Combinatorial Aspects of the Theory of Canonical Forms

by

Jozsef Losonczy, Jr.

B.A., Summa Cum Laude, New York University
(1989)

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Mathematics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1996
[June 1996]

© Massachusetts Institute of Technology 1996. All rights reserved.

Author . . . .
of Mathematics
January 11, 1996

Certified by. . . . .
Gian-Carlo Rota
Professor of Applied Mathematics and Philosophy
Thesis Supervisor

Accepted by . . . .
David A. Vogan, Jr.
Chairman, Departmental Committee on Graduate Students

SCIENCE
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY
[redacted] JUL 0 8 1996

Combinatorial Aspects of the Theory of Canonical Forms

by

Jozsef Losonczy, Jr.

B.A., Summa Cum Laude, New York University

(1989)

Submitted to the Department of Mathematics
on January 11, 1996, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Mathematics

Abstract

A geometric model for a class of bipartite graphs is introduced, and a type of perfect matching, called an acyclic matching, is defined and through geometric reasoning shown to exist for a subset of the bipartite graphs discussed. These acyclic matchings imply a nonvanishing determinant for a class of weighted biadjacency matrices.

This matching theory is applied to address a question raised by E. K. Wakeford in 1916, on the possible sets of monomials which can be removed from a generic homogeneous polynomial through linear changes in its variables.

The notion of essential rank for the p-th graded piece of the exterior algebra is given a geometric interpretation. It is shown that essential rank gives information about the Plücker embedding of the Grassmannian $G(p, V)$ in projective space over $\Lambda^p(V)$. The Lottery problem is then discussed, and its relationship to the essential rank of $\Lambda^p(V)$ is explained.

Thesis Supervisor: Gian-Carlo Rota
Title: Professor of Applied Mathematics and Philosophy

Acknowledgments

I am grateful to my thesis advisor, Gian-Carlo Rota, for introducing me to the theory of apolarity and canonical forms, but even more for offering his wise counsel on a very wide range of issues. His constant encouragement was a crucial ingredient in making this enterprise possible.

A portion of this thesis is joint work with C. Kenneth Fan and has been submitted for publication under the title, *Matchings and Canonical Forms for Symmetric Tensors.* My collaboration with Ken has been one of the most stimulating and satisfying experiences of my mathematical life. I have also enjoyed numerous helpful mathematical discussions with Richard Ehrenborg, Daniel Kleitman and Jim Propp while working on this thesis.

My warmest appreciation goes to my family–my sister Michele, my brother Steven, my father, and especially to my mother, Marianne, for her support and encouragement always. I am singularly indebted to James Kelly for his close and constant friendship since childhood.

Among those whose advice and friendship I have benefitted from during my years at MIT I wish to mention first my friend and roommate Daniel Zanger. I also want to thank Kirsten Bremke, Ludwig Chincarini, Dean Chou, Niev Duffy, David Ferguson, Bruce Fischer, Stefania Grippo, Edith Mooers, Lance Peters, Joshua Sher, Marina Tylo, Craig Wheeler, Robert West and Siman Wong.

Contents

Introduction

This thesis represents the author's attempt to answer some of the combinatorial questions which naturally arise in the theory of canonical forms. The canonical forms of which we speak are generic expressions for the representation of symmetric or skew-symmetric tensors over a finite-dimensional vector space. The theory of canonical forms itself was actively pursued near the end of the nineteenth century through the early part of the twentieth, chiefly by British and German algebraists, and mainly for symmetric tensors. In fact, it is a question raised by E. K. Wakeford in 1916 which motivated the first two chapters of the present work. We may interpret his question as: Which sets of monomials are removable from a generic homogeneous polynomial through a linear change in its variables? Using the notion of apolarity, which dates back to the time when E. Lasker, H. W. Richmond and Wakeford considered such questions, we easily transform the problem into the study of a certain class of weighted bipartite graphs. The matching theory which arises, due to the present author in collaboration with C. K. Fan, is geometric in flavor and seems interesting in its own right.

The first chapter contains a development of this matching theory and is presented without reference to the above question which inspired it. We briefly outline the theory as follows. Consider a finite subset B of $\mathbb{Z}^q$, whose elements we may think of as balls, and a set $D \subseteq \mathbb{Z}^q \setminus \{0\}$ satisfying $|D| = |B|$, which we think of as containing arrows or directions. Think of a pairing $(b, d) \in B \times D$ as an assignment of the initial point of the arrow d to the ball b. Is it always possible to assign the arrows to the balls in a one-to-one fashion so that for each assignment (b, d) we have $b + d \notin B$? The answer is yes, and the proof combines geometric reasoning with an application of the Marriage Theorem.

We then consider questions of existence of such one-to-one assignments (perfect matchings) with special uniqueness properties. In this vein, we define the notion of acyclic matching, prove that it exists and show how it implies a nonvanishing determinant for a class of matrices with entries from a polynomial ring. It is this property of acyclic matchings which we find useful in giving a partial answer to Wakeford's question in the following chapter.

It is in the second chapter that we discuss the concept of apolarity and show how it can be applied to the theory of canonical forms. Apolarity is useful in that it enables us to avoid working with large Jacobian-like matrices. We transform Wakeford's problem, via apolarity, into a question on the nonvanishing of the determinant of a certain weighted biadjacency matrix. Using our results from Chapter 1, we prove the following result: Let V be a q-dimensional vector space over $\mathbb{C}$. Any set B of $q(q-1)$ monomials in $S^p(V)$ of the form x^I, where each $i_k > 0$, may be removed from a generic element of $S^p(V)$ through a linear change in variables.

The theory of canonical forms for homogeneous elements of the exterior algebra $\Lambda(V)$, as developed recently by R. Ehrenborg, is described in Chapter 3 for the purpose of studying the notion of essential rank of a space $\Lambda^p(V)$ of skew-symmetric tensors. The essential rank of such a space is the minimum number n such that a generic skew-symmetric tensor can be written as a sum of n decomposables. It turns out that the essential rank gives information about the way Grassmannians $G(p, V)$ sit in projective space $\mathbb{P}(\Lambda^p(V))$ under the Plücker embedding. And it is once again the concept of apolarity, suitably generalized, which enables us to view the situation from a combinatorial perspective.

We can obtain upper bounds for the essential rank of $\Lambda^p(V)$, where $\dim(V) = q$, by considering the Lottery problem. The Lottery problem, stated combinatorially, asks us to find the smallest possible size of a collection S of p-element subsets of a q-element set T such that every p-element subset of T intersects some set in S in at least l elements. When $l = p - 1$, this minimum number is an upper bound for the essential rank of $\Lambda^p(V)$. We provide an integer programming formulation of the Lottery problem, which in principle provides a means for finding such minimum collections S. We conclude by deriving some lower and upper bounds for their sizes.

Chapter 1

Some Matching Theory

We begin by studying a matching problem in $\mathbb{Z}^q$. In the first section we introduce the bipartite graphs of interest to us and prove that they always admit perfect matchings. In Section 2, we study a subclass of perfect matchings which we call acyclic matchings. In Section 3, we discuss certain weighted biadjacency matrices arising from our bipartite graphs.

1.1 Perfect matchings

Let B be a finite subset of $\mathbb{Z}^q$ and let D be a subset of $\mathbb{Z}^q \setminus \{0\}$ satisfying $|D| = |B|$. We associate a bipartite graph $G = (N(G), E(G))$ to the pair of sets B and D, as follows. Since B and D are possibly nondisjoint we associate to each $b \in B$ the symbol x_b and to each $d \in D$ the symbol y_d, and then we set $X = \{x_b \mid b \in B\}$ and $Y = \{y_d \mid d \in D\}$. The nodes of G are given by the bipartition $N(G) = X \cup Y$, and there is an edge $e(x_b, y_d) \in E(G)$ joining $x_b \in X$ to $y_d \in Y$ if and only if $b + d \notin B$. A *matching* in an arbitrary graph is a collection of edges, of which no two are incident with a common node. A *perfect matching* is a matching which covers all the nodes. For the bipartite graph G associated to the pair of sets B and D we can, and will, identify perfect matchings with bijections $f : B \longrightarrow D$ satisfying $b + f(b) \notin B$ for all

$b \in B$. Our first objective is to prove that such bipartite graphs necessarily admit a perfect matching. To accomplish this, we introduce some geometry.

1.1.1 Geometric notions

We fix a positive definite, nondegenerate bilinear form $(\cdot,\cdot) : \mathbb{Z}^q \times \mathbb{Z}^q \longrightarrow \mathbb{Z}$.

Definition 1.1.1 Let $v \in \mathbb{Z}^q$ and let $w \in \mathbb{Z}^q \setminus \{0\}$. The *closed half-space* $H(v,w)$ *defined by* v *and* w is given by

$$H(v,w) = \{x \in \mathbb{Z}^q \mid (x-v,w) \leq 0\}.$$

Definition 1.1.2 Let S be a finite subset of $\mathbb{Z}^q$ and let $v \in \mathbb{Z}^q \setminus \{0\}$. The *wall defined by* S *and* v is given by

$$W(S,v) = \{x \in S \mid S \subseteq H(x,v)\}.$$

Observe that if $S \neq \emptyset$, then $W(S,v) \neq \emptyset$, by the finiteness of S.

Definition 1.1.3 Let B be a finite subset of $\mathbb{Z}^q$ and let D be a subset of $\mathbb{Z}^q \setminus \{0\}$ satisfying $|D| = |B|$. Let $S \subseteq B$ and $d \in D$.

1. We say that S *accepts* d if there exists a $b \in S$ such that $b + d \notin B$;

2. We say that S *hyperplane-accepts* d if there exists a $b \in S$ such that $b + d \notin B$ and $S \subseteq H(b,d)$.

If $S = \{b\}$, then we also say that b accepts d or that b hyperplane-accepts d.

Let B, D, S and d be as in Definition 1.1.3. Notice that the wall $W(S,d)$ may contain elements b such that $b + d \in B$.

Proposition 1.1.1 *Let B be a finite subset of $\mathbb{Z}^q$ and let D be a subset of $\mathbb{Z}^q \setminus \{0\}$ satisfying $|D| = |B|$. Given $S \subseteq B$, let $A \subseteq D$ denote the set of all $d \in D$ which are hyperplane-accepted by S. Then $|A| \geq |S|$.*

Proof: If $S = \emptyset$, then the result is trivial. Assume $S \neq \emptyset$. For each $d \in D \setminus A$ choose an element $b_d \in W(S,d)$. Observe that $b_d + d \in B \setminus S$. We claim that if $b_d + d = b_c + c$ then $d = c$. Let $p = b_d + d$. We have that b_d is the unique closest point to p in S, since $S \subseteq H(b_d, d)$. Likewise, b_c is the unique closest point to p in S, since $S \subseteq H(b_c, c)$. Hence, $b_d = b_c$, so that $d = c$.

From the above, we obtain $|B \setminus S| \geq |D \setminus A|$, which implies that $|A| \geq |S|$. This completes the proof.

1.1.2 Existence of perfect matchings

The following fundamental theorem is due to Frobenius [8]. For a discussion of its relationship to other results in matching theory, as well as a proof, see Lovász and Plummer [14].

