Berg Polytopes, Moment Maps, Transportation Fibers, and Contractions

Nikolai Mnëv*

23 September 2026

AI-generated† working report

Abstract

We begin from a simple observation about the left–right coordinate-torus action on matrix space: after projectivization, the moment map sends the squared matrix entries to their row and column marginals, and its fibers in the ambient coordinate simplex are transportation polytopes. A fixed positive diagonal-scaling orbit meets every admissible interior transportation fiber in one point; this is the matrix-scaling/Birch picture. We then compare this with the coordinate-torus action on a Grassmannian. The Grassmannian moment polytope is a matroid base polytope, while Berg's determinantal formula expresses the Moore–Penrose inverse as a convex combination of elementary basis inverses with the same squared-minor coefficients. This suggests the Berg polytope, the convex hull of elementary right inverses, together with its affine projection to the matroid base polytope by $X \mapsto \mathrm{diag}(XA)$. We determine its vertices, show that the matroid face lattice lifts canonically to the Berg polytope, identify rank-one basis-exchange edges, and give affine-dimension and interior-fiber formulas; for a connected matroid the interior fibers have the dimension $(r-1)(n-r-1)$ of the torus quotient of the Grassmannian, and their tangent spaces are dual to the corresponding deformation spaces. The fibers are linear images of "matroid transportation polytopes", and Birch's theorem produces a canonical entropy-minimizing section; its closed form is the classical weighted pseudoinverse, and its unweighted point is the Moore–Penrose inverse. We then extend

*PDMI RAS

†Claude Opus, GPT Sol, Lemma

the construction to finite based acyclic complexes. That the Moore–Penrose/Hodge contraction is the independent determinantal average of elementary basis contractions is essentially due to Catanzaro–Chernyak–Klein; here we study the whole polytope of averages under arbitrary joint laws of basis choices. The new information in its fibers is carried by correlations between basis choices in adjacent chain degrees: the normalization $h \mapsto hdh$ erases them, the square-zero locus is exactly the zero-correlation locus, and the Hodge contraction is the relative-entropy minimizer in its fiber. We identify the vertices with nondegenerate τ-chains, model the whole graded polytope as a linear image of a layered path-flow polytope, and obtain a hierarchy from the GKZ secondary polytope of the product basis configuration to a compatible-system fiber polytope and then to the graded Berg fiber polytope. We also record connections with the Cayley trick, iterated fiber polytopes, Plücker deformation theory, and higher matrix-tree weights, and a matroid refinement of the Dereziński–Warmuth variance formula for volume sampling.

Contents

1 Introduction

The starting observation of this note is elementary. Let

$$A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{C}),$$

and let the two coordinate tori act by left and right diagonal multiplication. The corresponding moment map records the squared norms of the rows and columns. After projectivization, the squared matrix entries become a probability distribution $q = (q_{ij})$, and the moment map records its two marginals. Consequently, the fibers of the linear moment map on the ambient coordinate simplex are precisely transportation polytopes.

At the same time, a fixed positive diagonal-scaling orbit of a matrix intersects every admissible interior transportation fiber in a unique point. This is one of the standard forms of the theorem usually called Birch's theorem in algebraic statistics; for full rectangular support it is also the geometry behind Sinkhorn scaling. The distinguished point is equivalently the unique relative-entropy minimizer in the transportation fiber; compare Birch [3], Sinkhorn [4], and Csiszár [5].

The analogous picture for the coordinate torus acting on a Grassmannian replaces the edge polytope of a bipartite graph by a matroid base polytope. If

$$A \in \mathbb{R}^{r\times n}, \qquad \operatorname{rank} A = r,$$

then the Grassmannian moment map of the row space of A is

$$\operatorname{diag}\big(A^{\mathsf T}(AA^{\mathsf T})^{-1}A\big).$$

The moment polytope of the torus-orbit closure is the matroid base polytope of the represented matroid, by the classical theorem of Gelfand–Goresky–MacPherson–Serganova [7].

The second observation is that the same squared-minor coefficients occurring in this Grassmannian moment map occur in Berg's determinantal formula for the Moore–Penrose inverse. Berg proved that the Moore–Penrose solution is a convex combination of the solutions of the nonsingular partial systems [1]; see also Miao–Ben-Israel [2]. In matrix form, for a full-row-rank matrix the pseudoinverse is a convex combination of embedded inverses of maximal nonsingular submatrices, with coefficients proportional to squared determinants. The same identity appears in Ben-Tal–Teboulle [19] and Ben-Israel [20]; for the incidence matrix of a graph it is Kirchhoff's classical description of electrical currents as averages over spanning trees [21].

This suggests introducing the convex hull of these elementary inverses and studying the map

$$X \longmapsto \operatorname{diag}(XA).$$

The map sends an elementary inverse indexed by a matroid basis B to its indicator vector 1_B. Thus one obtains an affine projection

$$\mathcal{B}(A) \longrightarrow P(M(A))$$

from a polytope of right inverses to the matroid base polytope. We call $\mathcal{B}(A)$ the *Berg polytope*.

The purpose of this note is to develop this observation and then extend it from a single matrix to a finite based acyclic chain complex. In the graded case a new phenomenon appears: besides the ordinary Berg geometry of the individual differentials, the fibers contain correlation data between basis choices in consecutive degrees. The Moore–Penrose/Hodge contraction is the distinguished zero-correlation point, and the minimizer of relative entropy to the independent determinantal law in its fiber; its description as the independent determinantal average is essentially due to Catanzaro–Chernyak–Klein [30].

1.1 Relation to prior work

Most ingredients are classical, and several statements below are known in other language. We list them so that the new material is clearly delimited.

- The determinantal convex-combination formula for the Moore–Penrose inverse is due to Berg [1]; it appears also in Ben-Tal–Teboulle [19], Ben-Israel [20] and Miao–Ben-Israel [2]. For graphs it is Kirchhoff's theorem [21].
- The weighted version, namely the closed form $WA^{\mathsf{T}}(AWA^{\mathsf{T}})^{-1}$ of the Birch–Berg section of Section 8 as a convex combination of basic solutions, is classical in the analysis of interior-point methods; see Dikin [22], Stewart [23], and the review and generalization by Forsgren [24].

- Leverage scores are the inclusion probabilities of the determinantal basis measure [25]. Unbiasedness and the closed-form covariance of the volume-sampled pseudoinverse are due to Dereziński–Warmuth [15]. Product-form maximum-entropy distributions on bases with prescribed marginals are the λ-uniform spanning trees of Asadpour et al. [27], in the general framework of Singh–Vishnoi [28].
- For cellular chain complexes, Catanzaro–Chernyak–Klein proved that the orthogonal projection onto cycles is a torsion-weighted average of tree projections [29], and that the (weighted) pseudoinverse of a boundary operator is an average over pairs of spanning co-trees and spanning trees with product weights [30, Thm. 6.1]. In the acyclic case this is the independent-average description of the Hodge contraction in Section 15. Determinantal measures on bases of cell complexes were studied by Lyons [26].
- The replacement $h \mapsto hdh$, which forces $h^2 = 0$, is the standard device for imposing side conditions in homological perturbation theory; see e.g. [31].

To our knowledge the following are new, although the literature search behind this list was not exhaustive: the Berg polytope and its projection to the matroid base polytope, with the face-lattice lift, the affine-hull and fiber-dimension theorems, and the identification of the tangent space of an interior fiber with the dual of the Grassmannian deformation space modulo the torus (Section 6); the graded Berg polytope as the polytope of averages under *arbitrary* joint laws, its covariance decomposition, the path-flow model, and the GKZ–compatible–Berg hierarchy (Sections 11 to 13); the probabilistic reading of the normalization and the identification of the square-zero locus with the zero-correlation locus (Section 14); and the matroid refinement of the Dereziński–Warmuth trace formula (Section 17). The examples of Section 19 should presently be regarded as a program of exact computations rather than as established results.

2 Matrices, the two-sided coordinate torus, and transportation polytopes

Let

$$\mathcal{M} = \mathrm{Mat}_{m\times n}(\mathbb{C})$$

with its standard Hermitian metric. Put

$$T_L = (S^1)^m, \qquad T_R = (S^1)^n,$$

and use the action

$$(D_L, D_R) \cdot A = D_L A D_R^{-1}.$$

Let $e_1, \ldots, e_m$ be the standard basis of $\mathfrak{t}_L^* \cong \mathbb{R}^m$ and $f_1, \ldots, f_n$ the standard basis of $\mathfrak{t}_R^* \cong \mathbb{R}^n$. The coordinate a_{ij} has weight

$$e_i - f_j.$$

Up to the conventional factor 1/2, the moment map is therefore

$$\mu(A) = \sum_{i,j} |a_{ij}|^2 (e_i - f_j),$$

or equivalently

$$\mu(A) = (\operatorname{diag}(AA^*), -\operatorname{diag}(A^*A)) \,. \tag{1}$$

Thus the left component consists of squared row norms,

$$r_i = \sum_j |a_{ij}|^2,$$

and the right component consists, up to sign, of squared column norms,

$$c_j = \sum_i |a_{ij}|^2.$$

The diagonal subgroup $(\lambda I_m, \lambda I_n)$ acts trivially, and correspondingly

$$\sum_i r_i = \sum_j c_j.$$

It is convenient from now on to change sign in the right factor and regard the moment coordinates as (r, c) with nonnegative coordinates.

2.1 Projectivization and the marginal map

For a nonzero matrix put

$$q_{ij} = \frac{|a_{ij}|^2}{\sum_{k,l} |a_{kl}|^2}.$$

Then $q_{ij} \geq 0$ and $\sum q_{ij} = 1$. If the allowed support is

$$S \subset [m] \times [n],$$

the vector q belongs to the simplex

$$\Delta_S = \left\{ q_{ij} \geq 0 : \sum_{(i,j)\in S} q_{ij} = 1 \right\}.$$

Define the marginal map

$$\pi_S : \Delta_S \longrightarrow \mathbb{R}^m \oplus \mathbb{R}^n$$

by

$$\pi_S(q) = (r, c), \qquad r_i = \sum_j q_{ij}, \qquad c_j = \sum_i q_{ij}. \tag{2}$$

Its image is

$$P_S = \operatorname{conv}\{e_i + f_j : (i, j) \in S\}. \tag{3}$$

For full support,

$$P_S = \Delta^{m-1} \times \Delta^{n-1}.$$

For a fixed $x = (r, c) \in P_S$, the fiber

$$\pi_S^{-1}(x) = \left\{ q_{ij} \geq 0 : \sum_j q_{ij} = r_i, \quad \sum_i q_{ij} = c_j \right\} \tag{4}$$

is the transportation polytope with marginals (r, c) and prescribed support S. Thus

$$\Delta_S \xrightarrow{\pi_S} P_S, \qquad \pi_S^{-1}(r, c) = \mathcal{T}_S(r, c). \tag{5}$$

The moment polytope is therefore *not* the transportation polytope. The moment polytope parametrizes possible marginals; the transportation polytope is a fiber over one fixed marginal point.

3 The positive torus orbit as a canonical section

Fix positive reference coefficients $a_{ij} > 0$ on S and put

$$u_{ij} = a_{ij}^2.$$

(We avoid the letter c here; it is reserved for column marginals.) Positive diagonal scaling gives, after squaring and projectivizing,

$$q_{ij}(\theta,\vartheta) = \frac{u_{ij}e^{\theta_i+\vartheta_j}}{\sum_{(k,l)\in S} u_{kl}e^{\theta_k+\vartheta_l}}, \qquad \theta \in \mathbb{R}^m,\ \vartheta \in \mathbb{R}^n. \tag{6}$$

Its marginals are $\pi_S(q(\theta,\vartheta))$.

The map $(\theta,\vartheta) \mapsto \pi_S(q(\theta,\vartheta))$ is invariant under adding a common constant to all θ_i and, for each connected component K of the bipartite support graph of S, under $(\theta,\vartheta) \mapsto (\theta + t1_K, \vartheta - t1_K)$, where 1_K denotes the indicator vector of the rows, respectively columns, of K. After dividing out this stabilizer, the map identifies the positive torus orbit with $\operatorname{relint} P_S$. Equivalently,

$$\{\text{positive diagonal scalings of } u\} \cap \mathcal{T}_S(r,c)$$

contains exactly one point for every admissible interior pair of marginals (r,c). This is a standard form of Birch's theorem [3]; for full rectangular support it is closely related to Sinkhorn's theorem [4].

