UNIVERSITÀ DI PISA

DIPARTIMENTO DI MATEMATICA

Corso di Laurea in Matematica

Tesi di Laurea Magistrale

Kronecker Decomposition of Pencils of Quadrics and Nonabelian Apolarity

with implementations in Macaulay2

Candidato:
Vincenzo Galgano

Relatore:
Prof. **Giorgio Ottaviani**

Controrelatore:
Prof. **Marco Franciosi**

Anno Accademico 2019/2020

Contents

List of Tables

List of Figures

Introduction

"Perchè ha due nomi (il teorema di) Rouchè-Capelli? Perchè quando l'hanno fatto l'algebra lineare non esisteva e la stavano inventando. L'algebra lineare di dimensione finita è l'esempio più riuscito di una teoria in cui tutte le difficoltà sono state sbriciolate, digerite nell'apparato definitorio. Cioè le definizioni sono date in modo così efficiente che tutto il resto è quasi un banale esempio. Cioè ogni cosa segue in modo diretto, banale, dall'apparato definitorio." - Prof. Riccardo Benedetti, 2012

The Kronecker-Weierstrass decomposition for matrix pencils is the equivalent of the Jordan form for matrices (actually, it is an extension of it since the latter is obtained by studying a matrix pencil of type $A + \lambda I$). Just as Rouchè-Capelli theorem, such decomposition is two-named since it was completely determined by Kronecker (1890, [23]) and Weierstrass (1868, [36]) separately: Weierstrass determined the decomposition of regular pencils with respect to their elementary divisors since he was interested in finding conditions for simultaneous congruence, and later Kronecker extended the classification to singular pencils by studying their minimal indices. The similarity with the Jordan form is in the nature of the invariants, but from a group action point of view the Kronecker-Weierstrass form extends the "left-right equivalence" for matrices, since its classes are given by left-right-multiplication and not by conjugacy (as for the Jordan form).

The Kronecker-Wierstrass decomposition has several applications in different areas of mathematics (and even other sciences), maybe the most known of which are the generalized eigenvalues problem in numerical analysis and the study of linear differential equations in analysis. However, by birthright, it finds applications in algebra and algebraic geometry too: indeed matrix pencils of size $m \times n$ have a natural structure of tensors of type $2 \times m \times n$, and this leads to derive several properties of given algebraic and geometric objects from the invariants of the corresponding matrix pencils, with the advantage of working with (multi)linear algebra tools.

In this work we investigate some of the above applications. In particular, we study how the Kronecker-Weierstrass decomposition (and invariants) applies to the Segre classification of intersections of quadrics and to tensor rank decomposition.

The classification of intersections of two quadrics is a classical result due to Corrado Segre (1883-1884, [34], [33]) and it may be reformulated in terms of matrix pencils since the pencils of quadrics may be represented by symmetric matrix pencils. In this perspective the Kronecker-Weierstrass decomposition helps to completely classify such intersections in terms of the algebraic Kronecker invariants. However this algebraic classification is strictly related to the geometric classification (due to Dimca in 1983, [10]) of some objects associated to the pencils of quadrics. Every such pencil defines a projective line in the space of all quadrics and a projective variety given by its base locus: the *position* of the line describes the regular part of the pencil, while the singularities of the base locus describe its singular part.

The *tensor rank decomposition* is a as-classical-as-modern topic in algebraic geometry with several applications to applied sciences. The Kronecker-Weierstrass decomposition plays a central role in the case of 2-*slice tensors*, that are tensors in $\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$: this comes from the correspondence between these tensors and the matrix pencils. This identification allows to explicitly determine the different ranks of such tensors and to classify their orbits with respect to the action by $\mathrm{GL}_2 \times \mathrm{GL}_m \times \mathrm{GL}_n$. A particular topic in tensor rank decomposition is the *Waring decomposition problem*, that is the problem of finding the minimal decomposition of a homogeneous polynomial as sum of decomposable symmetric tensors: the problem is trivial for tensors in $\mathbb{K}^m \otimes \mathbb{K}^n$ (since these tensors may be represented as matrices and the rank of such tensors is just the rank of the corresponding matrices) but it becomes more complicated for tensors of higher order. The main classical tool in tensor rank decomposition is given by secant varieties. However the symmetric tensors have the advantage of being seen as polynomials and this leads to tools which simplify the study of the symmetric decomposition: the most important of them is the apolarity lemma. Such result does not hold anymore for general tensors but it is generalized by the nonabelian apolarity due to Oeding and Ottaviani (2013, [30]), which translates the problem of decomposition in terms of eigenvectors of tensors.

We now give a detailed insight of our thesis work.

Chapter 1 is just preliminary: first we introduce λ-matrices and their invariants as a starting point to chapter 2, then we recall some basic properties about tensor products and symmetric tensors which are the main characters of chapter 4.

Chapter 2 is about the decomposition of matrix pencils (i.e. degree-1 homogeneous binomials with matrix coefficients) over a generic field $\mathbb{K}$, denoted by $\mathfrak{M}_{m\times n}(\mathbb{K}[\mu,\lambda]_1)$. In the first section we introduce the notion of strict equivalence which defines the orbits with respect to the group action

$$\begin{array}{ccc} \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K}) & \longrightarrow & \mathrm{Aut}\left(\mathfrak{M}_{m\times n}(\mathbb{K}[\mu,\lambda]_1)\right) \\ (P,Q) & \mapsto & \left(A(\mu,\lambda) \mapsto P\cdot A(\mu,\lambda)\cdot {}^tQ \right) \end{array}$$

Its invariants lead to the Kronecker-Weierstrass canonical form: we first classify the so-called *regular* pencils via the elementary divisors, then the *singular* ones via the minimal indices for columns and rows. This determines the representatives of the classes in the (double) quotient

$$\mathrm{GL}_m(\mathbb{K}) \diagdown \mathfrak{M}_{m\times n}(\mathbb{K}[\mu,\lambda]_1) \diagup \mathrm{GL}_n(\mathbb{K})$$

In the last section we focus on the symmetric pencils by exhibiting a symmetric Kronecker canonical form both over $\mathbb{C}$ and $\mathbb{R}$.

In chapter 3 we apply the Kronecker decomposition of complex symmetric pencils to intersections of complex projective quadrics. In the first section we determine the conditions for the simultaneous reduction of complex quadratic forms in terms of Kronecker invariants. In the second section we formulate the Segre classification of intersections of projective quadrics in terms of pencils of quadrics via the so-called *Segre symbol*: we explicitly classify all pencils of quadrics in $\mathbb{P}^2_{\mathbb{C}}$ and the regular pencils of quadrics in $\mathbb{P}^3_{\mathbb{C}}$, both schematically and geometrically (with the useful computation of the primary decompositions on `Macaulay2`). In the last section we give a geometric interpretation of the Kronecker form of pencils of quadrics in terms of projective lines and singularity of the base loci via projective bundles: for a concrete approach we study the two geometric objects relatively to the pencils of quadrics in $\mathbb{P}^2_{\mathbb{C}}$ and we list them in tables.

Chapter 4 offers a new perspective for matrix pencils and pencils of quadrics in terms of tensors. We start by briefly introducing different notions of rank (multilinear, border, symmetric) for generic tensors and the algebraic-geometric objects related to them (Segre varieties, secant varieties, Veronese varieties). In the second section we focus on tensors in $\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$, called 2-*slice tensors*: these are not a novelty since they are in correspondence with matrix pencils

$$\mathfrak{M}_{m\times n}(\mathbb{K}[\mu,\lambda]_1) \longleftrightarrow \mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$$

We introduce the GL-equivalence (which extends the strict equivalence for matrix pencils) defined by the group action

$$\begin{array}{ccc} \mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K}) & \longrightarrow & \mathrm{Aut}\left(\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n\right) \\ (M,P,Q) & \mapsto & \left(u\otimes v\otimes w \mapsto Mu\otimes Pv\otimes Qw \right) \end{array}$$

and we find out a canonical form: in small dimensions there are finitely many orbits and we compute their dimensions by studying the dimension of the Lie algebras of the stabilizers. Moreover, the Kronecker form allows to determine their ranks both by a direct combinatorial approach and by applying the *discrete Fourier transform*: these are the contents of the central sections. We also focus on the cases $m = n = 2, 3$: in these cases we list dimension, rank and border rank of all orbits, supported by an implementation on `Macaulay2`. We conclude this chapter by determining the *partially-symmetric rank* and the dimension of the orbits of tensors in $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$ with respect to the group action

$$\begin{array}{ccc} \mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_{m+1}(\mathbb{K}) & \longrightarrow & \mathrm{Aut}\left(\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})\right) \\ (M, P) & \mapsto & \left(u \otimes l^2 \mapsto Mu \otimes P \cdot l^2 \cdot {}^t P\right) \end{array}$$

which correspond to pencils of quadrics in $\mathbb{P}^m_{\mathbb{K}}$: again we implement such computation on `Macaulay2` and we list the results in tables.

Chapter 5 is about the abelian and nonabelian apolarity theory. In the first section we show the classical catalecticant method for solving the Waring problem: the minimal decomposition of a symmetric tensor $f \in \mathrm{Sym}^d V$ is to be found in the zeros of its *apolar* ideal $f^\perp$; the kernels of the *catalecticant maps*

$$C_{k,f} : \mathrm{Sym}^k V^\vee \to \mathrm{Sym}^{d-k} V$$

help to restrict the research loci. Since this method fails in many cases, in the second section we introduce a generalization in terms of vector bundles: this language not only allows to solve the gap of the previous method but it also suggests new methods for the decomposition of general tensors, even not symmetric ones. The first two sections are just introductory and motivational to the last one, where we recover the nonabelian apolarity for general tensors in $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$ by deriving it from the Kronecker decomposition of matrix pencils: in particular, the decomposition of the tensor $(B_1, B_2) \in \mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$ is given by the common eigenvectors of the kernel of the map

$$\begin{array}{cccc} C_{(B_1,B_2)} : & \mathfrak{sl}_{m+1}(\mathbb{C}) & \longrightarrow & \bigwedge^2 V \oplus \bigwedge^2 V \\ & A & \mapsto & \left(AB_1 - B_1({}^t A)\ ,\ AB_2 - B_2({}^t A)\right) \end{array}$$

We also exhibit an implementation of it on `Macaulay2` (see Ch.6, §4) which allows to get hand dirty in small dimension.

In chapter 6 we collect all *our* implementations on `Macaulay2`: they were made for two reasons, as support to lighten calculations and as cast out nines for the theorical statements. The implementations in the first section (about the Kronecker invariants) and in the last one (about the common eigenvectors for nonabelian apolarity) had the purpose of producing concrete examples which helped to get familiarity with the theory. Instead, the implemetations in the two sections in the middle (about the orbit dimensions) were specifically made to speed up a large number of calculations (and to avoid typical hand calculation errors as well!).

Chapter 1

Preliminaries

Let $\mathbb{K}$ be a field with characteristic 0, not necessarily algebraically closed.
We will work with finite dimensional vector spaces over $\mathbb{K}$.

1.1 λ-matrices and λ-equivalence

Let $\mathbb{K}[\lambda]$ be the ring of the polynomials in the variable λ with coefficients in $\mathbb{K}$.

Definition. A **λ-matrix**, or **polynomial matrix**, is a matrix $A(\lambda)$ whose entries are polynomials in $\mathbb{K}[\lambda]$.

We may represent a λ-matrix $A(\lambda)$ as a matrix polynomial[1] $A_0 + A_1\lambda + \cdots + A_k\lambda^k$, where k is the highest of the degrees of the elements of $A(\lambda)$. Thus we may think at the set of λ-matrices of size $m \times n$ as the matrix algebra $\mathfrak{M}_{m\times n}(\mathbb{K}[\lambda])$ or the polynomial algebra $(\mathfrak{M}_{m\times n}(\mathbb{K}))[\lambda]$. Consider the subalgebra

$$\mathrm{GL}_m(\mathbb{K}[\lambda]) = \{P(\lambda) \in \mathfrak{M}_m(\mathbb{K}[\lambda]) \mid \det P(\lambda) \in \mathbb{K}^\times\}$$

of invertible λ-matrices of size m and consider the group action

$$\begin{array}{ccc} \mathrm{GL}_m(\mathbb{K}[\lambda]) \times \mathrm{GL}_n(\mathbb{K}[\lambda]) & \longrightarrow & \mathrm{Aut}\left(\mathfrak{M}_{m\times n}(\mathbb{K}[\lambda])\right) \\ (P(\lambda), Q(\lambda)) & \mapsto & \left(A(\lambda) \mapsto P(\lambda) \cdot A(\lambda) \cdot {}^t Q(\lambda) \right) \end{array}$$

Definition. Two λ-matrices $A(\lambda)$ and $B(\lambda)$ of size $m \times n$ are said to be **λ-equivalent** if they are in the same orbit with respect to the above group action.

[1] a polynomial with matrix coefficients

Since $\mathbb{K}[\lambda]$ is a PID, from commutative algebra we know that every λ-matrix $A(\lambda)$ of size $m \times n$ and rank r is equivalent to a canonical diagonal λ-matrix of the form

$$\begin{bmatrix} d_1(\lambda) & & & & 0 & \dots & 0 \\ & d_2(\lambda) & & & \vdots & & \vdots \\ & & \ddots & & & & \\ & & & d_r(\lambda) & 0 & & \\ & & & & \vdots & & \vdots \\ & & & & 0 & \dots & 0 \end{bmatrix} \tag{1.1}$$

where $d_1(\lambda), \dots, d_r(\lambda) \in \mathbb{K}[\lambda]\backslash\{0\}$ are monics with the divisibility property $d_i(\lambda) \mid d_{i+1}(\lambda)$ and they are uniquely determined by $A(\lambda)$. Such form is said to be the **Smith canonical form** of the matrix.

Invariant polynomials. We want to find a complete system of invariants for λ-equivalence of λ-matrices. We recall the notion of *minor* of order h of a given matrix A: it is the determinant of a submatrix of A of size $h \times h$ of the form

$$\begin{bmatrix} a_{i_1 j_1} & \dots & a_{i_1 j_h} \\ \vdots & \ddots & \vdots \\ a_{i_h j_1} & \dots & a_{i_h j_h} \end{bmatrix}$$

If A is a matrix of rank r, then every minor of order $h > r$ of A is zero.

Let $g_j(\lambda)$ be the greatest common divisor of all minors of order j in $A(\lambda)$. By a simple count it turns out that each one of these polynomials is divided by the preceding one, i.e. $1 = g_0(\lambda) \mid g_1(\lambda) \mid \dots \mid g_r(\lambda)$.

Definition. The **invariant polynomials** of $A(\lambda)$ are the quotients

$$i_j(\lambda) = \frac{g_{r-j+1}(\lambda)}{g_{r-j}(\lambda)} \text{ for } \; j = 1 : r$$

These polynomials are said *invariants* because they are such under equivalent transformations of the λ-matrix. We may observe that, given $D(\lambda)$ the diagonal matrix in (1.1), its invariant polynomials are

$$i_1(\lambda) = d_r(\lambda) \; , \; \dots \; , \; i_r(\lambda) = d_1(\lambda)$$

Moreover by equivalence we deduce that these are exactly the invariant polynomials of every λ-matrix equivalent to such $D(\lambda)$. Thus the following holds.

Fact 1.1.1. The invariant polynomials form a complete system of invariants for λ-equivalence of λ-matrices, i.e. two λ-matrices of the same size are λ-equivalent if and only if they have the same invariant polynomials $i_j(\lambda)$.

By the divisibility property of the polynomials defining the Smith canonical form, it follows that in the sequence of invariant polynomials $i_1(\lambda), \ldots, i_r(\lambda)$ every polynomial divides the preceding one, i.e.

$$i_r(\lambda) \mid i_{r-1}(\lambda) \mid \ldots \mid i_1(\lambda)$$

Elementary divisors. We may wonder how invariant polynomials behave in a block-diagonal matrix

$$C(\lambda) = \begin{bmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{bmatrix}$$

We may hope that the invariant polynomials of $C(\lambda)$ are just the union of the ones of $A(\lambda)$ and $B(\lambda)$ but in general this does not hold: it is enough to think that the invariant polynomials have the property of divisibility but in general the invariant polynomials of $A(\lambda)$ and $B(\lambda)$ could not have any relations. We need to add one more hypothesis to make things work.

Fact 1.1.2. Given $C(\lambda) = \operatorname{diag}(A(\lambda), B(\lambda))$ a block-diagonal λ-matrix, if every invariant polynomial of $A(\lambda)$ divides every invariant polynomial of $B(\lambda)$, then the set of invariant polynomials of $C(\lambda)$ is the union of the invariant polynomials of the two diagonal blocks.

In the general case to determine the invariant polynomials of $C(\lambda)$ we need to introduce a new concept. Since $\mathbb{K}[\lambda]$ is a UFD, we can decompose each invariant polynomial into powers of irreducible factors over $\mathbb{K}$. Moreover the divisibility property allow us to consider the same irreducible factors for all the invariant polynomials and to let just their powers vary:

$$\begin{gathered} i_1(\lambda) = \varphi_1(\lambda)^{e_1^{(1)}} \cdot \ldots \cdot \varphi_s(\lambda)^{e_s^{(1)}} \\ \vdots \\ i_r(\lambda) = \varphi_1(\lambda)^{e_1^{(r)}} \cdot \ldots \cdot \varphi_s(\lambda)^{e_s^{(r)}} \end{gathered}$$

where $e_j^{(1)} \geq e_j^{(2)} \geq \ldots \geq e_j^{(r)} \geq 0$ for all $j = 1 : r$ and each irreducible factor is monic since each invariant polynomial is so. We observe that in the above notation there are powers which are just 1 because we have choosen to represent the invariant polynomials with the same set of irreducible factors.

Definition. The non-trivial powers $\varphi_j(\lambda)^{e_j^{(h)}}$ for all $j = 1 : s, h = 1 : r$ are called **elementary divisors** of the λ-matrix $A(\lambda)$ over the field $\mathbb{K}$.

Note: We consider 1 as elementary divisor only when there exists an invariant polynomial $i_{j_0}(\lambda) = 1$. Obviously such elementary divisors would change if we consider the λ-matrix over another field since prime factorizations could be different.

In the case of block-diagonal λ-matrix, the elementary divisors behave better than the invariant polynomials.

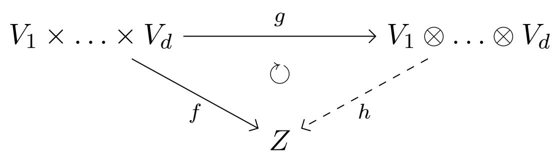
Fact 1.1.3. Given $C(\lambda) = \text{diag}(A(\lambda), B(\lambda))$ a block-diagonal λ-matrix, the set of elementary divisors of $C(\lambda)$ is always obtained by the union of the elementary divisors of $A(\lambda)$ with those of $B(\lambda)$.

In the second chapter we will focus on λ-matrices of degree 1, i.e. linear polynomials with coefficient in $\mathfrak{M}_{m\times n}(\mathbb{K})$ of the form $A + \lambda B$, and we will introduce a new equivalence, called *strict equivalence*, which is finer than the λ-equivalence and whose complete system of invariants includes the elementary divisors.

1.2 Basics on tensors

Theorem 1.2.1 (Definition)**.** Let $V_1, \ldots, V_d$ be $\mathbb{K}$-vector spaces (not necessarily finite dimensional). There exist (unique up to isomorphism) a $\mathbb{K}$-vector space denoted by $V_1 \otimes \ldots \otimes V_d$ and a multilinear map $g : V_1 \times \ldots \times V_d \to V_1 \otimes \ldots \otimes V_d$ such that the following universal property holds:

> *$\forall Z$ $\mathbb{K}$-vector space, $\forall f : V_1 \times \ldots \times V_d \to Z$ multilinear map there exists a unique linear map $h : V_1 \otimes \ldots \otimes V_d \to Z$ which makes the following diagram commute:*



Such vector space $V_1 \otimes \ldots \otimes V_d$ is called **tensor product** and its elements **tensors**: in particular the tensors in $\text{Im}(g : V_1 \times \ldots \times V_d \to V_1 \otimes \ldots \otimes V_d)$ (i.e. the ones of the form $v_1 \otimes \ldots \otimes v_d$) are said **decomposable**. Moreover, given $f : V_1 \times \ldots \times V_d \to Z$ a multilinear map, the corresponding linear map $h : V_1 \otimes \ldots \otimes V_d \to Z$ is defined on the decomposable tensors by $h(v_1 \otimes \ldots \otimes v_d) = f(v_1, \ldots, v_d)$ and extended by linearity to the whole space.

Note: Historically the tensor product was introduced to *linearize* a multilinear map from a cartesian product $V_1 \times \ldots \times V_d$ to another vector space Z: indeed we note that in the above theorem-definition the map h given by the universal property is linear and its *making diagram commute* is the same that *linearizing* the multilinear map f. The construction of the tensor product makes clear why it linearizes multilinear maps ([1, Proposition 2.12]).

Even if the definition holds in infinite dimension too, from now on we will assume to work with finite dimensional vector spaces: indeed in this case the isomorphism $V \simeq V^{\vee\vee}$ between a vector space and its bidual space holds and this leads to interesting results. Given U, V, W (finite dimensional) $\mathbb{K}$-vector spaces, we recall the following properties:

- (associativity) $U \otimes V \otimes W \simeq (U \otimes V) \otimes W \simeq U \otimes (V \otimes W)$;
- $U \otimes V \simeq V \otimes U$ and $U \otimes \mathbb{K} \simeq U$;
- $\exists\,!$ isomorphism $U^\vee \otimes V^\vee \otimes W \simeq \mathrm{Bil}(U, V; W)$ such that

$$f \otimes g \otimes w \mapsto \Big((u, v) \mapsto f(u)g(v)w\Big)$$

  and in particular $V \otimes W \simeq \mathrm{Bil}(V^\vee, W^\vee; \mathbb{K})$;
- if $V = \langle v_1, \dots v_m\rangle_{\mathbb{K}}$ and $W = \langle w_1, \dots, w_n\rangle_{\mathbb{K}}$, then a basis of $V \otimes W$ is given by $(v_i \otimes w_j \mid i = 1 : m, j = 1 : n)$ (and hence $\dim(V \otimes W) = \dim V \cdot \dim W$).

In general, the map $g : V_1 \times \ldots \times V_d \to V_1 \otimes \ldots \otimes V_d$ is not surjective, i.e. there are tensors which are not decomposable.

Coordinates description. Let $V \simeq \mathbb{K}^m$ and $W \simeq \mathbb{K}^n$ with basis $(v_1, \dots, v_m)$ and $(w_1, \dots, w_n)$ respectively. Let $T \in V \otimes W$ be a tensor. Then we can write T with respect to the basis $(v_i \otimes w_j \mid i, j)$:

$$T = \sum_{i,j} T_{ij} v_i \otimes w_j$$

where $T_{ij} \in \mathbb{K}$, hence we may look at T as a matrix (T_{ij}) of size $m \times n$ with coefficients in $\mathbb{K}$. This lead to a first identification of the tensor product as

$$V \otimes W \simeq \mathfrak{M}_{m\times n}(\mathbb{K})$$

Another way to look at the tensor T is as a $\mathbb{K}$-homomorphism $\varphi_T : V^\vee \to W$ defined as follows: given $(v^1, \dots, v^m)$ the $\mathbb{K}$-basis of $V^\vee$ dual to the basis $(v_1, \dots, v_m)$, we consider

$$\varphi_T(v^k) = \sum_{i,j} T_{ij} v^k(v_i) w_j = \sum_j T_{kj} w_j$$

Then

$$\varphi_T \cdot \left[v^1 \mid \dots \mid v^m\right] = \begin{bmatrix} T_{11} & \dots & T_{m1} \\ \vdots & \ddots & \vdots \\ T_{1n} & \dots & T_{mn} \end{bmatrix} = {}^tT$$

Similarly we may look at T as ${}^t\varphi_T : W^\vee \to V$ and we have ${}^t\varphi_T \cdot \left[w^1 \mid \dots \mid w^n\right] = T$. This other interpretation leads us to a second identification of the tensor product as

$$V \otimes W \simeq \mathrm{Hom}_{\mathbb{K}}(V^\vee, W)$$

Obviously these identifications hold for tensor products of three or more spaces too.

Flattenings. Let $V \simeq \mathbb{K}^m$ and $W \simeq \mathbb{K}^n$. We recall that $V \otimes W \simeq \mathfrak{M}_{m\times n}(\mathbb{K})$ and through this identification a tensor $T \in V \otimes W$ can be seen as a $m \times n$ matrix $T = (T_{ij})$. By *vectorizing* matrices we obtain the identification $\mathfrak{M}_{m\times n}(\mathbb{K}) \simeq \mathbb{K}^{mn}$, so we can look at a tensor $T \in V \otimes W$ as a vector

$$\mathrm{vect}(T) = {}^t\left[T_{11}, T_{21}, \dots, T_{m1}, T_{12}, \dots, T_{m2}, T_{13}, \dots, T_{m(n-1)}, T_{1n}, \dots, T_{mn}\right]$$

The tensor products of two spaces are in some sense poorer than the tensor products of three or more spaces because the more spaces the more the identifications we may consider.

For example, let U, V, W be $\mathbb{K}$-vector spaces of dimension m, n, p respectively and let $(u_i), (v_j), (w_k)$ be basis respectively. Then by associativity, with respect to these basis, we may identify

$$\begin{aligned} (U \otimes V) \otimes W &\simeq \mathfrak{M}_{mn\times p}(\mathbb{K}) &&\simeq \mathrm{Hom}((U \otimes V)^\vee, W) \\ U \otimes (V \otimes W) &\simeq \mathfrak{M}_{m\times np}(\mathbb{K}) &&\simeq \mathrm{Hom}(U^\vee, (V \otimes W)) \\ (U \otimes W) \otimes V &\simeq \mathfrak{M}_{mp\times n}(\mathbb{K}) &&\simeq \mathrm{Hom}((U \otimes W)^\vee, V) \end{aligned}$$

These different identifications lead to different expressions of a tensor $T \in U \otimes V \otimes W$: for example if we consider $T \in (U \otimes V) \otimes W$ then we can write it with respect to the above basis as

$$T = \sum_k \Big(\sum_{i,j} T_{ijk} u_i \otimes v_j\Big) \otimes w_k$$

or we can look at it in $\mathrm{Hom}((U \otimes V)^\vee, W)$ as

$$T(u^s \otimes v^t) = \sum_{i,j,k} T_{ijk}\Big((u^s \otimes v^t)(u_i \otimes v_j)\Big) w_k = \sum_k T_{stk} w_k \;\in W$$

We may think at $T \in U \otimes V \otimes W$ as a 3-dimensional matrix (T_{ijk}) (a *cube*-matrix): in this perspective, by choosing a different identification of the tensor product $U \otimes V \otimes W$

we choose a different face of the cube-matrix (T_{ijk}), then we may vectorize this cube-matrix with respect to one of these faces and obtain a 2-dimensional matrix (a *flat*-matrix). The idea is just the same as vectorizing a flat-matrix: we choose a side of the matrix, for istance the first column, and we queue the other columns at the end of the first one. We just gave the (quite intuitive) geometric idea of the following definition.

Definition. Let $T \in U \otimes V \otimes W$ be a tensor with coordinate (T_{ijk}) with respect to a given basis. The **flattening** of T *with respect to* W is the block-matrix

$$T_W = \left[T_{\cdot,1,\cdot} \;\middle|\; \dots \;\middle|\; T_{\cdot,n,\cdot} \right] \in \mathfrak{M}_{p\times nm}(\mathbb{K})$$

where

$$T_{\cdot,j,\cdot} = \begin{bmatrix} T_{1j1} & \dots & T_{mj1} \\ \vdots & \ddots & \vdots \\ T_{1jp} & \dots & T_{mjp} \end{bmatrix} \in \mathfrak{M}_{p\times m}(\mathbb{K})$$

Similarly we define the flattenings of T with respect to U and V respectively as

$$T_U = \left[T_{\cdot,\cdot,1} \;\middle|\; \dots \;\middle|\; T_{\cdot,\cdot,p} \right] \in \mathfrak{M}_{m\times pn}(\mathbb{K}) \quad , \quad T_{\cdot,\cdot,k} = \begin{bmatrix} T_{11k} & \dots & T_{1nk} \\ \vdots & \ddots & \vdots \\ T_{m1k} & \dots & T_{mnk} \end{bmatrix} \in \mathfrak{M}_{m\times n}(\mathbb{K})$$

$$T_V = \left[T_{1,\cdot,\cdot} \;\middle|\; \dots \;\middle|\; T_{m,\cdot,\cdot} \right] \in \mathfrak{M}_{n\times mp}(\mathbb{K}) \quad , \quad T_{i,\cdot,\cdot} = \begin{bmatrix} T_{i11} & \dots & T_{i1p} \\ \vdots & \ddots & \vdots \\ T_{in1} & \dots & T_{inp} \end{bmatrix} \in \mathfrak{M}_{n\times p}(\mathbb{K})$$

Faithfully to the intuitive geometric interpretation of the tensor T as cube-matrix, the submatrices

$$T_{i,\cdot,\cdot} : V^\vee \to W \quad , \quad T_{\cdot,j,\cdot} : W^\vee \to U \quad , \quad T_{\cdot,\cdot,k} : U^\vee \to V$$

are said **slices** of the tensor.

Example 1.2.2. Consider $U \simeq V \simeq W \simeq \mathbb{K}^2$ with basis $(u_1, u_2), (v_1, v_2), (w_1, w_2)$ respectively. Consider the tensor

$$T = u_1 \otimes v_1 \otimes w_1 + u_2 \otimes v_2 \otimes w_1 \quad (\in U \otimes V \otimes W)$$

Then its flattenings are

$$T_U = \left[T_{\cdot,\cdot,1} \;\middle|\; T_{\cdot,\cdot,2} \right] = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$$

$$T_V = \left[T_{1,\cdot,\cdot} \;\middle|\; T_{2,\cdot,\cdot} \right] = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

$$T_W = \left[T_{\cdot,1,\cdot} \;\middle|\; T_{\cdot,2,\cdot} \right] = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

1.3 Symmetric tensors

Let V be a m-dimensional $\mathbb{K}$-vector space and let $V^{\otimes d}$ be the tensor product of d copies of V: then $V^{\otimes d}$ has dimension m^d and, given $(v_1, \dots, v_m)$ a basis of V, $(v_{i_1} \otimes \dots \otimes v_{i_d} \mid i_j \in \{1, \dots, m\})$ is a basis of $V^{\otimes d}$. Then a tensor $T \in V^{\otimes d}$ can be written as

$$T = \sum_{i_j \in \{1,\dots,m\}} T_{i_1 \dots i_d}(v_{i_1} \otimes \dots \otimes v_{i_d})$$

The symmetric group $\mathfrak{S}_d$ acts linearly on $V^{\otimes d}$: the action is defined on the decomposable tensors as follows

$$\forall \sigma \in \mathfrak{S}_d,\ \sigma \cdot (v_1 \otimes \dots \otimes v_d) = v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(d)}$$

and extended to the whole tensor space by linearity. In particular, $\mathfrak{S}_d$ permutes the decomposable tensors, hence $V^{\otimes d}$ is a $\mathfrak{S}_d$-module.

Definition. A tensor $T \in V^{\otimes d}$ is **symmetric** if

$$\forall \sigma \in \mathfrak{S}_d \ , \quad T_{i_1 \dots i_d} = T_{\sigma(i_1) \dots \sigma(i_d)}$$

Note: If T is a symmetric tensor, we may look at its coordinates $T_{i_1 \dots i_d}$ as coefficients of a homogeneous polynomial of degree d in m variables. This perpective will be very useful when we will work with projective spaces.

We denote with

$$\operatorname{Sym}^d V = \{T \in V^{\otimes d} \mid T \text{ symmetric}\}$$

the subspace of all symmetric tensors, i.e. the subspace of $\mathfrak{S}_d$-invariant tensors: it is generated by $\{v^{\otimes d} \mid v \in V\}$ since

$$\forall \sigma \in \mathfrak{S}_d,\ \sigma \cdot (v_1 \otimes \dots \otimes v_d) = v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(d)} \iff v_1 = \dots = v_d$$

A monomial basis for $\operatorname{Sym}^d V$ is $(x_{i_1} \cdot \ldots \cdot x_{i_d} \mid i_1, \dots, i_d)$ where

$$x_{i_1} \cdot \ldots \cdot x_{i_d} = \frac{1}{d!} \sum_{\sigma \in \mathfrak{S}_d} v_{\sigma(i_1)} \otimes \dots \otimes v_{\sigma(i_d)}$$

hence

$$\dim_{\mathbb{K}} \operatorname{Sym}^d V = \binom{n+d-1}{d}$$

The above basis makes explicit the isomorphism of $\operatorname{Sym}^d(V)$ with the space of the degree-d homogeneous polynomials on V (however the isomorphism is independent from the basis). Moreover, with respect to this basis the coordinates of a symmetric tensor $T \in \mathrm{Sym}^d V$ are often denoted with

$$T_{(i_1,\dots,i_d)} = \frac{1}{d!} \sum_{\sigma \in \mathfrak{S}_d} T_{\sigma(i_1)\dots\sigma(i_d)}$$

Fact 1.3.1. Decomposable symmetric tensors $\{v^{\otimes d} \mid v \in V\}$ correspond to d-powers of linear forms $(\alpha_1 x_1 + \dots + \alpha_m x_m)^d$ on $V^\vee$ (where $\dim V = m$).

Note: To be honest we are cheating a little bit since we are identifying the symmetric algebra of degree d on V (properly denoted $\mathrm{Sym}^d V$) with the symmetric tensors. But we are fully allowed to do this since in characteristic 0 they can be identified through the map

$$\begin{array}{ccc} \mathrm{Sym}^d V & \longrightarrow & \{\text{symmetric tensors}\} \\ v_{i_1} \cdot \ldots \cdot v_{i_d} & \mapsto & \frac{1}{d!} \sum_{\sigma \in \mathfrak{S}_d} v_{\sigma(i_1)} \otimes \ldots \otimes v_{\sigma(i_d)} \end{array}$$

Chapter 2

Kronecker form of matrix pencils

In this chapter we study the decomposition of matrix pencils. In the first sections we introduce the notion of strict equivalence and we find out the Kronecker canonical form: to do so we first classify the so-called regular *pencils, then the* singular *ones. In the last sections we focus on the symmetric pencils by exhibiting a symmetric Kronecker canonical form both over $\mathbb{C}$ and $\mathbb{R}$.*

Let $\mathbb{K}$ be a field with characteristic 0, not necessarily algebraically closed.

2.1 Strict equivalence of pencils

Definition. A **matrix pencil** is a binomial of the form $A(\lambda) = A_0 + \lambda A_1$ whose coefficients are matrices.

Equivalently a matrix pencil is a λ-matrix of degree 1 with respect to the variable λ. We denote by $\mathfrak{M}_{m\times n}(\mathbb{K}[\lambda]_1)$ the set of matrix pencils of size $m \times n$. Consider the group action

$$\begin{array}{ccc} \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K}) & \longrightarrow & \mathrm{Aut}\left(\mathfrak{M}_{m\times n}(\mathbb{K}[\lambda]_1)\right) \\ (P,Q) & \mapsto & \left(A(\lambda) \mapsto P \cdot A(\lambda) \cdot {}^tQ \right) \end{array} \tag{2.1}$$

Definition. Two (matrix) pencils $A(\lambda)$ and $B(\lambda)$ are said to be **strictly equivalent** if they are in the same orbit with respect to the above group action.

We fix the notations

$$A(\lambda) \overset{\lambda}{\sim} B(\lambda) \quad , \quad A(\lambda) \overset{\circ}{\sim} B(\lambda)$$

for λ-equivalence and strict equivalence respectively.

We observe that λ-*equivalence* and *strict equivalence* differ in the dependence (or not) of the transformation matrices on the variable λ. Actually when the pencils are square these equivalences are somehow linked, but before showing this we share a preliminary useful lemma.

Lemma 2.1.1 (Generalized Euclidean Division)**.** Let R be a ring, not necessarily commutative, let $f(x), g(x) \in R[x]$ be polynomials of degree n, m (respectively) with coefficients in R such that $n \geq m$ and the leading coefficients g_m of g is invertible (i.e. $g_m \in R^\times$). Then there exist $p(x), q(x), r(x), s(x) \in R[x]$ polynomials such that:

- (*left-division of f by g*) $f(x) = g(x)p(x) + r(x)$;
- (*right-division of f by g*) $f(x) = q(x)g(x) + s(x)$;
- $\deg(r), \deg(s) < \deg(g)$.

Proof. Before proceeding with the proof we note that the notion “degree” is the one induced by the graded structure of the R-module $R[x]$. This enable us to work even if we are not in an euclidean domain.
We only prove the right-division (the left one is analogous) and we do so in an algorithmic way.
Let $f(x) = f_n x^n + \dots f_0$ and $g(x) = g_m x^m + \dots g_0$. Since $g_m \in R^\times$ we can left-multiply $g(x)$ by $f_n g_m^{-1} x^{n-m}$ and subtract the result to $f(x)$: we obtain

$$\tilde{f}^{(1)}(x) = f(x) - f_n g_m^{-1} x^{n-m} \cdot g(x) = \tilde{f}_{n-1} x^{n-1} + \dots + \tilde{f}_0$$

for suitable $\tilde{f}_i \in R$, with $\deg(\tilde{f}^{(1)}) \leq n-1$. We can repeat the right-division by $g(x)$ to $\tilde{f}^{(1)}$ and obtain $\tilde{f}^{(2)}(x)$ of degree $\leq n-2$. We iterate the right-division by $g(x)$ until we have $\deg(\tilde{f}^{(k)}) < m$. So this algorithm returns

$$f(x) = \tilde{f}^{(k-1)}(x) \cdot g(x) + \tilde{f}^{(k)}(x) \quad , \quad \deg(\tilde{f}^{(k)}) < \deg(g)$$

We note that at each iteration we strongly use the fact that g_m is a unit in R.
The same argument exactly holds for the left-division up to right-multiply g_m^{-1} and the whole divisor: for example in the first step we have

$$\tilde{f}^{(1)}(x) = f(x) - g(x) \cdot g_m^{-1} f_n x^{n-m}$$

□

Proposition 2.1.2. Let $A_0 + \lambda A_1$ and $B_0 + \lambda B_1$ be two square matrix pencils of size m with $\det A_1, \det B_1 \neq 0$. If they are λ-equivalent, then they are strictly equivalent. Moreover, if $P(\lambda)$ and $Q(\lambda)$ are such that $B_0 + \lambda B_1 = P(\lambda) \cdot (A_0 + \lambda A_1) \cdot {}^t Q(\lambda)$, then the matrices P and Q such that $B_0 + \lambda B_1 = P \cdot (A_0 + \lambda A_1) \cdot {}^t Q$ are obtained as left and right remainders respectively of $P(\lambda)$ and $Q(\lambda)$ in the division by $B_0 + \lambda B_1$.

Proof. Let $P(\lambda)$ and $Q(\lambda)$ be such that $B_0 + \lambda B_1 = P(\lambda) \cdot (A_0 + \lambda A_1) \cdot {}^t Q(\lambda)$. Let $P(\lambda)^{-1}$ be the inverse of $P(\lambda)$. Then

$$P(\lambda)^{-1} \cdot (B_0 + \lambda B_1) = (A_0 + \lambda A_1) \cdot {}^t Q(\lambda)$$

By regarding $P(\lambda)^{-1}$ and ${}^t Q(\lambda)$ as matrix polynomials, since $\det A_1, \det B_1 \neq 0$ (i.e. $A_1, B_1 \in \mathrm{GL}_m(\mathbb{K}) = \mathfrak{M}_m(\mathbb{K})^\times$) we can apply the previous lemma and divide this two matrix polynomials respectively by $A_0 + \lambda A_1$ (on the left) and $B_0 + \lambda B_1$ (on the right):

$$P(\lambda)^{-1} = (A_0 + \lambda A_1) \cdot S(\lambda) + R \quad , \quad {}^t Q(\lambda) = T(\lambda) \cdot (B_0 + \lambda B_1) + {}^t Q$$

where R,${}^t Q$ are constant matrices of size m. Hence

$$\Big((A_0 + \lambda A_1) \cdot S(\lambda) + R\Big) \cdot (B_0 + \lambda B_1) = (A_0 + \lambda A_1) \cdot \Big(T(\lambda) \cdot (B_0 + \lambda B_1) + {}^t Q\Big) \iff$$

$$\iff (A_0 + \lambda A_1) \cdot \Big(T(\lambda) - S(\lambda)\Big) \cdot (B_0 + \lambda B_1) = R \cdot (B_0 + \lambda B_1) - (A_0 + \lambda A_1) \cdot {}^t Q$$

By comparing the degrees of the two sides as matrix polynomials, since A_1 and B_1 are units (so there is no zero-divisor in the equation) it follows that $T(\lambda) - S(\lambda) = 0$ and

$$R \cdot (B_0 + \lambda B_1) = (A_0 + \lambda A_1) \cdot {}^t Q$$

Thus it is enough to prove that R is non-singular: in this case we have $P = R^{-1}$ and such P and Q satisfy our claim.
To do so we divide $P(\lambda)$ on the left by $B_0 + \lambda B_1$ and we obtain

$$P(\lambda) = (B_0 + \lambda B_1) \cdot U(\lambda) + P$$

By substitution in the previous equalities, we have

$$I = P(\lambda)^{-1} \cdot P(\lambda) = \Big((A_0 + \lambda A_1) \cdot S(\lambda) + R\Big) \cdot \Big((B_0 + \lambda B_1) \cdot U(\lambda) + P\Big) =$$

$$= (A_0 + \lambda A_1) \cdot \Big(S(\lambda) \cdot (B_0 + \lambda B_1) \cdot U(\lambda) + {}^t Q \cdot U(\lambda) + S(\lambda) \cdot P\Big) + R \cdot P$$

and by degrees arguments (and since A_1 is a unit) it necessarily follows that

$$S(\lambda) \cdot (B_0 + \lambda B_1) \cdot U(\lambda) + {}^t Q \cdot U(\lambda) + S(\lambda) \cdot P = 0 \quad , \quad RP = I$$

so $P = R^{-1}$ is non-singular, hence $B_0 + \lambda B_1 = P \cdot (A_0 + \lambda A_1) \cdot {}^t Q$.
We conclude by noticing that even Q is non-singular since

$$B_1 = P \cdot A_1 \cdot {}^t Q \implies \det B_1 = \det P \cdot \det A_1 \cdot \det {}^t Q$$

and $\det B_1, \det A_1 \neq 0$. □

In the preliminary chapter we exhibited a complete system of invariants for λ-equivalence (given by invariant polynomials). Now our goal is to find a complete system of invariants for strict equivalence: this would lead us to a canonical form for strict equivalence, that is to a canonical representative of the classes in the group

$$\mathrm{GL}_m(\mathbb{K}) \backslash \mathfrak{M}_{m\times n}(\mathbb{K}[\lambda]_1) / \mathrm{GL}_n(\mathbb{K})$$

Note: It is kind to underline that the strict equivalence extends the notion of (*left-right*) *equivalence* of matrices to matrix pencils.

We will reach our goal by steps and to do so we introduce the following definition.

Definition. A matrix pencil $A + \lambda B$ of size $m \times n$ is said to be:

- ◇ **regular** if $m = n$ and $\det(A + \lambda B)$ is a not-identically-zero polynomial;
- ◇ **singular** if $m \neq n$ or $\det(A + \lambda B)$ is identically zero.

Note: For square pencils, since $\det(A+\lambda B) = \det(B)\lambda^m + \cdots + \det(A)$, it follows that if the pencil is singular so are both A and B. Conversely, if A or B are non-singular, the pencil is regular.

2.1.1 From affine to projective matrix pencils

Let $A + \lambda B$ be a matrix pencil, not necessarily square. We can homogenize it and consider the homogeneous pencil $\mu A + \lambda B$. By repeating the construction of the invariant polynomials we did in the last section, we consider the (homogeneous) greatest common divisors $g_j(\mu, \lambda)$ of all the minors of order j of the matrix $\mu A + \lambda B$ and obtain the (homogeneous) invariant polynomials

$$i_j(\mu, \lambda) = \frac{g_{r-j+1}(\mu, \lambda)}{g_{r-j}(\mu, \lambda)} \quad \text{for} \quad j = 1 : r$$

Here again $\mathbb{K}[\mu, \lambda]$ is UFD, thus we can factorize the invariant polynomials as powers of homogeneous irreducible factors over $\mathbb{K}$ and obtain the elementary divisors $e_\alpha(\mu, \lambda)$ of the pencil $\mu A + \lambda B$ over $\mathbb{K}$.

We observe that by setting $\mu = 1$ we go back to the definitions we have given in the last section for a pencil of the form $A+\lambda B$ and from each $e_\alpha(\mu, \lambda)$ we obtain an elementary divisor $e_\alpha(1, \lambda) = \varphi_\alpha(\lambda)^{c_\alpha}$. Conversely, from each elementary divisor $\varphi_j(\lambda)^{c_j}$ of the pencil $A + \lambda B$ we can obtain an elementary divisor of $\mu A + \lambda B$ by homogenization $e_\alpha(\mu, \lambda) = \mu^\gamma \varphi_j(\frac{\lambda}{\mu})^{c_j}$.

In this way we obtain all the elementary divisors of $\mu A + \lambda B$ except the ones of the form μ^q: unfortunately this exception comes out from the fact that in general to dehomogenize and re-homogenize a polynomial gives not the identity on that polynomial. Nevertheless we will see that to dehomogenize and re-homogenize a matrix pencil gives right back the pencil we started from since we are just looking at a linear polynomial of the form $xa + by$.

In particular, elementary divisors of $\mu A + \lambda B$ of the form μ^q are called **infinite elementary divisors** of $A+\lambda B$ and, for square pencils, they exist if and only if $\det B = 0$: indeed

$$\mu^q \text{ is an elementary divisor of } \mu A + \lambda B \iff$$
$$\iff 1 \text{ is an invariant polynomial of } A + \lambda B \iff$$
$$\iff \text{the Smith form of } A + \lambda B \text{ is of the form } \operatorname{diag}(1, C + \lambda D) \iff$$
$$\iff \text{the Smith form of } B \text{ is (up to permutation) of the form } \operatorname{diag}(0, *) \iff$$
$$\iff \det B = 0$$

Note: In a geometric flavoured perspective, we may say that in the last section we just worked in the affine chart $\{\mu \neq 0\}$ of a projective space and what we missed were just the informations at the infinity $\{\mu = 0\}$. This perspective is very useful because by considering a pencil of the form $\mu A + \lambda B$ we are in fact considering a projective line parameterized by $[\mu : \lambda] \in \mathbb{P}^1_{\mathbb{K}}$ and defined by the two matrices A and B. Thus we are gratefully allowed to choose the generators A and B on this line as conveniently as we need. We may (and will) refer to the homogeneous pencils as *projective* and to the other ones as *affine*.

Since we obtain a homogeneous pencil by homogenizing an affine one, it follows:

Proposition 2.1.3. $A_0 + \lambda A_1$ and $B_0 + \lambda B_1$ are strictly equivalent so $\mu A_0 + \lambda A_1$ and $\mu B_0 + \lambda B_1$ are. In particular, if two pencils are strictly equivalent, then their elementary divisors (both finite and infinite ones) must coincide.

In general the elementary divisors are not enough to give a complete system of invariants of strict equivalence. Now it is important to distinguish between regular and singular pencils.

2.2 Canonical form for regular pencils

Let us now see that for regular pencils the converse to proposition 2.1.3 holds too.

Let $A_0 + \lambda A_1$ and $B_0 + \lambda B_1$ be two regular pencils of size m with same elementary divisors, both finite and infinite ones. Now we show that the homogenized pencils are strictly equivalent too.

