PENCILS OF QUADRICS: OLD AND NEW

CLAUDIA FEVOLA - YELENA MANDELSHTAM - BERND STURMFELS

Two-dimensional linear spaces of symmetric matrices are classified by Segre symbols. After reviewing known facts from linear algebra and projective geometry, we address new questions motivated by algebraic statistics and optimization. We compute the reciprocal curve and the maximum likelihood degrees, and we study strata of pencils in the Grassmannian.

1. Introduction

A pencil of quadrics is a two-dimensional linear subspace $\mathcal{L}$ in the space $\mathbb{S}^n$ of (real or complex) symmetric $n \times n$-matrices. It is a point in the Grassmannian $\mathrm{Gr}(2, \mathbb{S}^n)$, and it specifies a line $\mathbb{P}\mathcal{L}$ in the projective space $\mathbb{P}(\mathbb{S}^n) \simeq \mathbb{P}^{\binom{n+1}{2}-1}$. The group $\mathrm{GL}(n)$ acts on $\mathbb{S}^n$ by congruence and this induces an action on $\mathrm{Gr}(2, \mathbb{S}^n)$. We consider two pencils to be *isomorphic* if they are in the same $\mathrm{GL}(n)$-orbit.

Let $\mathcal{L}$ be a pencil with basis $\{A, B\}$. Its determinant is the binary form

$$\det(\mathcal{L}) \;=\; \det(\lambda A + \mu B).$$

This is well-defined up to the action of $\mathrm{GL}(2)$ by changing basis in $\mathcal{L}$. The zeros of $\det(\mathcal{L})$ form a multiset of size n in the line $\mathbb{P}^1$, well-defined up to isomorphisms of $\mathbb{P}^1$. We exclude pencils $\mathcal{L}$ that are *singular*, meaning that $\det(\mathcal{L}) = 0$.

The singular pencils form a subvariety $\mathrm{Gr}(2, \mathbb{S}^n)^{\mathrm{sing}}$ in the Grassmannian. We are interested in a natural stratification of the open set of all regular pencils:

$$\mathrm{Gr}(2, \mathbb{S}^n)^{\mathrm{reg}} \;=\; \mathrm{Gr}(2, \mathbb{S}^n) \setminus \mathrm{Gr}(2, \mathbb{S}^n)^{\mathrm{sing}}.$$

Each stratum is indexed by a *Segre symbol* σ. This is a multiset of partitions whose numbers of parts add up to n. One exception: the singleton $[(1, 1, \ldots, 1)]$ is not a Segre symbol. The number $S(n)$ of Segre symbols was already of interest to Arthur Cayley in 1855. In [2, p. 316], he derived the generating function

$$\sum_{n=1}^{\infty} S(n) x^n \;=\; \prod_{k \geq 1} \frac{1}{(1 - x^k)^{P(k)}} - \frac{1}{1-x} \;=\; 2x^2 + 5x^3 + 13x^4 + 26x^5 + 57x^6 + 110x^7 + \cdots,$$

where $P(k)$ is the number of partitions of the integer k. The two Segre symbols for $n = 2$ are $[(1), (1)]$ and $[(2)]$. For $n = 3$ and $n = 4$ they are shown in Figure 1.

The Segre symbol $\sigma = \sigma(\mathcal{L})$ of a given pencil $\mathcal{L}$ can be computed as follows. Pick a basis $\{A, B\}$ of $\mathcal{L}$, where B is invertible, and find the Jordan canonical form of AB^{-1}. Each eigenvalue of AB^{-1} determines a partition, according to the pattern of ones above the diagonal. Then σ is the associated multiset of partitions. It turns out that σ does not depend on the choice of basis $\{A,B\}$. For the relevant background in linear algebra see [9, 10] and Section 2 below.

The role of Segre symbols in projective geometry can be stated as follows.

Theorem 1.1 (Weierstrass-Segre)**.** Two pencils of quadrics in $\mathbb{S}^n$ are in the same $\mathrm{GL}(n)$-orbit if and only if their Segre symbols agree and their determinants define the same multiset of n points on the projective line $\mathbb{P}^1$, up to isomorphism.

Example 1.2 ($n = 2$)**.** All pencils $\mathcal{L}$ are regular, and there are two $\mathrm{GL}(2)$-orbits. These two orbits are distinguished by the rank of a matrix X that spans the one-dimensional space $\mathcal{L}^\perp = \{X \in \mathbb{S}^2 : \mathrm{trace}(AX) = \mathrm{trace}(BX) = 0\}$. If X has rank 2 then $\det(\mathcal{L})$ has two distinct roots in $\mathbb{P}^1$ and the Segre symbol is $\sigma(\mathcal{L}) = [(1),(1)]$. If X has rank 1 then $\det(\mathcal{L})$ has a double root in $\mathbb{P}^1$ and $\sigma(\mathcal{L}) = [(2)]$.

We learned about Theorem 1.1 from an unpublished note by Pieter Belmans, titled *Segre symbols*, which credits the 1883 PhD thesis of Corrado Segre. It appears in the textbooks on algebraic geometry by Dolgachev [4, §8.6.1] and Hodge-Pedoe [5, §XIII.10]. The idea goes back to at least the 1850s, in works of Cayley [2] and Sylvester [8]. The aim of this article is to revisit this topic.

We begin in Section 2 with a linear algebra perspective on Theorem 1.1. In Section 3 we study the reciprocal curve $\mathbb{P}\mathcal{L}^{-1}$ of a pencil $\mathcal{L} \in \mathrm{Gr}(2,\mathbb{S}^n)^{\mathrm{reg}}$. This curve lives in $\mathbb{P}(\mathbb{S}^n)$, and it is parametrized by the inverses of all invertible matrices in $\mathcal{L}$. We prove that $\mathbb{P}\mathcal{L}^{-1}$ is a rational normal curve. We express its degree in terms of the Segre symbol $\sigma(\mathcal{L})$, and we determine its prime ideal.

In Section 4 we turn to maximum likelihood estimation for Gaussians. A pencil $\mathcal{L}$ plays two different roles in statistics, depending on whether it lives in the space of concentration matrices (as in [7]) or in the space of covariance matrices (as in [3]). This yields two numerical invariants, the ML degree $\mathrm{mld}(\mathcal{L})$ and the reciprocal ML degree $\mathrm{rmld}(\mathcal{L})$. We compute these in Theorem 3.2.

In Section 5 we study the constructible set defined by a fixed Segre symbol:

$$\mathrm{Gr}_\sigma = \left\{ \mathcal{L} \in \mathrm{Gr}(2,\mathbb{S}^n)^{\mathrm{reg}} \; : \; \sigma(\mathcal{L}) = \sigma \right\}. \tag{1}$$

Its closure $\overline{\mathrm{Gr}_\sigma}$ is a variety. We study these varieties and their poset of inclusions, seen in Figure 1. This extends the stratification of $\mathrm{Gr}(2,\mathbb{R}^n)$ by matroids.

