SCHUBERT CALCULUS AND QUANTUM INFORMATION

ALLEN KNUTSON

ABSTRACT. These four lectures are for the "Workshop on Quantum Marginals and Density Matrices" at the Fields Institute in Toronto, July 27-31, 2009.

CONTENTS

1. Lecture 1: The Schur-Horn theorem and the Gel′fand-Cetlin system 1
1.1. The Schur-Horn problem 2
1.2. Flag manifolds 3
1.3. The Bruhat decomposition 3
1.4. The Gel′fand-Cetlin system 5
2. Lecture 2: The Weyl-Horn problem and Schubert calculus 5
2.1. The Johnson/Klyachko/Helmke-Rosenthal/Totaro inequalities 6
2.2. Schubert calculus 7
3. Lecture 3: Quantum marginals, and the Berenstein-Sjamaar/Ressayre theorems 8
3.1. Coadjoint orbits of Lie groups 9
3.2. Projections of coadjoint orbits 10
3.3. Relative Schubert calculus 10
4. Lecture 4: Schubert polynomials, Heckman's thesis 11
4.1. The barest minimum on cohomology rings 11
4.2. Cohomology rings of flag manifolds 12
4.3. Schubert polynomials 12
4.4. Heckman's thesis, and representation theory as a "quantum version" of coadjoint orbits 13
4.5. The state of knowledge of the Schur-Horn, Hermitian sum, and quantum marginal problems 14
References 14

1. LECTURE 1: THE SCHUR-HORN THEOREM AND THE GEL′FAND-CETLIN SYSTEM

Let V be a (finite-dimensional) complex vector space. For the purposes of these lectures, we will always work with the concrete spaces $\mathbb{C}^n$. A **state** S on V is an $n \times n$ Hermitian matrix with (necessarily real) eigenvalues between 0 and 1, adding up to 1. If S has only one nonzero eigenvalue (1), then S is a **pure state**, and otherwise it is a **mixed state**. We will never make use of the conditions on the eigenvalues, so we will usually just talk about Hermitian matrices.

1.1. **The Schur-Horn problem.** The *Schur-Horn problem* is the following: if S has the eigenvalues $\lambda = (\lambda_1 \geq \ldots \geq \lambda_n)$, what could the vector $(S_{11}, S_{22}, \ldots, S_{nn})$ of (also necessarily real) diagonal entries be? This is perhaps not an especially interesting problem, but it is the easiest analogue of the problem that really interests us, which will appear in the third and fourth lectures.

To see what shape the answer could take, let $\mathcal{O}_\lambda$ be the space of Hermitian matrices with spectrum λ, and consider the maps

$$U(n) \longrightarrow \mathcal{O}_\lambda \stackrel{diag}{\longrightarrow} \mathbb{R}^n$$

$$U \mapsto UD_\lambda U^{-1}, \quad S \mapsto (S_{11}, S_{22}, \ldots, S_{nn})$$

where D_λ is the diagonal matrix with entries λ, and $U(n)$ is the group of unitary $n \times n$ matrices.

This first map is *onto* – every matrix in $\mathcal{O}_\lambda$ is diagonalizable (to D_λ) by a unitary matrix. The unitary group is **compact**, meaning that it is closed (defined by equalities or weak inequalities, no strict inequalities) and bounded (no matrix entry has $|u_{ij}| > 1$). It is also connected. Consequently $\mathcal{O}_\lambda$ is compact and connected as well, as is its image under $diag$. So we have a very partial answer to the Schur-Horn problem – given λ, the set of possible diagonals is some compact, connected set in $\mathbb{R}^n$.

Let us compose further with a linear projection $\mathbb{R}^n \to \mathbb{R}$, i.e. $S \mapsto \mathrm{Tr}(SD_\mu)$ where D_μ is a diagonal matrix with entries $(\mu_1, \ldots, \mu_n)$. Now we have a compact, connected set in $\mathbb{R}$ – namely, a closed interval. (Put another way, a continuous real-valued function on a compact set attains its maximum and minimum values.) It is pretty easy to guess that the maximum is achieved when S is D_λ but with its entries permuted to be in the same order as D_μ's, and the minimum is achieved when the entries are permuted in the opposite order.

Lemma. *The critical points*[1] *of the* **Rayleigh trace** $\mathcal{O}_\lambda \to \mathbb{R}$, $S \mapsto \mathrm{Tr}(SD_\mu)$ *are where S commutes with D_μ, i.e. is block diagonal with blocks bounded by μ's multiplicities.*

This is pretty nearly a second-year calculus exercise.

Theorem (Schur's half of the theorem)**.** *The set $diag(\mathcal{O}_\lambda)$ is contained in the convex hull of the $n!$ permutations of the vector λ. (All those points are obviously in $diag(\mathcal{O}_\lambda)$, since each permutation of D_λ has the spectrum λ.)*

Proof. Let D_μ have diagonal entries 0 or 1, giving 2^n possibilities (though the all-0 or all-1 cases are dull). Applying the theorem above, we learn that any k diagonal entries of S have sum at most $\lambda_1 + \ldots + \lambda_k$, and at least $\lambda_{n-k+1} + \ldots + \lambda_n$. With some work one converts this description of the facets into the given description of the vertices. □

Horn proved the converse: *every* point in the convex hull is attained, so the set $diag(\mathcal{O}_\lambda)$ is determined by finitely many linear inequalities! We'll see a reason later to expect this sort of behavior.

[1]Meaning that first-order variations in S bring only second-order variations in $\mathrm{Tr}(SD_\mu)$. To move S by some small amount t, let H be any fixed Hermitian matrix, so $\exp(itH)$ is a family of unitary matrices, and look at $\exp(itH)S\exp(-itH)$.

1.2. **Flag manifolds.** It will be useful to have a more geometric understanding of the set $\mathcal{O}_\lambda$. If $\lambda = (1, \ldots, 1, 0, \ldots, 0)$ with k 1s, then this is pretty easy; $\mathcal{O}_\lambda$ can be corresponded with the space of k-dimensional subspaces of V, by $S \mapsto \mathrm{image}(S)$. This is called the **Grassmannian** $\mathrm{Gr}_k(V)$. The most familiar case is $k = 1$, where the Grassmannian is the **projective space** of lines through the origin in V, identified with pure states on V.

One reason that the Grassmannian is a nicer description than $\mathcal{O}_\lambda$ is that while we know how to apply unitary matrices to $\mathcal{O}_\lambda$, we know how to apply *arbitrary* invertible matrices to the Grassmannian (invertible transformations take k-planes to k-planes). In this sense it's a more symmetric description.