Given $\mu A_0 + \lambda A_1$ and $\mu B_0 + \lambda B_1$ their homogenizations, we consider the parameters transformation

$$\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} \tilde{\lambda} \\ \tilde{\mu} \end{bmatrix} = \begin{bmatrix} \lambda \\ \mu \end{bmatrix}$$

where $a_1 b_2 - a_2 b_1 \neq 0$. We write the pencils with respect to the new parameters:

$$\mu A_0 + \lambda A_1 = \tilde{\mu}(b_2 A_0 + a_2 A_1) + \tilde{\lambda}(b_1 A_0 + a_1 A_1) = \tilde{\mu}\tilde{A}_0 + \tilde{\lambda}\tilde{A}_1$$

$$\mu B_0 + \lambda B_1 = \tilde{\mu}(b_2 B_0 + a_2 B_1) + \tilde{\lambda}(b_1 B_0 + a_1 B_1) = \tilde{\mu}\tilde{B}_0 + \tilde{\lambda}\tilde{B}_1$$

Since $\mu A_0 + \lambda A_1$ and $\mu B_0 + \lambda B_1$ are regular, we can choose a_1, b_1 such that $\det \tilde{A}_1, \det \tilde{B}_1 \neq 0$. Hence we are in the case where λ-equivalence and strict equivalence coincide, so $\tilde{\mu}\tilde{A}_0 + \tilde{\lambda}\tilde{A}_1$ and $\tilde{\mu}\tilde{B}_0 + \tilde{\lambda}\tilde{B}_1$ are strictly equivalent, so $\mu A_0 + \lambda A_1$ and $\mu B_0 + \lambda B_1$ too.

We just proved the following result which answers our main problem in the case of regular pencils.

Theorem 2.2.1. Two regular pencils $A_0 + \lambda A_1$ and $B_0 + \lambda B_1$ are strictly equivalent if and only if their homogenized pencils $\mu A_0 + \lambda A_1$ and $\mu B_0 + \lambda B_1$ have the same elementary divisors.

Next we determine the canonical form of the strictly equivalence classes for regular pencils. Given a regular pencil $A + \lambda B$, since $\det(A + \lambda B)$ is not identically zero there exists $c \in \mathbb{K}$ such that $\det(A + cB) \neq 0$. By setting $A_1 = A + cB$ we rewrite the pencil

$$A + \lambda B = A_1 + (\lambda - c)B$$

with $\det A_1 \neq 0$, thus by left-multiplication for A_1^{-1} we obtain

$$I + (\lambda - c)A_1^{-1}B$$

Let $J = \operatorname{diag}(J_0, J_1)$ be the Jordan form of $A_1^{-1}B$ where J_0 is the nilpotent block (i.e. with all zero eigenvalues) and J_1 is the non-singular block (i.e. with all non-zero eigenvalues): then the pencil $I + (\lambda - c)A_1^{-1}B$ is similar to the pencil

$$I + (\lambda - c)\begin{bmatrix} J_0 & 0 \\ 0 & J_1 \end{bmatrix} = \begin{bmatrix} I + (\lambda - c)J_0 & 0 \\ 0 & I + (\lambda - c)J_1 \end{bmatrix}$$

Now we separately work on the two blocks.

- ⋄ By multiplying for $\operatorname{diag}((I - cJ_0)^{-1}, I)$ we obtain

$$\begin{bmatrix} I + \lambda(I - cJ_0)^{-1}J_0 & 0 \\ 0 & I + (\lambda - c)J_1 \end{bmatrix} \tag{2.2}$$

We observe that $(I - cJ_0)^{-1}J_0$ is nilpotent: indeed, if J_0 has nilpotence index h,

$$(I - cJ_0)^{-1} = I + cJ_0 + \ldots + (cJ_0)^h + \ldots \overset{J_0^h=0}{=} I + cJ_0 + \ldots + (cJ_0)^{h-1}$$

and $(I - cJ_0)^{-1}J_0 = J_0 + cJ_0^2 + \ldots + c^{h-1}J_0^h$, which is nilpotent.
Thus, if $\widehat{J_0}$ is the nilpotent Jordan form of $(I - cJ_0)^{-1}J_0$, it follows that (2.2) is similar to the block-diagonal pencil

$$\begin{bmatrix} I + \lambda\widehat{J_0} & \\ & I + (\lambda - c)J_1 \end{bmatrix} = \begin{bmatrix} N^{(u_1)} & & & \\ & \ddots & & \\ & & N^{(u_s)} & \\ & & & I + (\lambda - c)J_1 \end{bmatrix} \tag{2.3}$$

where $N^{(u_i)}$ is a block of size $u_i \times u_i$ of the form

$$N^{(u_i)} = \begin{bmatrix} 1 & \lambda & & \\ & \ddots & \ddots & \\ & & \ddots & \lambda \\ & & & 1 \end{bmatrix} \tag{2.4}$$

◇ Since $\det J_1 \neq 0$ we may multiply (2.3) for $\operatorname{diag}(I, J_1^{-1})$ and obtain

$$\begin{bmatrix} \operatorname{diag}(N^{(u_1)} \ldots N^{(u_s)}) & \\ & (J_1^{-1} - cI) + \lambda I \end{bmatrix}$$

and, if we set $\widehat{J_1}$ for the Jordan form of $(J_1^{-1} - cI)$, we obtain the similar pencil

$$\begin{bmatrix} \operatorname{diag}(N^{(u_1)} \ldots N^{(u_s)}) & \\ & \widehat{J_1} + \lambda I \end{bmatrix} \tag{2.5}$$

Next we want to find a canonical decomposition for the block $\widehat{J_1} + \lambda I$ too. One may propose to explicit the Jordan form $\widehat{J_1}$ but that would depend on the field $\mathbb{K}$ we are working over. Instead we are looking for a canonical form independently from the base field. To reach this we need to introduce new "canonical" blocks but first it is convenient to put ourselves in the homogeneous case.

Remark 2.2.2. Previously we remarked that to dehomogenize and re-homogenize a homogeneous matrix pencil gives back the homogeneous pencil we started from. One may wonder if to dehomogenize would give problems with respect to the elementary divisors since we lost informations on the multiplicity of the infinite divisors. However, from the point of view of the canonical form, these informations are kept hidden in the dimensions of the diagonal blocks. Thus it follows that by homogenizing the canonical form of the affine pencil we just obtain the canonical form of the projective one.

By homogenizing the (affine) blocks $N^{(u_i)}$ in (2.4) we obtain the (projective) blocks of size $u_i \times u_i$

$$H^{(u_i)} = \begin{bmatrix} \mu & \lambda & & \\ & \ddots & \ddots & \\ & & \ddots & \lambda \\ & & & \mu \end{bmatrix} \tag{2.6}$$

To lighten up the notation it is also convenient to introduce the next definition.

Definition. Let M, N be two matrix of size $m_1 \times m_2$ and $n_1 \times n_2$ respectively (not necessarily of the same size). We define the **block-direct sum** of M and N the $((m_1 + n_1) \times (m_2 + n_2))$-matrix

$$M \boxplus N = \begin{bmatrix} M & 0 \\ 0 & N \end{bmatrix}$$

Thus by homogeneizing the matrix in (2.5) we can rewrite it as

$$\left(\boxplus_{k=1}^{s} H^{(u_k)} \right) \boxplus (\mu \widehat{J_1} + \lambda I)$$

Let us introduce new canonical blocks. Let $p(t) \in \mathbb{K}[t]$ be a monic polynomial of the form

$$p(t) = t^n + p_{n-1}t^{n-1} + \ldots + p_0$$

and let $a \in \mathbb{K}$. We define:

- C_p the **companion matrix** of $p(t)$, i.e. the square matrix of size n of the form

$$C_p = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -p_0 & -p_1 & \ldots & -p_{n-2} & -p_{n-1} \end{bmatrix}$$

- $\mathfrak{F}_p$ the **Frobenius companion block** (or simply *Frobenius block*) of $p(t)$, i.e. the square pencil of size n of the form

$$\mathfrak{F}_p = \mu C_p + \lambda I_n = \begin{bmatrix} \lambda & \mu & & & \\ & \lambda & \mu & & \\ & & \ddots & \ddots & \\ & & & \lambda & \mu \\ -\mu p_0 & -\mu p_1 & \ldots & -\mu p_{n-2} & \lambda - \mu p_{n-1} \end{bmatrix}$$

- $J_{n,a}$ the **Jordan block** of size n *with respect to the eigenvalue* a, i.e.

$$J_{n,a} = \begin{bmatrix} a & 1 & & \\ & a & \ddots & \\ & & \ddots & 1 \\ & & & a \end{bmatrix}$$

- $\mathfrak{J}_{n,a}$ the **Jordan block** of size n *with respect to* $(\lambda + a\mu)^n$, i.e.

$$\mathfrak{J}_{n,a} = \mu J_{n,a} + \lambda I_n = \begin{bmatrix} \lambda + a\mu & \mu & & \\ & \lambda + a\mu & \ddots & \\ & & \ddots & \mu \\ & & & \lambda + a\mu \end{bmatrix}$$

Let $\mu A + \lambda B$ be a projective pencil of rank r with invariant polynomials

$$i_j(\mu, \lambda) = \frac{g_{r-j+1}(\mu, \lambda)}{g_{r-j}(\mu, \lambda)} \quad \text{for} \quad j = 1 : r \quad , \quad i_r(\mu, \lambda) \mid i_{r-1}(\mu, \lambda) \mid \ \dots \ \mid i_1(\mu, \lambda)$$

We factorize them as

$$i_j(\mu, \lambda) = \mu^{u_j} \cdot \prod_{k=1}^{r_j} \left(p_k^{(j)}(\mu, \lambda) \right)^{w_k}$$

where $u_j \geq 0$, $w_j \geq 0$ and $p_k^{(j)}(1, \lambda)$ is irreducible in $\mathbb{K}[\lambda]$ of degree > 0. The polynomials μ^{u_j} are the infinite elementary divisors of the affine pencil $A + \lambda B$ and the polynomials $\left(p_k^{(j)}(1, \lambda) \right)^{w_k}$ are the finite elementary divisors of $A + \lambda B$.

If $\mu A + \lambda B$ is regular of size m, then $\det(\mu A + \lambda B) = \prod_{j=1:m} i_j(\mu, \lambda)$. In particular, we recall that:

- $\det B$ is the leading coefficient with respect to λ of $\det(\mu A + \lambda B)$, i.e.

$$\det(\mu A + \lambda B) = \det B \cdot \lambda^m + D(\mu, \lambda) \quad \text{where} \ \deg_\lambda(D) < m$$

- $\mu A + \lambda B$ does not have elementary divisors of the form $\mu^{u_j} \iff \det B \neq 0$.

Proposition 2.2.3. Let $p(t) = t^n + p_{n-1}t^{n-1} + \ldots + p_0$ be a polynomial in $\mathbb{K}[t]$ and let $a \in \mathbb{K}$. Let $\mathfrak{F}_p$ and $\mathfrak{J}_{n,a}$ be the pencils defined above. Then:

(i) The only invariant polynomial of $\mathfrak{F}_p$ is

$$(-1)^n \mu^n p(-\frac{\lambda}{\mu}) = \lambda^n + \sum_{i=0}^{n-1} (-1)^{n-i} p_i \lambda^i \mu^{n-i}$$

(*ii*) The only invariant polynomial of $\mathfrak{J}_{n,a}$ is

$$(\lambda + a\mu)^n$$

Proof. The above polynomials are the determinants of the respective pencils, so it is enough to prove that the greatest common divisors of the minors of size less than n are all equal to 1.

(*i*) In $\mathfrak{F}_p$ for all $k < n$ there are the minors μ^k and λ^k which are coprime.

(*ii*) In $\mathfrak{J}_{n,a}$ for all $k < n$ there are the minors $(\lambda + a\mu)^k$ and μ^k which are coprime.

□

Since we are interested in the block $\mu\widehat{J_1} + \lambda I$, we now assume $\mu A + \lambda B$ regular with $\det B \neq 0$, hence there are not infinite elementary divisors, hence $\mu A + \lambda B \overset{\circ}{\sim} \mu\widehat{J_1} + \lambda I$.

Let $\left(p_k^{(j)}(1,\lambda)\right)^{w_k}$ be a finite elementary divisor of $A + \lambda B$ and let $q_{jk}(t)$ be the polynomial defined by

$$q_{jk}(t) = (-1)^{w_k \deg(p_k^{(j)})}\left(p_k^{(j)}(1,-t)\right)^{w_k}$$

By proposition 2.2.3 the only invariant polynomial of $\mathfrak{F}_{q_{jk}}$ is $\left(p_k^{(j)}(\mu,\lambda)\right)^{w_k}$: actually this is the only elementary divisor of the Frobenius block since $p_k^{(j)}(1,\lambda)$ is irreducible.

Remark 2.2.4. With abuse of notation we will write $\mathfrak{F}_{(p_k^{(j)})^{w_k}}$ instead of $\mathfrak{F}_{q_{jk}}$ to keep in mind its elementary divisor.

Thus the Frobenius block of $q_{jk}(t)$ is

$$\mathfrak{F}_{(p_k^{(j)})^{w_k}} = \mu C_{q_{jk}} + \lambda I = \begin{bmatrix} \lambda & \mu & & & \\ & \lambda & \mu & & \\ & & \ddots & \ddots & \\ & & & \lambda & \mu \\ (-1)^{n_{jk}}\mu\tilde{p}_0 & (-1)^{n_{jk}+1}\mu\tilde{p}_1 & \dots & \mu\tilde{p}_{n_{jk}-2} & \lambda - \mu\tilde{p}_{n_{kj}-1} \end{bmatrix}$$

where $\tilde{p}_i$ are the coefficients of $(p_k^{(j)}(1,\lambda))^{w_k}$ and $n_{kj} = w_k \deg(p_k^{(j)})$ its degree. Then

$$\boxplus_{j,k} \mathfrak{F}_{(p_k^{(j)})^{w_k}} = \boxplus_{j,k}(\mu C_{q_{jk}} + \lambda I) = \mu\Big(\boxplus_{j,k} C_{q_{jk}}\Big) + \lambda I$$

Since the elementary divisors are a complete system of invariants of strict equivalence for regular pencils, it follows that

$$\mu A + \lambda B \overset{\circ}{\sim} \mu\widehat{J_1} + \lambda I \overset{\circ}{\sim} \boxplus_{j,k} \mathfrak{F}_{(p_k^{(j)})^{w_k}}$$

This completes the classification for regular pencils. In the following we state the same result in two different ways: first by exhibiting the Kronecker form as a diagonal-block matrix, then as a (more compact) block-direct-sum.

Remark 2.2.5. We state the two theorems in terms of projective regular pencils: clearly they hold for affine ones too and it is enough to dehomogenize the canonical forms with respect to the variable μ, i.e. by setting $\mu = 1$.

Theorem 2.2.6 (Kronecker - regular form)**.** Every regular projective pencil $\mu A + \lambda B$ is strictly equivalent to a canonical regular projective pencil of the form

$$\begin{bmatrix} \operatorname{diag}\left(H^{(u_1)} \dots H^{(u_s)}\right) & \\ & \operatorname{diag}\left(\left\{\mathfrak{F}_{(p_k^{(j)})^{w_k}} \mid j, k\right\}\right) \end{bmatrix} \tag{2.7}$$

where the first s blocks $H^{(u_i)}$ are uniquely defined by the elementary divisors $\mu^{u_1} \dots \mu^{u_s}$ and the Frobenius blocks $\mathfrak{F}_{(p_k^{(j)})^{w_k}}$ are uniquely determined by the other elementary divisors $\{(p_k^{(j)})^{w_k} \mid j, k\}$.

Theorem 2.2.7 (Weierstrass - regular form)**.** Every projective pencil $\mu A + \lambda B$ is strictly equivalent to the canonical block-direct-sum

$$\left(\boxplus_{k=1}^{s} H^{(u_k)}\right) \boxplus \left(\boxplus_{l,z} \mathfrak{F}_{(p_z^{(l)})^{w_z}}\right) \tag{2.8}$$

where the blocks $H^{(u_k)}$ are uniquely determined by the elementary divisors of the form $\mu^{u_1} \dots \mu^{u_s}$ and the blocks $\mathfrak{F}_{(p_z^{(l)})^{w_z}}$ are uniquely determined by the elementary divisors of the form $(p_z^{(l)}(\mu, \lambda))^{w_z}$.

Note: We will indistinctly refer to both the above canonical forms as to the Kronecker form, Weierstrass form or even Kronecker-Weierstrass form.

2.2.1 Jordan blocks and $\mathrm{GL}_2(\mathbb{K})$-action

Before concluding this section we emphasyze two interesting facts. Previously we assumed that the pencil $\mu A + \lambda B$ was regular with $\det B \neq 0$, so it did not have elementary divisors of the form μ^{u_i}.

Decomposition with respect to invariant polynomials. Consider $i_j(\mu, \lambda)$ the j-th invariant polynomial of $\mu A + \lambda B$ and factorize it as

$$i_j(\mu, \lambda) = \prod_{k=1}^{r_j} \left(p_k^{(j)}(\mu, \lambda)\right)^{w_k}$$

where $p_k^{(j)}(1,\lambda)$ are irreducible in $\mathbb{K}[\lambda]$ of degree > 0 and their powers $\left(p_k^{(j)}(\mu,\lambda)\right)^{w_k}$ are (some of the) elementary divisors of $\mu A + \lambda B$.

Proposition 2.2.8. In the above assumptions,

$$\mathfrak{F}_{i_j} \overset{\circ}{\sim} \boxplus_k \mathfrak{F}_{(p_k^{(j)})^{w_k}}$$

Proof. Before proceeding we underline that $\mathfrak{F}_{i_j}$ is not the Frobenius block of $i_j(1,\lambda)$ but the one whose invariant polynomial is $i_j(\mu,\lambda)$ (same notation we fixed for $\mathfrak{F}_{(p_k^{(j)})^{w_k}}$). Since the elementary divisors give a complete system of invariants of strict equivalence for regular pencils, it is enough to show that the ones of the two pencils coincide.
The pencil $\mathfrak{F}_{i_j}$ has (one and only) invariant polynomial $i_j(\mu,\lambda)$, thus its elementary divisors are $\left\{\left(p_k^{(j)}(\mu,\lambda)\right)^{w_k}\right\}_k$. Each block $\mathfrak{F}_{(p_k^{(j)})^{w_k}}$ has (one and only) invariant polynomial $\left(p_k^{(j)}(\mu,\lambda)\right)^{w_k}$ which is also its only one elementary divisor. From the fact 1.1.3 it follows that the pencil $\boxplus_k \mathfrak{F}_{(p_k^{(j)})^{w_k}}$ has elementary divisors $\left\{\left(p_k^{(j)}(\mu,\lambda)\right)^{w_k}\right\}_k$, thus the thesis. □

This propostition gives us the perfect assist to our first goal:

Corollary 2.2.9. Let $\mu A + \lambda B$ be a regular pencil with $\det B \neq 0$. Then it is strictly equivalent to the block-direct-sum of the Frobenius blocks of its invariant polynomials, i.e.

$$\mu A + \lambda B \overset{\circ}{\sim} \boxplus_j \mathfrak{F}_{i_j}$$

Obviously the above decomposition is less fine than the one in Frobenius blocks of the elementary divisors. But it gains interest if the pencil has only one invariant polynomial (which factorizes in all the elementary divisors of the pencil) since it means that such pencil is strictly equivalent to the Frobenius block of one only polynomial. But there is more than this to unveil: in such case this polynomial is right the determinant of the pencil, i.e.

$$\mu A + \lambda B \overset{\circ}{\sim} \mathfrak{F}_{\det(\mu A + \lambda B)}$$

Remark 2.2.10. In general we cannot cluster the block-direct-sum $\boxplus_j \mathfrak{F}_{i_j}$ in one Frobenius block: we can do so only if the invariant polynomials $i_j(\mu,\lambda)$ are coprimes but by their divisibility property this is equivalent to say that there is only one invariant polynomial (i.e. $i_2 = \ldots i_r = 1$).

Jordan blocks. The next interesting fact we want to emphasyze concerns the Jordan block with respect to $(\lambda + a\mu)^n$, i.e.

$$\mathfrak{J}_{n,a} = \mu J_{n,a} + \lambda I_n = \begin{bmatrix} \lambda + a\mu & \mu & & \\ & \lambda + a\mu & \ddots & \\ & & \ddots & \mu \\ & & & \lambda + a\mu \end{bmatrix} \tag{2.9}$$

In particular we want to express some addends in the block-direct-sum of the Kronecker form as Jordan block. Before going on we remark that in general we are not working over an algebraically closed field, hence irreducible polynomials (in particular the elementary divisors) may have degree greater than 1.

Let $\big(p(\mu, \lambda)\big)^n$ be an elementary divisor of the regular pencil $\mu A + \lambda B$ and let us assume $p(\mu, \lambda)$ to be linear, i.e.

$$\big(p(\mu, \lambda)\big)^n = (\lambda + a\mu)^n$$

(of course we can always assume the polynomial to be monic in λ up to multiplication by a unit). Let $q(t)$ be the polynomial defined by

$$q(t) = (-1)^n \big(p(1, -t)\big)^m = (-1)^n (-t + a)^n$$

Then the Frobenius block $\mathfrak{F}_{p^n}$ of $q(t)$ has invariant polynomial $\big(p(\mu, \lambda)\big)^n = (\lambda + a\mu)^n$ and it is the only one. But this is the same (only one) invariant polynomial of the Jordan block with respect to $(\lambda + a\mu)^n$. Since such invariant polynomial is also the only one elementary divisor of each of the two blocks, it follows that:

Proposition 2.2.11. The Frobenius block corresponding to an elementary divisor of the form $(\lambda + a\mu)^n$ is strictly equivalent to the related Jordan block, i.e.

$$\mathfrak{F}_{(\lambda + a\mu)^n} \overset{\circ}{\sim} \mathfrak{J}_{n,a}$$

Thus the next result is just a corollary of what we found out.

Theorem 2.2.12. Every regular pencil $\mu A + \lambda B$ with $\det B \neq 0$ is strictly equivalent to the block-direct-sum of:

- the Jordan blocks $J_{w_k, a_{jk}}$ with respect to its completely-factorized elementary divisors $(\lambda + a_{jk}\mu)^{w_k}$;
- the Frobenius blocks $\mathfrak{F}_{(p_k^{(j)})^{w_k}}$ corresponding to the remaining elementary divisors.

In particular, if $\mathbb{K}$ is algebraically closed, the Kronecker-Weierstrass form is block-direct-sum of only Jordan blocks

$$\mu A + \lambda B \overset{\circ}{\sim} \boxplus_{j,k} \mathfrak{J}_{w_{jk},a_{jk}}$$

uniquely determined by the elementary divisors $(\lambda + a_{jk}\mu)^{w_{jk}}$.

$\mathrm{GL}_2(\mathbb{K})$**-action.** We conclude the section with a simple remark. Above we did the strong assumption that $\det B \neq 0$ in the regular pencil $\mu A + \lambda B$: we did so because we were interested in the last block $\mu\widehat{J_1} + \lambda I$.
Obviously not every regular pencil has such property, but we can always trace back to that case.

Proposition 2.2.13. Let $\mu A + \lambda B$ be a regular pencil. Then there exists a linear transformation over $\mathbb{K}^2$ for which the same pencil can be rewritten as $\tilde{\mu}\tilde{A} + \tilde{\lambda}\tilde{B}$ with $\det \tilde{B} \neq 0$.

Proof. We are looking for a linear transformation of the form

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \tilde{\mu} \\ \tilde{\lambda} \end{bmatrix} = \begin{bmatrix} \mu \\ \lambda \end{bmatrix}$$

with $ad - bc \neq 0$ (i.e. invertible) such that the pencil $\mu A + \lambda B = \tilde{\mu}(aA + cB) + \tilde{\lambda}(bA + dB)$ has $\det(bA + dB) \neq 0$. Since the linear system

$$\begin{cases} ad - bc = 0 \\ \det(bA + dB) = 0 \end{cases}$$

has a finite number of solutions and $\mathbb{K}$ is infinite, it follows that a linear trasformation satisfying our requests exists. □

This remark gains more interest if we look at the pencil $\mu A + \lambda B$ has a projective line: what we did was just to change the points at the infinity of the line, that is the infinite elemetary divisors μ^{u_i} of $A + \lambda B$ become finite elementary divisors of $\tilde{\mu}\tilde{A} + \tilde{\lambda}\tilde{B}$.

Let us denote with $\mathfrak{M}_{m\times n}(\mathbb{K}[\mu,\lambda]_1)$ the space of the projective pencils of size $m \times n$. In this chapter we just considered the action of $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ on $\mathfrak{M}_{m\times n}(\mathbb{K}[\mu,\lambda]_1)$. What we are doing now is to extend the acting group to $\mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ as follows:

$$\begin{array}{ccc} \mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K}) & \longrightarrow & \mathrm{Aut}\left(\mathfrak{M}_{m\times n}(\mathbb{K}[\mu,\lambda]_1)\right) \\ \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}, P, Q\right) & \mapsto & \left(\mu A + \lambda B \mapsto \tilde{\mu}(P \cdot A \cdot {}^t Q) + \tilde{\lambda}(P \cdot B \cdot {}^t Q) \right) \end{array} \tag{2.10}$$

where

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \mu \\ \lambda \end{bmatrix} = \begin{bmatrix} \tilde{\mu} \\ \tilde{\lambda} \end{bmatrix}$$

In this new perspective the strict equivalence is traduced in the action of the triples (I_2, P, Q) (where I_2 is the identity matrix of size 2). Later we will introduce a new equivalence defined by the orbits of this new group action.

2.3 Canonical form for singular pencils

Let $A + \lambda B$ be a singular pencil of size $m \times n$. Since it is singular, $\mathrm{Rk}(A + \lambda B) < \min\{m, n\}$: say $\mathrm{Rk}(A + \lambda B) < n$. Hence its columns are linearly dependent over $\mathbb{K}[\lambda]$, hence there exists $x(\lambda) \in (\mathbb{K}[\lambda])^n \setminus \{0\}$ such that $(A + \lambda B)x(\lambda) = 0$. Consider

$$\epsilon = \min \left\{ \deg x(\lambda) \mid x(\lambda) \in \ker(A + \lambda B) \setminus \{0\} \right\}$$

and $x(\lambda) = x_0 - x_1\lambda + \ldots + (-1)^\epsilon x_\epsilon \lambda^\epsilon$, where $x_i \in \mathbb{K}^n$ and $x_\epsilon \neq 0$.
Since $x(\lambda) \in \ker(A + \lambda B)$

$$\begin{cases} Ax_0 = 0 \\ Bx_{k-1} - Ax_k = 0 \quad \forall k = 1 : \epsilon \\ Bx_\epsilon = 0 \end{cases} \tag{2.11}$$

Such system may be represented by

$$M_\epsilon \begin{bmatrix} x_0 \\ -x_1 \\ \vdots \\ (-1)^{\epsilon-1}x_{\epsilon-1} \\ (-1)^\epsilon x_\epsilon \end{bmatrix} = \begin{bmatrix} A & & & \\ B & A & & \\ & B & \ddots & \\ & & \ddots & A \\ & & & B \end{bmatrix} \begin{bmatrix} x_0 \\ -x_1 \\ \vdots \\ (-1)^{\epsilon-1}x_{\epsilon-1} \\ (-1)^\epsilon x_\epsilon \end{bmatrix} = 0 \tag{2.12}$$

where M_ϵ is a block-matrix of size $(\epsilon + 2) \times (\epsilon + 1)$ with blocks of size $m \times n$.
For all $k = 0 : \epsilon$ let M_k be the submatrix of M_ϵ given by the first $k + 2$ block-rows and $k + 1$ block-columns and let $\rho_k = \mathrm{Rk}\, M_k$: surely $\rho_\epsilon < (\epsilon + 1)n$ and by minimality of ϵ

$$\begin{cases} \rho_0 = \mathrm{Rk} \begin{bmatrix} A \\ B \end{bmatrix} = n \\ \rho_k = (k + 1)m \quad \forall k = 1 : \epsilon - 1 \end{cases}$$

Thus

$$\epsilon = \min\{k \mid \rho_k \lneq (k + 1)n\}$$

Note: Let r_A, r_B, r be respectively the ranks of $A, B, A + \lambda B$. If $\epsilon = 0$, then we have a linear dependence over $\mathbb{K}$ (and not just over $\mathbb{K}[\lambda]$) for the columns of $A + \lambda B$, and for the identity principle of polynomials we have a linear dependence over $\mathbb{K}$ for A and B. Thus in this case $r_A, r_B \lneq n$.

Our goal is to determine a canonical form of strict equivalence for singular pencils, just as we did for the regular case. The main idea is to iteratively split a singular pencil into "subpencils" until one of them is regular. A first step in the good direction is given by the following result [14, Ch.XII, Theorem 4].

Theorem 2.3.1. If the equation $(A + \lambda B)x(\lambda) = 0$ has solution of minimum degree $\epsilon > 0$, then the pencil $A + \lambda B$ is strictly equivalent to the pencil

$$\begin{bmatrix} L_\epsilon & 0 \\ 0 & \tilde{A} + \lambda \tilde{B} \end{bmatrix}$$

where the matrix

$$L_\epsilon = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \end{bmatrix}$$

has size $\epsilon \times (\epsilon + 1)$ and $\tilde{A} + \lambda \tilde{B}$ is such that every solution of the equation $(\tilde{A} + \lambda \tilde{B})x(\lambda) = 0$ has degree greater than ϵ.

Remark 2.3.2. Until now we have worked with the columns of the pencil but we may repeat the same arguments by working with the rows of the pencil: it is enough to left-multiply the pencil by a row left-solution of length m, or equivalently to study the solution of the transposed pencil. Then the previous result still hold when working with rows up to considering the transposed matrix pencil and this leads us to a block of the form

$$^tL_\eta = \begin{bmatrix} \lambda & & \\ 1 & \ddots & \\ & \ddots & \lambda \\ & & 1 \end{bmatrix}$$

of size $(\eta + 1) \times \eta$.

Now we analyze two cases: in the previous notations, first we assume $\epsilon \gneq 0$ and so that there is no constant linear dependence (i.e. independent from λ) neither for rows or columns, then we admit the existence of such dependences.

Case 1: Let $r < n$ be the rank of the pencil and let $\epsilon_1 > 0$ be the minimum degree of the non-zero solutions of $(A + \lambda B)x = 0$. By theorem 2.3.1 it follows

$$A + \lambda B \overset{\circ}{\sim} \begin{bmatrix} L_{\epsilon_1} & 0 \\ 0 & A_1 + \lambda B_1 \end{bmatrix}$$

where $(A_1 + \lambda B_1)x^{(1)} = 0$ has no solution of degree $< \epsilon_1$.
We may repeat the above argument and consider ϵ_2 the minimum degree of the non-zero solutions of $(A_1 + \lambda B_1)x^{(1)} = 0$: obviously $0 < \epsilon_1 \le \epsilon_2$. By reapplying the theorem we obtain

$$A_1 + \lambda B_1 \overset{\circ}{\sim} \begin{bmatrix} L_{\epsilon_2} & 0 \\ 0 & A_2 + \lambda B_2 \end{bmatrix}$$

hence

$$A + \lambda B \overset{\circ}{\sim} \begin{bmatrix} L_{\epsilon_1} & & \\ & L_{\epsilon_2} & \\ & & A_2 + \lambda B_2 \end{bmatrix}$$

By iteration we have that

$$A + \lambda B \overset{\circ}{\sim} \begin{bmatrix} L_{\epsilon_1} & & & \\ & \ddots & & \\ & & L_{\epsilon_p} & \\ & & & A_p + \lambda B_p \end{bmatrix}$$

where $0 < \epsilon_1 \le \cdots \le \epsilon_p$ and $(A_p + \lambda B_p)x^{(p)} = 0$ has no non-zero solution, i.e. $A + \lambda B_p$ has linearly independent columns.

Note: If $\epsilon_1 + \cdots + \epsilon_p = m$, then $A_p + \lambda B_p$ is missing.

Now we focus on the rows of $A_p + \lambda B_p$. If they are linearly dependent, then the same arguments (together with remark 2.3.2) lead us to

$$A + \lambda B \overset{\circ}{\sim} \begin{bmatrix} \mathrm{diag}(L_{\epsilon_1} \dots L_{\epsilon_p}) & & \\ & \mathrm{diag}({}^tL_{\eta_1} \dots {}^t L_{\eta_q}) & \\ & & A_0 + \lambda B_0 \end{bmatrix}$$

where $0 < \epsilon_1 \le \cdots \le \epsilon_p$, $0 < \eta_1 \le \cdots \le \eta_q$ and $A_0 + \lambda B_0$ has linearly independent rows and columns, i.e. $A_0 + \lambda B_0$ is a regular pencil.

Remark 2.3.3. More precisely, the regular block $A_0 + \lambda B_0$ has size

$$\Big(m - (\sum_{i=1}^{p} \epsilon_i + \sum_{j=1}^{q} \eta_j + q)\Big) \times \Big(n - (\sum_{i=1}^{p} \epsilon_i + \sum_{j=1}^{q} \eta_j + p)\Big)$$

We note that $A_0 + \lambda B_0$ is indeed a square pencil, since the dimension of the span of linearly independent rows is always equal to the dimension of the span of the linearly independent columns, hence in this case necessarily p and q are such that $m-q = n-p$.

Case 2: Let us assume that there are constant linear dependences of rows and/or columns. We set

$$g = \max\{ \text{ number of independent constant solutions of } (A+\lambda B)x = 0 \ \}$$

$$h = \max\{ \text{ number of independent constant solutions of } ({}^tA+\lambda {}^tB)x = 0 \ \}$$

Let $e_1, \dots, e_g \in \mathbb{K}^n$ be linearly independent constant solutions of $(A+\lambda B)x = 0$. We extend them to basis of $\mathbb{K}^n$ and we rewrite the pencil with respect to the new basis: we obtain the pencil

$$\tilde{A} + \lambda \tilde{B} = \left[\begin{array}{c|c} 0_{m\times g} & \tilde{A}_1 + \lambda \tilde{B}_1 \end{array} \right]$$

In an analogous way by a basis change we may "annihilate" the first h rows of $\tilde{A}_1 + \lambda \tilde{B}_1$ and obtain the pencil

$$\hat{A} + \lambda \hat{B} = \begin{bmatrix} 0_{h\times g} & 0_{h\times(n-g)} \\ 0_{(m-h)\times g} & A^0 + \lambda B^0 \end{bmatrix}$$

where the diagonal block $A^0 + \lambda B^0$ has no constant linear dependences of rows and/or columns, so we are again in *Case 1.* What we just obtained is (a part of) the canonical form of a singular pencil:

$$\begin{bmatrix} 0_{h\times g} & & & \\ & \operatorname{diag}(L_{\epsilon_{g+1}} \dots L_{\epsilon_p}) & & \\ & & \operatorname{diag}({}^tL_{\eta_{h+1}} \dots {}^t L_{\eta_q}) & \\ & & & A_0 + \lambda B_0 \end{bmatrix} \tag{2.13}$$

where the last diagonal block is a regular pencil.

Note: We enumerated the diagonal blocks in the middle from $g+1$ to p and $h+1$ to q respectively just to keep the notation similar to the previous case. Moreover, $A_0 + \lambda B_0$ is actually a square pencil: we may have $g \neq h$ but we must have $m - q = n - p$.

Minimal indices. Let $x_1(\lambda), \dots, x_k(\lambda) \in (K[\lambda])^n$ be solutions of $(A+\lambda B)x(\lambda) = 0$. Such solutions are linearly dependent if the matrix $X(\lambda)$ given by these vectors as columns has rank less than k: in this case there exist $p_1(\lambda), \dots, p_k(\lambda) \in \mathbb{K}[\lambda]$ (not all zero) such that $p_1(\lambda)x_1(\lambda) + \ldots + p_k(\lambda)x_k(\lambda) = 0$. Else if the matrix $X(\lambda)$ has maximum rank k, such solutions are linearly independent.

Let $x_1(\lambda) \neq 0$ be a solution of minimum degree ϵ_1 and let $x_2(\lambda) \neq 0$ be a solution of minimum degree ϵ_2 among the solutions which are linearly independent with $x_1(\lambda)$. By iterating we obtain

$$x_1(\lambda)\ ,\ \dots\ ,\ x_p(\lambda)$$

(where $p \leq n$) with associated degrees $0 < \epsilon_1 \leq \ldots \leq \epsilon_p$: such set of solutions is called **fundamental series of solutions**.

In general such set is not uniquely determined by the pencil $A+\lambda B$ but the sequence of degrees is so. This justifies the following good definition.

Definition. The **minimal indices for columns** of the pencil $A+\lambda B$ are the minima[1] degrees of linearly independent solutions of the equation $(A+\lambda B)x(\lambda) = 0$.

Note: The solutions $x_1(\lambda), \ldots, x_p(\lambda)$ describes linear combinations of the columns of $A+\lambda B$ of degrees $\epsilon_1, \ldots, \epsilon_p$.

Definition. The **minimal indices for rows** of the pencil $A+\lambda B$ are the minima degrees of linearly independent solutions of the equation $({}^tA + \lambda {}^tB)x(\lambda) = 0$.

These minimal indices play a central role in the canonical form of singular pencil for strict equivalence as we realize by the following result.

Proposition 2.3.4. The minimal indices both for columns and rows are an invariant for strict equivalence, i.e. two strictly equivalent pencils have the same minimal indices.

Proof. Let $A+\lambda B$ and $A'+\lambda B'$ be two strictly equivalent pencils and let $P \in \mathrm{GL}_m(\mathbb{K})$ and $Q \in \mathrm{GL}_n(\mathbb{K})$ be such that

$$A' + \lambda B' = P \cdot (A+\lambda B) \cdot {}^tQ$$

Since

$$\begin{aligned}(A+\lambda B)x(\lambda) = 0 &\overset{P,Q\in\mathrm{GL}}{\Longleftrightarrow} P \cdot (A+\lambda B) \cdot {}^tQ \cdot ({}^tQ)^{-1}x(\lambda) = 0 \\ &\Longleftrightarrow (A'+\lambda B')(({}^tQ)^{-1}x(\lambda)) = 0\end{aligned}$$

for each $x_k(\lambda)$ solution of $(A+\lambda B)x(\lambda) = 0$ we have that $y_k(\lambda) = ({}^tQ)^{-1}x_k(\lambda)$ is a solution of $(A'+\lambda B')y(\lambda) = 0$, and conversely.
Moreover, since $({}^tQ)^{-1}$ does not depend on λ, $x_k(\lambda)$ and $y_k(\lambda)$ have the same degree. It follows that $A+\lambda B$ and $A'+\lambda B'$ have the same minimal indices for columns. By the same arguments on transposed matrices it follows that the minimal indices for rows are the same too. $\square$

[1] plural of the latin term *minimum*

Our next goal is to determine the minimal indices of the singular form

$$D(\lambda) = \begin{bmatrix} 0_{h\times g} & & & \\ & \mathrm{diag}(L_{\epsilon_{g+1}} \dots L_{\epsilon_p}) & & \\ & & \mathrm{diag}({}^tL_{\eta_{h+1}} \dots {}^t L_{\eta_q}) & \\ & & & A_0 + \lambda B_0 \end{bmatrix}$$

where $A_0 + \lambda B_0$ is a regular block.

Proposition 2.3.5. The complete system of minimal indices for columns (respectively for rows) of the pencil $D(\lambda)$ is given by the union of the systems of minimal indices for columns (respectively for rows) of each diagonal block.

Proof. We analyze the simpler case of a matrix with two diagonal blocks

$$C(\lambda) = \begin{bmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{bmatrix}$$

Given $x(\lambda)$ a solution with respect to $C(\lambda)$, we can rewrite the vector as

$$x(\lambda) = \begin{bmatrix} x_A(\lambda) \\ x_B(\lambda) \end{bmatrix}$$

so that

$$C(\lambda)x(\lambda) = \begin{bmatrix} A(\lambda)x_A(\lambda) \\ B(\lambda)x_B(\lambda) \end{bmatrix}$$

hence $C(\lambda)x(\lambda) = 0 \iff A(\lambda)x_A(\lambda) = B(\lambda)x_B(\lambda) = 0$.
If $x_A(\lambda)$ (resp. $x_B(\lambda)$) is a solution with respect to $A(\lambda)$ (resp. $B(\lambda)$) of minimum degree ϵ_A (resp. ϵ_B), then

$$x(\lambda) = \begin{bmatrix} x_A(\lambda) \\ 0 \end{bmatrix} \quad \left(\text{resp.} \quad x(\lambda) = \begin{bmatrix} 0 \\ x_B(\lambda) \end{bmatrix} \right)$$

is a solution with respect to $C(\lambda)$, thus the smallest minimal index for columns for $C(\lambda)$ is $\epsilon \le \epsilon_A$ (resp. $\le \epsilon_B$), thus $\epsilon \le \min\{\epsilon_A, \epsilon_B\}$.
On the other end, if $x(\lambda)$ is a solution with respect to $C(\lambda)$ of minimum degree ϵ and

$$x(\lambda) = \begin{bmatrix} x_A(\lambda) \\ x_B(\lambda) \end{bmatrix}$$

we have $\epsilon = \max\{\deg(x_A), \deg(x_B)\}$ and $\deg(x_A) \ge \epsilon_A$, $\deg(x_B) \ge \epsilon_B$, thus $\epsilon \ge \min\{\epsilon_A, \epsilon_B\}$. □

Now it only remains to determine the minimal indices of each block:

- Obviously the regular block $A_0 + \lambda B_0$ has no minimal index since the pencil is regular;
- Equally obviously the zero block $0_{h\times g}$ has g zero minimal indices for columns and h zero minimal indices for rows

$$\epsilon_1 = \ldots = \epsilon_g = \eta_1 = \ldots = \eta_h = 0$$

- For all $i = g+1 : p$ the blocks

$$L_{\epsilon_i} = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \end{bmatrix}$$

of size $\epsilon_i \times (\epsilon_i + 1)$ have linearly independent rows, thus there is just one minimal index for columns that is exactly ϵ_i.

- For $j = h+1 : q$ the blocks

$${}^tL_{\eta_j} = \begin{bmatrix} \lambda & & \\ 1 & \ddots & \\ & \ddots & \lambda \\ & & 1 \end{bmatrix}$$

of size $(\eta_j + 1) \times \eta_j$ have linearly independent columns, thus there is just one minimal index for rows that is exactly η_j.

Theorem 2.3.6. The pencil

$$D(\lambda) = \begin{bmatrix} 0_{h\times g} & & & \\ & \operatorname{diag}(L_{\epsilon_{g+1}} \ldots L_{\epsilon_p}) & & \\ & & \operatorname{diag}({}^tL_{\eta_{h+1}} \ldots {}^t L_{\eta_q}) & \\ & & & A_0 + \lambda B_0 \end{bmatrix}$$

(where the block $A_0 + \lambda B_0$ is regular) has minimal indices for columns $\epsilon_1 = \ldots = \epsilon_g = 0\ ,\ \epsilon_{g+1}, \ldots, \epsilon_p$ and minimal indices for rows $\eta_1 = \ldots = \eta_h = 0\ ,\ \eta_{h+1}, \ldots, \eta_q$.

We recall that the elementary divisors (both finite and infinite ones) are a complete system of invariants of strict equivalence for regular pencils. Moreover we have just found out that minimal indices (for rows and columns) are a complete system of invariants of strict equivalence for the singular block of singular pencils (i.e. the diagonal blocks except the regular one). Thus we have to expect that the minimal indices together with the elementary divisors give a complete system of invariants of strict equivalence for singular pencils. By fact 1.1.3: the elementary divisors of a diagonal-block matrix are given by the union of the ones of each block.

Remark 2.3.7. It is worth resuming their relations with the size of a pencil of size $m \times n$. Let p and q be the number of the minimal indices for columns and rows respectively: then

$$\begin{cases} m - q = n - p = \sum_{i=1}^{p} \epsilon_i + \sum_{j=1}^{q} \eta_j \\ p \leq n \\ q \leq m \end{cases} \tag{2.14}$$

where the last two inequalities hold since there are at most p (respectively q) independent solutions in $\mathbb{K}[\lambda]^n$ (respectively $\mathbb{K}[\lambda]^m$).

Remark 2.3.8. The blocks L_{ϵ_i} and ${}^tL_{\eta_j}$ have no elementary divisor: indeed they both have a minor of size ϵ_i and η_j respectively equal to 1 and a minor of the same size equal to λ^{ϵ_i} and λ^{η_j} respectively. It follows that the set of elementary divisors of the pencil $D(\lambda)$ is given by the set of elementary divisors of the regular block $A_0 + \lambda B_0$.

As for regular pencils, we want to reformulate what done so far in terms of projective pencils. Let $\mu A + \lambda B$ a singular projective pencil of size $m \times n$. In the regular case we noticed that we just need to homogenize the affine canonical form to obtain a projective one and the same arguments still hold in the singular case. Consider the homogenized blocks

$$\begin{gathered} R_{\epsilon_i} = \begin{bmatrix} \lambda & \mu & & \\ & \ddots & \ddots & \\ & & \lambda & \mu \end{bmatrix} \quad , \quad {}^tR_{\eta_j} = \begin{bmatrix} \lambda & & \\ \mu & \ddots & \\ & \ddots & \lambda \\ & & \mu \end{bmatrix} \\ H^{(u_i)} = \begin{bmatrix} \mu & \lambda & & \\ & \ddots & \ddots & \\ & & \ddots & \lambda \\ & & & \mu \end{bmatrix} \quad , \quad \mathfrak{F}_{(p_k^{(j)})^{w_k}} = \begin{bmatrix} \lambda & \mu & & & \\ & & \ddots & \ddots & \\ & & & \lambda & \mu \\ \mu * & \mu * & \dots & \mu * & \lambda - \mu * \end{bmatrix} \end{gathered} \tag{2.15}$$

where the $*$'s are for suitable coefficients.
The same arguments we have done so far prove the following theorems.

Theorem 2.3.9 (Kronecker)**.** Every projective pencil $\mu A + \lambda B$ is strictly equivalent to a canonical pencil of the form

$$\begin{bmatrix} 0_{h \times g} & & & & \\ & \operatorname{diag}\left(\{R_{\epsilon_i}\}_{i=g+1}^{p}\right) & & & \\ & & \operatorname{diag}\left(\{{}^tR_{\eta_j}\}_{j=h+1}^{q}\right) & & \\ & & & \operatorname{diag}\left(\{H^{(u_k)}\}_{k=1}^{s}\right) & \\ & & & & \operatorname{diag}\left(\left\{\mathfrak{F}_{(p_l^{(z)})^{w_l}}\right\}_{z,l}\right) \end{bmatrix} \tag{2.16}$$

where the blocks $0_{h\times g}$, R_{ϵ_i} and ${}^tR_{\eta_j}$ are uniquely determined by the minimal indices (for columns and rows) $\epsilon_1\dots\epsilon_p,\eta_1\dots\eta_q$, the blocks $H^{(u_k)}$ are uniquely determined by the elementary divisors $\mu^{u_1}\dots\mu^{u_s}$ and the Frobenius blocks $\mathfrak{F}_{(p_l^{(z)})^{w_l}}$ are uniquely determined by the other elementary divisors.

Theorem 2.3.10 (Weierstrass)**.** Every projective pencil $\mu A+\lambda B$ is strictly equivalent to the canonical block-direct-sum

$$0_{h\times g}\boxplus\left(\boxplus_{i=g+1}^{p} R_{\epsilon_i}\right)\boxplus\left(\boxplus_{j=h+1}^{q}{}^tR_{\eta_j}\right)\boxplus\left(\boxplus_{k=1}^{s} H^{(u_k)}\right)\boxplus\left(\boxplus_{l,z}\mathfrak{F}_{(p_z^{(l)})^{w_z}}\right) \tag{2.17}$$

where the blocks $0_{h\times g}$, R_{ϵ_i} and ${}^tR_{\eta_j}$ are uniquely determined by the minimal indices $\epsilon_1\dots\epsilon_p,\eta_1\dots\eta_q$, the blocks $H^{(u_k)}$ are uniquely determined by the elementary divisors $\mu^{u_1}\dots\mu^{u_s}$ and the Frobenius blocks $\mathfrak{F}_{(p_z^{(l)})^{w_z}}$ are uniquely determined by the other elementary divisors $(p_z^{(l)}(\mu,\lambda))^{w_z}$.

Remark 2.3.11. From the beginning we have worked (and we will continue to work) over an infinite field $\mathbb{K}$. However the above classification holds over finite fields too (we refer to the work of Mirwald in its PhD thesis [29]).

It is worth remarking how we can rewrite the Kronecker and Weierstrass forms in (2.16) and (2.17) when the field $\mathbb{K}$ is algebraically closed (e.g. $\mathbb{K}=\mathbb{C}$). We recall that $\mathfrak{J}_{w,a}$ is the Jordan block with respect to the elementary divisor $(\lambda+a\mu)^w$.