Theorem 1.1.1 (The Marriage Theorem). *A bipartite graph $G = (N(G), E(G))$ with bipartition $N(G) = X \cup Y$ admits a perfect matching if and only if $|X| = |Y|$ and for each $T \subseteq X$, we have $|T| \leq |\Gamma(T)|$, where $\Gamma(T)$ equals the set of elements of Y which are joined to some member of T.*

Corollary 1.1.1 *Let B be a finite subset of $\mathbb{Z}^q$ and let D be a subset of $\mathbb{Z}^q \setminus \{0\}$ satisfying $|D| = |B|$. Then there exists a perfect matching $f : B \longrightarrow D$.*

Proof: Let G be the bipartite graph associated with the pair of sets B and D. Let $T \subseteq X$ (recall the definition of $N(G) = X \cup Y$). The subset T of X corresponds to a subset S of B, which, according to Proposition 1.1.1, hyperplane-accepts the elements of a subset A of D of size $|A| \geq |S|$. Let U be the subset of Y corresponding to A. Since $\Gamma(T) \supseteq U$, we have $|\Gamma(T)| \geq |U| = |A| \geq |S| = |T|$. By the Marriage Theorem, the graph G admits a perfect matching, which we can identify with a bijection $f : B \longrightarrow D$ satisfying $b + f(b) \notin B$ for all $b \in B$, as required.

1.2 Acyclic matchings

Throughout this section, B will denote a finite subset of $\mathbb{Z}^q$ and D will denote a subset of $\mathbb{Z}^q \setminus \{0\}$ such that $|D| = |B|$.

Definition 1.2.1 For any perfect matching $f : B \longrightarrow D$ we define its *multiplicity map* $m_f : \mathbb{Z}^q \longrightarrow \mathbb{Z}$ by $m_f(v) = \#\{b \in B \mid b + f(b) = v\}$.

Definition 1.2.2 An *acyclic matching* is a perfect matching f with the property that for any perfect matching g satisfying $m_f = m_g$, we have $f = g$.

Definition 1.2.3 A *hyperplane chain C of length l* is a sequence of pairs $(b_i, d_i) \in B \times D$ for $i = 1, \ldots, l$ such that the following four properties hold:

1. The b_i are all distinct;
2. The d_i are all distinct;
3. Each b_i accepts d_i;
4. $B \setminus \{b_1, \ldots, b_{i-1}\} \subseteq H(b_i, d_i)$.

Given a hyperplane chain $C = \{(b_i, d_i)\}$ of length l, we define $B(C) = \{b_1, \ldots, b_l\}$ and $D(C) = \{d_1, \ldots, d_l\}$. Note that by definition, $|B(C)| = |D(C)| = l$. We also define $B^i(C) = b_i$ and $D^i(C) = d_i$.

Lemma 1.2.1 *Let $C = \{(b_i, d_i)\}$ be a hyperplane chain of length l. Assume that we have a vector $d \in D \setminus D(C)$, integers $0 \leq j < m \leq l$, and a hyperplane chain C^* of length $j + 1$ satisfying*

$$\begin{aligned} B(C^*) &= \{b_1, \ldots, b_j, b_m\}, \\ D(C^*) &= \{d_1, \ldots, d_j, d\}. \end{aligned}$$

Define the sequence $C' = \{(B^i(C'), D^i(C'))\}$ of length l in two steps, as follows.

First, set $B^i(C') = B^i(C^*)$ and $D^i(C') = D^i(C^*)$ for $i = 1, \ldots, j+1$. Second, let the sequence $B^{j+2}(C'), \ldots, B^l(C')$ equal the sequence $b_{j+1}, \ldots, \hat{b}_m, \ldots, b_l$, and similarly let the sequence $D^{j+2}(C'), \ldots, D^l(C')$ equal $d_{j+1}, \ldots, \hat{d}_m, \ldots, d_l$.

Then the sequence C' is a hyperplane chain of length l.

Proof: The first three parts of the definition of hyperplane chain are obviously satisfied by C'. To see that the fourth part holds as well, consider the set

$$B \setminus \{B^1(C'), \ldots, B^{i-1}(C')\}.$$

For $i \leq j+1$ we have $B \setminus \{B^1(C'), \ldots, B^{i-1}(C')\} = B \setminus \{B^1(C^*), \ldots, B^{i-1}(C^*)\} \subseteq H(B^i(C^*), D^i(C^*)) = H(B^i(C'), D^i(C'))$. If $i > j+1$ and $i-1 < m$ we have $B \setminus \{B^1(C'), \ldots, B^{i-1}(C')\} \subseteq B \setminus \{b_1, \ldots, b_{i-2}\} \subseteq H(b_{i-1}, d_{i-1}) = H(B^i(C'), D^i(C'))$. Finally, if $i > j+1$ and $i-1 \geq m$, then $B \setminus \{B^1(C'), \ldots, B^{i-1}(C')\} = B \setminus \{b_1, \ldots, b_{i-1}\} \subseteq H(b_i, d_i) = H(B^i(C'), D^i(C'))$, as desired.

Proposition 1.2.1 *Let $T \subseteq D$. There exists a hyperplane chain C of length $|T|$ such that $D(C) = T$.*

Proof: Let P_l denote the following statement:

> Given a hyperplane chain C of length l and any $d \in D \setminus D(C)$, there exists a hyperplane chain C' of length $l+1$ such that $D(C') = D(C) \cup \{d\}$ and $B(C) \subseteq B(C')$.

We prove P_l for $0 \leq l < |D|$ by induction on l. Once this is done, our result follows by successive applications of P_l, starting at $l = 0$.

We start with the base case. Given $d \in D$, choose any $b \in W(B, d)$. Put $C' = \{(b, d)\}$. This proves P_0. Let l satisfy $0 < l < |D|$ and assume P_s is true for all $s = 0, \ldots, l-1$. We show that P_l is true. Let $U = B \setminus B(C)$. We have that U hyperplane-accepts at least $|B| - l$ elements of D by Proposition 1.1.1. Hence, U must hyperplane-accept at least one element of $D(C) \cup \{d\}$.

We now construct by induction a sequence $(C_0, d_0), \ldots, (C_M, d_M)$ of hyperplane chains C_i of length l and vectors d_i such that for each i,

1. $B(C_i) = B(C)$;

2. $D(C_i) \cup \{d_i\} = D(C) \cup \{d\}$;

3. $\{d_0, \ldots, d_{i-1}\} = \{D^1(C_i), \ldots, D^i(C_i)\}$;

4. U does not hyperplane-accept any element of $\{D^1(C_i), \ldots, D^i(C_i)\}$.

Start by setting $C_0 = C$ and $d_0 = d$. Assume that C_j and d_j have been defined for an integer j, $0 \le j < l$. If U hyperplane-accepts d_j, then set $M = j$ and terminate the construction. If instead U does not hyperplane-accept d_j, we proceed to construct a hyperplane chain C_{j+1} of length l, and d_{j+1}, as follows. Consider the hyperplane chain of length j consisting of the pairs $\{(B^i(C_j), D^i(C_j))\}_{i=1,\ldots,j}$. With the vector d_j in hand, we can apply the induction hypothesis P_j to obtain a hyperplane chain C^* of length $j+1$ such that $D(C^*) = \{D^1(C_j), \ldots, D^j(C_j), d_j\}$ and $B(C^*) \supseteq \{B^1(C_j), \ldots, B^j(C_j)\}$. Moreover, since U does not hyperplane-accept d_j nor any $D^i(C_j)$ for $1 \le i \le j$, we have $B(C^*) \subseteq B(C)$. Thus, $B(C^*) = \{B^1(C_j), \ldots, B^j(C_j)\} \cup \{B^m(C_j)\}$ for some m, $j < m \le l$.

Define C_{j+1} as follows. For $i = 1, \ldots, j+1$, let $B^i(C_{j+1}) = B^i(C^*)$ and $D^i(C_{j+1}) = D^i(C^*)$. Let the sequence $B^{j+2}(C_{j+1}), \ldots, B^l(C_{j+1})$ equal $B^{j+1}(C_j), \ldots, B^l(C_j)$ with the term $B^m(C_j)$ removed. Similarly, let the sequence $D^{j+2}(C_{j+1}), \ldots, D^l(C_{j+1})$ equal $D^{j+1}(C_j), \ldots, D^l(C_j)$ with the term $D^m(C_j)$ removed. By Lemma 1.2.1, C_{j+1} is a hyperplane chain of length l. Finally, let $d_{j+1} = D^m(C_j)$. It is easy to see that the sequence $(C_0, d_0), \ldots, (C_{j+1}, d_{j+1})$ satisfies the four properties listed above.

Note that this construction must terminate because the d_i constitute a set of distinct elements from $D(C) \cup \{d\}$ and we have observed that U must hyperplane-accept at least one element in $D(C) \cup \{d\}$. Let M denote the largest integer for which C_M is defined (possibly $M = 0$). Then U hyperplane-accepts d_M. In other words, there exists a $b \in U$ which accepts d_M, and $U \subseteq H(b, d_M)$. We now define C'

by letting $B^i(C') = B^i(C_M)$ and $D^i(C') = D^i(C_M)$ for $1 \leq i \leq l$ and then setting $B^{l+1}(C') = b$ and $D^{l+1}(C') = d_M$. This construction proves that P_l is true. By the principle of mathematical induction, the proof is complete.

We now state and prove the main result of this chapter, which we shall call the *acyclic matching theorem.* We explain its significance for a particular class of weighted biadjacency matrices in the next section.

Theorem 1.2.1 *Assume that the elements of D are all of equal length. Then there exists an acyclic matching $f : B \longrightarrow D$.*

Proof: By Proposition 1.2.1, there exists a hyperplane chain $C = \{(b_i, d_i)\}$ of length $|B|$. Define $f : B \longrightarrow D$ by $f(b_i) = d_i$ for all $i = 1, \ldots, |B|$. Let g be any perfect matching satisfying $m_g = m_f$. We prove by induction on l that $f(b_l) = g(b_l)$ for all l.

For $l = 1$ we have $B \subseteq H(b_1, d_1)$, and b_1 is the unique closest point in $H(b_1, d_1)$ to $b_1 + d_1$. Since the elements of D all have equal length, we must have $m_f(b_1 + d_1) = 1$. Since $m_g(b_1 + d_1)$ must also equal 1, we obtain $g(b_1) = d_1$.

Let $l > 1$ and assume that $f(b_i) = g(b_i)$ for all $i < l$. We show that $f(b_l) = g(b_l)$. Since b_l is the unique closest element of $B \setminus \{b_1, \ldots, b_{l-1}\} \subseteq H(b_l, d_l)$ to $b_l + d_l$, we have $m_f(b_l + d_l) = 1 + \#\{i \mid 1 \leq i \leq l-1 \text{ and } b_i + d_i = b_l + d_l\} = m_g(b_l + d_l)$. Therefore, we have $g(b_l) = d_l$. We conclude that $f = g$, as desired.

Remark 1.2.1 It has recently been proved that the conclusion of Theorem 1.2.1 is true without the equal length assumption on D. See [1]. The strategy employed is necessarily different from ours, since a perfect matching f coming from a hyperplane chain of length $|B|$ is not necessarily acyclic. For our purposes in the sequel, however, the result stated and proved above is adequate.

1.3 Acyclicity and determinants

We continue to let B denote a finite subset of $\mathbb{Z}^q$ and D a subset of $\mathbb{Z}^q \setminus \{0\}$ such that $|D| = |B|$. Recall from the first section of this chapter the definition of the bipartite graph $G = (N(G), E(G))$ associated to the sets B and D.

