Applications of enumerative algebraic geometry: complete quadrics in statistics

Mateusz Michałek and Tim Seynnaeve

February 2021

Notation and conventions

We work over $\mathbb{C}$, unless stated otherwise. Varieties are not assumed to be irreducible.

$G(k,n)$ will denote the Grassmannian of k-dimensional vector subspaces of $\mathbb{C}^n$, and $\mathbb{G}(k-1,n-1)$ will denote the Grassmannian of $k-1$-dimensional projective linear subspaces of $\mathbb{P}^{n-1}$. In other words $G(k,n) \cong \mathbb{G}(k-1,n-1)$

We will occasionally use the word *scheme*, but don't be afraid; if you can accept that the intersection of a conic and a tangent line is a point with multiplicity 2, that suffices.

1 Chow rings and Grassmannians

1.1 Introduction

> Consider 4 randomly chosen lines L_1, L_2, L_3, L_4 in $\mathbb{P}^3$. How many lines pass through all 4 of them?

The approach we will follow to tackle this and similar questions is as follows: The set of all lines in $\mathbb{P}^3$ is the Grassmannian $\mathbb{G}(1,3)$. For each of our lines L_i, let $V_i \subset \mathbb{G}(1,3)$ be the set of all lines passing through L_i; this is a subvariety of the Grassmannian. Our problem can then be restated as: "how many points are in the intersection $\bigcap_{i=1}^4 V_i$?" Our next aim is to define the gadget that will help us solve this question.

1.2 The Chow group

In the upcoming 4 subsections we give a short introduction to Chow rings; see the first chapter of [EH16] for more. Given a (smooth, irreducible) variety X we define the *group of algebraic cycles* $Z(X)$ to be the free abelian group generated by all irreducible subvarieties of X. Given a (not necessarily irreducible) subvariety Y of X, we can associate to it an element $\langle Y \rangle$ of $Z(X)$ by taking the sum of its irreducible components.

The *Chow group* $A(X)$ of X is the quotient of $Z(X)$ by rational equivalence. Roughly speaking, two elements of $Z(X)$ are rationally equivalent if one can be algebraically deformed in one another. Here is a more precise definition.

Definition 1.2.1. For every irreducible subvariety Z of $\mathbb{P}^1 \times X$ that is not contained in one of the fibers $\{t\} \times X$, we write $Z_0 := Z \cap (\{0\} \times X) \subseteq X$, and $Z_\infty := Z \cap (\{\infty\} \times X) \subseteq X$ (where the intersections are taken scheme-theoretically). We think of Z as a one-parameter family interpolating between Z_0 and Z_∞, and say Z_0 and Z_∞ are *rationally equivalent*.

Let $\mathrm{Rat}(X) \subseteq Z(X)$ be the subgroup generated by differences $\langle Z_0\rangle - \langle Z_\infty\rangle$, where Z runs over the irreducible subvarieties of $\mathbb{P}^1 \times X$ that is not contained in one of the fibers $\{t\} \times X$. The *Chow group* of X is defined to be the quotient $Z(X)/\mathrm{Rat}(X)$.

For $Y \subseteq X$ a subvariety, we will denote the corresponding class in the Chow group $A(X)$ by $[Y]$.

Remark 1.2.2. The group $Z(X)$ is naturally graded by codimension: $Z(X) = \bigoplus_{k=0}^{\dim X} Z^k(X)$, where $Z^k(X)$ is the subgroup generated by subvarieties of codimension k. One can show ([EH16, Proposition 1.4]) that this grading is compatible with rational equivalence, so that the Chow group inherits this grading: $A(X) = \bigoplus_{k=0}^{\dim X} A^k(X)$. We also write $A_k(X) = A^{\dim X - k}(X)$.

Example 1.2.3. Let $X = \mathbb{P}^n$ and fix a codimension k. All codimension k linear subspaces of $\mathbb{P}^n$ are rationally equivalent[1], hence we can define a class $\zeta_k \in A^k(\mathbb{P}^n)$: the class of an $(n-k)$-plane. In fact, it turns out that for a subvariety Y of degree d and codimension k, $\langle Y\rangle$ is linearly equivalent to $d\langle L\rangle$, where L is an $(n-k)$-plane. So $A^k(\mathbb{P}^n) \cong \mathbb{Z}$, generated by the element ζ_k.

1.3 The Chow ring

The Chow group can be made into a ring via the intersection product.

Definition 1.3.1. Two subvarieties A, B of a smooth variety X intersect *transversely* at a point p if $T_pA + T_pB = T_pX$. We say that A and B intersect *generically transverse* if they intersect transversely at a general point of each component of $A \cap B$.

Theorem 1.3.2 (Chow-Fulton). *If X is a smooth quasi-projective variety, then there is a unique product structure on $A(X)$satisfying the condition:*

> *If two subvarieties A, B of X are generically transverse, then $[A] \cdot [B] = [A \cap B]$.*

This structure makes

$$A(X) = \bigoplus_{k=0}^{\dim X} A^k(X)$$

[1] this can be proved using that for 2 linear subspaces, there is a linear transformation transforming one in the other

into an associative, commutative ring, graded by codimension, called the Chow ring *of X.*

Elements of the degree 1 part $A^1(X)$ are divisors[2] up to rational equivalence; $A^1(X)$ is the divisor class group (which is also the Picard group).

Example 1.3.3. Let $X = \mathbb{P}^n$, and let A and B be general planes of codimensions a and b. Then A and B intersect transversely: if $a+b > n$ the intersection is empty, else it is a plane of codimension $a+b$. In other words, the intersection product on $A(\mathbb{P}^n)$ is described by

$$\zeta_a \cdot \zeta_b = \begin{cases} \zeta_{a+b} & \text{if } a+b \leq n \\ 0 & \text{if } a+b > n. \end{cases}$$

From this it follows that $A(\mathbb{P}^n) \cong \mathbb{Z}[\zeta]/(\zeta^{n+1})$ as rings, where $\zeta := \zeta_1$ and $\zeta^i = \zeta_i$.

1.4 Pushforward and pullback

Definition 1.4.1. For $f : X \to Y$ a proper morphism of varieties, there is a pushforward map $f_* : A(X) \to A(Y)$, defined as follows: For $Z \subseteq X$ an (irreducible) subvariety of dimension i, we put

$$f_*([Z]) = \begin{cases} 0 & \text{if } f|_Z \text{ is not generically finite (i.e. if } \dim(f(Z)) < \dim(Z)) \\ d[f(Z)] & \text{if } f|_Z \text{ is generically finite of degree } d \end{cases}$$

By linearly extending, we get a group morphism $f_* : A^{\dim(X)-i}(X) \to A^{\dim(Y)-i}(Y)$.

A proof that this is well-defined (i.e. is compatible with linear equivalence) can be found in [Ful98, Section 1.4]. f_* is only a morphism of groups, not of rings; it neither preserves the grading nor the product.

For every projective variety X, the unique map $\pi : X \to pt$ is a proper morphism. The corresponding pushforward map $\deg := \pi_* : A^{\dim(X)}(X) \to A^0(pt) \cong \mathbb{Z}$ is known as the *degree map*; it simply sends the class of each point in X to $1 \in \mathbb{Z}$. The degree map will play a fundamental role in solving enumerative problems: if $Z \subseteq X$ is a 0-dimensional subvariety (i.e. a collection of points in X), counting the number of points amounts to computing the degree $\deg([Z])$. In the literature, the degree map is often written as $\int_X$.

Definition 1.4.2. If $f : X \to Y$ is a morphism of varieties, it induces a *pullback map* $f^* : A(Y) \to A(X)$, defined in [EH16, Theorem 1.23(a)]. f^* is a morphism of graded rings. In principle, one defines $f^*(Z) = [f^{-1}(Z)]$. Formally, this holds as long as the map f is flat, or Z is Cohen-Macaulay and $f^{-1}(Z)$ has correct codimension. Note that the equality cannot hold without any assumptions: consider a blowup of $\mathbb{P}^2$.

[2]since X is smooth, we do not have to distinguish among Weil and Cartier divisors

Example 1.4.3. Consider $\mathbb{P}^2$ and two curves Z_1, Z_2 of degree d_1, d_2. We know $[Z_i] = d_i\xi$ in the ring $A(\mathbb{P}^2) = \mathbb{Z}[\xi]/(\xi^3)$. The computation $(d_1\xi)(d_2\xi) = d_1d_2\xi^2$ tells us that if Z_1 and Z_2 intersect transversally, then they intersect in d_1d_2 many points (Bézout's theorem).

1.5 Affine stratifications

Definition 1.5.1. A *stratification* of a variety X is a finite collection of irreducible, locally closed subvarieties U_i such that

1. $X = \bigsqcup U_i$.
2. The closure $\overline{U_i}$ of every U_i is a union of U_j's.

We call the U_i *open strata*, and their closures $Y_i = \overline{U_i}$ *closed strata*. If each U_i is isomorphic to an affine space $\mathbb{A}^k$, we say that the stratification is an *affine stratification*.

Theorem 1.5.2. *Let X be a quasiprojective variety with an affine stratification. Then:*

- *The Chow group $A(X)$ is generated (as a free abelian group) by the classes of the closed strata.*
- *In fact, the classes of the closed strata form a $\mathbb{Z}$-basis of $A(X)$.*

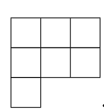
Proof. A proof of the first part can be found in [EH16, Proposition 1.17]. The second part is much harder, and was only proven in 2014 by Totaro [Tot14]. However, for all varieties we will meet, it is possible to prove the second part directly. □

Example 1.5.3. Let $X = \mathbb{P}^n$, and for $i \in \{0, \dots, n\}$, let $U_i := \{[0 : \dots : 0 : 1 : x_{i+1} : \dots : x_n]\} \subset \mathbb{P}^n$. Then the U_i form an affine stratification of $\mathbb{P}^n$, with closed strata $Y_i = \{[0 : \dots : 0 : x_i : x_{i+1} : \dots : x_n]\} \cong \mathbb{P}^{n-i} \subseteq \mathbb{P}^n$. Then Theorem 1.5.2 implies that $A^k(\mathbb{P}^n)$ is generated by $\zeta_i = [Y_i]$, without needing to explicitly show that every subvariety of $\mathbb{P}^n$ is rationally equivalent to a sum of linear subspaces.

1.6 Partitions and Young diagrams

A *partition* λ is a nonincreasing sequence $[\lambda_1, \dots, \lambda_n]$ of nonnegative integers. The *length* $\operatorname{len}(\lambda)$ of the partition is the length of the sequence, the *weight* $|\lambda| = \sum_{i=1}^n \lambda_i$. To every partition λ, we associate a *Young diagram*, with λ_i boxes in the i-th row. So $\operatorname{len}(\lambda)$ is the number of rows, and $|\lambda|$ is the number of boxes.

Example 1.6.1. Consider $\lambda = [3, 3, 1]$. The associated Young diagram is:



.

There is a bijection between partitions and sets of nonnegative integers: to a partition λ of length n, we associate a set

$$I(\lambda) := \{\lambda_n, \lambda_{n-1}+1, \ldots, \lambda_2+n-2, \lambda_1+n-1\} \subset \mathbb{N}^n.$$

Then given a set $I = \{i_1, \ldots, i_n\} \subset \mathbb{N}^n$, where we assume $i_1 < \ldots < i_n$, the corresponding partition is

$$\lambda(I) := [i_n-(n-1), i_{n-1}-(n-2), \ldots, i_2-1, i_1].$$

1.7 Affine stratification of the Grassmannian

Let $\Lambda \in G(k,n)$ be a k-plane in $\mathbb{C}^n$. Then Λ is the row span of a full rank $k \times n$ matrix A. Two such matrices give the same Λ precisely when one can be transformed in the other using elementary row operations. Hence out of all the matrices representing Λ, there is a unique one which is in reduced row echelon form. We will denote this matrix by A_Λ.

Definition 1.7.1. Let λ be a length k partition with $\lambda_1 \le n-k$ (i.e. the Young diagram fits inside a $k \times (n-k)$ box). We define the *open Schubert cell*

$$\Sigma^\circ_\lambda := \{\Lambda \in G(k,n) \mid \text{ the pivots of } A_\Lambda \text{ are in positions } I(\lambda)\}.$$

For instance, for G(2,4) we have:

$$\Sigma^\circ_\emptyset = \begin{pmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \end{pmatrix}, \Sigma^\circ_{\square} = \begin{pmatrix} 1 & * & 0 & * \\ 0 & 0 & 1 & * \end{pmatrix},$$
$$\Sigma^\circ_{\square\square} = \begin{pmatrix} 1 & * & * & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \Sigma^\circ_{\substack{\square\\\square}} = \begin{pmatrix} 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{pmatrix},$$
$$\Sigma^\circ_{\substack{\square\square\\\square}} = \begin{pmatrix} 0 & 1 & * & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \Sigma^\circ_{\substack{\square\square\\\square\square}} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The closed Schubert cells can be described by replacing in the above every 1 by a $*$.

Proposition 1.7.2. *The Schubert cells Σ_Λ form an affine stratification of $G(k,n)$.*

Sketch of proof. By construction, $G(k,n) = \bigsqcup_\lambda \Sigma^\circ_\lambda$, and $\Sigma^\circ_\lambda \cong \mathbb{A}^{k(n-k)-|\lambda|}$, and $\Sigma_\lambda = \bigsqcup_{\mu \ge \lambda} \Sigma^\circ_\mu$. Here $\mu \le \lambda$ means that $\mu_i \le \lambda_i$ for each i. $\square$

Corollary 1.7.3. *The Chow ring $A(G(k,n))$ has a $\mathbb{Z}$-basis given by the* Schubert classes $\sigma_\lambda := [\Sigma_\lambda]$.

Here is a more geometric description of the Schubert cells: consider the Grassmannian $G(k,V)$, where $\dim V = n$, and fix a complete flag $\mathcal{V} = (V_1 \subset \ldots \subset V_{n-1} \subset V_n = V)$. We define the Schubert cell

$$\Sigma_\lambda(\mathcal{V}) := \{\Lambda \in G(k,V) \mid \dim(V_{n-k+i-\lambda_i} \cap \Lambda) \ge i \ \forall i\}.$$

If we take $V = \mathbb{C}^n$ and $V_i = \mathrm{span}(e_{n-i+1}, \dots, e_n)$, this agrees with our previous definition.