There is an information-theoretic formulation. For $x = (r,c) \in \operatorname{relint} P_S$, the distinguished point is

$$q^x = \arg\min_{q\in\mathcal{T}_S(x)} D_{KL}(q\|u), \tag{7}$$

where, after normalizing u to a probability distribution,

$$D_{KL}(q\|u) = \sum_{ij} q_{ij} \log \frac{q_{ij}}{u_{ij}}.$$

The Lagrange multiplier equations give precisely the exponential form (6); see Csiszár [5] for the general I-projection framework.

The closure of the positive part of the corresponding projective toric orbit maps homeomorphically onto P_S. This is the positive-real form of the standard toric moment-map theorem; a convenient account is Sottile [6].

This transportation picture will serve as the model for the matroid and Berg constructions below.

4 Grassmannians and matroid moment polytopes

Let
$$A = [a_1, \dots, a_n] \in \mathbb{R}^{r\times n}, \qquad \operatorname{rank} A = r,$$
and let
$$V = \operatorname{row}(A) \in \operatorname{Gr}(r, n).$$
The coordinate torus acts by rescaling the columns. The orthogonal projection onto $V \subset \mathbb{R}^n$ is

$$\Pi_V = A^T(AA^T)^{-1}A. \tag{8}$$

Hence the Grassmannian moment map is

$$\mu_{\mathrm{Gr}}(V) = \operatorname{diag}\Pi_V. \tag{9}$$

Its coordinates are the leverage scores

$$\mu_i = a_i^T(AA^T)^{-1}a_i,$$

and satisfy
$$0 \le \mu_i \le 1, \qquad \sum_i \mu_i = r.$$

Let $M = M(A)$ be the represented matroid. Its bases are the r-subsets $B \subset [n]$ such that $\det A_B \neq 0$. The matroid base polytope is

$$P(M) = \operatorname{conv}\{1_B : B \in \mathfrak{B}(M)\}. \tag{10}$$

Under the Plücker embedding, the coordinate indexed by B has weight 1_B. The GGMS theorem identifies the moment polytope of the torus-orbit closure with $P(M)$ [7].

For a positive diagonal matrix

$$W = \operatorname{diag}(w_1, \dots, w_n),$$

one obtains the determinantal probability distribution

$$p_W(B) = \frac{\det(A_B)^2 \prod_{i\in B} w_i}{\det(AWA^T)}. \tag{11}$$

Cauchy–Binet gives the normalization. Its one-element marginals are

$$x_i = \sum_{B \ni i} p_W(B) = w_i\, a_i^{\mathsf T}(AWA^{\mathsf T})^{-1} a_i. \tag{12}$$

These are the inclusion probabilities of the determinantal basis measure [25]. Thus the Grassmannian moment map is itself a Birch map for the configuration

$$\{1_B : B \in \mathfrak{B}(M)\}.$$

The bipartite transportation picture has become

$$(i,j) \rightsquigarrow B, \qquad e_i + f_j \rightsquigarrow 1_B.$$

5 Berg's theorem and the Berg polytope

For a basis $B \in \mathfrak{B}(M)$, let

$$E_B : \mathbb{R}^r \longrightarrow \mathbb{R}^n$$

denote the coordinate inclusion into the coordinates indexed by B, and define

$$X_B = E_B A_B^{-1}. \tag{13}$$

Then

$$AX_B = I_r.$$

Thus X_B is the elementary right inverse supported on B.

If $g \in GL_r(\mathbb{R})$, then

$$X_B(gA) = X_B(A)g^{-1}, \qquad (gA)^{\dagger} = A^{\dagger}g^{-1}, \qquad X_B(gA)\, gA = X_B(A)A.$$

Hence $\mathcal{B}(gA) = \mathcal{B}(A)\, g^{-1}$ is linearly isomorphic to $\mathcal{B}(A)$, compatibly with the projection β introduced below, and the matroid and the determinantal weights $\det(A_B)^2/\det(AA^{\mathsf T})$ are unchanged. We use this freedom to perform row operations (Theorem 6.6) or to whiten A (Section 17).

Definition 5.1 (Berg polytope). The Berg polytope of A is

$$\mathcal{B}(A) = \operatorname{conv}\{X_B : B \in \mathfrak{B}(M)\}. \tag{14}$$

The following is the full-row-rank specialization of Berg's formula [1]; we include a short derivation because the same determinant identity will be used repeatedly.

Proposition 5.2 (Determinantal Berg identity). *One has*

$$\boxed{A^\dagger = \sum_{B\in\mathfrak{B}(M)} \nu_B X_B, \qquad \nu_B = \frac{\det(A_B)^2}{\det(AA^\mathsf{T})}.} \tag{15}$$

In particular the coefficients are nonnegative and sum to one.

Proof. Cauchy–Binet gives, for every $A' \in \mathrm{Mat}_{r\times n}(\mathbb{R})$,

$$Z(A') := \det(A'A'^\mathsf{T}) = \sum_{|B|=r} \det(A'_B)^2,$$

the sum running over all r-subsets of $[n]$; at $A' = A$ only the bases of M contribute. Let $H \in \mathrm{Mat}_{r\times n}(\mathbb{R})$ and $H_B = HE_B$. Jacobi's formula yields

$$DZ(A)[H] = 2Z(A)\, \left\langle (AA^\mathsf{T})^{-1}A, H\right\rangle_F.$$

Differentiating the Cauchy–Binet expansion term by term instead, and noting that $D(\det^2) = 2\det\cdot D\det$ vanishes at every r-subset B with $\det A_B = 0$, gives

$$\begin{aligned} DZ(A)[H] &= 2 \sum_{B\in\mathfrak{B}(M)} \det(A_B)^2\, \mathrm{tr}(A_B^{-1}H_B) \\ &= 2\left\langle \sum_{B\in\mathfrak{B}(M)} \det(A_B)^2 A_B^{-\mathsf{T}}E_B^\mathsf{T},\, H\right\rangle_F. \end{aligned}$$

Since this holds for every H,

$$\sum_B \det(A_B)^2 A_B^{-\mathsf{T}}E_B^\mathsf{T} = Z(A)(AA^\mathsf{T})^{-1}A.$$

Transposing and dividing by $Z(A)$ gives

$$\sum_B \nu_B E_B A_B^{-1} = A^\mathsf{T}(AA^\mathsf{T})^{-1} = A^\dagger.$$

The normalization $\sum_B \nu_B = 1$ is the first Cauchy–Binet identity above. $\square$

Thus the Moore–Penrose inverse is a distinguished determinantal average of the elementary inverses; see also [19, 20, 2] and the volume-sampling perspective of Dereziński–Warmuth [15].

Now define

$$\beta_A : \mathcal{B}(A) \longrightarrow \mathbb{R}^n, \qquad \beta_A(X) = \operatorname{diag}(XA). \tag{16}$$

For an elementary inverse, $X_B A$ is the projection onto the coordinate subspace $\mathbb{R}^B$ along $\ker A$. Therefore

$$\beta_A(X_B) = 1_B. \tag{17}$$

Consequently

$$\beta_A(\mathcal{B}(A)) = P(M). \tag{18}$$

Every $X \in \mathcal{B}(A)$ satisfies $AX = I$. Therefore

$$\Pi_X = XA$$

is idempotent:

$$\Pi_X^2 = X(AX)A = XA = \Pi_X,$$

and

$$\ker \Pi_X = \ker A.$$

Thus $\mathcal{B}(A)$ can equivalently be regarded as a polytope of oblique projections with fixed kernel, while β_A records their diagonals.

6 Vertices, edges, faces, and dimension

The projection β_A remembers enough of the matroid polytope that its face lattice appears canonically inside the face lattice of $\mathcal{B}(A)$. We begin with the elementary observation that the fibers over matroid vertices are singletons.

Proposition 6.1 (Elementary inverses are vertices). *For every basis $B \in \mathfrak{B}(M)$,*

$$\mathcal{B}(A) \cap \beta_A^{-1}(1_B) = \{X_B\}.$$

Consequently

$$\boxed{\operatorname{Vert} \mathcal{B}(A) = \{X_B : B \in \mathfrak{B}(M)\}.} \tag{19}$$

Proof. Write a point of $\mathcal{B}(A)$ as $X = \sum_C t_C X_C$, with $t_C \geq 0$ and $\sum_C t_C = 1$. If $\beta_A(X) = 1_B$, then

$$1_B = \sum_C t_C 1_C.$$

Since 1_B is a vertex of the matroid base polytope, every C with $t_C > 0$ equals B. Hence $X = X_B$. In particular X_B is a vertex. Conversely every vertex of a convex hull belongs to its displayed generating set. $\square$

Theorem 6.2 (Canonical lift of matroid faces). *Let F be a face of the matroid base polytope $P(M)$. Then*

$$\widehat{F} := \mathcal{B}(A) \cap \beta_A^{-1}(F)$$

is a face of $\mathcal{B}(A)$, with

$$\mathrm{Vert}(\widehat{F}) = \{X_B : \ 1_B \in F\}, \qquad \beta_A(\widehat{F}) = F.$$

Consequently

$$F \longmapsto \widehat{F}$$

is an injective order-preserving map from the face lattice of $P(M)$ to the face lattice of $\mathcal{B}(A)$.

Proof. Choose a linear functional ℓ exposing F:

$$F = \{x \in P(M) : \ell(x) = c_F\}, \qquad \ell \leq c_F \text{ on } P(M).$$

Then $\ell \circ \beta_A \leq c_F$ on $\mathcal{B}(A)$. Since $\mathcal{B}(A)$ is the convex hull of the X_B, the equality set of this supporting functional is

$$\mathrm{conv}\{X_B : \ell(1_B) = c_F\} = \mathrm{conv}\{X_B : 1_B \in F\}.$$

It is therefore a face. Proposition 6.1 shows that all displayed generators are vertices of $\mathcal{B}(A)$, hence exactly the vertices of this face. Their images are precisely the vertices of F, so $\beta_A(\widehat{F}) = F$. Injectivity and order preservation follow immediately. $\square$

Corollary 6.3 (Basis-exchange edges). *If two bases B, B' differ by one basis exchange, then*

$$[X_B, X_{B'}]$$

is an edge of $\mathcal{B}(A)$. Hence the basis-exchange graph of M occurs canonically as a spanning subgraph of the 1-skeleton of $\mathcal{B}(A)$.

Proof. The segment $[1_B, 1_{B'}]$ is an edge of the matroid base polytope. By Theorem 6.2, its lift is a face of $\mathcal{B}(A)$ with exactly the two vertices $X_B, X_{B'}$, hence is their segment. The description of the edges of a matroid base polytope is standard; see GGMS [7] or Oxley [8]. □

Remark 6.4. The Berg polytope can have additional faces, including additional edges, which do not descend from faces of $P(M)$. Faces not of the form $\widehat{F}$ are unavoidable whenever

$$\dim \mathcal{B}(A) > \dim P(M).$$

Indeed, if $F \mapsto \widehat{F}$ were onto, it would be an isomorphism of face lattices (its inverse is order-preserving because $\widehat{F}$ determines the vertex set of F), whereas the dimension of a polytope is determined by its face lattice as the length of a maximal chain of faces minus one. Thus Theorem 6.2 identifies a canonical matroidal skeleton inside a generally richer realization-dependent polytope.

The lifted basis-exchange edges have a particularly simple rank-one form. Write a_j for the jth column of A, and let $e_j \in \mathbb{R}^n$ denote the corresponding coordinate vector.

Proposition 6.5 (Rank-one exchange direction). *Fix a basis* B *and* $j \notin B$*. Put*

$$c(B, j) := e_j - X_B a_j \in \ker A.$$

For $b \in B$*, let* $\varphi_b^B \in (\mathbb{R}^r)^*$ *be the coordinate functional characterized by*

$$(X_B y)_b = \varphi_b^B(y).$$

If

$$B' = B - b + j$$

is a basis, then $c(B, j)_b \neq 0$ *and*

$$X_{B'} - X_B = -\frac{1}{c(B, j)_b} c(B, j) \otimes \varphi_b^B. \tag{20}$$

In particular every lifted basis-exchange edge has rank one. After multiplication by A*,*

$$(X_{B'} - X_B)A = -\frac{1}{c(B, j)_b} c(B, j) \otimes (\varphi_b^B A), \tag{21}$$

so the corresponding change of oblique projector is a fundamental-circuit–fundamental-cocircuit tensor.

Proof. The vector $c = c(B, j)$ is the fundamental circuit vector of j relative to B, normalized by $c_j = 1$. Thus c is supported on $B \cup \{j\}$, and $B - b + j$ is a basis exactly when $c_b \neq 0$.

