

# TENSORS AND THEIR ALGEBRAS

EMANUELE VENTURA

ABSTRACT. These are informal lecture notes of the five hours mini-course “Geometry of eigenvectors of tensors” given at the Introductory School for the AGATES semester at University of Warsaw and IMPAN, held in September 2022. This work is supported by the Thematic Research Programme “Tensors: geometry, complexity and quantum entanglement”, University of Warsaw, Excellence Initiative – Research University and the Simons Foundation Award No. 663281 granted to the Institute of Mathematics of the Polish Academy of Sciences for the years 2021-2023. I thank the organizers of the semester (W. Buczyńska, J. Buczyński, F. Galuppi, J. Jelisiejew), of the Introductory School (F. Galuppi, F. Gesmundo, J. Jelisiejew) and all the participants for a very welcoming and stimulating environment. Special thanks to Tomasz Mańdziuk for useful comments and corrections on an earlier version of these notes.

## NOTATION

- Let  $\mathbb{F} = \mathbb{R}$  or  $\mathbb{C}$ . Let  $\mathbb{P}(\mathbb{F}^m)$  be the corresponding projective space of one-dimensional subspaces of  $\mathbb{F}^m$ .
- For  $d \in \mathbb{N}$ , let  $[d] = \{1, \dots, d\}$ .
- For  $i \in [d]$ , let  $m_i \geq 2$  and  $\mathbf{m} = (m_1, \dots, m_d) \in \mathbb{N}^d$ .
- Let  $\text{Seg}_{\mathbb{F}}(\mathbf{m}) = \mathbb{P}(\mathbb{F}^{m_1}) \times \dots \times \mathbb{P}(\mathbb{F}^{m_d})$  be the *Segre product*. When  $\mathbb{F} = \mathbb{C}$ , we will drop the subscript and denote the Segre product simply by  $\text{Seg}(\mathbf{m})$ .
- Let  $\mathbb{F}^{\mathbf{m}} = \mathbb{F}^{m_1} \otimes \dots \otimes \mathbb{F}^{m_d}$ , where we fix a basis in each of the vector spaces  $\mathbb{F}^{m_i}$ . A *d-way tensor*  $T = [t_{i_1, \dots, i_d}]$  is an element of the tensor space  $\mathbb{F}^{\mathbf{m}}$ .

## 1. LECTURE I

The spectral theory of matrices pervades pure and applied mathematics. On the other hand, the spectral theory of tensors is not at a comparable level of depth and breadth of applications and impact. I wish we will witness several applications to pure and applied mathematics, and especially to combinatorics in the upcoming years. The aim of these short lectures is to present the starting point of a spectral theory of tensors that emerged and was established in the works of Lim, Qi, Friedland-Ottaviani, Cartwright-Sturmfels, and many other contributors.

The theory of eigenvalues for square matrices is canonical: the eigenvalues are the roots of the characteristic polynomial. Eigenvectors are not quite canonical, although eigenspaces are. For rectangular matrices the situation already drastically changes because there is no characteristic polynomial to employ. To establish an eigenvector theory, say in the context of real matrices in  $\mathbb{R}^{m_1} \otimes \mathbb{R}^{m_2}$ , we first define scalar products in  $\mathbb{R}^{m_1}$  and  $\mathbb{R}^{m_2}$ . The choice of scalar products inducing norms is equivalent to choosing symmetric non-degenerate positive-definite bilinear forms on each of the factors. The end result is that for any matrix  $A \in \mathbb{R}^{m_1} \otimes \mathbb{R}^{m_2}$  there exists an SVD decomposition which yields a rank-one decomposition of  $A$  of the form:

$$A = \sum_{i=1}^{\text{rk}(A)} \lambda_i u_i v_i^t,$$

where  $(u_i, v_i)$  is a pair of column vectors of norm one (in the respective norms) called *singular vector pair*, and the  $\lambda_i$ 's are the positive square roots of the eigenvalues of the symmetric matrices  $AA^t$  and  $A^t A$ . To imitate the theory of real singular vector pairs of real matrices,

we should make a choice and introduce norms on each of the spaces. This is exactly what is going on behind the following definition.

**Definition 1.1 (Contraction).** For  $\{1, \dots, s\} \subset [d]$ , let  $T \in \mathbb{F}^{\mathbf{m}}$  be as above and let  $T' = [t'_{i_1, \dots, i_s}] \in \mathbb{F}^{m_1} \otimes \dots \otimes \mathbb{F}^{m_s}$  be an  $s$ -way tensor. The *contraction* of  $T$  and  $T'$  is the  $(d - s)$ -way tensor defined as

$$(T \times T')_{i_{s+1}, \dots, i_d} = \sum_{i_1, \dots, i_s} t_{i_1, \dots, i_s, i_{s+1}, \dots, i_d} t'_{i_1, \dots, i_s},$$

where  $T \times T'$  is naturally regarded as a tensor in  $\mathbb{F}^{m_{s+1}} \otimes \dots \otimes \mathbb{F}^{m_d}$ . The same construction goes similarly for any subset of  $[d]$ .

**Example 1.2.** Let  $d = 2$  and  $A \in \mathbb{F}^{m_1} \otimes \mathbb{F}^{m_2}$ . For  $v_i \in \mathbb{F}^{m_i}$ , we have

$$A \times v_2 = Av_2 \text{ and } A \times v_1 = A^t v_1,$$

where the products on the right are usual matrix multiplications.

**Definition 1.3 (Singular vector tuples of tensors).** Let  $T \in \mathbb{F}^{\mathbf{m}}$ . A vector tuple

$$(v_1, \dots, v_d) \in (\mathbb{F}^{m_1} \setminus \{0\}) \times \dots \times (\mathbb{F}^{m_d} \setminus \{0\})$$

is a *singular vector tuple* of  $T$  if

$$(1) \quad T \times (\otimes_{j \in [d] \setminus \{i\}} v_j) = \lambda_i v_i, \text{ with } \lambda_i \in \mathbb{F}, \forall i \in [d].$$

Note that whenever  $(v_1, \dots, v_d)$  is a singular vector tuple of a tensor  $T$ , then  $(\alpha_1 v_1, \dots, \alpha_d v_d)$ , where  $\prod_{i=1}^d \alpha_i \neq 0$ , is a singular vector tuple of  $T$  as well. Therefore we shall identify all singular vector tuples  $(\alpha_1 v_1, \dots, \alpha_d v_d)$  and regard them as a *single point*  $([v_1], \dots, [v_d]) \in \text{Seg}_{\mathbb{F}}(\mathbf{m})$ .

**Example 1.4.** Let  $A \in \mathbb{R}^{m_1} \otimes \mathbb{R}^{m_2}$  be a real matrix. Then, according to Definition 1.3, a pair  $(v_1, v_2) \in (\mathbb{R}^{m_1} \setminus \{0\}) \times (\mathbb{R}^{m_2} \setminus \{0\})$  is a singular vector pair if

$$Av_2 = \lambda_1 v_1 \text{ and } A^t v_1 = \lambda_2 v_2 \text{ for } \lambda_1, \lambda_2 \in \mathbb{R}.$$

Now, equip each  $\mathbb{R}^{m_i}$  with the standard Euclidean norm  $\|v\| = \sqrt{v^t v}$ . Choosing  $v_1$  and  $v_2$  as unit vectors, we find  $\lambda_1 = \lambda_2$ : multiplying both sides of the second equation by  $v_2^t$ , we find  $v_2^t A^t v_1 = \lambda_2$ ; on the other hand,  $v_2^t A^t v_1 = (Av_2)^t v_1 = \lambda_1 v_1^t v_1 = \lambda_1$ . Thus this recovers the usual notion of singular vector pair in an SVD decomposition of  $A$ .

**Example 1.5 (Exercise 6.3).** The aim of this example is to make equations (1) concrete. Let  $\mathbf{m} = (2, 2, 3)$  and  $T = [t_{i_1, i_2, i_3}] \in \mathbb{C}^{\mathbf{m}}$  be given by  $t_{1,1,1} = 1$ ,  $t_{2,2,3} = 2$ , and zero otherwise.

Suppose  $a_i, b_j$ , and  $c_k$  are the fixed bases in the three factors. Then  $T = a_1 \otimes b_1 \otimes c_1 + 2a_2 \otimes b_2 \otimes c_3$ . Fix  $(v_1, v_2, v_3) = (x_1 a_1 + x_2 a_2, y_1 b_1 + y_2 b_2, z_1 c_1 + z_2 c_2 + z_3 c_3)$ , and define  $T' = v_2 \otimes v_3$ . Thus

$$(T \times T')_i = \sum_{i_2, i_3} t_{i, i_2, i_3} y_{i_2} z_{i_3} \in \mathbb{C}.$$

So its entries are  $(T \times T')_1 = y_1 z_1$  and  $(T \times T')_2 = 2y_2 z_3$ .

Now, we take another equivalent approach. The contraction operation introduced in Definition 1.1 gives an identification of any tensor space  $\mathbb{C}^{\mathbf{m}}$  with its dual  $(\mathbb{C}^{\mathbf{m}})^\vee$ . (Recall that the contraction was motivated by a choice of non-degenerate symmetric bilinear forms on each of the factors: these bilinear forms induce the identification we are referring to.)

Therefore we may regard  $T$  as an element of  $(\mathbb{C}^{\mathbf{m}})^\vee$ . In our example,  $T \in (\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^3)^\vee = \text{Hom}(\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^3, \mathbb{C}) \cong (\mathbb{C}^2)^\vee \otimes (\mathbb{C}^2)^\vee \otimes (\mathbb{C}^3)^\vee$ .

By the universal property of tensor spaces, there is a natural bijection between linear maps  $T \in \text{Hom}(\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^3, \mathbb{C})$  and multi-linear functions  $\mathbb{C}^2 \times \mathbb{C}^2 \times \mathbb{C}^3 \rightarrow \mathbb{C}$ . Thus  $T$  can be regarded as a multi-linear function and may be written as  $T = x_1 y_1 z_1 + 2x_2 y_2 z_3$ . Under this

identifications, the contractions appearing in the equations (1) are *derivations*. The three conditions for a vector triple

$$\left( \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \right) \in (\mathbb{C}^2 \setminus \{0\}) \times (\mathbb{C}^2 \setminus \{0\}) \times (\mathbb{C}^3 \setminus \{0\})$$

to be a singular vector triple for  $T$  are the following rank constraints

$$\begin{aligned} \operatorname{rk} \begin{pmatrix} \partial T / \partial x_1 & \partial T / \partial x_2 \\ x_1 & x_2 \end{pmatrix} &\leq 1, \\ \operatorname{rk} \begin{pmatrix} \partial T / \partial y_1 & \partial T / \partial y_2 \\ y_1 & y_2 \end{pmatrix} &\leq 1, \\ \operatorname{rk} \begin{pmatrix} \partial T / \partial z_1 & \partial T / \partial z_2 & \partial T / \partial z_3 \\ z_1 & z_2 & z_3 \end{pmatrix} &\leq 1. \end{aligned}$$

You shall find 8 singular vector tuples, as points of the Segre variety  $\mathbb{P}(\mathbb{C}^2) \times \mathbb{P}(\mathbb{C}^2) \times \mathbb{P}(\mathbb{C}^3)$ . A computer algebra check of this calculation can be done as follows. I shall use a code in Macaulay2:

```
R = QQ[x_1,x_2,y_1,y_2,z_1,z_2,z_3];
T = x_1*y_1*z_1 + 2*x_2*y_2*z_3;
M1 = matrix{{diff(x_1,T), diff(x_2,T)},{x_1,x_2}};
M2 = matrix{{diff(y_1,T), diff(y_2,T)},{y_1,y_2}};
M3 = matrix{{diff(z_1,T), diff(z_2,T), diff(z_3,T)},{z_1,z_2,z_3}};
I1 = minors(2,M1), I2 = minors(2,M2), I3 = minors(2,M3);
I = I1+I2+I3;
degree I
```

Here  $I$  is the ideal we are interested in. However, the degree of this ideal is 2! What is the issue? The problem is that  $I$  is *not saturated* as multi-homogeneous ideal defining a closed subscheme in  $\operatorname{Seg}((2, 2, 3))$ . Thus we continue:

```
J = saturate(I, ideal(x_1,x_2)*ideal(y_1,y_2)*ideal(z_1,z_2,z_3));
degree J, dim J
```

Now  $J$  is the the correct ideal to look at. Its degree is 8 indeed, but its affine dimension is 3 (that is correct, but why?).

**Definition 1.6.** Let  $Q_m = \{v \in \mathbb{C}^m \mid v^t v = 0\} \subset \mathbb{C}^m$  be the *isotropic quadric*. The points in  $Q_m$  are the *isotropic vectors*. Clearly, over  $\mathbb{R}$ ,  $Q_m$  contains only the zero vector.

There is an interesting dichotomy between isotropic and non-isotropic singular vector tuples corresponding to non-zero singular values.