To each $v \in \mathbb{Z}^q$ associate a symbol h_v. Let $H = \{h_v \mid v \in \mathbb{Z}^q\}$, and form the polynomial ring $\mathbb{C}[H]$. Associate a weight map $W_G : E(G) \longrightarrow \mathbb{C}[H]$ to the bipartite graph $G = (N(G), E(G))$, as follows:

$$W_G(e(x_b, y_d)) = h_{b+d}.$$

We may now consider the weighted biadjacency matrix $M(G)$ of the weighted bipartite graph $G = (N(G), E(G), W_G)$, indexed by $X \times Y$. Its entries are

$$M(G)_{x_b,y_d} = \begin{cases} W_G(e(x_b, y_d)) & \text{if } e(x_b, y_b) \text{ is an edge of } G; \\ 0 & \text{otherwise.} \end{cases}$$

Notice that the determinant of $M(G)$ is well-defined up to sign.

Proposition 1.3.1 *If there exists an acyclic matching $f : B \longrightarrow D$, then the determinant of $M(G)$ is nonzero.*

Proof: Let $f : B \longrightarrow D$ be an acyclic matching. The determinant of $M(G)$ equals, up to sign,

$$\sum_{\sigma} \left((-1)^{\sigma} \cdot \prod_{b \in B} M(G)_{b,\sigma f(b)} \right),$$

where σ ranges over all permutations of D. There is a one-to-one correspondence between the perfect matchings of G and the nonzero summands in the above expansion. The summand corresponding to $\sigma = 1$ is

$$\prod_{b \in B} M(G)_{b,f(b)} = \prod_{b \in B} W_G(e(x_b, y_{f(b)})) = \prod_{b \in B} h_{b+f(b)} = \prod_{v} h_v^{m_f(v)},$$

where in the last product v ranges over all vectors in $\mathbb{Z}^q$. Since f is an acyclic matching, this term is not cancelled in the above expansion. Therefore, the determinant of $M(G)$ is nonzero.

Chapter 2

Symmetric Tensors and Removability of Monomials

Let V be a q-dimensional vector space over the complex field $\mathbb{C}$ and consider the symmetric algebra over V:

$$S(V) = \bigoplus_{p \geq 0} S^p(V).$$

We are interested in finding canonical expressions for the elements of $S^p(V)$, the symmetric tensors of degree p. The canonical expressions we seek are of a special type, which we now describe. Call a product of p vectors from V (not necessarily distinct) a *monomial of degree p*. Given q elements $X_1, \ldots, X_q$ of V, we construct monomials $X_1^{i_1} \cdots X_q^{i_q}$ of degree p from the X_i according to the multi-indices $I = (i_1, \ldots, i_q)$ of nonnegative integers satisfying $i_1 + \cdots + i_q = p$. Let $T(q,p)$ denote the set of all such multi-indices. We ask: For which subsets $B \subseteq T(q,p)$ is it true that a generic element of $S^p(V)$ may be written as $\sum_{I \notin B} c_I \cdot X_1^{i_1} \cdots X_q^{i_q}$, for some $c_I \in \mathbb{C}$ and $X_1, \ldots, X_q \in V$? When a generic element of $S^p(V)$ can be written in this way we think of the monomials constructed from $X_1, \ldots, X_q$ corresponding to multi-indices $I \in B$ as having been *removed.* We also say that the set B itself is removable.

The above question was asked by E. K. Wakeford in 1916, in his dissertation on the possible canonical forms for homogeneous polynomials; see [20]. His question is largely answered by the main theorem of this chapter (Theorem 2.2.1), which states that any set of $q(q-1)$ monomials of the form $X_1^{i_1} \cdots X_q^{i_q}$, where each $i_j > 0$, is removable. Furthermore, any removable set contains at most $q(q-1)$ elements.

This chapter is divided into two sections. The first section provides the background material for the second. We start with a description of *apolarity*, a notion which dates back to the work of Clebsch, Lasker and Wakeford. The apolarity concept, including its relation to the theory of canonical forms, has been revisited and reworked by Richard Ehrenborg and Gian-Carlo Rota, and our treatment of the subject reflects their exposition in [6].

In Section 2, we show how apolarity theory transforms Wakeford's question into a question about certain weighted biadjacency matrices arising from a class of bipartite graphs. To each subset B of $T(q,p)$ satisfying $|B| = q(q-1)$ we associate such a matrix, and if this matrix has nonzero determinant, then the monomials corresponding to B are removable. The analysis of these matrices uses our results on matchings in $\mathbb{Z}^q$ from the previous chapter.

2.1 Apolarity and canonical forms

Let V be a q-dimensional vector space over $\mathbb{C}$. Since the space $S^p(V)$ of degree p symmetric tensors is finite dimensional, we may endow it with the Euclidean topology. We say that a property P holds *generically* in $S^p(V)$ if P holds for every element of some dense subset of $S^p(V)$. The canonical forms we arrive at will be generic canonical forms in the sense that each represents some dense subset of $S^p(V)$.

2.1.1 The apolar bilinear form

In this section, we introduce the apolar bilinear form. This bilinear form plays an important role in our development of the theory of canonical forms for symmetric tensors. Fix a basis $x_1, \ldots, x_q$ of V, and let $u_1, \ldots, u_q$ be the corresponding dual basis of V^*. Let $T(q,p)$ denote the set of all q-tuples $I = (i_1, \ldots, i_q)$ of nonnegative integers satisfying $i_1 + \cdots + i_q = p$. We define $x^I = x_1^{i_1} \cdots x_q^{i_q}$ and $I! = i_1! \cdots i_q!$. We define the *apolar bilinear form*

$$\langle \cdot, \cdot \rangle : S^p(V^*) \times S^p(V) \longrightarrow \mathbb{C},$$

by setting $\langle u^I, x^J \rangle = I! \cdot \delta_{I,J}$ and extending in a bilinear fashion. It is easy to see that this bilinear form is nondegenerate. It has another property which is important for our purposes; namely, it is invariant under the natural action of the general linear group $GL(V)$ on $S^p(V^*) \times S^p(V)$.

For completeness, we develop this property of the apolar form in some detail. We start with the natural linear actions of $GL(V)$ on the spaces V and V^*. If $A \in GL(V)$, $v \in V$ and $f \in V^*$, then these actions are given by $A \cdot v = A(v)$ and $A \cdot f = f \circ A^{-1}$. The latter is known as the contragredient action of $GL(V)$ on V^*. If $GL(V)$ acts linearly on a vector space W, we shall indicate this fact by the notation $GL(V) : W$. Thus, we have $GL(V) : V$ and $GL(V) : V^*$. Now, any linear action of $GL(V)$ on a vector space W induces a linear action of $GL(V)$ on the p-fold tensor product $W^{\otimes p}$. The resulting action is the unique linear action satisfying $A \cdot (w_1 \otimes \cdots \otimes w_p) = (A \cdot w_1) \otimes \cdots \otimes (A \cdot w_p)$ for all $A \in GL(V)$ and $w_1, \ldots, w_p \in W$. In this way, we obtain linear actions $GL(V) : V^{\otimes p}$ and $GL(V) : (V^*)^{\otimes p}$. Let U be the subspace of $V^{\otimes p}$ spanned by all elements of the form $v_1 \otimes \cdots \otimes v_p - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(p)}$, where $v_1, \ldots, v_p \in V$ and σ is a permutation of $\{1, \ldots, p\}$. Since U is an invariant subspace of $V^{\otimes p}$ under the action of $GL(V)$, we obtain a linear action of $GL(V)$ on the quotient space $V^{\otimes p}/U = S^p(V)$. Similarly, we obtain $GL(V) : S^p(V^*)$. Finally, this pair of linear actions induces a linear action $GL(V) : (S^p(V^*) \times S^p(V))$ in the obvious way.

The next proposition says that if $A \in GL(V)$, then the adjoint of the automorphism of $S^p(V)$ induced by A is the automorphism of $S^p(V^*)$ induced by A^{-1} (with respect to the apolar form).

Proposition 2.1.1 *For any $A \in GL(V)$, $g \in S^p(V^*)$ and $f \in S^p(V)$ we have*

$$\langle A^{-1} \cdot g, f \rangle \;=\; \langle g, A \cdot f \rangle.$$

Proof: Fix a basis $x_1, \ldots, x_q$ of V, and let $u_1, \ldots, u_q$ be the corresponding dual basis of V^*. It suffices to prove that $\langle A^{-1} \cdot u^I, x^J \rangle = \langle u^I, A \cdot x^J \rangle$. Write $A \cdot x_i = \sum_{j=1}^q A_{i,j} \cdot x_j$, where the $A_{i,j} \in \mathbb{C}$ are uniquely determined. Observe that $A^{-1} \cdot u_i = \sum_{j=1}^q A_{j,i} \cdot u_j$. We have

$$\begin{aligned}
\langle A^{-1} \cdot u^I, x^J \rangle &= \langle (A^{-1} \cdot u_1)^{i_1} \cdots (A^{-1} \cdot u_q)^{i_q}, x^J \rangle \\
&= \left\langle \left(\sum_{k_1=1}^q A_{k_1,1} \cdot u_{k_1} \right)^{i_1} \cdots \left(\sum_{k_q=1}^q A_{k_q,q} \cdot u_{k_q} \right)^{i_q}, x^J \right\rangle \\
&= J! \cdot \sum_M \prod_{l=1}^q \binom{i_l}{m_1^l, \ldots, m_q^l} \prod_{k=1}^q A_{k,l}^{m_k^l} \\
&= I! \cdot J! \cdot \sum_M \prod_{1 \le k,l \le q} \frac{A_{k,l}^{m_k^l}}{m_k^l!},
\end{aligned}$$

where in the last two sums M ranges over all q by q matrices $\{m_k^l\}$ with entries in $\mathbb{N}$ satisfying $\sum_k m_k^l = i_l$ and $\sum_l m_k^l = j_k$. On the other hand, one can verify that $\langle u^I, A \cdot x^J \rangle$ is equal to the same expression. This completes the proof.

Corollary 2.1.1 *The apolar form $\langle \cdot, \cdot \rangle$ is invariant under the action of $GL(V)$ on $S^p(V^*) \times S^p(V)$.*

Proof: Let $A \in GL(V)$, $g \in S^p(V^*)$ and $f \in S^p(V)$. By Proposition 2.1.1, we have

$$\langle A \cdot g, A \cdot f \rangle = \langle g, A^{-1} \cdot (A \cdot f) \rangle = \langle g, (A^{-1}A) \cdot f \rangle = \langle g, f \rangle,$$

as desired.

Definition 2.1.1 Let $g \in S^p(V^*)$ and $f \in S^r(V)$. We say that g is *apolar* to f if at least one of the following two conditions hold:

1. $r \le p$ and $\langle g, f \cdot h \rangle = 0$ for all $h \in S^{p-r}(V)$;

2. $r \ge p$ and $\langle g \cdot h, f \rangle = 0$ for all $h \in S^{r-p}(V^*)$.

