New constructions in classical invariant theory

PhD thesis

Péter E. Frenkel
Supervisors: Mátyás Domokos and András Szenes

Mathematics Institute
Budapest University of Technology and Economics
2007

Preface

The three topics discussed in the three chapters of this thesis are only loosely related.

Strictly speaking, only Chapter 1 is about invariant theory, i.e., about the structure of the ring of invariant polynomials under a certain group action. Namely, it is shown that the invariant theory of the orthogonal group acting on the direct sum of several copies of the standard vector representation differs drastically over fields of characteristic 2 from the well-known theory in all other characteristics. As a result, we encounter non-classical behaviour also over the ring of integers.

In Chapter 2, we work over the field of complex numbers. We obtain new formulae for the irreducible characters of the classical matrix groups, more specifically, we express them as fractions of polynomials in the entries of matrix powers. Characters are invariant under conjugation, but the numerators and denominators of our expressions will be far from being so. Thus, our formulae can be viewed as unexpected constructions of conjugation invariant functions of matrices. The explanation is that the numerators and denominators, though themselves non-invariant, can be thought of as linear functions of suitable tensors that are covariant under conjugation.

In Chapter 3, we work over the real field, and we prove inequalities for positive semi-definite matrices. The proofs start with interpreting the given positive semi-definite matrix as a Gram matrix, i.e., a matrix of inner products, which takes us back to orthogonal invariants. The proofs end by writing the expression we want to show is non-negative as a sum of squares. The difficulty in finding the expressions to square is caused by the fact that the individual terms cannot be chosen to be orthogonal invariants, though the sum is an invariant. The way out — this idea goes back more than 40 years to Marvin Marcus — is to think of the expressions under the square signs as coordinates of a suitable tensor which is covariant and therefore easier to

construct. This basic idea links Chapter 3 to Chapter 2. Chapter 3 is the most down-to-earth part of this thesis, it ends with an application to the problem of bounding from below the norm of a product of linear functionals.

Determinants will play an important role throughout. Another matrix function used in all three chapters is the *Pfaffian* of a $2n \times 2n$ matrix $C = (c_{i,j})$:

$$\text{pf } C = \frac{1}{n!2^n} \sum_{\pi \in \mathfrak{S}_{2n}} (-1)^\pi c_{\pi(1),\pi(2)} \cdots c_{\pi(2n-1),\pi(2n)},$$

where $\mathfrak{S}_{2n}$ is the symmetric group on $2n$ letters. When C is anti-symmetric, we may rewrite this as a polynomial with integer coefficients in the matrix entries $c_{i,j}$: we do not divide by $n!2^n$, but we restrict the sum to the permutations π with

$$\pi(1) < \pi(2), \quad \ldots, \quad \pi(2n-1) < \pi(2n)$$

and

$$\pi(1) < \pi(3) < \cdots < \pi(2n-1).$$

In the anti-symmetric case, we have $(\text{pf } C)^2 = \det C$. See e.g. [GW, Appendix B.2] for other basic properties of Pfaffians.

Notations, unless very standard, will be explained in the text. The $n \times n$ identity matrix will be written $\mathbf{1}_n$ or just $\mathbf{1}$. The rank of a matrix will be denoted by rk. Determinants will be denoted by vertical bars or by det. In Chapter 3, the transposed complex conjugate of a matrix A will be denoted by A^*.

Acknowledgements

I am greatly indebted to my supervisors Mátyás Domokos and András Szenes, without whom this thesis couldn't have been written. I am grateful to Elliott Lieb, Péter Major, Máté Matolcsi, Szilárd Révész and Endre Szabó for useful discussions, and to the anonymous referees of my papers for useful comments. I also thank my family for their constant support.

Contents

1 Orthogonal vector invariants in characteristic 2 **1**
- 1.1 Introduction . . . 1
  - 1.1.1 Summary of results . . . 1
  - 1.1.2 The orthogonal group . . . 2
  - 1.1.3 Invariants . . . 3
- 1.2 Constructing indecomposable invariants . . . 6
  - 1.2.1 Preservation of invariance . . . 6
  - 1.2.2 The basic calculations . . . 7
  - 1.2.3 The constructions . . . 10
- 1.3 Separation of orbits . . . 13
  - 1.3.1 The null-cone . . . 13
  - 1.3.2 Algebro-geometric lemmas . . . 14
  - 1.3.3 Rational invariants . . . 15
  - 1.3.4 The case $m \leq n$ . . . 21
- 1.4 Proofs of indecomposability . . . 30
- 1.5 The orthogonal group scheme . . . 34

2 Character formulae for classical groups **37**
- 2.1 Introduction . . . 37
- 2.2 General linear group . . . 39
- 2.3 Special linear group . . . 41
- 2.4 Odd special orthogonal group . . . 42
- 2.5 Symplectic group . . . 44
- 2.6 Even special orthogonal group . . . 45

3 Inequalities for positive semi-definite matrices **48**
- 3.1 Introduction . . . 48
- 3.2 Old inequalities on determinants and permanents . . . 49
- 3.3 New inequalities . . . 51
  - 3.3.1 Pfaffians . . . 51
  - 3.3.2 Hafnians . . . 53
- 3.4 Products of real linear functionals . . . 55

Chapter 1

Orthogonal vector invariants in characteristic 2

1.1 Introduction

1.1.1 Summary of results

This chapter is based on the two papers [DF1, DF2], which are joint with Mátyás Domokos. Over an algebraically closed base field $\mathbb{K}$ of characteristic 2, we study the ring $\mathbb{K}[n \times m]^G$ of invariants, where G is the orthogonal group $\mathrm{O}_n(\mathbb{K})$ or the special orthogonal group $\mathrm{SO}_n(\mathbb{K})$. The group G acts by left multiplication on the space $\mathbb{K}^{n\times m}$ of $n \times m$ matrices and thus acts also on the coordinate ring $\mathcal{O}(\mathbb{K}^{n\times m}) = \mathbb{K}[n \times m]$ of $\mathbb{K}^{n\times m}$, which is just a polynomial ring in $n \times m$ variables. We prove for $\mathrm{O}_n(\mathbb{K})$ $(n \geq 2)$ and for $\mathrm{SO}_n(\mathbb{K})$ $(n \geq 3)$ that there exist m–linear invariant polynomials with m arbitrarily large that are indecomposable (i.e., not expressible as polynomials in invariants of lower degree). This is in sharp contrast with the uniform description of the ring of invariant polynomials valid in all other characteristics. In fact, we shall explicitly construct indecomposable m–linear invariant polynomials for all possible values of m. Indecomposability of corresponding invariants of the complex (special) orthogonal group over $\mathbb{Z}$ will immediately follow. The constructions rely on analyzing the Pfaffian of the skew-symmetric matrix whose entries above the diagonal are the inner products of the vector variables. The constructions are given in Section 1.2, and indecomposability is proved in Section 1.4.

On the other hand, we will show that invariant rational functions in char-

acteristic 2 behave basically the same way as in all other characteristics (Theorem 1.3.5), and so do the invariant polynomials if $m \leq n$ (Theorems 1.3.16, 1.3.17 and 1.3.18). We shall see for all n and m that the algebra $\mathbb{K}[n \times m]^G$ is a finitely generated module over the subalgebra generated by the quadratic invariants (Corollary 1.3.2).

1.1.2 The orthogonal group

Let $\mathbb{F}$ stand for an algebraically closed field of arbitrary characteristic. Let n be a positive integer. Denote coordinates in $\mathbb{F}^n$ by x_1, y_1, ..., x_r, y_r if $n = 2r$ or by x_1, y_1, ..., x_r, y_r, z if $n = 2r + 1$. The *standard quadratic form* in n variables is

$$\begin{aligned} q &= x_1y_1 + \cdots + x_ry_r && \text{when } n = 2r, \\ q &= x_1y_1 + \cdots + x_ry_r + z^2 && \text{when } n = 2r + 1. \end{aligned}$$

The *polar form* β of q is the symmetric bilinear form given by

$$\beta\left(v^{(1)}, v^{(2)}\right) = q\left(v^{(1)} + v^{(2)}\right) - q\left(v^{(1)}\right) - q\left(v^{(2)}\right).$$

We define

$$\ker \beta = \{v \in \mathbb{F}^n \mid \beta(v, \cdot) = 0\}.$$

This is the trivial subspace unless $\operatorname{char} \mathbb{F} = 2$ and $n = 2r+1$, in which case it is the z axis. The quadratic form q is easily seen to be *non-degenerate*, i.e., if $v \in \ker \beta$ and $q(v) = 0$, then $v = 0$. Note that over any algebraically closed field, q is the only non-degenerate quadratic form up to change of basis.

We say that the linear transformation $g : \mathbb{F}^n \to \mathbb{F}^n$ *leaves q invariant* if $q \circ g = q$.

A vector $u \in \mathbb{F}^n$ is *singular* if $q(u) = 0$ and *non-singular* otherwise. A linear subspace is *totally singular* if all its vectors are singular.

Any non-singular vector u gives rise to the *reflection* g_u defined by

$$g_u v = v - \frac{\beta(v, u)}{q(u)} u. \tag{1.1}$$

All reflections leave q invariant. All reflections have determinant -1. The identity is a reflection if and only if $\ker \beta \neq 0$, i.e., if and only if $\operatorname{char} \mathbb{F} = 2$ and $n = 2r + 1$. In all other cases, the identity is not a product of an odd number of reflections [F0, T].

The *orthogonal group* $\mathrm{O}_n(\mathbb{F})$ is defined as the group of linear isomorphisms of $\mathbb{F}^n$ that leave the quadratic form q invariant. The group $\mathrm{O}_n(\mathbb{F})$ is a Zariski closed subvariety of the space $M_n(\mathbb{F})$ of linear transformations of $\mathbb{F}^n$. All elements of $\mathrm{O}_n(\mathbb{F})$ have determinant ± 1; in fact, they are all products of reflections [F0, T].

Recall that in the Zariski topology, the irreducible components of any algebraic group coincide with the connected components. The *special orthogonal group* $\mathrm{SO}_n(\mathbb{F})$ is defined as the component of $\mathrm{O}_n(\mathbb{F})$ containing the identity, which is the set of elements that are products of an even number of reflections.

Throughout this chapter $\mathbb{K}$ stands for an algebraically closed field of characteristic 2. The elements of $\mathrm{O}_n(\mathbb{K})$ all have determinant $\pm 1 = 1$. For $n = 2r$, the group $\mathrm{O}_{2r}(\mathbb{K})$ nevertheless has two components, due to the fact mentioned above that the identity is not a product of an odd number of reflections. For $n = 2r + 1$, the group $\mathrm{O}_{2r+1}(\mathbb{K})$ has only one component, thus

$$\mathrm{O}_{2r+1}(\mathbb{K}) = \mathrm{SO}_{2r+1}(\mathbb{K}).$$

This group carries a natural non-reduced scheme structure. We postpone the definition of this scheme $\mathrm{O}_{2r+1,\mathbb{K}}$ to Section 1.5, because we want to avoid scheme theory for the time being.

1.1.3 Invariants

Let $\mathbb{S} \ni 1$ be a subring of the algebraically closed field $\mathbb{F}$. We write $R = \mathbb{S}[n \times m]$ for the algebra of polynomials over $\mathbb{S}$ in $n \times m$ variables. We arrange the variables to form the $n \times m$ matrix V. The columns of V are n–dimensional vectors $v^{(1)}, \ldots, v^{(m)}$. A typical element of R can be written as

$$f = f(V) = f\left(v^{(1)}, \ldots, v^{(m)}\right).$$

The polynomial f is said to be *even* if it is a sum of homogeneous polynomials of even degree. The subalgebra of R formed by even polynomials is written R_0.

An $n \times n$ matrix $g \in \mathrm{O}_n(\mathbb{F})$ acts on $n \times m$ matrices by left multiplication (equivalently, acts on the column vectors by linear substitution):

$$g : V = \left(v^{(1)}, \ldots, v^{(m)}\right) \mapsto gV = \left(gv^{(1)}, \ldots, gv^{(m)}\right).$$

We now define rings of invariant polynomials. For $G = \mathrm{O}_n(\mathbb{F})$ or $G = \mathrm{SO}_n(\mathbb{F})$, we define the ring of G–invariants:

$$\mathbb{S}[n \times m]^G = \{f \in \mathbb{S}[n \times m] \mid f(gV) = f(V) \quad (g \in G)\}.$$

For even n, we define

$$\mathbb{S}[n \times m]^{\mathrm{O}_{n,\mathbb{F}}} = \mathbb{S}[n \times m]^{\mathrm{O}_n(\mathbb{F})}.$$

For all n, we define

$$\mathbb{S}[n \times m]^{\mathrm{SO}_{n,\mathbb{F}}} = \mathbb{S}[n \times m]^{\mathrm{SO}_n(\mathbb{F})}.$$

For odd n, we define

$$\mathbb{S}[n \times m]^{\mathrm{O}_{n,\mathbb{F}}} = \mathbb{S}[n \times m]_0 \cap \mathbb{S}[n \times m]^{\mathrm{SO}_n(\mathbb{F})}.$$

We might sometimes drop n or $\mathbb{F}$ from the subscript of the superscript if it is obvious from the context.

The definition of $\mathbb{S}[n \times m]^{\mathrm{O}_{n,\mathbb{F}}}$ for odd n may seem artificial. However, observe that if $\mathrm{char}\,\mathbb{F} \neq 2$, then

$$\mathbb{S}[(2r+1) \times m]^{\mathrm{O}_{\mathbb{F}}} = \mathbb{S}[(2r+1) \times m]^{\mathrm{O}(\mathbb{F})}.$$

In characteristic 2, this is no longer true, but we nevertheless stick to the above definition of $\mathbb{S}[n \times m]^{\mathrm{O}_{n,\mathbb{F}}}$. This terminology will be convenient for the formulation of some of our theorems, and it will be justified in Section 1.5 when we introduce the orthogonal group scheme.

Next we define some distinguished elements in R. Set

$$(kk) = (kk)_n = q\left(v^{(k)}\right) \qquad (1 \leq k \leq m)$$

and

$$(kl) = (kl)_n = \beta\left(v^{(k)}, v^{(l)}\right) \qquad (1 \leq k \leq m, \quad 1 \leq l \leq m, \quad k \neq l).$$

More explicitly, for $n = 2r$ we have

$$(kl) = x_1^{(k)} y_1^{(l)} + y_1^{(k)} x_1^{(l)} + \cdots + x_r^{(k)} y_r^{(l)} + y_r^{(k)} x_r^{(l)} \qquad (k \neq l),$$

whereas for $n = 2r + 1$ we have

$$(kl) = x_1^{(k)} y_1^{(l)} + y_1^{(k)} x_1^{(l)} + \cdots + x_r^{(k)} y_r^{(l)} + y_r^{(k)} x_r^{(l)} + 2z^{(k)} z^{(l)} \qquad (k \neq l).$$

Let

$$[i_1, \ldots, i_n] = \det \left[v^{(i_1)}, \ldots, v^{(i_n)} \right] \tag{1.2}$$

be the determinant of the matrix that has $v^{(i_1)}$, ..., $v^{(i_n)}$ as its columns. Then all (kl) are orthogonal invariants, and all $[i_1, \ldots, i_n]$ are special orthogonal invariants.

The classical "first fundamental theorem" for the (special) orthogonal group asserts that when $\mathbb{F}$ is of characteristic zero, the algebra $\mathbb{F}[n \times m]^{\mathrm{O}}$ is generated by the inner products (kl) of the indeterminate vectors under consideration, and the algebra $\mathbb{F}[n \times m]^{\mathrm{SO}}$ is generated by the inner products and the determinants. This has been discussed along with the analogous results for the other classical groups by Hermann Weyl in [W2]. These results are the most important specific examples of a general theorem of Weyl: if $\mathrm{char}\, \mathbb{F} = 0$, and we replace O_n and SO_n by any group G of linear transformations of $\mathbb{F}^n$, then the ring $\mathbb{F}[n \times m]^G$ is generated by elements each of which depends on at most n of the m vector variables.

De Concini and Procesi [CP, P], and later Richman [Ri], worked out characteristic free treatments of the theory of vector invariants of classical groups; in particular, they proved that the first fundamental theorem for the (special) orthogonal group remains unchanged in odd characteristic. In characteristic 2, De Concini and Procesi showed the same for the (non-reduced) group scheme G preserving the bilinear form $x_1^{(1)}x_1^{(2)} + \cdots + x_n^{(1)}x_n^{(2)}$. Richman worked with the corresponding reduced group, and showed that the invariant algebra is generated in degree 1 and 2. However, though this group G preserves the quadratic form $x_1^2 + \cdots + x_n^2$, it is not the orthogonal group in characteristic 2, not even up to change of basis: the quadratic form $x_1^2 + \cdots + x_n^2$ is the square of a linear form, hence is degenerate. So in characteristic 2 the question about vector invariants of the orthogonal group remains open. We shall see in Section 1.3 that the field of rational (special) orthogonal invariants is generated by obvious low-degree invariants even in characteristic 2. However, the behaviour of polynomial invariants will turn out in Sections 1.2 and 1.4 to be very much different. The general theorem of Hermann Weyl mentioned above fails drastically. This was known to happen for invariants of finite groups in positive characteristic, but it might be interesting to see it happen for a classical matrix group acting in its standard vector representation.

1.2 Constructing indecomposable invariants

An element of the polynomial algebra $R = \mathbb{S}[n \times m]$ is called *m–linear* if it is multi-linear in the vector variables $v^{(1)}, \ldots, v^{(m)}$. An element of the $\mathbb{S}$–algebra R^G of invariant polynomials with coefficients in $\mathbb{S}$ is said to be an *indecomposable* element if it is not contained in the subring of R^G generated by the elements of lower degree.

If n is odd or m is even, and $m \geq n \geq 3$, we shall construct an indecomposable m–linear $\mathrm{SO}_{n,\mathbb{K}}$–invariant. If m is even and $m > n \geq 2$, we shall construct an indecomposable m–linear $\mathrm{O}_{n,\mathbb{K}}$–invariant.

Note that for any $\mathrm{SO}_{n,\mathbb{K}}$–invariants of degree m to exist at all, it is necessary that n be odd or m be even — either see Theorem 1.3.5 or just consider the action of the r–dimensional torus $\mathrm{SO}_2(\mathbb{K})^r$ whose elements are the diagonal matrices in $\mathrm{SO}_{2r}(\mathbb{K})$. It follows that $\mathrm{O}_{n,\mathbb{K}}$–invariants can exist only in even degrees.

All of our indecomposable invariant polynomials shall be constructed as modulo 2 images of invariant polynomials with integer coefficients. It follows that the latter will have the corresponding indecomposability properties over $\mathbb{Z}$.

We give the constructions in this section. The proofs of indecomposability are postponed to Section 1.4.

The paper [DKZ] contained a more sophisticated proof of the existence of high-degree indecomposable invariants in the $\mathrm{SO}_{4,\mathbb{K}}$ case.

It seems very likely, but we are unable to prove, that our invariants, together with the standard quadratic invariants (kl) defined in the previous section, generate the ring of invariant polynomials.

1.2.1 Preservation of invariance

We shall need the following three lemmas. They say that the ring homomorphism $\mathbb{Z} \to \mathbb{S} \subset \mathbb{F}$ induces a homomorphism $\mathbb{Z}[n \times m] \to \mathbb{S}[n \times m] \subset \mathbb{F}[n \times m]$ which takes invariants to invariants. By a slight abuse of notation, a polynomial with integer coefficients might be denoted by the same letter as its image. Let $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ be the image of $\mathbb{Z}$ in $\mathbb{S}$; if $\mathrm{char}\, \mathbb{F} = 0$, then $\mathbb{F}_p = \mathbb{F}_0 = \mathbb{Z}$.

Lemma 1.2.1 *If $f \in \mathbb{Z}[2r \times m]^{\mathrm{O}_{2r,\mathbb{C}}}$, then its reduction mod p is contained in $\mathbb{F}_p[2r \times m]^{\mathrm{O}_{2r,\mathbb{F}}}$.*

Proof. Being invariant under O_{2r} is equivalent to being invariant under any reflection. We need to prove that if $f(g_u V) = f(V)$ holds over $\mathbb{C}$ for all $u \in \mathbb{C}^{2r}$, $q(u) \neq 0$, then it holds over $\mathbb{F}$ for all $u \in \mathbb{F}^{2r}$, $q(u) \neq 0$.