**Proposition 1.7.** *Suppose  $([v_1], \dots, [v_d]) \in \operatorname{Seg}(\mathbf{m})$  is a singular vector tuple of a tensor  $T$  such that the product  $\prod_{i \in [d]} \lambda_i \neq 0$ . Then exactly one of the following holds:*

- (a) *All  $v_i$  are isotropic;*
- (b) *All  $v_i$  are non-isotropic.*

*Proof.* The assumption says that  $T \times (\otimes_{j \in [d] \setminus \{i\}} v_j) = \lambda_i v_i$  with  $\lambda_i \neq 0$  for every  $i \in [d]$ . Multiplying both sides of this equality on the left by  $v_i^t$  gives:

$$T \times (\otimes_{j \in [d]} v_j) = \lambda_i v_i^t v_i \quad \forall i \in [d].$$

Note that the left-hand side is a quantity independent of  $i$ . So either for every  $i \in [d]$  one has  $v_i^t v_i \neq 0$  (which is alternative (b)), or not. In the latter case, there exists  $k \in [d]$  such that  $v_k^t v_k = 0$ . Since  $\lambda_i \neq 0$  for each  $i \in [d]$  and we have the equality

$$\lambda_1 v_1^t v_1 = \dots = \lambda_k v_k^t v_k = \dots = \lambda_d v_d^t v_d,$$

we find  $v_i^t v_i = 0 \quad \forall i \in [d]$ . □

**Remark 1.8.** Let  $T \in \mathbb{R}^{\mathbf{m}}$  have a real singular vector tuple  $([v_1], \dots, [v_d]) \in \text{Seg}_{\mathbb{R}}(\mathbf{m})$  such that  $\prod_{i \in [d]} \lambda_i = 0$ . Then for some  $k$  we must have  $\lambda_k = 0$ . On the other hand

$$\lambda_k = \frac{T \times (\otimes_{j \in [d]} v_j)}{v_k^t v_k} = 0.$$

The latter equality implies  $T \times (\otimes_{j \in [d]} v_j) = 0$ . Thus, in fact,  $\lambda_i = 0 \quad \forall i \in [d]$ .

This phenomenon does not work over  $\mathbb{C}$ , already for  $d = 2$ . Take

$$A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}, v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

Note that  $v_2$  is isotropic. Then  $Av_2 = 0$  and  $A^t v_1 = v_2$ . So  $\lambda_1 = 0$  and  $\lambda_2 = 1$ .

**Theorem 1.9** (Friedland-Ottaviani [9]). *Let  $h_1, \dots, h_d$  be commuting variables. Let  $\mathbf{m} = (m_1, \dots, m_d)$  and let  $c(\mathbf{m})$  be the coefficient of the monomial  $\prod_{i \in [d]} h_i^{m_i-1}$  in the polynomial*

$$\prod_{i \in [d]} \frac{\widehat{h}_i^{m_i} - h_i^{m_i}}{\widehat{h}_i - h_i}, \text{ where } \widehat{h}_i = \sum_{k \neq i} h_k.$$

*Let  $T \in \mathbb{C}^{\mathbf{m}}$  be general. Then  $T$  has exactly  $c(\mathbf{m})$  distinct singular vector tuples corresponding to nonzero singular values and non-isotropic. In particular, a general real-valued tensor  $T \in \mathbb{R}^{\mathbf{m}}$  has at most  $c(\mathbf{m})$  real singular vector tuples.*

**Remark 1.10.** Theorem 1.9 can be viewed as a generalization of the standard theory of singular vector pairs for real matrices. A general  $A \in \mathbb{R}^{m_1} \otimes \mathbb{R}^{m_2}$  has  $\min\{m_1, m_2\}$  distinct singular vector pairs corresponding to *nonzero* distinct singular values.

Let us check what we obtain from the Friedland-Ottaviani polynomial in Theorem 1.9. Let  $m_1 \leq m_2$ . The polynomial is

$$\begin{aligned} & \frac{h_2^{m_1} - h_1^{m_1}}{h_2 - h_1} \cdot \frac{h_1^{m_2} - h_2^{m_2}}{h_1 - h_2} = \\ & = \frac{h_1^{m_1-1} \left( \left( \frac{h_2}{h_1} \right)^{m_1} - 1 \right)}{\frac{h_2}{h_1} - 1} \cdot \frac{h_2^{m_2-1} \left( \left( \frac{h_1}{h_2} \right)^{m_2} - 1 \right)}{\frac{h_1}{h_2} - 1} = \\ & = h_1^{m_1-1} \left( \sum_{i=0}^{m_1-1} \frac{h_2^i}{h_1^i} \right) h_2^{m_2-1} \left( \sum_{j=0}^{m_2-1} \frac{h_1^j}{h_2^j} \right) = \left( \sum_{i=0}^{m_1-1} h_1^{m_1-1-i} h_2^i \right) \left( \sum_{j=0}^{m_2-1} h_1^j h_2^{m_2-1-j} \right). \end{aligned}$$

Thus a monomial in this product has the form  $h_1^{m_1-1-i+j} h_2^{m_2-1+i-j}$  which equals  $h_1^{m_1-1} h_2^{m_2-1}$  if and only if  $i = j$ . Since we have  $m_1$  values of  $i$ , the coefficient of  $h_1^{m_1-1} h_2^{m_2-1}$  is  $m_1 = \min\{m_1, m_2\}$ .

Let  $T \in S^d(\mathbb{C}^m)$  be a symmetric tensor. First we fix a natural bijection between  $S^d(\mathbb{C}^m)$  and the vector space  $\mathbb{C}[x_1, \dots, x_m]_d$ . This is defined such that the following equality holds:

$$T \times (v^{\otimes d}) = d \cdot f(v), \text{ where } v = (x_1, \dots, x_m).$$

In the case of a symmetric tensor, it is natural to consider singular vector tuples as in Definition 1.3, but with the extra requirement that all the  $v_i$ 's are the same vector  $v_i = v$ . (That makes sense because we are implicitly identifying all the factors of the tensor space.)

In this setting, the conditions for a singular vector tuple of a symmetric tensor  $T$  become

$$(2) \quad T \times (v^{\otimes(d-1)}) = \lambda v, \quad v \neq 0,$$

where the contraction is with respect to the *last*  $d-1$  indices. In coordinates, equations (2) are

$$(T \times (v^{\otimes(d-1)}))_i = \sum_{i_2, \dots, i_d} t_{i, i_2, \dots, i_d} x_{i_2} \cdots x_{i_d}.$$

In terms of  $f(v)$ , we may write the system of equations (2) as

$$(3) \quad \nabla f(v) = \lambda v, \quad v \neq 0,$$

where  $\nabla f(v)$  is the gradient of the homogeneous polynomial  $f$ . To see this, one notices that derivating both sides the defining equality  $T \times (v^{\otimes d}) = d \cdot f(v)$  we obtain  $d \cdot T \times (v^{\otimes(d-1)}) = d \cdot \nabla f(v)$ .

**Remark 1.11.** The system (2) also makes sense for tensors  $T \in \mathbb{C}^{\mathbf{m}}$  where  $\mathbf{m} = (m, m, \dots, m)$  and we implicitly identify the factors. Note that this system produces a partial symmetrization of such a tensor  $T$ : from the perspective of this system, it would be equivalent to start from a tensor  $T \in S^{d-1}(\mathbb{C}^m) \otimes \mathbb{C}^m$ .

**Definition 1.12 (E-eigenpairs).** A pair  $(\lambda, v)$  with  $v \neq 0$  satisfying (2) is an *E-eigenpair*:  $\lambda$  is an *E-eigenvalue* and  $v$  is an *E-eigenvector*. As for singular vector tuples, we will regard E-eigenvectors as points in  $\mathbb{P}(\mathbb{C}^m)$ .

**Theorem 1.13** (Cartwright-Sturmfels [4]). *If  $T \in \mathbb{C}^{\mathbf{m}}$  has finitely many eigenpairs, then counted with multiplicities there are exactly*

$$\frac{(d-1)^m - 1}{d-2} = \sum_{i=0}^{m-1} (d-1)^i.$$

*For general  $T$ , all multiplicities are one.*

**Remark 1.14.** Some remarks are now in order:

- The multiplicity in the statement refers to the scheme structure that the set of E-eigenvectors naturally possess as a subscheme of  $\mathbb{P}(\mathbb{C}^m)$ .
- For  $d = 2$ ,  $T$  is an  $m \times m$  matrix and evaluating the right-hand side we find the  $m$  eigenvalues.
- For general symmetric tensors, Theorem 1.13 is a consequence of Theorem 1.9. The same vector bundle techniques in [9] were also employed by Oeding-Ottaviani [14] to obtain the same counting as in Theorem 1.13. We shall indicate some of these techniques later but for non-symmetric tensors.
- Theorem 1.13 was also known in the complex dynamics community thanks to the work of Fornæss-Sibony [8].
- We can make heuristically sense of the number in Theorem 1.13 as follows. In the system (2), substitute  $\lambda = \tilde{\lambda}^{d-2}$ . Then we have  $m$  homogeneous polynomial equations of degree  $d-1$  in  $m+1$  variables (including  $\tilde{\lambda}$ ). When the system is general and assuming that  $\tilde{\lambda} = 0$  is not a solution, then Bezout's theorem gives  $(d-1)^m$  solutions in  $\mathbb{P}^m = \mathbb{P}(\mathbb{C}^{m+1})$ . From them we have to remove the trivial solution ( $x_i = 0$  and  $\tilde{\lambda} \neq 0$ ), and so now we have  $(d-1)^m - 1$ . Since multiplying  $\tilde{\lambda}$  by the root  $e^{2\pi i/(d-2)}$  gives the same value of  $\lambda$ , we obtain

$$\frac{(d-1)^m - 1}{d-2}$$

solutions.

## 2. LECTURE II

In this lecture, we will look at some examples and properties of E-eigenpairs.

**Example 2.1** (Exercise 6.4). We show that the number in Theorem 1.13 is tight for symmetric tensors. Let  $T = [t_{i_1, \dots, i_d}] \in \mathbb{C}^{\mathbf{m}}$  with  $t_{i_1, \dots, i_d} = 1$  for all  $1 \leq i \leq m$  and zero otherwise. Note that  $T$  is a symmetric tensor and its corresponding homogeneous polynomial is  $f(v) = \frac{1}{d} (\sum_{i=1}^m x_i^d)$ . An E-eigenpair  $(\lambda, v)$  is a solution to the equations

$$(4) \quad x_i^{d-1} = \lambda x_i, \quad \forall i \in [m].$$

Equations defining an eigenpair can be regarded as equations in the *weighted projective space*  $\mathbb{P}(1, \dots, 1, d-2) = \text{Proj}(R)$ , where  $R = \mathbb{C}[x_1, \dots, x_m, \lambda]$ , where the variables are graded  $\deg(x_i) = 1$  and  $\deg(\lambda) = d-2$ . It is by doing intersection theory on this toric variety

that Cartwright-Sturmfels proved their result. All non-trivial solutions to (4) have  $\lambda \neq 0$ . By rescaling, we may assume  $\lambda = 1$ . Fix  $\xi = e^{2\pi i/(d-2)}$  and define the set of symbols  $S = \{0, \dots, d-3, *\}$ . For a word  $s \in S^{\times m}$ , define  $x_i = \xi^{s_i}$  if  $s_i \in \mathbb{Z}$  and  $x_i = 0$  if  $s_i = *$ . This gives  $(d-1)^m - 1$  eigenpairs. Some of them are equivalent though: for each  $v \neq 0$ , we have the equivalent E-eigenvectors  $v, \xi v, \dots, \xi^{d-3}v$ . Thus there are  $\frac{(d-1)^m - 1}{d-2}$  in total. (Check that these are all.)

The next example shows that there exist real-valued tensors *without* real E-eigenvalues.

**Example 2.2** (Exercise 6.5). Let  $d$  be even and  $m = 2$ . Let  $T = [t_{i_1, \dots, i_d}] \in \mathbb{C}^m$  be such that  $t_{1,2, \dots, 2} = 1$  and  $t_{2,1, \dots, 1} = -1$ . Then the eigenpairs of  $T$  satisfy:

$$\begin{aligned} x_2^{d-1} &= \lambda x_1, \\ -x_1^{d-1} &= \lambda x_2. \end{aligned}$$

Hence  $x_1^d + x_2^d = 0$ , which has no nonzero solutions over  $\mathbb{R}$ .

**Corollary 2.3** (Exercise 6.6). *Suppose  $T$  has real entries and has finitely many E-eigenpairs. If either  $d$  or  $m$  is odd, then  $T$  has a real eigenpair.*

The situation contemplated in Example 2.2 does not happen for real symmetric tensors. This is shown in the next result.

**Theorem 2.4** (Qi [17]). *Let  $f$  be a real homogeneous polynomial of degree  $d$ . Then:*

- (a)  *$f$  has always real E-eigenvectors.*
- (b) *Suppose  $d$  is even. Then  $f$  is positive semidefinite (i.e.  $f(v) \geq 0 \ \forall v \in \mathbb{R}^m$ ) if and only if every real eigenpair of  $f$  has nonnegative E-eigenvalue.*

*Proof.* (a). We equip  $\mathbb{R}^m$  with the Euclidean topology. The polynomial  $f$  defines a continuous function on  $\mathbb{R}^m$  and so does its restriction to the unit sphere  $\{\sum_{i=1}^m x_i^2 = 1\} = \mathbb{S}^{m-1} \subset \mathbb{R}^m$ . By Lagrange multipliers, the critical points of  $f|_{\mathbb{S}^{m-1}}$  are exactly the E-eigenvectors of  $f$  (Exercise 6.7). Since  $\mathbb{S}^{m-1}$  is compact and  $f$  continuous, then there exist maxima and minima of  $f$  on the sphere. These are critical points for this function and therefore are E-eigenvectors. So real E-eigenvectors exist.

(b). First, we claim the following equivalence is true:  $f$  is positive semidefinite if and only if  $\min_{v \in \mathbb{S}^{m-1}} f(v) \geq 0$ . One direction is clear. For the converse, suppose  $\min_{v \in \mathbb{S}^{m-1}} f(v) \geq 0$ . For the sake of contradiction, assume that there exists  $y \in \mathbb{R}^m$  such that  $f(y) < 0$ . Write  $y = tv$  where  $t > 0$  and  $v \in \mathbb{S}^{m-1}$ . Then:

$$0 > f(tv) = t^d f(v) \geq 0,$$

which is a contradiction.