Example 1.3 ($n = 3$)**.** There are five strata Gr_σ in the Grassmannian $\mathrm{Gr}(2,\mathbb{S}^3)$:

symbol	codim	degrees	P	Q	variety in $\mathbb{P}^2$
$[1,1,1]$	0	$(2,2,3)$	$ax^2+by^2+cz^2$	$x^2+y^2+z^2$	*four reduced points*
$[2,1]$	1	$(2,1,2)$	$2axy+y^2+bz^2$	$2xy+z^2$	*one double point, two others*
$[3]$	2	$(2,0,1)$	$2axz+ay^2+2yz$	$2xz+y^2$	*one triple point, one other*
$[(1,1),1]$	2	$(1,1,1)$	$ax^2+ay^2+bz^2$	$x^2+y^2+z^2$	*two double points*
$[(2,1)]$	3	$(1,0,0)$	$2axy+y^2+az^2$	$2xy+z^2$	*quadruple point*

For each Segre symbol σ, we display $\mathrm{codim}(\mathrm{Gr}_\sigma)$, the triple of degrees $\big(\deg(\mathcal{L}^{-1}), \mathrm{mld}(\mathcal{L}), \mathrm{rmld}(\mathcal{L})\big)$, the basis $\{P,Q\}$ from Section 2, and its variety in $\mathbb{P}^2$. Here, x,y,z are coordinates on $\mathbb{P}^2$, and a,b,c are distinct nonzero reals. This accounts for all regular pencils. A pencil is singular if P and Q share a linear factor. One such $\mathcal{L}$ is spanned by xy and xz. This defines a line and a point in $\mathbb{P}^2$. This shows that $\mathrm{Gr}(2,\mathbb{S}^3)^{\mathrm{sing}}$ is an irreducible variety of dimension 4.

2. Canonical Representatives

We identify symmetric $n\times n$ matrices A with quadratic forms $\mathbf{x}A\mathbf{x}^T$ in unknowns $\mathbf{x} = (x_1,\dots,x_n)$. We fix the field to be $\mathbb{C}$. The $\binom{n+1}{2}$-dimensional vector space $\mathbb{S}^n$ is equipped with the trace inner product $(A,B) \mapsto \mathrm{trace}(AB)$. The group $\mathrm{GL}(n)$ acts on quadratic forms by linear changes of coordinates, via $\mathbf{x} \mapsto \mathbf{x}g$. This corresponds to the action of $\mathrm{GL}(n)$ on symmetric matrices by congruence:

$$\mathrm{GL}(n)\times\mathbb{S}^n \to \mathbb{S}^n,\ (g,A) \mapsto gAg^T.$$

Let $\mathcal{L} = \mathbb{C}\{A,B\}$ be a regular pencil in $\mathrm{Gr}(2,\mathbb{S}^n)$, with $\det(B) \neq 0$. The polynomial ring $\mathbb{C}[\lambda]$ in one variable λ is a principal ideal domain. The cokernel of the matrix $A-\lambda B$ is a module over this PID. Consider its *elementary divisors*

$$(\lambda-\alpha_1)^{e_1},\,(\lambda-\alpha_2)^{e_2},\,\dots,\,(\lambda-\alpha_s)^{e_s}. \tag{2}$$

Here $e_1,\dots,e_s$ are positive integers whose sum equals n. The list (2) is unordered and its product equals $\det(\mathcal{L}) = \pm\det(A-\lambda B)$. The complex numbers α_i are the *eigenvalues* of (A,B). They form multiset of cardinality n in $\mathbb{P}^1$.

Suppose there are r distinct eigenvalues α_i. We have $r\le s\le n$. The exponents e_i corresponding to one fixed eigenvalue form a partition. This gives a multiset of r partitions, with s parts in total, where the sum of all parts is n. This multiset of partitions is the Segre symbol $\sigma = \sigma(\mathcal{L})$. It is thus visible in (2). We now paraphrase Theorem 1.1 using the elementary divisors of the matrix $A-\lambda B$.

Corollary 2.1. Consider two quadrics $\mathbf{x}A\mathbf{x}^T$ and $\mathbf{x}B\mathbf{x}^T$ with $\det(B)\neq 0$. There exists a change of coordinates $\mathbf{x}\mapsto\mathbf{x}g$ which transforms them to $\mathbf{x}C\mathbf{x}^T$ and $\mathbf{x}D\mathbf{x}^T$ if and only if the matrices $A-\lambda B$ and $C-\lambda D$ have the same elementary divisors.

Proof. For a textbook proof of this classical fact see [5, Theorem 1, p. 278]. □

Corollary 2.1 is used to construct a canonical form for pencils. For $e\in\mathbb{N}$ and $\alpha\in\mathbb{C}$, we define a pair of symmetric $e\times e$ matrices by filling their antidiagonals:

$$P_e(\alpha) = \begin{pmatrix} 0 & 0 & \cdots & 0 & \alpha \\ 0 & 0 & \cdots & \alpha & 1 \\ \vdots & \vdots & ⋰ & ⋰ & \vdots \\ 0 & \alpha & 1 & \vdots & 0 \\ \alpha & 1 & \cdots & 0 & 0 \end{pmatrix} \quad\text{and}\quad Q_e = \begin{pmatrix} 0 & \cdots & 0 & 0 & 1 \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 1 & 0 & 0 \\ \vdots & ⋰ & \vdots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 & 0 \end{pmatrix}. \tag{3}$$

The $e\times e$ matrix $P_e(\alpha)-\lambda Q_e$ has only one elementary divisor, namely $(\lambda-\alpha)^e$.

Let us now start with the list in (2). For each elementary divisor $(\lambda-\alpha_i)^{e_i}$ we form the $e_i\times e_i$ matrices in (3), and we aggregate these blocks as follows:

$$P=\begin{pmatrix} P_{e_1}(\alpha_1) & 0 & \cdots & 0 \\ 0 & P_{e_2}(\alpha_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_{e_s}(\alpha_s)\end{pmatrix} \quad\text{and}\quad Q=\begin{pmatrix} Q_{e_1} & 0 & \cdots & 0 \\ 0 & Q_{e_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_{e_s}\end{pmatrix}. \quad (4)$$

The matrices $A-\lambda B$ and $P-\lambda Q$ have the same elementary divisors. Hence, by Corollary 2.1, the pair $(\mathbf{x}A\mathbf{x}^T,\mathbf{x}B\mathbf{x}^T)$ is equivalent to $(\mathbf{x}P\mathbf{x}^T,\mathbf{x}Q\mathbf{x}^T)$ under the action by $\mathrm{GL}(n)$. As in Example 1.3, every regular pencil $\mathcal{L}\in\mathrm{Gr}(2,\mathbb{S}^n)$ has a normal form $\mathbb{C}\{P,Q\}$, where the matrices P and Q are uniquely defined by the unordered list (2). Given any Segre symbol σ, its canonical representative is $\mathcal{L}=\mathbb{C}\{P,Q\}$ where $\alpha_1,\ldots,\alpha_r$ are parameters. Here, and in what follows, we use index-free notation for unknowns, like $\mathbf{x}=(x,y,z)$ and $(\alpha_1,\alpha_2,\alpha_3)=(a,b,c)$.