Set
$$X' := X_B - \frac{1}{c_b} c \otimes \varphi_b^B.$$
Since $Ac = 0$, one has $AX' = I_r$. For $y \in \mathbb{R}^r$, the bth coordinate of $X'y$ is
$$\varphi_b^B(y) - \frac{c_b}{c_b}\varphi_b^B(y) = 0.$$
Outside $B \cup \{j\}$ all coordinates vanish because both X_B and c vanish there. Hence X' is a right inverse supported on $B - b + j$. Since $A_{B'}$ is invertible, there is only one such right inverse, namely $X_{B'}$. This proves (20). Multiplication by A gives (21); the row vector $\varphi_b^B A$ is the corresponding fundamental cocircuit. □

We next determine the affine hull. Let
$$M = M_1 \oplus \cdots \oplus M_\kappa$$
be the decomposition into connected components, with ground sets E_α, sizes $n_\alpha = |E_\alpha|$, and ranks r_α.

Theorem 6.6 (Component decomposition and affine hull). *After a change of basis in* $\mathbb{R}^r$*, the representation splits as a direct sum*
$$A_\alpha : \mathbb{R}^{E_\alpha} \longrightarrow V_\alpha, \qquad \mathbb{R}^r = \bigoplus_{\alpha=1}^{\kappa} V_\alpha,$$
and there is an affine product decomposition
$$\mathcal{B}(A) \cong \prod_{\alpha=1}^{\kappa} \mathcal{B}(A_\alpha). \tag{22}$$
For a connected component of rank r_α *on* n_α *elements,*
$$\operatorname{aff} \mathcal{B}(A_\alpha) = \{X_\alpha : A_\alpha X_\alpha = I_{V_\alpha}\}.$$
Consequently
$$\boxed{\dim \mathcal{B}(A) = \sum_{\alpha=1}^{\kappa} r_\alpha(n_\alpha - r_\alpha).} \tag{23}$$
In particular, if $M(A)$ *is connected,*
$$\operatorname{aff} \mathcal{B}(A) = \{X : AX = I_r\}, \qquad \dim \mathcal{B}(A) = r(n-r). \tag{24}$$

Proof. For a direct sum of matroids, the column spans of the components have additive dimensions and therefore form a direct sum. Choosing a basis of $\mathbb{R}^r$ adapted to these spans makes A block diagonal by components. Every matroid basis is a disjoint union $B = \bigsqcup B_\alpha$, and X_B is the corresponding direct sum of the X_{B_α}. Taking convex hulls gives (22). A rank-zero loop component or a rank-one singleton coloop component contributes a single point and hence zero dimension, so it remains to treat a nontrivial connected component, for which $0 < r < n$.

Fix a basis B_0 and, after row operations (which change $\mathcal{B}(A)$ only by the linear isomorphism $X \mapsto Xg^{-1}$ of Section 5) and a column permutation, write

$$A = [I_r \ C].$$

Let $I = B_0$ and $J = E \setminus B_0$. For $j \in J$ put

$$k_j = e_j - \sum_{i \in I} C_{ij} e_i \in \ker A.$$

The vectors k_j, $j \in J$, form a basis of $\ker A$. Let ε_i, $i \in I$, be the standard dual basis of $(\mathbb{R}^r)^*$, and let

$$W := \operatorname{span}\{X_B - X_{B_0} : B \in \mathfrak{B}(M)\}.$$

We claim that

$$k_j \otimes \varepsilon_i \in W \qquad (i \in I,\ j \in J). \tag{25}$$

Since these tensors form a basis of $\operatorname{Hom}(\mathbb{R}^r, \ker A)$, the claim proves the theorem.

Consider the fundamental bipartite graph $G_{B_0}(M)$, with vertex classes I and J and an edge $i - j$ when $C_{ij} \neq 0$. By Proposition 6.5, (25) holds whenever $i - j$ is an edge. For a connected matroid this fundamental graph is connected; see Oxley [8, Prop. 4.3.2]. We prove (25) for arbitrary i, j by induction on their distance in this graph.

Let a shortest path from i to j be

$$i_0 - j_1 - i_1 - j_2 - \cdots - i_{s-1} - j_s, \qquad i_0 = i, \quad j_s = j.$$

Its length is $2s-1$. The case $s = 1$ is precisely the fundamental-exchange case above. Assume $s > 1$ and that the claim is known for all pairs at smaller graph distance. Because the path is shortest, it is induced: if two nonconsecutive path vertices were adjacent, that edge would shorten the path.

Set

$$I' = \{i_1, \dots, i_{s-1}\}, \qquad J' = \{j_1, \dots, j_{s-1}\}, \qquad B' = B_0 - I' + J'.$$

Order the rows of $C[I', J']$ as $i_1, \dots, i_{s-1}$ and the columns as $j_1, \dots, j_{s-1}$. Inducedness says that, inside these selected rows and columns, j_t can meet only i_{t-1} and i_t. Since i_0 is not among the selected rows, the matrix has the form

$$\begin{pmatrix} C_{i_1 j_1} & C_{i_1 j_2} & 0 & \cdots & 0 \\ 0 & C_{i_2 j_2} & C_{i_2 j_3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & C_{i_{s-2} j_{s-2}} & C_{i_{s-2} j_{s-1}} \\ 0 & \cdots & \cdots & 0 & C_{i_{s-1} j_{s-1}} \end{pmatrix},$$

with every diagonal entry nonzero. Hence B' is a basis. Similarly

$$B'' = B' - i_0 + j_s = B_0 - \{i_0, \dots, i_{s-1}\} + \{j_1, \dots, j_s\}$$

is a basis: the corresponding square submatrix has nonzero diagonal entries $C_{i_{t-1} j_t}$ and is triangular after the same ordering. Thus B'' is obtained from B' by the single exchange $i_0 \leftrightarrow j_s$.

Apply Proposition 6.5 to this exchange. Since $c(B', j_s)$ is supported on $B' \cup \{j_s\}$ and has j_s-coordinate 1, its expansion in the original kernel basis $\{k_j : j \in J\}$ can involve only $j_s, j_1, \dots, j_{s-1}$:

$$c(B', j_s) = k_{j_s} + \sum_{t=1}^{s-1} \alpha_t k_{j_t}.$$

Likewise $i_0 \in B'$, and solving for the coefficient of the unchanged basis column e_{i_0} can depend only on the rows $i_0, i_1, \dots, i_{s-1}$ replaced along the path. Hence

$$\varphi_{i_0}^{B'} = \varepsilon_{i_0} + \sum_{p=1}^{s-1} \beta_p \varepsilon_{i_p}.$$

Proposition 6.5 therefore places in W a nonzero scalar multiple of

$$\left(k_{j_s} + \sum_{t<s} \alpha_t k_{j_t} \right) \otimes \left(\varepsilon_{i_0} + \sum_{p>0} \beta_p \varepsilon_{i_p} \right).$$

The desired term is $k_{j_s} \otimes \varepsilon_{i_0}$. Every other term has strictly smaller graph distance: indeed

$$\mathrm{dist}(i_0, j_t) = 2t - 1 < 2s - 1, \qquad \mathrm{dist}(i_p, j_s) = 2(s-p) - 1 < 2s - 1,$$

and for $p, t > 0$ the path segment joining i_p to j_t is shorter still. By the induction hypothesis all those other elementary tensors already lie in W. Subtracting them isolates $k_{j_s} \otimes \varepsilon_{i_0} \in W$. This completes the induction and proves (25).

Thus the differences of Berg vertices span

$$\mathrm{Hom}(\mathbb{R}^r, \ker A) = \{Y : AY = 0\},$$

which is exactly the translation space of the affine right-inverse space $\{X : AX = I_r\}$. Its dimension is $r(n-r)$. □

The dimension formula should be compared with

$$\dim P(M) = n - \kappa.$$

The difference is the dimension of every fiber over a relative interior point.

Lemma 6.7 (Relative-interior fibers of affine polytope maps). *Let* $f : P \to Q = f(P)$ *be an affine surjection of polytopes and let* $q \in \mathrm{relint}\, Q$. *Then*

$$\mathrm{aff}\big(P \cap f^{-1}(q)\big) = \mathrm{aff}(P) \cap f^{-1}(q).$$

Equivalently, the translation space of the fiber is the kernel of the linear part of f *restricted to the translation space of* $\mathrm{aff}(P)$.

Proof. Let $p_1, \dots, p_N$ be all vertices of P and $q_i = f(p_i)$. Put

$$\bar{q} = \frac{1}{N} \sum_{i=1}^{N} q_i.$$

Because the q_i affinely span Q and every coefficient in this average is positive, $\bar{q} \in \mathrm{relint}\, Q$. Since also $q \in \mathrm{relint}\, Q$, the relative interior is open in $\mathrm{aff}\, Q$; hence for sufficiently small $\varepsilon > 0$ the point

$$q' := \frac{q - \varepsilon \bar{q}}{1 - \varepsilon} = q + \frac{\varepsilon}{1 - \varepsilon}(q - \bar{q})$$

still belongs to Q. Choose any convex representation

$$q' = \sum_i \mu_i q_i, \qquad \mu_i \geq 0, \qquad \sum_i \mu_i = 1.$$

Then

$$q = \sum_i \lambda_i q_i, \qquad \lambda_i = \frac{\varepsilon}{N} + (1-\varepsilon)\mu_i > 0.$$

Thus

$$p_0 := \sum_i \lambda_i p_i$$

is a strictly positive convex combination of all vertices of P, hence lies in $\operatorname{relint} P$, and $f(p_0) = q$.

Let L be the translation space of $\operatorname{aff}(P)$ and let $v \in L$ lie in the kernel of the linear part of f. Because $p_0 \in \operatorname{relint} P$, for sufficiently small $t > 0$ both $p_0 \pm tv$ lie in P. They also lie in the fiber over q. Thus every such v belongs to the translation space of the affine hull of the fiber. The reverse inclusion is immediate. □

Theorem 6.8 (Affine hull and dimension of an interior fiber). *Let $x \in \operatorname{relint} P(M)$. If M is connected, then*

$$\operatorname{aff} \beta_A^{-1}(x) = \{X : AX = I_r,\ \operatorname{diag}(XA) = x\}, \tag{26}$$

and

$$\dim \beta_A^{-1}(x) = r(n-r) - (n-1) = (r-1)(n-r-1). \tag{27}$$

For a matroid with connected components (n_α, r_α)*,*

$$\boxed{\dim \beta_A^{-1}(x) = \sum_{\alpha=1}^{\kappa} (r_\alpha - 1)(n_\alpha - r_\alpha - 1).} \tag{28}$$

Proof. Apply Lemma 6.7 to

$$\beta_A : \mathcal{B}(A) \longrightarrow P(M)$$

and use Theorem 6.6. In the connected case the affine fiber of the linear equations is exactly the right-hand side of (26). Since $\dim P(M) = n-1$, subtraction gives (27). The component formula follows from (22) and $\dim P(M) = n - \kappa$. □

Remark 6.9. For a connected matroid the generic fiber is zero-dimensional exactly in rank 1 or corank 1. The first genuinely positive-dimensional case is therefore $r \geq 2$ and $n - r \geq 2$.

Finally, the tangent space of an interior fiber has a useful dual interpretation. Put

$$\mathcal{K}_A = \{Y \in \mathrm{Mat}_{n\times r}(\mathbb{R}) : AY = 0,\ \mathrm{diag}(YA) = 0\}. \tag{29}$$

For connected M, this is the translation space of (26).

Proposition 6.10 (Cotangent space and configuration moduli). *Assume $M(A)$ is connected and nontrivial. With the trace pairing between $\mathrm{Mat}_{r\times n}$ and $\mathrm{Mat}_{n\times r}$,*

$$\mathcal{K}_A^* \cong \frac{\mathrm{Mat}_{r\times n}(\mathbb{R})}{\mathfrak{gl}_r A + A\mathfrak{d}_n}, \tag{30}$$

where $\mathfrak{d}_n$ denotes the diagonal matrices. Equivalently,

$$\boxed{\mathcal{K}_A \cong \left(T_{[A]}\,\mathrm{Gr}(r,n)/T_{[A]}(T^{n-1}\cdot [A])\right)^*.} \tag{31}$$

Thus an interior Berg fiber is infinitesimally dual to the local deformation space of the represented Grassmannian point modulo coordinate rescalings.

Proof. Let

$$U = \{Y : AY = 0\}, \qquad V = \{Y : \mathrm{diag}(YA) = 0\}.$$

The adjoint, for the trace pairing, of $Y \mapsto AY$ has image $\mathfrak{gl}_r A$, so

$$U^\perp = \mathfrak{gl}_r A.$$

Likewise the adjoint of $Y \mapsto \mathrm{diag}(YA)$ sends a diagonal matrix D to AD, hence

$$V^\perp = A\mathfrak{d}_n.$$

Therefore

$$\mathcal{K}_A^* = (U \cap V)^* \cong \mathrm{Mat}_{r\times n}/(U^\perp + V^\perp),$$

which proves (30).

