1 Finite Groups

1.1 Generalities

Let V be a vector space over $\mathbb{C}$, and let $GL(V)$ denote the group of all isomorphisms on V. For a fixed basis of V, $GL(V)$ is isomorphic to the group $GL_n(\mathbb{C})$, the group of all $n \times n$ invertible matrices.

Let G be a group and $\rho : G \to GL(V)$ be a representation. We also denote this by the tuple (ρ, V) or say that V is a G-module. Let $Z \subseteq V$ be a subspace such that $\rho(g)(Z) \subseteq Z$ for all $g \in G$. Then, we say that Z is an **invariant subspace**. We say that (ρ, V) is **irreducible** if there is no proper subspace $W \subset V$ such that $\rho(g)(W) \subseteq W$ for all $g \in G$. We say that (ρ, V) is **indecomposable** is there is no expression $V = W_1 \oplus W_2$ such that $\rho(g)(W_i) \subseteq W_i$, for all $g \in G$.

For a point $v \in V$, the **orbit** $O(v)$, and the **stabilizer** $Stab(v)$ are defined as:

$$\begin{array}{rcl} O(v) & = & \{v' \in V | \exists g \in G \, with \, \rho(g)(v) = v'\} \\ Stab(v) & = & \{g \in G | \rho(g)(v) = v\} \end{array}$$

One may also define $v \sim v'$ if there is a $g \in G$ such that $\rho(g)(v) = v'$. It is then easy to show that $[v]_\sim = O(v)$.

Let V^* be the dual-space of V. The representation (ρ, V) induces the **dual** representation (ρ^*, V^*) defined as $\rho^*(v^*)(v) = v^*(\rho(g^{-1})(v))$. It will be convenient for ρ^* to act on the right, i.e., $((v^*)(\rho^*))(v) = v^*(\rho(g^{-1})(v))$.

When ρ is fixed, we abbrieviate $\rho(g)(v)$ as just $g \cdot v$. Along with this, there are the standard constructions of the **tensor** $T^d(V)$, the **symmetric power** $Sym^d(V)$ and the **alternating power** $\wedge^d(V)$ representations.

Of special significance is $Sym^d(V^*)$, the space of homogeneous polynomial functions on V of degree d. Let $dim(V) = n$ and let $X_1, \ldots, X_n$ be a basis of V^*. We define as follows:

$$R = \mathbb{C}[X_1, \ldots, X_n] = \oplus_{d=0}^{\infty} R_d = \oplus_{d=0}^{\infty} Sym^d(V^*)$$

Thus R is the ring of all polynomial functions on V and is isomorphic to the algebra (over $\mathbb{C}$) of n indeterminates. Since G acts on the domain V, G also acts on all functions $f : V \to \mathbb{C}$ as follows:

$$(f \cdot g)(v) = f(g^{-1} \cdot v)$$

This action of G on all functions extends the action of G on polynomial functions above. Indeed, for any $g \in G$, the map $t_g : R \to R$ given by $f \to f \cdot g$ is an algebra isomorphism. This is called the **translation map**.

For an $f \in R$, we say that f is an **invariant** if $f \cdot g = f$ for all $g \in G$. The following are equivalent:

- $f \in R$ is an invariant.
- $Stab(f) = G$.

- $f(g \cdot v) = f(v)$ for all $g \in G$ and $v \in V$.
- For all v, v' such that $v' \in Orbit(v)$, we have $f(v) = f(v')$.

If W_1 and W_2 are two modules of G and $\phi : W_1 \to W_2$ is a linear map such that $g \cdot \phi(w_1) = \phi(g \cdot w_1)$ for all $g \in G$ and $w_1 \in W_1$ then we say that ϕ is **G-equivariant** or that ϕ is a **morphism of G-modules**.

1.2 The finite group action

Let G be a finite group and (μ, W) be a representation.

Recall that a **complex inner product** on W is a map $h : W \times W \to \mathbb{C}$ such that:

- $h(\alpha w + \beta w', w'') = \overline{\alpha} h(w, w'') + \overline{\beta} h(w', w'')$ for all $\alpha, \beta \in \mathbb{C}$ and all $w, w', w'' \in W$.
- $h(w'', \alpha w + \beta w') = \alpha h(w'', w) + \beta h(w'', w')$ for all $\alpha, \beta \in \mathbb{C}$ and all $w, w', w'' \in W$.
- $h(w, w) > 0$ for all $w \neq 0$.

