11

Invariants

1 Applications to Invariant Theory

In this chapter we shall make several computations of invariants of a group G acting linearly on a space V. In our examples we have that G is algebraic and often we are working over $\mathbb{C}$.

In the general case, V is a vector space over a general field F, and $G \subset GL(V) = GL(n, F)$ is a linear group. One should consider the following basic principles:

(1) Let $\overline{F}$ be the algebraic closure of F and let $\overline{G}$ be the Zariski closure of G in $GL(n, \overline{F})$. Then a polynomial on $V \otimes_F \overline{F}$ is invariant under $\overline{G}$ if and only if it is invariant under G.
(2) Suppose F is infinite. If we find a set of generating polynomial invariants f_i for the action of $\overline{G}$ on $V \otimes_F \overline{F}$, and we suppose that the f_i have coefficients in F, then the f_i form a set of generators for the polynomial invariants under the action of G on V.
(3) If $F = \mathbb{R}$, $\overline{F} = \mathbb{C}$, then the real and the imaginary part of a polynomial invariant with complex coefficients are invariant.

These remarks justify the fact that we often work geometrically over $\mathbb{C}$.

1.1 Cauchy Formulas

Assume now that we have an action of a group G on a space U of dimension m and we want to compute the invariants of n copies of U. Assume first $n \geq m$.

We think of n copies of U as $U \otimes \mathbb{C}^n$ and the linear group $GL(n, \mathbb{C})$ acts on this vector space by tensor action on the second factor, commuting with the G action on U. As we have seen in Chapter 3, the Lie algebra of $GL(n, \mathbb{C})$ acts by polarization operators. The ring of G-invariants is stable under these actions.

From Chapter 9, §6.3.2, the ring $\mathcal{P}(U^n)$ of polynomial functions on $U \otimes \mathbb{C}^n$ equals

$$S(U^* \otimes \mathbb{C}^n) = \bigoplus\nolimits_{\lambda,\, ht(\lambda) \leq m} S_\lambda(U^*) \otimes S_\lambda(\mathbb{C}^n).$$

For the invariants under G we clearly have that

(1.1.1) $$S(U^* \otimes \mathbb{C}^n)^G = \bigoplus_{\lambda,\, ht(\lambda)\le m} S_\lambda(U^*)^G \otimes S_\lambda(\mathbb{C}^n).$$

This formula describes the ring of invariants of G acting on the polynomial ring of U^n as a representation of $GL(n, \mathbb{C})$. In particular we see that the multiplicity of $S_\lambda(\mathbb{C}^n)$ in the ring $\mathcal{P}(U^n)^G$ equals the dimension of the space of invariants $S_\lambda(U^*)^G$.

The restriction on the height implies that when we restrict to $m \le n$ copies we again have the same formula, and hence we deduce the following by comparing isotypic components and by Chapter 9, §6.3.3:

$$S_\lambda(U^*)^G \otimes S_\lambda(\mathbb{C}^n) \cap S(U^* \otimes \mathbb{C}^m) = S_\lambda(U^*)^G \otimes S_\lambda(\mathbb{C}^m).$$

We deduce:

Theorem 1. *If* $\dim U = m$*, the ring of invariants of* $S(U^* \otimes \mathbb{C}^n)$ *is generated, under polarization, by the invariants of* m *copies of* U.

Proof. Each isotypic component (under $GL(n, \mathbb{C})$) $S_\lambda(U^*)^G \otimes S_\lambda(\mathbb{C}^n)$ is generated under this group by $S_\lambda(U^*)^G \otimes S_\lambda(\mathbb{C}^m)$ since $ht(\lambda) \le m$. □

There is a useful refinement of this theorem.

If λ is a partition of height exactly m, we can write it as $\lambda = \mu + k\,1^m$ with $ht(\mu) \le m-1$. Then $S_\lambda(U) \otimes S_\lambda(\mathbb{C}^m) = (\bigwedge^m U \otimes \bigwedge^m \mathbb{C}^m)^k S_\mu(U) \otimes S_\mu(\mathbb{C}^m)$. If the group G is contained in the special linear group the *determinant* of m-vectors, i.e., a generator of the 1-dimensional space $\bigwedge^m U \otimes \bigwedge^m \mathbb{C}^m$ is invariant and we have

$$S_\lambda(U)^G \otimes S_\lambda(\mathbb{C}^m) = \left(\bigwedge^m U \otimes \bigwedge^m \mathbb{C}^m\right)^k S_\mu(U)^G \otimes S_\mu(\mathbb{C}^m).$$

Thus we obtain by the same reasoning:

Theorem 2. *If* $\dim U = m$ *and* $G \subset SL(U)$*, the ring of* G*-invariants of* $S(U^* \otimes \mathbb{C}^n)$ *is generated, under polarization, by the determinant and invariants of* $m-1$ *copies of* U.

Alternatively, we could use the $G \times U^+$ invariants $\oplus S_\lambda(U^*)^G \otimes S_\lambda(\mathbb{C}^n)^{U^+}$. They are contained in the polynomial ring $S(U^* \otimes C^{m-1})[d]$, where d is the determinant of the first m vectors u_i. By the theory of the highest weight, they generate, under polarization, all the invariants.[102]

[102] Every subspace $W \subset S(U^* \otimes \mathbb{C}^n)$ which is stable under polarization is generated by $W \cap S(U^* \otimes \mathbb{C}^m)$. A subspace $W \subset S(U^* \otimes \mathbb{C}^n)$ stable under polarization and multiplication by the determinant d is generated (under polarizations and multiplication by d) by $W \cap S(U^* \otimes \mathbb{C}^{m-1})$.

1.2 FFT for $SL(n, \mathbb{C})$

We shall now discuss the first fundamental theorem of the special linear group $SL(V)$, $V = \mathbb{C}^n$, acting on m copies of the fundamental representation.

In the computation of invariants we will often use two simple ideas.

(1) If two invariants, under a group G coincide on a set X, then they coincide on the set $GX := \{gx \mid g \in G, x \in X\}$.
(2) If two polynomial invariants, under a group G coincide on a set X dense in Y, then they coincide on Y.

We identify V^m with the space of $n \times m$ matrices, with $SL(n, \mathbb{C})$ acting by left multiplication.

We consider the polynomial ring on V^m as the ring $\mathbb{C}[x_{ij}]$, $i = 1, \dots, n$; $j = 1, \dots, m$, of polynomials in the entries of $n \times m$ matrices. Set $X := (x_{ij})$ the matrix whose entries are the indeterminates.

Given n indices $i_1, \dots, i_n$ between $1, \dots, m$, we shall denote by $[i_1, \dots, i_n]$ the determinant of the maximal minor of X extracted from the corresponding columns.

Theorem 1. *The ring of invariants $\mathbb{C}[x_{ij}]^{SL(n,\mathbb{C})}$ coincides with the ring $A := \mathbb{C}[[i_1, \dots, i_n]]$ generated by the $\binom{m}{n}$ elements $[i_1, \dots, i_n]$, $1 \le i_1 < \cdots < i_n \le m$.*

Proof. We can apply the previous theorem using the fact that the proposed ring A is certainly made of invariants, closed under polarization, and by definition it contains the determinant on the first n copies which is $[1, 2, \dots, n]$. Thus it suffices to show that A coincides with the ring of invariants for $n - 1$ copies of V.

Now it is clear that, given any $n - 1$ linearly independent vectors of V, they can be completed to a basis with determinant 1. Hence this set of $n - 1$ tuples, which is open, forms a unique orbit under $SL(V)$. Therefore the invariants of $n - 1$ copies are just the constants, and the theorem is proved. □

It is often convenient to formulate statements on invariants in a geometric way, that is, in the language of quotients (cf. Chapter 14, §3 for details).

Definition. Given an affine variety V with the action of a group G and a map $\pi : V \to W$, we say that π is a quotient under G, and denote $W = V//G$ if the comorphism $\pi^* : k[W] \to k[V]$ is an isomorphism of $k[W]$ to $k[V]^G$.

We consider the space of m-tuples of vectors of dimension n as the space of n-tuples $u_1, \dots, u_n$ of m-dimensional vectors, then the FFT of $SL(n)$ becomes:

Theorem 2. *The map $(u_1, \dots, u_n) \mapsto u_1 \wedge \dots \wedge u_n$ with image the decomposable vectors of $\bigwedge^n \mathbb{C}^m$ is the quotient of $(\mathbb{C}^m)^n$ under $SL(n, \mathbb{C})$.*[103]

Let us understand this ring of invariants as a representation of $GL(m, \mathbb{C})$. Notice that we recover a special case of Theorem 6.6 of Chapter 10.

---

[103] We have not really proved everything; the reader should check that the decomposable vectors are a subvariety.

Corollary. *The space of polynomials of degree k in the elements $[i_1, \dots, i_n]$ as a representation of $GL(m, \mathbb{C})$ equals $S_{k^n}(\mathbb{C}^m)$.*

Proof. In the general formula $S(U^* \otimes \mathbb{C}^m)^{SL(U)} = \bigoplus_\lambda S_\lambda(U^*)^{SL(U)} \otimes S_\lambda(\mathbb{C}^m)$. We see that $S_\lambda(U^*)$ is always an irreducible representation of $SL(U)$, and hence it contains invariants only if it is the trivial representation. This happens if and only if $\lambda := k^n$ and $S_{k^n}(\mathbb{C}^n) = \mathbb{C}$. Thus we have the result that

$$S(U^* \otimes \mathbb{C}^m)^{SL(U)} = \bigoplus_k S_{k^n}(U^*) \otimes S_{k^n}(\mathbb{C}^m) = \bigoplus_k S_{k^n}(\mathbb{C}^m).$$

By comparing degrees we have the result. □

Using the results of Chapter 10, §6.6 we could also describe the quadratic equations satisfied by the invariants. We prefer to leave this task to Chapter 13, where it will be approached in a combinatorial way.

One can in fact combine this theorem with the first fundamental theorem of Chapter 9 §1.4 to get a theorem of invariants for $SL(U)$ acting on $U^m \oplus (U^*)^p$. We describe this space as pairs of matrices $X \in M_{m,n}(\mathbb{C}), Y \in M_{n,p}(\mathbb{C})$, X of rows $\phi_1, \dots, \phi_m$ and Y of columns $u_1, \dots, u_p$. A matrix $A \in SL(n, \mathbb{C})$ acts as (XA^{-1}, AY) and we have invariants

$$\langle \phi_i | u_j \rangle, \ [u_{i_1}, \dots, u_{i_n}], \ [\phi_{j_1}, \dots, \phi_{j_n}].$$

The $\langle \phi_i | u_j \rangle$ are the entries of XY, and the $[u_{i_1}, \dots, u_{i_n}]$ are the determinants of the maximal minors of Y, while $[\phi_{j_1}, \dots, \phi_{j_n}]$ are the determinants of the maximal minors of X.

Theorem 3. *The ring of invariants $\mathcal{P}[U^m \oplus (U^*)^p]^{SL(U)}$ is the ring generated by*

$$\langle \phi_i | u_j \rangle, \ [u_{i_1}, \dots, u_{i_n}], \ [\phi_{j_1}, \dots, \phi_{j_n}].$$

Proof. We apply the methods of §1.1 to U, U^* and reduce to $n-1$ copies of both U and U^*.

In this case, of the invariants described we only have the $\langle \phi_i | u_j \rangle$. Consider the open set in which the $n-1$ vectors $u_1, \dots, u_{n-1}$ are linearly independent. Acting with $SL(U)$ we can transform these vectors to be $e_1, \dots, e_{n-1}$. Let us thus consider the subset

$$W := \{(\phi_1, \dots, \phi_{n-1}; e_1, \dots, e_{n-1})\}.$$

We can still think of the linear forms $\phi_1, \dots, \phi_{n-1}$ as rows of an $n-1 \times n$ matrix X. This set W is stable under the subgroup of $SL(n, \mathbb{C})$:

$$H := \left\{ \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & a_1 \\ 0 & 1 & 0 & \dots & 0 & a_2 \\ & \dots & \dots & \dots & & \dots \\ 0 & 0 & 0 & \dots & 1 & a_{n-1} \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix} \right\}.$$

A nonzero $SL(n, \mathbb{C})$ invariant function restricts to W to a nonzero H-invariant function.

The action of an element of H on a matrix X consists of subtracting, from the last column of X, the linear combination with the coefficients a_i of the first $n-1$ columns.

Thus, on the open set where the first $n-1$ columns are linearly independent, we can make the last column 0 by acting with H. Hence we see that an H-invariant function is independent of the coordinates of the last column. Now notice that the $(n-1)^2$ functions $\langle \phi_i | u_j \rangle$ restrict on W to the functions $\langle \phi_i | e_j \rangle$, $j \leq n-1$, which give all the coordinates of the first $n-1$ columns of X. It follows then that, given any invariant function f, there is a polynomial $g(\langle \phi_i | u_j \rangle)$ which coincides with f on W. By the initial discussion we must have that $f = g(\langle \phi_i | u_j \rangle)$. □

2 The Classical Groups

2.1 FFT for Classical Groups

We start here the description of the representation theory of other classical groups, in particular the orthogonal and the symplectic group; again we relate invariant theory with representation theory by the same methods used for the linear group.

The proofs work over any field F of characteristic 0. In practice we think of $F = \mathbb{R}, \mathbb{C}$.

We fix a vector space V (over F) with a nondegenerate invariant bilinear form. Let us denote by (u, v) a symmetric bilinear form, and by $O(V)$ the orthogonal group fixing it.

Let us denote by $[u, v]$ an antisymmetric form and by $Sp(V)$ the symplectic group fixing it.

We start by remarking that for the fundamental representation of either one of these two groups we have a nondegenerate invariant bilinear form which identifies this representation with its dual. Thus for the first fundamental theorem it suffices to analyze the invariants of several copies of the fundamental representation. When convenient we identify vectors, or exterior products of vectors, with functions.

The symplectic group $Sp(V)$ on a $2n$-dimensional vector space is formed by unimodular matrices (Chapter 5, §3.6.2). Fix a symplectic basis for which the matrix of the skew form is $J = \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix}$. Given $2n$ vectors u_i which we write as column vectors of a $2n \times 2n$ matrix A, the determinant of A equals the Pfaffian of the matrix $A^t J A$ which has as entries the skew products $[u_i, u_j]$, (Chapter 5, §3.6.2). Hence

Lemma. *The determinant of $A = [u_1, \dots, u_{2n}]$ equals the Pfaffian of the matrix $A^t J A$.*

The orthogonal group instead contains a subgroup of index 2, the special linear group $SO(V)$ formed by the orthogonal matrices of determinant 1.

When we discuss either the symplectic or the special orthogonal group we assume we have chosen a trivialization $\bigwedge^{\dim V} V = \mathbb{C}$ of the top exterior power of V.

If $m = \dim V$ and $v_1, \dots, v_m$ are m vector variables, the element $v_1 \wedge \dots \wedge v_m$ is to be understood as a function on $V^{\oplus m}$ invariant under the special linear group.

Given a direct sum $V^{\oplus k}$ of copies of the fundamental representation, we denote by $u_1, \dots, u_k$ a typical element of this space.

Theorem. *Let* $\dim V = n$*:*

(i) The ring of invariants of several copies of the fundamental representation of $SO(V)$ is generated by the scalar products (u_i, u_j) and by the determinants $u_{i_1} \wedge u_{i_2} \wedge \dots \wedge u_{i_n}$.
(ii) The ring of invariants of several copies of the fundamental representation of $O(V)$ is generated by the scalar products (u_i, u_j).
(iii) The ring of invariants of several copies of the fundamental representation of $Sp(V)$ is generated by the functions $[u_i, u_j]$.

Before proving this theorem we formulate it in the language of matrices and quotients.

Consider the group $O(n, \mathbb{C})$ of $n \times n$ matrices X with $X^t X = X X^t = 1$ and consider the space of $n \times m$ matrices Y with the action of $O(n, \mathbb{C})$ given by multiplication XY. Then:

The mapping $Y \mapsto Y^t Y$ from the variety of $n \times m$ matrices to the symmetric $m \times m$ matrices of rank $\leq n$ is a quotient under the orthogonal group.

Similarly, let

$$J_n := \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix}$$

be the standard $2n \times 2n$ skew-symmetric matrix and $Sp(2n, \mathbb{C})$ the standard symplectic group of the matrices X such that $X^t J_n X = J_n$. Then consider the space of $2n \times m$ matrices Y with the action of $Sp(2n, \mathbb{C})$ given by multiplication XY. Then:

The mapping $Y \to Y^t J_n Y$ from the space of $2n \times m$ matrices to the variety of antisymmetric $m \times m$ matrices of rank $\leq 2n$ is a quotient under the symplectic group.

Proof of the Theorem for $SO(V)$, $O(V)$. We prove first that the theorem for $SO(V)$ implies the theorem for $O(V)$.

One should remark that since $SO(V)$ is a normal subgroup of index 2 in $O(V)$, we have a natural action of the group $O(V)/SO(V) \cong \mathbb{Z}/(2)$ on the ring of $SO(V)$ invariants.

Let τ be the element of $\mathbb{Z}/(2)$ corresponding to the orthogonal transformations of determinant -1 (*improper transformations*). The elements (u_i, u_j) are invariants of this action while $\tau(u_{i_1} \wedge u_{i_2} \wedge \dots u_{i_m}) = -u_{i_1} \wedge u_{i_2} \wedge \dots u_{i_m}$. It follows that the orthogonal invariants are polynomials in the special orthogonal invariants in which every monomial contains a product of an even number of elements of type $u_{i_1} \wedge u_{i_2} \wedge \dots u_{i_n}$.

Thus it is enough to verify the following identity:

$$(u_{i_1} \wedge u_{i_2} \wedge \dots u_{i_n})(u_{j_1} \wedge u_{j_2} \wedge \dots u_{j_n})$$

$$= \det \begin{pmatrix} (u_{i_1}, u_{j_1}) & (u_{i_1}, u_{j_2}) & \dots & (u_{i_1}, u_{j_n}) \\ (u_{i_2}, u_{j_1}) & (u_{i_2}, u_{j_2}) & \dots & (u_{i_2}, u_{j_n}) \\ \vdots & \vdots & \vdots & \vdots \\ (u_{i_n}, u_{j_1}) & (u_{i_n}, u_{j_2}) & \dots & (u_{i_n}, u_{j_n}) \end{pmatrix}.$$

This is easily verified, since in an orthonormal basis the matrix having as rows the coordinates of the vectors u_i times the matrix having as columns the coordinates of the vectors u_j yields the matrix of scalar products.

Now we discuss $SO(V)$. Let A be the proposed ring of invariants. From the definition, this ring contains the determinants and it is closed under polarization operators.

From §1.1 we deduce that it is enough to prove the Theorem for $n-1$ copies of the fundamental representation. We work by induction on n and can assume $n > 1$.

We have to use one of the two possible reductions.

First We first prove the theorem for the case of real invariants on a real vector space $V := \mathbb{R}^n$ with the standard Euclidean norm.

This method is justified by the following analysis that we leave to the reader to justify. Suppose that we have an algebraic group G acting on a complex space $V_{\mathbb{C}}$ which is the complexification of a real space V. Assume also that we have a real subgroup H of G which acts on V and which is Zariski dense in G. Given a polynomial f on $V_{\mathbb{C}}$, we have that f is invariant under G if and only if it is invariant under H. Such a polynomial can be uniquely decomposed into $f_0(v) + if_1(v)$ where both f_0 and f_1 have real coefficients. Moreover, f is H-invariant if and only if both f_0 and f_1 are H-invariant. Finally, a polynomial with real coefficients is H-invariant as a function on $V_{\mathbb{C}}$ if and only if it is invariant as a function on V. The previous setting applies to $SO(n, \mathbb{R}) \subset SO(n, \mathbb{C})$ and $(\mathbb{R}^n)^m \subset (\mathbb{C}^n)^m$.

Let e_i denote the canonical basis of $\mathbb{R}^n$ and consider $\overline{V} := \mathbb{R}^{n-1}$ formed by the vectors with the last coordinate 0 (and spanned by e_i, $i < n$).

We claim that any special orthogonal invariant E on V^{n-1} restricted to $\overline{V}^{n-1}$ is an invariant under the orthogonal group of $\overline{V}$: in fact, it is clear that every orthogonal transformation of $\overline{V}$ can be extended to a special orthogonal transformation of V.

By induction, therefore, we have a polynomial $F((u_i, u_j))$ which, restricted to $\overline{V}^{n-1}$, coincides with E. We claim that F, E coincide for every choice of $n-1$ vectors $u_1, \dots, u_{n-1}$.

For any such choice there is a vector u of norm 1 and orthogonal to these vectors. There is a special orthogonal transformation which brings this vector u into e_n and thus the vectors u_i into the space $\mathbb{R}^{n-1}$. Since both F, E are invariant and they coincide on $\mathbb{R}^{n-1}$, the claim follows.

Second If one does not like the reduction to the real case, one can argue as follows. Prove first that the set of $n-1$ tuples of vectors which span a nondegenerate subspace in V are a dense open set of V^{n-1}, and then argue as before. □

Proof of the Theorem for $Sp(V)$. Again let R be the proposed ring of invariants for $Sp(V)$, $\dim V = 2m$. From the remark on the Pfaffian (Lemma 2.1) we see that R contains the determinants, and it is closed under polarization operators.

From §1.1 we deduce that it is enough to prove the theorem for $2m-1$ copies of the fundamental representation. We work by induction on m. For $m = 1$, $Sp(2, \mathbb{C}) = SL(2, \mathbb{C})$ and the theorem is clear. Assume we have chosen a symplectic basis e_i, f_i, $i = 1, \dots, m$, and consider the space of vectors $\overline{V}$ having coordinate 0 in e_1.

On this space the symplectic form is degenerate with kernel spanned by f_1 and it is again nondegenerate on the subspace W where both the coordinates of e_1, f_1 vanish.

Claim A symplectic invariant $F(u_1, \dots, u_{2m-1})$, when computed on elements $u_i \in \overline{V}$, is a function which depends only on the coordinates in e_i, f_i, $i > 1$.

To prove the claim, consider the symplectic transformations $e_1 \mapsto te_1$, $f_1 \mapsto t^{-1}f_1$ and identity on W. These transformations preserve $\overline{V}$, induce multiplication by t on the coordinate of f_1, and fix the other coordinates. If a polynomial is invariant under this group of transformations, it must be independent of the coordinate of f_1, hence the claim.

Since $F(u_1, \dots, u_{2m-1})$ restricted to W^{2m-1} is invariant under the symplectic group of W, by induction there exists a polynomial $G([u_i, u_j])$ which coincides with F on W^{2m-1} and by the previous claim, also on $\overline{V}^{2m-1}$. We claim that $G([u_i, u_j]) = F(u_1, \dots, u_{2m-1})$ everywhere.

It is enough by continuity to show this on the set of $2m-1$ vectors which are linearly independent. In this case such a set of vectors generates a subspace where the symplectic form is degenerate with a 1-dimensional kernel. Hence, by the theory of symplectic forms, there is a symplectic transformation which brings this subspace to $\overline{V}$ and the claim follows. □

We can now apply this theory to representation theory.

3 The Classical Groups (Representations)

3.1 Traceless Tensors

We start from a vector space with a nondegenerate symmetric or skew form, denoted $(\ ,\ )$, $[\ ,\ ]$ respectively. Before we do any further computations we need to establish a basic dictionary deduced from the identification $V = V^*$ induced by the form.

We first want to study the identification $\mathrm{End}(V) = V \otimes V^* = V \otimes V$ and treat $V \otimes V$ as operators. We make explicit some formulas in the two cases (easy verification):

$$(u \otimes v)(w) = u(v, w), \ (u \otimes v) \circ (w \otimes z) = u \otimes (v, w)z,$$
$$\operatorname{tr}(u \otimes v) = (u, v), \tag{3.1.1}$$

$$(u \otimes v)(w) = u[v, w], \ (u \otimes v) \circ (w \otimes z) = u \otimes [v, w]z,$$
$$\operatorname{tr}(u \otimes v) = -[u, v]. \tag{3.1.2}$$

Furthermore, for the adjoint case we have $(u \otimes v)^* = v \otimes u$ in the orthogonal case and $(u \otimes v)^* = -v \otimes u$ in the symplectic case. Now we enter a more interesting area: we want to study the tensor powers $V^{\otimes n}$ under the action of $O(V)$ or of $Sp(V)$.

We already know that these groups are linearly reductive (Chapter 7, §3.2). In particular all the tensor powers are completely reducible and we want to study these decompositions.

Let us denote by G one of the two previous groups. We use the notation of the symmetric case but the discussion is completely formal and it applies also to the skew case.

First, we have to study $\hom_G(V^{\otimes h}, V^{\otimes k})$. From the basic principle of Chapter 9, §1.1 we identify

$$\hom_G(V^{\otimes h}, V^{\otimes k}) = [V^{\otimes h} \otimes (V^*)^{\otimes k}]^G = [(V^{\otimes h+k})^*]^G. \tag{3.1.3}$$

Thus the space of intertwiners between $V^{\otimes h}$, $V^{\otimes k}$ can be identified with the space of multilinear invariants in $h + k$ vector variables.

Explicitly, on each $V^{\otimes p}$ we have the scalar product

$$(w_1 \otimes \cdots \otimes w_p, z_1 \otimes \cdots \otimes z_p) := \prod_i (w_i, z_i)$$

and we identify $A \in \hom_G(V^{\otimes h}, V^{\otimes k})$ with the invariant function

$$\psi_A(X \otimes Y) := (A(X), Y), \ X \in V^{\otimes h}, \ Y \in V^{\otimes k}. \tag{3.1.4}$$

It is convenient to denote the variables as $(u_1, \dots, u_h, v_1, \dots, v_k)$. Theorem 2.1 implies that these invariants are spanned by suitable monomials (the multilinear ones) in the scalar or skew products between these vectors. In particular there are nontrivial intertwiners if and only if $h + k = 2n$ is even.

It is necessary to identify some special intertwiners.

Contraction The map $V \otimes V \to \mathbb{C}$, given by $u \otimes v \to (u, v)$, is called an *elementary contraction.*

Extension By duality in the space $V \otimes V$, the space of G-invariants is one-dimensional; a generator can be exhibited by choosing a pair of dual bases $(e_i, f_j) = \delta_{ij}$ and setting $I := \sum_i e_i \otimes f_i$. The map $\mathbb{C} \to V \otimes V$ given by $a \to aI$ is an *elementary extension.*

We remark that since $u := \sum_i (u, f_i)e_i = \sum_i (u, e_i) f_i$, we have[104]

$$(I, u_1 \otimes u_2) = (u_1, u_2). \tag{3.1.5}$$

[104] In the skew case $u := \sum_i (u, f_i)e_i = -\sum_i (u, e_i) f_i$.

So I is identified with the given bilinear form. One can easily extend these maps to general tensor powers and consider the contractions and extensions

(3.1.6) $$c_{ij} : V^{\otimes k} \to V^{\otimes k-2}, \ e_{ij} : V^{\otimes k-2} \to V^{\otimes k}$$

given by contracting in the indices i, j or inserting in the indices i, j (e.g., $e_{13} : V \to V^{\otimes 3}$ is $v \mapsto \sum_i e_i \otimes v \otimes f_i$).

Remark. Notice in particular that c_{ij} is surjective and e_{ij} is injective.

We have $c_{ij} = c_{ji}, \ e_{ij} = e_{ji}$ in the symmetric case, while $c_{ij} = -c_{ji}, \ e_{ij} = -e_{ji}$ in the skew-symmetric case.

In order to keep some order in these maps it is useful to consider the two symmetric groups S_h and S_k which act on the two tensor powers commuting with the group G and as orthogonal (or symplectic) transformations.

Thus $S_h \times S_k$ acts on $\hom_G(V^{\otimes h}, V^{\otimes k})$ with $(\sigma, \tau)A := \tau A \sigma^{-1}$. We also have an action of $S_h \times S_k$ on the space $[(V^{\otimes h+k})^*]^G$ of multilinear invariants by the inclusion $S_h \times S_k \subset S_{h+k}$.

We need to show that the identification $\psi : \hom_G(V^{\otimes h}, V^{\otimes k}) = [(V^{\otimes h+k})^*]^G$ is $S_h \times S_k$-equivariant. The formula $\psi_A(X \otimes Y) := (A(X), Y)$ gives

$$\begin{aligned} ((\sigma, \tau)\psi_A)(X \otimes Y) &:= \psi_A(\sigma^{-1}X \otimes \tau^{-1}Y) = (A(\sigma^{-1}X), \tau^{-1}Y) \\ &= (\tau A(\sigma^{-1}X), Y), \end{aligned} \tag{3.1.7}$$

as required.

Consider now a multilinear monomial in the elements $(u_i, u_j), \ (v_h, v_k), \ (u_p, v_q)$. In this monomial the $h + k = 2n$ elements $(u_1, \dots, u_h, v_1, \dots, v_k)$ each appear once and the monomial itself is described by the combinatorics of the n pairings.

Suppose we have exactly a pairings of type (u_i, v_j).[105] Then $h - a, \ k - a$ are both even and the remaining pairings are all *homosexual.*

It is clear that under the action of the group $S_h \times S_k$, this invariant can be brought to the following canonical form:

(3.1.8)
$$J_a := (u_1, v_1)(u_2, v_2) \dots (u_a, v_a) \prod_{i=1}^{(h-a)/2} (u_{a+2i-1}, u_{a+2i}) \prod_{j=1}^{(k-a)/2} (v_{a+2j-1}, v_{a+2j}).$$

Lemma. *The invariant J_a corresponds to the intertwiner*

(3.1.9)
$$C_a : u_1 \otimes u_2 \otimes \cdots \otimes u_h \mapsto \prod_{i=1}^{(h-a)/2} (u_{a+2i-1}, u_{a+2i}) u_1 \otimes \cdots \otimes u_a \otimes I^{\otimes (h-a)/2}.$$

[105] Weyl calls these pairings heterosexual and the others homosexual.

Proof. We compute explicitly

$$\Big(\prod_{i=1}^{(h-a)/2} (u_{a+2i-1}, u_{a+2i}) u_1 \otimes \cdots \otimes u_a \otimes I^{\otimes (h-a)/2}, v_1 \otimes v_2 \otimes \cdots \otimes v_k \Big) = J_a$$

from iteration of 3.1.5. □

The main consequence of the lemma is:

Theorem 1. *Any intertwiner between $V^{\oplus p}$ and $V^{\oplus q}$ is a composition of symmetries, contractions and extensions.*

Remark. It is convenient, in order to normalize this composition, to perform first all the contractions and after all the extensions.

We have seen that there are no intertwiners between $V^{\otimes h}$, $V^{\otimes k}$ if $h + k$ is odd but there are injective G-equivariant maps $V^{\otimes h-2a} \to V^h$ for all $a \leq h/2$.

In particular we can define the subspace $T^0(V^h)$ as the sum of all the irreducible representations of G which do not appear in the lower tensor powers $V^{\otimes h-2a}$. We claim:

Theorem 2. *The space $T^0(V^h)$ is the intersection of all the kernels of the maps c_{ij}. It is called the space of* ***traceless tensors****.*

Proof. If an irreducible representation M of G appears both in $V^{\otimes h-2a}$, $a > 0$, and in V^h, by semisimplicity an isomorphism between these two submodules can be extended to a nonzero intertwiner between V^h, $V^{\otimes h-2a}$.

From the previous theorem, these intertwiners vanish on $T^0(V^h)$, and so all the irreducible representations in $T^0(V^h)$ do not appear in $V^{\otimes h-2a}$, $a > 0$. The converse is also clear: if a contraction does not vanish on an irreducible submodule N of $V^{\otimes h}$, then the image of N is an isomorphic submodule of $V^{\otimes h-2}$. Thus we may say that $T^0(V^h)$ contains all the *new* representations of G in $V^{\otimes h}$. □

In particular, $T^0(V^h)$ is a sum of isotypic components and we may study the restriction of the centralizer of G in $V^{\otimes h}$ to $T^0(V^h)$.

Proposition. *$T^0(V^h)$ is stable under the action of the symmetric group S_h, which spans the centralizer of G in $T^0(V^h)$.*

Proof. Since clearly $\sigma c_{ij} \sigma^{-1} = c_{\sigma(i)\sigma(j)}$, $\forall i, j, \sigma \in S_h$, the first claim is clear.

Since the group is linearly reductive, any element of the centralizer of G acting on $T^0(V^h)$ is the restriction of an element of the centralizer of G acting on $V^{\otimes h}$.

From Theorem 1, these elements are products of symmetries, contractions and extensions. As soon as at least one contraction appears, the operator vanishes on $T^0(V^h)$. Hence only the symmetries are left. □

Thus, we have again a situation similar to the one for the linear group, except that the space $T^0(V^h)$ is a more complicated object to describe. Our next task is to decompose

$$T^0(V^h) = \bigoplus\nolimits_{\lambda \vdash h} U_\lambda \otimes M_\lambda,$$

where the M_λ are the irreducible representations of S_h, which are given by the theory of Young symmetrizers, and the U_λ are the corresponding new representations of G. We thus have to discover which λ appear. In order to do this we will have to work out the second fundamental theorem.

3.2 The Envelope of $O(V)$

Let us complete the analysis with a simple presentation of the algebra U_p spanned by the orthogonal group acting on tensor space $V^{\otimes p}$, for some p.

By definition this algebra is contained in the algebra A_p spanned by the linear group acting on $V^{\otimes p}$ and this, in turn, is formed by the symmetric elements of $\mathrm{End}(V)^{\otimes p}$.

We have already, by complete reducibility, that U_p is the centralizer of $\mathrm{End}_{O(V)}(V^{\otimes p})$, and by the analysis of the previous section one sees immediately that this centralizer is generated by the symmetric group S_p and a single operator:

(3.2.1) $$C : v_1 \otimes v_2 \otimes \cdots \otimes v_p \to (v_1, v_2) I \otimes v_3 \otimes \cdots \otimes v_p.$$

Thus we want to understand the condition for an element of A_p to commute with C. We claim that this is equivalent to the following system of linear equations. Define the following linear map of $\mathrm{End}(V)^{\otimes p}$ to $\mathrm{End}(V)^{\otimes (p-1)}$:

$$\pi : A_1 \otimes A_2 \otimes A_3 \otimes \cdots \otimes A_p \to A_1 A_2^t \otimes A_3 \otimes \cdots \otimes A_p.$$

Theorem. *An element $X \in A_p$ is also in U_p if and only if $\pi(X)$ is of the form $\alpha 1_V \otimes Y$, $Y \in \mathrm{End}(V)^{\otimes(p-2)}$, $\alpha \in F$ a scalar.*

Proof. First, it is clear that the set of elements satisfying the given condition forms a linear space containing the elements $A^{\otimes p}$, $A \in O(V)$. Let us verify indeed that the condition is just the condition to commute with C. As usual we can work on decomposable elements $A_1 \otimes A_2 \otimes Y$. The condition to commute with C is that $(A_1 u, A_2 v) I = (u, v) A_1 \otimes A_2 I$, that is, taking vectors a, b, that

$$\begin{aligned}(a, b)(A_1 u, A_2 v) &= (a \otimes b, (A_1 u, A_2 v) I) = (u, v)(a \otimes b, A_1 \otimes A_2 I) \\ &= (u, v)(A_1^t a, A_2^t b)\end{aligned}$$

implies that $B := A_1 A_2^t$ satisfies $(a, b)(u, B^t v) = (u, v)(a, Bb)$. Taking the a, b, u, v from an orthonormal basis, we deduce the identities for the entries x_{ij} of B:

$$x_{h,k}\delta_i^j = x_{i,j}\delta_h^k \implies x_{i,i} = x_{h,h}, \quad x_{i,j} = 0 \quad \text{if} \quad i \neq j.$$

That is, B is a scalar matrix. $\square$

4 Highest Weights for Classical Groups

4.1 Example: Representations of $SL(V)$

Let V be a vector space with a given basis $e_1, \dots, e_n$. To V we associate the torus of diagonal matrices and the Borel subgroup B of upper triangular matrices.