Now, fix any real eigenpair of  $f$ , say  $(\lambda, v)$ . Rescaling, we obtain an E-eigenpair of  $f$  on the sphere:  $(\lambda', v')$  where  $\lambda' = t^{d-2}\lambda$  and  $v' = tv$ , where  $v'^t v' = 1$ . Thus  $\nabla f(v') = \lambda' v'$ . By Euler's identify, we have

$$v'^t \nabla f(v') = d \cdot f(v') = \lambda'.$$

Any real eigenpair has  $\lambda \geq 0$  if and only if  $\lambda' \geq 0$  for any corresponding eigenpair on the sphere.

Now,  $\lambda' \geq 0$  for any eigenpair on the sphere if and only if  $\min_{v \in \mathbb{S}^{m-1}} f(v) \geq 0$ . Indeed, if  $\min_{v \in \mathbb{S}^{m-1}} f(v) \geq 0$ , then, since  $d \cdot f(v') = \lambda'$  for each E-eigenvector, we have  $\lambda' \geq 0$ . Conversely, if  $\lambda' \geq 0$  for any eigenpair on the sphere, since  $d \cdot f(v') = \lambda' \geq 0$  and the E-eigenvalues are all the critical points by (a), it follows that the minima of  $f$  on the unit sphere satisfy the same inequality. This means  $\min_{v \in \mathbb{S}^{m-1}} f(v) \geq 0$ . The previous claim finishes the proof.  $\square$

The following is a beautiful result over the reals. This shows that there exist real symmetric tensors with finitely many E-eigenpairs having a number of real eigenpairs equal to the number featured in Theorem 1.13, i.e. the maximum number.

**Theorem 2.5** (Kozhasov [12]). *For any  $d$  and  $m$ , there exist tensors in  $S^d(\mathbb{R}^m)$  whose eigenpairs are  $\frac{(d-1)^{m-1}}{d-2}$  and all real. Such symmetric tensors correspond to harmonic polynomials.*

**Remark 2.6.** Kozhasov produced *spherical harmonics*, i.e. functions in the set

$$\mathcal{H}_{d,m} = \{f|_{\mathbb{S}^{m-1}} \text{ where } f \text{ is such that } \left( \sum_{i=1}^m \frac{\partial^2}{\partial x_i^2} \right) \circ f = 0\},$$

possessing the maximum number of critical points on  $\mathbb{S}^{m-1}$ .

For  $m = 3$ , another beautiful construction providing  $d^2 - d + 1$  real E-eigenvectors is due to Abo-Seigal-Sturmfels [1].

**Theorem 2.7** (Abo-Seigal-Sturmfels, Exercise 6.8). *Let  $\mathcal{A}$  be a generic arrangement of  $d$  lines  $\ell_1, \dots, \ell_d$  in the real projective plane  $\mathbb{P}_{\mathbb{R}}^2$ . Then  $f(v) = \ell_1(v) \cdots \ell_d(v)$  has  $d^2 - d + 1$  real E-eigenvectors.*

**Remark 2.8.** The E-eigenvectors in the previous construction are the following: there are  $\binom{d}{2}$  intersection points among the real lines. By genericity, these are all distinct and the lines divide the real projective plane into  $\binom{d}{2} + 1$  connected regions. In each of them, they find a unique E-eigenvector of  $f$ . This gives the claimed number.

As a corollary of this construction, these authors prove the following.

**Corollary 2.9.** *There exists a smooth curve of degree  $d$  in the real projective plane whose complex E-eigenvectors are  $d^2 - d + 1$  real points.*

Given  $T \in \mathbb{C}^m$ , we can define a rational map  $\psi_T : \mathbb{P}^{m-1} \dashrightarrow \mathbb{P}^{m-1}$  by the formula  $\varphi_T([v]) = [T \times (v^{\otimes(d-1)})]$ . The fixed points of  $\psi_T$  are the E-eigenvectors of  $T$  with non-zero eigenvalue and the base locus of  $\psi_T$  is the set of E-eigenvectors with zero eigenvalue.

Note that the map  $\psi_T$  is defined everywhere if and only if 0 is not an eigenvalue of  $T$ . For  $d = 2$ , the condition that every eigenvalue of the matrix  $T$  is zero is equivalent to saying that  $T$  is nilpotent, i.e.  $T^m = 0$ . That means that the  $m$ th iterate of the corresponding rational map  $\psi_T$  is defined nowhere on  $\mathbb{P}^{m-1}$ . Based on these observations, it is natural to introduce the following definition.

**Definition 2.10.** A tensor  $T$  is *nilpotent* if some iterate of  $\psi_T$  is defined nowhere.

**Proposition 2.11** (Cartwright-Sturmfels [4]). *If a tensor  $T$  is nilpotent, then 0 is the only eigenvalue of  $T$ . The converse is not true: there exist tensors with only eigenvalue 0 that are not nilpotent.*

*Proof.* Suppose  $\lambda \neq 0$  is an eigenvalue corresponding to an E-eigenvector  $v \in \mathbb{C}^m \setminus \{0\}$ . Then  $[v]$  is a point fixed by  $\psi_T$  and so is a fixed point for any iterate of  $\psi_T$ . In particular,  $\psi_T^r$  for any  $r$  is defined on an open neighborhood of  $[v] \in \mathbb{P}^{m-1}$ . Thus  $T$  is not nilpotent.

To show the second statement, we will give an example. Let  $T = [t_{i_1, i_2, i_3}] \in \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$  be the 3-way tensor with  $t_{1,1,1} = t_{2,1,1} = t_{2,1,2} = 1$ , and zero otherwise. The eigenpairs of  $T$  are the solutions to

$$\begin{aligned} x_1^2 &= \lambda x_1, \\ x_1^2 + x_1 x_2 &= \lambda x_2. \end{aligned}$$

Up to equivalence, the only eigenpair is given by  $v = (0, 1)$  and  $\lambda = 0$  (check this). However, the self-map  $\psi_T$  on  $\mathbb{P}^1$  is dominant. Indeed,  $\psi_T$  acts by translation on the affine line  $\mathbb{A}^1 = \{x_1 \neq 0\}$  because on that copy of  $\mathbb{A}^1$  we have  $(x_1^2 : x_1^2 + x_1 x_2) = (x_1 : x_1 + x_2)$ . Thus  $T$  is not nilpotent but has 0 as only eigenvalue.  $\square$

Another direction of current research asks for recognizing subschemes of  $\mathbb{P}(\mathbb{C}^m)$  that are sets of E-eigenvectors of a symmetric tensor. More precisely, we first introduce a definition.

**Definition 2.12.** Let  $f \in S^d(\mathbb{C}^m)$  be a symmetric tensor. The *eigenscheme* or *eigenconfiguration* of  $f$  is the scheme  $E(f) \subset \mathbb{P}(\mathbb{C}^m) = \mathbb{P}^{m-1}$  defined by the ideal

$$I_f = 2 \times 2 \text{ minors of the matrix } \begin{pmatrix} x_1 & \cdots & x_m \\ \partial f / \partial x_1 & \cdots & \partial f / \partial x_m \end{pmatrix}.$$

Recall from Theorem 1.13 that, for  $m = 3$ ,  $E(f)$  consists of  $d^2 - d + 1$  distinct points for general  $f$  (in scheme theory terminology,  $E(f)$  is a reduced zero-dimensional scheme of degree  $d^2 - d + 1$ ).

**Remark 2.13.** The ideal  $I_f$  need not be saturated. However, suppose  $m = 3$  and  $\dim(E(f)) = 0$ , then  $I_f$  is saturated. Let  $S = \mathbb{C}[x_1, x_2, x_3]$  and  $R = S/I_f$ . Then  $\dim(E(f)) = 0$  is equivalent to the Krull dimension of  $R$  being  $\dim(R) = 1$ . Then  $\text{ht}(I_f) = \text{depth}(I_f) = 2$ . By Hilbert-Burch theorem, one has that the projective dimension of the  $S$ -module  $R$  is  $\text{pd}_S(R) \leq 2$ . By Auslander-Buchsbaum formula,  $\text{pd}_S(R) + \text{depth}(R) = \text{depth}(S) = 3$ . So  $\text{depth}(R) = 1 = \dim(R)$ . Thus  $R$  is a Cohen-Macaulay (CM) ring. Since  $R$  is CM and 1-dimensional, then  $I_f$  is saturated.

**Lemma 2.14.** Suppose  $d = 2k$  and  $f \in S^d(\mathbb{C}^m)$ . Let  $q = x_1^2 + \cdots + x_m^2$ . Then  $E(f) = E(\lambda f + \mu q^k)$  for all  $\lambda \in \mathbb{C}^*$  and  $\mu \in \mathbb{C}$ .

*Proof.* Let  $h = \lambda f + \mu q^k$ . Thus

$$x_j \frac{\partial h}{\partial x_i} - x_i \frac{\partial h}{\partial x_j} = \lambda \left( x_j \frac{\partial f}{\partial x_i} - x_i \frac{\partial f}{\partial x_j} \right).$$

Thus  $I_f = I_h$ . Hence  $E(f) = E(h)$ . □

**Definition 2.15.** Let  $s = d^2 - d + 1$  and let  $(\mathbb{P}^2)^{(s)}$  be the  $s$ th symmetric power of  $\mathbb{P}^2$ : the points of  $(\mathbb{P}^2)^{(s)}$  are the orbits of  $s$  tuples of points in  $\mathbb{P}^2$  under the symmetric group  $\mathfrak{S}_s$ . (This exists as a projective variety.) Define the rational map

$$\begin{aligned} \varphi_d : \mathbb{P}(S^d(\mathbb{C}^3)) &\dashrightarrow (\mathbb{P}^2)^{(s)}, \\ f &\mapsto E(f). \end{aligned}$$

This is well-defined at those  $f$  whose  $E(f)$  is a reduced zero-dimensional scheme. Let  $\text{Eig}_d = \overline{\text{Im}(\varphi_d)}$ . We may alternatively use the Hilbert scheme of zero-dimensional schemes of degree  $s$  on  $\mathbb{P}^2$  in the construction above. This Hilbert scheme maps surjectively on  $(\mathbb{P}^2)^{(s)}$  through a desingularization map. However, for our purposes the two constructions retain the same amount of information.

**Remark 2.16.** Define the vector space of  $2 \times 3$  matrices

$$U = \left\{ \begin{pmatrix} x_1 & x_2 & x_3 \\ \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \frac{\partial f}{\partial x_3} \end{pmatrix} \mid f \in S^d(\mathbb{C}^3) \right\}.$$

Let

$$H = \left\{ \begin{pmatrix} 1 & 0 \\ g & \lambda \end{pmatrix} \mid g \in S^{d-2}(\mathbb{C}^3), \lambda \in \mathbb{C}^*, \begin{pmatrix} 1 & 0 \\ g & \lambda \end{pmatrix} \cdot U \subseteq U \right\}.$$

This is a group acting from the left on  $U$ . Abo-Seigal-Sturmfels [1] prove that:  $\omega_1, \omega_2 \in U$  are such that  $\omega_1 = h \circ \omega_2$  for some  $h \in H$  if and only if the schemes defined by the  $2 \times 2$  minors of  $\omega_1$  and  $\omega_2$  are the same. Thus  $\dim(\text{Eig}_d) = \dim U - \dim H$ .

**Lemma 2.17** (Beorchia-Galuppi-Venturello [2]). *If  $d$  is odd, then  $H = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}, \lambda \in \mathbb{C}^* \right\}$ .*

*If  $d = 2k$ , then  $H = \left\{ \begin{pmatrix} 1 & 0 \\ g & \lambda \end{pmatrix}, g = \mu q^{k-1}, \lambda \in \mathbb{C}^*, \mu \in \mathbb{C} \right\}$ .*

*Proof.* Let  $g \in S^{d-2}(\mathbb{C}^3)$  such that  $\begin{pmatrix} 1 & 0 \\ g & \lambda \end{pmatrix} \in H$ . Then for every  $f \in S^d(\mathbb{C}^3)$ , there exists  $F \in S^d(\mathbb{C}^3)$  such that

$$\begin{pmatrix} 1 & 0 \\ g & \lambda \end{pmatrix} \cdot \begin{pmatrix} x_1 & x_2 & x_3 \\ \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \frac{\partial f}{\partial x_3} \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \\ \frac{\partial F}{\partial x_1} & \frac{\partial F}{\partial x_2} & \frac{\partial F}{\partial x_3} \end{pmatrix},$$

where  $\frac{\partial F}{\partial x_i} = x_i \cdot g + \lambda \frac{\partial f}{\partial x_i}$ . Write  $\partial_i h = \frac{\partial h}{\partial x_i}$  for any homogeneous polynomial  $h$  for simplicity of notation.

By Euler's identity, we have

$$F = \frac{1}{d}(x_1 \partial_1 F + x_2 \partial_2 F + x_3 \partial_3 F) = \frac{q \cdot g}{d} + \lambda f.$$

Thus  $x_i \cdot g + \lambda \partial_i f = \partial_i F = \frac{2x_i \cdot g + q \cdot \partial_i g}{d} + \lambda \partial_i f$ . Then

$$(5) \quad (d-2)x_i \cdot g = q \cdot \partial_i g.$$

If  $d = 2$ , then  $g$  is constant and we are done. If  $d = 3$ , then  $x_i \cdot g = q \cdot \partial_i g$ ; so  $g = 0$  because  $q$  is irreducible and  $g$  is linear, and we are done. If  $d \geq 4$ , then from equation (5) we find that there exists  $g_1 \in S^{d-4}(\mathbb{C}^3)$  such that  $g = q \cdot g_1$ . Substituting we find

$$(d-4)x_i g_1 = q \cdot \partial_i g_1,$$

and we repeat the procedure just described. In conclusion, one checks that: if  $d$  is odd then the procedure yields  $g = 0$ ; if  $d$  is even then the procedure yields  $g = \mu q^{k-1}$ .  $\square$

We are ready to state the last theorem for the lecture. The proof is an easy application of what we have shown so far.