What if λ is general? Then for each k where $\lambda_k > \lambda_{k+1}$, we get a map $\mathcal{O}_\lambda \to \mathrm{Gr}_k(V)$. Putting them together, we get a correspondence of $\mathcal{O}_\lambda$ with the set of increasing chains $(V_{k_1} < V_{k_2} < \ldots < V)$ of subspaces, called (partial) **flags**. (If λ has no repeats, they are called **full flags**.) Again, the full general linear group $GL(n)$ of invertible matrices acts on these "flag manifolds".

1.3. **The Bruhat decomposition.** Given a smooth real-valued function f on a compact manifold M (like $\mathrm{Tr}(\bullet D_\mu)$ on $\mathcal{O}_\lambda$), we can study its gradient flow.[2] Under long-term gradient flow, almost every point in M falls down to the minimum value of f. But a few rare unlucky points settle down at other critical points (meaning, where $\Delta f = \vec{0}$).

The first thing to determine, then, is the set of critical points. Say we're lucky and there are only finitely many such.[3] A natural second step is to ask, given a critical point $p \in M$, which points m fall into p under long-term gradient flow? Call this set M_p, so we have the disjoint decomposition

$$M = \bigcup_p M_p, \qquad p \in \{m : \Delta f(m) = \vec{0}\}.$$

Take μ strictly decreasing (this will be to ensure that there are finitely many critical points). The "pure states" case of projective space, $\lambda = (1, 0, \ldots, 0)$, is easy to analyze. As in the lemma above, the critical points are the basis lines $L_k = (0, \ldots, 0, * \text{ in } k\text{th spot}, 0, \ldots, 0)$. Most lines $\mathbb{C} \cdot \vec{v} \leq V$ will flow downhill to the minimum eigenline (for D_μ), namely L_n. *However, if L is contained in the subspace $(*, *, \ldots, *, 0, \ldots, 0)$ of V where the last k coordinates are 0, then it will stay there during gradient flow.* Hence

$$M_{L_k} = \{\mathbb{C} \cdot \vec{v} : \text{the last nonzero entry of } \vec{v} \text{ is the } k\text{th}\}.$$

Note that $\vec{v}$ has *some* nonzero entry, since we're just using it to generate some 1-dimensional space in V. So we can rescale to get

$$M_{L_k} = \{\mathbb{C} \cdot \vec{v} : \text{the } k\text{th entry of } \vec{v} \text{ is } 1, \text{ with } 0\text{s after}\}$$

letting us view it as a copy of $\mathbb{C}^{k-1}$.

The next level of complexity is $\mathcal{O}_\lambda$ a k-Grassmannian, i.e. $\lambda = (1^{k \text{ times}}, 0^{n-k \text{ times}})$. The critical points correspond to the $\binom{n}{k}$ *coordinate* k-planes. To study a k-plane, we can pick a basis for it, make those vectors the rows of a $k \times n$ matrix, and then row-reduce it to put it in reduced row-echelon form; the result is then independent of the choice of basis.

[2]Technically, we need a metric on the manifold. There are a couple of natural choices, and they lead to the same result, so we won't detail this.

[3]If they're isolated, then compactness implies there are only finitely many.

Theorem. *Let $W \in \mathrm{Gr}_k(\mathbb{C}^n)$. Then under the gradient flow of the Rayleigh trace $f : W \mapsto \mathrm{Tr}(D_\mu \pi_W)$, the point W flows down to the coordinate k-plane whose k coordinates correspond to the columns of the k pivots in W's reduced row-echelon matrix. (Here π_W is the orthogonal projection onto the subspace W.)*

One can determine this k-element subset of $\{1,\ldots,n\}$ as follows: look at the subspaces

$$0 = (W \cap 0) \leq (W \cap \mathbb{C}^1) \leq (W \cap \mathbb{C}^2) \leq \cdots \leq (W \cap \mathbb{C}^n) = W$$

where $\mathbb{C}^i$ denotes the coordinate subspace $\{(*,\ldots,*,0,\ldots,0)\}$, then compute their dimensions,

$$0 \leq \dim(W \cap \mathbb{C}^1) \leq \dim(W \cap \mathbb{C}^2) \leq \cdots \leq \dim(W \cap \mathbb{C}^n) = k.$$

These numbers jump by either 0 or 1, and jump by 1 exactly k times out of n. This gives the same subset as the reduced row-echelon matrix calculation.

Note again that if we fix the locations of the pivots, the remaining freedom in the reduced row-echelon matrix is just to choose some independent complex numbers. So again, the decomposition is into a union of vector spaces.

It is interesting to note that any two such reduced row-echelon matrices with the same pivots are related by left multiplication by **lower unipotent matrices**, meaning lower triangular matrices with 1s on the diagonal. So this decomposition of the Grassmannian comes as orbits of a subgroup of the general linear group (but not of the unitary group, which it intersects trivially).[4]

Given a subset $I \subseteq \{1,\ldots,n\}$ of size k, define the **Bruhat cell** $\mathrm{Gr}_k(\mathbb{C}^n)^\circ_I$ to be the set of W that flow down to the coordinate k-plane using the coordinates I, or equivalently, the set of W whose reduced row-echelon matrices have pivots in columns I. Its closure (points not necessarily in it, but reachable by taking limits) $\mathrm{Gr}_k(\mathbb{C}^n)_I$ is called a **Schubert variety**, and is not usually smooth nearby the added points. The codimension (meaning, number of degrees of freedom loft) of $\mathrm{Gr}_k(V)_I$ is the number of pairs $\{(a,b) : a < b, a \notin I, b \in I\}$.

Corollary (Hersch-Zwahlen 1962). *If $W \in \mathrm{Gr}_k(\mathbb{C}^n)_I$, and π_W is the orthogonal projection onto W, then $\mathrm{Tr}(\pi_W D_\mu) \geq \sum_{i \in I} \mu_i$.*

Proof. It is enough to prove it for $W \in \mathrm{Gr}_k(\mathbb{C}^n)^\circ_I$, since weak inequalities continue to hold in limits. As we follow gradient flow downward, the Rayleigh trace only decreases, until it reaches its limiting value $\sum_{i \in I} \mu_i$. □

The final case, of general λ, is not too bad because the flag manifold $\mathcal{O}_\lambda$ includes into a product of Grassmannians $\prod_{k:\, \lambda_k > \lambda_{k+1}} \mathrm{Gr}_k(V)$, and we can infer the gradient flow and the Bruhat decompositions from there. Instead of coordinate subspaces indexed by subsets of $\{1,\ldots,n\}$, we have coordinate flags indexed by chains of subsets. If λ has all distinct eigenvalues, so that $\mathcal{O}_\lambda$ is the manifold of full flags, then the maximal chains correspond exactly to permutations of $\{1,\ldots,n\}$. The analogue of picking a basis of the k-plane, and to do row operations, is to build a basis of the whole space by gradually extending it to larger subspaces in the flag, and only allow downward row operations. Then one must remember both the columns and rows that the pivots appear in, which again is encoded by a permutation.