Proposition. *The vector $e_1 \wedge e_2 \wedge \dots \wedge e_k$ is a highest weight vector, for B, of the exterior power $\bigwedge^k V$, as an $SL(V)$ module, of weight ω_k.*

Proof. Apply

$$e_{i,j}(e_1 \wedge e_2 \wedge \dots \wedge e_k) = \sum_{h=1}^{k} e_1 \wedge e_2 \wedge \dots e_{i,j}(e_h) \wedge \dots \wedge e_k$$

$$= \sum_{h=1}^{k} e_1 \wedge e_2 \wedge \dots \delta_{j,h} e_i \wedge \dots \wedge e_k,$$

which is 0, since, if $\delta_{j,h} \neq 0$ we obtain two factors e_i in the product. When we apply a diagonal matrix $\sum \alpha_i e_{i,i}$ to the vector we obtain the weight $\omega_k = \sum_{i=1}^k \alpha_i$. □

For $sl(n, \mathbb{C})$ consider the representation on a tensor power associated to a partition $\lambda := h_1, h_2, \dots, h_n$ and dual partition $n_1, n_2, \dots, n_t = 1^{a_1} 2^{a_2} \dots n^{a_n}$. In Chapter 9, §3.1, the formulas 3.1.1, 3.1.2 produce the tensor

$$a_T U = (e_1 \wedge e_2 \wedge \dots \wedge e_{n_1}) \otimes (e_1 \wedge e_2 \wedge \dots \wedge e_{n_2}) \otimes \dots \otimes (e_1 \wedge e_2 \wedge \dots \wedge e_{n_t})$$

in this representation. One easily verifies that $a_T U$ is a highest weight vector of weight

$$\omega_\lambda := \sum_{j=1}^{n-1} a_j \omega_j. \tag{4.1.1}$$

Set $V = \mathbb{C}^n$. The previous statement is a consequence of the previous proposition and Chapter 10, §5.2 Proposition 3. In fact by construction $a_T U$ is the tensor product of the highest weight vectors in the factors of $V^{\otimes a_1} \otimes (\bigwedge^2 V)^{\otimes a_2} \otimes \dots (\bigwedge^n V)^{\otimes a_n}$. The factors $\bigwedge^n V$ do not intervene since they are the trivial representation of $sl(n, \mathbb{C})$.

Remark. Notice that $h_i = \sum_{j=i}^n a_j$, so that the sum in the lattice of weights correspond to the sum of the sequences $\lambda = (h_1, h_2, \dots, h_n)$ as vectors with n coordinates. This should not be confused with our use of the direct sum of partitions in Chapter 8, §4.2.

Remark. If we think of the weights of the Cartan subalgebra also as weights for the maximal torus, we can see that the highest weight gives the leading term in the character, which for $SL(n, \mathbb{C})$ is the Schur function.

One should remark that when we take a $2n$-dimensional symplectic space with basis $e_1, \dots, e_n, f_1, \dots, f_n$, the elements $e_1 \wedge e_2 \wedge \dots \wedge e_k$, $k \leq n$ are still highest weight vectors of weight ω_i but, as we shall see in §6.6, the exterior powers are no longer irreducible. Moreover, for the orthogonal Lie algebras, besides the exterior powers, we need to discuss the spin representations.

4.2 Highest Weights and U-Invariants

We now pass to groups. Let G be a semisimple group. Fix a Cartan subalgebra $\mathfrak{t}$ a set of positive roots and the corresponding Lie algebras $\mathfrak{u}^+$, $\mathfrak{b}^+$. The three corresponding algebraic subgroups of G are denoted by T, U^+, B^+. One has that $B^+ = U^+T$ as a semidirect product (in particular $T = B^+/U^+$) (Chapter 10, §6.4).

As an example in $SL(n, \mathbb{C})$ we have that T is the subgroup of diagonal matrices, B^+ the subgroup of upper triangular matrices and U^+ the subgroup of strictly upper triangular matrices, that is, the upper triangular matrices with 1 on the diagonal or, equivalently, with all eigenvalues 1 (unipotent elements).

For an irreducible representation of G, the highest weight vector relative to B^+ is the unique (up to scalars) vector v invariant under U^+; it is also an eigenvector under B^+ and hence T.

Thus v determines a multiplicative character on B^+.[106] Notice that any multiplicative algebraic character of B^+ is trivial on U^+ (by Chapter 7, §1.5) and it is just induced by an (algebraic) character of $T = B^+/U^+$.

In the following, when we use the word character, we mean an algebraic multiplicative character.

The geometric interpretation of highest weights is obtained as follows.

In Chapter 10, §6.9 we extended the notion of highest weight to reductive groups and described the representations. Consider an action of a reductive group G on an affine algebraic variety V. Let $A[V]$ be the coordinate ring of V which is a rational representation under the induced action of G. Thus $A[V]$ can be decomposed into a direct sum of irreducible representations.

If $f \in A[V]$ is a highest weight vector of some weight λ, we have for every $b \in B^+$ that $bf = \lambda(b)f$, and conversely, a function f with this property is a highest weight vector.

Notice that if f_1, f_2 are highest weight vectors of weight λ_1, λ_2, then $f_1 f_2$ is a highest weight vector of weight $\lambda_1 + \lambda_2$. Now unless f is invertible, the set

$$S_f := \{x \in V \mid f(x) = 0\}$$

is a hypersurface of V and it is clearly stable under B^+. Conversely, if V satisfies the property that every hypersurface is defined by an equation (for instance if V is an affine space) we have that to a B^+-stable hypersurface is associated a highest weight vector.

Of course, as usual in algebraic geometry, this correspondence is not bijective, but we have to take into consideration multiplicities.

Lemma. *If $A[V]$ is a unique factorization domain, then we have a 1–1 correspondence between irreducible B^+-stable hypersurfaces and irreducible (as polynomials) highest weight vectors (up to a multiplicative scalar factor).*

106 Recall that we define multiplicative character of a group G to be any homomorphism of G to the multiplicative group $\mathbb{C}^*$.

Proof. It is enough to show that if $f = \prod_i g_i$ is a highest weight vector, say of weight χ, factored into irreducible polynomials, then the g_i are also highest weight vectors.

For this take an element $b \in B^+$. We have $\chi(b)f = (bf) = \prod_i (bg_i)$. Since B^+ acts as a group of automorphisms, the bg_i are irreducible, and thus the elements bg_i must equal the g_j up to some possible permutation and scalar multiplication. Since the action of B^+ on $A[V]$ is rational there is a B^+-stable subspace $U \subset A[V]$ containing the elements g_i.

Consider the induced action of B^+ on the projective space of lines of U. By assumption the lines through the elements g_i are permuted by B^+. Since B^+ is connected, the only possible algebraic actions of B^+ on a finite set are trivial. It follows that the g_i are eigenvectors under B^+. One deduces, from the previous remarks, that they are U^+-invariant and $bg_i = \chi_i(b)g_i$, where χ_i are characters of B^+ and in fact, for the semisimple part of G, are dominant weights. □

4.3 Determinantal Loci

Let us analyze now, as a first elementary example, the orbit structure of some basic representations.

We start with $\hom(V, W)$ thought of as a representation of $GL(V) \times GL(W)$. It is convenient to introduce bases and use the usual matrix notation.

Let n, m be the dimensions of V and W. Using bases we identify $\hom(V, W)$ with the space $M_{m\,n}$ of rectangular matrices. The group $GL(V) \times GL(W)$ is also identified to $GL(n) \times GL(m)$ and the action on $M_{m\,n}$ is $(A, B)X = BXA^{-1}$.

The notion of the rank of an operator is an invariant notion. Furthermore we have:

Proposition. *Two elements of* $\hom(V, W)$ *are in the same* $GL(V) \times GL(W)$ *orbit if and only if they have the same rank.*

Proof. This is an elementary fact. One can give an abstract proof as follows.

Given a matrix C of rank k, choose a basis of V such that the last $n - k$ vectors are a basis of its kernel. Then the image of the first k vectors are linearly independent and we can complete them to a basis of W. In these bases the operator has matrix (in block form):

$$\begin{pmatrix} 1_k & 0 \\ 0 & 0 \end{pmatrix},$$

where 1_k is the identity matrix of size k. This matrix is obtained from C by the action of the group, and so it is in the same orbit. We have a canonical representative for matrices of rank k. In practice this abstract proof can be made into an effective algorithm, for instance, by using Gaussian elimination on rows and columns. □

As a consequence we also have: Consider $V \otimes W$ as a representation of $GL(V) \times GL(W)$. Then there are exactly $\min(n, m) + 1$ orbits, formed by the tensors which can be expressed as sum of k decomposable tensors (and not less), $k = 0, \ldots, \min(m, n)$.

This is left to the reader using the identification $V \otimes W = \hom(V^*, W)$. We remark that these results are quite general and make no particular assumptions on the field F.

We suggest to the reader a harder exercise which is in fact quite interesting and has far-reaching generalizations.

Exercise. Consider again the space of $m \times n$ matrices but restrict the action to $B^+(m) \times B^-(n)$ where $B^+(m)$ (resp. $B^-(n)$) is the group of upper (resp. of lower) triangular matrices, and prove that also in this case there are finitely many orbits. This is a small generalization of the Bruhat decomposition (Chapter 10, §6.4):

The orbits of $B^+(n) \times B^-(n)$ acting on $GL(n)$ by $(A, B)X := AXB^{-1}$ are in 1–1 correspondence with the symmetric group S_n.

4.4 Orbits of Matrices

Let us now consider the action on bilinear forms, restricting to ϵ symmetric forms on a vector space U over a field F. Representing them as matrices, the action of the linear group is $(A, X) \mapsto AXA^t$.

For antisymmetric forms on U, the only invariant is again the rank, which is necessarily even. The rank classifies symmetric forms if F is algebraically closed, otherwise there are deeper arithmetical invariants. For instance, in the special case of the real numbers, the signature is a sufficient invariant.[107]

The proof is by induction. If an antisymmetric form is nonzero we can find a pair of vectors e_1, f_1 on which the matrix of the form is $\begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix}$. If V is the span of these vectors, the space U decomposes into the orthogonal sum $V \oplus V^\perp$. Then one proceeds by induction on $V^\perp$ until one reaches a complement where the form is 0 and one chooses an arbitrary basis v_j of it, getting a basis $e_i, f_i, i = 1, \dots, k, v_j$. In the dual basis we have a canonical representative of the form as $\sum_{i=1}^k e^i \wedge f^i$.

For a nonzero symmetric form (over an algebraically closed field) instead, one can choose a vector e_1 of norm 1 and proceed with the orthogonal decomposition until one has a space where the form is 0. With obvious notation the form is $\sum_{i=1}^k (e^i)^2$.

Summarizing, we have seen in particular the orbit structure for the following representations:

1. The space $M_{n,m}(\mathbb{C})$ of $n \times m$ matrices X with the action of $GL(n, \mathbb{C}) \times GL(m, \mathbb{C})$ given by $(A, B)X := AXB^{-1}$.
2. The space $M_n^+(\mathbb{C})$ of symmetric $n \times n$ matrices X with the action of $GL(n, \mathbb{C})$ given by $A.X := AXA^t$.
3. The space $M_n^-(\mathbb{C})$ of skew-symmetric $n \times n$ matrices X with the action of $GL(n, \mathbb{C})$ given by $A.X := AXA^t$.

In each case there are only finitely many orbits under the given group action, and two matrices are in the same orbit if and only if they have the same rank.

[107] The theory of quadratic forms over $\mathbb{Q}$ or $\mathbb{Z}$ is a rather deep part of arithmetic.

Exercise. (i) If V_k denotes the set of matrices of rank k (where k must be even in case 3), we have that the closure $\overline{V}_k$ is

$$\overline{V}_k := \cup_{j \leq k} V_j. \tag{4.4.1}$$

(ii) The varieties $\overline{V}_k$ are the only varieties invariant under the given group action.
(iii) The varieties $\overline{V}_k$ are irreducible.

From the correspondence between varieties and ideals we deduce that the ideals defining them are the only invariant ideals equal to their radical and they are all prime ideals.

We shall deduce from this the second fundamental theorem of invariant theory in §6.

More interesting is the fact that there are also finitely many orbits under the action of a Borel subgroup. We will not compute all the orbits, but we will restrict ourselves to analyzing the invariant hypersurfaces. We discuss the 3 cases.

1. To distinguish between $GL(n, \mathbb{C})$ and $GL(m, \mathbb{C})$ we let $T(n)$, $T(m)$, $U^+(n)$, $U^+(m), \ldots$, etc., denote the torus, unipotent, etc. of the two groups.

We take as a Borel subgroup of $GL(n, \mathbb{C}) \times GL(m, \mathbb{C})$ the subgroup $B(n)^- \times B(m)^+$ of pairs (A, B) where A is a lower and B is an upper triangular matrix. We may assume $n \leq m$ (or we transpose).

If $(A, B) \in B(n)^- \times B(m)^+$, $X \in M_{n,m}(\mathbb{C})$, the matrix AXB^{-1} is obtained from X by elementary row and column operations of the following types:

(a) multiply a row or a column by a nonzero scalar,
(b) add to the i^{th} row the j^{th} row, with $j < i$, multiplied by some number,
(c) add to the i^{th} column the j^{th} column, with $j < i$, multiplied by some number.

This is the usual Gaussian elimination on rows and columns of X without performing any exchanges.

The usual remark about these operations is that, for every $k \leq n$, they do not change the rank of the $k \times k$ minor X_k of X extracted from the first k rows and the first k columns. Moreover, if we start from a matrix X with the property that for every $k \leq n$ we have $\det(X_k) \neq 0$, then the standard algorithm of Gaussian elimination proves that, under the action of $B(n)^- \times B(m)^+$, this matrix is equivalent to the matrix I with entries 1 on the diagonal and 0 elsewhere. We deduce

Theorem 1. *The open set of matrices $X \in M_{n,m}(\mathbb{C})$ with $\det(X_k) \neq 0$, $k = 1, \ldots, n$, is a unique orbit under the group $B(n)^- \times B(m)^+$.*

The only $B(n)^- \times B(m)^+$ stable hypersurfaces of $M_{n,m}(\mathbb{C})$ are the ones defined by the equations $\det(X_k) = 0$, which are irreducible.

Proof. We have already remarked that the first part of the theorem is a consequence of Gaussian elimination. As for the second, since the complement of this open orbit is the union of the hypersurfaces of equations $\det(X_k) = 0$, it is clearly enough to prove that these equations are irreducible. The main property of the functions $\det(X_k)$ is the fact that they are highest weight vectors for $B(n)^- \times B(m)^+$.

Thus we compute the weight of $\det(X_k)$ directly. Given a pair $D_1, D_2 \in T(n) \times T(m)$ of diagonal matrices with entries x_i, y_j respectively, we have

$$(4.4.2) \qquad (D_1, D_2)d_k(X) := d_k(D_1^{-1}XD_2) = \prod_{i=1}^{k} x_i^{-1}y_i d_k(X).$$

Hence the weight of d_k is

$$\prod_{i=1}^{k} x_i^{-1} \prod_{i=1}^{k} y_i$$

which is the highest weight of $(\bigwedge^k \mathbb{C}^n)^* \otimes \bigwedge^k \mathbb{C}^m$, a fundamental weight.

We can now prove that the functions $\det(X_k)$ are irreducible. If $\det(X_k)$ is not irreducible, it is a product of highest weight vectors g_i and its weight is the sum of the weights of the g_i which are dominant weights. We have seen that $d_k := \det(X_k)$ is a fundamental weight, hence we are done once we remark that the only polynomials which belong to the 0 weight are constant. □

2. and 3. are treated as follows. One thinks of a symmetric or skew-symmetric matrix as the matrix of a form.

Again we choose as Borel subgroup $B(n)^-$ and Gaussian elimination is the algorithm of putting the matrix of the form in normal form by a triangular change of basis.

The generic orbit is obtained when the given form has maximal rank on all the subspaces spanned by the first k basis vectors $k \leq n$, which, in the symmetric case, means that the form is nondegenerate on these subspaces, while in the skew case it means that the form is nondegenerate on the even-dimensional subspaces.

2. On symmetric matrices this is essentially the Gram–Schmidt algorithm. We get that

Theorem 2. *The open set of symmetric matrices $X \in M_n^+(\mathbb{C})$ with $\det(X_k) \neq 0,\ \forall k$ is a unique orbit under the group $B(n)^-$.*

The only $B(n)^-$ stable hypersurfaces of $M_n^+(\mathbb{C})$ are the ones defined by the equations $s_k := \det(X_k) = 0$, which are irreducible.

Proof. Here we can proceed as in the linear case except at one point, when we arrive at the computation of the character of s_k, we discover that it is $\prod_{i=1}^k x_i^{-2}$, which is twice a fundamental weight.

Hence a priori we could have $s_k = ab$ with a, b with weight $\prod_{i=1}^k x_i^{-1}$. To see that this is not possible, set the variables $x_{ij} = 0,\ i \neq j$ getting $s_k = \prod_{i=1}^k x_{ii}$; hence a, b should specialize to two factors of $\prod_{i=1}^k x_{ii}$, but clearly these factors never have as weight $\prod_{i=1}^k x_i^{-1}$. Hence s_k is irreducible. □

3. Choose as a Borel subgroup $B(n)^-$ and perform Gaussian elimination on skew-symmetric matrices. For every k with $2k \leq n$, consider the minor X_{2k}. The condition that this skew-symmetric matrix be nonsingular is that the Pfaffian $p_k := Pf(X_{2k})$ is nonzero.

Theorem 3. *The open set of skew-symmetric matrices $X \in M_n^-(\mathbb{C})$ with $Pf(X_{2k}) \neq 0$, $\forall k$ is a unique orbit under the group $B(n)^-$.*

The only $B(n)^-$-stable hypersurfaces of $M_n^-(\mathbb{C})$ are the ones defined by the equations $p_k := Pf(X_{2k}) = 0$, which are irreducible.

Proof. If $Pf(X_{2k}) \neq 0$ for all k, we easily see that we can construct in a triangular form a symplectic basis for the matrix X, hence the first part. For the rest, again it suffices to prove that the polynomials p_k are irreducible. In fact we can compute their weight which, by the formula 3.6.2 in Chapter 5, is $\prod_{i=1}^{2k} x_i^{-1}$, a fundamental weight. □

From Lemma 4.2 we can describe, in the three previous examples, the highest weight vectors as monomials in the irreducible ones. This gives an implicit description of the polynomial ring as representation. We shall discuss this in more detail in the next sections.

4.5 Cauchy Formulas

We can deduce now the Cauchy formulas that are behind this theory. We do it in the symmetric and skew-symmetric cases, which are the new formulas, using the notations of the previous section.

In the symmetric case we have that the full list of highest weight vectors is the set of monomials $\prod_{k=1}^n s_k^{m_k}$ with weight $\prod_{k=1}^n \prod_{i=1}^k x_i^{-2m_k}$.

If we denote by V the fundamental representation of $GL(n,\mathbb{C})$, we have that

$$\prod_{k=1}^{n}\prod_{i=1}^{k} x_i^{-2m_k} = \prod_{i=1}^{n} x_i^{-\sum_{i\le k} 2m_k}$$

is the highest weight of $S_\lambda(V)^*$, where λ is the partition $\lambda_k = 2\sum_{i\le k} m_k$.

In the skew case we obtain the monomials $\prod_{2k\le n} p_k^{m_k}$ with weight $\prod_{i=1}^n x_i^{-\sum_{i\le 2k} m_k}$.

We deduce the *special plethystic formulas* (Chapter 9, §7.3).[108]

Theorem. *As a representation of $GL(V)$ the ring $S(S^2(V))$ decomposes as*

$$S(S^2(V)) := \bigoplus\nolimits_\lambda S_{2\lambda}(V). \tag{4.5.1}$$

As a representation of $GL(V)$ the ring $S(\bigwedge^2(V))$ decomposes as

$$S\Big(\bigwedge{}^2(V)\Big) := \bigoplus\nolimits_\lambda S_{\widetilde{2\lambda}}(V). \tag{4.5.2}$$

Proof. We have that the space of symmetric forms on V is, as a representation, $S^2(V)^*$ and the action is, in matrix notation, given by $(A, X) \to (A^{-1})^t X A^{-1}$, so $S(S^2(V))$ is the ring of polynomials on this space. We can apply the previous theorem and the analysis following it. The considerations for the skew case are similar. □

[108] Recall these are formulas which describe the composition $S_\lambda(S_\mu(V))$ of Schur functors.

In more pictorial language, thinking of λ as the shape of a diagram we can say that the diagrams appearing in $S(S^2(V))$ have even rows, while the ones appearing in $S(\bigwedge^2(V))$ have even columns.[109] It is convenient to have a short notation for these concepts and write

$$\lambda \vdash^{er} n, \ \lambda \vdash^{ec} n$$

to express the fact that the diagram λ has even rows, resp., even columns.

We should remark that the previous theorem corresponds to identities of characters.

According to Molien's formula (Chapter 9, §4.3.3), given a linear operator A on a vector space U, its action on the symmetric algebra has as graded character $\frac{1}{\det(1-tA)}$.

If $e_1, \dots, e_n$ is a basis of V and X is the matrix $Xe_i = x_i e_i$, we have that $e_i e_j, \ i \leq j$, is a basis of $S^2(V)$ and $e_i \wedge e_j, \ i < j$ a basis of $\bigwedge^2(V)$. Thus from Molien's formula and the previous theorem we therefore deduce

$$\frac{1}{\prod_{i\leq j}(1 - x_i x_j)} = \sum_m \sum_{\lambda \vdash^{er} m} S_\lambda(x_1, \dots, x_n) \tag{4.5.3}$$

$$\frac{1}{\prod_{i<j}(1 - x_i x_j)} = \sum_m \sum_{\lambda \vdash^{ec} m} S_\lambda(x_1, \dots, x_n). \tag{4.5.4}$$

These are the formulas $(C2)$ and $(C3)$ stated in Chapter 2, §4.1.

4.6 Bruhat Cells

We point out also the following fact, which generalizes the discussion for the linear group and whose details we leave to the reader. If G is a semisimple simply connected algebraic group, we know that its coordinate ring $\mathbb{C}[G] = \bigoplus_{\lambda \in \Lambda^+} V_\lambda \otimes V_\lambda^*$. Given a Borel subgroup B of G, the highest weight vectors (for the $B \times B$) action are the elements $f_\lambda := v_\lambda \otimes \phi_\lambda$ where v_λ resp. ϕ_λ are the highest weight vectors of V_λ, resp. V_λ^*. If $\lambda = \sum_i m_i \omega_i$ with the ω_i fundamental, we have $f_\lambda = \prod_i f_{\omega_i}^{m_i}$. $f_{\omega_i} = 0$ is a hypersurface of G which is stable under $B \times B$. It is not difficult to see that it is the closure of the double coset $Bs_i w_0 B$. Here s_i denotes the simple reflection by the simple root of the same index i as the fundamental weight.

5 The Second Fundamental Theorem (SFT)

5.1 Determinantal Ideals

We need to discuss now the relations among invariants. We shall take a geometric approach reserving the combinatorial approach for the section on tableaux. The

[109] Unfortunately the row or column encoding is not really canonical. In Chapter 13 we will be obliged to switch the convention.

study of the relations among invariants proceeds as follows. We express the first fundamental theorem in matrix form and deduce a description of the invariant prime ideals, which is equivalent to the SFT, from the highest weight theory.

For the general linear group we have described invariant theory through the multiplication map $f : M_{p,m} \times M_{m,q} \to M_{p,q}$, $f(X, Y) := XY$ (Chapter 9, §1.4.1).

The ring of polynomial functions on $M_{p,m} \times M_{m,q}$ which are $Gl(m, \mathbb{C})$-invariant is given by the polynomial functions on $M_{p,q}$ composed with the map f. We have remarked that by elementary linear algebra, the multiplication map f has as its image the subvariety of $p \times q$ matrices of rank $\leq m$. This is the whole space if $m \geq \min(p, q)$. Otherwise, it is a proper subvariety defined, at least set theoretically, by the vanishing of the determinants of the $(m + 1) \times (m + 1)$ minors of the matrix of coordinate functions x_{ij} on $M_{p,q}$.

For the group $O(n, \mathbb{C})$ we have considered the space of $n \times m$ matrices Y with the action of $O(n, \mathbb{C})$ given by multiplication XY. Then the mapping $Y \to Y^tY$ from the space of $n \times m$ matrices to the symmetric $m \times m$ matrices of rank $\leq n$ is a quotient under the orthogonal group. Again, the determinants of the $(m+1) \times (m+1)$ minors of these matrices define this subvariety set-theoretically.

Similarly, for the symplectic group we have considered the space of $2n \times m$ matrices Y with the action of $Sp(2n, \mathbb{C})$ given by multiplication XY. Then the mapping $Y \to Y^t J_n Y$ (with J_n the standard $2n \times 2n$ skew-symmetric matrix), from the space of $2n \times m$ matrices to the antisymmetric $m \times m$ matrices of rank $\leq 2n$ is a quotient under the symplectic group. In this case the correct relations are not the determinants of the minors but rather the Pfaffians of the principal minors of order $2(n + 1)$.

We have thus identified three types of *determinantal varieties* for each of which we want to determine the ideal of relations. We will make use of the plethystic formulas developed in the previous section.

According to §4.3 we know that the determinantal varieties are the only varieties that are invariant under the appropriate group action. According to the matrix formulation of the first fundamental theorem they are also the varieties which have rings of invariants as coordinate rings. We want to describe the ideals of definition and their coordinate rings as representations.

In Chapter 9, §7.1 we have seen that, given two vector spaces V and W, we have the decomposition

$$\mathcal{P}[\hom(V, W)] = \bigoplus_\lambda S_\lambda(W^*) \otimes S_\lambda(V) = \bigoplus_\lambda \hom(S_\lambda(V), S_\lambda(W))^*.$$

Moreover if we think of $\mathcal{P}[\hom(V, W)]$ as the polynomial ring $\mathbb{C}[x_{ij}]$, its subspace D_k, spanned by the determinants of the minors of order k of the matrix $X := (x_{ij})$, is identified with the subrepresentation $D_k = \bigwedge^k W^* \otimes \bigwedge^k V$. We define

$$I_k := \mathcal{P}[\hom(V, W)]D_k \tag{5.1.1}$$

to be the determinantal ideal generated by the determinants of all the $k \times k$ minors of X.

Consider now $S(S^2(V)) = \mathcal{P}[S^2(V)^*]$ as the polynomial ring $\mathbb{C}[x_{ij}]$, $x_{ij} = x_{ji}$. Let $X := (x_{ij})$ be a symmetric matrix of variables.

We want to see how to identify, along the same lines as Chapter 9, §7.1, the subspace D_k^s of $\mathcal{P}[S^2(V)^*]$ spanned by the determinants of the minors of order k of the matrix X.

Let $\lambda \vdash m$ be a partition. Given a symmetric form A on V, it induces a symmetric form $A^{\otimes m}$ on $V^{\otimes m}$ by

$$A^{\otimes m}(u_1 \otimes \cdots \otimes u_m, v_1 \otimes \cdots \otimes v_m) := \prod_{i=1}^{m} A(u_i, v_i).$$

Thus by restriction it induces a symmetric form, which we will denote by $S_\lambda(A)$, on $S_\lambda(V)$, or equivalently, an element of $S^2(S_\lambda(V)^*)$ which we identify with $S^2(S_\lambda(V))^*$.

In other words we can identify $S_\lambda(A)$ with a linear form on $S^2(S_\lambda(V))$.

According to the Corollary in Chapter 10, §5.2, $S_{2\lambda}(V)$ appears with multiplicity 1 in $S^2(S_\lambda(V))$. So we deduce that $S_\lambda(A)$ induces a linear function $\langle S_\lambda(A) \mid v\rangle$ on $S_{2\lambda}(V)$.

As a function of A, v this function $\langle S_\lambda(A) \mid v\rangle$ is $GL(V)$-invariant, linear in v, and a homogeneous polynomial of degree $2m$ in A. Thus we have a dual map

$$S_{2\lambda}(V) \to \mathcal{P}[S^2(V)^*].$$

We claim that this map is nonzero. Since $S_{2\lambda}(V)$ is irreducible, it then identifies $S_{2\lambda}(V)$ with its corresponding factor in the decomposition 4.5.1.

To see this, we compute the linear form $S_\lambda(A)$ on the highest weight vector $U \otimes U$ of $S_{2\lambda}(V)$, $U = (e_1 \wedge e_2 \wedge \ldots \wedge e_{n_1}) \otimes (e_1 \wedge e_2 \wedge \ldots \wedge e_{n_2}) \otimes \cdots \otimes (e_1 \wedge e_2 \wedge \ldots \wedge e_{n_t})$. By definition we get a product of determinants of suitable minors of A, so in general a nonzero element.

In particular, we apply this to $\lambda = 1^k$, $S_\lambda(V) = \bigwedge^k(V)$, $S_\lambda(A) = \wedge^k A$. We see that if $A = (a_{ij})$ is the matrix of the given form in a basis $e_1, \ldots, e_n$, the matrix of $\wedge^k A$ in the basis $e_{i_1} \wedge e_{i_2} \ldots \wedge e_{i_k}$ is given by the formula

(5.1.2)

$$\wedge^k A(e_{i_1} \wedge e_{i_2} \wedge \ldots \wedge e_{i_k}, e_{j_1} \wedge e_{j_2} \wedge \ldots \wedge e_{j_k}) = \det(a_{i_r, j_s}), \ r, s = 1, \ldots, k.$$

The determinant 5.1.2 is the determinant of the minor extracted from the rows $i_1, \ldots, i_k$ and the columns $j_1, \ldots, j_k$ of A.

We have thus naturally defined a map $j_k : S^2(\bigwedge^k V) \to \mathcal{P}[S^2(V)^*]$ with the image the space D_k^s (spanned by the determinants of the minors of order k).

We need to prove that j_k, restricted to $S_{2^k}(V)$, gives an isomorphism to D_k^s, which is therefore irreducible with highest weight vector $2\omega_k$.

To see this we analyze the decomposition of $S^2(\bigwedge^k V)$ into irreducible representations. We have $S^2(\bigwedge^k V) \subset \bigwedge^k V \otimes \bigwedge^k V$; the decomposition of $\bigwedge^k V \otimes \bigwedge^k V$ is given by Pieri's formula. Since $\bigwedge^k V$ consists of just one column of length k, $\bigwedge^k V \otimes \bigwedge^k V = \bigoplus_{i=0}^k S_{1^{2i}2^{k-i}}(V)$, the sum of the $S_\mu(V)$ where μ has at most two rows, of lengths $2k - i, i$ respectively ($i \leq k$). Of these partitions, the only one with even rows is $S_{2^k}(V)$. Hence all the other irreducible representation appearing in

$\bigwedge^k V \otimes \bigwedge^k V$ do not appear in $\mathcal{P}[S^2(V)^*]$ and so the map j_k on them must be 0, proving the claim.

Consider now $\mathcal{P}[\bigwedge^2(V)^*]$ as the polynomial ring $\mathbb{C}[x_{ij}]$, $x_{ij} = -x_{ji}$. Let $X := (x_{ij})$ be a skew-symmetric matrix of variables. We carry out a similar analysis for the subspace P_k of $\mathcal{P}[\bigwedge^2(V)^*]$ spanned by the Pfaffians of the principal minors of order $2k$ of X.

In this case the analysis is simpler. The exterior power map $\bigwedge^2 V^* \to \bigwedge^{2k} V^*$, $A \to A^k := \wedge^k A$ gives for $A = \sum_{i<j} a_{ij} e_i \wedge e_j$ that

$$A^k = k! \sum_{i_1 < i_2 < \ldots i_{2k}} [i_1, i_2, \ldots, i_{2k}] e_{i_1} \wedge e_{i_2} \wedge \ldots \wedge e_{i_{2k}}, \tag{5.1.3}$$

where $[i_1, i_2, \ldots, i_{2k}]$ denotes the Pfaffian of the principal minor of A extracted from the row and column indices $i_1 < i_2 < \ldots i_{2k}$. One has immediately by duality the required map with image the space P_k:

$$\bigwedge^{2k} V \to S\Big(\bigwedge^2 (V)\Big).$$

Of course $\bigwedge^{2k} V$ corresponds to a single even column of length $2k$.

Using the results of the previous section and the previous discussion we have the

Second Fundamental Theorem.

(1) The only $GL(V) \times GL(W)$-invariant prime ideals in $\mathcal{P}[\hom(V, W)]$ are the ideals I_k. As representations we have that

$$I_k = \bigoplus\nolimits_{\lambda, ht(\lambda) \geq k} S_\lambda(W^*) \otimes S_\lambda(V),$$

$$\mathcal{P}[\hom(V, W)]/I_k = \bigoplus\nolimits_{\lambda, ht(\lambda) < k} S_\lambda(W^*) \otimes S_\lambda(V). \tag{5.1.4}$$

(2) The only $GL(V)$-invariant prime ideals in $\mathcal{P}[S^2(V)^*]$ are the determinantal ideals

$$I_k^+ := \mathcal{P}[S^2(V)^*] J_k$$

generated by the determinants of the $k \times k$ minors of the symmetric matrix X. As representations we have that

$$I_k^+ = \bigoplus\nolimits_{\lambda \vdash^{er}, ht(\lambda) \geq k} S_\lambda(V), \quad \mathcal{P}[S^2(V)^*]/I_k^+ = \bigoplus\nolimits_{\lambda \vdash^{er}, ht(\lambda) < k} S_\lambda(V). \tag{5.1.5}$$

(3) The only $GL(V)$-invariant prime ideals in $\mathcal{P}[\bigwedge^2(V)^*]$ are the Pfaffian ideals

$$I_k^- := \mathcal{P}\Big[\bigwedge^2 (V)^*\Big] P_k$$

generated by the Pfaffians of the principal $2k \times 2k$ minors. As representations we have that

(5.1.6) $I_k^- = \bigoplus_{\lambda\vdash^{ec},ht(\lambda)\geq 2k} S_\lambda(V),\ \mathcal{P}\left[\bigwedge^2(V)^*\right]/I_k^- = \bigoplus_{\lambda\vdash^{ec},ht(\lambda)<2k} S_\lambda(V).$

Proof. In each of the 3 cases we know that the set of highest weights is the set of monomials $\prod_{i=1}^n d_i^{h_i}$ in a certain number of elements d_i (determinants or Pfaffians) highest weights of certain representations N_i (the determinantal or Pfaffian spaces described before). Specifically, in case 1 we have $n = \min(\dim V, \dim W)$ and d_i is the determinant of the principal $i \times i$ minor of the matrix $x_{i,j}$. In case 2, again $n = \dim V$ with d_i the same type of determinants, while in type 3 we have $\dim V = 2n$ and the d_i is the Pfaffian of the principal $2i \times 2i$ submatrix.

Every invariant subspace M is identified by the set I of highest weight vectors that it contains.

If M is an ideal, and $\prod_{i=1}^n d_i^{h_i} \in I$, then certainly $\prod_{i=1}^n d_i^{k_i} \in I$ if $k_i \geq h_i$ for each i.

If M is a prime ideal it follows that, if a monomial $\prod d_i^{m_i} \in I$, then at least one d_i appearing with nonzero exponent must be in I.