2.1.2 The apolarity theorem

Fix $p \ge 0$. Let $d_1, \ldots, d_r$ be nonnegative integers and let Ω be a finite subset of $\mathbb{N}^r$. For each r-tuple $I \in \Omega$ let t_I be a homogeneous symmetric tensor over V. Assume that for all $I \in \Omega$ we have

$$i_1 d_1 + i_2 d_2 + \cdots + i_r d_r + \deg(t_I) = p.$$

The above data determine a proposed canonical form; that is, we may propose that a generic element of $S^p(V)$ can be written as

$$\sum_{I \in \Omega} t_I \cdot s_1^{i_1} \cdots s_r^{i_r},$$

where each $s_j \in S^{d_j}(V)$.

The following theorem, due to Ehrenborg and Rota [6], characterizes those proposed canonical forms which are, in fact, canonical. For completeness, we supply a proof here.

Theorem 2.1.1 *A generic element of $S^p(V)$ can be written as*

$$F = \sum_{I \in \Omega} t_I \cdot s_1^{i_1} \cdots s_r^{i_r}$$

for some $s_j \in S^{d_j}(V)$ if and only if there exist $s'_j \in S^{d_j}(V)$ such that the only $g \in S^p(V^*)$ apolar to all the $\left(\frac{\partial F}{\partial s_j}\right)_{s_j = s'_j}$ is zero.

Proof: Assume that F is canonical. Fix a basis $x_1, \ldots, x_q$ of V. Expanding each variable symmetric tensor s_j in terms of the monomials x^J, that is, writing $s_j = \sum_{J \in T(q,d_j)} \alpha_{j,J} \cdot x^J$, we obtain variable coefficients $\alpha_{j,J}$, which we call parameters. Let

P be the set of all parameters coming from F. Notice that since F is canonical we must have

$$\dim S^p(V) = \binom{q+p-1}{p} \leq |P| = \binom{q+d_1-1}{d_1} + \cdots + \binom{q+d_r-1}{d_r}.$$

Expand the variable symmetric tensor $\sum_{I\in\Omega} t_I \cdot s_1^{i_1} \cdots s_r^{i_r}$ in terms of the monomials x^J, where J ranges over all elements of $T(q,p)$. We obtain

$$\sum_{I\in\Omega} t_I \cdot s_1^{i_1} \cdots s_r^{i_r} = \sum_{J\in T(q,p)} \psi_J \cdot x^J,$$

where the ψ_J are elements of the polynomial ring $\mathbb{C}[P]$. The statement that F is canonical is equivalent to the statement that the polynomial map

$$\Psi : \mathbb{C}^P \longrightarrow \mathbb{C}^{T(q,p)},$$

where $\Psi = \{\psi_J\}_{J\in T(q,p)}$, has dense range. This in turn is equivalent to the statement that the polynomials ψ_J are algebraically independent. The condition of algebraic independence of the set $\{\psi_J | J \in T(q,p)\}$ is equivalent to the condition that the Jacobian matrix

$$\left\{\frac{\partial \psi_J}{\partial \alpha}\right\}_{(J,\alpha)\in T(q,p)\times P}$$

has full rank over the quotient field $\mathbb{C}(P)$. Therefore, we can choose values $\alpha' \in \mathbb{C}$ for the parameters $\alpha \in P$ so that the resulting evaluated matrix $\left\{\frac{\partial \psi_J}{\partial \alpha}(\alpha')\right\}$ has full rank over $\mathbb{C}$. Notice that choosing values for the parameters is tantamount to selecting symmetric tensors s_j' of degree d_j for $j = 1, \ldots, r$.

Since the matrix $\left\{\frac{\partial \psi_J}{\partial \alpha}(\alpha')\right\}$ has full rank, its columns (indexed by the $\alpha \in P$) span the space $\mathbb{C}^{T(q,p)}$. This space is naturally isomorphic to $S^p(V)$. Under this isomorphism, the column corresponding to α, $\left(\frac{\partial \psi_J}{\partial \alpha}(\alpha')\right)_{J\in T(q,p)}$, is sent to the symmetric tensor

$$\sum_{J \in T(q,p)} \frac{\partial \psi_J}{\partial \alpha}(\alpha') \cdot x^J = \frac{\partial F}{\partial \alpha}(\alpha').$$

Thus, the symmetric tensors $\frac{\partial F}{\partial \alpha}(\alpha')$, where α ranges over P, span $S^p(V)$. Each parameter α occurs in precisely one of the variable symmetric tensors $s_1, \ldots, s_r$. Suppose α occurs in s_j. Since $\frac{\partial s_j}{\partial \alpha} = x^I$ for some $I \in T(q, d_j)$, we have by the chain rule

$$\frac{\partial F}{\partial \alpha} = \frac{\partial F}{\partial s_j} \cdot \frac{\partial s_j}{\partial \alpha} = \frac{\partial F}{\partial s_j} \cdot x^I.$$

We therefore have that the symmetric tensors $\left(\frac{\partial F}{\partial s_j}\right)_{s_j = s_j'} \cdot x^I$, where I ranges over $T(q, d_j)$ and j ranges from 1 to r, span $S^p(V)$. Owing to the nondegeneracy of the apolar form, there is no nonzero $g \in S^p(V^*)$ satisfying

$$\left\langle g, \left(\frac{\partial F}{\partial s_j}\right)_{s_j = s_j'} \cdot x^I \right\rangle = 0$$

for all $I \in T(q, d_j)$ and $j = 1, \ldots, r$. But this is the same as saying that the only $g \in S^p(V^*)$ apolar to all the $\left(\frac{\partial F}{\partial s_j}\right)_{s_j = s_j'}$ is zero. Observing that our argument is valid in reverse, our proof is complete.

2.2 Removability of monomials

We say that a subset B of $T(q,p)$ is *removable* if a generic element of $S^p(V)$ can be written as

$$\sum_{I \in T(q,p) \setminus B} c_I \cdot X_1^{i_1} \cdots X_q^{i_q},$$

for some $c_I \in \mathbb{C}$ and $X_1, \ldots, X_q \in V$. In this section, we attempt to describe as completely as possible those subsets B of $T(q,p)$ which are removable. Assume that $p, q > 1$ throughout.

2.2.1 The size of removable sets

Let $e_1, \ldots, e_q \in \mathbb{Z}^q$ be defined by $e_i = (\delta_{1i}, \ldots, \delta_{qi})$.

Proposition 2.2.1 *If $B \subseteq T(q,p)$ is removable, then $|B| \leq q(q-1)$.*

Proof: Let $B \subseteq T(q,p)$ satisfy $|B| > q(q-1)$. We show that

$$F = \sum_{I \in T(q,p) \setminus B} c_I \cdot X_1^{i_1} \cdots X_q^{i_q}$$

is noncanonical. Let $c'_I \in \mathbb{C}$ and $X'_r \in V$ be arbitrarily chosen. If the X'_r are linearly dependent, we may assume, by applying an element $A \in GL(V)$ if necessary, that each X'_r lies in the span of $x_1, \ldots, x_{q-1}$. Then $g = u_q^p$ is apolar to $\frac{\partial F}{\partial c_I}(c'_I; X'_r)$, $\frac{\partial F}{\partial X_r}(c'_I; X'_r)$ for all $I \in T(q,p) \setminus B$ and $r = 1, \ldots, q$. If instead the X'_r are linearly independent, we may assume, again applying an element $A \in GL(V)$ if necessary, that each $X'_r = x_r$. Now, any $g \in S^p(V^*)$ apolar to the $\frac{\partial F}{\partial c_I}(c'_I; x_r)$ for $I \in T(q,p) \setminus B$ must have the form

$$g = \sum_{I \in B} a_I \cdot u^I.$$

We aim to show that there exists a nonzero such g apolar to the

$$\frac{\partial F}{\partial X_r}(c'_I; x_r) = \sum_{I \in T(q,p) \setminus B} i_r \cdot c'_I \cdot x^{I - e_r}$$

for all r. Observe that g is apolar to $\frac{\partial F}{\partial X_r}(c'_I; x_r)$ for all r if and only if g is apolar to $\frac{\partial F}{\partial X_r}(c'_I; x_r) \cdot x_s$ for $r \neq s$. These apolarity conditions produce $q(q-1)$ homogeneous linear equations in the unknowns a_I, of which there are $|B| > q(q-1)$. Hence, there is a nonzero solution, i.e., a nonzero g apolar to $\frac{\partial F}{\partial c_I}(c'_I; x_r)$, $\frac{\partial F}{\partial X_r}(c'_I; x_r)$ for all $I \in T(q,p) \setminus B$ and $r = 1, \ldots, q$. Applying Corollary 2.1.1 and Theorem 2.1.1, we obtain that F is noncanonical. Hence B is nonremovable, as desired.

2.2.2 The weighted bipartite graph

We turn our attention now to subsets B of $T(q,p)$ of size $q(q-1)$. To each such subset B we associate a weighted bipartite graph $G_B = (N(G_B), E(G_B), W_{G_B})$, as follows. Let $D = \{e_i - e_j \mid 1 \leq i,j \leq q,\ i \neq j\}$. The nodes of G_B are given by the bipartition $N(G_B) = (B, D)$, and there is an edge $e(I,J)$ joining $I \in B$ to $J \in D$ if and only if $I + J \in T(q,p) \setminus B$. To each $I \in T(q,p) \setminus B$ associate a symbol c_I, and, letting $H_B = \{c_I \mid I \in T(q,p) \setminus B\}$, form the polynomial ring $\mathbb{C}[H_B]$. The weight map $W_{G_B} : E(G_B) \longrightarrow \mathbb{C}[H_B]$ is then given by

$$W_{G_B}(e(I,J)) = (i_k + 1) \cdot I! \cdot c_{I+J},$$

where $J = e_k - e_m$. As with any bipartite graph, we may define its weighted biadjacency matrix, $M(G_B)$. We index $M(G_B)$ with the set $B \times D$, so that

$$M(G_B)_{I,J} = \begin{cases} W_{G_B}(e(I,J)) & \text{if } e(I,J) \text{ is an edge of } G_B; \\ 0 & \text{otherwise.} \end{cases}$$

We remark that the determinant of $M(G_B)$ is well-defined up to sign.

Example 2.2.1 Fix $q = 2$. We may represent the elements of $T(2,p)$ as locations along a straight line, as indicated in Figure 1.

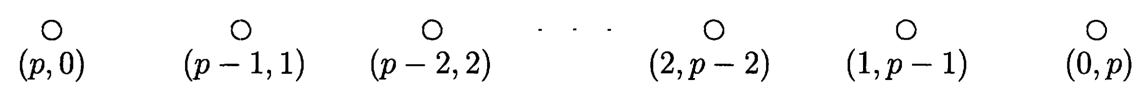

Figure 1

Let $B \subseteq T(2,p)$ satisfy $|B| = 2$. For $q = 2$, $D = \{(-1,1),(1,-1)\}$. We may indicate in our diagrammatic representation the elements of B by blackening the corresponding locations; also, we may represent $(-1,1)$ as a unit length rightward pointing arrow and $(1,-1)$ as a unit length leftward pointing arrow. Then the graph G_B has an edge joining $I \in B$ to $J \in D$ if and only if in our diagrammatic representation, when we place the initial point of the arrow J onto I, the arrowhead points to a location in our diagram which has not been blackened. The case $B = \{(3,1),(2,2)\} \subseteq T(2,4)$ is depicted in Figure 2, wherein the unique perfect matching of the bipartite graph G_B is indicated by the assignment of the two arrows to the two blackened locations.