Also recall that if $Z \subseteq W$ is a subspace, then $Z^\perp$ is defined as:

$$Z^\perp = \{w \in W | h(w, z) = 0 \, \forall z \in Z\}$$

Also recall that $W = Z \oplus Z^\perp$.

We say that an inner product h is **G-invariant** if $h(g \cdot w, g \cdot w') = h(w, w')$ for all $w, w' \in W$ and $g \in G$.

Proposition 1 *Let W be as above, and Z be an invariant subspace of W. Then $Z^\perp$ is also an invariant subspace. Thus every reducible representation of G is also decomposable.*

Proof: Let $x \in Z^\perp, z \in Z$ and let us examine $(g \cdot x, z)$. Applying g^{-1} to both sides, we see that:

$$h(g \cdot x, z) = h(g^{-1} \cdot g \cdot x, g^{-1} \cdot z) = h(x, g^{-1} \cdot z) = 0$$

Thus, G preserves $Z^\perp$ as claimed. $\square$

Let h be a complex inner product on W. We define the inner product h^G as follows:

$$h^G(w, w') = \frac{1}{|G|} \sum_{g' \in G} h(g' \cdot w, g' \cdot w')$$

Lemma 2 *h^G is a G-invariant inner product.*

Proof: First we see that

$$h^G(w, w) = \frac{1}{|G|} \sum_{g' \in G} h(w', w')$$

where $w' = g' \cdot w$. Thus $h^G(w, w) > 0$ unless $w = 0$. Secondly, by the linearity of the action of G, we see that h^G is indeed an inner product. Finally, we see that:

$$h^G(g \cdot w, g \cdot w') = \frac{1}{|G|} \sum_{g' \in G} h(g' \cdot g \cdot w, g' \cdot g \cdot w')$$

Since as g' ranges over G, so will $g' \cdot g$ for any fixed g, we have that h^G is G-invariant. □

Theorem 3
- *Let G be a finite group and (ρ, V) be an indecomposable representation, then it is also irreducible.*
- *Every representation (ρ, V) may be decomposed into irreducible representations V_i. Thus $V = \oplus_i V_i$, where (ρ_i, V_i) is an irreducible representation.*

Proof: Suppose that $Z \subseteq V$ is an invariant subspace, then $V = Z \oplus Z^\perp$ is a non-trivial decomposition of V contradicting the hypothesis. The second part is proved by applying the first, recursively. □

We have seen the operation of **averaging over the group** in going from the inner product h to the G-invariant inner product h^G. A similar approach may be used for constructing invariant polynomials functions. So let $p(X) \in R = \mathbb{C}[X_1, \ldots, X_n]$ be a polynomial function. We define the function $p^G : V \to \mathbb{C}$ as:

$$p^G(v) = \frac{1}{|G|} \sum_{g \in G} p(g \cdot v)$$

The transition from p to p^G is called the **Reynold's operator**.

Proposition 4 *Let $p \in R$ be of degree atmost d, then p^G is also a polynomial function of degree atmost d. Next, p^G is an invariant.*

Let R^G denote the set of all invariant polynomial functions on the space V. It is easy to see that $R^G \subseteq R$ is actually a *subring* of R.

Let $Z \subseteq V$ be an arbitrary subset of V. We say that Z is G-closed if $g \cdot z \in Z$ for all $g \in G$ and $z \in Z$. Thus Z is a union of orbits of points in V.

Lemma 5 *Let $p \in R^G$ be an invariant and let $Z = V(p)$ be the variety of p. Then Z is G-closed.*

We have already seen that $O(v)$, the orbit of v arises from the equivalence class $\sim$ on V. Since the group is finite, $|O(v)| \leq |G|$ for any v. Let O_1 and O_2 be disjoint orbits. It is essential to determine if elements of R^G can separate O_1 and O_2.