It follows, for a prime ideal M, that M is necessarily generated by the subspaces N_i contained in it. To conclude, it is enough to remark, in the case of determinantal ideals, that d_i is in the ideal generated by d_k as soon as $i \geq k$. It clearly suffices to show this for d_{k+1}, i.e., to show that

(5.1.7) $$D_{k+1} \subset I_k,\ D_{k+1}^s \subset I_k^+,\ P_{k+1} \subset I_k^-.$$

The first two statements follow immediately by a row or column expansion of a determinant of a $(k+1) \times (k+1)$ minor in terms of determinants of $k \times k$ minors. As for the Pfaffians, recall the defining formula 3.6.1 of Chapter 5. $A = \sum_{i<j} x_{ij} e_i \wedge e_j$:

$$A^{k+1} = (k+1)! \sum_{i_1<i_2<\dots i_{2(k+1)}} [i_1, i_2, \dots, i_{2(k+1)}] e_{i_1} \wedge e_{i_2} \wedge \dots \wedge e_{i_{2(k+1)}}$$

$$= A \wedge k! \sum_{i_1<i_2<\dots<i_{2k}} [i_1, i_2, \dots, i_{2k}] e_{i_1} \wedge e_{i_2} \wedge \dots \wedge e_{i_{2k}}.$$

From the above we deduce the typical Laplace expansion for a Pfaffian:

$$(k+1)[1, 2, \dots, 2(k+1)]$$
$$= \sum_{i<j} (-1)^{i+j-1} x_{ij} [1, 2, \dots, i-1, i+1, \dots, j-1, j+1, \dots, 2(k+1)].$$

Therefore in each case we have that for a given prime invariant ideal M there is an integer k such that $\prod_{i=1}^n d_i^{m_i} \in M$ if and only if $m_i > 0$ for at least one $i \geq k$. Thus the ideal is generated by D_k, J_k^+, P_k^- for the minimal index k for which this space is contained in the ideal. The remaining statements are just a reinterpretation of what we said. □

We called the previous theorem the Second Fundamental Theorem since it describes the ideals of matrices (in the three types) of rank $< k$ (or $< 2k$ in the skew-symmetric case). We have already seen that these varieties are the varieties corresponding to the three algebras of invariants we considered, so this theorem is a description of the ideal of relations among invariants.

5.2 Spherical Weights

There is a rather general theory of spherical subgroups H of a given group G and the orthogonal and symplectic group are spherical in the corresponding group $GL(V)$.

The technical definition of a spherical subgroup H of a reductive group G is the following: We assume that a Borel subgroup B has an open orbit in G/H, or equivalently, that there is a Borel subgroup B with BH dense in G. There is a deep theory of these pairs which we will not treat.

One of the first properties of spherical subgroups is that given an irreducible representation of G, the space of invariants under H is at most 1-dimensional. When it is exactly 1, the corresponding dominant weight is then called a *spherical weight* relative to H. We want to illustrate this phenomenon for the orthogonal and symplectic groups.

1. Orthogonal group. Take an n-dimensional orthogonal space V and consider $n-1$ copies of V which we display as $V \otimes W$, $\dim W = n-1$.

We know that the orthogonal or special orthogonal invariants of $n-1$ copies of V are the scalar products (u_i, u_j) and that they are algebraically independent. As a representation of $GL(n-1, \mathbb{C}) = GL(W)$, these basic invariants transform as $S^2(W)$ and so the ring they generate is $S(S^2(W)) = \bigoplus_\mu S_{2\mu}(W)$. On the other hand, by Cauchy's formula,

$$S(V \otimes W)^{O(V)} = \bigoplus_\lambda S_\lambda(V)^{O(V)} \otimes S_\lambda(W).$$

We deduce an isomorphism as $GL(W)$-representations:

$$\bigoplus_\lambda S_\lambda(V)^{O(V)} \otimes S_\lambda(W) = \bigoplus_\mu S_{2\mu}(W). \tag{5.2.1}$$

Corollary of 5.2.1. $\dim S_\lambda(V)^{O(V)} = 1$ *if* $\lambda = 2\mu$ *and* 0 *otherwise.*

There is a simple explanation for this invariant which we leave to the reader.

The scalar product induces a nondegenerate scalar product on $V^{\otimes p}$ for all p and thus a nondegenerate scalar product on $S_\lambda(V)$ for each V. This induces a canonical invariant element in $S^2(S_\lambda(V))$, and its projection to $S_{2\lambda}(V)$ is the required invariant (cf. §5.1).

Since $SO(V) \subset SL(V)$ the restriction to dimension $n-1$ is not harmful since all representations of $SL(V)$ appear. For $O(V)$ we can take n-copies and leave the details to the reader.

2. Symplectic group. Take a $2n$-dimensional symplectic space V and consider $2n$ copies of V which we display as $V \otimes W$, $\dim W = 2n$.

We know that the symplectic invariants of $2n$ copies of V are the scalar products $[u_i, u_j]$ and that they are algebraically independent. As a representation of $GL(2n, \mathbb{C}) = GL(W)$ these basic invariants transform as $\bigwedge^2(W)$ and so the ring they generate is $S(\bigwedge^2(W)) = \bigoplus_\mu S_{\widetilde{2\mu}}(W)$. On the other hand by Cauchy's formula,

$$S(V \otimes W)^{Sp(V)} = \bigoplus_\lambda S_\lambda(V)^{Sp(V)} \otimes S_\lambda(W),$$

we deduce an isomorphism as $GL(W)$ representations:

(5.2.2) $$\bigoplus_\lambda S_\lambda(V)^{Sp(V)} \otimes S_\lambda(W) = \bigoplus_\mu S_{\widetilde{2\mu}}(W).$$

Corollary of 5.2.2. $\dim S_\lambda(V)^{Sp(V)} = 1$ *if* $\lambda = \widetilde{2\mu}$ *and* 0 *otherwise.*

There is a similar explanation for this invariant. Start from the basic invariant $J \in \bigwedge^2 V$. We get by exterior multiplication $J^k \in \bigwedge^{2k} V$, and then by tensor product, the invariant $J^{k_1} \otimes J^{k_2} \otimes \cdots \otimes J^{k_m} \in \bigwedge^{2k_1} V \otimes \cdots \otimes \bigwedge^{2k_r} V$, which projects to the invariant in $S_\lambda(V)$, where $\tilde{\lambda} = 2k_1, 2k_2, \ldots, 2k_r$.

6 The Second Fundamental Theorem for Intertwiners

6.1 Symmetric Group

We want to apply the results of the previous section to intertwiners. We need to recall the discussion of multilinear characters in §6.4 of Chapter 9. Start from $\mathcal{P}[\hom(V, W)] = S(W^* \otimes V) = \bigoplus_\lambda S_\lambda(W^*) \otimes S_\lambda(V)$, with $\dim V = n,\ W = \mathbb{C}^n$.

Using the standard basis e_i of $\mathbb{C}^n$, we have identified $V^{\otimes n}$ with a subspace of $S(\mathbb{C}^n \otimes V)$ by encoding the element $\prod_{i=1}^n e_i \otimes v_i$ as $v_1 \otimes v_2 \otimes \cdots \otimes v_n$.

The previous mapping identifies $V^{\otimes n} \subset S(\mathbb{C}^n \otimes V)$ with a weight space under the torus of diagonal matrices X with entries x_i. We have $Xe_i = x_i e_i$ and

$$X \prod_{i=1}^n e_i \otimes v_i = \prod_i x_i \prod_{i=1}^n e_i \otimes v_i.$$

Now we extend this idea to various examples. First, we identify the group algebra $\mathbb{C}[S_n]$ with the space P^n of polynomials in the variables x_{ij} which are multilinear in right and left indices, that is, we consider the span P^n of those monomials $x_{i_1,j_1} x_{i_2,j_2} \ldots x_{i_n,j_n}$ such that both $i_1, i_2, \ldots, i_n$ and $j_1, j_2, \ldots, j_n$ are permutations of $1, 2, \ldots, n$.

Of course a monomial of this type can be uniquely displayed as

$$x_{1,\sigma^{-1}(1)} x_{2,\sigma^{-1}(2)} \ldots x_{n,\sigma^{-1}(n)} = x_{\sigma(1),1} x_{\sigma(2),2} \ldots x_{\sigma(n),n}$$

which defines the required map

(6.1.1) $$\Phi : \sigma \mapsto x_{\sigma(1),1} x_{\sigma(2),2} \ldots x_{\sigma(n),n}.$$

Now remark that this space of polynomials is a weight space with respect to the product $T \times T$ of two maximal tori of diagonal matrices under the induced $GL(n, \mathbb{C}) \times GL(n, \mathbb{C})$ action on $\mathbb{C}[x_{ij}]$. Let us denote by χ_1, χ_2 the two weights.

We remark also that this mapping is equivariant under the left and right action of S_n on $\mathbb{C}[S_n]$ which correspond, respectively, to

(6.1.2) $$x_{i,j} \to x_{\sigma(i),j},\ x_{i,j} \to x_{i,\sigma(j)}.$$

Fix an m-dimensional vector space V and recall the basic symmetry homomorphism

$$(6.1.3) \qquad i_n : \mathbb{C}[S_n] \to \mathrm{End}(V^{\otimes n}).$$

We have a homomorphism given by the FFT. With the notations of Chapter 9, §1.4 it is $f : \mathbb{C}[x_{ij}] \to \mathbb{C}[\langle\alpha_i|v_j\rangle]$. It maps P^n onto the space of multilinear invariants of n covector and n vector variables.

Let us denote by $\mathcal{I}_n$ this space, which is spanned by the elements $\prod_{i=1}^n \langle\alpha_{\sigma(i)}|v_i\rangle$.

Finally, we have the canonical isomorphism $j : \mathrm{End}_{GL(V)} V^{\otimes n} \xrightarrow{j} \mathcal{I}_n$. It maps the endomorphism induced by a permutation σ to $\prod_{i=1}^n \langle\alpha_{\sigma(i)}|v_i\rangle$ (cf. Chapter 9, §1.3.2).

We have a commutative diagram

$$(6.1.4) \qquad \begin{array}{ccc} \mathbb{C}[S_n] & \xrightarrow{\Phi} & P_n \\ {\scriptstyle i_n}\downarrow & & {\scriptstyle f}\downarrow \\ \mathrm{End}_{GL(V)} V^{\otimes n} & \xrightarrow{j} & \mathcal{I}_n \end{array}$$

from which we deduce that the kernel of i_n can be identified, via the linear isomorphism Φ, with the intersection of P_n with the determinantal ideal I_{m+1} in $\mathbb{C}[x_{ij}]$.

Given a matrix X we denote by $(i_1, i_2, \dots, i_{m+1}|j_1, j_2, \dots, j_{m+1})(X)$ or, if no confusion arises, by $(i_1, i_2, \dots, i_{m+1}|j_1, j_2, \dots, j_{m+1})$ the determinant of the minor of X extracted from the rows of indices i and the columns of indices j.

Clearly $(i_1, i_2, \dots, i_{m+1}|j_1, j_2, \dots, j_{m+1})$, multiplied by any monomial in the x_{ij}, is a weight vector for $T \times T$.

In order to get a weight vector in P_n we must consider the products

$$(6.1.5) \qquad (i_1, i_2, \dots, i_{m+1}|j_1, j_2, \dots, j_{m+1})x_{i_{m+2},j_{m+2}} \cdots x_{i_n,j_n}$$

with $i_1, i_2, \dots, i_n$ and $j_1, j_2, \dots, j_n$, both permutations of $1, 2, \dots, n$.

Theorem. *Under the isomorphism Φ the space $P_n \cap I_{m+1}$ is 0 if $m \geq n$. If $m < n$ it corresponds to the two-sided ideal of $\mathbb{C}[S_n]$ generated by the element*

$$(6.1.6) \qquad A_{m+1} := \sum_{\sigma \in S_{m+1}} \epsilon_\sigma \sigma,$$

the antisymmetrizer on $m+1$ elements (chosen arbitrarily from the given n elements).

Proof. From 6.1.2 it follows that the element corresponding to a typical element of type 6.1.5 is of the form $\sigma A \tau^{-1}$ where

$$\begin{aligned} A &= (1, 2, \dots, m+1|1, 2, \dots, m+1)x_{m+2,m+2} \cdots x_{n,n} \\ &= \Big(\sum_{\sigma \in S_{m+1}} \epsilon_\sigma x_{\sigma(1),1}x_{\sigma(2),2} \cdots x_{\sigma(m+1),m+1}\Big)x_{m+2,m+2} \cdots x_{n,n}. \end{aligned}$$

This element clearly corresponds to A_{m+1}. $\square$

Now we want to recover in a new form the result of Chapter 9, §3.1.3. First, recall that, as any group algebra, $\mathbb{C}[S_n]$ decomposes into the sum of its minimal ideals corresponding to irreducible representations

$$\mathbb{C}[S_n] = \bigoplus_{\lambda \vdash n} M_\lambda^* \otimes M_\lambda.$$

We first decompose

$$\mathbb{C}[x_{ij}] = \bigoplus_\lambda S_\lambda(\mathbb{C}^n)^* \otimes S_\lambda(\mathbb{C}^n), \quad I_{m+1} = \bigoplus_{ht\lambda \geq m+1} S_\lambda(\mathbb{C}^n)^* \otimes S_\lambda(\mathbb{C}^n).$$

Then we pass to the weight space

$$P^n = \mathbb{C}[x_{ij}]^{\chi_1, \chi_2} = \bigoplus_\lambda S_\lambda(\mathbb{C}^n)^{*\chi_1} \otimes S_\lambda(\mathbb{C}^n)^{\chi_2},$$

$$P^n \cap I_{m+1} = \bigoplus_{ht\lambda \geq m+1} S_\lambda(\mathbb{C}^n)^{*\chi_1} \otimes S_\lambda(\mathbb{C}^n)^{\chi_2}$$

$$(6.1.8) \qquad = \bigoplus_{\lambda \vdash n,\, ht\lambda \geq m+1} M_\lambda^* \otimes M_\lambda.$$

As a consequence the image of $\mathbb{C}[S_n]$ in $\mathrm{End}(V^{\otimes n})$ is $\bigoplus_{\lambda \vdash n,\, ht\lambda \leq m} M_\lambda^* \otimes M_\lambda$.

6.2 Multilinear Spaces

Now we go on to the orthogonal and symplectic group. First, we formulate some analogue of P^n. In the symmetric or skew-symmetric cases, we take a (symmetric or skew-symmetric) matrix $Y = (y_{ij})$ of even size $2n$, and consider the span of the *multilinear monomials*

$$y_{i_1,j_1} y_{i_2,j_2} \cdots y_{i_n,j_n}$$

such that $i_1, i_2, \dots, i_n, j_1, j_2, \dots, j_n$, is a permutation of $1, 2, \dots, 2n$. This is again the weight space of the weight χ of the diagonal group T of $GL(2n, \mathbb{C})$.

Although at first sight the space looks the same in the two cases, it is not so, and we will call these two cases P_+^{2n} and P_-^{2n}. In fact, the symmetric group S_{2n} acts on the polynomial ring $\mathbb{C}[y_{ij}]$ (as a subgroup of $GL(2n, \mathbb{C})$) by $\sigma(y_{ij}) = y_{\sigma(i)\sigma(j)}$.

It is clear that S_{2n} acts transitively on the given set of multilinear monomials.

First, we want to understand it as a representation on $P_\pm^n$ and for this we consider the special monomial

$$M_0 := y_{1,2} y_{3,4} \cdots y_{2n-1,2n}.$$

Let H_n be the subgroup of S_{2n} which fixes the partition of $\{1, 2, \dots, 2n\}$ formed by the n sets with 2 elements $\{2i-1, 2i\}$, $i = 1, \dots, n$. Clearly, $H_n = S_n \ltimes \mathbb{Z}/(2)^n$ is the obvious semidirect product, where S_n acts in the diagonal way on odd and even numbers $\sigma(2i-1) = 2\sigma(i) - 1$, $\sigma(2i) = 2\sigma(i)$ and $\mathbb{Z}/(2)^n$ is generated by the transpositions $(2i-1, 2i)$.

In either case H_n is the stabilizer of the line through M_0. In the symmetric case H_n fixes M_0, while in the skew-symmetric case it induces on this line the sign of the permutation (remark that $S_n \subset S_{2n}$ is made of even permutations). We deduce:

Proposition. *P_+^n, as a representation of S_{2n}, coincides with the permutation representation $\mathrm{Ind}_{H_n}^{S_{2n}} \mathbb{C}$ associated to S_{2n}/H_n.*

P_-^n, as a representation of S_{2n}, coincides with the representation $\mathrm{Ind}_{H_n}^{S_{2n}} \mathbb{C}(\epsilon)$ induced to S_{2n} from the sign representation $\mathbb{C}(\epsilon)$ of H_n.

Next, one can describe these representations in terms of irreducible representations of S_{2n}. Using the formulas 4.5.1,2 and Chapter 9, §6.4 we get

Theorem. *As a representation of S_{2n} the space P_+^n decomposes as*

$$P_+^n = S(S^2(V))^\chi := \bigoplus_{\lambda \vdash^{er} 2n} S_\lambda(V)^\chi = \bigoplus_{\lambda \vdash^{er} 2n} M_\lambda. \tag{6.2.1}$$

P_-^n decomposes as

$$P_-^n = S\Big(\bigwedge^2(V)\Big)^\chi := \bigoplus_{\lambda \vdash^{ec} 2n} S_\lambda(V)^\chi = \bigoplus_{\lambda \vdash^{ec} 2n} M_\lambda. \tag{6.2.2}$$

Remark. By Frobenius reciprocity (Chapter 8, §5.2), this theorem computes the multiplicity of the trivial and the sign representation of H_n in any irreducible representation of S_{2n}. Both appear with multiplicity at most 1, and in fact the trivial representation appears when λ has even columns and the sign representation when it has even rows.

One can think of H_n as a *spherical subgroup* of S_{2n}.

Next we apply the ideas of 6.1 to intertwiners. We do it only in the simplest case.

6.3 Orthogonal and Symplectic Group

1. Orthogonal case. Let V be an orthogonal space of dimension m, and consider the space $\mathcal{I}_{2n}^+$ of multilinear invariants in $2n$ variables $u_i \in V$, $i = 1, \dots, 2n$. $\mathcal{I}_{2n}^+$ is spanned by the monomials $(u_{i_1}, u_{i_2})(u_{i_3}, u_{i_4}) \dots (u_{i_{2n-1}}, u_{i_{2n}})$, where $i_1, i_2, \dots, i_{2n}$ is a permutation of $1, 2, \dots, 2n$.

Let $y_{ij} = y_{ji}$ be symmetric variables. Under the map

$$\mathbb{C}[y_{ij}] \to \mathbb{C}[(u_i, u_j)], \quad y_{ij} \mapsto (u_i, u_j),$$

the space P_+^n maps surjectively onto $\mathcal{I}_{2n}^+$ with kernel $P_+^n \cap I_{m+1}^+$.

The same proof as in 6.1 shows that

Theorem 1. *As a representation of S_{2n} we have*

$$P_+^n \cap I_{m+1}^+ = \bigoplus_{\lambda \vdash^{er} 2n,\ ht(\lambda) \geq m+1} M_\lambda, \quad \mathcal{I}_{2n}^+ = \bigoplus_{\lambda \vdash^{er} 2n,\ ht(\lambda) \leq m} M_\lambda.$$

The interpretation of the relations in the algebras of intertwiners $\mathrm{End}_{O(V)} V^{\otimes n}$ is more complicated and we shall not describe it in full detail.

2. Symplectic case. Let V be a symplectic space of dimension $2m$ and consider the space $\mathcal{I}_{2n}^-$ of multilinear invariants in $2n$ variables $u_i \in V$, $i = 1, \ldots, 2n$. $\mathcal{I}_{2n}^-$ is spanned by the monomials $[u_{i_1}, u_{i_2}][u_{i_3}, u_{i_4}] \ldots [u_{i_{2n-1}}, u_{i_{2n}}]$, where $i_1, i_2, \ldots, i_{2n}$ is a permutation of $1, 2, \ldots, 2n$.

Let $y_{ij} = -y_{ji}$ be antisymmetric variables. Under the map

$$\mathbb{C}[y_{ij}] \to \mathbb{C}[[u_i, u_j]], \ y_{ij} \mapsto [u_i, u_j],$$

the space P_-^n maps surjectively onto $\mathcal{I}_{2n}^-$ with kernel $P_-^n \cap I_{m+1}^-$.

The same proof as in 6.1 shows that

Theorem 2. *As a representation of S_{2n}, we have*

$$P_-^n \cap I_{m+1}^- = \bigoplus\nolimits_{\lambda \vdash^{ec} 2n, \ ht(\lambda) \geq 2m+2} M_\lambda, \ \mathcal{I}_{2n}^- = \bigoplus\nolimits_{\lambda \vdash^{ec} 2n, \ ht(\lambda) \leq 2m} M_\lambda.$$

The interpretation of the relations in the algebras of intertwiners $\mathrm{End}_{Sp(V)} V^{\otimes n}$ is again more complicated and we shall not describe it in full detail. We want nevertheless to describe part of the structure of the algebra of intertwiners, the one relative to traceless tensors $T^0(V^{\otimes n})$. In both the orthogonal and symplectic cases the idea is similar. Let us first develop the symplectic case which is simpler, and leave the discussion of the orthogonal case to §6.5.

Consider a monomial $M := y_{i_1,j_1} y_{i_2,j_2} \ldots y_{i_n,j_n}$. We can assume by symmetry $i_k < j_k$ for all k. We see by the formulas 3.1.8, 3.1.9 that the operator $\phi_M : V^{\otimes n} \to V^{\otimes n}$ involves at least one contraction unless all the indices i_k are $\leq n$ and $j_k > n$.

Let us denote by $\overline{\phi}_M$ the restriction of the operator ϕ_M to $T^0(V^{\otimes n})$. We have seen that $\overline{\phi}_M = 0$ if the monomial contains a variable $y_{ij}, \ y_{n+i,n+j}, \ i, j \leq n$. Thus the map $M \to \overline{\phi}_M$ factors through the space $\overline{P}_-^n$ of monomials obtained, setting to zero one of the two previous types of variables.

The only monomials that remain are of type $M_\sigma := \prod_{i=1}^n y_{i,n+\sigma(i)}, \sigma \in S_n$, and M_σ corresponds to the invariant

$$\prod [u_i, u_{n+\sigma(i)}] = [u_{\sigma^{-1}(1)} \otimes u_{\sigma^{-1}(2)} \otimes \cdots \otimes u_{\sigma^{-1}(n)}, u_{n+1} \otimes u_{n+2} \otimes \cdots \otimes u_{2n}]$$

which corresponds to the map induced by the permutation σ on $V^{\otimes n}$.

We have just identified $\overline{P}_-^n$ with the group algebra $\mathbb{C}[S_n]$ and the map

$$\rho : \overline{P}_-^n = \mathbb{C}[S_n] \to \mathrm{End}_{Sp(V)}(T^0(V^{\otimes n})), \quad \rho(M) := \overline{\phi}_M,$$

with the canonical map to the algebra of operators induced by the symmetric group. Since $T^0(V^{\otimes m})$ is a sum of isotypic components in $V^{\otimes m}$, the map ρ is surjective.

The image of $P_-^n \cap I_{m+1}^-$ in $\overline{P}_-^n = \mathbb{C}[S_n]$ is contained in the kernel of ρ. To identify this image, take the Pfaffian of the principal minor of Y of indices $i_1, i_2, \ldots, i_{2m+2}$ and evaluate after setting $y_{ij} = y_{n+i,n+j} = 0, \ i, j \leq n$. Let us say that h of these indices are $\leq n$ and $2m + 2 - h$ are $> n$.

The specialized matrix has block form $\begin{pmatrix} 0 & Z \\ -Z^t & 0 \end{pmatrix}$ and the minor extracted from the indices $i_1, i_2, \ldots, i_{2m+2}$ contains a square block matrix, made of 0's, whose size is the larger of the two numbers h and $2m + 2 - h$.

Since the maximum dimension of a totally isotropic space, for a nondegenerate symplectic form on a $2m+2$-dimensional space, is $m+1$, we deduce that the only case in which this Pfaffian can be nonzero is when $h = 2m+2-h = m+1$. In this case Z is a square $(m+1) \times (m+1)$ matrix, and the Pfaffian equals $\det(Z)$.

Thus, arguing as in the linear case, we see that the image of $P_-^n \cap I_{m+1}^-$ in $\overline{P}_-^n = \mathbb{C}[S_n]$ is the ideal generated by the antisymmetrizer on $m+1$ elements.

Theorem 3. *The algebra* $\mathrm{End}_{Sp(V)}(T^0(V^{\otimes n}))$ *equals the algebra* $\mathbb{C}[S_n]$ *modulo the ideal generated by the antisymmetrizer on* $m+1$ *elements.*

Proof. We have already seen that the given algebra is a homomorphic image of the group algebra of S_n modulo the given ideal. In order to prove that there are no further relations, we observe that if $U \subset V$ is the subspace spanned by $e_1, \dots, e_m$, it is a (maximal) totally isotropic subspace and thus $U^{\otimes n} \subset T^0(V^{\otimes n})$. On the other hand, by the linear theory, the kernel of the action of $\mathbb{C}[S_n]$ on $U^{\otimes n}$ coincides with the ideal generated by the antisymmetrizer on $m+1$ elements, and the claim follows. □

6.4 Irreducible Representations of $Sp(V)$

We are now ready to exhibit the list of irreducible rational representations of $Sp(V)$. First, using the double centralizer theorem, we have a decomposition

$$T^0(V^{\otimes n}) = \bigoplus_{\lambda \vdash n,\ ht(\lambda) \le m} M_\lambda \otimes T_\lambda(V), \tag{6.4.1}$$

where we have indicated by $T_\lambda(V)$ the irreducible representation of $Sp(V)$ paired to M_λ. We should note then that we can construct, as in 4.1, the tensor

(6.4.2)
$$(e_1 \wedge e_2 \wedge \dots \wedge e_{n_1}) \otimes (e_1 \wedge e_2 \wedge \dots \wedge e_{n_2}) \otimes \cdots \otimes (e_1 \wedge e_2 \wedge \dots \wedge e_{n_t}) \in T_\lambda(V)$$

where the partition λ has columns $n_1, n_2, \dots, n_t$. We ask the reader to verify that it is a highest weight vector for $T_\lambda(V)$ with highest weight $\sum_{j=1}^t \omega_{n_j}$.

Since $Sp(V)$ is contained in the special linear group, from Proposition 1.4 of Chapter 7, all irreducible representations appear in tensor powers of V. Since $T^0(V^{\otimes m})$ contains all the irreducible representations appearing in $V^{\otimes n}$ and not in $V^{\otimes k}$, $k < n$, we deduce:

Theorem. *The irreducible representations* $T_\lambda(V)$, $ht(\lambda) \le m$, *constitute a complete list of inequivalent irreducible representations of* $Sp(V)$.

Since $Sp(V)$ is simply connected (Chapter 5, §3.10), from Theorem 6.1 of Chapter 10 it follows that this is also the list of irreducible representations of the Lie algebra $sp(2m, \mathbb{C})$. Of course, we are recovering in a more explicit form the classification by highest weights.

6.5 Orthogonal Intertwiners

We want to describe also in the orthogonal case part of the structure of the algebra of intertwiners, the one relative to traceless tensors $T^0(V^{\otimes n})$.

We let V be an m-dimensional orthogonal space. Consider a monomial $M := y_{i_1,j_1}y_{i_2,j_2}\dots y_{i_n,j_n}$. We can assume by symmetry that $i_k < j_k$ for all k. We see by the formulas 3.1.8 and 3.1.9 that the operator $\phi_M : V^{\otimes n} \to V^{\otimes n}$ involves at least one contraction unless all the indices i_k are $\leq n$ and $j_k > n$.

Let us denote by $\overline{\phi}_M$ the restriction of the operator ϕ_M to $T^0(V^{\otimes n})$. We have seen that $\overline{\phi}_M = 0$ if the monomial contains a variable $y_{ij},\ y_{n+i,n+j},\ i,j \leq n$. Thus the map $M \to \overline{\phi}_M$ factors through the space $\overline{P}^n_+$ of monomials obtained by setting to 0 one of the two previous types of variables. The only monomials that remain are of type

$$M_\sigma := \prod_{i=1}^n y_{i,n+\sigma(i)}, \qquad \sigma \in S_n,$$

and M_σ corresponds to the invariant

$$\prod (u_i, u_{n+\sigma(i)}) = (u_{\sigma^{-1}(1)} \otimes u_{\sigma^{-1}(2)} \otimes \cdots \otimes u_{\sigma^{-1}(n)}, u_{n+1} \otimes u_{n+2} \otimes \cdots \otimes u_{2n}),$$

which corresponds to the map induced by the permutation σ on $V^{\otimes n}$.

We have just identified $\overline{P}^n_+$ with the group algebra $\mathbb{C}[S_n]$ and the map

$$\rho : \overline{P}^n_+ = \mathbb{C}[S_n] \to \mathrm{End}_{Sp(V)}(T^0(V^{\otimes n})), \quad \rho(M) := \overline{\phi}_M,$$

with the canonical map to the algebra of operators induced by the symmetric group. Since $T^0(V^{\otimes n})$ is a sum of isotypic components, the map ρ is surjective.

Let us identify the image of $P^n_+ \cap I^+_{m+1}$ in $\overline{P}^n_+ = \mathbb{C}[S_n]$ which is in the kernel of ρ.

Take the determinant D of an $(m+1) \times (m+1)$ minor of Y extracted from the row indices $i_1, i_2, \dots, i_{m+1}$, the column indices $j_1, j_2, \dots, j_{m+1}$, and evaluate after setting $y_{ij} = y_{n+i,n+j} = 0,\ i,j \leq n$. Let us say that h of the row indices are $\leq n$ and $m+1-h$ are $> n$ and also k of the column indices are $\leq n$ and $m+1-k$ are $> n$.

The specialized matrix has block form $\begin{pmatrix} 0 & Z \\ W & 0 \end{pmatrix}$ where Z is an $h \times m+1-k$ and W an $m+1-h \times k$ matrix. If this matrix has nonzero determinant, the image of the first k basis vectors must be linearly independent. Hence $m+1-h \geq k$, and similarly $h \leq m+1-k$. Hence $h+k = m+1$, i.e., Z is a square $h \times h$ and W is a square $k \times k$ matrix.

Up to sign the determinant of this matrix is $D = \det(W)\det(Z)$.

This determinant is again a weight vector which, multiplied by a monomial M in the y_{ij}, can give rise to an element of $DM \in P^n_+$ if and only if the indices $i_1, i_2, \dots, i_{m+1}$ and $j_1, j_2, \dots, j_{m+1}$ are all distinct.

Up to a permutation in $S_n \times S_n$ we may assume then that these two sets of indices are $1, 2, \dots, h, n+h+1, n+h+2, \dots, n+h+k$ and $h+1, h+2, \dots, h+k$, $n+1, n+2, \dots, n+h$ so that using the symmetry $y_{n+h+i,h+j} = y_{h+j,n+h+i}$, we obtain that $\det(W)\det(Z)$ equals

$$\sum_{\sigma \in S_h} \epsilon_\sigma y_{\sigma(1),n+1} y_{\sigma(2),n+2} \cdots y_{\sigma(h),n+h}$$
$$\times \sum_{\sigma \in S_k} \epsilon_\sigma y_{h+\sigma(1),n+h+1} y_{h+\sigma(2),n+h+2} \cdots y_{h+\sigma(k),n+h+k}.$$

We multiply this by $M := \prod_{t=1}^{n-h-k} y_{h+k+t,n+h+k+t}$.

This element corresponds in $\mathbb{C}[S_n]$ to the antisymmetrizer relative to the partition consisting of the parts $(1, 2, \dots, h)$, $(h+1, \dots, h+k)$, 1^{n-h-k}. This is the element $\sum_{\sigma \in S_h \times S_k} \epsilon_\sigma \sigma$. A determinant of a minor of order $> m+1$ can be developed into a linear combination of monomials times determinants of order $m+1$; thus the ideal of $\mathbb{C}[S_n]$ corresponding to $P_+^n \cap I_{m+1}^+$ contains the ideal generated by the products of two antisymmetrizers on two disjoint sets whose union has $\geq m+1$ elements.

From the description of Young symmetrizers, it follows that each Young symmetrizer, relative to a partition with the first two columns adding to a number $\geq m+1$ is in the ideal generated by such products. Thus we see that:

The image of $P_+^n \cap I_{m+1}^+$ in $\overline{P}_+^n = \mathbb{C}[S_n]$ contains the ideal generated by all the Young symmetrizers relative to diagrams with the first two columns adding to a number $\geq m+1$.

Theorem 1. *The algebra $\mathrm{End}_{O(V)}(T^0(V^{\otimes n}))$ equals the algebra $\mathbb{C}[S_n]$ modulo the ideal generated by all the Young symmetrizers relative to diagrams with the first two columns adding to a number $\geq m+1$.*

Proof. We have already seen that the given algebra is a homomorphic image of the group algebra of S_n modulo the given ideal. In order to prove that there are no further relations it is enough to show that if λ is a partition with the first two columns adding up to at most m, then we can find a nonzero tensor u with $c_T u$ traceless and $c_T u \neq 0$, where c_T is a Young symmetrizer of type λ.

For this we cannot argue as simply as in the symplectic case: we need a variation of the theme. First, consider the diagram of λ filled in increasing order from up to down and right to left with the numbers $1, 2, \dots, n$, e.g.,

$$\begin{matrix} 1 & 5 & 8 & 11 \\ 2 & 6 & 9 & \\ 3 & 7 & 10 & \\ 4 & & & \end{matrix}$$

Next suppose we fill it as a tableau with some of the basis vectors e_i, f_j, e.g.,

$$\begin{matrix} e_2 & f_3 & e_1 & e_1 \\ f_2 & e_4 & f_4 & \\ e_3 & e_3 & f_1 & \\ f_4 & & & \end{matrix}$$

To this display we associate a tensor u in which we place in the i^{th} position the vector placed in the tableau in the case labeled with i, e.g., in our previous example:

$$u := e_2 \otimes f_2 \otimes e_3 \otimes f_4 \otimes f_3 \otimes e_4 \otimes e_3 \otimes e_1 \otimes f_4 \otimes f_1 \otimes e_1.$$

If the vectors that we place on the columns are all distinct, the antisymmetrization $a_T u \neq 0$.

Assume first $m = 2p$ is even. If the first column has $\leq p$ elements, we can work in the totally isotropic subspace $U := \langle e_1, \dots, e_p \rangle$. As for the symplectic group, $U^{\otimes n}$ is made of traceless tensors. On $U^{\otimes n}$ the group algebra $\mathbb{C}S_n$ acts with the kernel generated by the antisymmetrizer on $p+1$ elements. We construct the display with e_i on the i^{th} row. The associated tensor u is a highest weight vector as in 6.4.2.

Otherwise, let $p+s,\ p-t,\ s \leq t$ be the lengths of the first two columns.

We first fill the diagram with e_i in the i^{th} row, $i \leq p$, and we are left with s rows with just 1 element, which we fill with $f_p, f_{p-1}, \dots, f_{p-s+1}$. This tensor is symmetric in the row positions, so when we apply to it the corresponding Young symmetrizer we only have to antisymmetrize the columns.

When we perform a contraction on this tensor there are s possible contractions that are nonzero (in fact have value 1) and that correspond to the indices of the first column occupied by the pairs e_{p-i}, f_{p-i}. Notice that if we exchange these two positions, the contraction does not change.

It follows that when we antisymmetrize this element, we get a required nonzero traceless tensor in M_λ. Explicitly, up to a permutation we have the tensor:

$$\begin{aligned} &(e_1 \wedge \dots \wedge e_p \wedge f_p \wedge f_{p-1} \wedge \dots \wedge f_{p-s+1}) \otimes (e_1 \wedge \dots \wedge e_{p-t}) \\ &\otimes \cdots \otimes (e_1 \wedge \dots \wedge e_{n_t}) \in T_\lambda(V). \end{aligned} \tag{6.5.1}$$

The odd case is similar. □

Proposition. *The tensor* 6.5.1 *is a highest weight vector of weight* $\omega_{p-s} + \omega_{p-t} + \sum \omega_{n_i}$.