It remains only to identify the common infinitesimal stabilizer. Suppose

$$RA = AD$$

with $R \in \mathfrak{gl}_r$ and $D = \operatorname{diag}(d_1, \dots, d_n)$. Then every nonzero column a_i is an eigenvector of R with eigenvalue d_i. If C is a circuit and

$$\sum_{i \in C} \gamma_i a_i = 0, \qquad \gamma_i \neq 0,$$

is its minimal dependence, applying R gives

$$\sum_{i \in C} \gamma_i d_i a_i = 0.$$

The dependence space on a circuit is one-dimensional, hence the d_i are equal on C. In a connected matroid any two elements lie in a common circuit (equivalently, are linked by a chain of circuits), so all d_i equal one scalar λ. Since the columns span $\mathbb{R}^r$, $R = \lambda I_r$. Thus

$$\mathfrak{gl}_r A \cap A\mathfrak{d}_n = \mathbb{R}A,$$

and the effective diagonal action is the $(n-1)$-dimensional coordinate torus. Quotienting first by $\mathfrak{gl}_r A$ gives $T_{[A]} \operatorname{Gr}(r, n)$, and the image of $A\mathfrak{d}_n$ is exactly the tangent space of that torus orbit, proving (31). The dimension is

$$rn - r^2 - n + 1 = (r-1)(n-r-1),$$

in agreement with Theorem 6.8. □

7 The Berg fiber as a matroid transportation quotient

Let

$$\Delta_{\mathfrak{B}} = \left\{ (p_B) : p_B \geq 0, \ \sum_B p_B = 1 \right\}$$

be the probability simplex on the matroid bases. There are two affine maps

$$L_A : \Delta_{\mathfrak{B}} \longrightarrow \mathcal{B}(A), \qquad L_A(p) = \sum_B p_B X_B, \tag{32}$$

$$\pi_M : \Delta_{\mathfrak{B}} \longrightarrow P(M), \qquad \pi_M(p) = \sum_B p_B 1_B. \tag{33}$$

They satisfy

$$\pi_M = \beta_A \circ L_A. \tag{34}$$

For $x \in P(M)$ define the *matroid transportation polytope*

$$\mathcal{T}_M(x) = \left\{ p \in \Delta_{\mathfrak{B}} : \sum_B p_B 1_B = x \right\}. \tag{35}$$

It consists of probability distributions on matroid bases with prescribed one-element marginals

$$x_i = \Pr(i \in B).$$

The Berg fiber is exactly

$$\boxed{\beta_A^{-1}(x) = L_A\big(\mathcal{T}_M(x)\big).} \tag{36}$$

Thus the Berg fiber is a linear image of a transportation-type polytope. This is the precise matroid analogue of (5).

8 The Birch–Berg section

Take the reference determinantal measure

$$\nu_B = \frac{\det(A_B)^2}{\det(AA^T)}. \tag{37}$$

For $x \in \operatorname{relint} P(M)$ consider

$$p^x = \arg\min_{p \in \mathcal{T}_M(x)} D_{KL}(p \| \nu). \tag{38}$$

Birch's theorem gives

$$p_B^x = \frac{\det(A_B)^2 \prod_{i \in B} w_i}{\det(AWA^T)} \tag{39}$$

for a positive diagonal matrix $W = \operatorname{diag}(w_i)$. The distribution p^x is unique. The matrix W is unique up to multiplication by an independent positive scalar on each connected component E_α of M; such rescalings leave every $p_W(B)$ unchanged because each basis meets E_α in exactly r_α elements.

Push this distinguished transportation point down to the Berg polytope:

$$\sigma_A(x) = L_A(p^x). \tag{40}$$

The closed formula follows directly from Proposition 5.2.

Proposition 8.1 (Closed form of the Birch–Berg section). *For the diagonal matrix W determined by x through Birch's theorem,*

$$\boxed{\sigma_A(x) = WA^\top(AWA^\top)^{-1}.} \tag{41}$$

Moreover

$$\beta_A(\sigma_A(x)) = \operatorname{diag}\big(WA^\top(AWA^\top)^{-1}A\big) = x. \tag{42}$$

Hence $\sigma_A : \operatorname{relint} P(M) \to \mathcal{B}(A)$ *is a section of* β_A*.*

Proof. Put

$$\widetilde{A} = AW^{1/2}.$$

For a basis B,

$$\det(\widetilde{A}_B)^2 = \det(A_B)^2 \prod_{i\in B} w_i,$$

and the elementary right inverse of $\widetilde{A}$ supported on B is

$$\widetilde{X}_B = E_B W_B^{-1/2} A_B^{-1}.$$

Proposition 5.2, applied to $\widetilde{A}$, therefore gives

$$\widetilde{A}^\dagger = \sum_B p_B^x \, E_B W_B^{-1/2} A_B^{-1}.$$

Multiplying on the left by $W^{1/2}$ and using

$$W^{1/2} E_B W_B^{-1/2} = E_B$$

gives

$$W^{1/2}\widetilde{A}^\dagger = \sum_B p_B^x E_B A_B^{-1} = L_A(p^x) = \sigma_A(x).$$

Because $\widetilde{A}$ has full row rank,

$$\widetilde{A}^\dagger = \widetilde{A}^\top(\widetilde{A}\widetilde{A}^\top)^{-1} = W^{1/2}A^\top(AWA^\top)^{-1},$$

which proves (41). The identity $A\sigma_A(x) = I_r$ is immediate. Finally,

$$\beta_A(\sigma_A(x)) = \sum_B p_B^x \beta_A(X_B) = \sum_B p_B^x 1_B = x,$$

where the last equality is precisely the prescribed marginal condition in Birch's theorem. This proves (42) without any further matrix calculation. □

At $W = I$ one recovers

$$\sigma_A(x_0) = A^\dagger, \qquad x_0 = \operatorname{diag}(A^\dagger A). \tag{43}$$

Thus the Moore–Penrose inverse is the unscaled Birch point.

Remark 8.2. The identity $WA^\top(AWA^\top)^{-1} = \sum_B p_W(B) X_B$, that is, the fact that weighted minimum-norm and least-squares solutions are convex combinations of basic solutions with the weights $p_W(B)$ of (11), is classical in the analysis of interior-point methods [22, 23, 24]. What is added here is the interpretation of W as the Birch scaling attached to a prescribed point $x \in \operatorname{relint} P(M)$, which turns $x \mapsto \sigma_A(x)$ into a section of β_A. The product form (39) of the relative-entropy minimizer on bases with prescribed marginals is the determinantal analogue of the λ-uniform spanning-tree distributions of [27, 28].

8.1 An entropy potential on the Berg polytope

Define

$$J_A(X) = \min_{\substack{p \in \Delta_{\mathfrak{B}} \\ L_A(p) = X}} D_{KL}(p \| \nu). \tag{44}$$

Proposition 8.3 (Entropy characterization of the Birch–Berg section). *The function J_A is convex on $\mathcal{B}(A)$. For every $x \in \operatorname{relint} P(M)$,*

$$\boxed{\sigma_A(x) = \arg \min_{\substack{X \in \mathcal{B}(A) \\ \beta_A(X) = x}} J_A(X).} \tag{45}$$

The minimizer is unique.

Proof. The minimum in (44) exists because, for $X \in \mathcal{B}(A)$, its feasible set is a nonempty compact polytope (a slice of the simplex $\Delta_{\mathfrak{B}}$) and

$D_{KL}(\cdot\|\nu)$ is continuous on $\Delta_{\mathfrak{B}}$, since $\nu_B > 0$ for every basis. Let $X_0, X_1 \in \mathcal{B}(A)$ and let p_0, p_1 be minimizing lifts. For $0 \le t \le 1$, the convex combination

$$p_t = (1-t)p_0 + tp_1$$

lifts $(1-t)X_0 + tX_1$. Convexity of relative entropy therefore gives

$$J_A((1-t)X_0 + tX_1) \le D_{KL}(p_t\|\nu) \le (1-t)J_A(X_0) + tJ_A(X_1).$$

Hence J_A is convex.

Now use the factorization $\pi_M = \beta_A \circ L_A$. Then

$$\begin{aligned} \min_{\beta_A(X)=x} J_A(X) &= \min_{\substack{p\in\Delta_{\mathfrak{B}} \\ \beta_A(L_A(p))=x}} D_{KL}(p\|\nu) \\ &= \min_{p\in\mathcal{T}_M(x)} D_{KL}(p\|\nu). \end{aligned}$$

Birch's theorem identifies the unique minimizer on the right as p^x. Its image under L_A is $\sigma_A(x)$, proving (45). If another X attained the same minimum, a minimizing lift of X would be another minimizer of the strictly convex relative-entropy problem on $\mathcal{T}_M(x)$, hence would equal p^x; therefore $X = \sigma_A(x)$. $\square$

Thus the Birch–Berg section selects a canonical entropy center in every interior Berg fiber.

The Legendre-dual description is the standard one for finite exponential families. With $W = \operatorname{diag}(e^{\theta_i})$, the log-partition function of the family (39) is

$$\begin{aligned} \Phi_A(\theta) &= \log \sum_B \nu_B e^{\langle \theta, 1_B\rangle} \\ &= \log\det\left(A \operatorname{diag}(e^{\theta_i})A^T\right) - \log\det(AA^T), \qquad \theta \in \mathbb{R}^n. \end{aligned} \tag{46}$$

It is convex, and its gradient is the basis-marginal map,

$$\nabla\Phi_A(\theta) = \operatorname{diag}\big(WA^T(AWA^T)^{-1}A\big),$$

which equals x exactly when W is the Birch scaling attached to x. The function on the base polytope is the Legendre conjugate: for $x \in \operatorname{relint} P(M)$,

$$\Phi_A^*(x) = \sup_\theta\big(\langle\theta, x\rangle - \Phi_A(\theta)\big) = \min_{p\in\mathcal{T}_M(x)} D_{KL}(p\|\nu) = \min_{\substack{X\in\mathcal{B}(A) \\ \beta_A(X)=x}} J_A(X).$$

This places the Berg section directly in the standard finite exponential-family/Legendre-duality framework.

9 Fiber polytopes and infinitesimal Plücker geometry

The factorization

$$\Delta_{\mathfrak{B}} \xrightarrow{L_A} \mathcal{B}(A) \xrightarrow{\beta_A} P(M)$$

has an immediate consequence in the sense of Billera–Sturmfels fiber polytopes [9]. The simplex projection

$$\Delta_{\mathfrak{B}} \longrightarrow P(M)$$

has as its fiber polytope, up to scaling and translation, the secondary polytope of the basis configuration

$$\{1_B : B \in \mathfrak{B}(M)\}.$$

Since Minkowski integration commutes with linear maps,

$$\Sigma\big(\mathcal{B}(A) \to P(M)\big) = L_A\left(\Sigma(\Delta_{\mathfrak{B}} \to P(M))\right), \tag{47}$$

up to the standard normalization (scaling and translation) conventions.

There is also an infinitesimal Plücker interpretation. Let

$$H \in \mathrm{Mat}_{r\times n}.$$

Let $\ell_H(X) = \mathrm{tr}(HX)$ be the corresponding linear functional on $\mathrm{Mat}_{n\times r}$. On elementary inverses,

$$\ell_H(X_B) = \mathrm{tr}(H_B A_B^{-1}), \qquad H_B = HE_B.$$

But

$$\mathrm{tr}(H_B A_B^{-1}) = \left.\frac{d}{dt}\right|_{t=0} \log|\det(A_B + tH_B)|. \tag{48}$$

Hence height functions on the basis configuration arising from linear functionals on the Berg polytope are precisely infinitesimal logarithmic variations of Plücker coordinates.

Perturbations $H = RA$ change every height by the same constant. Perturbations

$$H = A\,\mathrm{diag}(\lambda_i)$$

add the affine function

$$B \longmapsto \sum_{i \in B} \lambda_i,$$

which is invisible to the secondary fan. Thus the nontrivial linear geometry of Berg fibers is naturally related to infinitesimal projective deformation data of the represented configuration.

The iterated-fiber-polytope construction of Billera–Sturmfels [10] becomes relevant below, where a natural flag of polytope projections appears.