Lemma 6 *Let O_1 and O_2 be as above, and I_1 and I_2 be their ideals in R. Then there are $p_1 \in I_1$ and $p_2 \in I_2$ so that $p_1 + p_2 = 1$.*

Proof: This follows from the Hilbert Nullstellensatz. Since the point sets are finite, there is an explicit construction based on Lagrange interpolation. □

Let G be a finite group and (ρ, V) be a representation as above. We have see that this induces an action on $\mathbb{C}[X_1, \ldots, X_n]$. Also note that this action is **homogeneous**: for a $g \in G$ and $p \in R_d$, we have that $p \cdot g \in R_d$ as well. Thus R^G, the ring of invariants, is a *homogeneous subring* of R. In other words:

$$R^G = \oplus_{d=0}^{\infty} R_d^G$$

where R_d^G are invariants which are homogeneous of degree d. The existence of the above decomposition implies that every invariant is a sum of homogeneous invariants. Now, since $R_d^G \subseteq R_d$ as a vector space over $\mathbb{C}$. Thus

$$dim_{\mathbb{C}}(R_d^G) \leq dim_{\mathbb{C}}(R_d) \leq \binom{n+d-1}{n-1}$$

We define the **hilbert function** $h(R^G)$ of R^G (or for that matter, of any homogeneous ring) as:

$$h(R^G) = \sum_{d=0}^{\infty} dim_{\mathbb{C}}(R_d^G) z^d$$

We will see now that $h(R^G)$ is actually a rational function which is easily computed. We need a lemma.

Let (ρ, W) be a representation of the finite group G. Let

$$W^G = \{w \in W | g \cdot w = w\}$$

be the set of all vector invariants in W, and this is a subspace of W. .

Lemma 7 *Let (ρ, W) be as above. We have:*

$$dim_{\mathbb{C}}(W^G) = \frac{1}{|G|} \sum_{g \in G} trace(\rho(g))$$

Proof: Define $P = \frac{1}{|G|} \sum_{g \in G} \rho(g)$, as the average of the representation matrices. We see that $\rho(g) \cdot P = P \cdot \rho(g)$ and that $P^2 = P$. Thus P is diagonalizable and the eigenvalues of P are in the set $\{1, 0\}$. Let W_1 and W_0 be the corresponding eigen-spaces. It is clear that $W^G \subseteq W_1$ and that W_1 is fixed by each $g \in G$. We now argue that every $w \in W_1$ is actually an invariant. For that, let $w_g = g \cdot w$. We then have that $Pw = w$ implies that

$$w = \frac{1}{|G|} \sum_{g \in G} w_g$$

Note that a change-of-basis does not affect the hypothesis nor the assertion. We may thus assume that each $\rho(g)$ is unitary, we have that $w_g = w$ for all $g \in G$. Now, the claim follows by computing $trace(P)$. $\square$

We are now ready to state **Molien's Theorem**:

Theorem 8 *Let (ρ, W) be as above. We have:*

$$h(R^G) = \frac{1}{|G|}\sum_{g\in G}\frac{1}{det(I - z\rho(g))}$$

Proof: Let $dim_{\mathbb{C}}(W) = n$ and let $\{X_1, \ldots, X_n\}$ be a basis of W^*. Since $R^G = \sum_d R^G_d$ and each $R^G_d \subseteq \mathbb{C}[X_1, \ldots, X_n]_d$. Note that each $C[X_1, \ldots, X_n]_d$ is also a representation ρ_d of G. Furthermore, it is easy to see that if $\{\lambda_1, \ldots, \lambda_n\}$ are the eigenvalues of $\rho(g)$, then the eigen-values of the matrix $\rho_d(g)$ are precisely (including multiplicity)

$$\{\prod_i \lambda_i^{d_i} | \sum_i d_i = d\}$$

Thus

$$trace(\rho_d(g)) = \sum_{\overline{d}:|\overline{d}|=d}\prod_i \lambda_i^{d_i}$$

We then have:

$$\begin{array}{rcl} h(R^G) & = & \sum_d z^d dim_{\mathbb{C}}(R^G_d) \\ & = & \sum_d z^d[\frac{1}{|G|}\sum_g trace(\rho_d(g))] \\ & = & \frac{1}{|G|}\sum_g \frac{1}{(1-\lambda_1(g)z)\ldots(1-\lambda_n(g)z)} \\ & = & \frac{1}{|G|}\sum_g \frac{1}{det(I-z\rho(g))} \end{array}$$

This proves the theorem. □

1.3 The Symmetric Group

S_n will denote the symmetric group of all bijections on the set $[n]$. The standard representation of S_n is obviously on $V = \mathbb{C}^n$ with

$$\sigma \cdot (v_1, \ldots, v_n) = (v_{\sigma(1)}, \ldots, v_{\sigma(n)})$$

Thus, regarding V as column vectors, and S_n as the group of $n \times n$-permutation matrices, we see that the action of permutation P on vector v is given by the matrix multiplication $P \cdot v$.