Example 2.2. Let $n=5$ and $\sigma=[(2,1),2]$. The list of elementary divisors equals

$$(\lambda-a)^2,\ (\lambda-a),\ (\lambda-b)^2.$$

Our canonical representatives (4) for this class of pencils $\mathcal{L}$ is the matrix pair

$$P=\begin{pmatrix} 0 & a & 0 & 0 & 0 \\ a & 1 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 \\ 0 & 0 & 0 & 0 & b \\ 0 & 0 & 0 & b & 1 \end{pmatrix} \quad\text{and}\quad Q=\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

The quadrics $P=2axy+y^2+az^2+2buv+v^2$ and $Q=2xy+z^2+2uv$ define a degenerate del Pezzo surface of degree four in $\mathbb{P}^4$. This surface has two singular points, $(0:0:0:1:0)$ and $(1:0:0:0:0)$; their multiplicities are one and three.

Remark 2.3. To appreciate Theorem 1.1 and Corollary 2.1, it helps to distinguish the two geometric figures associated with a pencils of quadrics, and how the groups $\mathrm{GL}(2)$ and $\mathrm{GL}(n)$ act on these. First, there is the configuration of n points in $\mathbb{P}^1$ defined by $\det(\mathcal{L})$. This configuration undergoes projective transformations via $\mathrm{GL}(2)$ but it is left invariant by $\mathrm{GL}(n)$. Second, there is the codimension 2 variety in $\mathbb{P}^{n-1}$ defined by the quadrics in $\mathcal{L}$. This variety undergoes projective transformations via $\mathrm{GL}(n)$ but it is left invariant by $\mathrm{GL}(2)$. We here derived a normal form with respect to both group actions simultaneously.

In this section, linear algebra over a PID is used to study pencils $\mathcal{L}=\mathbb{C}\{A,B\}$. To understand this better, recall the relationship between *elementary divisors* and *invariant factors*, seen in a first course in abstract algebra. A practical tool for computing these is the *Smith normal form* algorithm over $\mathbb{C}[\lambda]$. We here apply this to a specific torsion module, namely the cokernel of our matrix $A-\lambda B$.

Fix n and a Segre symbol $\sigma=[\sigma_1,\dots,\sigma_r]$, where each entry is now a weakly decreasing vector $\sigma_i=(\sigma_{i1},\sigma_{i2},\dots,\sigma_{in})$ of nonnegative integers. With this convention, the Segre symbol $\sigma=[\sigma_1,\sigma_2]$ in Example 2.2, with $n=5,s=3,r=2$, has $\sigma_1=(2,1,0)$ and $\sigma_2=(2,0,0)$. Write $\alpha_1,\dots,\alpha_r\in\mathbb{C}$ for the distinct roots of $\det(A-\lambda B)$. Then the elementary divisors are $(\lambda-\alpha_i)^{\sigma_{ij}}$ for $i=1,\dots,r$ and $j=1,\dots,n$. Only s of these are different from 1. The invariant factors are

$$d_j \;:=\; \prod_{i=1}^{r}(\lambda-\alpha_i)^{\sigma_{ij}} \quad \text{for } j=1,\dots,n.$$

Note that $d_n|d_{n-1}|\cdots|d_2|d_1$. The number of non-trivial invariant factors is the maximum number of parts among the r partitions σ_i. For instance, in Example 2.2, the invariant factors are $d_1=(\lambda-a)^2(\lambda-b)^2$, $d_2=\lambda-a$, $d_3=d_4=d_5=1$.

The ideal of $k\times k$-minors of $A-\lambda B$ is generated by the greatest common divisor D_k of these minors. The theory of modules over a PID tells us that

$$D_k \;:=\; \prod_{j=1}^{k} d_{n+1-j} \;=\; \prod_{i=1}^{r}(\lambda-\alpha_i)^{\sigma_{i,n-k+1}+\cdots+\sigma_{i,n-1}+\sigma_{in}}. \tag{5}$$

We can thus find the Segre symbol of a given pencil $\mathcal{L}=\mathbb{C}\{A,B\}$ by inspecting the ideal of $k\times k$-minors of $A-\lambda B$ for $k=1,\dots,n$. This computation can be done by transforming $A-\lambda B$ to its Smith normal form. In the Introduction we proposed a different method, namely computing the *Jordan canonical form* of AB^{-1}. This operation uses only linear algebra over $\mathbb{C}$ and it outperforms the Smith normal form for larger n. To see that the Jordan canonical form of AB^{-1} reveals the Segre symbol, consider the transformation from (A,B) to (P,Q) in Corollary 2.1. This preserves the conjugacy class of AB^{-1}. Therefore, AB^{-1} and PQ^{-1} have the same Jordan canonical form. We now see in (4) that Q is a permutation matrix, and hence so is Q^{-1}. Furthermore, P is already in Jordan canonical form, after permuting rows and columns, and σ is clearly visible in P.

3. The Reciprocal Curve

For any nonsingular pencil $\mathcal{L}$, we are interested in the reciprocal curve $\mathbb{P}\mathcal{L}^{-1}$. We write $\deg(\mathcal{L}^{-1})$ for the degree of this curve in $\mathbb{P}(\mathbb{S}^n)$. In Example 1.3, $\deg(\mathcal{L}^{-1})=2$ in three cases, so $\mathbb{P}\mathcal{L}^{-1}$ is a plane conic. In the other two cases,

$\mathbb{P}\mathcal{L}^{-1}$ is a line in $\mathbb{P}^5$. Here are the homogeneous prime ideals of these curves:

Segre symbol	Ideal of the reciprocal curve $\mathbb{P}\mathcal{L}^{-1}$	mingens
$[1,1,1]$	$\langle x_{12},x_{13},x_{23},(c-b)x_{11}x_{22}+(a-c)x_{11}x_{33}+(b-a)x_{22}x_{33}\rangle$	$(3,1)$
$[2,1]$	$\langle x_{13}, x_{22}, x_{23}, x_{12}^2+(c-a)x_{11}x_{33}-2x_{12}x_{33}\rangle$	$(3,1)$
$[3]$	$\langle x_{23}, x_{33}, x_{13}-2x_{22}, x_{12}^2-x_{11}x_{22}\rangle$	$(3,1)$
$[(1,1),1]$	$\langle x_{12}, x_{13}, x_{23}, x_{11}-x_{22}\rangle$	$(4,0)$
$[(2,1)]$	$\langle x_{13}, x_{22}, x_{23}, x_{12}-2x_{33}\rangle$	$(4,0)$

The column "mingens" gives the numbers of linear and quadratic generators.