Theorem 2.3.12 (Kronecker)**.** Let $\mathbb{K}=\overline{\mathbb{K}}$. In the same notations as above, every projective pencil $\mu A+\lambda B$ is strictly equivalent to a canonical pencil of the form

$$\begin{bmatrix} 0_{h\times g} &&&& \\ & \operatorname{diag}\left(\{R_{\epsilon_i}\}_{i=g+1}^{p}\right) &&& \\ && \operatorname{diag}\left(\{{}^tR_{\eta_j}\}_{j=h+1}^{q}\right) && \\ &&& \operatorname{diag}\left(\{H^{(u_k)}\}_{k=1}^{s}\right) & \\ &&&& \operatorname{diag}\left(\left\{\mathfrak{J}_{w_{lz},a_{lz}}\right\}_{z,l}\right)\end{bmatrix} \tag{2.18}$$

or equivalently to a canonical block-direct-sum of the form

$$0_{h\times g}\boxplus\left(\boxplus_{i=g+1}^{p} R_{\epsilon_i}\right)\boxplus\left(\boxplus_{j=h+1}^{q}{}^tR_{\eta_j}\right)\boxplus\left(\boxplus_{k=1}^{s} H^{(u_k)}\right)\boxplus\left(\boxplus_{l,z}\mathfrak{J}_{w_{lz},a_{lz}}\right) \tag{2.19}$$

This concludes our research for canonical representatives of the classes in

$$\mathrm{GL}_m(\mathbb{K})\backslash\mathfrak{M}_{m\times n}(\mathbb{K}[\mu,\lambda]_1)/\mathrm{GL}_n(\mathbb{K})$$

Example 2.3.13. Let $\mu A+\lambda B$ be a projective pencil with minimal indices $\epsilon_1 = 0, \epsilon_2 = 1, \epsilon_3 = 2, \eta_1 = 0, \eta_2 = 0, \eta_3 = 2$ and elementary divisors $\mu^3, (\lambda + \mu)^2$. Then the pencil is strictly equivalent to the Kronecker form

$$\begin{bmatrix}
0 & & & & & & & & & & & & \\
0 & & & & & & & & & & & & \\
 & \lambda & \mu & & & & & & & & & & \\
 & & & \lambda & \mu & 0 & & & & & & & \\
 & & & 0 & \lambda & \mu & & & & & & & \\
 & & & & & & \lambda & 0 & & & & & \\
 & & & & & & \mu & \lambda & & & & & \\
 & & & & & & 0 & \mu & & & & & \\
 & & & & & & & & \mu & \lambda & 0 & & \\
 & & & & & & & & 0 & \mu & \lambda & & \\
 & & & & & & & & 0 & 0 & \mu & & \\
 & & & & & & & & & & & \lambda+\mu & \mu \\
 & & & & & & & & & & & 0 & \lambda+\mu
\end{bmatrix}$$

Next consider the pencil $\mu C + \lambda D$ with minimal indices $\epsilon_1 = 0, \epsilon_2 = 0, \epsilon_3 = 2, \eta_1 = 0, \eta_2 = 1, \eta_3 = 2$ and elementary divisors $\mu^2, (\lambda + \mu)^2, \lambda + \mu$. Then its Kronecker canonical form is

$$\begin{bmatrix}
0 & 0 & & & & & & & & & & & \\
 & & \lambda & \mu & 0 & & & & & & & & \\
 & & 0 & \lambda & \mu & & & & & & & & \\
 & & & & & \lambda & & & & & & & \\
 & & & & & \mu & & & & & & & \\
 & & & & & & \lambda & 0 & & & & & \\
 & & & & & & \mu & \lambda & & & & & \\
 & & & & & & 0 & \mu & & & & & \\
 & & & & & & & & \mu & \lambda & & & \\
 & & & & & & & & 0 & \mu & & & \\
 & & & & & & & & & & \lambda+\mu & & \\
 & & & & & & & & & & & \lambda+\mu & \mu \\
 & & & & & & & & & & & 0 & \lambda+\mu
\end{bmatrix}$$

By comparing the pencils $\mu A + \lambda B$ and $\mu C + \lambda D$ we focus on the elementary divisors of the form $(\lambda+\mu)^k$: in the first pencil we have just one elementary divisor of this form and indeed there is just one corresponding block of size $k = 2$; instead in the second pencil there are two elementary divisors which are powers of $\lambda+\mu$, of multiplicity 1 and 2 respectively, and indeed we have two corresponding blocks of size 1 and 2 respectively. This underlines that in the elementary divisors there is all the necessary information to reconstruct the regular block of the canonical form. Same arguments hold for the singular part but with minimal indices instead of elementary divisors.

Remark 2.3.14. Given a projective pencil $\mu A + \lambda B$, to compute its Kronecker form (or at least its Kronecker invariants) we need to:

- compute the greatest common divisor of all the minor $k \times k$ and compute their consecutive quotients: these are the invariant polynomials of the pencil;
- factorize the invariant polynomials: each factor (counted with multiplicity) is an elementary divisor and it uniquely determines a block of size its degreee;
- compute the dimension of $\ker A \cap \ker B$ and $\ker({}^tA) \cap \ker({}^tB)$ to obtain the sizes g and h respectively of the zero-block;
- compute the minima[2] degrees of linearly independent vectors in $\ker(\mu A + \lambda B) \setminus (\ker A \cap \ker B)$ and in $\ker(\mu {}^tA + \lambda {}^tB) \setminus (\ker({}^tA) \cap \ker({}^tB))$ to obtain the minimal indices $\epsilon_{g+1}, \dots, \epsilon_p$ and $\eta_{h+1}, \dots, \eta_q$ respectively.

In Chapter 6 we exhibit the implementations on `Macaulay2` for the above steps.

2.4 Symmetric matrix pencils

Symmetric matrix pencils have a geometric interest: they describe pencils of projective quadrics (we will better investigate this aspect in the next chapter). Consider two quadrics $\mathcal{A}$ and $\mathcal{B}$ in $\mathbb{P}^{m-1}_{\mathbb{K}}$ described by the equations

$$\mathcal{A}:\ A(X,X) = \sum_{i,j=1}^{m} a_{ij} X_i X_j \quad \text{and} \quad \mathcal{B}:\ B(X,X) = \sum_{i,j=1}^{m} b_{ij} X_i X_j$$

Let $A = (a_{ij})_{i,j}$ and $B = (b_{ij})_{i,j}$ be the *symmetric* matrices associated to the equations of $\mathcal{A}$ and $\mathcal{B}$ respectively: they generate a pencil of square matrices $\mu A + \lambda B$ which describes a pencil of quadratic forms of equations $\mu A(X,X) + \lambda B(X,X)$. We remark that a square matrix pencil is *symmetric* if it is defined by two symmetric matrices and we denote the set of symmetric pencils of size m over $\mathbb{K}$ by

$$\operatorname{Sym}^2 \mathbb{K}^m[\mu, \lambda]_1$$

Now consider $X = TZ$ a non-singular linear transformation T (with $\det T \neq 0$) of the projective variables $X_1, \dots, X_m$ into the new variable $Z_1, \dots, Z_m$: we obtain two quadrics $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{B}}$ described by the *congruent* matrices $\tilde{A} = {}^tTAT$ and $\tilde{B} = {}^tTBT$ respectively, hence we have a pencil

$$\mu\tilde{A} + \lambda\tilde{B} = {}^tT(\mu A + \lambda B)T$$

As for matrices, we say that the two above pencils are **congruent** and we denote it by

$$\mu A + \lambda B \equiv \mu\tilde{A} + \lambda\tilde{B}$$

[2] plural of the latin term *minimum*

Obviously *congruence* is a particular case of strict equivalence where the acting matrices are a non-singular matrix T and its transpose ${}^tT$: indeed the transformation T is constant (i.e. independent from μ, λ). Thus we may reformulate the above definition of congruence in terms of the group action

$$\begin{array}{ccc} \mathrm{GL}_m(\mathbb{K}) & \longrightarrow & \mathrm{Aut}\left(\mathrm{Sym}^2\,\mathbb{K}^m[\mu,\lambda]_1\right) \\ T & \mapsto & \left(\,\mu A + \lambda B \mapsto \mu({}^tTAT) + \lambda({}^tTBT)\,\right) \end{array}$$

Definition. Two symmetric pencils are **congruent** if they are in the same orbit with respect to the above group action.

In general the converse does not hold, but it does over the complex field $\mathbb{C}$: to prove so we need the following lemma.

Lemma 2.4.1. Let $U \in \mathfrak{M}_m(\mathbb{C})$ be a non-singular complex matrix ($\det U \neq 0$). Then there exists a polynomial $f(t) \in \mathbb{C}[t]$ such that

$$f(U)^2 = U$$

Proof. Let $m_U(t)$ be the minimal polynomial of U: we recall it vanishes only for all the eigenvalues of U.
The condition the polynomial $f(t)$ we are looking for must satisfy is equivalent to the condition $m_U(t) \mid f(t)^2 - t$, hence $f(\lambda)^2 - \lambda = 0$ for all $\lambda \in \mathbb{C}$ eigenvalue of U. In particular we can look for a polynomial $f(t)$ such that

$$\forall \lambda \in \mathrm{Eigval}(U), \quad f(\lambda) = \sqrt{\lambda}$$

Let $\mathrm{Eigval}(U) = \{\lambda_1, \dots \lambda_s\}$ and for all $i = 1 : s$ let k_i be the multiplicity of λ_i as root of the characteristic polynomial of U (i.e. its algebraic multiplicity). We put $f(t)$ the polynomial interpolating the function $g(t) = \sqrt{t}$ and its derivatives until the $(k_i - 1)$-th in the points λ_i, i.e.

$$f(\lambda_i) = g(\lambda_i) \quad , \quad f^{(j)}(\lambda_i) = g^{(j)}(\lambda_i) \quad \forall i = 1 : s \ , \ \forall j = 0 : k_i - 1$$

We note that we are allowed to consider it since the only singularity of the function is in 0 but $0 \notin \mathrm{Eigval}(U)$ since U is non-singular. Thus such $f(t)$ satisfies $m_U(t) \mid f(t)^2 - t$ and this concludes the proof. □

Theorem 2.4.2. Two symmetric complex pencils are strictly equivalent if and only if they are congruent:

$$\mathcal{P}_1 \overset{\circ}{\sim} \mathcal{P}_2 \iff \mathcal{P}_1 \equiv \mathcal{P}_2$$

Proof. The "only if" implication is obvious since the matrix which defines the congruence is constant. Let us see the "if".
Let $\mathcal{P}_1 = \mu A + \lambda B$ and $\mathcal{P}_2 = \mu\tilde{A} + \lambda\tilde{B}$ be two strictly equivalent symmetric pencils of size m: then there exist two constant non-singular matrices $P, Q \in \mathrm{GL}(m)$ such that $\mathcal{P}_2 = P \cdot \mathcal{P}_1 \cdot {}^tQ$. By transposing the whole equation we obtain

$$\mathcal{P}_2 = {}^t\mathcal{P}_2 = Q \cdot {}^t\mathcal{P}_1 \cdot {}^tP = Q \cdot \mathcal{P}_1 \cdot {}^tP$$

Since ${}^tP$ and Q are non-singular too, we obtain

$$P \cdot \mathcal{P}_1 \cdot {}^tQ = Q \cdot \mathcal{P}_1 \cdot {}^tP \implies \mathcal{P}_1 \cdot ({}^tQ) \cdot ({}^tP)^{-1} = P^{-1} \cdot Q \cdot \mathcal{P}_1$$

We set $U = ({}^tQ) \cdot ({}^tP)^{-1}$ (which is still non-singular) and we have $\mathcal{P}_1 \cdot U = {}^tU \cdot \mathcal{P}_1$. This equality extends to every polynomial in U: given $f(t)$ a polynomial, it holds

$$\mathcal{P}_1 \cdot f(U) = {}^tf(U) \cdot \mathcal{P}_1$$

Let $f(t)$ be such that $\det(f(U)) \neq 0$ (at least the polynomial $f(t) = t$ satisfies it since U is non-singular): then

$$\mathcal{P}_1 = {}^tf(U) \cdot \mathcal{P}_1 \cdot f(U)^{-1} \implies \mathcal{P}_2 = P \cdot \mathcal{P}_1 \cdot {}^tQ = P \cdot \left({}^tf(U) \cdot \mathcal{P}_1 \cdot f(U)^{-1} \right) \cdot {}^tQ$$

The last equality would be a congruence transformation between $\mathcal{P}_1$ and $\mathcal{P}_2$ if

$${}^t\!\left(P \cdot {}^tf(U)\right) = f(U)^{-1} \cdot {}^tQ \iff f(U)^2 = ({}^tQ) \cdot ({}^tP)^{-1} = U$$

But from the previous lemma we can choose $f(t)$ such that the last equality holds. Thus the matrix $T = f(U)^{-1} \cdot {}^tQ$ gives the congruence

$$\mathcal{P}_2 = {}^tT \cdot \mathcal{P}_1 \cdot T \equiv \mathcal{P}_1$$

□

Corollary 2.4.3. Two pencils of complex quadratic forms $\mu\mathcal{A} + \lambda\mathcal{B}$ and $\mu\tilde{\mathcal{A}} + \lambda\tilde{\mathcal{B}}$ can be carried into one another by a non-singular transformation $X = TZ$ if and only if the symmetric complex pencils $\mu A + \lambda B$ and $\mu\tilde{A} + \lambda\tilde{B}$ describing them have the same minimal indices and the same elementary divisors.

Remark 2.4.4. In a symmetric pencil the minimal indices for columns and rows coincide, i.e. $p = q$ and $\epsilon_1 = \eta_1, \ldots, \epsilon_p = \eta_q$, since by transposing the pencil it remains the same. Hence from now on we will talk about minimal indices without distinction between rows and columns and we will denote them $\epsilon_1 = \ldots = \epsilon_g = 0, \epsilon_{g+1}, \ldots, \epsilon_p$.

2.4.1 The symmetric Kronecker form over $\mathbb{C}$

Let us fix $\mathbb{K} = \mathbb{C}$. We are now interested in a canonical form of strict equivalence (or equivalently, of congruence) which is symmetric too since the canonical form we referred to until now is not so. Let W_k be the *inverse diagonal* matrix[3] of size k

$$W_k = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}$$

We will use this matrix to *symmetrize* the blocks in the Kronecker form.
Let $\mu A + \lambda B$ be a complex symmetric pencil with minimal indices $\epsilon_1 = \ldots = \epsilon_g = 0, \epsilon_{g+1}, \ldots, \epsilon_p$ and elementary divisors $\mu^{u_1}, \ldots, \mu^{u_s}, (\lambda + \mu a_1)^{c_1}, \ldots, (\lambda + \mu a_t)^{c_t}$.

- For all $i = g+1 : p$ we join the singular blocks R_{ϵ_i} and ${}^tR_{\epsilon_i}$ in the symmetric singular block of size $2\epsilon_i + 1$

$$S_{\epsilon_i} = \begin{bmatrix} 0 & R_{\epsilon_i} \\ {}^tR_{\epsilon_i} & 0 \end{bmatrix} \tag{2.20}$$

  uniquely determined by the minimal indices.

- For all $j = 1 : s$ we right-multiply the regular block $H^{(u_j)}$ by the block W_{u_j} and we obtain the symmetric regular block of size u_j

$$K^{(u_j)} = H^{(u_j)} W_{u_j} = \begin{bmatrix} & & & \lambda & \mu \\ & & \iddots & \iddots & \\ & & & & \\ \lambda & \iddots & & & \\ \mu & & & & \end{bmatrix}$$

  uniquely determined by the elementary divisor μ^{u_j}.

- For all $k = 1 : t$ we right-multiply the regular Jordan block $\mathfrak{J}_{c_k,a_k}$ by the block W_{c_k} and we obtain the symmetric regular block of size c_k

$$\check{\mathfrak{J}}_{c_k,a_k} = \mathfrak{J}_{c_k,a_k} W_{c_k} = \begin{bmatrix} & & & \mu & \lambda + a_k\mu \\ & & \iddots & \iddots & \\ & & & & \\ \mu & \iddots & & & \\ \lambda + a_k\mu & & & & \end{bmatrix} \tag{2.21}$$

  uniquely determined by the elementary divisor $(\lambda + a_k\mu)^{c_k}$.

[3] the only non zero elements are the ones on the inverse diagonal which are 1

By joining the singular blocks in one we are just re-ordering the basis, hence we are acting by coniugacy and this preserves the strict equivalence classes. The right-multiplication by W_k preserves the strict equivalence classes too. Thus the symmetric matrix we obtain is strictly equivalent to the pencil we are working with.

Theorem 2.4.5 (Kronecker - complex symmetric form)**.** Every symmetric complex pencil $\mu A + \lambda B$ is strictly equivalent to a canonical symmetric complex pencil of the form

$$\begin{bmatrix} 0_g & & & \\ & \operatorname{diag}(S_{\epsilon_{g+1}} \dots S_{\epsilon_p}) & & \\ & & \operatorname{diag}(K^{(u_1)}, \dots, K^{(u_s)}) & \\ & & & \operatorname{diag}(\check{\mathfrak{J}}_{c_1,a_1}, \dots, \check{\mathfrak{J}}_{c_t,a_t}) \end{bmatrix} \tag{2.22}$$

where the blocks 0_g and S_{ϵ_i} are uniquely determined by the minimal indices ϵ_i, the blocks $K^{(u_j)}$ are uniquely determined by the elementary divisors of the form μ^{u_j} and the blocks $\check{\mathfrak{J}}_{c_k,a_k}$ are uniquely determined by the elementary divisors of the form $(\lambda+\mu a_k)^{c_k}$.

2.4.2 The symmetric Kronecker form over $\mathbb{R}$

Let us now fix $\mathbb{K} = \mathbb{R}$. Obviously the symmetric canonical form we found out over the complex field does not hold anymore since we may have elementary divisors which are powers of quadratic polynomials, hence we need other regular symmetric blocks than the Jordan ones.

Remark 2.4.6. It is worth noting that, given $\mu A+\lambda B$ a symmetric pencil over $\mathbb{R}$ such that B is definite positive, then all roots of $\det(\mu A+\lambda B)$ are real [15, Ch.IX, §13], that is all the elementary divisors are powers of linear real polinomials and we just need the Jordan blocks (as in the complex case).

We recall that for general pencils (not necessarily symmetric) these blocks are the Frobenius blocks with respect to suitable polynomials (so that their determinants are exactly the elementary divisors). Clearly we can not symmetrize a Frobenius block $\mathfrak{F}_{(p_k^{(j)})^{w_k}}$ by the inverse diagonal matrix $W_{w_k \cdot \deg(p_k^{(j)})}$.
Since the characteristic polynomial of a real symmetric matrix is completely factorisable (i.e. its eigenvalues are all real), one at first may hope that even the determinant of a real symmetric pencil is completely factorisable, but unfortunately this does not hold. An elementary counterxample is given by the real symmetric regular pencil

$$\mu \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

whose determinant is $-(\lambda^2 + \mu^2)$ which is irreducible over $\mathbb{R}$. To solve this we resort to the real Jordan form of a matrix.

First of all we note that the only blocks we still need to fix are the ones related to elementary divisors of the form $(\lambda^2 + a_k\mu\lambda + b_k\mu)^{w_k}$ where $a_k, b_k \in \mathbb{R}$, since all the other blocks are symmetrizable just as in the complex case. Thus we restrict ourselves to study the case of a real symmetric regular pencil with one only elementary divisor.

Let $\mu R + \lambda S$ be a real symmetric regular pencil with only elementary divisor $(\lambda^2 + a_k\mu\lambda + b_k\mu)^{h_k}$ and let $\lambda^2 + a_k\mu\lambda + b_k\mu = (\lambda - \alpha_k\mu)(\lambda - \overline{\alpha}_k\mu)$ be the factorization over the algebraic closure $\mathbb{C}$. We recall that the real Jordan form with respect to a complex eigenvalue $\alpha \in \mathbb{C}$ of multiplicity h_α is the $2h_\alpha$ block-matrix

$$J^{\mathbb{R}}_{\alpha,h_\alpha} = \begin{bmatrix} \Re(\alpha) & \Im(\alpha) & 1 & 0 & & & & \\ -\Im(\alpha) & \Re(\alpha) & 0 & 1 & & & & \\ & & & \ddots & \ddots & & & \\ & & & & \ddots & & 1 & 0 \\ & & & & & & 0 & 1 \\ & & & & & & \Re(\alpha) & \Im(\alpha) \\ & & & & & & -\Im(\alpha) & \Re(\alpha) \end{bmatrix}$$

with h_α diagonal blocks of size 2.
We define $\mathfrak{R}_{\alpha,h_\alpha}$ the *real Jordan block* of size $2h_\alpha$ with respect to $((\lambda - \alpha\mu)(\lambda - \overline{\alpha}\mu))^{h_\alpha}$ as the pencil of size $2h_\alpha$

$$\mathfrak{R}_{\alpha,h_\alpha} = \mu J^{\mathbb{R}}_{\alpha} - \lambda I = \begin{bmatrix} \Re(\alpha)\mu - \lambda & \Im(\alpha)\mu & \mu & 0 & & & & \\ -\Im(\alpha)\mu & \Re(\alpha)\mu - \lambda & 0 & \mu & & & & \\ & & & \ddots & \ddots & & & \\ & & & & \ddots & & \mu & 0 \\ & & & & & & 0 & \mu \\ & & & & & & \Re(\alpha)\mu - \lambda & \Im(\alpha)\mu \\ & & & & & & -\Im(\alpha)\mu & \Re(\alpha)\mu - \lambda \end{bmatrix}$$

Note: A simple count shows that the pencil $\mathfrak{R}_{\alpha,h_\alpha}$ is regular with only elementary divisor $((\lambda - \alpha\mu)(\lambda - \overline{\alpha}\mu))^{h_\alpha}$, hence it is strictly equivalent to the Frobenius block defined by the same elementary divisor.

The advantage we take from the above construction is that we can symmetrize the real Jordan block by right-multiplation by the inverse diagonal: thus we obtain the symmetric real regular block of size $2h_\alpha$

$$\check{\mathfrak{R}}_{\alpha,h_\alpha} = \mathfrak{R}_{\alpha,h_\alpha} W_{2h_\alpha} = \begin{bmatrix} & & & & 0 & \mu & \Im(\alpha)\mu & \Re(\alpha)\mu - \lambda \\ & & & & \mu & 0 & \Re(\alpha)\mu - \lambda & -\Im(\alpha)\mu \\ & & & \iddots & \iddots & & & \\ & 0 & \mu & \iddots & & & & \\ & \mu & 0 & & & & & \\ \Im(\alpha)\mu & & \Re(\alpha)\mu - \lambda & & & & & \\ \Re(\alpha)\mu - \lambda & & -\Im(\alpha)\mu & & & & & \end{bmatrix}$$

uniquely determined by the elementary divisor $((\lambda - \alpha\mu)(\lambda - \overline{\alpha}\mu))^{h_\alpha}$.

Theorem 2.4.7 (Kronecker - real symmetric form)**.** Every real symmetric pencil $\mu A + \lambda B$ is strictly equivalent to a canonical real symmetric pencil of the form

$$\begin{bmatrix} 0_g & & & & \\ & \mathrm{diag}(\{S_{\epsilon_i}\}_{i=g+1}^{p}) & & & \\ & & \mathrm{diag}(\{K^{(u_j)}\}_{j=1}^{s}) & & \\ & & & \mathrm{diag}(\{\check{\mathfrak{J}}_{c_k,a_k}\}_{k=1}^{t}) & \\ & & & & \mathrm{diag}(\{\check{\mathfrak{R}}_{\alpha_f,h_f}\}_{f=1}^{l}) \end{bmatrix} \quad (2.23)$$

where the blocks 0_g and S_{ϵ_i} are uniquely determined by the minimal indices ϵ_i, the blocks $K^{(u_j)}$ are uniquely determined by the elementary divisors of the form μ^{u_j}, the blocks $\check{\mathfrak{J}}_{c_k,a_k}$ are uniquely determined by the completely factorisable elementary divisors of the form $(\lambda+\mu a_k)^{c_k}$ and the blcoks $\check{\mathfrak{R}}_{\alpha_f,h_f}$ are uniquely determined by the elementary divisors of the form $((\lambda - \alpha_f\mu)(\lambda - \overline{\alpha}_f\mu))^{h_f}$.

Chapter 3

Classification of pencils of projective quadrics

In this chapter we study two applications of the Kronecker form for symmetric pencils to complex quadratic forms: the simultaneous reduction of complex quadratic forms and the Segre classification of intersections of two projective quadrics. In the latter we define two new invariants for symmetric pencils (the Segre symbol *and the continuous moduli) and give the complete classification in $\mathbb{P}^2_{\mathbb{C}}$ and $\mathbb{P}^3_{\mathbb{C}}$. In the last section we give a geometric interpretation of the Kronecker form in terms of projective lines and singularity of the base loci via projective bundles.*

In this chapter we work over $\mathbb{K} = \mathbb{C}$.

3.1 Simultaneous reduction of complex quadratic forms

Our goal in this section is to answer to the following question: *given two complex quadratic forms*

$$A(X,X) = \sum_{i,j=1}^{m} a_{ij} X_i X_j \quad , \quad B(X,X) = \sum_{i,j=1}^{m} b_{ij} X_i X_j$$

under which conditions can they be reduced simultaneously *to sums of squares*

$$\tilde{A}(X,X) = \sum_{i=1}^{m} \tilde{a}_i Z_i^2 \quad , \quad \tilde{B}(X,X) = \sum_{i=1}^{m} \tilde{b}_i Z_i^2$$

by a non singular transformation $X = TZ$?

We recall that two complex quadratic forms define a complex symmetric pencil. To ask for two complex quadratic forms to be simultaneously reducible to other two is to ask for the related complex symmetric pencils to be strictly equivalent. Thus to answer the question it is enough to study the Kronecker form of the pencil defined by two complex quadratic forms which are sums of squares.

Consider the complex quadratic forms

$$\tilde{A}(X,X) = \sum_{i=1}^{m} \tilde{a}_i Z_i^2 \quad , \quad \tilde{B}(X,X) = \sum_{i=1}^{m} \tilde{b}_i Z_i^2$$

and let $\tilde{\mathcal{P}} = \mu\tilde{A} + \lambda\tilde{B}$ the complex symmetric pencil they define. Let $r_A = \operatorname{Rk}\tilde{A}$ and $r_B = \operatorname{Rk}\tilde{B}$: since $\tilde{A}$ and $\tilde{B}$ are diagonal matrices, the rank of the pencil is $r \max\{r_A, r_B\}$. Up to a basis re-ordering we may assume $\tilde{a}_{r_A+1} = \ldots = \tilde{a}_m = \tilde{b}_{r_B+1} = \ldots = \tilde{b}_m = 0$. Thus the pencil is of the form

$$\tilde{\mathcal{P}} = \operatorname{diag}(\mu\tilde{a}_1 + \lambda\tilde{b}_1, \ldots, \mu\tilde{a}_r + \lambda\tilde{b}_r, 0, \ldots, 0)$$

Clearly the above pencil has minimal indices identically zero $\epsilon_1 = \ldots = \epsilon_{m-r} = 0$ and linear elementary divisors $\mu\tilde{a}_1 + \lambda\tilde{b}_1, \ldots, \mu\tilde{a}_r + \lambda\tilde{b}_r$: we remark that we are not assuming $\tilde{a}_i$ (resp. $\tilde{b}_j$) to be distinct but the elementary divisors to have multiplicity 1.
Thus the following theorem is just a corollary and it completely answers to our question.

Theorem 3.1.1. Two complex quadratic forms $A(X,X)$ and $B(X,X)$ can be reduced simultaneously to sums of squares by a non singular transformation of the variables if and only if the pencil $\mu A + \lambda B$ they describe has identically-zero minimal indices and linear elementary divisors.

Corollary 3.1.2. Let $\mathcal{P} = \mu A + \lambda B$ be a complex symmetric pencil of size m and let r_A, r_B be the ranks of A and B respectively. If $\mathcal{P}$ has identically zero minimal indices and linear elementary divisors, then

$$\mathcal{P} \equiv \mu \begin{bmatrix} a_1 & & & \\ & \ddots & & \\ & & a_{r_A} & \\ & & & 0 \end{bmatrix} + \lambda \begin{bmatrix} I_{r_B} & \\ & 0 \end{bmatrix}$$

3.2 Segre classification of intersections of two quadrics

Let $X_0, \ldots, X_m$ be coordinates for $\mathbb{C}^{m+1}$. The quadratic forms

$$A(X,X) = \sum_{i,j=0}^{m} a_{ij} X_i X_j \quad , \quad B(X,X) = \sum_{i,j=0}^{m} b_{ij} X_i X_j$$

(where $A = (a_{ij})$ and $B = (b_{ij})$ are symmetric matrices) define two quadrics $\mathcal{A}, \mathcal{B}$ in $\mathbb{P}^m_{\mathbb{C}}$. As in the previous section we may consider the matrix pencil $\mathcal{P} = \mu A + \lambda B$ which describes a pencil of quadrics $\mathcal{P}$ in $\mathbb{P}^m_{\mathbb{C}}$. For each elementary divisor $(x_i\mu + y_i\lambda)^{e^i_j}$ we may consider its root $[y_i : -x_i] \in \mathbb{P}^1$ with multiplicity e^i_j. We refer to these roots as to the **roots of the pencil**.

Definition. Let $\mathcal{P}$ be a pencil of quadrics with minimal indices $\epsilon_1 = \ldots = \epsilon_g = 0$, $\epsilon_{g+1}, \ldots, \epsilon_p$ and elementary divisors $\{(x_i\mu+y_i\lambda)^{e^i_1}, \ldots, (x_i\mu+y_i\lambda)^{e^i_{r_i}} \mid i = 1 : k\}$, where the k roots $[y_i : -x_i]$ are all distinct. We define its **Segre symbol** to be the ordered sequence of its invariants

$$\Sigma(\mathcal{P}) = \Big[(e^1_1, \ldots, e^1_{r_1}) \ldots (e^k_1, \ldots, e^k_{r_k}); \epsilon_{g+1}, \ldots, \epsilon_p; g\Big] \tag{3.1}$$

with the ordering

$$r_1 \geq \ldots \geq r_k \quad , \quad e^i_1 \geq \ldots \geq e^i_{r_i} \quad , \quad \epsilon_{g+1} \leq \ldots \leq \epsilon_p$$

In particular, we put *semicolons* when we pass from the (multiplicities of the) roots to the non-zero minimal indices and from the latter to the number of the zero minimal indices; the *round brackets* distinguish the multiplicities of different roots and they are omitted if $r_i = 1$.

Example 3.2.1.

- The pencil $\begin{bmatrix} \lambda & & \\ & \mu & \\ & & 0 \end{bmatrix}$ has Segre symbol [1 1; ; 1].
- The pencil $\begin{bmatrix} \mu & \lambda & \\ \lambda & & \\ & & 0 \end{bmatrix}$ has Segre symbol [2; ; 1].
- The pencil $\begin{bmatrix} \mu & \lambda & \\ \lambda & & \\ & & \lambda \end{bmatrix}$ has Segre symbol [(2 1)].

However the Segre symbol does not uniquely define the pencil even up to GL-action (that is up to strict equivalence and to $\mathrm{GL}_2(\mathbb{C})$-action on $\mathbb{P}^1$): indeed it contains the informations about the multiplicities of the roots but not the ones about their *position* in the projective space. The latter information is actually important from a geometric point of view, thus to completely determine a pencil of quadrics we need one more object.

Let k be the number of the distinct roots $[y_i : -x_i] \in \mathbb{P}^1$ of the pencil. Up to acting by $\mathrm{GL}_2(\mathbb{C})$ on $\mathbb{P}^1$ we may always assume such roots to be of the form $[1 : -\frac{x_i}{y_i}]$, hence represent them as scalars $z_i \in \mathbb{C}$. Moreover, since they are all distinct, we may look at them in the space

$$\mathbb{C}^{(k)} = \Big\{z \in \mathbb{C}^k \mid z_i \neq z_j \ \forall i \neq j\Big\}$$

To simultaneously change them by $\mathrm{GL}_2(\mathbb{C})$-action on $\mathbb{P}^1$ defines on $\mathbb{C}^{(k)}$ the following equivalence relation

$$z \sim w \iff \exists \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{GL}_2(\mathbb{C}) \ : \ \forall i = 1:k, \ w_i = \frac{az_i + b}{cz_i + d} \tag{3.2}$$

Hence the quotient $\mathbb{C}^{(k)}/_\sim$ parametrizes all the possible k-tuples of roots (up to $\mathrm{GL}_2(\mathbb{C})$-action on $\mathbb{P}^1$) of a pencil with k distinct roots.

Definition. A class $[v] \in \mathbb{C}^{(k)}/_\sim$ is called a **continuous modulus**.

The next result follows straightforward.

Proposition 3.2.2. Let $\mathcal{P} = \mu Q_1 + \lambda Q_2$ be a pencil of quadrics such that $[Q_1] \neq [Q_2] \in \mathbb{P}W$. Then $\mathcal{P}$ is uniquely determined up to GL-action by its Segre symbol as in (3.1) and by a continuous modulus $[v] \in \mathbb{C}^{(k)}/_\sim$.

The classification of intersections of complex quadrics is quite immediate because of the next result [20, Ch.XIII, §10] which directly follows by corollary 2.4.3.

Theorem 3.2.3. Let $\mathcal{P}$ and $\mathcal{Q}$ be two pencils of quadrics in $\mathbb{P}^m$ with roots $[\mu_i^{\mathcal{P}} : \lambda_i^{\mathcal{P}}]$ and $[\mu_i^{\mathcal{Q}} : \lambda_i^{\mathcal{Q}}]$ for $i = 1 : k$. Then $\mathcal{P}$ and $\mathcal{Q}$ are projectively equivalent in $\mathbb{P}^m$ if and only if they have the same Segre symbols. In such case there is an automorphism of $\mathbb{P}^1$ taking $[\mu_i^{\mathcal{P}} : \lambda_i^{\mathcal{P}}]$ to $[\mu_i^{\mathcal{Q}} : \lambda_i^{\mathcal{Q}}]$ for all i with same multiplicities.

Remark 3.2.4. Theorem 3.2.3 allows us to start from the Kronecker form and this helps to avoid equivalent or unattainable cases. For example in $\mathbb{P}^1$ the quadrics X_0^2 and αX_0^2 do not define a pencil, while the pairs of quadrics $(2X_0X_1, 2\alpha X_0X_1 + X_1^2)$ and $(X_0^2, 2X_0X_1)$ define equivalent pencils. Moreover, we avoid the zero pencil and the pencils depending on one only variable, since they do not describe intersections of quadrics.

3.2.1 Classification in $\mathbb{P}^2_{\mathbb{C}}$

We fix $\mathbb{P}^2 = \mathbb{P}^2_{\mathbb{C}}$ with coordinates x, y, z. Since we are working with symmetric pencils, we use the symmetric Kronecker form (2.22). To lighten up the notation, we do not make distinctions between the quadrics and their equation, that is we will write $\mathcal{A} = A(X, X)$. We denote by $\mathcal{A}$ and $\mathcal{B}$ the quadrics defining the pencil and by $V(\mathcal{P})$ the base locus of their intersection. Remark 3.2.4 ensures us that the only possible pencils (or equivalently, intersections) are the ones listed below.

Regular pencil	Segre sym.	$\mathcal{A}$	$\mathcal{B}$	$V(\mathcal{P})$
$\begin{bmatrix} \lambda & & \\ & \lambda+\mu & \\ & & \mu \end{bmatrix}$	[1 1 1]	y^2-z^2	x^2-y^2	four distinct points
$\begin{bmatrix} \mu & \lambda & \\ \lambda & & \\ & & \mu \end{bmatrix}$	[2 1]	x^2-z^2	$2xy$	a double point and two other points
$\begin{bmatrix} \lambda & & \\ & \lambda & \\ & & \mu \end{bmatrix}$	[(1 1) 1]	z^2	x^2-y^2	two double points
$\begin{bmatrix} & \mu & \lambda \\ \mu & \lambda & \\ \lambda & & \end{bmatrix}$	[3]	$2xy$	y^2+2xz	a curvilinear triple point and another point
$\begin{bmatrix} \mu & \lambda & \\ \lambda & & \\ & & \lambda \end{bmatrix}$	[(2 1)]	x^2	$2xy+z^2$	a curvilinear quadruple point

(3.3)

Singular pencil	Segre sym.	$\mathcal{A}$	$\mathcal{B}$	$V(\mathcal{P})$
$\begin{bmatrix} & \lambda & \mu \\ \lambda & & \\ \mu & & \end{bmatrix}$	[; 1;]	$2xz$	$2xy$	a line and a disjoint point
$\begin{bmatrix} \lambda & & \\ & \mu & \\ & & 0 \end{bmatrix}$	[1 1; ; 1]	y^2	x^2	a non-curvilinear quadruple point
$\begin{bmatrix} \mu & \lambda & \\ \lambda & & \\ & & 0 \end{bmatrix}$	[2; ; 1]	x^2	$2xy$	a line and an embedded double point

(3.4)

For all 8 cases one may find out the base locus with a direct calculus of the intersecting equations or by computing on `Macaulay2` the primary decomposition of the ideal generated by the the equations of the two quadrics: to avoid approximated calculus we chose the equations of the quadrics such that the primary decomposition in $\mathbb{Q}[x, y, x]$ was the same than over $\mathbb{C}$ (for instance, $y^2 - z^2$ instead of $y^2 + z^2$).

Example 3.2.5. Consider the case with Segre symbol [2 1]:

```
i1:  R=QQ[x,y,z];
i2:  I=ideal(x^2-z^2,x*y);
i3:  primaryDecomposition I
o3:  {ideal (y,x+z), ideal (y,x-z), ideal (x,z^2)}
```

The primary ideal (x, z^2) determines the double point $[0 : 1 : 0]$, while the primary ideal $(y, x \pm z)$ describes the distinct (simple) point $[1 : 0 : \mp 1]$.

Remark 3.2.6. Regular and singular pencils geometrically differ in the condition

$$\mathcal{P} \text{ is regular} \iff V(\mathcal{P}) \text{ is contained in a smooth curve}$$

Moreover, every regular case (table (3.3)) gives a 0-dimensional zero locus. Neverthless not all singular cases give a 1-dimensional zero locus: indeed the case [1 1; ; 1] has a quadruple point as zero locus, that is 0-dimensional.

To better geometrically classify the base loci in tables 3.3 and 3.4 we have to underline what being a double or triple point means. To do this it is helpful to introduce *affine schemes*.

Affine schemes, 0-dimensional nonreduced. In the following, given the functor

$$\text{Spec} : \Big\{\text{rings}\Big\} \longrightarrow \Big\{\text{schemes}\Big\}$$

which associates to every ring R the scheme $\text{Spec}(R)$, we focus on the case $R = \mathbb{C}[x_1 \dots x_m]$. Consider the affine scheme $\mathbb{A}^m = \mathbb{A}^m_{\mathbb{C}} = \text{Spec}(\mathbb{C}[x_1 \dots x_m])$. Every ideal $I \subset \mathbb{C}[x_1 \dots x_m]$ defines an affine subscheme

$$X = \text{Spec}(\mathbb{C}[x_1 \dots x_m]/I)$$

which is 0-dimensional if and only if $\mathbb{C}[x_1 \dots x_m]/I$ has finite dimension as $\mathbb{C}$-vector space: in this case, one refers to such dimension as the *length* of X.
Since we are interested in *multiple points*, we have to investigate 0-dimensional *nonreduced* affine schemes, that is schemes where the ring $\mathbb{C}[x_1 \dots x_m]/I$ has nilpotents or, equivalently, the ideal I is not radical.
The first case to analyze is given by length-2 nonreduced schemes: every such scheme is isomorphic to

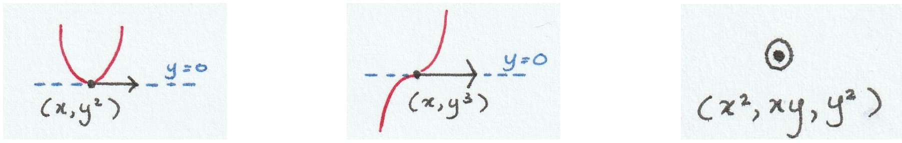
$$\text{Spec}\left(\mathbb{C}[x]/(x^2)\right)$$

and it corresponds to the double point[1] (x^2). Geometrically, a double point is the intersection point between a smooth curve and the tangent line to the curve in that point: in particular, it may be represented by a point equipped with an outgoing arrow describing the direction of the tangent line (Fig.3.1(a)).
A length-3 nonreduced scheme is isomorphic to either

$$\text{Spec}\left(\mathbb{C}[x]/(x^3)\right) \quad \text{or} \quad \text{Spec}\left(\mathbb{C}[x,y]/(x^2, xy, y^2)\right) \simeq \text{Spec}\left(\mathbb{C}[x,y]/(x,y)^2\right)$$

(see [12, Ch.II, §3]) and it correspond to the *curvilinear* triple point (x^3) or to the *fat* triple point $(x, y)^2$ respectively. Geometrically, the first type may be seen as the intersection point between a smooth curve and a line in a flex, while the latter as a point to which two other points approach from different directions: in particular, a curvilinear triple point may be represented by a point equipped with an outgoing arrow (Fig.3.1(b)) while a fat triple one by a circled point (Fig.3.1(c)).

[1] here we do not make distinction between the ideal and its variety



(a) double point (b) curvilinear triple point (c) fat triple point

Figure 3.1: Double and triple points

Since the base loci we deal with are projective, we also need to introduce projective schemes and show how to pass from them to the affine ones to recover the geometric informations we have just seen. We spoiler that the passage from the projective to the affine will be by *localization* which is the algebraic equivalent to the geometric passage to affine chart.

Projective schemes, localizations. Consider the functor

$$\mathrm{Proj} : \Big\{\text{graded rings}\Big\} \longrightarrow \Big\{\text{schemes}\Big\}$$

which associates to every graded ring $A = A_0 \oplus A_+$ the projective scheme $\mathrm{Proj}(A)$ given by the homogeneous prime ideals which do not contain the positive component A_+

$$\mathrm{Proj}(A) = \{\mathfrak{p} \in \mathrm{Spec}(A) \mid \mathfrak{p} \not\supseteq A_+,\ \mathfrak{p} \text{ homogeneous}\}$$

with the topology defined by the closed sets

$$V_+(J) = \{\mathfrak{p} \in \mathrm{Proj}(A) \mid J \subset \mathfrak{p}\} \quad , \forall J \subset A \text{ ideal}$$

Let us fix the graded ring $A = \mathbb{C}[x_1 \dots x_{m+1}]$ and the projective scheme $\mathbb{P}^m = \mathbb{P}^m_{\mathbb{C}} = \mathrm{Proj}(A)$. Given a homogeneous ideal $I \subset A$, we consider the subscheme $\mathrm{Proj}(A/I)$: for every homogeneous element $f \in (A/I)_+$ it holds [28, Ch.II, Proposition 3.38]

$$\mathrm{Proj}(A/I) \setminus V_+(f) = \mathrm{Spec}\Big((A/I)_f\Big) \tag{3.5}$$

that is we may endow $\mathrm{Proj}(A/I)$ with an affine-scheme structure on the open sets of the form $\mathrm{Proj}(A/I) \setminus V_+(f)$. Thus by (3.5) we can study the geometric properties of a 0-dimensional projective scheme by studying a 0-dimensional affine scheme.

Let us underline the geometric properties of the zero loci in table (3.3).

- For [1 1 1], the zero locus is given by the four distinct (simple) points $(y \pm z, x \pm z)$ with the property that three of them are never collinear (by Bèzout theorem).
- For [2 1], neither two distinct (simple) points $(y, x+z)$ and $(y, x-z)$ lie on the tangent line of the double point (x, z^2) (Fig.3.2(a)).

- For [(1 1) 1], the tangent lines of the two double points $(z^2, x+y)$ and $(z^2, x-y)$ do not intersect eachothers (Fig.3.2(b)).

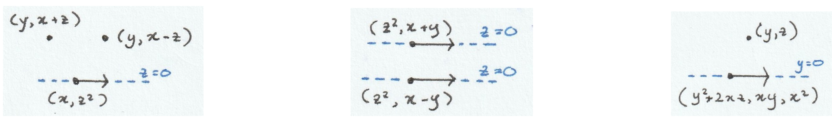
- For [3], the triple point (y^2+2xz, xy, x^2) is curvilinear since by (3.5) we have in the chart $\{z \neq 0\}$ it holds

$$
\begin{aligned}
\operatorname{Proj}\left(\mathbb{C}[x,y,z]/(y^2+2xz, xy, x^2)\right) &= \operatorname{Spec}\left(\left(\mathbb{C}[x,y,z]\big/(y^2+2xz, xy, x^2)\right)_z\right) \\
&= \operatorname{Spec}\left(\mathbb{C}[u,v]/(v^2+2u, uv, u^2)\right) \\
&\simeq \operatorname{Spec}(\mathbb{C}[v]/(v^3))
\end{aligned}
$$

  where $u = \frac{x}{z}$ and $v = \frac{y}{z}$ (or equivalently, since the curve y^2+2xz is smooth). The simple point (y, z) does not lie on its tangent line (Fig.3.2(c)).

- For [(2 1)], the quadriple point $(2xy+z^2, x^2)$ is curvilinear since by (3.5) in the chart $\{y \neq 0\}$ it holds (for $u = \frac{x}{y}$ and $w = \frac{z}{y}$)

$$
\begin{aligned}
\operatorname{Proj}\left(\mathbb{C}[x,y,z]/(2xy+z^2, x^2)\right) &= \operatorname{Spec}\left(\left(\mathbb{C}[x,y,z]\big/(2xy+z^2, x^2)\right)_y\right) \\
&= \operatorname{Spec}\left(\mathbb{C}[u,w]/(2u+w^2, u^2)\right) \\
&\simeq \operatorname{Spec}(\mathbb{C}[w]/(w^4))
\end{aligned}
$$



(a) [2 1] (b) [(1 1) 1] (c) [3]

Figure 3.2: Some base loci of pencils of quadrics in $\mathbb{P}^2$

The real case. Tables (3.3) and (3.4) make it clear that in $\mathbb{P}^2_{\mathbb{C}}$ there are 5 possible non-singular intersections and 3 singular ones, for a total of 8 cases. The complex case allows to determine the classfication in $\mathbb{P}^2_{\mathbb{R}}$ too: indeed, since the invariant for congruence over $\mathbb{R}$ is the *signature* of the quadratic form, at worst each complex invariance class splits into different real invariance classes. More precisely, the splitting classes are the ones whose subdiscriminant has at least two different roots, that is

- the class represented by the Segre symbol [1 1 1] splits in 3 real classes;

- each class of the ones represented by [2 1], [(1 1) 1] and [1 1; ; 1] splits in 2 real classes.

Thus in $\mathbb{P}^2_{\mathbb{R}}$ there are 13 cases of which 9 are non-singular and 4 are singular.

3.2.2 Classification in $\mathbb{P}^3_{\mathbb{C}}$

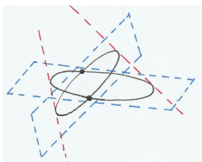
We fix $\mathbb{P}^3 = \mathbb{P}^3_{\mathbb{C}}$ with coordinates x, y, z, w. In this case we only list the regular pencils (or equivalently, non-singular intersections). The same remarks (in the same notations) done in $\mathbb{P}^2_{\mathbb{C}}$ lead to the following non-singular intersections.