[4]In the unitary setting, it is more natural to pick *orthonormal* bases of the k-planes, but then there is no simple analogue of reduced row-echelon form.

1.4. **The Gel′fand-Cetlin system.** There is a refinement of the Schur-Horn problem: given the spectrum λ of a Hermitian matrix S, what possible diagonals μ can occur, *with what probability density?*

To answer this, we factor the map $\mathrm{diag} : \mathcal{O}_\lambda \to \mathbb{R}^n$ into

$$\mathcal{O}_\lambda \to \mathbb{R}^{\binom{n}{2}} \to \mathbb{R}^n$$

where the first map takes

$$S \mapsto (\text{the } i\text{th eigenvalue of the upper left } j \times j \text{ square})_{i \leq j}$$

recorded as a triangle with j entries in the jth row, the bottom row containing the constant λ. The second map takes the differences in the row sums (the traces).

Theorem.
(1) *(Weyl) The entries in row j are interspersed between those in row $j+1$, for $j = 1, \ldots, n-1$. These linear inequalities define* **Gel′fand-Cetlin patterns** *(used in representation theory).*
(2) *(Guillemin-Sternberg) These are the only conditions; every Gel'fand-Cetlin pattern with λ as its bottom row is in the image.*
The $U(n)$-invariant measure on $\mathcal{O}_\lambda$ pushes forward to (a constant multiple of) Lebesgue measure on the Gel′fand-Cetlin polytope.

Hence the probability density of obtaining μ as the diagonal is given by the volume of the polytope of Gel′fand-Cetlin patterns with λ on the bottom, and μ as the differences of the row sums (up to an unimportant global constant).

Note that the words "upper left" completely break the permutation-invariance that the Schur-Horn question, and its answer, possess. For example, the Gel′fand-Cetlin polytopes for μ and $w \cdot \mu$ (for w a permutation in S_n) are usually not isomorphic, though they have the same volume.

There's a $1:1$ correspondence between *integral* Gel′fand-Cetlin patterns (with given bottom row and row sums) and semistandard Young tableaux (with given shape and content). It's much easier to talk about the volume of a polytope of real GC patterns than it is to give a definition of nonintegral semistandard Young tableaux, so I won't really focus on them.

2. Lecture 2: The Weyl-Horn problem and Schubert calculus

We turn now to a richer problem, even older than the Schur-Horn problem, but only wholly solved in the 1990s. In 1912 Weyl asked: given two Hermitian matrices H_λ and H_μ with spectra λ and μ, what can the possible spectra ν be of their sum?

There are a couple of simple hints that this, like the Schur-Horn problem, might have a polyhedral answer. One hint is that if H_λ is taken (without loss of generality) to be diagonal, and λ's eigenvalues are so spread apart that H_μ may be considered a tiny perturbation, then the spectrum of $H_\lambda + H_\mu$ is really determined by the diagonal entries of H_μ. So in this limiting case, Weyl's problem reduces[5] to the Schur-Horn problem.

[5]Approximately. Or so it would seem; but in fact the corresponding Schur-Horn polytope gives *exactly* the right answer whenever all $\lambda_i - \lambda_{i+1}$ are more than $\mu_1 - \mu_n$.

The other hint is that $\nu_1 \leq \lambda_1 + \mu_1$. Proof: the largest eigenvalue of a Hermitian matrix H is the maximum value of $\langle H\vec{v}|\vec{v}\rangle$ taken over unit vectors $\vec{v}$ (as discussed in the last lecture). So

$$\begin{aligned}
\nu_1 &= \max_{\vec{v}} \langle (H_\lambda + H_\mu)\vec{v}|\vec{v}\rangle \\
&= \max_{\vec{v}} \left(\langle H_\lambda\vec{v}|\vec{v}\rangle + \langle H_\mu\vec{v}|\vec{v}\rangle\right) \\
&= \max_{\vec{v}_1,\vec{v}_2:\ \vec{v}_1=\vec{v}_2} \left(\langle H_\lambda\vec{v}_1|\vec{v}_1\rangle + \langle H_\mu\vec{v}_2|\vec{v}_2\rangle\right) \\
&\leq \max_{\vec{v}_1,\vec{v}_2} \left(\langle H_\lambda\vec{v}_1|\vec{v}_1\rangle + \langle H_\mu\vec{v}_2|\vec{v}_2\rangle\right) \\
&= \max_{\vec{v}_1} \langle H_\lambda\vec{v}_1|\vec{v}_1\rangle + \max_{\vec{v}_2} \langle H_\mu\vec{v}_2|\vec{v}_2\rangle \\
&= \lambda_1 + \mu_1.
\end{aligned}$$

QED.

2.1. **The Johnson/Klyachko/Helmke-Rosenthal/Totaro inequalities.**

We generalize this latter argument in two ways. One is to go beyond 1-dimensional spaces to k-dimensional, which leads to inequalities about sums of k eigenvalues from each of λ, μ, ν instead of just one. But the much subtler ingredient is to maximize not over all subspaces, but just ones in certain Schubert varieties.