To prove the proposition, one applies the elements $\underline{e}_i$[110] defined in Chapter 10, §4.1.7 and 4.1.21 and one uses the formula 5.1.1 of the same chapter.

We are now ready to exhibit the list of irreducible rational representations of $O(V)$. First, using the double centralizer theorem, we have a decomposition

$$T^0(V^{\otimes n}) = \bigoplus_{\lambda \vdash n,\ h_1 + h_2 \leq m} M_\lambda \otimes T_\lambda(V),$$

where we have indicated by $T_\lambda(V)$ the irreducible representation of $O(V)$ paired to M_λ and h_1, h_2 the first two columns of λ.

The determinant representation of $O(V)$ is contained in $V^{\otimes m}$ and it is equal to its inverse. Hence, from Theorem 1.4 of Chapter 7, all irreducible representations appear in tensor powers of V. Since $T^0(V^{\otimes n})$ contains all the irreducible representations appearing in $V^{\otimes n}$ and not in $V^{\otimes k},\ k < n$, we deduce

Theorem 2. *The irreducible representations $T_\lambda(V)$, $h_1 + h_2 \leq m$, form a complete list of inequivalent irreducible representations of $O(V)$.*

[110] We underline to indicate the Chevalley generators, in order not to confuse them with the basis elements.

6.6 Irreducible Representations of $SO(V)$

We can now pass to $SO(V)$. In this case there is one more invariant $[v_1, \dots, v_m]$ which gives rise to new intertwiners.

Moreover, since $SO(V)$ has index 2 in $O(V)$, we can apply Clifford's Theorem (Chapter 8, §5.1), and deduce that each irreducible representation M of $O(V)$ remains irreducible under $SO(V)$ or splits into two irreducibles according whether it is or it is not isomorphic to $M \otimes \epsilon$. The sign representation is the one induced by the determinant.

The theory is nontrivial only when $\dim V = 2n$ is even. In fact in the case of $\dim V$ odd, we have that $O(V) = SO(V) \times \mathbb{Z}/(2)$ where $\mathbb{Z}/(2)$ is ± 1 (plus or minus the identity matrix). Therefore an irreducible representation of $O(V)$ remains irreducible when restricted to $SO(V)$.

To put together these two facts we start with:

Lemma 1. *Let $T : V^{\otimes p} \to V^{\otimes q}$ be an $SO(V)$-invariant linear map.*

We can decompose $T = T_1 + T_2$ so that:

T_1 is $O(V)$-linear, and T_2 is $O(V)$-linear provided we twist $V^{\otimes q}$ by the determinant representation.

Given any irreducible representation $N \subset V^{\otimes p}$, if $T_2(N) \neq 0$, then $T_2(N)$ is isomorphic to $N \otimes \epsilon$ (ϵ is the determinant or sign representation of $O(V)$).

Proof. $\mathbb{Z}/(2) = O(V)/SO(V)$ acts on the space of $SO(V)$-invariant linear maps. Any such map is canonically decomposed as a sum $T = T_1 + T_2$ for the two eigenvalues ± 1 of $\mathbb{Z}/(2)$. Then T_1 is $O(V)$-invariant while for $X \in O(V)$, we have $T_2(Xa) = \det(X)X(T_2(a))$, the required condition.

The second part is an immediate consequence of the first. □

Let us use the notation $[v_1, \dots, v_m] := v_1 \wedge v_2 \wedge \dots \wedge v_m$. First, let us analyze, for $k \leq m = \dim(V)$, the operator

$$* : \otimes^k V \to \otimes^{m-k} V,$$

defined by the implicit formula (using the induced scalar product on tensor):

$$(*(v_1 \otimes \cdots \otimes v_k), v_{k+1} \otimes \cdots \otimes v_m) = [v_1, \dots, v_m]. \tag{6.6.1}$$

Remark that if $\sigma \in S_{m-k}$ we have, by symmetry of the scalar product,

$$\begin{aligned}
&(\sigma * (v_1 \otimes \cdots \otimes v_k), v_{k+1} \otimes \cdots \otimes v_m) \\
&= (*(v_1 \otimes \cdots \otimes v_k), \sigma^{-1}(v_{k+1} \otimes \cdots \otimes v_m)) \\
&= [v_1, \dots, v_k, v_{\sigma(k+1)}, \dots, v_{\sigma(m)}] \\
&= \epsilon_\sigma [v_1, \dots, v_m].
\end{aligned}$$

This implies that $*(v_1 \otimes \cdots \otimes v_k) \in \bigwedge^{m-k} V$. Similarly

$$*(\sigma(v_1 \otimes \cdots \otimes v_k)) = \epsilon_\sigma * (v_1 \otimes \cdots \otimes v_k)$$

implies that $*$ factors through a map

$$* : \bigwedge^k V \to \bigwedge^{m-k} V$$

still defined by the implicit formula

$$(6.6.2) \qquad (*(v_1 \wedge \ldots \wedge v_k), v_{k+1} \wedge \ldots \wedge v_m) = [v_1, \ldots, v_m].$$

In particular exterior powers are pairwise isomorphic under $SO(V)$.

In an orthonormal oriented basis u_i we have $*(u_{i_1} \wedge \ldots \wedge u_{i_k}) = \epsilon u_{j_1} \wedge \ldots \wedge u_{j_{m-k}}$, where $i_1, \ldots, i_k, j_1, \ldots, j_{m-k}$ is a permutation of $1, \ldots, m$, and ϵ is the sign of this permutation.

It is better to show one case explicitly in a hyperbolic basis, which explains in a simple way some of the formulas we are using. We do the even case in detail. Let V be of even dimension $2n$ with the usual hyperbolic basis e_i, f_i. Then,

$$(6.6.3) \qquad *(e_1 \wedge e_2 \wedge \ldots \wedge e_k) = f_{k+1} \wedge f_{k+2} \wedge \ldots \wedge f_n \wedge e_1 \wedge e_2 \wedge \ldots \wedge e_n.$$

Proof. By definition, taking the exterior products of the hyperbolic basis as the basis of $\bigwedge^{2n-k}(V)$ we see that the only nonzero scalar product of $*(e_1 \wedge e_2 \wedge \ldots \wedge e_k)$ with the basis elements is $(*(e_1 \wedge e_2 \wedge \ldots \wedge e_k), e_{k+1} \wedge e_{k+2} \wedge \ldots \wedge e_n \wedge f_1 \wedge \ldots \wedge f_n) = e_1 \wedge e_2 \wedge \ldots \wedge \ldots \wedge e_n \wedge f_1 \wedge \ldots \wedge f_n = 1$. □

In particular, notice that this implies that $f_{k+1} \wedge f_{k+2} \wedge \ldots \wedge f_n \wedge e_1 \wedge e_2 \wedge \ldots \wedge e_n$ is a highest weight vector of $\bigwedge^{2n-k} V$.

In general let us understand what type of intertwiners we obtain on traceless tensors using this new invariant. By the same principle of studying only *new invariants* we start by looking at $SO(V)$ equivariant maps $\gamma : V^{\otimes p} \to V^{\otimes q}$ which induce nonzero maps $\gamma : T^0(V^{\otimes p}) \to T^0(V^{\otimes q})$.

We may restrict to a map γ corresponding to an $SO(V)$ but not to $O(V)$-invariant $(\gamma(u_1 \otimes \cdots \otimes u_p), v_1 \otimes \cdots \otimes v_q)$, which is the product of one single determinant and several scalar products. We want to restrict to the ones that give rise to operators $T : V^{\otimes p} \to V^{\otimes q}$ which do not vanish on traceless tensors and also which cannot be factored through an elementary extension. This implies that the invariant $(T(u_1 \otimes \cdots \otimes u_p), v_1 \otimes \cdots \otimes v_q)$ should not contain any *homosexual contraction.* Thus, up to permuting the $u's$ and $v's$ separately, we are reduced to studying the maps γ_k from $T^0(V^{k+t})$ to $T^0(V^{2n-k+t})$ induced by the invariants:

$$(\gamma_k(u_1 \otimes \cdots \otimes u_{k+t}), v_1 \otimes \cdots \otimes v_{2n+t-k})$$

$$(6.6.4) \qquad = \prod_{i=1}^{t} (u_{k+i}, v_{2n-k+i})[u_1, \ldots, u_k, v_1, \ldots, v_{2n-k}].$$

Lemma 2. *If k is the length of the first column of λ, then γ_k maps $T_\lambda(V)$ to $T_{\lambda'}(V)$, where λ' is obtained from λ substituting the first column k with $2n - k$.*

Thus $T_\lambda(V) \otimes \epsilon = T_{\lambda'}(V)$.

Proof. First, it is clear, from 6.6.1, that we have

$$\gamma(u_1 \otimes u_2 \otimes \cdots \otimes u_k \otimes \cdots \otimes u_{k+t})$$
$$= *(u_1 \otimes u_2 \otimes \cdots \otimes u_k) \otimes u_{k+1} \otimes \cdots \otimes u_{k+t}.$$

In the lemma we have two *dual* cases, $k \leq 2n - k$ or $k \geq 2n - k$. For instance in the second case, let $k = n + s,\ s \leq n$. Let us compute using 6.6.3:

$$\begin{aligned}
&\gamma[(e_1 \wedge e_2 \wedge \ldots \wedge e_n \wedge f_n \wedge f_{n-1} \wedge \ldots \wedge f_{n-s+1}) \otimes (e_1 \wedge e_2 \wedge \ldots \wedge e_{n-t}) \\
&\quad \otimes \cdots \otimes (e_1 \wedge e_2 \wedge \ldots \wedge e_{h_r})] \\
&= [*(e_1 \wedge e_2 \wedge \ldots \wedge e_n \wedge f_n \wedge f_{n-1} \wedge \ldots \wedge f_{n-s+1})] \\
&\quad \otimes (e_1 \wedge e_2 \wedge \ldots \wedge e_{n-t}) \otimes \ldots \otimes (e_1 \wedge e_2 \wedge \ldots \wedge e_{h_r}) \\
&= (e_1 \wedge e_2 \wedge \ldots \wedge e_s) \otimes (e_1 \wedge e_2 \wedge \ldots \wedge e_{n-t}) \otimes \ldots \otimes (e_1 \wedge e_2 \wedge \ldots \wedge e_{h_r}).
\end{aligned}$$

Thus γ maps a highest weight vector of $T_\lambda(V)$ to one of $T_{\lambda'}(V)$. Since these two representations are irreducible as $O(V)$-modules, the claim follows. □

We will say that the two partitions λ, λ' are *associated*. By definition $\lambda = \lambda'$ if and only only if the first column has length n. If the two associated partitions are distinct, one of them has the first column of length $< n$ and the other $> n$. Thus we have:

Proposition 1. *$T_\lambda(V)$ is isomorphic (as an $O(V)$ module) to $T_\lambda(V) \otimes \epsilon$ if and only if λ is self-associated, i.e., $\lambda = \lambda'$.*

Proof. We have already seen that $T_\lambda(V) \otimes \epsilon = T_{\lambda'}(V)$. □

In the case $\lambda = \lambda'$, from Clifford's theorem $T_\lambda(V)$ decomposes as the direct sum of two irreducible representations under $SO(V)$. It is interesting to write explicitly the two highest weight vectors of $SO(V)$.

Proposition 2. *In $\bigwedge^n V$, $\dim V = 2n$ we have the two highest weight vectors:*

$$e_1 \wedge e_2 \wedge \ldots \wedge e_n, \qquad e_1 \wedge e_2 \wedge \ldots \wedge e_{n-1} \wedge f_n,$$

with weights: $2s_\pm = \sum_{i=1}^{n-1} \alpha_i \pm \alpha_n$. (cf. Chapter 10, §5.1.1)

If $\dim V = 2n + 1$, choose a hyperbolic basis $e_1, \ldots, e_n, f_1, \ldots, f_n, u$.

$\bigwedge^n V \equiv \bigwedge^{n+1} V$ are irreducible under $SO(V)$ with highest weight vectors $e_1 \wedge e_2 \wedge \ldots \wedge e_n$, $e_1 \wedge e_2 \wedge \ldots \wedge e_n \wedge u$ and weight $2s = \sum_{i=1}^n \alpha_i$. (cf. Chapter 10, §5.1.2).

Proof. By definition we have to check that the two vectors are killed by the Chevalley generators of $SO(V)$. We then use the formulas in 4.1.21 of Chapter 10 (unfortunately we have a certain overlap of symbols). The analysis in the odd case is similar. □

For general $\lambda = \lambda'$ we construct the highest weight vectors as follows.

We build a tensor as in the pattern of 6.5. We insert e_i in the i^{th} row for $i \leq n-1$, then the n^{th} row; in one case we insert e_n, in the other f_n.

Summarizing we have:

Theorem. *If* $\dim V = 2n+1$ *is odd, the irreducible representations of* $SO(V)$ *are the* $T_\lambda(V)$ *indexed by partitions* λ *of height* $\leq n$.

If $\dim V = 2n$ *is even, the irreducible representations of* $SO(V)$ *are of two types:*

(1) The restriction to $SO(V)$ *of the irreducible representations* $T_\lambda(V)$ *of* $O(V)$ *indexed by partitions* λ *of height* $\leq n$ *which are not self-associated.*
(2) For each self-associated partition $\lambda = \lambda'$, *the two irreducible representations in which* $T_\lambda(V)$ *splits.*

In both cases their highest weights are linear combinations of the fundamental weights where the spin weights appear with even coefficients.

Remark. From the formulas of Chapter 10, §4.1.3 and our discussion it follows that, in the odd case $2n+1$, the exterior powers $\bigwedge^i V$, $i < n$, are fundamental representations. For $SO(2n+1)$, $\bigwedge^n V$ is irreducible corresponding to twice the spin representation which is fundamental.

On the other hand, when the dimension is even $2n$, the exterior powers $\bigwedge^i V$, $i < n-1$, are fundamental representations. The other two fundamental weights $\pm s$ belong to spin representations which do not appear in tensor powers. The exterior power $\bigwedge^n V$ decomposes as a direct sum of two irreducible representations corresponding to twice the half-spin representations. The exterior power $\bigwedge^{n-1} V$ appears as the leading term of the tensor product of the two half-spin representations (cf. §7.1).

The explanation is that in tensor powers we find the Lie algebra representations which, integrated to the spin group, factor through the special orthogonal group.

6.7 Fundamental Representations

We give here a complement to the previous theory by analyzing the action of the symplectic group of a space V on the exterior algebra in order to describe the *fundamental* representations.

We start with the symplectic case which in some way is more interesting. Assume $\dim V = 2n$.

First, by Theorem 6.4 we have that the traceless tensors $T^0(V^{\otimes m})$ contain a representation associated to the full antisymmetrizer (the sign representation of S_m) if and only if $m \leq n$. This is a new representation in $V^{\otimes m}$, hence it appears only in the traceless tensors and, by 6.4.1, with multiplicity 1.

Let us denote by $\bigwedge_0^m(V)$ this representation which, by what we have just seen, appears with multiplicity 1 also in $\bigwedge^m(V)$. The element $e_1 \wedge e_2 \wedge \ldots \wedge e_m$ is its highest weight vector. Its weight is the fundamental weight ω_m (cf. Chapter 10, §5.1.3).

Remark next that, by definition, $\bigwedge^2 V$ contains a canonical bivector

$$\psi := \sum_{i=1}^{n} e_i \wedge f_i$$

invariant under $Sp(V)$.

We want to compute now the $Sp(V)$ equivariant maps between $\bigwedge^k V$ and $\bigwedge^h V$.

Since the skew-symmetric tensors are direct summands in tensor space, any $Sp(V)$ equivariant map between $\bigwedge^k V$ and $\bigwedge^h V$ can be decomposed as $\bigwedge^k V \xrightarrow{i} V^{\otimes k} \xrightarrow{p} V^{\otimes h} \xrightarrow{A} \bigwedge^h V$ where i is the canonical inclusion, p is some equivariant map and A is the antisymmetrizer.

We have seen, in 3.1, how to describe equivariant maps $V^{\otimes k} \xrightarrow{p} V^{\otimes h}$ up to the symmetric group action.

If we apply the symmetric groups to i or to A, it changes at most the sign. In particular we see that the insertion maps $\bigwedge^k V \to \bigwedge^{k+2} V$ can, up to sign, be identified with $u \to \psi \wedge u$, which we shall also call ψ. For the contraction, we have a unique map (up to constant) which can be normalized to

$$(6.7.1)\quad c : v_1 \wedge v_2 \wedge \ldots \wedge v_k \to \sum_{i<j} (-1)^{i+j-1} \langle v_i, v_j \rangle v_1 \wedge v_2 \wedge \check{v}_i \ldots \check{v}_j \ldots \wedge v_k.$$

The general formula of 3.1 gives in this case that

Lemma. *All the maps between $\bigwedge^k V$ and $\bigwedge^h V$ are linear combinations of $\psi^i c^j$.*

Definition. The elements

$$\bigwedge\nolimits_0(V) := \left\{ a \in \bigwedge V \mid c(a) = 0 \right\}$$

are called *primitive*.

In order to understand the commutation relations between these two maps, let us set $h := [c, \psi]$. Go back to the spin formalism of Chapter 5, §4.1 and recall the formulas of the action of the Clifford algebra on the exterior power:

$$i(v)(u) := v \wedge u, \quad j(\varphi)(v_1 \wedge v_2 \ldots \wedge v_k)$$
$$:= \sum_{t=1}^{k} (-1)^{t-1} \langle \varphi | v_t \rangle v_1 \wedge v_2 \ldots \check{v}_t \ldots \wedge v_k$$

together with the identity

$$i(v)^2 = j(\varphi)^2 = 0, \quad i(v)j(\varphi) + j(\varphi)i(v) = \langle \varphi | v \rangle.$$

Now clearly as an operator we have $\psi := \sum_i i(e_i)i(f_i)$. Using the dual basis let us show that $c = \sum_i j(f^i)j(e^i)$. Let us drop the symbols i, j and compute directly in the Clifford algebra of $V \oplus V^*$ with the standard hyperbolic form:

$$\sum_i f^i e^i v_1 \wedge v_2 \wedge \ldots \wedge v_k = \sum_i \sum_{s<t} (-1)^{t+s} (\langle e^i | v_t \rangle \langle f^i | v_s \rangle - \langle f^i | v_t \rangle \langle e^i | v_s \rangle$$

$$\times v_1 \wedge v_2 \wedge \check{v}_s \ldots \check{v}_t \ldots \wedge v_k$$

$$= \sum_{s<t} (-1)^{t+s-1} \langle v_s | v_t \rangle v_1 \wedge v_2 \wedge \check{v}_s \ldots \check{v}_t \ldots \wedge v_k.$$

Now we can use the commutation relations

$$e_i e^j + e^j e_i = 0,\ i \neq j,\ e_i f^j + f^j e_i = 0,\ f_i e^j + e^j f_i = 0,$$

$$f_i f^j + f^j f_i = 0,\ i \neq j,\ e_i e^i + e^i e_i = 1,\ f_i f^i + f^i f_i = 1$$

to deduce that

$$h = [c, \psi] = \sum_{i,j} [f^j e^j, e_i f_i] = \sum_i [f^i e^i, e_i f_i] = \sum_i (f^i e^i e_i f_i - e_i f_i f^i e^i)$$

$$= \sum_i (-f^i e_i e^i f_i + f^i f_i - e_i e^i + e_i f^i f_i e^i)$$

$$= \sum_i (-f^i e_i e^i f_i + f^i f_i + e^i e_i + f^i e_i e^i f_i) \sum_i (f^i f_i + e^i e_i).$$

Now we claim that on $\bigwedge^k V$ the operator $\sum_i (f^i f_i + e^i e_i)$ acts as $2n - k$. In fact when we consider a vector $u := v_1 \wedge v_2 \ldots \wedge v_k$ with the v_i from the symplectic basis, the operators $f^i f_i, e^i e_i$ annihilate u if e_i, f_i is one of the vectors $v_1, \ldots, v_k$. Otherwise they map u into u itself.

Proposition. *The elements c, ψ, h satisfy the commutation relations of the standard generators e, f, h of $sl(2, \mathbb{C})$.*

Proof. We need only show that $[h, c] = 2c,\ [c, \psi] = -2\psi$. This follows immediately from the fact that c maps $\bigwedge^k V$ to $\bigwedge^{k-2} V$, ψ maps $\bigwedge^k V$ to $\bigwedge^{k+2} V$, while h has eigenvalue $2n - k$ on $\bigwedge^k V$. □

We can apply now the representation theory of $sl(2, \mathbb{C})$ to deduce

Theorem. *We have the direct sum decomposition:*

(1) $$\bigwedge V = \bigwedge\nolimits_0(V) \oplus \psi \wedge \left(\bigwedge V\right),$$

(2) $$\bigwedge\nolimits_0(V) = \bigoplus\nolimits_{m \leq n} \bigwedge\nolimits_0^m(V),$$

(3) $$\bigwedge^k V = \bigoplus\nolimits_{2i \geq k-n} \bigwedge\nolimits_0^{k-2i}(V) \wedge \psi^i.$$

Proof. For every finite-dimensional representation M of $sl(2, \mathbb{C})$, we have a decomposition $M = M^e \bigoplus_{i>0} f^i M,\ M^e := \{m \in M \mid em = 0\}$, by highest weight theory. Let $M = \bigwedge V$. Since all the contractions reduce to c on skew-symmetric tensors, the traceless skew-symmetric tensors $\bigoplus_{m \leq n} \bigwedge_0^m(V)$ are the kernel of c or $\bigwedge_0(V) = \bigoplus_{m \leq n} \bigwedge_0^m(V)$. Thus, $\bigwedge^k V = \bigoplus_i \bigwedge_0(V) \wedge \psi^i$. Comparing degrees in this formula finally gives all the statements. □

We can interpret the first part of the previous theorem as saying that:

Corollary. *The quotient algebra* $\bigwedge V/\psi \wedge (\bigwedge V) = \bigoplus_{m \leq n} \bigwedge_0^m (V)$ *is the direct sum of the trivial and the fundamental representations of* $Sp(V)$ *in the same way as the exterior powers* $\bigwedge^i V$, $1 \leq i < \dim V$, *are the fundamental representations of* $SL(V)$.

6.8 Invariants of Tensor Representations

When one tries to study invariants of representations different from sums of the defining representation, one finds quickly that these rings tend to become extremely complicated. The classical theory of $SL(2, \mathbb{C})$, which is essentially the theory of binary forms, (Chapter 15) shows this clearly. Nevertheless, at least in a rather theoretical sense, one could from the theory developed compute invariants of general representations for classical groups.

One starts from the general remark that for a linearly reductive group G and an equivariant surjective map $U \to V$, we have a surjective map of invariants $U^G \to V^G$. Thus if $M \supset N$ are two linear representations, it follows that the invariants of degree m on N are the restriction to N of the invariants of degree m on M.

For classical groups, up to the problem of the determinant, one can embed representations into a sum $\bigoplus_i V^{\otimes h_i}$ of tensor powers of the defining vector space V. An invariant of degree m on $\bigoplus_i V^{\otimes h_i}$ is thus an invariant linear form on $S^m(\bigoplus_i V^{\otimes h_i})$, which we may think of as being embedded into $(\bigoplus_i V^{\otimes h_i})^{\otimes m}$. We are thus led to study linear invariants on such a tensor power. This last space is clearly a (possibly very large) direct sum of tensor powers and we know, for $SL(V)$, $O(V)$, $Sp(V)$ all the linear invariants, given by various kinds of contractions. These contractions are usually expressed in symbolic form on decomosable tensors, giving rise to the symbolic method (Chapter 15). For $GL(V)$ one will have to work with direct sums of tensor powers $V^{\otimes p} \otimes (V^*)^{\otimes q}$ and the relative contractions.

In principle, these expressions give formulas of invariants which span the searched-for space of invariants for the given representation. In fact, it is almost impossible to make computations in this way due to several obstacles. First, to control the linear dependencies among the invariants given by different patterns of contraction is extremely hard. It is even worse to understand the multiplicative relations and in particular which symbolic invariants generate the whole ring of invariants, what the relations among them are, and so on. In fact, in order to give a theoretical answer to these questions Hilbert, in an impressive series of papers, laid down the foundations of modern commutative algebra (cf. Chapter 14).

7 Spinors

In this section we discuss some aspects of the theory of spinors, both from the point of view of representations as well as from the view of the theory of pure spinors.

7.1 Spin Representations

We now discuss the spin group, the spin representations and the conjugation action on the Clifford algebra of some orthogonal space W. We use the notations of Chapter 5, §4 and §5. Since $\bigwedge^2 W$ is the Lie algebra of the spin group and clearly $\bigwedge^2 W$ generates the even Clifford algebra $C^+(W)$, we have that any irreducible representation of $C^+(W)$ remains irreducible on its restriction to $\mathrm{Spin}(W)$.

We treat the case over $\mathbb{C}$, and use the symbol $C(n)$ to denote the Clifford algebra of an n-dimensional complex space. Recall that

$$C^+(n) \equiv C(n-1), \dim C(n) = 2^n, \ \ C(2m) = \mathrm{End}(\mathbb{C}^{2^m}),$$
$$C(2m+1) = \mathrm{End}(\mathbb{C}^{2^m}) \oplus \mathrm{End}(\mathbb{C}^{2^m}).$$

In the even-dimensional case we identify

$$W = V \oplus V^*, \quad C^+(W) = \mathrm{End}\big(\mathbb{C}^{2^{m-1}}\big) \oplus \mathrm{End}\big(\mathbb{C}^{2^{m-1}}\big).$$

Since $C^+(V \oplus V^*) \subset \mathrm{End}(\bigwedge V)$ and it has only two nonisomorphic irreducible representations each of dimension 2^{m-1}, the only possibility is that $\bigwedge_{\mathrm{ev}} V$, $\bigwedge_{\mathrm{odd}} V$, are exactly these two nonisomorphic representations.

Definition 1. These two representations are called the two *half-spin representations* of the even spin group, denoted S_0, S_1.

The odd case is similar; $C^+(2m+1) = C(2m) = \mathrm{End}(\mathbb{C}^{2^m})$ has a unique irreducible module of dimension 2^m. To analyze it let us follow the spin formalism for $C(2m)$ followed by the identification of $C^+(2m+1) = \mathbb{C}(2m)$ as in Chapter 5, §4.4. From that discussion, starting from a hyperbolic basis $e_1, \dots, e_n, f_1, \dots, f_n, u$, we have that $C^+(2m+1)$ is the Clifford algebra over the elements $a_i := e_i u, \ b_i := u f_i$. Now apply the spin formalism and, if V is the vector space spanned by the a_i, we consider $\bigwedge V$ as a representation of $\mathbb{C}^+(2m+1)$ and of the spin group.

Definition 2. This representation is called the *spin representation* of the odd spin group.

In the even case, we have the two half-spin representations.

Recall that the formalism works with $1/2$ of the standard hyperbolic form on $W = V \oplus V^*$. We take $e_1, \dots, e_n, f_1, \dots, f_n$ to be the standard basis with $(e_i, e_j) = (f_i, f_j) = 0$, $(e_i, f_j) = \delta_i^j/2$.

Let us now compute the highest weight vectors. We have to interpret the Chevalley generators:

$$\text{(7.1.1)} \qquad \underline{e}_i := e_{i,i+1} - e_{n+i+1,n+i}, \ i = 1, \dots, n-1, \ \underline{e}_n := e_{n-1,2n} - e_{n,2n-1}$$
$$\underline{f}_i := e_{i+1,i} - e_{n+i,n+i+1}, \ i = 1, \dots, n-1, \ \underline{f}_n := e_{2n,n-1} - e_{2n-1,n}$$

in terms of the spin formalism and formulas 4.4.1, 4.4.2 of Chapter 5:

$$[a \wedge b, c] = (b, c)a - (a, c)b,$$

$$[c \wedge d, a \wedge b] = (b, d)a \wedge c + (a, d)c \wedge b - (a, c)d \wedge b - (b, c)a \wedge d.$$

We deduce (recall we are using $1/2$ the standard form):

(7.1.2)
$$e_i \wedge e_j = e_i e_j = e_{i,n+j} - e_{j,n+i}, \quad e_i \wedge f_j = e_i f_j = e_{i,j} - e_{n+j,n+i}, \ i \neq j,$$

(7.1.3)
$$f_i \wedge f_j = f_i f_j = e_{n+i,j} - e_{n+j,i}, \quad e_i \wedge f_i = e_i f_i - 1/2 = e_{i,i} - e_{n+i,n+i}.$$

Hence

$$\underline{e}_i = e_i f_{i+1}, \qquad \underline{f}_i = e_{i+1} f_i, \qquad \underline{e}_n = e_{n-1} e_n, \qquad \underline{f}_n = f_n f_{n-1}.$$

We deduce that

$$e_1 \wedge e_2 \wedge \ldots \wedge e_n, \quad e_1 \wedge e_2 \wedge \ldots \wedge e_{n-1}$$

are highest weight vectors. In fact it is clear that they are both killed by $\underline{e}_n$ but also by the other $\underline{e}_i$, $i < n$, by a simple computation.

For the weights, apply

$$\sum_i^n x_i e_i \wedge f_i = \sum_i x_i (e_{i,i} - e_{n+i,n+i}) \text{ to } e_1 \wedge e_2 \wedge \ldots \wedge e_n \text{ and } e_1 \wedge e_2 \wedge \ldots \wedge e_{n-1}$$

to obtain the weights $1/2 \sum_i^n x_i$, $1/2(\sum_i^{n-1} x_i - x_n)$ respectively. We have the two half-spin weights $s_\pm$ determined in Chapter 10, §5.1.1.

In the odd case consider the Chevalley generators in the form

$$\underline{e}_i := e_{i,i+1} - e_{n+i+1,n+i} = e_i f_{i+1} = a_i b_{i+1}, \underline{e}_n = e_{n,2n+1}$$

(7.1.4)
$$- e_{2n+1,2n} = e_n u = a_n,$$

from which it follows that $a_1 \wedge a_2 \wedge \ldots \wedge a_n$ is a highest weight vector.

As for its weight, apply, as in the even case,

$$\sum_i^n x_i \frac{[a_i, b_i]}{2} = \sum_i^n x_i e_i \wedge f_i = \sum_i x_i (e_{i,i} - e_{n+i,n+i})$$

and obtain the weight $1/2 \sum_{i=1}^n x_i$, which is the fundamental spin weight for the odd orthogonal group.

7.2 Pure Spinors

It may be slightly more convenient to work with the lowest weight vector 1 killed by $e_i \wedge f_j, \forall i \neq j,\ f_i \wedge f_j$, with $(e_1 \wedge f_1)1 = -1/2$.

Definition. A vector in one of the two half-spin representations is called a *pure spinor* if it is in the orbit, under the spin group, of the line through the highest weight vector.

This is a special case of the theory discussed in Chapter 10, §6.2. If G is the spin group and P is the parabolic subgroup fixing the line through 1, in projective space a pure spinor corresponds to a point of the variety G/P.

We now do our computations in the exterior algebra on the generators e_i, which is the sum of the two half-spin representations.

Theorem 1. *The two varieties G/P for the two half-spin representations correspond to the two varieties of maximal totally isotropic subspaces. If a is a nonzero pure spinor, the maximal isotropic subspace to which it corresponds is $N_a := \{x \in W \mid xa = 0\}$.*

Proof. Let us look for instance at the even half-spin representation. Here 1 is a lowest weight vector and $x1 = 0$ if and only if x is in the span of $f_1, f_2, \dots, f_n$, so that we have $N_1 = \langle f_1, f_2, \dots, f_n \rangle$. If a is any nonzero pure spinor, $a = g\lambda 1$, $g \in \mathrm{Spin}(2n)$, $\lambda \in \mathbb{C}$, and hence the set of $x \in W$ with $xa = 0$ is $g(N_1)$; the claim follows. □

The parabolic subalgebra $\mathfrak{p}$ stabilizing 1 is the span of the elements $f_i \wedge f_j$, $e_i \wedge f_j$, which in matrix notation, from 7.1.2 and 7.1.3 coincides with the set of block matrices $\begin{vmatrix} A & 0 \\ B & -A^t \end{vmatrix}$, $B^t = -B$. It follows that the parabolic subgroup fixing 1 coincides with the stabilizer of the maximal isotropic space N_1.

Thus the unipotent group, opposite to $\mathfrak{p}$, has its Lie algebra spanned by the elements $e_i \wedge e_j$. Applying the exponential of an element $\omega_X := \sum_{i<j} x_{i,j} e_i \wedge e_j$ to 1, one obtains the open cell in the corresponding variety of pure spinors. This is described by formula 3.6.3 of Chapter 5, involving the Pfaffians $[i_1, i_2, \dots, i_{2k}]$:

$$(7.2.1) \qquad \exp(\omega_X) = \sum_k \sum_{i_1<i_2<\dots<i_{2k}} [i_1, i_2, \dots, i_{2k}] e_{i_1} \wedge e_{i_2} \wedge \dots \wedge e_{i_{2k-1}} \wedge e_{i_{2k}}.$$

Remark that exponentiating the toral subalgebra $\sum_i^n x_i e_i \wedge f_i$ we have a torus, whose coordinates we may call t_i such that the vector $e_{i_1} \wedge e_{i_2} \wedge \dots \wedge e_{i_k}$ has weight $t_{i_1} t_{i_2} \dots t_{i_k}$.

To obtain all the pure spinors we may apply the Weyl group. To understand it take one of the $sl_i(2, \mathbb{C})$, $i < n$, generated by $\underline{e}_i = e_i f_{i+1}$, $\underline{f}_i = e_{i+1} f_i$, $\underline{h}_i = e_i f_i - e_{i+1} f_{i+1}$.

We see by direct computation that all the vectors are killed by this Lie algebra except for the pairs $e_i \wedge u$, $e_{i+1} \wedge u$, with u a product not containing e_i, e_{i+1}. These pair of vectors transform as the standard basis of the basic representation of $sl(2, \mathbb{C})$. It follows that the reflection s_i maps $e_i \wedge u$ to $-e_{i+1} \wedge u$, and $e_{i+1} \wedge u$, to $e_i \wedge u$.

In the special linear group $SL(n, \mathbb{C})$ which acts on $\bigwedge(\sum_i \mathbb{C} e_i)$, the elements s_i defined by $s_i(e_j) = e_j$, $j \neq i, j \neq i+1$, $s_i(e_i) = -e_{i+1}$, $s_i(e_{i+1}) = e_i$ are the usual elements we choose to generate a group $\tilde{W}$ which induces, by adjoint action, the symmetric group S_n.

The last case is $sl_n(2,\mathbb{C})$ generated by $\underline{e}_n = e_{n-1}e_n$, $\underline{f}_n = f_n f_{n-1}$, $\underline{h}_n = e_{n-1}f_{n-1} - f_n e_n$. Now the vectors killed by $\underline{e}_n$ are all the vectors that contain either e_{n-1} or e_n or both.

We are left with two types of nontrivial representations: $e_{n-1} \wedge u, e_n \wedge u$ or $u, e_{n-1} \wedge e_n \wedge u$. The simple reflection s_n sends (u does not contain n):

$$e_{n-1} \wedge u \mapsto -e_n \wedge u,\ e_n \wedge u \mapsto e_{n-1} \wedge u, u \mapsto -e_{n-1} \wedge e_n \wedge u,\ e_{n-1} \wedge e_n \wedge u \mapsto u.$$

At this point we can easily check:

Proposition.