In particular, if $\lambda = [a, 0, \dots, 0]$, the class $\sigma_\lambda =: \sigma_a$ is given by the locus of all k-planes intersecting a given $(n-k+1-a)$-plane nontrivially. The degree of σ_λ in $A(X)$ (i.e. the codimension of the variety above) is given by $|\lambda| := \lambda_1 + \dots + \lambda_k$.

1.8 Interlude: symmetric polynomials and Schur polynomials

We consider the ring $\Lambda_k = \mathbb{Z}[x_1, \dots, x_k]^{\mathfrak{S}_k}$ of symmetric polynomials in k variables. One of the most important classes of symmetric functions are the *Schur polynomials*. They play a central role in the representation theory of the general linear group, and as we will see later, also in Schubert calculus.

There are several ways of defining Schur polynomials, but non of them is very easy. The definition below, via Jacobi's bialternant formula, is perhaps the most direct one:

Definition 1.8.1. For a partition λ (or Young diagram) of length k, the Schur polynomial s_λ is equal to the following quotient of determinants:

$$s_\lambda(x_1, \dots, x_k) = \frac{\begin{vmatrix} x_1^{\lambda_1+k-1} & x_2^{\lambda_1+k-1} & \dots & x_k^{\lambda_1+k-1} \\ x_1^{\lambda_2+k-2} & x_2^{\lambda_2+k-2} & \dots & x_k^{\lambda_2+k-2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{\lambda_k} & x_2^{\lambda_k} & \dots & x_k^{\lambda_k} \end{vmatrix}}{\begin{vmatrix} x_1^{k-1} & x_2^{k-1} & \dots & x_k^{k-1} \\ x_1^{k-2} & x_2^{k-2} & \dots & x_k^{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{vmatrix}}.$$

Note that the determinant in the denominator is the Vandermonde determinant $\prod_{i<j}(x_i - x_j)$, and that the degree of s_λ is equal to the weight $|\lambda|$ of the partition λ. Two important special cases are

$$s_{\underbrace{\square\square\square}_{a}} = h_a := \sum_{i_1 \le \dots \le i_a} x_{i_1} \cdots x_{i_a},$$

the a'th *complete homogeneous symmetric polynomial*, and

$$s_{\left.\begin{smallmatrix}\square\\\square\\\square\end{smallmatrix}\right\}a} = e_a := \sum_{i_1 < \dots < i_a} x_{i_1} \cdots x_{i_a},$$

the a'th *elementary symmetric polynomial*.

The Schur polynomials s_λ form a $\mathbb{Z}$-basis of Λ_k. In particular, every product of Schur polynomials can be written as a $\mathbb{Z}$-linear combination of Schur polynomials:

$$s_\lambda s_\mu = \sum_\nu c^\nu_{\lambda\mu} s_\mu,$$

where the sum runs over all partitions ν of length at most k and weight $|\nu| = |\lambda| + |\mu|$. The coefficients $c^\nu_{\lambda\mu} = c^\nu_{\mu\lambda} \in \mathbb{Z}$ are known as the *Littlewood-Richardson coefficients*[3]. A combinatorial formula for computing them can be found e.g. in [FH91, (A.8)], or on Wikipedia.

There are 2 cases in which the Littlewood-Richardson coefficients are easy to describe:

Theorem 1.8.2 (Pieri's rule)**.** *For any partition λ and integer n we have:*

$$s_\lambda(x_1, \dots, x_k) \cdot s_{[n]}(x_1, \dots, x_k) = \sum_\mu s_\mu(x_1, \dots, x_k),$$

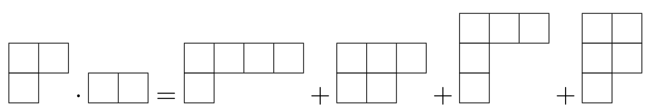
where the sum is over all Young diagrams μ obtained from λ by adding n boxes, at most 1 in each column.

Similarly, for $[1, \dots, 1]$ a partition of n we have:

$$s_\lambda(x_1, \dots, x_k) \cdot s_{[1,\dots,1]}(x_1, \dots, x_k) = \sum_\mu s_\mu(x_1, \dots, x_k),$$

where the sum is over all Young diagrams μ obtained from λ by adding n boxes, at most 1 in each row.

To simplify notation one may write λ instead of s_λ.

Example 1.8.3. 

1.9 Exercises

Exercise 1.9.1.

1. Let $v_1, \dots v_r$ be r linearly independent vectors in $\mathbb{C}^n$. Consider the following subvariety of $G(k, \mathbb{C}^n)$:

$$\{\Lambda \mid v_1, \dots, v_r \text{ are linearly dependent in } \mathbb{C}^n/\Lambda\}.$$

   What is the class of this subvariety in the Chow ring $A(G(k, n))$?
   Hint: it is a Schubert class σ_λ, where λ is some Young diagram.

2. Let $\beta_1, \dots \beta_r$ be r linearly independent vectors in $(\mathbb{C}^n)^*$. Consider the following subvariety of $G(k, \mathbb{C}^n)$:

$$\{\Lambda \mid \beta_1, \dots, \beta_r \text{ become linearly dependent when restricted to } \Lambda\}.$$

   What is the class of this subvariety in the Chow ring $A(G(k, n))$?

3. Congratulations! You just computed the Chern classes of the universal quotient bundle and the dual universal subbundle of the Grassmannian. We'll explain what all these words mean in the next lectures.

[3] In fact, it turns out $c^\nu_{\lambda\mu}$ is always nonnegative.

Exercise 1.9.2. Using Pieri's rule, show that, after setting to zero all (Schur polynomials corresponding to) diagrams with either more than three rows or columns one has:

$$(1 + t\,\ydiagram{1} + t^2\,\ydiagram{2} + t^3\,\ydiagram{3})(1 - t\,\ydiagram{1} + t^2\,\ydiagram{1,1} - t^3\,\ydiagram{1,1,1}) = 1$$

Exercise 1.9.3.
- Let λ be a Young diagram that fits in an $r \times (n-r)$ box, and let $\lambda^c = [n-r-\lambda_r, \ldots, n-r-\lambda_1]$ be its complement. Show that $\sigma_\lambda \sigma_{\lambda^c} = 1 \in A(G(r,V))$, by fixing two general flags $\mathcal{V} = (V_1 \subset \ldots \subset V_n = V)$ and $\mathcal{W} = (W_1 \subset \ldots \subset W_n = V)$ in V and showing that $\Sigma_\lambda(\mathcal{V})$ and $\Sigma_{\lambda^c}(\mathcal{W})$ intersect in a single point. If you like working in coordinates, you can take $V = \mathbb{C}^n$, $V_i = \operatorname{span}(e_{n-i+1}, \ldots, e_n)$, and $W_i = \operatorname{span}(e_1, \ldots, e_i)$. Then the cells $\Sigma_\mu(\mathcal{V})$ are the ones from Definition 1.7.1, and the cells $\Sigma_\mu(\mathcal{W})$ look similar but with the order of the coordinates reversed. Do try this out on some examples first!

- Let λ and μ be two Young diagrams inside an $r \times (n-r)$ box such that $|\lambda| + |\mu| = r(n-r)$. If $\mu \neq \lambda^c$, show that $\sigma_\lambda \sigma_\mu = 0 \in A(G(r,V))$. Again, do try this out on some examples first!

2 K-theory, vector bundles and Grassmannians

2.1 Ring structure on the Chow ring of the Grassmannian

Theorem 2.1.1. *The product of two Schubert classes in $A(G(k,n))$ is given by*

$$\sigma_\lambda \sigma_\mu = \sum_\nu c^\nu_{\lambda\mu} \sigma_\nu,$$

where $c^\nu_{\lambda\mu}$ are the Littlewood-Richardson coefficients.

We can restate the above as

Theorem 2.1.2. *The Chow ring of the Grassmannian is isomorphic to the ring of symmetric polynomials in k variables, modulo the ideal generated by all Schur polynomials s_λ for which $\lambda_1 > n-k$. The isomorphism sends the Schubert class σ_λ to the Schur polynomial s_λ.*

One way of proving the above theorems is by showing that the Pieri rule holds in $A(G(k,n))$, see for instance [EH16, Proposition 4.9], or Proposition 3.8 in *these lecture notes*. The theorem then follows from the fact that the ring Λ_k of symmetric polynomials is generated by the complete homogeneous symmetric polynomials.

Remark 2.1.3. The Grassmannian $Gr(k,V)$ is naturally isomorphic to $Gr(n-k, V^*)$, by sending a subspace $\Lambda \subset V$ to $(V/\Lambda)^* \subset V^*$. This isomorphism induces an isomorphism of Chow rings: the Schubert class $\sigma_\lambda \in A(Gr(k,V))$ gets sent to the Schubert class $\sigma_{\lambda^T} \in A(Gr(k,V))$, where λ^T is the conjugate partition.

In summary: here is a general strategy for solving intersection problems on the Grassmannian:

Problem: given a number of conditions $C_1, \dots, C_m$ on k-planes in $\mathbb{C}^n$, determine how many k-planes simultaneously satisfy all of the conditions.

Strategy:

1. For each condition C_i, let $Y_i \subset G(k,n)$ be the variety of all k-planes satisfying the condition, and determine the class $[Y_i] \in A(G(k,n))$ as a linear combination of the Schubert classes.

2. Verify that the Y_i intersect generically transversely.

3. Compute the product $[Y_1] \cdots [Y_m] \in A(G(k,n))$.

Step 3 is now completely solved, as Theorem 2.1.1, together with the Littlewood-Richardson rule, tells us how to compute products in the Chow ring. Step 2 (verifying transversality) can be hard in general, but in many concrete examples our conditions are generic (e.g. "passing through a *general* point" or "being tangent to a *general* quadric") and then transversality will follow from this genericity. One can make this precise using *Kleiman's transversality theorem* (see [EH16, Theorem 1.7]).

Theorem 2.1.4. *Suppose a group G acts on a variety X transversally (like $GL(V)$ on a Grassmannian $G(k,V)$). For any two subvarieties $Y_1, Y_2 \subset X$ and general $g \in G$ the varieties gY_1 and Y_2 are generically transverse.*

The main challenge remaining is step 1: given a condition on k-planes in $\mathbb{C}^n$, how to determine the corresponding class?

In some sense you were solving such a problem in Exercise 1.9.1 from the last lecture. Below we present one more example.

Example 2.1.5. What is the locus of points in $G(2,4)$ that correspond to lines intersecting a given line l?

Note that here, as $GL(4)$ acts transitively on $G(2,4)$ it does not matter which l we fix. Of course the explicit subvarieties of $G(2,4)$ may differ (depending on l), but they always define the same class in $A(G(2,4))$. Considering a complete flag that goes through l this variety, by definition, is the Schubert variety $\Sigma_{\square}$.

So, *how many lines intersect four general lines*? We need to intersect four times $\Sigma_{\square}$, obtaining:

$$\square^4 = 2\,\yng(2,2).$$

In other words, by Theorem 2.1.4, the intersection consists of two points, i.e. there are two such lines.

For more examples we refer to `https://fulges.github.io/docs/Teaching/2021_MPI_EnumerativeGeometry/IntroEnumerativeGeometry.pdf`, Chapter 2.

Next, we provide more tools for solving step 1 above: often, the classes we are interested in can be interpreted as *Chern classes* of a *vector bundle*.

2.2 Vector bundles

Intuitively, a vector bundle is a family of vector spaces V_x parameterized by points $x \in X$ of an algebraic variety.

Definition 2.2.1 (Vector bundle)**.** A vector bundle is a morphism of algebraic varieties $f : V \to X$ such that:

1. There is an open covering U_i of X such that $m_i : f^{-1}(U_i) \simeq \mathbb{C}^r \times U_i$ and the map f restricted to $f^{-1}(U_i)$ is the composition of m_i with the projection to U_i.

2. For $x \in U_{i_1} \cap U_{i_2}$ the induced map: $\mathbb{C}^r \simeq \mathbb{C}^r \times \{x\} \simeq f^{-1}(x) \simeq \mathbb{C}^r \times \{x\} \simeq \mathbb{C}^r$ is an isomorphism of vector spaces. (Here the second and third isomorphisms come respectively from m_{i_1} and m_{i_2}.)

The maps $U_{i_1} \cap U_{i_2} \to GL(r)$ from point 2 are known as *transition* functions. Note that one can reconstruct the bundle V from the transition functions, by glueing procedures.

The fiber $f^{-1}(x)$ is often denoted by V_x. It comes with a natural structure of a vector space (but not with a basis!).

A *section* of a vector bundle is such a morphism $s : X \to V$ that $f \circ s = id_X$. Every vector bundle has a (zero) section.

A *morphism* of vector bundles V_1, V_2 is such a morphism (of varieties) that the following diagram commutes:

$$\begin{array}{ccc} V_1 & \longrightarrow & V_2 \\ & \searrow & \downarrow \\ & & X \end{array}$$

and over each point of X we obtain a linear map of vector spaces.

(For people familiar with sheaves: think about locally free sheaf of modules, where the elements of the module are sections of f)

To every vector bundle V we associate the dual vector bundle V^*, where each vector space V_x over $x \in X$ is replaced by V_x^*. Formally, we change the transition functions by their inverse transpose and construct V^* by glueing.

Example 2.2.2. For any variety X we have the trivial rank r vector bundle $\mathbb{C}^r \times X$. The dual of the trivial bundle is the (same) trivial bundle.

On the projective space $\mathbb{P}(V)$ we have the trivial vector bundle $V \times \mathbb{P}(V)$. It contains the tautological line (i.e. rank one) bundle $\mathcal{O}(-1)$ given by considering over a point $[x] \in \mathbb{P}(V)$ the line in V that it represents.

On the open set $U_i \subset \mathbb{P}(V)$ given by $x_i \neq 0$ we have the trivialization of $\mathcal{O}(-1)$ that assigns to a point $(a_0, \dots, a_n) \in V$ on a line through $[a_0 : \dots : a_n] \in \mathbb{P}(V)$ the number $a_i \in \mathbb{C}$.

This line bundle has no sections, apart from the zero one. The dual of $\mathcal{O}(-1)$ is $\mathcal{O}(1)$. The latter has sections: every homogeneous linear polynomial gives a map from fibers of $\mathcal{O}(-1)$ to $\mathbb{C}$.

On the Grassmannian $G(k, V)$ we also have the trivial bundle $V \times G(k, V)$ that for simplicity we will denote by $\underline{V}$. It contains the tautological bundle $\mathcal{U}$, where over $[W] \in G(k, V)$ we take $W \subset V$. We also have the bundle Q, where over $[W] \in G(k, V)$ we consider V/W. Formally, it is easier to construct first the dual by considering $(V/W)^*$ as a subspace of V^* and then dualizing.