Let $\mathrm{Gr}_k(V)_{H,I}$ denote the Schubert variety in $\mathrm{Gr}_k(V)$ defined using the Rayleigh trace $\mathrm{Tr}(\bullet H)$, where H is no longer necessarily diagonal. This can be equally well described by adding H's eigenspaces for its j largest eigenvalues to obtain an increasing chain of subspaces, then determine the Bruhat cell of $W \in \mathrm{Gr}_k(V)$ by fixing the jumps in the dimension of the intersections of W with this chain. [6] The Hersch-Zwahlen theorem has two analogues, for $W \in \mathrm{Gr}_k(V)_{H,I} \cap \mathrm{Gr}_k(V)_{-H,J}$:

$$\sum_{i\in I} \mu_i \leq \mathrm{Tr}(H\pi_W) \leq \sum_{j\in J} \mu_{n+1-j}.$$

Theorem (Johnson 1979 (unpublished thesis), Klyachko ~1993, Totaro 1994, Helmke-Rosenthal 1995)**.** *Let* $H_\nu = H_\lambda + H_\mu$*, with spectrum* ν*. Let* I, J, K *be three* k*-element subsets of* $1, \ldots, n$*, and assume that the intersection* $\mathrm{Gr}_k(V)_{H_\lambda,I} \cap \mathrm{Gr}_k(V)_{H_\mu,J} \cap \mathrm{Gr}_k(V)_{-H_\nu,K}$ *is nonempty. Then*

$$\sum_I \lambda_i + \sum_J \mu_j \leq \sum_K \nu_{n+1-k}.$$

(Any similar inequality with $\geq$ *can be turned into one of these, using the equality* $\mathrm{Tr}(H_\lambda) + \mathrm{Tr}(H_\mu) = \mathrm{Tr}(H_\nu)$*.)*

Proof. Let W be a point in the intersection. Then

$$\mathrm{Tr}((H_\lambda + H_\mu)\pi_W) = \mathrm{Tr}(H_\nu\pi_W) = -\mathrm{Tr}((-H_\nu)\pi_W).$$

Using the Hersch-Zwahlen inequality, we get

$$\mathrm{Tr}(H_\lambda\pi_W) + \mathrm{Tr}(H_\mu\pi_W) \geq \sum_I \lambda_i + \sum_J \mu_j, \qquad \mathrm{Tr}((-H_\nu)\pi_W) \geq \sum_K -\nu_{n+1-k}.$$

[6] If H has repeated eigenvalues, perturb it to some H' commuting with H where H' has simple spectrum. Only certain Schubert varieties will be independent of this choice. It's a worrisome-looking annoyance but doesn't actually affect anything.

(The eigenvalues of $-H_\nu$, *in decreasing order*, are $-\nu_n \geq \ldots - \nu_1$, hence the $n+1-k$.) Combining these, we get the desired inequality. $\square$

Note that while the proof uses W, the statement only uses the existence of W. Many partial results on Weyl's problem were given by people who constructed explicit subspaces W from the flags corresponding to H_λ, H_μ, H_ν. For example, Weyl's inequality $\lambda_i + \mu_j \geq \nu_{i+j-1}$ comes from adding λ's i-plane, μ's j-plane, and ν's $(n-i-j-1)$-plane, and picking W to be a hyperplane containing that sum.

Horn studied the map $\mathcal{O}_\lambda \times \mathcal{O}_\mu \to \mathbb{R}^n$, taking (H_λ, H_μ) to the spectrum of their sum. *Away from the walls where ν has repeated eigenvalues,* this map is smooth and one can do multivariable calculus to determine its boundary. Horn used this to show that the image is locally polyhedral, and that the inequalities all have the approximate form of the theorem above. He also gave in 1962 an explicit conjecture for a set of triples (I, J, K), that Terry Tao and I proved in 1999.

2.2. **Schubert calculus.** The interesting bit about the theorem above is that one can often guarantee that these Schubert varieties intersect *for purely topological reasons*. Think of two circles drawn on the surface of a doughnut, one passing through the hole and the other going around the outside. One can wiggle them around on the doughnut but they cannot be thereby made to avoid one another.

Theorem (Kleiman 1973)**.** *Let I, J, K be three k-element subsets of $1, \ldots, n$, and let g_1, g_2, g_3 vary over elements of $GL(V)$. Then for almost all g_1, g_2, g_3, the intersection $(g_1 \cdot Gr_k(V)_I) \cap (g_2 \cdot Gr_k(V)_J) \cap (g_3 \cdot Gr_k(V)_K)$ has the same dimension, and if finite, the same number of points.*

Its codimension is the sum of the three codimensions.

Theorem.
1. *(Horn 1962) It is enough to consider triples (I, J, K) such that this intersection has finitely many points; the other Hersch-Zwahlen inequalities are implied by those. (This occurs when the sum of the three codimensions is $k(n-k)$.)*
2. *(Klyachko ~1993) These inequalities are not only necessary, but sufficient, for the existence of a triple $(H_\lambda, H_\mu, H_\nu)$.*
3. *(Belkale 1999) It is enough to consider triples (I, J, K) such that this intersection has exactly one point.*
4. *(Knutson-Tao-Woodward 2004) Belkale's list is irredundant.*
5. *(Bercovici et al. 2009) For each inequality on Belkale's list, the subspace W can be constructed from the eigenspaces of H_λ, H_μ, H_ν using sums and intersections.*
   (So Weyl's technique would have found them all, eventually!)

The question of computing the number of points in these (generic) intersections is called **Schubert calculus on Grassmannians**. Schubert's calculus motivated the development of "cohomology rings" of topological spaces; the Schubert varieties provide a basis for the cohomology ring of the Grassmannian, and these intersection numbers are then the coefficients in the multiplication.

Schubert's calculus doesn't let one calculate these positive numbers in a positive way. Today there are many positive ways to calculate them, like the Littlewood-Richardson rule; my favorite version is with the "puzzles" in [KnTao03].

Horn's conjectured list of triples was specified in a very curious *recursive* manner, reducing the study of Hermitian matrices of size n with arbitrary eigenvalues to the study of Hermitian matrices of each size $k < n$ with *integer* eigenvalues. This has only received a really satisfactory explanation more recently, in work of Belkale, Purbhoo, and Sottile.

Tao and I approached this problem by starting with the harder one – what is the probability measure on the space of ν? – which has a Gel′fand-Cetlin-like answer due to Berenstein and Zelevinsky (1991), though was actually found first by Johnson (1979). Linear programming arguments led us to the puzzles, and comparison with Klyachko's theorem suggested that the puzzles were actually computing Schubert calculus.

3. Lecture 3: Quantum marginals, and the Berenstein-Sjamaar/Ressayre theorems

The **quantum marginals** problem is the following: given a state H_{AB} on $A \otimes B$ with known spectrum, whose partial traces H_A, H_B are thus states on A,B, what can the pair of spectra of the **marginals** H_A, H_B be?

I'll first give a Helmke-Rosenthal-style family of inequalities. (As far as I know this proof does not appear in the literature. Some discussion of the history will follow.)