(i) *The exterior products of the elements e_i are extremal vectors (for the two representations).*
(ii) *The stabilizer in the Weyl group of the weight of 1 is S_n and has index 2^{n-1}.*
(iii) *The half-spin representations are minuscule.*

Proof. Let us denote by $\tilde{W}_n$ the group generated by the elements s_i, $i = 1, \dots, n$. It contains the subgroup W generated by the s_i, $i = 1, \dots, n-1$, which acts on weights as the symmetric group. Let us compute the orbit of 1 under $\tilde{W}$. We claim that it is formed by all the vectors $\pm e_{i_1} \wedge e_{i_2} \wedge \dots \wedge e_{i+2k}$. In fact, this set of vectors is stable under the action of $\tilde{W}$. The elements s_i^2, $i = 1, \dots, n-1$, change the sign of the elements e_i, e_{i+1}, while s_n^2 changes the sign of e_{n-1}, e_n. Thus the subgroup generated by the s_i^2 is the group of diagonal transformations which changes an even number of signs. The group W permutes transitively the vectors $\pm e_{i_1} \wedge e_{i_2} \wedge \dots \wedge e_{i+2k}$ for a given k up to sign. Finally, s_n allows us to pass from a suitable product with k factors to one with $k-2$ or $k+2$ factors. This shows that the 2^{n-1} vectors $e_{i_1} \wedge e_{i_2} \wedge \dots \wedge e_{i+2k}$ are extremal.

For the other half-spin representation we may start from e_1 instead of 1. Thus (i) and (iii) follow. Also (ii) follows from the fact that the Weyl group of Spin$(2n)$ is a quotient of $\tilde{W}$ and has order $n!2^{n-1}$. □

It is useful to do some computations directly inside the Clifford algebra. We consider the even case and the usual basis e_i, f_i. Set $f := f_n f_{n-1} \cdots f_1$, $e := e_1 e_2 \dots e_n$. Set furthermore $A = \bigoplus_{k=0}^n A_k$ to be the (graded) subalgebra generated by the e_i's, isomorphic to the exterior algebra $\bigwedge V$, $V = \langle e_1, \dots, e_n \rangle$.

Proposition 1. *The left ideal in $C(2n)$ generated by f is minimal and equals Af. The mapping $a \mapsto af$, $a \in \bigwedge V$ is an isomorphism of representations of $C(2n)$.*

Proof. Since $f_i f = 0$, for all i we see that $C(2n)f = Af$. Since A is irreducible and the map $a \mapsto af$ nonzero, we must have that A maps isomorphically to Af. □

Notice in particular that $fe_{i_1}e_{i_2}\dots e_{i_k}f = 0$ unless $k = n$ and $fef = fe_1e_2\dots e_n f = f$. This allows us to define a nondegenerate bilinear form on $\bigwedge V = Af$ invariant under the spin group. We use the principal involution which is the identity on the generators e_i, f_i and consider, given $a, b \in A$, the element $(af)^*bf = f^*a^*bf$. We can assume both elements a and b are homogeneous. Since

$f^* = (-1)^{n(n-1)/2} f$ we deduce that $(af)^* bf = 0$ unless a^*b has degree n, hence it is a multiple, $\beta(a, b)e$ of e. Then $(af)^* bf = \beta(a, b) f$. The number $\beta(a, b)$ is a bilinear form on $\bigwedge V$.

Theorem 2.

(1) The form β is nondegenerate. It pairs in a nondegenerate way $\bigwedge^k V$ and $\bigwedge^{n-k} V$.
(2) β satisfies the symmetry condition $\beta(a, b) = (-1)^{n(n-1)/2} \beta(b, a)$.
(3) β is stable under the spin group.
(4) The principal involution is the adjunction with respect to the form β.

Proof. (1) On the *decomposable* elements $e_{i_1} e_{i_2} \dots e_{i_k}$ it is clear that β pairs, with some sign, elements of complementary indices, proving 1.

(2) If a^*b has degree n, then $b^*a = (a^*b)^* = (-1)^{n(n-1)/2} a^*b$.

(3) If $s \in \mathrm{Spin}(2n)$, we have $\beta(sa, sb)$ is computed by

$$(saf)^* sbf = f^* a^* s^* sbf = f^* a^* bf.$$

(4) If $c \in C(2n)$ we see that $\beta(ca, b) f = (caf)^* bf = \beta(a, c^*b) f$. So the principal involution is the adjunction with respect to the form β. $\square$

There are several interesting consequences of the previous theorem.

Corollary.

(1) Under the form β, the even and odd spinors are maximal totally isotropic if n is odd and instead orthogonal to each other if n is even.
(2) The two half-spin representations are self-dual if n is even, and each is the dual of the other if n is odd.
(3) If $n = 2k$ the form β is a nondegenerate form on each of the two half-spin representations. It is symmetric if k is even, skew symmetric if k is odd.

Proof. If a, b are homogeneous, and a^*b has degree n, we must have that a, b have different parity if n is odd, and the same if n is even, proving (1). (2) follows from (1). (3) follows from (2) and part (2) of Theorem 2. $\square$

Remark. We could have worked also with $B = \bigoplus_{k=0}^n B_k$, the algebra generated by the f_i, and the minimal left ideal Be.

Now that we have seen that the spin representation $\bigwedge V = Af$ is self-dual we can identify $\bigwedge V \otimes \bigwedge V$ with $\mathrm{End}(\bigwedge V) = C(2n)$, thought of as a representation of the spin group under conjugation. This is best done using the internal map:

Theorem. *The map $i : Af \otimes Be \to C(2n)$, $af \otimes be \mapsto af(be)^*$, is an isomorphism as representations of the spin group.*

Proof. Since Af, Be are minimal left ideals, $(Be)^* = eB$ is a minimal right ideal. $Af(Be)^* = AfeB$, a two-sided ideal. Since $C(2n)$ is a simple algebra and the two spaces have the same dimension, i is an isomorphism. If $s \in \mathrm{Spin}(2n)$, we have $s(ae(bf)^*)s^{-1} = s(ae(bf)^*)s^* = (sae)(sbf)^*$. $\square$

Remark. Notice that $Be = Afe$ so we can use the isomorphism i in the form $a_1 \otimes a_2 \mapsto a_1 fef a_2^* = a_1 f a_2^*$.

For a set $J := \{i_1 < i_2 \dots < i_k\}$, set $e_J := e_{i_1} e_{i_2} \dots e_{i_k}$, $f_J := e_{i_k} e_{i_{k-1}} \dots e_{i_1}$.

Lemma 1. *We have*

$$fe = \sum_{J \subset [1,2,\dots,n]} (-1)^{|J|} e_J f_J.$$

Proof. By induction

$$fe = f_n \Bigg(\sum_{J \subset [1,2,\dots,n-1]} (-1)^{|J|} e_J f_J \Bigg) e_n = \sum_{J \subset [1,2,\dots,n-1]} (-1)^{|J|} e_J f_n e_n f_J$$

$$= \sum_{J \subset [1,2,\dots,n-1]} (-1)^{|J|} e_J (-e_n f_n + 1) f_J = \sum_{J \subset [1,2,\dots,n]} (-1)^{|J|} e_J f_J. \qquad \square$$

We use two bases for $C(2n)$, $e_J f_K, e_J f e_K$, $J, K \subset [1, 2, \dots, n]$. The first basis $e_J f_K$ is adapted to the filtration by the degree of the monomials in the e_i, f_j.

Lemma 2. *The spaces $C_i := \bigoplus_{|J|+|K| \le i} \mathbb{C} e_J f_K$ are representations under the conjugation action of the spin group. C_i / C_{i-1} is isomorphic to $\bigwedge^i (V \oplus V^*)$.*

Proof. Since the spin group conjugates the space spanned by all the e_i, f_i, $i = 1, \dots, n$, into itself inducing the standard representation of the orthogonal group, the lemma follows by the commutation relations. $\square$

The second basis $e_J f e_K$ is adapted to the tensor product decomposition.

Under the tensor product the two summands $S_0 \otimes S_0 \oplus S_1 \otimes S_1$, $S_0 \otimes S_1 \oplus S_1 \otimes S_0$, correspond to $Cl^+(2n)$, $Cl^-(2n)$ when n is even and the other way if n is odd.

It is useful to describe the base change explicitly. Let J^c define the complement of J in $[1, 2, \dots, n]$ and ϵ_J the sign of the permutation such that $f = \epsilon_J f_J f_{J^c}$. Thus:

$$e_J f e_K = \epsilon_K e_J f_K f_{K^c} e_K = (-1)^{n(n-|K|)} \epsilon_K e_J f_K e_K f_{K^c}$$

$$= (-1)^{n(n-|K|)} \epsilon_K e_J \Bigg(\sum_{A \subset K} (-1)^{|A|} e_A f_A \Bigg) f_{K^c}$$

$$(7.2.2) \qquad = (-1)^{n(n-|K|)} \epsilon_K \sum_{A \subset K \cap J^c} (-1)^{|A|} e_J e_A f_A f_{K^c}.$$

Notice that the leading term of this sum is

$$(-1)^{n(n-|K|)} \epsilon_K (-1)^{|K \cap J^c|} e_J e_{K \cap J^c} f_{K \cap J^c} f_{K^c}.$$

Proposition 2. *The conjugation action of the spin group on $Cl(W)$ factors through an action of the special orthogonal group and it is isomorphic to $\bigoplus_{k=0}^{\dim W} \bigwedge^k W$.*

Proof. We have a homomorphism $T(W) \to Cl(W)$ which, restricted to the direct sum $\bigoplus_k \bigwedge^k(W)$ of antisymmetric tensors, is a linear isomorphism of representations of the orthogonal group. □

If $\dim W = 2n, 2n+1$, the even Clifford algebra $Cl^+(W)$ decomposes as $\bigoplus_{k=0}^n \bigwedge^{2k} W$. If $v_1, v_2, \dots, v_k \in W$, we have $v_1 \wedge v_2 \wedge \dots \wedge v_k = \frac{1}{k!}\sum_{\sigma \in S_k} \epsilon_\sigma v_{\sigma(1)} v_{\sigma(2)} \cdots v_{\sigma(k)}$.

Lemma 3. *If $v_1, v_2, \dots, v_k \in W$ are orthogonal, we have $v_1 \wedge v_2 \wedge \dots \wedge v_k = v_1 v_2 \dots v_k$.*

Proof. From the commutation relations $v_{\sigma(1)} v_{\sigma(2)} \dots v_{\sigma(k)} = \epsilon_\sigma v_1 v_2 \dots v_k$. □

If the elements are not orthogonal, this is only the leading term in the filtration.

In particular if $v_1, v_2, \dots, v_{2n}$ form an orthogonal basis of W the element $v_1 v_2 \dots v_{2n}$ is a generator of the 1-dimensional vector space $\bigwedge^{2n} W$ and is, under the conjugation action, invariant under $SO(W)$.

Choose as basis $u_{2i-1} := e_i + f_i$, $u_{2i} := e_i - f_i$, $i = 1, \dots, n$. These vectors are an orthogonal basis of W, and so anticommute. Moreover $u_{2i-1}^2 = 1$, $u_{2i}^2 = -1$.

The element $z := u_1 u_2 \dots u_{2n}$ anticommutes with each element of W, commutes with the even Clifford algebra, and generates $\bigwedge^{2n} W$.

Theorem 3. $z^2 = 1$, $fz = f$ *and* $\bigwedge^k Wz = \bigwedge^{2n-k} W$.

Proof. The permutation $2n, 2n-1, \dots, 2, 1$ has sign $(-1)^n$, so

$$z^2 = (-1)^n u_1 u_2 \dots u_{2n} u_{2n} u_{2n-1} \dots u_2 u_1 = (-1)^n (-1)^n = 1.$$

For fz let us see the first multiplications. From $f_i(e_i + f_i)(e_i - f_i) = -f_i$ we have

$$\begin{aligned} fz &= (-1)^n f u_{2n} u_{2n-1} \dots u_2 u_1 = (-1)^{n-1} f_1 \dots f_{n-1} f_n u_{2n-2} u_{2n-3} \dots u_2 u_1 \\ &= (-1)^{n-1} f_1 \dots f_{n-1} u_{2n-2} u_{2n-3} \dots u_2 u_1 f_n. \end{aligned}$$

Then work by induction. Since z is an invariant, the map $u \mapsto uz$ is an $SO(W)$-equivariant map, so it maps $\bigwedge^k W$, which is irreducible if $k < n$, to an isomorphic irreducible representation. Thus to prove that $z \bigwedge^k W = \bigwedge^{2n-k} W$, it is enough to see that there is an element $v \in \bigwedge^k W$ with $vz \in \bigwedge^{2n-k} W$. We can take $v = u_1 u_2 \dots u_k$ for which $vz = (-1)^k u_{k+1} u_{k+2} \dots u_{2n}$. We know, by §6.6, that $\bigwedge^n W$ decomposes as the direct sum of two irreducible representations of highest weights twice the two half-spin representations which are not isomorphic to the representations $\bigwedge^k W$ for $k < n$. It follows that z induces on $\bigwedge^n W$ a linear map with eigenvalues $+1$ and -1, preserving these two irreducibles. □

We will see presently that the two summands $\bigwedge^n W_+$, $\bigwedge^n W_-$ of $\bigwedge^n W$ relative to the eigenvalues $+1, -1$ coincide with the irreducibles of weights $2s_+, 2s_-$ respectively.

For $k < n$ we also decompose the direct sum $\bigwedge^k W \oplus \bigwedge^{2n-k} W = L_k^+ \oplus L_k^-$ as a sum of the two eigenspaces of eigenvalues ± 1 for z. Notice that, as representations, both L_k^+, L_k^- are isomorphic to $\bigwedge^k W \equiv \bigwedge^{2n-k} W$.

Notice that we also have $\bigwedge^k W \oplus \bigwedge^{2n-k} W = L_k^+ \oplus \bigwedge^k W = L_k^- \oplus \bigwedge^k W$.

We can now analyze $S_0 \otimes S_0$, $S_1 \otimes S_1$ by analyzing their images in $C(2n)$.

Take an element $afb^*, a, b \in S$. If $b \in S_0$, we have $afb^*z = afzb^* = afb^*$. If $b \in S_1$ is odd, we have $afb^*z = -afzb^* = -afb^*$. It follows that for each k we do have $\bigwedge^k W \not\subset SfS_0$ and $\bigwedge^k W \not\subset SfS_1$. Otherwise multiplication by z on $\bigwedge^k W$ would equal ± 1, while $\bigwedge^k Wz = \bigwedge^{2n-k} W$.

Theorem 4.

(i) The image of $S_0 \otimes S_0$ *in* $C(2n)$ *is* $\bigwedge^n W_+ \bigoplus_k L_k^+$, $k < n$ *and* $k \equiv n$, mod 2.
(ii) $\bigwedge^n W_+$ *is the irreducible representation with highest weight* $2s_+$.
(iii) The image of $S_1 \otimes S_1$ *in* $C(2n)$ *is* $\bigwedge^n W_- \bigoplus_k L_k^-$, $k < n$ *and* $k \equiv n$, mod 2.
(iv) $\bigwedge^n W_-$ *is the irreducible representation with highest weight* $2s_-$.

Proof. Under the equivariant map $i : S \otimes S \to C(2n)$, $i(a \otimes b) = afb^*$ we have that $S_0 \otimes S_0 \oplus S_1 \otimes S_1$ maps to the even or odd part of the Clifford algebra, depending on whether n is even or odd.

Conversely, $S_0 \otimes S_1 \oplus S_1 \otimes S_0$ maps to the even or odd part of the Clifford algebra, depending on whether n is odd or even.

In both cases, the odd or even part of the image belonging to $S \otimes S_0$ is the set of elements u with $uz = u$, and the image belonging to $S \otimes S_1$ is the set of elements u with $uz = -u$. The first claim follows now from the definitions. Next we know that $S_0 \otimes S_0$ contains (as leading term) the irreducible representation of highest weight $2s_+$ which appears in $\bigwedge^n W$ so it must coincide with $\bigwedge^n W_+$. A similar argument on $S_1 \otimes S_1$ shows that $\bigwedge^n W_-$ must be the irreducible representation of highest weight $2s_-$ which appears in $\bigwedge^n W$. □

We have that the image of $S_0 \otimes S_0$, $S_1 \otimes S_1$ is in the even or odd Clifford algebra according to whether n is even or odd. Let us assume n is even to simplify the notation; the other case is similar. Let us denote by C_{n-1}^+ the part of $C(2n)^+$ of filtration degree $\leq n-1$. Then:

Proposition 3. *We have a direct sum decomposition*

$$C(2n)^+ = S_0 \otimes S_0 \oplus C_{n-1}^+ \oplus \bigwedge^n W_-$$

and also

$$C(2n)^+ = S_1 \otimes S_1 \oplus C_{n-1}^+ \oplus \bigwedge^n W_+.$$

Proof. Let us prove one of the two statements, the other being similar.

We have $S_0 \otimes S_0 = \bigwedge^n W_+ \bigoplus_{2i<n} L_{2i}^+$.

$$C_{n-1}^{+} = \bigoplus_{2i<n} \bigwedge^{2i} W,$$
$$C(2n)^{+} = \bigoplus_{i\leq n} \bigwedge^{2i} W = \bigwedge^{n} W \bigoplus_{2i<n} \Big(\bigwedge^{2i} W \oplus \bigwedge^{2n-2i} W\Big)$$
$$= \bigwedge^{n} W_{+} \oplus \bigwedge^{n} W_{-} \bigoplus_{2i<n} \Big(\bigwedge^{2i} W \oplus L_{2i}^{+}\Big).$$

The claim follows. □

From this corollary we have a more explicit way of computing the quadratic equations defining pure spinors. We do it for S_0; for S_1 it is similar. Given $u = \sum_J x_J e_J \in S_0$ we know, by Chapter 10, §6.6, that the property of being a pure spinor is equivalent to the fact that $u \otimes u \in \bigwedge^n W_+$. This means, using the previous proposition, that when we project ufu^* to $C(2n)/C_n$ the image must be 0. This is quite an explicit condition since projecting to $C(2n)/C_n$ is the same as imposing the condition that in the expansion of ufu^*, all terms of degree $> n$ must vanish. Now

$$ufu^* = \Big(\sum_J x_J e_J\Big) f \Big(\sum_J (-1)^{|J|} x_J e_J\Big).$$

This sum can be computed using formulas 7.2.3. Notice that when applied to a pure spinor given by formula 7.2.1, one obtains some identical quadratic equations satisfied by Pfaffians. The theme of quadratic equations will be revised in Chapter 13 when we use such equations to develop standard monomial theory for tableaux.

Remarks.

(1) All the representations of $\mathrm{Spin}(V)$ which do not factor through $SO(V)$ appear in the spaces $S \otimes V^{\otimes p}$.
(2) In order to study intertwiners between these representations one must use the structure of $\mathrm{End}(S)$.

7.3 Triality

There is one special case to be noticed: $\mathrm{Spin}(8)$. In this case we have three fundamental representations of dimension 8. One is the defining representation, the other two are the half-spin representations. Since in each of the half-spin representations we have a nondegenerate quadratic form β preserved by the spin group, and since the Lie algebra of the spin group is simple, we get:

Proposition. *In each of the 3 fundamental 8-dimensional representations* $\mathrm{Spin}(8)$ *induces the full special orthogonal group.*

There is a simple explanation of this phenomenon which is called *triality* and it is specific to dimension 8. This comes from the external automorphism group of the spin group or its associated Lie algebra. We have that the symmetric group S_3 of permutations of 3 elements acts as symmetry group of the Dynkin diagram. Thus (Chapter 10, §6.10) it induces a group of external automorphisms which permutes transitively the 3 outer vertices, which correspond to the 3 fundamental representations described.

Notice the simple:

Corollary. *The set of pure spinors in a half-spin representation equals the set of isotropic vectors for the form β.*

8 Invariants of Matrices

In this section we will deduce the invariant theory of matrices from our previous work.

8.1 FFT for Matrices

We are interested now in the following problem: describe the ring of invariants of the action of the general linear group $GL(n, \mathbb{C})$ acting by simultaneous conjugation on m copies of the space $M_n(\mathbb{C})$ of square matrices.

In intrinsic language, we have an n-dimensional vector space V and $GL(V)$ acts on $\mathrm{End}(V)^{\oplus m}$.

We will denote by $(X_1, X_2, \dots, X_m)$ an m-tuple of $n \times n$ matrices.

Before we formulate and prove the main theorem let us recall the results of Chapter 9, §1. The theorem proved there can be reformulated as follows. Suppose we are interested in the multilinear invariants of m matrices, i.e., the invariant elements of the dual of $\mathrm{End}(V)^{\otimes m}$.

First, remark that the dual of $\mathrm{End}(V)^{\otimes m}$ can be identified, in a $GL(V)$-equivariant way, with $\mathrm{End}(V)^{\otimes m}$ by the pairing formula:

$$\begin{aligned}&\langle A_1 \otimes A_2 \dots \otimes A_m | B_1 \otimes B_2 \dots \otimes B_m\rangle \\ &:= \mathrm{tr}(A_1 \otimes A_2 \dots \otimes A_m \circ B_1 \otimes B_2 \dots \otimes B_m) = \prod \mathrm{tr}(A_i B_i).\end{aligned}$$

Therefore under this isomorphism the multilinear invariants of matrices are identified with the $GL(V)$-invariants of $\mathrm{End}(V)^{\otimes m}$ which in turn are spanned by the elements of the symmetric group. We deduce from the theory of Chapter 9 and formula 6.1.3 there:

Lemma. *The multilinear invariants of m matrices are linearly spanned by the functions:*

$$\phi_\sigma(X_1, X_2, \dots, X_m) := \mathrm{tr}(\sigma^{-1} \circ X_1 \otimes X_2 \otimes \dots \otimes X_m), \ \sigma \in S_m.$$

Recall that if $\sigma = (i_1 i_2 \dots i_h)(j_1 j_2 \dots j_k) \dots (s_1 s_2 \dots s_m)$ is the cycle decomposition of σ, then from Chapter 9, Theorem 6.1.3, we have that

$$\phi_\sigma(X_1, X_2, \dots, X_m) = \mathrm{tr}(X_{i_1} X_{i_2} \dots X_{i_h}) \, \mathrm{tr}(X_{j_1} X_{j_2} \dots X_{j_k}) \dots \mathrm{tr}(X_{s_1} X_{s_2} \dots X_{s_m}).$$

This explains our convention in defining ϕ_σ.

Theorem (FFT for Matrices). *The ring of invariants of matrices under simultaneous conjugation is generated by the elements*

$$\operatorname{tr}(X_{i_1}X_{i_2}\dots X_{i_{k-1}}X_{i_k}),$$

where the formula means that we take all possible noncommutative monomials in the X_i and form their traces.

Proof. The proof of this theorem is an immediate consequence of the previous lemma and the Aronhold method.

The proposed ring of invariants is in fact clearly closed under polarizations and coincides with the invariants, by the previous lemma, for the multilinear elements. The claim follows. □

Exercise. It is easy to generalize this theorem when one considers as a representation $\operatorname{End}(V)^{\otimes p} \oplus V^q \oplus (V^*)^s$.

One gets as generators of invariants besides the elements $\operatorname{tr}(M)$ also the elements $\langle \phi \mid Mv\rangle = \langle M^t\phi \mid v\rangle$, $\phi \in V^*$, $v \in V$ and M a monomial.

One can compute also the $SL(V)$-invariants and then add the invariants

$$M_1^t\phi_1 \wedge \dots \wedge M_n^t\phi_n, \quad M_1v_1 \wedge \dots \wedge M_nv_n, \quad \phi_i \in V^*, v_i \in V$$

and M a monomial.

8.2 FFT for Matrices with Involution

There is a similar theorem for the orthogonal and symplectic groups.

Assume that V is equipped with a nondegenerate form $\langle u, v\rangle$ (symmetric or skew symmetric). Then we can identify $\operatorname{End}(V) = V \otimes V^* = V \otimes V$ by

$$u \otimes v(w) := \langle v, w\rangle u\,.$$

Let $\epsilon = \pm 1$ according to the symmetry, i.e., $\langle a, b\rangle = \epsilon\langle b, a\rangle$. We then get the formulas

$$(a \otimes b) \circ (c \otimes d) = a \otimes \langle b, c\rangle d, \operatorname{tr}(a \otimes b) = \langle b, a\rangle$$

$$\langle (a \otimes b)c, d\rangle = \langle b, c\rangle\langle a, d\rangle = \epsilon\langle c, (b \otimes a)d\rangle.$$

In particular, using the notion of adjoint X^* of an operator, $\langle Xa, b\rangle := \langle a, X^*b\rangle$ we see that $(a \otimes b)^* = \epsilon(b \otimes a)$.

We can now analyze first the multilinear invariants of m matrices under the group G (orthogonal or symplectic) fixing the given form, recalling the FFT of invariant theory for such groups.

Compute such an invariant function on $\operatorname{End}(V)^{\otimes m} = V^{\otimes 2m}$. Write a decomposable element in this space as

$$X_1 \otimes X_2 \dots \otimes X_m = u_1 \otimes v_1 \otimes u_2 \otimes v_2 \otimes \dots u_m \otimes v_m, \quad X_i = u_i \otimes v_i.$$

Then the invariants are spanned by products of m scalar products $\langle x_i, y_i\rangle$ such that the $2m$ elements x_i, y_i exhaust the list of the u_i, v_j.

Of course in these scalar products a vector u can be paired with another u or a v (homosexual or heterosexual pairings, according to Weyl).

The previous formulas show that, up to a sign, such an invariant can be expressed in the form

$$\phi_{\overline{\sigma}}(X_1, \dots, X_m) := \mathrm{tr}(Y_{i_1} Y_{i_2} \dots Y_{i_h}) \,\mathrm{tr}(Y_{j_1} Y_{j_2} \dots Y_{j_k}) \dots \mathrm{tr}(Y_{s_1} Y_{s_2} \dots Y_{s_m}),$$

where, for each i, the element Y_i is either X_i or X_i^*.

Combinatorially this can be pictured by a *marked permutation* $\overline{\sigma}$, e.g.,

$$\overline{\sigma} = \{3, 2, \overline{1}, \overline{5}, 4\}, \quad \phi_{\overline{\sigma}}(X_1, \dots, X_5) := \mathrm{tr}(X_3 X_1^* X_2)\,\mathrm{tr}(X_4 X_5^*).$$

We deduce then the:

Theorem (FFT for Matrices). *The ring of invariants of matrices under simultaneous conjugation by the group G (orthogonal or symplectic) is generated by the elements*

$$\mathrm{tr}(Y_{i_1} Y_{i_2} \dots Y_{i_{k-1}} Y_{i_k}), \ Y_{i_h} = X_{i_h} \ or \ X_{i_h}^*.$$

Proof. Same as before. □

Exercise. Generalize this theorem when one considers as a representation $\mathrm{End}(V)^{\otimes p} \oplus V^q$.

One gets as generators of invariants, besides the elements $\mathrm{tr}(M)$, also the elements $(u\,,\ Mv) = (M^*u\,,\ v)$, $v \in V$ and M a monomial in the X_i, X_j^*.

The computation of the $SO(V)$-invariants is more complicated but can also be performed. In this case one should observe that the only new case is when $\dim V$ is even. In fact, when $\dim V$ is odd, $-1 \in SO(V)$ is an improper orthogonal transformation which acts as 1 on matrices. Thus, in this case we have no further invariants when we restrict to $SO(V)$.

Instead, when $\dim(V) = 2n$, given a skew-symmetric matrix Y we have that $Pf(Y)$ is invariant under $SO(V)$, but for $A \in O(V)$ we have $Pf(AYA^{-1}) = Pf(AYA^t) = \det(A)Pf(Y)$, from Chapter 5, §3.6.2.

Of course we may think of $Pf\left(\frac{X-X^t}{2}\right)$ as an invariant of matrices of degree n.

When we polarize it we obtain a symmetric multilinear invariant $Q(X_1, \dots, X_n)$ which under an improper orthogonal transformation is multiplied by -1. When we specialize the matrices $X_i := u_i \otimes v_i$ we get a special orthogonal, but not an orthogonal invariant of the $2n$ vectors u_i, v_j. We know that, up to a scalar, this invariant must be equal to $u_1 \wedge v_1 \wedge u_2 \wedge v_2 \wedge \dots \wedge u_n \wedge v_n$. If we set $u_i := e_i,\ v_i := f_i$ we have that the constant is 1. Now we have more general invariants, given by the functions $Q(M_1, M_2, \dots, M_n)$ where the M_i's are monomials in the X_i, X_i^t, and we see this, by the same symbolic method.

Theorem (FFT for Matrices). *The ring of invariants of matrices under simultaneous conjugation by the group $SO(V)$, $\dim V = 2n$ is generated by the elements*

$$\mathrm{tr}(M),\ Q(M_1, M_2, \dots, M_n),\ M,\ M_i \ \textit{are monomials in} \ X_i, X_i^t.$$

8.3 Algebras with Trace

It is useful to formalize the previous analysis, in the spirit of universal algebra, as follows.

Definition. An algebra with trace is an associative algebra R equipped with a unary (linear) map $\mathrm{tr} : R \to R$ satisfying the following axioms:

(1) $\mathrm{tr}(a)b = b\,\mathrm{tr}(a), \ \forall a, b \in R$.
(2) $\mathrm{tr}(\mathrm{tr}(a)b) = \mathrm{tr}(a)\,\mathrm{tr}(b), \ \forall a, b \in R$.
(3) $\mathrm{tr}(ab) = \mathrm{tr}(ba), \ \forall a, b \in R$.

An algebra with involution is an associative algebra R equipped with a unary (linear) map $* : R \to R, \ x \to x^*$ satisfying the following axioms:

$$(x^*)^* = x, \ (xy)^* = y^*x^*, \ \forall x, y \in R.$$

For an algebra with involution and a trace we shall assume the compatibility condition $\mathrm{tr}(x^*) = \mathrm{tr}(x), \forall x$.

As always happens in universal algebra, for these structures one can construct free algebras on a set S (or on the vector space spanned by S).

Explicitly, in the category of associative algebras, the free algebra on S is obtained by introducing variables x_s, $s \in S$ and considering the algebra having as basis all the *words* or monomials in the x_s. For algebras with involution we have to add also the *adjoints* x_s^* as an independent set of variables.

When we introduce a trace we have to add to these free algebras a set of commuting indeterminates $t(M)$ as M runs over the set of monomials.

Here, in order to preserve the axioms, we also need to impose the identifications given by *cyclic equivalence*: $t(AB) = t(BA)$, and *adjoint symmetry*: $t(A^*) = t(A)$, for all monomials A, B.

In all cases the free algebra with trace is the tensor product of the free algebra (without trace) and the polynomial ring in the elements $t(A)$. This polynomial ring will be called the *free trace ring* (with or without involution).

The free algebras F_S, by definition, have the universal property that given elements $f_s \in F_S$, $s \in S$, there is a unique homomorphism $F_S \to F_S$ (compatible with the structures of trace, involution) which maps x_s to f_s for all s.

In particular we can rescale independently the x_s and thus speak about multihomogeneous, in particular multilinear, elements in F_S.

Let $S = \{1, 2, \dots, m\}$. We describe the multilinear elements in the various cases.

1. For the free associative algebra in m variables, the multilinear monomials (in all the variables) correspond to permutations in S_m as $x_{i_1}x_{i_2}\dots x_{i_m}$.
2. For the free algebra with involution the multilinear monomials (in all the variables) correspond to marked permutations in S_m as: $y_{i_1}y_{i_2}\dots y_{i_m}$ where $y_i = x_i$ if i is unmarked, while $y_i = x_i^*$ if i is marked.

3. For the free associative algebra with trace in m variables the multilinear monomials (in all the variables) correspond to permutations in S_{m+1} according to the following rule. Take such a permutation and decompose it into cycles isolating the cycle containing $m+1$ as follows: if $\sigma = (i_1 i_2 \dots i_h)(j_1 j_2 \dots j_k)\dots(r_1 \dots r_p)$ $(s_1 s_2 \dots s_q m+1)$, set

$$\psi_\sigma(x_1, x_2, \dots, x_m) = \operatorname{tr}(x_{i_1} x_{i_2} \dots x_{i_h}) \operatorname{tr}(x_{j_1} x_{j_2} \dots x_{j_k}) \dots \\ \times \operatorname{tr}(x_{r_1} x_{r_2} \dots x_{r_p}) x_{s_1} x_{s_2} \dots x_{s_q}.$$

4. For the free associative algebra with trace and involution in m variables the multilinear elements (in all the variables) correspond to marked permutations in S_{m+1}, but there are some equivalences due to the symmetry of trace under involution.

In fact it may be interesting to isolate, in the case of trace algebras, the part T_{m+1} of the multilinear elements in $m+1$ variables lying in the trace ring and compare it with the full set A_m of multilinear elements in m variables using the map $c_m : A_m \to T_{m+1}$, given by $c_m(M) := t(Mx_{m+1})$. We leave it to the reader to verify that this map is a linear isomorphism.

Example. The correspondence between A_2 and T_3, in the case of involutions:

$$x_1x_2, t(x_1x_2x_3);\ x_2x_1, t(x_2x_1x_3);\ x_1x_2, t(x_1x_2x_3);\ x_2x_1, t(x_2x_1x_3);$$

$$x_1^*x_2, t(x_1^*x_2x_3);\ x_2^*x_1, t(x_2^*x_1x_3);\ x_1x_2^*, t(x_1x_2^*x_3);\ x_2x_1^*, t(x_2x_1^*x_3);$$

$$t(x_1)x_2, t(x_1)t(x_2x_3);\ t(x_2)x_1, t(x_2)t(x_1x_3);\ t(x_1)x_2^*, t(x_1)t(x_2^*x_3);$$

$$t(x_2)x_1^*, t(x_2)t(x_1^*x_3);\ t(x_1)t(x_2), t(x_1)t(x_2)t(x_3);\ t(x_1x_2), t(x_1x_2)t(x_3);$$

$$t(x_1x_2^*), t(x_1x_2^*)t(x_3).$$

Let us also denote for convenience by R the free algebra with trace in infinitely many variables and by Tr the polynomial ring of traces in R. The formal trace in R is denoted by $t : R \to \mathrm{Tr}$.

We formalize these remarks as follows. We define maps

$$\Psi : \mathbb{C}[S_{m+1}] \to R, \quad \Psi(\sigma) := \psi_\sigma; \quad \Phi : \mathbb{C}[S_m] \to \mathrm{Tr}, \quad \Phi(\sigma) := \phi_\sigma.$$

We have a simple relation between these maps. If $\sigma \in S_{m+1}$, then

$$\Phi(\sigma) = t(\Psi(\sigma)x_{m+1}).$$

We remark finally that for a permutation $\sigma \in S_m$ and an element $a \in \mathbb{C}[S_m]$ we have

$$\Phi(\sigma a \sigma^{-1}) = \Phi(a)(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(m)}).$$

8.4 Generic Matrices

We shall now take advantage of this formalism and study more general algebras. We want to study the set of G-equivariant maps

$$f : \mathrm{End}(V)^{\oplus m} \to \mathrm{End}(V).$$

Here $n = \dim V$ and G is either $GL(V)$ or the orthogonal or symplectic group of a form.

We shall denote this space by $R_m(n)$ in the $GL(V)$ case, and by $R_m^o(n)$, $R_m^s(n)$, respectively, in the orthogonal and symplectic cases.

First, observe that the scalar multiples of the identity $\mathbb{C}1_V \subset \mathrm{End}(V)$ form the trivial representation, and the equivariant maps with values in $\mathbb{C}1_V$ can be canonically identified with the ring of invariants of matrices. We shall denote this ring by $T_m(n)$, $T_m^o(n)$, $T_m^s(n)$ in the 3 corresponding cases.

Next remark that $\mathrm{End}(V)$ has an algebra structure and a trace, both compatible with the group action. We deduce that under pointwise multiplication of the values the space $R_m(n)$ is a (possibly noncommutative) algebra, and moreover, applying the trace function, we deduce that:

$R_m(n)$ is an algebra with trace and the trace takes values in $T_m(n)$.