2.3 K-theory

A sequence of morphisms of vector bundles

$$0 \to V_1 \to V_2 \to V_3 \to 0$$

is called exact, if it is an exact sequence of vector spaces over each point $x \in X$.

Definition 2.3.1. Let X be a smooth variety. The K-theory (or Grothendieck group) $K^0(X)$ is defined as the free abelian group on isomorphisms classes of vector bundles on X modulo the subgroup generated by $[V_1] + [V_3] - [V_2]$ for every exact sequence:

$$0 \to V_1 \to V_2 \to V_3 \to 0.$$

Another operation on vector bundles is their tensor product $V_1 \otimes V_2$, which over each $x \in X$ is the tensor product of the two vector spaces coming from the two bundles. Note that formally, one needs to say what are the transition functions. This product gives $K^0(X)$ a structure of a ring, where identity is the (class of the) trivial rank one (line) bundle.

In a similar way, we may take exterior and symmetric powers of a vector bundle.

Example 2.3.2. When X is a point, a vector bundle is just a vector space. We have $K^0(X) = \mathbb{Z}$ (with the usual ring structure).

Example 2.3.3. We have $K^0(\mathbb{P}^n) = \mathbb{Z}[x]/(x^{n+1})$. This will be a special case of more general series of examples, but first let us try to understand the above equality. The first question is: what is x?

We claim that $x = [\mathcal{O}(-1)] - 1$. Why would we have $x^{n+1} = 0$? Consider the case $n = 2$. We have the Euler sequence:

$$0 \to \Omega^1 \to \mathcal{O}(-1)^{\oplus 3} \to \mathcal{O} \to 0.$$

By looking at the second exterior power we obtain the exact sequence:

$$0 \to \Omega^2 \to \bigwedge^2 \mathcal{O}(-1)^{\oplus 3} \to \Omega^1 \to 0.$$

But Ω^2 corresponds on $\mathbb{P}^2$ to the canonical line bundle $\mathcal{O}(-3)$. Hence by the second sequence:

$$[\mathcal{O}(-3)] + [\Omega^1] - [\bigwedge^2 \mathcal{O}(-1)^{\oplus 3}] = 0.$$

We also have an exact sequence:

$$0 \to \mathcal{O}(-1)^{\oplus 2} \to \mathcal{O}(-1)^{\oplus 3} \to \mathcal{O}(-1) \to 0.$$

Taking the second exterior power we obtain:

$$0 \to \bigwedge^2 \mathcal{O}(-1)^{\oplus 2} \to \bigwedge^2 \mathcal{O}(-1)^{\oplus 3} \to \mathcal{O}(-2)^{\oplus 2} \to 0.$$

Note that $\bigwedge^2 \mathcal{O}(-1)^{\oplus 2}$ is a line bundle equal to $\mathcal{O}(-2)$. Computing $[\Omega^1]$ from the first exact sequence we obtain:

$$[\mathcal{O}(-3)] + 3[\mathcal{O}(-1)] - 1 - 3[\mathcal{O}(-2)] = 0,$$

which precisely tells us $x^3 = 0$.

If you are confused about exterior powers for exact sequences look in: Hartshorne, exercise 5.16.

If you are confused about Euler sequence look in Wikipedia.

If you are confused how we see the type of line bundle from an expression like: $\bigwedge^2 \mathcal{O}(-1)^{\oplus 2}$ - we will introduce general tools soon.

If you are not confused at all - it must have been very boring for you so far!

It is often more convenient to look at rational coefficients for the K-theory, i.e. consider $K_{\mathbb{Q}}(X) := K(X) \otimes_{\mathbb{Z}} \mathbb{Q}$. It is not a coincidence that the K-theory ring for $\mathbb{P}^n$ was the same as the Chow ring (and the cohomology ring). Later we will define the Chern map that will give an isomorphism $K^0_{\mathbb{Q}}(X) \to A_{\mathbb{Q}}(X)$.

2.4 Exercises

Exercise 2.4.1.

Can you present $A(G(2,4))$ as a polynomial ring modulo an ideal?

Exercise 2.4.2. Extend the formula $x^{n+1} = 0$ in Example 2.3.3 to arbitrary n.

Exercise 2.4.3. (*) How many lines are tangent to four general quadrics in $\mathbb{P}^3$?

3 Chern classes, Chern roots and Chern character

3.1 Chern classes

We will now describe a way to associate (classes of) subvarieties to a vector bundle. Let us start with line bundles.

The group of isomorphism classes of line bundles is denoted by $\mathrm{Pic}(X)$. If you are familiar with the definition of the Picard group as equivalence classes of Cartier divisors, note that the data of Cartier divisor is exactly the data needed to construct a line bundle. Now given a line bundle $L \to X$, there is always a nonzero, *rational* section $s : X \dashrightarrow L$ (for example $s(x) = (x, 1)$ on one trivialization $U \times \mathbb{C}$). Such a section s gives a rational function on any $U_i \subset X$ on which we fix a trivialization of L. Looking at the order of the poles and zeros of s we obtain $D(s)$ a linear combination of codimension one subvarieties of X:

$$D(s) = \sum_{Y \subset Z} (ord_Y(s))Y.$$

Example 3.1.1. Consider $\mathbb{P}^1 = \mathbb{P}(V)$ and the line bundle $L = \mathcal{O}(-1)$. On the open set $U_1 := \{(1 : x)\}$ we have the trivialization of the line bundle, $U_1 \times \mathbb{C} \to U_1 \times V$ given by $((1 : x), t) \mapsto ((1 : x), (t, tx))$. On this set we consider the rational section $U_1 \to L_{|U_1} \subset U_1 \times V$ given by $(1 : x) \mapsto ((1 : x), (1, x))$. Clearly, this section has neither poles nor zeros on U_1.

There is the second open set $U_2 := \{(x : 1)\}$ over which L also trivializes to $U_2 \times \mathbb{C} \to U_2 \times V$ given by $((x : 1), t) \mapsto ((x : 1), (tx, t))$. Consider the rational map induced by our section

$$U_2 \dashrightarrow U_2 \times \mathbb{C} \to \mathbb{C},$$

where the last map is simply the projection. We have:

$$(x : 1) = (1 : 1/x) \mapsto ((1 : 1/x), (1, 1/x)) \in U_2 \times V.$$

Hence, over $(x : 1)$ our section, as an element of V takes value $(1, 1/x)$. This is the image of $((x : 1), (1/x)) \in U_2 \times \mathbb{C}$ under the trivialization map. Thus s defines a function $U_2 \to \mathbb{C}$ given by $(x : 1) \to 1/x$. This has a simple pole. Thus $D(s) = -pt$ where pt is the unique point of $U_2 \setminus U_1$.

The above example may suggest that it is very easy to look at zeros and poles of a function, note however that our variety X may be glueing of spectra of strange rings, not just polynomial rings. Fortunately we assume that X is smooth, in which case the order of poles and zeros is well-defined by valuation theory. (The correct algebra statement is that a one dimensional regular local ring is a DVR.)

Theorem 3.1.2. *The map $L \to D(s)$ descends to a map*

$$c_1 : \mathrm{Pic}(X) \to A_{\dim X - 1}(X).$$

In other words, the choice of the section s changes the divisor $D(s)$, but not its equivalence class.

Surprisingly, in general, this map needs not be injective or surjective. Fortunately, under the assumption that X is smooth it is an isomorphism. (More precisely it is an isomorphism for locally factorial varieties, and the correct algebraic statement is that regular local ring is UFD).

The map c_1 is known as the *first Chern class* of a line bundle. We also say that all higher Chern classes of any line bundle are zero, or even more generally $c_i(\mathcal{E}) = 0$ for any vector bundle $\mathcal{E}$ of rank strictly smaller than i. Next we will define Chern classes $c_i(\mathcal{E})$ for $i = 0, \ldots, \mathrm{rk}\, \mathcal{E}$. There are a few ways to do this.

(Intuitive way) Suppose you want to globally trivialize a line bundle L. This means you need to say where is 1 in each line over each $x \in X$. Equivalently you want to find a nonzero section. If you pick any section, the zeros tell you "how much" you failed.

Now we would like to do the same for a vector bundle $\mathcal{E}$ of rank r. To trivialize it we would like to choose r sections that are *linearly independent* over each point $x \in X$. How much we failed is now encoded by the locus of points where the sections fail to be linearly dependent. The r sections could be encoded as the section of the *line bundle* $\bigwedge^r \mathcal{E}$. By definition $c_1(\mathcal{E}) := c_1(\bigwedge^r \mathcal{E})$.

Let us now choose $s_1, \dots, s_{r-i+1}$ general, global sections of $\mathcal{E}$. We *assume that the set* $D := \{x \in X : s_1(x), \dots, s_{r-i+1}(x)$ *fail to be linearly independent*$\}$ *has codimension* i. Then we can define $c_i(\mathcal{E}) \in A^i(X)$ as the sum of all the (classes of) varieties that are components of D. The assumption we make is quite strong, but gives a very good intuition what Chern classes measure.

Note that if we want to work with global sections, the same problem arises with line bundles. If we pick the zero section, it is not linearly independent with itself (i.e. zero) in codimension 0 (instead of wanted 1), and hence we cannot define c_1 with that section.

(Projectivisation of line bundle) In what follows we will use the variety $\mathbb{P}(\mathcal{E})$. This variety is obtained from $\mathcal{E}$ by first replacing each $U_i \times \mathbb{C}^r$ with $U_i \times \mathbb{P}(\mathbb{C}^r)$ and then using the transition function to glue the pieces together. The advantage is that $\mathbb{P}(\mathcal{E})$ comes with a natural line bundles $\mathcal{O}(-1)$ and its dual $\mathcal{O}(1)$ and the map $\pi : \mathbb{P}(\mathcal{E}) \to X$.

Definition 3.1.3. The i-th Segre class $\mathrm{Seg}_i(\mathcal{E})$ is defined as:

$$\mathrm{Seg}_i(\mathcal{E}) := \pi_*(c_1(\mathcal{O}(1))^{r+i-1}) \in A^i(X).$$

Note that Segre classes may be nonzero beyond the rank of the vector bundle!

Theorem 3.1.4. *There exist unique* $c_i(\mathcal{E}) \in A_{\dim X - i}(X)$ *(known as Chern classes) such that:*

$$(1 + \mathrm{Seg}_1(\mathcal{E})t + \mathrm{Seg}_2(\mathcal{E})t^2 + \dots)(1 + c_1(\mathcal{E})t + c_2(\mathcal{E})t^2 + \dots) = 1 \in A(X)[t].$$

In this sense the Chern classes are "inverses" of Segre classes. More explicitely by expanding: $c_1(\mathcal{E}) = -\mathrm{Seg}_1(\mathcal{E})$, $c_2(\mathcal{E}) = \mathrm{Seg}_1(\mathcal{E})^2 - \mathrm{Seg}_2(\mathcal{E})$ *and*

$$c_n(\mathcal{E}) = -\mathrm{Seg}_1(\mathcal{E})c_{n-1}(\mathcal{E}) - \mathrm{Seg}_2(\mathcal{E})c_{n-2}(\mathcal{E}) - \dots - \mathrm{Seg}_n(\mathcal{E}).$$

We always write $c_0(\mathcal{E}) = \mathrm{Seg}_0(\mathcal{E}) = 1$.

We note that the pull-back π^* makes $A(\mathbb{P}(\mathcal{E}))$ into a module over $A(X)$. This module is free and generated by $c_1(\mathcal{O}(1))^i$ for $i = 0, \dots, r-1$. Using this and Theorem 3.1.4 we obtain that the $c_i(\mathcal{E})$ are the unique elements of $A(X)$ that allow to decompose $c_1(\mathcal{O}(1))^r$ in the basis above. Indeed let $\xi = c_1(\mathcal{O}(1))$. There must exist classes $x_i \in A^i(X)$ which satisfy

$$\xi^r + \sum_{i=0}^{r-1} \pi^*(x_i)\xi^i = 0.$$

Theorem 3.1.5 (Projection formula)**.** *For a proper morphism* π *we have:*

$$\pi_*(\pi^*(A)B) = A\pi_*(B).$$

Multiplying the equation above the theorem by ξ^{r-1}, applying π_* and the projection formula, we obtain the formula as in Theorem 3.1.4 for $n = r$ and $x_i = c_{\dim X - i}(\mathcal{E})$.

Remark 3.1.6. One needs to be very careful about signs conventions, as all possible problems appear in the literature. The first one is with the construction of $\mathbb{P}(\mathcal{E})$. For some authors the fibers are $\mathbb{P}(\mathcal{E}_x)$, for some $\mathbb{P}(\mathcal{E}_x^*)$. This also depends on if the projective space parameterizes lines or hyperplanes. Second is sign convention for Segre classes. Some authors set $\mathrm{Seg}_1(\mathcal{E}) = c_1(\mathcal{E})$. Even Fulton himself uses different conventions is his different books.

3.2 Chern roots

Definition 3.2.1. The *Chern polynomial* is defined as:

$$c_t(\mathcal{E}) := \sum_{i=0}^{\infty} c_i(\mathcal{E})t^i \in A(X)[t].$$

The total Chern class is defined as:

$$c(\mathcal{E}) := \sum c_i(\mathcal{E}) \in A(X).$$

Theorem 3.2.2. *For an exact sequence of vector bundles:*

$$0 \to V_1 \to V_2 \to V_3 \to 0,$$

we have $c_t(V_2) = c_t(V_1)c_t(V_3)$.

Note in particular that if $\mathcal{E}$ is a direct sum of line bundles $\mathcal{E} = \bigoplus_{i=1}^r L_i$, then the Chern polynomial factors $c_t(\mathcal{E}) = \prod_{i=1}^r (1 + c_1(L_i)t)$. Although in general we are interested in cases when $\mathcal{E}$ does not split, it turns out that, when proving equalities among Chern classes it is enough to restrict to such cases. Formally the *splitting principle* is presented in [EH16, Section 5.4]). In general, one can 'pretend' that $c_t(\mathcal{E})$ factors.