Coefficients of both sides may be viewed as rational functions with coefficients in $\mathbb{Z}$ resp. $\mathbb{F}_p$ of the vector variable u, and O_{2r}-invariance of f boils down to formal equality of pairs of such rational functions. Since formal equality over $\mathbb{Z}$ implies that over $\mathbb{F}_p$, the lemma is proved. □

Lemma 1.2.2 *If $f \in \mathbb{Z}[n \times m]^{\mathrm{SO}_{n,\mathbb{C}}}$, then its reduction mod p is contained in $\mathbb{F}_p[n \times m]^{\mathrm{SO}_{n,\mathbb{F}}}$.*

Proof. The proof is analogous to the previous one. Being invariant under SO_n is equivalent to being invariant under the product of any two reflections. We need to prove that if $f(g_u g_w V) = f(V)$ holds over $\mathbb{C}$ for all $u, w \in \mathbb{C}^n$, $q(u)q(w) \neq 0$, then it holds over $\mathbb{F}$ for all $u \in \mathbb{F}^n$, $q(u)q(w) \neq 0$.

Coefficients of both sides may be viewed as rational functions with coefficients in $\mathbb{Z}$ resp. $\mathbb{F}_p$ of the vector variables u and w, and SO_n-invariance of f boils down to formal equality of pairs of such rational functions. Since formal equality over $\mathbb{Z}$ implies that over $\mathbb{F}_p$, the lemma is proved. □

Lemma 1.2.3 *If $f \in \mathbb{Z}[(2r+1) \times m]^{\mathrm{O}_{2r+1,\mathbb{C}}}$, then its reduction mod p is contained in $\mathbb{F}_p[(2r+1) \times m]^{\mathrm{O}_{2r+1,\mathbb{F}}}$.*

Proof. By the previous lemma, the image of f is invariant under $\mathrm{SO}_{2r+1}(\mathbb{F})$. We need to prove that the image of f is a sum of homogeneous polynomials of even degrees. It suffices to prove that f itself is such. This follows from the definition of $\mathrm{O}_{2r+1,\mathbb{C}}$–invariance. □

1.2.2 The basic calculations

We shall use the symbol $*$ to mean any one of the two letters x and y. We define the sign $\operatorname{sgn} w$ of an m–linear monomial

$$w = z^{(i_{01})} \dots z^{(i_{0m_0})} *_1^{(i_{11})} \dots *_1^{(i_{1m_1})} \dots\dots *_r^{(i_{r1})} \dots *_r^{(i_{rm_r})} \tag{1.3}$$

with $i_{j1} < \dots < i_{jm_j}$ for each j to be the sign of the permutation

$$i_{01}, \dots, i_{0m_0}, i_{11}, \dots, i_{1m_1}, \dots\dots, i_{r1}, \dots, i_{rm_r}$$

of the indices $1, \dots, m$.

In our next two propositions we shall be working over $\mathbb{Z}$. We write

$$Pf(i_1, \ldots, i_{2\mu}) = Pf(i_1, \ldots, i_{2\mu})_n =$$
$$= \mathrm{pf} \begin{pmatrix} 0 & (i_1 i_2) & (i_1 i_3) & \ldots & (i_1 i_{2\mu}) \\ -(i_1 i_2) & 0 & (i_2 i_3) & \ldots & (i_2 i_{2\mu}) \\ -(i_1 i_3) & -(i_2 i_3) & 0 & \ldots & (i_3 i_{2\mu}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -(i_1 i_{2\mu}) & -(i_2 i_{2\mu}) & -(i_3 i_{2\mu}) & \ldots & 0 \end{pmatrix} \tag{1.4}$$

for the Pfaffian of the $2\mu \times 2\mu$ anti-symmetric matrix whose upper half consists of the $(\cdot, \cdot)$'s corresponding to the indices $i_1, \ldots, i_{2\mu}$. All entries of the matrix, and therefore also its Pfaffian, are invariant under $\mathrm{O}_n(\mathbb{C})$. The Pfaffian (1.4) is multi-linear in the vectors $v^{(i_1)}, \ldots, v^{(i_{2\mu})}$ if there is no repetition in the indices.

Proposition 1.2.4 *The coefficient in $Pf(1, \ldots, 2\mu)$ of a 2μ–linear monomial w as in (1.3) above is zero unless, for each j, $m_j = 2\mu_j$ is even and the x_j and the y_j occur in an alternating order. In this case, the coefficient is*

$$2^{\mu - |\{j>0 : \mu_j \geq 1\}|} \mathrm{sgn}\ w.$$

Proof. The coefficient of w is given by substituting the corresponding vectors of the standard basis into $Pf(1, \ldots, 2\mu)$. This amounts to calculating the Pfaffian of the skew-symmetric matrix given by the β's of these vectors. Permuting rows and columns turns it into a direct sum of matrices and shows that the Pfaffian is $\mathrm{sgn}\ w$ times a product of $m_j \times m_j$ Pfaffians, one factor for each index j. So, assuming that w has a non-zero coefficient, each m_j must be even. Set $\mu_j = m_j/2$.

If, for some $j > 0$, we have two occurrences of x_j with no y_j in between (or *vice versa*), then the j–th Pfaffian has two identical (adjacent) rows, which makes it zero. So the x_j and y_j must occur in an alternating order for each $j > 0$. Therefore, the j–th Pfaffian, for $j > 0$, is that of the $m_j \times m_j$ checkerboard matrix with entries $b^{(kl)} = \mathrm{sgn}(l - k)$ for odd $l - k$ and zero otherwise. Expanding by the first row and using induction on μ_j shows that this is $2^{\mu_j - 1}$ if $\mu_j \geq 1$. It is $1 = 2^{\mu_j}$ if $\mu_j = 0$.

For odd n and $j = 0$, we have the Pfaffian of the $m_0 \times m_0$ matrix with entries $b^{(kl)} = 2\mathrm{sgn}(l - k)$. Expanding by the first row and using induction on μ_0 shows that this is 2^{μ_0}.

Adding up the exponents of 2 yields the result. □

Proposition 1.2.5 *Let $m \geq n$ with $m \equiv n \mod 2$, and set $\mu = \lfloor m/2 \rfloor$. Consider the m–linear polynomial*

$$\sum (-1)^{\pi} \cdot [\pi(1), \ldots, \pi(n)] Pf(\pi(n+1), \ldots, \pi(m)), \tag{1.5}$$

the sum being extended over those permutations $\pi \in \mathfrak{S}_m$ that satisfy $\pi(1) < \cdots < \pi(n)$ and $\pi(n+1) < \cdots < \pi(m)$. The coefficient in (1.5) of an m–linear monomial w as in (1.3) above is zero unless, for each $j > 0$, m_j is even and strictly positive, and the x_j and the y_j occur in an alternating order. In this case, the coefficient is

$$2^{\mu - r} (-1)^{\left| \left\{ j > 0 \,:\, *_j^{(i_{j1})} = y_j^{(i_{j1})} \right\} \right|} \operatorname{sgn} w.$$

Proof. Assuming that w has a non-zero coefficient, it follows trivially that for each $j > 0$, m_j must be even and strictly positive. Set $\mu_j = m_j/2$. We call a choice of one x_j and one y_j from w a *good* choice if the remaining $2(\mu_j - 1)$ of the x_j and y_j occur in an alternating order. Any choice of one z from w shall be called a good choice. Now Proposition 1.2.4 shows that the terms in the sum (1.5) in which w has non-zero coefficients correspond to the simultaneous good choices (for each j) of one of each letter from w, and the coefficient of w in such a term is

$$2^{\mu - r - |\{j > 0 : \mu_j \geq 2\}|} \operatorname{sgn} w \cdot \prod_r \pm 1,$$

where the j–th ± 1 is the sign of that permutation of the letters x_j and y_j in w (resp. the letters z in w) which puts the chosen one(s) in front and in alphabetical order, leaving the rest in the order they had in w.

It follows that the coefficient of w in (1.5) is

$$2^{\mu - r - |\{j > 0 : \mu_j \geq 2\}|} \operatorname{sgn} w \cdot \prod_j \sum \pm 1,$$

where the j–th summation is over the good choices of one x_j and one y_j (resp. one z) from w.

If we have two occurrences of x_j with no y_j in between (or *vice versa*), then the j–th sum is zero, for we can pair off the choices by interchanging the role of the two adjacent x_j's, and the two ± 1's in each pair will cancel.

So the x_j and y_j must occur in an alternating order for each $j > 0$. In this case, the j–th sum is

$$\pm 1 \quad \text{if} \quad \mu_j = 1 \qquad \text{and} \qquad \pm\left(1 + \sum_{l=1}^{m_j - 1} (-1)^{l-1}\right) = \pm 2 \quad \text{if} \quad \mu_j \geq 2.$$

The sign is $+$ if x_j comes first and it is $-$ if y_j comes first.

For odd n, the 0–th sum is

$$\sum_{l=1}^{m_0} (-1)^{l-1} = 1.$$

Adding up the exponents of 2 and (-1) respectively, we arrive at the result. $\square$

1.2.3 The constructions

For $\mu \geq r$, divide the Pfaffian in Proposition 1.2.4 by $2^{\mu - r}$ to get a 2μ–linear $\mathrm{O}_n(\mathbb{C})$–invariant with integer coefficients, call it

$$(1, \ldots, 2\mu) = (1, \ldots, 2\mu)_n \in \mathbb{Z}[n \times 2\mu]^{\mathrm{O}_{n,\mathbb{C}}}.$$

Apply the ring homomorphism $\mathbb{Z} \to \mathbb{F}$ to get a 2μ–linear $\mathrm{O}_{n,\mathbb{F}}$–invariant with coefficients in $\mathbb{F}_p$, it will be still called

$$(1, \ldots, 2\mu) = (1, \ldots, 2\mu)_n \in \mathbb{F}_p[n \times 2\mu]^{\mathrm{O}_{n,\mathbb{F}}}.$$

Analogously, for $m \geq n$ with $m \equiv n \mod 2$, divide the sum (1.5) of Proposition 1.2.5 by $2^{\mu - r}$ to get an m–linear $\mathrm{SO}_n(\mathbb{C})$–invariant with integer coefficients, call it

$$[1, \ldots, m] = [1, \ldots, m]_n \in \mathbb{Z}[n \times m]^{\mathrm{SO}_{n,\mathbb{C}}}.$$

Apply $\mathbb{Z} \to \mathbb{F}$ to get an m–linear $\mathrm{SO}_n(\mathbb{F})$–invariant with coefficients in $\mathbb{F}_p$, call it

$$[1, \ldots, m] = [1, \ldots, m]_n \in \mathbb{F}_p[n \times m]^{\mathrm{SO}_{n,\mathbb{F}}}.$$

Note that invariance over $\mathbb{F}$ in each case follows from that over $\mathbb{C}$ by the lemmas above. Note also that the square bracket notation introduced here generalizes the one defined in formula (1.2).

Now consider the case when $\mathbb{F} = \mathbb{K}$ has characteristic 2. By Proposition 1.2.4, the invariant

$$(1, \dots, m) \in \mathbb{F}_2[n \times m]^{\mathrm{O}_{n,\mathbb{K}}},$$

when it is defined, is the sum of those m–linear monomials w that, when written in the form (1.3), have strictly positive and even m_j for all $j > 0$, and the x_j and y_j occur in an alternating order. By Proposition 1.2.5, the same holds for

$$[1, \dots, m] \in \mathbb{F}_2[n \times m]^{\mathrm{SO}_{n,\mathbb{K}}}.$$

This shows in particular that for $m \geq n$ both even, $(1, \dots, m) = [1, \dots, m]$ in characteristic 2.

For odd n and any $m \geq n$, exactly one of $(1, \dots m)$ and $[1, \dots, m]$ is defined. We shall prove in Section 1.4 that it is an indecomposable $\mathrm{SO}_{n,\mathbb{K}}$–invariant if $n \geq 3$. This will rely on the following observation. We define the *multiplicity in x, y* of a polynomial to be the infimum of the total degrees of its monomials in the x, y variables. (For a non-zero homogeneous polynomial, this equals the difference between the degree and the degree in the z variables.)

Fact 1.2.6 *Let $\mu \geq r$. Then the $\mathrm{O}_{2r+1,\mathbb{K}}$–invariant $(1, \dots, 2\mu)_{2r+1}$ and the $\mathrm{SO}_{2r+1,\mathbb{K}}$–invariant $[1, \dots, 2\mu+1]_{2r+1}$ have multiplicity $2r$ in x, y.*

Proof. This is obvious from the description above of the monomials that appear. □

We shall also prove in Section 1.4 that for $n = 2r$ and $\mu > r$, the $\mathrm{O}_{2r,\mathbb{K}}$–invariant $(1, \dots, 2\mu)$ is indecomposable. This will rely on the following, slightly more subtle multiplicity observation. We pass to the new linear coordinates $t_j = x_j$ and $s_j = x_j + y_j$ $(j = 1, \dots, r)$. We define the *multiplicity in s* of a polynomial to be the infimum of the total degrees of its monomials in the s variables.

Fact 1.2.7 *Let $\mu \geq r$. Then the multiplicity in s of the $\mathrm{O}_{2r,\mathbb{K}}$–invariant $(1, \dots, 2\mu)_{2r}$ is at most r.*

Proof. The 2μ–linear monomial $s_1^{(1)} \cdots s_r^{(r)} t_1^{(r+1)} \cdots t_r^{(2r)} t_1^{(2r+1)} \cdots t_1^{(2\mu)}$ occurs with coefficient 1. □

In the remainder of this section, let $m \geq n$ both be even. We turn to the construction of an m–linear $\mathrm{SO}_n(\mathbb{F})$–invariant which is indecomposable if char $\mathbb{F} = 2$ and $n \geq 4$. Subtract the sum (1.5) of Proposition 1.2.5 from the Pfaffian in Proposition 1.2.4 and divide by $2^{\mu-r+1}$ to get an m–linear $\mathrm{SO}_n(\mathbb{C})$–invariant with integer coefficients. Call it

$$\langle 1, \ldots, m \rangle = \langle 1, \ldots, m \rangle_n \in \mathbb{Z}[n \times m]^{\mathrm{SO}_n(\mathbb{C})},$$

noting that

$$\langle 1, \ldots, m \rangle = ((1, \ldots, m) - [1, \ldots, m])/2.$$

Apply $\mathbb{Z} \to \mathbb{F}$ to get an m–linear $\mathrm{SO}_n(\mathbb{F})$–invariant

$$\langle 1, \ldots, m \rangle = \langle 1, \ldots, m \rangle_n \in \mathbb{F}_p[n \times m]^{\mathrm{SO}_n(\mathbb{F})}.$$

Invariance of $\langle 1, \ldots, m \rangle$ under $\mathrm{SO}_n(\mathbb{F})$ again follows from that under $\mathrm{SO}_n(\mathbb{C})$ by Lemma 1.2.2.

The indecomposability proof for $\langle 1, \ldots, m \rangle$, to be given in Section 1.4, will rely on

Fact 1.2.8 *Let $\mu \geq r$. Substituting z in place of x_r and y_r in the $\mathrm{SO}_{2r}(\mathbb{C})$–invariant $\langle 1, \ldots, 2\mu \rangle_{2r}$, we get the $\mathrm{O}_{2r-1}(\mathbb{C})$–invariant $(1, \ldots, 2\mu)_{2r-1}$.*

Proof. Following the definitions, we have

$$\begin{aligned} \langle 1, \ldots, 2\mu \rangle_{2r} \mid_{x_r = y_r = z} &= 2^{-1}(1, \ldots, 2\mu)_{2r} \mid_{x_r = y_r = z} = \\ &= 2^{r-\mu-1} Pf(1, \ldots, 2\mu)_{2r} \mid_{x_r = y_r = z} = \\ &= 2^{r-\mu-1} Pf(1, \ldots, 2\mu)_{2r-1} = (1, \ldots, 2\mu)_{2r-1}, \end{aligned}$$

as claimed. □

We shall also need

Fact 1.2.9 *Let $\mu \geq r$. Then the $\mathrm{SO}_{2r}(\mathbb{F})$–invariant $\langle 1, \ldots, 2\mu \rangle$ is not invariant under $\mathrm{O}_{2r}(\mathbb{F})$.*

Proof. The orthogonal group contains the reflection $x_1 \leftrightarrow y_1$, under which $\langle 1, \ldots, 2\mu \rangle \in \mathbb{F}[n \times m]$ is not invariant. Indeed, remember that $(1, \ldots, 2\mu) \in \mathbb{Z}[2r \times 2\mu]$ is invariant under $\mathrm{O}_n(\mathbb{C})$, therefore also under $x_1 \leftrightarrow y_1$. If char $\mathbb{F} \neq 2$, we only need that $[1, \ldots, 2\mu] \in \mathbb{F}[2r \times 2\mu]$ is not invariant, which is obvious from Proposition 1.2.5 and the definition of $[1, \ldots, 2\mu]$, since $1 \neq -1$ in $\mathbb{F}$. For the characteristic 2 case, we need that $[1, \ldots, 2\mu] \in \mathbb{Z}[2r \times 2\mu]$ is not invariant mod 4, which is again obvious. □

Remark 1.2.10 Using a formula of Wick [Ga, L1, Wi] on Pfaffians, all the invariant polynomials constructed in this section can be tied up with certain basic Clifford algebra representations. We do not pursue this connection here.

1.3 Separation of orbits

The theorems in this section are valid over any algebraically closed field $\mathbb{F}$, but most of them are well known if char $\mathbb{F} \neq 2$. The interesting part is that they are valid also in characteristic 2. The proofs use Witt's theorem [T, Theorem 7.4], standard facts concerning reductive groups, and basic algebraic geometry.

1.3.1 The null-cone

Recall that the null-cone corresponding to a graded algebra of polynomials is defined to be the locus of common zeros of its homogeneous elements of positive degree.

Let us introduce the notation

$$A = \mathbb{F}\left[(kl) : 1 \leq k \leq l \leq m\right].$$

This is a subalgebra of $\mathbb{F}[n \times m]^{\mathrm{O}_{\mathbb{F}}}$. The (kl) are not in general algebraically independent, so A is in general not a polynomial algebra.

Theorem 1.3.1 *The null-cones corresponding to the three algebras*

$$\mathbb{F}[n \times m]^{\mathrm{SO}} \geq \mathbb{F}[n \times m]^{\mathrm{O}_{\mathbb{F}}} \geq A$$

are the same.

Proof. If char $\mathbb{F} \neq 2$, then the theorem follows from the result in [CP] saying that

$$\mathbb{F}[n \times m]^{\mathrm{SO}} = A\left[[i_1, \ldots, i_n] : 1 \leq i_1 < \cdots < i_n \leq m\right],$$

where $[i_1, \ldots, i_n]^2 \in A$. So let us assume that $\mathbb{F} = \mathbb{K}$ has characteristic 2.

Suppose that the point $\left(v^{(1)}, \ldots, v^{(m)}\right)$ belongs to the null-cone of A; that is, the vectors $v^{(1)}$, $\ldots$, $v^{(m)}$ satisfy the equations $(kl) = 0$. The subspace they span is then totally singular (i.e. $q = 0$ everywhere on the subspace).

Let W be a maximal totally singular subspace containing them. It follows from Witt's theorem that the dimension of W is $r = \lfloor n/2 \rfloor$, and that there exists a maximal totally singular subspace W_1 such that

$$\mathbb{K}^n = W \oplus W_1 \oplus \ker \beta.$$

For $0 \neq t \in \mathbb{K}$, let $g_t \in \mathrm{SO}_n(\mathbb{K})$ stand for the special orthogonal transformation that multiplies vectors in W by t, vectors in W_1 by $1/t$, and vectors in $\ker \beta$ by 1. Any $f \in \mathbb{K}[n \times m]^{\mathrm{SO}}$ is invariant under g_t, so

$$f\left(tv^{(1)}, \ldots, tv^{(m)}\right) = f\left(v^{(1)}, \ldots, v^{(m)}\right).$$

This holds for arbitrary $t \neq 0$, so it must also hold for $t = 0$. This implies that the point $\left(v^{(1)}, \ldots, v^{(m)}\right)$ is contained in the null-cone of $\mathbb{K}[n \times m]^{\mathrm{SO}}$. □

Corollary 1.3.2 *The algebras $\mathbb{F}[n \times m]^{\mathrm{SO}}$ and $\mathbb{F}[n \times m]^{\mathrm{O}_{\mathbb{F}}}$ are finitely generated as A–modules.*

Proof. Since A is of finite type over a field, it is Noetherian. Thus it will suffice to prove the Corollary for the larger algebra $\mathbb{F}[n \times m]^{\mathrm{SO}}$.