For the orthogonal and symplectic case $\mathrm{End}(V)$ has also a canonical involution, so $R_m^o(n)$ and $R_m^s(n)$ are algebras with trace and involution.

Finally, observe that the *coordinate maps* $(X_1, X_2, \dots, X_m) \to X_i$ are clearly equivariant. We shall denote them (by the usual abuse of notations) by X_i. We have:

Theorem. *In the case of $GL(V)$, $R_m(n)$ is generated as an algebra over $T_m(n)$ by the variables X_i.*

$R_m^o(n)$ and $R_m^s(n)$ are generated as algebras over $T_m^o(n)$, $T_m^s(n)$ by the variables X_i, X_i^*.

Proof. Let us give the proof of the first statement; the others are similar.

Given an equivariant map $f(X_1, \dots, X_m)$ in m variables, we construct the invariant function of $m+1$ variables $g(X_1, \dots, X_m, X_{m+1}) := \mathrm{tr}(f(X_1, \dots, X_m)X_{m+1})$.

By the structure theorem of invariants, g can be expressed as a linear combination of elements of the form $\mathrm{tr}(M_1)\,\mathrm{tr}(M_2)\dots\mathrm{tr}(M_k)$, where the M_i's are monomials in the variables X_i, $i = 1, \dots, m+1$.

By construction, g is linear in X_{m+1}; thus we can assume that each term

$$\mathrm{tr}(M_1)\,\mathrm{tr}(M_2)\dots\mathrm{tr}(M_k)$$

is linear in X_{m+1} and in particular (using the cyclic equivalence of trace) we may assume that X_{m+1} appears only in M_k and $M_k = N_k X_{m+1}$. Then N_k does not contain X_{m+1} and

$$\mathrm{tr}(M_1)\,\mathrm{tr}(M_2)\dots\mathrm{tr}(M_{k-1})\,\mathrm{tr}(M_k) = \mathrm{tr}((\mathrm{tr}(M_1)\,\mathrm{tr}(M_2)\dots\mathrm{tr}(M_{k-1})N_k)X_{m+1}).$$

It follows that we can construct a polynomial $h(X_1, \dots, X_m)$ (noncommutative) in the variables X_i, $i \le m$, with invariant coefficients and such that

$$\operatorname{tr}(h(X_1, \dots, X_m)X_{m+1}) = \operatorname{tr}(f(X_1, \dots, X_m)X_{m+1}).$$

Now we can use the fact that the trace is a nondegenerate bilinear form and that X_{m+1} is an independent variable to deduce that $h(X_1, \dots, X_m) = f(X_1, \dots, X_m)$, as desired. $\square$

Definition. The algebra $R_m(n)$ is called the *algebra of m-generic $n \times n$ matrices with trace.*

The algebra $R^o_m(n)$ (resp. $R^s_m(n)$) is called the *algebra of m-generic $n \times n$ matrices with trace and orthogonal* (resp. *symplectic*) *involution.*

8.5 Trace Identities

Our next task is to use the second fundamental theorem to understand the relations between the invariants that we have constructed. For this we will again use the language of universal algebra. We have to view the algebras constructed as quotients of the corresponding free algebras, and we have to deduce some information on the kernel of this map.

Recall that in an algebra R with some extra operations an ideal I must be stable under these operations, so that R/I can inherit the structure. In an algebra with trace or involution, we require that I be stable under the trace or the involution.

Let us call by F_m, F^i_m the free algebra with trace in m-variables and the free algebra with trace and involution in m-variables. We have the canonical maps (compatible with trace and, when it applies, with the involution)

$$\pi : F_m \to R_m(n), \ \pi^o : F^i_m \to R^o_m(n), \ \pi^s : F^i_m \to R^s_m(n).$$

We have already seen that in the free algebras we have the operation of substituting the variables x_s by any elements f_s. Then one has the following:

Definition. An ideal of a free algebra, stable under all the substitutions of the variables, is called a *T-ideal* (or an *ideal of polynomial identities*).

The reason for this notation is the following. Given an algebra R, a morphism of the free algebra in R consists of evaluating the variables x_s in some elements r_s. The intersection of all the kernels of all possible morphisms are those expressions of the free algebra which vanish identically when evaluated in R, and it is clear that they form a T-ideal, the ideal of polynomial identities of R. Conversely, if $I \subset F_S$ is a T-ideal, it is easily seen that it is the ideal of polynomial identities of F_S/I.

Of course an intersection of T-ideals is again a T-ideal and thus we can speak of the T-ideal generated by a set of elements (polynomial identities, or trace identities).

Going back to the algebra $R_m(n)$ (or $R^o_m(n)$, $R^s_m(n)$) we also see that we can compose any equivariant map $f(X_1, \dots, X_m)$ with any m maps $g_i(X_1, \dots, X_m)$ getting a new map $f(g_1(X_1, \dots, X_m), \dots, g_m(X_1, \dots, X_m))$. Also in $R_m(n)$, we have the morphisms given by substitutions of variables.

Clearly substitution in the free algebra is compatible with substitution in the algebras $R_m(n)$, $R_m^o(n)$, $R_m^s(n)$, and thus the kernels of the maps π, π^o, π^s are all T-ideals. They are called respectively:

The ideal of trace identities of matrices.

The ideal of trace identities of matrices with orthogonal involution.

The ideal of trace identities of matrices with symplectic involution.

We can apply the language of substitutions in the free algebra and thus define polarization and restitution operators, and we see immediately (working with infinitely many variables) that

Lemma. *Two T-ideals which contain the same multilinear elements coincide.*

Thus we can deduce that the kernels of π, π^o, π^s are generated as T-ideals by their multilinear elements. We are going to prove in fact that they are generated as T-ideals by some special identities. So let us first analyze the case of $R(n) := \cup_m R_m(n)$, $T(n) := \cup_m T_m(n)$.

The multilinear elements of degree m (in the first m variables) are contained in $R_m(n)$ and are of the form

$$\sum_{\sigma \in S_{m+1}} a_\sigma \psi_\sigma(X_1, X_2, \dots, X_m).$$

From the theory developed we know that the map

$$\Psi : \mathbb{C}[S_{m+1}] \to R_m(n), \quad \sum_{\sigma \in S_{m+1}} a_\sigma \sigma \mapsto \sum_{\sigma \in S_{m+1}} a_\sigma \psi_\sigma(X_1, X_2, \dots, X_m)$$

has as its kernel the ideal generated by the antisymmetrizer in $n+1$ elements. Thus, the multilinear identities appear only for $m \geq n$ and for $m = n$ there is a unique identity (up to scalars). So the first step consists of identifying the identity $A_n(x_1, \dots, x_n)$ corresponding to the antisymmetrizer

$$A_n(x_1, \dots, x_n) := \sum_{\sigma \in S_{n+1}} \epsilon_\sigma \psi_\sigma(x_1, x_2, \dots, x_n)$$

with ϵ_σ the sign of the permutation.

For this, recall that there is a canonical identity, homogeneous of degree n in 1 variable, called *the Cayley–Hamilton identity*. Consider the characteristic polynomial of X, $\chi_X(t) := \det(t-X) = t^n - \operatorname{tr}(X)t^{n-1} + \cdots + (-1)^n \det(X)$, we have $\chi_X(X) = 0$.

We want to interpret this as a trace identity. Remark that $\operatorname{tr}(X^i)$ is the i^{th} Newton function in the eigenvalues of X. Hence, by Chapter 2, §1.1.3 we can interpret each coefficient of the characteristic polynomial as a well-defined polynomial in the elements $\operatorname{tr}(X^i)$.

Thus we can consider $CH_n(x) := x^n - t(x)x^{n-1} + \cdots + (-1)^n \det(x)$ as a formal element of the free algebra. If we fully polarize this element we get a multilinear trace identity $CH(x_1, \dots, x_n)$ for $n \times n$ matrices, whose terms not containing traces arise from the polarization of x^n and are thus of the form $\sum_{\tau \in S_n} x_{\tau(1)} x_{\tau(2)} \dots x_{\tau(n)}$.

By the uniqueness of the identities in degree n we must have that the polarized Cayley–Hamilton identity is a multiple of the identity corresponding to the antisymmetrizer. To compute the scalar we may look at the terms not containing a trace in the two identities.

Clearly $x_1x_2\dots x_n = \psi_{(1\,2\dots n\,n+1)}$ and $\epsilon_{(1\,2\dots n\,n+1)} = (-1)^n$, and thus we have finally:

Proposition. $A_n(x_1,\dots,x_n) = (-1)^n CH(x_1,\dots,x_n)$.

Example. $n = 2$ (*polarize* $P_2(x)$)

$$A_2 = x_1x_2 + x_2x_1 - t(x_1)x_2 - t(x_1)x_2 - t(x_1x_2) + t(x_1)t(x_2);$$

$$P_2(x) = x^2 - t(x)x + \det(x) = x^2 - t(x)x + \frac{1}{2}(t(x)^2 - t(x^2)).$$

Exercise. Using this formal description (and restitution) write combinatorial expressions for the coefficients of the characteristic polynomial of X in terms of $\operatorname{tr}(X^i)$ (i.e., expressions for elementary symmetric functions in terms of Newton power sums).

It is particularly interesting to see what the polarized form is of the determinant. The terms corresponding to the determinant correspond exactly to the sum over all the permutations which fix $n+1$. Therefore we deduce:

Corollary. *The polarized form of* $\det(X)$ *is the expression*

$$\sum_{\sigma\in S_n} \epsilon_\sigma \phi_\sigma(x_1,\dots,x_n). \tag{8.5.1}$$

Let us look at some implications for relations among traces.

According to the general principle of correspondence between elements in R and T, we deduce the trace relation

$$\begin{aligned} T_{n+1}(x_1,\dots,x_n,x_{n+1}) &:= \sum_{\sigma\in S_{n+1}} \epsilon_\sigma\phi_\sigma(x_1,x_2,\dots,x_n,x_{n+1}) \\ &= t(A_n(x_1,\dots,x_n)x_{n+1}). \end{aligned}$$

Recall that

$$\phi_\sigma(x_1,x_2,\dots,x_m) = \operatorname{tr}(x_{i_1}x_{i_2}\dots x_{i_h})\operatorname{tr}(x_{j_1}x_{j_2}\dots x_{j_k})\dots\operatorname{tr}(x_{s_1}x_{s_2}\dots x_{s_m}).$$

Example. ($n = 2$)

$$\begin{aligned} T_3 := {} & t(x_1x_2x_3) + t(x_2x_1x_3) - t(x_1)t(x_2x_3) - t(x_1)t(x_2x_3) - t(x_1x_2)t(x_3) \\ & + t(x_1)t(x_2)t(x_3). \end{aligned}$$

Observe also that when we apply the operator of restitution to T_{n+1} we get $n!t(P_n(x)x) = T_{n+1}(x,x,\dots,x)$; the vanishing of this expression for $n\times n$ matrices is precisely the identity which expresses the $n+1$ power sum of n indeterminates as a polynomial in the lower power sums, e.g., $n = 2$:

$$t(x^3) = \frac{3}{2}t(x)t(x^2) - \frac{1}{2}t(x)^3.$$

Before we pass to the general case we should remark that any substitution map in R sends the trace ring Tr into itself. Thus it also makes sense to speak of a T-ideal in Tr. In particular we have the T-ideal of trace relations, the kernel of the evaluation map of Tr into $T(n)$. We wish to prove:

Theorem. *The T-ideal of trace relations is generated (as a T-ideal) by the trace relation T_{n+1}.*

The T-ideal of (trace) identities of $n \times n$ matrices is generated (as a T-ideal) by the Cayley–Hamilton identity.

Proof. From all the remarks made it is sufficient to prove that a multilinear trace relation (resp. identity) (in the first m variables) is in the T-ideal generated by T_{n+1} (resp. A_n).

Let us first look at trace relations. By the description of trace identities it is enough to look at the relations of the type $\Phi(\tau(\sum_{\sigma\in S_{n+1}} \epsilon_\sigma \sigma)\gamma)$, $\tau, \gamma \in S_{m+1}$.

We write such a relation as $\Phi(\gamma^{-1}(\gamma\tau(\sum_{\sigma\in S_{n+1}} \epsilon_\sigma \sigma))\gamma)$. We have seen that conjugation corresponds to permutation of variables, an operation allowed in the T-ideal; thus we can assume that $\gamma = 1$.

We start with the remark ($m \geq n$):

Splitting the cycles. Every permutation τ in S_{m+1} can be written as a product $\tau = \alpha \circ \beta$ where $\beta \in S_{n+1}$ and, in each cycle of α, there is at most 1 element in $1, 2, \dots, n+1$.

This is an exercise on permutations. It is based on the observation that

$$(xxxxx\ a\ yyyyy\ b\ zzzz\ c\ \dots) = (xxxxx\ a)(yyyyy\ b)(zzzz\ c)(\dots)(abc\dots).$$

From the remark it follows that (up to a sign) we can assume that τ has the property that in each cycle of τ there is at most 1 element in $1, 2, \dots, n+1$.

Assume that τ satisfies the previous property and let $\sigma \in S_{n+1}$. The cycle decomposition of the permutation $\tau\sigma$ is obtained by formally substituting in the cycles of σ, for every element $a \in [1, \dots, n+1]$, the cycle of τ containing a as a word (written formally with a at the end) and then adding all the cycles of τ not containing elements $\leq n+1$.

If we interpret this operation in terms of the corresponding trace element $\phi_\sigma(x_1, \dots, x_{m+1})$ we see that the resulting element is the product of a trace element corresponding to the cycles of τ not containing elements $\leq n+1$ and an element obtained by substituting monomials in $\phi_\sigma(x_1, \dots, x_{m+1})$ for the variables x_i (which one reads off from the cycle decomposition of τ).

As a result we have proved that a trace relation is in the T-ideal generated by T_{n+1}.

Let us pass now to trace identities. First we remark that, by the definition of an ideal in a trace ring, the relation T_{n+1} is a consequence of Cayley–Hamilton.

A trace polynomial $f(x_1, x_2, \dots, x_m)$ is a trace identity for matrices if and only if $t(f(x_1, x_2, \dots, x_m)x_{m+1})$ is a trace relation. Thus from the previous proof we have that $t(f(x_1, x_2, \dots, x_m)x_{m+1})$ is a linear combination of elements of type

$$A\,T_{n+1}(M_1, \dots, M_{n+1}) = A\,t(A_n(M_1, \dots, M_n)M_{n+1}).$$

Now we have to consider two cases: the variable x_{m+1} appears either in A or in one of the M_i. In the first case we have

$$A\,T_{n+1}(M_1, \dots, M_{n+1}) = t(T_{n+1}(M_1, \dots, M_{n+1})Bx_{m+1}),$$

and $T_{n+1}(M_1, \dots, M_{n+1})B$ is a consequence of Cayley–Hamilton. In the second, due to the antisymmetry of T_{n+1}, we can assume that x_{m+1} appears in $M_{n+1} = Bx_{m+1}C$. Hence

$$\begin{aligned} A\,t(A_n(M_1, \dots, M_n)M_{n+1}) &= A\,t(A_n(M_1, \dots, M_n)Bx_{m+1}C) \\ &= t(CAA_n(M_1, \dots, M_n)Bx_{m+1}). \end{aligned}$$

$CAA_n(M_1, \dots, M_n)B$ is also clearly a consequence of Cayley–Hamilton. □

8.6 Polynomial Identities

We now want to discuss polynomial identities of matrices. A polynomial identity is a special type of trace identity in which no traces appear. In fact these identities were studied before the trace identities, although they turn out to be harder to describe. Under the map $\Psi : \mathbb{C}[S_{m+1}] \to R(m)$ the elements that correspond to polynomials without traces are the full cycles $(i_1, i_2, \dots, i_m, m+1)$ and

$$\Psi((i_1, i_2, \dots, i_m, m+1)) = x_{i_1}x_{i_2}\dots x_{i_m}.$$

Thus:

Theorem. *The space $M_n(m)$ of multilinear polynomial identities of $n \times n$ matrices of degree m can be identified with $P_{m+1} \cap I_n$, where P_{m+1} is the subspace of $\mathbb{C}[S_{m+1}]$ spanned by the full cycles, and I_n is the ideal of $\mathbb{C}[S_{m+1}]$ generated by an antisymmetrizer on $n+1$ elements.*

A more precise description of this space is not known. Indeed we have only asymptotic information about its dimension [RA]. One simple remark is useful. Consider the automorphism τ of the group algebra defined on the group elements by $\tau(\sigma) := \epsilon_\sigma \sigma$ (Chapter 9, §2.5). τ is multiplication by $(-1)^m$ on P_{m+1} and transforms I_n into the ideal J_n of $\mathbb{C}[S_{m+1}]$ generated by a symmetrizer on $n+1$ elements. It follows that

$$M_n(m) = P_m \cap I_n \cap J_n\,.$$

The ideal I_n is the sum of all blocks corresponding to partitions λ with height $\geq n+1$, while J_n is the sum of all blocks corresponding to partitions λ with height of $\tilde{\lambda}$

$\geq n+1$. Thus $I_n \cap J_n$ is the sum of blocks relative to partitions which contain the hook $n+1, 1^n$. It follows that $M_n(m) = 0$ if $m < 2n$.[111] As for degree $2n$, we have the famous Amitsur–Levitski identity, given by the vanishing of the *standard polynomial* S_{2n}:

$$S_{2n}(x_1, \dots, x_{2n}) = \sum_{\sigma \in S_{2n}} \epsilon_\sigma x_{\sigma(1)} \dots x_{\sigma(2n)}.$$

This can be deduced from the Cayley–Hamilton identity as follows: consider

$$\sum_{\sigma \in S_{2n}} \epsilon_\sigma CH(x_{\sigma(1)} x_{\sigma(2)}, \dots, x_{\sigma(2n-1)} x_{\sigma(2n)}).$$

The terms of CH which do not contain trace expressions give rise to a multiple of the Amitsur–Levitski identity, but the terms which contain trace expressions can be grouped so as to contain each a factor of the form $\mathrm{tr}(S_{2n}(x_1, \dots, x_{2m}))$. Now we claim that $\mathrm{tr}(S_{2n}(x_1, \dots, x_{2m})) = 0$ identically as a formal expression. This can be seen, for instance, by the fact that $\mathrm{tr}(x_{i_1} x_{i_2} \dots x_{i_{2m}}) = \mathrm{tr}(x_{i_2} \dots x_{i_{2m}} x_{i_1})$, and these two permutations have opposite sign and so cancel in the expression.

It can be actually easily seen that $M_n(2n)$ is 1-dimensional, generated by S_{2n}.

8.7 Trace Identities with Involutions

One can also deduce a second fundamental theorem for invariants of matrices and trace identities in the case of involutions. To do so, it is just a question of translating the basic determinants or Pfaffians into trace relations.

As usual we will describe the multilinear relations and deduce the general relations through polarization and restitution. In order to study the multilinear relations involving $m-1$ matrix variables X_i we may restrict to decomposable matrices $X_i := u_i \otimes v_i$. A monomial of degree $m-1$ in these variables in the formal trace algebra is of the type $\mathrm{tr}(M_1)\,\mathrm{tr}(M_2) \dots \mathrm{tr}(M_{k-1}) N_k$, where M_i, N_k are monomials in X_i, X_i^* and each index $i = 1, \dots, m-1$, appears exactly once. To this monomial we can, as usual, associate the trace monomial $\mathrm{tr}(M_1)\,\mathrm{tr}(M_2) \dots \mathrm{tr}(M_{k-1})\,\mathrm{tr}(N_k X_m)$. Extending from monomials to linear combinations, a multilinear trace identity in $m-1$-variables corresponds to a multilinear trace relation in m variables.

Let us substitute in a given monomial, for the matrices X_i, $i = 1, \dots, m$, their values $u_i \otimes v_i$ and expand the monomial according to the rules in 8.2.

To each trace monomial we associate a product $\prod_{i=1}^m \langle x_i, y_i \rangle$, where the list of variables $x_1, \dots, x_m, y_1, \dots, y_m$ is just a reordering of $u_1, \dots, u_m, v_1, \dots, v_m$. Conversely, given such a product, it is easily seen that it comes from a unique trace monomial (in this correspondence one takes care of the formal identities between expressions).

For us it may be convenient to consider the scalar products between the u_i, v_j as evaluations of variables $x_{ij} = (i|j)$ which satisfy the symmetry condition $(i|j) =$

[111] This fact can be proved directly.

$\epsilon(j|i)$, with $\epsilon = 1$ in the orthogonal case and -1 in the symplectic case. We consider the variables $(i|j)$ as entries of an ϵ-symmetric $2m \times 2m$ matrix.

Consider next the space M_m spanned by the *multilinear monomials* $x_{i_1 i_2} x_{i_3 i_4} \dots x_{i_{2m} i_{2m}} = (i_1|i_2)(i_3|i_4)\dots(i_{2m}|i_{2m})$, where $i_1, i_2, i_3, i_4, \dots, i_{2m}, i_{2m}$ is a permutation of $1, 2, \dots, 2m$.

With a different notation[112] this space has been studied in §6.2 as a representation of S_{2m}.

We have a *coding map* from M_m to the vector space of multilinear pure trace expressions $f(X_1, \dots, X_{m+1})$ of degree $m+1$ in the variables $X_i,\ X_i^*$ or the noncommutative trace expressions $g(x_1, \dots, X_m)$ of degree m, with $f(X_1, \dots, X_{m+1}) = \operatorname{tr}(g(X_1, \dots, X_m)X_{m+1})$.

It is given formally by considering the numbers i as vector symbols, the basis of a free module F over the polynomial ring in the variables $(i|j)$, $X_i := (2i-1)\otimes 2i \in F \otimes F \subset \operatorname{End}(F)$, and $(i|j)$ as the scalar product $\langle i|j\rangle$, with the multiplication, trace and adjoint as in 3.1.1., deduced by the formulas:

$$i \otimes j \circ h \otimes k = i \otimes (j|h)k, \quad \operatorname{tr}(i \otimes j) = \epsilon\,(i|j), \quad (i \otimes j)^* = \epsilon\, j \otimes i.$$

For instance, for $m = 2$ we have 15 monomials. We list a few of them for $\epsilon = -1$:

$$(1|2)(3|4)(5|6) = -\operatorname{tr}(X_1)\operatorname{tr}(X_2)\operatorname{tr}(X_3),$$
$$(1|2)(3|5)(4|6) = -\operatorname{tr}(X_1)\operatorname{tr}(X_2^* X_3),$$
$$(1|2)(3|6)(4|5) = \operatorname{tr}(X_1)\operatorname{tr}(X_2X_3),$$
$$(1|3)(2|5)(4|6) = -\operatorname{tr}(X_1^* X_2 X_3^*) = -\operatorname{tr}(X_3 X_2^* X_1)$$

and in noncommutative form:

$$(1|2)(3|4)(5|6) = -\operatorname{tr}(X_1)\operatorname{tr}(X_2), \qquad (12)(35)(46) = -\operatorname{tr}(X_1)X_2^*,$$
$$(1|2)(3|6)(4|5) = \operatorname{tr}(X_1)X_2, \qquad (1|3)(2|5)(4|6) = -X_2^* X_1.$$

Under this coding map the subspace of relations of invariants corresponds to the trace identities. Thus one can reinterpret the second fundamental theorem as giving basic identities from which all the others can be deduced.

The reinterpretation is particularly simple for the symplectic group. In this case the analogue of the Cayley–Hamilton identity is a *characteristic Pfaffian* equation.

Thus consider M_{2n} the space of $2n \times 2n$ matrices, $J := \begin{vmatrix} 0 & 1_n \\ -1_n & 0 \end{vmatrix}$ the standard matrix of the symplectic involution and $A^* := -JA^tJ$ the symplectic involution.

Let M_{2n}^+ denote the space of symmetric matrices under this involution. Under the map $A \to B := AJ$ we identify M_{2n}^+ with the space of skew-symmetric matrices, and we can define the *characteristic Pfaffian* as

[112] We called it $P_+^{2n},\ P_-^{2n}$.

(8.7.1) $$Pf_A(\lambda) := Pf((\lambda I - A)J).$$

We leave it to the reader to verify that $Pf_A(A) = 0$.

Hint. A generic symmetric matrix is semisimple and one can decompose the $2n$-dimensional symplectic space into an orthogonal sum of stable 2-dimensional subspaces. On each such subspace A is a scalar.

If we polarize the identity $Pf_A(A) = 0$, we obtain a multilinear identity $P(X_1, \dots, X_n)$ satisfied by symmetric symplectic matrices. In another form we may write this identity as an identity $P(X_1 + X_1^*, \dots, X_n + X_n^*)$ satisfied by all matrices. This is an identity in degree n. On the other hand, there is a unique canonical pure trace identity of degree $n+1$ which, under the coding map, comes from the Pfaffian of a $(2n+2) \times (2n+2)$ matrix of skew indeterminates; thus, up to a scalar, the polarized characteristic Pfaffian must be a multiple of the noncommutative identity coded by that Pfaffian. For instance, in the trivial case $n = 1$ we get the identity $\mathrm{tr}(X_1) = X_1 + X_1^*$. As for the Cayley–Hamilton one may prove:

Theorem. *All identities for matrices with symplectic involution are a consequence of* $P(X_1 + X_1^*, \dots, X_n + X_n^*)$.

Proof. We will give the proof for the corresponding trace relations. Consider the matrix of (skew) scalar products:

$$S := \begin{vmatrix} 0 & [u_1, u_2] & \dots & [u_1, u_m] & [u_1, v_1] & [u_1, v_2] & \dots & [u_1, v_m] \\ [u_2, u_1] & 0 & \dots & [u_2, u_m] & [u_2, v_1] & [u_1, v_2] & \dots & [u_2, v_m] \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ [u_m, u_1] & [u_m, u_2] & \dots & 0 & [u_m, v_1] & [u_m, v_2] & \dots & [u_m, v_m] \\ [v_1, u_1] & [v_1, u_2] & \dots & [v_1, u_m] & 0 & [v_1, v_2] & \dots & [v_1, v_m] \\ [v_2, u_1] & [v_2, u_2] & \dots & [v_2, u_m] & [v_2, v_1] & 0 & \dots & [v_2, v_m] \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ [v_m, u_1] & [v_m, u_2] & \dots & [v_m, u_m] & [v_m, v_1] & [v_m, v_2] & \dots & 0 \end{vmatrix}.$$

A typical identity in $m - 1$ variables corresponds to a polynomial

$$[w_1, w_2, \dots, w_{2n+2}][x_1, y_1] \dots [x_s, y_s]$$

of the following type. $m = n+1+s$, the elements $w_1, w_2, \dots, w_{2n+2}, x_1, y_1, \dots, x_s, y_s$ are given by a permutation of $u_1, \dots, u_m, v_1, \dots, v_m$ and $[w_1, w_2, \dots, w_{2n+2}]$ denotes the Pfaffian of the principal minor of S whose row and column indices correspond to the symbols $w_1, w_2, \dots, w_{2n+2}$. To understand the corresponding trace relation we argue as follows:

1. Permuting the variables X_i corresponds to permuting simultaneously the u_i and v_i with the same permutation.
2. Exchanging X_i with X_i^* corresponds to exchanging u_i with v_i. Performing these operations we can reduce our expression to one of type

$$[u_1, \dots, u_k, v_1, \dots, v_k, v_{k+1}, \dots, v_{2n+2-k}][x_1, y_1] \dots [x_s, y_s].$$

If $k = n + 1$, the Pfaffian $[u_1, \dots, u_{n+1}, v_1, \dots, v_{n+1}]$ corresponds to the identity of the characteristic Pfaffian. The given expression is a product of this relation by trace monomials. Otherwise we look at the symbol u_{k+1} which does not appear in the Pfaffian but appears in one of the terms $[x_i, y_i]$. We have that u_{k+1} is paired with either a term u_i or a term v_j. If u_{k+1} is paired with v_j, let us use the formula

$$X_j X_{k+1} = u_j \otimes [v_j, u_{k+1}] v_{k+1}.$$

If u_{k+1} is paired with u_j, we use the formula

$$X_j^* X_{k+1} = -v_j \otimes [u_j, u_{k+1}] v_{k+1}.$$

In the first case we introduce a new variable $\overline{X}_j := X_j X_{k+1}$; in the second $\overline{X}_j := X_j^* X_{k+1}$

$$\overline{X}_j := \overline{u}_j \otimes \overline{v}_j, \quad \overline{u}_j := u_j, \ \overline{v}_j := [v_j, u_{k+1}] v_{k+1},$$

or

$$\overline{u}_j := -v_j, \quad \overline{v}_j := [u_j, u_{k+1}] v_{k+1}.$$

Let us analyze for instance the case $[u_j, u_{k+1}]$, the other being similar. We substitute $[v_j, u_{k+1}]$ inside the Pfaffian:

$$\begin{aligned}
&[u_1, \dots, u_k, v_1, \dots, v_k, v_{k+1}, \dots, v_{2n+2-k}][u_j, u_{k+1}] \\
&= [u_1, \dots, u_k, v_1, \dots, v_k, [u_j, u_{k+1}] v_{k+1}, \dots, v_{2n+2-k}] \\
&= [u_1, \dots, u_k, v_1, \dots, v_k, \overline{v}_j, \dots, v_{2n+2-k}].
\end{aligned}$$

We obtain a new formal expression in $m - 1$-variables. The variable X_{k+1} has been suppressed, while the variable X_j has been substituted with a new variable $\overline{X}_j$. In order to recover the old m-variable expression from the new one, one has to substitute $\overline{X}_j$ with $X_j X_{k+1}$ or $X_j^* X_{k+1}$. By induction the theorem is proved. □

8.8 The Orthogonal Case

The orthogonal case is more complicated, due to the more complicated form of the second fundamental theorem for orthogonal invariants.

Consider the matrix of (symmetric) scalar products:

$$S := \begin{vmatrix} (u_1,u_1) & (u_1,u_2) & \dots & (u_1,u_m) & (u_1,v_1) & (u_1,v_2) & \dots & (u_1,v_m) \\ (u_2,u_1) & (u_2,u_2) & \dots & (u_2,u_m) & (u_2,v_1) & (u_1,v_2) & \dots & (u_2,v_m) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ (u_m,u_1) & (u_m,u_2) & \dots & (u_m,u_m) & (u_m,v_1) & (u_m,v_2) & \dots & (u_m,v_m) \\ (v_1,u_1) & (v_1,u_2) & \dots & (v_1,u_m) & (v_1,v_1) & (v_1,v_2) & \dots & (v_1,v_m) \\ (v_2,u_1) & (v_2,u_2) & \dots & (v_2,u_m) & (v_2,v_1) & (v_2,v_2) & \dots & (v_2,v_m) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ (v_m,u_1) & (v_m,u_2) & \dots & (v_m,u_m) & (v_m,v_1) & (v_m,v_2) & \dots & (v_m,v_m) \end{vmatrix}.$$

A relation may be given by the determinant of any $(n+1) \times (n+1)$ minor times a monomial. Again we have that:

1. Permuting the variables X_i corresponds to permuting simultaneously the u_i and v_i with the same permutation.
2. Exchanging X_i with X_i^* corresponds to exchanging u_i with v_i.

In this case the basic identities to consider correspond to the determinants of scalar products, which we may denote by

$$(u_{i_1}u_{i_2}\dots u_{i_k}v_{j_1}v_{j_2}\dots v_{j_{n+1-k}} \mid u_{i_{k+1}}u_{i_{k+2}}\dots u_{i_{n+1}}v_{j_{n+2-k}}v_{j_{n+3-k}}\dots v_{j_{n+1}}),$$

where the condition that they represent a multilinear relation in the matrix variables $X_1,\dots,X_{n+1}$ means that the indices $i_1,\dots,i_{n+1}$ and also the indices $j_1,\dots,j_{n+1}$ are permutations of $1,2,\dots,n+1$.

Using the previous symmetry laws we may first exchange all u_i, v_i for all the indices for which u_i appears on the left but v_i does not appear. After reordering the indices we end up with a determinant

$$(u_1u_2\dots u_kv_1v_2\dots v_{n+1-k} \mid u_{k+1}u_{k+2}\dots u_{n+1}v_{n+2-k}v_{n+3-k}\dots v_{n+1}),$$
$$n+1-k \geq k.$$

From the coding map, the above determinant corresponds to some trace relation which we denote by $F_k(X_1,\dots,X_{n+1})$. Unlike the symplectic case these basic identities do not have enough symmetry to consider them as polarizations of 1-variable identities.

We have, exchanging u_i, v_i,

$$F_k(X_1,\dots,X_i,\dots,X_{n+1})$$

(a) $$= -F_k(X_1,\dots,X_i^t,\dots,X_{n+1}), \text{ if } \quad i > n+k-1, \text{ or } i \leq k.$$

If σ is a permutation of $1,2,\dots,n+1$ which preserves the subsets $(1,2,\dots,k)$; $(k+1,\dots,n+1-k)$; $(n_1-k+1,\dots,n+1)$ we have

(b) $$F_k(X_1,\dots,X_i,\dots,X_{n+1}) = F_k(X_{\sigma(1)},\dots,X_{\sigma(i)},\dots,X_{\sigma(n+1)}).$$

This means that F_k can be viewed as polarization of an identity in 3 variables $F_k(X, Y, Z)$. Moreover, from the first relation, the first and third variable can be taken to be skew-symmetric variables $\frac{X-X'}{2}$, $\frac{Z-Z'}{2}$.

Finally, exchanging all u_i, v_i and transposing the determinant we get

$$\begin{aligned} &(v_{k+1}v_{k+2}\dots v_{n+1}u_{n+2-k}u_{n+3-k}\dots u_{n+1} \mid v_1v_2\dots v_ku_1u_2\dots u_{n+1-k}) \\ &= (u_{n+1}u_{n+3-k}\dots u_{n+2-k}v_{n+1}\dots v_{k+2}v_{k+1} \mid u_{n+1-k}\dots u_2u_1v_k\dots v_2v_1) \\ \text{(c)}\quad &= F_k(X_1^t, \dots, X_i^t, \dots, X_{n+1}^t) = F_k(X_{n+1}, \dots, X_{n+1-i}, \dots, X_1). \end{aligned}$$

We need to make a remark when we pass from the trace relation to the noncommutative trace relation. We pick a variable X_i and write

$$F_k(X_1, \dots, X_i, \dots, X_{n+1}) = \operatorname{tr}(G_k^i(X_1, \dots, \check{X}_i, \dots, X_{n+1})X_i).$$

It is clear that G_k^i depends on i only up to the symmetry embedded in the previous statements (a), (b), (c).

From the symmetry (b) it appears that we obtain 3 inequivalent identities G_k^1, G_k^2, G_k^3 from the 3 subsets. From (c) we obtain a different deduction formula from the identity in the first and the third subset.

$$F_k(X_1, \dots, X_{n+1}) = \operatorname{tr}(G_k^1(X_2, \dots, X_{n+1})X_1),$$

gives

$$\begin{aligned} F_k(X_1^t, \dots, X_{n+1}^t) &= \operatorname{tr}(G_k^1(X_2^t, \dots, X_{n+1}^t)X_1^t) \\ &= \operatorname{tr}(G_k^1(X_2^t, \dots, X_{n+1}^t)^t X_1) = \operatorname{tr}(G_k^3(X_{n+1}, \dots, X_2)X_1). \end{aligned}$$

Hence

$$G_k^1(X_2^t, \dots, X_{n+1}^t)^t = G_k^3(X_{n+1}, \dots, X_2).$$

In this sense only G_k^1 and G_k^2 should be taken as generating identities.