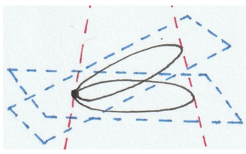
Regular pencil	Segre sym.	$\mathcal{A}$	$\mathcal{B}$	$V(\mathcal{P})$
$\begin{bmatrix} \lambda & & & \\ & \lambda+\mu & & \\ & & \lambda-\mu & \\ & & & \mu \end{bmatrix}$	[1 1 1 1]	$y^2 - z^2 + w^2$	$x^2 + y^2 + z^2$	elliptic curve
$\begin{bmatrix} \mu & \lambda & & \\ \lambda & & & \\ & & \mu & \\ & & & \lambda+\mu \end{bmatrix}$	[2 1 1]	$x^2 + z^2 + w^2$	$2xy + w^2$	nodal curve
$\begin{bmatrix} \lambda & & & \\ & \lambda & & \\ & & \mu & \\ & & & \lambda+\mu \end{bmatrix}$	[(1 1) 1 1]	$z^2 + w^2$	$x^2 + y^2 + w^2$	two conics in general position (Fig.3.3(a))
$\begin{bmatrix} & \mu & \lambda & \\ \mu & \lambda & & \\ \lambda & & & \\ & & & \mu \end{bmatrix}$	[3 1]	$2xy + w^2$	$y^2 + 2xz$	cuspidal curve
$\begin{bmatrix} \mu & \lambda & & \\ \lambda & & & \\ & & \lambda & \\ & & & \mu \end{bmatrix}$	[(2 1) 1]	$x^2 + w^2$	$2xy + z^2$	two tangent conics (Fig.3.3(b))
$\begin{bmatrix} \lambda & & & \\ & \lambda & & \\ & & \lambda & \\ & & & \mu \end{bmatrix}$	[(1 1 1) 1]	w^2	$x^2 + y^2 + z^2$	a double conic
$\begin{bmatrix} \mu & \lambda & & \\ \lambda & & & \\ & & \lambda & \mu \\ & & \mu & \end{bmatrix}$	[2 2]	$x^2 + 2zw$	$2xy + z^2$	a twisted cubic and a bisecant
$\begin{bmatrix} \lambda & & & \\ & \lambda & & \\ & & \lambda & \mu \\ & & \mu & \end{bmatrix}$	[(1 1) 2]	$2zw$	$x^2 + y^2 + z^2$	a conic and two lines in triangle (Fig.3.3(c))
$\begin{bmatrix} \lambda & & & \\ & \lambda & & \\ & & \mu & \\ & & & \mu \end{bmatrix}$	[(1 1) (1 1)]	$z^2 + w^2$	$x^2 + y^2$	four lines in two plane-pairs intersecting eachothers (Fig.3.3(d))

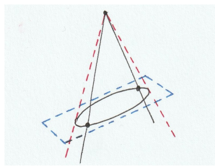
(3.6)

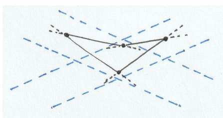
Regular pencil	Segre sym.	$\mathcal{A}$	$\mathcal{B}$	$V(\mathcal{P})$	
$\begin{bmatrix} & & \mu & \lambda \\ & \mu & \lambda & \\ \mu & \lambda & & \\ \lambda & & & \end{bmatrix}$	[4]	$y^2 + 2xz$	$2(xw + yz)$	a twisted cubic and a tangent line	
$\begin{bmatrix} & \mu & \lambda & \\ \mu & \lambda & & \\ \lambda & & & \\ & & & \lambda \end{bmatrix}$	[(3 1)]	$2xy$	$y^2 + 2xz + w^2$	a conic and two lines intersecting in one point (Fig.3.3(e))	(3.7)
$\begin{bmatrix} \mu & \lambda & & \\ \lambda & & & \\ & & \mu & \lambda \\ & & \lambda & \end{bmatrix}$	[(2 2)]	$x^2 + z^2$	$2(xy + zw)$	a double line meeting two disjoint lines	
$\begin{bmatrix} \mu & \lambda & & \\ \lambda & & & \\ & & \lambda & \\ & & & \lambda \end{bmatrix}$	[(2 1 1)]	x^2	$2xy + z^2 + w^2$	two double lines intersecting	

All the base loci in tables (3.6) and (3.7) can be obtained in `Macaulay2` by computing the primary decomposition of the polynomial ideal generated by the equations of the two corresponding quadrics. For a detailed geometric description we refer to [20, pgg. 305-308].


(a) [(1 1) 1 1]


(b) [(2 1) 1]


(c) [(1 1) 2]


(d) [(1 1) (1 1)]

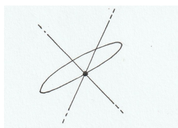

(e) [(3 1)]

Figure 3.3: Some base loci of pencils of quadrics in $\mathbb{P}^3$

3.3 A geometric interpretation

The aim of this section is to give a geometric interpretation to the equivalence classes of pencils of complex quadrics, that is to describe some geometric properties which uniquely determine their Kronecker class. To do so we first need to introduce the geometric objects which are related to a pencil of quadratic forms.

Let W be the vector space of quadratic forms on $\mathbb{C}^{m+1}$, i.e.

$$W = \Big\{ Q : \mathbb{C}^{m+1} \to \mathbb{C} \text{ quadratic form} \Big\}$$

and let $\mathbb{P}W$ be its projective space. We also consider for all $r = 1 : m+1$ the subset

$$W_r = \{ Q \in W \mid \operatorname{Rk}(Q) = r \} \subset W$$

of the quadratic forms of fixed rank r and the corresponding projectivized $\mathbb{P}W_r$: the latter are irreducible smooth quasi-projective varieties. These varieties have the following properties [10, Proposition 1.1], where $\overline{\mathbb{P}W_r}$ denotes the Zariski closure of the variety and $\operatorname{Sing}(\overline{\mathbb{P}W_r})$ its singular part.

Proposition 3.3.1. *(i)* $\overline{\mathbb{P}W_r} = \bigcup_{i=1}^{r} \mathbb{P}W_i$ for all $r = 1 : m+1$;

(ii) $\operatorname{codim} \mathbb{P}W_r = \frac{(m+1-r)(m+2-r)}{2}$ for all $r = 1 : m+1$;

(iii) $\forall 1 < r < m+1$, $\overline{\mathbb{P}W_{r-1}} = \operatorname{Sing}(\overline{\mathbb{P}W_r})$.

Next we introduce the geometric objects related to a pencil of quadrics. Given a pencil $\mathcal{P} = \mu Q_1 + \lambda Q_2$ defined by two linearly independent quadratic forms $Q_1, Q_2 \in W \setminus \{0\}$, it determines a *projective line* $L_{\mathcal{P}} \subset \mathbb{P}W$ and a *projective variety* $V(\mathcal{P}) \subset \mathbb{P}^m$ given by the base locus of the intersection of the two quadrics, that is $V(\mathcal{P}) = \{Q_1 = Q_2 = 0\}$.

Our claim is to prove that the Kronecker class of a pencil of quadrics $\mathcal{P}$ is uniquely determined by the "position" of the line $L_{\mathcal{P}}$ with respect to the subvarieties $\overline{\mathbb{P}W_r}$ and by the singular part $\operatorname{Sing}(V(\mathcal{P}))$ of the base locus $V(\mathcal{P})$: actually one may bet that the position of the line $L_{\mathcal{P}}$ determines the regular part of the pencil (or equivalently the multiplicities of its elementary divisors) while $\operatorname{Sing}(V(\mathcal{P}))$ the singular part (that is its minimal indices).

Remark 3.3.2. In this section by "singular part" of the base locus we refer not only to its schematically-singular one (i.e. subvarieties of multiplicity greater than 1 such as multiple points or multiple lines) but also to its part of dimension greater than the expected one. For instance, in $\mathbb{P}^2$ the singular part of the base locus of the pencil of quadrics $[2;;1]$ is not only the double point (x^2, y) but also the line (x) (see also example 3.3.24).

The thread of this section will be the following: first we define the notion of *similar position* of two lines in $\mathbb{P}W$ and we show how it determines the regular part of the pencil; then we briefly introduce the notion of *projective bundle* and we explicitly build some of them over $\mathbb{P}^1$; finally we study how $\operatorname{Sing}(V(\mathcal{P}))$ determines the minimal indices of $\mathcal{P}$.

3.3.1 Similar position of lines in $\mathbb{P}W$

Let $L \subset \mathbb{P}W$ be a projective line of quadrics. We define

$$m_0(L) = \min\{r \mid L \subset \overline{\mathbb{P}W_r}\}$$

and, given $\{P_1, \dots, P_{q_L}\} = L \cap \overline{\mathbb{P}W_{m_0(L)-1}}$, for all $i = 1 : q_L$ and $j = 1 : m_0(L) - 1$ we set

$$m_{ij}(L) = \mathrm{mult}_{P_i}(L \cap \overline{\mathbb{P}W_{m_0(L)-j}})$$

to be the multiplicity of intersection in the point P_i. Moreover, for all $i = 1 : q_L$ we set

$$k_i(L) = \max\{k \mid P_i \in \overline{\mathbb{P}W_{m_0(L)-k}}\}$$

thus for all $i = 1 : q_L$ we have $m_{ij}(L) = 0$ for all $j > k_i(L)$ and we may restrict to consider $j = 1 : k_i(L)$.

Definition. Two lines $L, L' \subset \mathbb{P}W$ are said to have **similar position** with respect to the varieties $\{\overline{\mathbb{P}W_r} \mid r = 1 : m + 1\}$ if

- $m_0(L) = m_0(L')$;
- $q(L) = q(L')$;
- there exists an isomorphism $h : L \to L'$ such that
  - ◇ $h(P_i) = P_i'$;
  - ◇ $k_i(L) = k_i(L')$ for all $i = 1 : q(L)$;
  - ◇ $m_{ij}(L) = m_{ij}(L')$ for all $i = 1 : q(L)$ and $j = 1 : k_i(L)$.

In particular, we define the **position** of a line $L \subset \mathbb{P}W$ as the set of the above values.

The next result shows which informations are hidden in these values when considering a line associated to a pencil of quadrics: for a proof we refer to [10, Lemma 2.3].

Lemma 3.3.3. Let $\mathcal{P}$ be a pencil of quadrics with Segre symbol

$$\Sigma(\mathcal{P}) = \left[(e_1^1, \dots, e_{r_1}^1) \dots (e_1^k, \dots, e_{r_k}^k); \epsilon_{g+1}, \dots, \epsilon_p; g\right]$$

and let $L_{\mathcal{P}} \subset \mathbb{P}W$ be the corresponding line. Then:

(*i*) $m_0(L_{\mathcal{P}}) = m + 1 - p$;

(*ii*) $q(L_{\mathcal{P}}) = k$, that is the number of intersection points P_i is equal to the number of the roots of the pencil;

(*iii*) $k_i(L_{\mathcal{P}}) = r_i$ for all $i = 1 : k$;

(*iv*) $m_{ij}(L_{\mathcal{P}}) = \sum_{l=1}^{r_i-j+1} e^i_{r_i-l+1}$ for all $i = 1 : k$ and $j = 1 : r_i$.

We may reformulate lemma 3.3.3 in terms of similar position.

Proposition 3.3.4. Let $\mathcal{P}, \mathcal{P}'$ be two pencil of quadrics with corresponding lines $L, L' \subset \mathbb{P}W$ respectively. Then L, L' have similar position if and only if the pencils have Segre symbols with the same multiplicities (i.e $k = k'$ and $e^i_j = (e')^i_j$) and same number of minimal indices (i.e. same $p = p'$), other than same continuous modulus $[v] \in \mathbb{C}^{(k)}/_\sim$.

Corollary 3.3.5. If $\mathcal{P}$ is a regular pencil of quadrics, then it is uniquely determined by the position of $L_{\mathcal{P}}$.

Proof. The position of $L_{\mathcal{P}}$ is given by the values $m_0(L_{\mathcal{P}}), q(L_{\mathcal{P}}), k_i(L_{\mathcal{P}})$ and $m_{ij}(L_{\mathcal{P}})$. If the pencil is regular (i.e. $p = 0$), it is completely determined by the multiplicities of the roots and by the continuous moduli, hence by proposition 3.3.4 and lemma 3.3.3 we conclude. □

Remark 3.3.6. We underline that if the pencil is singular, the position of $L_{\mathcal{P}}$ is not enough to uniquely determine it: indeed in proposition 3.3.4 we do not ask to have the same minimal indices but only the same number of them.

By proposition 3.3.4 it is clear that the position of the line $L_{\mathcal{P}}$ completely determine the regular part of the pencil.

Example 3.3.7 (Projective plane, position of lines)**.** Consider the Segre classification of regular pencils of quadrics in $\mathbb{P}^2$. First of all we note that in this case $\mathbb{P}W \simeq \mathbb{P}^5$ with *determinant hypersurface* $\overline{\mathbb{P}W_2}$.
For each pencil $\mathcal{P}$ in table (3.3) we have $p = g = 0$, thus by lemma 3.3.3 it follows $m_0(L_{\mathcal{P}}) = 3$: indeed by proposition 3.3.1(*ii*) we know that $\dim(\mathbb{P}W_2) = 5 - 1 = 4$ but none of such pencils is contained in a hypersurface. In particular, the regular pencils of quadrics in $\mathbb{P}^2$ are never contained in the determinant hypersurface $\overline{\mathbb{P}W_2}$, hence they cut it in a finite number of points. But $\overline{\mathbb{P}W_2} = \overline{\mathbb{P}W_{m_0(L_{\mathcal{P}})-1}}$, hence the number of such intersection points is exactly the value $q(L_{\mathcal{P}})$. However the value $q(L_{\mathcal{P}})$ is still not enough to distinguish all the cases (see table 3.8) and actually neither the multiplications of intersection with $\overline{\mathbb{P}W_2}$ are enough. One has to study the intersections with $\overline{\mathbb{P}W_1}$ too: indeed the pencils [2 1] and [3] do not intersect $\overline{\mathbb{P}W_1}$ while [(1 1) 1] and [(2 1)] do since for $\lambda = 0, \mu = 1$ the latter pencils have rank 1.

The same arguments may be reformulated for singular pencils.

$\Sigma(\mathcal{P})$	$L_{\mathcal{P}}$	$\det(\mathcal{P})$	$q(L_{\mathcal{P}})$	$L_{\mathcal{P}} \cap \overline{\mathbb{P}W_2}$	$L_{\mathcal{P}} \cap \overline{\mathbb{P}W_1}$
[1 1 1]	$\lambda x^2 + (\mu - \lambda)y^2 - \mu z^2$	$\lambda(\lambda+\mu)\mu$	3	$1+1+1$	$\emptyset$
[2 1]	$\mu x^2 - \mu z^2 + 2\lambda xy$	$\lambda^2\mu$	2	$2+1$	$\emptyset$
[(1 1) 1]	$\lambda x^2 - \lambda y^2 + \mu z^2$	$\lambda^2\mu$	2	$2+1$	1
[3]	$\lambda y^2 + 2\lambda xz + 2\mu xy$	λ^3	1	3	$\emptyset$
[(2 1)]	$\mu x^2 + 2\lambda xy + \lambda z^2$	λ^3	1	3	1
[; 1;]	$\mu xz + \lambda xy$	0	0	$L_{\mathcal{P}}$	$\emptyset$
[1 1; ; 1]	$\mu y^2 + \lambda x^2$	0	2	$L_{\mathcal{P}}$	$1+1$
[2; ; 1]	$\mu x^2 + \lambda xy$	0	1	$L_{\mathcal{P}}$	2

(3.8)

Remark 3.3.8. Actually for $m = 2, 3$ the position of the line $L_{\mathcal{P}}$ completely determines the singular part too, hence the whole Segre symbol: this comes from combinatorial restraints such as (2.14) and from the fact that each non-zero minimal index ϵ_i defines a square block of size $2\epsilon_i + 1$ (see (2.20)). Let $\mathcal{P}$ be a singular pencil of quadrics with Segre symbol $\Sigma(\mathcal{P}) = \left[(e_1^1, \dots, e_{r_1}^1) \dots (e_1^k, \dots, e_{r_k}^k); \epsilon_{g+1}, \dots, \epsilon_p; g\right]$ with $p > 0$.

For $\mathbb{P}^2$ it is clear from table (3.4): if there are no roots at all (i.e. no elementary divisors), then $p = m + 1 = 3$ and the only possible case is [; 1;] (since $g = m + 1 = 3$ would give the identically zero pencil); if there is at least one root, then either there is another root or the only one root has multiplicity 2 and in both cases it must be $p = g = 1$.

In $\mathbb{P}^3$, if the pencil has two or three roots counted with multiplicity, then it must be either $p = g = 1$ or $p = g = 2$ (since a block R_{ϵ_i} has at least size 3). If the pencil has exactly one simple root, then it must be $p = 1 > g = 0$ and $\epsilon_1 = 1$ (since the pencil $\mathrm{diag}(\lambda, 0, 0, 0)$ does not describe a pencil of quadrics by remark 3.2.4). If the pencil has no root, the only possible case is $g = 1$ and $\epsilon_{g+1} = \epsilon_p = 1$ (from size arguments as well).

For $\mathbb{P}^4$ the position of $L_{\mathcal{P}}$ is not enough in general: for instance, the pencils

$$\begin{bmatrix} 0 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & \mu & \lambda \\ & & & \lambda & \end{bmatrix}, \quad \begin{bmatrix} & \lambda & \mu & & \\ \lambda & & & & \\ \mu & & & & \\ & & & \mu & \lambda \\ & & & \lambda & \end{bmatrix}$$

(with Segre symbols [2; 0; 3] and [2; 1; 0] respectively) have corresponding lines in similar position but they are not equivalent.

At this moment it remains to determine the singular part of a pencil of quadrics $\mathcal{P}$ and this is possible by analyzing $\mathrm{Sing}(V(\mathcal{P}))$[2]. As mentioned before, to do so we first need to introduce some projective bundles over $\mathbb{P}^1$.

[2]in the sense of remark 3.3.2

3.3.2 Projective bundles over $\mathbb{P}^1$

We define a **projective bundle** to be a bundle whose fibers are projective spaces. In the following we restrict to consider projective bundles of vector bundles: given a vector bundle $\mathcal{E}$ over a variety X, we consider the projective bundle $\mathbb{P}\mathcal{E}$ given by the projective spaces of the fibers of $\mathcal{E}$.

Remark 3.3.9. Every vector bundle $\mathcal{E}$ defines a projective bundle $\mathbb{P}\mathcal{E}$ but not every projective bundle arises from a vector bundle: however the converse holds for X to be a smooth variety, e.g. a compact Riemann surface. For more details we refer to [18, Ch.II, §7].

We now give a formal definition for projective bundles of the form $\mathbb{P}\mathcal{E}$, that is over a smooth variety X. Let $p : \mathcal{E} \to X$ be a rank-r vector bundle over X: then there exist a open covering $\{U_i\}_I$ of X and isomorphisms $g_i : p^{-1}(U_i) \to U_i \times \mathbb{C}^r$ for all $i \in I$ such that the transition maps are of the form

$$g_{ij} = g_j \circ g_i^{-1} : \begin{array}{ccc} (U_i \cap U_j) \times \mathbb{C}^r & \to & (U_i \times U_j) \times \mathbb{C}^r \\ (P, v) & \mapsto & (P, G_{ij} v) \end{array}$$

where $G_{ij} \in \mathrm{GL}_r(\mathbb{C})$.

Definition. In the above notations, the **projective bundle** $\rho : \mathbb{P}\mathcal{E} \to X$ is the bundle over X defined by the trivializations $\psi_i : \rho^{-1}(U_i) \to U_i \times \mathbb{P}^{r-1}_{\mathbb{C}}$ and the transition maps

$$\psi_{ij} = \psi_j \circ \psi_i^{-1} : \begin{array}{ccc} (U_i \cap U_j) \times \mathbb{P}^{r-1} & \to & (U_i \cap U_j) \times \mathbb{P}^{r-1} \\ (P, [v]) & \mapsto & (P, \Psi_{ij}[v]) \end{array}$$

where $\Psi_{ij} = [G_{ij}] \in \mathbb{P}\mathrm{GL}_r(\mathbb{C})$.

Projective bundles of the form $\mathbb{P}\mathcal{E}$ well behave with twistings of $\mathcal{E}$ by line bundles [18, Ch.II, Lemma 7.9]. We recall that the isomorphism classes of line bundles on X form the *Picard group* $\mathrm{Pic}(X)$.

Lemma 3.3.10. For all $\mathcal{L} \in \mathrm{Pic}(X)$ there is a natural isomoprhism

$$\mathbb{P}\mathcal{E} \simeq \mathbb{P}(\mathcal{E} \otimes \mathcal{L})$$

Proof. To tensorize $\mathcal{E}$ by $\mathcal{L}$ just multiplies the transition functions of $\mathcal{E}$ by a scalar function but this does not affect the morphism in projective coordinates. □

Corollary 3.3.11. For all $\mathcal{L}, \mathcal{L}' \in \mathrm{Pic}(X)$ it holds

$$\mathbb{P}\mathcal{L} \simeq \mathbb{P}\mathcal{L}' \simeq X$$

Our interest is in projective bundles over $\mathbb{P}^1$: by remark 3.3.9 we know that every such projective bundle is of the form $\mathbb{P}\mathcal{E}$ for a certain vector bundle $\mathcal{E}$. Such vector bundles are completely classified as follows.

Theorem 3.3.12 (Segre-Grothendieck splitting theorem)**.** Every holomorphic rank-r vector bundle $\mathcal{E}$ over $\mathbb{P}^1_{\mathbb{C}}$ splits into a direct sum of holomorphic line bundles

$$\mathcal{E} \simeq \bigoplus_{i=1}^{r} \mathcal{O}(e_i)$$

for certain $e_1, \ldots, e_r \in \mathbb{Z}$.

Remark 3.3.13. Grothendieck [16] explicitly proved it in 1957 as particular case of a more general result, but its formulation is equivalent to the *Birkhoff factorization* due to Birkhoff [3] in 1909: a simplification of the latter proof was given by Hazewinkel and Martin [19] in terms of linear algebra. However in 1884 Segre [34] proved this theorem too even if without the bundle language: this is why we refer to it as the *Segre-Grothendieck theorem.*

The splitting theorem 3.3.12 allows to determine when two projective bundles (of the same rank of course) over $\mathbb{P}^1$ are isomorphic[3].

Proposition 3.3.14. Let $\mathcal{E} = \bigoplus_{i=1}^{r} \mathcal{O}(a_i)$ and $\mathcal{F} = \bigoplus_{j=1}^{r} \mathcal{O}(b_j)$ be vector bundles over $\mathbb{P}^1$. Then

$$\mathbb{P}\mathcal{E} \overset{\text{bundles}}{\simeq} \mathbb{P}\mathcal{F} \iff \exists m \in \mathbb{Z},\ \exists \sigma \in \mathfrak{S}_k \ :\ a_i - b_{\sigma(i)} = m \ \forall i$$

Proof. First of all we note that the right condition is equivalent to

$$\exists m \in \mathbb{Z} \ :\ \mathcal{E} \simeq \mathcal{F} \otimes \mathcal{O}(m)$$

Thus the "if" condition directly follows by lemma 3.3.10.
For the "only if" condition, we assume there exists an isomorphism of projective bundles $\varphi : \mathbb{P}\mathcal{E} \xrightarrow{\sim} \mathbb{P}\mathcal{F}$ and we consider the transition maps for $\mathbb{P}\mathcal{E}$ and $\mathbb{P}\mathcal{F}$

$$[G_{01}] = \mathbb{P}\begin{bmatrix} (\frac{x_0}{x_1})^{a_1} & & \\ & \ddots & \\ & & (\frac{x_0}{x_1})^{a_k} \end{bmatrix} \quad , \quad [H_{01}] = \mathbb{P}\begin{bmatrix} (\frac{x_0}{x_1})^{b_1} & & \\ & \ddots & \\ & & (\frac{x_0}{x_1})^{b_k} \end{bmatrix}$$

respectively, where the "$\mathbb{P}$" denotes we are considering the classes in $\mathbb{P}\mathrm{GL}_k(\mathbb{C})$.
We may assume that φ is (uniquely) defined by the matrix

$$\widetilde{\Phi_{01}} = \mathbb{P}\begin{bmatrix} (\frac{x_0}{x_1})^{b_1-a_1} & & \\ & \ddots & \\ & & (\frac{x_0}{x_1})^{b_k-a_k} \end{bmatrix}$$

[3] as projective bundles over $\mathbb{P}^1$

in $\mathbb{P}\mathrm{GL}_k(\mathbb{C})$: indeed the automorphisms of $\mathcal{E}$ which fix $\mathbb{P}^1$ are

$$\ker\Big(\mathrm{Aut}(\mathbb{P}\mathcal{E}) \to \mathrm{Aut}(\mathbb{P}^1)\Big) = \mathbb{P}\mathrm{Aut}(\mathcal{E})$$

and, given $\Phi_{01} \in \mathbb{P}\mathrm{GL}_k(\mathbb{C})$ the matrix defining φ, there are $\psi_M \in \mathbb{P}\mathrm{Aut}(\mathcal{E})$ and $\psi_N \in \mathbb{P}\mathrm{Aut}(\mathcal{F})$ such that $\widetilde{\Phi_{01}} = M \cdot \Phi_{01} \cdot N$, hence we may consider the isomorphism of projective bundles $\widetilde{\phi} = \psi_N \circ \varphi \circ \psi_M : \mathbb{P}\mathcal{E} \to \mathbb{P}\mathcal{F}$. But such $\widetilde{\Phi}_{01}$ gives an isomoprhism if and only if $\widetilde{\Phi}_{01} = [I] \in \mathbb{P}\mathrm{GL}_k(\mathbb{C})$ and this holds if and only if there exists $m \in \mathbb{Z}$ such that $a_i - b_i = m$ for all $i = 1 : k$. □

Moreover, projective bundles over $\mathbb{P}^1$ are uniquely determined (up to isomorphism) by their structure as varieties, as follows by the next result.

Lemma 3.3.15. If two projective bundles over $\mathbb{P}^1$ are isomorphic as (abstract) varieties, then they are isomoprhic as projective bundles.

Proof. Let $\mathbb{P}\mathcal{E}$ and $\mathbb{P}\mathcal{F}$ be two projective bundles over $\mathbb{P}^1$ of rank $r-1$ that are isomorphic as varieties. We prove the thesis by induction on the rank of the bundles.
If $r = 1$, by corollary 3.3.11 $\mathbb{P}\mathcal{E} \overset{\text{bundles}}{\simeq} \mathbb{P}\mathcal{F} \simeq \mathbb{P}^1$.
If $r = 2$, the thesis follows by the classification of rational ruled surfaces ([18, pg. 372]).
Let $r \geq 3$ and let $\varphi : \mathbb{P}\mathcal{E} \to \mathbb{P}\mathcal{F}$ be an isomorphism as varieties: it is enough to prove that φ maps the fibers of the bundle projection $\rho_{\mathcal{E}}$ onto the ones of $\rho_{\mathcal{F}}$. Given $P \in \mathbb{P}^1$, we consider the fiber $(\mathbb{P}\mathcal{E})_P$ of $\rho_{\mathcal{E}}$ in P and the composition

$$(\mathbb{P}\mathcal{E})_P \xrightarrow{\varphi} \mathbb{P}\mathcal{F} \xrightarrow{\rho_{\mathcal{F}}} \mathbb{P}^1$$

But $(\mathbb{P}\mathcal{E})_P \simeq \mathbb{P}^{r-1}$, hence the above map is constant, since every map $\mathbb{P}^N \to \mathbb{P}^1$ with $N > 1$ is so. This concludes. □

We now exhibit some embeddings of projective bundles over $\mathbb{P}^1$ in a projective space $\mathbb{P}^N$. Let $\mathbb{P}\mathcal{E}$ a projective bundle of rank $r-1$ over $\mathbb{P}^1$ where $\mathbb{E} = \bigoplus_{i=1}^r \mathcal{O}(a_i)$.
Let us assume $0 < a_1 \leq \ldots \leq a_r$. Consider $N \in \mathbb{N}$ large enough such that:

- $\mathbb{P}^N$ contains the Veronese variety $\nu_{a_i}(\mathbb{P}^1)$ for all $i = 1 : r$;
- given $\psi_i : \mathbb{P}^1 \overset{\nu_{a_i}}{\to} \nu_{a_i}(\mathbb{P}^1) \to \mathbb{P}^N$ for all $i = 1 : r$, it holds

$$L(\operatorname{Im}\psi_i) \cap L\Bigg(\bigcup_{j\neq i} \operatorname{Im}\psi_j\Bigg) = \emptyset \ , \forall i \tag{3.9}$$

  where $L(\operatorname{Im}\psi_i)$ denotes the linear span of $\operatorname{Im}\psi_i$ in $\mathbb{P}^N$.

Remark 3.3.16. From (3.9) it follows that with a suitable choise of coordinates in $\mathbb{P}^N$ we may assume

$$\psi_i([x:y]) = [\underbrace{0:\ldots:0}_{i+\sum_{j<i} a_j} : x^{a_i} : \ldots : y^{a_i} : 0 : \ldots : 0] \ , \forall i = 1:r$$

We set

$$P(a_1 \ldots a_r) = \bigcup_{z\in\mathbb{P}^1} L\Big(\psi_1(z), \ldots, \psi_r(z)\Big) \tag{3.10}$$

Clearly $P(a_1 \ldots a_r)$ is a projective bundle of rank $r-1$ over $\mathbb{P}^1$ with fibers

$$\Big(P(a_1 \ldots a_r)\Big)_z = L\Big(\psi_1(z), \ldots, \psi_r(z)\Big) \simeq \mathbb{P}^{r-1} \ , \forall z \in \mathbb{P}^1$$

Moreover, by remark 3.3.16 it can be covered by the trivialization

$$\begin{array}{cccc} \varphi_0 : U_0 \times \mathbb{P}^{r-1} \simeq & \mathbb{C}\times\mathbb{P}^{r-1} & \longrightarrow & \chi^{-1}(U_0) \\ & (s,[t_1:\ldots:t_r]) & \mapsto & [t_1 : t_1 s : \ldots : t_1 s^{a_1} : t_2 : \ldots : t_r s^{a_r}] \end{array}$$

and the complementary one $\varphi_1 : U_1 \times \mathbb{P}^{r-1} \to \chi^{-1}(U_1)$. Thus it immediately follows:

Corollary 3.3.17. $P(a_1 \ldots a_r) \overset{\text{bundles}}{\simeq} \mathbb{P}\mathcal{E}$.

By starting from the bundle $P(a_1 \ldots a_r)$ we can define a new variety which will help us in the study of the singular part $\mathrm{Sing}(V(\mathcal{P}))$. Let $L_0 \subset \mathbb{P}^N$ be a linear subspace of dimension $a_0 \geq 0$ such that

$$L_0 \cap L\Big(P(a_1 \ldots a_r)\Big) = \emptyset$$

We define the **join variety** of dimension $r + a_0 + 1$

$$J(a_1 \ldots a_r; a_0) = \bigcup_{\substack{x\in P(a_1\ldots a_r)\\ y\in L_0}} L(x,y) \subseteq \mathbb{P}^N \tag{3.11}$$

Remark 3.3.18 (Schematical singularity in a join variety)**.**

- $J(a_1 \ldots a_r; a_0)$ is (schematically) non-singular if and only if $J(a_1 \ldots a_r; a_0)$ is linear if and only if $r = a_1 = 1$, that is $J(1; a_0) = L(\mathbb{P}^1, L_0)$.
- If $J(a_1 \ldots a_r; a_0)$ is (schematically) singular, then it has singularity in L_0.

We recall that we make distinction between being schematically singular and being singular in the sense of remark 3.3.2: indeed in proposition 3.3.20 we will see that $\mathrm{Sing}(V(\mathcal{P}))$[4] can have components isomorphic to a join variety but not necessarily *schematically* singular (see the pencil [; 1;] in example 3.3.24).

[4] in the sense of remark 3.3.2

We can determine when two join varieties are isomorphic [10, Lemma 1.5].

Lemma 3.3.19. $J(a_1 \dots a_r; a_0) \simeq J(b_1 \dots b_s; b_0) \iff r = s$ and $a_i = b_i\ \forall i = 0 : r$.

Note: A join variety $J(a_1 \dots a_r; a_0)$ is never a projective bundle over $\mathbb{P}^1$.

3.3.3 From $\mathrm{Sing}(V(\mathcal{P}))$ to minimal indices

Let $\mathcal{P} = \mu A_1 + \lambda Q_2$ be a pencil of quadrics with $[Q_1] \neq [Q_2] \in \mathbb{P}W$ and Segre symbol

$$\Sigma(\mathcal{P}) = \left[(e^1_1, \dots, e^1_{r_1}) \dots (e^k_1, \dots, e^k_{r_k}); \epsilon_{g+1}, \dots, \epsilon_p; g\right]$$

Let $\mathrm{Sing}(V(\mathcal{P}))$ be the singular part of the complete intersection $Q_1 = Q_2 = 0$ in the sense of remark 3.3.2 (we do not get tired to underline it!). The following result tells us that the geometric informations with respect to the minimal indices of the pencil are hidden into the components of $\mathrm{Sing}(V(\mathcal{P}))$ [10, Lemma 2.5].

Proposition 3.3.20. In the above notations, let $\overline{k} = k - \#\{i \mid r_i = e^i_{r_i} = 1\}$. Then $\mathrm{Sing}(V(\mathcal{P}))$ has at least t components $\mathcal{S}_1, \dots, \mathcal{S}_t$ (with reduced structure) where

$$t = \begin{cases} \overline{k} & \text{if } p = g = 0 \text{ (no minimal indices)} \\ \max\{\overline{k}, 1\} & \text{if } p = g > 0 \text{ (only zero minimal indices)} \\ \overline{k} + 1 & \text{if } p > g \text{ (there are non-zero minimal indices)} \end{cases}$$

Moreover, up to permutation of the $\mathcal{S}_i$'s, it holds:

(i) each $\mathcal{S}_i$ is either a linear subspace of dimension $d_i = r_i + p - 1$ (for $e^i_{r_i} > 1$) or a quadrics of dimension $d_i - 1$ and corank $d_i + 1 - \#\{j \mid e^i_j = 1\}$ (for $e^i_{r_i} = 1$).

(ii) If $p > g$ (i.e. there are non-zero minimal indices), then in addition $\mathcal{S}_t = \mathcal{S}_{\overline{k}+1}$ is either a projective bundle of type $P(\epsilon_{g+1} \dots \epsilon_p)$ (for $g = 0$) or a join variety of type $J(\epsilon_{g+1} \dots \epsilon_p; g - 1)$ (for $g > 0$).

Remark 3.3.21. In [10] Dimca stated that in the above proposition the t components were the irreducible ones but this is not true in general: Newstead[5] remarked that such components could be reducible. In this work we also notice this when studying the pencils of quadrics in $\mathbb{P}^2$ (see example 3.3.24): indeed by proposition 3.3.20 the pencil [(1 1) 1] has $t = 1$ but from table (3.3) we know that its base locus has two irreducible singular components (namely, the two double points). So it is important to keep in mind that the components $\mathcal{S}_1, \dots, \mathcal{S}_t$ in proposition 3.3.20 may be not irreducible, hence one may refine them. However, as Newstead noticed, the next final result still holds.

[5] Mathematical Review MR0708627 (85e:14011), report to [10]

Proposition 3.3.20 fills out all the informations in the Segre symbol of the pencil with respect to the minimal indices. Thus we can conclude the geometric classification of pencils of quadrics.

Theorem 3.3.22. Two pencils of quadrics $\mathcal{P}$ and $\mathcal{P}'$ are equivalent (up to GL-action) if and only if

(i) the lines $L_{\mathcal{P}}, L_{\mathcal{P}'} \subset \mathbb{P}W$ have similar positions;

(ii) the irreducible components of $\mathrm{Sing}(V(\mathcal{P}))$ and $\mathrm{Sing}(V(\mathcal{P}'))$ are isomorphic.

Proof. The “if” condition directly follows by propositions 3.3.4 and 3.3.20.
Let us prove the “only if” condition. By proposition 3.3.4 the hypothesis (i) implies that the two pencils have same continuous moduli, same multiplicities (of the roots) and same number of minimal indices $p = p'$, hence they have Segre symbols

$$\Sigma(\mathcal{P}) = \left[(e_1^1, \ldots, e_{r_1}^1)\ldots(e_1^k, \ldots, e_{r_k}^k); \epsilon_{g+1}, \ldots, \epsilon_p; g\right]$$

$$\Sigma(\mathcal{P}') = \left[(e_1^1, \ldots, e_{r_1}^1)\ldots(e_1^k, \ldots, e_{r_k}^k); \epsilon'_{g'+1}, \ldots, \epsilon'_p; g'\right]$$

By hypothesis (ii) the singular parts of the base loci of the two pencils have the same number of irreducible components, hence by proposition 3.3.20 either $p-g = p-g' = 0$ (i.e. both pencils have no non-zero minimal indices) or $p-g, p-g' > 0$ (i.e. both pencils have non-zero minimal indices). In the first case we have $p = g = g'$ and we conclude; in the latter case we have $p > g, g'$, hence by proposition 3.3.20(ii) it holds either (for $g = g' = 0$)

$$P(\epsilon_{g+1}\ldots\epsilon_p) \simeq \mathcal{S}_t \overset{\text{hp}}{\simeq} \mathcal{S}'_t \simeq P(\epsilon'_{g'+1}\ldots\epsilon'_p) \tag{3.12}$$

or (for $g, g' > 0$)

$$J(\epsilon_{g+1}\ldots\epsilon_p; g-1) \simeq \mathcal{S}_t \overset{\text{hp}}{\simeq} \mathcal{S}'_t \simeq J(\epsilon'_{g'+1}\ldots\epsilon'_p; g'-1) \tag{3.13}$$

If $g = g' = 0$, by lemma 3.3.15 the isomorphism (3.12) is actually an isomorphism of projective bundles and by corollary 3.3.17 we have

$$\mathbb{P}\left(\bigoplus_{i=1}^{p} \mathcal{O}(\epsilon_i)\right) \simeq \mathcal{S}_t \simeq \mathcal{S}'_t \simeq \mathbb{P}\left(\bigoplus_{i=1}^{p} \mathcal{O}(\epsilon'_i)\right)$$

hence by 3.3.14 there exists $M \in \mathbb{Z}$ such that $\epsilon_i - \epsilon'_i = M$ for all i, but by the restraints in (2.14) it must be $\epsilon_i = \epsilon'_i$ for all i, that is the thesis.

Else if $g, g' > 0$, by lemma 3.3.19 the equation (3.13) implies $g = g'$ and $\epsilon_i = \epsilon'_i$ for all i and this concludes. □

Remark 3.3.23. It is worth underlining that regular pencils may have a singular part $\mathrm{Sing}(V(\mathcal{P}))$ too. The number of its irreducible components is lower-bounded by $t = \overline{k}$.

Example 3.3.24 (Projective plane, singularities of base loci)**.** By remark 3.3.8 in $\mathbb{P}^2$ the position of the line $L_{\mathcal{P}}$ is enough to determine the whole Segre symbol even for singular pencils (table 3.4). However we now analyze $\mathrm{Sing}(V(\mathcal{P}))$ for all pencils in the plane.

Let $\mathcal{P}$ be the pencil with Segre symbol $[;1;]$: from table 3.4 we know its base locus $V(\mathcal{P})$ is given by the line (x) and the simple point (y, z). By proposition 3.3.20 we know that $\mathrm{Sing}(V(\mathcal{P}))$ has at least one component (since $p = 1 > g = 0, k = 0$ and $t = \overline{k} + 1 = 1$): let $\mathrm{Sing}(V(\mathcal{P})) = \mathcal{S}_1$. Since $g = 0$, by proposition 3.3.20(ii) it follows that $\mathcal{S}_1$ is isomorphic to the projective bundle

$$\mathcal{S}_1 \simeq P(\epsilon_1) = P(1) = \mathbb{P}^1$$

hence it is a projective line (actually, the line (x)).

The two remaining pencils have both Segre symbol of type $[*;;1]$, that is $p = g = 1$: by proposition 3.3.20 the number of irreducible components of $\mathrm{Sing}(V(\mathcal{P}))$ is at least $t = \max\{\overline{k}, 1\} = 1$ for both pencils. Neverthless the two varieties are not isomorphic by proposition 3.3.20(i): for $[1\ 1;;1]$ the singular part has dimension $r_i + p - 2 = 0$, in fact it is the quadruple point (x^2, y^2); for $[2;;1]$ the singular part has dimension $r_i + p - 1 = 1$ and it is the line (x).

By remark 3.3.23 we know that every regular pencil in $\mathbb{P}^2$ has at least $t = \overline{k}$ singular irreducible components. By table 3.3 we know that, when a singularity exists, it is 0-dimensional: indeed by computing such dimensions by proposition 3.3.20(i) one always gets dimension 0. Moreover, the components obtained by proposition 3.3.20 are irreducible for all regular pencils but [(1 1) 1] (see remark 3.3.21).
In the following table we resume the informations about $\mathrm{Sing}(V(\mathcal{P}))$ for all pencils in $\mathbb{P}^2$: the last column comes from a schematic study as in tables 3.3 and 3.4.

$\Sigma(\mathcal{P})$	$\overline{k}$	t	d_i	$\mathrm{Sing}(V(\mathcal{P}))$
[1 1 1]	0	0		$\emptyset$
[2 1]	1	1 (irred.)	0	one double point
[(1 1) 1]	1	1 (reducible)	0	two double points
[3]	1	1 (irred.)	0	one triple point
[(2 1)]	1	1 (irred.)	0	one (curv.) quadruple point
$[;1;]$	0	1 (reducible)	1	a line and a disjoint point
$[1\ 1;;1]$	0	1 (irred.)	0	one (non-curv.) quadruple point
$[2;;1]$	1	1 (reducible)	1	a line with embedded double point

(3.14)

Remark 3.3.25. By remark 3.3.8 we know that in $\mathbb{P}^2$ and $\mathbb{P}^3$ the position of the line $L_{\mathcal{P}}$ of a given pencil of quadrics $\mathcal{P}$ completely determines its Segre symbol $\Sigma(\mathcal{P})$. Since the components of $\mathrm{Sing}(V(\mathcal{P}))$ given by proposition 3.3.20 are (schematically) reduced, they are not enough to completely determine the pencil up to equivalence: indeed from table (3.14) it is clear that the components Dimca refers to are just 0-dimensional and by proposition 3.3.20 we can not distinguish neither the number nor the multiplicity of such singular points.
But if one looks at the base locus $V(\mathcal{P})$ with its projective scheme structure (as we did for table (3.3)), then for $\mathbb{P}^2$ the base locus is enough as well to completely determine the pencil up to equivalence: indeed we already discussed the multiplicities of the singular points in (3.3) and (3.4). The same holds for regular pencils in $\mathbb{P}^3$ as it follows by tables (3.6) and (3.7).

Chapter 4

Tensor rank decomposition

In this chapter we introduce different notions of rank for tensors and the algebraic-geometric objects related to them. In the first section we show how the notion of rank changes from $V_1 \otimes V_2$ to $V_1 \otimes \ldots \otimes V_d$ for $d \geq 3$ and we define border *and* symmetric rank*. In the second section we focus on tensors in $\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$, called* 2-slice tensors*, we introduce the* GL*-equivalence and we unveil that these tensors correspond to matrix pencils. In this perspective the Kronecker form allows to determine their ranks and we show this in two different ways: by a direct combinatorial approach and by applying the* discrete Fourier transform*. In the last section we determine the* partially-symmetric rank *of tensors in $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$ which are pencils of quadrics in $\mathbb{P}^m_{\mathbb{K}}$.*

Let $\mathbb{K}$ be an algebraically closed field with characteristic 0.
We will work with finite dimensional vector spaces over $\mathbb{K}$.

4.1 Ranks of tensors

4.1.1 Rank, multilinear rank and Segre varieties

The case $V \otimes W$. Let us fix $V \simeq \mathbb{K}^m$ and $W \simeq \mathbb{K}^n$ and let us consider their tensor product $V \otimes W$. We recall that the *decomposable* tensors are the ones in $\mathrm{Im}(g : V \times W \to V \otimes W)$. We use these tensors to define the rank over $V \otimes W$.

Proposition 4.1.1. Let $T = \sum_{i,j} a_{ij} v_i \otimes w_j \in V \otimes W \setminus \{0\}$. Then

$$T \text{ is decomposable} \iff \mathrm{Rk}(a_{ij}) = 1$$

Proof. We note that $\mathrm{Rk}(a_{ij}) = 1$ if and only if $a_{ij} = x_i y_j$ for $i = 1 : m$ and $j = 1 : n$ since rank-1-matrices can be always written as product of a column-vector for a row-vector:

$$\mathrm{Rk}(a_{ij}) = 1 \iff (a_{ij}) = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix}$$

If T is decomposable then it can be written as $T = v \otimes w$ whose corresponding matrix M_T has rank 1 and by acting with $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ on $V \otimes W$ we have $M_T = P(a_{i,j})Q$: since the rank is invariant under this action we have $\mathrm{Rk}(a_{ij}) = 1$. Conversely, if $(a_{i,j})$ has rank 1, then $a_{i,j} = x_i y_j$ for all i, j and for suitable $x_i, y_j \in \mathbb{K}$, hence $T = \sum_{i,j} a_{ij} v_i \otimes w_j = (\sum_i x_i v_i) \otimes (\sum_j y_j w_j)$ is decomposable. □

Since matrices of rank r are sum of r matrices of rank 1, the previous lemma makes sure the following definition is a good one.

Definition. The **rank** of a tensor in $V \otimes W$ is the minimum number of decomposable tensors in which it can be written as sum.

Proposition 4.1.2. The set $\mathrm{Im}(g)$ of decomposable tensors in $V \otimes W$ is an algebraic variety.

Proof. Decomposable tensors have rank 1, thus they are described by polynomial equations given by their 2×2 minors. □

Remark 4.1.3. Since $g^{-1}(0) = \Big(V \times \{0\}\Big) \cup \Big(\{0\} \times W\Big)$ we can (well) define the projective version of tensor product as

$$\begin{array}{rccc} \mathbb{P}g: & \mathbb{P}V \times \mathbb{P}W & \to & \mathbb{P}(V \otimes W) \\ & ([v], [w]) & \mapsto & [v \otimes w] \end{array}$$

Moreover, since $\forall v \otimes w \neq 0$ it holds $g^{-1}(v \otimes w) \simeq \mathbb{K}^\times$, the map $\mathbb{P}g$ is injective: actually this is an embedding, called **Segre embedding**, and its image

$$\mathrm{Im}(\mathbb{P}g) = \mathrm{Seg}(\mathbb{P}V \times \mathbb{P}W) \subset \mathbb{P}(V \otimes W)$$

is a projective variety, called **Segre variety**, defined by the homogeneous quadrics given by 2×2 minors. These definitions will be extended to a general tensor product $V_1 \otimes \ldots \otimes V_d$.

Consider the isomorphism of $\mathbb{K}$-vector spaces $V \otimes W \simeq \mathrm{Hom}_{\mathbb{K}}(V^\vee, W)$: let $f \in \mathrm{Hom}_{\mathbb{K}}(V^\vee, W)$ and $T \in V \otimes W$ corresponding each other via the above isomorphism.

Proposition 4.1.4. Let $\mathrm{Rk}(f)$ be the dimension of $\mathrm{Im}(f)$ as linear map (or equivalently the rank of f seen as matrix) and let $\mathrm{Rk}(T)$ be the rank of the tensor as minimum number of decomposable summands. Then

$$\mathrm{Rk}(f) = \mathrm{Rk}(T)$$

Proof. Let $\mathrm{Rk}(T) = R$ and let $T = \sum_{i=1}^{R} v_i \otimes w_i$ be sum of R decomposable tensors. Since each decomposable tensor $v_i \otimes w_j$ corresponds to a rank-1 matrix, the tensor T corresponds to a sum of R rank-1 matrices, hence $\mathrm{Rk}(f) \leq R$. Conversely, by Gaussian elimination $A \cdot f \cdot B = \begin{bmatrix} I_{\mathrm{Rk}(f)} & 0 \\ 0 & 0 \end{bmatrix}$ for suitable A, B matrices, hence f is sum of $\mathrm{Rk}(f)$ rank-1 matrices, hence $R \leq \mathrm{Rk}(f)$. □

The case $V_1 \otimes \ldots \otimes V_d$ for $d \geq 3$. Let $V_1, \ldots, V_d$ be $\mathbb{K}$-vector spaces with $\dim_{\mathbb{K}} V_i = m_i$ and let $V_1 \otimes, \ldots \otimes V_d$ be their tensor product. We start by defining the rank of a tensor and by characterizing the decomposable tensors.

Definition. The **rank** of a tensor in $V_1 \otimes \ldots \otimes V_d$ is the minimum number of decomposable tensors in which it can be written as sum. In particular, decomposable tensors have rank 1.

We recall that $V_1 \otimes \ldots \otimes V_d \simeq \mathrm{Hom}_{\mathbb{K}}(V_1^\vee \otimes \ldots \otimes \widehat{V_i^\vee} \otimes \ldots \otimes V_d^\vee, V_i)$ for all $i = 1 : d$, where the *hat* denotes the factor is missing: this identification is given by corresponding to a tensor its flattening with respect to V_i of the coordinate matrix. Hence for all $i = 1 : d$ each tensor $T \in V_1 \otimes \ldots \otimes V_d$ corresponds to a linear map $\varphi_{T,i} \in \mathrm{Hom}_{\mathbb{K}}(V_1^\vee \otimes \ldots \otimes \widehat{V_i^\vee} \otimes \ldots \otimes V_d^\vee, V_i)$.