Let $X_1, \ldots, X_n$ be a basis of V^*. S_n acts on $R = \mathbb{C}[X_1, \ldots, X_n]$ by $X_i \cdot \sigma = X_{\sigma(i)}$. The orbit of any point $v = (v_1, \ldots, v_n)$ is the collection of all permutation of the entries of the vector v and thus the size of the orbit is bounded by $n!$.

The invariants for this action are given by the elementary symmetric polynomials $e_k(X)$, for $k = 1, \ldots, n$, where

$$e_k(X) = \sum_{i_1<i_2<\ldots<i_k} X_{i_1}X_{i_2}\ldots X_{i_k}$$

Given two vector v and w, if $w \notin O(v)$, then there is a k such that $e_k(v) \neq e_k(w)$. This follows from the theory of equations in one variable.

The ring R^G equals $\mathbb{C}[e_1, \ldots, e_n]$ and has no algebraic dependencies. The hilbert function of R_G may then be expressed as:

$$h(R^G) = \frac{1}{(1-z)(1-z^2)\ldots(1-z^n)}$$

It is an exercise to verify that Molien's expression agrees with the above.

A related action of S_n is the **diagonal action**: Let $X = \{X_1, \ldots, X_n\}$, $Y = \{Y_1, \ldots, Y_n\}$, and so on upto $W = \{W_1, \ldots, W_n\}$ be a family of r (disjoint) variables. Let $B = \mathbb{C}[X, Y, \ldots, W]$ be the ring of polynomials in the variables of the disjoint union.

We define the action of S_n on $X \cup Y \cup \ldots \cup W$ as $X_i \cdot \sigma = X_{\sigma(i)}$, $Y_i \cdot \sigma = Y_{\sigma(i)}$, and so on.

The matrix equivalence of this action is the action of the permutation matrices on $n \times r$ matrices A, where the action of P on A is given by $P \cdot A$.

The invariants B^G is obtained from the $r = 1$ case by a *curious* operation: Let D_{XY} denote the operator:

$$D_{XY} = Y_1 \frac{\partial}{\partial X_1} + \ldots + Y_n \frac{\partial}{\partial X_n}$$

The ring B^G is obtained from R^G by applying the operators D_{XY}, D_{XW}, D_{WX} and so on, to elements of R^G. As an example, we have

$$e_2(X) = X_1X_2 + X_1X_3 + \ldots + X_{n-1}X_n$$

We have $D_{XY}(e_2)$ as:

$$D_{XY}(e_2(X)) = \sum_{i \neq j} X_i Y_j$$

This is clearly an element of B^G.

2 The Group SL_n

2.1 The Canonical Representation

Let V be a vector space of dimension n, and let $x_1, \ldots, x_n$ be a basis for V. Let $X_1, \ldots, X_n$ be the dual basis of V^*.

$SL(V)$ will denote the group of all unimodular linear transformations on V. In the above basis, this group is isomorphic to that of all $n \times n$ matrices of determinant 1, or in other words $SL_n(\mathbb{C})$.

The standard representation of $SL(V)$ is obviously V itself: Given $\phi \in SL(V)$ and $v \in V$, we have $\phi \cdot v = \phi(v)$ is the action of ϕ on v.

In terms of the basis x above we may write $v = [x_1, \ldots, x_n][\alpha_1, \ldots, \alpha_n]^T$ and thus $\phi \cdot v$ as $[\phi \cdot x_1, \ldots, \phi \cdot x_n][\alpha_1, \ldots, \alpha_n]^T$. If $\phi \cdot x_i = [x_1, \ldots, x_n][a_{1i}, \ldots, a_{ni}]^T$, then we have

$$\phi \cdot v = [x_1, \ldots, x_n] \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

We denote this matrix as A_ϕ.

We may now work with $SL_n(\mathbb{C})$ or simply SL_n. Given a vector $a = [\alpha_1, \ldots, \alpha_n]^T$, we see that the matrix multiplication $A \cdot a$ is the action of A on the column vector a.

Let us now understand the orbits of typical elements in the column space $\mathbb{C}^n$. The typical $v \in \mathbb{C}^n$ is a non-zero column vector. For any non-zero vector w, we see that there is an element $A \in SL_n$ such that $w = Av$. Furthermore, for any $B \in SL_n$, clearly $Bv \neq 0$. Thus we see that $\mathbb{C}^n$ has exactly two orbits:

$$O_0 = \{0\} \quad O_1\{v \in \mathbb{C}^n | v \neq 0\}$$

Note that O_1 is dense in $V = \mathbb{C}^n$ and its closure includes the orbit O_0 and therefore, the whole of V.