Theorem 1. *Let $H_{A\otimes B}$ be a Hermitian operator on $A\otimes B$, with marginals H_A, H_B. Assume that there exist subspaces $W_A \leq A$, $W_B \leq B$ of dimensions i, j, such that*

(1) $W_A \in Gr_i(A)_{H_A,I}$ *where* $I \subseteq \{1,\ldots,\dim A\}, |I| = i$
(2) $W_B \in Gr_j(B)_{H_B,J}$ *where* $J \subseteq \{1,\ldots,\dim B\}, |J| = j$
(3) $W_A\otimes W_B \in Gr_{ij}(A\otimes B)_{-H_{AB},K}$ *where* $K \subseteq \{1,\ldots,\dim A \dim B\}, |K| = ij$
(4) $(W_A\otimes B) + (A\otimes W_B) \in Gr_{i\dim B + j \dim A - ij}(A\otimes B)_{-H_{AB},L}$
where $L \subseteq \{1,\ldots,\dim A\dim B\}, |L| = i\dim B + j\dim A - ij$.

(Since $(W_A\otimes B) + (A\otimes W_B) \geq W_A\otimes W_B$, it is automatic that $L \supseteq K$.)

Let $\varepsilon_{Q,p}$ denote the pth largest eigenvalue of H_Q. Then

$$\begin{aligned}\sum_I \varepsilon_{A,i} + \sum_J \varepsilon_{B,j} &\leq \sum_K \varepsilon_{AB,\dim A\dim B+1-k} + \sum_L \varepsilon_{AB,\dim A\dim B+1-l}\\ &= \sum_{m=1}^{\dim A\dim B} \varepsilon_{AB,\dim A\dim B+1-m}\begin{cases}2 & \text{if } m\in L\\ 1 & \text{if } m\in K, m\notin L\\ 0 & \text{if } m\notin K.\end{cases}\end{aligned}$$

Proof.

$$\begin{aligned}\sum_I \varepsilon_{A,i} + \sum_J \varepsilon_{B,j} &\leq \mathrm{Tr}_A(H_A\pi_{W_A}) + \mathrm{Tr}_B(H_B\pi_{W_B})\\ &= \mathrm{Tr}_{A\otimes B}(H_{AB}(\pi_{W_A}\otimes 1_B)) + \mathrm{Tr}_{A\otimes B}(H_{AB}(1_A\otimes\pi_{W_B}))\\ &= \mathrm{Tr}_{A\otimes B}(H_{AB}(\pi_{W_A}\otimes 1_B + 1_A\otimes\pi_{W_B}))\\ &= \mathrm{Tr}_{A\otimes B}(H_{AB}(\pi_{W_A\otimes B + A\otimes W_B} + \pi_{A\otimes B}))\\ &= \mathrm{Tr}_{A\otimes B}(H_{AB}\,\pi_{W_A\otimes B + A\otimes W_B}) + \mathrm{Tr}_{A\otimes B}(H_{AB}\,\pi_{A\otimes B})\\ &\leq \sum_K \varepsilon_{AB,\dim A\dim B+1-k} + \sum_L \varepsilon_{AB,\dim A\dim B+1-l}\end{aligned}$$

The other equality is obvious. $\square$

It is easy to generalize this to $A_1 \otimes \cdots \otimes A_m$, in which case the left side looks much the same while the right side has coefficients between 1 and m.

Unfortunately, these are not all the conditions, but the others are all also linear inequalities, described by Klyachko (2004) in the general case. We get some more using chains of subspaces.

Theorem 2. *Given chains of subspaces* $0 < A_1 < A_2 < \ldots < A_{d_A} = A$ *and* $0 < B_1 < B_2 < \ldots < B_{d_B} = B$, *we can define a chain* $0 < C_2 < \ldots < C_{d_A+d_B} = A \otimes B$ *by* $C_k = \sum_{i+j=k} A_i \otimes B_j$, *of dimension* $c_k = \sum_{i+j \leq k} \dim(A_i/A_{i-1}) \dim(B_j/B_{j-1})$.

Let H_{AB} *be a state on* $A \otimes B$, *with marginals* H_A, H_B. *Assume there exist chains of subspaces* $(A_i), (B_j)$ *and chains of subsets* $(P_{i=1,\ldots,d_A} \subseteq \{1, \ldots, \dim A\})$, $(Q_{j=1,\ldots,d_B} \subseteq \{1, \ldots, \dim B\})$, $(R_{k=1,\ldots,d_A+d_B}) \subseteq \{1, \ldots, \dim A \dim B\})$, *such that*

- $A_i \in Gr_{a_i}(A)_{H_A, P_i}$ *where* $|P_i| = \dim A_i$
- $B_j \in Gr_{b_j}(B)_{H_B, Q_j}$ *where* $|Q_i| = \dim B_i$
- $C_k \in Gr_{c_k}(A \otimes B)_{-H_{AB}, R_k}$.

Then

$$\sum_{i=1}^{d_A} \sum_{p \in P_i} \varepsilon_{A,p} + \sum_{j=1}^{d_B} \sum_{q \in Q_j} \varepsilon_{A,q} \quad \leq \quad \sum_{k=2}^{d_A+d_B} \sum_{q \in Q_k} \varepsilon_{AB, \dim A \dim B + 1 - q}$$

Proof. The proof is the same, starting from the equality

$$\sum_i (\pi_{A_i} \otimes 1_B) + \sum_j (1_A \otimes \pi_{B_j}) = \sum_k \pi_{C_k}$$

instead of

$$\pi_{W_A} \otimes 1_B + 1_A \otimes \pi_{W_B} = \pi_{W_A \otimes B + A \otimes W_B} + \pi_{A \otimes B}$$

in the previous theorem. ∎

In the rest of this lecture we consider the question: what is the family of problems that we expect to have such lovely, polyhedral solutions? This requires developing some technology.

In the next lecture we take up the problem of determining when these magic subspaces W exist. (Unfortunately, there is in general no known analogue of the puzzles we had in the Horn problem.)

3.1. **Coadjoint orbits of Lie groups.** Let G be a "compact connected Lie group", which you may think of a set of unitary matrices, closed under multiplication or taking limits, all connected to the identity. Then G acts on itself by conjugation, holding the identity $1 \in G$ in place. Therefore G acts on its "Lie algebra" $\mathfrak{g} := T_1 G$, and also on the dual vector space $\mathfrak{g}^*$. We can also use a G-invariant metric[7] to identify $\mathfrak{g}$ with $\mathfrak{g}^*$, which is convenient since $\mathfrak{g}$ is easier to comprehend. (But $\mathfrak{g}^*$ will be more natural for our application to come.)