We have to remark one interesting thing: if in the case $d = 2$ the isomorphisms $V \otimes W \simeq \mathrm{Hom}_{\mathbb{K}}(W^\vee, V) \simeq \mathrm{Hom}_{\mathbb{K}}(V^\vee, W)$ leave the rank unchanged (since in the second isomorphism we are just transposing), in the more general case $d \geq 3$ the isomoprhisms $V_1 \otimes \ldots \otimes V_d \simeq \mathrm{Hom}_{\mathbb{K}}(V_1^\vee \otimes \ldots \otimes \widehat{V_i^\vee} \otimes \ldots \otimes V_d^\vee, V_i)$ for all $i = 1 : d$ may let the rank change between the different flattenings. Hence it makes sense to give another definition of rank of a tensor.

Definition. Let $T \in V_1 \otimes \ldots \otimes V_d$ be a tensor and for all $i = 1 : d$ let T_i be the flattening of T with respect to V_i, i.e. the hypermatrix corresponding to $\varphi_{T,i} : V_1^\vee \otimes \ldots \otimes \widehat{V_i^\vee} \otimes \ldots \otimes V_d^\vee \to V_i$. We define the **multilinear rank** (or *multi-rank*) of T as

$$\mathrm{multRk}(T) = (\mathrm{Rk}\, T_1, \ldots, \mathrm{Rk}\, T_d) = \Big(\dim \mathrm{Im}(\varphi_{T,1}), \ldots, \dim \mathrm{Im}(\varphi_{T,d})\Big)$$

Example 4.1.5. Consider the tensor in the example 1.2.2:

$$T = u_1 \otimes v_1 \otimes w_1 + u_2 \otimes v_2 \otimes w_1$$

We explicited its flattenings and it is clear that its multi-rank is $\mathrm{multRk}(T) = (2, 2, 1)$ while its rank is 2.

We note that, while for $d = 2$ it holds $\mathrm{Rk}(T) \leq \min(\dim V_1, \dim V_2)$, for $d \geq 3$ the previous inequality does not hold anymore: in general $\mathrm{Rk}(T)$ may be larger than $\dim V_i$ for all $i = 1 : d$ (see table (4.2)). However if $\mathrm{multRk}(T) = (1, \ldots, 1)$, then $\mathrm{Rk}(T) = 1$ as declared in the following result.

Proposition 4.1.6. A tensor $T \in V_1 \otimes \ldots \otimes V_d$ is decomposable if and only if $\forall i = 1 : d$ the corresponding map $\varphi_{T,i} : V_1^\vee \otimes \ldots \otimes \widehat{V_i^\vee} \otimes \ldots \otimes V_d^\vee \to V_i$ has rank 1.

Proof. We note that, if the condition of having rank 1 is verified for $d-1$ values of i, then it is verified for the d-th value.
If T is decomposable, then $T = v_1 \otimes \ldots \otimes v_d$ and for all $i = 1 : d$ we have $\mathrm{Im}(\varphi_{T,i}) \subseteq \langle v_i \rangle$.
Conversely, if $\dim \mathrm{Im}(\varphi_{T,i}) = 1$ for all $i = 1 : d$, we get $v_i \in V_i$ such that $\langle v_i \rangle = \mathrm{Im}(\varphi_{T,i})$.
We claim that $T = v_1 \otimes \ldots \otimes v_d$ up to scalar. For all i we extend v_i to a basis of V_i and we represent each $\varphi_{T,i}$ as flattening of T with respect to these basis: by construction of v_i each flattening has only one non-zero coordinate in the first entry, hence T has rank 1 and $T = v_1 \otimes \ldots \otimes v_d$. □

Corollary 4.1.7 (Segre theorem)**.** The set of decomposable tensors

$$\mathrm{Seg}(\mathbb{P}V_1 \times \ldots \times \mathbb{P}V_d) = \mathrm{Im}\left(\mathbb{P}g : \mathbb{P}V_1 \times \ldots \times \mathbb{P}V_d \to \mathbb{P}(V_1 \otimes \ldots \otimes V_d)\right)$$

is an algebraic variety, called **Segre variety**.

Proof. The set of decomposable tensors is determined by the equations given by the vanishing of all 2×2 minors in the flattenings. □

Even if in general the multi-rank does not tell us the exact rank of a given tensor, it gives us some bounds.

Lemma 4.1.8. Let $T \in V_1 \otimes \ldots \otimes V_d$ be a tensor with $\mathrm{multRk}(T) = (r_1, \ldots, r_d)$. Then for all $i, j = 1 : d$ (even $i \neq j$) it holds

$$r_i \leq \mathrm{Rk}(T) \leq \prod_{k \neq j} r_k$$

Proof. Let $r = \mathrm{Rk}(T)$. We write T as sum of decomposables

$$T = \sum_{h=1:r} v_h^{(1)} \otimes \ldots \otimes v_h^{(d)}$$

where $v_h^{(i)} \in V_i$. For all $i = 1 : d$ let T_i be the flattening of T with respect to V_i

$$T_i : V_1^\vee \otimes \ldots \otimes \widehat{V_i^\vee} \otimes \ldots \otimes V_d^\vee \longrightarrow V_i$$

Then $\mathrm{Im}(T_i) \subseteq \langle v_1^{(i)}, \ldots, v_r^{(i)} \rangle$, hence $r_i \leq r$. Moreover, the above restraint tells us we can express T using only r_i elements of a basis in V_i for all $i = 1 : d$: hence, up to fixing all $i = 1 : d$ except $j \in \{1, \ldots, d\}$ and joining the summands in V_j with respect to the $\prod_{k \neq j} r_k$ elements in $V_1 \otimes \ldots \otimes \widehat{V_j} \otimes \ldots \otimes V_d$, we can write T as sum of $\prod_{k \neq j} r_k$ decomposable tensors, that is $r \leq \prod_{k \neq j} r_k$. □

Note: We said that for $d = 2$ the rank of T is invariant when T is seen as homomorphism from the dual of a space to the other space. This comes out from the above lemma too: indeed if $d = 2$ the lemma says that $r_1 \leq r \leq r_2$ and $r_2 \leq r \leq r_1$, hence $r = r_1 = r_2$.

Corollary 4.1.9. If $\mathrm{Rk}(T_i) = 1$ for all i but a $j \in \{1, \ldots, d\}$, then $\mathrm{Rk}(T_j) = \mathrm{Rk}(T) = 1$.

The case $\mathbb{K}^a \otimes \mathbb{K}^b \otimes \mathbb{K}^c$. We now focus on the case $d = 3$: it is rich enough to analyze differences from the case $d = 2$ but even simple enough to make counts by dirting our hands.
Given $T \in \mathbb{K}^a \otimes \mathbb{K}^b \otimes \mathbb{K}^c$, by lemma 4.1.8 we know that its multi-rank satisfies

$$\mathrm{Rk}(T_i) \leq \mathrm{Rk}(T_j)\,\mathrm{Rk}(T_k) \quad \forall i, j, k \in \{a, b, c\} \tag{4.1}$$

Obviously we have the additional conditions

$$\mathrm{Rk}(T_a) \leq a\ ,\ \mathrm{Rk}(T_b) \leq b\ ,\ \mathrm{Rk}(T_c) \leq c$$

One natural question is: *given* $(r_1, r_2, r_3) \in \mathbb{N}^3$ *a triple of natural numbers satisfying the condition* (4.1), *does exist a tensor* T *(in a suitable tensor product space) such that* $\mathrm{multRk}(T) = (r_1, r_2, r_3)$*?* In 2011 Carlini and Kleppe answered affirmatively [8]:

Theorem. The conditions (4.1) are necessary and sufficient for the existence of a T having prescribed multi-rank.

4.1.2 Border rank and secant varieties

We start from an example.

Example 4.1.10. Consider in $\mathbb{K}^2 \otimes \mathbb{K}^2 \otimes \mathbb{K}^2$ the tensor

$$T = (a_1 + a_2) \otimes b_1 \otimes c_1 + a_1 \otimes b_1 \otimes c_2 + a_1 \otimes b_2 \otimes c_1$$

T has rank 3 but it can be approximated by rank-2 tensors:

$$T(\epsilon) = \frac{1}{\epsilon}\Big[(\epsilon - 1)a_1 \otimes b_1 \otimes c_1 + (a_1 + \epsilon a_2) \otimes (b_1 + \epsilon b_2) \otimes (c_1 + \epsilon c_2)\Big]$$

This means that T is not a rank-2 tensor but it is the limit of rank-2 tensors, hence it is in some closure of the latter ones.

Definition. A tensor T has **border rank** r if it is limit of rank-r tensors but it is not a limit of rank-s tensors for any $s < r$. We denote the border rank of T by $\underline{\mathrm{Rk}}(T)$. In particular, $\underline{\mathrm{Rk}}(T) \leq \mathrm{Rk}(T)$.

The notion of border rank is related to the following variety.

Definition. Let $X \subset \mathbb{P}^N$ be an algebraic variety. Its **r-th secant variety** $\sigma_r(X)$ is the Zariski closure of all points in $\mathbb{P}^N$ which are linear combination of r points from X:

$$\sigma_r(X) = \overline{\{\sum_{i=1}^{r} P_i \mid P_i \in X\}} = \overline{\bigcup_{P_1,\dots,P_r \in X} \langle P_1, \dots, P_r \rangle}$$

The secant varieties of an algebraic variety $X \subset \mathbb{P}^N$ form an ascending chain

$$X = \sigma_1(X) \subseteq \dots \subseteq \sigma_R(\mathcal{X}) \subseteq \dots$$

If the chain stabilizes at $\mathbb{P}^N$, the minimum R such that $\sigma_R(X) = \mathbb{P}^N$ is said the **generic rank** of X.

Now let $X = \mathrm{Seg}(\mathbb{P}V_1 \times \dots \times \mathbb{P}V_d)$ be the Segre variety of $\mathbb{P}V_1 \otimes \dots \otimes \mathbb{P}V_d$. Then its r-th secant variety is

$$\sigma_r(\mathrm{Seg}(\mathbb{P}V_1 \times \dots \times \mathbb{P}V_d)) = \overline{\{[T] \in \mathbb{P}(V_1 \otimes \dots \otimes V_d) \mid \underline{\mathrm{Rk}}(T) \leq r\}}$$

As observed in the previous example, by working with secant varieties we work with a Zariski closure, hence there are tensors which are limit of rank-r tensors but that can have a higher rank.

The case $\mathbb{K}^2 \otimes \mathbb{K}^2 \otimes \mathbb{K}^2$. We now present the classification of tensors in $\mathbb{K}^2 \otimes \mathbb{K}^2 \otimes \mathbb{K}^2$ with respect to their rank r and multi-rank (r_1, r_2, r_3).

symbols	r_1	r_2	r_3	r	representatives
A	1	1	1	1	$a_0 \otimes b_0 \otimes c_0$
$B1$	1	2	2	2	$a_0 \otimes b_0 \otimes c_0 + a_0 \otimes b_1 \otimes c_1$
$B2$	2	1	2	2	$a_0 \otimes b_0 \otimes c_0 + a_1 \otimes b_0 \otimes c_1$
$B3$	2	2	1	2	$a_0 \otimes b_0 \otimes c_0 + a_1 \otimes b_1 \otimes c_0$
W	2	2	2	3	$a_0 \otimes b_0 \otimes c_1 + a_0 \otimes b_1 \otimes c_0 + a_1 \otimes b_0 \otimes c_0$
G	2	2	2	2	$a_0 \otimes b_0 \otimes c_0 + a_1 \otimes b_1 \otimes c_1$

(4.2)

Clearly the tensors of types $A, B1, B2$ and $B3$ are forced to be the ones of such multi-ranks and ranks because of the conditions (4.1). We note that the tensor in example 4.1.10 is a representative for the class W. This classification can also be obtained by the Kronecker form of pencils as we will see in the next sections (see table (4.12)).

4.1.3 Symmetric rank and Veronese varieties

Proposition 4.1.11. The locus of decomposable symmetric tensors

$$\nu_d(\mathbb{P}V) = \Big(\mathrm{Seg}(\mathbb{P}V \times \dots \times \mathbb{P}V)\Big) \cap \mathbb{P}(\mathrm{Sym}^d V)$$

is isomorphic to $\mathbb{P}V$ and it is called **d-Veronese variety** of $\mathbb{P}V$.

Proof. The isomorphism is given by the *d-th Veronese map*

$$\begin{array}{rccc} \nu_d: & \mathbb{P}V & \longrightarrow & \mathbb{P}(\operatorname{Sym}^d V) \\ & [v] & \mapsto & [v^{\otimes d}] \end{array}$$

which in coordinates is $[x_0 : \ldots : x_m] \mapsto [x_0^d : x_0^{d-1}x_1 : \ldots : x_{m-1}x_m^{d-1} : x_m^d]$. □

Since hyperplanes in $\mathbb{P}(\operatorname{Sym}^d V)$ define degree-d polynomials on $\mathbb{P}V$, the Veronese varieties give the following geometric reinterpretation:

Proposition 4.1.12. The d-th Veronese variety is not contained in a hyperplane.

Proof. If it was so, the hyperplane would induce the zero polynomial, hence the coefficients over any infinite field would be zero, but this is a contraddiction in the projective space. □

Corollary 4.1.13. Any homogeneous polynomial $f \in \operatorname{Sym}^d V$ has decomposition of the form

$$f = \sum_{i=1}^{r} c_i l_i^d$$

where l_i are linear forms and $c_i \in \mathbb{K}$. When the number r of summands is as minimum as possible, the decomposition is said to be the **Waring decomposition** of f.

Note: If $\mathbb{K} = \mathbb{C}$ we may assume $f = \sum_{i=1}^{r} l_i^d$, that is $c_i = 1$ for all i.

Definition. The **symmetric rank** of a homogeneous polynomial (or equivalently, of a symmetric tensor) $f \in \operatorname{Sym}^d V$ is

$$\operatorname{symRk}(f) = \min \left\{ r \;\middle|\; f = \sum_{i=1}^{r} c_i l_i^d \right\}$$

Note: $\operatorname{symRk}(f) = 1 \iff f \in \nu_d(\mathbb{P}V) \iff f = l^d$.

Question: *given $f \in \operatorname{Sym}^d V$, how to detect if $f = l^d$?*
We observe that $f = l^d$ if and only if $\frac{\partial f}{\partial x_i}$ are two by two linearly dependent, that is if and only if the matrix whose rows are given by the coefficients of these partials has rank 1.
To light up the notation let us write ∂_i instead of $\frac{\partial}{\partial x_i}$. Then ∂_i are the coordinates on $(\operatorname{Sym}^d V)^\vee$ dual to the coordinates x_i on V, that is $\partial_i x_j = \delta_{ij}$.
Consider the map

$$\begin{array}{ccc} \left(\operatorname{Sym}^e V\right)^\vee \otimes \operatorname{Sym}^d V & \longrightarrow & \operatorname{Sym}^{d-e} V \\ (\partial, f) & \mapsto & \partial \cdot f \end{array} \tag{4.3}$$

hence, by fixing f, the map

$$\begin{array}{cccc} C_{e,f}: & \left(\operatorname{Sym}^e V\right)^\vee & \longrightarrow & \operatorname{Sym}^{d-e} V \\ & \partial & \mapsto & \partial \cdot f \end{array} \tag{4.4}$$

Definition. The map

$$\begin{array}{cccc} C_{e,f}: & \operatorname{Sym}^e V^\vee & \longrightarrow & \operatorname{Sym}^{d-e} V \\ & g & \mapsto & g \cdot f \end{array}$$

is called e-th **catalecticant map** for f.

Note: For $e = 1$ we have the flattening $C_{1,f} : V^\vee \to \operatorname{Sym}^{d-1} V$.

Proposition 4.1.14. Let $f \in \operatorname{Sym}^d V$. Then $f = l^d \iff \operatorname{Rk}(C_{1,f}) = 1$.

Proof. If $f = l^d$, then $C_{1,f}(\partial_i) = \partial_i(l^d) = d(\partial_i l)l^{d-1}$, hence $\operatorname{Im}(C_{1,f}) = \langle l^{d-1} \rangle$ has dimension 1. Conversely, if $\operatorname{Rk}(C_{1,f}) = 1$, then there exist $\lambda_0, \dots, \lambda_m \in \mathbb{K}$ such that $0 = \sum_{i=0}^m \lambda_i \partial_i f = \left(\sum \lambda_i \partial_i\right) f$, then by transposing $C_{1,f}$ we have $\partial f = c \cdot l$ for some linear form l and some $c \in \mathbb{K}$. Up to acting by $\operatorname{GL}(V)$ we may assume $l = x_0$: then only the monomials x_0^d can appear in f otherwise by applying ∂ there would be more than x_0 in l. Then $f = l^d$. □

The rank of $C_{1,f}$ is the minimum number of variables such that f can be written in after a linear change of coordinates: it is said **essential rank** of f.

Corollary 4.1.15. Let $f \in \operatorname{Sym}^d V$. Then $\operatorname{Rk}(C_{1,f}) \leq \operatorname{symRk}(f)$.

Remark 4.1.16. We have to emphasize that *symmetric rank* and *rank of a symmetric tensor* are different: the first one is about the decomposition of f in $\operatorname{Sym}^d V$, the latter in $V^{\otimes d}$. In general, given $f \in \operatorname{Sym}^d V$, it holds $\operatorname{Rk}(f) \leq \operatorname{symRk}(f)$. However for $\dim V = 2$ equality always holds.

4.2 GL-equivalence in $\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$

Now we focus on tensors in $\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$. In the perspective of tensors as cube-matrices, the $2 \times m \times n$-tensors are said **2-slice tensors** of size $m \times n$ and so they can be represented (with respect to a given basis) by a cube-matrix given by two matrices of size $m \times n$ suitably put parallel to each other: for instance in the previous example we considered a 2-slice tensor of size 2. But these 2-slice tensors are not a novelty for us since we met them throughout the previous chapters: indeed 2-slice tensors can be seen as projective matrix pencils. Let us better investigate this identification.

Let $T \in \mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$ be a 2-slice tensor of size $m \times n$. Through the identification

$$\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n \xrightarrow{\simeq} \mathrm{Bil}(\mathbb{K}^m, \mathbb{K}^n; \mathbb{K}^2)$$

T corresponds to a unique bilinear map $\phi_T : (v, w) \mapsto \phi_T(v, w) = (a, b) \in \mathbb{K}^2$. With respect to some basis, such bilinear map ϕ_T can be represented by a pair of matrices (A, B) in $(\mathbb{K}^{m\times n})^2$ such that

$$^t v \cdot A \cdot w = a \quad , \quad ^t v \cdot B \cdot w = b$$

Then we consider the pencil of size $m \times n$ defined by this pair of matrices, i.e. $\mu A + \lambda B$.

In Chapter 2 we observed that we could extend the group action in (2.1) to

$$\begin{array}{ccc} \mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K}) & \longrightarrow & \mathrm{Aut}\Big(\mathfrak{M}_{m\times n}(\mathbb{K}[\mu, \lambda]_1)\Big) \\ \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}, P, Q\right) & \mapsto & \Big(\ \mu A + \lambda B \mapsto P \cdot \Big((a\mu + b\lambda)A + (c\mu + d\lambda)B\Big) \cdot {}^t Q\ \Big) \end{array} \tag{4.5}$$

by also acting on the variables μ and λ by linear transformation of $\mathbb{K}^2$.

Now that we know that matrix pencils can be identified by 2-slice tensors, the action of the group $\mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ gains the different perspective

$$\begin{array}{ccc} \mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K}) & \longrightarrow & \mathrm{Aut}\Big(\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n\Big) \\ (M, P, Q) & \mapsto & \Big(\ u \otimes v \otimes w \mapsto Mu \otimes Pv \otimes Qw\ \Big) \end{array} \tag{4.6}$$

or

$$\begin{array}{ccc} \mathrm{SL}_2(\mathbb{K}) \times \mathrm{SL}_m(\mathbb{K}) \times \mathrm{SL}_n(\mathbb{K}) & \longrightarrow & \mathrm{Aut}\Big(\mathbb{P}(\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n)\Big) \\ ([M], [P], [Q]) & \mapsto & \Big(\ [u \otimes v \otimes w] \mapsto [Mu \otimes Pv \otimes Qw]\ \Big) \end{array} \tag{4.7}$$

where the above actions are defined on the decomposable tensors and extended by linearity.

Let us show that the actions defined on tensors and on pencils respectively are compatible with the correspondence between these two objects: actually it is enough to restrict our attention only to the action of $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ on $\mathbb{K}^m \otimes \mathbb{K}^n$ since $\mathrm{GL}_2(\mathbb{K})$ acts just on the component in $\mathbb{K}^2$ of the tensor (and just on the variables defining the pencil). Moreover we can choose to fix the canonical basis $(e_i)_{i=1:m}$ and $(e_i)_{i=1:n}$ for $\mathbb{K}^m$ and $\mathbb{K}^n$ respectively, and the canonical basis (E_{ij}) for $\mathfrak{M}_{m\times n}(\mathbb{K})$. Thus we ask for the following diagram to commute:

$$\begin{array}{ccc} u\otimes w=\sum_{i,j}T_{ij}e_i\otimes e_j & \longleftrightarrow & T=(T_{ij}) \\ \curvearrowright\big\downarrow & & \big\downarrow\curvearrowright \\ Pu\otimes Qw & \longleftrightarrow & P\cdot T\cdot{}^tQ \end{array}$$

Let $v\otimes w=(\sum_s v_se_s)\otimes(\sum_t w_te_t)=\sum_{s,t}v_sw_te_s\otimes e_t$, hence

$$v\otimes w\longleftrightarrow T=(T_{st})=(v_sw_t)=\sum_{s,t}v_sw_tE_{st}$$

By acting on T by (P,Q) we obtain

$$\begin{aligned} P\cdot T\cdot{}^tQ &= P\cdot\Big(\sum_{s,t}v_sw_tE_{st}\Big)\cdot{}^tQ=\sum_{s,t}v_sw_t(P\cdot E_{st}\cdot{}^tQ) \\ &= \sum_{s,t}v_sw_t\Big(\sum_{i,j}(P\cdot E_{ij}\cdot{}^tQ)_{ij}E_{ij})\Big) \\ &= \sum_{s,t}v_sw_t\Big(\sum_{i,j}\Big(\sum_{p,q}P_{ip}(E_{st})_{pq}({}^tQ)_{qj}\Big)E_{ij}\Big) \\ &= \sum_{s,t}v_sw_t\Big(\sum_{i,j}(P_{is}({}^tQ)_{tj})E_{ij}\Big) \\ &= \sum_{i,j}\Big(\sum_{s,t}P_{is}v_sQ_{jt}w_t\Big)E_{ij} \\ &= \sum_{i,j}\Big((Pv)_i(Qw)_j\Big)E_{ij} \quad\longleftrightarrow\quad Pv\otimes Qw \end{aligned}$$

and this concludes.

Definition. Two 2-slice tensors T_1 and T_2 of size $m\times n$ are said to be GL**-equivalent** if they are in the same orbit with respect to the group action (4.6). Equivalently, two pencils $\mu_1A_1+\lambda_1B_1$ and $\mu_2A_2+\lambda_2B_2$ of size $m\times n$ are said to be GL**-equivalent** if they are in the same orbit with respect to the group action (4.5). We denote this equivalence by

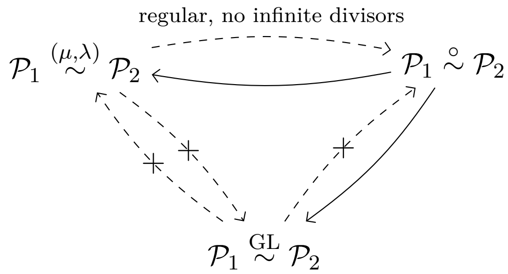
$$T_1\overset{\mathrm{GL}}{\sim}T_2 \quad , \quad \mu_1A_1+\lambda_1B_1\overset{\mathrm{GL}}{\sim}\mu_2A_2+\lambda_2B_2$$

Note: We defined λ-equivalence for λ-matrices. By homogenizing we may easily define (μ,λ)-equivalence in the same way: two (μ,λ)-matrices of size $m\times n$ are said to be (μ,λ)-equivalent if they are in the same orbit of the group action

$$\begin{array}{ccc} \mathrm{GL}_m(\mathbb{K}[\mu,\lambda])\times\mathrm{GL}_n(\mathbb{K}[\mu,\lambda]) & \longrightarrow & \mathrm{Aut}\Big(\mathfrak{M}_{m\times n}(\mathbb{K}[\mu,\lambda])\Big) \\ (P(\mu,\lambda),Q(\mu,\lambda)) & \mapsto & \Big(\ A(\mu,\lambda)\mapsto P(\mu,\lambda)\cdot A(\mu,\lambda)\cdot{}^tQ(\mu,\lambda)\ \Big) \end{array} \tag{4.8}$$

It is quite natural to ask ourselves how the three equivalences we defined for pencils behave with respect to each other: of course strict equivalence implies (μ, λ)-equivalence, but we have also already noticed that strict equivalence implies GL-equivalence since the equivalent classes with respect to strict equivalence are defined by the orbit of the action of the subgroup $\{I_2\} \times \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$; we also recall that by proposition 2.1.2 if $\mu A_1 + \lambda B_1 \overset{(\mu,\lambda)}{\sim} \mu A_2 + \lambda B_2$ are square pencils such that $\det B_i \neq 0$ (i.e. both pencils have no elementary divisor of the form μ^t), then $\mu A_1 + \lambda B_1 \overset{\circ}{\sim} \mu A_2 + \lambda B_2$.

Let us convince ourselves the following implications hold for $\mathcal{P}_1, \mathcal{P}_2$ projective pencils:



Clearly in general $\mathcal{P}_1 \overset{\mathrm{GL}}{\sim} \mathcal{P}_2$ doesn not imply $\mathcal{P}_1 \overset{\circ}{\sim} \mathcal{P}_2$: this holds only if the two pencils are defined by the same variables μ, λ, that is their equivalence is given by an element of $\{I_2\} \times \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$, but this hypothesis is too strong.
And even if the pencils are regular with no elementary divisors of the form μ^t, their being GL-equivalent still does not imply they are strictly equivalent: indeed in general by acting on $\mathbb{K}^2$ (even only by switching μ and λ by $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$) elementary divisors change, hence there is not strict equivalence. This also suggests us that the implication $\mathcal{P}_1 \overset{\mathrm{GL}}{\sim} \mathcal{P}_2 \Rightarrow \mathcal{P}_1 \overset{(\mu,\lambda)}{\sim} \mathcal{P}_2$ fails.

Finally, since invariant polynomials give a complete system of invariants of (μ, λ)-equivalence and the action of $\mathrm{GL}_2(\mathbb{K})$ changes them, being GL-equivalent does not even imply being (μ, λ)-equivalent.

4.2.1 Canonical form of GL-equivalence

In chapter 3 we determined a complete system of invariants of GL-equivalence for pencils of quadrics (or equivalently for symmetric 2-slice tensors), given by the Segre symbol of the pencil and a continuous modulus.
Now our goal is to determine a canonical form of GL-equivalence for general 2-slice tensors and to do this we start again from the Kronecker form of a pencil corresponding to its coordinate matrix (with respect to a given basis).

Let $T \in \mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$ be a 2-slice tensor and let $\mathcal{P}_T : \mu A + \lambda B$ be a pencil corresponding to T. Let $\mathcal{D}_T$ be the Kronecker form of the pencil $\mathcal{P}_T$ as in (2.16): since strict equivalence is a particular case of GL-equivalence, the pencil $\mathcal{D}_T$ is a good starting point for our goal; actually the pencil $\mathcal{D}_T$ gives us the main structure of the canonical form we are looking for since it comes from the action of the group $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ on the last two factors of the tensor space.
What is left is to understand how the action of $\mathrm{GL}_2(\mathbb{K})$ on the first factor may change the Kronecker form. But the action of $\mathrm{GL}_2(\mathbb{K})$ is just a linear transformation of the variables μ, λ which define the pencil $\mathcal{D}_T$: we already engaged this perspective in proposition 2.2.13, when we noticed that we can always assume a pencil to be without elementary divisors of the form μ^{u_i} (i.e. $\det B \neq 0$) up to apply a linear transformation over $\mathbb{K}^2$.

Let us resume what we have just observed: the canonical form of GL-equivalence of the 2-slice tensor T is determined by the Kronecker form of the pencil $\mathcal{P}_T$ (concerning the action of $\mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$) and the possible transformations over $\mathbb{P}^1$ (concerning the action of $\mathrm{GL}_2(\mathbb{K})$); the minimal indices (both for rows and columns) are invariant under the action of $\mathrm{GL}_2(\mathbb{K})$ since it preserves the singularity of the pencil; moreover the action of $\mathrm{GL}_2(\mathbb{K})$ changes the irreducible polynomials which define the elementary divisors (since it moves the roots of the determinant of $\mathcal{P}_T$ in $\mathbb{P}^1$) but without changing their multiplicity.

Then we are allowed to talk about the **minimal indices** of a 2-slice tensor. It only remains to find an invariant which replaces the elementary divisors since they are clearly not preserved by the action of $\mathrm{GL}_2(\mathbb{K})$.
For what observed above, let us assume $\mathcal{P}_T$ with no elementary divisor of the form μ^u. Let $\{(\lambda + a_k\mu)^{w_j^{(k)}}\}_{j,k}$ be its elementary divisors (where $w_j^{(k)}$ is the multiplicity of $\lambda + a_k\mu$ in the j-th invariant polynomial). We know that the integers $w_j^{(k)}$ are invariant under the action of $\mathrm{GL}_2(\mathbb{K})$, hence under the action of the whole group $\mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$. This allows us to give the following definition.

Definition. The **invariant degrees** of a 2-slice tensor are the integers $\{w_j^{(k)} \mid j, k\}$.

Corollary 4.2.1. Let $T \in \mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$ be a 2-slice tensor. Then the minimal indices and the invariant degrees of the tensor are GL-invariant.

In particular, we can define an extended version of the Segre symbol for general 2-slice tensors simply by making distinction between minimal indices for columns and for rows.

Definition. The GL-**symbol** of a 2-slice tensor T (or equivalently, of a matrix pencil $\mathcal{P}$) is the ordered sequence

$$\mathscr{S}(\mathcal{P}) = \Big[(w_1^{(1)}, \ldots, w_{r_1}^{(1)}) \ldots (w_1^{(k)}, \ldots, w_{r_k}^{(k)}); \epsilon_{g+1}, \ldots, \epsilon_p; \eta_{h+1}, \ldots, \eta_q; g; h\Big] \tag{4.9}$$

with the ordering

$$r_1 \geq \ldots \geq r_k \quad , \quad w_1^{(i)} \geq \ldots \geq w_{r_i}^{(i)} \quad , \quad \epsilon_{g+1} \leq \ldots \leq \epsilon_p \quad , \quad \eta_{h+1} \leq \ldots \leq \eta_q$$

where k is the number of distinct roots of the tensor, $w_i^{(j)}$ are their multiplicities, ϵ_s and g are the minimal indices for columns and η_t and h are the ones for rows.

Nevertheless the minimal indices and the invariant degrees are not enough to form a complete system of invariants for GL-equivalence for 2-slice tensors: indeed in general some relations between the roots (such as their disposition in the projective space) must be preserved. We solve this as we did for pencils of quadrics, that is by assigning to the pencil the **continuous modulus** of its roots (defined by (3.2)).

However it follows that a canonical form of GL-equivalence of a 2-slice tensor has associated pencil

$$\begin{bmatrix} 0_{h\times g} & & \\ & \operatorname{diag}(R_{\epsilon_{g+1}},\ldots,R_{\epsilon_p},{}^t R_{\eta_{h+1}},\ldots,{}^t R_{\eta_q}) & \\ & & \operatorname{diag}(\mathfrak{J}_{w_1^{(1)},a_1},\ldots\mathfrak{J}_{w_r^{(t)},a_t}) \end{bmatrix} \tag{4.10}$$

where R_{ϵ_i} and ${}^tR_{\eta_j}$ are as in (2.15) and $\mathfrak{J}_{w_h^{(k)},a_k}$ is the Jordan block of size $w_h^{(k)}$

$$\mathfrak{J}_{w_h^{(k)},a_k} = \begin{bmatrix} \lambda + a_k\mu & \mu & & \\ & \lambda + a_k\mu & \ddots & \\ & & \ddots & \mu \\ & & & \lambda + a_k\mu \end{bmatrix}$$

Remark 4.2.2. We say *a* canonical form and not *the* canonical form since in general we can not find a canonical sequence of roots $\{a_1,\ldots,a_s\}$ (we can always change them by a transformation on $\mathbb{P}^1$) but only give a class which represent them (i.e. the continuous modulus).

By remark 4.2.2 we deduce that in general there are infinitely many GL-orbits in $\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$. However in small dimensions things are so much easier.

4.2.2 GL-orbits in small dimensions

Theorem 4.2.3. The tensor space $\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$ has finitely many GL-orbits if and only if $m \leq 3$ or $n \leq 3$. In particular, in these cases a canonical sequence of roots is given by $\{0, 1, -1\}$.

Proof. Let us assume $m \leq 3$: then the worst case would be to have three different roots a_1, a_2, a_3 but we can always find a transformation of $\mathbb{P}^1$ which sends two of them to the values $0, 1$ and the third to the infinite point, hence the three of them to $0, 1, -1$. Conversely, let us assume $m, n \geq 4$: then it would be enough to consider the case of four different roots a_1, a_2, a_3, a_4 to convince ourselves[1] that there are infinite GL-orbits. □

We may also wonder which dimension a given GL-orbit in $\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$ has. Since by theorem 4.2.3 the number of orbits is finite only for $m \leq 3$ or $n \leq 3$, we focus on tensors in $\mathbb{K}^2 \otimes \mathbb{K}^2 \otimes \mathbb{K}^n$ and $\mathbb{K}^2 \otimes \mathbb{K}^3 \otimes \mathbb{K}^n$. Moreover, from the correspondence between 2-slice tensors and matrix pencils it makes sense to consider only the cases

$$m = 2 \;\wedge\; n = 2, 3, 4 \quad , \quad m = 3 \;\wedge\; n = 3, 4, 5, 6 \tag{4.11}$$

Given $\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$ as in (4.11) and T a tensor in it, the action (4.6) induces the map

$$\begin{array}{rccc} \gamma_T : & \mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K}) & \longrightarrow & \mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n \\ & G & \mapsto & G \cdot T \end{array}$$

whose image is exactly the GL-orbit of T. By deriving γ_T in the identity $I = (I_2, I_m, I_n)$, we get the linear map

$$\begin{array}{rccc} d(\gamma_T)_I : & \mathfrak{gl}_2(\mathbb{K}) \times \mathfrak{gl}_m(\mathbb{K}) \times \mathfrak{gl}_n(\mathbb{K}) & \longrightarrow & \mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n \\ & g & \mapsto & g \cdot T \end{array}$$

whose kernel is the Lie algebra of the stabilizer $\mathrm{stab}_{\mathrm{GL}}(T)$ and whose rank is the dimension of $\mathrm{orb}_{\mathrm{GL}}(T)$. Hence to compute the dimension of the orbit one may compute

$$\dim\big(\mathrm{orb}_{\mathrm{GL}}(T)\big) = \mathrm{Rk}\left(d(\gamma_T)_I\right) = 4 + m^2 + n^2 - \dim\Big(\ker\left(d(\gamma_T)_I\right)\Big)$$

Note: In Ch.6, §2.1 we implement the computation of the orbit dimensions on `Macaulay2`.

Example 4.2.4. Consider $T = \lambda \otimes b_1 \otimes c_1 \in \mathbb{K}^2 \otimes \mathbb{K}^2 \otimes \mathbb{K}^2$ with pencil $\begin{bmatrix} 0 & \\ & \lambda \end{bmatrix}$. It is decomposable, thus we already know its orbit is $\mathrm{Seg}(\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1)$ of dimension 4. Moreover, if the generic image $d(\gamma_T)_I(T) = (M, P, Q) \cdot T$ goes to 0, it means

$$(M_{21}\mu + M_{22}\lambda) \otimes b_1 \otimes c_1 + \lambda \otimes (P_{21}b_0 + P_{22}b_1) \otimes c_1 + \lambda \otimes b_1 \otimes (Q_{21}c_0 + Q_{22}c_1) = 0$$

that is only if $M_{21} = P_{21} = Q_{21} = (M_{22} + P_{22} + Q_{22}) = 0$, hence $\dim(\ker(d(\gamma_T)_I)) = 8$.

Remark 4.2.5. One can also geometrically determine the projective orbit closure of each tensor and its projective dimension in $\mathbb{P}(\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n)$ (see [6, §6] for details). The orbits in $\mathbb{K}^2 \otimes \mathbb{K}^3 \otimes \mathbb{K}^3$ and their dimensions are deeply studied by Parfenov [31].

[1] with a *cross-ratio* and homographies argument

We end this section with a description of all GL-orbits in the cases $m = n = 2$ (table (4.12)) and $m = n = 3$ (table (4.13)): ranks and border ranks in (4.12) follow by table (4.2) and example 4.1.10, while ranks in (4.13) follow by theorem 4.3.9.

$\mathcal{P}_T$	$\dim(\mathrm{orb}_{\mathrm{GL}}(T))$	Rk	$\underline{\mathrm{Rk}}$	T	symbols
$\begin{bmatrix} 0 & \\ & \lambda \end{bmatrix}$	4	1	1	$\lambda \otimes b_1 \otimes c_1$	A
$\begin{bmatrix} \lambda & \\ & \lambda \end{bmatrix}$	5	2	2	$\lambda \otimes b_0 \otimes c_0 + \lambda \otimes b_1 \otimes c_1$	$B1$
$\begin{bmatrix} \lambda & \mu \\ & \end{bmatrix}$	5	2	2	$\lambda \otimes b_0 \otimes c_0 + \mu \otimes b_0 \otimes c_1$	$B2$
$\begin{bmatrix} \lambda & \\ \mu & \end{bmatrix}$	5	2	2	$\lambda \otimes b_0 \otimes c_0 + \mu \otimes b_1 \otimes c_0$	$B3$
$\begin{bmatrix} \lambda & \mu \\ & \lambda \end{bmatrix}$	7	3	2	$\lambda \otimes b_0 \otimes c_0 + \lambda \otimes b_1 \otimes c_1 + {}$ $+ \mu \otimes b_0 \otimes c_1$	W
$\begin{bmatrix} \lambda & \\ & \mu \end{bmatrix}$	8	2	2	$\lambda \otimes b_0 \otimes c_0 + \mu \otimes b_1 \otimes c_1$	G

(4.12)

$\mathcal{P}_T$	$\dim(\mathrm{orb}_{\mathrm{GL}}(T))$	Rk	$\underline{\mathrm{Rk}}$	T
$\begin{bmatrix} \lambda & & \\ & \mu & \\ & & \lambda+\mu \end{bmatrix}$	18	3	3	$a_2 \otimes b_1 \otimes c_1 + a_1 \otimes b_2 \otimes c_2 + {}$ $+(a_2+a_1) \otimes b_3 \otimes c_3$
$\begin{bmatrix} \lambda & & \\ & \lambda & \\ & & \mu \end{bmatrix}$	15	3	3	$a_2 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + {}$ $+ a_1 \otimes b_3 \otimes c_3$
$\begin{bmatrix} \lambda & \mu & \\ & \lambda & \\ & & \mu \end{bmatrix}$	17	4	3	$a_2 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + {}$ $+(a_1+a_2) \otimes b_3 \otimes c_3 + a_1 \otimes b_1 \otimes c_2$
$\begin{bmatrix} \lambda & & \\ & \lambda & \\ & & \lambda \end{bmatrix}$	10	3	3	$a_2 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + {}$ $+ a_2 \otimes b_3 \otimes c_3$
$\begin{bmatrix} \lambda & \mu & \\ & \lambda & \\ & & \lambda \end{bmatrix}$	14	4	3	$a_2 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + {}$ $+ a_2 \otimes b_3 \otimes c_3 + a_1 \otimes b_1 \otimes c_2$
$\begin{bmatrix} \lambda & \mu & \\ & \lambda & \mu \\ & & \lambda \end{bmatrix}$	16	4	3	$a_2 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + {}$ $+ a_2 \otimes b_3 \otimes c_3 + a_1 \otimes b_1 \otimes c_2 + a_1 \otimes b_2 \otimes c_3$
$\begin{bmatrix} \lambda & \mu & \\ & \lambda & \mu \\ & & \end{bmatrix}$	14	3	3	$a_2 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + {}$ $+ a_1 \otimes b_1 \otimes c_2 + a_1 \otimes b_2 \otimes c_3$
$\begin{bmatrix} \lambda & & \\ \mu & \lambda & \\ & \mu & \end{bmatrix}$	14	3	3	$a_2 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + {}$ $+ a_1 \otimes b_2 \otimes c_1 + a_1 \otimes b_3 \otimes c_2$
$\begin{bmatrix} \lambda & \mu & \\ & & \lambda \\ & & \mu \end{bmatrix}$	14	4	3	$a_2 \otimes b_1 \otimes c_1 + a_1 \otimes b_1 \otimes c_2 + {}$ $+ a_2 \otimes b_2 \otimes c_3 + a_1 \otimes b_3 \otimes c_3$

(4.13)

Note: In table (4.13) the tensors which lies in $\mathbb{K}^2 \otimes \mathbb{K}^2 \otimes \mathbb{K}^2$ embedded in $\mathbb{K}^2 \otimes \mathbb{K}^3 \otimes \mathbb{K}^3$ are missing: clearly the dimensions of these orbits changes with respect to the space. The decompositions are all minimal but the last one, where we have five summands but the rank is 4 (see examples 4.3.12 and 4.4.4). Moreover, the border rank is always 3 since the 3-rd secant variety $\sigma_3(\mathrm{Seg}(\mathbb{P}^1 \times \mathbb{P}^2 \times \mathbb{P}^2))$ fills up the ambient space $\mathbb{P}^{17}$.

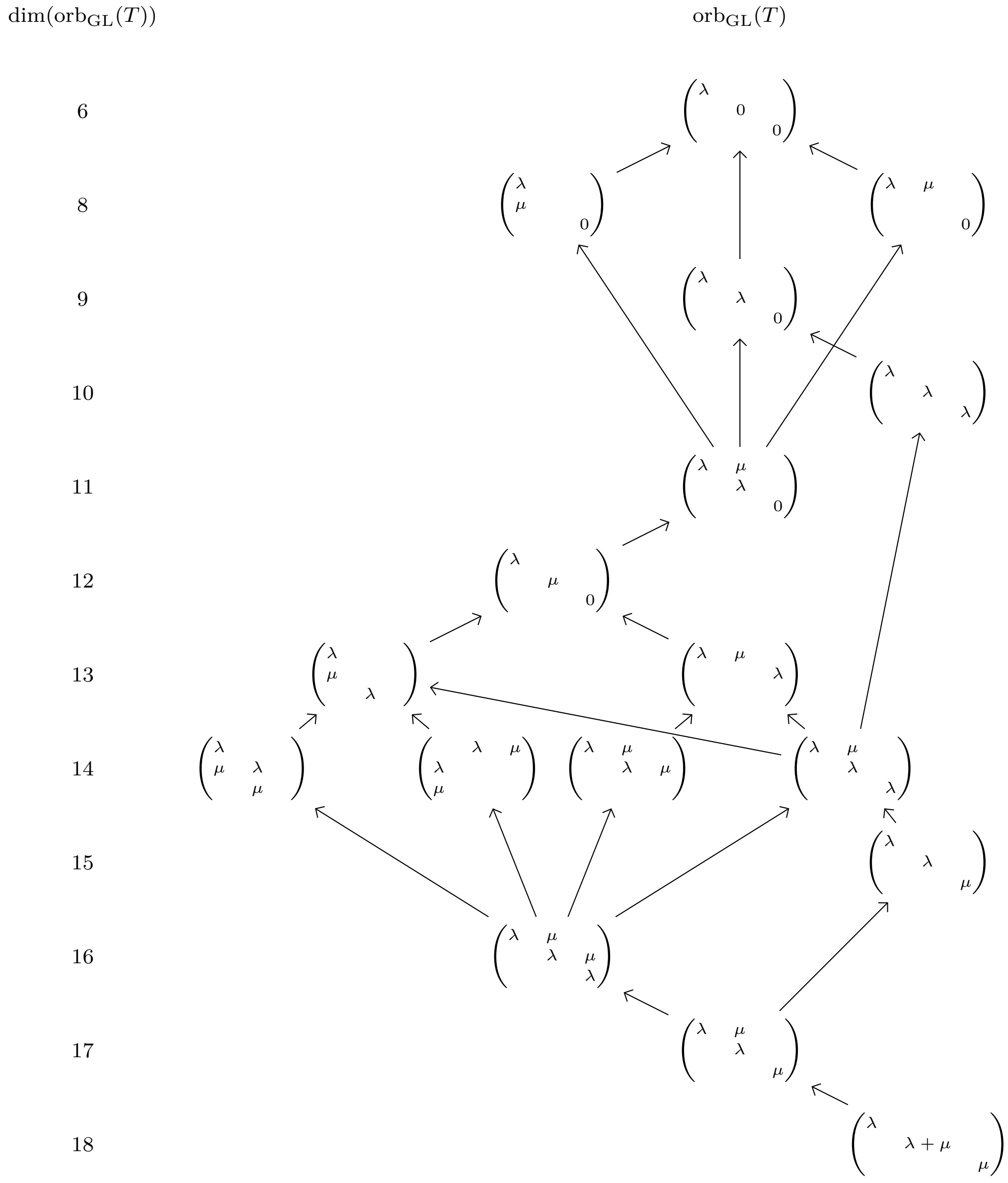


Figure 4.1: Oriented graph of GL-orbits in $\mathbb{K}^2 \otimes \mathbb{K}^3 \otimes \mathbb{K}^3$

The oriented graph represents the degeneracy relations between the GL-orbits in $\mathbb{K}^2 \otimes \mathbb{K}^3 \otimes \mathbb{K}^3$: each arrow goes from an orbit to a degenerate one of its. For details we refer to [31].

Remark 4.2.6. The case $\mathbb{K}^2 \otimes \mathbb{K}^2 \otimes \mathbb{K}^2$ has a remarkable difference with respect to other cases $\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^m$ for $m \neq 2$: indeed in the latter ones we have just a symmetry with respect to the two last factors of the tensor product, while in the case $\mathbb{K}^2 \otimes \mathbb{K}^2 \otimes \mathbb{K}^2$ such symmetry involves all three factors. Namely, for $\mathbb{K}^2 \otimes \mathbb{K}^2 \otimes \mathbb{K}^2$ we may consider the action by $\mathfrak{S}_3$ which permutes the three factors: of course this action does not give a GL-equivalence as the one we previously mentioned, but however it preserves ranks. Thus $B1, B2, B3$ in (4.12) are not GL-equivalent but they are all conjugated by $\mathfrak{S}_3$-action.

4.3 Rank in $\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$

We are now interested into study the rank of a tensor in $\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$.

Let $T \in \mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$ be a 2-slice tensor and let $\mathcal{P}_T$ be the associated matrix pencil. Of course the rank of the tensor is invariant under GL-action, hence we may freely work with the Kronecker form of the pencil. We may even talk about the rank of the pencil to be the rank of the corresponding tensor.
First we note that we may assume $\mathcal{P}_T = \mu A + \lambda B$ with no identically-zero minimal indices (i.e. $g = h = 0$ in (2.17)) otherwise we may restrict to consider T in $\mathbb{K}^2 \otimes \mathbb{K}^M \otimes \mathbb{K}^N$ for suitable $M < m$ and $N < n$. Under this hypothesis the pencil (or the tensor) is said to be **concise**: to be precise the definition of *conciseness* holds more in general for bilinear forms and the one we gave is just an equivalence when working with pencils [7]. Clearly the blocks R_ϵ and $\mathfrak{J}_{w,a}$ in (2.17) are concise.

Remark 4.3.1. By lemma 4.1.8, if $T \in \mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$ is concise, then $\mathrm{Rk}(T) \geq m, n$.