Let $R = \mathbb{C}[X_1, \ldots X_n]$ be the ring of polynomial functions on $\mathbb{C}^n$. We examine the action of SL_n on $\mathbb{C}[X]$. Recall that the action of A on X should be such that the evaluations $X_i(x_j) = \delta_{ij}$ must be conserved. Thus if the column vectors $x_i = [0, \ldots, 0, 1, 0 \ldots, 0]^T$ are the basis vectors of V and the row vectors $X_i = [0, \ldots, 0, 1, 0 \ldots, 0]$ that of V^*, then a matrix $A \in SL_n$ transforms x_i to Ax_i and X_j to $X_j A^{-1}$. Thus $X_j(x_i) = X_j / cdot x_i$ goes to $X_j A^{-1} A x_i = X_j \cdot x_i$.

Next, we examine $R = \mathbb{C}[X]$ for invariants. First note that the action of SL_n is homogeneous and thus we may assume that an invariant p is actually homogeneous of degree d. Next we see that p must be constant on O_1 and O_0. In this case, if $p(O_1) = \alpha$ then by the density of O_1 in V, we see that p is actually constant on V. Thus $R^G = R_0 = \mathbb{C}$.

2.2 The Diagonal Representation

Let us now consider the diagonal representation of the above representation. In other words, let V^r be the space of all complex $n \times r$ matrices x. The action of A on x is obvious $A \cdot x$. Let X be the $r \times n$-matrix dual to x. It transforms according to $X \to XA^{-1}$.

Let us examine the case when $r < n$ and compute the orbits in V^r. Let $x \in V^r$ be a matrix and y be a column vector such that $x \cdot y = 0$. We see that $(A \cdot x) \cdot y = 0$ as well. Thus if $ann(x) = \{y \in \mathbb{C}^r | x \cdot y = 0\}$ is the annihilator space of x, then we see that $ann(x) = ann(A \cdot x)$.

We show that the converse is also true:

Proposition 9 *Let $r < n$ and $x, x' \in V^r$ be such that $ann(x) = ann(x')$. Then there is an $A \in SL_n$ such that $x = A \cdot x'$.*

Proof: Use the row-echelon form construction. Make the pivots as 1 using appropriate diagonal matrices. Since $r < n$ these can be chosen from SL_n. □

This decides the orbit structure of V^r for $r < n$. The largest orbit O_r is of course when $ann(x) = 0$. Thus, when $x \in V_r$ is a mtrix of rank r, we see that $ann(x) = 0$ is trivial. The *generic* element of V^r is of this form. Thus O_r is dense in V^r. For this reason, there are no non-constant invariants.

Another calculation is the computation of the closure of orbits. Let O, O' be two arbitrary orbits of V^r. We have:

Proposition 10 *O' lies in the closure of O if and only if $ann(O') \supseteq ann(O)$.*

Proof: One direction is clear. We prove the other direction when $ann(O) \subseteq ann(O')$ and $dim(ann(O)) = dim(ann(O')) - 1$. Thus, we may assume that $O = [x]$ and $O' = [x']$ with both x and x' such that $rowspace(x) \supseteq rowspace(x')$ with $rank(x') = rank(x) - 1$. Then, upto SL_n, we may assume that the rows $x'[1], \ldots, x'[k]$ match the first k rows of x, and that $x'[k+1] = x'[k+2] = \ldots = x'[n] = 0$. Note that $x[k+1]$ is non-zero and $x[k+2]$ exists and is zero. We then construct the matrix $A(t)$ as follows:

$$A(t) = \left[\begin{array}{c|cc|c} I_k & 0 & 0 & 0 \\ \hline 0 & t & 0 & 0 \\ 0 & 0 & t^{-1} & 0 \\ \hline 0 & 0 & 0 & I_{n-k-2} \end{array} \right]$$

We see that $A(t) \in SL_n$ for all $t \neq 0$. Next, if we let $x(t) = A(t) \cdot x$, then we see that:

$$\lim_{t \to 0} x(t) = x'$$

This shows that x' lies in the closure of O. $\Box$

We now look at the case when $r \geq n$. We have:

Proposition 11 *Let $r \geq n$ and $x, x' \in V^r$ be such that $ann(x) = ann(x')$. If (i) $rank(x) < n$ then there is an $A \in SL_n$ such that $x' = Ax$, (ii) if $rank(x) = n$ there is a unique $A \in SL_n$ and a $\lambda \in \mathbb{C}^*$ such that if $z = A \cdot x$, then the first $n-1$ rows of z equal those of x' and $z[n] = \lambda x'[n]$.*

The proof is easy. Let x be matrix in V^r of rank n, and C be a subset of $[r] = \{1, 2, \ldots, r\}$ such that $det(x[C]) \neq 0$.