We now describe this powerful method to compute Chern classes. One formally writes:

$$c_t(\mathcal{E}) = \prod_{i=1}^{r} (1 + a_i t). \tag{1}$$

The a_i are just formal expressions - they do not need to be some classes! Only elementary symmetric polynomials in those classes give the Chern classes. Explicitly $e_1(a_i) = \sum a_i = c_1(\mathcal{E})$, $e_2(a_i) = c_2(\mathcal{E})$. But now note that any symmetric polynomial in a_i's makes sense! Indeed, the ring of symmetric polynomials is generated by the elementary symmetric polynomials.

Chern roots give an easy way to describe how Chern classes change under basic operations. Let $\mathcal{E}, \mathcal{F}$ be vector bundles with chern roots a_i, b_j.

Proposition 3.2.3. *We have:*

- *the roots for $\mathcal{E}^*$ are $-a_i$;*
- *the roots for $\mathcal{E} \otimes \mathcal{F}$ are $a_i + b_j$ (note that there are $(\operatorname{rk}\mathcal{E}) \cdot (\operatorname{rk}\mathcal{F})$ many of those);*
- *the roots of $\wedge^p \mathcal{E}$ are $a_{i_1} + \cdots + a_{i_p}$ for all sequences $i_1 < \cdots < i_p$;*
- *the roots of $S^p\mathcal{E}$ are $a_{i_1} + \cdots + a_{i_p}$ for all sequences $i_1 \leq \cdots \leq i_p$;*

Example 3.2.4. On $\mathbb{P}^2$, let $-x := c_1(\mathcal{O}(-1))$. The vector bundle $(\mathcal{O}(-1))^{\oplus 2}$ has (two, the same) Chern roots: $-x, -x$. The line bundle $\bigwedge^2(\mathcal{O}(-1))^{\oplus 2}$ has one Chern root: $-x + (-x) = -2x$, i.e. $c_1(\bigwedge^2(\mathcal{O}(-1))^{\oplus 2}) = -2x$. In the same way we see that $c_1(\mathcal{O}(a)) = c_1(\mathcal{O}(1)^{\otimes a}) = ax$. As these are the only line bundles on $\mathbb{P}^2$, we must have $\bigwedge^2(\mathcal{O}(-1))^{\oplus 2} = \mathcal{O}(-2)$.

Example 3.2.5. Consider $G(r, V)$ where $\dim V = n$. Clearly $c_t(\underline{V}) = 1$.

We also note that $\mathcal{U}^*$ has natural sections, coming from linear functions in V^*. When do $e_1^*, \ldots, e_{r-i+1}^*$ fail to be independent on an r-dimensional subspace $W \subset V$? (cf. Exercise 1.9.1) Precisely when W intersects the space $\langle e_{r-i+2}, \ldots, e_n \rangle$ in dimension larger than expected. This gives us the Schubert variety $\Sigma_{(1,\ldots,1,0,\ldots,0)}$, where 1 appears r times!

Hence:

$$c_t(\mathcal{U}^*) = 1 + \ydiagram{1}\, t + \ydiagram{1,1}\, t^2 + \cdots + \ydiagram{1,1,1,1,1}\, t^r,$$

where the last Young diagram has r boxes.

Let's recall the isomorphism $A(G(r,V)) \cong \mathbb{Z}[x_1, \ldots, x_r]^{S_r}/I$ from Theorem 2.1.2: we can write

$$c_t(\mathcal{U}^*) = 1 + e_1(x_1, \ldots, x_r)t + e_2(x_1, x_2, \ldots, x_r)t^2 + \cdots = \prod_{i=1}^{r}(1 + x_i t).$$

Comparing with (1), we see that *under the isomorphism of Theorem 2.1.2, the variables x_i are precisely the Chern roots of the dual universal subbundle $\mathcal{U}^*$.*

Using (or even better not using, cf. Exercise 1.9.1) the exact sequence:

$$0 \to \mathcal{U} \to \underline{V} \to \mathcal{Q} \to 0$$

one can get:

$$c_t(\mathcal{Q}) = 1 + \ydiagram{1}\, t + \ydiagram{2}\, t^2 + \cdots + \ydiagram{7}\, t^{n-r},$$

where the last Young diagram has $n - r$ boxes.

We would like to now compute $c_t(S^2\mathcal{U})$. We know that $\mathcal{U}$ has Chern roots $-x_1, \ldots, -x_r$ and $e_j(x_1, \ldots, x_r) = \ydiagram{1,1,1,1,1}$. The Chern roots of $S^2\mathcal{U}$ are $-x_i - x_j$

and we want to compute the elementary symmetric polynomials in those. Let us carry this out for $G(2,4)$. We have:

$$x_1 + x_2 = \square, \quad x_1 \cdot x_2 = \begin{array}{|c|}\hline \ \\ \hline \ \\ \hline\end{array}.$$

Hence:

$$c_1(S^2\mathcal{U}) = e_1(-2x_1, -x_1 - x_2, -2x_2) = -3(x_1 + x_2) = -3\square$$

and

$$c_2(S^2\mathcal{U}) = e_2(-2x_1, -x_1 - x_2, -2x_2) = 2(x_1 + x_2)^2 + 4x_1 \cdot x_2 = 2\begin{array}{|c|c|}\hline \ & \ \\ \hline\end{array} + 6\begin{array}{|c|}\hline \ \\ \hline \ \\ \hline\end{array}$$

and

$$c_3(S^2\mathcal{U}) = e_3(-2x_1, -x_1-x_2, -2x_2) = -4x_1x_2(x_1+x_2) = -4\begin{array}{|c|}\hline \ \\ \hline \ \\ \hline\end{array} \cdot \square = -4\begin{array}{|c|c|}\hline \ & \ \\ \hline \ & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}.$$

Proposition 3.2.6. *For $\mathcal{E}$ a vector bundle with Chern roots $a_1, \dots, a_r$, (hence $c_i(\mathcal{E}) = e_i(a_1, \dots, a_r)$) the i-th Segre class is given by $\mathrm{Seg}_i(\mathcal{E}) = (-1)^i s_{[i]}(a_1, \dots, a_r)$, the i'th complete homogeneous symmetric polynomial in the Chern roots. This follows by combining Theorem 3.1.4 and the following identity of symmetric polynomials:*

$$\left(\sum_{i=0}^{r} e_i(x_1, \dots, x_r)t^i\right)\left(\sum_{j=0}^{\infty} s_{[j]}(x_1, \dots, x_r)t^j\right) = 1.$$

3.3 Chern character

Once we know the Chern roots it is easy to define the Chern character.

Definition 3.3.1. Let $\mathcal{E}$ be a vector bundle of rank r with Chern roots $x_1, \dots, x_r$. We define the *Chern character* $ch(\mathcal{E})$ by:

$$ch(\mathcal{E}) = \sum_{i=1}^{r} e^{x_i}.$$

Here, e^{x_i} should be understood as $\sum_{n=0}^{\infty} x_i^n/n!$. We note that each single summand does not make sense. However, after summing them up, we obtain, in each degree, a symmetric polynomial in the x_i's, i.e. a class in $A_{\mathbb{Q}}(X)$.

It is also possible to write down explicitly:

$$ch(\mathcal{E}) = r + c_1(\mathcal{E}) + \frac{1}{2}(c_1(\mathcal{E})^2 - 2c_2(\mathcal{E})) + \frac{1}{6}(c_1(\mathcal{E})^3 - 3c_1(\mathcal{E})c_2(\mathcal{E}) + c_3(\mathcal{E})) + \dots.$$

The expression above comes from writing the power sum polynomials in terms of elementary symmetric polynomials, which may be done from Newton's formula:

$$P_n - e_1P_{n-1} + e_2P_{n-2} - \dots + (-1)^{n-1}e_{n-1}P_1 + (-1)^n ne_n = 0,$$

where P_i is the i-th power sum polynomial and e_i is the i-th elementary symmetric polynomial.

Example 3.3.2. For the line bundle L and $D = c_1(L)$ we have: $ch(L) = 1 + D + D^2/2! + \dots$.

Theorem 3.3.3. *For an exact sequence of vector bundles*

$$0 \to V_1 \to V_2 \to V_3 \to 0$$

we have:

$$ch(V_2) = ch(V_1) + ch(V_3).$$

For any two vector bundles V, V' we have:

$$ch(V \otimes V') = ch(V)ch(V').$$

Proof. For the first equality: we already know that $c_t(V_2) = c_t(V_1)c_t(V_3)$, i.e. the set of Chern roots of V_2 is the union of the set of Chern roots of V_1 and V_3. Hence, the k-th power sum of Chern roots for V_2 is the sum of the k-th power sums of Chern roots for V_1 plus k-th power sum of Chern roots for V_3.

For the second equality:

$$\sum_{i,j} e^{x_i+y_j} = (\sum_i e^{x_i})(\sum_j e^{y_j}).$$

□

Corollary 3.3.4. *The Chern character is a well-defined map*

$$ch : K_{\mathbb{Q}}(X) \to A_{\mathbb{Q}}(X)$$

preserving the ring structure. This is an isomorphism!

Remark 3.3.5. We also have a map $A(X) \to H^*(X)$ and the induced map $K(X) \otimes \mathbb{Q} \to H^*(X, \mathbb{Q})$ taking its values in the even part of the cohomology ring. (Not an isomorphism in general!)

Example 3.3.6. Let $X = \mathbb{P}^n$ and $H = c_1(\mathcal{O}(1))$. We have:

$$ch(\mathcal{O}(1)) = 1 + H + H^2/2! + \cdots + H^n/n!.$$

Hence:

$$ch(\mathcal{O}(-1) - 1) = -H + H^2/2! + \cdots + (-H)^n/n!.$$

Recall that: $K(\mathbb{P}^n) \otimes \mathbb{Q} = \mathbb{Q}[x]/(x^{n+1})$, where $x = \mathcal{O}(-1) - 1$ and $A(X) \otimes \mathbb{Q} = \mathbb{Q}[H]/H^{n+1}$. The isomorphism of the two rings given by the Chern character satisfies the equality above.

3.4 Exercises

Exercise 3.4.1. Prove the formula for $c_t(\mathcal{Q})$.

Exercise 3.4.2. Compute the Segre classes of $S^2\mathcal{U}$ on $G(2,4)$.

Exercise 3.4.3. Prove Proposition 3.2.6. Possible hint:

$$0 \to \mathcal{U} \to \underline{V} \to \mathcal{Q} \to 0.$$

Also redo Exercise 1.9.1 in one sentence.

4 The variety of complete quadrics (including ML-degrees in algebraic statistics)

4.1 ML-degrees in algebraic statistics

In the one-dimensional case, in order to determine a Gaussian distribution on $\mathbb{R}$, one needs to specify its mean $\mu \in \mathbb{R}$ and its variance $\sigma \in \mathbb{R}_{>0}$. In the n-dimensional case, the mean is a vector $\mu \in \mathbb{R}^n$, and the second parameter is a positive-definite $n \times n$ covariance matrix Σ. The corresponding Gaussian distribution on $\mathbb{R}^n$ is given by

$$f_{\mu,\Sigma}(x) := \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)},$$

where ^{T} denotes the transpose.

A sample of such a distribution is just a point in $\mathbb{R}^n$. Given many samples, we may want to reconstruct μ and Σ. But what does that mean? Theoretically, any μ and Σ could have lead to the samples we obtained.

In statistics one looks for the parameters that *maximize the likelihood function.* In our case this is the product $\prod_i f_{\mu,\Sigma}(s_i)$, where $s_i \in \mathbb{R}^n$ are the samples we observed. Finding μ and Σ that maximize this function exactly means finding the parameters that best explain our observation. As intuition dictates, it turns out that the best μ is the average of the s_i's. The optimal Σ_0 is also determined (as an average of rank one covariance matrices obtained from each sample).

Let V be an n-dimensional vector space. We can identify $S^2(V)$ with the space of symmetric $n \times n$-matrices. In what follows we will focus on $\Sigma \in S^2V$, but we will not allow it to be an arbitrary symmetric matrix. In statistics this is called choosing a model, and we will study the linear concentration model. The symmetric matrix $K := \Sigma^{-1} \in S^2V^*$ is know as the *concentration matrix.* In the linear concentration model we want K to belong to a fixed linear subspace $\mathcal{L} \subset S^2V^*$ of symmetric matrices. In other words $\Sigma \in \mathcal{L}^{-1} = \{K^{-1} : K \in \mathcal{L}\} \subset S^2V$. In fact in the previous definition we only invert the invertible matrices in $\mathcal{L}$ and most often we abuse the notation and let $\mathcal{L}^{-1}$ be the closure of the set $\{K^{-1} : K \in \mathcal{L}\}$. Hence our aim now is not to maximize the likelihood function over the space S^2V, but over the variety $\mathcal{L}^{-1}$.

Theorem 4.1.1. *The maximum likelihood estimate for the linear concentration model is the unique positive definite matrix Σ' that has the same image as Σ_0 under the projection $\pi : S^2V \to S^2V/\mathcal{L}^\perp$. Here $\mathcal{L}^\perp$ is the (linear) space of linear forms that vanish on $\mathcal{L}$.*

The map $\pi_{|\mathcal{L}^{-1}} : \mathcal{L}^{-1} \to S^2V/\mathcal{L}^\perp$ is finite but not bijective, i.e. there could be many matrices in $\pi^{-1}(\pi(\Sigma_0)) \cap \mathcal{L}^{-1}$, but the one that is positive definite is the maximum likelihood estimate. However it is the cardinality of the fiber that measures the complexity of describing Σ' in terms of Σ_0.

Definition 4.1.2. The maximum likelihood degree is the degree of the (finite) map $\pi_{|\mathcal{L}^{-1}}$.

For $\mathcal{L}$ general, the ML-degree depends on: $d = \dim \mathcal{L}$ and n. It is denoted by $\phi(n, d)$.

Theorem 4.1.3 (Teissier [Tei73, LTT88])**.** $\phi(n, d) = \deg \mathcal{L}^{-1}$

Rather than proving this theorem it is very important (exercise!) to realize why it is not obvious! For more elementary proofs of this theorem we refer to [SU10, KWJ20].

Passing to the complex projective space $\phi(n, d)$ is the number of pairs $(K, \Sigma) \in \mathbb{P}(S^2V) \times \mathbb{P}(S^2V^*)$ with

$$\Sigma \cdot K = Id_n,\ K \in \mathcal{L},\ \Sigma \in \mathcal{M},$$

where $\mathcal{L} \subset \mathbb{P}(S^2V)$ and $\mathcal{M} \subset \mathbb{P}(S^2V^*)$ are general linear subspaces, of dimension $(d-1)$ respectively codimension $(d-1)$.