Moreover up to acting by $\mathrm{GL}_2(\mathbb{K})$ we may assume that there are no elementary divisors of the form μ^u. So we consider the Kronecker-Weierstrass form of $\mathcal{P}_T$

$$\Big(\boxplus_{i=1}^{p} R_{\epsilon_i} \Big) \boxplus \Big(\boxplus_{j=1}^{q} {}^tR_{\eta_j} \Big) \boxplus \Big(\boxplus_{j,k} \mathfrak{J}_{w_{jk}, a_{jk}} \Big) \tag{4.14}$$

The following lemma allows us to compute the rank of the blocks in (4.14): for a proof we refer to [7, Ch.19, §2].

Lemma 4.3.2. Let $\mathcal{P}$ be a pencil of size $m \times n$ with block decomposition $\begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ 0 & \mathcal{P}_{22} \end{bmatrix}$, where $\mathcal{P}_{11}, \mathcal{P}_{12}, \mathcal{P}_{22}$ are of size $a \times b$, $a \times (n-b)$, $(m-a) \times (n-b)$ respectively.

(1) If $\mathcal{P}_{22}$ is concise, then $\mathrm{Rk}(\mathcal{P}) \geq \mathrm{Rk}(\mathcal{P}_{11}) + (m-a)$;

(2) $\mathrm{Rk}(\mathcal{P}_{11} \boxplus \mathcal{P}_{22}) \leq \mathrm{Rk}(\mathcal{P}_{11}) + \mathrm{Rk}(\mathcal{P}_{22})$;

(3) If $\mathcal{P}_{22}$ is concise and $\mathrm{Rk}(\mathcal{P}_{22}) = \max\{m-a, n-b\}$, then equality holds.

Let $\mathcal{P}$ be a matrix pencil with non-zero minimal indices $\epsilon_1, \ldots, \epsilon_p, \eta_1, \ldots, \eta_q$ and regular block $\mathcal{K}$ of size N. We define $\delta(\mathcal{K})$ to be the number of invariant polynomials of $\mathcal{K}$ which are not squarefree. We define

$$\chi(\mathcal{P}) = \sum_{i=1}^{p}(\epsilon_i + 1) + \sum_{j=1}^{q}(\eta_j + 1) + N + \delta(\mathcal{K})$$

Our claim is that $\chi(\mathcal{P}) = \mathrm{Rk}(\mathcal{P})$.

Proposition 4.3.3. Let $T \in \mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$ be a 2-slice tensor with associated pencil $\mathcal{P}_T = \mu A + \lambda B$. Then $\mathrm{Rk}(T) \leq r$ if and only if there exist D_1, D_2 diagonal of size r and P, Q of size $m \times r$ and $r \times n$ respectively such that $\mathcal{P}_T = \mu(P \cdot D_1 \cdot {}^tQ) + \lambda(P \cdot D_2 \cdot {}^tQ)$.

Proof. Assume $\mathrm{Rk}(T) \leq r$: given (with respect to some basis)

$$T = \sum_{i,j,k} T_{ijk} u_i \otimes v_j \otimes w_k$$

it follows that there exist vectors $\alpha_1, \ldots, \alpha_r \in \mathbb{K}^2$ and $\beta_1, \ldots, \beta_r \in \mathbb{K}^m$ and $\gamma_1, \ldots, \gamma_r \in \mathbb{K}^n$ such that

$$T_{ijk} = \sum_{l=1}^{r} \alpha_l^{(i)} \beta_l^{(j)} \gamma_l^{(k)}$$

where $\alpha_l^{(i)}$ is the i-th component of the vector α_l (idem for $\beta_l^{(j)}, \gamma_l^{(k)}$). Then we set

$$D_1 = \mathrm{diag}(\alpha_1^{(1)}, \ldots, \alpha_r^{(1)}) \ , \ D_2 = \mathrm{diag}(\alpha_1^{(2)}, \ldots, \alpha_r^{(2)}) \ , \ P = (\beta_l^{(j)})_{j,l} \ , \ {}^tQ = (\gamma_l^{(k)})_{l,k}$$

Conversely, it is enough to follow backforward the above argument to conclude. □

Proposition 4.3.4. Let $\mathfrak{F}_p$ be the Frobenius block of size N associated to the invariant polynomial $p(\mu, \lambda)$. Then $\mathrm{Rk}(\mathfrak{F}_p) = \chi(\mathfrak{F}_p)$.

Proof. Set $\mathcal{F} = \mathfrak{F}_{p^w}$. By remark 4.3.1, since $\mathcal{F}$ is concise, we know that $\mathrm{Rk}(\mathcal{F}) \geq N$. By the above proposition we know that equality holds if and only if there exist D_1, D_2 diagonal of size N and U, V of size N such that

$$\mathcal{F} = \mu C_p + \lambda I_N = \mu(U \cdot D_1 \cdot {}^tV) + \lambda(U \cdot D_2 \cdot {}^tV)$$

We may assume $D_2 = I_N$. Then $\mathrm{Rk}(\mathcal{F}) = N$ if and only if C_p is diagonalizable, but since $\mathcal{F}$ has unique invariant polynomial $p(\mu, \lambda)$ this is equivalent to ask for $p(\mu, \lambda)$ to be squarefree, that is

$$\mathrm{Rk}(\mathcal{F}) = N \iff \delta(\mathcal{F}) = 0$$

Now we claim that $\mathrm{Rk}(\mathcal{F}) \leq N + 1$. Since $\mathbb{K}$ is infinite we may shift the polynomial $q(t)$ defining $\mathcal{F}$ by a suitable polynomial $c(t)$ such that $q(t) + c(t)$ has N distinct roots, that is $\delta(\mathfrak{F}_{q+c}) = 0$ (hence $\mathrm{Rk}(\mathfrak{F}_{q+c}) = N$ by the above argument). We may rewrite the pencil as

$$\mathcal{F} = \mathfrak{F}_{q+c} + \begin{bmatrix} 0 & \dots & 0 & c_0\lambda \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & c_{N-1}\lambda \end{bmatrix}$$

But the latter pencil has clearly rank 1, hence by lemma 4.3.2(2) $\mathrm{Rk}(\mathcal{F}) \leq N + 1$. Thus we conclude that $\mathrm{Rk}(\mathcal{F}) = N + \delta(\mathcal{F}) = \chi(\mathcal{F})$. □

Corollary 4.3.5. Let $\mathfrak{J}_{w,a}$ be the Jordan block associated to the elementary divisor $(\lambda + a\mu)^w$. Then $\mathrm{Rk}(\mathfrak{J}_{w,a}) = \chi(\mathfrak{J}_{w,a}) = w + (1 - \delta_{w1})$ (δ_{w1} is the Kronecker symbol).

Proposition 4.3.6. Let $\epsilon \geq 1$ and let R_ϵ be a minimal column block of size $\epsilon \times (\epsilon + 1)$. Then $\mathrm{Rk}(R_\epsilon) = \chi(R_\epsilon) = \epsilon + 1$.

Proof. Since R_ϵ is concise, by remark 4.3.1 we have $\mathrm{Rk}(R_\epsilon) \geq \epsilon + 1$. Let $q(t)$ be a polynomial of degree ϵ with ϵ distinct roots in $\mathbb{K}$. Then

$$R_\epsilon = \begin{bmatrix} & & \\ & \mathfrak{F}_q & \\ & & \\ 0 & \dots & 0 \end{bmatrix} + \begin{bmatrix} 0 & \dots & 0 & q_0\lambda \\ \vdots & & \vdots & \vdots \\ & & & q_{\epsilon-1}\lambda \\ 0 & \dots & 0 & \lambda \end{bmatrix}$$

Since $\mathrm{Rk}(\mathfrak{F}_q) = \epsilon$ and the latter summand has rank 1, by lemma 4.3.2(2) it follows $\mathrm{Rk}(R_\epsilon) \leq \epsilon + 1$, hence the equality holds. □

Corollary 4.3.7. Let $\eta \geq 1$ and let ${}^tR_\eta$ be a minimal row block of size $(\eta + 1) \times \eta$. Then $\mathrm{Rk}({}^tR_\eta) = \chi({}^tR_\eta) = \eta + 1$.

Our goal is to prove that for all 2-slice tensor T it holds $\mathrm{Rk}(T) = \chi(\mathcal{P}_T)$. In the previous propositions we computed the rank of each block of the Kronecker form, now we have to glue them together (by direct-block-sum) and prove that the above equality still holds. The idea is to reduce to the case of $\boxplus_{i\leq l} \mathfrak{J}_{2,a}$ whose rank may be computed by hand: for a proof we refer to [7, Ch.19, §2].

Proposition 4.3.8. For all $l \geq 1$ it holds $\mathrm{Rk}(\boxplus_{i=1:l} \mathfrak{J}_{2,a}) = 2l + l$.

Theorem 4.3.9 (Grigoriev-JàJà)**.** Let $T \in \mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n$ and let $\mathcal{P}_T$ be the associated pencil with minimal indices $\epsilon_1, \dots, \epsilon_p, \eta_1, \dots, \eta_q$ and regular part $\mathcal{K}$ of size N. Let $\delta(\mathcal{K})$ be the number of its non-squarefree invariant polynomials. Then

$$\mathrm{Rk}(T) = \sum_{i=1}^{p}(\epsilon_i + 1) + \sum_{j=1}^{q}(\eta_j + 1) + N + \delta(\mathcal{K}) \tag{4.15}$$

Proof. We may assume $\mathcal{P}_T$ to be in the Kronecker form. We may even assume T to be concise, thus

$$\mathcal{P}_T = \Big(\boxplus_{i=1}^{p} R_{\epsilon_i} \Big) \boxplus \Big(\boxplus_{j=1}^{q} {}^t R_{\eta_j} \Big) \boxplus \Big(\boxplus_{j,k} \mathfrak{J}_{w_{jk}, a_{jk}} \Big)$$

By proposition 4.3.6, $\mathrm{Rk}(R_{\epsilon_i}) = \epsilon_i + 1$ (idem for ${}^t R_{\eta_j}$), thus by lemma 4.3.2(3)

$$\begin{aligned}\mathrm{Rk}(T) &= \mathrm{Rk}\Big(\boxplus_{i=1}^{p} R_{\epsilon_i} \Big) + \mathrm{Rk}\Big(\boxplus_{j=1}^{q} {}^t R_{\eta_j} \Big) + \mathrm{Rk}\Big(\boxplus_{j,k} \mathfrak{J}_{w_{jk}, a_{jk}} \Big) = \\ &= \sum_{i=1}^{p} (\epsilon_i + 1) + \sum_{j=1}^{q} (\eta_j + 1) + \mathrm{Rk}\Big(\boxplus_{j,k} \mathfrak{J}_{w_{jk}, a_{jk}} \Big)\end{aligned}$$

Hence we may assume $\mathcal{P}_T = \boxplus_{j,k} \mathfrak{J}_{w_{jk}, a_{jk}}$ to be regular of size N. By lemma 4.3.2(2) and proposition 4.3.4 we already know

$$\mathrm{Rk}(\mathcal{P}_T) \leq N + \delta(\mathcal{P}_T)$$

We want to prove that actually equality holds. Pick $a \in \mathbb{K}$ such that $(\lambda + a\mu)^2$ divides the first $\delta(\mathcal{P}_T)$ invariant polynomials $p_{\delta(\mathcal{P}_T)} | \dots | p_1$. Then

$$\boxplus_{w} \mathfrak{J}_{w,a} = \Big(\boxplus_{w_k \geq 2, k=1}^{\delta(\mathcal{P}_T)} \mathfrak{J}_{w,a} \Big) \boxplus \Big(\boxplus_{i=1}^{s} \mathfrak{J}_{1,a} \Big)$$

By corollary 4.3.5 $\mathrm{Rk}\Big(\boxplus_{i=1}^{s} \mathfrak{J}_{1,a} \Big) = \mathrm{size}\Big(\boxplus_{i=1}^{s} \mathfrak{J}_{1,a} \Big) = s$, so by lemma 4.3.2(3)

$$\mathrm{Rk}\Big(\boxplus_{w} \mathfrak{J}_{w,a} \Big) = \mathrm{Rk}\Big(\boxplus_{w_k \geq 2, k=1}^{\delta(\mathcal{P}_T)} \mathfrak{J}_{w,a} \Big) + \mathrm{Rk}\Big(\boxplus_{i=1}^{s} \mathfrak{J}_{1,a} \Big)$$

By iteratively applying lemma 4.3.2(1) we have

$$\mathrm{Rk}\Big(\boxplus_{w \geq 2} \mathfrak{J}_{w,a} \Big) \geq \Big(\boxplus_{w_k \geq 2, k=1}^{\delta(\mathcal{P}_T)} \mathfrak{J}_{2,a} \Big) + \sum_{k=1}^{\delta(\mathcal{P}_T)} (w_k - 2) \overset{(4.3.8)}{=} \delta(\mathcal{P}_T) + \sum_{k=1}^{\delta(\mathcal{P}_T)} w_k$$

thus

$$\mathrm{Rk}\Big(\boxplus_{w} \mathfrak{J}_{w,a} \Big) \geq \delta(\mathcal{P}_T) + \sum_{k=1}^{\delta(\mathcal{P}_T)} w_k + s = \delta(\mathcal{P}_T) + \mathrm{size}\Big(\boxplus_{w} \mathfrak{J}_{w,a} \Big)$$

Finally by lemma 4.3.2(1) we conclude that

$$\mathrm{Rk}(\mathcal{P}_T) \geq \mathrm{Rk}\Big(\boxplus_{w} \mathfrak{J}_{w,a} \Big) + \mathrm{size}\Big(\boxplus_{v, b \neq a} \mathfrak{J}_{v,b} \Big) \geq N + \delta(\mathcal{P}_T)$$

□

Remark 4.3.10. In particular, in $\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^m$ the maximum possible rank of a tensor is $\lfloor \frac{3m}{2} \rfloor$: in [5] the case for m even is explicitly studied.

Note: It is important to underline that the weight $\delta(\mathcal{P}_T)$ depends on the number of non-squarefree invariant polynomials and not on the number of non-squarefree elementary divisors. We will observe this subtlety in the case of pencils of quadrics in $\mathbb{P}^3$ (see remark 4.5.8).

The Kronecker-Weierstrass form of a pencil does not only allow to compute the rank of the associated 2-slice tensor but even to give a decomposition which is unique up to $\mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$-action (or $\mathrm{SL}_2(\mathbb{K}) \times \mathrm{SL}_m(\mathbb{K}) \times \mathrm{SL}_n(\mathbb{K})$ if we are working with $\mathbb{P}(\mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^n)$). The decomposition of a 2-slice tensor is particularly aestetically beautiful if we assume the tensor to be regular with linear elementary divisors (i.e. without minimal indices and $\delta(\mathcal{P}_T) = 0$).

Corollary 4.3.11. Let $T \in \mathbb{K}^2 \otimes \mathbb{K}^m \otimes \mathbb{K}^m$ be regular (in the sense of its pencil) and such that $\delta(\mathcal{P}_T) = 0$. Let $\{\lambda + a_i\mu \mid i = 1 : m\}$ be its elementary divisors (not necessarily distincts). With respect to a suitable basis, its minimal decomposition is

$$T = \sum_{i=1}^{m} (\lambda + a_i\mu) \otimes v_i \otimes w_i$$

Example 4.3.12. Let $T \in \mathbb{K}^2 \otimes \mathbb{K}^3 \otimes \mathbb{K}^3$ be the 2-slice tensor with associated pencil

$$\mathcal{P}_T = \begin{bmatrix} \lambda + a\mu & \mu & \\ & \lambda + a\mu & \mu \\ & & \lambda + a\mu \end{bmatrix}$$

With respect to a suitable basis we may write a first decomposition of T in five terms:

$$(\lambda + a\mu) \otimes v_1 \otimes w_1 + (\lambda + a\mu) \otimes v_2 \otimes w_2 + (\lambda + a\mu) \otimes v_3 \otimes w_3 + \mu \otimes v_1 \otimes w_2 + \mu \otimes v_2 \otimes w_3$$

But by theorem 4.3.9 we know that $\mathrm{Rk}(T) = 3 + \delta(\mathcal{P}_T)$ and, since $\mathcal{P}_T$ has unique invariant polynomial $(\lambda + a\mu)^3$, we have $\delta(\mathcal{P}_T) = 1$ and $\mathrm{Rk}(T) = 3 + 1 = 4$. Then we know that we can optimize the above decomposition in a shorter one of four terms. However it is not trivial to obtain a decomposition in four summands of the above tensor: to do so we need to use a decomposition technique based on the *discrete Fourier transform*. We will show the solution in the next section (see example 4.4.4).

4.4 Discrete Fourier transform in $\mathbb{K}^{\alpha+1} \otimes \mathbb{K}^{\beta+1} \otimes \mathbb{K}^{\alpha+\beta+1}$

The decomposition of some tensors in $\mathbb{K}^{\alpha+1} \otimes \mathbb{K}^{\beta+1} \otimes \mathbb{K}^{\alpha+\beta+1}$ is strictly connected to the multiplication of two polynomials of degree α and β. Let us see how.

Let $a(X) = a_\alpha X^\alpha + \ldots + a_0$ and $b(X) = b_\beta X^\beta + \ldots + b_0$ be two polynomials and let $c(X) = a(X) \cdot b(X) = c_{\alpha+\beta} X^{\alpha+\beta} + \ldots + c_0$ be their product. Let $\underline{a} \in \mathbb{K}^{\alpha+1}, \underline{b} \in \mathbb{K}^{\beta+1}, \underline{c} \in \mathbb{K}^{\alpha+\beta+1}$ be the coordinate vectors of $a(X), b(X), c(X)$ respectively. Consider the polynomial multiplication map

$$\begin{array}{rccc} \mathcal{M}: & \mathbb{K}^{\alpha+1} \times \mathbb{K}^{\beta+1} & \longrightarrow & \mathbb{K}^{\alpha+\beta+1} \\ & (\underline{a}, \underline{b}) & \mapsto & \underline{c} \end{array}$$

By the universal property of tensor product we have the linear map

$$\begin{array}{rccc} \tilde{\mathcal{M}}: & \mathbb{K}^{\alpha+1} \otimes \mathbb{K}^{\beta+1} & \longrightarrow & \mathbb{K}^{\alpha+\beta+1} \\ & \underline{a} \otimes \underline{b} & \mapsto & \underline{c} \end{array}$$

and the mutiplication map can be encoded (up to dual) as a tensor

$$T = \sum_{i=0}^{\alpha} \sum_{j=0}^{\beta} a_i \otimes b_j \otimes c_{i+j} \in \mathbb{K}^{\alpha+1} \otimes \mathbb{K}^{\beta+1} \otimes \mathbb{K}^{\alpha+\beta+1}$$

called **multiplication tensor**.

Fast Fourier multiplication. Before analyzing how the polynomial multiplication affects the tensor decomposition let us recall some of its computational-algebraic properties. The standard polynomial multiplication has complexity of order $O(N^2)$ (where N is the maximum degree of the factors).

A method for polynomial multiplication is the one of *evaluation-interpolation*: given $a(X)$ and $b(X)$ of degree α, β respectively and given $n = \max\{\alpha, \beta\}$, one chooses (at least) $N = 2n + 1$ distinct values $\lambda_k \in \mathbb{K}$, evaluates $A_k = a(\lambda_k)$ and $B_k = b(\lambda_k)$, computes the values $C_k = a(\lambda_k) \cdot b(\lambda_k)$ and computes the polynomial $c(X)$ interpolating the values C_k in the point λ_k; the interpolating polynomial $c(X)$ is the the product $a(X) \cdot b(X)$.

The complexity of this method is still $O(n^2)$ for general values λ_k, but things get a better taste if one properly chooses the values λ_k: indeed by choosing $\lambda_k = \zeta_N^{k-1}$ (where ζ_N is a primitive N-th root of the unity) the complexity decreases to $O(n \log n)$. This faster method is called *fast Fourier multiplication* and the processes of evaluation and interolation are known as *fast Fourier transform* and *fast Fourier interpolation* [27, Ch.IX].

The algorithm of fast Fourier multiplication is the following. Let $N = 2^m$ be the smaller power of 2 greater than $2n$, ζ_N be a primitive N-th root of unity and $V[\zeta_N] = V(1, \zeta_N, \ldots, \zeta_N^{N-1})$ be the Vandermonde matrix. Let $\underline{A}, \underline{B}, \underline{C}$ be the vectors with coordinates $A_k = a(\zeta_N^k), B_k = b(\zeta_N^k), C_k = A_k \cdot B_k$ respectively. Then:

- *evaluation*: $\underline{A} = V[\zeta_N] \cdot \underline{a}$, $\underline{B} = V[\zeta_N] \cdot \underline{b}$;

- *interpolation*: $\underline{c} = N^{-1} V[\zeta_N^{-1}] \cdot \underline{C}$.

Remark 4.4.1. In this algorithm the evaluation and interpolation processes are called *fast* because they leverage the choice $N = 2^m$ by applying a binary splitting of the polynomials in their even and odd terms and by iterating the algorithms to them. Without this splitting we just refer to these processes as *discrete* (instead of *fast*).

Remark 4.4.2. The fast Fourier multiplication holds over finite fields $\mathbb{F}_q$ with characteristic $p \gneq 2$ too: the hypothesis $p \neq 2$ ensures the existence of N^{-1} (since $N = 2^m$) and one may always find a primitive N-th roots of unity ζ_N in $\mathbb{F}_q$ thanks to the Dirichlet theorem on arithmetic progressions: an algorithmic way is to find a primitive element in the field $\mathbb{F}_q$.

Application to tensor decomposition. Consider the multiplication tensor

$$T = \sum_{i=0}^{\alpha} \sum_{j=0}^{\beta} a_i \otimes b_j \otimes c_{i+j} \in \mathbb{K}^{\alpha+1} \otimes \mathbb{K}^{\beta+1} \otimes \mathbb{K}^{\alpha+\beta+1}$$

As we observed at the beginning of the section we may reinterpret it as the polynomial multiplication

$$(a_0 + \ldots + a_\alpha X^\alpha) \cdot (b_0 + \ldots + b_\beta X^\beta) = c_0 + \ldots + c_{\alpha+\beta} X^{\alpha+\beta}$$

Idea: We want to find a decomposition of T in $\alpha + \beta + 1$ summands by applying the fast Fourier multiplication at $a(X)$ and $b(X)$.

Let us set $N = \alpha+\beta+1$ and $\zeta = \zeta_{\alpha+\beta+1}$ a primitive N-th root of unity. Let $A_k = a(\zeta^k)$ and $B_k = b(\zeta^k)$ be the evaluations of $a(X)$ and $b(X)$ in the points $\{\zeta^k \mid k = 0 : \alpha + \beta\}$. Let $\underline{C}$ be the vector with k-th coordinate $A_k \cdot B_k$. By the fast Fourier multiplication we have

$$\underline{c} = \frac{1}{\alpha + \beta + 1} V[\zeta^{-1}] \cdot \underline{C}$$

then for $i = 0 : \alpha + \beta$

$$c_i = \frac{1}{\alpha + \beta + 1} \sum_{j=0}^{\alpha+\beta} \zeta^{-ij} A_j B_j = \frac{1}{\alpha + \beta + 1} \sum_{j=0}^{\alpha+\beta} \zeta^{-ij} \Big(\sum_{k=0}^{\alpha} a_k \zeta^{kj} \Big) \Big(\sum_{k=0}^{\beta} b_k \zeta^{kj} \Big)$$

By going back to the tensor interpretation we have

$$\sum_{i=0}^{\alpha} \sum_{j=0}^{\beta} a_i \otimes b_j \otimes c_{i+j} = \frac{1}{\alpha + \beta + 1} \sum_{j=0}^{\alpha+\beta} \left[\Big(\sum_{i=0}^{\alpha+\beta} c_i \zeta^{-ij} \Big) \otimes \Big(\sum_{i=0}^{\alpha} a_i \zeta^{ij} \Big) \otimes \Big(\sum_{i=0}^{\beta} b_i \zeta^{ij} \Big) \right] \quad (4.16)$$

that is a decomposition of T in $\alpha + \beta + 1$ decomposable tensors. In particular, we have proven the following result.

Corollary 4.4.3. The multiplication tensor in $\mathbb{K}^{\alpha+1}\otimes\mathbb{K}^{\beta+1}\otimes\mathbb{K}^{\alpha+\beta+1}$ has rank $\alpha+\beta+1$.

Example 4.4.4. Let us go back to example 4.3.12, where we exhibited a decomposition of a tensor $T \in \mathbb{K}^2 \otimes \mathbb{K}^3 \otimes \mathbb{K}^3$ with associated pencil a Jordan block $\mathfrak{J}_{3,a}$ in five decomposable summands. However theorem 4.3.9 tells us that the rank of T is four. First of all we *extend* the tensor T to the tensor $\tilde{T} \in \mathbb{K}^2 \otimes \mathbb{K}^3 \otimes \mathbb{K}^4$ with associated pencil

$$\mathcal{P}_{\tilde{T}} = \begin{bmatrix} \lambda + a\mu & \mu & & \\ & \lambda + a\mu & \mu & \\ & & \lambda + a\mu & \mu \end{bmatrix}$$

The idea is to use the fast Fourier multiplication to decompose $\tilde{T}$ in four decomposable tensors and then *cut* the last column to obtain a decomposition of T.
We are in the following setting:

$$\alpha = 1 \ , \ \beta = 2 \ , \ N = 4 \ , \ \zeta = \zeta_4 = i$$

$$a_0 = (\lambda + a\mu) \ , \ a_1 = \mu \ , \ b_k = v_{i-1} \ \forall k = 1 : 3 \ , \ c_k = w_{k-1} \ \forall k = 1 : 4$$

and so by (4.16)

$$\begin{aligned}
\tilde{T} = \sum_{k=0}^{1}\sum_{j=0}^{2} a_k \otimes b_j \otimes c_{k+j} &= \frac{1}{4}\sum_{j=0}^{3}\left[\Big(\sum_{k=0}^{3} c_k(i)^{-kj}\Big) \otimes \Big(\sum_{k=0}^{1} a_k(i)^{kj}\Big) \otimes \Big(\sum_{k=0}^{2} b_k(i)^{kj}\Big)\right] \\
&= \frac{1}{4}\Big[(a_0 + a_1) \otimes (b_0 + b_1 + b_2) \otimes (c_0 + c_1 + c_2 + c_3) + \\
&\quad + (a_0 + ia_1) \otimes (b_0 + ib_1 - b_2) \otimes (c_0 - ic_1 - c_2 + ic_3) + \\
&\quad + (a_0 - a_1) \otimes (b_0 - b_1 + b_2) \otimes (c_0 - c_1 + c_2 - c_3) + \\
&\quad + (a_0 - ia_1) \otimes (b_0 - ib_1 - b_2) \otimes (c_0 + ic_1 - c_2 - ic_3)\Big]
\end{aligned}$$

By rewriting the decomposition of $\tilde{T}$ in terms of λ, μ, v_i, w_j and by setting $c_3 = 0$, we have the rank-4 decomposition of T

$$\begin{aligned}
T = &\frac{1}{4}\Big[(\lambda + (a+1)\mu) \otimes (v_1 + v_2 + v_3) \otimes (w_1 + w_2 + w_3) + \\
&+ (\lambda + (a+i)\mu) \otimes (v_1 + iv_2 - v_3) \otimes (w_1 - iw_2 - w_3) + \\
&+ (\lambda + (a-1)\mu) \otimes (v_1 - v_2 + v_3) \otimes (w_1 - w_2 + w_3) + \\
&+ (\lambda + (a-i)\mu) \otimes (v_1 - iv_2 - v_3) \otimes (w_1 + iw_2 - w_3)\Big]
\end{aligned}$$

Multiplication tensor in $\mathbb{K}^2\otimes\mathbb{K}^{\beta+1}\otimes\mathbb{K}^{\beta+2}$. Now we fix $\alpha = 1$ and let β be generic. Consider the multiplication tensor

$$T = \sum_{i=0}^{1}\sum_{j=0}^{\beta} a_i \otimes b_j \otimes c_{i+j} \in \mathbb{K}^2 \otimes \mathbb{K}^{\beta+1} \otimes \mathbb{K}^{\beta+2}$$

with associated pencil

$$\mathcal{P}_T = \begin{bmatrix} a_0 & a_1 & & \\ & \ddots & \ddots & \\ & & a_0 & a_1 \end{bmatrix}$$

We know this pencil! Indeed by substituting a_0 and a_1 with λ and μ respectively we have a Kronecker singular block R_ϵ with respect to the minimal index for columns $\epsilon = \beta + 1$.
By corollary 4.4.3 and (4.16) we know that $\mathrm{Rk}(T) = \beta + 2$ and a decomposition of T is given by

$$\frac{1}{\beta+2} \sum_{j=0}^{\beta+1} \left[\Big(\sum_{i=0}^{\beta+1} c_i \zeta_{\beta+2}^{-ij} \Big) \otimes \Big(\sum_{i=0}^{1} a_i \zeta_{\beta+2}^{ij} \Big) \otimes \Big(\sum_{i=0}^{\beta} b_i \zeta_{\beta+2}^{ij} \Big) \right]$$

Then the Kronecker singular block $R_{\beta+1}$ has rank $\beta+2$, as also confirmed by proposition 4.3.6.

Remark 4.4.5 (Cherry-on-the-cake!)**.** By the last two paragraphs it follows that theorem 4.3.9 may also be proved by the discrete Fourier transform. Given the Kronecker form

$$0_{h\times g} \boxplus \Big(\boxplus_{i=g+1}^{p} R_{\epsilon_i} \Big) \boxplus \Big(\boxplus_{j=h+1}^{q} {}^tR_{\eta_j} \Big) \boxplus \Big(\boxplus_{l,z} \mathfrak{J}_{w_{lz},a_{lz}} \Big)$$

one may looks at each singular block R_{ϵ_i} (similarly for ${}^tR_{\eta_j}$) as a multiplication tensor $T_{\epsilon_i} \in \mathbb{K}^2 \otimes \mathbb{K}^{\epsilon_i} \otimes \mathbb{K}^{\epsilon_i+1}$ and at each regular block $\mathfrak{J}_{w_{lz},a_{lz}}$ as the troncation of a multiplication tensor $T_{w_{lz},a_{lz}} \in \mathbb{K}^2 \otimes \mathbb{K}^{w_{lz}} \otimes \mathbb{K}^{w_{lz}+1}$ (as in example 4.4.4). Thus by the discrete Fourier decomposition we know the rank of each block and by lemma 4.3.2(3) we get the rank of the pencil.

4.5 Ranks and GL-orbits in $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$

In the previous sections we showed that 2-slice tensors correspond to matrix pencils and that the Kronecker form of the latters allows to determine the rank of the former ones. In this section we work with tensors in $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$, called **symmetric 2-slice tensors**, corresponding to pencils of quadrics.

Note: In chapter 3 we determined the canonical form for GL-equivalence of pencils of quadrics and we found out that a complete system of invariants is given by the Segre symbol and the continuous modulus. We also studied such pencils with a particular focus on the ones in $\mathbb{P}^2$ and $\mathbb{P}^3$ both algebraically and geometrically. In this perspective one may reinterpret the Segre classification as the classification of tensors in $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$ up to GL-equivalence.

4.5.1 Partially-symmetric rank

For tensors in $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$ we have to adapt the notion of symmetric rank to a weaker form. First of all the decomposable tensors in $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$ are the ones of the form $u \otimes l^2$ where $u \in \mathbb{K}^2$ and $l \in \mathbb{K}^{m+1}$ is a linear form. Geometrically, the decomposable symmetric 2-slice tensors are the ones in $\mathbb{P}^1 \times \mathbb{P}^m$ embedded in $\mathbb{P}(\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1}))$ with the sheaf $\mathcal{O}(1,2)$, or equivalently the ones in

$$\mathrm{Seg}(\mathbb{P}^1 \times \nu_2(\mathbb{P}^m)) \subset \mathbb{P}\Big(\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})\Big)$$

Definition. The **partially-symmetric rank** of a symmetric 2-slice tensor $T \in \mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$ is the minimum number of decomposable summands of type $u \otimes l^2$ in which T decomposes, that is

$$\mathrm{symRk}_p(T) = \left\{ r \;\middle|\; T = \sum_{i=1}^{r} u_i \otimes l_i^2 \right\}$$

Note: Since $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1}) \subset \mathbb{K}^2 \otimes \mathbb{K}^{m+1} \otimes \mathbb{K}^{m+1}$, as well as for symmetric rank it surely holds

$$\mathrm{Rk} \leq \mathrm{symRk}_p \tag{4.17}$$

In the following we will go back to the setting of Segre classification (see Chapter 3) and we will study rank and partially-symmetric rank of symmetric 2-slice tensors. The thread of our study is the following: first we determine the different ranks for the symmetric blocks

$$S_{\epsilon_i} = \begin{bmatrix} 0 & R_{\epsilon_i} \\ {}^tR_{\epsilon_i} & 0 \end{bmatrix} \quad , \quad \check{\mathfrak{J}}_{w,a} = \begin{bmatrix} & & & \mu & \lambda + a\mu \\ & & \iddots & \iddots & \\ & & \iddots & & \\ & \mu & & & \\ \lambda + a\mu & & & & \end{bmatrix} \tag{4.18}$$

of a Kronecker form, then we deduce the partially-symmetric rank of a symmetric 2-slice tensor.

Ranks of symmetric singular blocks. Let $S_\epsilon \in \mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{2\epsilon+1})$ be the symmetric 2-slice tensor defined by a Kronecker singular block as in (4.18). By proposition 4.3.6 we know that both the blocks R_ϵ and ${}^tR_\epsilon$ have rank $\epsilon + 1$, hence

$$\mathrm{Rk}(S_\epsilon) = 2\epsilon + 2$$

By inequality (4.17) it follows that $\mathrm{symRk}_p(S_\epsilon) \geq 2\epsilon + 2$. Actually equality holds.

Lemma 4.5.1. $\mathrm{symRk}_p(S_\epsilon) = \mathrm{Rk}(S_\epsilon) = 2\epsilon + 2$.

Proof. Since $\mathrm{Rk}(R_\epsilon) = \epsilon + 1$ there exist $A_0, \ldots, A_\epsilon$ rank-1 matrices such that

$$R_\epsilon(\overline{\lambda}, \overline{\mu}) \subset \langle A_0, \ldots, A_\epsilon \rangle$$

for all $\overline{\lambda}, \overline{\mu}$. For all $i = 0 : \epsilon$ the $B_i = \begin{bmatrix} 0 & A_i \\ {}^t A_i & 0 \end{bmatrix}$ is symmetric of rank 2 and it holds

$$S_\epsilon(\overline{\lambda}, \overline{\mu}) \subset \langle B_0, \ldots, B_\epsilon \rangle$$

for all $\overline{\lambda}, \overline{\mu}$. Thus $\mathrm{symRk}_p(S_\epsilon) \leq \mathrm{Rk}(B_0) + \ldots + \mathrm{Rk}(B_\epsilon) = 2(\epsilon + 1)$. □

Ranks of symmetric Jordan blocks. Let $(\lambda + a\mu)^w$ be an elementary divisor of a symmetric 2-slice tensor in $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$ and let $\check{\mathfrak{J}}_{w,a}$ be the corresponding Kronecker Jordan block as in (4.18). By corollary 4.3.5 we know that

$$\mathrm{Rk}(\check{\mathfrak{J}}_{w,a}) = w + (1 - \delta_{w1})$$

where δ_{w1} is the Kronecker symbol. The case $w = 1$ is trivial, thus we assume $w > 1$, hence $\mathrm{Rk}(\check{\mathfrak{J}}_{w,a}) = w + 1$. By inequality (4.17) it follows that $\mathrm{symRk}_p(\check{\mathfrak{J}}_{w,a}) \geq w + 1$. Actually the next lemma will show that equality holds, but first we need some preliminary.

First of all we note that every symmetric Jordan block $\check{\mathfrak{J}}_{w,a}$ as in (4.18) is a matrix of *Hankel* type, that is every skew-diagonal is constant:

$$(\check{\mathfrak{J}}_{w,a})_{i,j} = (\check{\mathfrak{J}}_{w,a})_{i+k,j-k} \quad \forall i \leq j,\ k = 0 : j - i$$

We may represent a Hankel matrix H of size N in the compact form

$$H = \mathrm{Hank}(h_1, \ldots, h_{N-1}, h_N, h_{N+1}, \ldots, h_{2N-1})$$

where h_i is each entry of the i-th ascending skew-diagonal. For instance,

$$\check{\mathfrak{J}}_{w,a} = \mathrm{Hank}\Big(\underbrace{0, \ldots, 0}_{w-2}, \mu, \lambda + a\mu, \underbrace{0, \ldots, 0}_{w-1}\Big)$$

The only two non-zero skew-diagonals of $\check{\mathfrak{J}}_{w,a}$ have entries μ and $\lambda + a\mu$ respectively, which are relatively prime: then, for any fixed finite subset $\Gamma \subset \mathbb{K}$ and for any general vector $v \in \mathbb{K}^w$, the polynomial ${}^t v \check{\mathfrak{J}}_{w,a} v$ has distinct linear factors none of which equal to $\lambda + \gamma\mu$ for any $\gamma \in \Gamma$.

Fact 4.5.2. The above property holds for every Hankel matrix whose entries are relatively prime homogeneous polynomials of fixed degree in two variables.

For all square invertible matrix A of size m and $u, v \in \mathbb{K}^m$ it holds [11, Lemma 1.1]

$$\det(A + u \cdot {}^t v) = (1 + {}^t v A^{-1} u)\det(A)$$

and its symmetric form

$$\det(A + u \cdot {}^t u) = (1 + {}^t u A^{-1} u)\det(A) = \det(A) + {}^t u A^* u \tag{4.19}$$

where A^* is the *cofactor matrix* of A defined by $(A^*)_{i,j} = (-1)^{i+j}\det(A_{ij})$ for A_ij the submatrix of A obtained by deleting the i-th row and the j-th column.

The cofactor matrix of $\check{\mathfrak{J}}_{w,a}$ is the lower skew-triangular Hankel matrix

$$(\check{\mathfrak{J}}_{w,a})^* = -\operatorname{Hank}\Big(\underbrace{0,\dots,0}_{w-1}, (-\lambda - a\mu)^{w-1}, \dots, (-\lambda - a\mu)^{w-k}\mu^{k-1}, \dots, \mu^{w-1}\Big)$$

Lemma 4.5.3. $\operatorname{symRk}_p(\check{\mathfrak{J}}_{w,a}) = \operatorname{Rk}(\check{\mathfrak{J}}_{w,a}) = w + \delta_{w1}$.

Proof. The case $w = 1$ is trivial, thus we assume $w > 1$. The cofactor matrix $(\check{\mathfrak{J}}_{w,a})^*$ has the property in 4.5.2, thus for a general $u \in \mathbb{K}^w$ the polynomial ${}^t u(\check{\mathfrak{J}}_{w,a})^* u$ has distinct factors none of which are equal to $(\lambda + a\mu)$. Moreover, up to rescale $u \in \mathbb{K}^w$, the polynomial

$$\det\Big(\check{\mathfrak{J}}_{w,a} + \mu(u \cdot {}^t u)\Big) \overset{(4.19)}{=} \det(\check{\mathfrak{J}}_{w,a}) + \mu\Big({}^t u(\check{\mathfrak{J}}_{w,a})u\Big) = (\lambda + a\mu)^w + \mu\Big({}^t u(\check{\mathfrak{J}}_{w,a})u\Big)$$

has distinct linear factors, hence the pencil $\check{\mathfrak{J}}_{w,a} + \mu(u \cdot {}^t u)$ is diagonalizable. Since $\mu(u \cdot {}^t u)$ has rank 1 by construction, it holds

$$\operatorname{symRk}_p(\check{\mathfrak{J}}_{w,a}) \le \operatorname{symRk}_p(\check{\mathfrak{J}}_{w,a} + \mu(u \cdot {}^t u)) + \operatorname{symRk}_p(\mu(u \cdot {}^t u)) \le w + 1$$

and by inequality (4.17) we conclude. □

By lemmas 4.5.1 and 4.5.3 we can conclude the following result.

Theorem 4.5.4. For all pencil of quadrics $T \in \mathbb{K}^2 \otimes \operatorname{Sym}^2(\mathbb{K}^{m+1})$ it holds

$$\operatorname{symRk}_p(T) = \operatorname{Rk}(T)$$

Proof. Consider the symmetric Kronecker-Weierstrass decomposition of T

$$T = \left(\boxplus_{i=1}^{p} S_{\epsilon_i}\right) \boxplus \left(\boxplus_{l,k} \check{\mathfrak{J}}_{w_{lk}, a_{lk}}\right)$$

By theorem 4.3.9 we know that

$$\operatorname{symRk}_p(T) \ge \operatorname{Rk}(T) = \sum_{i=1}^{p} 2(\epsilon_i + 1) + \sum_{l,k} w_{lk} + \delta$$

where δ is the number of the non-squarefree invariant polynomials of T. Clearly

$$\begin{aligned}\mathrm{symRk}_p(T) &\leq \sum_{i=1}^{p} \mathrm{symRk}_p(S_{\epsilon_i}) + \mathrm{symRk}_p\left(\boxplus_{l,k} \check{\mathfrak{J}}_{w_{lk},a_{lk}}\right)\\ &= \sum_{i=1}^{p} 2(\epsilon_i + 1) + \mathrm{symRk}_p\left(\boxplus_{l,k} \check{\mathfrak{J}}_{w_{lk},a_{lk}}\right)\end{aligned}$$

Let us focus on the regular part of T. In the same perspective of corollary 2.2.9, we decompose the regular part $\boxplus_{lk} \check{\mathfrak{J}}_{w_{lk},a_{lk}}$ in

$$\left(\boxplus_{\alpha\in I_1} \check{\mathfrak{J}}_{w_\alpha,a_\alpha}\right) \boxplus \ldots \boxplus \left(\boxplus_{\beta\in I_r} \check{\mathfrak{J}}_{w_\beta,a_\beta}\right) \boxplus \left(\boxplus_{\gamma\in J} \check{\mathfrak{J}}_{1,a_\gamma}\right)$$

where the block-direct-sum indexed by I_j is given by the Jordan blocks of the non-squarefree factors of the j-th invariant polynomial and the block indexed by J is given by the Jordan blocks of size 1 (hence it is diagonal and we denote it by D).
Let us denote $P_j = \boxplus_{I_j} \check{\mathfrak{J}}_{w_s,a_s}$ and $N_j = \mathrm{size}(P_j)$. Every block-direct-sum P_j is a block-diagonal pencil with Hankel blocks and it has distinct roots $a_s \neq a_t$ for all $s \neq t \in I_j$: in particular, it has the same property in 4.5.2 and same arguments [6, Lemma 7.3] in lemma 4.5.3 lead to

$$\mathrm{symRk}_p(P_j) = \mathrm{Rk}(P_j) = N_j + 1$$

It follows

$$\begin{aligned}\mathrm{symRk}_p(T) &\leq \sum_{i=1}^{p} 2(\epsilon_i + 1) + \sum_{j=1}^{r} \mathrm{symRk}_p(P_j) + \mathrm{symRk}_p(D)\\ &= \sum_{i=1}^{p} 2(\epsilon_i + 1) + \sum_{lk} w_{lk} + \delta\\ &= \mathrm{Rk}(T)\end{aligned}$$

□

Note: In tables (4.21) and (4.22) we list the pencils of quadrics we studied in Ch.3, §2 by exhibiting their partially-symmetric rank (equivalent to the rank by theorem 4.5.4), their border rank and their minimal decomposition.

4.5.2 GL-orbits and their dimensions

By theorem 2.4.2 the action (4.6) for $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$ simplifies to the action

$$\begin{array}{ccc}\mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_{m+1}(\mathbb{K}) & \longrightarrow & \mathrm{Aut}\left(\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})\right)\\ (M, P) & \mapsto & \left(u \otimes l^2 \mapsto Mu \otimes P \cdot l^2 \cdot {}^t P\right)\end{array} \quad (4.20)$$

Hence GL-orbits in the partially-symmetric case are with respect to the action (4.20).

Remark 4.5.5. By theorem 4.2.3 it follows that even for pencils of quadrics (or equivalently, tensors in $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$) the GL-orbits are finitely many if and only if $m+1 \leq 3$. In particular, the representatives in table (4.21) cover all the GL-orbits of pencils of quadrics in $\mathbb{P}^2$ but the representatives in table (4.22) do not even cover all the GL-orbits of regular pencils in $\mathbb{P}^3$: for the latter ones we need the continuous moduli too.

Given a pencil of quadrics $T \in \mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$, the action (4.20) induces the map

$$\begin{array}{rccc} \gamma_T: & \mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_{m+1}(\mathbb{K}) & \longrightarrow & \mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1}) \\ & G & \mapsto & G \cdot T \end{array}$$

with image $\mathrm{orb}_{\mathrm{GL}}(T)$. By deriving γ_T in the identity $I = (I_2, I_{m+1})$, we get the linear map

$$\begin{array}{rccc} d(\gamma_T)_I: & \mathfrak{gl}_2(\mathbb{K}) \times \mathfrak{gl}_{m+1}(\mathbb{K}) & \longrightarrow & \mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1}) \\ & g & \mapsto & g \cdot T \end{array}$$

whose kernel is the Lie algebra of the stabilizer $\mathrm{stab}_{\mathrm{GL}}(T)$ and whose rank is the dimension of $\mathrm{orb}_{\mathrm{GL}}(T)$. As for general tensors, to compute the dimension of the orbit we may compute

$$\dim\big(\mathrm{orb}_{\mathrm{GL}}(T)\big) = \mathrm{Rk}\left(d(\gamma_T)_I\right) = 4 + (m+1)^2 - \dim\Big(\ker\left(d(\gamma_T)_I\right)\Big)$$

Note: In Ch.6, §2.2 we implement the computation of the orbit dimensions on `Macaulay2`.

Example 4.5.6. Consider $T = \lambda \otimes x^2 \in \mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^3)$ with pencil $\begin{bmatrix} \lambda & & \\ & 0 & \\ & & 0 \end{bmatrix}$. Then $[T] \in \mathrm{Seg}(\mathbb{P}^1 \times \nu_2(\mathbb{P}^2))$ is decomposable and the SL-action sends decomposable tensors in decomposable tensors, hence its projective SL-orbit is $\mathrm{Seg}(\mathbb{P}^1 \times \nu_2(\mathbb{P}^2))$ of dimension 3, that is the affine GL-orbit has dimension 4.

Remark 4.5.7. In all tables describing pencils of quadrics in chapter 3, we always considered only the pencils defined by non-degenerate intersections: for instances, we did not consider the pencils $\begin{bmatrix} \lambda & & \\ & 0 & \\ & & 0 \end{bmatrix}$ and $\begin{bmatrix} \lambda & & \\ & \lambda & \\ & & \lambda \end{bmatrix}$ (with Segre symbols [1] and [(1 1 1)] respectively) since one of the two quadrics defining them are identically zero, hence they give trivial intersections. For this reason in the following we do not consider all orbits in $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^{m+1})$ but only the ones corresponding to (non-trivial) pencils of quadrics.

By keeping in mind remark 4.5.7, we now list the GL-orbits of all non-trivial pencils of quadrics in $\mathbb{P}^2$ and the ones of the regular pencils of quadrics in $\mathbb{P}^3$ by exhibiting orbit dimensions, partially-symmetric ranks, border ranks and minimal decompositions.

Segre symbol	dim	symRk_p	$\underline{\mathrm{Rk}}$	T
[1 1 1]	12	3	3	$\lambda{\otimes}x^2{+}(\lambda{+}\mu){\otimes}y^2{+}\mu{\otimes}z^2$
[2 1]	11	4	3	$\lambda{\otimes}(x{+}y)^2{+}(\mu{-}\lambda){\otimes}x^2{-}\lambda{\otimes}y^2{+}\mu{\otimes}z^2$
[(1 1) 1]	10	3	3	$\lambda{\otimes}x^2{+}\lambda{\otimes}y^2{+}\mu{\otimes}z^2$
[3]	10	4	3	
[(2 1)]	9	4	3	$\lambda{\otimes}(x{+}y)^2{+}(\mu{-}\lambda){\otimes}x^2{-}\lambda{\otimes}y^2{+}\lambda{\otimes}z^2$
[; 1;]	8	4	3	
[1 1; ; 1]	8	2	2	$\lambda{\otimes}x^2{+}\mu{\otimes}y^2$
[2; ; 1]	7	3	2	$\lambda{\otimes}(x{+}y)^2{+}(\mu{-}\lambda){\otimes}x^2{-}\lambda{\otimes}y^2$

(4.21)

Note: The regular cases in $\mathbb{P}^2$ where (partially-symmetric) rank and border rank do not coincide are exactly the ones where the regular block is not diagonal. The same can be observed in $\mathbb{P}^3$.