Example 3.1 ($n=4$)**.** Two quadrics P and Q in $\mathbb{P}^3$ meet in a quartic curve. There are 13 cases, one for each Segre symbol. Here, x,y,z,u are coordinates on $\mathbb{P}^3$.

symbol	codims	degrees	mingens	quadrics P,Q	variety in $\mathbb{P}^3$
$[1,1,1,1]$	$0,0,0$	$(\mathbf{3},3,5)$	$(6,3)$	$ax^2+by^2+cz^2+du^2$ $x^2+y^2+z^2+u^2$	*elliptic curve*
$[2,1,1]$	$1,1,1$	$(\mathbf{3},2,4)$	$(6,3)$	$2axy+y^2+cz^2+du^2$ $2xy+z^2+u^2$	*nodal curve*
$[(1,1),1,1]$	$3,2,2$	$(\mathbf{2},2,3)$	$(7,1)$	$a(x^2+y^2)+cz^2+du^2$ $x^2+y^2+z^2+u^2$	*two conics meet twice*
$[3,1]$	$2,2,2$	$(\mathbf{3},1,3)$	$(6,3)$	$2axz+ay^2+2yz+du^2$ $2xz+y^2+u^2$	*cuspidal curve*
$[2,2]$	$2,2,2$	$(\mathbf{3},1,3)$	$(6,3)$	$2axy+y^2+2bzu+u^2$ $2xy+2zu$	*twisted cubic with secant*
$[(2,1),1]$	$4,3,3$	$(\mathbf{2},1,2)$	$(7,1)$	$2axy+y^2+az^2+du^2$ $2xy+z^2+u^2$	*two tangent conics*
$[4]$	$3,3,3$	$(\mathbf{3},0,2)$	$(6,3)$	$2axu+2ayz+2yu+z^2$ $2xu+2yz$	*twisted cubic with tangent*
$[2,(1,1)]$	$4,3,3$	$(\mathbf{2},1,2)$	$(7,1)$	$2axy+y^2+c(z^2+u^2)$ $2xy+z^2+u^2$	*conic meets two lines*
$[(3,1)]$	$5,4,4$	$(\mathbf{2},0,1)$	$(7,1)$	$2axz+ay^2+2yz+au^2$ $2xz+y^2+u^2$	*conic and two lines concur*
$[(1,1),(1,1)]$	$6,4,4$	$(\mathbf{1},1,1)$	$(8,0)$	$a(x^2+y^2)+c(z^2+u^2)$ $x^2+y^2+z^2+u^2$	*quadrangle of lines*
$[(1,1,1),1]$	$8,5,5$	$(\mathbf{1},1,1)$	$(8,0)$	$a(x^2+y^2+z^2)+du^2$ $x^2+y^2+z^2+u^2$	*double conic*
$[(2,2)]$	$7,5,5$	$(\mathbf{1},0,0)$	$(8,0)$	$2axy+y^2+2azu+u^2$ $2xy+2zu$	*double line and two lines*
$[(2,1,1)]$	$9,6,6$	$(\mathbf{1},0,0)$	$(8,0)$	$2axy+y^2+a(z^2+u^2)$ $2xy+z^2+u^2$	*two double lines*

We see that $\mathbb{P}\mathcal{L}^{-1}\subset\mathbb{P}^9$ is either a line, or a plane conic, or a twisted cubic curve. This is explained by the next theorem, which is our main result in Section 3.

Theorem 3.2. Let $\mathcal{L}$ be a regular pencil in $\mathbb{S}^n$ with Segre symbol $\sigma=[\sigma_1,\ldots,\sigma_r]$. Then $\mathbb{P}\mathcal{L}^{-1}$ is a rational normal curve of degree d in $\mathbb{P}(\mathbb{S}^n)$, where $d=\sum_{i=1}^r\sigma_{i1}-1$ is one less than the sum of the first parts of the partitions in σ. The ideal of $\mathbb{P}\mathcal{L}^{-1}$ is generated by $\binom{n+1}{2}-d-1$ linear forms and $\binom{d}{2}$ quadrics in $\binom{n+1}{2}$ unknowns.

Proof. The curve $\mathbb{P}\mathcal{L}^{-1}$ is parametrized by $\binom{n+1}{2}$ rational functions in one unknown λ, namely the entries in the inverse of matrix $P-\lambda Q$ in Section 2. We scale each entry by $D_n = \det(P-\lambda Q)$ to get a polynomial parametrization by the adjoint of $P-\lambda Q$. This is the $n\times n$-matrix whose entries are the $(n-1)\times(n-1)$-minors of $P-\lambda Q$. These are polynomials of degree $\leq n-1$ in λ, which are divisible by the invariant factor D_{n-1}. Note that D_{n-1} has degree $\sum_{i=1}^r \sum_{j=2}^n \sigma_{ij}$ in λ. Subtracting this from the expected degree $n-1$, we obtain $d = \sum_{i=1}^r \sigma_{i1} - 1$. We remove the factor D_{n-1} from each entry of the adjoint. The resulting matrix $(D_n/D_{n-1})\cdot(P-\lambda Q)^{-1}$ also parametrizes $\mathbb{P}\mathcal{L}^{-1}$. The entries of that matrix are polynomials in λ of degree $\leq d$. As a key step, we will show that these span the $(d+1)$-dimensional space $\mathbb{C}[\lambda]_{\leq d}$ of all polynomials in λ of degree $\leq d$.

The inverse of $P-\lambda Q$ is a block matrix, where the blocks are the inverses of the $e\times e$-matrices $P_e(\alpha)-\lambda Q_e$ in (3), one for each elementary divisor. A computation shows that the entry of $(P_e(\alpha)-\lambda(Q_e))^{-1}$ in row i and column j is

$$-(\lambda-\alpha)^{i+j-e-2} \quad \text{if } i+j\leq e+1 \qquad \text{and} \qquad 0 \quad \text{if } i+j\geq e+2.$$

It follows that the distinct nonzero entries in the $n\times n$-matrix $(P-\lambda Q)^{-1}$ are

$$\pm(\lambda-\alpha_i)^{-k} \quad \text{where } 1\leq k\leq \sigma_{i1} \text{ and } 1\leq i\leq r. \tag{6}$$

The common denominator of these $d+1 = \sum_{i=1}^r \sigma_{i1}$ rational functions in λ is equal to $D_n/D_{n-1} = \prod_{i=1}(\lambda-\alpha_i)^{\sigma_{i1}}$. Multiplying by that common denominator, we obtain $d+1$ polynomials in λ of degree $\leq d$. Lemma 3.3 below tells us that these polynomials are linearly independent. Hence they span $\mathbb{C}[\lambda]_{\leq d}\simeq \mathbb{C}^{d+1}$.