Let $X \subset \mathbb{P}(S^2V) \times \mathbb{P}(S^2V^*)$ be the variety parametrized by (K, K^{-1}). The points of X are all pairs $(K, \Sigma) \in \mathbb{P}(S^2V) \times \mathbb{P}(S^2V^*)$ with $K \cdot \Sigma = Id_n$ or $K \cdot \Sigma = 0$.

The computation of $\phi(n, d)$ now reduces to computation in the Chow ring of X. However, a tachinical difficulty arises: X is singular. We will replace it with a smooth variety: *the space of complete quadrics.*

One could ask why we cannot do the computation of the quadrics directly on $\mathbb{P}(S^2V)$. Let us present the following example.

Example 4.1.4. Consider the five dimensional projective space of plane quadrics in three variables $\mathbb{P}(S^2\mathbb{C}^3) = \mathbb{P}(S^2V)$.

First we ask how many quadrics pass through five general points? Passing through one point is a hyperplane condition in $\mathbb{P}(S^2V)$. After intersecting five such, we obtain precisely one quadric. Thus, as we know, there is precisely one quadric passing through five general points.

Now let us ask how many quadrics are tangent to five general lines. First note that a quadric $Q \in \mathbb{P}(S^2V)$ is tangent to a line (which is also a hyperplane) $l \subset \mathbb{P}(V)$ if and only if the dual quadric $Q^\vee \in \mathbb{P}(S^2V^*)$ passes through a point corresponding to $l \in \mathbb{P}(V^*)$. This is a *quadratic* condition in the entries of Q, i.e. there is a *quadric* $S_l \subset \mathbb{P}(S^2V)$, which points correspond to quadrics that are tangent to l. Let us now provide two answers to the original question.

Answer 1: We have to intersect five quadrics in the space $\mathbb{P}(S^2V)$, thus by Bezout's theorem we get $2^5 = 32$ points.

Answer 2: We do the computation in $\mathbb{P}(S^2V^*)$, i.e. instead of counting Q's that satisfy the condition we count the $Q^\vee$'s. Here, the five lines correspond to five general points. Hence, it is a problem that we already answered: there is one $Q^\vee$.

Clearly $1 \neq 32$. What went wrong? Which answer is correct?

The problem is that the five quadrics in Answer 1 do not intersect in 32 points. These quadrics are not general. They are (general) linear combinations of 2×2 minors. In particular they all vanish on the variety $V_1 \subset \mathbb{P}(S^2V)$ of rank one matrices. What actually happens is that if we intersect five of those, we will get the variety V_1 and... one point, corresponding to the unique quadric Q that is tangent to five lines, which we found in Answer 2.

Why didn't we have this problem when computing the number of quadrics going through five points? After all here the hyperplane conditions are also not complete general. However, in this case our hyperplanes do not have the base locus (i.e. they do not vanish on some variety). Put yet differently: in the first case of points the intersection is transversal, while in the second case of lines it is not.

Doing such computations, when the system has base locus belongs to the domian know as *excess intersection theory.* This is hard. What best works in practice is to *change* the space on which we perform intersection in order to avoid the base locus. This is precisely what we did in Answer 2 and what we will do to compute $\phi(n, d)$.

4.2 The space of complete quadrics

For $A \in S^2(V)$, we write $\wedge^k A \in S^2(\wedge^k V)$ for the k-th *compound matrix.* In coordinates, $\wedge^k A$ is the $\binom{n}{k} \times \binom{n}{k}$-matrix whose entries are the $k \times k$-minors of A. The coordinate-free description is as follows: if we view A as a linear map $V^* \to V$, then $\wedge^k A : \wedge^k(V)^* \cong \wedge^k(V^*) \to \wedge^k V$. Note that $\wedge^{n-1} A = \operatorname{adj}(A)$, the adjugate matrix of A.

Definition 4.2.1. The variety $\mathcal{CQ}(V)$ of *complete quadrics* is the closure of the image of the set of invertible matrices under the map

$$\varphi : \mathbb{P}(S^2V) \dashrightarrow \mathbb{P}\left(S^2V\right) \times \mathbb{P}\left(S^2(\bigwedge^2 V)\right) \times \cdots \times \mathbb{P}\left(S^2(\bigwedge^{n-1} V)\right),$$

sending a matrix A to $(A, \bigwedge^2 A, \dots, \bigwedge^{n-1} A)$.

So a point in $\mathcal{CQ}(V)$ is given by a tuple $(A_1, \dots, A_{n-1})$ of symmetric matrices $A_i \in \mathbb{P}\left(S^2(\bigwedge^k V)\right)$. For a general point, all these matrices will be invertible and $A_i = \wedge^k A_1$. However, since in the definition we took a closure, there are other points in $\mathcal{CQ}(V)$.

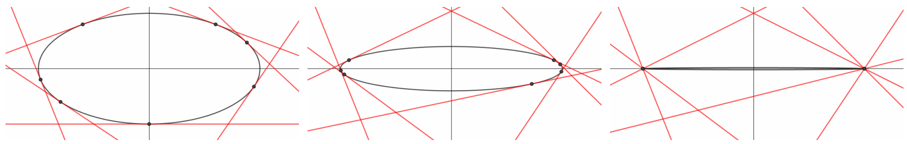


Figure 1: A conic degenerating to a double line with two marked points.

Example 4.2.2. Let $V = \mathbb{C}^3$. For every $\varepsilon \in \mathbb{C}^*$, the point

$$(\begin{pmatrix} -\varepsilon & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} -\frac{1}{\varepsilon} & 0 & 0 \\ 0 & \frac{1}{\varepsilon} & 0 \\ 0 & 0 & 1 \end{pmatrix}) = (\begin{pmatrix} -\varepsilon & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \varepsilon \end{pmatrix}) \in \mathbb{P}(S^2 V) \times \mathbb{P}(S^2(V^*))$$

is contained in $\mathcal{CQ}(V)$. Taking the limit $\varepsilon \to 0$, we find that

$$(\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}) \in \mathcal{CQ}(V).$$

The name *complete quadrics* is explained by the following geometric interpretation: a matrix $A \in S^2(V)$ defines a quadric hypersurface Q in $\mathbb{P}(V^*)$; cut out by the equation $x^T Ax$. More generally, the matrix $\wedge^k A$ defines a quadric hypersurface Q_k in the Grassmannian $\mathbb{G}(k-1, \mathbb{P}(V^*))$, cut out by the equation $P^T \wedge^k AP$ in the Plücker coordinates.

Lemma 4.2.3. *Geometrically, Q_k is the space of $(k-1)$-planes tangent to $Q = Q_0$.*

Proof. The space $W = \langle w_1, \dots, w_k \rangle$ is tangent to the quadric Q if and only if there exist $p_1, \dots, p_k \in \mathbb{C}$ not all equal to zero, such that $Q(\sum_{i=1}^k p_i w_i, w) = 0$ for any $w \in W$ (where in the last equality we treat Q as a 2-form). This happens if and only if the $k \times k$ matrix with (i,j)-entry $Q(w_i, w_j)$ is degenerate, i.e. the determinant of that matrix equals zero. But that determinant equals $(\bigwedge^k Q)(\bigwedge_{i=1}^k w_i, \bigwedge_{i=1}^k w_i)$, i.e. the evaluation of the quadric $\bigwedge^k Q$ on the point of the Grassmannian , $\bigwedge_{i=1}^k w_i$ corresponding to the space W. □

So a point $\mathcal{A} = (A_1, \dots, A_{n-1}) \in \mathcal{CQ}(V)$ is given by a collection of (possibly nonsmooth) quadrics $Q_k = Q(A_k) \subset \mathbb{G}(k-1, \mathbb{P}V^*)$. For a general $\mathcal{A}$, all Q_k are smooth, and Q_k is the space of $(k-1)$-planes tangent to Q_1.

Example 4.2.4. Let us revisit Example 4.2.2: if $A_0 = \begin{pmatrix} -\varepsilon & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & 1 \end{pmatrix}$, then $Q_0 \subset \mathbb{P}(V^*)$ is the smooth conic with equation $\varepsilon x_0^2 = \varepsilon x_1^2 + x_2^2$. $Q_1 \subseteq \mathbb{P}(V)$ is the dual conic; it is the space of lines tangent to Q_0, and has equation $\beta_0^2 = \beta_1^2 + \varepsilon \beta_2^2$, where β_i are the coordinates on V.

For people familiar with blow-ups, we give an alternative construction of the space of complete quadrics. Start with the space $\mathbb{P}(S^2V)$ and consider the following sequence of blow-ups

$$\mathbb{P}(S^2V) = X_0 \longleftarrow X_1 \longleftarrow \cdots \longleftarrow X_{n-2} =: \mathcal{CQ}(V), \tag{2}$$

where X_i is the blow-up of X_{i-1} at the strict transform of the locus of rank i matrices.

Below we present one more point of view on points of $\mathcal{CQ}(V)$. The general idea is the following inductive description: a point $p \in \mathcal{CQ}(V)$ is a quadric Q on V and a point on $\mathcal{CQ}(V/\operatorname{im} Q)$.

Given a flag $\emptyset = U_0 \subset U_1 \subset \ldots \subset U_k = V$, and for every $j \in \{1, \ldots, k\}$ a smooth quadric (i.e. a full symmetric matrix) $M_j \in \mathbb{P}(S^2(U_j/U_{j-1}))$, we can construct a complete quadric as follows. We will write $d_j := \dim U_j$. For every $j \in \{1, \ldots, k\}$ and $\ell \leq d_j - d_{j-1}$, we have a linear map

$$\varphi_{j,\ell} : S^2\left(\bigwedge^{\ell}(U_j/U_{j-1})\right) \to S^2\left(\bigwedge^{d_{j-1}+\ell} V\right)$$

induced from the map $\bigwedge^{\ell}(U_j/U_{j-1}) \to \bigwedge^{d_{j-1}+\ell} V$ sending $v_1 \wedge \ldots \wedge v_\ell$ to $u_1 \wedge \ldots \wedge u_{d_{j-1}} \wedge v_1 \wedge \ldots \wedge v_\ell$, where $u_1, \ldots, u_{d_{j-1}}$ is a basis of U_{j-1}. Then the point

$$\begin{aligned}\big(\varphi_{1,1}(M_1), \ldots, \varphi_{1,d_1}(M_1), \varphi_{2,1}(M_2), \ldots, \varphi_{2,d_2-d_1}(M_2), \ldots, \varphi_{k,n-d_{k-1}-1}(M_k)\big) \\ \in \mathbb{P}(V) \times \mathbb{P}(\wedge^2 V) \times \cdots \times \mathbb{P}(\wedge^{n-1}V) \end{aligned} \tag{3}$$

lies in $\mathcal{CQ}(V)$. One can verify that this gives rise to a one-to-one correspondence

$$\left\{ \begin{array}{l} \text{flags } \emptyset = U_0 \subset U_1 \subset \ldots \subset U_k = V \\ \text{and } \forall j \in \{1, \ldots, k\} : M_j \in \mathbb{P}(S^2(U_j/U_{j-1})) \end{array} \right\} \overset{1:1}{\longleftrightarrow} \mathcal{CQ}(V). \tag{4}$$

Here is a geometric interpretation: the matrix $M_j \in \mathbb{P}(S^2(U_j/U_{j-1}))$ correponds to a quadric $q_j \subseteq V^*$ which is contained in $U_{j-1}^{\perp} \subseteq V^*$ and is a cone over $U_j^{\perp}$. The corresponding complete quadric is then a tuple $(Q_1, \ldots, Q_{n-1})$, where for $d_j < k \leq d_{j+1}$, $Q_i \subseteq \mathbb{G}(k-1, \mathbb{P}(V^*))$ is a the quadric of $(k-1)$-planes Λ for which $\Lambda \cap U_{j-1}^{\perp}$ is tangent to q_j.

Example 4.2.5. Consider $n = 2$, ie. 2×2 symmetric matrices. The space $\mathbb{P}(S^2V)$ is two dimensional. The map $\mathbb{P}(S^2V) \to \mathbb{P}(S^2V^*)$ is regular and in fact a linear isomorphism. Hence, in this case the graph of the map and the space of complete quadrics are simply $\mathcal{CQ}(V) = \mathbb{P}^2$. The pull-backs of the hyperplanes in $\mathbb{P}(S^2V)$ and $\mathbb{P}(S^2V^*)$ are simply hyperplanes in $\mathcal{CQ}(V)$. The rank $n-1=1$ matrices form a quadric (defined by the determinant of the 2×2 symmetric matrix).

Example 4.2.6. Consider $n = 3$, ie. 3×3 symmetric matrices. The space $\mathbb{P}(S^2V)$ is five dimensional. The map $\mathbb{P}(S^2V) \dashrightarrow \mathbb{P}(S^2V^*)$ is rational, not

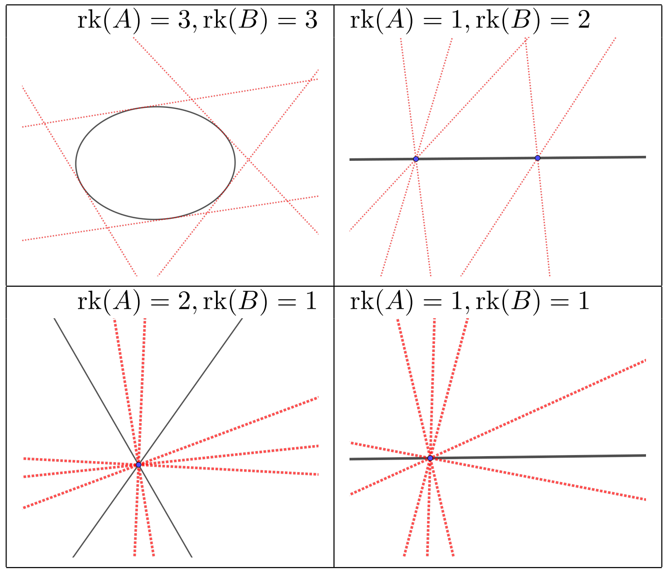


Figure 2: The four types of complete conics

defined on the locus of rank $n-2=1$ matrices. In this case the graph of the map and the space of complete quadrics are isomorphic and are the blow-up of rank 1 matrices. A point of $\mathcal{CQ}(V)$ is a pair of matrices $(A,B) \in \mathbb{P}(S^2V) \times \mathbb{P}(S^2V^*)$ such that either:

- A has rank 3 or 2 and B is the adjugate matrix. Note that if A has rank 3 then so does B and hence A may be reconstructed from B. However, if A has rank 2 then B has rank 1 and all matrices with the same image as A may be paired with B.