Proposition 12 *Let x be as above. Then $O(x)$, the orbit of x, equals all points $x' \in V^r$ such that (i) $ann(x) = ann(x')$ and (ii) $det(X'[C]) = det(x[C])$. The set of all rank n points in V^r is dense in V^r.*

The proof is easy.

Proposition 13 *The orbit $O(x)$ as above is closed.*

Proof: Notice that if $A \in SL_n$ and $z = Ax$ then $det(x[C]) = det(z[C])$. We may rewrite condition (i) above as $ann(x') \supseteq ann(x)$. Condition (ii) ensures that $rank(x') = n$ and that condition (i) holds with equality. Thus is $y_1, \ldots, y_{r-n}$ are column vectors generating $ann(x)$, then the equations $x'y_s = 0$ and $det(x'[C]) = det(x[C])$ determines the orbit $O(x)$. Thus $O(x)$ is the zero-set of some algebraic equations and thus is closed. $\Box$

Proposition 14 *Let $x \in V^r$ be such that $rank(x) < n$. Then (i) $O(x)$ equals those x' such that $ann(x) = ann(x')$, (ii) $O(x)$ is not closed and its closure $\overline{O(x)}$ equals points x' such that $ann(x') \supseteq ann(x)$.*

We now move to the computation of invariants. The space $\mathbb{C}[V^r]$ equals the space of all polynomials in the variable matrix $X = (X_{ij})$ where $i = 1, \ldots, n$ and $j = 1, \ldots, r$. It is clear that for any set C of n columns of X, we see that $det(X[C])$ is an invariant. We will denote C as $C = c_1 < c_2 < \ldots < c_n$ and $det(X[C])$ as p_C, the C-th **Plucker** coordinate. We aim to show now that these are the only invariants.

Let $C_0 = \{1 < 2 < \ldots < n\}$ and $W \subseteq V^r$ be the space of all matrices $x \in V^r$ such that $x[C_0] = diag(1, \ldots, 1, \lambda)$. Let W' be those elements of W for which $\lambda \neq 0$.

Lemma 15 *Let W' be as above. (i) If $x \in W'$, then $O(x) \cap W' = x$, (ii) for any $x \in V^r$ such that $det(x[C_0]) \neq 0$, there is a unique $A \in SL_n$ such that $Ax \in W'$.*

Let us call $Z' \subseteq V^r$ as those x such that $det(x[C_0]) \neq 0$. We then have the projection map

$$\pi : Z' \to W'$$

given by the above lemma. Note that Z' is SL_n-invariant: if $x \in Z'$ and $A \in SL_n$ then $A \cdot x \in Z'$ as well.

The ring $\mathbb{C}[Z']$ of regular functions on Z' is precisely $\mathbb{C}[X]_{det([X[C_0])}$, the localization of $\mathbb{C}[X]$ at $det(X[C_0])$.

We may parametrize W' as:

$$W = \left[\begin{array}{ccc|ccc} 1 & \ldots & 0 & w_{1,n+1} & \ldots & w_{1,r} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \ldots & w_{n,n} & w_{n,n+1} & \ldots & w_{n,r} \end{array} \right]$$

Let

$$W = \{W_{i,j} | i = 1, \ldots, n, \; j = n+1, \ldots, r\} \cup \{W_{n,n}\}$$

be a set of $(n-r) \cdot r + 1$ variables. The ring $\mathbb{C}[W']$ of regular function on W' is precisely $\mathbb{C}[W]_{W_{n,n}}$.