The proof of Theorem 3.2 is now completed as follows. By recording which entries of $(P-\lambda Q)^{-1}$ are zero, and which pairs of entries are equal to each other, we obtain $\binom{n+1}{2}-d-1$ linearly independent linear forms that vanish on $\mathbb{P}\mathcal{L}^{-1}$. We know from the previous paragraph that there exist linear forms u_i in the matrix entries which evaluate to λ^i for $i=0,1,2,\ldots,d-1$. The $\binom{d}{2}$ quadratic forms that vanish on our curve are the 2×2 minors of the $2\times d$-matrix

$$\begin{pmatrix} u_0 & u_1 & u_2 & \cdots & u_{d-1} \\ u_1 & u_2 & u_3 & \cdots & u_d \end{pmatrix}. \tag{7}$$

We have thus constructed an isomorphism between our curve $\mathbb{P}\mathcal{L}^{-1}$ and the rational normal curve $\{(1:\lambda:\cdots:\lambda^{d-1})\}$, whose prime ideal is given by (7). $\square$

Lemma 3.3. A finite set of distinct rational functions $(\lambda-\alpha_j)^{-s_{ij}}$, each a negative power of one of the expressions $\lambda-\alpha_1,\ldots,\lambda-\alpha_r$, is linearly independent.

Proof. We use induction on r, with base case $r=1$. We claim that $(\lambda-\alpha)^{-s_1},\ldots,$ $(\lambda-\alpha)^{-s_n}$ are linearly independent when $0<s_1<\cdots<s_n$. Suppose

$$k_1(\lambda-\alpha)^{-s_1}+\cdots+k_n(\lambda-\alpha)^{-s_n} = 0 \qquad \text{for some } k_1,\ldots,k_n\in\mathbb{C}.$$

Clearing denominators, we obtain $k_1(\lambda-\alpha)^{s_n-s_1}+\cdots+k_n=0$. Setting $\lambda=\alpha$ we find $k_n=0$. Repeating this computation n times, we conclude $k_1=k_2=\cdots=k_n=0$.

For the induction step from $r-1$ to r, we consider distinct negative powers

$$\begin{array}{cccc} (\lambda-\alpha_1)^{-s_{11}}, & (\lambda-\alpha_1)^{-s_{12}}, & \ldots, & (\lambda-\alpha_1)^{-s_{1,n_1}}, \\ \vdots & \vdots & & \vdots \\ (\lambda-\alpha_r)^{-s_{r1}}, & (\lambda-\alpha_r)^{-s_{r2}}, & \ldots, & (\lambda-\alpha_r)^{-s_{r,n_r}}, \end{array} \tag{8}$$

where $0\le s_{ij}<s_{i,j+1}$ for $i=1,\ldots,r$ and $j=1,\ldots,n_i$. Consider a linear combination of (8) with coefficients $k_{11},\ldots,k_{r,n_r}$. Multiplying by $(\lambda-\alpha_r)^{s_{rn}}$ and setting $\lambda=\alpha_r$, we find $k_{rn}=0$. Repeating with $(\lambda-\alpha_r)^{s_{ri}}$ for $i=n_r-1,n_r-2,\ldots,1$, we get $k_{r1}=\cdots=k_{rn}=0$. By the induction hypothesis, the first $r-1$ rows of (8) are linearly independent. This proves that all k_{ij} are zero. Lemma 3.3 follows. □

The last paragraph in the proof of Theorem 3.2 gives an algorithm for computing generators for the ideal of $\mathbb{P}\mathcal{L}^{-1}$. We show this for our running example.

Example 3.4. Let $\sigma=[(2,1),2]$ as in Example 2.2. We have $d=\sigma_{11}+\sigma_{21}-1=2+2-1=3$, so $\mathbb{P}\mathcal{L}^{-1}$ is a twisted cubic curve in $\mathbb{P}^{14}$. The inverse of $P-\lambda Q$ satisfies the $\binom{6}{2}-3-1=11$ linear forms $x_{13},x_{14},x_{15},x_{22},x_{23},x_{24},x_{25},x_{34},x_{35},x_{55}$ and $x_{12}-x_{33}$. The three quadratic ideal generators are relations between $(\lambda-a)^{-1}=-x_{12}$, $(\lambda-a)^{-2}=-x_{11}$, $(\lambda-b)^{-1}=-x_{45}$, and $(\lambda-b)^{-2}=-x_{44}$. Following (7), we write these three quadrics as $u_0u_2-u_1^2$, $u_0u_3-u_1u_2$ and $u_1u_3-u_2^2$, where

$$\begin{array}{rcc} u_0 & = & (a-b)x_{11}-2x_{12}+(a-b)x_{44}+2x_{45}, \\ u_1 & = & (a^2-ab)x_{11}-(a+b)x_{12}+(ab-b^2)x_{44}+(a+b)x_{45}, \\ u_2 & = & (a^3-a^2b)x_{11}-2abx_{12}+(ab^2-b^3)x_{44}+2abx_{45}, \\ u_3 & = & (a^4-a^3b)x_{11}+(a^3-3a^2b)x_{12}+(ab^3-b^4)x_{44}+(3ab^2-b^3)x_{45}. \end{array}$$

4. Maximum Likelihood Degrees

Let $\mathbb{S}^n_{\succ 0}$ denote the open convex cone of positive definite real symmetric $n\times n$ matrices. For any fixed $S\in\mathbb{S}^n$, we consider the following *log-likelihood function*:

$$\ell_S:\mathbb{S}^n_{\succ 0}\to\mathbb{R},\ M\mapsto\log(\det(M))-\mathrm{trace}(S\cdot M). \tag{9}$$

We seek to compute the critical points of ℓ_S restricted to a smooth subvariety of $\mathbb{S}^n$. Here, by a *critical point* we mean a nonsingular matrix M in the subvariety whose normal space contains the gradient vector of ℓ_S at M. This is an algebraic problem because the $\binom{n+1}{2}$ partial derivatives of ℓ_S are rational functions.

The determinant and the trace of a square matrix are invariant under conjugation. This implies the following identity for all invertible $n\times n$-matrices g:

$$\ell_{g^{-1}S(g^{-1})^T}(g^TMg)\;=\;\log(\det(g^TMg))-\mathrm{trace}(g^{-1}SMg)\;=\;\ell_S(M). \tag{10}$$

Let $\mathcal{L} \subset \mathbb{S}^n$ be a linear space of symmetric matrices (LSSM). The *ML degree* $\mathrm{mld}(\mathcal{L})$ is the number of complex critical points of ℓ_S on $\mathcal{L}$ for generic $S \in \mathbb{S}^n$. The *reciprocal ML degree* $\mathrm{rmld}(\mathcal{L})$ of $\mathcal{L}$ is the number of complex critical points of ℓ_S on $\mathcal{L}^{-1}$ for generic $S \in \mathbb{S}^n$. Both ML degrees do not depend on the choice of S, as long as S is generic. If $\mathcal{L}$ is a generic pencil then it is known that

$$\mathrm{mld}(\mathcal{L}) = \deg(\mathcal{L}^{-1}) = n-1 \qquad \text{and} \qquad \mathrm{rmld}(\mathcal{L}) = 2n-3. \tag{11}$$

The formula on the left appears in [7, Section 2.2] and that on the right is due to Coons, Marigliano and Ruddy [3]. We shall present a formula for all pencils $\mathcal{L}$.