8.9 Some Estimates

Both the first and second fundamental theorem for matrices are not as precise as the corresponding theorems for vectors and forms. Several things are lacking; the first is a description of a minimal set of generators for the trace invariants. At the moment the best result is given by the following:

Theorem (Razmyslov). *The invariants of $n \times n$ matrices are generated by the elements* $\operatorname{tr}(M)$ *where M is a monomial of degree $\leq n^2$.*

Proof. The precise question is: for which values of m, can we express the trace monomial $\operatorname{tr}(X_1X_2\dots X_{m+1})$ as a product of shorter trace monomials? This in turn is equivalent to asking for which values of m, in the free algebra with trace $R(n)$, the monomial $X_1X_2 \dots X_m$ lies in the ideal $T(n)^+R(n)$, where as usual $T(n)^+$ denotes the trace invariants without constant term.

If F denotes the free algebra with trace and T the trace expressions, by the general correspondence between the symmetric group and the trace expressions, we have that the elements of S_{m+1} which map to T^+F are not full cycles. Let Q_{m+1} denote the subspace that they span in $\mathbb{C}[S_{m+1}]$. Let P_{m+1} be the subspace of $\mathbb{C}[S_{m+1}]$ spanned by the full cycles so that $Q_{m+1} \oplus P_{m+1} = \mathbb{C}[S_{m+1}]$. By symmetry, the requirement that modulo the identities of $n \times n$ matrices we have $X_1X_2 \dots X_m \in T(n)^+R(n)$ is equivalent to requiring that $P_{m+1} \subset Q_{m+1}+I_n$; in other words, that $\mathbb{C}[S_{m+1}] = Q_{m+1}+I_n$.

So let us prove that this equality is true for $m = n^2$. Consider the automorphism $\tau(\sigma) = \epsilon_\sigma \sigma$. τ is $(-1)^m$ on P_{m+1}. It preserves Q_{m+1} and maps I_n to J_n the ideal generated by a symmetrizer on $n+1$ elements. Therefore $\tau(Q_{m+1}+I_n) = Q_{m+1}+J_n$ so that $Q_{m+1} + I_n = Q_{m+1} + I_n + J_n$. Finally, if $m \geq n^2$, every diagram with $m+1$ cases belongs either to I_n or to J_n so that $S_{n^2+1} = I_n + J_n$, and the theorem is proved. □

If one computes explicitly for small values of n, one realizes that n^2 does not appear to be the best possible estimate. A better possible estimate should be $\binom{n+1}{2}$. It is interesting that these estimates are related to a theorem in noncommutative algebra known as the Dubnov–Ivanov–Nagata–Higman Theorem.

8.10 Free Nil Algebras

To explain the noncommutative algebra we start with a definition:

Definition. (1) An algebra R (without 1) is said to be nil of exponent n if for every $r \in R$ we have $r^n = 0$

(2) An algebra R (without 1) is said to be nilpotent of exponent m if $R^m = 0$.

To say that R is nilpotent of exponent m means that every product $r_1r_2 \dots r_m = 0$, $r_i \in R$. Of course an algebra R which is nilpotent of exponent m is also nil of exponent m. The converse is not true.

Theorem 1.

(i) The algebra $R(n)^+/T(n)^+R(n)$ is the free nil algebra of exponent n. In other words, it is the free algebra without 1*, modulo the T-ideal generated by the polynomial identity $z^n = 0$.*

(ii) In characteristic 0 *an algebra R, nil of exponent n, is nilpotent of exponent $\leq n^2$.*

Proof. (i) In the free algebra with trace without 1 we impose the identity $t(z) = 0$, that is the trace of every element is 0. We obtain the free algebra without trace. Thus the free algebra without 1 modulo the polynomial identity $z^n = 0$ can be seen as the free algebra with trace without 1 modulo the two identities $t(z) = z^n = 0$. Now the Cayley–Hamilton identity plus the identity $t(z) = 0$ becomes $z^n = 0$. Hence the free algebra without one modulo the polynomial identity $z^n = 0$ can be seen also as the free algebra with trace without 1 modulo the two identities $t(z) = CH_n(z) = 0$.

But clearly this last algebra is $R(n)^+/T(n)^+R(n)$.

(ii) To say that $X_1X_2\dots X_m$ lies in the ideal $T(n)^+R(n)$ is the same as saying that $X_1X_2\dots X_m = 0$ in $R(n)^+/T(n)^+R(n)$. Since this is the free algebra modulo $z^n = 0$, to say that $X_1X_2\dots X_m = 0$ is equivalent to saying that $X_1X_2\dots X_m = 0$ is an identity in $R(n)^+/T(n)^+R(n)$ or that $R(n)^+/T(n)^+R(n)$ is nilpotent of exponent $\leq m$.

If the free algebra modulo $z^n = 0$ is nilpotent of exponent m, then so is every nil algebra of exponent n, since it is a quotient of the free algebra. Finally, from the previous theorem we know that $m = n^2$ satisfies the condition that $X_1X_2\dots X_m$ lies in the ideal $T(n)^+R(n)$. □

Remark. We have in fact proved, for fixed n, (over fields of characteristic 0) that for an integer m the following conditions are equivalent:

(i) $\mathbb{C}[S_{m+1}] = Q_{m+1} + I_n$.
(ii) The invariants of $n \times n$ matrices are generated by the elements $\mathrm{tr}(M)$ where M is a monomial of degree $\leq m$.
(iii) An algebra R that is nil of exponent n is nilpotent of exponent $\leq m$.

We have shown that there is a minimum value m_0 for which these three conditions are satisfied and $m_0 \leq n^2$.

It is known that the minimum value of m is $\geq \binom{n+1}{2}$. One may conjecture that it is exactly $\binom{n+1}{2}$. This is true by explicit computations for $n \leq 5$.

One can also interpret Theorem 8.7 on symplectic involution in a similar way and obtain:

Theorem 2.

(i) The algebra $R^s(2n)^+/T^s(2n)^+R^s(2n)$ is the free nil algebra with involution in which every symmetric element is nilpotent of exponent n; in other words, it is the free algebra without 1 *with involution modulo the polynomial identity* $(z + z^*)^n = 0$.
(ii) In characteristic 0 *if an algebra R with involution satisfies* $(z + z^*)^n = 0$, *then it satisfies* $z^{2n} = 0$.

Proof. The first part is like the case of nil algebras. For the second part, it is enough to remark that $2n \times 2n$ symplectic matrices satisfy the Cayley–Hamilton theorem in degree $2n$, which then can be deduced from the Pfaffian identity. □

8.11 Cohomology

We can apply the theory developed to the computation of the cohomology of the classical Lie algebras. Given a semisimple Lie algebra L, we apply Cartan's theorem that the cohomology space $H^i(L, \mathbb{C})$ with constant coefficients can be identified to the space of invariant multilinear and alternating functions of k variables in L. These spaces are the summands of a graded algebra under $\wedge$ product.

Let us start thus from $sl(n, \mathbb{C})$, and even start computing the space of invariant multilinear and alternating functions of k matrix variables X_i (we restrict to trace 0 only later).

By definition one can obtain such a function by alternating any given multilinear function $f(x_1, \dots, x_k)$. Let us call $Af(x_1, \dots, x_k) := \frac{1}{k!}\sum_{\sigma\in S_k} \epsilon_\sigma f(x_{\sigma(1)}, \dots, x_{\sigma(k)})$. Moreover, if a multilinear function is a product

$$h(x_1, \dots, x_{a+b}) = f(x_1, \dots, x_a)g(x_{a+1}, \dots, x_{a+b}),$$

one has by definition that $Ah = Af \wedge Ag$.

We want to apply these to invariant functions of matrices. A multilinear invariant function of matrices is up to permuting the variables a product of functions

$$\operatorname{tr}(X_1X_2 \dots X_a)\operatorname{tr}(X_{a+1} \dots X_{a+b}) \dots tr(X_sX_s{+}1 \dots X_{s+t}).$$

Therefore we see:

Lemma. *The algebra of multilinear and alternating functions of $n \times n$ matrices is generated (under $\wedge$) by the elements* $\operatorname{tr}(S_{2h+1}(X_1, \dots, X_{2h+1}))$, $h = 1, \dots, n-1$. *S_{2h+1} is the standard polynomial.*

Proof. When we alternate one of the factors $\operatorname{tr}(X_1X_2 \dots X_a)$ we then obtain $\operatorname{tr}(S_a(X_1, \dots, X_a))$. We have noticed in 8.6 that this function is 0 if a is even, and it vanishes on $n \times n$ matrices if $a \geq 2n$ by the Amitsur–Levitski theorem. The claim follows. □

Theorem. *The algebra $A(n)$ of multilinear and alternating functions of $n \times n$ matrices is the exterior algebra $\wedge(c_1, \dots, c_{2n-1})$ generated by the elements*

$$c_h := \operatorname{tr}(S_{2h+1}(X_1, \dots, X_{2h+1})),\ h = 0, \dots, n-1,\ \textit{of degree } 2h+1.$$

Proof. We have seen that these elements generate the given algebra. Moreover, since they are all of odd degree they satisfy the basic relations $c_h \wedge c_k = -c_k \wedge c_h$. Thus we have a surjective homomorphism $\pi : \bigwedge(c_1, \dots, c_{2n-1}) \to A(n)$. In order to prove the claim, since every ideal of $\bigwedge(c_1, \dots, c_{2n-1})$ contains the element $c_1 \wedge \dots \wedge c_{2n-1}$, it is enough to see that this element does not map to 0 under π. If not, we would have that in $A(n)$ there are no nonzero elements of degree $\sum_{h=1}^{n}(2h-1) = n^2$. Now the determinant $X_1 \wedge X_2 \wedge \dots \wedge X_{n^2}$ of n^2 matrix variables is clearly invariant alternating, of degree n^2, hence the claim. □

In fact the elements c_i have an important property which comes from the fact that the cohomology $H^*(G, \mathbb{R})$ of a group G has extra structure. From the group properties and the properties of cohomology it follows that $H^*(G, \mathbb{R})$ is also a Hopf algebra. The theorem of Hopf, of which we mentioned the generalization by Milnor–Moore in Chapter 8, implies that $H^*(G, \mathbb{R})$ is an exterior algebra generated by the primitive elements (cf. Chapter 8, §7). The comultiplication Δ in our case can be viewed as follows, we map a matrix X_i to $X_i \otimes 1 + 1 \otimes X_i$, and then

$$\begin{aligned}&\Delta\,\mathrm{tr}(S_{2h+1}(X_1,\dots,X_{2h+1}))\\&\quad=\mathrm{tr}(S_{2h+1}(X_1\otimes 1+1\otimes X_1,\dots,X_{2h+1}\otimes 1+1\otimes X_{2h+1})).\end{aligned}$$

When we expand the expression

$$\mathrm{tr}(S_{2h+1}(X_1\otimes 1+1\otimes X_1,\dots,X_{2h+1}\otimes 1+1\otimes X_{2h+1}))$$

we obtain a sum:

$$\begin{aligned}&\mathrm{tr}(S_{2h+1}(X_1,\dots,X_{2h+1}))\otimes 1+1\otimes\mathrm{tr}(S_{2h+1}(X_1,\dots,X_{2h+1}))\\&\quad+\sum_{a+b=2h-1} c_a\,\mathrm{tr}(S_a(X_{i_1},,\dots,X_{i_a}))\otimes\mathrm{tr}(S_b(X_{j_1},\dots,X_{j_b})).\end{aligned}$$

The first two terms give $c_h\otimes 1+1\otimes c_h$; the other terms with $a,b>0$ vanish since either a or b is even. We thus see that the c_i are primitive.

In a similar way one can treat the symplectic and odd orthogonal Lie algebras. We only remark that in the odd case, the element -1 acts trivially on the Lie algebra, hence instead of $SO(2n+1,\mathbb{C})$-invariants we can work with $O(2n+1,\mathbb{C})$-invariants, for which we have the formulas $\mathrm{tr}(M)$ with M a monomial in X_i, X_i^*. Since the variables are antisymmetric the variables X_i^* disappear and we have to consider the same expressions c_h. In this case we have a further constraint. Applying the involution we have

$$\begin{aligned}\mathrm{tr}(S_{2h+1}(X_1,\dots,X_{2h+1}))&=\mathrm{tr}(S_{2h+1}(X_1,\dots,X_{2h+1})^*)\\&=\mathrm{tr}(-S_{2h+1}(X_{2h+1},\dots,X_1))\\&=\mathrm{tr}(-(-1)^h S_{2h+1}(X_1,\dots,X_{2h+1})).\end{aligned}$$

Therefore $c_h=0$ unless h is odd. We deduce that for $Sp(2n,\mathbb{C}),\ SO(2n+1,\mathbb{C})$ the corresponding cohomology is an exterior algebra in the primitive elements

$$d_h:=\mathrm{tr}(S_{4h+3}(X_1,\dots,X_{4h+3})),\ h=0,\dots,n-1,\ \text{ of degree } 4h+3.$$

The even case $SO(2n,\mathbb{C})$ is more complicated, since in this case we really need to compute with $SO(2n,\mathbb{C})$-invariants.

It can be proved, and we leave it to the reader, that besides the elements $d_h:=\mathrm{tr}(S_{4h+3}(X_1,\dots,X_{4h+3})),\ h=0,\dots,n-2$ of degree $4h+3$, we also have an element of degree $2n-1$ which can be obtained by antisymmetrizing the invariant $Q(X_1,X_2X_3,X_4X_5,\dots,X_{2n-2}X_{2n-1})$ (cf. 8.2). In order to understand this new invariant, and why we construct it in this way, let us recall the discussion of §8.7. Let $J=\begin{vmatrix}0&1_n\\-1_n&0\end{vmatrix}$. We have seen that if X is skew-symmetric, XJ is symplectic-symmetric. Generically XJ has n-distinct eigenvalues $\lambda_1,\dots,\lambda_n$, each counted with multiplicity 2. Then $Pf(X)=\prod_i\lambda_i$. To see this in a concrete way, let Λ be the diagonal $n\times n$ matrix with the λ_i as eigenvalues. Consider the matrix $X_\Lambda:=\begin{vmatrix}0&\Lambda\\-\Lambda&0\end{vmatrix}$. Then $X_\Lambda J=\begin{vmatrix}-\Lambda&0\\0&-\Lambda\end{vmatrix}$. Finally $Pf(X)=\det(X)=\prod_i\lambda_i$.

Next observe that

$$\mathrm{tr}((XJ)^h) = 2(-1)^h \sum_{i=1}^{n} \lambda_i^h.$$

We deduce that if $P_n(t_1, \dots, t_n)$ is the polynomial expressing the n^{th} elementary symmetric function in terms of the Newton functions, we have

$$Pf(X) = P_n(\mathrm{Tr}(-XJ)/2, tr((-XJ)^2/2, \dots, \mathrm{tr}((-XJ)^n/2).$$

It is important to polarize this identity. Recall that in Corollary 8.5.1 we have seen that the polarization of the determinant is the expression

$$\sum_{\sigma \in S_n} \epsilon_\sigma \phi_\sigma(x_1, \dots, x_n).$$

We deduce that the polarization of $Pf(X)$ is the expression

$$Q(X_1, \dots, X_n) = \sum_{\sigma \in S_n} \epsilon_\sigma 2^{-|\sigma|} \phi_\sigma(-X_1 J, \dots, -X_n J), \tag{8.11.1}$$

where we denote by $|\sigma|$ the number of cycles into which σ decomposes. As in the proof of the Amitsur–Levitski theorem we see that when we perform the antisymmetrization of $Q(X_1, X_2X_3, \dots, X_{2n-2}X_{2n})$ only the terms corresponding to full cycles survive, and we obtain $\mathrm{tr}(S_{2n-1}(-X_1J, \dots, -X_{2n-1}J))$. It is easy to see, as we did before, that $\mathrm{tr}(S_{2n-1}(-X_1J, \dots, -X_{2n-1}J))$ is a primitive element. One can now take advantage of the following easy fact about Hopf algebras:

Proposition. *In a graded Hopf algebra, linearly independent primitive elements generate an exterior algebra.*

We have thus found primitive elements for the cohomology of $so(2n, \mathbb{C})$ of degrees $2n-1$, $4h+3$, $h = 0, \dots, n-2$. Their product is a nonzero element of degree $2n^2 - n = \dim so(2n, \mathbb{C})$. Therefore there cannot be any other primitive elements since there are no skew-symmetric invariants in degree $> \dim so(2n, \mathbb{C})$. We have thus completed a description of cohomology and primitive generators.

Remark. A not-so-explicit description of cohomology can be given for any simple Lie algebra (Borel). We have generators in degrees $2h_i - 1$ where the numbers h_i, called *exponents*, are the degrees of the generators of the polynomials, invariant under W in the reflection representation.

9 The Analytic Approach to Weyl's Character Formula

9.1 Weyl's Integration Formula

The analytic approach is based on the idea of applying the orthogonality relations for class functions in the compact case and obtaining directly a formula for the characters.

We illustrate this for the unitary group $U(n, \mathbb{C})$, leaving some details to the reader for the general case of a compact Lie group (see [A]).

The method is based on Weyl's integration formula for class functions.

Let $d\mu$, $d\tau$, respectively, denote the normalized Haar measures on $U(n, \mathbb{C})$ and on T, the diagonal unitary matrices. We have

Theorem (Weyl's Integration Formula). *For a class function f on $U(n, \mathbb{C})$ we have*

$$\int_{U(n,\mathbb{C})} f(g)d\mu = \frac{1}{n!}\int_T f(t)V(t)\overline{V}(t)d\tau, \tag{9.1.1}$$

where $V(t_1, \dots, t_n) := \prod_{i<j}(t_i - t_j)$ is the Vandermonde determinant.

Assuming this formula for a moment we have:

Corollary. *The Schur functions $S_\lambda(y)$ are irreducible characters.*

Proof. By Weyl's formula, we compute

$$\frac{1}{n!}\int_T S_\lambda(y)\overline{S}_\mu(y)V(y)\overline{V}(y)d\tau = \frac{1}{n!}\int_T A_\lambda(y)\overline{A}_\mu(y)d\tau = \delta_{\lambda,\mu},$$

where the last equality follows from the usual orthogonality of characters for the torus. It follows that the class functions on $U(n, \mathbb{C})$, which restricted to T give the Schur functions, are orthonormal with respect to Haar measure.

The irreducible characters restricted to T are symmetric functions which are sums of monomials with positive integer coefficients, and we have proved (Chapter 10, §6.7) that they span the symmetric functions. It follows that the Schur functions are $\pm$ irreducible characters. The sign must be plus since their leading coefficient is 1.

By the description of the ring of symmetric functions with the function $e_n = \prod y_i$ inverted, it follows immediately that the characters $S_\lambda e_n^k$, $ht(\lambda) < n$, $k \in \mathbb{Z}$, are a basis of this ring, and so they exhaust all possible irreducible characters. □

Let us explain the proof of 9.1.1. The idea is roughly the following. Decompose the unitary group into its conjugacy classes. Two unitary matrices are conjugate if and only if they have the same eigenvalues; hence each conjugacy class intersects the torus in an orbit under S_n. Generically, the eigenvalues are distinct and a conjugacy class intersects the torus in $n!$ points. Therefore if we use the torus T as a parameter space for conjugacy classes we are counting each class $n!$ times. Now perform the integral by first integrating on each conjugacy class the function $f(g)$, which now is a constant function, then on the set of conjugacy classes, or rather on T, dividing by $n!$.

If we keep track of this procedure correctly we see that the various conjugacy classes do not all have the same volume and the factor $V(t)\overline{V}(t)$ arises in this way.

Remark that given an element $t \in T$, its conjugacy class can be identified with the set $U(n, \mathbb{C})/Z_t$ where $Z_t = \{g \in U(n, \mathbb{C}) \mid gt = tg\}$. If t has distinct eigenvalues we have $Z_t = T$, and so the generic conjugacy classes can be identified with

$U(n,\mathbb{C})/T$. In any case we have a global mapping $\pi : U(n,\mathbb{C})/T \times T \to U(n,\mathbb{C})$ given by $\pi(gT, y) := gyg^{-1}$.

Under this mapping, given a class function $f(g)$ on $U(n,\mathbb{C})$, we have that $f(\pi(gT, y))$ depends only on the coordinate y and not on gT.

Now the steps that we previously described are precisely the following:

1. $U(n,\mathbb{C})/T$ is a compact manifold over which $U(n,\mathbb{C})$ acts differentiably and transitively, that is, it is a homogeneous space. $U(n,\mathbb{C})/T$ has a measure invariant under the action of $U(n,\mathbb{C})$, which we normalize to have volume 1 and still call a Haar measure.
2. Consider the open sets in T^0, $U(n,\mathbb{C})^0$ made of elements with distinct eigenvalues in T, $U(n,\mathbb{C})$, respectively.
   (i) The complements of these sets have measure 0.
   (ii) π induces a mapping $\pi^0 : U(n,\mathbb{C})/T \times T^0 \to U(n,\mathbb{C})^0$ which is an unramified covering with exactly $n!$ sheets.
3. Let $d\mu$ denote the Haar measure on $U(n,\mathbb{C})$ and $d\nu$, $d\tau$ the normalized Haar measures on $U(n,\mathbb{C})/T$ and T, respectively. Let $\pi^*(d\mu)$ denote the measure induced on $U(n,\mathbb{C})/T \times T^0$ by the covering π^0. Then

$$\pi^*(d\mu) = V(t)\overline{V}(t)d\nu \times d\tau. \tag{9.1.2}$$

From these 3 steps Weyl's integration formula follows immediately. From 1 and 2:

$$\int_{U(n,\mathbb{C})} f(g)d\mu = \int_{U(n,\mathbb{C})^0} f(g)d\mu = \frac{1}{n!}\int_{U(n,\mathbb{C})/T\times T^0} f(gtg^{-1})\pi^*(d\mu).$$

From 3 and Fubini's theorem we have

$$\begin{aligned}
&\int_{U(n,\mathbb{C})/T\times T^0} f(gtg^{-1})V(t)\overline{V}(t)d\nu \times d\tau \\
&= \int_{T^0} V(t)\overline{V}(t)\int_{U(n,\mathbb{C})/T} f(gtg^{-1})d\nu d\tau \\
&= \int_{T^0} V(t)\overline{V}(t)f(t)d\tau.
\end{aligned}$$

Let us now explain the previous 3 statements.

1. We shall use some simple facts of differential geometry. Given a Lie group G and a closed Lie subgroup H, the coset space G/H has a natural structure of differentiable manifold over which G acts in an analytic way (Chapter 4, §3.7).

On an oriented manifold M of dimension n a measure can be defined by a differential form of degree n, which in local coordinates is $f(x)dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n$.

In order to define on the manifold G/H a G-invariant metric in this way we need a G-invariant top differential form. This form is determined by the nonzero value it takes at a given point, which can be the coset H. Given such a value ψ_0 the form exists if and only if ψ_0 is invariant under the action of H on the tangent space of G/H at H itself. This means that H acts on this space by matrices of determinant 1. If H is compact, then it will act by orthogonal matrices, hence of determinant ± 1. If, furthermore, H is connected, the determinant must be 1. This is our case.[113]

This tangent space of G/H at H is $\mathfrak{g}/\mathfrak{h}$ where $\mathfrak{g}, \mathfrak{h}$ are the Lie algebras of G, H.[114]

In our case the Lie algebra of $U(n, \mathbb{C})$ consists of the matrices iA with A Hermitian, the Lie algebra of T corresponds to the subspace where A is diagonal, and we can take as the complement the set H_0 of anti-Hermitian matrices with 0 in the diagonal.

H_0 has basis the elements $i(e_{hj} + e_{jh}),\ e_{hj} - e_{jh}$ with $h < j$.

2. (i) Since $T, U(n, \mathbb{C})$ are differentiable manifolds and the Haar measures are given by differential forms, to verify that a set S has measure 0, it is sufficient to work in local coordinates and see that S has measure 0 for ordinary Lebesgue measure. For T, using angular coordinates, the condition on our set S is that two coordinates coincide, clearly a set of measure 0. For $U(n, \mathbb{C})$ it is a bit more complicated but similar, and we leave it as an exercise.

(ii) We have $U(n, \mathbb{C})/T \times T^0 = \pi^{-1}U(n, \mathbb{C})^0$ and since clearly π is proper, also π^0 is proper. The centralizer of a given matrix X in T^0 is T itself and the conjugacy class of X intersects T at exactly $n!$ elements obtained by permuting the diagonal entries. This implies that the preimage of any element in $U(n, \mathbb{C})^0$ under π^0 has $n!$ elements. Finally in the proof of 3 we show that the Jacobian of π^0 is nonzero at every element. Hence π^0 is a local diffeomorphism. These facts are enough to prove (i).

3. This requires a somewhat careful analysis. The principle is that when we compare two measures on two manifolds given by differential forms, under a local diffeomorphism, we are bound to compute a Jacobian. Let us denote $\pi^*(d\mu) = F(g, t)dv\, d\tau$ and let $\omega_1, \omega_2, \omega_3$ represent, respectively, the value of the form defining the normalized measure in the class of 1 for $U(n, \mathbb{C})/T, T, U(n, \mathbb{C})$, respectively.

Locally $U(n, \mathbb{C})/T$ can be parameterized by the classes $e^A T,\ A \in H_0$.

Given $g \in U(n, \mathbb{C}), y \in T$ consider the map $\ell_{g,y} : U(n, \mathbb{C})/T \times T \to U(n, \mathbb{C})/T \times T$ defined by multiplication $\ell_{g,y} : (hT, z) \mapsto (ghT, yz)$. It maps $(T, 1)$ to (gT, y). Its differential induces a linear map of the tangent spaces at these points. Denote this map $d_{g,y}$.

By construction $\omega_1 \wedge \omega_2$ is the pullback under the map $d_{g,y}$ of the value of $dv\, d\tau$ at the point (gT, y); similarly for $U(n, \mathbb{C})$, the value of $d\mu$ at an element h is the pullback of ω_3 under the map $r_h : x \to xh^{-1}$.

On the other hand $d\pi_{(gT,y)}$ induces a linear map (which we will see is an isomorphism) between the tangent space of $U(n, \mathbb{C})/T \times T$ at (gT, y) and the tan-

---

[113] The orientability follows.

[114] There is also a more concrete realization of $U(n, \mathbb{C})/T$ for which one can verify all the given statements. It is enough to take the set of $n \times n$ Hermitian matrices with n prescribed distinct eigenvalues.

gent space of $U(n, \mathbb{C})$ at gyg^{-1}. By definition $\pi^*(d\mu)$ at the point (gT, y) is $d\pi^*_{(gT,y)}\mu$.

Consider the composition of maps

$$\psi : U(n, \mathbb{C})/T \times T \xrightarrow{\ell_{g,y}} U(n, \mathbb{C})/T \times T \xrightarrow{\pi} U(n, \mathbb{C})$$
$$\xrightarrow{r_{gyg^{-1}}} U(n, \mathbb{C}), \quad \psi(T, 1) = 1.$$

We get that the pullback of ω_3, (at 1) under this map ψ is $F(g, t)\omega_1 \wedge \omega_2$.

In order to compute the function $F(g, t)$, we fix a basis of the tangent spaces of $U(n, \mathbb{C})/T, T, U(n, \mathbb{C})$ at 1 and compute the Jacobian of $r_{gyg^{-1}}\pi\ell_{g,y}$ in this basis. This Jacobian is $F(g, t)$ up to some constant, independent of g, t, which measures the difference between the determinants of the given bases and the normalized invariant form.

At the end we will compute the constant by comparing the integrals of the constant function 1. Take as local coordinates in $U(n, \mathbb{C})/T \times T$ the parameters $\exp(A)T, \exp(D)$ where $A \in H_0$ and D is diagonal. Since we need to compute the linear term (differential) of a map, we can compute everything in power series, saving at each step only the linear terms (set $\mathcal{O}$ to be the higher order terms):

$$\psi((1 + A)T, 1 + D) = g(1 + A)y(1 + D)(g(1 + A))^{-1}gy^{-1}g^{-1}$$
$$= 1 + g[D + A - yAy^{-1}]g^{-1} + \mathcal{O}.$$

The required Jacobian is the determinant of $(D, A) \mapsto g[D + A - yAy^{-1}]g^{-1}$. Since conjugating by g has determinant 1 we are reduced to the map $(D, A) \to (D, A - yAy^{-1})$. This is a block matrix with one block the identity. We are reduced to $\gamma : A \mapsto A - yAy^{-1}$.

To compute this determinant we complexify the space, obtaining as basis the elements e_{ij}, $i \neq j$. We have $\gamma(e_{i,j}) = (1 - y_i y_j^{-1})e_{i,j}$. In this basis the determinant is

$$\prod_{i \neq j}(1 - y_i y_j^{-1}) = \prod_{i<j}(1 - y_i y_j^{-1})(1 - y_j y_i^{-1})$$
$$= \prod_{i<j}(y_j - y_i)(y_j^{-1} - y_i^{-1}) = V(y)\overline{V}(y)$$

since y is unitary.[115]

At this point formula 9.1.1 is true, possibly up to some multiplicative constant.

The constant is 1 since if we take $f = 1$ the left-hand side of 9.1.1 is 1. As for the right-hand side, remember that the monomials in the y_i (coordinates of T) are the irreducible characters and so they are orthonormal. It follows that $\frac{1}{n!}\int_T V(y)\overline{V}(y)d\tau = 1$, since $V(y) = \sum_{\sigma \in S_n} \epsilon_\sigma y_{\sigma(1)}^{n-1} y_{\sigma(2)}^{n-2} \cdots y_{\sigma(n-1)}$, and the proof is complete.

[115] Note that this formula shows that $U(n, \mathbb{C})/T \times T^0$ is the set of points of $U(n, \mathbb{C})/T \times T$ where $d\pi$ is invertible.

We have used the group $U(n, \mathbb{C})$ to illustrate this theorem, but the proof we gave is actually quite general. Let us see why. If K is a connected compact Lie group and T a maximal torus, we know from Chapter 10, §7.3 that $K = \cup_{g \in K} gTg^{-1}$ and the normalizer N_T modulo T is the Weyl group W. The nonregular elements are the union of a finite number of submanifolds, thus a set of measure 0. The quotient K/T is the flag variety (Chapter 10, §7.3) and, when we complexify the tangent space of K/T in the class of T we have the direct sum of all the root spaces. Therefore, the determinant of the corresponding Jacobian is $\prod_{\alpha \in \Phi}(1-t^\alpha)$. When t is in the compact torus, $t^{-\alpha}$ is conjugate to t^α, so setting

$$V(t) := \prod_{\alpha \in \Phi^+} (1 - t^\alpha),$$

one has the general form of Weyl's integration formula:

Theorem (Weyl's Integration Formula). *For a class function f on K we have*

$$\int_K f(g) d\mu = \frac{1}{|W|} \int_T f(t) V(t) \overline{V}(t) d\tau \tag{9.1.3}$$

where $V(t) := \prod_{\alpha \in \Phi^+}(1 - t^\alpha)$.

10 Characters of Classical Groups

10.1 The Symplectic Group

From the theory developed in Chapter 10 we easily see the following facts for $Sp(2n, \mathbb{C})$ or of its compact form, which can be proved directly in an elementary way.

Fix a symplectic basis $e_i, f_i,\ i = 1, \dots, n$. A maximal torus in $Sp(2n, \mathbb{C})$ is formed by the diagonal matrices such that if x_i is the eigenvalue of e_i, x_i^{-1} is the eigenvalue of f_i.

Besides the standard torus there are other symplectic matrices which preserve the basis by permuting it or changing sign. In particular

(i) The permutations S_n where $\sigma(e_i) = e_{\sigma(i)}, \quad \sigma(f_i) = f_{\sigma(i)}$.

(ii) For each i, the exchange ϵ_i which fixes all the elements except $\epsilon_i(e_i) = f_i, \epsilon_i(f_i) = -e_i$.

These elements form a semidirect product (or rather a wreath product) $S_n \ltimes \mathbb{Z}/(2)^n$.

The normalizer N_T of the standard torus T is the semidirect product of T with the Weyl group $S_n \ltimes \mathbb{Z}/(2)^n$.

Proof. If a symplectic transformation normalizes the standard torus it must permute its eigenspaces. Moreover, if it maps the eigenspace of α to that of β, it maps the eigenspace of α^{-1} to that of β^{-1}. Therefore the group of permutations on the set of the $2n$ eigenspaces $\mathbb{C}e_i, \mathbb{C}f_i$ induced by the elements of N_T is the group $S_n \ltimes \mathbb{Z}/(2)^n$, hence every element of N_T can be written in a unique way as the product of an element of $S_n \ltimes \mathbb{Z}/(2)^n$ and a diagonal matrix that is an element of T. $\square$

It is important to remark that the same analysis applies to the compact form of the symplectic group $Sp(n, \mathbb{H}) = Sp(2n, \mathbb{C}) \cap U(2n, \mathbb{C})$. The compact maximal torus T_c is formed by the unitary matrices of T which are just the diagonal matrices with $|\alpha_i| = 1$. The matrices in $S_n \ltimes \mathbb{Z}/(2)^n$ are unitary and we have the analogues of the previous facts for the compact group $Sp(n, \mathbb{H})$.

The normalizer of T acts on T by conjugation. In particular the Weyl group acts as $\sigma(x_i) = x_{\sigma(i)}$ and $\epsilon_i(x_j) = x_j,\ i \neq j,\ \epsilon_i(x_i) = x_i^{-1}$. One can thus look at the action on the coordinate ring of T. As in the theory of symmetric functions we can even work in an arithmetic way, considering the ring of Laurent polynomials $A := \mathbb{Z}[x_i, x_i^{-1}],\ i = 1, \ldots, n$, with the action of $S_n \ltimes \mathbb{Z}/(2)^n$.

When we look at the invariant functions we can proceed in two steps. First, look at the functions invariant under $\mathbb{Z}/(2)^n$. We claim that an element of A is invariant under $\mathbb{Z}/(2)^n$ if and only if it is a polynomial in the elements $x_i + x_i^{-1}$. This follows by simple induction. If R is any commutative ring, consider $R[x, x^{-1}]$. Clearly an element in $R[x, x^{-1}]$ is invariant under $x \mapsto x^{-1}$ if and only if it is of the form $\sum_j r_j(x^j + x^{-j})$. Now it suffices to prove by simple induction that, for every j, the element $x^j + x^{-j}$ is a polynomial with integer coefficients in $x + x^{-1}$ (*hint*: $x^j + x^{-j} = (x + x^{-1})^j +$ *lower terms*). The next step is to apply the symmetric group to the elements $x_i + x_i^{-1}$ and obtain:

Proposition 1. *The invariants of $\mathbb{Z}[x_i, x_i^{-1}],\ i = 1, \ldots, n$, under the action of $S_n \ltimes \mathbb{Z}/(2)^n$ form the ring of polynomials $\mathbb{Z}[e_1, e_2, \ldots, e_n]$ where the e_i are the elementary symmetric functions in the variables $x_i + x_i^{-1}$.*

Let us consider now invariant functions on $Sp(2n, \mathbb{C})$ under conjugation. Let us first comment on the characteristic polynomial $p(t) = t^{2n} + \sum_{i=1}^{2n}(-1)^i s_i t^{2n-i}$ of a symplectic matrix X with eigenvalues $x_1, \ldots, x_n, x_1^{-1}, \ldots, x_n^{-1}$.

It is $\prod_{i=1}^n (t - x_i)(t - x_i^{-1}) = \prod_{i=1}^n (t^2 - (x_i + x_i^{-1})t + 1)$. Thus it is a *reciprocal* polynomial, $t^{2n} p(t^{-1}) = p(t)$. It is easy to give explicit relations between the first n coefficients s_i of $t^{2n-i},\ i = 1, \ldots, n$, and the elementary symmetric functions e_i of the previous proposition, showing that the ring of invariants is also generated by the elements s_i. We deduce the following.