10 From one matrix to a based contractible complex

Let

$$0 \longrightarrow C_m \xrightarrow{d_m} C_{m-1} \longrightarrow \cdots \xrightarrow{d_1} C_0 \xrightarrow{d_0} C_{-1} \longrightarrow 0 \tag{49}$$

be a finite exact based real chain complex, and let E_k denote the distinguished coordinate basis of C_k. Over a field, exact, acyclic and contractible (chain homotopy equivalent to zero) are equivalent conditions, and we use the three words interchangeably. For a contractible cell complex one should take the augmented cellular chain complex.

Put

$$Z_k = \ker d_k = \operatorname{im} d_{k+1}.$$

For each k, let $M_k = M(d_k)$ be the column matroid of d_k, now regarded as the surjection

$$d_k : C_k \longrightarrow Z_{k-1}.$$

If

$$S_k \in \mathfrak{B}(M_k),$$

then restriction of d_k to the coordinate subspace $\mathbb{R}^{S_k} \subset C_k$ is an isomorphism onto Z_{k-1}. Denote its inverse by

$$s_k^{S_k} : Z_{k-1} \longrightarrow C_k. \tag{50}$$

Define

$$\rho_k^{S_k} = I - s_k^{S_k} d_k. \tag{51}$$

This is the projection of C_k onto $Z_k = \ker d_k$ along the coordinate subspace $\mathbb{R}^{S_k}$.

For a tuple $S = (S_0, \dots, S_m)$ of bases define

$$h^{S}_{k-1} = s_k^{S_k} \rho_{k-1}^{S_{k-1}}. \tag{52}$$

We use the following conventions at the two ends of the complex. Since $d_{-1} = 0$, the only basis of M_{-1} is $S_{-1} = \emptyset$, and $\rho_{-1} = I_{C_{-1}}$; thus $h^S_{-1} = s_0^{S_0}$. Since d_m is injective, M_m is free, so $S_m = E_m$ is forced, $Z_m = 0$ and $\rho_m = 0$; we put $h_m = 0$. With these conventions the identities below hold in every degree.

Proposition 10.1 (Elementary coordinate contractions). *For every tuple of bases* S *one has*

$$dh^S + h^S d = I, \qquad (h^S)^2 = 0. \tag{53}$$

Moreover, if $R_k = E_k \setminus S_k$*, then every block*

$$d_k[R_{k-1}, S_k] \tag{54}$$

is square and nonsingular. Hence, up to the standard complementary-index convention, the tuples S *are exactly the nondegenerate matrix* τ*-chains of the based acyclic complex.*

Proof. Suppress the superscripts S_k. Since $d_k s_k = I_{Z_{k-1}}$, the operator

$$\rho_k = I - s_k d_k$$

is the projection of C_k onto $Z_k = \ker d_k$ along $\operatorname{im} s_k$. In particular

$$d_k \rho_k = 0, \qquad \rho_k|_{Z_k} = I_{Z_k}, \qquad \rho_k s_k = 0. \tag{55}$$

On C_{k-1} we therefore have

$$\begin{aligned} d_k h_{k-1} + h_{k-2} d_{k-1} &= d_k s_k \rho_{k-1} + s_{k-1} \rho_{k-2} d_{k-1} \\ &= \rho_{k-1} + s_{k-1} d_{k-1} \\ &= I_{C_{k-1}}. \end{aligned}$$

Here $\rho_{k-2} d_{k-1} = d_{k-1}$ because $\operatorname{im} d_{k-1} = Z_{k-2}$. This proves $dh + hd = I$ degree by degree. Likewise

$$h_k h_{k-1} = s_{k+1} \rho_k s_k \rho_{k-1} = 0$$

by (55), proving $h^2 = 0$.

It remains to justify the τ-chain statement. The kernel of ρ_{k-1} is $\operatorname{im} s_{k-1} = \mathbb{R}^{S_{k-1}}$, while its image is Z_{k-1}. Hence the coordinate projection

$$\mathrm{pr}_{R_{k-1}} : Z_{k-1} \longrightarrow \mathbb{R}^{R_{k-1}}$$

is injective. Exactness gives

$$|R_{k-1}| = \dim C_{k-1} - \operatorname{rank} d_{k-1} = \dim Z_{k-1} = \operatorname{rank} d_k = |S_k|,$$

so this projection is in fact an isomorphism. Since

$$d_k|_{\mathbb{R}^{S_k}} : \mathbb{R}^{S_k} \longrightarrow Z_{k-1}$$

is also an isomorphism by definition of S_k, their composite

$$\mathbb{R}^{S_k} \xrightarrow{d_k} Z_{k-1} \xrightarrow{\mathrm{pr}_{R_{k-1}}} \mathbb{R}^{R_{k-1}}$$

is an isomorphism. Its matrix is exactly (54), so that block is nonsingular. Conversely, suppose that all blocks $d_k[R_{k-1}, S_k]$ are square and nonsingular. Since $R_{-1} = E_{-1}$, induction on k and exactness give $|S_k| = |R_{k-1}| = \dim C_{k-1} - \operatorname{rank} d_{k-1} = \operatorname{rank} d_k$, and nonsingularity of the block shows that the columns of d_k indexed by S_k are linearly independent. Hence each S_k is a basis of M_k, which recovers the tuple of degreewise bases. This is the usual matrix τ-chain condition; compare [11]. □

Definition 10.2 (Graded Berg polytope). Define

$$\mathcal{B}_{\mathrm{gr}}(C_\bullet) = \operatorname{conv}\{h^S\}. \tag{56}$$

Every point of $\mathcal{B}_{\mathrm{gr}}(C_\bullet)$ still satisfies

$$dh + hd = I,$$

although a convex combination need not satisfy $h^2 = 0$.

11 The graded matroid projection

Define

$$\beta_{\mathrm{gr}}(h) = (\operatorname{diag}(h_{k-1} d_k))_k\,. \tag{57}$$

At an elementary vertex,

$$h^{S}_{k-1} d_k = s_k^{S_k} d_k,$$

so

$$\mathrm{diag}(h^{S}_{k-1} d_k) = 1_{S_k}. \tag{58}$$

Therefore

$$\beta_{\mathrm{gr}} : \mathcal{B}_{\mathrm{gr}}(C_\bullet) \longrightarrow \prod_k P(M_k) \tag{59}$$

is an affine surjection.

This is the graded analogue of $\mathcal{B}(A) \to P(M(A))$.

12 The graded phenomenon: correlations between degrees

Let

$$\Omega = \prod_k \mathfrak{B}(M_k)$$

and let $p(S)$ be a probability distribution on tuples of bases. Then

$$h = \mathbb{E}_p h^S$$

has components

$$h_{k-1} = \mathbb{E}_p \left[s_k^{S_k} \rho_{k-1}^{S_{k-1}} \right]. \tag{60}$$

Thus h_{k-1} depends only on the joint distribution of the adjacent pair (S_{k-1}, S_k).

If

$$\bar{s}_k = \mathbb{E} s_k^{S_k}, \qquad \bar{\rho}_k = I - \bar{s}_k d_k,$$

then

$$\boxed{h_{k-1} = \bar{s}_k \bar{\rho}_{k-1} + \mathbb{E}\left[(s_k^{S_k} - \bar{s}_k)(\rho_{k-1}^{S_{k-1}} - \bar{\rho}_{k-1}) \right].} \tag{61}$$

The second term is an operator-valued covariance. Therefore the graded geometry has two qualitatively different parts:

$$\text{degreewise Berg geometry} + \text{cross-degree correlation geometry}. \tag{62}$$

The graded moment map sees the one-degree marginals; its fibers retain correlations between basis choices in adjacent chain dimensions. The independent case, in which the covariance term vanishes, is the setting of the Kirchhoff–Boltzmann formula of Catanzaro–Chernyak–Klein [30, Thm. 6.1]; the correlation term is what the present construction adds.

Since the degree graph

$$0 - 1 - \cdots - m$$

is a tree, a global probability distribution is determined, for the purposes of h, by compatible node marginals and adjacent pair marginals. Thus the graded Berg polytope is a linear image of the local marginal polytope of a path graphical model. For fixed node distributions, the pair-marginal fibers are ordinary transportation polytopes. This gives a literal graded analogue of the transportation geometry with which the note began.

13 Compatible τ-chains and graded GKZ geometry

The preceding observation can be made completely polyhedral. It is useful because it separates three levels of combinatorics: full systems of basis choices, the adjacent-degree compatibility data actually seen by a contraction, and the still coarser data retained by the graded Berg polytope.

For brevity write

$$\mathfrak{B}_k = \mathfrak{B}(M_k), \qquad \Omega = \prod_{k=0}^{m} \mathfrak{B}_k.$$

Thus Ω is the set of nondegenerate τ-chains of the based acyclic complex. Let

$$P_{\mathrm{gr}} = \prod_{k=0}^{m} P(M_k).$$

For $S = (S_0, \dots, S_m)$ put

$$\nu_S = (1_{S_0}, \dots, 1_{S_m}) \in P_{\mathrm{gr}}.$$

The configuration

$$\mathcal{A}_{\mathrm{gr}} = \{\nu_S : S \in \Omega\} = \mathcal{A}_0 \times \cdots \times \mathcal{A}_m, \tag{63}$$

where $\mathcal{A}_k = \{1_S : S \in \mathfrak{B}_k\}$, is the product basis configuration. Its convex hull is P_{gr}.

13.1 Vertices, lifted faces, and basis-exchange edges

The graded moment map immediately gives the vertex set.

Proposition 13.1 (Vertices of the graded Berg polytope). *Every elementary contraction* h^S *is a vertex of* $\mathcal{B}_{gr}(C_\bullet)$, *different tuples give different vertices, and*

$$\boxed{\operatorname{Vert}\mathcal{B}_{gr}(C_\bullet) = \{h^S : S \in \Omega\}.} \tag{64}$$

Hence the vertices are canonically indexed by nondegenerate τ*-chains.*

Proof. The point ν_S is a vertex of the product polytope P_{gr}, and $\beta_{gr}(h^S) = \nu_S$. If h^S were a nontrivial convex combination of other elementary contractions, applying β_{gr} would express the vertex ν_S as a nontrivial convex combination of other vertices of P_{gr}. This is impossible. □

Proposition 13.2 (Lifted graded faces and exchange edges). *Every face*

$$F = F_0 \times \cdots \times F_m \leq P_{gr}$$

lifts canonically to the face

$$\widehat{F} = \mathcal{B}_{gr}(C_\bullet) \cap \beta_{gr}^{-1}(F) = \operatorname{conv}\{h^S : \nu_S \in F\}. \tag{65}$$

Its vertices are exactly the h^S *with* $\nu_S \in F$, *and* $\beta_{gr}(\widehat{F}) = F$. *Consequently the face lattice of* P_{gr} *embeds canonically in that of the graded Berg polytope.*

In particular,

$$G(M_0) \,\square\, G(M_1) \,\square \cdots \square\, G(M_m) \subseteq \operatorname{Skel}_1 \mathcal{B}_{gr}(C_\bullet) \tag{66}$$

is a spanning subgraph of the graded Berg 1*-skeleton.*

Proof. Choose a supporting functional ℓ of P_{gr} exposing F, say $\ell \leq c_F$ with equality exactly on F. Then $\ell \circ \beta_{gr} \leq c_F$ on the graded Berg polytope. Since that polytope is the convex hull of the elementary contractions and

$$\beta_{gr}(h^S) = \nu_S,$$

the equality set is exactly the convex hull displayed in (65). The preceding vertex proposition shows that all its displayed generators are vertices. Their images are precisely the vertices of F, so the image is all of F.

An edge of P_{gr} changes exactly one factor, and in that factor it is a basis-exchange edge of a matroid base polytope. Its lifted face therefore has exactly two vertices and is an edge of the graded Berg polytope. Taking all such edges yields (66). Additional graded Berg edges may occur; these are correlation phenomena invisible in P_{gr}. □

13.2 The layered path-flow polytope

Form the layered directed graph Γ_C whose vertices in layer k are the bases $S \in \mathfrak{B}_k$, with every vertex in layer $k-1$ joined to every vertex in layer k. Add a source before layer 0 and a sink after layer m. A source–sink path is exactly a tuple $S \in \Omega$.

Let $\mathfrak{F}_C$ be the unit source–sink flow polytope of Γ_C. Equivalently,

$$\mathfrak{F}_C = \mathrm{conv}\{\chi_S : S \in \Omega\}, \tag{67}$$

where χ_S is the edge-incidence vector of the corresponding path. Since the constraint matrix is a network matrix, $\mathfrak{F}_C$ is integral and its vertices are exactly these path vectors.