We begin by observing that the two ML degrees are invariant under the action of $\mathrm{GL}(n)$ by congruence on $\mathbb{S}^n$. The following invariance property holds:

Lemma 4.1. The ML degree and the reciprocal ML degree of an LSSM $\mathcal{L}$ are determined by its congruence class. In particular, this holds for pencils $\mathcal{L}$.

Proof. Fix g and $\mathcal{L}$. If S is a generic matrix in $\mathbb{S}^n$ then so is $g^{-1}S(g^{-1})^T$. The image of $\mathcal{L}$ under congruence by g^T consists of all matrices $g^T M g$ where $M \in \mathcal{L}$. By (10), the likelihood function of S on $\mathcal{L}$ agrees with the likelihood function of $g^{-1}S(g^{-1})^T$ on $g^T\mathcal{L}g$. The two functions have the same number of critical points, so the linear spaces $\mathcal{L}$ and $g^T\mathcal{L}g$ have the same ML degree. The same argument works if $\mathcal{L}$ is replaced by any nonlinear variety in $\mathbb{S}^n$, such as $\mathcal{L}^{-1}$. □

We now focus on pencils ($m = 2$), and we state our main result in Section 4.

Theorem 4.2. Let $\mathcal{L}$ be a pencil with Segre symbol $\sigma = [\sigma_1, \ldots, \sigma_r]$. Then

$$\mathrm{mld}(\mathcal{L}) = r-1 \;\text{ and }\; \mathrm{rmld}(\mathcal{L}) = \sum_{i=1}^{r} \sigma_{i1} + r - 3 \;=\; \deg(\mathcal{L}^{-1}) + \mathrm{mld}(\mathcal{L}) - 1. \tag{12}$$

In the generic case, when $\sigma = [(1), \ldots, (1)]$, this recovers the formulas in (11). By Lemma 4.1, it suffices to prove (12) for the pencils $\mathcal{L} = \mathbb{C}\{P, Q\}$ in (4).

Example 4.3 ($n = 5$)**.** Let $\sigma = [(2,1), 2]$ as in Example 2.2, fix a matrix $S = (s_{ij})$ of format 5×5, and parametrize the pencil $\mathcal{L}$ by $(x, y) \mapsto xP - yQ$. Here, $\mathbb{P}\mathcal{L}^{-1}$ is a twisted cubic curve in $\mathbb{P}^{14}$. The ML degrees are $\mathrm{mld}(\mathcal{L}) = 1$ and $\mathrm{rmld}(\mathcal{L}) = 3$. To verify this, we consider the restriction of the log-likelihood function ℓ_S to $\mathcal{L}$:

$$\log\big((ax-y)^3(bx-y)^2\big) + 2s_{12}(ax-y) + s_{22}x + s_{33}(ax-y) + 2s_{45}(bx-y) + s_{55}x.$$

Its two partial derivatives are rational functions in x and y. Equating these to zero, we find that ℓ_S has a unique critical point (x^*, y^*) in $\mathcal{L}$. Its coordinates are

$$\begin{aligned}
x^* &= \big(4(a-b)s_{12} + 5s_{22} + 2(a-b)s_{33} - 6(b-a)s_{45} + 5s_{55}\big)/\Delta,\\
y^* &= \big(4a(a-b)s_{12} + (2a+3b)s_{22} + 2a(a-b)s_{33} + 6b(b-a)s_{45} + (2a+3b)s_{55}\big)/\Delta,\\
\Delta &= \big(-s_{22} + 2(a-b)s_{45} - s_{55}\big) \cdot \big(2(a-b)s_{12} + s_{22} + (a-b)s_{33} + s_{55}\big).
\end{aligned}$$

The coordinates of x^* and y^* are rational functions, i.e. $\mathrm{mld}(\mathcal{L}) = 1$. Consider now the restriction of the log-likelihood function to the reciprocal variety $\mathcal{L}^{-1}$:

$$\ell_S(x,y) = \log\big((ax-y)^3(bx-y)^2\big) - \frac{s_{11}x}{(ax-y)^2} + \frac{2s_{12}}{ax-y} + \frac{s_{33}}{ax-y} - \frac{s_{44}x}{(bx-y)^2} + \frac{2s_{45}}{bx-y}.$$

The critical equations $\frac{\partial \ell_S}{\partial x} = \frac{\partial \ell_S}{\partial y} = 0$ have three solutions (x^*, y^*). With Cardano's formula we can write their coordinates in radicals in $a, b, s_{11}, s_{12}, s_{33}, s_{44}, s_{45}$.

Sketch of proof of Theorem 4.2. After clearing denominators in the derivatives of $\ell_S(x,y)$, we obtain a system of two polynomial equations in two variables. We need to count the number of intersection points of these two curves in $\mathbb{C}^2$. This requires care because we must exclude intersection points that lie on the lines defined by the denominators, like $ax-y$, $bx-y$. We adapt the method introduced in [3], which starts from Bézout's Theorem, to carry out the desired count. □

The log-likelihood function (9) is important for applications to statistics. The matrix S is the sample covariance matrix that encodes data points in $\mathbb{R}^n$. The matrix M plays the role of the concentration matrix. Its inverse M^{-1} is the covariance matrix. These two matrices represent Gaussian distributions on $\mathbb{R}^n$. The LSSM $\mathcal{L}$ encodes linear constraints, either on the former or on the latter. For the former, we get the ML degree. For the latter, we get the reciprocal ML degree. These degrees are measures for the algebraic complexity of maximum likelihood estimation. In the language in [3, 6], $\mathrm{mld}(\mathcal{L})$ refers to the *linear concentration model,* while $\mathrm{rmld}(\mathcal{L})$ refers to the *linear covariance model*.

A real LSSM $\mathcal{L}$ is a statistical model if and only if it contains a positive definite quadric. If this holds and $\mathcal{L}$ has dimension 2 then we say that $\mathcal{L}$ is a *d-pencil* [11]. From an applied point of view, studying the maximum likelihood degrees $\mathrm{mld}(\mathcal{L})$ and $\mathrm{rmld}(\mathcal{L})$ is meaningful only for pencils $\mathcal{L}$ that are d-pencils. Thus, in statistics, we can take advantage of the following classical linear algebra fact.

Lemma 4.4. Every d-pencil $\mathcal{L}$ can be simultaneously diagonalized over $\mathbb{R}$. After a change of coordinates, $\mathcal{L}$ is spanned by the quadrics $\sum_{i=1}^n a_i x_i^2$ and $\sum_{i=1}^n x_i^2$.