**Theorem 2.18** ([1, 2]). *Let  $f \in S^d(\mathbb{C}^3)$  be general. Then:*

(a)  $d = 2k$ .

$$E(f) = E(f_1) \iff f_1 = \lambda f + \mu q^k, \quad \lambda \in \mathbb{C}^*, \mu \in \mathbb{C}.$$

(b)  $d = 2k + 1$ .

$$E(f) = E(f_1) \iff f_1 = \lambda f \quad \lambda \in \mathbb{C}^*.$$

*In particular,  $\dim(\text{Eig}_d) = \binom{d+2}{2} - 1$  if  $d$  is odd and  $\dim(\text{Eig}_d) = \binom{d+2}{2} - 2$  if  $d$  is even.*

### 3. LECTURE III

In this lecture we talk about *best approximations*. We will start with a small detour and apply the general results to tensors.

Let  $\langle \cdot, \cdot \rangle$  and  $\|\cdot\|$  be the standard inner product and norm on  $\mathbb{R}^m$ , that comes equipped with the standard Lebesgue measure. (Every other inner product or norm on this space will work just as well.) For  $U \subset \mathbb{R}^m$ , let  $U^\perp$  be the orthogonal of  $U$  with respect to  $\langle \cdot, \cdot \rangle$ .

Let  $C \subset \mathbb{R}^m$  be a proper closed subset. For  $x \in \mathbb{R}^m$ , consider the infimum of the *distance function* on  $C$ :

$$d(x, C) = \inf_{y \in C} \|x - y\|.$$

**Remark 3.1.** The distance function achieves its minimum at some point  $y^* \in C$ . To see this, let  $z \in C$  be an arbitrary point. Then  $d(x, C) \leq \|x - z\|$  and to search for a minimizer of the distance function we have to look for the points  $y$  such that  $\|x - y\| \leq \|x - z\|$ . Thus, by triangular inequality, we find that  $\|y\| \leq \|x\| + \|x - z\|$ . Thus we may restrict the function  $d(x, C)$  on the compact subset  $S = \{y \in C : \|y\| \leq \|x\| + \|x - z\|\}$  of  $C$  and therefore achieves a minimum on  $S$ .

**Definition 3.2.** A minimizer  $y^*$  of  $d(x, C)$  on  $C$  is called a *best approximation* of  $x$  to  $C$ .

**Lemma 3.3.** [9, Lemma 17] *Let  $C \subset \mathbb{R}^m$  be a proper closed subset. Then:*

(a)  $|d(x, C) - d(z, C)| \leq \|x - z\| \quad \forall x, z \in \mathbb{R}^m;$

- (b) The function  $d(\cdot, C)$  is differentiable almost everywhere in  $\mathbb{R}^m$ , i.e. it is differentiable in the complement of a set of Lebesgue measure zero;
- (c) Let  $x \in \mathbb{R}^m \setminus C$ . Suppose  $d(\cdot, C)$  is differentiable at  $x$  with differential  $\partial d(x)$ . Let  $y^* \in C$  be a best approximation of  $x$ . Then

$$\partial d(x)(u) = \left\langle u, \frac{1}{d(x, C)}(x - y^*) \right\rangle,$$

where  $\partial d(x)$  is regarded as a functional on  $\mathbb{R}^m$ .

**Remark 3.4.** Statement (b) is a Rademacher theorem about 1-Lipschitz functions (point (a) says that  $d(\cdot, C)$  is 1-Lipschitz). Statement (c) implies that if  $z^*$  is another best approximation of  $x$ , then:

$$\left\langle u, \frac{1}{d(x, C)}(x - y^*) \right\rangle = \left\langle u, \frac{1}{d(x, C)}(x - z^*) \right\rangle.$$

So  $\langle z^* - y^*, u \rangle = 0$  for all  $u \in \mathbb{R}^m$ . Then  $z^* - y^* \in (\mathbb{R}^m)^\perp = 0$  and so they are equal. The conclusion is that *a best approximation is unique almost everywhere*.

We now return to tensors. We equip the real tensor space  $\mathbb{R}^{\mathbf{m}}$  with the inner product  $\langle T, T' \rangle = T \times T'$  and the norm  $\|T\| = \sqrt{\langle T, T \rangle}$ . This is sometimes called *Frobenius norm* or *Bombier-Weyl norm*. For  $d = 2$ , this is the usual Frobenius norm  $\|A\| = \sqrt{\text{trace}(A^t A)}$ .

**Theorem 3.5** (Friedland-Ottaviani [9]). *Almost all tensors  $T \in \mathbb{R}^{\mathbf{m}}$  have a unique best rank-one approximation.*

*Proof.* Let  $C$  be the closed subset in  $\mathbb{R}^{\mathbf{m}}$  of all rank at most-one tensors (this is the affine cone over  $\text{Seg}(\mathbf{m})$ ). Then Lemma 3.3 yields the statement.  $\square$

Now one may ask: given a tensor  $T \in \mathbb{R}^{\mathbf{m}}$ , how does one construct a best rank-one approximation (unique or not)? To answer this question, let  $\mathbb{S}(\mathbf{m}) = \mathbb{S}^{m_1-1} \times \dots \times \mathbb{S}^{m_d-1}$  be the product of unit spheres in  $\mathbb{R}^{m_1} \times \dots \times \mathbb{R}^{m_d}$ .

Observe that for a tuple of unit vectors  $(v_1, \dots, v_d) \in \mathbb{S}(\mathbf{m})$  we have  $\|\otimes_{j \in [d]} v_j\| = \prod_{j \in [d]} \|v_j\| = 1$ . So  $\otimes_{j \in [d]} v_j$  is a unit vector in the given norm. Thus we can extend it to an orthonormal basis of  $\mathbb{R}^{\mathbf{m}}$ . Let  $U = \langle \otimes_{j \in [d]} v_j \rangle \subset \mathbb{R}^{\mathbf{m}}$ . The *projection onto  $U$*  of  $T$ , called  $P_U(T)$ , is a vector of the form

$$f_T(v_1, \dots, v_d) \cdot (\otimes_{j \in [d]} v_j),$$

where  $f_T(v_1, \dots, v_d) = \langle T, \otimes_{j \in [d]} v_j \rangle$ . Let  $P_{U^\perp}$  be the projection onto the span of the rest of the orthonormal basis. In this orthonormal basis, we have the identity

$$\|T\|^2 = \|P_U(T)\|^2 + \|P_{U^\perp}(T)\|^2.$$

Since we can scale, a best rank-one approximation of  $T$  is a solution to the minimization problem

$$(6) \quad \min_{(v_1, \dots, v_d) \in \mathbb{S}(\mathbf{m})} \min_{\alpha \in \mathbb{R}} \|T - \alpha \cdot (\otimes_{j \in [d]} v_j)\|.$$

The minimization in (6) is equivalent to the *maximization* of  $\|P_U(T)\|$  on  $\mathbb{S}(\mathbf{m})$ . Equivalently, a best rank-one approximation corresponds to the maximal value

$$(7) \quad \lambda_{\max}(T) = \max_{(v_1, \dots, v_d) \in \mathbb{S}(\mathbf{m})} |f_T(v_1, \dots, v_d)|.$$

We translated a best rank-one approximation of a tensor  $T$  into a maximization problem of a real continuous function  $f_T$  on a product of spheres. Thus we can solve this problem using Lagrange multipliers on this product manifold. This was one of the original motivations of the introduction of singular vector tuples of tensors by Lim [13] and Qi [17], viewed as *critical points* on the spheres, as in the following result.

**Proposition 3.6** (Lim, Qi). *For  $T \in \mathbb{R}^{\mathbf{m}}$ , the critical points of  $f_{T_{|\mathbb{S}(\mathbf{m})}}$  are the singular vector tuples of  $T$  satisfying*

$$(8) \quad T \times (\otimes_{j \in [d] \setminus \{i\}} v_j) = \lambda v_i \quad \forall i \in [d].$$

*Proof.* To determine the critical points of  $f_{T_{|\mathbb{S}(\mathbf{m})}}$  we use Lagrange multipliers. Let

$$g(v_1, \dots, v_d) = f_T(v_1, \dots, v_d) - \sum_{j \in [d]} \frac{\lambda_j}{2} (v_j^t v_j - 1) = \langle T, \otimes_{j \in [d]} v_j \rangle \sum_{j \in [d]} \frac{\lambda_j}{2} (v_j^t v_j - 1).$$

The critical points of  $f_{T_{|\mathbb{S}(\mathbf{m})}}$  are those of  $g(v_1, \dots, v_d)$ . These are the zero of the derivatives of  $g$ , i.e. the solutions to the following equations:

$$T \times (\otimes_{j \in [d] \setminus \{i\}} v_j) - \lambda_i v_i = 0 \quad \forall i \in [d].$$

Note that  $T \times (\otimes_{j \in [d]} v_j) = v_i^t \cdot T \times (\otimes_{j \in [d] \setminus \{i\}} v_j) = \lambda_i v_i^t v_i = \lambda_i$ . Since the left-most quantity is independent of  $i$ , all the  $\lambda_i$ 's coincide.  $\square$

**Remark 3.7.** Note that if  $(v_1, \dots, v_d)$  is a singular vector tuple of  $T$  on  $\mathbb{S}(\mathbf{m})$  then so is any vector of the form  $(\pm v_1, \dots, \pm v_d)$ . It is easy to see that we can then change signs to the  $v_i$ 's to make sure that the corresponding value of the function  $f_T(v_1, \dots, v_d) = \lambda$  is nonnegative. So the in the maximization problem (7) we may remove the absolute value.

**Remark 3.8.** Let  $m_1 \leq m_2$ . Recall that any real  $m_1 \times m_2$  matrix  $A$  has an SVD decomposition, i.e.

$$A = U \Sigma V^t,$$

where  $U, V$  are orthogonal matrices and  $\Sigma$  is an  $m_1 \times m_2$  pseudo-diagonal of the form

$$\Sigma = \text{diag}(\lambda_1, \dots, \lambda_{m_1}) = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & \dots & 0 \\ 0 & 0 & \ddots & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda_{m_1} & \dots & 0 \end{pmatrix},$$

where the singular values satisfy  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{m_1} \geq 0$ . The largest index  $r$  such that  $\lambda_r \neq 0$  is  $r = \text{rk}(A)$ . The squares of  $\lambda_i$ 's are the eigenvalues of  $AA^t$  and  $A^t A$ .

The factorization  $A = U \Sigma V^t$  is equivalent to writing

$$A v_i = \lambda u_i \quad \text{and} \quad A^t u_i = \lambda v_i \quad \text{for } 1 \leq i \leq \text{rk}(A),$$

where  $u_i$  and  $v_i$  are the  $i$ th columns of  $U$  and  $V$  respectively. For a general real matrix  $A$ , the SVD is unique up to sign change of the singular vector pair  $(u_i, v_i)$ .

Proposition 3.6 applies to matrices. So we see that the singular values in a SVD decomposition correspond to the critical points of the function  $f_T(v_1, v_2)$  on the product of the two unit spheres. In particular, best rank-one approximation for matrices corresponds to the largest singular value, i.e.  $\lambda_1$ , as we shall see in the Eckart-Young Theorem 3.9.

The result on best approximations for matrices is a beautiful result originally due to Eckart and Young.

**Theorem 3.9** (Eckart-Young, 1936). *Let  $A = U \Sigma V^t$  be an SVD decomposition of a real matrix  $A$ . Then*

- (a) *Let  $\Sigma_1 = \text{diag}(\lambda_1, 0, \dots, 0)$  and  $X_1 = U \Sigma_1 V^t$ . Then  $X_1$  is a best rank-one approximation of  $A$ , i.e.*

$$\|A - X_1\| \leq \|A - X\|,$$

*for all rank-one matrices  $X$ .*

- (b) *For any  $1 \leq s \leq \text{rk}(A)$ , let  $X_s = U \Sigma_1 V^t + \dots + U \Sigma_s V^t$ , where  $\Sigma_j = \text{diag}(0, \dots, \lambda_j, \dots, 0)$ . Then  $X_s$  is the best rank- $s$  approximation of  $A$ .*

**Remark 3.10.** For generalizations of the Eckart-Young theorem to the context of tensors, see e.g. the work by Draisma, Ottaviani and Tocino [6]. Let us just remark in passage that, for tensors, it is generally false that the singular vector tuples span the tensor itself or its best rank- $s$  approximations. One of the motivations behind the cited work is precisely trying to establish to which extent the considerations for real matrices are still true for tensors.

We shall prove a stronger result than Theorem 3.9 about matrices in Theorem 3.19. Before we proceed, we record another interesting result by Friedland-Ottaviani which requires more ideas to be proven [9, Theorem 24].

**Theorem 3.11** (Friedland-Ottaviani [9]). *Let  $\omega = (d_1, \dots, d_p) \vdash d$  and let  $T \in S^\omega(\mathbb{R}) = \bigotimes_{i \in [p]} S^{d_i}(\mathbb{R}^{m_i})$  be a partially symmetric tensor. Then every such a tensor admits a best rank-one approximation that is  $\omega$ -symmetric (i.e. a rank-one tensor in  $S^\omega(\mathbb{R})$ ). Moreover, for almost all tensors, a best rank-one approximation is unique and  $\omega$ -symmetric.*

**Remark 3.12.** Theorem 3.11 for  $\omega = (d)$ , i.e. for real symmetric tensors, was originally proven by the polish mathematician Stefan Banach.

Best approximantions are typical types of problems one faces in *geometric optimization*. We may generalize what we saw for matrices or tensors to any real algebraic variety as follows.

We may assume we have a real algebraic variety  $X \subset \mathbb{R}^m$ , viewed as a *mathematical model*. Doing experiments we have access to data points  $u \in \mathbb{R}^n$ . The goal is to find the points  $u^* \in X$  that are the *best fit* from our model to the data we measure. Often this can be formulated in terms of finding the points  $u^* \in X$  *minimizing* the squared Euclidean distance (ED)  $d(u, x) = \|u - x\|^2$  for  $x \in X$ .