Now $G = \mathrm{SO}_n(\mathbb{F})$ is a reductive algebraic group, so Nagata's theorem [N, Theorem 3.4] says that $\mathbb{F}[n \times m]^G$ is finitely generated as an algebra.

Consider a homogeneous element $f \in \mathbb{F}[n \times m]^G$. By Theorem 1.3.1 and the Nullstellensatz, f has a power in the ideal of $\mathbb{F}[n \times m]$ generated by the (kl). It follows by [N, Lemma 3.4.2] that f has a power in the ideal of $\mathbb{F}[n \times m]^G$ generated by the (kl).

Applying that to each element f of a finite system of homogeneous generators of the algebra $\mathbb{F}[n \times m]^G$ shows that the ideal of $\mathbb{F}[n \times m]^G$ generated by the (kl) contains all elements of $\mathbb{F}[n \times m]^G$ that are homogeneous of high enough degree. So $\mathbb{F}[n \times m]^G$, as an A–module, is generated by elements of degree lower than some finite number. These elements form a finite-dimensional vector space, so a finite number of them will suffice. □

1.3.2 Algebro-geometric lemmas

We recall some well-known facts from algebraic geometry. The word 'variety' below stands for an irreducible affine algebraic variety over the arbitrary algebraically closed field $\mathbb{F}$. Write $\mathcal{O}(X)$ for the algebra of polynomial functions on X, and write $K(X)$ for the field of rational functions on X. Let

$\phi : X \to Y$ be a morphism of varieties. The morphism ϕ is said to be *dominant* if $\phi(X)$ is dense in Y. In this case, the comorphism ϕ^* identifies $K(Y)$ with the subfield $\phi^*K(Y)$ of $K(X)$. The morphism ϕ is said to be *separable* if $K(X) \geq \phi^*K(\phi(X))$ is a separable field extension. We need the following criterion, see for example [Bo, (17.3) Theorem]: The morphism ϕ is dominant and separable if and only if there is a non-singular point x on X such that $\phi(x)$ is non-singular in Y, and the differential $d_x\phi : T_xX \to T_{\phi(x)}Y$ at x is surjective.

Lemma 1.3.3 *Let $\phi : X \to Y$ be a dominant, separable morphism of varieties. Suppose that f is a rational function on X such that for some non-empty Zariski open subset U of X, the restriction $f|_U$ is regular and is constant along the fibers of $\phi|_U$. Then f is the pull-back of a rational function on Y, that is, $f \in \phi^*K(Y)$.*

Proof. We may assume that $X \setminus U$ is a hyper-surface in X. Then f is purely inseparable over $\phi^*K(Y)$ by [Bo, (18.2), p.78]; that is, f^{p^s} is contained in $\phi^*K(Y)$ for some natural number s. Thus f itself is contained in $\phi^*K(Y)$ because ϕ is separable by our assumption. □

More can be said when Y is normal. See for example [Bo, (18.3), p.79]:

Lemma 1.3.4 *Let $\phi : X \to Y$ be a surjective morphism of varieties, and assume that Y is normal. Suppose that h is a polynomial function on X, such that h is the pull-back of a rational function on Y, i.e., h is contained in $\phi^*K(Y)$. Then h is the pull-back of a polynomial function on Y, that is, $h \in \phi^*\mathcal{O}(Y)$.*

Proof. See for example [Bo, (18.3), p.79], and note that since we are dealing with affine varieties, 'regular functions' in the sense of [Bo] (i.e. everywhere defined rational functions) are the same as 'polynomial functions'. □

1.3.3 Rational invariants

Let $G = \mathrm{SO}_n(\mathbb{F})$ or $G = \mathrm{O}_n(\mathbb{F})$. We define $K = \mathbb{F}(n \times m)$ to be the fraction field of the polynomial algebra $R = \mathbb{F}[n \times m]$. The elements of K are called rational functions in $n \times m$ variables. Invariant rational functions are defined in the same manner as invariant polynomials. If $\mathrm{char}\ \mathbb{F} \neq 2$, a rational function is $\mathrm{O}_{2r+1}(\mathbb{F})$–invariant if and only if it is $\mathrm{SO}_{2r+1}(\mathbb{F})$–invariant

and *even* (i.e., contained in the field generated by homogeneous polynomials of even degree). We accept this as the definition of $\mathrm{O}_{2r+1,\mathbb{F}}$–invariance in arbitrary characteristic: for odd n, we define

$$K^{\mathrm{O}_{\mathbb{F}}} = K_0 \cap K^{\mathrm{SO}(\mathbb{F})},$$

where K_0 is the field of even rational functions.

We now look at the field K^G of invariant rational functions, which is much easier to deal with than the algebra R^G of invariant polynomials. Note that K^G is the fraction field of R^G. This follows easily from the fact that the group $\mathrm{SO}_n(\mathbb{F})$ is perfect, i.e., it is generated by commutators of its elements.

We shall use the notation $D = [1, \ldots, n]_n$ and $\Delta = \langle 1, \ldots, 2r \rangle_{2r}$. These are SO–invariants we have constructed in Section 1.2. We define

$$\langle 1, \ldots, 2r-1, i \rangle = \langle 1, \ldots, 2r-1, i \rangle_{2r} = \Delta\left(v^{(1)}, \ldots, v^{(2r-1)}, v^{(i)}\right).$$

Theorem 1.3.5 *(a) The field $\mathbb{F}(2r \times m)^{\mathrm{O}}$ is generated by the algebraically independent invariants*

$$(kl) \qquad (1 \le k \le l \le m, \quad k \le 2r).$$

(b) The field $\mathbb{F}((2r+1) \times m)^{\mathrm{SO}}$ is generated by the algebraically independent invariants

$$(kl) \qquad (1 \le k \le l \le m, \quad k \le 2r)$$

and

$$[1, \ldots, 2r, i] \qquad (2r+1 \le i \le m).$$

(c) The field $\mathbb{F}(2r \times m)^{\mathrm{SO}}$ is generated by the algebraically independent invariants

$$(kl) \qquad (1 \le k \le l \le m, \quad k \le 2r-1)$$

and

$$\langle 1, \ldots, 2r-1, i \rangle \qquad (2r \le i \le m).$$

(d) The field $\mathbb{F}((2r+1) \times m)^{\mathrm{O}_{\mathbb{F}}}$ is generated by the algebraically independent invariants

$$(kl) \qquad (1 \le k \le l \le m, \quad k \le 2r)$$

and

$$D \cdot [1, \ldots, 2r, i] \qquad (2r+1 \le i \le m).$$

(e) For $m < n$ we have $\mathbb{F}(n \times m)^{\mathrm{SO}} = \mathbb{F}(n \times m)^{\mathrm{O}_{\mathbb{F}}}$. For $m \geq n$, the field $\mathbb{F}(n \times m)^{\mathrm{SO}}$ is a quadratic extension of $\mathbb{F}(n \times m)^{\mathrm{O}_{\mathbb{F}}}$, generated for example by the invariant Δ if n is even and by D if n is odd.

The description in (e) will be made complete in Subsection 1.3.4 where we determine the quadratic polynomials over $\mathbb{F}(n \times m)^{\mathrm{O}_{\mathbb{F}}}$ that Δ and D satisfy.

For the proof, we introduce the following notion.

Definition 1.3.6 (a) A $2r \times m$ matrix is O–regular if the first $\min(m, 2r)$ columns are linearly independent. Set E to be the $2r \times m$ matrix with 1 on the main diagonal and zero elsewhere.

(b) A matrix $V \in \mathbb{F}^{(2r+1)\times m}$ is SO–regular if

- the images of the first $\min(m, 2r)$ columns are linearly independent in $\mathbb{F}^n / \ker \beta$, and
- there exists a non-singular vector which is β–orthogonal to the first $\min(m, 2r)$ columns.

Set E to be the $(2r+1) \times m$ matrix with 1 on the main diagonal, except for the position corresponding to the z coordinate, and zero elsewhere.

(c) A matrix $V \in \mathbb{F}^{2r \times m}$ is SO–regular if

- its first $\min(m, 2r-1)$ columns are linearly independent, and
- there exists a non-singular vector which is β–orthogonal to the first $\min(m, 2r-1)$ columns, and
- in the case $m \geq 2r$, there exists a vector $v \in \mathbb{F}^{2r}$ which is β–orthogonal to the first $2r-1$ columns and has

$$\Delta\left(v^{(1)}, \ldots, v^{(2r-1)}, v\right) \neq 0.$$

We let E be the leftmost $2r \times m$ submatrix of the $2r \times \infty$ matrix

$$\begin{pmatrix} \mathbf{1}_{2r-2} & 0 & 0 \cdots \\ 0 & 1 & 0 \cdots \\ 0 & 1 & 0 \cdots \end{pmatrix}.$$

Proposition 1.3.7 *In all three cases of Definition 1.3.6, regular matrices form a Zariski open set containing the point E.*

Proof. Openness follows easily from the fact that totally singular subspaces form a closed set in the Grassmannian.

Regularity of E is obvious for (a) and (b). For (c), observe that $v = (0, \dots, 0, 1, -1)^\top \in \mathbb{F}^{2r}$ is non-singular, is orthogonal to all columns of E, and has

$$\Delta\left(e^{(1)}, \dots, e^{(2r-1)}, v\right) =$$
$$= \frac{1}{2}\left(\mathrm{pf}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{\oplus(r-1)} \oplus \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right) - \det\left(\mathbf{1}_{r-2} \oplus \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\right)\right) = 1.$$

□

We shall need the following consequence of Witt's theorem.

Proposition 1.3.8 *(a) If $V', V'' \in \mathbb{F}^{2r\times m}$ are* O*–regular and satisfy*

$$(kl)(V') = (kl)(V'') \qquad (1 \le k \le l \le m, \quad k \le 2r),$$

then there exists a transformation $g \in \mathrm{O}_{2r}(\mathbb{F})$ such that $gV' = V''$.

(b) If $V', V'' \in \mathbb{F}^{(2r+1)\times m}$ are SO*–regular and satisfy*

$$\begin{aligned} (kl)(V') &= (kl)(V'') && (1 \le k \le l \le m, \quad k \le 2r), \\ [1, \cdots, 2r, i](V') &= [1, \cdots, 2r, i](V'') && (2r+1 \le i \le m), \end{aligned}$$

then there exists a special orthogonal transformation $g \in \mathrm{SO}_{2r+1}(\mathbb{F})$ such that $gV' = V''$.

(c) If $V', V'' \in \mathbb{F}^{2r\times m}$ are SO*–regular and satisfy*

$$\begin{aligned} (kl)(V') &= (kl)(V'') && (1 \le k \le l \le m, \quad k \le 2r-1), \\ \langle 1, \cdots, 2r-1, i\rangle(V') &= \langle 1, \cdots, 2r-1, i\rangle(V'') && (2r \le i \le m), \end{aligned}$$

then there exists a special orthogonal transformation $g \in \mathrm{SO}_{2r}(\mathbb{F})$ such that $gV' = V''$.

Proof. The first $\min(m, 2r)$ resp. $\min(m, 2r)$ resp. $\min(m, 2r-1)$ columns of V' and also of V'' always span a subspace W with $W \cap \ker \beta = 0$. Witt's theorem provides a $g \in M_n(\mathbb{F})$ that leaves q invariant and

$$gv^{(i)\prime} = v^{(i)\prime\prime} \tag{1.6}$$

for $1 \leq i \leq \min(m, 2r)$ resp. $1 \leq i \leq \min(m, 2r)$ resp. $1 \leq i \leq \min(m, 2r-1)$. In cases (b) resp. (c), there exists a non-singular vector u orthogonal to ${v^{(i)}}'$ for $i = 1, \ldots, 2r$ resp. $i = 1, \ldots, 2r-1$. The reflection g_u defined in (1.1) fixes these ${v^{(i)}}'$. So, by replacing g with gg_u if necessary, we have a $g \in \mathrm{SO}$ such that (1.6) holds for the same indices i as said before.

We need to show that (1.6) also holds for $2r < i \leq m$ resp. $2r < i \leq m$ resp. $2r - 1 < i \leq m$. Let us assume that m is large enough for such i to exist.

(a) As β is non-degenerate and ${v^{(1)}}'', \ldots, {v^{(2r)}}''$ is a basis of $\mathbb{F}^{2r}$, it suffices to show that

$$\beta\left({v^{(k)}}'', gv^{(i)\prime}\right) = \beta\left({v^{(k)}}'', {v^{(i)}}''\right) \tag{1.7}$$

for $1 \leq i \leq m$ and $1 \leq k \leq 2r$. This is equivalent to

$$\beta\left(gv^{(k)\prime}, gv^{(i)\prime}\right) = \beta\left({v^{(k)}}', {v^{(i)}}'\right),$$

which follows from $g \in \mathrm{O}_{2r}(\mathbb{F})$.

(b) Equality (1.7) is proved as above, and shows that the difference

$$gv^{(i)\prime} - {v^{(i)}}'' \tag{1.8}$$

is orthogonal to the vectors ${v^{(k)}}''$ for $1 \leq k \leq 2r$, but it is also spanned by them because of

$$\begin{aligned}\det\left[{v^{(1)}}'', \ldots, {v^{(2r)}}'', gv^{(i)\prime}\right] = \det\left[gv^{(1)\prime}, \ldots, gv^{(2r)\prime}, gv^{(i)\prime}\right] \overset{*}{=} \\ \overset{*}{=} \det\left[{v^{(1)}}', \ldots, {v^{(2r)}}', {v^{(i)}}'\right] = \det\left[{v^{(1)}}'', \ldots, {v^{(2r)}}'', {v^{(i)}}''\right],\end{aligned}$$

so the difference (1.8) is zero. Note that equality $*$ above follows from $g \in \mathrm{SO}_{2r+1}(\mathbb{F})$.

(c) We can proceed as in (b). We now have equality (1.7) for $1 \leq i \leq m$ and $1 \leq k \leq 2r-1$. Thus the difference (1.8) is orthogonal to the vectors

$v^{(k)\prime\prime}$ for $1 \leq k \leq 2r-1$, therefore, it is a scalar multiple of the vector v provided by the last condition in Definition 1.3.6.(c). We also have

$$\Delta\left(v^{(1)\prime\prime}, \ldots, v^{(2r-1)\prime\prime}, gv^{(i)\prime}\right) = \Delta\left(gv^{(1)\prime}, \ldots, gv^{(2r-1)\prime}, gv^{(i)\prime}\right) \stackrel{*}{=}$$
$$\stackrel{*}{=} \Delta\left(v^{(1)\prime}, \ldots, v^{(2r-1)\prime}, v^{(i)\prime}\right) = \Delta\left(v^{(1)\prime\prime}, \ldots, v^{(2r-1)\prime\prime}, v^{(i)\prime\prime}\right),$$

so the difference (1.8) is zero. Note that equality $*$ above follows from $g \in \mathrm{SO}_{2r}(\mathbb{F})$. □

Proof of Theorem 1.3.5. (a), (b) and (c): Write ϕ for the regular map defined on $\mathbb{F}^{n\times m}$ that has the invariants in the theorem as its coordinates. We show that ϕ is dominant and separable. We claim that the differential of ϕ at the point E given in Definition 1.3.6 is onto. The partial derivatives are as follows.

$$\frac{\partial(kl)}{\partial x_j^{(k)}} = y_j^{(l)}, \qquad \frac{\partial(kl)}{\partial y_j^{(k)}} = x_j^{(l)}, \qquad \frac{\partial(kl)}{\partial x_j^{(l)}} = y_j^{(k)}, \qquad \frac{\partial(kl)}{\partial y_j^{(l)}} = x_j^{(k)}, \tag{1.9}$$

and in case (b),

$$\frac{\partial(kl)}{\partial z^{(l)}} = 2z^{(k)}, \qquad \frac{\partial(kl)}{\partial z^{(k)}} = 2z^{(l)}. \tag{1.10}$$

All other partial derivatives of (kl) are zero.

For case (b), we observe that $\partial(kl)/\partial z^{(i)} = 0$ at E for all k, l and i.

For case (c), we observe that

$$\left(\frac{\partial}{\partial x_r^{(i)}} - \frac{\partial}{\partial y_r^{(i)}}\right)(kl) = 0$$

at E for all k, l and i.

In all three cases, the $n \times m$ matrix formed by the partial derivatives at E of (kl) has $e^{(k)}$ with x and y coordinates interchanged as its l–th column and has $e^{(l)}$ with x and y coordinates interchanged as its k–th column, all other columns being zero. We easily see that all these $n \times m$ matrices are linearly independent. Our claim follows in case (a). For (b), it suffices to prove that the $1 \times m$ matrices formed by the partial derivatives at E of each

$[1, \ldots, 2r, i]$ with respect to the variables $z^{(i')}$ are linearly independent. This is obvious, since at E,

$$\frac{\partial [1, \ldots, 2r, i]}{\partial z^{(i')}} = \delta^{(i)}_{(i')}$$

for $2r + 1 \leq i, i' \leq m$. For (c), we use the fact that at E,

$$\left(\frac{\partial}{\partial x_r^{(i')}} - \frac{\partial}{\partial y_r^{(i')}} \right) \langle 1, \ldots, 2r-1, i \rangle = \delta^{(i)}_{(i')}$$

for $2r \leq i, i' \leq m$.

Now, to prove (a), (b) or (c), suppose that $f \in \mathbb{F}(n \times m)^G$, where G is the group in the statement we want to prove. We think of f as a rational function on $\mathbb{F}^{n \times m}$. Then f is constant along the orbits of G, so Proposition 1.3.8 shows that f is constant along the fibers of ϕ (at least on some non-empty open set). By Lemma 1.3.3, f is the pull-back of a rational function.

The statements (d) and (e) easily follow; note that

- for $n = 2r + 1$, D is not O–invariant because it is of odd degree;
- for $n = 2r$, Δ is not O–invariant by Fact 1.2.9.

□

1.3.4 The case $m \leq n$

In this subsection, we describe the invariant algebras $\mathbb{F}[n \times m]^{\mathrm{SO}}$ and $\mathbb{F}[n \times m]^{\mathrm{O}}$ for $m \leq n$. We shall find that their behaviour in characteristic 2 is essentially the same as in all other characteristics, i.e., there are no 'exotic' invariant polynomials for $m \leq n$.

This will be very easy for $m < n$, but to treat the case $m = n$, we shall need a few simple lemmas. In these, we think of the symbols (kl) for $1 \leq k \leq l \leq m$ as independent variables, i.e., we temporarily forget their definition given in Subsection 1.1.3. We define $(lk) = (kl)$.

Lemma 1.3.9 *Suppose that i_1, …, $i_{2\mu}$ are distinct. Then the polynomial $Pf(i_1, \ldots, i_{2\mu})$, as defined by formula (1.4) in Section 1.2, is irreducible in the polynomial ring $\mathbb{F}[(kl) : 1 \leq k \leq l \leq m]$.*

Note that $Pf(i_1, \ldots, i_{2\mu})$ has integer coefficients (see the Preface), so it is defined also over $\mathbb{F}$.