Proof. We assume $n \geq 3$. A pencil is a d-pencil if and only if it has no zeros in the real projective space $\mathbb{P}^{n-1}$. This is the Main Theorem in [11]. It was also proved by Calabi in [1]. The fact that pencils without real zeros in $\mathbb{P}^{n-1}$ can be diagonalized is [11, page 221, (PM)]. It is also Remark 2 in [1, page 846]. □

Suppose there are r distinct elements in $\{a_1, a_2, \ldots, a_n\}$. The integer r is the main invariant of a d-pencil $\mathcal{L}$. If A and B are positive definite matrices that span $\mathcal{L}$, then r is the number of distinct eigenvalues of AB^{-1}. Theorem 4.2 implies:

Corollary 4.5. If $\mathcal{L}$ is a d-pencil then $\mathrm{mld}(\mathcal{L}) = \deg(\mathcal{L}^{-1}) = r-1$ and $\mathrm{rmld}(\mathcal{L}) = 2r-3$, where $\mathcal{L}$ has r distinct eigenvalues. This holds for all statistical models.

The log-likelihood function on $\mathcal{L}$ is the following function in two variables:

$$\ell_S(x,y) \;=\; \sum_{i=1}^{n} \Big(\log(a_i x + y) - s_i(a_i x + y) \Big).$$

Here $s_1,\ldots,s_n \in \mathbb{R}$ represent data. The MLE is the maximizer of $\ell_S(x,y)$ over the cone $\{(x,y) \in \mathbb{R}^2 : a_i x + y > 0 \text{ for } i = 1,\ldots,n\}$. Corollary 4.5 says that $\ell_S(x,y)$ has $r-1$ critical points. One of them is the MLE. The reciprocal log-likelihood is

$$\ell_S(x,y) \;=\; \sum_{i=1}^{n} \Big(-\log(a_i x + y) \;-\; \frac{s_i}{a_i x + y} \Big). \tag{13}$$

The invariant $\mathrm{rmld}(\mathcal{L})$ is the number of critical points (x^*,y^*) of this function with $\prod_{i=1}^{n}(a_i x^* + y^*) \neq 0$, provided $s = (s_1,\ldots,s_n)$ is generic in $\mathbb{R}^n$. Corollary 4.5 states that $\ell_S(x,y)$ has $2r-3$ complex critical points. One of them is the MLE.

In the generic case, when $r = n$, Coons et al. [3] had already shown that $\mathrm{rmld}(\mathcal{L}) = 2n-3$, but their approach left open the question whether the complex critical points are distinct for generic s, and whether they can all be real.

Conjecture 4.6. Let $\mathcal{L}$ be a d-pencil with r distinct eigenvalues. There exists $s = (s_1,\ldots,s_n) \in \mathbb{R}^n$ such that the function (13) has $2r-3$ distinct real critical points.

We can prove this conjecture for small values of n by explicit computation.

Example 4.7. Fix the pencil $\mathcal{L}$ with $n = k$ and $(a_1,\ldots,a_n) = (1,\ldots,n)$. For $n \le 7$ we found $s \in \mathbb{R}^n$ such that the reciprocal log-likelihood function ℓ_s has $2n-3$ distinct real critical points. For $n = 7$ we can take $s = (-\frac{74}{39}, \frac{13}{47}, \frac{61}{40}, \frac{1}{7}, \frac{23}{18}, -73, -\frac{27}{43})$.

5. Strata in the Grassmannian

We define a partial order on the set Segre_n of all Segre symbols for fixed n. If σ and τ are in Segre_n then we say that σ *is above* τ if $|\sigma| > |\tau|$ and τ is obtained from σ by replacing two partitions σ_i, σ_j by their sum, or if $|\sigma| = |\tau|$ if σ and τ differ in precisely one partition, with index i, and $\tau_i \lhd \sigma_i$ in the dominance order on partitions. The partial order on Segre_n is the transitive closure of the relation "is above". The top element of our poset is $[(1),(1),\ldots,(1)]$, and the bottom element is $[(2,1,\ldots,1)]$. The Hasse diagrams for $n = 3,4$ are shown in Figure 1.

We wish to study the strata Gr_σ in (1). Recall that Gr_σ is the constructible subset of $\mathrm{Gr}(2,\mathbb{S}^n)$ whose points are the pencils $\mathcal{L}$ with $\sigma(\mathcal{L}) = \sigma$. Its closure $\overline{\mathrm{Gr}_\sigma}$ is a subvariety of the Grassmannian $\mathrm{Gr}(2,\mathbb{S}^n)$. Its defining equations can

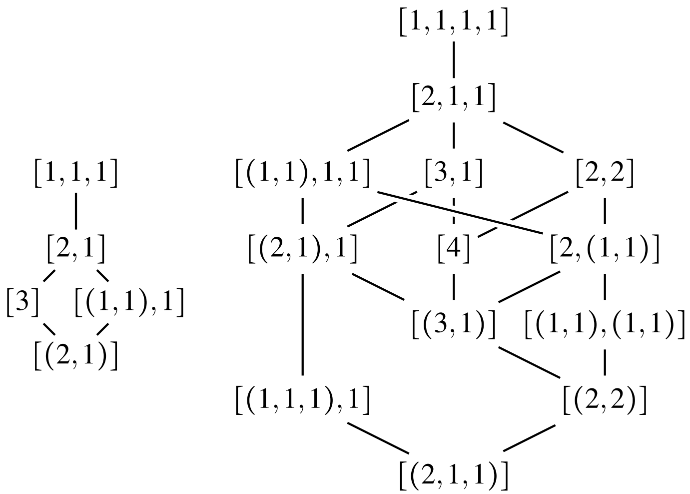

Figure 1: The posets of all Segre symbols for $n=3$ (left) and $n=4$ (right).

be written either in the $\frac{1}{8}(n+2)(n+1)n(n-1)$ Plücker coordinates, or in the $(n+1)n$ Stiefel coordinates which are the matrix entries of a basis A,B of $\mathcal{L}$.

Consider the related *Jordan stratification*. For each $\sigma\in\text{Segre}_n$, the Jordan stratum Jo_σ is the set of $n\times n$ matrices whose Jordan canonical form has pattern σ. Its closure $\overline{\text{Jo}_\sigma}$ is an affine variety in $\mathbb{C}^{n\times n}$. Its defining prime ideal consists of homogeneous polynomials in the entries of an $n\times n$ matrix $X=(x_{ij})$.

Theorem 5.1. Our poset models inclusions of both Grassmann strata and Jordan strata. That is, $\sigma\geq\tau$ in Segre_n if and only if $\overline{\text{Gr}_\sigma}\supseteq\overline{\text{Gr}_\tau}$ if and only if $\overline{\text{Jo}_\sigma}\supseteq\overline{\text{Jo}_\tau}$.

The codimensions of the Jordan strata generally differ from those of the Grassmann strata. While the $\overline{\text{Jo}_\sigma}$ are familiar from linear algebra, the $\overline{\text{Gr}_\sigma}$ capture the geometry of the varieties listed on the right in Examples 1.3 and 3.1. The codimensions are ≥ 1, unless $\sigma=[(1),\ldots,(1)]$ where both strata are dense.