Segre symbol	dim	symRk_p	$\underline{\mathrm{Rk}}$	T
[1 1 1 1]	19	4	4	$\lambda{\otimes}x^2{+}(\lambda{+}\mu){\otimes}y^2{+}(\lambda{-}\mu){\otimes}z^2{+}\mu{\otimes}w^2$
[2 1 1]	19	5	4	$\lambda{\otimes}(x{+}y)^2{+}(\mu{-}\lambda){\otimes}x^2{-}\lambda{\otimes}y^2{+}\mu{\otimes}z^2{+}(\lambda{+}\mu){\otimes}w^2$
[(1 1) 1 1]	18	4	4	$\lambda{\otimes}x^2{+}\lambda{\otimes}y^2{+}\mu{\otimes}z^2{+}(\lambda{+}\mu){\otimes}w^2$
[3 1]	18	5	4	
[(2 1) 1]	17	5	4	$\lambda{\otimes}(x{+}y)^2{+}(\mu{-}\lambda){\otimes}x^2{-}\lambda{\otimes}y^2{+}\lambda{\otimes}z^2{+}\mu{\otimes}w^2$
[(1 1 1) 1]	15	4	4	$\lambda{\otimes}x^2{+}\lambda{\otimes}y^2{+}\lambda{\otimes}z^2{+}\mu{\otimes}w^2$
[2 2]	18	5	4	
[(1 1) 2]	17	5	4	$\lambda{\otimes}x^2{+}\lambda{\otimes}y^2{+}\mu{\otimes}(z{+}w)^2{+}(\lambda{-}\mu){\otimes}z^2{-}\mu{\otimes}w^2$
[(1 1) (1 1)]	16	4	4	$\lambda{\otimes}x^2{+}\lambda{\otimes}y^2{+}\mu{\otimes}z^2{+}\mu{\otimes}w^2$
[4]	17	5	4	
[(3 1)]	17	5	4	
[(2 2)]	15	6	4	$\lambda{\otimes}(x{+}y)^2{+}(\mu{-}\lambda){\otimes}x^2{-}\lambda{\otimes}y^2{+}$ $+\lambda{\otimes}(z{+}w)^2{+}(\mu{-}\lambda){\otimes}z^2{-}\lambda{\otimes}w^2$
[(2 1 1)]	14	5	4	$\lambda{\otimes}(x{+}y)^2{+}(\mu{-}\lambda){\otimes}x^2{-}\lambda{\otimes}y^2{+}\lambda{\otimes}z^2{+}\lambda{\otimes}w^2$

(4.22)

Remark 4.5.8. The pencils of quadrics [2 2] and [(2 2)] in $\mathbb{P}^3$ make fully appreciate theorem 4.3.9 and the weight δ in formula (4.15). Indeed both pencils have two Jordan blocks of size 2 but in [2 2] they are related to different roots while in [(2 2)] to the same root: this means that in [2 2] there is one only invariant polynomial and it is non-squarefree, hence the weight is $\delta = 1$; instead in [(2 2)] there are two invariant polynomials both non-squarefree, hence the weight is $\delta = 2$. This is why $\mathrm{symRk}_p([2\ 2]) = 5$ while $\mathrm{symRk}_p([(2\ 2)]) = 6$.

Remark 4.5.9. All orbits in table (4.22) have border rank 4 since the 4-th secant variety $\sigma_4(\mathbb{P}^1 \times \nu_2(\mathbb{P}^3))$ fills up the ambient space $\mathbb{P}^{19}$. Actually, every orbit in $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^4)$ has border rank ≤ 3 if and only if the corresponding pencil is singular.

Chapter 5

Abelian and nonabelian apolarity

In this chapter we introduce the abelian and nonabelian apolarity theory. In the first section we show the classical catalecticant method for decomposing a given $f \in \operatorname{Sym}^d V$: the Waring decomposition of f is to be found in the zeros of its apolar *ideal $f^\perp$ and the kernels of the catalecticant maps $C_{e,f}$ help to restrict the research loci. Since this method fails in many cases, in the second section we introduce a generalization in terms of vector bundles: this language not only allows to solve the gap of the previous method but it also suggests new methods for the decomposition of general tensors (even not symmetric). The first two sections are just introductory and motivational to the last one, where we recover the nonabelian apolarity for symmetric 2-slice tensors by deriving it from the Kronecker decomposition of matrix pencils.*

Let $\mathbb{K}$ be an algebraically closed field with characteristic 0.
We will work with finite dimensional vector spaces over $\mathbb{K}$.

5.1 Catalecticant method in $\operatorname{Sym}^d V$ and apolarity

Goal: Express $f \in \operatorname{Sym}^d V$ as sum of powers of linear form $\sum_{i=0}^r l_i^d$.
Let $x_0, \dots, x_m$ and $\partial_0, \dots \partial_m$ be coordinates on V and $V^\vee$ respectively, where $\partial_i = \frac{\partial}{\partial x_i}$.
A given $g \in \operatorname{Sym}^e V^\vee$ acting on $\operatorname{Sym}^d V$ defines the **contraction** map

$$\begin{array}{cccc} \lrcorner g : & \operatorname{Sym}^d V & \longrightarrow & \operatorname{Sym}^{d-e} V \\ & f & \mapsto & g \cdot f \end{array} \tag{5.1}$$

Remark 5.1.1.

- if $e > d$, then $\lrcorner g \equiv 0$; else if $e = d$, then $\lrcorner g(\operatorname{Sym}^d V) \subseteq \mathbb{K}$;
- the $\operatorname{Sym}^d V \times \operatorname{Sym}^e V^\vee \to \operatorname{Sym}^{d-e} V$ such that $(f, g) \mapsto g \cdot f$ is bilinear.

Proposition 5.1.2. Let $e \leq d$. Let $g \in \operatorname{Sym}^e V^\vee$ and $l^d \in \nu_d(\mathbb{P}V) \subseteq \operatorname{Sym}^d V$. Then

$$g(\partial_0, \dots, \partial_m)(l^d) = \frac{d!}{(d-e)!} g(l) l^{d-e}$$

where $g(l)$ is the evaluation of g at l. In particular, if $e \leq d$

$$g \cdot l^d = 0 \iff g(l) = 0$$

Proof. It is enough to prove for $g = \varphi^e$ where $\varphi = \sum_{i=0}^m \varphi_i \partial_i$ and we do it by induction on e. Our claim is

$$\varphi^e(l^d) = \frac{d!}{(d-e)!} \varphi^e(l) l^{d-e}$$

If $e = 0$, then $l^d = l^d$. Let $e \geq 1$: then

$$\begin{aligned}
\varphi^{e+1}(l^d) &= \varphi \cdot \varphi^e(l^d) \\
&= \varphi \cdot \Big(\frac{d!}{(d-e)!} \varphi^e(l) l^{d-e} \Big) \\
&= \frac{d!}{(d-e)!} \varphi^e(l) \cdot \varphi(l^{d-e}) \\
&= \frac{d!}{(d-e)!} \varphi^e(l) \frac{(d-e)!}{(d-e-1)!} \varphi(l) l^{d-e-1} \\
&= \frac{d!}{(d-(e+1))!} \varphi^{e+1}(l) l^{d-(e+1)}
\end{aligned}$$

□

Corollary 5.1.3. If $e = d$, then $g \cdot l^d = d! g(l) \in \mathbb{K}$.

Now we introduce an important tool to reach our goal.

Definition. The **apolar ideal** of a given $f \in \operatorname{Sym}^d V$ is the polynomial ideal

$$f^\perp = \{ g \in \operatorname{Sym}^\bullet V^\vee \mid g \cdot f = 0 \} \subset \mathbb{K}[\partial_0, \dots, \partial_m]$$

Definition. The **apolar ring** of a given $f \in \operatorname{Sym}^d V$ is the quotient ring

$$A_f = \mathbb{K}[\partial_0, \dots, \partial_m] / f^\perp$$

Example 5.1.4. If $f = x_0 x_1 x_2 \in \operatorname{Sym}^3 V$, then $f^\perp = (\partial_0^2, \partial_1^2, \partial_2^2) \subset \mathbb{K}[\partial_0, \partial_1, \partial_2]$.

Let us underline many basic properties of these two *apolar* objects: given $f \in \operatorname{Sym}^d(\mathbb{K}^{m+1})$, it holds

- $f^\perp$ is a homogeneous ideal in $\mathbb{K}[\partial_0, \dots, \partial_m]$;

- when $e > d = \deg(f)$, then $(f^{\perp})_e = \mathrm{Sym}^e V^{\vee}$;
- A_f is artinian (hence noetherian and 0-dimensional);
- $(A_f)_e = {\mathrm{Sym}^e V^{\vee}}\big/{(f^{\perp})_e}$;
- $(f^{\perp})_d$ is a 1-codimensional subspace of $\mathrm{Sym}^d V^{\vee}$: indeed

$$(f^{\perp})_d = \ker(\mathrm{Sym}^d V^{\vee} \to \mathbb{K})$$

Next we investigate some more advanced apolar properties.

Proposition 5.1.5. $(f^{\perp})_d$ determines the whole apolar ideal $f^{\perp}$. More precisely, given $\mathfrak{m} = (\partial_0, \ldots, \partial_m)$ the irrelevant (maximal) ideal in $\mathbb{K}[\partial_0, \ldots, \partial_m]$, for all $i \leq d$

$$(f^{\perp})_i = \Big((f^{\perp})_d \ : \ \mathfrak{m}^{d-i}\Big)_i = \Big\{g \in \mathrm{Sym}^i V^{\vee} \mid g \cdot \partial^{\alpha} \in (f^{\perp})_d,\ \forall \partial^{\alpha} \in \mathbb{K}[\partial_0, \ldots, \partial_m]_{d-i}\Big\}$$

Definition. Let V, W be finite dimensional $\mathbb{K}$-vector spaces. A **perfect pairing** of V and W is a bilinear map

$$A : V \times W \to \mathbb{K}$$

such that

$(P1)$ $A(v, w) = 0\ \forall v \in V \implies w = 0$;

$(P2)$ $A(v, w) = 0\ \forall w \in W \implies v = 0$.

Note: One may say that V and W are a perfect pairing if there exists a bilinear map A that satisfies $(P1)$ and $(P2)$. If V and W are a perfect pairing, then $V \simeq W^{\vee}$ (in particular, $\dim V = \dim W$) since

$$\begin{array}{ccc} V & \to & W^{\vee} \\ v & \mapsto & A(v, \cdot) \end{array} \qquad \begin{array}{ccc} W & \to & V^{\vee} \\ w & \mapsto & A(\cdot, w) \end{array}$$

are injective for $(P1)$ and $(P2)$ respectively.

Definition. Let $\mathcal{A}$ be a graded algebra over $\mathbb{K}$, finite dimensional ad $\mathbb{K}$-vector space (i.e. $\mathcal{A} = \bigoplus_{i=0}^{N} \mathcal{A}_i$). Then $\mathcal{A}$ is **Gorenstein of socle** N (or N-Gorenstein) if its graded components satisfy the Poincarè duality, that is $\mathcal{A}_N \simeq \mathbb{K}$ and for all $i \leq N$

$$\mathcal{A}_i \times \mathcal{A}_{N-i} \longrightarrow \mathcal{A}_N \simeq \mathbb{K}$$

is a perfect pairing. In particular, $\dim \mathcal{A}_i = \dim \mathcal{A}_{N-i}$.

Theorem 5.1.6. The apolar ring A_f is Gorenstein of socle d.

Proof. Consider $[l] \in (A_f)_i$ and $[m] \in (A_f)_{d-i}$: then $[m \cdot l] \in (A_f)_d \simeq \mathbb{K}$. We want to prove that

$$[m \cdot l] = [0] \ \forall l \implies [m] = [0]$$

But $[m \cdot l] = [0] \ \forall l$ means that $m \cdot l \in (f^\perp)_d \ \forall l \in \operatorname{Sym}^i V^\vee$, that is $m \in (f^\perp)_{d-i}$, hence $[m] = [0]$. □

Actually a more general result holds [21, Lemma 2.14]:

Theorem 5.1.7 (Macaulay, 1916)**.**

$$\frac{\mathbb{K}[x_0, \dots, x_m]}{I} \text{ is artinian and } d\text{-Gorenstein} \iff I = f^\perp \text{ for some } f \in \operatorname{Sym}^d(\mathbb{K}^{m+1})$$

The next result is the main bridge between the Waring decomposition of a polynomial and its apolar ideal.

Lemma 5.1.8 (Apolarity)**.** Let $\mathcal{Z}$ be a finite set of linear forms on V (or equivalently a finite set of hyperplanes on $V^\vee$) and let $f \in \operatorname{Sym}^d V$. Then

$$f = \sum_{l \in \mathcal{Z}} l^d \iff \mathcal{I}_{\mathcal{Z}} \subseteq f^\perp$$

where $\mathcal{I}_{\mathcal{Z}} = \{g \in \operatorname{Sym}^\bullet V^\vee \mid g(l) = 0 \ \forall l \in \mathcal{Z}\}$.

Proof. Given $g \in \mathcal{I}_{\mathcal{Z}}$, we have $g \cdot f = \sum_{\mathcal{Z}} g \cdot l^d = \sum_{\mathcal{Z}} g(l) l^{d-\deg(g)} = 0$, hence $\mathcal{I}_{\mathcal{Z}} \subseteq f^\perp$. Conversely, by hypothesis $(\mathcal{I}_{\mathcal{Z}})_d \subseteq (f^\perp)_d$: we may work with this graded component since by proposition 5.1.5 it determines $f^\perp$. Then $(\bigcap_{l \in \mathcal{Z}} \mathcal{I}_l)_d \subseteq (f^\perp)_d$ and by applying the $\perp$-operator we have

$$(f^\perp)_d \subseteq \sum_{l \in \mathcal{Z}} (l^d) = \sum_{l \in \mathcal{Z}} (\mathcal{I}_l)_d^\perp$$

hence $f = \sum_{l \in \mathcal{Z}} l^d$. □

Remark 5.1.9. By apolarity lemma 5.1.8 if we look for a decomposition of f we may compute $f^\perp$ and look in there for 0-dimensional ideals (indeed these are related to finite sets).

Example 5.1.10. Consider $f = xyz$. Then $(xyz)^\perp = (\partial_x^2, \partial_y^2, \partial_z^2)$. The ideal $(\partial_x^2 - \partial_y^2, \partial_y^2 - \partial_z^2) \subset (xyz)^\perp$ is the ideal of the four points $(\pm 1, \pm 1, 1)$ and we may consider

$$l_1 = x + y + z \ , \ l_2 = x - y + z \ , \ l_3 = -x + y + z \ , \ l_4 = -x - y + z$$

Then $l_1^3 - l_2^3 - l_3^3 + l_4^3 = 24xyz$ gives a decomposition of f (up to scalar).

The 2-dimensional case. When $\dim V = 2$ the apolarity lemma 5.1.8 allows to compute both rank and border-rank. We may split the apolar ideal $f^{\perp}$ as

$$\begin{aligned} f^{\perp} &= \{g \in \mathrm{Sym}^{\bullet} V^{\vee} \mid g \cdot f = 0\} = \\ &= \bigoplus_k \{g \in \mathrm{Sym}^k V^{\vee} \mid g \cdot f = 0\} = \\ &= \sum_k \left\{ \ker(C_{k,f} : \mathrm{Sym}^k V^{\vee} \to \mathrm{Sym}^{d-k} V) \right\} \end{aligned}$$

Remark 5.1.11. When f is of the form $f = \sum_i a_i \binom{d}{i} x^{d-i} y^i$, then the k-th catalecticant map for f, with respect to the basis $(\frac{\partial}{\partial x^i \partial y^{k-i}})$ and $(x^j y^{d-k-j})$, corresponds to a *Hostle* matrix (i.e. a *Toeplitz* with respect to the inverse diagonal). Hence apolarity lemma 5.1.8 underlines the importance of the kernels above.

We showed that, given $f \in \mathrm{Sym}^d V$, it holds $\mathrm{Rk}(C_{1,f}) \leq \mathrm{symRk}(f)$; moreover, we recall that $\mathrm{Rk}(f) \leq \mathrm{symRk}(f)$. But we can say more.

Lemma 5.1.12. If $\mathrm{Rk}(f) \leq r$, then $\mathrm{Rk}(C_{k,f}) \leq r$.

Proof. If $f = x^d$ the $C_{k,f} = \langle x^{d-k} \rangle$ has dimension 1. Now let $f = \sum_{i=1}^r l_i^d$: then

$$\mathrm{Rk}\left(C_{k,\sum_{i=1}^r l_i^d}\right) = \mathrm{Rk}\left(\sum_{i=1}^r C_{k,l_i^d}\right) \leq \sum_{i=1}^r \mathrm{Rk}(C_{k,l_i^d})$$

but the latter ranks are equally 1, so $\mathrm{Rk}\left(C_{k,\sum_{i=1}^r l_i^d}\right) \leq r$. □

The catalecticant maps are also linked to the secant variety of the *Veronese parabola* $\nu_2\mathbb{P}^1$: the following result was at first proven by Gundelfinger [17] but for a proof using apolarity theory we refer to Kung [24].

Theorem 5.1.13 (Gundelfinger)**.** $X = \sigma_k(\nu_2\mathbb{P}^1)$ is schematically given by the $(k+1)$-minors of the catalecticant $C_{[\frac{d}{2}],f}$. Moreover the singular locus is $\mathrm{Sing}(X) = \sigma_{k-1}(\nu_2\mathbb{P}^1)$ (in general it just holds the restraint $\supseteq$).

By the above theorem it follows an important characterization of the border rank in terms of catalecticant maps:

Corollary 5.1.14. Let $f \in \mathrm{Sym}^d \mathbb{C}^2$. Then $\underline{\mathrm{Rk}}(f) = \mathrm{Rk}(C_{[\frac{d}{2}],f})$.

We recall that the **generic rank** in $\mathrm{Sym}^d \mathbb{C}^2$ is the first k such that the k-th secant variety $\sigma_k(\nu_d(\mathbb{P}^1))$ fills the ambient space. By the previous corollary it follows that for d even the generic rank is

$$\frac{d}{2} + 1 = \dim \mathrm{Sym}^{\frac{d}{2}} \mathbb{C}^2$$

But for d odd the generic rank is

$$\frac{d+1}{2} = \mathrm{Rk}\left(\mathrm{Sym}^{\frac{d+1}{2}}\mathbb{C}^{2\vee} \to \mathrm{Sym}^{\frac{d-1}{2}}\mathbb{C}^2\right)$$

Theorem 5.1.15 (Sylvester)**.** Let d be odd. The generic $f \in \mathrm{Sym}^d\mathbb{C}^2$ has a unique decomposition with $\frac{d+1}{2}$ summands. In particular, its canonical form is

$$f = \sum_{i=1}^{\frac{d+1}{2}} l_i^d$$

Proof. $\mathrm{Sym}^{\frac{d+1}{2}}\mathbb{C}^2 \to \mathrm{Sym}^{\frac{d-1}{2}}\mathbb{C}^2$ has maximal rank for f general. If the rank is maximum, the catalecticant map has 1-dimensional kernel generated by a unique polynomial of degree $\frac{d+1}{2}$ of the form

$$\prod_i(-\beta_i\partial_x + \alpha_i\partial_y)$$

Then by apolarity lemma 5.1.8 $f = \sum_i(\alpha_i x + \beta_i y)^d$, hence we conclude. □

Remark 5.1.16. For d even the generic rank is $\frac{d}{2}+1$ but there are infinitely many decompositions: indeed the catalecticant

$$\mathrm{Sym}^{\frac{d}{2}+1}\mathbb{C}^2 \to \mathrm{Sym}^{\frac{d}{2}-1}\mathbb{C}^2$$

has kernel of dimension 2.

Next we go back to consider a general $(m+1)$-dimensional vector space V. As we have seen for $m = 1$, the catalecticant maps allow to completely determine the rank of a given homogeneous polynomial. But for $m \geq 2$ this method is not always successful. We conclude the study of the catalecticant method by exhibiting an algorithm for computing the rank of a homogeneous polynomial and (unfortunately) by analyzing when this method fails: this *impass* leads us to introduce new methods which will be the main characters of the next section.

Catalecticant algorithm. This algorithm is due to Iarrobino and Kanev (1999, [21]). Before listing it we remark some properties which derive from what we said so far: given $f \in \mathrm{Sym}^d V$ we have

- for all k-th catalecticant map for f, $\mathrm{Rk}(C_{k,f}) \leq \mathrm{Rk}(f)$;
- for all k-th catalecticant map for f, $\ker C_{k,f} \subset f^{\perp}$;
- by abelian apolarity lemma 5.1.8 we may find the linear summands of f in finite sets whose ideal is contained in $f^{\perp}$.

We are ready for the algorithm. Let $f \in \mathrm{Sym}^d V$ where $\dim V = m+1$.

(1) Construct the $\lceil \frac{d}{2} \rceil$-th catalecticant map for f

$$C_{\lceil \frac{d}{2} \rceil, f} : \mathrm{Sym}^{\lceil \frac{d}{2} \rceil} V^\vee \to \mathrm{Sym}^{d - \lceil \frac{d}{2} \rceil} V$$

(2) Compute $\ker C_{\lceil \frac{d}{2} \rceil, f}$;

(3) Compute the Krull dimension $\dim_{\mathrm{Krull}}(\ker C_{\lceil \frac{d}{2} \rceil, f})$:

 (a) if it is ≥ 1, the method fails!

 (b) else compute $\mathcal{Z} = \mathcal{Z}(\ker C_{\lceil \frac{d}{2} \rceil, f}) = \{[l_1], \ldots, [l_r]\}$;

(4) Solve the linear system $f = \sum_{i=1}^r c_i l_i^d$ where c_i are the indeterminates.

The next result gives us a sufficient condition for its success: it also motivates the choice of the value $\lceil \frac{d}{2} \rceil$ and for a complete proof we refer to [30, Theorem 2.4].

Theorem 5.1.17. Let $f = \sum_{i=1}^r l_i^d \in \mathrm{Sym}^d V$ be a general form of rank r, let $z_i = [l_i] \in \mathbb{P}V$ and $\mathcal{Z} = \{z_1, \ldots, z_r\}$.

(i) If d is even and $r \leq \binom{m + \lceil \frac{d}{2} \rceil}{m} - m - 1$ or d is odd and $r \leq \binom{m + \lceil \frac{d}{2} \rceil - 1}{m}$, then

$$\ker C_{\lceil \frac{d}{2} \rceil, f} = \mathcal{I}_{\mathcal{Z}} \cap \mathrm{Sym}^{\lceil \frac{d}{2} \rceil} V$$

and the above algorithm outputs the Waring decomposition of f.

(ii) If d is even but $r = \binom{m + \lceil \frac{d}{2} \rceil}{m} - m$, then it is possible that the finite set $\mathcal{Z}'$ produced by the algorithm is such that $\mathcal{Z} \subsetneq \mathcal{Z}'$. In particular,

- if $m = 2$ the algorithm succeed;
- if $m \geq 3$ the algorithm will succeed after iterating step (4) finitely many times using $\mathcal{Z}'' \subset \mathcal{Z}$ of size $\mathrm{Rk}(C_{\lceil \frac{d}{2} \rceil, f})$.

Failing cases. The catalecticant method can work for a given $f \in \mathrm{Sym}^d V$ only if

$$\mathrm{Rk}(f) \lneq \binom{m + \lfloor \frac{d}{2} \rfloor}{\lfloor \frac{d}{2} \rfloor}$$

Indeed we have that the maximum rank of a catalecticant map is

$$\mathrm{Rk}\left(C_{\lceil \frac{d}{2} \rceil, f} : \mathrm{Sym}^{\lceil \frac{d}{2} \rceil} V^\vee \to \mathrm{Sym}^{\lfloor \frac{d}{2} \rfloor} V\right) = \binom{m + \lfloor \frac{d}{2} \rfloor}{\lfloor \frac{d}{2} \rfloor}$$

and if r is greater or equal either (for d even) there is no kernel to work with or (for d odd) the equality $\ker C_{\lceil \frac{d}{2} \rceil, f} = \mathcal{I}_{\mathcal{Z}} \cap \mathrm{Sym}^{\lceil \frac{d}{2} \rceil} V$ fails. Usually the general rank $\frac{\binom{m+d}{d}}{m+1}$ is larger than $\binom{m + \lfloor \frac{d}{2} \rfloor}{\lfloor \frac{d}{2} \rfloor}$, then cases where the method fails are not rare.

5.2 Modern vector bundles tools and nonabelian apolarity

The dictionary of the modern language for tensor rank decomposition is given by vector bundles and their sections. But before introducing its general setting it is worth underlining that actually the catalecticant method hides the vector bundles language behind the isomoprhism

$$\mathrm{Sym}^d(\mathbb{C}^{m+1})^\vee \simeq H^0(\mathbb{P}^m, \mathcal{O}(d))$$

So let us reinterpret the catalecticant method in terms of vector bundles. Given $e \leq d$, the contraction map (4.3) gives the linear map

$$H^0(\mathbb{P}^m, \mathcal{O}(e)) \otimes H^0(\mathbb{P}^m, \mathcal{O}(d))^\vee \longrightarrow H^0(\mathbb{P}^m, \mathcal{O}(d-e))^\vee = H^0\Big(\mathbb{P}^m, \mathcal{O}(e)^\vee \otimes \mathcal{O}(d)\Big)^\vee$$

and, by fixing $f \in H^0(\mathbb{P}^m, \mathcal{O}(d))^\vee$, we have the "cohomological" catalecticant map

$$C_{e,f} : H^0(\mathbb{P}^m, \mathcal{O}(e)) \longrightarrow H^0\Big(\mathbb{P}^m, \mathcal{O}(e)^\vee \otimes \mathcal{O}(d)\Big)^\vee \tag{5.2}$$

Thus we may reformulate apolarity lemma 5.1.8 as follows [26]:

Lemma 5.2.1 (Abelian apolarity)**.** Let $f = \sum_{i=1}^r z_i \in H^0(\mathbb{P}^m, \mathcal{O}(d))^\vee$ and let $\mathcal{Z} = \{[z_1], \dots, [z_r]\} \subseteq \mathbb{P}(H^0(\mathbb{P}^m, \mathcal{O}(d))^\vee)$. Then

$$H^0(\mathbb{P}^m, \mathcal{I}_\mathcal{Z} \otimes \mathcal{O}(e)) \subseteq \ker(C_{e,f}) \quad , \quad H^0(\mathbb{P}^m, \mathcal{I}_\mathcal{Z} \otimes \mathcal{O}(d-e)) \subseteq (\operatorname{Im} C_{e,f})^\perp$$

Moreover, if $H^0(\mathbb{P}^m, \mathcal{O}(d-e)) \to H^0(\mathbb{P}^m, \mathcal{O}(d-e)_{|\mathcal{Z}})$ is surjective (resp. $H^0(\mathbb{P}^m, \mathcal{O}(e)) \to H^0(\mathbb{P}^m, \mathcal{O}(e)_{|\mathcal{Z}})$), then the first inclusion is an equality (resp. the second).

Let us explain what the above theorem states. We refer to the *base locus* of $H^0(X, \mathcal{E})$ (for a given vector bundle $\mathcal{E}$ on a variety X) as the locus of common zeros to all the sections of the space. Now let $f = \sum_{[z_i] \in \mathcal{Z}} z_i \in H^0(\mathbb{P}^m, \mathcal{O}(d))^\vee$ and assume $H^0(\mathbb{P}^m, \mathcal{O}(d-e)) \twoheadrightarrow H^0(\mathbb{P}^m, \mathcal{O}(d-e)_{|\mathcal{Z}})$ to be surjective: then $\ker(C_{e,f}) = H^0(\mathbb{P}^m, \mathcal{I}_\mathcal{Z} \otimes \mathcal{O}(e))$ and this means that the base locus of $H^0(\mathbb{P}^m, \mathcal{I}_\mathcal{Z} \otimes \mathcal{O}(e))$ is $\mathcal{Z}$ itself, that is the decomposition of f can be computed from the base locus of $\ker C_{e,f}$.

Remark 5.2.2. The advantage of the reformulation of the apolarity due to Oeding and Ottaviani [30] is that $\ker(C_{e,f})$ is computable by an explicit matrix construction via the *presentation* of a bundle. Moreover, by studying $\ker(C_{e,f})$ one works with polynomials of lower degree than the starting polynomial f and this is computationally more efficient.

We may generalize the map (5.2) to a general vector bundle $\mathcal{E}$ on an algebraic variety X (instead of $\mathcal{O}(e)$ and $\mathbb{P}^m$ respectively) and get a line bundle $\mathcal{L}$ which gives the embedding $X \subset \mathbb{P}(H^0(X, \mathcal{L})^\vee) = \mathbb{P}^N$ (it exists by *Kodaira embedding*). By doing this we extend the abelian apolarity (lemma 5.2.1) to a *nonabelian apolarity*.

Remark 5.2.3. The term "nonabelian" may be misleading in the following sense: we refer to apolarity in lemma 5.2.1 as "abelian" since $\mathcal{E} = \mathcal{O}(e)$ gives the abelian object $H^0(\mathbb{P}^m, \mathcal{O}(e)) \simeq \mathrm{Sym}^e V^\vee$, while we refer to "nonabelian" apolarity for general vector bundles $\mathcal{E}$ (which do not necessarily lead to abelian objects). In this sense the abelian apolarity is actually a particular case of the nonabelian one, despite the name. Maybe one may target as nonabelian only when working with vector bundles of rank ≥ 2 but this would not emphasize how the nonabelian apolarity naturally arises from the abelian one.

Then the natural contraction map

$$H^0(X, \mathcal{E}) \otimes H^0(X, \mathcal{E}^\vee \otimes \mathcal{L}) \to H^0(X, \mathcal{L}) \tag{5.3}$$

leads to the linear map

$$H^0(X, \mathcal{E}) \otimes H^0(X, \mathcal{L})^\vee \to H^0(X, \mathcal{E}^\vee \otimes \mathcal{L})^\vee$$

and, by fixing $f \in H^0(X, \mathcal{L})^\vee$, to the linear map

$$C_{\mathcal{E},f} : H^0(X, \mathcal{E}) \to H^0(X, \mathcal{E}^\vee \otimes \mathcal{L})^\vee \tag{5.4}$$

which depends linearly on f. Then lemma 5.2.1 generalizes to the following result.

Proposition 5.2.4. Let $f = \sum_{i=1}^r z_i \in H^0(X, \mathcal{L})^\vee$ and let $\mathcal{Z} = \{[z_1], \ldots, [z_r]\} \subseteq \mathbb{P}(H^0(X, \mathcal{L})^\vee)$. Then

$$H^0(X, \mathcal{I}_\mathcal{Z} \otimes \mathcal{E}) \subseteq \ker(C_{\mathcal{E},f}) \quad , \quad H^0(X, \mathcal{I}_\mathcal{Z} \otimes \mathcal{E}^\vee \otimes \mathcal{L}) \subseteq (\mathrm{Im}\, C_{\mathcal{E},f})^\perp$$

Moreover, if $H^0(X, \mathcal{E}^\vee \otimes \mathcal{L}) \to H^0(X, \mathcal{E}^\vee \otimes \mathcal{L}_{|\mathcal{Z}})$ is surjective (resp. $H^0(X, \mathcal{E}) \to H^0(X, \mathcal{E}_{|\mathcal{Z}})$), then the first inclusion is an equality (resp. the second).

Now let $f = \sum_{i=1}^r z_i$ be a minimal decomposition and let $\mathcal{Z} = \{[z_1], \ldots, [z_r]\} \subseteq \mathbb{P}(H^0(X, \mathcal{L})^\vee)$. Let us denote with $\mathrm{Rk}(\mathcal{E})$ be the rank of the vector bundle $\mathcal{E}$ (i.e. the dimension of its fibers).

Lemma 5.2.5 (Nonabelian apolarity)**.** If $\mathrm{Rk}(C_{\mathcal{E},f}) = r \cdot \mathrm{Rk}(\mathcal{E})$, then

$$H^0(X, \mathcal{I}_\mathcal{Z} \otimes \mathcal{E}) = \ker(C_{\mathcal{E},f}) \quad , \quad H^0(X, \mathcal{I}_\mathcal{Z} \otimes \mathcal{E}^\vee \otimes \mathcal{L}) = (\mathrm{Im}\, C_{\mathcal{E},f})^\perp$$

Proof. By proposition 5.2.4 we already know the inclusions $\subseteq$ hold. But by the hypothesis on the ranks it follows

$$\mathrm{codim}\, H^0(X, \mathcal{I}_\mathcal{Z} \otimes \mathcal{E}) \leq r \cdot \mathrm{Rk}(\mathcal{E}) = \mathrm{Rk}(C_{\mathcal{E},f}) = \mathrm{codim} \ker(C_{\mathcal{E},f})$$

$$\mathrm{codim}\, H^0(X, \mathcal{I}_\mathcal{Z} \otimes \mathcal{E}^\vee \otimes \mathcal{L}) \leq r \cdot \mathrm{Rk}(\mathcal{E}) = \mathrm{Rk}(C_{\mathcal{E},f}) = \mathrm{codim}(\mathrm{Im}\, C_{\mathcal{E},f})^\perp$$

□

Now we need to introduce a class of projective varieties and a result about them due to Chiantini and Ciliberto [9, Theorem 1.4].

Definition. A projective variety X is said to be **k-weakly defective** if its intersection with a general k-tangent hyperplane has no isolated singularities at the k points of tangency.

Theorem 5.2.6 (Chiantini, Ciliberto)**.** Let $X \subseteq \mathbb{P}^N$ be a projective variety of dimension n and let k be such that $N \geq (n+1)k$. If X is not k-weakly defective, then, given $P_1, \dots, P_k$ general points on X, the general k-tangent hyperplane $H \in \mathcal{H}(-2P_1 - \dots - 2P_k)$ is tangent to X only at exactly those k points.

The following theorem allows to completely determine the set $\mathcal{Z}$ as base locus of sections.

Theorem 5.2.7. Let $\mathrm{Rk}(C_{\mathcal{E},f}) = r \cdot \mathrm{Rk}(\mathcal{E})$, let X be not r-weakly defective and let the contraction map

$$H^0(X, \mathcal{I}_{\mathcal{Z}} \otimes \mathcal{E}) \otimes H^0(X, \mathcal{I}_{\mathcal{Z}} \otimes \mathcal{E}^\vee \otimes \mathcal{L}) \to H^0(X, \mathcal{I}_{\mathcal{Z}}^2 \otimes \mathcal{L})$$

be surjective. Then the common base locus of $\ker(C_{\mathcal{E},f})$ and $(\operatorname{Im} C_{\mathcal{E},f})^\perp$ is $\mathcal{Z}$ itself.

Proof. We know that the equalities in lemma 5.2.5 hold. If the common base locus contained $\mathcal{Z} \cup \{\hat{z}\}$, then the whole $H^0(X, \mathcal{I}_{\mathcal{Z}}^2 \otimes \mathcal{L})$ would vanish doubly on $\hat{z}$, thus the general hyperplane section of $X \subseteq \mathbb{P}(H^0(X, \mathcal{L})^\vee)$ is singular at both $\mathcal{Z}$ and $\hat{z}$, hence in $r+1$ tangent points in contraddiction to theorem 5.2.6. □

5.2.1 Eigenvectors of tensors and presentations

We are now interested in to exhibit an algorithm for tensor decomposition based on the nonabelian apolarity we have introduced so far. We want to leverage the advantage we underlined in remark 5.2.2: in particular this matricial perspective allows to interpret the *base loci* of sections of a suitable vector bundle as (a sort of) *eigenvectors of a tensor*. First of all we introduce the vector bundle $\mathcal{E}$ we will work with.

Definition. Let V be a $m+1$-dimensional $\mathbb{K}$-vector space. The **quotient bundle** Q of $\mathbb{P}V$ is the cokernel bundle of the monomorphism $\mathcal{O}_{\mathbb{P}V}(-1) \hookrightarrow \mathcal{O}_{\mathbb{P}V} \otimes V$. The **Euler exact sequence** for $\mathbb{P}V$ is the SES

$$0 \longrightarrow \mathcal{O}_{\mathbb{P}V}(-1) \longrightarrow \mathcal{O}_{\mathbb{P}V} \otimes V \longrightarrow Q \longrightarrow 0 \tag{5.5}$$

By taking wedge a-powers and by tensoring by $\mathcal{O}(e)$ we get the SES

$$0 \longrightarrow \bigwedge^{a-1} V \otimes \mathcal{O}(e-1) \longrightarrow \bigwedge^{a} V \otimes \mathcal{O}(e) \longrightarrow \bigwedge^{a} Q(e) \longrightarrow 0 \tag{5.6}$$

The vector bundle we will work with is $\mathcal{E} = \bigwedge^a Q(e)$.

Eigenvectors for tensors. The next step is to introduce the eigenvectors for a tensor and to find a relation with sections of $\bigwedge^a Q(e)$. Since

$$H^0\Big(\mathbb{P}V, \bigwedge^h V \otimes \mathcal{O}(k)\Big) \simeq \bigwedge^h V \otimes \operatorname{Sym}^k V^\vee \simeq \operatorname{Hom}\Big(\operatorname{Sym}^k V, \bigwedge^h V\Big)$$

the previous SES gives in cohomology

$$0 \longrightarrow \operatorname{Hom}\Big(\operatorname{Sym}^{e-1} V, \bigwedge^{a-1} V\Big) \longrightarrow \operatorname{Hom}\Big(\operatorname{Sym}^{e} V, \bigwedge^{a} V\Big) \xrightarrow{\psi} H^0\Big(\mathbb{P}V, \bigwedge^a Q(e)\Big) \longrightarrow 0$$

Let us fix the following notation: given a tensor $M \in \operatorname{Hom}(\operatorname{Sym}^e V, \bigwedge^a V)$, we denote the corresponding global section $\psi(M)$ by $s_M \in H^0(\mathbb{P}V, \bigwedge^a Q(e))$.

Definition. Given a tensor $M \in \operatorname{Hom}(\operatorname{Sym}^e V, \bigwedge^a V)$, a vector $v \in V$ is said to be **eigenvector of the tensor** M if

$$M(v^e) \wedge v = 0$$

Note: For $e = a = 1$ the above definition coincides with the classical one of eigenvector of a linear map: indeed $M(v) \wedge v = 0$ if and only if $M(v) \in \langle v \rangle_{\mathbb{C}}$.

Lemma 5.2.8. *(i)* For all $[v] \in \mathbb{P}V$ the fiber

$$\Big(\bigwedge^a Q(e)\Big)_{[v]} \simeq \operatorname{Hom}\left(\langle v^e \rangle, \frac{\bigwedge^a V}{\langle v \wedge \bigwedge^{a-1} V\rangle}\right) \tag{5.7}$$

(ii) $\forall M \in \operatorname{Hom}(\operatorname{Sym}^e V, \bigwedge^a V)$, $s_M([v]) = 0$ if anf only if v is eigenvector for M.

Proof. Consider the composition

$$\langle v^e \rangle \hookrightarrow \operatorname{Sym}^e V \xrightarrow{M} \bigwedge^a V \overset{\pi}{\twoheadrightarrow} \frac{\bigwedge^a V}{\langle v \wedge \bigwedge^{a-1} V\rangle}$$

which on the fiber in $[v]$ corresponds to the section s_M. Then $s_M([v]) = 0$ if and only if $\pi \circ M(v^e) = 0$ if and only if $M(v^e) \wedge v = 0$. □

Corollary 5.2.9. The common base locus of $H^0(\mathbb{P}V, \bigwedge^a Q(e))$ corresponds to the common eigenvectors for $\operatorname{Hom}(\operatorname{Sym}^e V, \bigwedge^a V)$.

Presentations of bundles. Before clarifying how sections and eigenvectors for tensors apply to tensor decomposition, we briefly introduce the *presentation* of a bundle $\mathcal{E}$ on $\mathbb{P}V$. Consider the (finite) *minimal resolution* of the bundle $\mathcal{E}$

$$\ldots \to L_2 \to L_1 \to \mathcal{E} \to 0$$

where each L_i is direct sum of line bundles on $\mathbb{P}V$: since it is obtained by sheafificating the corresponding minimal free resolution of the graded module $\bigoplus_k H^0(\mathbb{P}V, \mathcal{E}(k))$, the map $H^0(\mathbb{P}V, L_1) \stackrel{\alpha}{\to} H^0(\mathbb{P}V, \mathcal{E})$ is surjective.
We may also consider the (finite) minimal resolution of $\mathcal{E}^\vee$

$$\ldots \to L_{-1}^\vee \to L_0^\vee \to \mathcal{E}^\vee \to 0$$

where each $L_{-i}^\vee$ is a direct sum of line bundles: in this case, for all line bundle $\mathcal{L}$, the map $H^0(\mathbb{P}V, L_0^\vee \otimes \mathcal{L}) \stackrel{{}^t\beta}{\to} H^0(\mathbb{P}V, E^\vee \otimes \mathcal{L})$ is surjective.
Dualizing we have the resolution

$$\ldots \to L_2 \to L_1 \stackrel{\rho}{\to} L_0 \to L_{-1} \to \ldots$$

where $\mathrm{Im}(\rho) = \mathcal{E}$. We define $\rho : L_1 \to L_0$ to be the **presentation** of $\mathcal{E}$.

The bridge. We are now ready to build the bridge from the sections of the vector bundle $\mathcal{E} = \bigwedge^a Q(e)$ (or equivalently the eigenvectors for $M \in \mathrm{Hom}(\mathrm{Sym}^e V, \bigwedge^a V)$) to the tensor decomposition of a given $f \in H^0(\mathbb{P}^m, \mathcal{L})^\vee = \mathrm{Sym}^d V$, where $\mathcal{L} = \mathcal{O}(d)$.
In the new hypothesis the map (5.4) becomes

$$C_{e,f} : H^0\Big(\mathbb{P}^m, \bigwedge^a Q(e)\Big) \longrightarrow H^0\Big(\mathbb{P}^m, \bigwedge^a Q(e)^\vee \otimes \mathcal{L}\Big)^\vee \tag{5.8}$$

where $f \in H^0(\mathbb{P}^m, \mathcal{L})^\vee$. Given $\rho : L_1 \to L_0$ the presentation of $\bigwedge^a Q(e)$, consider the composition

$$P_{e,f} : H^0(\mathbb{P}^m, L_1) \stackrel{\alpha}{\to} H^0(\mathbb{P}^m, \bigwedge^a Q(e)) \xrightarrow{C_{e,f}} H^0(\mathbb{P}^m, \bigwedge^a Q(e)^\vee \otimes \mathcal{L})^\vee \stackrel{\beta}{\to} H^0(\mathbb{P}^m, L_0^\vee \otimes \mathcal{L})^\vee \tag{5.9}$$

where α is surjective and β is injective beacuse of what we said on presentations of bundles.

Corollary 5.2.10. In the above notations:

(*i*) $\mathrm{Rk}(C_{e,f}) = \mathrm{Rk}(P_{e,f})$;

(*ii*) $\ker(C_{e,f})$ and $\ker(P_{e,f})$ have the same base locus.

Remark 5.2.11. The advantage of working with $P_{e,f}$ is that it can be computed from a matrix whose entries are homogeneous polynomials. In particular the matrix representing $P_{e,f}$ can be constructed by the presentation ρ by replacing the variable x_i by the catalecticant $C_{\lceil \frac{d}{2} \rceil, \partial_i f}$ if ρ has a linear entry depending on x_i, otherwise by replacing each monomial $g(x_i)$ in the x_i by the catalecticant of $g(\partial_i) \cdot f$ [26, §8.3].

The presentation $\rho : L_1 \to L_0$ of $\mathcal{E} = \bigwedge^a Q(e)$ on $\mathbb{P}V = \mathbb{P}^m$ is

$$\rho : \bigwedge^{a} V \otimes \mathcal{O}(e) \longrightarrow \bigwedge^{m-a} V^\vee \otimes \mathcal{O}(e+1)$$

Since $H^0(\mathbb{P}^m, L_1) = \mathrm{Hom}(\mathrm{Sym}^e V, \bigwedge^a V)$ and

$$\begin{aligned} H^0(\mathbb{P}^m, L_0^\vee \otimes \mathcal{L})^\vee &= H^0\Big(\mathbb{P}^m, \bigwedge^{m-a} V \otimes \mathcal{O}(e+1)^\vee \otimes \mathcal{O}(d)\Big)^\vee = H^0\Big(\mathbb{P}^m, \bigwedge^{m-a} V \otimes \mathcal{O}(d-e-1)\Big)^\vee \\ &= \mathrm{Hom}\left(\mathrm{Sym}^{d-e-1} V, \bigwedge^{m-a} V\right)^\vee = \mathrm{Hom}\left(\bigwedge^{m-a} V, \mathrm{Sym}^{d-e-1} V\right) \end{aligned}$$

from (5.9) we have

$$P_{e,f} : \mathrm{Hom}\left(\mathrm{Sym}^e V, \bigwedge^{a} V\right) \longrightarrow \mathrm{Hom}\left(\bigwedge^{m-a} V, \mathrm{Sym}^{d-e-1} V\right) \tag{5.10}$$

In particular, (5.10) is defined for $f = v^d$ by

$$P_{e,v^d}(M)(w) = \Big(M(v^e) \wedge v \wedge w\Big) \cdot v^{e-m-1} \tag{5.11}$$

where $M \in \mathrm{Hom}(\mathrm{Sym}^e V, \bigwedge^a V)$, $w \in \bigwedge^{m-a} V$ and $(M(v^e) \wedge v \wedge w) \in \bigwedge^{m+1} V \simeq \mathbb{K}$, then extendend by linearity to any $f \in \mathrm{Sym}^d V$.

Proposition 5.2.12. Let $M \in \mathrm{Hom}(\mathrm{Sym}^e V, \bigwedge^a V)$.

(*i*) $v \in V$ is eigenvector for M if and only if $M \in \ker(P_{e,v^d})$.

(*ii*) Given $f = \sum_{i=1}^r v_i^d$, if each v_i is eigenvector for M, then $M \in \ker(P_{e,f})$.

Proof. (*i*) By (5.11) we get

$$M(v^e) \wedge v = 0 \iff \forall w \in \bigwedge^{m-a} V,\ M(v^e) \wedge v \wedge w = 0 \iff M \in \ker(P_{e,v^d})$$

(*ii*) The thesis follows by linearity of the map (5.10) with respect to f.

□

Now we are ready to draw our conclusions. Let $f = \sum_{i=1}^r v_i^d \in \mathrm{Sym}^d V$ and $\mathcal{Z} = \{[v_1^d], \ldots, [v_r^d]\}$. We assume to work with the vector bundle $\mathcal{E} = \bigwedge^a Q(e)$ for suitable a, e. By theorem 5.2.7, we look for $\mathcal{Z}$ in the common base locus of $\ker(C_{e,f})$, thus by corollary 5.2.10 in the common base locus of $\ker(P_{e,f})$, hence by proposition 5.2.12 between the common eigenvectors for $\ker(P_{e,f})$.

5.3 Nonabelian apolarity for symmetric 2-slice tensors

We are now interested into finding relations between the nonabelian apolarity and the Kronecker form of symmetric 2-slice tensors in $\mathbb{C}^2 \otimes \mathrm{Sym}^2 V$. The vector bundles over $\mathbb{P}^m$ we will work with are the tangent bundle $T\mathbb{P}^m$ and the bundle of holomorphic 1-forms $\Omega^1_{\mathbb{P}^m}$:we obtain both of them by starting from $\mathcal{E} = \bigwedge^a Q(e)$ for $a = e = 1$.
Let V be a $m+1$-dimensional $\mathbb{C}$-vector space and $\mathbb{P}^m = \mathbb{P}V$ its projective space. We recall the Euler sequence (5.5) and we twist it by tensoring by $\mathcal{O}(1)$: thus we have the SES

$$0 \to \mathcal{O} \to \mathcal{O}(1) \otimes V \to Q(1) \to 0$$

Remark 5.3.1. The above SES is actually the SES (5.6) for $a = e = 1$.