Proof. We proceed by induction on μ. We have $Pf(i_1 i_2) = (i_1 i_2)$, which is irreducible. We shall prove the statement for $\mu > 1$ assuming that it is true for $\mu - 1$. We may assume that $i_1 = 1$, $\ldots$, $i_{2\mu} = 2\mu$. Observe that $Pf(1, \ldots, 2\mu)$ is a polynomial of degree 1 in $(2\mu - 1, 2\mu)$, with leading coefficient $Pf(1, \ldots, 2\mu - 2)$, which is irreducible. We are done unless $Pf(1, \ldots, 2\mu-2)$ divides $Pf(1, \ldots, 2\mu)$. In fact, using this argument for other pairs of indices, we are done unless $Pf(i_1, \ldots, i_{2\mu-2})$ divides $Pf(1, \ldots, 2\mu)$ for all $1 \leq i_1 < \cdots < i_{2\mu-2} \leq 2\mu$. As these divisors are essentially distinct irreducible polynomials, we would then get the degree inequality

$$\binom{2\mu}{2}(2\mu - 2) \leq 2\mu,$$

which is absurd. □

We define the *Gramian*

$$Gr(i_1, \ldots, i_m) = \begin{vmatrix} 2(i_1 i_1) & (i_1 i_2) & \cdots & (i_1 i_m) \\ (i_2 i_1) & 2(i_2 i_2) & \cdots & (i_2 i_m) \\ \vdots & \vdots & \ddots & \vdots \\ (i_m i_1) & (i_m i_2) & \cdots & 2(i_m i_m) \end{vmatrix}.$$

For Gramians of odd size, we have

Lemma 1.3.10

$$Gr(i_1, \ldots, i_{2\mu+1}) \in 2\mathbb{Z}[(kl) : 1 \leq k \leq l \leq m].$$

Proof. Each expansion term in the determinant either has a factor from the diagonal and therefore has an even coefficient, or is a product of off-diagonal entries and can be paired with the transposed term; note that $(kl) = (lk)$. □

Lemma 1.3.11 *(a) Assume that all indices i_1, $\ldots$, $i_{2\mu+1}$ are distinct. Then the image of the polynomial*

$$Gr(i_1, \ldots, i_{2\mu+1})/2 \in \mathbb{Z}[(kl) : 1 \leq k \leq l \leq m]$$

is irreducible in the polynomial ring $\mathbb{F}[(kl) : 1 \leq k \leq l \leq m]$.

(b) Assume that char $\mathbb{F} \neq 2$ *and all indices* $i_1, \ldots, i_{2\mu}$ *are distinct. Then the polynomial* $Gr(i_1, \ldots, i_{2\mu})$ *is irreducible in the polynomial ring* $\mathbb{F}[(kl) : 1 \leq k \leq l \leq m]$.

Proof. We proceed by induction on μ. Case (a) for $\mu = 0$ is obvious. Let us assume (a) for $\mu - 1$ and prove (b) for μ. We may assume that $i_1 = 1$, $i_2 = 2$ etc. The polynomial $Gr(1, \ldots, 2\mu)$ has degree 1 in the variable $(2\mu, 2\mu)$, with leading coefficient $2Gr(1, \ldots, 2\mu - 1)$. Since char $\mathbb{F} \neq 2$, this leading coefficient is irreducible by the induction hypothesis. We are done unless $Gr(1, \ldots, 2\mu - 1)$ divides $Gr(1, \ldots, 2\mu)$. In fact, we are done unless $Gr(i_1, \ldots, i_{2\mu-1})$ divides $Gr(1, \ldots, 2\mu)$ for all $1 \leq i_1 < \cdots < i_{2\mu-1} \leq 2\mu$, which leads to $2\mu(2\mu - 1) \leq 2\mu$, thus $\mu = 1$ and $Gr(12)$ is a scalar multiple of $(11)(22)$, which is absurd.

Now let us assume (b) for μ and prove (a) for μ. We may assume that $i_1 = 1$, $i_2 = 2$ etc. The polynomial $Gr(1, \ldots, 2\mu + 1)/2$ has degree 1 in the variable $(2\mu + 1, 2\mu + 1)$, with leading coefficient $Gr(1, \ldots, 2\mu)$.

When char $\mathbb{F} \neq 2$, this leading coefficient is irreducible by the induction hypothesis. We are done unless $Gr(1, \ldots, 2\mu)$ divides $Gr(1, \ldots, 2\mu + 1)/2$. In fact, we are done unless $Gr(i_1, \ldots, i_{2\mu})$ divides $Gr(1, \ldots, 2\mu + 1)$ for all $1 \leq i_1 < \cdots < i_{2\mu} \leq 2\mu + 1$, which leads to $(2\mu + 1)2\mu \leq 2\mu + 1$, thus $\mu = 0$, in which case $Gr(1)/2 = (11)$ is indeed irreducible.

When $\mathbb{F}$ is of characteristic 2, we have

$$Gr(1, \ldots, 2\mu) = \begin{vmatrix} 2(11) & (12) & \cdots & (1, 2\mu) \\ (21) & 2(22) & \cdots & (2, 2\mu) \\ \vdots & \vdots & \ddots & \vdots \\ (2\mu, 1) & (2\mu, 2) & \cdots & 2(2\mu, 2\mu) \end{vmatrix} =$$

$$= \begin{vmatrix} 0 & (12) & \cdots & (1, 2\mu) \\ -(21) & 0 & \cdots & (2, 2\mu) \\ \vdots & \vdots & \ddots & \vdots \\ -(2\mu, 1) & -(2\mu, 2) & \cdots & 0 \end{vmatrix} = Pf(1, \ldots, 2\mu)^2.$$

Using Lemma 1.3.9, we are done unless $Pf(i_1, \ldots, i_{2\mu})$ divides $Gr(1, \ldots, 2\mu + 1)$ for all $1 \leq i_1 < \cdots < i_{2\mu} \leq 2\mu + 1$, which leads to $(2\mu + 1)\mu \leq 2\mu + 1$, thus $\mu = 1$ and $Gr(1, 2, 3)$ is a scalar multiple of $(12)(23)(31)$, which is absurd. □

For $m = n$ odd, we now introduce an additional independent variable D and we define

$$P = (-1)^r D^2 - Gr(1, \ldots, n)/2 \in \mathbb{Z}[D, (kl) : 1 \leq k \leq l \leq n]. \qquad (1.11)$$

The following proposition deals with the hyper-surface $\{P = 0\}$ in the vector space $\mathbb{F}^{\binom{n+1}{2}+1}$ where coordinates will be denoted by D and by (kl) with $1 \leq k \leq l \leq n$.

Lemma 1.3.12 *The hyper-surface $\{P = 0\}$ in $\mathbb{F}^{\binom{n+1}{2}+1}$ is normal.*

Proof. A hyper-surface H (the zero locus of a single polynomial in an affine space) is normal if and only if the set of singular points has codimension ≥ 2 in H; this follows for example from Seidenberg's criterion for normality [Se, Theorem 3].

Calculate

$$\frac{\partial P}{\partial (kk)} = -Gr(1, \ldots, \hat{k}, \ldots, n),$$

where the hatted index is omitted. This derivative does not involve the variable D. We claim that the locus of common zeros of the three polynomials P, $\partial P/\partial(11) = -Gr(2, \ldots, n)$ and $\partial P/\partial(nn) = -Gr(1, \ldots, n-1)$ is of codimension 3 in $\mathbb{F}^{\binom{n+1}{2}+1}$. Equivalently, the locus of common zeros of $\partial P/\partial(11)$ and $\partial P/\partial(nn)$ is of codimension 2 in the hyperplane $\{D = 0\}$. To see this equivalence note that projection from the direction of the D coordinate axis defines a finite-to-one map from the hyper-surface $\{P = 0\}$ onto the hyperplane $\{D = 0\}$. The vanishing of $\partial P/\partial(11)$ and $\partial P/\partial(nn)$ on a common hyper-surface in the affine space $\{D = 0\}$ would mean having the defining polynomial of that hyper-surface as a common factor. But, by Lemma 1.3.11, these two polynomials are essentially distinct irreducibles if char $\mathbb{F} \neq 2$. If $\mathbb{F} = \mathbb{K}$ is of characteristic 2, they are negative squares of essentially distinct irreducibles (namely, Pfaffians) by Lemma 1.3.9. They have no common factors in either case. So the locus of common zeros of P, $\partial P/\partial(11)$ and $\partial P/\partial(nn)$ is of codimension 3. The singular locus of $\{P = 0\}$ is contained in that locus, so it has codimension ≥ 2 in $\{P = 0\}$, which is therefore normal. □

We now return to the original interpretation where the (kl) and D are elements of $\mathbb{Z}[n \times m]$, and therefore also of $\mathbb{F}[n \times m]$. The significance of the expression P discussed above is given by

Proposition 1.3.13 *For odd n, the polynomials (kl) and D satisfy the relation $P = 0$ over $\mathbb{Z}$ and thus also over $\mathbb{F}$.*

Proof. Working over $\mathbb{Q}$, the matrix of the polar form β of the quadratic form

$$q = x_1y_1 + \cdots + x_ry_r + z^2$$

is

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \oplus (2).$$

For arbitrary $V \in \mathbb{Q}^{n\times n}$ with ith column $v^{(k)}$, we have

$$V^TMV = \left(\beta\left(v^{(k)}, v^{(l)}\right)\right)_{k,l=1}^{n}.$$

Taking determinants gives

$$(-1)^r \cdot 2 \cdot (\det V)^2 = \det\left(\beta\left(v^{(k)}, v^{(l)}\right)\right)_{k,l=1}^{n}.$$

The proposition follows, since $\beta\left(v^{(k)}, v^{(k)}\right) = 2q\left(v^{(k)}\right)$. □

We will also need the following propositions. Recall that a matrix B is defined to be *alternating* if it is anti-symmetric (i.e., in characteristic 2, symmetric), and all its diagonal elements are zero. It is well known that then B is cogredient to

$$J \oplus 0 = \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix} \oplus 0$$

with J of size equal to the rank of B (which is thus always even).

Let

$$\omega(v', v'') = \sum_{i=1}^{r} (x_i'y_i'' - y_i'x_i'')$$

be the standard alternating bilinear form on $\mathbb{F}^n$. It is non-degenerate if and only if n is even. When $n = 2r$, the symplectic group $\mathrm{Sp}_{2r}(\mathbb{F})$ is defined to be the group of linear transformations g of $\mathbb{F}^{2r}$ that leave ω invariant.

Proposition 1.3.14 *Let n and m be any positive integers, and let $B = \left(\beta^{(kl)}\right)$ be any alternating $m \times m$ matrix of rank $\leq n$. Then there exist vectors $u^{(1)}, \ldots, u^{(m)} \in \mathbb{F}^n$ with*

$$\omega\left(u^{(k)}, u^{(l)}\right) = \beta^{(kl)} \qquad (k, l = 1, \ldots, m).$$

Proof. We may assume (by performing a linear change of basis) that

$$\left(\beta^{(kl)}\right) = J \oplus 0 = \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix} \oplus 0$$

with J of size $\leq n$. The proposition obviously holds in this case, and the general case follows. □

Proposition 1.3.15 *Let $m \leq n$. Then the map*

$$((kl) : 1 \leq k \leq l \leq m) : \mathbb{F}^{n \times m} \to \mathbb{F}^{\binom{m+1}{2}}$$

is surjective.

Proof. This is well known in characteristic $\neq 2$, so assume that $\mathbb{F} = \mathbb{K}$ is of characteristic 2. Then $\beta = \omega$.

Let $\left(\beta^{(kl)}\right)$ be any $m \times m$ alternating matrix over $\mathbb{K}$, and let $q^{(1)}, \ldots, q^{(m)} \in \mathbb{K}$. We must prove that there exist vectors $v^{(1)}, \ldots, v^{(m)} \in \mathbb{K}^n$ with

$$\beta\left(v^{(k)}, v^{(l)}\right) = \beta^{(kl)} \qquad (k, l = 1, \ldots, m)$$

and

$$q\left(v^{(k)}\right) = q^{(k)} \qquad (k = 1, \ldots, m).$$

As always, we set $r = \lfloor n/2 \rfloor$. Choose vectors $u^{(1)}, \ldots, u^{(m)} \in \mathbb{K}^{2r}$ as in the previous proposition.

Consider $n = 2r+1$ first. Note that the standard quadratic form q is onto $\mathbb{K}$ on any line parallel to $\ker \beta$ (the z-axis). Therefore, there exist vectors $v^{(k)} \in \mathbb{K}^{2r+1}$ that are mapped to the $u^{(k)}$ by the projection

$$\mathbb{K}^{2r+1} \to \mathbb{K}^{2r+1} / \ker \beta = \mathbb{K}^{2r}$$

and have $q\left(v^{(k)}\right) = q^{(k)}$.

Now let $n = 2r$. First suppose that $m = n$ and $u^{(1)}, \ldots, u^{(m)}$ is a basis of $\mathbb{K}^n$. Define a new quadratic form q^* by the formula

$$q^*\left(\sum_{i=1}^{m} \lambda_i u^{(k)}\right) = \sum_{i=1}^{m} \lambda_i^2 q^{(k)} + \sum_{1 \leq i < l \leq m} \lambda_i \lambda_l \beta^{(kl)}.$$

Let β^* stand for the polar form of q^*. Then

$$\beta^*\left(u^{(k)}, u^{(l)}\right) = q^*\left(u^{(k)} + u^{(l)}\right) - q^*\left(u^{(k)}\right) - q^*\left(u^{(l)}\right) = \beta^{(kl)} = \beta\left(u^{(k)}, u^{(l)}\right),$$

therefore, $\beta^* = \beta$. It follows that q^* is non-degenerate. Since $\mathbb{K}$ is algebraically closed, all non-degenerate quadratic forms are equivalent. So there is a linear isomorphism $g : \mathbb{K}^n \to \mathbb{K}^n$ such that $q(gu) = q^*(u)$ for all $u \in \mathbb{K}^n$. It of course follows that

$$\beta(gu', gu'') = \beta^*(u', u'') = \beta(u', u'')$$

for all $u', u'' \in \mathbb{K}^n$. That is, $g \in \mathrm{Sp}_n(\mathbb{K})$. Define $v^{(k)} = gu^{(k)}$ $(k = 1, \ldots, m)$. Then

$$\beta\left(v^{(k)}, v^{(l)}\right) = \beta\left(u^{(k)}, u^{(l)}\right) = \beta^{(kl)}$$

and

$$q\left(v^{(k)}\right) = q^*\left(u^{(k)}\right) = q^{(k)},$$

i. e., $v^{(1)}, \ldots, v^{(m)}$ have the desired properties.

Suppose finally that $n = 2r$ but $u^{(1)}, \ldots, u^{(m)}$ do not span $\mathbb{K}^n$. Choose some non-zero vector $0 \neq u^{(0)}$ orthogonal to each of $u^{(1)}, \ldots, u^{(m)}$. Choose a linear function $f : \mathbb{K}^n \to \mathbb{K}$ with $f\left(u^{(0)}\right) \neq 0$. Define the new quadratic form q^* by the formula

$$q^* = q + \lambda f^2,$$

with some $\lambda \in \mathbb{K}$ that gives $q^*\left(u^{(0)}\right) \neq 0$. The quadratic form f^2 has 0 as its polar form, so q^* has β as its polar form. It follows that q^* is non-degenerate. We therefore have a linear isomorphism $g : \mathbb{K}^n \to \mathbb{K}^n$ such that $q(gu) = q^*(u)$ for all $u \in \mathbb{K}^n$. Of course $g \in \mathrm{Sp}_n(\mathbb{K})$. The vectors $gu^{(k)}$ have

$$\beta\left(gu^{(k)}, gu^{(l)}\right) = \beta\left(u^{(k)}, u^{(l)}\right) = \beta^{(kl)}.$$

Note also that $gu^{(0)}$ is β–orthogonal to each of $gu^{(1)}, \ldots, gu^{(m)}$ and that $q\left(gu^{(0)}\right) \neq 0$. The latter ensures that q is onto $\mathbb{K}$ on any line parallel to $\mathbb{K}gu^{(0)}$. So there are vectors $v^{(k)} \in gu^{(k)} + \mathbb{K}gu^{(0)}$ with $q\left(v^{(k)}\right) = q^{(k)}$. They have all desired properties. □

Theorem 1.3.16 *For $m \leq n$, the algebra $\mathbb{F}[n \times m]^{\mathrm{O}_\mathbb{F}}$ is generated by the $\binom{m+1}{2}$ algebraically independent invariant polynomials (kl). When $m < n$, we have $\mathbb{F}[n \times m]^{\mathrm{SO}} = \mathbb{F}[n \times m]^{\mathrm{O}_\mathbb{F}}$.*

Proof. The second claim follows from Theorem 1.3.5(e). We prove the first claim. Let $\phi : \mathbb{F}^{n \times m} \to \mathbb{F}^{\binom{m+1}{2}}$ stand for the regular map that has the (kl) as its coordinates. Choose any element $f \in \mathbb{F}[n \times m]^{\mathrm{O}_\mathbb{F}}$. Theorem 1.3.5 says that

f is a rational function in the (kl) — in the case $m = n = 2r + 1$, note that D^2 is a polynomial in the (kl) by Proposition 1.3.13. So f is the pull-back under ϕ of a rational function. But f is a polynomial, and Proposition 1.3.15 says that ϕ is surjective. By Lemma 1.3.4, f is the pull-back of a polynomial function. □

Theorem 1.3.17 *Let $m = n = 2r + 1$. Let D stand for $[1, \dots, n]$. Then the algebra $\mathbb{F}[n \times m]^{\mathrm{SO}}$ is generated by the $\binom{n+1}{2} + 1$ invariant polynomials (kl) and D, the ideal of algebraic relations between whom is generated by the single element P defined by formula* (1.11).

Proof. Consider the map

$$\phi : \mathbb{F}^{n \times n} \to \mathbb{F}^{\binom{n+1}{2}+1}$$

that has the (kk), the (kl), and D as its coordinates. It follows from Propositions 1.3.13 and 1.3.15 that the image of ϕ is the hyper-surface $\{P = 0\}$.

Choose any $f \in \mathbb{F}[n \times n]^{\mathrm{SO}}$. Theorem 1.3.5 says that f is the pull-back of a rational function on $\{P = 0\}$. But f is a polynomial, so, by Lemma 1.3.4 and Proposition 1.3.12, f is the pull-back of a polynomial. □

We now turn to the description of the algebra of special orthogonal invariants in the case $m = n = 2r$. We write $Pf = Pf(1, \dots, 2r)$, $D = [1, \dots, 2r]$, $\Delta = \langle 1, \dots, 2r \rangle = (Pf - D)/2$ and $\bar{\Delta} = (Pf + D)/2$.

Theorem 1.3.18 *Let $m = n = 2r$. Then the algebra $\mathbb{F}[n \times m]^{\mathrm{SO}}$ is generated by the $\binom{n+1}{2} + 1$ invariants (kl) and Δ, the ideal of algebraic relations between whom is generated by the single element Π defined as*

$$\Pi = \Delta^2 - \Delta Pf + \left(Pf^2 - (-1)^r Gr(1, \dots, n)\right)/4. \tag{1.12}$$

(See the proof for the meaning of 1/4 here.)

Proof. The proof is rather similar to that of Theorem 1.3.17. Write L for the difference in parentheses in (1.12) above. Then L is a polynomial in the (kl) with integral coefficients.

First interpret the (kl) and Δ as polynomials over $\mathbb{Z}$ in the x, y variables. We have

$$\Delta\bar{\Delta} = \left(Pf^2 - D^2\right)/4 = L/4,$$

where the second equality is proved in the same manner as Proposition 1.3.13. Note that $\Delta + \bar{\Delta} = Pf$. It follows that the (kl) and Δ (considered as polynomials over $\mathbb{Z}$ in the x, y variables) satisfy the relation

$$\Delta^2 - \Delta Pf + L/4 = 0. \tag{1.13}$$

We claim that the coefficients of L are divisible by four, so $L/4$ is a polynomial in the variables (kl) with integer coefficients. Indeed, multiply the relation $\Delta\bar{\Delta} = L/4$ by 4 and consider it modulo 2: the left hand side becomes zero, so we obtain on the right hand side an algebraic relation over $\mathbb{K}$ holding between the (kl) (defined over $\mathbb{K}$). But the (kl) are algebraically independent elements of $\mathbb{K}[n \times m]^{\mathrm{O}}$ by Theorem 1.3.5, so this relation must be trivial. This means that all coefficients of L (as a polynomial in the (kl)) are even. Taking now the relation $2\Delta\bar{\Delta} = L/2$ modulo 2 and repeating the same argument we obtain our claim. So (1.13) is an algebraic relation with integral coefficients holding between the (kl) and Δ (considered as polynomials over $\mathbb{Z}$ in the x, y variables).

It follows immediately that (1.13) makes sense and holds as a relation over $\mathbb{F}$; that is, the relation $\Pi = 0$ makes sense and holds in $\mathbb{F}[n \times m]^{\mathrm{SO}}$.