By Lagrange multipliers, given a data point  $u \in \mathbb{R}^n$ , the problem is equivalent to finding:

$$x \in X_{\text{reg}} = \{\text{smooth points of } X\} \text{ and } (u - x) \perp T_x X,$$

since the vector  $(u - x)$  has the same direction as the gradient  $\nabla_x d(u, x) = 2(u - x)$ . We look at all complex solutions to this minimization problem originally over  $\mathbb{R}$ .

**Remark 3.13.** Let  $c = \text{codim}(X)$ . For  $x \in X_{\text{reg}}$ , let  $J_x$  be the Jacobian matrix calculated at  $x$ . Then  $J_x$  has rank  $c$  and  $\text{Ker}(J_x) = T_x X$ . Note that since  $(u - x) \in (T_x X)^\perp$ , the vector  $(u - x)$  is linearly dependent to the rows of  $J_x$ . It follows that, given  $x \in X_{\text{reg}}$ , a data point  $u \in \mathbb{C}^m$  has  $x$  as critical point if and only if all  $(c + 1) \times (c + 1)$  minors of the matrix  $\begin{pmatrix} u - x \\ J_x \end{pmatrix}$  vanish at  $u$ .

**Lemma 3.14** ([5]). *Let  $X \subset \mathbb{C}^m$  be a reduced and irreducible algebraic variety. For general  $u \in \mathbb{C}^m$ , the variety of critical points on  $X_{\text{reg}}$  is finite.*

*Proof.* Given  $x \in X_{\text{reg}}$ , all the  $u \in \mathbb{C}^m$  such that  $(u - x) \perp T_x X$  form an affine space of dimension  $c$  because  $\dim(T_x X) = m - c$ . Hence the variety of pairs (data point, critical point) =  $(u, x) \in X_{\text{reg}} \times \mathbb{C}^m$  is irreducible of dimension  $m$ . Hence the fiber of the projection onto the second factor must be finite for a general  $u \in \mathbb{C}^m$ .  $\square$

**Definition 3.15 (The ED degree).** The finite number established in Lemma 3.14 is the *Euclidean Distance degree* (ED degree for short) of  $X$  and it is denoted  $\text{EDdegree}(X)$ .

We may reformulate some of the previous results on the number of singular vector tuples/eigenvectors for general tensors in terms of ED degrees.

**Theorem 3.16** ([5]). *The ED degree of the real Segre product (in its Segre embedding)  $\text{Seg}_{\mathbb{R}}(\mathbf{m})$  is the quantity in Theorem 1.9. The ED degree of the real Veronese variety  $\nu_d(\mathbb{P}(\mathbb{R}^m))$  is the quantity in Theorem 1.13.*

We now find the ED degree of the variety of rank  $r$  matrices  $\mathcal{M}_r \setminus \mathcal{M}_{r-1}$ , revisiting the Eckart-Young Theorem 3.9. For this, we follow [15, Theorem 2.9]. We first establish two observations.

**Lemma 3.17.** *if  $A_1 = u_1 \otimes v_1, A_2 = u_2 \otimes v_2$ , then the Frobenius inner product is*

$$\langle A_1, A_2 \rangle = \langle u_1, u_2 \rangle \cdot \langle v_1, v_2 \rangle,$$

*where the inner products on the left are the standard Euclidean inner product.*

This is left as an exercise.

**Lemma 3.18.** *Let  $B \in \mathbb{R}^{m_1} \otimes \mathbb{R}^{m_2}$ . If  $\langle B, \mathbb{R}^{m_1} \otimes v \rangle = 0$ , then  $\langle \text{Row}(B), v \rangle = 0$ , where  $\text{Row}(B)$  is the row space of  $B$ . If  $\langle B, u \otimes \mathbb{R}^{m_2} \rangle = 0$ , then  $\langle \text{Col}(B), u \rangle = 0$ , where  $\text{Row}(B)$  is the column space of  $B$ .*

*Proof.* Let  $\{e_1, \dots, e_{m_1}\}$  be the canonical basis of  $\mathbb{R}^{m_1}$ . Then  $\langle B, e_k \otimes v \rangle = 0$  for  $1 \leq k \leq m_1$  by assumption. Thus

$$0 = \text{trace}(B \cdot (v \cdot e_k^t)) = \text{trace}(B \cdot V_k) = \langle B_k, v \rangle,$$

where  $V_k$  is the matrix whose  $k$ th column is the vector  $v$  and is zero otherwise, whereas  $B_k$  is the  $k$ th row of  $B$ . The equality implies  $\langle \text{Row}(B), v \rangle = 0$ .  $\square$

**Theorem 3.19** ([15]). *Let  $\mathcal{M}_r$  be the variety of real  $m_1 \times m_2$  matrices with  $m_1 \leq m_2$  of rank  $\leq r$ . Let  $A = U\Sigma V^t$  and let  $1 \leq r \leq \text{rk}(A)$ . All the critical points with respect to the Frobenius distance function from  $A$  to the smooth variety  $\mathcal{M}_r \setminus \mathcal{M}_{r-1}$ , are the matrices*

$$U(\Sigma_{i_1} + \dots + \Sigma_{i_r})V^t,$$

*where  $\Sigma_{i_\ell} = \text{diag}(0, \dots, \lambda_{i_\ell}, \dots, 0)$  and the  $\lambda_j$ 's are the singular values of  $A$ . If the nonzero singular values of  $A$  are distinct, then the number of critical points is  $\binom{\text{rk}(A)}{r}$ . In particular,  $\text{EDdegree}(\mathcal{M}_r) = \binom{m_1}{r}$ .*

*Proof.* The matrix  $U(\Sigma_{i_1} + \dots + \Sigma_{i_r})V^t$  is a critical point of the distance function from  $A$  to  $\mathcal{M}_r \setminus \mathcal{M}_{r-1}$  if and only if the vector  $A - U(\Sigma_{i_1} + \dots + \Sigma_{i_r})V^t$  is orthogonal to the tangent space

$$T_{U(\Sigma_{i_1} + \dots + \Sigma_{i_r})V^t} \mathcal{M}_r = (\mathbb{R}^{m_1} \otimes v_{i_1} + u_{i_1} \otimes \mathbb{R}^{m_2}) + \dots + (\mathbb{R}^{m_1} \otimes v_{i_r} + u_{i_r} \otimes \mathbb{R}^{m_2}).$$

From the given SVD of  $A$ , we have

$$A - U(\Sigma_{i_1} + \dots + \Sigma_{i_r})V^t = U(\Sigma_{j_1} + \dots + \Sigma_{j_\ell})V^t = \lambda_{j_1} u_{j_1} \otimes v_{j_1} + \dots + \lambda_{j_\ell} u_{j_\ell} \otimes v_{j_\ell},$$

where  $\{j_1, \dots, j_\ell\} = \{1, \dots, \text{rk}(A)\} \setminus \{i_1, \dots, i_r\}$ . Let  $\{e_1, \dots, e_{m_1}\}$  be the canonical basis of  $\mathbb{R}^{m_1}$ . By Lemma 3.17, we have

$$\langle \lambda_{j_h} u_{j_h} \otimes v_{j_h}, e_\ell \otimes v_{i_k} \rangle = \lambda_{j_h} \langle u_{j_h}, e_\ell \rangle \langle v_{j_h}, v_{i_k} \rangle = 0,$$

because  $\langle v_{j_h}, v_{i_k} \rangle = 0$  by orthogonality. Similarly,  $\langle \lambda_{j_h} u_{j_h} \otimes v_{j_h}, u_{i_k} \otimes e_\ell \rangle = 0$ . Hence  $A - U(\Sigma_{i_1} + \dots + \Sigma_{i_r})V^t$  is orthogonal to  $T_{U(\Sigma_{i_1} + \dots + \Sigma_{i_r})V^t} \mathcal{M}_r$ . So  $U(\Sigma_{i_1} + \dots + \Sigma_{i_r})V^t$  is a rank  $r$  critical point.

Conversely, let  $B \in \mathcal{M}_r \setminus \mathcal{M}_{r-1}$  be a critical point for the Frobenius distance function from  $A$  to  $\mathcal{M}_r \setminus \mathcal{M}_{r-1}$ . Then  $A - B$  is orthogonal to the tangent space  $T_B \mathcal{M}_r$ . Let  $B' = U'(\Sigma'_1 + \dots + \Sigma'_r)V^t$  and  $A - B = U''(\Sigma''_1 + \dots + \Sigma''_r)V^t$  be SVD for  $B$  and  $A - B$ .

Since  $A - B$  is orthogonal to  $T_B \mathcal{M}_r = (\mathbb{R}^{m_1} \otimes v'_1 + u'_1 \otimes \mathbb{R}^{m_2}) + \dots + (\mathbb{R}^{m_1} \otimes v'_r + u'_r \otimes \mathbb{R}^{m_2})$ . By Lemma 3.18, we have

$$\langle \text{Col}(A - B), u'_k \rangle = 0 \text{ and } \langle \text{Row}(A - B), v'_k \rangle = 0 \quad \forall k \in [r].$$

In particular,  $\text{Col}(A - B)$  is a subspace of  $\langle u'_1, \dots, u'_r \rangle^\perp$ ;  $\text{Row}(A - B)$  is a subspace of  $\langle v'_1, \dots, v'_r \rangle^\perp$ . Note that  $\ell \leq m_1 - r$ . Let  $\text{Col}(A - B) = \langle u''_1, \dots, u''_\ell \rangle$  and  $\text{Row}(A - B) = \langle v''_1, \dots, v''_\ell \rangle$ . Thus the orthonormal columns  $u''_1, \dots, u''_\ell, u'_1, \dots, u'_r$  can be completed with other  $m_1 - \ell - r$  orthonormal columns in  $\mathbb{R}^{m_1}$  to obtain an  $m_1 \times m_1$  orthogonal matrix  $U$ . Similarly, one obtains an  $m_2 \times m_2$  orthogonal matrix  $V$ .

Thus

$$A - B = U \begin{pmatrix} \Sigma'' & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} V^t \quad \text{and} \quad B = U \begin{pmatrix} 0 & 0 & 0 \\ 0 & \Sigma' & 0 \\ 0 & 0 & 0 \end{pmatrix} V^t,$$

where  $\Sigma'' = \text{diag}(\lambda''_1, \dots, \lambda''_\ell)$  and  $\Sigma' = \text{diag}(\lambda'_1, \dots, \lambda'_r)$ . Therefore

$$A = (A - B) + B = U \begin{pmatrix} \Sigma'' & 0 & 0 \\ 0 & \Sigma' & 0 \\ 0 & 0 & 0 \end{pmatrix} V^t,$$

where the last factorization can be transformed into an SVD of  $A$  by simply reordering the positive entries in  $\Sigma'$  and  $\Sigma''$  (and the corresponding columns in  $U$  and  $V$ ). Hence the critical point  $B$  then has the desired form. The last two statements are simple consequences.  $\square$

**Remark 3.20.** Note that the lengths of the rank-one critical points are the  $\lambda_i$ 's, i.e. the corresponding singular values.

#### 4. LECTURE IV

- Some of the main references for vector bundles and intersection theory in algebraic geometry (Chern classes, Chern polynomials, and much more): [7, 10, 11]; see also [3, Chapter IV] for a topological introduction.

In this lecture we first recall the notion of a vector bundle. By a *variety*  $X$  we shall mean an integral separated scheme of finite type  $X$  over  $\mathbb{C}$  (or any algebraically closed field of characteristic zero). An example of such a scheme is a projective variety  $X \subset \mathbb{P}(\mathbb{C}^m)$ . We will consider only the closed points of  $X$ , i.e. those  $p \in X$  such that the residue field  $k(p) \cong \mathbb{C}$ .

**Definition 4.1 (Vector bundle).** A variety  $\mathcal{E}$  is a *vector bundle of rank  $r$*  over a variety  $X$  if there is a morphism

$$\pi : \mathcal{E} \longrightarrow X,$$

and an open cover  $\{U_i\}_{i \in I}$  (that can be taken to be finite) of  $X$  such that:

- (a) For every  $i \in I$ , there is an isomorphism

$$\varphi_i : \pi^{-1}(U_i) \cong U_i \times \mathbb{C}^r,$$

inducing an isomorphism  $\pi^{-1}(p) \cong \{p\} \times \mathbb{C}^r \cong \mathbb{C}^r$  such that  $\text{pr}_{U_i} \circ \varphi_i = \pi|_{\pi^{-1}(U_i)}$ . (This says that, locally,  $\mathcal{E}$  looks like a product space.)

- (b) For every  $i, j$ , there is  $g_{ij} \in \text{GL}_r(\Gamma(U_i \cap U_j, \mathcal{O}_X))$  such that the induced isomorphism

$$\varphi_j^{-1} \circ \varphi_i : (U_i \cap U_j) \times \mathbb{C}^r \cong (U_i \cap U_j) \times \mathbb{C}^r$$

coincides with the linear isomorphism  $\text{Id}_{|U_i \cap U_j} \times g_{ij}$ . These matrices  $g_{ij}$  are called *transition functions* and they satisfy the following compatibilities:  $g_{ij} = g_{ji}^{-1}$  on  $U_i \cap U_j$  and  $g_{ik} = g_{ij} \circ g_{jk}$  on triple intersections  $U_i \cap U_j \cap U_k$ .

The data  $\{U_i\}_{i \in I}$  (where, for simplicity, we drop the maps  $\varphi_i$ ) is called a *trivialization* of  $\mathcal{E}$ . The linear isomorphisms specified in (b) allow us to give a well-defined vector space structure to  $\pi^{-1}(p) \cong \mathbb{C}^r$  for each  $p \in X$ . This vector space is denoted  $\mathcal{E}_p$  and called the *fibre* at  $p \in X$ .