[7] since G is compact, so we can do integrals over it and get finite numbers; in particular we can average any randomly chosen metric to get a G-invariant one

For example, if G is actually $U(n)$, then the Lie algebra is

$$\begin{aligned}\mathfrak{g} &= \{M : (1+\varepsilon M) \text{ satisfies } U(n)\text{'s equation to first order}\} \\ &= \{M : (1+\varepsilon M)(1+\varepsilon M)^* = 1 + O(\varepsilon^2)\} \\ &= \{M : M + M^* = 0\}\end{aligned}$$

AKA skew-Hermitian matrices. We'll identify its dual with Hermitian matrices using the form $\langle H|S\rangle := i\mathrm{Tr}(HS)$.

Under this identification, the **coadjoint** action of G on $\mathfrak{g}^*$ is just the usual one of conjugating Hermitian matrices by unitary matrices. The orbits of this action are just the $\{\mathcal{O}_\lambda\}$ we were studying in the first lecture.

We already know how to list the orbits of $U(n)$ on Hermitian matrices: there is exactly one for each weakly decreasing list λ of real numbers. More specifically, any Hermitian matrix can be diagonalized, then put in decreasing order, and the result is unique. It turns out there is a similar result for any compact connected Lie group; any element of $\mathfrak{g}^*$ can be conjugated into $\mathfrak{t}^*$, and then uniquely into the "positive Weyl chamber", a polyhedral cone. We will only need this result for G a product of unitary groups.

3.2. **Projections of coadjoint orbits.**

Now let H be a compact, connected subgroup of G. Then $\mathfrak{h} \leq \mathfrak{g}$, so we get a linear projection $\mathfrak{g}^* \twoheadrightarrow \mathfrak{h}^*$. Given a G-coadjoint orbit $\mathcal{O}_\lambda \subseteq \mathfrak{g}^*$, its projection to $\mathfrak{h}^*$ will be H-invariant, so a union of H-coadjoint orbits.

(1) Which ones are in the image?
(2) What is the induced measure on the image?

Example: $G = U(n)$, $H = U(1)^n$, the diagonal unitary matrices. Then the map $\mathfrak{g}^* \to \mathfrak{h}^*$ becomes, under the identification with Hermitians, the map diag from the first lecture. So the first question becomes the Schur-Horn problem, and the second is answered by the Gel'fand-Cetlin technology.

Example: $G = U(n) \times U(n)$, $H = U(n)$ sitting inside diagonally $\{(U_1, U_2) : U_1 = U_2\}$. Then the map $\mathfrak{g}^* \to \mathfrak{h}^*$ is the map $(H_\lambda, H_\mu) \mapsto H_\lambda + H_\mu$. So the first question becomes Weyl's problem (as first answered by Klyachko), and the second question is answered by the Johnson/Berenstein-Zelevinsky technology.

Example: $G = U(mn)$, $H = U(m) \times U(n)$.[8] Then the map $\mathfrak{u}(mn)^* \to \mathfrak{u}(m)^* \oplus \mathfrak{u}(n)^*$ takes a state of Alice+Bob to the pair (state of Alice, state of Bob), by tracing over the two components, the **quantum marginals**. No useful answer is known to the second question.

With modern technology (the Guillemin-Sternberg-Kirwan convexity theorem), it is easy to prove that the first question always has a polyhedral answer. In general, the second question (about measures) is easy to answer *as an alternating sum*, which makes it very hard to determine its support. The first question received a general answer by Berenstein and Sjamaar, following Klyachko's lead, and I will describe it below.

3.3. **Relative Schubert calculus.**

There is another problem whose input is a pair $(G \geq H)$ of Lie groups.

[8]Technically, there is only a map $H \to G$, not an inclusion, but that's good enough. It just says that the map $\mathfrak{g}^* \to \mathfrak{h}^*$ is not onto, which we didn't use for anything.

We saw already that for $G = U(n)$, the coadjoint orbits of G have a Bruhat decomposition, defined using gradient flow. There is an analogous decomposition for the coadjoint orbits of any compact Lie group. Separately, one can prove that each smallest-dimensional orbit of H on a coadjoint orbit $\mathcal{O}_1$ of G is itself naturally isomorphic to some coadjoint orbit $\mathcal{O}_2$ of H.

(1) For which pairs $(\mathcal{O}_1)_\pi$, $(\mathcal{O}_2)_\rho$ of Schubert varieties in these two coadjoint orbits, is a generic translate $g \cdot (\mathcal{O}_1)_\pi \cap h \cdot (\mathcal{O}_2)_\rho$ nonempty?
(2) Kleiman's theorem lets one figure out the dimension of the intersection. When it is 0-dimensional, how many elements does it have?

As before, the second question has a solution that is not obviously positive, and hence is of limited use in answering the first. We'll discuss this in the next lecture. The first was given some very nice necessary conditions and (separately) sufficient conditions by Purbhoo.

Theorem (Berenstein-Sjamaar 1998). *If one can solve (a slight generalization of) the first question, then one can give a complete (albeit redundant) list of the linear inequalities determining the image of $\mathcal{O}_\lambda \subseteq \mathfrak{g}^*$ projected to $\mathfrak{h}^*$.*

With this, one can e.g. recover the Schur-Horn result and Klyachko's 1993 results (concerning spectra of sums).

Klyachko (2004) applied their results to the quantum marginal problem, giving a complete (but complicated) list of inequalities. Ressayre (2007) gave an improved version of Berenstein-Sjamaar that automatically computes the minimal list of inequalities. Guillemin and Sjamaar (2005; section 3.8) have given a Hersch-Zwahlen-style proof of the general Berenstein-Sjamaar theorem, so the proof in theorem 1 is not much of a surprise.

In the quantum marginals problem, we still have a question: for which quadruples I, J, K, L of Schubert conditions is there a unique pair of subspaces (W, V) satisfying those conditions?

4. Lecture 4: Schubert polynomials, Heckman's thesis

4.1. **The barest minimum on cohomology rings.** *Algebraic topology* is about studying continuous maps of topological spaces using algebra, i.e. integers, or elements of more general groups and rings. Since it's hard for an integer to vary continuously, algebraic topology is mostly concerned with properties of continuous maps that don't vary under deformation; a good example is the winding number of a function $f : S^1 \to S^1$ (where S^1 denotes the circle.

Schubert's calculational techniques were not rigorous in his day, and one of Hilbert's problems was to make them so. That was answered by the development of *cohomology rings* of spaces. Every topological space X has an associated cohomology ring $H^*(X)$, and every continuous map $f : X \to Y$ has an associated map $f^* : H^*(Y) \to H^*(X)$. (The "co" is because this map is backwards.)