Consider now the map

$$\phi : \mathbb{F}^{n \times n} \to \mathbb{F}^{\binom{n+1}{2}+1}$$

that has the (kl) and Δ as its coordinates. It follows from the relation $\Pi = 0$ and Proposition 1.3.15 that the image of ϕ is the hyper-surface $\{\Pi = 0\}$ in $\mathbb{F}^{\binom{n+1}{2}+1}$. (For surjectivity, we also need that the coset $\mathrm{O}_{2r}(\mathbb{F})\backslash\mathrm{SO}_{2r}(\mathbb{F})$ interchanges Δ and $\bar{\Delta}$, so a point with coordinates (kl) and Δ is in the image of ϕ if and only if the point with coordinates (kl) and $\bar{\Delta}$ is in the image of ϕ.) Choose any $f \in \mathbb{F}[n \times n]^{\mathrm{SO}}$. Theorem 1.3.5 says that f is the pull-back of a rational function on the hyper-surface $\{\Pi = 0\}$. But f is a polynomial, so, by Lemma 1.3.4 and Proposition 1.3.19 below, f is the pull-back of a polynomial. □

Lemma 1.3.19 *Consider the vector space $\mathbb{F}^{\binom{n+1}{2}+1}$, with coordinates denoted by Δ and (kl) $(1 \leq k \leq l \leq n)$. Then the hyper-surface $\{\Pi = 0\}$ in $\mathbb{F}^{\binom{n+1}{2}+1}$ is normal.*

Proof. Just as in Lemma 1.3.12, it suffices to prove that the singular locus has codimension ≥ 2 in the hyper-surface.

Calculate

$$\frac{\partial \Pi}{\partial (kk)} = (-1)^{r+1} Gr(1, \ldots, \hat{k}, \ldots, n)/2,$$

which does not involve Δ.

We claim that the locus of common zeros of Π, $\partial\Pi/\partial(11)$ and $\partial\Pi/\partial(nn)$ is of codimension 3 in $\mathbb{F}^{\binom{n+1}{2}+1}$. Equivalently, the locus of common zeros of $\partial\Pi/\partial(11)$ and $\partial\Pi/\partial(nn)$ is of codimension 2 in the hyperplane $\{\Delta = 0\}$. It suffices to show that $\partial\Pi/\partial(11)$ and $\partial\Pi/\partial(nn)$ have no common factors as polynomials in the (kl). This follows from Lemma 1.3.11.

So the locus of common zeros of Π, $\partial\Pi/\partial(11)$ and $\partial\Pi/\partial(nn)$ is of codimension 3. The singular locus of $\{\Pi = 0\}$ is contained in that locus, so it has codimension ≥ 2 in $\{\Pi = 0\}$, which therefore is normal. $\square$

1.4 Proofs of indecomposability

In this section we shall prove that the high degree multi-linear invariants $(1, \ldots, 2\mu)$, $[1, \ldots, m]$ and $\langle 1, \ldots, 2\mu \rangle$ defined in Section 1.2 are indecomposable over the algebraically closed field $\mathbb{K}$ of characteristic 2.

First let $n = 2r + 1$. Recall from Section 1.2 that we have defined the *multiplicity in x, y* of a polynomial to be the infimum of the total degrees of its monomials in the x, y variables.

Lemma 1.4.1 *The multiplicity in x, y of any m–linear $\mathrm{SO}_{2r+1}(\mathbb{K})$–invariant is at least* $\min(m, 2r)$.

Proof. For $m \leq 2r$ we have Theorem 1.3.16 that says that the algebra $\mathbb{K}[(2r+1) \times m]^{\mathrm{SO}_{2r+1}(\mathbb{K})}$ is generated by the (kl). Multi-linear invariants are spanned by products of the (kl) with $k \neq l$, which do not involve the z variables in characteristic 2, so the multiplicity in x, y of any m–linear invariant is m. The case $m \geq 2r$ is reduced to this as follows.

Indirectly assume that an m–linear $\mathrm{SO}_{2r+1}(\mathbb{K})$–invariant has a monomial of the form

$$*_{j_1}^{(1)} \cdots *_{j_d}^{(d)} z^{(d+1)} \cdots z^{(m)}$$

with $0 \leq d < 2r$. Recall that the z axis in the space $\mathbb{K}^{2r+1}$ is the kernel of the symmetric bilinear form β, so its unit vector e is stabilized by $\mathrm{SO}_{2r+1}(\mathbb{K})$. It follows that substituting e for the vector variables $v^{(d+2)}, \ldots, v^{(m)}$ in our

$\mathrm{SO}_{2r+1}(\mathbb{K})$–invariant yields a $(d+1)$–linear $\mathrm{SO}_{2r+1}(\mathbb{K})$–invariant, having a monomial of the form

$$*_{j_1}^{(1)} \cdots *_{j_d}^{(d)} z^{(d+1)}.$$

But $d+1 \leq 2r$, so we have a contradiction. □

Since the multiplicity in x, y of a product of homogeneous polynomials is the sum of the multiplicities of the factors, it follows for $r \geq 1$ that the multiplicity in x, y of any decomposable m–linear invariant with $m > 2r$ is strictly greater than $2r$, so we get

Theorem 1.4.2 *Let $n \geq 3$ be odd and $m \geq n$. Then the m–linear $\mathrm{SO}_n(\mathbb{K})$–invariant $(1, \ldots, m)_n$ or $[1, \ldots, m]_n$ (whichever one is defined) is indecomposable.*

Proof. This follows from Fact 1.2.6. □

Corollary 1.4.3 *Let $n \geq 3$ be odd and $m \geq n$. Then the m–linear invariant $(1, \ldots, m)$ or $[1, \ldots, m]$ (whichever one is defined) is indecomposable in the ring $\mathbb{Z}[n \times m]^{\mathrm{SO}_n(\mathbb{C})}$. When m is even, $(1, \ldots, m)$ is therefore indecomposable also in the ring $\mathbb{Z}[n \times m]^{\mathrm{O}_n(\mathbb{C})}$.*

Another, independent proof of Theorem 1.4.2 for $m > n$ will be given in a remark following the discussion of the O_{2r} case. Note that the case $m = n$ is immediate from Theorem 1.3.16 cited in the proof above.

We now turn to the case $n = 2r$. We observe that for $\mu \geq r$, the $\mathrm{O}_{2r}(\mathbb{K})$–invariant $(1, \ldots, 2\mu)_{2r}$ has the smallest possible multiplicity in s. Indeed, we have

Lemma 1.4.4 *The multiplicity in s of any 2μ–linear $\mathrm{O}_{2r}(\mathbb{K})$–invariant $f = f\left(v^{(1)}, \ldots, v^{(2\mu)}\right)$ is at least $\min(\mu, r)$.*

Proof. For $\mu \leq r$ we have Theorem 1.3.16 which tells us that the algebra $\mathbb{K}[2r \times 2\mu]^{\mathrm{O}}$ is generated by the (kk), which are quadratic in the corresponding $v^{(k)}$, and by the

$$(kl) = \sum_{j=1}^{r} \left(t_j^{(k)} \left(t_j^{(l)} + s_j^{(l)} \right) + \left(t_j^{(k)} + s_j^{(k)} \right) t_j^{(l)} \right) =$$

$$= \sum_{j=1}^{r} \left(t_j^{(k)} s_j^{(l)} + s_j^{(k)} t_j^{(l)} \right) \qquad (k < l),$$

which are bilinear in $v^{(k)}$, $v^{(l)}$ and have multiplicity 1 in s, so the multiplicity in s of any 2μ–linear invariant f is μ.

Now let $\mu \geq r$. It suffices to prove that if $0 \leq d < r$, then the coefficient in f of the monomial

$$s_{j_1}^{(1)} \cdots s_{j_d}^{(d)} t_{j_{d+1}}^{(d+1)} \cdots t_{j_{2\mu}}^{(2\mu)} \tag{1.14}$$

is zero. This coefficient, expressed using the original x, y coordinates, is the sum of the coefficients in f of the $2^{2\mu-d}$ monomials

$$y_{j_1}^{(1)} \cdots y_{j_d}^{(d)} *_{j_{d+1}}^{(d+1)} \cdots *_{j_{2\mu}}^{(2\mu)} .$$

If $\{j_{d+1}, \ldots, j_{2\mu}\} \not\subseteq \{j_1, \ldots, j_d\}$, say $j_{d+1} \notin \{j_1, \ldots, j_d\}$, then these monomials can be paired off via the reflection

$$x_{j_{d+1}} \leftrightarrow y_{j_{d+1}},$$

and the two coefficients in each pair are equal due to the O_{2r}–invariant property of f.

Now suppose that $\{j_{d+1}, \ldots, j_{2\mu}\} \subseteq \{j_1, \ldots, j_d\}$. Since $d < 2\mu - d$, the number of occurrences of at least one of the indices $1,\ldots, r$ among $j_1,\ldots, j_d$ is less than among $j_{d+1}, \ldots, j_{2\mu}$. We may assume

$$j_1 = \cdots = j_a = 1 \neq j_{a+1}, \ldots, j_d$$

and

$$j_{d+1} = \cdots = j_{d+a+1} = 1.$$

Consider the $\mathrm{O}_{2r}(\mathbb{K})$–invariant

$$f\left(u^{(j_1)}, \ldots, u^{(j_d)}, v^{(d+1)}, \ldots, v^{(d+a+1)}, w^{(j_{d+a+2})}, \ldots, w^{(j_{2\mu})}\right)$$

depending on a new set of vector variables whose cardinality is

$$|\{j_1, \ldots, j_d\}| + a + 1 + |\{j_{d+a+2}, \ldots, j_{2\mu}\}| \leq d - a + 1 + a + 1 + d = 2(d+1) \leq 2r.$$

By Theorem 1.3.16 again, this invariant must be a polynomial in the q's and β's of its vector variables. As it is linear in each of $v^{(d+1)},\ldots, v^{(d+a+1)}$, these can be involved only via their β's with each other or with the u's and w's. To get the coefficient of the monomial (1.14), we substitute 1 for the s_j coordinate of each $u^{(j)}$, for the $t_{j_{d+1}}$ coordinate of $v^{(d+1)},\ldots$, for the $t_{j_{d+a+1}}$ coordinate of $v^{(d+a+1)}$, and for the t_j coordinate of each $w^{(j)}$, and we substitute zero for all other s and t coordinates. After this substitution, each of $v^{(d+1)},\ldots, v^{(d+a+1)}$ will be β–orthogonal to all substituted vectors except $u^{(1)}$, but our polynomial has only degree a in $u^{(1)}$, so the value we get is zero. $\square$

Since the multiplicity in s of a product of polynomials is the sum of the multiplicities of the factors, it follows for $r \geq 1$ that the multiplicity in s of any decomposable 2μ–linear invariant with $\mu > r$ is strictly greater than r. We arrive at

Theorem 1.4.5 *Let $m > n \geq 2$ both be even. Then the m–linear $\mathrm{O}_n(\mathbb{K})$–invariant $(1, \ldots, m)_n$ is indecomposable.*

Proof. This follows from Fact 1.2.7. □

Corollary 1.4.6 *Let $m > n \geq 2$ both be even. Then the m–linear $\mathrm{O}_n(\mathbb{C})$–invariant $(1, \ldots, m)_n$ is indecomposable in the ring $\mathbb{Z}[n \times m]^{\mathrm{O}_n(\mathbb{C})}$.*

Remark 1.4.7 Theorem 1.4.2 for $m > n$ follows from Theorem 1.4.5. Indeed, identify $\mathrm{O}_{n-1}(\mathbb{K})$ with the subgroup of $\mathrm{O}_n(\mathbb{K})$ acting on the x, y coordinate hyperplane in the standard way and fixing z. Then any $\mathrm{O}_n(\mathbb{K})$–invariant polynomial may be viewed as an $\mathrm{O}_{n-1}(\mathbb{K})$–invariant polynomial in just the x and y variables (regarding the z variables as constants). View $(1, \ldots m)$ or $[1, \ldots, m]$ that way, and break it up into its multi-homogeneous components. One of these is $(1, \ldots, m)_{n-1}$ or $(1, \ldots, m-1)_{n-1} z^{(m)}$, which is an indecomposable $\mathrm{O}_{n-1}(\mathbb{K})$–invariant by Theorem 1.4.5. It follows that $(1, \ldots, 2\mu)_n$ and $[1, \ldots, 2\mu+1]_n$ are not in the subalgebra of $\mathbb{K}[n \times m]^{\mathrm{SO}_n(\mathbb{K})}$ generated by the elements of degree less than 2μ. Since no $\mathrm{SO}_n(\mathbb{K})$–invariants of degree 1 exist, indecomposability follows for $[1, \ldots, 2\mu+1]_n$ as well as for $(1, \ldots, 2\mu)_n$.

Theorem 1.4.8 *Let $m \geq n \geq 4$ both be even. Then the m–linear $\mathrm{SO}_n(\mathbb{K})$–invariant $\langle 1, \ldots, m \rangle$ is indecomposable.*

Corollary 1.4.9 *Let $m \geq n \geq 4$ both be even. Then the m–linear $\mathrm{SO}_n(\mathbb{C})$–invariant $\langle 1, \ldots, m \rangle$ is indecomposable in the ring $\mathbb{Z}[n \times m]^{\mathrm{SO}_n(\mathbb{C})}$.*

Proof. Substituting z for each occurrence of x_r and y_r in an $\mathrm{SO}_{2r}(\mathbb{K})$–invariant yields an $\mathrm{O}_{2r-1,\mathbb{K}}$–invariant — this follows easily from Witt's Theorem [T, Theorem 7.4]. Degrees are not increased. By Fact 1.2.8, the image of $\langle 1, \ldots, m \rangle_n$ is $(1, \ldots, m)_{n-1}$, which is of the same degree and is indecomposable by Theorem 1.4.2. The image of a non-trivial decomposition of $\langle 1, \ldots, m \rangle_n$ would be a non-trivial decomposition of $(1, \ldots, m)_{n-1}$, so $\langle 1, \ldots, m \rangle_n$ must also be indecomposable. □

1.5 The orthogonal group scheme

In this section, the reader is assumed to be familiar with basic notions concerning group schemes. For these, see e.g. [H, Ja].

The *orthogonal group scheme*

$$\mathrm{O}_n = \mathrm{O}_{n,\mathbb{Z}}$$

is defined to be the stabilizer of the standard quadratic form q with respect to the action of the general linear group scheme $\mathrm{GL}_{n,\mathbb{Z}}$ on the $\mathbb{Z}$–space of quadratic forms. Thus O_n is the spectrum of the ring $\mathbb{Z}[g]/(q \circ g - q)$. Here $\mathbb{Z}[g]$ is a polynomial ring in $n \times n$ variables $g_{i,j}$, and the ideal $(q \circ g - q)$ is generated by the coefficients of the quadratic form $q \circ g - q$, where $g = (g_{i,j})$, thus the coefficients of the quadratic form are degree 2 polynomials over $\mathbb{Z}$ in the $g_{i,j}$.

As before, let $\mathbb{F}$ be an algebraically closed field. We define

$$\mathrm{O}_{n,\mathbb{F}} = \mathrm{O}_n \times \mathrm{Spec}\ \mathbb{F}.$$

The group of $\mathbb{F}$–points of the scheme O_n is the orthogonal group $\mathrm{O}_n(\mathbb{F})$ defined in Subsection 1.1.2. The scheme $\mathrm{O}_{n,\mathbb{F}}$ is reduced unless $\mathrm{char}\ \mathbb{F} = 2$ and $n = 2r + 1$. This is well known and easy.

Left multiplication gives an action of the group scheme $\mathrm{O}_{n,\mathbb{F}}$ on the affine space $\mathrm{Spec}\ \mathbb{F}[n \times m]$. A polynomial $f \in \mathbb{F}[n \times m]$ is defined to be an *invariant of the orthogonal group scheme* if its pull-back under the group scheme action map

$$\mathrm{O}_{n,\mathbb{F}} \times \mathrm{Spec}\ \mathbb{F}[n \times m] \to \mathrm{Spec}\ \mathbb{F}[n \times m]$$

coincides with its pull-back under the projection to $\mathrm{Spec}\ \mathbb{F}[n \times m]$. This is the same as saying that

$$f \circ g - f \in (q \circ g - q)\mathbb{F}[g][n \times m].$$

Clearly, invariants of the orthogonal group scheme form a subalgebra of the polynomial algebra $\mathbb{F}[n \times m]$.

Lemma 1.5.1 *Let $f, s \in \mathbb{F}[n \times m]$ with $s \neq 0$. If both s and sf are invariants of the orthogonal group scheme $\mathrm{O}_{n,\mathbb{F}}$, then so is f.*

Proof. We need to prove that

$$f \circ g - f \in (q \circ g - q)\mathbb{F}[g][n \times m].$$

Since $0 \neq s \in \mathbb{F}[n \times m]$, it suffices to prove that

$$s \cdot (f \circ g - f) \in (q \circ g - q)\mathbb{F}[g][n \times m].$$

Modulo $(q \circ g - q)$, we have

$$s \cdot (f \circ g - f) = s \cdot (f \circ g) - sf = (s \circ g) \cdot (f \circ g) - sf = sf \circ g - sf = 0.$$

□

The algebra of invariants of the orthogonal group scheme is a subalgebra of $\mathbb{F}[n \times m]^{\mathrm{O}_n(\mathbb{F})}$, and it is equal to it (by the Nullstellensatz) whenever the group scheme is reduced, i.e., always except when $\operatorname{char} \mathbb{F} = 2$ and $n = 2r+1$.

We now prove that invariants under the orthogonal group scheme $\mathrm{O}_{2r+1,\mathbb{F}}$ coincide with even invariants of the special orthogonal group $\mathrm{SO}_{2r+1}(\mathbb{F})$.

Lemma 1.5.2 *Let* $n = 2r + 1$ *and*

$$k, l, i_1, \ldots, i_n, j_1, \ldots j_n \in \{1, \ldots m\}.$$

Then the polynomials (kl) *and* $[i_1, \ldots, i_n][j_1, \ldots j_n]$ *are invariant under the orthogonal group scheme* $\mathrm{O}_{2r+1,\mathbb{F}}$.

Proof. If $k = l$, then $(kl) = (kk) = q(v^{(k)})$ is obviously invariant. If $k \neq l$, then

$$(kl) = \beta\left(v^{(k)}, v^{(l)}\right) = q\left(v^{(k)} + v^{(l)}\right) - q\left(v^{(k)}\right) - q\left(v^{(l)}\right)$$

is obviously invariant.

Recall the notation $D = [1, \ldots, n]$. By Proposition 1.3.13, D^2 can be expressed as a polynomial with integer coefficients in the (kl), thus D^2 is invariant under the orthogonal group scheme $\mathrm{O}_{2r+1,\mathbb{F}}$.

We have

$$D^2 \circ g = (D \circ g)^2 = \det g^2 \cdot D^2.$$

It follows that all coefficients of the polynomial $(\det g^2 - 1)D^2 \in \mathbb{F}[g][n \times n]$ are contained in the ideal $(q \circ g - q) \lhd \mathbb{F}[g]$. Therefore

$$\det g^2 - 1 \in (q \circ g - q).$$

We have

$$[i_1, \dots, i_n][j_1, \dots j_n] \circ g = ([i_1, \dots, i_n] \circ g) \cdot ([j_1, \dots j_n] \circ g) = \\ = \det g^2 \cdot [i_1, \dots, i_n][j_1, \dots j_n].$$

Thus

$$[i_1, \dots, i_n][j_1, \dots j_n] \circ g - [i_1, \dots, i_n][j_1, \dots j_n] = \\ = (\det g^2 - 1) \cdot [i_1, \dots, i_n][j_1, \dots j_n] \in (q \circ g - q)\mathbb{F}[g][n \times m],$$

as claimed. □

Proposition 1.5.3 *A polynomial $f \in \mathbb{F}[n \times m]$ is invariant under the orthogonal group scheme $\mathrm{O}_{2r+1,\mathbb{F}}$ precisely if it is $\mathrm{SO}_{2r+1}(\mathbb{F})$–invariant and* even *(i.e., a sum of homogeneous polynomials of even degree).*

Proof. Assume first that f is invariant under the orthogonal group scheme $\mathrm{O}_{2r+1,\mathbb{F}}$.