Write

$$Q_k(S,T) \geq 0, \qquad S \in \mathfrak{B}_{k-1}, \quad T \in \mathfrak{B}_k,$$

for the flow on an edge between consecutive layers. Flow conservation produces node marginals η_k satisfying

$$\sum_S Q_k(S,T) = \eta_k(T) = \sum_U Q_{k+1}(T,U) \tag{68}$$

for interior layers, with the evident endpoint formulas.

There is a linear map

$$(\mathcal{L}_C Q)_{k-1} = \sum_{S \in \mathfrak{B}_{k-1}} \sum_{T \in \mathfrak{B}_k} Q_k(S,T)\, s_k^T \rho_{k-1}^S. \tag{69}$$

At a path vertex this gives h^S. Consequently

$$\mathcal{L}_C(\mathfrak{F}_C) = \mathcal{B}_{\mathrm{gr}}(C_\bullet). \tag{70}$$

The moment projection from the flow polytope is

$$\Pi_C(Q) = \left(\sum_{S \in \mathfrak{B}_k} \eta_k(S) 1_S \right)_{k=0}^{m}. \tag{71}$$

By construction,

$$\Pi_C = \beta_{gr} \circ \mathcal{L}_C. \tag{72}$$

Thus we obtain the graded analogue of the simplex–Berg–matroid factorization:

$$\Delta_\Omega \longrightarrow \mathfrak{F}_C \xrightarrow{\mathcal{L}_C} \mathcal{B}_{gr}(C_\bullet) \xrightarrow{\beta_{gr}} P_{gr}. \tag{73}$$

The first arrow forgets all higher-order information in a distribution on full τ-chains and retains only adjacent pair marginals.

13.3 Compatible transportation systems

Fix probability distributions η_k on $\mathfrak{B}_k$. Once the node marginals are fixed, the adjacent flow matrices may be chosen independently subject only to their two marginals. Hence

$$\varpi_C^{-1}(\eta_0, \dots, \eta_m) \cong \prod_{k=1}^{m} \mathcal{T}(\eta_{k-1}, \eta_k), \tag{74}$$

where ϖ_C is the node-marginal map and $\mathcal{T}(\eta, \eta')$ is the ordinary transportation polytope of couplings of η and η'.

The adjective "compatible" here has a literal meaning: two consecutive transport plans are compatible when the right marginal of one equals the left marginal of the next. On a path this local condition is already sufficient for a global distribution.

Lemma 13.3 (Gluing compatible adjacent couplings). *Let Q_k be a coupling of η_{k-1} and η_k for every $1 \le k \le m$. Then there exists a probability distribution p on $\Omega = \prod_k \mathfrak{B}_k$ whose $(k-1, k)$ pair marginal is Q_k for every k. If the intermediate marginals $\eta_1, \dots, \eta_{m-1}$ are strictly positive, one such distribution is*

$$p(S_0, \dots, S_m) = \frac{\prod_{k=1}^{m} Q_k(S_{k-1}, S_k)}{\prod_{k=1}^{m-1} \eta_k(S_k)}. \tag{75}$$

Consequently, for fixed node marginals, the compatible-flow fiber is canonically the product of ordinary transportation polytopes in (74).

Proof. Restrict first to the supports of the η_k, so that all denominators are positive. Define the transition kernels

$$K_k(T \mid S) = \frac{Q_k(S, T)}{\eta_{k-1}(S)}.$$

The left-marginal condition on Q_k gives

$$\sum_T K_k(T \mid S) = 1,$$

while the right-marginal condition gives

$$\sum_S \eta_{k-1}(S) K_k(T \mid S) = \eta_k(T).$$

Hence

$$p(S_0, \dots, S_m) = \eta_0(S_0) \prod_{k=1}^{m} K_k(S_k \mid S_{k-1})$$

is a probability law. Cancelling the successive factors $\eta_{k-1}(S_{k-1})$ gives exactly (75). Iterated summation shows that its node marginal in layer k is η_k. Summing over all variables except S_{k-1}, S_k leaves

$$\eta_{k-1}(S_{k-1}) K_k(S_k \mid S_{k-1}) = Q_k(S_{k-1}, S_k),$$

so the prescribed pair marginals are recovered.

Conversely, every global law produces adjacent pair marginals with the stated common node marginals. Since the path-flow variables are precisely these pair marginals and the only linear compatibility equations are their common marginals, fixing (η_k) leaves the independent Cartesian product of coupling polytopes in (74). If some η_k has zeros, the same argument applies after deleting the zero-mass states and then extending by zero. $\square$

Thus no higher compatibility obstruction occurs on a path.

Now fix only a graded moment point

$$x = (x_0, \dots, x_m) \in P_{\mathrm{gr}}.$$

For each degree, the possible node laws form the matroid transportation polytope $\mathcal{T}_{M_k}(x_k)$ from Section 7. Therefore the flow fiber $\Pi_C^{-1}(x)$ projects onto

$$\prod_k \mathcal{T}_{M_k}(x_k), \tag{76}$$

and the fiber over a point $(\eta_k)_k$ is exactly the product in (74). In this precise sense a graded Berg fiber is a linear image of an *iterated system of transportation polytopes*: first choose a basis distribution in every degree with prescribed matroid marginals, and then choose compatible couplings between consecutive degrees.

13.4 The graded secondary polytope and compatible systems

The full configuration $\mathcal{A}_{\mathrm{gr}}$ has the ordinary GKZ secondary polytope

$$\Sigma(\mathcal{A}_{\mathrm{gr}}) = \Sigma(\Delta_\Omega \longrightarrow P_{\mathrm{gr}}),$$

whose normal fan is the secondary fan of regular subdivisions of the product basis configuration [14, 12].

Let

$$R_{\mathrm{pair}} : \Delta_\Omega \longrightarrow \mathfrak{F}_C$$

be the adjacent-pair-marginal map. Define

$$\Sigma_{\mathrm{comp}}(C_\bullet) := \Sigma(\mathfrak{F}_C \xrightarrow{\Pi_C} P_{\mathrm{gr}}). \tag{77}$$

By functoriality of Minkowski integration under linear maps,

$$\Sigma_{\mathrm{comp}}(C_\bullet) = R_{\mathrm{pair}}\big(\Sigma(\mathcal{A}_{\mathrm{gr}})\big) \tag{78}$$

up to the usual translation convention for fiber polytopes. Thus Σ_{comp} is a canonical linear image of the full GKZ secondary polytope.

There is a useful dual description. A linear functional on $\mathfrak{F}_C$ pulls back to a height on full τ-chains of the form

$$\omega(S_0, \dots, S_m) = \sum_{k=1}^{m} \omega_k(S_{k-1}, S_k). \tag{79}$$

Hence the normal fan of Σ_{comp} classifies exactly those coherent subdivisions of $\mathcal{A}_{\mathrm{gr}}$ that are visible to pairwise-additive, path-local heights. We shall call these *coherent compatible τ-chain systems*. This terminology is descriptive; the underlying statement is the standard Billera–Sturmfels correspondence between the normal fan of a fiber polytope and coherent subdivisions [9].

Finally define the graded Berg fiber polytope

$$\Sigma^{\mathrm{gr}}_{\mathrm{Berg}}(C_\bullet) := \Sigma\left(\mathcal{B}_{\mathrm{gr}}(C_\bullet) \xrightarrow{\beta_{\mathrm{gr}}} P_{\mathrm{gr}}\right). \tag{80}$$

The same functoriality gives

$$\Sigma(\mathcal{A}_{\mathrm{gr}}) \xrightarrow{R_{\mathrm{pair}}} \Sigma_{\mathrm{comp}}(C_\bullet) \xrightarrow{\mathcal{L}_C} \Sigma^{\mathrm{gr}}_{\mathrm{Berg}}(C_\bullet). \tag{81}$$

Thus the graded Berg secondary geometry is a two-step quotient of ordinary GKZ secondary geometry:

$$\begin{aligned}\text{all coherent systems of } \tau\text{-chains} &\longrightarrow \text{adjacent-compatible systems}\\ &\longrightarrow \text{systems distinguishable by contractions}\end{aligned}$$

This is the precise graded analogue of the one-matrix identity expressing the Berg fiber polytope as a linear image of the secondary polytope of the matroid basis configuration.

The flag of projections (73) also puts the construction naturally in the setting of iterated fiber polytopes of Billera–Sturmfels [10]: each stage records exactly the coherent information lost at the next quotient.

13.5 Where the Cayley trick enters

For an adjacent pair of degrees the local state configuration is

$$\mathcal{A}_{k-1,k} = \mathcal{A}_{k-1} \times \mathcal{A}_k, \tag{82}$$

and local coherent structures are regular subdivisions of this product configuration. In general this is product-secondary geometry rather than literally a Cayley configuration. However, whenever one of the two factors is affinely independent, say

$$\mathcal{A}_{k-1} \cong \Delta^{s-1},$$

we have

$$\Delta^{s-1} \times \mathcal{A}_k = \mathrm{Cayley}(\mathcal{A}_k, \ldots, \mathcal{A}_k)$$

with s copies. The classical Cayley trick then identifies regular subdivisions of the adjacent product with coherent mixed subdivisions of

$$\underbrace{P(M_k) + \cdots + P(M_k)}_{s \text{ copies}},$$

and triangulations with fine mixed subdivisions [14, 12, 13].

This applies in particular to the bottom block $(0,1)$ of any augmented complex whose matroid M_0 has rank one, for instance the augmented chain complex of a contractible simplicial complex: the bases of M_0 are single vertices and $P(M_0)$ is a simplex. At the top, M_m has a single basis, so the block $(m-1, m)$ carries no correlation at all. For the augmented chain complex of the N-simplex, the block $(N-2, N-1)$ is also of Cayley type, because M_{N-1} is the corank-one uniform matroid $U_{N,N+1}$, whose base polytope is again a simplex. Thus the outermost nontrivial correlation blocks of a simplex lie exactly in the classical Cayley–GKZ regime. The middle blocks, where neither factor is a simplex, are genuinely product-secondary objects; the path-flow polytope and the compatible secondary polytope above provide the appropriate replacement for a naive Cayley description.

This distinction is useful. It prevents one from forcing the ordinary Cayley trick where it does not literally apply, while retaining a precise statement: *the graded theory is governed globally by a secondary polytope of the product basis configuration, locally by secondary polytopes of adjacent products, and in simplex-factor blocks by the usual Cayley correspondence with mixed subdivisions.*

14 Canonical normalization: erasing correlations

The algebraic normalization used in the graded theory is worth isolating because it explains exactly which part of a contraction is correlation data. The first part of the following proposition is the standard device for imposing the side condition $h^2 = 0$ in homological perturbation theory (see e.g. [31]); the probabilistic interpretation and the description of the square-zero locus are what we add.

Proposition 14.1 (Normalization and the zero-correlation locus). *Let* h *be any degree-*$+1$ *operator satisfying* $dh + hd = I$*, and define*

$$N(h) = hdh. \tag{83}$$

Then

$$Nd = hd, \qquad dN = dh, \qquad dN + Nd = I, \qquad N^2 = 0. \tag{84}$$

In particular N *is a square-zero contraction with the same two complementary projections as* h.

If $h \in \mathcal{B}_{\mathrm{gr}}(C_\bullet)$ *is represented by a probability law on tuples of bases and*

$$\bar{s}_k = \mathbb{E} s_k^{S_k}, \qquad \bar{\rho}_k = I - \bar{s}_k d_k,$$

then

$$N(h)_{k-1} = \bar{s}_k \bar{\rho}_{k-1}. \tag{85}$$

Consequently $N(h)$ *lies again in the graded Berg polytope, has the same graded moment point as* h, *and is obtained probabilistically by replacing the joint law by the product of its one-degree marginals.*

Finally, if

$$\mathcal{S}_k := \mathrm{conv}\{s_k^S : S \in \mathfrak{B}(M_k)\}$$

is the ordinary degree-k *section polytope, then the map*

$$\Upsilon : \prod_k \mathcal{S}_k \longrightarrow \mathcal{B}_{\mathrm{gr}}(C_\bullet), \qquad \Upsilon((s_k))_{k-1} = s_k(I - s_{k-1} d_{k-1}), \tag{86}$$

is injective and

$$\boxed{\mathcal{B}_{\mathrm{gr}}(C_\bullet) \cap \{h : h^2 = 0\} = \Upsilon\left(\prod_k \mathcal{S}_k\right).} \tag{87}$$

Thus all genuinely new graded directions are correlation directions.