- A has rank 1 and B has rank 2. This is dual to the case of rank 2 matrix above.

- A has rank 1 and B has rank 1. Here B is a rank one matrix that vanishes on the image of A.

Remark 4.2.7. There is an isomorphism $\mathcal{CQ}(V) \cong \mathcal{CQ}(V^*)$, induced from

$$\mathbb{P}\left(S^2V\right) \times \cdots \times \mathbb{P}\left(S^2(\bigwedge^{n-1} V)\right) \to \mathbb{P}\left(S^2V^*\right) \times \cdots \times \mathbb{P}\left(S^2(\bigwedge^{n-1} V^*)\right)$$
$$(A_1, \ldots, A_{n-1}) \mapsto (A_{n-1}, \ldots, A_1),$$

where we used $S^2 \bigwedge^k V \cong S^2 \bigwedge^{n-k} V^*$.

4.3 Degeneration spaces

For each $r \in \{1, \ldots, n-1\}$, we define a subvariety $S_r \subset \mathcal{CQ}(V)$:

$$\begin{aligned} S_r =& \{(A_1, \ldots, A_{n-1}) \in \mathcal{CQ}(V) \mid \operatorname{rk} A_r = 1\} \\ =& \overline{\{(A_1, \ldots, A_{n-1}) \in \mathcal{CQ}(V) \mid \operatorname{rk} A_1 = r\}} \\ =& \overline{\{(A_1, \ldots, A_{n-1}) \in \mathcal{CQ}(V) \mid \operatorname{rk} A_{n-1} = n-r\}}. \end{aligned}$$

Under the correspondence (4), S_r consists of all complete quadrics whose flag contains a space U_j of dimension r. Alternatively, S_r is the strict transform of the exceptional locus of the blow-up $X_{r-1} \longleftarrow X_r$ in (2). From this last description it follows that $S_r \subset \mathcal{CQ}(V)$ is irreducible of codimension one. A general point in S_r is determined by a rank one matrix A_1 and rank $n-r$ matrix A_{n-1} that vanishes on the image of A_1. The image of A_1 (resp. kernel of A_{n-1}) is parameterized by the Grassmannian $G(r,n)$. The following computation also shows that S_r is a divisor:

$$(n-r)r + \binom{r+1}{2} - 1 + \binom{n-r+1}{2} - 1 = \binom{n+1}{2} - 1 - 1.$$

4.4 ML-degree via complete quadrics

$\mathcal{CQ}(V)$ is a smooth variety that fits in the diagram

$$\begin{array}{ccc} & \mathcal{CQ}(V) & \\ {\scriptstyle \pi_1} \swarrow & & \searrow {\scriptstyle \pi_{n-1}} \\ \mathbb{P}(S^2V) & \underset{A \mapsto A^{-1}}{\dashrightarrow} & \mathbb{P}(S^2V^*) \end{array}$$

We can express our ML-degrees $\phi(n,d)$ as follows: $\phi(n,d) = |\pi_1^{-1}(\mathcal{L}) \cap \pi_{n-1}^{-1}(\mathcal{M})|$, where $\mathcal{L} \subseteq \mathbb{P}(S^2V)$ and $\mathcal{M} \subseteq \mathbb{P}(S^2V^*)$ are generic linear subspaces with $\dim(\mathcal{L}) = d-1$ and $\operatorname{codim}(\mathcal{M}) = d-1$.

Thus, we need to compute the product $[\pi_1^{-1}(\mathcal{L})][\pi_{n-1}^{-1}(\mathcal{M})]$ in the Chow ring of $\mathcal{CQ}(V)$. In this lecture, we will not give a complete description of this Chow ring. Instead, we will only describe the Picard group (which is the degree 1 part of the Chow ring), and then explain how our computation on $A(\mathcal{CQ}(V))$ can be reduced to a computation on the Chow ring of the Grassmannian, which we know.

Remark 4.4.1. It is possible to construct an affine stratification of $\mathcal{CQ}(V)$ [Str86], and hence by Theorem 1.5.2 give a $\mathbb{Z}$-basis of its Chow ring. Describing the ring structure is more complicated, see for instance [DCGMP88].

For $i \in \{1, \ldots, n-1\}$, we define a class $L_i \in A^1(\mathcal{CQ}(V))$ as $L_i = \pi_i^*([H])$, where $H \subseteq \mathbb{P}(S^2 \bigwedge^i V)$ is a general hyperplane. Abusing notation, we will denote the class of the degeneration space $S_r \subset \mathcal{CQ}(V)$ also by $S_r \in A^1(\mathcal{CQ}(V))$.

Proposition 4.4.2. *The classes $L_1, \dots, L_{n-1}$ form a basis of* $\mathrm{Pic}(\mathcal{CQ}(V))$, *in which the classes $S_1, \dots, S_{n-1}$ are given by the relations*

$$S_i = -L_{i-1} + 2L_i - L_{i+1},$$

with $L_0 = L_n := 0$.

Proof. These relations were already known to Schubert [Sch94]. For a modern treatment, see for example [Mas20, Proposition 3.6 and Theorem 3.13]. □

Proposition 4.4.3.

$$\phi(n,d) = \deg(L_1^{\binom{n+1}{2}-d} L_{n-1}^{d-1})$$

Let us use the relations from Proposition 4.4.2 to rewrite this:

$$\begin{aligned}
\phi(n,d) &= \int_{CQ_n} L_1^{\binom{n+1}{2}-d} L_{n-1}^{d-1} \\
&= \frac{1}{n} \int_{CQ_n} L_1^{\binom{n+1}{2}-d-1} L_{n-1}^{d-1} \sum_{s=1}^{n-1} s S_{n-s} \qquad (5) \\
&= \frac{1}{n} \sum_{s=1}^{n-1} s \int_{CQ_n} L_1^{\binom{n+1}{2}-d-1} L_{n-1}^{d-1} S_{n-s}
\end{aligned}$$

Perhaps surprisingly, the products $\int_{CQ_n} L_1^{\binom{n+1}{2}-d-1} L_{n-1}^{d-1} S_{n-s}$ will be easier to understand than $\int_{CQ_n} L_1^{\binom{n+1}{2}-d} L_{n-1}^{d-1}$. We give them a name:

Definition 4.4.4. We define:

$$\delta(m,n,r) = \int_{S_r} L_1^{\binom{n+1}{2}-m-1} L_{n-1}^{m-1} = \int_{CQ_n} S_r L_1^{\binom{n+1}{2}-m-1} L_{n-1}^{m-1}.$$

Now (5) becomes

$$\phi(n,d) = \frac{1}{n} \sum_{s=1}^{n-1} s\delta(d, n, n-s) \qquad (6)$$

Remark 4.4.5. The number δ is known as the algebraic degree of semidefinite programming (SDP). It appears while studying the degrees of algebraic extensions of numbers that are solutions to optimization problems of maximizing a linear function over spectrahedra. There is also a purely geometric definition of $\delta(m,n,r)$: is the dual degree of the variety obtained by intersecting the locus of rank $\leq r$ symmetric $n \times n$ matrices with a projective m-plane. See [MMM+20, Definition 1.4], [NRS10], [GvBR09].

Our next aim is to get a better understanding of $\delta(m,n,r)$. For that we will change the variety S_r to something more familiar. This is similar to passing from $\mathbb{P}(S^2V)$ to $\mathcal{CQ}(V)$. However currently our problem is not the base locus, but the fact that S_r is quite complicated. Thus, we will build a variety that is birational with S_r and such that the systems corrsponding to L_1 and L_{n-1} are also basepoint free. As we are only interested in the number of points we get in the intersection we will solve the problem on the new variety instead of S_r.

The first indication how to carry this out, was by describing the general point of S_r as a pair of matrices (A_1, A_{n-1}). First we needed to choose the image of A_1/kernel of A_{n-1} which is a point $[W] \in G(r,V)$. Next, we need to choose A_1. It is a symmetric linear map $V^* \to V$ with image contained in W. In particular, it descends to (and is characterized by) the symmetric map $W^* \to W$. In other words, we choose $A_1 \in \mathbb{P}(S^2\mathcal{U})_{[W]}$. Dually (i.e. replacing V by V^*) we see that the choice of A_{n-1} corresponds to choosing a point in $\mathbb{P}(S^2\mathcal{Q}^*)_{[W]}$.

What we have described above is a birational map

$$CQ(V) \dashrightarrow \mathbb{P}(S^2\mathcal{U}) \times_{G(r,V)} \mathbb{P}(S^2\mathcal{Q}^*) := X.$$

The variety X is not a (projectivized) vector bundle. Rigorously it may be defined as a subvariety of $\mathbb{P}(S^2\mathcal{U} \otimes S^2\mathcal{Q}^*)$ consisting of rank one tensors, or as a fiber product of the two projective bundles.

The next step is to understand what L_1 and L_{n-1} give on X. The divisor L_1 is a hyperplane condition on $\mathbb{P}(S^2V)$, i.e. it is a section of $\mathcal{O}(1)$. Restricting $\mathcal{O}(1)$ from $\mathbb{P}(\underline{S^2V}) = \mathbb{P}(S^2V) \times G(r,V)$ to $S^2\mathcal{U}$ we see that L_1 is the divisor corresponding to $c_1(\mathcal{O}(1))$. Consider the following commutative diagram:

$$\begin{array}{ccc} X & \xrightarrow{\pi_1} & \mathbb{P}(S^2\mathcal{U}) \\ \downarrow{\scriptstyle \pi_2} & & \downarrow{\scriptstyle i_1} \\ \mathbb{P}(S^2\mathcal{Q}^*) & \xrightarrow{i_2} & Gr(r,V) \end{array}$$

The discussion above proves the following lemma.

Lemma 4.4.6.

$$\delta(m,n,r) = \pi_1^*(c_1(\mathcal{O}_{S^2\mathcal{U}}(1))^{\binom{n+1}{2}-m-1})\pi_2^*(c_1(\mathcal{O}_{S^2\mathcal{Q}^*}(1))^{m-1})$$

Formally the class on the right needs to be pushed-forward to a point. In particular, we may first push it forward to $G(r,n)$:

$$(i_1)_*(\pi_1)_*(\pi_1^*(c_1(\mathcal{O}_{S^2\mathcal{U}}(1))^{\binom{n+1}{2}-m-1})\pi_2^*(c_1(\mathcal{O}_{S^2\mathcal{Q}^*}(1))^{m-1}))$$

In order to provide a more closed formula we present two general lemmas [Ful98, Proposition 1.7, Theorem 3.2].

Lemma 4.4.7. *Consider the fiber product diagram, where g_1, g_2 are flat and f_1, f_2 are proper:*

$$\begin{array}{ccc} X & \xrightarrow{f_2} & Y \\ \downarrow{\scriptstyle g_2} & & \downarrow{\scriptstyle g_1} \\ Z & \xrightarrow{f_1} & W \end{array}$$

Then $(f_2)_* g_2^* = g_1^*(f_1)_*$ as functions of Chow groups $A(Z) \to A(Y)$.

Lemma 4.4.8. *For a proper map $f : Y \to X$ we have:*

$$f_*(f^*(x)y) = xf_*(y),$$

for any $x \in A(X)$ and $y \in A(Y)$.

We now have:

$$\begin{aligned} (i_1)_*(\pi_1)_*\left(\pi_1^*(c_1(\mathcal{O}_{S^2\mathcal{U}}(1))^{\binom{n+1}{2}-m-1})\pi_2^*(c_1(\mathcal{O}_{S^2\mathcal{Q}^*}(1))^{m-1})\right) = \\ (i_1)_*\left((c_1(\mathcal{O}_{S^2\mathcal{U}}(1))^{\binom{n+1}{2}-m-1})((\pi_1)_*\pi_2^*(c_1(\mathcal{O}_{S^2\mathcal{Q}^*}(1))^{m-1}))\right) = \\ (i_1)_*\left((c_1(\mathcal{O}_{S^2\mathcal{U}}(1))^{\binom{n+1}{2}-m-1})(i_1^* i_{2*}(c_1(\mathcal{O}_{S^2\mathcal{Q}^*}(1))^{m-1}))\right) = \\ \left((i_1)_*(c_1(\mathcal{O}_{S^2\mathcal{U}}(1))^{\binom{n+1}{2}-m-1}\right)\left(i_{2*}(c_1(\mathcal{O}_{S^2\mathcal{Q}^*}(1))^{m-1}))\right) \end{aligned} \tag{7}$$

Using the definition of the Segre class we obtain the following proposition.

Proposition 4.4.9. *If $\binom{n-r+1}{2} \le m \le \binom{n+1}{2} - \binom{r+1}{2}$* [4]*, then*

$$\delta(m,n,r) = \int_{G(r,n)} \mathrm{Seg}_{(\binom{n+1}{2}-m-\binom{r+1}{2})}(S^2\mathcal{U})\,\mathrm{Seg}_{(m-\binom{n-r+1}{2})}(S^2\mathcal{Q}^*),$$

and otherwise $\delta(m,n,r) = 0$

4.5 Exercises

Exercise 4.5.1. Express (multiplicities of) S_i's in terms of L_i's.

Exercise 4.5.2. Compute $\delta(3,3,2)$.

Exercise 4.5.3. Compute $\phi(4,3)$.

5 Lascoux coefficients, polynomiality

In the previous lecture, we introduced the ML-degree $\phi(n,d)$, interpreted it as an intersection number on the space of complete quadrics, and reduced the computation of $\phi(n,d)$ to a computation in the Chow ring of the Grassmannian. In this final lecture, we will proceed to give an explicit combinatorial formula for $\phi(n,d)$. This can be used to compute the ML-degree effectively, and will also allow us to prove a conjecture by Sturmfels and Uhler that for fixed d, $\phi(n,d)$ is a polynomial in n.

[4]these are known as *Pataki's inequalities*

5.1 The Lascoux coefficients

In order to compute the product (4.4.9), we first consider the Segre classes of $S^2\mathcal{U}$, where $\mathcal{U}$ is the universal subbundle of the Grassmannian $G(r,n)$. Like any element of $A(G(k,n))$, we can expand $\mathrm{Seg}_d(S^2\mathcal{U})$ in Schubert classes:

$$\mathrm{Seg}_d(S^2\mathcal{U}) = \sum_\lambda c_\lambda \sigma_\lambda \tag{8}$$

where the sum is over all Young diagrams λ of weight d that fit inside an $r \times (n-r)$ box, and the coefficients $c_\lambda \in \mathbb{Z}$ (which a priori could depend on n and r) remain to be determined.