Let $g \in \mathrm{SO}_{2r+1}(\mathbb{F})$. Then $q \circ g = q$, so g satisfies all polynomials in the ideal defining the orthogonal group scheme, thus $f \circ g = f$, i.e., f is invariant under g.

We need to prove that f is even. This follows easily from the fact that $\mathrm{O}_{2r+1,\mathbb{F}}$ contains as diagonal subgroup scheme the (possibly non-reduced) Abelian group scheme

$$\mathrm{T} = \mathrm{Spec}\, \mathbb{F}[\xi_1, \eta_1, \dots, \xi_r, \eta_r, \zeta]/(\xi_1\eta_1 - 1, \dots, \xi_r\eta_r - 1, \zeta^2 - 1).$$

All invariants of T are even.

To prove the converse, assume now that f is $\mathrm{SO}_{2r+1}(\mathbb{F})$–invariant and even, i.e, $f \in \mathbb{F}[n \times m]^{\mathrm{O}_{\mathbb{F}}}$ in the notation used throughout Sections 1.1–1.4. By Theorem 1.3.5(d) and Lemmas 1.5.2 and 1.5.1, it follows that f is invariant under the orthogonal group scheme $\mathrm{O}_{2r+1,\mathbb{F}}$. □

Chapter 2

Character formulae for classical groups

2.1 Introduction

This chapter is essentially identical to the paper [F1]. We give formulae relating the value $\chi_\lambda(g)$ of an irreducible character of a classical group G to entries of powers of the matrix $g \in G$. This yields a far-reaching generalization of a result of J. L. Cisneros-Molina concerning the GL_2 case [C].

The Weyl character formula [FH, GW, W1, W2] tells us how to compute the character χ_λ of an irreducible finite dimensional representation V_λ with highest weight λ of a (complex, semisimple, connected) Lie group G:

$$\sum_{\sigma \in \mathfrak{W}} (-1)^\sigma \, z^{\sigma(\lambda+\rho)} = \chi_\lambda \cdot \sum_{\sigma \in \mathfrak{W}} (-1)^\sigma \, z^{\sigma\rho}.$$

Here, for each weight $\ell \in \mathfrak{h}^*$ of the Lie algebra $\mathfrak{g}$ of G, the exponential $z^\ell : \tilde{H} \to \mathbb{C}^*$ is the corresponding multiplicative character of the preimage $\tilde{H}$ of a maximal torus $H \leq G$ in a universal covering $\tilde{G} \to G$. The weight $\rho \in \mathfrak{h}^*$ is the half-sum of the positive roots, $\mathfrak{W}$ is the Weyl group. Note that both sides of the formula are $\mathfrak{W}$–antisymmetric characters of $\tilde{H}$, but χ_λ is well-defined as a $\mathfrak{W}$–symmetric character of H. Note also that the Weyl denominator

$$\sum_{\sigma \in \mathfrak{W}} (-1)^\sigma \, z^{\sigma\rho}$$

is not identically zero, so χ_λ is expressed as a ratio of Laurent polynomials in the coordinates on $\tilde{H}$. (We know *a priori* that χ_λ is itself a Laurent

polynomial in the coordinates on H, but with many more terms in general than the numerator and denominator.)

The Weyl character formula expresses the value of the character χ_λ at a group element $g \in G$ in terms of a conjugate of the semisimple part of g in the maximal torus H, i.e., in the case of the classical matrix groups, in terms of the eigenvalues of g. There are equally explicit expressions, called determinantal identities or Giambelli formulae [FH, Section A.1, formulae (A.5), (A.6) and Section A.3], in terms of the elementary resp. the complete symmetric polynomials in the eigenvalues.

Below, we shall consider the connected classical groups and we shall prove variants of the Weyl character formula that express the value $\chi_\lambda(g)$ in many different, explicit rational ways in terms of the entries of powers of the generic matrix $g \in G$. It seems likely that these formulae provide the fastest and most straightforward way of calculating $\chi_\lambda(g)$ for generic g.

This work was motivated by J. L. Cisneros-Molina's paper [C] whose main result is the following. Let $\omega \neq 0$ be a linear function on the space M_2 of 2×2 matrices, such that $\omega(\mathbf{1}) = 0$. For example, ω could be one of the two off-diagonal entries. Then, for $\lambda = 0, 1, \ldots$, we have

$$\omega\left(g^{\lambda+1}\right)/\omega(g) = \text{tr } S^\lambda g, \tag{2.1}$$

where the right hand side is the trace of the action of g on the λ-th symmetric power of the standard vector representation. This is surprising for several reasons. Firstly, there is no obvious connection between matrix power and symmetric power. Secondly, the right hand side is invariant under conjugation, which is not obvious for the left hand side, so we can view this as an unexpected construction of a classical matrix invariant. Thirdly, one would not expect such a quick and straightforward method to calculate the trace of the symmetric power.

Our results below can be considered as far-reaching generalizations of (2.1). In particular, for $g \in M_{r+1}$, and $\lambda = 0, 1, \ldots$, the trace tr $S^\lambda g$ equals the ratio of the r–dimensional volumes of the two parallel-epipeda spanned in M_{r+1}/M_1 by the images of g, g^2, $\ldots$, g^{r-1}, $g^{\lambda+r}$ and of g, g^2, $\ldots$, g^{r-1}, g^r, respectively (except when both volumes are zero, i.e., g has a minimal polynomial of degree $< r+1$). This is a particular case of Corollary 2.2.3 below.

Our proofs are motivated by the first of the four proofs given in [C], which is due to Jeremy Rickard.

2.2 General linear group

Let $G = \mathrm{GL}_{r+1}(\mathbb{C})$ with $H = (\mathbb{C}^*)^{r+1}$ the maximal torus consisting of all invertible diagonal matrices, $\mathrm{Hom}(H, \mathbb{C}^*) = \mathbb{Z}^{r+1}$ the weight lattice, and $\mathfrak{W} = \mathfrak{S}_{r+1}$ the Weyl group. Write

$$\rho = \left(\frac{r}{2}, \frac{r-2}{2}, \ldots, \frac{2-r}{2}, \frac{-r}{2}\right)$$

for the half-sum of the positive roots. Set $\rho_t = \rho + (t, t, \ldots, t, t)$ for $t \in \mathbb{C}$. For $\lambda = (\lambda_0, \ldots, \lambda_r) \in \mathbb{Z}^{r+1}$, write $z^\lambda : H \to \mathbb{C}^*$ for the corresponding multiplicative character of the torus H, and, when λ is dominant, i.e. $\lambda_0 \geq \cdots \geq \lambda_r$, write $\chi_\lambda : G \to \mathbb{C}$ for the character of the irreducible representation with highest weight λ. The Weyl character formula

$$\sum_{\sigma \in \mathfrak{W}} (-1)^\sigma \, z^{\sigma(\lambda+\rho_t)} = \chi_\lambda \cdot \sum_{\sigma \in \mathfrak{W}} (-1)^\sigma \, z^{\sigma \rho_t}$$

holds with any $t \in \mathbb{C}$. The freedom in the choice of t comes from the central $\mathbb{C}^*$ in G. Both sides of the formula are $\mathfrak{W}$–antisymmetric characters of the infinite cover $\tilde{H}$. When $\rho_t \in \mathbb{Z}^{r+1}$, both sides descend to H.

Recall that, for $g \in G$, a value of $\log g$ is defined to be any matrix $X \in \mathfrak{g}$ such that $\exp X \stackrel{\mathrm{def}}{=} \sum_{n=0}^{\infty} X^n/n! = g$. Such values X exist for any $g \in G$. Then we define

$$g^{\ell_i} = \exp(\ell_i \log g).$$

When $\ell_i \in \mathbb{Z}$, this does not depend on the chosen value of $\log g$ and coincides with the elementary definition of the matrix power.

For $\ell \in \mathbb{C}^{r+1}$ and $g \in G$, define

$$g^\ell = \bigotimes_{i=0}^{r} g^{\ell_i} \in M_{r+1}(\mathbb{C})^{\otimes(r+1)}.$$

This is multi-valued, it depends on a choice of the value of $\log g \in \mathfrak{g} = \mathfrak{gl}_{r+1}(\mathbb{C})$. When $\ell \in \mathbb{Z}^{r+1}$, it is single-valued. When $\ell \in \mathbb{Z}^{r+1}_{\geq 0}$, we may allow $g \in M_{r+1}(\mathbb{C})$ rather than $g \in G$.

We have

$$\sum_{\sigma \in \mathfrak{W}} (-1)^\sigma \, g^{\sigma \ell} = \bigwedge_{i=0}^{r} g^{\ell_i} \in M_{r+1}(\mathbb{C})^{\wedge(r+1)}.$$

Theorem 2.2.1 *Let $\lambda = (\lambda_0 \geq \cdots \geq \lambda_r) \in \mathbb{Z}^{r+1}$ and $g \in \mathrm{GL}_{r+1}(\mathbb{C})$. Then, for any $t \in \mathbb{C}$,*

$$\sum_{\sigma \in \mathfrak{W}} (-1)^\sigma \, g^{\sigma(\lambda+\rho_t)} = \chi_\lambda(g) \cdot \sum_{\sigma \in \mathfrak{W}} (-1)^\sigma \, g^{\sigma \rho_t};$$

equivalently,

$$\bigwedge_{i=0}^{r} g^{\lambda_i + r/2 - i + t} = \chi_\lambda(g) \cdot \bigwedge_{i=0}^{r} g^{r/2-i+t},$$

where the powers are defined using any (but always the same) value of $\log g$. *When $\rho_t \in \mathbb{Z}^{r+1}$, the powers are single-valued. In particular, for $t = r/2$, we get*

$$\bigwedge_{i=0}^{r} g^{\lambda_i + r - i} = \chi_\lambda(g) \cdot \bigwedge_{i=0}^{r} g^{r-i}.$$

When ρ_t and λ are both in $\mathbb{Z}^{r+1}_{\geq 0}$, we may allow $g \in M_{r+1}(\mathbb{C})$.

Proof. The set of diagonalizable invertible matrices is dense in $M_{r+1}(\mathbb{C})$, so we may assume that g is such. The statement of the theorem is invariant under conjugation, so we may assume that $g = \mathrm{diag}(z_0, \ldots, z_r) \in H$. Then

$$\sum_{\sigma \in \mathfrak{W}} (-1)^\sigma \, g^{\sigma \ell} = \bigwedge_{i=0}^{r} g^{\ell_i} = \left| z_j^{\ell_i} \right| \cdot \bigwedge_{j=0}^{r} e_{jj},$$

where e_{jj} is the diagonal matrix with a single 1 at the j-th position. The theorem now follows from the Weyl character formula. □

Corollary 2.2.2 *Let Ω be an alternating $(r+1)$–linear form on the space $M_{r+1}(\mathbb{C})$. Then, for $\lambda = (\lambda_0 \geq \cdots \geq \lambda_r) \in \mathbb{Z}^{r+1}$, we have*

$$\Omega\left(g^{\lambda_0 + r}, g^{\lambda_1 + r - 1}, \ldots, g^{\lambda_r}\right) = \chi_\lambda(g) \cdot \Omega\left(g^r, g^{r-1}, \ldots, \mathbf{1}\right).$$

To express $\chi_\lambda(g)$ as a rational function in entries of powers of g, we must choose Ω such that the right hand side is not identically zero. For example, $\Omega(g_0, \ldots, g_r)$ could be the determinant of the matrix formed by the diagonals, or by the first rows, etc. of the argument matrices. To calculate $\chi_\lambda(g)$ for a numerically given g, we need to choose Ω such that $\Omega(g^r, g^{r-1}, \ldots, \mathbf{1}) \neq 0$. This is possible if and only if $\bigwedge_{i=0}^{r} g^{r-i} \neq 0$, i.e., the minimal polynomial of g is its characteristic polynomial. When g has a minimal polynomial of lower degree, we can use l'Hospital's rule.

Corollary 2.2.3 *On the space $M_{r+1}(\mathbb{C})$, consider an alternating r–linear form ω such that ω vanishes if an argument is $\mathbf{1}$. Then, for λ as above and with $\lambda_r = 0$, we have*

$$\omega\left(g^{\lambda_0+r}, g^{\lambda_1+r-1}, \ldots, g^{\lambda_{r-1}+1}\right) = \chi_\lambda(g) \cdot \omega\left(g^r, g^{r-1}, \ldots, g\right).$$

To express $\chi_\lambda(g)$ as a rational function in entries of powers of g, we must choose ω such that the right hand side is not identically zero. For example, $\omega(g_0, \ldots, g_{r-1})$ could be the determinant of the $r \times r$ matrix formed by the truncated (i.e., leftmost entry omitted) first rows of the argument matrices. To calculate $\chi_\lambda(g)$ for a numerically given g, we need to choose ω such that $\omega(g^r, g^{r-1}, \ldots, g) \neq 0$. This is possible if and only if the minimal polynomial of g is its characteristic polynomial.

Corollary 2.2.3, for $r = 1$, is the result of J. L. Cisneros-Molina's paper [C] mentioned in the Introduction.

To derive Corollary 2.2.3 from Corollary 2.2.2, simply set $\Omega = d\omega$, defined as usual by

$$\Omega(g_0, \ldots, g_r) = \sum_{i=0}^{r} (-1)^i \omega(g_0, \ldots, g_{i-1}, g_{i+1}, \ldots, g_r).$$

Then $\Omega(g_0, \ldots, g_{r-1}, \mathbf{1}) = (-1)^r \omega(g_0, \ldots, g_{r-1})$ and Corollary 2.2.3 follows.

2.3 Special linear group

Let $G = \mathrm{SL}_{r+1}(\mathbb{C})$ with $H \simeq (\mathbb{C}^*)^r$ the maximal torus consisting of all unimodular diagonal matrices, $\mathrm{Hom}(H, \mathbb{C}^*) = \mathbb{Z}^{r+1}/\mathbb{Z}$ the weight lattice, and $\mathfrak{W} = \mathfrak{S}_{r+1}$ the Weyl group. Write $\rho = (r, r-1, \ldots, 0) + \mathbb{Z} \cdot (1, \ldots, 1) \in \mathbb{Z}^{r+1}/\mathbb{Z}$ for the half-sum of the positive roots. When $\lambda = (\lambda_0, \ldots, \lambda_r) \in \mathbb{Z}^{r+1}/\mathbb{Z}$, write $z^\lambda : H \to \mathbb{C}^*$ for the corresponding multiplicative character of the torus H, and, when λ is dominant, write $\chi_\lambda : G \to \mathbb{C}$ for the character of the irreducible representation with highest weight λ. The Weyl character formula is valid as stated in the introduction. Both sides are $\mathfrak{W}$–antisymmetric characters of H.

For $\ell \in \mathbb{Z}^{r+1}/\mathbb{Z}$ and $g \in G$, the antisymmetric tensor

$$\bigwedge_{i=0}^{r} g^{\ell_i} \in M_{r+1}(\mathbb{C})^{\wedge(r+1)}$$

is well defined because either g has a minimal polynomial of degree $< r+1$, in which case the algebra $\mathbb{C}[g]$ has dimension $< r+1$ and the antisymmetric tensor above is zero, or else g has its characteristic polynomial as minimal polynomial, in which case $\dim \mathbb{C}[g] = r+1$ and multiplication by g on it has determinant $\det g = 1$, so the tensor is independent of the chosen representative of ℓ.

Theorem 2.3.1 *Let* $\lambda = (\lambda_0 \geq \cdots \geq \lambda_r) + \mathbb{Z} \cdot (1, \ldots, 1) \in \mathbb{Z}^{r+1}/\mathbb{Z}$ *and* $g \in \mathrm{SL}_{r+1}(\mathbb{C})$. *Then*

$$\bigwedge_{i=0}^{r} g^{\ell_i} = \chi_\lambda(g) \cdot \bigwedge_{i=0}^{r} g^{r-i},$$

where $\ell_i = \lambda_i + r - i$.

Proof. The theorem trivially follows from Theorem 2.2.1. □

2.4 Odd special orthogonal group

As in Chapter 1, let $G = \mathrm{SO}_{2r+1}(\mathbb{C})$ be the connected group preserving the quadratic form

$$x_1 y_1 + \cdots + x_r y_r + z^2.$$

We take the maximal torus $H = (\mathbb{C}^*)^r$ consisting of all special orthogonal diagonal matrices

$$\mathrm{diag}\left(z_1, z_1^{-1}, \ldots, z_r, z_r^{-1}, 1\right).$$

In the weight lattice $\mathrm{Hom}(H, \mathbb{C}^*) = \mathbb{Z}^r$, we take $\lambda = (\lambda_1, \ldots, \lambda_r)$ to correspond to the monomial

$$z^\lambda = \prod_{j=1}^{r} z_j^{\lambda_j}.$$

The Weyl group $\mathfrak{W}$ is the semidirect product of $\mathfrak{S}_r$ and Z_2^r. Write $\rho = (r - 1/2, r - 3/2, \ldots, 3/2, 1/2)$ for the half-sum of the positive roots. The Weyl character formula is valid as stated in the introduction. Both sides are $\mathfrak{W}$–antisymmetric characters of the double cover $\tilde{H}$.

For $\ell \in \left(\mathbb{Z} + \frac{1}{2}\right)^r$ and $g \in G$, define

$$g^\ell = \bigotimes_{i=1}^{r} g^{\ell_i} \in M_{2r}(\mathbb{C})^{\otimes r}.$$

This is multi-valued, it depends on a choice of $\sqrt{g}$. We may choose any matrix $\sqrt{g}$ whose square is g. Such matrices $\sqrt{g} \in G$ always exist. Then write

$$g^{\ell_i} = \sqrt{g}^{2\ell_i}$$

to define the matrix power. We have

$$\sum_{\sigma \in \mathfrak{W}} (-1)^{\sigma} g^{\sigma \ell} = \bigwedge_{i=1}^{r} \left(g^{\ell_i} - g^{-\ell_i} \right) \in M_{2r}(\mathbb{C})^{\wedge r}.$$

Theorem 2.4.1 *Let $\lambda = (\lambda_1 \geq \cdots \geq \lambda_r) \in \mathbb{Z}^r_{\geq 0}$ and $g \in \mathrm{SO}_{2r+1}(\mathbb{C})$. Then*

$$\sum_{\sigma \in \mathfrak{W}} (-1)^{\sigma} g^{\sigma \ell} = \chi_{\lambda}(g) \cdot \sum_{\sigma \in \mathfrak{W}} (-1)^{\sigma} g^{\sigma \rho};$$

equivalently,

$$\bigwedge_{i=1}^{r} \left(g^{\ell_i} - g^{-\ell_i} \right) = \chi_{\lambda}(g) \cdot \bigwedge_{i=1}^{r} \left(g^{r+1/2-i} - g^{-(r+1/2-i)} \right),$$

where $\ell = \lambda + \rho$, i.e. $\ell_i = \lambda_i + r + 1/2 - i$, and the powers are defined using any, but always the same value of $\sqrt{g} \in \mathrm{SO}_{2r+1}(\mathbb{C})$.

Proof. The set of diagonalizable matrices is dense in G, so we may assume that $\sqrt{g}$ is such. The statement of theorem is invariant under conjugation, so we may assume that

$$\sqrt{g} = \mathrm{diag}\left(z_1^{1/2}, z_1^{-1/2}, \ldots, z_r^{1/2}, z_r^{-1/2}, 1 \right) \in H.$$

Then

$$\sum_{\sigma \in \mathfrak{W}} (-1)^{\sigma} g^{\sigma \ell} = \bigwedge_{i=1}^{r} \left(g^{\ell_i} - g^{-\ell_i} \right) = \left| z_j^{\ell_i} - z_j^{-\ell_i} \right| \cdot \bigwedge_{j=1}^{r} (e_{jj} - f_{jj}),$$

where e_{jj} resp. f_{jj} is the diagonal matrix with a single 1 at the position corresponding to the x_j resp. y_j coordinate. The theorem now follows from the Weyl character formula. $\square$

2.5 Symplectic group

Let $G = \mathrm{Sp}_{2r}(\mathbb{C})$ be the group preserving the skew bilinear form

$$\sum_{i=1}^{r} (x_i' y_i'' - y_i' x_i'')$$

on $\mathbb{C}^{2r}$. We take the maximal torus $H = (\mathbb{C}^*)^r$ consisting of all symplectic diagonal matrices

$$\mathrm{diag}\left(z_1, z_1^{-1}, \ldots, z_r, z_r^{-1}\right).$$

In the weight lattice $\mathrm{Hom}(H, \mathbb{C}^*) = \mathbb{Z}^r$, we take $\lambda = (\lambda_1, \ldots, \lambda_r)$ to correspond to the monomial

$$z^\lambda = \prod_{j=1}^{r} z_j^{\lambda_j}.$$

The Weyl group $\mathfrak{W}$ is the semidirect product of $\mathfrak{S}_r$ and Z_2^r. Write $\rho = (r, r-1, \ldots, 1)$ for the half-sum of the positive roots. The Weyl character formula is valid as stated in the introduction. Both sides are $\mathfrak{W}$–antisymmetric characters of H.