Proof. Put

$$P = hd, \qquad Q = dh.$$

The relation $dh + hd = I$ says $P + Q = I$. Using $d^2 = 0$,

$$P^2 = hdhd = h(dh)d = h(I - hd)d = hd = P,$$

and similarly

$$Q^2 = dhdh = d(hd)h = d(I - dh)h = dh = Q.$$

Moreover $PQ = hddh = 0$, hence also $QP = (I - P)P = 0$. Since

$$N = hdh = PhQ,$$

we get

$$N^2 = PhQPhQ = 0.$$

Also

$$Nd = hdhd = (hd)^2 = hd, \qquad dN = dhdh = (dh)^2 = dh,$$

which proves all identities in (84). In particular the diagonal operators $h_{k-1}d_k$ that enter β_{gr} are unchanged, so

$$\beta_{gr}(N(h)) = \beta_{gr}(h). \tag{88}$$

Now let h be an average of elementary contractions. Since $d_k(C_k) \subset Z_{k-1}$ and every $\rho_{k-1}^{S_{k-1}}$ acts as the identity on Z_{k-1},

$$h_{k-1}d_k = \mathbb{E}\left[s_k^{S_k}\rho_{k-1}^{S_{k-1}}d_k\right] = \bar{s}_k d_k.$$

Likewise

$$d_k h_{k-1} = \mathbb{E}\left[d_k s_k^{S_k}\rho_{k-1}^{S_{k-1}}\right] = \bar{\rho}_{k-1}.$$

Therefore

$$N(h)_{k-1} = h_{k-1}d_k h_{k-1} = \bar{s}_k d_k h_{k-1} = \bar{s}_k\bar{\rho}_{k-1},$$

proving (85). If the one-degree marginals of the original law are η_k, then the product law $\bigotimes_k \eta_k$ has expectation

$$\mathbb{E}_{\otimes\eta_k}\left[s_k^{S_k}\rho_{k-1}^{S_{k-1}}\right] = \bar{s}_k\bar{\rho}_{k-1};$$

hence $N(h)$ belongs to the graded Berg polytope and literally erases the adjacent correlations.

For the last assertion, choose barycentric laws realizing arbitrary $s_k \in \mathcal{S}_k$ and take their product. Its average elementary contraction is $\Upsilon((s_k))$, so Υ lands in the graded Berg polytope; the elementary calculation of Proposition 10.1, or directly

$$(I - s_k d_k)s_k = 0,$$

shows that $\Upsilon((s_k))^2 = 0$. Conversely, if $h \in \mathcal{B}_{gr}$ and $h^2 = 0$, then multiplying $dh + hd = I$ on the left by h gives

$$hdh + h^2 d = h,$$

so $N(h) = h$. Formula (85) therefore writes $h = \Upsilon((\bar{s}_k))$. Finally, Υ is injective because on the cycle space Z_{k-1} one has

$$\Upsilon((s_j))_{k-1}|_{Z_{k-1}} = s_k,$$

as d_{k-1} vanishes there. This proves (87). $\square$

15 The Hodge contraction as the independent Birch point

For each differential d_k, take the determinantal basis measure

$$\nu_k(S) \propto \det(d_{k,S})^2, \tag{89}$$

where the determinant is computed after choosing an orthonormal basis of $Z_{k-1} = \operatorname{im} d_k$.

Proposition 15.1 (Independent determinantal average). *Let*

$$\boldsymbol{\nu} = \bigotimes_k \nu_k. \tag{90}$$

Then

$$\boxed{\mathbb{E}_{\boldsymbol{\nu}} h^S_{k-1} = d_k^\dagger} \tag{91}$$

for every k. *Consequently*

$$h_H = (d_k^\dagger)_k \tag{92}$$

is the expectation of the elementary contractions under independent determinantal basis choices and satisfies

$$d_{k+1} d_{k+1}^\dagger + d_k^\dagger d_k = I_{C_k}. \tag{93}$$

Proof. Apply Proposition 5.2 to the surjection

$$d_k : C_k \longrightarrow Z_{k-1}$$

after choosing an orthonormal basis of its image. This gives

$$\mathbb{E}_{\nu_k} s_k^{S_k} = d_k^\dagger|_{Z_{k-1}}.$$

Likewise

$$\mathbb{E}_{\nu_{k-1}} \rho_{k-1}^{S_{k-1}} = I - d_{k-1}^\dagger d_{k-1},$$

which is the orthogonal projector onto

$$\ker d_{k-1} = Z_{k-1}.$$

Independence of S_k and S_{k-1} therefore gives

$$\begin{aligned}\mathbb{E}_\nu h_{k-1}^S &= \left(\mathbb{E}s_k^{S_k}\right)\left(\mathbb{E}\rho_{k-1}^{S_{k-1}}\right)\\ &= d_k^\dagger(I - d_{k-1}^\dagger d_{k-1})\\ &= d_k^\dagger,\end{aligned}$$

because $d_k^\dagger$ vanishes on $Z_{k-1}^\perp$ and hence already factors through the orthogonal projection onto Z_{k-1}. This proves (91).

Finally, $d_{k+1}d_{k+1}^\dagger$ is the orthogonal projector onto $\operatorname{im} d_{k+1} = Z_k$, whereas $d_k^\dagger d_k$ is the orthogonal projector onto $(\ker d_k)^\perp = Z_k^\perp$. The two subspaces are orthogonal complements, proving (93). □

Hence the Moore–Penrose/Hodge contraction is characterized probabilistically as the independent determinantal average of elementary basis contractions. Theorem 15.1 is the acyclic, unweighted case of the Kirchhoff–Boltzmann projection formula of Catanzaro–Chernyak–Klein [30, Thm. 6.1], which writes the pseudoinverse of a boundary operator as an average over pairs (spanning co-tree, spanning tree) with product weights. In an acyclic based complex a spanning co-tree in degree $k-1$ is a basis S_{k-1} of M_{k-1}, the associated projection is $\rho_{k-1}^{S_{k-1}}$, and the normalized product weights are $\nu_{k-1} \otimes \nu_k$. We include the short proof above because it is the unweighted instance of the section constructed next.

After positive diagonal rescaling in each chain degree, the degreewise measures become Birch exponential families. Prescribing

$$x_k \in \operatorname{relint} P(M_k)$$

selects a unique degreewise Birch distribution $p_k^{x_k}$. Their product gives a canonical section

$$\sigma_{\mathrm{gr}} : \prod_k \operatorname{relint} P(M_k) \longrightarrow \mathcal{B}_{\mathrm{gr}}(C_\bullet). \tag{94}$$

Explicitly, by Theorem 8.1 applied degree by degree, $\sigma_{\mathrm{gr}}(x)_{k-1} = \bar{s}_k\bar{\rho}_{k-1}$ with $\bar{s}_k = W_k d_k^T(d_k W_k d_k^T)^{-1}$ (in an orthonormal basis of Z_{k-1}), where W_k is the Birch scaling for x_k; thus $\sigma_{\mathrm{gr}}(x)$ is the Hodge contraction of $C_\bullet$ for the diagonal inner products $\langle u, v\rangle_k = u^T W_k^{-1} v$.

The section has the following variational characterization. If p is any joint law on tuples of bases with one-degree marginals p_k, then

$$\boxed{D_{KL}\left(p \,\middle\|\, \bigotimes_k \nu_k\right) = D_{KL}\left(p \,\middle\|\, \bigotimes_k p_k\right) + \sum_k D_{KL}(p_k \| \nu_k).} \tag{95}$$

The first term is the multi-information and is nonnegative, with equality exactly for the product law. Under prescribed graded moment data x_k, the second term is minimized degree by degree by $p_k^{x_k}$, by Birch's theorem. Therefore $\bigotimes_k p_k^{x_k}$ is the unique relative-entropy minimizer among all joint laws with those graded moment constraints. At the unweighted point this is precisely the product determinantal law of Proposition 15.1. In this precise sense the graded Hodge/Birch section is the minimum-relative-entropy law with the prescribed graded moments. Among couplings with fixed one-degree marginals, minimizing $D_{KL}(\cdot \| \bigotimes_k \nu_k)$ is the same as maximizing Shannon entropy, so it is also the maximum-entropy coupling of its marginals.

16 Cellular trees and torsion weights

For an acyclic regular cell complex, for instance a regular cellular ball, bases of the boundary matroids are higher-dimensional analogues of spanning trees and forests. In higher dimension determinant weights need not be 1. Kalai's higher-dimensional tree enumeration [16] and the simplicial matrix-tree theorem of Duval–Klivans–Martin [17] show that higher spanning trees naturally occur with the square of the order of their torsion homology.

Thus, in non-unimodular cellular examples, the reference measure in the Berg/Hodge construction is naturally a torsion-squared determinantal measure. The Moore–Penrose contraction does not merely average combinatorial trees: it averages elementary inverse operators with exactly the determinant/torsion weights already singled out by higher matrix-tree theory. At the operator level this torsion weighting already appears in the higher projection formula of Catanzaro–Chernyak–Klein [29] and in their Kirchhoff–Boltzmann formula [30]; see also Lyons [26] for determinantal measures on bases of cell complexes.

17 Exchange geometry and fluctuations of inverses

There is another useful invariant already for a single matrix. First whiten A, replacing it by $(AA^{\mathsf{T}})^{-1/2}A$, so that

$$AA^{\mathsf{T}} = I_r.$$

For a basis B, let $X_B = E_B A_B^{-1}$ and let

$$d_M(B) = \#\{(i,j) : i \in B,\ j \notin B,\ B - i + j \in \mathfrak{B}(M)\} \tag{96}$$

be the degree of B in the basis-exchange graph. By Section 5, whitening changes $\mathcal{B}(A)$ only by a linear isomorphism and changes neither the matroid nor the determinantal measure.

Under the determinantal measure $\nu_B = \det(A_B)^2$ (which is already normalized because $AA^{\mathsf{T}} = I_r$), one has the following exact fluctuation identity.

Proposition 17.1 (Mean exchange degree equals the variance of elementary inverses). *For whitened* A*,*

$$\boxed{\mathbb{E}_\nu \|X_B - A^\dagger\|_F^2 = \mathbb{E}_\nu d_M(B).} \tag{97}$$

Proof. Put $\Pi_B = X_B A$. The jth column of Π_B is e_j for $j \in B$, whereas for $j \notin B$ it is supported on B, with entries the coefficients c_{ij}, $i \in B$, in

$$A_j = \sum_{i \in B} c_{ij} A_i.$$

Hence

$$\|\Pi_B\|_F^2 = r + \sum_{j \notin B} \sum_{i \in B} c_{ij}^2. \tag{98}$$

Cramer's rule gives

$$\det(A_B)^2 c_{ij}^2 = \det(A_{B-i+j})^2,$$

where the right-hand side is interpreted as zero when $B - i + j$ is not a basis. Multiplying (98) by ν_B and summing over B therefore gives

$$\mathbb{E}_\nu \|\Pi_B\|_F^2 = r + \sum_B \sum_{\substack{i \in B,\ j \notin B \\ B-i+j \in \mathfrak{B}(M)}} \det(A_{B-i+j})^2.$$

The map

$$(B,i,j) \longmapsto (B-i+j,j,i)$$

is an involution on directed basis exchanges. Reindexing the last sum by this involution yields

$$\mathbb{E}_\nu \|\Pi_B\|_F^2 = r + \sum_B \nu_B d_M(B) = r + \mathbb{E}_\nu d_M(B).$$

Because $AA^T = I_r$,

$$\|\Pi_B\|_F^2 = \mathrm{tr}(X_B A A^T X_B^T) = \|X_B\|_F^2.$$

Moreover Proposition 5.2 gives

$$\mathbb{E}_\nu X_B = A^\dagger = A^T, \qquad \|A^\dagger\|_F^2 = r.$$

Therefore

$$\begin{aligned}
\mathbb{E}_\nu \|X_B - A^\dagger\|_F^2 &= \mathbb{E}_\nu \|X_B\|_F^2 - 2\left\langle \mathbb{E}_\nu X_B, A^\dagger \right\rangle_F + \|A^\dagger\|_F^2 \\
&= \mathbb{E}_\nu \|\Pi_B\|_F^2 - r \\
&= \mathbb{E}_\nu d_M(B),
\end{aligned}$$

which proves (97). □

Remark 17.2. If M is uniform, i.e. for A in general position, every basis has $d_M(B) = r(n-r)$, and (97) gives $\mathbb{E}_\nu \|X_B - A^\dagger\|_F^2 = r(n-r)$ and $\mathbb{E}_\nu \|X_B\|_F^2 = r(n-r+1)$. This agrees with the trace of the closed-form covariance of the volume-sampled pseudoinverse obtained by Dereziński–Warmuth [15] under a general-position assumption. If M is not uniform, the mean exchange degree is strictly smaller than $r(n-r)$: for a dependent r-set D and a basis B with $|B \cap D| = \mathrm{rank}(D) < r$, every exchange $B - i + j$ with $i \in B \setminus D$ and $j \in D \setminus B$ is dependent. Thus (97) is a matroid-sensitive refinement of the general-position formula.