Tangent bundle. Our first claim is to prove that the above twisted quotient bundle $Q(1)$ is neverthless that the *tangent bundle* $T\mathbb{P}^m$ on $\mathbb{P}^m$.
First it is worth understanding which are the fibers of $T\mathbb{P}^m$. Let us fix $[x] \in \mathbb{P}^m$. Since $\mathrm{GL}_{m+1}(\mathbb{C})$ acts transitively on $\mathbb{P}^m$, the map $\mathrm{GL}_{m+1}(\mathbb{C}) \to \mathbb{P}^m$ such that $g \mapsto g \cdot [x] = [g \cdot x]$ and $I \mapsto [x]$ is surjective and *differentiates* in the identity I to the surjection

$$\mathfrak{gl}_{m+1}(\mathbb{C}) = T_I\Big(\mathrm{GL}_{m+1}(\mathbb{C})\Big) \longrightarrow T_{[x]}\mathbb{P}^m$$

whith kernel $\{g \mid g\langle x\rangle \subseteq \langle x\rangle\}$. Thus it follows that the fibers of the tangent bundle are

$$T_{[x]}\mathbb{P}^m \simeq \frac{\mathfrak{gl}_{m+1}(\mathbb{C})}{\{g \mid g\langle x\rangle \subseteq \langle x\rangle\}} \tag{5.12}$$

Remark 5.3.2. Given $g \in \mathfrak{gl}_{m+1}(\mathbb{C})$, it defines for all $[x] \in \mathbb{P}^m$ an element in $T_{[x]}\mathbb{P}^m$ which vanishes if and only if x is eigenvector for g: indeed $g \mapsto 0 \in T_{[x]}\mathbb{P}^m$ if and only if $g\langle x\rangle \subset \langle x\rangle$ if and only if x is eigenvector for g.

Moreover, by (5.7) we know that the fibers of the twisted quotient bundle $Q(1)$ are

$$\Big(Q(1)\Big)_{[x]} \simeq \mathrm{Hom}\left(\langle x\rangle, \frac{V}{\langle x\rangle}\right) \tag{5.13}$$

Proposition 5.3.3. The isomorphism $Q(1) \simeq T\mathbb{P}^m$ of vector bundles holds.

Proof. We prove the isomorphism (of vector spaces) for the fibers $(Q(1))_{[x]} \simeq T_{[x]}\mathbb{P}^m$. Let us fix $[x] \in \mathbb{P}^m$. We consider the natural map

$$\begin{array}{ccc} \mathfrak{gl}_{m+1}(\mathbb{C}) & \longrightarrow & \mathrm{Hom}\left(\langle x\rangle, \frac{V}{\langle x\rangle}\right) \\ g & \mapsto & \pi_{\langle x\rangle} \circ g_{|_{\langle x\rangle}} \end{array}$$

where $\pi_{\langle x\rangle}$ is the quotient projection: this map is surjective for $x \neq 0$ and its kernel is $\{g \mid g\langle x\rangle \subseteq \langle x\rangle\}$, hence by (5.12) and (5.13) it follows $T_{[x]}\mathbb{P}^m \simeq (Q(1))_{[x]}$. $\square$

Corollary 5.3.4. The following short sequence is exact

$$0 \longrightarrow \mathcal{O} \longrightarrow \mathcal{O}(1) \otimes V \longrightarrow T\mathbb{P}^m \longrightarrow 0 \tag{5.14}$$

The SES (5.14) leads to the SES in cohomology

$$0 \to H^0(\mathbb{P}^m, \mathcal{O}) \to H^0(\mathbb{P}^m, \mathcal{O}(1) \otimes V) \to H^0(\mathbb{P}^m, T\mathbb{P}^m) \to 0$$

(since $H^1(\mathbb{P}^m, \mathcal{O}) = 0$). Now $H^0(\mathbb{P}^m, \mathcal{O}) \simeq \mathbb{C}$ and

$$H^0(\mathbb{P}^m, \mathcal{O}(1) \otimes V) \simeq H^0(\mathbb{P}^m, \mathcal{O}(1))^{\oplus m+1} \simeq \left((\mathbb{C}^{m+1})^\vee\right)^{\oplus m+1} \simeq \mathfrak{gl}_{m+1}(\mathbb{C})$$

By exactness we have $h^0(\mathbb{P}^m, T\mathbb{P}^m) = (m+1)^2 - 1$. Moreover, $H^0(\mathbb{P}^m, T\mathbb{P}^m)$ has a natural Lie algebra structure by (5.12). Since the map $H^0(\mathbb{P}^m, \mathcal{O}) \to H^0(\mathbb{P}^m, \mathcal{O}(1) \otimes V)$ corresponds to

$$\begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathfrak{gl}_{m+1}(\mathbb{C}) \\ \lambda & \mapsto & \lambda I \end{array}$$

whose cokernel is the Lie algebra $\mathfrak{sl}_{m+1}(\mathbb{C})$ of traceless matrices, it follows

$$H^0(\mathbb{P}^m, T\mathbb{P}^m) \simeq \mathfrak{sl}_{m+1}(\mathbb{C})$$

Remark 5.3.5. By remark 5.3.2, if $g \in \mathfrak{gl}_{m+1}(\mathbb{C})$ vanishes in $T_{[x]}\mathbb{P}^m$ for all $[x] \in \mathbb{P}^m$, then it must be of the form $g = \lambda I$, that is it induces the zero section $s_g = 0$ in $H^0(\mathbb{P}^m, T\mathbb{P}^m)$. In particular, the common base locus of $H^0(\mathbb{P}^m, T\mathbb{P}^m)$ coincides with the common eigenvectors for $\mathfrak{gl}_{m+1}(\mathbb{C})$, as corollary (5.2.9) exactly states.

Holomorphic 1-forms. Since $(T\mathbb{P}^m)^\vee \simeq \Omega^1_{\mathbb{P}^m}$, if we dualize and tensorize (5.14) by $\mathcal{O}(2)$ we get the SES

$$0 \longrightarrow \Omega^1(2) \longrightarrow \mathcal{O}(1) \otimes V^\vee \longrightarrow \mathcal{O}(2) \longrightarrow 0$$

(the functor $\mathrm{Hom}(\bullet, \mathbb{C})$ is exact since $\mathbb{C}$ is injective as $\mathbb{C}$-module), thus in cohomology

$$0 \to H^0(\mathbb{P}^m, \Omega^1(2)) \to V^\vee \otimes V^\vee \to \mathrm{Sym}^2 V^\vee \to 0$$

(since $H^1(\mathbb{P}^m, \Omega^1(2)) = 0$ by Bott's formula [4]). Moreover, $H^0(\mathbb{P}^m, \Omega^1(2)) \simeq \bigwedge^2 V^\vee$ is (isomorphic to) the space of *skewsymmetric* matrices of size $m+1$.

Nonabelian apolarity in $\mathbb{C}^2 \otimes \operatorname{Sym}^2 V$. Let $T \in \mathbb{C}^2 \otimes \operatorname{Sym}^2 V$ be a symmetric 2-slice tensor with associated symmetric pencil defined by the pair (B_1, B_2) of symmetric matrices of size $m+1$. We recall that the map in (5.4) was defined by fixing $f \in H^0(\mathbb{P}^m, \mathcal{L})^\vee$: for $\mathcal{L} = \mathcal{O}(2)$ we may see f as a symmetric matrices, so if we have a pair of symmetric matrices (B_1, B_2) it comes quite natural to pick $(B_1, B_2) \in H^0(\mathbb{P}^m, \mathcal{O}(2))^\vee \oplus H^0(\mathbb{P}^m, \mathcal{O}(2))^\vee$. Thus we set $\mathcal{L} = \mathcal{O}(2)$.
We now *extend* the natural contraction map (5.3) to

$$H^0(\mathbb{P}^m, \mathcal{E}) \otimes \Big(H^0(\mathbb{P}^m, \mathcal{E}^\vee(2)) \oplus H^0(\mathbb{P}^m, \mathcal{E}^\vee(2))\Big) \to H^0(\mathbb{P}^m, \mathcal{O}(2)) \oplus H^0(\mathbb{P}^m, \mathcal{O}(2))$$

which leads, by fixing $(B_1, B_2) \in H^0(\mathbb{P}^m, \mathcal{O}(2))^\vee \oplus H^0(\mathbb{P}^m, \mathcal{O}(2))^\vee$, to

$$C_{(B_1,B_2)} : H^0(\mathbb{P}^m, \mathcal{E}) \longrightarrow H^0(\mathbb{P}^m, \mathcal{E}^\vee(2))^\vee \oplus H^0(\mathbb{P}^m, \mathcal{E}^\vee(2))^\vee \tag{5.15}$$

which *extends* (5.8). For $\mathcal{E} = T\mathbb{P}^m \simeq Q(1)$ we obtain

$$C_{(B_1,B_2)} : H^0(\mathbb{P}^m, T\mathbb{P}^m) \longrightarrow H^0(\mathbb{P}^m, \Omega^1(2))^\vee \oplus H^0(\mathbb{P}^m, \Omega^1(2))^\vee \tag{5.16}$$

Fortunately we may reintepret the above map in terms of matrices and this leads to an explicit homomorphism.

Proposition 5.3.6. Up to scalars, the map $C_{(B_1,B_2)}$ in (5.16) is equivalent to

$$\begin{array}{rccc} C_{(B_1,B_2)} : & \mathfrak{sl}_{m+1}(\mathbb{C}) & \longrightarrow & \bigwedge^2 V \oplus \bigwedge^2 V \\ & A & \mapsto & \Big(AB_1 - B_1({}^tA)\ ,\ AB_2 - B_2({}^tA)\Big) \end{array} \tag{5.17}$$

Proof. First of all we underline that the above map is well defined since for all matrices A, B the matrix $AB - ({}^tB)({}^tA)$ is skewsymmetric and the pair (B_1, B_2) is symmetric. Clearly it is enough to prove the thesis for

$$\begin{array}{rccc} C_i : & \mathfrak{sl}_{m+1}(\mathbb{C}) \otimes \operatorname{Sym}^2 V & \to & \bigwedge^2 V \\ & A \otimes B_i & \mapsto & AB_i - B_i({}^tA) \end{array}$$

We consider the following actions of $\mathrm{GL}_{m+1}(\mathbb{C})$:

- on $\bigwedge^2 V$ and $\operatorname{Sym}^2 V$ it acts by congruence;
- on $\mathfrak{sl}_{m+1}(\mathbb{C})$, given $G \in \mathrm{GL}_{m+1}(\mathbb{C})$, by $A \mapsto {}^tGA({}^tG)^{-1}$.

The map C_i is a homomorphism of $\mathrm{GL}_{m+1}(\mathbb{C})$-modules: indeed

$$\begin{aligned} {}^tG\Big(AB_i - B_i({}^tA)\Big)G &= ({}^tG)AB_iG - ({}^tG)B_i({}^tA)G \\ &= ({}^tG)A({}^tG)^{-1}({}^tG)B_iG - ({}^tG)B_iG(G^{-1})({}^tA)G \\ &= \Big({}^tGA({}^tG)^{-1}\Big)\Big({}^tGB_iG\Big) - \Big({}^tGB_iG\Big){}^t\Big({}^tGA({}^tG)^{-1}\Big) \end{aligned}$$

As $\mathrm{GL}_{m+1}(\mathbb{C})$-module, $\mathfrak{sl}_{m+1}(\mathbb{C}) \otimes \mathrm{Sym}^2 V$ is completely reducible and it has $\bigwedge^2 V$ as submodule, hence we have the decomposition $\mathfrak{sl}_{m+1}(\mathbb{C}) \otimes \mathrm{Sym}^2 V = \bigwedge^2 V \oplus W$. Moreover, by *Pieri's formula* [13, Proposition 15.25] the irreducible submodule $\bigwedge^2 V$ appears with multiplicity 1 in $\mathfrak{sl}_{m+1}(\mathbb{C}) \otimes \mathrm{Sym}^2 V$.
Now we note that by *Schur's lemma* for all $\varphi_1, \varphi_2 : \bigwedge^2 V \to \bigwedge^2 V$ it holds $\varphi_1 = \lambda\varphi_2$ for some scalar $\lambda \in \mathbb{C}^\times$, hence $C_{i|\bigwedge^2 V}$ is the only one $\mathrm{GL}_{m+1}(\mathbb{C})$-endomorphism (up to scalars) of $\bigwedge^2 V$. Moreover, always by Schur's lemma the only $\mathrm{GL}_{m+1}(\mathbb{C})$-homomorphism $W \to \bigwedge^2 V$ is the identically zero one, thus there is only one extension of $C_{i|\bigwedge^2 V}$ to $\mathfrak{sl}_{m+1}(\mathbb{C}) \otimes \mathrm{Sym}^2 V$ and it has to be C_i. □

Remark 5.3.7. Another possible proof is *by hand*: one explicits the isomorphisms $H^0(\mathbb{P}^m, T\mathbb{P}^m) \simeq \mathfrak{sl}_{m+1}(\mathbb{C})$ and $H^0(\mathbb{P}^m, \Omega^1(2))^\vee \simeq \bigwedge^2 V$ and the map (5.16) on the global sections. We chose a more theoretical proof to underline the $\mathrm{GL}_{m+1}(\mathbb{C})$-module structure of $\mathfrak{sl}_{m+1}(\mathbb{C})$ and $\bigwedge^2 V$ (and to avoid counts as well!).

Remark 5.3.8. One may wonder why we just need (5.17) up to scalars. The reason is because we are interested in studying the kernel of the map to recover its eigenvectors, and the kernel is invariant for scalar multiplication.

Before stating the main result of this section we need to prove some preliminary lemmas to make the proof more natural. To light up the notation we will not make distinction between a symmetric pencil of size $m+1$, its corresponding tensor in $\mathbb{C}^2 \otimes \mathrm{Sym}^2 V$ and the pair of matrices which defines it: hence we will improperly write $(B_1, B_2) \in \mathbb{C}^2 \otimes \mathrm{Sym}^2 V$.

Note: In the following we will work with *general* pencils. In algebraic geometry an element is "general" if it belongs to a dense open subset of the space, or equivalently it does not belong to a closed subset. This notion is a double-edged sword: on one hand it allows to describe properties which hold outside a null subset of the space, so that we may assume for almost all elements those properties; on the other hand there is not an exact characterization for general elements since, if we have a dense open subset whose elements respect a given property $\mathscr{P}$ and we add one single element without that property, the new subset is still a dense open but not all of its elements satisfy $\mathscr{P}$.

We need to clarify as far as possible what it means for a pencil in $\mathbb{C}^2 \otimes \mathrm{Sym}^2 V$ to be *general*. We start by assuming that a general pencil is regular since this condition is equivalent to ask its determinant not to be identically zero, hence it defines a dense open subset.

Lemma 5.3.9. Let $(B_1, B_2) \in \mathbb{C}^2 \otimes \mathrm{Sym}^2 V$ be a general symmetric pencil of size $m+1$. Then we may assume (B_1, B_2) to be regular with no infinite elementary divisors and distinct linear finite elementary divisors, i.e. its Kronecker form to be

$$\mathrm{diag}(\lambda + a_1\mu, \dots, \lambda + a_{m+1}\mu) = \mu\, \mathrm{diag}(a_1, \dots, a_{m+1}) + \lambda I_{m+1}$$

where $a_i \in \mathbb{C}^\times$ are all distincts.

Proof. As mentioned right before stating the lemma, we may assume the pencil to be regular. Up to acting with $\mathrm{GL}_2(\mathbb{C})$ we may also assume the pencil to have no infinite elementary divisors.
Let us focus on the (finite) elementary divisors: since they are defined by the eigenvalues of $B_1B_2^{-1}$, it is enough to prove that a general $B_1B_2^{-1}$ has distinct eigenvalues. But the characteristic polynomial of a general $B_1B_2^{-1}$ is a general polynomial, hence it has distinct roots (otherwise the variety it defines would be contained in a hypersurface, against the hypothesis to be general). By the same argument, if the pencil has non-zero minimal indices, then the characteristic polynomial of $B_1B_2^{-1}$ would vanish in 0, but a general polynomial does not. □

Lemma 5.3.10. The kernel $\ker(C_{(B_1,B_2)})$ of the map in (5.17) is invariant for $\mathrm{GL}_2(\mathbb{C})$-action on $\mathbb{C}^2 \otimes \mathrm{Sym}^d V$. In particular it does not depend only on the matrices B_1, B_2 but on the pencil they define.

Proof. Let $\mu B_1 + \lambda B_2$ be the pencil defined by the pair (B_1, B_2) and let $\tilde{\mu} B_1 + \tilde{\lambda} B_2 = \mu(\alpha B_1 + \gamma B_2) + \lambda(\beta B_1 + \delta B_2) = \mu \tilde{B_1} + \lambda \tilde{B_2}$ be the pencil obtained by acting by GL(2). If $A \in \ker(C_{(B_1,B_2)})$ then

$$A(\alpha B_1 + \gamma B_2) - (\alpha B_1 + \gamma B_2)({}^tA) = \alpha(AB_1 - B_1({}^tA)) + \gamma(AB_2 - B_2({}^tA)) = 0$$

that is $A \in \ker(C_{(\tilde{B_1},\tilde{B_2})})$ and since the linear transformation is invertible the converse holds too. □

Now we state and prove our result for a *general* Kronecker form: the case of a generic *general* pencil is an immediate corollary.

Theorem 5.3.11. Let $(D, I) \in \mathbb{C}^2 \otimes \mathrm{Sym}^2 V$ be a general symmetric pencil in Kronecker form as in lemma 5.3.9. Then:

(i) all matrices in $\ker(C_{(D,I)})$ have the same common (linearly independent) eigenvectors $v_1, \dots, v_{m+1}$; in particular, these vectors are the ones which give the Kronecker form, i.e.

$$T_{(D,I)} = \sum_{i=1}^{m+1} \alpha_i \otimes v_i \otimes v_i$$

where $\alpha_i = (D_{ii}, 1) \in \mathbb{C}^2$;

(*ii*) $\ker(C_{(D,I)})$ has dimension $m+1$ in $\mathfrak{gl}_{m+1}(\mathbb{C})$ and m in $\mathfrak{sl}_{m+1}(\mathbb{C})$;

(*iii*) for $C \in \ker(C_{(D,I)})$ general, in $\mathfrak{gl}_{m+1}(\mathbb{C})$ it holds $\ker(C_{(D,I)}) = \langle I, C, \dots, C^m \rangle_{\mathbb{C}}$. In particular, in $\mathfrak{sl}_{m+1}(\mathbb{C})$ it holds $\ker(C_{(D,I)}) = \langle I, C, \dots, C^m \rangle_{\mathbb{C}} \cap \mathfrak{sl}_{m+1}(\mathbb{C})$.

Proof. Since the pencil is general and it already is in Kronecker form, by lemma 5.3.9 we know that $D = \operatorname{diag}(a_1, \dots, a_{m+1})$ where $a_i \in \mathbb{C}^\times$ are all distinct. Consider the map in (5.17)

$$A \mapsto \Big(AD - D({}^tA)\ ,\ AI - I({}^tA)\Big)$$

(*i*) Let $A \in \ker(C_{(D,I)})$: then from the second component of its image we know that $A = {}^tA$, that is A is symmetric. Then

$$\ker(C_{(D,I)}) \subseteq \mathfrak{sl}_{m+1}(\mathbb{C}) \cap \operatorname{Sym}^2(\mathbb{C}^{m+1})$$

Then from the first component of its image we have $0 = AD - D({}^tA) = AD - DA$, that is

$$\ker(C_{(D,I)}) \subseteq Z_{\mathfrak{sl}}(D)$$

where $Z_{\mathfrak{sl}}(D)$ is the centralizer of D in $\mathfrak{sl}_{m+1}(\mathbb{C})$. Since $Z_{\mathfrak{sl}}(D)$ is an abelian Lie subalgebra of $\mathfrak{sl}_{m+1}(\mathbb{C})$, for all $A, B \in \ker(C_{(D,I)})$ we have the zero bracket $[A, B] = 0$. In particular, $\ker(C_{(D,I)})$ is an abelian Lie subalgebra of $\mathfrak{sl}_{m+1}(\mathbb{C})$. But $Z_{\mathfrak{sl}}(D)$ is a maximal toral subalgebra since D is a semisimple element, hence all matrices in $Z_{\mathfrak{sl}}(D)$ are simultaneously diagonalizable to D with same common eigenvectors $v_1, \dots, v_{m+1}$: in particular, the ones in $\ker(C_{(D,I)})$ are so.

(*ii*) Since $\dim \mathfrak{sl}_{m+1}(\mathbb{C}) = (m+1)^2 - 1$ and $\dim \bigwedge^2 V = \binom{m+1}{2}$, by the first group homomorphism theorem it follows

$$\dim_{\mathfrak{sl}} \ker(C_{(D,I)}) \geq (m+1)^2 - 1 - (m+1)m = m$$

But $\ker(C_{(D,I)}) \subseteq Z_{\mathfrak{sl}}(D)$ and, since D is a semisimple element in the semisimple Lie algebra $\mathfrak{sl}_{m+1}(\mathbb{C})$, $Z_{\mathfrak{sl}}(D)$ is actually a Cartan subalgebra of $\mathfrak{sl}_{m+1}(\mathbb{C})$: since all Cartan subalgebras are conjugated, $\dim Z_{\mathfrak{sl}}(D) = m$, thus

$$\dim_{\mathfrak{sl}} \ker(C_{(D,I)}) \leq m$$

It follows than $\ker(C_{(D,I)})$ has dimension m in $\mathfrak{sl}_{m+1}(\mathbb{C})$, hence $m+1$ in $\mathfrak{gl}_{m+1}(\mathbb{C})$.

(*iii*) Let $C_{(D,I)}$ be defined on $\mathfrak{gl}_{m+1}(\mathbb{C})$. Clearly, if $C \in \ker(C_{(D,I)})$, then $C^k \in \ker(C_{D,I})$ for all $k \geq 0$, but by *Hamilton-Cayley* at least C^{m+1} is linearly dependent on $I, C, \dots, C^m$: it follows that $\langle I, C, \dots, C^m \rangle_{\mathbb{C}} \subseteq \ker(C_{D,I})$. For C general, the $I, C, \dots, C^m$ are linearly independent and, since the kernel has dimension $m+1$ in $\mathfrak{gl}_{m+1}(\mathbb{C})$, the equality holds.

□

Theorem 5.3.12. Let $(B_1, B_2) \in \mathbb{C}^2 \otimes \operatorname{Sym}^2 V$ be a general symmetric pencil as in lemma 5.3.9. Then:

(*i*) all matrices in $\ker(C_{(B_1,B_2)})$ have the same common eigenvectors $v_1, \dots, v_{m+1}$ which are induced by the vectors $\tilde{v}_1, \dots, \tilde{v}_{m+1}$ defining the Kronecker form

$$T_{(B_1,B_2)} \overset{\mathrm{GL}}{\sim} \sum_{i=1}^{m+1} \alpha_i \otimes \tilde{v}_i \otimes \tilde{v}_i$$

where $\alpha_i \in \mathbb{C}^2$;

(*ii*) $\ker(C_{(B_1,B_2)})$ has dimension $m+1$ in $\mathfrak{gl}_{m+1}(\mathbb{C})$ and m in $\mathfrak{sl}_{m+1}(\mathbb{C})$;

(*iii*) for $C \in \ker(C_{(B_1,B_2)})$ general, in $\mathfrak{gl}_{m+1}(\mathbb{C})$ it holds $\ker(C_{(B_1,B_2)}) = \langle I, C, \dots, C^m \rangle_{\mathbb{C}}$. In particular, in $\mathfrak{sl}_{m+1}(\mathbb{C})$ it holds $\ker(C_{(B_1,B_2)}) = \langle I, C, \dots, C^m \rangle_{\mathbb{C}} \cap \mathfrak{sl}_{m+1}(\mathbb{C})$.

Proof. We recall that two complex symmetric pencils are strictly equivalent if and only if they are congruent (by theorem 2.4.2).
It is enough to prove that the kernels of all strictly equivalent pencils are all conjugated: more precisely, for all $P \in \mathrm{GL}_{m+1}(\mathbb{C})$ it holds

$$\ker(C_{(PB_1({}^tP), PB_2({}^tP))}) = P^{-1} \cdot \ker(C_{(B_1,B_2)}) \cdot P$$

Clearly it is enough to prove it with respect to the Kronecker form of the pencil. Since the pencil defined by (B_1, B_2) is general, by lemma 5.3.9 we know that its Kronecker form is defined by the pair (D, I) where D is diagonal with all distinct diagonal elements $a_i \in \mathbb{C}^\times$. Let $P \in \mathrm{GL}_{m+1}(\mathbb{C})$ be such that $(PB_1({}^tP), PB_2({}^tP)) = (D, I)$. Then

$$\begin{aligned} A \in \ker(C_{(D,I)}) &\iff \begin{cases} AD = D({}^tA) \\ AI = I({}^tA) \end{cases} \iff \begin{cases} APB_1({}^tP) = PB_1({}^tP)({}^tA) \\ APB_2({}^tP) = PB_2({}^tP)({}^tA) \end{cases} \\ &\iff \begin{cases} P^{-1}APB_1 = B_1({}^tP)({}^tA)({}^tP)^{-1} \\ P^{-1}APB_2 = B_2({}^tP)({}^tA)({}^tP)^{-1} \end{cases} \\ &\iff \begin{cases} (P^{-1}AP)B_1 = B_1({}^t(P^{-1}AP)) \\ (P^{-1}AP)B_2 = B_2({}^t(P^{-1}AP)) \end{cases} \iff P^{-1}AP \in \ker(C_{B_1,B_2}) \end{aligned}$$

This correspondence makes things easier.

(*i*) By theorem 5.3.11 we know that $\ker(C_{D,I})$ admits the same common eigenvectors $\tilde{v}_1, \dots, \tilde{v}_{m+1}$ and these give the Kronecker form (D, I). Let $A \in \ker(C_{B_1,B_2})$: then

$PAP^{-1} \in \ker(C_{D,I})$, hence for all $i = 1 : m+1$ we have $PAP^{-1}\tilde{v}_i = \lambda_{A,i}\tilde{v}_i$. By setting $v_i = P^{-1}\tilde{v}_i$ for all $i = 1 : m+1$ we have

$$Av_i = A(P^{-1}\tilde{v}_i) = \lambda_{A,i}(P^{-1}\tilde{v}_i) = \lambda_{A,i}v_i$$

Since this works for all $A \in \ker(C_{B_1,B_2})$ it follows that $v_1, \ldots, v_{m+1}$ are common eigenvectors for $\ker(C_{(B_1,B_2)})$.

(*ii*) Since $\ker(C_{B_1,B_2}) = P \cdot \ker(C_{D,I}) \cdot P^{-1}$ and the dimension is invariant for conjugacy, we have that $\ker(C_{(B_1,B_2)})$ has dimension m in $\mathfrak{sl}_{m+1}(\mathbb{C})$ and $m+1$ in $\mathfrak{gl}_{m+1}(\mathbb{C})$.

(*iii*) It follows by the same arguments in theorem 5.3.11(*iii*).

□

Remark 5.3.13. We proved the theorem by assuming the symmetric pencil (B_1, B_2) to be general as in lemma 5.3.9. Actually we may weaken such hypotheses of "generality" by considering regular pencils with no infinite elementary divisors and linear finite elementary divisors not necessarily distincts: in this case theorem 5.3.12 still holds since, given (D, I) the Kronecker form of such a pencil, $\ker(C_{(D,I)}) \subseteq Z_{\mathfrak{sl}}(D)$ and the latter is a maximal toral subalgebra of $\mathfrak{sl}_{m+1}(\mathbb{C})$, hence all its elements are simultaneously diagonalizable.

Chapter 6

Implementations on Macaulay2

6.1 Kronecker invariants for matrix pencils

Remark 6.1.1 (True Sad Story). Unfortunately the author found out the existence of an already-implemented package[1] in Macaulay2 only after he independently implemented some functions by himself: this is why in the following a final function exhibiting the Kronecker form of a pencil is missing and we only show the implementations made by the author.

Remark 6.1.2. We set the polynomial ring we will work in to be `R=QQ[x,y]`. We work over $\mathbb{Q}$ to be sure to not have any problem between *exact arithmetic* and *approximated arithmetic*: indeed we are not interested into find the roots of the invariant polynomials but just into factorize them into the elementary divisors and these will be in $\mathbb{Q}[x,y]$ too.

Computing invariant polynomials:

```
--input:  M matrix with coefficients in R
--output:  L list of invariant polynomials of M

invpoly=(M)->(
  r=rank(M);
  d1=(gens saturate minors(r,M))_(0,0);
  -- computation of the gcd of the minors of size r
  for i from 0 to r-1 list (
    d1=(gens saturate minors(r-i,M))_(0,0);
    d2=(gens saturate minors(r-i-1,M))_(0,0);
    -- iterative division of the gcd of (r-i)-minors by the gcd of (r-i-1)-minors
    continue (r-i,d1//d2) )
  )
```

[1] http://www2.macaulay2.com/Macaulay2/doc/Macaulay2/share/doc/Macaulay2/Kronecker/html/

```
L=invpoly(M)
List
```

Comment: The function `invpoly` outputs the list of the invariant polynomials of the pencil in the following order: the invariant polynomial `(L_j)_1` divides the preceeding one `(L_{j-1})_1` .

Computing elementary divisors:

```
--input L=invpoly(M) list of invariant polynomials of M
--output:  E list of elementary divisors of M

elemdiv=(L)->(
  r=length(L);
  i=0;
  P=(L_0)_1;
  -- fix the bigger invariant polynomial (the other ones divide this)
  while P%x^(i+1)!=0 do i=i+1;
  for i from 0 to r-1 list (
    if (L_i)_1!=1 and gcd(x^i,(L_i)_1)!=1 then continue gcd(x^i,(L_i)_1) ),
    -- the infinite elementary divisors has been found
  P=P//x^i;
  F=factor P;
  -- the elementary divisors which are maxima powers have been found
  l=#F;
  for j from 0 to l-1 list (
    for i from 0 to r-1 do (
      if gcd(value F#j,(L_i)_1)!=1 then continue gcd(value F#j,(L_i)_1) ))
      -- the other elementary divisors are obtained as gcd between the first
      -- ones and the invariant polynomials of lower degree
  )
E=elemdiv(M)
List
```

Comment: The function `elemdiv` outputs the list of the elementary divisors of the pencil in the following order: first the infinite ones with decreasing exponents, second the finite ones which divide the bigger invariant polynomial, then the others with decreasing degree.

Computing minimal indices:

```
--input:  M matrix with coefficients in R
--output:  I list of minimal indices for columns of M, J list of minimal indices
    for rows of M

minindxcol=(M)-> (
```

```
  k=mingens ker M;
  C=degrees k;
  C1=C_1;
  c1=length C1;
  for i from 0 to c1-1 list (
    if (C1_i)_0==0 then continue (C1_i)_0
    else continue (C1_i)_0-1 )
  )
I=minindxcol(M)
List

minindxrow=(M)-> (
  N=transpose M;
  h=mingens ker N;
  D=degrees h;
  D1=D_1;
  r1=length D1;
  for j from 0 to r1-1 list (continue (D1_j)_0)
  )
J=minindxrow(M)
List
```

Comment: The functions `minindxcol` and `minindxrow` output the list of minimal indices for columns and rows respectively. We just show the algorithm for columns since the one for rows is similar up to transposing the pencil: the algorithm first computes the minimal generators of the kernel and the vectors of their degrees via the functions `mingens` and `degrees`, then it outputs the degrees in decreasing order and outputs them as a list.

6.2 Dimension of GL-orbits in $\mathbb{K}^2 \otimes \mathbb{K}^n \otimes \mathbb{K}^n$

In the following we implement the computation of the orbit dimension of a given tensor $t \in \mathbb{K}^2 \otimes \mathbb{K}^n \otimes \mathbb{K}^n$ (actually in $\mathbb{K}^n \otimes \mathbb{K}^n \otimes \mathbb{K}^2$). Here we set $n = 3$ and $t = a_2 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + a_2 \otimes b_3 \otimes c_3 + a_1 \otimes b_1 \otimes c_2 + a_1 \otimes b_2 \otimes c_3$ (see table (4.13)).
First we define the polynomial ring we will work on, the generic matrices acting on the tensor space and the tensor whose orbit dimension we want to compute:

```
n=3
A=a_(0,0)..a_(n-1,n-1)
B=b_(0,0)..b_(n-1,n-1)
C=c_(0,0)..c_(1,1)
R2=QQ[A,B,C,x_0..x_(n-1),y_0..y_(n-1),z_0,z_1]
pp=transpose genericMatrix(R2,a_(0,0),n,n)
```

```
qq=transpose genericMatrix(R2,b_(0,0),n,n)
cc=transpose genericMatrix(R2,c_(0,0),2,2)
t=z_1*x_0*y_0+z_1*x_1*y_1+z_1*x_2*y_2+z_0*x_0*y_1+z_0*x_1*y_2
```

We now define the row-vector of length 22 defining a general element in the acting group $\mathfrak{gl}_3(\mathbb{K}) \times \mathfrak{gl}_3(\mathbb{K}) \times \mathfrak{gl}_2(\mathbb{K})$ and the row-vector of length 18 defining a general element in the tensor product $\mathbb{K}^3 \otimes \mathbb{K}^3 \otimes K^2$:

```
abc=matrix{{A,B,C}}
xyz=matrix{{x_0..x_(n-1)}}**matrix{{y_0..y_(n-1)}}**matrix{{z_0,z_1}}
```

Next we define the group action grp and its jacobian jac, then we compute their substitutions at the identity $I = (pp^0, qq^0, cc^0)$ and finally the orbit dimension of t in $\mathbb{K}^3 \otimes \mathbb{K}^3 \otimes \mathbb{K}^2$:

```
Gx=apply(n,i->(x_i=>sum(n,j->x_j*a_(i,j))))
Gy=apply(n,i->(y_i=>sum(n,j->y_j*b_(i,j))))
Gz=apply(2,i->(z_i=>sum(2,j->z_j*c_(i,j))))
grp=Gx|Gy|Gz
jac=diff(xyz,diff(transpose abc,sub(t,grp)))
ident=apply(n,j->(a_(0,j)=>(pp^0)_(0,j)))
for i from 1 to n-1 do
  ident=ident|apply(n,j->(a_(i,j)=>(pp^0)_(i,j)))
for i from 0 to n-1 do
  ident=ident|apply(n,j->(b_(i,j)=>(qq^0)_(i,j)))
for i from 0 to 1 do
  ident=ident|apply(2,j->(c_(i,j)=>(cc^0)_(i,j)))
rank sub(jac,ident)
```

6.3 Dimension of GL-orbits in $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^n)$

In the following we readapt the previous implementation to compute the orbit dimension for symmetric 2-slice tensors. Here we set $n = 3$ and $t = \lambda \otimes x_0^2$.
The idea is the same as the previous one with few modifications since we are in the symmetric case. First of all the acting group we work with is $\mathfrak{gl}_2(\mathbb{K}) \times \mathfrak{gl}_n(\mathbb{K})$: if we represent an element $f \in \mathrm{Sym}^2(\mathbb{K}^n)$ by a symmetric matrix F, given a tensor $u \otimes F \in \mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^n)$, the action on $u \otimes F$ is given by

$$(M, P) \mapsto Mu \otimes P \cdot F \cdot {}^t P$$

So we need to define only one generic matrix and transpose it, that is we need to introduce less variables than previously.
We start by defining the polynomial ring and the generic acting matrices:

```
n=3
A=a_(0,0)..a_(n-1,n-1)
C=c_(0,0)..c_(1,1)
R2=QQ[A,C,x_0..x_(n-1),z_0,z_1]
pp=transpose genericMatrix(R2,a_(0,0),n,n)
cc=transpose genericMatrix(R2,c_(0,0),2,2)
t=z_1*x_0^2
```

Next we define the generic symmetric matrix with quadratic monomial entries and we compute the minima generators of its 1×1 minors. With the latter command we get the quadratic monomials as a base of $\mathrm{Sym}^2(\mathbb{K}^n)$ and we can compute the basis of the spaces $\mathfrak{gl}_2 \times \mathfrak{gl}_n(\mathbb{K})$ and $\mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^n)$:

```
A=transpose matrix{{x_0..x_(n-1)}}*matrix{{x_0..x_(n-1)}}
S=mingens minors(1,A)
abc=matrix{{a_(0,0)..a_(n-1,n-1),c_(0,0)..c_(1,1)}}
xyz=S**matrix{{z_0,z_1}}
```

Now we define the action on the basis and compute its jacobian and the substitution at the identity $I = (pp^0, cc^0)$. We underline that by acting on the variables x_i instead of the variables $x_i x_j$ we hide the right-multiplication by ${}^tP$ in the sums $\sum_k x_k P_{ik}$. Finally we compute the orbit dimension by computing the rank:

```
Gn=apply(n,i->(x_i=>sum(n,j->x_j*pp_(i,j))))
G2=apply(2,i->(z_i=>sum(2,j->z_j*cc_(i,j))))
grp=Gn|G2
jac=diff(xyz,diff(transpose abc,sub(t,grp)))
ident=apply(n,j->(a_(0,j)=>(pp^0)_(0,j)))
for i from 1 to n-1 do
  ident=ident|apply(n,j->(a_(i,j)=>(pp^0)_(i,j)))
for i from 0 to 1 do
  ident=ident|apply(2,j->(c_(i,j)=>(cc^0)_(i,j)))
rank sub(jac,ident)
```

6.4 Nonabelian apolarity for symmetric 2-slice tensors

The following implementation was born by an unpublished research work of Prof. Giorgio Ottaviani (University of Florence).

We will work in dimension 3. First we set the polynomial ring whose variables correspond to the entries of a generic square matrix of size 3:

```
n=3
R=QQ[a_(0,0)..a_(n-1,n-1)]
aa=transpose genericMatrix(R,a_(0,0),n,n)
```

Next we define a pencil $(b1s, b2s) \in \mathbb{K}^2 \otimes \mathrm{Sym}^2(\mathbb{K}^3)$ of general type as in lemma 5.3.9 so that it will decompose in three summands:

```
b1s=matrix{{1,0,0_R},{0,1,0},{0,0,0}}
b2s=matrix{{0,0,0_R},{0,5,0},{0,0,1}}
```

Now we define the map $C_{(b1s,b2s)}$ (5.17) extended to $\mathfrak{gl}_3(\mathbb{K})$ and we call it `cont`:

```
D1=diff(transpose basis(1,R),mingens minors(1,aa*b1s-b1s*transpose(aa)))
D2=diff(transpose basis(1,R),mingens minors(1,aa*b2s-b2s*transpose(aa)))
cont=D1|D2
```

The map has size 9×6: indeed $\dim \mathfrak{gl}_3 = 9$ and $\dim \bigwedge^2(\mathbb{C}^3) = 3$. Next we compute its rank and the dimension r of its kernel in $\mathfrak{gl}_3(\mathbb{K})$: by theorem 5.3.12 we expect $r = 3$.

```
rank cont
r=numcols gens kernel transpose cont
```

We now want to compute the simultaneous eigenvectors of the kernel since by theorem 5.3.12 they give the decomposition of the pencil. To do so we first compute the matrices k_0, k_1, k_2 which generate $\ker(C_{(b1s,b2s)})$ in $\mathfrak{gl}_3(\mathbb{K})$:

```
for i from 0 to r-1 do
  k_i=diff(aa,(basis(1,R)*(gens kernel transpose cont))_(0,i))
```

The three elements k_0, k_1, k_2 in the kernel are $\mathrm{diag}(1,0,0), \mathrm{diag}(0,1,0), \mathrm{diag}(0,0,1)$ respectively and they actually commute, as expected. Moreover, they give the simultaneous eigenvectors e_1, e_2, e_3 (vectors in the canonical basis): these actually give the Kronecker form of the pencil $(b1s, b2s)$ (it was already in such form!).

We now rewrite the previous implementation starting from a random symmetric pencil $(b1s, b2s)$: clearly by randomly generating it we risk to work with a symmetric pencil which is not in general form as in lemma 5.3.9, hence theorem 5.3.12 may fail. So in the new setting the previous implementation becomes:

```
n=3
R=QQ[a_(0,0)..a_(n-1,n-1)]
aa=transpose genericMatrix(R,a_(0,0),n,n)
b1=random(R^{n:0},R^{n:0})
b1s=b1+transpose(b1)
b2=random(R^{n:0},R^{n:0})
b2s=b2+transpose(b2)
D1=diff(transpose basis(1,R),mingens minors(1,aa*b1s-b1s*transpose(aa)))
D2=diff(transpose basis(1,R),mingens minors(1,aa*b2s-b2s*transpose(aa)))
cont=D1|D2
r=numcols gens kernel transpose cont
for i from 0 to r-1 do
  k_i=diff(aa,(basis(1,R)*(gens kernel transpose cont))_(0,i))
```

A generic eigenvector v for k_i may be computed by setting $v = \mathtt{aa}_0$ (the first column of the generic matrix aa) and by imposing that the matrix $N_i = \left[k_i \cdot v \;\middle|\; v\right]$ has rank 1, that is $k_i \cdot v = \alpha_i v$ for a suitable $\alpha_i \in \mathbb{K}$. But asking such rank to be 1 is equivalent to asking v to be in the base locus of the ideal I generated by the 2×2 minors of N_i. Even better, asking a given v to be common eigenvector for $k_0, \ldots, k_{r-1}$ simultaneously is equivalent to asking v to be in the base locus of the ideal I generated by the 2×2 minors of all N_i for $i = 0 : r - 1$. Hence we compute such ideal:

```
I=minors(2,k_0*aa_{0}|aa_{0})
for i from 1 to (r-1) do
  I=I+minors(2,k_i*aa_{0}|aa_{0})
```

However we may explicitly compute the common eigenvectors by determining the eigenvectors of one of the matrices which generate $\ker(C_{(b1s,b2s)})$, since from the hypothesis of "generality" in lemma 5.3.9 we are considering matrices with distinct eigenvalues. In the following v_i are the common eigenvectors and M_i the matrices corresponding to

$v_i \otimes v_i$: we underline that we need to work in $\mathbb{C}$ to compute the eigenvectors and this unfortunately leads to approximations.

```
for i from 0 to r-1 do
  (v_i=matrix ((eigenvectors(sub(k_1,CC)))_1)_i, M_i=v_i*transpose v_i)
```

The matrices M_i actually decompose the pencil $(b1s, b2s)$, that is there exist vectors $(\alpha_i, \beta_i) \in \mathbb{C}^2$ such that $(b1s, b2s) = \left(\sum_{i=0}^{r-1} \alpha_i M_i, \sum_{i=0}^{r-1} \beta_i M_i \right)$.

Bibliography

[1] M.F. Atiyah and I.G. Macdonald. *Introduction to commutative algebra.* Addison-Wesley Series in Mathematics. Westview Press, Boulder, CO, 2016.

[2] P. Belmans. “Segre symbols”. *preprint* (2016).

[3] G.D. Birkhoff. “Singular points of ordinary linear differential equations”. *Trans. Amer. Math. Soc.* 10.4 (1909), pp. 436–470. ISSN: 0002-9947.

[4] R. Bott. “Homogeneous vector bundles”. *Annals of mathematics* (1957), pp. 203–248.

[5] J. Buczyński, K. Han, M. Mella, and Z. Teitler. “On the locus of points of high rank”. *Eur. J. Math.* 4.1 (2018), pp. 113–136. ISSN: 2199-675X.

[6] J. Buczyński and J.M. Landsberg. “Ranks of tensors and a generalization of secant varieties”. *Linear Algebra and its Applications* 438.2 (2013), pp. 668–689.

[7] P. Bürgisser, M. Clausen, and M.A. Shokrollahi. *Algebraic complexity theory.* Vol. 315. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1997.

[8] E. Carlini and J. Kleppe. “Ranks derived from multilinear maps”. *Journal of Pure and Applied Algebra* 215.8 (2011), pp. 1999–2004.

[9] L. Chiantini and C. Ciliberto. “Weakly defective varieties”. *Transactions of the American Mathematical Society* 354.1 (2002), pp. 151–178.

[10] A. Dimca. “A geometric approach to the classification of pencils of quadrics”. *Geometriae Dedicata* 14.2 (1983), pp. 105–111.

[11] J. Ding and A. Zhou. “Eigenvalues of rank-one updated matrices with some applications”. *Applied Mathematics Letters* 20.12 (2007), pp. 1223–1226.

[12] D. Eisenbud and J. Harris. *The geometry of schemes.* Vol. 197. Springer Science & Business Media, 2006.

[13] W. Fulton and J. Harris. *Representation theory: a first course.* Vol. 129. Springer Science & Business Media, 2013.

[14] F.R. Gantmacher. *Theory of Matrices, Vol. II.* AMS Chelsea Publishing Co., New York, 1959.

[15] F.R. Gantmacher. *The theory of matrices. Vol. 1.* AMS Chelsea Publishing, Providence, RI, 1998.

[16] A. Grothendieck. "Sur la classification des fibrés holomorphes sur la sphère de Riemann". *Amer. J. Math.* 79 (1957), pp. 121–138. ISSN: 0002-9327.

[17] S. Gundelfinger. "Zur Theorie der binären Formen." *Journal für die reine und angewandte Mathematik* 100 (1887), pp. 413–424.

[18] R. Hartshorne. *Algebraic geometry.* Vol. 52. Springer Science & Business Media, 2013.

[19] M. Hazewinkel and C.F. Martin. "A short elementary proof of Grothendieck's theorem on algebraic vectorbundles over the projective line". *J. Pure Appl. Algebra* 25.2 (1982), pp. 207–211. ISSN: 0022-4049.

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

[21] A. Iarrobino and V. Kanev. *Power sums, Gorenstein algebras, and determinantal loci.* Springer Science & Business Media, 1999.

[22] J. Ja'Ja'. "Optimal evaluation of pairs of bilinear forms". *SIAM J. Comput.* 8.3 (1979), pp. 443–462. ISSN: 0097-5397.

[23] L. Kronecker. "Algebraische Reduktion der Scharen quadratischen Formen". *Sitzungsber. Berl. Akad.* (1890), pp. 763–776, 1225–1375.

[24] J.P.S. Kung. "Gundelfinger's theorem on binary forms". *Studies in Applied Mathematics* 75.2 (1986), pp. 163–169.

[25] J.M. Landsberg. *Tensors: geometry and applications.* Vol. 128. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2012.

[26] J.M. Landsberg and G. Ottaviani. "Equations for secant varieties of Veronese and other varieties". *Annali di Matematica Pura ed Applicata* 192.4 (2013), pp. 569–606.

[27] J.D. Lipson. *Elements of algebra and algebraic computing.* Addison Wesley Publishing Company, 1981.

[28] Q. Liu. *Algebraic geometry and arithmetic curves.* Vol. 6. Oxford University Press on Demand, 2002.

[29] R. Mirwald. *Rang von Matrizenpaaren über $\mathbb{Z}/2\mathbb{Z}$ und Komplexität von Paaren Boolescher quadratischer Formen.* Johann Wolfgang Goethe Universität, 1991.

[30] L. Oeding and G. Ottaviani. "Eigenvectors of tensors and algorithms for Waring decomposition". *J. Symbolic Comput.* 54 (2013), pp. 9–35. ISSN: 0747-7171.

[31] P.G. Parfenov. "Orbits and their closures in the spaces $\mathbb{C}^{k_1} \otimes \ldots \otimes \mathbb{C}^{k_r}$". *Sbornik: Mathematics* 192.1 (2001), pp. 89–112.

[32] B.N. Parlett. "Symmetric matrix pencils". In: *Proceedings of the International Symposium on Computational Mathematics (Matsuyama, 1990).* Vol. 38. 1-3. 1991, pp. 373–385.

[33] C. Segre. *Studio sulle quadriche in uno spazio lineare ad un numero qualunque di dimensioni.* Memorie della Reale Accademia delle Scienze di Torino, serie II, tomo XXXVI, pp. 3-86, 1883.

[34] C. Segre. *Sulla teoria e sulla classificazione delle omografie in uno spazio lineare ad un numero qualunque di dimensioni.* Atti della R. Accademia dei Lincei, Memorie, serie III, vol. XIX, pp. 127-148, 1884.

[35] J.P. Serre. "Géométrie algébrique et géométrie analytique". *Ann. Inst. Fourier (Grenoble)* 6 (1955/56), pp. 1–42. ISSN: 0373-0956.

[36] K. Weierstrass. *Zur Theorie der bilinearen und quadratischen Formen.* Königl. Akademie der Wissenschaften zu Berlin, 1868.