For $\ell \in \mathbb{Z}^r$ and $g \in G$, define

$$g^\ell = \bigotimes_{i=1}^{r} g^{\ell_i} \in M_{2r}(\mathbb{C})^{\otimes r}.$$

Then

$$\sum_{\sigma \in \mathfrak{W}} (-1)^\sigma g^{\sigma\ell} = \bigwedge_{i=1}^{r} \left(g^{\ell_i} - g^{-\ell_i}\right) \in M_{2r}(\mathbb{C})^{\wedge r}.$$

Theorem 2.5.1 *Let $\lambda = (\lambda_1 \geq \cdots \geq \lambda_r) \in \mathbb{Z}^r_{\geq 0}$ and $g \in \mathrm{Sp}_{2r}(\mathbb{C})$. Then*

$$\sum_{\sigma \in \mathfrak{W}} (-1)^\sigma g^{\sigma\ell} = \chi_\lambda(g) \cdot \sum_{\sigma \in \mathfrak{W}} (-1)^\sigma g^{\sigma\rho};$$

equivalently,

$$\bigwedge_{i=1}^{r} \left(g^{\ell_i} - g^{-\ell_i}\right) = \chi_\lambda(g) \cdot \bigwedge_{i=1}^{r} \left(g^{r+1-i} - g^{-(r+1-i)}\right),$$

where $\ell = \lambda + \rho$, i.e., $\ell_i = \lambda_i + r + 1 - i$.

Proof. The set of diagonalizable matrices is dense in G, so we may assume that g is such. The statement of theorem is invariant under conjugation, so we may assume that $g = \operatorname{diag}\left(z_1, z_1^{-1}, \ldots, z_r, z_r^{-1}\right) \in H$. Then

$$\sum_{\sigma \in \mathfrak{W}} (-1)^{\sigma} g^{\sigma \ell} = \bigwedge_{i=1}^{r} \left(g^{\ell_i} - g^{-\ell_i}\right) = \left|z_j^{\ell_i} - z_j^{-\ell_i}\right| \cdot \bigwedge_{j=1}^{r} \left(e_{jj} - f_{jj}\right),$$

where e_{jj} resp. f_{jj} is the diagonal matrix with a single 1 at the position corresponding to the x_j resp. y_j coordinate. The theorem now follows from the Weyl character formula. □

2.6 Even special orthogonal group

As in Chapter 1, let $G = \mathrm{SO}_{2r}(\mathbb{C})$ be the connected group preserving the quadratic form

$$q = x_1 y_1 + \cdots + x_r y_r.$$

We take the maximal torus $H = (\mathbb{C}^*)^r$ consisting of all special orthogonal diagonal matrices

$$\operatorname{diag}\left(z_1, z_1^{-1}, \ldots, z_r, z_r^{-1}\right).$$

In the weight lattice $\operatorname{Hom}(H, \mathbb{C}^*) = \mathbb{Z}^r$, we take $\lambda = (\lambda_1, \ldots, \lambda_r)$ to correspond to the monomial

$$z^{\lambda} = \prod_{j=1}^{r} z_j^{\lambda_j}.$$

The Weyl group $\mathfrak{W}$ is the semidirect product of $\mathfrak{S}_r$ and Z_2^{r-1}. It acts by permuting the indices and by performing an even number of sign changes. Write $\widetilde{\mathfrak{W}} > \mathfrak{W}$ for the Weyl group in the full orthogonal group $\mathrm{O}_{2r}(\mathbb{C})$. It is the semidirect product of $\mathfrak{S}_r$ and Z_2^r. If $\sigma \in \widetilde{\mathfrak{W}}$, we write $[\sigma]$ for its image in $\mathfrak{S}_r$. Write $\rho = (r-1, r-2, \ldots, 1, 0)$ for the half-sum of the positive roots. Write $\epsilon = (1, 1, \ldots, 1, 1)$ so that $e = \epsilon + \rho = (r, r-1, \ldots, 2, 1)$ is regular for $\widetilde{\mathfrak{W}}$. The Weyl character formula is valid as stated in the introduction. Both sides are $\mathfrak{W}$–antisymmetric characters of H.

For $\ell \in \mathbb{Z}^r$ and $g \in G$, define

$$g^{\ell} = \bigotimes_{i=1}^{r} g^{\ell_i} \in M_{2r}(\mathbb{C})^{\otimes r}.$$

We have

$$2\sum_{\sigma\in\mathfrak{W}}(-1)^{\sigma}g^{\sigma\ell}=\sum_{\sigma\in\widetilde{\mathfrak{W}}}\left((-1)^{[\sigma]}+(-1)^{\sigma}\right)g^{\sigma\ell}=$$
$$=\bigwedge_{i=1}^{r}\left(g^{\ell_i}+g^{-\ell_i}\right)+\bigwedge_{i=1}^{r}\left(g^{\ell_i}-g^{-\ell_i}\right)\in M_{2r}\left(\mathbb{C}\right)^{\wedge r}.$$

Note that the second term is zero if any ℓ_i is zero.

Theorem 2.6.1 *Let $\lambda=(\lambda_1,\ldots,\lambda_r)\in\mathbb{Z}^r$ with $\lambda_1\geq\cdots\geq\lambda_{r-1}\geq|\lambda_r|$. Set $\bar{\lambda}=(\lambda_1,\ldots,\lambda_{r-1},-\lambda_r)$. Let $g\in\mathrm{SO}_{2r}(\mathbb{C})$. Then*

$$2\sum_{\sigma\in\widetilde{\mathfrak{W}}}(-1)^{[\sigma]}g^{\sigma\ell}=\left(\chi_{\lambda}+\chi_{\bar{\lambda}}\right)(g)\cdot\sum_{\sigma\in\widetilde{\mathfrak{W}}}(-1)^{[\sigma]}g^{\sigma\rho};$$

equivalently,

$$2\bigwedge_{i=1}^{r}\left(g^{\ell_i}+g^{-\ell_i}\right)=\left(\chi_{\lambda}+\chi_{\bar{\lambda}}\right)(g)\cdot\bigwedge_{i=1}^{r}\left(g^{r-i}+g^{-(r-i)}\right).$$

Also,

$$\left(\chi_{\epsilon}-\chi_{\bar{\epsilon}}\right)(g)\cdot\sum_{\sigma\in\widetilde{\mathfrak{W}}}(-1)^{\sigma}g^{\sigma\ell}=\left(\chi_{\lambda}-\chi_{\bar{\lambda}}\right)(g)\cdot\sum_{\sigma\in\widetilde{\mathfrak{W}}}(-1)^{\sigma}g^{\sigma e};$$

equivalently,

$$\sqrt{-1}^{r}\mathrm{pf}_{q}\left(g-g^{-1}\right)\cdot\bigwedge_{i=1}^{r}\left(g^{\ell_i}-g^{-\ell_i}\right)=$$
$$=\left(\chi_{\lambda}-\chi_{\bar{\lambda}}\right)(g)\cdot\bigwedge_{i=1}^{r}\left(g^{r+1-i}-g^{-(r+1-i)}\right).$$

Throughout, $\ell=\lambda+\rho$, i.e., $\ell_i=\lambda_i+r-i$.

Note that

$$\left(\chi_{\epsilon}-\chi_{\bar{\epsilon}}\right)(g)=\sqrt{-1}^{r}\mathrm{pf}_{q}\left(g-g^{-1}\right).$$

Here the Pfaffian with respect to q of the linear transformation

$$g-g^{-1}\in\mathfrak{so}_{2r}(\mathbb{C})$$

is a square root of the determinant. It is defined by computing the linear transformation's matrix with respect to a positively oriented ordered q–orthonormal basis of the standard vector representation $\mathbb{C}^{2r}$, and taking the Pfaffian, as defined in the Preface, of that anti-symmetric matrix. We declare the ordered q–orthonormal bases of the standard vector representation $\mathbb{C}^{2r}$ with determinant $\left(2\sqrt{-1}\right)^r$ to be of positive orientation, as opposed to those with determinant $-\left(2\sqrt{-1}\right)^r$.

Proof. The set of diagonalizable matrices is dense in G, so we may assume that g is such. The statement of the theorem is invariant under conjugation, so we may assume that $g = \operatorname{diag}\left(z_1, z_1^{-1}, \ldots, z_r, z_r^{-1}\right) \in H$. Then

$$\sum_{\sigma \in \widetilde{\mathfrak{W}}} (-1)^{[\sigma]}\, g^{\sigma\ell} = \bigwedge_{i=1}^{r} \left(g^{\ell_i} + g^{-\ell_i}\right) = \left|z_j^{\ell_i} + z_j^{-\ell_i}\right| \cdot \bigwedge_{j=1}^{r} (e_{jj} + f_{jj})$$

and

$$\sum_{\sigma \in \widetilde{\mathfrak{W}}} (-1)^{\sigma}\, g^{\sigma\ell} = \bigwedge_{i=1}^{r} \left(g^{\ell_i} - g^{-\ell_i}\right) = \left|z_j^{\ell_i} - z_j^{-\ell_i}\right| \cdot \bigwedge_{j=1}^{r} (e_{jj} - f_{jj})\,,$$

where e_{jj} resp. f_{jj} is the diagonal matrix with a single 1 at the position corresponding to the x_j resp. y_j coordinate. The theorem now follows from the Weyl character formula. □

Chapter 3

Inequalities for positive semi-definite matrices

3.1 Introduction

This chapter is a slightly extended version of the paper [F2]. Its contents are as follows. In Section 3.2, we sketch one part of the historic background: classical inequalities on determinants and permanents of positive semi-definite matrices. The most interesting permanental inequalities mentioned are only conjectures. In Section 3.3, we prove new Pfaffian and Hafnian versions of these inequalities. Also, we formulate Conjecture 3.3.5, another Hafnian inequality. In Section 3.4, we apply the Hafnian inequality of Theorem 3.3.4 to our main goal: improving the lower bound of Révész and Sarantopoulos on the norm of a product of linear functionals on a real Euclidean space (this subject is sometimes called the 'real linear polarization constant' problem, its history is sketched at the end of the chapter). This is achieved in Theorem 3.4.3. We point out that Conjecture 3.3.5 would be sufficient to completely settle the real linear polarization constant problem.

3.2 Old inequalities on determinants and permanents

Recall that the determinant and the permanent of an $m \times m$ matrix $C = (c_{i,j})$ are defined by

$$\det C = \sum_{\pi \in \mathfrak{S}_m} (-1)^\pi \prod_{i=1}^m c_{i,\pi(i)}, \qquad \text{per } C = \sum_{\pi \in \mathfrak{S}_m} \prod_{i=1}^m c_{i,\pi(i)},$$

where $\mathfrak{S}_m$ is the symmetric group on m elements. Throughout this section, we assume that C is a positive semi-definite Hermitian $m \times m$ matrix (we write $C \geq 0$). For such C, Hadamard proved that

$$\det C \leq \prod_{i=1}^m c_{i,i}, \tag{3.1}$$

with equality if and only if C has a zero row or is a diagonal matrix. Fischer generalized this to

$$\det C \leq \det B' \cdot \det B'' \tag{3.2}$$

for

$$C = \begin{pmatrix} B' & A \\ A^* & B'' \end{pmatrix} \geq 0, \tag{3.3}$$

with equality if and only if $\det B' \cdot \det B'' \cdot A = 0$.

Concerning the permanent of a positive semi-definite matrix, Marcus [Mar1, Mar2] proved that

$$\text{per } C \geq \prod_{i=1}^m c_{i,i}, \tag{3.4}$$

with equality if and only if C has a zero row or is a diagonal matrix. Lieb [L] generalized this to

$$\text{per } C \geq \text{per } B' \cdot \text{per } B'' \tag{3.5}$$

for C as in (3.3), with equality if and only if C has a zero row or $A = 0$. Moreover, he proved that in the polynomial $P(\lambda)$ of degree n (=size of B') defined by

$$P(\lambda) = \text{per} \begin{pmatrix} B' & A \\ \lambda A^* & B'' \end{pmatrix} = \sum_{t=0}^n c_t \lambda^t,$$

all coefficients c_t are real and non-negative. This is indeed a stronger theorem since it implies

$$\text{per } C = P(1) = \sum_{t=0}^{n} c_t \geq c_0 = \text{per } B' \cdot \text{per } B''.$$

If $m = 2n$, then the inequalities $c_t \geq 0$ even imply

$$\text{per } C \geq \text{per } B' \cdot \text{per } B'' + |\text{per } A|^2, \tag{3.6}$$

since the right hand side is $c_0 + c_n$. Inequality (3.6) is case $p = 2$ of the following conjecture of Marcus: If C is a positive semi-definite Hermitian $pn \times pn$ matrix partitioned into $p \times p$ blocks $A_{i,j}$, each of size $n \times n$, then

$$\text{per } C \geq \text{per}\,((\text{per } A_{i,j})_{i,j}).$$

This is itself a special case of the so-called permanental dominance conjecture, which we do not state here.

Đoković [D, Mi] gave a simple proof of Lieb's above inequalities, and showed also that if B' and B'' are positive definite then $c_t = 0$ if and only if all sub-permanents of A of order t vanish. Lieb [L] also states an analogous (and analogously provable) theorem for determinants: for C as in (3.3), let

$$D(\lambda) = \det \begin{pmatrix} B' & A \\ -\lambda A^* & B'' \end{pmatrix} = \sum_{t=0}^{n} d_t \lambda^t. \tag{3.7}$$

If $\det B' \cdot \det B'' = 0$, then $D(\lambda) = 0$. If B' and B'' are positive definite, then d_t is positive for $t \leq \text{rk } A$ and is zero for $t > \text{rk } A$. In contrast to the above deduction of the permanental Lieb inequality (3.5) from $c_t \geq 0$, there is no obvious way of deducing the Fischer inequality (3.2) from $d_t \geq 0$. Instead of (3.2), we get

$$D(1) = \det \begin{pmatrix} B' & A \\ -A^* & B'' \end{pmatrix} \geq \det B' \cdot \det B''. \tag{3.8}$$

Remark 3.2.1 In all of Lieb's inequalities mentioned above, the condition that the matrix C is positive semi-definite can be replaced by the weaker condition that the diagonal blocks B' and B'' are positive semi-definite. The proof goes through virtually unchanged. Alternatively, this stronger form of the inequalities can be easily deduced from the seemingly weaker form above.

3.3 New inequalities

Marvin Marcus and his school have recognized the usefulness of the inner product on spaces of various types of tensors as a tool for proving inequalities for positive semi-definite matrices. The proofs in this section will be greatly simplified by introducing inner products of tensors, despite the fact that no tensors will appear in the theorems.

By a real Euclidean space, we mean a real vector space endowed with a positive definite inner product $(\cdot, \cdot)$. The Euclidean norm is defined by $|v|^2 = (v, v)$. Euclidean inner products on the exterior tensor algebra $\bigwedge V$ and the symmetric tensor algebra SV are defined by

$$\left(\bigwedge v_i,\ \bigwedge w_j\right) := \det((v_i, w_j)),$$

and

$$\left(\prod v_i, \prod w_j\right) := \text{per } ((v_i, w_j)),$$

respectively, cf. [Mar1, Mar2, MN, Mi]. Homogeneous tensors of distinct degrees are considered to be orthogonal, in accordance with the convention that non-square determinants and permanents are zero. The positive definiteness of these inner products is easily verified by choosing arbitrary orthogonal bases in the original space V, then constructing orthogonal bases in the tensor spaces in the obvious way, and checking $(v, v) > 0$ for each basis tensor v.

We shall use the following notation for submatrices. For an $n \times n$ matrix $A = (a_{i,j})$ and subsets S, T of $N := \{1, \ldots, n\}$, we write

$$A_{S,T} := (a_{i,j})_{i \in S, j \in T}.$$

3.3.1 Pfaffians

As far as the applications in Section 3.4 are concerned, this subsection may be skipped.

We define the sign of a subset T of $N := \{1, \ldots, n\}$ to be

$$(-1)^T := (-1)^{\lceil |T|/2 \rceil + \sum_{j \in T} j}.$$

If $T = \{i_1 < i_2 < \ldots\}$ and $N \setminus T = \{j_1 < j_2 < \ldots\}$, then $(-1)^T$ is the sign of the permutation $i_1, i_2, \ldots, j_1, j_2 \ldots$.

Recall from the Preface that the Pfaffian of a $2n \times 2n$ matrix $C = (c_{i,j})$ is defined by

$$\text{pf } C = \frac{1}{n!2^n} \sum_{\pi \in \mathfrak{S}_{2n}} (-1)^\pi c_{\pi(1),\pi(2)} \cdots c_{\pi(2n-1),\pi(2n)}.$$

When C is anti-symmetric, we have $(\text{pf } C)^2 = \det C$.

For A and B both of size $n \times n$, we consider the polynomial

$$(-1)^{\lfloor n/2 \rfloor} \text{pf} \begin{pmatrix} -\lambda A & B \\ -B & A \end{pmatrix} = \sum_{t=0}^{\lfloor n/2 \rfloor} p_t \lambda^t.$$

Theorem 3.3.1 *Let A and B be real $n \times n$ matrices with A anti-symmetric and B symmetric. If B is positive semi-definite, then $p_t \geq 0$ for all t. If B is positive definite, then $p_t > 0$ for $t \leq (\text{rk } A)/2$ and $p_t = 0$ for $t > (\text{rk } A)/2$.*

Proof. If $B = (b_{i,j})$ is positive semi-definite, then there exist vectors $x_1, \ldots, x_n$ in a real Euclidean space V such that $(x_i, x_j) = b_{i,j}$. We have

$$p_t = \sum_{|S|=2t} \sum_{|T|=2t} (-1)^S (-1)^T \text{pf } A_{S,S} \cdot \text{pf } A_{T,T} \cdot \det B_{N \setminus S, N \setminus T} =$$

$$= \sum_{|S|=2t} \sum_{|T|=2t} \left((-1)^S \text{pf } A_{S,S} \cdot \bigwedge_{i \notin S} x_i,\ (-1)^T \text{pf } A_{T,T} \cdot \bigwedge_{j \notin T} x_j \right) =$$

$$= \left| \sum_{|S|=2t} (-1)^S \text{pf } A_{S,S} \cdot \bigwedge_{i \notin S} x_i \right|^2 \geq 0.$$

Assume that B is positive definite. Then the vectors x_i are linearly independent. It follows that the tensors $\bigwedge_{i \notin S} x_i$ are also linearly independent as S runs over the subsets of N. Thus $p_t = 0$ if and only if pf $A_{S,S} = 0$ for all $|S| = 2t$, i.e., if and only if $2t > \text{rk } A$. □

Theorem 3.3.2 *Let A and B be real $n \times n$ matrices with A anti-symmetric and B symmetric. Let $\lambda \geq 0$. If B is positive semi-definite, then*

$$(-1)^{\lfloor n/2 \rfloor} \text{pf} \begin{pmatrix} -\lambda A & B \\ -B & A \end{pmatrix} \geq \det B + \lambda^{n/2} \det A.$$

If B is positive definite, then equality occurs if and only if $\lambda A = 0$ or $n = 2$.