$H^*(X)$ can be a horribly complicated commutative ring, but things are very nice if X has a decomposition into even-real-dimensional vector spaces (think the Bruhat decomposition); then $H^*(X)$ is a vector space with a natural basis indexed by the pieces in the decomposition. If $X = \mathcal{O}_\lambda$ then these are called the *Schubert classes* in $H^*(X)$.

Just to keep the threads together, I'll restate the input to the Berenstein-Sjamaar/Ressayre theorems. Let $\mathcal{O}_\lambda \to \mathcal{O}_\mu$ is an inclusion of H's coadjoint orbit into G's, and let S_π be a Schubert class in $H^*(\mathcal{O}_\mu)$. Then its image (backwards) in $H^*(\mathcal{O}_\lambda)$ is some linear combination of Schubert classes in $H^*(\mathcal{O}_\lambda)$. What are the coefficients? This is what I'll call the **Schubert restriction problem** hereafter.

Note that any space X has a diagonal inclusion $X \to X \times X$, so there's a map $H^*(X\times X) \to H^*(X)$. This is eventually the source of the multiplicative structure on $H^*(X)$, and why people prefer cohomology to homology (since they prefer multiplication to comultiplication). It means also that in Horn's problem, one is interested in the coefficients of multiplication on $H^*(Gr_k(\mathbb{C}^n))$, which is solved by puzzles and many other things.

4.2. **Cohomology rings of flag manifolds.** Borel's theorem gives a presentation of $H^*(\mathcal{O}_\lambda)$ when λ is generic; it is the ring of polynomial functions on $\mathfrak{t}$, modulo Weyl-group-symmetric polynomials. In the $U(n)$ case it comes to

$$H^*(\mathcal{O}_\lambda) \cong \mathbb{Z}[x_1, \ldots, x_n]/\langle \text{symmetric polynomials with no constant coefficient}\rangle$$

e.g. in the $U(3)$ case, $H^*(\mathcal{O}_\lambda) \cong \mathbb{Z}[x_1,x_2,x_3]/\langle x_1+x_2+x_3, x_1x_2+x_1x_3+x_2x_3, x_1x_2x_3\rangle$.

In geometric terms, recall that $\mathcal{O}_\lambda$ is the manifold of full flags $(0 < W_1 < W_2 < \ldots < W_n = \mathbb{C}^n)$, and if λ' is not generic, then $\mathcal{O}_{\lambda'}$ is a manifold of partial flags, skipping some subspaces. In particular there is a forgetful map $\mathcal{O}_\lambda \twoheadrightarrow \mathcal{O}_{\lambda'}$, which happens to induce an inclusion

$$H^*(\mathcal{O}_{\lambda'}) \hookrightarrow H^*(\mathcal{O}_\lambda) \cong \mathbb{Z}[x_1, \ldots, x_n]/\langle \text{symmetrics w/o constant coefficient}\rangle$$

whose image is generated by polynomials that are symmetric in two variables if the corresponding elements of λ' coincide.

In the quantum marginals case, the map $U(m) \times U(n) \to U(mn)$ induces the maps

$$T^m \times T^n \to T^{mn}, \qquad (D,E) \mapsto D \otimes E$$

$$\mathfrak{t}^m \oplus \mathfrak{t}^n \to \mathfrak{t}^{mn} \qquad (D,E) \mapsto (D\otimes 1_n) + (1_m \otimes E)$$

$$(\mathfrak{t}^m)^* \oplus (\mathfrak{t}^n)^* \longleftarrow (\mathfrak{t}^{mn})^*, \qquad z_{ij} \mapsto x_i + y_j$$

which plugged into Borel's theorem, gives

$$\mathbb{Z}[z_{11}, \ldots, z_{mn}]/\langle \text{sym}\rangle \to \mathbb{Z}[x_1, \ldots, x_m]/\langle \text{sym}\rangle \otimes \mathbb{Z}[y_1, \ldots, x_n]/\langle \text{sym}\rangle, \qquad z_{ij} \mapsto x_i + y_j.$$

4.3. **Schubert polynomials.** Borel's presentation doesn't help one answer the Schubert restriction problem, because Borel doesn't tell you where to find the Schubert classes inside his presentation. Note that it's kind of annoying to specify an element of a quotient ring; it would be nice if each Schubert class had an actual polynomial associated, even though it doesn't look natural to privilege one choice over another.

In fact there is a natural choice, found by Lascoux and Schützenberger in 1982, called **Schubert polynomials**. I won't explain the multiple reasons why their choice is natural (though see [KnMil05]), but I will give a nice formula for them.

Define a **pipe dream** as a tiling of the fourth quadrant by two kinds of tiles: crosses and elbows. There are two conditions; there should be only finitely many crosses, and no two pipes should cross twice. To a pipe dream P whose last cross occurs before the nth antidiagonal, one can associate a permutation $\mathrm{perm}(P)$ of $\{1, \ldots, n\}$: write the numbers down the left column, and convey them along the pipes to the top. Then read off the numbers.

To P we also associate a monomial $\prod_{+\in P} x_{row} \in \mathbb{Z}[x_1,\ldots,x_n]$. Then the formula (very far from Lascoux and Schützenberger's definition) is

$$S_\pi = \sum_{P:\ \mathrm{perm}(P)=\pi} \prod_{+\in P} x_{row}.$$

(Note that if $\mathrm{perm}(P)$ only moves the numbers up to n, then it is easy to see that P has no crosses after the nth antidiagonal, so this sum is finite.) For example, 132 has two pipe dreams, and $S_{132} = x_1 + x_2$. The papers [BB, KnMil05] give ways of producing all the pipe dreams for a given permutation.

Theorem. *If $\pi(i) < \pi(i+1)$, i.e. π has an* **ascent** *at i, then S_π is symmetric in x_i, x_{i+1}. (And only if.)*

In particular, the last variables x_d occurring in S_π is for d the largest **descent** *of π.*

If π has only one descent, after the kth place, then S_π is the **Schur polynomial** *Schur_λ in k variables associated to the partition λ whose ith smallest row has $\pi(i)-i$ boxes. This arises also as the character of the irreducible representation V_λ of $U(k)$ (and in the representation theory of the symmetric group, and elsewhere); the pipe dreams can be easily corresponded with semistandard Young tableaux in this case.*

4.4. **Heckman's thesis, and representation theory as a "quantum version" of coadjoint orbits.** The possible spectra of Hermitian matrices correspond uniquely to weakly decreasing lists of real numbers. The irreducible representations of $U(n)$ correspond to weakly decreasing lists of *integers*.[9] In this and many other senses, one should think of irreducible representations as the "quantum analogues" of coadjoint orbits.

