susan Lazarus, *Two-generated commutative matrix subalgebras*, Linear Algebra Appl. **147** (1991), 249–273. MR 1088666

[MS05] Ezra Miller and Bernd Sturmfels, *Combinatorial commutative algebra*, Graduate Texts in Mathematics, vol. 227, Springer-Verlag, New York, 2005. MR 2110098 (2006d:13001)

[MT55] T. S. Motzkin and Olga Taussky, *Pairs of matrices with property L. II*, Trans. Amer. Math. Soc. **80** (1955), 387–401. MR 0086781

[NS99] Michael G. Neubauer and B. A. Sethuraman, *Commuting pairs in the centralizers of* 2*-regular matrices*, J. Algebra **214** (1999), no. 1, 174–181. MR 1684884

[NŠ14] Nham V. Ngo and Klemen Šivic, *On varieties of commuting nilpotent matrices*, Linear Algebra Appl. **452** (2014), 237–262. MR 3201099

[OCV11] Kevin C. O'Meara, John Clark, and Charles I. Vinsonhaler, *Advanced topics in linear algebra*, Oxford University Press, Oxford, 2011, Weaving matrix problems through the Weyr form. MR 2849857

[O'M20] Kevin C. O'Meara, *The Gerstenhaber problem for commuting triples of matrices is "decidable"*, Comm. Algebra **48** (2020), no. 2, 453–466. MR 4068883

[RS18] Jenna Rajchgot and Matthew Satriano, *New classes of examples satisfying the three matrix analog of Gerstenhaber's theorem*, J. Algebra **516** (2018), 245–270. MR 3863478

[RSS20] Jenna Rajchgot, Matthew Satriano, and Wanchun Shen, *Some combinatorial cases of the three matrix analog of Gerstenhaber's theorem (research)*, Advances in mathematical sciences, Assoc. Women Math. Ser., vol. 21, Springer, Cham, 2020, pp. 181–201. MR 4143269

[Set11] B. A. Sethuraman, *The algebra generated by three commuting matrices*, Math. Newsl. **21** (2011), no. 2, 62–67. MR 3013206

[Š12] Klemen Šivic, *On varieties of commuting triples II*, Linear Algebra Appl. **437** (2012), no. 2, 461–489. MR 2921711