**Definition 4.2.** We give some more definitions that are useful later.

- Let  $\mathcal{E}$  be a vector bundle of rank  $r$  on a variety  $X$ . A *section*  $\sigma$  of  $\mathcal{E}$  over an open subset  $U \subset X$  is a morphism

$$\sigma : U \longrightarrow \mathcal{E}$$

such that  $\pi \circ \sigma(p) = p$  for every  $p \in U$ .

- Let  $\{U_i\}_{i \in I}$  be a trivialization of  $\mathcal{E}$ . Then  $\sigma(p) = \sum_{j=1}^r \sigma_{ji}(p) e_{ji}$  for each  $p \in U \cap U_i$ , where we identify the fibre  $\mathcal{E}_p$  with a fixed vector space with basis  $\{e_{ji}\}_{j \in [r]}$  at each  $p \in U_i$ .
- When  $U = X$ ,  $\sigma$  is called a *global section*. Global sections of  $\mathcal{E}$  change as we move around  $X$  according to the transition functions prescribed in a trivialization of  $\mathcal{E}$ . More precisely, define  $\sigma_i = \sigma|_{U_i}$ . Then  $\sigma_i|_{U_i \cap U_j} = g_{ij} \circ \sigma_j|_{U_i \cap U_j}$ .

- Given transition functions  $g_{ij}$  on an open cover of  $X$  satisfying the compatibilities we mentioned, there is a vector bundle  $\pi : \mathcal{E} \rightarrow X$ , whose transition functions are precisely the ones we prescribed.
- Sections of a vector bundle form a locally free sheaf. Conversely, given a locally free sheaf, one can construct a vector bundle: let  $\mathcal{F}$  be a locally free sheaf; the fibre of the corresponding vector bundle at  $p \in X$  is  $\mathcal{F}_p \otimes k(p)$ , where  $\mathcal{F}_p$  is the stalk of the sheaf at  $p$  and  $k(p)$  is the residue field of  $p$ , i.e.  $\mathcal{O}_{X,p}/\mathfrak{m}_p$ .
- The vector space of global sections of  $\mathcal{E}$  is  $H^0(\mathcal{E})$ . A subspace  $V \subset H^0(\mathcal{E})$  *generates*  $\mathcal{E}$  if, for every  $p \in X$ ,

$$V(p) = \{\text{the value of all sections in } V \text{ at } p\} = \mathcal{E}_p$$

Such a vector bundle  $\mathcal{E}$  is said to be *globally generated*.

**Remark 4.3.** On a projective variety  $X$ ,  $H^0(\mathcal{E})$  is a finite-dimensional vector space over the ground field  $\mathbb{C}$ , whereas on an arbitrary variety it might not be.

- A vector bundle of rank one is called *line bundle*.
- A vector bundle  $\mathcal{E}$  is *trivial*  $\iff$  it has  $r$  everywhere independent global sections  $\sigma_0, \dots, \sigma_{r-1}$ . This means that at each  $p \in X$  their values are independent vectors in the fibre  $\mathcal{E}_p$ .

**Definition 4.4.** Some new vector bundles from old ones:

- Given vector bundles  $\mathcal{E}_1$  and  $\mathcal{E}_2$  over a variety  $X$  we may form their direct sum  $\mathcal{S} = \mathcal{E}_1 \oplus \mathcal{E}_2$ . The fibre at  $p \in X$  of  $\mathcal{S}$  is  $\mathcal{S}_p = \mathcal{E}_{1,p} \oplus \mathcal{E}_{2,p}$ .
- Given vector bundles  $\mathcal{E}_1$  and  $\mathcal{E}_2$  over a variety  $X$  we may form their tensor product  $\mathcal{F} = \mathcal{E}_1 \otimes \mathcal{E}_2$ . The fibre at  $p \in X$  of  $\mathcal{F}$  is  $\mathcal{F}_p = \mathcal{E}_{1,p} \otimes \mathcal{E}_{2,p}$ .
- Given vector bundles  $\mathcal{E}_1$  and  $\mathcal{E}_2$  over a variety  $X$  we may define  $\mathcal{G} = \text{Hom}(\mathcal{E}_1, \mathcal{E}_2)$ . The fibre at  $p \in X$  of  $\mathcal{G}$  is  $\mathcal{G}_p = \text{Hom}(\mathcal{E}_{1,p}, \mathcal{E}_{2,p})$ . In particular, when  $\mathcal{E}_2$  is a trivial line bundle, then  $\mathcal{G} = \mathcal{E}_1^\vee$  is the *dual vector bundle* of  $\mathcal{E}_1$ .
- Let  $\mathcal{E}_i$  be vector bundles on varieties  $X_i$ , for  $i = 1, 2$ . Given the projection

$$\pi_1 : X_1 \times X_2 \rightarrow X_1,$$

one can define the *pull-back*  $\pi_1^* \mathcal{E}_1$  of  $\mathcal{E}_1$  as the vector bundle on  $X_1 \times X_2$  whose fibre at  $(p, q) \in X_1 \times X_2$  is  $\mathcal{E}_{1,p}$ .

In the following, we will implicitly identify vector bundles with their locally free sheaves of sections.

**Definition 4.5 (Some bundles on projective space).** We shall work with three bundles on projective space  $\mathbb{P}(\mathbb{C}^m)$ .

- The trivial bundle  $\mathcal{F}$  on  $\mathbb{P}(\mathbb{C}^m)$  of rank  $m$ . Its fibre at  $[v] \in \mathbb{P}(\mathbb{C}^m)$  is  $\mathcal{F}_{[v]} = \mathbb{C}^m$ .
- The *tautological bundle*  $\mathcal{T}$  on  $\mathbb{P}(\mathbb{C}^m)$  has rank one and  $\mathcal{T}_{[v]} = \langle v \rangle \subset \mathbb{C}^m$ . It is a subbundle of  $\mathcal{F}$  and its dual is called *hyperplane line bundle*  $\mathcal{H} = \mathcal{T}^\vee$ . These bundles are usually denoted  $\mathcal{T} = \mathcal{O}_{\mathbb{P}^{m-1}}(-1)$  and  $\mathcal{H} = \mathcal{O}_{\mathbb{P}^{m-1}}(1)$ . However, we will use the former notation.
- There is a short exact sequence of locally free sheaves (called *Euler's sequence*) on  $\mathbb{P}(\mathbb{C}^m)$ :

$$0 \rightarrow \mathcal{T} \rightarrow \mathcal{F} \rightarrow \mathcal{Q} \rightarrow 0,$$

where  $\mathcal{Q}$  is a vector bundle called *quotient bundle*. Its fibre at  $[v] \in \mathbb{P}(\mathbb{C}^m)$  is the quotient  $\mathcal{Q}_{[v]} = \mathcal{F}_{[v]} / \mathcal{T}_{[v]}$ .

We will introduce a measure of *non-triviality* of a globally generated vector bundle  $\mathcal{E}$  on  $X$ . Such a measure is the size of the locus where general global sections  $\sigma_0, \dots, \sigma_{r-i}$  are dependent vectors. For a trivial bundle of rank  $r$ , we will have  $r$  global sections that are always independent and so these loci are all empty.

The locus above is the zero subscheme of  $X$  of the section

$$\sigma_0 \wedge \cdots \wedge \sigma_{r-i} \in \bigwedge^{r-i+1} \mathcal{E}.$$

This is denoted  $D = V(\sigma_0 \wedge \cdots \wedge \sigma_{r-i})$  and called *degeneracy locus* of  $\sigma_0, \dots, \sigma_{r-i}$ . Locally on  $X$ ,  $D$  is cut out by the  $(r-i+1) \times (r-i+1)$  minors of an  $(r-i+1) \times r$  matrix. So, as for determinantal ideals, one expects these loci to have codimension  $i$  in  $X$ .

**Lemma 4.6** (Kleiman, Exercise 6.19). *Suppose  $\mathcal{E}$  is a globally generated vector bundle of rank  $r$  on a variety  $X$  and let  $1 \leq i \leq r$ . Let  $\sigma_0, \dots, \sigma_{r-i}$  be general elements in  $H^0(\mathcal{E})$ . Then, when non-empty, the degeneracy locus  $D = V(\sigma_0 \wedge \cdots \wedge \sigma_{r-i})$  is generically reduced and has pure codimension  $i$  in  $X$ .*

Generically reducedness holds because our  $X$  is defined in characteristic zero, otherwise it is generally false.

**Theorem 4.7.** *There is a way of assigning in a graded ring  $A^\bullet(X)$ , called the Chow ring, to each vector bundle  $\mathcal{E}$  on a smooth variety  $X$  a polynomial*

$$c(t, \mathcal{E}) = 1 + c_1(\mathcal{E})t + \cdots + c_r(\mathcal{E})t^r \in A^\bullet(X)[t]$$

called Chern polynomial and whose coefficients  $c_i(\mathcal{E})$  are the Chern classes, such that the following properties hold:

(a) *If  $\mathcal{L}$  is a line bundle on  $X$ , the Chern polynomial is*

$$c(t, \mathcal{L}) = 1 + c_1(\mathcal{L})t,$$

where  $c_1(L) \in A^1(X) \cong \text{Pic}(X)$  is the class of the divisor of zeros and poles of any rational section of  $L$ .

(b) *If  $\mathcal{E}$  is globally generated and  $\sigma_0, \dots, \sigma_{r-i}$  are general global sections of  $\mathcal{E}$ , then the  $i$ th Chern class  $c_i(\mathcal{E}) = [D] \in A^i(X)$ , where  $[D]$  is the class of the corresponding degeneracy locus.*

(c) **(Whitney's formula)** *Given a short exact sequence*

$$0 \longrightarrow \mathcal{E} \longrightarrow \mathcal{F} \longrightarrow \mathcal{G} \longrightarrow 0$$

of vector bundles on  $X$ , then  $c(t, \mathcal{F}) = c(t, \mathcal{E}) \cdot c(t, \mathcal{G}) \in A^\bullet(X)[t]$ .

(d) *The dual vector bundle  $\mathcal{E}^\vee$  satisfies  $c_i(\mathcal{E}^\vee) = (-1)^i c_i(\mathcal{E})$ . (Exercise 6.13.)*

(e) *If  $r = \text{rank}(\mathcal{E}) = \dim(X)$  and  $\mathcal{E}$  is globally generated, then the zero locus of a general  $\sigma \in H^0(\mathcal{E})$  consists of  $\deg(c_r(\mathcal{E}))$  simple points on  $X$ . The class  $c_r(\mathcal{E})$  is called top Chern class and the number  $\deg(c_r(\mathcal{E}))$  is called the top Chern number, often denoted by  $\int_X c_r(\mathcal{E})$ .*

**Remark 4.8.** Statement (e) in Theorem 4.7 is false for non-globally generated vector bundles. See Exercise 6.20 for a counterexample.

**Example 4.9 (Chow ring of a projective space).** The Chow ring of  $\mathbb{P}(\mathbb{C}^m)$  is isomorphic to its singular cohomology ring:

$$A^\bullet(\mathbb{P}(\mathbb{C}^m)) \cong \mathbb{Z}[h]/(h^m),$$

where  $h = c_1(\mathcal{H}) = c_1(\mathcal{O}_{\mathbb{P}^{m-1}}(1)) \in A^1(\mathbb{P}(\mathbb{C}^m))$  is the hyperplane class. The monomial  $h^{m-1} \in A^{m-1}(\mathbb{P}(\mathbb{C}^m)) \cong \mathbb{Z}$  is the class of a point.

**Example 4.10 (Chow ring of a Segre product).** The Segre product  $\text{Seg}(\mathbf{m})$  comes equipped with projections maps

$$\pi_i : \text{Seg}(\mathbf{m}) \longrightarrow \mathbb{P}(\mathbb{C}^{m_i}).$$

Put  $h_i = c_1(\pi_i^* \mathcal{H}_i)$ , where  $\mathcal{H}_i$  is the hyperplane bundle on  $\mathbb{P}(\mathbb{C}^{m_i})$ . Then

$$A^\bullet(\text{Seg}(\mathbf{m})) \cong \mathbb{Z}[h_1, \dots, h_d]/(h_1^{m_1}, \dots, h_d^{m_d}).$$

Here the class of a point is the monomial  $\prod_{i=1}^d h_i^{m_i-1}$ .

**Example 4.11 (Chern polynomial of the quotient bundle).** Recall that on  $\mathbb{P}(\mathbb{C}^m)$  we have the Euler's exact sequence:

$$0 \longrightarrow \mathcal{T} \longrightarrow \mathcal{F} \longrightarrow \mathcal{Q} \longrightarrow 0,$$

where  $\mathcal{T}^\vee = \mathcal{H}$ , the hyperplane line bundle. As before,  $h = c_1(\mathcal{H}) \in A^1(\mathbb{P}(\mathbb{C}^m))$  and so  $c(t, \mathcal{H}) = 1 + ht$ . Hence  $c(t, \mathcal{T}) = 1 - ht$ . By Whitney's formula (Theorem 4.7(c)), we find

$$1 = c(t, \mathcal{F}) = c(t, \mathcal{T}) \cdot c(t, \mathcal{Q}).$$

Thus

$$(9) \quad c(t, \mathcal{Q}) = \frac{1}{1 - ht} = 1 + \sum_{j=1}^{m-1} h^j t^j \quad (h^m = 0).$$

## 5. LECTURE V

In this lecture we continue with vector bundles on Segre product and we shall see how to interpret singular vector tuples in terms of vector bundles.