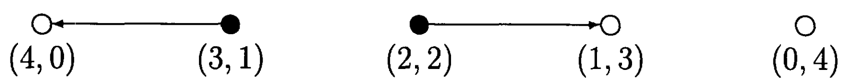

Figure 2

The bipartite graph G_B itself, together with the weights of its edges, is depicted in Figure 3.

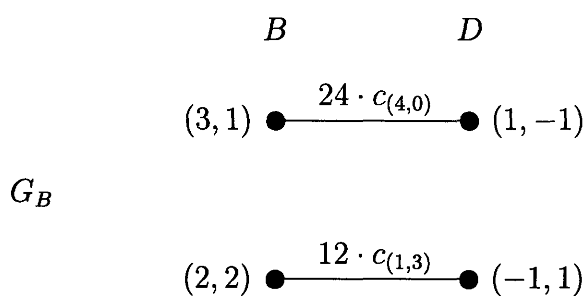

Figure 3

Example 2.2.2 Fix $q = 3$. We may represent the elements of $T(3,p)$ as locations in a triangular configuration such as the one depicting $T(3,3)$ in Figure 4.

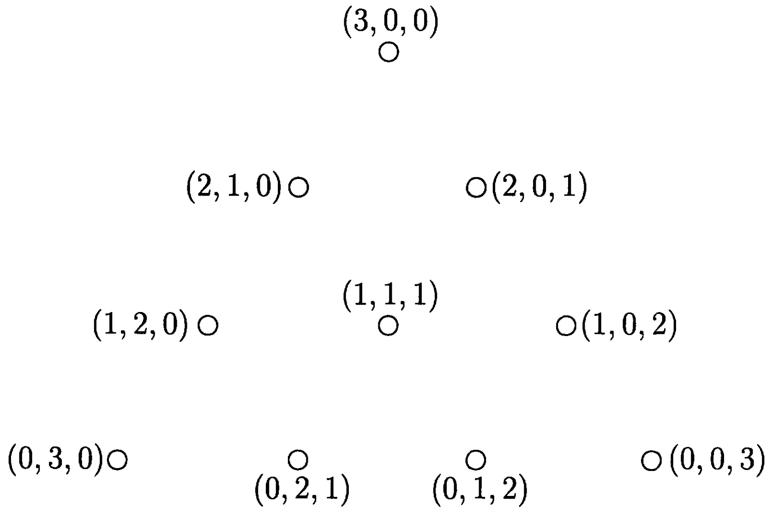

Figure 4

Let $B \subseteq T(3,p)$ satisfy $|B| = 6$. For $q = 3$, we have

$$D = \{(1,-1,0),(0,-1,1),(-1,0,1),(-1,1,0),(0,1,-1),(1,0,-1)\}.$$

We indicate in our triangular configuration the elements of B by blackening the corresponding locations. The element $(1,-1,0)$ of D is represented by a unit length arrow which points in the northeast direction. The element $(0,-1,1)$ points east, $(-1,0,1)$ points southeast, $(-1,1,0)$ points southwest, $(0,1,-1)$ points west and $(1,0,-1)$ points northwest. Just as in Example 2.2.1, the graph G_B has an edge joining $I \in B$ to $J \in D$ if and only if in our diagrammatic representation, when we place the initial point of the arrow J onto I, the arrowhead points to a location in our diagram which has not been blackened. The case

$$B = \{(2,1,0),(2,0,1),(1,2,0),(1,0,2),(0,2,1),(0,1,2)\} \subseteq T(3,3)$$

is depicted in Figure 5; a perfect matching of the bipartite graph G_B is also indicated in this figure by the assignment of each of the six elements of D to a different blackened location. There is one other perfect matching which would have all arrows pointing to the interior location.

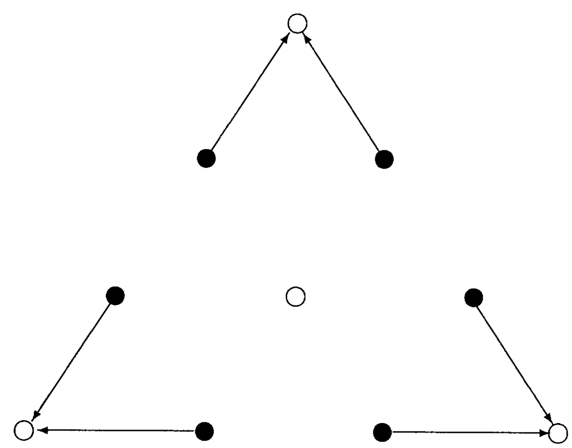

Figure 5

The relevance of the graph G_B to our discussion of removability of monomials is made explicit by the following result.

Proposition 2.2.2 *Let $B \subseteq T(q,p)$ be of size $q(q-1)$. Then B is removable if and only if the determinant of $M(G_B)$ is nonzero.*

Proof: We have the proposed canonical form

$$F(c_I; X_r) = \sum_{I \in T(q,p) \setminus B} c_I \cdot X_1^{i_1} \cdots X_q^{i_q}.$$

By Theorem 2.1.1 and a change of variables argument, F is canonical if and only if we can set each $X_r = x_r$ and choose $c_I' \in \mathbb{C}$ so that the only $g \in S^p(V^*)$ apolar to all the $\frac{\partial F}{\partial c_I}(c_I'; x_r)$, $\frac{\partial F}{\partial X_r}(c_I'; x_r)$ is zero. So set each $X_r = x_r$, leaving the c_I to be determined.

Any $g \in S^p(V^*)$ apolar to $\frac{\partial F}{\partial c_I}(c_I; x_r)$ for all $I \in T(q,p) \setminus B$ is of the form

$$g = \sum_{I \in B} a_I \cdot u^I.$$

Further, g is apolar to $\frac{\partial F}{\partial X_r}(c_I; x_r)$ for all r if and only if g is apolar to $\frac{\partial F}{\partial X_r}(c_I; x_r) \cdot x_s$ for all $r \neq s$. These apolarity conditions produce a system of $q(q-1)$ homogeneous linear equations in the $q(q-1)$ unknowns a_I. We claim that the coefficient matrix of this system is the weighted biadjacency matrix $M(G_B)$. To prove this claim, we look at the apolarity condition

$$\begin{aligned}
0 &= \left\langle g, \frac{\partial F}{\partial X_r}(c_I; x_r) \cdot x_s \right\rangle \\
&= \left\langle \sum_{I \in B} a_I \cdot u^I, \frac{\partial F}{\partial X_r}(c_I; x_r) \cdot x_s \right\rangle \\
&= \sum_{I \in B} a_I \left\langle u^I, \frac{\partial F}{\partial X_r}(c_I; x_r) \cdot x_s \right\rangle \\
&= \sum_{I \in B} a_I \cdot I! \cdot \left\{ \text{coefficient of } x^I \text{ in } \frac{\partial F}{\partial X_r}(c_I; x_r) \cdot x_s \right\} \\
&= \sum_{I \in B} a_I \cdot I! \cdot \left\{ \begin{array}{ll} (i_r + 1) \cdot c_{I+e_r-e_s} & \text{if } I + e_r - e_s \in T(q,p) \setminus B \\ 0 & \text{otherwise} \end{array} \right\} \\
&= \sum_{I \in B} a_I \cdot M(G_B)_{I, e_r - e_s}.
\end{aligned}$$

Our claim is justified. The result follows.

2.2.3 Relation to matchings

Assume that the bipartite graph G_B admits a perfect matching. Such a perfect matching can be identified with a bijection $f : B \longrightarrow D$ such that $I \in B$ is joined to $f(I) \in D$ for all $I \in B$. Given such a perfect matching f, we define, in analogy with Definition 1.2.1, its multiplicity map $m_f : T(q,p) \longrightarrow \mathbb{Z}$ by letting $m_f(I) = \#\{J \in B \mid I = J + f(J)\}$. Observe that m_f vanishes on B.

Definition 2.2.1 We say that a perfect matching f of G_B is *finite-acyclic* if for any perfect matching g with $m_f = m_g$, we have $f = g$.

Remark 2.2.1 Not every G_B admits a perfect matching. For example, take $B = \{(2,0),(1,1)\} \subseteq T(2,2)$. Even if G_B admits a perfect matching, it may not admit a finite-acyclic matching. For instance, if one takes

$$B = \{(3,0,0),(2,1,0),(1,1,1),(1,0,2),(0,3,0),(0,0,3)\} \subseteq T(3,3),$$

then one can verify that G_B admits precisely two perfect matchings f and g which satisfy $m_f = m_g$.

Proposition 2.2.3 *Let $B \subseteq T(q,p)$ have size $q(q-1)$. If G_B admits a finite-acyclic matching, then B is removable.*

Proof: The proof follows that of Proposition 1.3.1. Let $f : B \longrightarrow D$ be a finite-acyclic matching. The determinant of $M(G_B)$ is, up to sign,

$$\sum_{\sigma} \left((-1)^{\sigma} \cdot \prod_{I \in B} M(G_B)_{I,\sigma f(I)} \right),$$

where σ ranges over all permutations of D. There is a one-to-one correspondence between the perfect matchings of G_B and the nonzero summands in the above expansion. The summand corresponding to $\sigma = 1$ is

$$\prod_{I \in B} M(G_B)_{I,f(I)} = \prod_{I \in B} W_{G_B}(e(I, f(I))) = K \cdot \prod_{I \in B} c_{I+f(I)} = K \cdot \prod_{J \in T(q,p) \setminus B} c_J^{m_f(J)},$$

where K is some positive constant. Since f is a finite-acyclic matching, this term is not cancelled in the above expansion. Hence the determinant of $M(G_B)$ is nonzero. By Proposition 2.2.2, B is removable.

2.2.4 Interior q-tuples and removability

Definition 2.2.2 Let $I \in T(q,p)$. We say that I is an *interior q-tuple* if $I + J \in T(q,p)$ for all $J \in D$.

The following theorem is the main result of this chapter. Its proof relies on the matching theory in $\mathbb{Z}^q$ which we developed in the previous chapter.

Theorem 2.2.1 *Let $B \subseteq T(q,p)$ consist of $q(q-1)$ interior q-tuples. Then B is removable.*

Proof: The set $D = \{e_i - e_j \mid 1 \leq i \neq j \leq q\} \subseteq \mathbb{Z}^q \setminus \{0\}$ has size $q(q-1)$. Moreover, with respect to the standard bilinear form $(\cdot,\cdot) : \mathbb{Z}^q \times \mathbb{Z}^q \longrightarrow \mathbb{Z}$ given by $(e_i, e_j) = \delta_{i,j}$, the elements of D all have length $\sqrt{2}$. By Theorem 1.2.1, there exists an acyclic matching $f : B \longrightarrow D$. Since B consists of interior q-tuples we have $b + f(b) \in T(q,p) \setminus B$ for all $b \in B$, so that f is in fact finite-acyclic. By Proposition 2.2.3, B is removable.

2.2.5 A closer look at the case $\dim(V) = 2$

Assume that $p > 1$. In the language of homogeneous polynomials, the following theorem states that any pair of terms of a generic binary form is removable through a linear change of variables, except the two pairs $\{x^p, x^{p-1}y\}$ and $\{xy^{p-1}, y^p\}$.