This suggests the determinant-weighted exchange polynomial

$$\Psi_A(x;q) = \sum_{B \in \mathfrak{B}(M)} \det(A_B)^2 x^B q^{d_M(B)}. \tag{99}$$

At $q = 1$,

$$\Psi_A(x;1) = \det\left(A\,\mathrm{diag}(x) A^T\right), \tag{100}$$

the usual configuration polynomial. Moreover

$$q\frac{\partial}{\partial q}\log\Psi_A(1;q)\bigg|_{q=1} = \mathbb{E}_\nu d_M(B) = \mathbb{E}_\nu\|X_B - A^\dagger\|_F^2. \tag{101}$$

Thus basis-exchange geometry measures the mean-square fluctuation of elementary inverses around the Moore–Penrose inverse.

18 Relation to secondary geometry

The statistic

$$B \longmapsto d_M(B)$$

is a height function on the basis configuration $\{1_B\}$. Consequently it determines a regular subdivision of the basis configuration, namely the one indexed by the cone of the secondary fan containing this height.

Introducing $q = e^\eta$ and $x_i = e^{\theta_i}$ gives the log-partition function

$$\begin{aligned}\Phi_A(\theta,\eta) &= \log\sum_B \nu_B e^{\langle\theta,1_B\rangle+\eta d_M(B)}\\ &= \log\Psi_A(e^\theta;e^\eta) - \log\det(AA^T),\end{aligned} \tag{102}$$

which extends (46): $\Phi_A(\theta,0) = \Phi_A(\theta)$. Its gradient is

$$\nabla\Phi_A(\theta,\eta) = \left(\mathbb{E}_{\theta,\eta}1_B, \mathbb{E}_{\theta,\eta}d_M(B)\right), \tag{103}$$

the expectations being taken for the tilted law

$$p_{\theta,\eta}(B) \propto \nu_B\, e^{\langle\theta,1_B\rangle+\eta d_M(B)}.$$

Thus the exchange energy itself is another mean parameter. The corresponding lifted polytope is

$$\widehat{P}_{\mathrm{ex}}(M) = \mathrm{conv}\{(1_B, d_M(B)) : B \in \mathfrak{B}(M)\}. \tag{104}$$

The ordinary Birch family is the slice $\eta = 0$. At $(\theta,\eta) = (0,0)$ and for whitened A, the η-derivative is the fluctuation $\mathbb{E}_\nu\|X_B - A^\dagger\|_F^2$ of Theorem 17.1; away from that point it is only the mean exchange degree of the tilted law. This gives another connection between the Berg construction and GKZ secondary geometry [14].

19 Small examples and computational directions

The smallest graded example is the augmented chain complex of a triangle. The degreewise Berg fibers are trivial, but the graded fiber has one dimension. Indeed, $M_0 = U_{1,3}$ and $M_1 = U_{2,3}$ have trivial matroid transportation polytopes, so the node laws are forced by the graded moment point, and M_2 has a single basis. The 3×3 coupling polytope of the block $(0, 1)$ is 4-dimensional, but by (61) only the covariance term $\mathbb{E}[(s_1^{S_1} - \bar{s}_1)(\rho_0^{S_0} - \bar{\rho}_0)]$ varies; it is an operator with image in Z_1, which is one-dimensional, and a direct computation shows that it is not constant. Its only freedom is therefore a correlation between the choice of a vertex basis and the choice of an edge basis; the Hodge contraction is the independent point in the middle of this interval.

For a tetrahedron, ordinary degreewise Berg-fiber directions coexist with genuinely graded correlation directions. The following observations come from computations that are not reproduced here and should be treated as conjectural. In the 4-simplex, exact computations indicate a relation between an outer correlation body and the linear-ordering polytope. The relevant six-dimensional representation is

$$\Lambda^2(1^{\perp}),$$

the same representation in which the hidden S_5-symmetry of LO_4 lives. Katthän [18] obtains this S_5-action from an S_5-equivariant projection of LO_5 onto LO_4; we expect the cyclic cosets appearing in our computations to be explained by this representation-theoretic picture, but this has not been checked.

In the middle degree of even-dimensional simplices, complementation of simplices produces a signed Kneser/odd-graph operator. In the 4-simplex this is a signed Petersen operator; in the 6-simplex it becomes, after normalization, a complex structure on the middle cycle space. These examples suggest that discrete Hodge-star operators interact naturally with graded Berg correlation geometry.

These detailed examples should be separated in a final version into a computational section, with exact scripts or independently checked certificates for assertions depending on enumeration.

20 Summary

The construction may be read as a sequence of increasingly refined versions of the same diagram.

For ordinary matrices,

$$\text{coordinate simplex} \longrightarrow \text{row/column moment polytope},$$

with transportation polytopes as fibers and a positive toric/Birch section.

For a represented matroid,

$$\Delta_{\mathfrak{B}(M)} \longrightarrow \mathcal{B}(A) \longrightarrow P(M),$$

where the first map averages elementary inverses and the second takes the diagonal of the associated oblique projection. The fibers of $\mathcal{B}(A) \to P(M)$ are linear images of matroid transportation polytopes. Birch's theorem selects an entropy-minimizing distribution in the simplex fiber; its image is the canonical Berg section

$$x \longmapsto WA^{T}(AWA^{T})^{-1}.$$

At the unweighted point this is the Moore–Penrose inverse.

For a based contractible chain complex,

$$\mathcal{B}_{\mathrm{gr}}(C_\bullet) \longrightarrow \prod_k P(M_k),$$

and the new information in the fiber consists of correlations between basis choices in adjacent dimensions. The normalization

$$h \longmapsto hdh$$

erases these correlations without changing the graded moment point. The independent determinantal distribution then gives the Hodge contraction, as in [30]; the independent Birch distributions give the weighted Hodge contractions.

In this sense the guiding principle is:

$$\begin{array}{c}\text{transportation geometry}\\ \downarrow\\ \text{toric/Birch section}\\ \downarrow\\ \text{matroid moment geometry}\\ \downarrow\\ \text{Berg polytope of inverses}\\ \downarrow\\ \text{graded Berg polytope of contractions.}\end{array}$$

The ordinary moment map remembers marginal data. The Berg polytope remembers inverse or splitting data above those marginals. In a chain complex, the graded Berg polytope remembers additionally how splittings in consecutive dimensions are correlated.

This places Moore–Penrose and Hodge theory, determinantally weighted spanning trees, matroid moment polytopes, transportation geometry, and fiber and secondary polytopes in one common framework.

References

[1] L. Berg, *Three results in connection with inverse matrices*, Linear Algebra Appl. **84** (1986), 63–77. DOI: 10.1016/0024-3795(86)90307-1.

[2] J. Miao and A. Ben-Israel, *Minors of the Moore–Penrose inverse*, Linear Algebra Appl. **195** (1993), 191–207. DOI: 10.1016/0024-3795(93)90264-O.

[3] M. W. Birch, *Maximum likelihood in three-way contingency tables*, J. Roy. Statist. Soc. Ser. B **25** (1963), 220–233.

[4] R. Sinkhorn, *Diagonal equivalence to matrices with prescribed row and column sums*, Amer. Math. Monthly **74** (1967), 402–405. DOI: 10.2307/2314570.

[5] I. Csiszár, I-*divergence geometry of probability distributions and minimization problems*, Ann. Probab. **3** (1975), 146–158. DOI: 10.1214/aop/1176996454.

[6] F. Sottile, *Toric ideals, real toric varieties, and the moment map*, in *Topics in Algebraic Geometry and Geometric Modeling*, Contemp. Math. **334**, Amer. Math. Soc., 2003, 225–240. Preprint: arXiv:math/0212044.

[7] I. M. Gelfand, R. M. Goresky, R. D. MacPherson, and V. V. Serganova, *Combinatorial geometries, convex polyhedra, and Schubert cells*, Adv. Math. **63** (1987), 301–316. DOI: 10.1016/0001-8708(87)90059-4.

[8] J. Oxley, *Matroid Theory*, 2nd ed., Oxford Graduate Texts in Mathematics, Oxford University Press, 2011. DOI: 10.1093/acprof:oso/9780198566946.001.0001.

[9] L. J. Billera and B. Sturmfels, *Fiber polytopes*, Ann. of Math. (2) **135** (1992), 527–549. DOI: 10.2307/2946575.

[10] L. J. Billera and B. Sturmfels, *Iterated fiber polytopes*, Mathematika **41** (1994), 348–363. DOI: 10.1112/S0025579300007440.

[11] V. G. Turaev, *Introduction to Combinatorial Torsions*, Lectures in Mathematics ETH Zürich, Birkhäuser, Basel, 2001.

[12] J. A. De Loera, J. Rambau, and F. Santos, *Triangulations: Structures for Algorithms and Applications*, Algorithms and Computation in Mathematics, Vol. 25, Springer, Berlin, 2010. DOI: 10.1007/978-3-642-12971-1.

[13] F. Santos, *The Cayley trick and triangulations of products of simplices*, in *Integer Points in Polyhedra—Geometry, Number Theory, Algebra, Optimization*, Contemp. Math. **374**, Amer. Math. Soc., 2005, 151–177. Preprint: arXiv:math/0312069.

[14] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky, *Discriminants, Resultants, and Multidimensional Determinants*, Birkhäuser, Boston, 1994.

[15] M. Dereziński and M. K. Warmuth, *Unbiased estimates for linear regression via volume sampling*, in *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, 3084–3093. Preprint: arXiv:1705.06908.

[16] G. Kalai, *Enumeration of $\mathbb{Q}$-acyclic simplicial complexes*, Israel J. Math. **45** (1983), 337–351. DOI: 10.1007/BF02804017.

[17] A. M. Duval, C. J. Klivans, and J. L. Martin, *Simplicial matrix-tree theorems*, Trans. Amer. Math. Soc. **361** (2009), 6073–6114.

[18] L. Katthän, *The linear ordering polytope via representations*, 2011. Preprint: arXiv:1109.5040.

[19] A. Ben-Tal and M. Teboulle, *A geometric property of the least squares solution of linear equations*, Linear Algebra Appl. **139** (1990), 165–170.

[20] A. Ben-Israel, *A volume associated with $m \times n$ matrices*, Linear Algebra Appl. **167** (1992), 87–111. DOI: 10.1016/0024-3795(92)90340-G.

[21] G. Kirchhoff, *Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird*, Ann. Phys. Chem. **72** (1847), 497–508.

[22] I. I. Dikin, *On the speed of an iterative process*, Upravlyaemye Sistemy **12** (1974), 54–60 (in Russian).

[23] G. W. Stewart, *On scaled projections and pseudoinverses*, Linear Algebra Appl. **112** (1989), 189–193.

[24] A. Forsgren, *On linear least-squares problems with diagonally dominant weight matrices*, SIAM J. Matrix Anal. Appl. **17** (1996), 763–788. DOI: 10.1137/S0895479895284014.

[25] R. Lyons, *Determinantal probability measures*, Publ. Math. Inst. Hautes Études Sci. **98** (2003), 167–212.

[26] R. Lyons, *Random complexes and ℓ^2-Betti numbers*, J. Topol. Anal. **1** (2009), 153–175.

[27] A. Asadpour, M. X. Goemans, A. Mądry, S. Oveis Gharan, and A. Saberi, *An $O(\log n/\log\log n)$-approximation algorithm for the asymmetric traveling salesman problem*, in *Proc. 21st ACM–SIAM Symposium on Discrete Algorithms (SODA 2010)*, 379–389.

[28] M. Singh and N. K. Vishnoi, *Entropy, optimization and counting*, in *Proc. 46th ACM Symposium on Theory of Computing (STOC 2014)*, 50–59.

[29] M. J. Catanzaro, V. Y. Chernyak, and J. R. Klein, *Kirchhoff's theorems in higher dimensions and Reidemeister torsion*, Homology Homotopy Appl. **17** (2015), no. 1, 165–189. DOI: 10.4310/HHA.2015.v17.n1.a8.

[30] M. J. Catanzaro, V. Y. Chernyak, and J. R. Klein, *A higher Boltzmann distribution*, J. Appl. Comput. Topol. **1** (2017), no. 2, 215–240. Preprint: arXiv:1506.06775.

[31] M. Crainic, *On the perturbation lemma, and deformations*, preprint, 2004. arXiv:math/0403266.