**Definition 5.1 (Some vector bundles on Segre products).** We introduce the following vector bundles on  $\text{Seg}(\mathbf{m})$ :

- Let  $\widehat{\mathcal{T}}_i = \bigotimes_{j \in [d] \setminus \{i\}} \mathcal{T}_j$ . Here  $\mathcal{T}_j$  is the tautological bundle on the  $j$ th factor  $\mathbb{P}(\mathbb{C}^{m_j})$ . Note that  $(\widehat{\mathcal{T}}_i)^\vee \cong \bigotimes_{j \in [d] \setminus \{i\}} \mathcal{H}_j$ .
- Let  $R_i = \text{Hom}(\widehat{\mathcal{T}}_i, \pi_i^* \mathcal{Q}_i) \cong (\widehat{\mathcal{T}}_i)^\vee \otimes \pi_i^* \mathcal{Q}_i$ , where  $\mathcal{Q}_i$  is the quotient bundle on the  $i$ th factor  $\mathbb{P}(\mathbb{C}^{m_i})$ .
- Let  $R(\mathbf{m}) = \bigoplus_{i \in [d]} R_i$ . Notice that  $\text{rank}(R(\mathbf{m})) = \sum_{i=1}^d \text{rank}(R_i) = \sum_{i=1}^d (m_i - 1) = \dim(\text{Seg}(\mathbf{m}))$ .

**Remark 5.2.** By Whitney's formula, we find

$$c(t, R(\mathbf{m})) = \prod_{i \in [d]} c(t, R_i).$$

Since  $\text{rank}(R(\mathbf{m})) = \dim(\text{Seg}(\mathbf{m}))$ , the top Chern number of  $R(\mathbf{m})$  is the coefficient of the monomial  $\prod_{i=1}^d h_i^{m_i-1}$  in  $c(1, R(\mathbf{m})) \in A^\bullet(\text{Seg}(\mathbf{m}))$ .

Recall the Friedland-Ottaviani Theorem 1.9: a general  $T \in \mathbb{C}^{\mathbf{m}}$  has exactly  $c(\mathbf{m})$  simple singular vector tuples corresponding to nonzero singular vector tuples and that are non-isotropic. The number  $c(\mathbf{m})$  is the top Chern number of  $R(\mathbf{m})$ . The detailed proof of Theorem 1.9 is technical and beyond the scope of these short lectures. Nevertheless, we shall answer these questions:

- (a) How to calculate  $c(\mathbf{m})$  and why do the polynomials in the statement of Theorem 1.9 appear?
- (b) How the vector bundle  $R(\mathbf{m})$  plays a role?

In the following lemma, we give an answer to (a).

**Lemma 5.3.** *The top Chern number  $c(\mathbf{m})$  of  $R(\mathbf{m})$  is the coefficient of the monomial  $\prod_{i=1}^d h_i^{m_i-1}$  in the polynomial*

$$(10) \quad \prod_{i \in [d]} \frac{\widehat{h}_i^{m_i} - h_i^{m_i}}{\widehat{h}_i - h_i}, \quad \text{where } \widehat{h}_i = \sum_{k \neq i} h_k.$$

*Proof.* If we factor the Chern polynomial of a vector bundle  $\mathcal{E}$  as

$$c(t, \mathcal{E}) = \prod_{j=1}^{\text{rank}(\mathcal{E})} (1 + \xi_j(\mathcal{E})t),$$

the classes  $\xi_j(\mathcal{E})$  are called *Chern roots*. These are classes in the Chow ring of the projective bundle  $\mathbb{P}(\mathcal{E})$ . From these classes, one defines the *Chern character* of  $\mathcal{E}$  as  $ch(\mathcal{E}) = \sum_{j=1}^{\text{rank}(\mathcal{E})} e^{\xi_j(\mathcal{E})}$ , where  $e^{\xi_j(\mathcal{E})}$  is a class (formally expanding the exponential) in the rational Chow ring of the projective bundle of  $\mathcal{E}$ .

Let  $\xi = e^{2\pi i/m_i}$ . Then

$$\frac{\prod_{k=0}^{m_i-1} (1 - \xi_i^k x)}{1 - x} = \frac{1 - x^{m_i}}{1 - x} = \sum_{k=0}^{m_i-1} x^k.$$

By equation (9), which describes the Chern polynomial of the quotient bundle  $\mathcal{Q}_i$  on  $\mathbb{P}(\mathbb{C}^{m_i})$ , we find

$$c(t, \pi_i^* \mathcal{Q}_i) = \prod_{k \in [m_i-1]} (1 - \xi_i^k \cdot h_i \cdot t).$$

The product formula for Chern characters (Exercise 6.15) yields that

$$ch(\otimes_{j \in [d] \setminus \{i\}} \pi_j^* \mathcal{H}_j) = e^{\widehat{h}_i},$$

and, by definition, we derive

$$ch(R_i) = \sum_{k \in [m_i-1]} e^{\widehat{h}_i - \xi_i^k h_i}.$$

Hence

$$c(t, R_i) = \prod_{k \in [m_i-1]} (1 + (\widehat{h}_i - \xi_i^k h_i)t).$$

From this, we easily deduce that

$$(11) \quad c(1, R_i) = \sum_{j=0}^{m_i-1} (1 + \widehat{h}_i)^{m_i-1-j} h_i^j,$$

and

$$(12) \quad c(1, R(\mathbf{m})) = \prod_{i \in [d]} \left( \sum_{j=0}^{m_i-1} (1 + \widehat{h}_i)^{m_i-1-j} h_i^j \right).$$

The degree  $(m_i - 1)$  part of (11) is

$$\sum_{j=0}^{m_i-1} \widehat{h}_i^{m_i-1-j} h_i^j = \frac{\widehat{h}_i^{m_i} - h_i^{m_i}}{\widehat{h}_i - h_i},$$

where on left-hand side we are ignoring that  $h_i^{m_i} = 0$  in the Chow ring. Hence, in (12), the degree  $\sum_{i=1}^d (m_i - 1)$  part is the product

$$\prod_{i \in [d]} \frac{\widehat{h}_i^{m_i} - h_i^{m_i}}{\widehat{h}_i - h_i}.$$

In the Chow ring  $A^\bullet(\text{Seg}(\mathbf{m}))$ , this homogeneous polynomial reduces to the monomial  $\prod_{i=1}^d h_i^{m_i-1}$  along with its integral coefficient. This shows the statement.  $\square$

We now come to question (b) concerning the role of  $R(\mathbf{m})$ . First, we interpret tensors as sections of  $R(\mathbf{m})$ ; this is one of the most beautiful ideas in [9]. For a tensor  $T \in \mathbb{C}^{\mathbf{m}}$ , we define a section of  $R_i$ . Define  $L_i : \mathbb{C}^{\mathbf{m}} \rightarrow H^0(R_i)$  with  $L_i(T)(([v_1], \dots, [v_d])) \in R_{i,([v_1], \dots, [v_d])}$  as

$$L_i(T)(([v_1], \dots, [v_d]))(\otimes_{j \in [d] \setminus \{i\}} v_j) = [T \times (\otimes_{j \in [d] \setminus \{i\}} v_j)],$$

where the left-hand side denotes the class of the resulting vector in the quotient vector space  $\mathbb{C}^{m_i}/\langle v_i \rangle$ . This is indeed a global section of  $R_i$ .

**Proposition 5.4.** *Let  $R_i$  and  $R(\mathbf{m})$  be the vector bundles on  $\text{Seg}(\mathbf{m})$  defined above. Then:*

(a) For each  $i \in [d]$ , the map

$$L_i : \mathbb{C}^{\mathbf{m}} \longrightarrow H^0(R_i)$$

defined above is injective and its image generates  $R_i$ .

(b)  $L = (L_1, \dots, L_d)$  is an injective map  $L : \bigoplus_{i=1}^d \mathbb{C}^{\mathbf{m}} \longrightarrow H^0(R(\mathbf{m}))$ , whose image generates  $R(\mathbf{m})$ .

*Proof.* (a). Assume  $T \neq 0$ . Then there exist  $w_j \in \mathbb{C}^{m_j}$  such that  $T \times (\otimes_{j \in [d]} w_j) \neq 0$ . Hence  $u_i = T \times (\otimes_{j \in [d] \setminus \{i\}} w_j) \in \mathbb{C}^{m_i} \setminus \{0\}$ . Let  $v_j = w_j$  for  $j \neq i$  and choose  $v_i \in \mathbb{C}^{m_i} \setminus \langle u_i \rangle$ . Then  $L_i(T)(([v_1], \dots, [v_d])) \neq 0$  and so  $L_i$  is injective.

To show that  $L_i(\mathbb{C}^{m_i})$  generates  $R_i$ : let  $y_i \in \mathbb{C}^{m_i}$  and choose  $g_j \in \mathbb{C}^{m_j}$  such that  $g_j^t v_j = 1$  for  $j \in [d]$ . Define  $T = (\otimes_{j \leq i-1} g_j) \otimes y_i \otimes (\otimes_{j \geq i+1} g_j)$ . Then

$$L_i(T)(([v_1], \dots, [v_d])) = [y_i].$$

Since  $y_i$  was arbitrary, this shows that the image of  $L_i$  generates  $R_i$ . The proof for statement (b) is analogous.  $\square$

**Remark 5.5.** Using Künneth formula for sheaf cohomology, one finds that  $L_i(\mathbb{C}^{\mathbf{m}}) = H^0(R_i)$  and  $L(\bigoplus_{i=1}^d \mathbb{C}^{\mathbf{m}}) = H^0(R(\mathbf{m}))$ .

**Proposition 5.6 (Singular vector tuples as zeros).** Let  $T \in \mathbb{C}^{\mathbf{m}}$  and consider the section  $\widehat{T} = L \circ \delta(T) \in H^0(R(\mathbf{m}))$ , where

$$\begin{aligned} \delta : \mathbb{C}^{\mathbf{m}} &\longrightarrow \bigoplus_{i=1}^d \mathbb{C}^{\mathbf{m}} \\ T &\mapsto (T, T, \dots, T), \end{aligned}$$

and  $L$  is the map defined in Proposition 5.4. Then  $([v_1], \dots, [v_d]) \in \text{Seg}(\mathbf{m})$  is a zero of  $\widehat{T}$  if and only if  $([v_1], \dots, [v_d])$  is a singular vector tuple of  $T$ .

*Proof.* By definition,  $\widehat{T}(([v_1], \dots, [v_d])) = 0$  if and only if  $L_i(T)(([v_1], \dots, [v_d]))$  is a zero vector in the fiber  $R_{i,([v_1], \dots, [v_d])}$  for each  $i \in [d]$ . This is equivalent to the conditions

$$T \times (\otimes_{j \in [d] \setminus \{i\}} v_j) = \lambda_i v_i, \quad \lambda_i \in \mathbb{C}, \quad \forall i \in [d],$$

which is the definition of a singular vector tuple of  $T$ .  $\square$

**Remark 5.7.** By Proposition 5.6, singular vector tuples are zeros of elements in a subspace of global sections of  $R(\mathbf{m})$ , i.e.  $V = L \circ \delta(\mathbb{C}^{\mathbf{m}})$ . The issue is that  $V$  does not generate  $R(\mathbf{m})$  at all points of  $X$  but only at points outside products of the form  $(\prod_{i \in I} Q_i) \times (\prod_{j \notin I} \mathbb{P}(\mathbb{C}^{m_j}))$  for  $|I| \geq 2$  inside  $\text{Seg}(\mathbf{m})$ , i.e. does not generate at points where at least two entries are isotropic vectors. Friendland-Ottaviani overcome this problem by proving that a general  $\sigma \in V$  has zero locus consisting of  $c(\mathbf{m})$  (the top Chern number of  $R(\mathbf{m})$ ) simple points that have nonzero singular values and are non-isotropic. To prove this, they provide a similar statement to Lemma 4.6 for *almost generated vector bundles*. We refer to their paper [9] for the fine details.

## 6. EXERCISES

**6.1. Singular vector tuples and critical points.** The exercises below are largely inspired by the foundational references [1, 4, 9, 13, 14, 16, 17].

**Exercise 6.1.** With the help of a computer algebra software, use defining equations – rank constraints – to calculate the number of singular vector triples of a general (random) 3-way tensor  $T \in \mathbb{C}^{\mathbf{m}}$  when  $\mathbf{m} = (2, 2, 2)$ ,  $\mathbf{m} = (2, 2, 3)$ ,  $\mathbf{m} = (2, 2, 4)$ ,  $\mathbf{m} = (3, 3, 3)$ . How these numbers compare?

**A little Summer game.** What is the missing number in the following sequence?

$$10 \quad 11 \quad 12 \quad 13 \quad 20 \quad ? \quad 1000$$

The number of singular vector triples of a general 3-way tensor for  $\mathbf{m} = (2, 2, 3)$  is a hint.

**Exercise 6.2.** Using Friedland-Ottaviani's formula, calculate the number of singular vector tuples of a general  $d$ -way tensor  $T \in \mathbb{C}^{\mathbf{m}}$  for  $\mathbf{m} = (2, \dots, 2)$ .

**Exercise 6.3.** Let  $\mathbf{m} = (2, 2, 3)$  and  $T = [t_{i,j,k}] \in \mathbb{C}^{\mathbf{m}}$  be such that  $t_{1,1,1} = 1, t_{2,2,3} = 2$  and zero otherwise. Compute its singular vector triples.

**Exercise 6.4.** Let  $m \geq 2$ ,  $\mathbf{m} = (m, m, \dots, m)$ . Let  $T = [t_{i_1, \dots, i_d}] \in \mathbb{C}^{\mathbf{m}}$  be such that  $t_{i, i, \dots, i} = \alpha_i \neq 0$  for all  $1 \leq i \leq m$  and zero otherwise. Determine all the  $E$ -eigenpairs of  $T$ .

**Exercise 6.5.** Let  $d, m \geq 2$  be even and  $\mathbf{m} = (m, \dots, m)$ . Find a  $d$ -way tensor  $T \in \mathbb{C}^{\mathbf{m}}$  with no real  $E$ -eigenpairs.