To see a little of this, consider the "fundamental cases" where $\lambda = (1^k, 0^{n-k})$. Then the corresponding coadjoint orbit is the Grassmannian of k-planes (or rather, rank k projections), and the corresponding irrep is $\mathrm{Alt}^k\mathbb{C}^n$. These are famously related by the Plücker embedding $\mathrm{Gr}_k(\mathbb{C}^n) \to \mathbb{P}(\mathrm{Alt}^k\mathbb{C}^n)$, $W \mapsto \mathrm{Alt}^k W$. (In coordinates, we pick a basis of W and wedge it together.)

More generally, we can construct the irrep V_λ by looking in the very big representation

$$V^{\otimes(\lambda_1-\lambda_2)} \otimes (\mathrm{Alt}^2 V)^{\otimes(\lambda_2-\lambda_3)} \otimes \cdots \otimes (\mathrm{Alt}^n V)^{\otimes(\lambda_n - 0)}$$

taking the $U(n)$-orbit of the high weight vector, and either projectivizing it (to get the coadjoint orbit) or taking its linear span (to get the irrep).

G. Heckman's thesis included the following:

Theorem. *(Heckman 1980?) Let $G \geq H$ be a pair of Lie groups, and λ, μ dominant weights of them.*

1. *Let $f_\lambda^\mu(n) =$ the multiplicity of $V_{n\mu}$ in the restriction of $V_{n\lambda}$ from G to H.*
2. *Let $\mathcal{M}_\lambda^\mu$ be the preimage under the linear projection $\mathfrak{g}^* \to \mathfrak{h}^*$ of $\mathcal{O}_\mu$ in $\mathcal{O}_\lambda$, divided by H.*

[9] It is frequently asserted that they correspond to partitions of height at most n, i.e. weakly decreasing lists of *naturals*. Those are the ones that extend continuously to noninvertible matrices, leaving out e.g. the representation $U \mapsto \det(U)^{-1}$. Any one of these more general representations can be tensored with $U \mapsto \det(U)^N$ for some large N in order to fall into this smaller set, which is why people get away with this.

Then for sufficiently divisible n, $f_\lambda^\mu(n)$ *is a polynomial, whose leading term* cn^d *has*

$$d = \dim \mathcal{M}_\lambda^\mu, \quad c = \operatorname{vol} \mathcal{M}_\lambda^\mu.$$

The subleading terms in f_λ^μ can be read from subtler invariants of the space $\mathcal{M}_\lambda^\mu$.

Corollary.
- *If* V_μ *appears in* V_λ*, then this* $\mathcal{M}$ *is nonempty.*
- *If this* $\mathcal{M}$ *is nonempty, then* $V_{N\mu}$ *appears in* $V_{N\lambda}$ *for some large* N*.*

Using the Borel-Weil construction of irreps (not described here), it is easy to prove that the set of pairs $\{(\lambda, \mu) : V_\mu \text{ appears in } V_\lambda\}$ is closed under addition and finitely generated. Using the Corollary, this implies that the set of pairs $\{(\lambda, \mu) : \mathcal{M}_\lambda^\mu \neq \emptyset\}$ is a convex polyhedral cone. This is my favorite way of seeing that these $H \leq G$ problems have polyhedral answers (one also has the "Kirwan convexity theorem" that doesn't use the representation theory side).

In good cases one can dispense with the N in the second statement; this phenomenon is called *saturation*. It holds for the pair $T \to U(n)$ and for the pair $U(m) \to U(m) \times U(m)$ (which I proved with Tao), but not for $U(m) \times U(n) \to U(mn)$.

Note that this coefficient c, considered as a function of μ, is the probability density we were asking about in lecture 1 (unless d drops from its maximum, in which case the probability density is 0). So the very *best* thing to compute is the function f. This can be done in the $T \leq U(n)$ and $U(n) \leq U(n) \times U(n)$ cases by counting *integral* Gel′fand-Cetlin or Berenstein-Zelevinsky patterns.

4.5. **The state of knowledge of the Schur-Horn, Hermitian sum, and quantum marginal problems.** In the case $T \leq U(n)$, we know not only the $\sim 2^n$ Schur-Horn inequalities but the Gel′fand-Cetlin cone. This cone has only $2\binom{n+1}{2}$ facets; it is in this sense much simpler than its projection. It is easy to list its edges; they correspond to G-C patterns of 1s and 0s, with the dividing line given by some path from bottom to top. There are $2^n - 1$ such (the all-0 case doesn't give an edge); on the representation theory side, they correspond to the basis vectors in all the fundamental representations $\{\mathrm{Alt}^k \mathbb{C}^n\}$.

In the cse $U(n) \leq U(n) \times U(n)$, we know not only the Klyachko inequalities but the Berenstein-Zelevinsky cone. This cone has only $2\binom{n+1}{2}$ facets; it is in this sense much simpler than its projection. It is *not* easy to list its edges. Moreover, it is not easy in general to locate the corresponding copy of V_ν inside $V_\lambda \otimes V_\mu$.

In the case $U(m) \times U(n) \to U(mn)$, all we have are the Berenstein-Sjamaar-Ressayre inequalities, as elucidated by Klyachko. It would be much nicer to have a polyhedral way to compute the Kronecker coefficients.

References

[BB] N. Bergeon, S. Billey

[Br03] Bravyi

[Bri05] ______: Lectures on the Geometry of Flag Varieties, in *Topics in cohomological studies of algebraic varieties*, ed. Piotr Pragacz, Trends in Mathematics. Birkhäuser Verlag, 2005, 33–85.

[KnMil05] ______, E. Miller: Gröbner geometry of Schubert polynomials, Ann. of Math. (2) 161 (2005), no. 3, 1245–1318. `http://arxiv.org/abs/math/0110058`

[KnTao03] ______, T. Tao: Puzzles and (equivariant) cohomology of Grassmannians, Duke Math. J. 119 (2003), no. 2, 221–260. `http://arxiv.org/abs/math.AT/0112150`

[Re00] Geometric invariant theory and generalized eigenvalue problem http://arxiv.org/abs/0704.2127

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NY 14853 USA

E-mail address: allenk@math.cornell.edu