Definition 5.1.1. The *Lascoux coefficients* ψ_λ (where λ is a partition, possibly with trailing zeroes) are the coefficients appearing in the Schur expansion of the complete homogeneous symmetric polynomial in sums of variables $x_i + x_j$:

$$s_{[d]}(\{x_i + x_j \mid 1 \le i \le j \le r\}) = \sum_\lambda \psi_\lambda s_\lambda(x_1, \dots, x_r), \tag{9}$$

where the sum is over all partitions of weight d and length r.

Proposition 5.1.2. *The coefficients appearing in (8) are the Lascoux coeffients. Precisely, we have*

$$\mathrm{Seg}_d(S^2\mathcal{U}) = \sum_\lambda \psi_\lambda \sigma_\lambda, \tag{10}$$

where the sum is over all partitions λ of weight d and length r with $\lambda_1 \le n-r$.

Proof. The Chern roots of $\mathcal{U}$ are $-x_1, \dots, -x_r$, hence the Chern roots of $S^2\mathcal{U}$ are $\{-x_i - x_j \mid 1 \le i \le j \le r\}$. Hence by Proposition 3.2.6, the d'th Segre class of $S^2\mathcal{U}$ equals

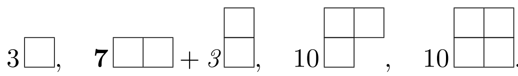
$$\mathrm{Seg}_d(S^2\mathcal{U}) = (-1)^d s_{[d]}(\{-x_i - x_j \mid 1 \le i \le j \le r\}) = s_{[d]}(\{x_i + x_j \mid 1 \le i \le j \le r\}). \tag{11}$$

The result now follows from the definition of ψ_λ and the identification of $A(G(k,n))$ with a quotient of the ring of symmetric polynomials. □

Remark 5.1.3.

- Recall that we allowed partitions to have trailing zeroes. This is important: in contrast to for example the Littlewood-Richardson coefficients, adding trailing zeroes to a partition λ might change the Lascoux coefficient ψ_λ.
- It will often be useful to use the bijection $\lambda \mapsto I(\lambda)$ from Section 1.6 and index the Lascoux coefficients by subsets of $\mathbb{N}$ instead of partitions. We will simply write ψ_I instead of $\psi_{\lambda(I)}$. This is the notation usually used in the literature [LLT89, MMM$^+$20, Sey].
- We note that Lascoux coefficients appear in many publications with different notations. In particular one needs to be careful with the shift: $\psi_{\{j_1,\dots,j_r\}}$ as defined above equals $\psi_{\{j_1+1,\dots,j_r+1\}}$ in [GvBR09]. On the other hand our notation is consistent with [LLT89, NRS10].

Example 5.1.4. In Exercise 3.4.2 we computed the Segre classes of $S^2\mathcal{U}$ on $G(2,4)$:



Their coefficients are the Lascoux coefficients, namely:

$$\psi_{0,2} = 3, \quad \psi_{0,3} = \mathbf{7}, \quad \psi_{1,2} = \mathit{3}, \quad \psi_{1,3} = 10, \quad \psi_{2,3} = 10.$$

We use boldface and emphasis above and below to indicate the same numbers. We may also compute them by expanding complete symmetric polynomials, where now x_1, x_2 are simply formal variables.

$$s_{[2]}(2x_1, x_1 + x_2, 2x_2) = 7x_1^2 + 7x_2^2 + 10x_1x_2 =$$

$$= 7(x_1^2 + x_1x_2 + x_2^2) + 3x_1x_2 = \mathbf{7}s_{[2,0]}(x_1,x_2) + \mathit{3}s_{[1,1]}(x_1,x_2).$$

5.2 Combinatorial formulas for ϕ and δ

Recall there is an isomorphism $\Phi : G(r,V) \cong G(n-r,V^*)$, sending an r-plane $\Lambda \subset V$ to the plane $\Lambda^\perp$ of linear functions vanishing on Λ. It identifies the dual quotient bundle $\mathcal{Q}^*$ on $G(r,V)$ with the universal subbundle $\mathcal{U}$ on $G(n-r,V^*)$. Moreover, Φ sends the Schubert cell Σ_λ of $G(r,V)$ to the Schubert cell Σ_{λ^T} of $G(n-r,V^*)$.

Hence we can express the Segre class of $\mathrm{Seg}_d(S^2\mathcal{Q}^*)$ in terms of our Lascoux coefficients:

$$\mathrm{Seg}_d(S^2\mathcal{Q}^*) = \sum_\lambda \psi_{\lambda^T}\sigma_\lambda$$

where the sum is over all Young diagrams of weight d that fit inside an $(n-r)\times r$ box.

We can now write down a combinatorial formula for δ:

$$\begin{aligned}\delta(m,n,r) &= \int_{G(r,n)} \mathrm{Seg}_{\left(\binom{n+1}{2} - m - \binom{r+1}{2}\right)}(S^2\mathcal{U})\,\mathrm{Seg}_{\left(m - \binom{n-r+1}{2}\right)}(S^2\mathcal{Q}^*) \\ &= \int_{G(r,n)} \left(\sum_{\lambda \vdash \binom{n+1}{2} - m - \binom{r+1}{2}} \psi_\lambda\sigma_\lambda\right)\left(\sum_{\mu \vdash m - \binom{n-r+1}{2}} \psi_{\mu^T}\sigma_\mu\right) \\ &= \sum_{\mu \vdash m - \binom{n-r+1}{2}} \psi_{\mu^c}\psi_{\mu^T},\end{aligned}$$

where in the last equality we used Exercise 1.9.3. It is convenient to switch our notation to subsets: letting $I = I(\mu^T)$, we have $\#(I) = n-r$ and $\sum I = m - \binom{n-r+1}{2} + \binom{n-r}{2} = m - n + r$. Moreover one can easily verify that for λ a partition fitting inside an $r \times (n-r)$ box, we have $I((\lambda^c)^T) = [n] \setminus I(\lambda)$. Hence our formula becomes

Theorem 5.2.1 ([GvBR09, Theorem 1.1]). *For $0 < m < \binom{n+1}{2}$ and $0 < r < n$,*

$$\delta(m,n,r) = \sum_{\substack{I\subset[n] \\ \#I=n-r \\ \sum I=m-n+r}} \psi_I \psi_{[n]\setminus I}.$$

As an immediate corollary, we get a formula for $\phi(n,d)$:

Corollary 5.2.2.

$$\phi(n,d) = \frac{1}{n} \sum_{1\le\binom{s+1}{2}\le d} s \sum_{\substack{I\subset[n] \\ \#I=s \\ \sum I=d-s}} \psi_I \psi_{[n]\setminus I}$$

Proof.

$$\begin{aligned}
\phi(n,d) &= \frac{1}{n}\sum_{s=1}^{n-1} s\delta(d,n,n-s) \\
&= \frac{1}{n} \sum_{1\le\binom{s+1}{2}\le d} s\delta(d,n,n-s) \\
&= \frac{1}{n} \sum_{1\le\binom{s+1}{2}\le d} s \sum_{\substack{I\subset[n] \\ \#I=s \\ \sum I=d-s}} \psi_I \psi_{[n]\setminus I}
\end{aligned}$$

where the second equality follows from Pataki's inequality: $\delta(d,n,n-s) = 0$ whenever $\binom{s+1}{2} \ge d$. ☐

5.3 More on the Lascoux coefficients

There are several combinatorial formulas and recursive relations between the Lascoux coefficients.

Theorem 5.3.1. *1. For $\#I = 1$, we have $\psi_I = \psi_{\{i\}} = 2^i$.*

2. For $\#I = 2$, we have $\psi_I = \psi_{\{i,j\}} = \sum_{k=i+1}^{j} \binom{i+j}{k}$.

3. For any $I = \{i_0, \ldots, i_r\}$

$$(r+1)\psi_{\{i_0,\ldots,i_r\}} - 2\sum_{\ell=0}^{r} \psi_{\{i_0,\ldots,i_\ell-1,\ldots,i_r\}} = \begin{cases} \psi_{\{i_1,\ldots,i_r\}} & \text{if } i_0 = 0, \\ 0 & \text{else.} \end{cases} \tag{12}$$

where the summation is over all ℓ for which $i_\ell - 1 > i_{\ell-1}$. Together with 1. , these recursive relations determine ψ_I uniquely.

4. For $I = \{i_1 < \ldots < i_r\}$:

$$\psi_I = \sum_{J<I} \det \binom{i_k}{j_l}_{1\le k,l\le r},$$

where the sum is over all $J = \{j_1 < \ldots < j_r\} \neq I$ with $j_k \le i_k \forall k$.

5. For $r > 2$, ψ_I can be computed as

$$\psi_I = \mathrm{Pf}(\psi_{\{i_k,i_l\}})_{0<k<l\le n} \textit{ for even } r,$$

$$\psi_I = \mathrm{Pf}(\psi_{\{i_k,i_l\}})_{0\le k<l\le n} \textit{ for odd } r,$$

where $\psi_{\{i_0,i_k\}} := \psi_{\{i_k\}}$, and $\mathrm{Pf}(\psi_{\{i_k,i_l\}})$ is the Pfaffian of the skew-symmetric matrix with (k,l)th entry $\psi_{\{i_k,i_l\}}$ if $i_k < i_l$ and $-\psi_{\{i_l,i_k\}}$ if $i_k > i_l$.

6. For $j_1 = 0$ we have:

$$\psi_{\{j_1,j_2,\ldots,j_r\}} = \sum_{j_\ell \le j'_\ell < j_{\ell+1}} \psi_{\{j'_1,\ldots,j'_{r-1}\}}. \tag{13}$$

We will only prove the formulas that we will use later. For the other ones, see [LLT89, Appendix].

Proof of 1. Follows immediately from $s_{[i]}(2x_1) = 2^i s_{[i]}$. □

Proof of 3. The following proof is adapted from [Pra96, p. 163-166]. We will make use of the symmetrizing operator

$$\Delta : \mathbb{Z}[x_1,\ldots,x_{r+1}]^{\mathfrak{S}_r\times\mathfrak{S}_1} \to \mathbb{Z}[x_1,\ldots,x_{r+1}]^{\mathfrak{S}_{r+1}}$$
$$f(x_1,\ldots,x_{r+1}) \mapsto \sum_{i=1}^{r+1} \left(\frac{f(x_1,\ldots,\widehat{x_i},\ldots,x_{r+1},x_i)}{\prod_{j\neq i}(x_i - x_j)} \right)$$

We will need the following properties of Δ, which a reader with some background in symmetric polynomials can verify as an exercise:

0. Δ is linear and well-defined, i.e. the expression above is really a polynomial, and it is symmetric.

1. If $f \in \mathbb{Z}[x_1,\ldots,x_{r+1}]^{\mathfrak{S}_{r+1}}$ and $g \in \mathbb{Z}[x_1,\ldots,x_{r+1}]^{\mathfrak{S}_r\times\mathfrak{S}_1}$, then $\Delta(fg) = f\Delta(g)$.

2. $\Delta(x_{r+1}^b) = \begin{cases} 0 & \text{if } b < r, \\ 1 & \text{if } b = r, \\ x_1 + \cdots + x_{r+1} & \text{if } b = r+1. \end{cases}$

3. $\Delta(s_\lambda(x_1,\ldots,x_r)) = \begin{cases} (-1)^r s_{[\lambda_1-1,\ldots,\lambda_r-1,0]}(x_1,\ldots,x_{r+1}) & \text{if } \lambda_r > 0, \\ 0 & \text{if } \lambda_r = 0. \end{cases}$

Now for the actual proof: let's write

$$H_r = \prod_{1 \le i \le j \le r} \frac{1}{1-(x_i+x_j)} = \sum_{d=0}^{\infty} s_{[d]}(\{x_i + x_j \mid 1 \le i \le j \le r\}).$$

then we have

$$\begin{aligned} H_r &= H_{r+1} \cdot \prod_{i=1}^{r+1} (1 - x_i - x_{r+1}) \\ &= H_{r+1} \cdot \left(\sum_{d=0}^{r+1} (-1)^d e_d(x_1, \ldots, x_{r+1})(1 - x_{r+1})^{r+1-d} \right) \end{aligned}$$

by using 1 and 2 above, we find

$$\Delta(H_r) = H_{r+1} \left((r+1) - 2(x_1 + \cdots + x_{r+1}) \right).$$

Our recursive formula now follows by expanding both sides in the Schur basis (where for the left hand side we use 3 above, and for the right hand side we use Pieri's rule) and comparing coefficients. □

Proof of 6. Recall that $s_{[d]}$ is the complete homogeneous symmetric polynomial of degree d, and that we have:

$$s_{[d]}(\{x_i + x_j \mid 1 \le i \le j \le r\}) = \sum_{\substack{\lambda(I) \vdash d \\ \#I = r}} \psi_I s_{\lambda(I)}(x_1, \ldots, x_r).$$

Substituting $x_r = 0$ we obtain:

$$\sum_{i=0}^{d} s_{[i]}(\{x_i + x_j \mid 1 \le i \le j \le r-1\}) s_{[d-i]}(x_1, \ldots, x_{r-1}) =$$

$$s_{[d]}(\{x_i + x_j \mid 1 \le i \le j \le r-1\}, x_1, \ldots, x_{r-1}) = \sum_{\substack{\lambda(I) \vdash d \\ \mathrm{len}(\lambda(I)) \le r-1}} \psi_I s_{\lambda(I)}(x_1, \ldots, x_{r-1}).$$

We note that $\mathrm{len}(\lambda(I)) \le r-1$ if and only if $0 \in I$. On the other hand we may apply Pieri's rule to

$$\sum_{i=0}^{d} s_{[i]}(\{x_i + x_j \mid 1 \le i \le j \le r-1\}) s_{[d-i]}(x_1, \ldots, x_{r-1}) =$$

$$\sum_{i=0}^{d} \left(\sum_{\substack{\lambda(I) \vdash i \\ \#I = r-1}} \psi_I s_{\lambda(I)}(x_1, \ldots, x_{r-1}) \right) s_{[d-i]}(x_1, \ldots, x_{r-1}).$$

Comparing the coefficients of Schur polynomials in both expressions gives the formula. □

5.4 Polynomiality results

Our goal is to prove the following polynomiality result for the Lascoux coefficients.