**Exercise 6.6.** Let  $\mathbf{m} = (m, \dots, m)$  and  $T \in \mathbb{C}^{\mathbf{m}}$  be a  $d$ -way tensor. Show that if  $T$  has real entries, it has finitely many  $E$ -eigenpair and either  $d$  or  $m$  is odd, then  $T$  has a real eigenpair.

**Hint:** Use [10, Corollary 13.2] and the observation that  $\mathbb{R}$  is a 2-field in the sense explained just before the referred statement in *loc. cit.*

**Exercise 6.7.** Let  $\mathbb{S}^{m-1} \subset \mathbb{R}^m$  be the unit sphere and let  $f \in S^d \mathbb{R}^m$  be a symmetric tensor.

- What are the  $E$ -eigenvectors of  $f$  whose corresponding eigenvalue is zero?
- Show that a vector  $v \in \mathbb{S}^{m-1}$  is an  $E$ -eigenvector of  $f$  if and only if  $v$  is a critical point of the function  $f$  naturally defined on  $\mathbb{S}^{m-1}$ . (Use Lagrange multipliers.)
- What is the eigenvalue corresponding to an  $E$ -eigenvector  $v \in \mathbb{S}^{m-1}$  of  $f$ ?

**Exercise 6.8.** Let  $f(x) = x_1 x_2 x_3 (x_1 + x_2 + x_3)$ . Compute its  $E$ -eigenvectors and check that all of them are real. (Khozasov [12] proved that for any  $d$  and  $m$  there exist real homogeneous polynomials with the property of having  $\frac{(d-1)^m - 1}{d-2}$  real  $E$ -eigenvectors. His construction is based on harmonic polynomials.)

**Exercise 6.9.** Find a nonzero complex homogeneous polynomial  $f$  such that every normalized vector  $v$ , i.e.  $v^t v = 1$ , is an  $E$ -eigenvector of  $f$ .

**Exercise 6.10.** Let  $\lambda \neq 0$  and  $d \geq 2$ , and let  $f(x)$  be a complex homogenous polynomial in  $m$  variables of degree  $d$ . Then  $v \in \mathbb{C}^m$  is a normalized  $E$ -eigenvector of  $f$  (i.e.  $v^t v = 1$ ) with eigenvalue  $\lambda$  if and only if  $v$  is a singular point of the affine hypersurface defined by

$$f(x) - \frac{\lambda}{2} x^t \cdot x - \left( \frac{1}{d} - \frac{1}{2} \right) \lambda = 0.$$

**Exercise 6.11.** Choose your favorite positive semidefinite real homogeneous polynomial  $f$  (i.e.  $f(x) \geq 0$  for all  $x \in \mathbb{R}^m$ ) of degree  $d \geq 4$  and in  $m \geq 3$  variables. Calculate its  $E$ -eigenpairs and verify Qi's result (Theorem 2.4 or [17, Theorem 5(b)]).

**Exercise 6.12.** Let  $p_1, \dots, p_d$  be  $d$  distinct points on  $\mathbb{P}^1$ . Is there a  $d$ -way tensor  $T \in \mathbb{C}^{(2, \dots, 2)}$  such that its  $E$ -eigenvectors are the  $p_i$ 's? Is there a symmetric tensor  $f \in S^d \mathbb{C}^2$  such that its  $E$ -eigenvectors are the  $p_i$ 's? (See [1, Theorem 2.7].)

**6.2. Chern classes.** Recall that by a *variety* we mean an irreducible and reduced separated scheme of finite type over an algebraically closed field of characteristic zero (e.g. a projective variety over that field).

The *splitting principle* for vector bundles is a very useful tool to compute Chern classes. It is the content of the following result: it roughly says that every relation among Chern classes can be established assuming that the bundles we are working with are split.

**Theorem.** Let  $X$  be a smooth variety and  $\mathcal{E}$  a vector bundle of rank  $r$  on  $X$ . Then there exists a smooth variety  $Y$  and a morphism  $\varphi : Y \rightarrow X$  such that the pullback  $\varphi^* \mathcal{E}$  has a filtration  $0 = \mathcal{E}_0 \subset \mathcal{E}_1 \subset \mathcal{E}_2 \subset \dots \subset \mathcal{E}_r = \varphi^* \mathcal{E}$  where  $\mathcal{L}_i = \mathcal{E}_i / \mathcal{E}_{i-1}$  is a line bundle and the pullback map  $\varphi^* : A(X) \rightarrow A(Y)$  is an injective ring homomorphism. Moreover,  $c(t, \varphi^* \mathcal{E}) = \prod_{i=1}^r c(t, \mathcal{L}_i)$ , i.e.  $\varphi^* \mathcal{E}$  has the same Chern classes of  $\mathcal{L}_1 \oplus \dots \oplus \mathcal{L}_r$ .

**Exercise 6.13.** Using the splitting principle and that  $c_1(\mathcal{L}^\vee) = -c_1(\mathcal{L})$  for any line bundle  $\mathcal{L}$  on  $X$ , check that

$$c_i(\mathcal{E}^\vee) = (-1)^i c_i(\mathcal{E}).$$

**Exercise 6.14.** Let  $X$  be a smooth variety. Let  $\mathcal{L}$  and  $\mathcal{E}$  be a line bundle and a vector bundle of rank  $r$  on  $X$ , respectively. Using the splitting principle, find a formula for the  $k$ th Chern class  $c_k(\mathcal{E} \otimes \mathcal{L})$  in terms of Chern classes of  $\mathcal{E}$  and  $\mathcal{L}$ . (When  $r = 1$ ,  $c_1(\mathcal{E} \otimes \mathcal{L}) = c_1(\mathcal{E}) + c_1(\mathcal{L})$ .)

**Exercise 6.15.** If  $\mathcal{E}$  is vector bundle of rank  $r$  on  $X$ , factor its Chern polynomial  $c(t, \mathcal{E}) = \prod_{j=1}^r (1 + \xi_j(\mathcal{E})t)$ , where  $\xi_j(\mathcal{E})$  are the Chern roots. The Chern character of  $\mathcal{E}$  is  $ch(\mathcal{E}) = \sum_{j=1}^r e^{\xi_j(\mathcal{E})}$ .

Keep the assumptions from **Exercise 6.14**. Let  $ch(\mathcal{E})$  and  $ch(\mathcal{L})$  be the Chern characters of  $\mathcal{E}$  and  $\mathcal{L}$ . Using the splitting principle, prove that

$$ch(\mathcal{E} \otimes \mathcal{L}) = ch(\mathcal{E}) \cdot ch(\mathcal{L}).$$

**Exercise 6.16.** Calculate the Chern polynomial of  $\Omega_{\mathbb{P}^n}^1(2)$  where  $\Omega_{\mathbb{P}^n}^1$  is the cotangent bundle on  $\mathbb{P}^n$ . Show that  $c_n(\Omega_{\mathbb{P}^n}^1(2)) = 0$  if and only if  $n$  is odd.

**Exercise 6.17.** Let  $X$  be a smooth variety of dimension  $n$ . Suppose the tangent bundle  $\mathcal{T}_X$  splits as  $\mathcal{T}_X = \mathcal{L}_1 \oplus \cdots \oplus \mathcal{L}_n$ , where the  $\mathcal{L}_i$  are line bundles with  $a_i = c_1(\mathcal{L}_i)$ . Show that the canonical bundle of  $X$  is isomorphic to  $\mathcal{O}_X(-a_1 - \cdots - a_n)$ . Check that  $c_1(\mathcal{T}_X)$  is the class of the anticanonical divisor  $-K_X$  indeed. (This last statement is always true.)

We continue with an important theorem.

**Theorem.** Let  $X$  be a smooth complex variety of dimension  $n$ . Then the degree of  $c_n(\mathcal{T}_X)$ , i.e. the top Chern number of the tangent bundle  $\mathcal{T}_X$  (usually denoted  $\int_X c_n(\mathcal{T}_X)$ ) equals its topological Euler characteristic  $\chi_{\text{top}}(X)$ .

**Exercise 6.18.** Let  $X \subset \mathbb{P}^{n+1}$  be a smooth complex hypersurface of degree  $d$ . Use the tangent exact sequence

$$0 \longrightarrow \mathcal{T}_X \longrightarrow \mathcal{T}_{\mathbb{P}^{n+1}|X} \longrightarrow \mathcal{O}_X(d) \longrightarrow 0$$

to find the topological Euler characteristic of  $X$ . What does the formula read for a degree  $d$  smooth curve in  $\mathbb{P}^2$ ?

**Exercise 6.19.** Let  $\mathcal{E}$  be a vector bundle of rank  $r$  on a variety  $X$ . Let  $\sigma_0, \dots, \sigma_{r-i}$  be general elements of a vector space  $W \subset H^0(\mathcal{E})$  of global sections generating  $\mathcal{E}$ . Prove that, when non-empty, the degeneracy locus  $D = V(\sigma_0 \wedge \cdots \wedge \sigma_{r-i})$  is generically reduced of codimension  $i$  in  $X$ .

**Hints:**

- (a) Let  $\dim W = m$ . Define a morphism  $\varphi$  from  $X$  to the Grassmannian of  $(m - r)$ -dimensional linear subspaces of  $W$ .
- (b) Let  $U$  be a given subspace of dimension  $r - i + 1$  generated by global sections  $\sigma_0, \dots, \sigma_{r-i}$ . The idea is to define, for a general such  $U$ , a locus  $\Sigma \subset \mathbb{G}(m - r, W)$  of codimension  $i$  such that  $\varphi^{-1}(\Sigma) = D$ . Kleiman's transversality theorem in characteristic zero [7, Theorem 1.7] applied to the Grassmannian  $\mathbb{G}(m - r, W)$  shows that  $D$  has pure codimension  $\text{codim}(D) = \text{codim}(\Sigma) = i$  and it is generically reduced. See also the proof of [7, Lemma 5.2].

**Exercise 6.20.** Let  $\mathcal{E}$  be a globally generated vector bundle of rank  $r$  on a variety  $X$ . Then its top Chern class  $c_r(\mathcal{E}) = 0$  if and only if the zero locus of a general global section of  $\mathcal{E}$  is empty.

The statement is false when one drops the assumption on global generation: let  $X = \mathbb{P}^1 \times \mathbb{P}^1$  and  $\mathcal{E} = \mathcal{O}_X(-a, 0) \oplus \mathcal{O}_X(a, 0)$  for some  $a > 0$ . Show that  $\mathcal{E}$  has vanishing top Chern class but its general global section has nonempty zero locus.

## REFERENCES

- [1] H. Abo, A. Seigal, and B. Sturmfels, *Eigenconfigurations of tensors*. Algebraic and geometric methods in discrete mathematics, 1–25, Contemp. Math., 685, Amer. Math. Soc., Providence, RI, 2017.
- [2] V. Beorchia, F. Galuppi, and L. Venturello. *Eigenschemes of ternary tensors*. SIAM J. Appl. Algebra Geom. 5 (4), 620–650.
- [3] R. Bott and L. W. Tu, *Differential forms in algebraic topology*. Graduate Texts in Mathematics, 82. Springer-Verlag, New York-Berlin, 1982. xiv+331 pp.
- [4] D. Cartwright and B. Sturmfels, *The number of eigenvalues of a tensor*, Linear Algebra Appl. 438 (2013), no. 2, 942–952.
- [5] J. Draisma, E. Horobeț, G. Ottaviani, B. Sturmfels, and R.R. Thomas, *The Euclidean distance degree of an algebraic variety*. Found. Comput. Math. 16 (2016), no. 1, 99–149.
- [6] J. Draisma, G. Ottaviani, A. Tocino, *Best rank- $k$  approximations for tensors: generalizing Eckart-Young*, Res. Math. Sci. 5 (2) (2018) 27.
- [7] D. Eisenbud and J. Harris, *3264 and all that—a second course in algebraic geometry*. Cambridge University Press, Cambridge, 2016. xiv+616 pp.
- [8] J.E. Fornæss and N. Sibony, *Complex dynamics in higher dimension, I*, in: Complex Analytic Methods in Dynamical Systems, Rio de Janeiro, 1992, in: Astérisque, vol. 222, 1994, pp. 201–231.
- [9] S. Friedland and G. Ottaviani, *The number of singular vector tuples and uniqueness of best rank-one approximation of tensors*. Found. Comput. Math. 14 (2014), no. 6, 1209–1242.
- [10] W. Fulton, *Intersection theory*. Second edition, Springer-Verlag, Berlin, 1998. xiv+470 pp.
- [11] P. Griffiths and J. Harris, *Principles of algebraic geometry*. Pure and Applied Mathematics, Wiley-Interscience, New York, 1978. xii+813 pp.
- [12] K. Khozasov, *On fully real eigenconfigurations of tensors*, SIAM J. Appl. Algebra Geom. 2 (2018), no. 2, 339–347.
- [13] L.-H. Lim. *Singular values and eigenvalues of tensors: a variational approach*. Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP '05), vol. 1 (2005), 129–132.
- [14] L. Oeding and G. Ottaviani, *Eigenvectors of tensors and algorithms for Waring decomposition*. J. Symbolic Comput. 54 (2013), 9–35.
- [15] G. Ottaviani and R. Paoletti, *A geometric perspective on the singular value decomposition*, Rend. Ist. Mat. Univ. Trieste 47 (2015) 107–125.
- [16] L. Qi and Z. Luo, *Tensor analysis. Spectral theory and special tensors*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2017. xiii+305 pp.
- [17] L. Qi, *Eigenvalues of a real supersymmetric tensor*, J. Symbolic Comput. 40 (2005), no. 6, 1302–1324.

POLITECNICO DI TORINO, DIPARTIMENTO DI SCIENZE MATEMATICHE “G. L. LAGRANGE”, CORSO  
 DUCA DEGLI ABRUZZI 24, 10129 TORINO, ITALY  
 Email address: emanuele.ventura@polito.it