Theorem 2.2.2 *Suppose* $\dim(V) = 2$. *Let $B \subseteq T(2,p)$. Then B is removable if and only if $|B| \leq 2$ and $B \neq \{(p,0),(p-1,1)\}, \{(1,p-1),(0,p)\}$.*

Proof: Assume that $|B| = 2$ and consider the weighted bipartite graph G_B. The condition that B equal neither of the sets $\{(p,0),(p-1,1)\}$, $\{(1,p-1),(0,p)\}$ is equivalent to the condition that G_B admit some perfect matching $f : B \longrightarrow D$. Observe that the existence of a perfect matching $f : B \longrightarrow D$ is a necessary condition for the nonvanishing of the determinant of $M(G_B)$. Also, it is easy to see that since $|B| = 2$, every perfect matching of G_B is finite-acyclic. Therefore, by Proposition 2.2.2 and Proposition 2.2.3, we have that for $|B| = 2$, the condition that B equal neither of the sets $\{(p,0),(p-1,1)\}$, $\{(1,p-1),(0,p)\}$ is equivalent to the condition that B be removable. To complete the proof, it is enough to observe that by Proposition 2.2.1, any removable $B \subseteq T(2,p)$ has at most two elements, and that any $B \subseteq T(2,p)$ satisfying $|B| \leq 1$ is removable.

Chapter 3

Results on Skew-Symmetric Tensors

Let V be a q-dimensional vector space over the complex field $\mathbb{C}$ and consider the exterior algebra over V:

$$\Lambda(V) = \bigoplus_{p \geq 0} \Lambda^p(V).$$

One can study the problem of finding canonical expressions for the elements in $\Lambda^p(V)$, the skew-symmetric tensors of step p, and the apolarity framework developed in Chapter 2 can be adapted for this purpose. In fact, Ehrenborg [4] has recently extended the theory to handle the study of generic canonical forms in a rather broad class of objects, called $\mathcal{S}$-algebras; we will prove our results within the $\mathcal{S}$-algebra setting.

We now describe the basic question of interest to us. In accordance with current terminology, we call a wedge product of p vectors $v_1, \ldots, v_p \in V$ a *decomposable* skew-symmetric tensor, written $v_1 \wedge \cdots \wedge v_p$. We ask: What is the smallest integer n such that a generic element of $\Lambda^p(V)$ can be written as a sum of n decomposables? This integer n is known as the *essential rank* of the space $\Lambda^p(V)$, and in this chapter we describe what this number means geometrically, as well as show how one can obtain upper bounds for the essential rank by considering the Lottery Problem.

In the first section, we outline Ehrenborg's general theory of apolarity and canonical forms. We then show how the theory can be applied to skew-symmetric tensors with an example.

In Section 2, we discuss the Plücker embedding of the Grassmannian $G(p, V)$ in projective space over $\Lambda^p(V)$ and describe the connection between this embedding and the concept of essential rank.

Finally, in Section 3, we discuss the Lottery Problem and how it relates to essential rank. We conclude with a conjecture on the asymptotics of the Lottery Problem.

3.1 General theory of canonical forms

For proofs of the results in this section, we refer the reader to [4].

3.1.1 $\mathcal{S}$-algebras

Definition 3.1.1 An *$\mathcal{S}$-algebra* is a pair $(A, \mathcal{S})$ consisting of a vector space A and a collection $\mathcal{S}$ of multilinear forms on A. For each $M \in \mathcal{S}$ we let $k(M)$ denote the integer k such that $M : A^k \longrightarrow A$.

The simplest examples of $\mathcal{S}$-algebras include all associative and nonassociative algebras. In these cases, the set $\mathcal{S}$ consists of a single bilinear form.

In what follows, A will always denote a Hausdorff topological vector space over $\mathbb{C}$. If B is a finite-dimensional linear subspace of such an A, then B is closed and the induced topology on B is Euclidean. For the basic facts about topological vector spaces, see [16]. In addition, we will assume that the multilinear forms contained in $\mathcal{S}$ are continuous.

Definition 3.1.2 Let $A\{x_1, \ldots, x_n\}$ denote the smallest set containing A and the symbols $x_1, \ldots, x_n$, and which is closed under the following two operations:

1. if $p, q \in A\{x_1, \ldots, x_n\}$ and $\alpha, \beta \in \mathbb{C}$, then $\alpha p + \beta q \in A\{x_1, \ldots, x_n\}$;

2. if $M \in \mathcal{S}$ has $k(M) = k$ and $p_1, \ldots, p_k \in A\{x_1, \ldots, x_n\}$, then $M(p_1, \ldots, p_k) \in A\{x_1, \ldots, x_n\}$.

We call the elements of $A\{x_1, \ldots, x_n\}$ *$\mathcal{S}$-polynomials* in $x_1, \ldots, x_n$.

Lemma 3.1.1 *There is a unique mapping $A\{x_1, \ldots, x_n\} \times A^n \longrightarrow A$, which we denote by* eval, *such that for all $(a_1, \ldots, a_n) \in A^n$,*

1. $\mathrm{eval}(x_i; a_1, \ldots, a_n) = a_i$;

2. $\mathrm{eval}(a; a_1, \ldots, a_n) = a \quad$ for all $a \in A$;

3. $\mathrm{eval}(\alpha p + \beta q; a_1, \ldots, a_n) = \alpha \cdot \mathrm{eval}(p; a_1, \ldots, a_n) + \beta \cdot \mathrm{eval}(q; a_1, \ldots, a_n)$ for all $\alpha, \beta \in \mathbb{C}$ and $p, q \in A\{x_1, \ldots, x_n\}$;

4. $\mathrm{eval}(M(p_1, \cdots, p_k); a_1, \ldots, a_n) = M(\mathrm{eval}(p_1; a_1, \ldots, a_n), \ldots, \mathrm{eval}(p_k; a_1, \ldots, a_n))$ for $M \in \mathcal{S}$, $k = k(M)$ and $p_1, \ldots, p_k \in A\{x_1, \ldots, x_n\}$.

3.1.2 Polarizations and homogeneity

Lemma 3.1.2 *There is a unique linear map $A\{x_1, \ldots, x_n\} \longrightarrow A\{t, x_1, \ldots, x_n\}$, which we denote by D_{t,x_i}, satisfying*

1. *$D_{t,x_i}(a) = 0 \quad$ for all $a \in A$;*

2. *$D_{t,x_i}(x_j) = \delta_{ij} \cdot t$;*

3. *$D_{t,x_i}(M(p_1, \ldots, p_k)) = \sum_{j=1}^{k} M(p_1, \ldots, p_{j-1}, D_{t,x_i}(p_j), p_{j+1}, \ldots, p_k)$ for $M \in \mathcal{S}$, $k = k(M)$ and $p_1, \ldots, p_k \in A\{x_1, \ldots, x_n\}$.*

We call the linear map D_{t,x_i} polarization of x_i to t.

Definition 3.1.3 Let $V, W_1, \ldots, W_n$ be finite-dimensional linear subspaces of A. We say that an $\mathcal{S}$-polynomial $p \in A\{x_1, \ldots, x_n\}$ is *homogeneous* with respect to the spaces $V, W_1, \ldots, W_n$ if for all $w_1 \in W_1, \ldots, w_n \in W_n$ we have $\mathrm{eval}(p; w_1, \ldots, w_n) \in V$.

Proposition 3.1.1 *Let $p \in A\{x_1, \ldots, x_n\}$ and let $a_1, \ldots, a_n \in A$. The following mapping $A \longrightarrow A$ is linear:*

$$y \longmapsto \mathrm{eval}(D_{t,x_i}(p); y, a_1, \ldots, a_n).$$

Moreover, if the $\mathcal{S}$-polynomial p is homogeneous with respect to the finite-dimensional linear spaces $V, W_1, \ldots, W_n$, then $D_{t,x_i}(p)$ is homogeneous with respect to the spaces $V, W_i, W_1, \ldots, W_n$.

3.1.3 Apolarity theorem: general case

Definition 3.1.4 Let V and W be vector spaces over $\mathbb{C}$, and let $f : W \longrightarrow V$ be a linear map. We say that $L \in V^*$ is *apolar* to f if for all $w \in W$,

$$\langle L \mid f(w) \rangle = 0.$$

Definition 3.1.5 We say that a property P holds *generically* in a finite-dimensional vector space V if P holds for every element of some dense subset of V in the Euclidean topology.

Theorem 3.1.1 *Let A be an $\mathcal{S}$-algebra, and let $V, W_1, \ldots, W_n$ be finite-dimensional linear subspaces of A. Suppose $p \in A\{x_1, \ldots, x_n\}$ is homogeneous with respect to $V, W_1, \ldots, W_n$. Then a generic element $v \in V$ can be written in the form*

$$v = \mathrm{eval}(p; w_1, \ldots, w_n)$$

for some $w_i \in W_i$ if and only if there exist $w_i' \in W_i$ such that the only element of V^* apolar to the linear maps $W_j \longrightarrow V$,

$$y_j \longmapsto \mathrm{eval}(D_{t,x_j}(p); y_j, w_1', \ldots, w_n') \quad \textit{for all } j = 1, \ldots, n$$

is zero. When this is the case, we will say that p is canonical.

3.1.4 A demonstration

It is well known that every step 2 skew-symmetric tensor over an n-dimensional vector space can be written as a sum of m decomposables, where $m = \lfloor \frac{n}{2} \rfloor$; see [5]. The following is an odd-dimensional alternative to the above classical canonical form; the proof is a demonstration of how Theorem 3.1.1 can be applied in practice.

Proposition 3.1.2 *Let V be a q-dimensional vector space over $\mathbb{C}$. A generic element of $\Lambda^2(V)$ may be written as*

$$\sum_{1 \le i < j \le q-1} c_{i,j} \cdot v_i \wedge v_j,$$

for some $c_{i,j} \in \mathbb{C}$ and $v_1, \ldots, v_{q-1} \in V$ if and only if q is odd.

Proof: Let $A = \mathbb{C} \oplus V_1 \oplus V_2 \oplus \Lambda^2(V)$, where $V_1 = V_2 = V$, and let $\mathcal{S} = \{M\}$, where $M : A \times A \times A \longrightarrow A$ is the unique trilinear map satisfying

$$M(a,b,c) = \begin{cases} a \cdot b \wedge c & \text{if } a \in \mathbb{C},\ b \in V_1,\ c \in V_2; \\ 0 & \text{otherwise.} \end{cases}$$

We endow the finite-dimensional space A with the Euclidean topology and observe that the trilinear form M is continuous. Our proposed canonical form is the $\mathcal{S}$-polynomial $p \in A\{x_{i,j}, y_1, \ldots, y_{q-1}\}$ $(1 \le i < j \le q-1)$ given by

$$p = \sum_{1 \le i < j \le q-1} M(x_{i,j}, y_i, y_j).$$

Think of the $x_{i,j}$ as listed left to right according to the lexicographic order on the pairs (i,j). Notice that p is homogeneous with respect to

$$\Lambda^2(V),\ \underbrace{\mathbb{C},\dots,\mathbb{C}}_{\frac{(q-1)(q-2)}{2}},\ \underbrace{V,\dots,V}_{q-1}.$$

Let $w_1,\dots,w_{q-1}\in V$ and let $c_{i,j}\in\mathbb{C}$ for $1\le i<j\le q-1$. Consider the two groups of linear maps

$$\mathbb{C}\longrightarrow\Lambda^2(V),\quad \alpha_{k,l}\longmapsto \mathrm{eval}(D_{t,x_{k,l}}(p);\alpha_{k,l},c_{i,j},w_1,\dots,w_{q-1})$$

and

$$V\longrightarrow\Lambda^2(V),\quad v_k\longmapsto \mathrm{eval}(D_{t,y_i}(p);v_k,c_{i,j},w_1,\dots,w_{q-1}).$$

The first group can be written down explicitly as

$$\alpha_{k,l}\longmapsto\alpha_{k,l}\cdot w_k\wedge w_l,$$

and the second as

$$v_k\longmapsto\sum_{i<j\le q-1}c_{i,j}\cdot v_k\wedge w_j+\sum_{1\le j<i}c_{j,i}\cdot w_j\wedge v_k.$$

Observe that if $w_1,\dots,w_{q-1}$ are linearly dependent, then there exists a nonzero element of $\Lambda^2(V)^*$ apolar to all the above linear maps. Therefore, by Theorem 3.1.1, p is canonical if and only if we can choose linearly independent vectors $w_1,\dots,w_{q-1}\in V$ and constants $c_{i,j}\in\mathbb{C}$ such that the only element of $\Lambda^2(V)^*$ apolar to all the above linear maps (with respect to these choices) is zero.