Example 5.2 ($n=3$)**.** We computed the prime ideals for the Jordan strata in $\mathbb{C}^{3\times3}$, for the Plücker strata in $\text{Gr}(2,\mathbb{S}^3)\subset\mathbb{P}^{14}$, and for the Stiefel strata in $\mathbb{P}^5\times\mathbb{P}^5$:

symbol	Jordan	Plücker	Stiefel	codims	degrees
$[2,1]$	6_1	6_1	$(6,6)_1$	$1,1,1$	$6,6,[6,6]$
$[3]$	$2_1,3_1$	4_{21}	$(2,4)_1,(3,3)_1,(4,2)_1$	$2,2,2$	$6,99,[6,15,6]$
$[(1,1),1]$	3_{20}	3_{20}	$(3,3)_{20}$	$3,2,2$	$6,36,[4,4,4]$
$[(2,1)]$	2_9	2_6	$(2,2)_6$	$4,3,3$	$6,56,[4,12,12,4]$

The sextic in the first row is the discriminant of the characteristic polynomial of X. We shall explain the last row, indexed by $\sigma=[(2,1)]$. The Jordan stratum

Jo_σ has codimension 4 and degree 6. Its ideal is generated by nine quadrics, like $x_{11}x_{31} - 2x_{22}x_{31} + 3x_{21}x_{32} + x_{31}x_{33}$. Under the substitution $X = AB^{-1}$, these transform into six quadrics in Plücker coordinates, like $p_{04}p_{14} + p_{12}p_{14} - p_{03}p_{15} - p_{12}p_{23} - 3p_{02}p_{34} + 2p_{01}p_{35}$. Here $p_{01}, p_{02}, \ldots, p_{45}$ denote the 2×2 minors of

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{22} & a_{23} & a_{33} \\ b_{11} & b_{12} & b_{13} & b_{22} & b_{23} & b_{33} \end{pmatrix}.$$

The stratum Gr_σ has codimension 3 in $\mathrm{Gr}(2, \mathbb{S}^3)$ and degree 56 in the ambient $\mathbb{P}^{14}$. The six Plücker quadrics give six polynomials of bidegree $(2,2)$ in (A, B). These define a variety of multidegree $4a^3 + 12a^2b + 12ab^2 + 4b^3 \in H^*(\mathbb{P}^5 \times \mathbb{P}^5)$.

Example 5.3 ($n = 4$)**.** The column "codims" in Example 3.1 gives the codimensions of Jordan strata, Plücker strata and Stiefel strata. The last two agree; they quantify the moduli of quartic curves in $\mathbb{P}^3$ listed on the right. We found equations of low degree for the 13 strata. For instance, $\mathrm{Jo}_{[4]}$ lies on a unique quadric:

$$\begin{aligned} &3x_{11}^2 - 2x_{11}x_{22} - 2x_{11}x_{33} - 2x_{11}x_{44} + 8x_{12}x_{21} + 8x_{13}x_{31} + 8x_{14}x_{41} + 3x_{22}^2 \\ &-2x_{22}x_{33} - 2x_{22}x_{44} + 8x_{23}x_{32} + 8x_{24}x_{42} + 3x_{33}^2 - 2x_{33}x_{44} + 8x_{34}x_{43} + 3x_{44}^2. \end{aligned}$$

Sketch of proof for Theorem 5.1. For Segre symbols σ with one partition σ_1, the Jordan strata Jo_σ are the *nilpotent orbits* of Lie type A_{n-1}. Gerstenhaber's Theorem states that inclusion of nilpotent orbit closures corresponds to the dominance order $\triangleleft$ among the partitions σ_1. This explains the second condition in our definition of "is above" for the poset Segre_n. The other condition captures the degeneration that occurs when two eigenvalues come together. Generally, this leads to a fusion of Jordan blocks, made manifest by adding partitions σ_i and σ_j. The inclusions of orbit closures are preserved under the map $X \mapsto AB^{-1}$ that links Stiefel strata to Jordan strata. Furthermore, the Plücker stratification is obtained from the Stiefel stratification by taking the quotient modulo $\mathrm{GL}(2)$. This operation also preserves the combinatorics of orbit closure inclusions. □

The connection to nilpotent orbits leads to the following formula for the codimension of Jo_σ. Namely, for each partition σ_i occurring in the Segre symbol $\sigma = [\sigma_1, \ldots, \sigma_r]$, consider the conjugate partition $\sigma_i^* = (\sigma_{i1}^*, \ldots, \sigma_{in}^*)$. We have

$$\mathrm{codim}(\mathrm{Jo}_\sigma) \;=\; \sum_{i=1}^{r} \sum_{j=1}^{n} (\sigma_{ij}^*)^2 \;-\; r.$$

At present we do not know a formula for the codimension of the strata Gr_σ.

REFERENCES

[1] E. Calabi: *Linear systems of real quadratic forms*, Proc. Amer. Math. Soc **15** (1964) 844-846.

[2] A. Cayley: *Récherches sur les matrices dont les termes sont des fonctions linéaires d'une seule indéterminée*, J. Reine angew. Math. **50** (1855) 313–317.

[3] J. Coons, O. Marigliano and M. Ruddy: *Maximum likelihood degree of the two-dimensional linear Gaussian covariance model*, Algebraic Statistics **1** (2020).

[4] I. Dolgachev: *Classical Algebraic Geometry: A Modern View*, Cambridge Univ. Press, 2012.

[5] W. Hodge and D. Pedoe: *Methods of Algebraic Geometry*, Vol. II, Book IV: Quadrics and Grassmann Varieties, Cambridge University Press, 1952.

[6] B. Sturmfels, S. Timme and P. Zwiernik: *Estimating linear covariance models with numerical nonlinear algebra*, Algebraic Statistics **1** (2020).

[7] B. Sturmfels and C. Uhler: *Multivariate Gaussians, semidefinite matrix completion, and convex algebraic geometry*, Annals of the Institute of Statistical Mathematics **62** (2010) 603–638.

[8] J. J. Sylvester: *An enumeration of the contacts of lines and surfaces of the second order*, Phil. Mag. **1** (1851) 119–140.

[9] R.C. Thompson: *Pencils of complex and real symmetric and skew matrices*, Linear Algebra and its Applications **147** (1991) 323-371.

[10] F. Uhlig: *A canonical form for a pair of real symmetric matrices that generate a nonsingular pencil*, Linear Algebra and its Applications **14** (1976) 189–209.

[11] F. Uhlig: *A recurring theorem about pairs of quadratic forms and extensions: a survey*, Linear Algebra and its Applications **25** (1979) 219–237.

CLAUDIA FEVOLA
MPI for Mathematics in the Sciences, Leipzig
e-mail: Claudia.Fevola@mis.mpg.de

YELENA MANDELSHTAM
University of California, Berkeley
e-mail: yelenam@berkeley.edu

BERND STURMFELS
MPI for Mathematics in the Sciences, Leipzig
e-mail: bernd@mis.mpg.edu