Theorem 5.4.1. *Let $I = \{i_1, \ldots, i_r\}$ be a set of strictly increasing nonnegative integers. For $n \geq 0$ the function:*

$$LP_I(n) := \begin{cases} \psi_{[n]\setminus I} & \text{if } I \subseteq [n], \\ 0 & \text{otherwise.} \end{cases}$$

is a polynomial.

As corollaries, we obtain polynomiality of δ and ϕ:

Corollary 5.4.2. *For any fixed $m, s > 0$, the function $\delta(m, n, n-s)$ is a polynomial in n. Moreover this polynomial vanishes at $n = 0$.*

Proof. By Theorem 5.2.1, we have

$$\delta(m, n, n-s) = \sum_{\substack{I \subset [n] \\ \#I = s \\ \sum I = m-s}} \psi_I \psi_{[n]\setminus I} = \sum_{\substack{\#I = s \\ \sum I = m-s}} \psi_I LP_I(n)$$

By Theorem 5.4.1, each of the summands is a polynomial in n that vanishes for $n = 0$. Thus $\delta(m, n, n-s)$ is also a polynomial in n, which proves Theorem 5.4.2, and hence Theorem 5.4.3. □

Corollary 5.4.3. *For any fixed $d > 0$, the function $\phi(n, d)$ is a polynomial for $n > 0$.*

Proof. For all $n, d > 0$, we have:

$$\phi(n, d) = \frac{1}{n} \sum_{1 \leq \binom{s+1}{2} \leq d} s\delta(d, n, n-s). \tag{14}$$

By Corollary 5.4.2 every term in the right hand side of (14) is a polynomial divisible by n, hence the theorem follows. □

Before we prove Theorem 5.4.1, let us first give a few examples.

Example 5.4.4. By induction, one can check the following formulas for LP_I, when I has cardinality one or two:

$$LP_{(i)}(n) = \binom{n}{j+1}, \qquad LP_{(0,j)}(n) = j\binom{n+1}{j+2},$$

and more generally, for $i < j$,

$$LP_{(i,j)}(n) = \frac{(j-i)[n+1]_{j+2}}{(i+1)!(j+1)!(i+j+2)!} \sum_{d=0}^{i} (-1)^d a_{i,d} (i+j+1-d)![n]_{i-d},$$

where $a_{i,d} = \prod_{k=0}^{d-1}(i-k)(i-k+1)$ and $[n]_d = n(n-1)\cdots(n-d+1)$.

Proof of Theorem 5.4.1. We proceed by induction first on $\#I$, then on $\sum I := \sum_{i_j \in I} i_j$. The base case is $I = \emptyset$, when $\psi_{\{0,\dots,n-1\}} = 1$.

For the induction step, fix I, and assume the theorem has been proven for all I' with $\#I' < \#I$, and for all I' with $\#I' = \#I$ and $\sum I' < \sum I$. We consider two cases:

Case 1. $i_1 = 0$. We claim that for every $n \geq 0$,

$$LP_I(n) = (n - r + 1)LP_{I \setminus \{0\}}(n) - 2 \sum_{\ell : i_{\ell+1} > i_\ell + 1} LP_{I \setminus \{0, i_\ell\} \sqcup \{i_\ell + 1\}}(n), \tag{15}$$

where for summation we formally assume $i_{r+1} = +\infty$. Indeed: if $n \leq i_r$ then both sides are 0, and if $n > i_r$ then the equation is precisely (12).

Case 2. $i_1 > 0$. We claim that for every $n \geq 0$,

$$LP_I(n) - LP_I(n-1) = \sum_J LP_J(n-1), \tag{16}$$

where the sum is over all $J \neq I$ of the form $\{i_1 - \epsilon_1, \dots, i_r - \epsilon_r\}$ with $\epsilon_\ell \in \{0, 1\}$. Again, if $n \leq i_r$ then both sides are 0, and if $n > i_r$ then the equation is precisely (13).

In both cases, it follows that LP_I is a polynomial. □

Let's try to determine the degree of this polynomial LP_I. From the recursions in the proof above, we get a candidate: $\sum I + \#I$. More precisely: if we can show that in the right hand sides of (15) and (16) there is no cancellation in the leading coefficients, then it follows inductively that $\deg(LP_I) = \sum I + \#I$. One way of showing there is no cancellation is by finding and proving a formula the leading coefficient and seeing that it is never 0. Such a formula was actually obtained in [MMM$^+$20, Theorem 4.2], but the proof is quite lengthy technical and uses a different description of the Lascoux coefficients. We believe it can be done more directly, leading to our first open problem:

Problem 5.4.5. *Use the recursive relations (15) and (16) to prove a formula for the leading coefficient of LP_I, and deduce that this leading coefficient is never 0.*

6 Further directions

6.1 Skew-symmetric and general matrices

We now know that for a generic d-dimensional linear space $\mathcal{L}$ of symmetric $n \times n$ matrices, the degree $\phi(n, d)$ of $\mathcal{L}^{-1}$ is a polynomial in d. We can ask the same question also for general square matrices, or for skew-symmetric matrices.

Definition 6.1.1. Let M_n be the space of $n \times n$-matrices over $\mathbb{C}$, and $A_{2n} \subset M_{2n}$ be the space of skew-symmetric matrices.

- The number $\phi_A(n, d)$ is the degree of the variety $\mathcal{L}^{-1}$, where $\mathcal{L} \subseteq \mathbb{P}(M_n)$ is a general linear subspace of dimension $d - 1$.

- The number $\phi_D(n,d)$ is the degree of the variety $\mathcal{L}^{-1}$, where $\mathcal{L} \subseteq \mathbb{P}(A_{2n})$ is a general linear subspace of dimension $d-1$.

Remark 6.1.2. A skew-symmetric matrix of odd size is always singular, which is why we only consider the even case.

Computing these numbers can again be reduced to an intersection problem on a suitable space. One way of defining this suitable space is by replacing in Definition 4.2.1 symmetric matrices by general matrices or by skew-symmetric matrices.

Definition 6.1.3. Let V and W be two vector spaces of equal dimension n. The space $\mathbb{P}(V^* \otimes W)$ represents linear maps from V to W; the open subset of rank n linear maps is denoted by $\mathbb{P}(V^* \otimes W)^\circ$. Then the space of complete collineations $\mathcal{CC}(V,W)$ is defined as the closure of the image of the map

$$\phi : \mathbb{P}(V^*\otimes W)^\circ \to \mathbb{P}(V^*\otimes W)\times\mathbb{P}\left(\bigwedge^2 V^* \otimes \bigwedge^2 W\right)\times\ldots\times\mathbb{P}\left(\bigwedge^{n-1} V^* \otimes \bigwedge^{n-1} W\right),$$

given by

$$[A] \mapsto ([A], [\wedge^2 A], \ldots, [\wedge^{n-1} A]).$$

As before, in coordinates this map sends a matrix to its minors of various sizes.

Definition 6.1.4. Let V be a $2n$-dimensional vector space. The space of complete skew-forms $\mathcal{CS}(V)$ is defined as the closure of $\phi(\mathbb{P}(\bigwedge^2(V))^\circ)$, where

$$\phi : \mathbb{P}\left(\bigwedge^2 V\right)^\circ \to \mathbb{P}\left(\bigwedge^2 V\right) \times \mathbb{P}\left(\bigwedge^4 V\right) \times \ldots \times \mathbb{P}\left(\bigwedge^{2n-2} V\right),$$

given by

$$[A] \mapsto ([A], [\wedge^2 A], \ldots, [\wedge^{n-1} A]).$$

We note that here $\wedge^i A$ is viewed as an element of $\bigwedge^{2i} V$, see also [Ber97, Section 3]. In coordinates, the map $\bigwedge^2 V \to \bigwedge^{2i} V$ sends the entries of a skew-symmetric matrix to the Pfaffians of its principal $2i \times 2i$ submatrices.

Similar to before, one can express $\phi_A(n,d)$ (resp. $\phi_D(n,d)$) in terms of the so-called *type A* (resp. *type D*) *Lascoux coefficients*, which allows us to show that $\phi_A(n,d)$ (resp. $\phi_D(n,d)$) is a polynomial in n. See [MMM+20, Section 6 and 7] for details.

Problem 6.1.5. *The type A Lascoux polynomials $LP^A_{I,J}(n)$ (where $I, J \subset \mathbb{N}$ are subsets of the same size) were introduced in [MMM+20, Theorem 6.11]. From the recursive relations they satisfy, it follows that $\deg(LP^A_{I,J}(n)) \leq \#I + \sum I + \sum J$. Show that this is an equality, for instance by finding and proving a formula for the leading coefficient. Same question for the type D Lascoux quasipolynomials $LP^D_I(n)$, defined in [MMM+20, Theorem 7.10].*

6.2 Projective duality

The algebraic degree of semidefinite programming $\delta(m,n,r)$ has a natural geometric definition:

Definition 6.2.1. For $0 < m < \binom{n+1}{2}$ and $0 < r < n$, let $\mathcal{L} \subset S^2\mathbb{C}^n$ be a general linear space of symmetric matrices, of (affine) dimension $m+1$, and let $SD_m^{r,n} \subset \mathbb{P}(\mathcal{L})$ denote the projectivization of the cone of matrices of rank at most r in $\mathcal{L}$. Then $\delta(m,n,r)$ is the degree of the projective dual $(SD_m^{r,n})^*$ of $SD_m^{r,n}$ if this dual is a hypersurface, and zero otherwise.

The equivalence of the above definition with our Definition 4.4.4 is proven in [GvBR09, Proposition 4.1] (see also [MMM$^+$20, Proposition 3.5]).

The above definition makes sense also for general (not necessarily square) matrices, and for skew-symmetric matrices. For general square matrices and skew-symmetric matrices of even size these were studied in [MMM$^+$20, Sections 6 and 7]. In particular, we have analogues of Definition 4.4.4 and 5.2.1.

Problem 6.2.2. *For general (non-square) matrices, interpret the dual degree δ as an intersection problem on the space of complete collineations. Introduce Lascoux coefficients ψ in this case, write δ in terms of ψ, and prove polynomiality of ψ and δ. Same question for skew-symmetric matrices of odd size.*

In Definition 6.2.1, it is known when $(SD_m^{r,n})^*$ is a hypersurface: namely precisely when *Pataki's inequalities* hold:

$$\binom{n-r+1}{2} \leq m \leq \binom{n+1}{2} - \binom{r+1}{2}. \tag{17}$$

If $m < \binom{n-r+1}{2}$ then $SD_m^{r,n} = \emptyset$, but for $m > \binom{n+1}{2} - \binom{r+1}{2}$ our dual variety $(SD_m^{r,n})^*$ is nonempty and of codimension greater than one.

Problem 6.2.3. *Study the dimension and degree of $(SD_m^{r,n})^*$ for $m > \binom{n+1}{2} - \binom{r+1}{2}$. See [MMM$^+$20, Section 8.A] for some precise conjectures.*

References

[Ber97] Aaron Bertram. An application of a log version of the kodaira vanishing theorem to embedded projective varieties. *arXiv preprint alg-geom/9707001*, 1997.

[DCGMP88] C. De Concini, M. Goresky, R. MacPherson, and C. Procesi. On the geometry of quadrics and their degenerations. *Comment. Math. Helv.*, 63(3):337–413, 1988.

[EH16] David Eisenbud and Joe Harris. *3264 and all that—a second course in algebraic geometry.* Cambridge University Press, Cambridge, 2016.

[FH91] William Fulton and Joe Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991. A first course, Readings in Mathematics.

[Ful98] William Fulton. *Intersection theory*, volume 2 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, second edition, 1998.

[GvBR09] Hans-Christian Graf von Bothmer and Kristian Ranestad. A general formula for the algebraic degree in semidefinite programming. *Bull. Lond. Math. Soc.*, 41(2):193–197, 2009.

[KWJ20] Kathlen Kohn, Rosa Winter, and Yuhan Jiang. Linear spaces of symmetric matrices with non-maximal maximum likelihood degree. *arXiv preprint arXiv:2012.00145*, 2020.

[LLT89] D. Laksov, A. Lascoux, and A. Thorup. On Giambelli's theorem on complete correlations. *Acta Math.*, 162(3-4):143–199, 1989.

[LTT88] Dũng Lê Tráng and Bernard Teissier. Limites d'espaces tangents en géométrie analytique. *Commentarii Mathematici Helvetici*, 63(1):540–578, 1988.

[Mas20] Alex Massarenti. On the birational geometry of spaces of complete forms I: collineations and quadrics. *Proc. Lond. Math. Soc. (3)*, 121(6):1579–1618, 2020.

[MMM$^+$20] Laurent Manivel, Mateusz Michałek, Leonid Monin, Tim Seynnaeve, and Martin Vodička. Complete quadrics: Schubert calculus for gaussian models and semidefinite programming. *arXiv preprint arXiv:2011.08791*, 2020.

[NRS10] Jiawang Nie, Kristian Ranestad, and Bernd Sturmfels. The algebraic degree of semidefinite programming. *Math. Program.*, 122(2, Ser. A):379–405, 2010.

[Pra96] Piotr Pragacz. Symmetric polynomials and divided differences in formulas of intersection theory. In *Parameter spaces (Warsaw, 1994)*, volume 36 of *Banach Center Publ.*, pages 125–177. Polish Acad. Sci. Inst. Math., Warsaw, 1996.

[Sch94] H. Schubert. Allgemeine Anzahlfunctionen für Kegelschnitte, Flächen und Räume zweiten Grades in n Dimensionen. *Math. Ann.*, 45(2):153–206, 1894.

[Sey] Tim Seynnaeve. Complete quadrics: Schubert calculus for gaussian models and semidefinite programming. `https://mathrepo.mis.mpg.de/MLdegreeCompleteQuadrics/index.html`.

[Str86] Elisabetta Strickland. Schubert-type cells for complete quadrics. *Adv. in Math.*, 62(3):238–248, 1986.

[SU10] Bernd Sturmfels and Caroline Uhler. Multivariate gaussians, semidefinite matrix completion, and convex algebraic geometry. *Annals of the Institute of Statistical Mathematics*, 62(4):603–638, 2010.

[Tei73] Bernard Teissier. Cycles évanescents, sections planes et conditions de whitney. *Astérisque*, 7(8):285–362, 1973.

[Tot14] Burt Totaro. Chow groups, Chow cohomology, and linear varieties. *Forum Math. Sigma*, 2:Paper No. e17, 25, 2014.