Let $w_1,\dots,w_{q-1}\in V$ be linearly independent, leaving the $c_{i,j}$ to be determined. Let $L\in\Lambda^2(V)^*=\Lambda^2(V^*)$ be apolar to the two groups of linear maps. The first group causes L to have the form

$$L = \sum_{1 \le i \le q-1} \beta_{i,q} \cdot x_i^* \wedge x_q^*.$$

The apolarity condition on the second group produces the following system of linear equations:

$$\begin{pmatrix} 0 & -c_{1,2} & -c_{1,3} & \cdots & -c_{1,q-1} \\ c_{1,2} & 0 & -c_{2,3} & \cdots & -c_{2,q-1} \\ c_{1,3} & c_{2,3} & 0 & \cdots & -c_{3,q-1} \\ \vdots & \vdots & \vdots & & \vdots \\ c_{1,q-1} & c_{2,q-1} & c_{3,q-1} & \cdots & 0 \end{pmatrix} \cdot \begin{pmatrix} \beta_{1,q} \\ \beta_{2,q} \\ \beta_{3,q} \\ \vdots \\ \beta_{q-1,q} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

The $(q-1) \times (q-1)$ coefficient matrix, which we may denote by $\mathcal{C}$, is skew-symmetric. If q is even, then $\det \mathcal{C}$ vanishes. Since our choice of independent vectors $w_1, \ldots, w_{q-1}$ was arbitrary, we can conclude in this case that p is noncanonical. On the other hand, if q is odd, then in the expansion of $\det \mathcal{C}$ the term $\pm c_{1,2}^2 c_{3,4}^2 \cdots c_{q-2,q-1}^2$ is not cancelled. Hence, for odd q, the $\mathcal{S}$-polynomial p is canonical. This completes the proof.

3.2 Grassmannians and essential rank

Throughout this section, V denotes a q-dimensional vector space over $\mathbb{C}$.

3.2.1 Grassmannians

Let $\mathbb{P}(V)$ denote the collection of all one-dimensional subspaces of V. The set $\mathbb{P}(V)$ is known as *projective space* over V. For any $v \in V \setminus \{0\}$ we let $[v]$ denote the line spanned by v, so that $\mathbb{P}(V) = \{[v] \mid v \in V \setminus \{0\}\}$. Given a basis $v_1, \ldots, v_q$ of V, the dual basis $v_1^*, \ldots, v_q^*$ of V^* is sometimes called a *system of homogeneous* *coordinates* on $\mathbb{P}(V)$.

Consider the space of degree p symmetric tensors over V^*, $S^p(V^*)$. An element $f \in S^p(V^*)$ may be written in terms of the dual basis $v_1^*, \ldots, v_q^*$ as

$$f = \sum_{I \in T(q,p)} c_I \cdot (v_1^*)^{i_1} \cdots (v_q^*)^{i_q},$$

for some $c_I \in \mathbb{C}$. Let $x \in \mathbb{P}(V)$. The condition $f(x) \neq 0$ is well-defined, since $f(\lambda v) = \lambda^p \cdot f(v)$ for all $\lambda \in \mathbb{C}$ and $v \in V$. It therefore makes sense to speak of the common zero locus of a collection of homogeneous symmetric tensors over V^*. Such a common zero locus is known as a *projective variety.*

By a *linear variety* in $\mathbb{P}(V)$ we mean the common zero locus of some collection S of elements from $S^1(V^*)$. Every such projective variety is the image in $\mathbb{P}(V)$ of a linear subspace W of V, the linear subspace W being merely the affine variety defined by S. Conversely, if W is a linear subspace of V, then its image $\{[w] \mid w \in W \setminus \{0\}\}$ in $\mathbb{P}(V)$ is a linear variety.

The collection of linear varieties in $\mathbb{P}(V)$ is closed under the operation of taking arbitrary intersections. Dually, we define the *span* of a family of linear varieties to be the smallest linear variety in $\mathbb{P}(V)$ containing their union.

Let $G(p, V)$ denote the set of p-dimensional linear subspaces of V. In order to give the *Grassmannian* $G(p, V)$ the structure of a projective variety, we embed $G(p, V)$ in projective space, as follows. Let $W \in G(p, V)$ and let $w_1, \ldots, w_p$ be a basis for W. Consider the decomposable skew-symmetric tensor $w_1 \wedge \cdots \wedge w_p$. If $w_1', \ldots, w_p'$ is any other basis for W, then

$$w_1' \wedge \cdots \wedge w_p' = c \cdot w_1 \wedge \cdots \wedge w_p,$$

for some $c \neq 0$. Therefore, there is a well-defined map $G(p, V) \longrightarrow \mathbb{P}(\Lambda^p(V))$, which sends a p-dimensional subspace W spanned by $w_1, \ldots, w_p$ to $[w_1 \wedge \cdots \wedge w_p]$. This map is injective, and is known as the *Plücker embedding* of $G(p, V)$ in projective space. We identify $G(p, V)$ with its image in $\mathbb{P}(\Lambda^p(V))$, which is a projective variety of dimension $p(q-p)$; see Harris [11].

3.2.2 Essential rank

Definition 3.2.1 Let V be a q-dimensional vector space over $\mathbb{C}$. We define the *essential rank* of the space $\Lambda^p(V)$ to be the smallest integer n with the property that there exists a dense subset D of $\Lambda^p(V)$ such that if $f \in D$ then

$$f = \sum_{i=1}^{n} v_{i,1} \wedge \cdots \wedge v_{i,p},$$

for some $v_{i,j} \in V$. We denote the essential rank by 'ess rank $\Lambda^p(V)$'.

The following theorem gives a geometric interpretation of essential rank. It states that the essential rank of the space $\Lambda^p(V)$ equals the minimum size of a collection of projective tangent planes to the Grassmannian $G(p, V)$ whose span is the entire ambient space $\mathbb{P}(\Lambda^p(V))$. Thus, it relates essential rank to the way the Grassmannian sits in projective space under the Plücker embedding.

Theorem 3.2.1 *We have the following equality:*

$$\text{ess rank } \Lambda^p(V) \;=\; \min\left\{ n \;\middle|\; \begin{array}{l} \text{there exist } x_1, \ldots, x_n \in G(p, V) \text{ such that} \\ T_{x_1}G(p, V), \ldots, T_{x_n}G(p, V) \text{ span } \mathbb{P}(\Lambda^p(V)) \end{array} \right\}.$$

Proof: Let $A = V_1 \oplus \cdots \oplus V_p \oplus \Lambda^p(V)$, where each $V_i = V$, and let $\mathcal{S} = \{M\}$, where $M : \underbrace{A \times \cdots \times A}_{p} \longrightarrow A$ is the unique multilinear map satisfying

$$M(a_1, \ldots, a_p) = \begin{cases} a_1 \wedge \cdots \wedge a_p & \text{if each } a_i \in V_i; \\ 0 & \text{otherwise.} \end{cases}$$

Endow the finite-dimensional space A with the Euclidean topology and observe that M is continuous. Let H_n denote the set of all pairs (i, j) with $1 \leq i \leq n$ and $1 \leq j \leq p$.

Define the $\mathcal{S}$-polynomial $p \in A\{x_{i,j} | (i,j) \in H_n\}$ by

$$p = \sum_{i=1}^{n} M(x_{i,1}, \ldots, x_{i,p}).$$

Think of the $x_{i,j}$ as listed left to right according to the lexicographic order on the pairs (i,j). Notice that p is homogeneous with respect to the spaces

$$\Lambda^p(V), \underbrace{V_1, \ldots, V_p, \ldots, V_1, \ldots, V_p}_{np}.$$

Choose vectors $w_{i,j} \in V$ for each $(i,j) \in H_n$. With respect to these choices, we define a collection of linear maps $f_{k,l} : V \longrightarrow \Lambda^p(V)$, for $(k,l) \in H_n$, by

$$\begin{aligned} f_{k,l}(v) &= \text{eval}(D_{t,x_{k,l}}(p); v, w_{i,j}) \\ &= w_{k,1} \wedge \cdots \wedge w_{k,l-1} \wedge v \wedge w_{k,l+1} \wedge \cdots \wedge w_{k,p}. \end{aligned}$$

For each $(k,l) \in H_n$ let $\Omega_{k,l}$ denote the linear variety in $\mathbb{P}(\Lambda^p(V))$ corresponding to the linear subspace Image$(f_{k,l})$ of $\Lambda^p(V)$. Note that $\Omega_{k,l} \subseteq G(p,V)$. By Theorem 3.1.1, the $\mathcal{S}$-polynomial p is canonical if and only if for some choice of vectors $w_{i,j}$ the linear maps $f_{k,l}$ satisfy $\sum_{(k,l)} \text{Image}(f_{k,l}) = \Lambda^p(V)$. The projective formulation of this goes: p is canonical if and only if we can choose vectors $w_{i,j} \in V$ such that the span of the corresponding projective varieties $\Omega_{k,l}$ is all of $\mathbb{P}(\Lambda^p(V))$. We need the following simple lemma.

Lemma 3.2.1 *Assume* $p \leq q = \dim(V)$. *Let* $v_1, \ldots, v_p \in V$, *and for each* $i = 1, \ldots, p$, *let* $W_i = \{v_1 \wedge \cdots \wedge v_{i-1} \wedge v \wedge v_{i+1} \wedge \cdots \wedge v_p \mid v \in V\}$. *There exist linearly independent vectors* $v'_1, \ldots, v'_p \in V$ *such that* $\Sigma_i W_i \subseteq \Sigma_i W'_i$, *where* $W'_i = \{v'_1 \wedge \cdots \wedge v'_{i-1} \wedge v \wedge v'_{i+1} \wedge \cdots \wedge v'_p \mid v \in V\}$.