Proof. The left hand side is

$$p_0 + p_1\lambda + \cdots + p_{\lfloor n/2 \rfloor}\lambda^{\lfloor n/2 \rfloor}.$$

The right hand side is $p_0 + p_{n/2}\lambda^{n/2}$. (When n is odd, we define $p_{n/2} = 0 = \det A$.) □

I am grateful to the anonymous referee of [F2] for the idea of the following alternative proof, in the spirit of [J], of Theorems 3.3.1 and 3.3.2. We may assume $B > 0$, since every positive semi-definite matrix is a limit of positive definite ones. We define $\sqrt{B} = B^{1/2}$ to be the unique positive definite symmetric real matrix whose square is B. We define $B^{-1/2} = \sqrt{B}^{-1}$. The matrix $B^{-1/2}AB^{-1/2}$ being real and anti-symmetric, there exists a unitary matrix U such that $D := U^{-1}B^{-1/2}AB^{-1/2}U$ is diagonal with purely imaginary eigenvalues $a_1\sqrt{-1}$, ..., $a_n\sqrt{-1}$. The real multi-set $\{a_1, \ldots, a_n\}$ is invariant under $a \leftrightarrow -a$. We have

$$\det \begin{pmatrix} -\lambda A & B \\ -B & A \end{pmatrix} = \det \begin{pmatrix} -\lambda\sqrt{B}UDU^{-1}\sqrt{B} & B \\ -B & \sqrt{B}UDU^{-1}\sqrt{B} \end{pmatrix} =$$

$$= \det \left(\begin{pmatrix} \sqrt{B}U & 0 \\ 0 & \sqrt{B}U \end{pmatrix} \begin{pmatrix} -\lambda D & \mathbf{1} \\ -\mathbf{1} & D \end{pmatrix} \begin{pmatrix} U^{-1}\sqrt{B} & 0 \\ 0 & U^{-1}\sqrt{B} \end{pmatrix} \right) =$$

$$= \det \sqrt{B}^4 \cdot \prod_{i=1}^{n} \det \begin{pmatrix} -\lambda a_i\sqrt{-1} & 1 \\ -1 & a_i\sqrt{-1} \end{pmatrix} = \det B^2 \cdot \prod_{i=1}^{n}(1 + a_i^2\lambda),$$

We may extract square roots. Choosing the sign of the square root in accordance with $p_0 = +\det B$, we get

$$\sum_{t=0}^{\lfloor n/2 \rfloor} p_t\lambda^t = (-1)^{\lfloor n/2 \rfloor}\mathrm{pf} \begin{pmatrix} -\lambda A & B \\ -B & A \end{pmatrix} = \det B \cdot \prod_{a_i > 0}(1 + a_i^2\lambda),$$

whence theorems 3.3.1 and 3.3.2 immediately follow.

3.3.2 Hafnians

Recall that the Hafnian of a $2n \times 2n$ matrix $C = (c_{i,j})$ is defined by

$$\mathrm{haf}\, C = \frac{1}{n!2^n} \sum_{\pi \in \mathfrak{S}_{2n}} c_{\pi(1),\pi(2)} \cdots c_{\pi(2n-1),\pi(2n)}.$$

For A and B, both of size $n \times n$, we consider the polynomial

$$\text{haf} \begin{pmatrix} \lambda A & B \\ B & A \end{pmatrix} = \sum_{t=0}^{\lfloor n/2 \rfloor} h_t \lambda^t.$$

Theorem 3.3.3 *Let A and B be symmetric real $n \times n$ matrices. If B is positive semi-definite, then $h_t \geq 0$ for all t. If B is positive definite, then $h_t = 0$ if and only if all $2t \times 2t$ diagonal sub-Hafnians of A vanish.*

Proof. If $B = (b_{i,j})$ is positive semi-definite, then there exist vectors $x_1, \ldots, x_n$ in a real Euclidean space V such that $(x_i, x_j) = b_{i,j}$. We have

$$h_t = \sum_{|S|=2t} \sum_{|T|=2t} \text{haf } A_{S,S} \cdot \text{haf } A_{T,T} \cdot \text{per } B_{N \setminus S, N \setminus T} =$$

$$= \left| \sum_{|S|=2t} \text{haf } A_{S,S} \cdot \prod_{i \notin S} x_i \right|^2 \geq 0.$$

Assume that B is positive definite. Then the vectors x_i are linearly independent. It follows that the tensors $\prod_{i \notin S} x_i$ are also linearly independent as S runs over the subsets of N. Thus $h_t = 0$ if and only if haf $A_{S,S} = 0$ for all $|S| = 2t$. □

Theorem 3.3.4 *Let A and B be symmetric real $n \times n$ matrices. Let $\lambda \geq 0$. If B is positive semi-definite, then*

$$\text{haf} \begin{pmatrix} \lambda A & B \\ B & A \end{pmatrix} \geq \text{per } B + \lambda^{n/2} (\text{haf } A)^2.$$

If B is positive definite, then equality occurs if and only if λA is a diagonal matrix or $n = 2$.

(For odd n, we define haf $A = 0$.)

Proof. The left hand side is

$$h_0 + h_1 \lambda + \cdots + h_{\lfloor n/2 \rfloor} \lambda^{\lfloor n/2 \rfloor}.$$

The right hand side is $h_0 + h_{n/2} \lambda^{n/2}$. (When n is odd, we define $h_{n/2} = 0 = (\text{haf } A)^2$.) □

Setting $A = B$ and $\lambda = 1$, dropping the second term on the right hand side, and combining with Marcus's inequality (3.4), we arrive at case $p = 1$ of

Conjecture 3.3.5 *If $A = (a_{i,j})$ is a positive semi-definite symmetric real $n \times n$ matrix, then the Hafnian of the $2pn \times 2pn$ matrix consisting of $2p \times 2p$ blocks A is at least*

$$(2p-1)!!^n \prod_{i=1}^{n} a_{i,i}^p,$$

with equality if and only if A has a zero row or is a diagonal matrix.

3.4 Products of real linear functionals

In this section, we apply Theorem 3.3.4 to products of jointly normal random variables and then to products of real linear functionals, which was the main motivation for the work in this chapter. The ideas in this section are analogous to those that Arias-de-Reyna [A] used in the complex case.

Let $\xi_1, \ldots, \xi_d$ denote independent random variables with standard Gaussian distribution, i.e., with the joint density function $(2\pi)^{-d/2} \exp(-|\xi|^2/2)$, where $|\xi|^2 = \sum \xi_k^2$. We write $Ef(\xi)$ for the expectation of a function $f = f(\xi) = f(\xi_1, \ldots, \xi_d)$. Recall that

$$E\xi_k^{2p} = (2p-1)!! = (2p-1)(2p-3)\cdots 3 \cdot 1$$

for $k = 1, \ldots, d$ (easy inductive proof via integration by parts), and thus

$$E \prod_{k=1}^{d} \xi_k^{2p_k} = \prod_{k=1}^{d} (2p_k - 1)!!.$$

On $\mathbb{R}^d$, we write $(\cdot, \cdot)$ for the standard Euclidean inner product. We recall the well-known [G, S, Z]

Wick formula *Let $x_1, \ldots, x_n$ be vectors in $\mathbb{R}^d$ with Gram matrix $A = ((x_i, x_j))$. Then*

$$E \prod_{i=1}^{n} (x_i, \xi) = \text{haf } A. \tag{3.9}$$

(For odd n, we define haf $A = 0$.)

Proof. Both sides are multilinear in the x_i, so we may assume that each x_i is an element of the standard orthonormal basis $e_1, \ldots, e_d$. If there is an e_k that occurs an odd number of times among the x_i, then both sides are zero. If each e_k occurs $2p_k$ times, then (3.9) becomes

$$E \prod_{k=1}^{d} \xi_k^{2p_k} = \prod_{k=1}^{d} (2p_k - 1)!!,$$

which is true. □

The following theorems are easy corollaries of Theorem 3.3.4 together with the Wick formula (3.9) and Marcus's theorem (3.4).

Theorem 3.4.1 *If $X_1, \ldots, X_n$ are jointly normal random variables with zero expectation, then*

$$E\left(X_1^2 \cdots X_n^2\right) \geq EX_1^2 \cdots EX_n^2.$$

Equality holds if and only if the X_i are independent or at least one of them is almost surely zero.

Proof. The variables can be written as $X_i = (x_i, \xi)$ with ξ of standard normal distribution and the x_i constant vectors with a positive semi-definite Gram matrix $A = (a_{i,j}) = ((x_i, x_j))$. Then

$$\begin{aligned} E \prod_{i=1}^{n} X_i^2 = E \prod_{i=1}^{n} (x_i, \xi)^2 = \\ = \operatorname{haf} \begin{pmatrix} A & A \\ A & A \end{pmatrix} \geq \operatorname{per} A \geq \prod_{i=1}^{n} a_{i,i} = \\ = \prod_{i=1}^{n} E(x_i, \xi)^2 = \prod_{i=1}^{n} EX_i^2, \end{aligned}$$

with equality if and only if A is a diagonal matrix or has a zero row, i.e., the x_i are pairwise orthogonal or at least one of them is zero. □

The generalization of Theorem 3.4.1 to an arbitrary even exponent $2p$ is equivalent to Conjecture 3.3.5.

Theorem 3.4.2 *For any $x_1, \ldots, x_n \in \mathbb{R}^d$, $|x_i| = 1$, the average of*

$$\prod_{i=1}^{n} (x_i, \xi)^2$$

on the unit sphere $\{\xi \in \mathbb{R}^d : |\xi| = 1\}$ is at least

$$\frac{\Gamma(d/2)}{2^n \Gamma(d/2+n)} = \frac{(d-2)!!}{(d+2n-2)!!} = \frac{1}{d(d+2)(d+4)\ldots(d+2n-2)},$$

with equality if and only if the vectors x_i are pairwise orthogonal.

Proof. The average on the unit sphere is the constant in the theorem times the expectation with respect to the standard Gaussian measure (see e.g. [B1]). By Theorem 3.4.1, the latter expectation is minimal if and only if the x_i are pairwise orthogonal, in which case it is 1. □

Theorem 3.4.3 *For real linear functionals f_i on a real Euclidean space,*

$$||f_1 \cdots f_n|| \geq \frac{||f_1|| \cdots ||f_n||}{\sqrt{n(n+2)(n+4)\cdots(3n-2)}}.$$

Here $||\cdot||$ means supremum of the absolute value on the unit sphere. In the infinite-dimensional case, functionals with infinite norm may be allowed. Then the convention $0 \cdot \infty = 0$ should be used on the right hand side.

Proof. We may assume that the space is $\mathbb{R}^d$ with $d \leq n$, and the functionals are given by $f_i(\xi) = (x_i, \xi)$ with $||f_i|| = |x_i| = 1$. Then $||f_1 \cdots f_n||^2$ is at least the average of

$$\prod_{i=1}^{n} f_i^2(\xi) = \prod_{i=1}^{n} (x_i, \xi)^2$$

on the unit sphere, which by Theorem 3.4.2 and $d \leq n$ is at least

$$1/(n(n+2)(n+4)\cdots(3n-2)).$$

□

It is an unsolved problem, raised by Benítez, Sarantopoulos and Tonge [BST] (1998), whether Theorem 3.4.3 is true with n^n under the square root sign in the denominator on the right hand side. This is called the 'real linear polarization constant' problem.

For the same question in the complex case, the affirmative answer was proved by Arias-de-Reyna [A] in 1998, based on the complex analog of the Wick formula [A, B2, G] and on Lieb's inequality (3.5). The anonymous referee of [F2] called my attention to the fact that Arias-de-Reyna used only the special case of (3.5) where the matrix B' is of rank 1. This is much simpler than (3.5) in general, it can be proved essentially by the argument Marcus used in [Mar1, Mar2] to prove the even more special case of (3.5) where B' is of size 1×1, which is still stronger than Marcus's inequality (3.4). The affirmative answer in the complex case also follows from Keith Ball's solution to the complex plank problem [Ball] (2001).

In the real case, the affirmative answer to the [BST] question above for $n \leq 5$ was proved by Pappas and Révész [PR] in 2004. For general n, the best result known before [F2] was that of Révész and Sarantopoulos [RS] (2004), based on results of [MST], with $(2n)^n/4$ under the square root sign. See [Mat1, Mat2, MM, R] for accounts on this and related questions. Note that

$$\begin{aligned} n(n+2)(n+4)\cdots(3n-2) = \\ = \exp\left(\log n + \log(n+2) + \log(n+4) + \cdots + \log(3n-2)\right) < \\ < \exp\left(\frac{1}{2}\int_n^{3n} \log u \cdot \mathrm{d}u\right) = \\ = \exp\left([u(\log u - 1)]_n^{3n}/2\right) = \\ = \exp((3n\log 3n - 3n - n\log n + n)/2) = \\ = \exp\frac{n(2\log n + 3\log 3 - 2)}{2} = \left(\frac{3\sqrt{3}}{e}n\right)^n, \end{aligned}$$

and $3\sqrt{3}/e < 3 \cdot 1.8/2.7 = 2$, so Theorem 3.4.3 is an improvement on [RS]. Note also that the statement with n^n under the square root sign would follow from Conjecture 3.3.5.

References

[A] J. Arias-de-Reyna, Gaussian variables, polynomials and permanents, Lin. Alg. Appl. 285 (1998), 107–114.

[Ball] K. M. Ball, The complex plank problem, Bull. London Math. Soc. 33 (2001), 433–442.

[B1] A. Barvinok, Estimating L^∞ norms by L^{2k} norms for functions on orbits, Found. Comput. Math. 2 (2002), 393–412.

[B2] A. Barvinok, Integration and optimization of multivariate polynomials by restriction onto a random subspace, arXiv preprint: math.OC/0502298

[BST] C. Benítez, Y. Sarantopoulos, A. Tonge, Lower bounds for norms of products of polynomials, Math. Proc. Camb. Phil. Soc. 124 (1998), 395–408.

[Bo] A. Borel, Linear Algebraic Groups, W. A. Benjamin Inc., New York, 1969.

[C] J. L. Cisneros-Molina, An invariant of 2×2 matrices, Electronic Journal of Linear Algebra 13 (2005), 146–152.

[CP] C. De Concini and C. Procesi, A characteristic free approach to invariant theory, Adv. Math. 21 (1976), 330–354.

[DF1] M. Domokos, P. E. Frenkel, On orthogonal invariants in characteristic 2, J. Algebra 274 (2004), 662–688. Published online: `http://authors.elsevier.com/sd/article/S0021869303005131`, arXiv preprint: `math.RA/0303106`

[DF2] M. Domokos, P. E. Frenkel, Mod 2 indecomposable orthogonal invariants, Adv. Math. 192/1 (2005), 209–217.

[DKZ] M. Domokos, S. G. Kuzmin, A. N. Zubkov, Rings of matrix invariants in positive characteristic, J. Pure Appl. Alg. 176 (2002), 61–80.

[D] D. Ž. Đoković, Simple proof of a theorem on permanents, Glasgow Math. J. 10 (1969), 52–54.

[F0] P. E. Frenkel, Vector invariants of the orthogonal group in characteristic 2 (in Hungarian), MSc thesis,
http://www.cs.elte.hu/math/diploma/math/index.html

[F1] P. E. Frenkel, Character formulae for classical groups, Central European Journal of Mathematics 4 (2006), no. 2, 242–249.

[F2] P. E. Frenkel, Pfaffians, hafnians and products of real linear functionals, Math. Res. Lett., to appear.

[FH] W. Fulton and J. Harris, Representation theory, GTM, Springer, New York, 1991.

[Ga] M. Gaudin, Une démonstration simplifiée du théorème de Wick en Mécanique statistique, Nuclear Phys. 15 (1960), 89–91.

[GW] R. Goodman and N. R. Wallach, Representations and invariants of the classical groups, Cambridge University Press, Cambridge, 1998.

[G] L. Gurvits, Classical complexity and quantum entanglement, J. Comput. System Sci. 69 (2004), no. 3, 448–484.

[H] R. Hartshorne, Algebraic geometry, Graduate texts in mathematics 52, Springer, 1977.

[Ja] J. C. Jantzen, Representations of algebraic groups, Mathematical surveys and monographs 107, American Math. Soc., 2003.

[J] C. R. Johnson, Inequalities for a complex matrix whose real part is positive definite, Trans. Amer. Math. Soc. 212 (1975), 149–154.

[L] E. H. Lieb, Proofs of some conjectures on permanents, J. Math. Mech. 16 (1966), 127–134.

[L1] E. H. Lieb, A theorem on Pfaffians, J. Combinatiorial Theory 5 (1968), 313–319.

[Mar1] M. Marcus, The permanent analogue of the Hadamard determinant theorem, Bull. Amer. Math. Soc. 69 (1963), 494–496.

[Mar2] M. Marcus, The Hadamard theorem for permanents, Proc. Amer. Math. Soc. 15 (1964), 967–973.

[MN] M. Marcus, M. Newman, The permanent function as an inner product, Bull. Amer. Math. Soc. 67 (1961), 223–224.

[Mat1] M. Matolcsi, A geometric estimate on the norm of product of functionals, Lin. Alg. Appl. 405 (2005), 304–310.

[Mat2] M. Matolcsi, The linear polarization constant of $\mathbb{R}^n$, Acta Math. Hungar. 108 (2005), no. 1-2, 129–136.

[MM] M. Matolcsi, G. A. Muñoz, On the real linear polarization constant problem, Math. Inequal. Appl. 9 (2006), no. 3, 485–494.

[Mats] H. Matsumura, Commutative Ring Theory, Cambridge Univ. Press, 1986.

[Mi] H. Minc, Permanents, Encyclopedia of Mathematics and its Applications, Addison-Wesley, 1978

[MST] G. A. Muñoz, Y. Sarantopoulos, A. Tonge, Complexifications of real Banach spaces, polynomials and multilinear maps, Studia Math. 134 (1999), no. 1, 1–33.

[N] P. Newstead, Introduction to moduli problems and orbit spaces, Tata Inst. Lecture Notes, Springer-Verlag, 1978.

[P] C. Procesi, Lie groups — An approach through invariants and representations, Springer, 2007.

[PR] A. Pappas, Sz. Révész, Linear polarization constants..., J. Math. Anal. Appl. 300 (2004), 129–146.

[R] Sz. Gy. Révész, Inequalities for multivariate polynomials, Annals of the Marie Curie Fellowships 4 (2006), `http://www.mariecurie.org/annals/`, arXiv preprint: `math.CA/0703387`

[RS] Sz. Gy. Révész, Y. Sarantopoulos, Plank problems, polarization and Chebyshev constants, J. Korean Math. Soc. 41 (2004) 157–174.

[Ri] D. R. Richman, The fundamental theorems of vector invariants, Adv. Math. 73 (1989), 43–78.

[Se] A. Seidenberg, The hyperplane sections of normal varieties, Trans. Amer. Math. Soc. 69 (1950), 357–386.

[S] B. Simon, The $P(\phi)_2$ Euclidean (Quantum) Field Theory, Princeton Series in Physics, Princeton University Press, 1974

[T] D. E. Taylor, The Geometry of the Classical Groups, Heldermann Verlag, Berlin, 1992.

[W1] H. Weyl, Theorie der Darstellung kontinuerlicher halbeinfacher Gruppen durch lineare Transformationen, I, II, III, und Nachtrag, Math. Zeitschrift 23 (1925) 271–309 and 24 (1925) 328–376, 377–395, 789-791; reprinted in Selecta Hermann Weyl, 262–366, Birkhäuser, Basel, 1956.

[W2] H. Weyl, The classical groups, Their invariants and representations, Princeton University Press, Princeton, 1946.

[Wi] G. C. Wick, The evaluation of the collision matrix, Phys. Rev. 80 (1950), 268–272.

[Z] A. Zvonkin, Matrix integrals and map enumeration: an accesible introduction, Combinatorics and physics (Marseille, 1995), Math. Comput. Modelling 26 (1997), 281–304.

Nyilatkozat önálló munkáról, hivatkozások átvételéről

Alulírott Frenkel Péter kijelentem, hogy ezt a doktori értekezést magam készítettem és abban csak a megjelölt forrásokat használtam fel. Minden olyan részt, amelyet szó szerint, vagy azonos tartalomban, de átfogalmazva más forrásból átvettem, egyértelműen, a forrás megadásával megjelöltem.

Budapest, 2008. január 3.

Nyilatkozat nyilvánosságra hozatalról

Alulírott Frenkel Péter hozzájárulok a doktori értekezésem interneten történő nyilvánosságra hozatalához korlátozás nélkül.

Budapest, 2008. január 3.