Proposition 2. *The ring of invariant polynomials under conjugation on $Sp(2n, \mathbb{C})$ is generated by the coefficients $s_i(X),\ i = 1, \ldots, n$, of the characteristic polynomial of X.*

Proof. One can give two simple proofs. One is essentially the same as in Chapter 2, §5.1. Another is as follows. Using the FFT of §8.2, an invariant of any matrix X under conjugation of the symplectic group is a polynomial in traces of monomials in X and X^s. If the matrix is in the symplectic group, these monomials reduce to $X^k,\ k \in \mathbb{Z}$, and then one uses the theory of symmetric functions we have developed. $\square$

We come now to the more interesting computation of characters. First, let us see how the Weyl integration formula of 9.1 takes shape in our case for the compact group $Sp(n, \mathbb{H})$ and its compact torus T_c. The method of the proof carries over with the following change. The covering now has $2^n n!$ sheets, and the Jacobian J expressing the change of measure is always given by the same method. One takes the Lie algebra $\mathfrak{k}$ of $Sp(n, \mathbb{H})$ which one decomposes as the direct sum of the Lie algebra of T_c and the T_c-invariant complement $\mathfrak{p}$ (under adjoint action). Then J is the determinant of $1 - \mathrm{Ad}(t)$, where $\mathrm{Ad}(t)$ is the adjoint action of T_c on $\mathfrak{p}$. Again, in order to compute this determinant, it is convenient to complexify $\mathfrak{p}$, obtaining the direct sum of all root spaces. Hence the Jacobian is the product $\prod(1 - \alpha(t))$ where $\alpha(t)$ runs over all the roots. As usual, now we are replacing the additive roots for the Cartan subalgebra with the multiplicative roots for the torus. Looking at the formulas 4.1.12 of Chapter 10 we see that the multiplicative roots are

$$(10.1.1) \qquad x_i x_j^{-1}, \quad (x_i x_j)^{\pm 1}, \ i \neq j, \qquad x_i^{\pm 2}.$$

$$J(x_1, \dots, x_n) = \prod_{i \neq j}(1 - x_i x_j^{-1})(1 - x_i x_j)(1 - (x_i x_j)^{-1}) \prod_{i=1}^{n}(1 - x_i^2)(1 - x_i^{-2}).$$

We collect the terms from the positive roots and obtain

$$\tilde{\Delta}(x_1, \dots, x_n) = \prod_{i<j}(1 - x_i x_j^{-1})(1 - x_i x_j) \prod_{i=1}^{n}(1 - x_i^2)$$
$$= \prod_{i<j} x_i (x_i + x_i^{-1} - (x_j + x_j^{-1})) \prod_{i=1}^{n}(-x_i)(x_i - x_i^{-1}).$$

When we compute J for the compact torus where $|x_i| = 1$, we have that the factors x_i cancel and $J = \Delta_C \overline{\Delta}_C$ with

$$\Delta_C(x_1, \dots, x_n) = \prod_{i<j}(x_i + x_i^{-1} - (x_j + x_j^{-1})) \prod_{i=1}^{n}(x_i - x_i^{-1}).$$

We have:

Theorem (Weyl's Integration Formula). *For a class function f on $Sp(n, \mathbb{H})$ we have*

$$(10.1.2) \qquad \int_{Sp(n,\mathbb{H})} f(g) d\mu = \frac{1}{2^n n!} \int_{T_c} f(x) \Delta_C(x) \overline{\Delta}_C(x) d\tau$$

where $\Delta_C(x_1, \dots, x_n) = \prod_{i<j}(x_i + x_i^{-1} - (x_j + x_j^{-1})) \prod_{i=1}^{n}(x_i - x_i^{-1})$.

Let us denote for simplicity by Q_n the group $S_n \ltimes \mathbb{Z}/(2)^n$. As for ordinary symmetric functions, we also have a sign character *sign* for the group Q_n, which is the ordinary sign on S_n, and $sgn(\epsilon_i) = -1$. It is the usual determinant when we consider the group in its natural reflection representation. We can consider again polynomials which are antisymmetric, i.e., which transform as this sign representation. Clearly $\Delta_C(x_1, \dots, x_n)$ is antisymmetric. The same argument given in Chapter 2, 3.1 shows that:

Proposition 3.

(1) A Q_n-antisymmetric Laurent polynomial $f(x_1, \dots, x_n, x_1^{-1}, \dots, x_n^{-1})$ is of the form $g(x)\Delta_C(x)$ with $g(x)$ symmetric.

(2) A basis of antisymmetric polynomials is given by the polynomials:

$$A^C_{\ell_1,\ell_2,\dots,\ell_n} := \sum_{\sigma \in Q_n} sign(\sigma)\sigma(x_1^{\ell_1} x_2^{\ell_2} \dots x_n^{\ell_n}), \quad \ell_1 > \ell_2 > \dots > \ell_n > 0.$$

(3) $A^C_{n,n-1,n-2,\dots,1}(x) = \Delta_C(x)$.

(4) $A^C_{\ell_1,\ell_2,\dots,\ell_n}$ is the determinant of the matrix with $x_i^{\ell_j} - x_i^{-\ell_j}$ in the i, j position.

Proof. (1) Let $f(x_1, \dots, x_n, x_1^{-1}, \dots, x_n^{-1})$ be antisymmetric, and expand it as a sum $\sum_k a_k x_i^k + b_k x_i^{-k}$ where a_k, b_k do not depend on x_i. We then see that $a_0 = b_0 = 0$ and $a_k = -b_k$. This shows that f is divisible by $\prod_i (x_i - x_i^{-1})$. If we write $f(x) = \tilde{f}(x) \prod_i (x_i - x_i^{-1})$ we have that $\tilde{f}(x)$ is symmetric with respect to the group of exchanges $x_i \to x_i^{-1}$. Hence, by a previous argument, $\tilde{f}(x)$ is a polynomial in the variables $x_i + x_i^{-1}$. Looking at the action of S_n we notice that $\prod_i (x_i - x_i^{-1})$ is symmetric, hence $\tilde{f}(x)$ is still antisymmetric for this action. Then, by the usual theory, it is divisible by the Vandermonde determinant of the variables $x_i + x_i^{-1}$ which is the remaining factor of $\Delta_C(x)$.

(2) Since the group Q_n permutes the Laurent monomials, the usual argument shows that in order to give a basis of antisymmetric polynomials, we have to check for each orbit if the alternating sum on the orbit is 0.

In each orbit there is a unique monomial $M = x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$, $k_1 \geq k_2 \geq \dots \geq k_n \geq 0$. Unless all the inequalities are strict there is an odd element[116] of Q_n which fixes M. In this case M and its entire orbit cannot appear in an antisymmetric polynomial. The other orbits correspond to the elements $A^C_{\ell_1,\ell_2,\dots,\ell_n}$.

(3) By part (1) each $A^C_{\ell_1,\ell_2,\dots,\ell_n}$ is divisible by Δ_C, so δ, up to some constant, must equal the one of lowest degree. To see that the constant factor is 1 it is enough to remark that the leading monomial of Δ_C is indeed $x_1^n x_2^{n-1} \dots x_n$.

(4)
$$\begin{aligned} A^C_{\ell_1,\ell_2,\dots,\ell_n} &= \sum_{\sigma \in Q_n} sign(\sigma)\sigma(x_1^{\ell_1} x_2^{\ell_2} \dots x_n^{\ell_n}) \\ &= \sum_{\sigma \in S_n} sign(\sigma)\sigma\Bigg(\sum_{\tau \in \mathbb{Z}/(2)^n} sign(\tau)\tau(x_1^{\ell_1} x_2^{\ell_2} \dots x_n^{\ell_n})\Bigg) \\ &= \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^n (x_{\sigma(i)}^{\ell_i} - x_{\sigma(i)}^{-\ell_i}). \end{aligned}$$
□

[116] We use "odd" to mean that its sign is -1.

We can now define the analogues of the Schur functions for the symplectic group:

Proposition 4. *The elements*

$$S_\lambda^C(x) := A^C_{\ell_1,\ell_2,\dots,\ell_n}/\Delta_C(x), \quad \lambda := \ell_1 - n, \ell_2 - n + 1, \dots, \ell_n - 1$$

are an integral basis of the ring of Q_n-invariants of $\mathbb{Z}[x_i, x_i^{-1}]$.

The leading monomial of $S_\lambda^C(x)$ is $x_1^{\ell_1-n}x_2^{\ell_2-n+1}\dots x_n^{\ell_n-1}$.

Theorem 2. *The functions $S_\lambda^C(x)$ (as functions of the eigenvalues) are the irreducible characters of the symplectic group. $S_\lambda^C(x)$ is the character of the representation $T_\lambda(V)$ of §6.4.*

Proof. We proceed as for the unitary group. First, we prove that these characters are orthonormal. This follows immediately by the Weyl integration formula. Next we know that they are an integral basis of the invariant functions, hence we deduce that they must be $\pm$ the irreducible characters. Finally, a character is a positive sum of monomials, so by the computation of the leading term we can finally conclude that they must coincide with the characters.

Part (2) follows from the highest weight theory. In §6.4 the partition λ indexing the representation is the sequence of exponents of the character of the highest weight. □

10.2 Determinantal Formula

Let us continue this discussion developing an analogue for the functions $S_\lambda^C(x)$ of the determinantal formulas for Schur functions described in Chapter 9, §8.3.

Let us use the following notation. Given n-Laurent polynomials $f_i(x)$, in a variable x we denote by

$$\begin{aligned}|f_1(x), f_2(x), \dots, f_n(x)| &:= \det(f_i(x_j)) \\ &= \sum_{\sigma\in S_n} sgn(\sigma) f_1(x_{\sigma(1)}), f_2(x_{\sigma(2)}), \dots, f_n(x_{\sigma(n)}).\end{aligned}$$

The symbol introduced is antisymmetric and multilinear in the obvious sense. Moreover if $g(x)$ is another polynomial

$$\prod_{i=1}^n g(x_i)|f_1(x), f_2(x), \dots, f_n(x)| = |g(x)f_1(x), g(x)f_2(x), \dots, g(x)f_n(x)|.$$

Let us start from the step of the Cauchy formula transformed as

$$\frac{V(x)V(y)}{\prod_{i,j=1}^{n}(x_i - y_j)} = \det\left(\frac{1}{x_i - y_j}\right).$$

Substitute for x_i the element $x_i + x_i^{-1}$, and for y_i the element $y_i + y_i^{-1}$. Observe that

$$x_i + x_i^{-1} - y_j - y_j^{-1} = x_i^{-1}(1 - x_i y_j)(1 - x_i y_j^{-1})$$

and that the Vandermonde determinant in the elements $x_i + x_i^{-1}$ equals the determinant $|x^{n-1} + x^{-n+1}, x^{n-1} + x^{-n+2}, \dots, 1|$ to deduce the identity:

$$\prod_i x_i^{n-1} \frac{|x^{n-1} + x^{-n+1}, x^{n-2} + x^{-n+2}, \dots, 1|\,|y^{n-1} + y^{-n+1}, y^{n-2} + y^{-n+2}, \dots, 1|}{\prod_{i,j}(1 - y_j x_i)(1 - y_j^{-1} x_i)}$$

$$= \det\left(\frac{1}{(1 - y_j x_i)(1 - y_j^{-1} x_i)}\right)$$

$$= \frac{|x^{2n-2} + 1, x^{2n-3} + x, \dots, x^{n-1}|\,|y^{n-1} + y^{-n+1}, y^{n-1} + y^{-n+2}, \dots, 1|}{\prod_{i,j}(1 - y_j x_i)(1 - y_j^{-1} x_i)}.$$

Notice that

$$\frac{1}{(1 - yx)(1 - y^{-1}x)} = \sum_{k=0}^{\infty}(y^k + y^{k-2} + \dots + y^{-k})x^k = \sum_{k=0}^{\infty} \frac{y^{k+1} - y^{-k-1}}{y - y^{-1}} x^k.$$

Hence the determinant $\det\left|\frac{1}{(1-y_j x_i)(1-y_j^{-1} x_i)}\right|$ expands as

$$\text{(10.2.1)} \qquad \sum_{k_1 > k_2 > \dots > k_n \geq 0} A_{k_1,k_2,\dots,k_n}(x) A^C_{k_1+1,k_2+1,\dots,k_n+1}(y) \prod_{i=1}^{n} (y_i - y_i^{-1})^{-1}.$$

Next, consider the polynomial $\prod_{j=1}^{n}(1 - y_j z)(1 - y_j^{-1} z) = \det(1 - zA)$, where A is a symplectic matrix with eigenvalues y_i, y_i^{-1}, $i = 1, \dots, n$.

Let us define the symmetric functions $p_f(A) = p_f(y)$ by the formula

$$\text{(10.2.2)} \qquad \frac{1}{\det(1 - zA)} = 1 + \sum_{k=1}^{\infty} p_k(A) z^k.$$

Dividing both terms of the identity by $|y^{n-1} + y^{-n+1}, y^{n-2} + y^{-n+2}, \dots, 1|$ we get

$$\sum_{\lambda} A_{\lambda+\rho}(x)S_{\lambda}^{C}(y) = |x^{2n-2}+1, x^{2n-3}+x, \ldots, x^{n-1}| \prod_{i=1}^{n}\left(1+\sum_{k=1}^{\infty} p_k(y)x_i^k\right)$$

$$= \left|\left(1+\sum_{k=1}^{\infty} p_k(y)x^k\right)(x^{2n-2}+1), \left(1+\sum_{k=1}^{\infty} p_k(y)x^k\right)\right.$$

$$\left.\times(x^{2n-3}+x), \ldots, \left(1+\sum_{k=1}^{\infty} p_k(y)x^k\right)x^{n-1}\right|$$

$$= \sum_{k_1,k_2,\ldots,k_n} p_{k_1}p_{k_2}\cdots p_{k_n}|x^{k_1}(x^{2n-2}+1), x^{k_2}(x^{2n-3}+x), \ldots, x^{k_n}x^{n-1}|$$

$$= \sum_{k_1,k_2,\ldots,k_n} p_{k_1}p_{k_2}\cdots p_{k_n}|x^{k_1+2n-2}+x^{k_1}, x^{k_2+2n-3}+x^{k_2+1}, \ldots, x^{k_n+n-1}|$$

$$= \sum_{k_1,k_2,\ldots,k_n} p_{k_1-2n+2}p_{k_2}\cdots p_{k_n}|x^{k_1}, x^{k_2+2n-3}+x^{k_2+1}, \ldots, x^{k_n+n-1}|$$

$$+ \sum_{k_1,k_2,\ldots,k_n} p_{k_1}p_{k_2}\cdots p_{k_n}|x^{k_1}, x^{k_2+2n-3}+x^{k_2+1}, \ldots, x^{k_n+n-1}|$$

$$\times \sum_{k_1,k_2,\ldots,k_n} (p_{k_1-2n+2}+p_{k_1})p_{k_2}\cdots p_{k_n}|x^{k_1}, x^{k_2+2n-3}+x^{k_2+1}, \ldots, x^{k_n+n-1}|.$$

Iterating the procedure,

$$\sum_{k_1,k_2,\ldots,k_n} (p_{k_1-2n+2}+p_{k_1})(p_{k_2-2n+3}+p_{k_2-1})\cdots$$

$$(p_{k_i-2n+i}+p_{k_i-i+2})\cdots p_{k_n-n+1}|x^{k_1}, \ldots, x^{k_n}|$$

$$= \sum_{k_1>k_2>\ldots>k_n} |p_{k_1-2n+2}+p_{k_1}, p_{k_2-2n+3}+p_{k_2-1}, \ldots, p_{k_i-2n+i}$$

$$+ p_{k_i-i+2}, \ldots, p_{k_n-n+1}||x^{k_1}, \ldots, x^{k_n}|,$$

where the symbol $|p_{k_1-2n+2}+p_{k_1}, p_{k_2-2n+3}+p_{k_2-1}, \ldots, p_{k_i-2n+i+1}+p_{k_i-i+1}, \ldots, p_{k_n-n+1}|$ equals the determinant of the matrix in which the element in the i^{th} row and j^{th} column is $p_{k_i-2n+j}+p_{k_i-j+2}$.

Thus, with the notations of partitions, the function $S_{\lambda}^{C}(y)$ with $\lambda = h_1, h_2, \ldots, h_n$ corresponds to the coefficient where $k_1 = h_1+n-1, k_2 = h_2+n-2, \ldots k_n = h_n$ and so it is

$$S_{\lambda}^{C}(y) =$$
$$|p_{h_1-n+1}+p_{h_1+n-1}, p_{h_2-n+1}+p_{h_2+n-3}, \ldots, p_{h_i-n+1}+p_{h_i+n-2i+1}, \ldots, p_{h_n-n+1}|.$$

Notice that when we evaluate for $y_i = 1$:

(10.2.3)

$$1+\sum_{k=1}^{\infty} p_k(1)z^k = \frac{1}{\det(1-zI_{2n})} = \frac{1}{(1-z)^{2n}} = 1+\sum_{k=1}^{\infty}\binom{2n+k-1}{k} z^k$$

gives a determinantal expression in terms of binomial coefficients for $S_\lambda^C(1) = \dim T_\lambda(V)$.

10.3 The Spin Groups: Odd Case

Let $\mathrm{Spin}(2n+1)$ be the compact spin group.

Let us work with the usual hyperbolic basis e_i, f_i, u. Consider the improper orthogonal transformation $J : J(e_i) = e_i, J(f_i) = f_i, J(u) = -u$, so that $O(V) = SO(V) \cap SO(V)J$.

Its roots are $\Phi^+ := \alpha_i - \alpha_j, i < j,\ \alpha_i + \alpha_j,\ i \neq j, \alpha_i$, cf. Chapter 10, §4.1.5. Its Weyl group is a semidirect product $S_n \ltimes \mathbb{Z}/(2)^n$ as for the symplectic group.

The character group X of the corresponding maximal torus T of the special orthogonal group is the free group in the variables x_i. X must be of index 2 in the character group X' of the corresponding torus T' for the spin group, since T' is the preimage of T in the spin group, hence it is a double covering of T. X' contains the fundamental weight of the spin representation which we may call s with $s^2 = \prod_i x_i$, so it is convenient to consider X as the subgroup of the free group in variables $x_i^{1/2}$ with the usual action of the Weyl group, generated by the elements x_i and $s = (\prod_{i=1}^n x_i^{1/2})$. It is of course itself free on the variables $x_i,\ i < n,\ s$.

For the Weyl integration formula we obtain

(10.3.1)

$$\int_{O(V)} f(g)d\mu = \frac{1}{2^n n!}\left(\int_{T_c} f(x)\Delta_B(x)\overline{\Delta}_B(x)d\tau + \int_{T_cJ} f(x)\Delta_B(x)\overline{\Delta}_B(x)d\tau\right),$$

where $\Delta_B(x_1,\dots,x_n) = \prod_{i<j}(x_i + x_i^{-1} - (x_j + x_j^{-1}))\prod_i(x_i^{1/2} - x_i^{-1/2})$. A function

$$A^B_{\ell_1,\ell_2,\dots,\ell_n} := \sum_{\sigma\in Q_n} sign(\sigma)\sigma(x_1^{\ell_1}x_2^{\ell_2}\dots x_n^{\ell_n}), \qquad \ell_1 > \ell_2 > \cdots > \ell_n > 0.$$

is in the character group X' if the ℓ_i are either all integers or all half integers.

$\Delta_B(x_1,\dots,x_n) = \prod_{i<j}(x_i + x_i^{-1} - (x_j + x_j^{-1}))\prod_i(x_i^{1/2} - x_i^{-1/2})$ is skew-symmetric with leading term $x_1^{n-1/2}\dots x_n^{1/2}$ so $\Delta_B = A^B_{n-1/2,n-1-1/2,\dots,1/2}$. Notice that

$$\Delta_B\prod_i(x_i^{1/2} + x_i^{-1/2}) = \prod_{i<j}(x_i + x_i^{-1} - (x_j + x_j^{-1}))\prod_i(x_i - x_i^{-1}),$$

which is the formula we found for the symplectic group (10.1). If the ℓ_i are all integers, then $A^B_{\ell_1,\ell_2,\dots,\ell_n}$ is divisible by $\Delta_B\prod_i(x_i^{1/2} + x_i^{-1/2})$ with the quotient being a linear combination of characters all with integer exponents. Otherwise, if all the ℓ_i are half integers, we multiply by $\prod_i x_i^{1/2}$ and still see that all the A^B are multiples of Δ_B. We get a theorem similar to the one for the symplectic group.

Proposition. *The elements*

$$S^B_\lambda(x) := A^B_{\ell_1,\ell_2,\dots,\ell_n}/\Delta_B(x), \quad \lambda := \ell_1 - n, \ell_2 - n + 1, \dots, \ell_n - 1$$

are an integral basis of the ring of Q_n invariants of $\mathbb{Z}[x_i, x_i^{-1}, \prod_i x_i^{1/2}]$.

The leading monomial of $S^B_\lambda(x)$ is $x_1^{\ell_1-n+1/2}x_2^{\ell_2-n+3/2}\dots x_n^{\ell_n-1/2}$.

Theorem. *The functions $S^B_\lambda(x)$ (as functions of the eigenvalues) are the irreducible characters of the spin group. The functions $S^B_\lambda(x)$ with integral weight are the irreducible characters of the special orthogonal group.*

If λ is integral, $S^B_\lambda(x)$ is the character of the representation $T_\lambda(V)$ of §6.6.

Proof. Same as for the symplectic group. □

10.4 The Spin Groups: Even Case

Let Spin$(2n)$ be the compact spin group.

Its positive roots are $\Phi^+ := \alpha_i - \alpha_j, i < j,\ \alpha_i + \alpha_j,\ i \neq j$, Chapter 10, §4.1.19. Its Weyl group is a semidirect product $S_n \ltimes S$. S_n is the symmetric group permuting the coordinates. S is the subgroup of the *sign group* $\mathbb{Z}/(2)^n := (\pm 1, \pm 1, \dots, \pm 1)$ which changes the signs of the coordinates, formed by only even number of sign changes.

For the Weyl integration formula we obtain

$$\text{(10.4.1)} \qquad \int_{\mathrm{Spin}(2n)} f(g)d\mu = \frac{1}{2^{n-1}n!}\int_{T_c} f(x)\Delta^D(x)\overline{\Delta}^D(x)d\tau,$$

where $\Delta^D(x_1, \dots, x_n) = \prod_{i<j}(x_i + x_i^{-1} - (x_j + x_j^{-1}))$, with leading monomial $x_1^{n-1}x_2^{n-2}\dots x_{n-1}$.

The character ring of the maximal torus of the special orthogonal group is the polynomial ring $\mathbb{Z}[x_i^{\pm 1}]$, while that of the spin group has also the element $\prod_i x_i^{1/2}$.

Now it is not always true in the orbit of a monomial under the action of the Weyl group W there is a monomial with all exponents positive. Thus we have a different notion of leading monomial. In fact it is easy to verify that, given $h_1 > h_2 > \dots > h_n \geq 0$, if $h_n > 0$ the two monomials $x_1^{h_1}x_1^{h_2}\dots x_n^{h_n}$ and $x_1^{h_1}x_1^{h_2}\dots x_{n-1}^{h_{n-1}}x_n^{-h_n}$ are in two different orbits and leading in each of them. In the language of weights, an element is leading in an orbit if it is in the fundamental chamber (cf. Chapter 10, Theorem 2.4). Therefore each of these two elements will be the leading term of an antisymmetric function. Dividing these functions by $\Delta^D(x_1, \dots, x_n)$ we get the list of the irreducible characters. We can understand to which representations they correspond using the theory of highest weights.

10.5 Weyl's Character Formula

As should be clear from the examples, Weyl's character formula applies to all simply connected semisimple groups, as well as the integration formula for the corresponding compact groups. In the language of roots and weights let us use the notation that if α is a weight of the Cartan subalgebra, e^α is the corresponding weight for the torus according to the formula $e^\alpha(e^t) = e^{\alpha(t)}$, so that $e^{\alpha+\beta} = e^\alpha e^\beta$.

The analogue of Δ used in the integration formula is

$$(10.5.1) \qquad \rho := \sum_{\alpha\in\Phi^+} \alpha/2, \qquad \Delta := e^\rho \prod_{\alpha\in\Phi^+}(1-e^{-\alpha}) = \prod_{\alpha\in\Phi^+}(e^{\alpha/2}-e^{-\alpha/2}).$$

We have seen (Chapter 10, §2.4) the notions of weight lattice, dominant, regular dominant and fundamental weight. We denote by Λ the weight lattice. Λ is identified with the character group of the maximal torus T of the associated simply connected group. We also denote by Λ^+ the *dominant weights*, Λ^{++} the *regular dominant weights* and ω_i the *fundamental weights*.

The character ring for T is the group ring $\mathbb{Z}[\Lambda]$. Choosing as generators the fundamental weights ω_i this is also a Laurent polynomial ring in the variables e^{ω_i}.

From Chapter 10, Theorems 2.4 and 2.3 (6) it follows that there is one and only one dominant weight in the orbit of any weight. The stabilizer in W of a dominant weight $\sum_{i=1}^n a_i\omega_i$ is generated by the simple reflections s_i for which $a_i = 0$. Let $\rho = \sum_i \omega_i$.

It follows that the skew-symmetric elements of the character ring have a basis obtained by antisymmetrizing the dominant weights for which all $a_i > 0$, i.e., the regular dominant weights. Observe that ρ is the minimal regular dominant weight and we have the 1–1 correspondence $\lambda \mapsto \lambda + \rho$ between dominant and regular dominant weights.

We have thus a generalization of the ordinary theory of skew-symmetric polynomials:

Proposition 1. *The group $\mathbb{Z}[\Lambda]^-$ of skew-symmetric elements of the character ring has as basis the elements*

$$A_{\lambda+\rho} := \sum_{w\in W} sign(w) e^{w(\lambda+\rho)}, \qquad \lambda \in \Lambda^+.$$

First, we want to claim that $A_{\lambda+\rho}$ has the *leading term* $e^{\lambda+\rho}$.

Lemma 1. *Let λ be a dominant weight and w an element of the Weyl group. Then $\lambda - w(\lambda)$ is a linear combination with positive coefficients of the positive roots $s_{i_1}s_{i_2}\dots s_{i_{h-1}}(\alpha_{i_h})$ sent to negative roots by w^{-1}. If $w \neq 1$ and λ is regular dominant, this linear combination is nonzero and $w(\lambda) \prec \lambda$ in the dominance order.*

Proof. We will use the formula of Chapter 10, §2.3 with $w = s_{i_1}s_{i_2}\dots s_{i_k}$ a reduced expression, $u := s_{i_2}\dots s_{i_k}$, $\gamma_h := s_{i_2}\dots s_{i_{h-1}}(\alpha_{i_h})$, $\beta_h := s_{i_1}s_{i_2}\dots s_{i_{h-1}}(\alpha_{i_h})$. We have by induction

$$\begin{aligned}\lambda - w(\lambda) &= \lambda - s_{i_1}u(\lambda) = \lambda + s_{i_1}(\lambda - u(\lambda)) - s_{i_1}\lambda \\ &= \lambda + s_{i_1}\Big(\sum_h n_h\gamma_h\Big) - \lambda + \langle\lambda\,|\,\alpha_{i_1}\rangle\alpha_{i_1} \\ &= \sum_h n_h\beta_h + \langle\lambda\,|\,\alpha_i\rangle\alpha_i.\end{aligned}$$

Since λ is dominant, $\langle\lambda\,|\,\alpha_i\rangle \geq 0$. If, moreover, λ is regular $\langle\lambda\,|\,\alpha_i\rangle > 0$. □

In order to complete the analysis we need to be able to work as we did for Schur functions. Notice that choosing a basis of the weight lattice, the ring $\mathbb{Z}[\Lambda] = \mathbb{Z}[x_i^{\pm 1}]$, $i = 1, \dots, n$, is a unique factorization domain. Its invertible elements are the elements $\pm e^\lambda$, $\lambda \in \Lambda$. It is immediate that the elements $1 - x_i, 1 - x_i^{-1}$ are irreducible elements.

Lemma 2. *For any root α the element $1 - e^{-\alpha}$ is irreducible in $\mathbb{Z}[\Lambda]$.*

Proof. Under the action of the Weyl group every root α is equivalent to a simple root; thus we can reduce to the case in which $\alpha = \alpha_1$ is a simple root.

We claim that α_1 is part of an integral basis of the weight lattice. This suffices to prove the lemma. In fact let ω_i be the fundamental weights (a basis). Consider $\alpha_1 = \sum_j \langle \alpha_1, \alpha_i \rangle \omega_j$ (Chapter 10, §2.4.2). By inspection, for all Cartan matrices we have that each column is formed by a set of relatively prime integers. This implies that α_1 can be completed to an integral basis of the weight lattice. □

One clearly has $s_i(\Delta) = -\Delta$ which shows the antisymmetry of Δ. Moreover one has:

Proposition 2. $\mathbb{Z}[\Lambda]^- = \mathbb{Z}[\Lambda]^W \Delta$ *is the free module over the ring of invariants* $\mathbb{Z}[\Lambda]^W$ *with generator* Δ.

(10.5.2)
$$\Delta = \sum_{w \in W} sign(w) e^{w(\rho)} = A_\rho.$$

The elements $S_\lambda := A_{\lambda+\rho}/\Delta$ form an integral basis of $\mathbb{Z}[\Lambda]^W$.

Proof. Given a positive root α let s_α be the corresponding reflection. By decomposing W into cosets with respect to the subgroup with two elements 1 and s_α, we see that the alternating function $A_{\lambda+\rho}$ is a sum of terms of type $e^\mu - s_\alpha(e^\mu) = e^\mu - e^{\mu - \langle \mu | \alpha \rangle \alpha} = e^\mu(1 - e^{-\langle \mu | \alpha \rangle \alpha})$. Since $\langle \mu | \alpha \rangle$ is an integer, it follows that $(1 - e^{-\langle \mu | \alpha \rangle \alpha})$ is divisible by $1 - e^{-\alpha}$; since all terms of $A_{\lambda+\rho}$ are divisible, $A_{\lambda+\rho}$ is divisible by $1 - e^{-\alpha}$ as well. Now if $\alpha \neq \beta$ are positive roots, $1 - e^{-\alpha}$ and $1 - e^{-\beta}$ are not associate irreducibles, otherwise $1 - e^{-\alpha} = \pm e^\lambda(1 - e^{-\beta})$ implies $\alpha = \pm\beta$. It follows that each function $A_{\lambda+\rho}$ is divisible by Δ and the first part follows as for symmetric functions.

The second formula $\Delta = A_\rho$ can be proved as follows. In any skew-symmetric element there must appear at least one term of type e^λ with λ strongly dominant. On the other hand, ρ is minimal in the dominance order among strongly dominant weights. The alternating function $\Delta - A_\rho$ does not contain any such term, hence it is 0. The last statement follows from the first. □

Now let us extend the leading term theory also to the S_λ.

Proposition 3. $S_\lambda = e^\lambda$ *plus a linear combination of elements e^μ where μ is less than λ in the dominance order.*

Proof. We know the statement for $A_{\lambda+\rho}$, so it follows for

$$S_\lambda = A_{\lambda+\rho}/\Delta = A_{\lambda+\rho}e^{-\rho}\prod_{\alpha\in\Phi^+}(1-e^{-\alpha})^{-1} = A_{\lambda+\rho}e^{-\rho}\prod_{\alpha\in\Phi^+}\left(\sum_{k=0}^{\infty}e^{-k\alpha}\right). \qquad \square$$

The analysis culminates with:

Weyl's Character Formula. *The element S_λ is the irreducible character of the representation with highest weight λ.*

Proof. The argument is similar to the one developed for $U(n,\mathbb{C})$. First, we have the orthogonality of these functions. Let G be the corresponding compact group, T a maximal torus of G. Apply Weyl's integration formula:

$$\int_G S_\lambda\overline{S}_\mu d\mu = \frac{1}{|W|}\int_T S_\lambda\overline{S}_\mu\Delta\overline{\Delta}d\tau = \frac{1}{|W|}\int_T A_{\lambda+\rho}\overline{A}_{\mu+\rho}d\tau = \delta_\lambda^\mu.$$

The last equality comes from the fact that if $\lambda \neq \mu$, the two W-orbits of $\lambda+\rho$, $\mu+\rho$ are disjoint, and then we apply the usual orthogonality of characters for the torus.

Next, we know also that irreducible characters are symmetric functions of norm 1. When we write them in terms of the S_λ, we must have that an irreducible character must be $\pm S_\lambda$. At this point we can use the highest weight theory and the fact that e^λ with coefficient 1 is the leading term of S_λ and also the leading term of the irreducible character of the representation with highest weight λ. To finish, since all the dominant weights are highest weights the theorem is proved. $\square$

We finally deduce also:

Weyl's Dimension Formula. *The value $S_\lambda(1)$, dimension of the irreducible representation with highest weight λ, is*

$$\prod_{\alpha\in\Phi^+}\frac{(\lambda+\rho,\alpha)}{(\rho,\alpha)} = \prod_{\alpha\in\Phi^+}\left(1+\frac{(\lambda,\alpha)}{(\rho,\alpha)}\right). \tag{10.5.3}$$

Proof. Remark first that if $\alpha = \sum_i n_i\alpha_i \in \Phi^+$, we have $(\rho,\alpha) = \sum_i n_i(\rho,\alpha_i) > 0$, so the formula makes sense. One cannot evaluate the fraction of the Weyl Character formula directly at 1. In fact the denominator vanishes exactly on the nonregular elements. So we compute by the usual method of calculus. We take a regular vector $v \in \mathfrak{t}$ and compute the formula on $\exp(v)$, then take the limit for $v \to 0$ by looking at the linear terms in the numerator and denominator. By definition, $e^\beta(e^v) = e^{\beta(v)}$, so we analyze the quotient of (use 10.5.1):

$$\sum_{w\in W} sign(w)e^{w(\lambda+\rho)(v)} \text{ and } \sum_{w\in W} sign(w)e^{w(\rho)(v)} = e^{\rho(v)}\prod_{\alpha\in\Phi^+}(1-e^{-\alpha(v)}).$$

By duality, we work directly on the root space and analyze equivalently the quotient of

$$\sum_{w\in W} sign(w)e^{(w(\lambda+\rho),\beta)} \text{ and } \sum_{w\in W} sign(w)e^{(w(\rho),\beta)} = e^{(\rho,\beta)} \prod_{\alpha\in\Phi^+} (1-e^{-(\alpha,\beta)}).$$

Take $\beta = s\rho$. We have $\sum_{w\in W} sign(w)e^{(w(\lambda+\rho),s\rho)} = \sum_{w\in W} sign(w)e^{s(\lambda+\rho,w(\rho))}$. Now substituting for $\beta = s(\lambda+\rho)$ in the second identity we get

$$\begin{aligned}\sum_{w\in W} sign(w)e^{s(\lambda+\rho,w(\rho))} &= \sum_{w\in W} sign(w)e^{(w(\rho),s(\lambda+\rho))} \\ &= e^{s(\rho,\lambda+\rho)} \prod_{\alpha\in\Phi^+} (1-e^{-s(\alpha,\lambda+\rho)}).\end{aligned}$$

The character computed in $s\rho$ is thus

$$\frac{e^{s(\rho,\lambda+\rho)}}{e^{(\rho,s\rho)}} \prod_{\alpha\in\Phi^+} \frac{(1-e^{-s(\alpha,\lambda+\rho)})}{(1-e^{-(\alpha,s\rho)})}.$$

Its limit when $s \to 0$ is clearly $\prod_{\alpha\in\Phi^+} \frac{(\lambda+\rho,\alpha)}{(\rho,\alpha)}$. □

A final remark. The computation of the character on a maximal torus determines the character on the entire group. This is obvious for the compact form since every element is conjugate to one in the maximal torus. For the algebraic group we have seen in Chapter 10, §6.7 that the set of elements conjugate to one in the maximal torus is dense and we have proved that the ring of regular functions on a linearly reductive group, invariant under conjugation, has as basis the irreducible characters.