Proof: If $v_1, \ldots, v_p$ are independent, then the conclusion is obvious. If the space spanned by these vectors has dimension less than $p-1$, then each $W_i = \{0\}$ and any independent set $v'_1, \ldots, v'_p$ will do. Thus, we may assume without loss of generality that $v_1, \ldots, v_{p-1}$ are independent and that $v_p = \Sigma_{j=1}^{p-1} \alpha_j \cdot v_j$, for some $\alpha_j \in \mathbb{C}$. Put $v'_i = v_i$ for $i = 1, \ldots, p-1$ and let v'_p be any vector in $V \setminus \text{span}\{v'_1, \ldots, v'_{p-1}\}$. Observe that $W_p = W'_p$. Let $i < p$, and let $v_1 \wedge \cdots \wedge v_{i-1} \wedge v \wedge v_{i+1} \wedge \cdots \wedge v_p \in W_i$. We rewrite this vector as $v_1 \wedge \cdots \wedge v_{i-1} \wedge v \wedge v_{i+1} \wedge \cdots \wedge (\Sigma_{j=1}^{p-1} \alpha_j \cdot v_j)$

$$\begin{aligned} &= \Sigma_{j=1}^{p-1} \alpha_j \cdot v_1 \wedge \cdots \wedge v_{i-1} \wedge v \wedge v_{i+1} \wedge \cdots \wedge v_{p-1} \wedge v_j \\ &= -\alpha_i \cdot v_1 \wedge \cdots \wedge v_{p-1} \wedge v \\ &\in W'_p. \end{aligned}$$

The lemma follows.

Returning to our Theorem, we now fix k and consider the vectors $w_{k,1}, \ldots w_{k,p}$. By Lemma 3.2.1, there exist vectors $w'_{k,1}, \ldots w'_{k,p}$ such that if $f'_{k,1}, \ldots, f'_{k,p}$ are the corresponding linear maps, then $\sum_{(k,l)} \text{Image}(f_{k,l}) \subseteq \sum_{(k,l)} \text{Image}(f'_{k,l})$. Let $\Omega'_{k,l}$ denote the image in $\mathbb{P}(\Lambda^p(V))$ of $\text{Image}(f'_{k,l})$. Letting Ω_k denote the span of $\Omega_{k,1}, \ldots, \Omega_{k,p}$ and Ω'_k denote the span of $\Omega'_{k,1}, \ldots, \Omega'_{k,p}$, we have that each $\Omega_{k,l} \subseteq \Omega_k \subseteq \Omega'_k$.

Let $x_k = [w'_{k,1} \wedge \cdots \wedge w'_{k,p}] \in G(p, V)$. We have, for each $l = 1, \ldots, p$,

$$\Omega'_{k,l} = T_{x_k} \Omega'_{k,l} \subseteq T_{x_k} G(p, V).$$

As $T_{x_k} G(p, V)$ is a linear variety, we have $\Omega_k \subseteq \Omega'_k \subseteq T_{x_k} G(p, V)$.

Now observe that if p is canonical, then the span of $\Omega_1, \ldots \Omega_n$ is all of $\mathbb{P}(\Lambda^p(V))$. Since each Ω_k is contained in some projective tangent space $T_{x_k} G(p, V)$, we have that there exist $x_1, \ldots, x_n \in G(p, V)$ such that $T_{x_1} G(p, V), \ldots, T_{x_n} G(p, V)$ span $\mathbb{P}(\Lambda^p(V))$. Conversely, if we are given such $x_1, \ldots, x_n \in G(p, V)$, it is easy to see that p is canonical. In fact, a dimension count shows that the inclusion $\Omega'_k \subseteq T_{x_k} G(p, V)$ above is equality. The result follows.

3.3 The Lottery problem

Let $[n]$ denote the set $\{1, 2, \ldots, n\}$, and let $[n]_k$ denote the set of all k-subsets of $[n]$.

Definition 3.3.1 The Lottery number $L(n, k, l) = \min|S|$, where S ranges over all subsets of $[n]_k$ with the property that, for each $T \in [n]_k$ there exists $U \in S$ satisfying $|U \cap T| \geq l$.

The numbers $L(n, k, l)$ have the following popular interpretation. Consider a game involving a lottery authority and yourself in which the authority selects k numbers from $1, 2, \ldots, n$. Before this selection is made you may purchase tickets from the authority, on each of which you select k numbers from $1, 2, \ldots, n$. Then $L(n, k, l)$ equals the minimum number of tickets required to guarantee that on at least one there will be at least l of the k numbers selected by the authority; i.e, the collection of tickets bought ensures at least an l-hit.

The Lottery problem is that of determining the values $L(n, k, l)$ for all possible triples (n, k, l).

Example 3.3.1 It is easy to see that $L(n, k, 1) = \lfloor \frac{n}{k} \rfloor$. In [10], Hanani, Ornstein and Sós prove that

$$L(n, k, 2) \geq \frac{n(n-k+1)}{k(k-1)^2}$$

and

$$\lim_{n \to \infty} \frac{L(n, k, 2)}{\frac{n(n-k+1)}{k(k-1)^2}} = 1.$$

In particular, then,

$$\frac{n(n-2)}{12} \leq L(n, 3, 2) \leq \frac{n(n-2)}{12}(1 + o(1)).$$

3.3.1 Integer programming formulation

The numbers $L(n, k, l)$ can in principle be computed using an integer programming formulation of the Lottery problem. The integer programming approach is applied to a similar problem in [18].

Let $W_k = \{x = (x_1, \ldots, x_n) \in \{0,1\}^n \mid \sum_i x_i = k\}$, the binary vectors of length n and weight k. Notice that W_k is in natural bijective correspondence with the collection of k-subsets of $[n]$. Let $\pi_1, \ldots, \pi_n : \{0,1\}^n \longrightarrow \{0,1\}$ denote the coordinate maps $\pi_i(x_1, \ldots, x_n) = x_i$, and let $\Pi_k = \{(\pi_{i_1}, \ldots, \pi_{i_k}) \mid 1 \leq i_1 < \cdots < i_k \leq n\}$. Let $P_{k,l} = \{x = (x_1, \ldots, x_k) \in \{0,1\}^k \mid \sum_i x_i \geq l\}$. A collection $S \subseteq W_k$ is specified by setting

$$\begin{aligned} v_x &= 1 \quad \text{if } x \in S; \\ v_x &= 0 \quad \text{if } x \notin S. \end{aligned}$$

Finding a minimal collection S of k-subsets of $[n]$ ensuring at least an l-hit is equivalent to the following integer program:

$$\begin{array}{lll} \text{minimize} & \displaystyle\sum_{x \in W_k} v_x & \text{subject to} \\ & \displaystyle\sum_{f(x) \in P_{k,l}} v_x \geq 1 & \text{for all } f \in \Pi_k \\ & v_x \in \{0,1\} & \text{for all } x \in W_k. \end{array}$$

Example 3.3.2 Solving the above program for $(n, k, l) = (9, 3, 2), (9, 4, 3)$, we find that $L(9, 3, 2) = 7$ and $L(9, 4, 3) = 9$. Since $L(n, k, k-1) = L(n, n-k, n-k-1)$, we also get $L(9, 6, 5) = 7$ and $L(9, 5, 4) = 9$. And from the previous Example 3.3.1, we have $L(9, 2, 1) = 4$ and $L(9, 7, 6) = 4$.

3.3.2 Some bounds for $L(n, k, l)$

Definition 3.3.2 The Turán number $T(n, k, l) = \min|S|$, where S ranges over all subsets of $[n]_l$ with the property that, for each $T \in [n]_k$ there exists $U \in S$ such that $U \subseteq T$.

Proposition 3.3.1 *We have* $L(n,k,l) \geq \dfrac{T(n,k,l)}{\binom{k}{l}}$.

Proof: Let S be a collection of k-subsets of $[n]$ such that each member of $[n]_k$ intersects some member of S in at least l elements. Let $S' = \{T \in [n]_l \mid T \subseteq U,\text{ for some } U \in S\}$. We have $|S'| \leq |S| \cdot \binom{k}{l}$. Every element of $[n]_k$ contains at least one member of S'. Hence, $T(n,k,l) \leq |S'| \leq |S| \cdot \binom{k}{l}$. The conclusion follows.

Remark 3.3.1 In [3], de Caen proves that $T(n,k,l) \geq \frac{n-k+1}{l \cdot \binom{k-1}{l-1}} \cdot \binom{n}{l-1}$. We therefore obtain, after some simplification, the lower bound

$$L(n,k,l) \geq \frac{(n-k+1) \cdot \binom{n}{l-1}}{k \cdot \binom{k-1}{l-1}^2}.$$

One way to bound the number $L(n,k,l)$ from above is to consider the *covering number* $C(n,k,l)$. This number is the minimum possible size of a collection S of elements of $[n]_k$ with the property that every element of $[n]_l$ is contained in some member of S. It is obvious that $L(n,k,l) \leq C(n,k,l)$. Tables containing upper bounds for $C(n,k,l)$ can be found in [9].

3.3.3 Relation to essential rank

Using Theorem 3.1.1, it is easy to prove the following result. See [4].

Proposition 3.3.2 *We have* ess rank $\Lambda^p(\mathbb{C}^q) \leq L(q,p,p-1)$.

Corollary 3.3.1 ess rank $\Lambda^3(\mathbb{C}^q) \leq \frac{q(q-2)}{12}(1+o(1))$.

Corollary 3.3.2 ess rank $\Lambda^p(\mathbb{C}^9) \leq 1,4,7,9,9,7,4,1$ *for* $p = 1,2,3,4,5,6,7,8$, *respectively.*

Bibliography

[1] N. Alon, C. K. Fan, and D. Kleitman. Aclyclic matchings. To appear.

[2] M. Barnabei, A. Brini, and G.-C. Rota. On the exterior calculus of invariant theory. *J. Algebra*, **96**:120–160, 1985.

[3] D. de Caen. Extension of a theorem of Moon and Moser on complete subgraphs. *Ars Combinatoria*, **16**:5–10, 1985.

[4] R. Ehrenborg. On apolarity and generic canonical forms. To appear.

[5] R. Ehrenborg, G. Hetyei, and G.-C. Rota. Lectures on polynomials. To appear.

[6] R. Ehrenborg and G.-C. Rota. Apolarity and canonical forms for homogeneous polynomials. *Europ. J. Combinatorics*, **14**:157–191, 1993.

[7] C. K. Fan and J. Losonczy. Matchings and canonical forms for symmetric tensors. To appear in *Advances in Mathematics.*

[8] G. Frobenius. Über zerlegbare Determinanten. *Sitzungsber. König. Preuss. Akad. Wiss.*, **XVIII**:274–277, 1917.

[9] D. Gordon, G. Kuperberg, and O. Patashnik. New constructions for covering designs. To appear.

[10] H. Hanani, D. Ornstein, and V. T. Sós. On the lottery problem. *Magyar Tud. Akad. Kutato Int.*, **9**:155–158, 1964.

[11] J. Harris. *Algebraic Geometry.* Springer-Verlag, 1992.

[12] S. Lang. *Algebra.* Addison-Wesley, 1993.

[13] E. Lasker. Zur theorie der kanonischen formen. *Math. Annalen*, **58**:434–440, 1904.

[14] L. Lovász and M. D. Plummer. *Matching Theory.* North-Holland, 1986.

[15] H. W. Richmond. On canonical forms. *Q. J. Math.*, **33**:331–340, 1902.

[16] W. Rudin. *Functional Analysis.* McGraw-Hill, 1991.

[17] J. Schönheim. On coverings. *Pacific J. Math.*, **14**:1405–1411, 1964.

[18] N. J. A. Sloane. Covering arrays and intersecting codes. *J. Combinatorial Designs*, **1**:51–63, 1993.

[19] H. W. Turnbull. *The Theory of Determinants, Matrices and Invariants.* Dover, 1960.

[20] E. K. Wakeford. On canonical forms. *Proc. London Math. Soc.*, **18**(2):403–410, 1918-19.