Let $x \in Z'$ be an arbitrary point and $A = x[C_0]$. Let $C_{i,j}$ be the set $C_0 - i + j$. The map π is given by:

$$\begin{array}{rcl} \pi(x)_{n,n} & = & det(A) \\ \pi(x)_{i,j} & = & \left\{ \begin{array}{ll} det(x[C_{i,j}])/det(A) & for\, i \neq n \\ det(x[C_{i,j}]) & otherwise \end{array} \right. \end{array}$$

The map π causes the map:

$$\pi^* : \mathbb{C}[W'] \to \mathbb{C}[Z']$$

given by:

$$\begin{array}{rcl} \pi^*(W_{n,n}) & = & det(X[C_0]) \\ \pi^*(W_{i,j}) & = & \left\{ \begin{array}{ll} det(X[C_{i,j}])/det(X[C_0]) & for\, i \neq n \\ det(X[C_{i,j}]) & otherwise \end{array} \right. \end{array}$$

Now we note that Z' is dense in V^r. Let $p \in \mathbb{C}[X]^{SL_n}$ be an invariant. Clearly p restricted to W' defines an element $p_{W'}$ of $\mathbb{C}[W']$. This then extends to Z' via π^*. Clearly $\pi^*(p_{W'})$ must match p on Z'. Thus we have that p is a polynomial in $det(X[C_{i,j}])$ possibly localized at $det(X[C_0])$.

Note that for a general C, $det(X[C])$ is already expressible in $det(X[C_0])$ and $det(X[C_{i,j}])$.

2.3 Other Representations

We discuss two other representations:

The Conjugate Action: Let $\mathcal{M}$ be the space of all $n \times n$ matrices with complex entries. we define the action of $A \in SL_n$ on $\mathcal{M}$ as follows. For an $M \in \mathcal{M}$, A acts on it by conjugation.

$$A \cdot M = AMA^{-1}$$

Note that $dim_{\mathbb{C}}(\mathcal{M}) = n^2$. Let $X = \{X_{ij} | 1 \leq i,j \leq n\}$ be the dual space to $\mathcal{M}$. The invariants for this action are $Tr(X), \ldots, Tr(X^i), \ldots, Tr(X^n)$. That these are invariants ic clear, for $Tr(AM^iA^{-1}) = Tr(M)$ for any A, M.

Also note that $Tr(X) = X_{11} + \ldots + X_{nn}$ is linear in the X's. This indicates an invariant hyperplane in $\mathcal{M}$. This is precisely the *trace zero matrices.* Thus $\mathcal{M} = \mathcal{M}_0 \oplus \mathcal{M}_1$ where $\mathcal{M}_0$ are all matrices M such that $Tr(M) = 0$. The one-dimensional complementary space $\mathcal{M}_1$ is composed of multiples of the identity matrix.

The orbits are parametrized by the **Jordan canonical form** $JCF(M)$. In other words, if $JCF(M) = JCF(M')$ them $M' \in O(M)$. Furthermore, if $JCF(M)$ is diagonal then the orbit of M is **closed** in $\mathcal{M}$.

The Space of Forms: Let V be a complex vector space of dimension n andV^* its dual. Let $X_1, \ldots, X_n$ be the dual basis. The space $Sym^d(V^*)$ consists of degree d forms are formal linear combinations of the monomials $X_1^{i_1} \ldots X_n^{i_n}$ with $i_1 + \ldots + i_n = d$. The typical form may be written as:

$$f(X) = \sum_{i_1 + \ldots + i_n = d} B_{i_1 \ldots i_n} X_1^{i_1} \ldots X_n^{i_n}$$

Let $A \in SL_n$ be such that

$$A^{-1} = \begin{bmatrix} a_{11} & \ldots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \ldots & a_{nn} \end{bmatrix}$$

The space $Sym^d(V^*)$ is the formal linear combination of the generaic coefficients $\{B_{\bar{i}} | |\bar{i}| = d\}$. The action of A is obtained by substituting

$$X_i \rightarrow \sum_j a_{ij} X_i$$

in $f(X)$ and recoomputing the coefficients. we illustrate this for $n = 2$ and $d = 2$. Then, the generic form is given by:

$$f(X_1, X_2) = B_{20}X_1^2 + B_{11}X_1^1X_2^1 + B_{02}X_2^2$$

Upon substitution, we get $f(a_{11}X_1 + a_{12}X_2, a_{21}X_1 + a_{22}X_2)$: Thus the new coefficients are:

$$\begin{array}{llll} B_{20} \rightarrow & B_{20}a_{11}^2+ & B_{11}a_{11}a_{21}+ & B_{02}a_{21}^2 \\ B_{11} \rightarrow & B_{20}2a_{11}a_{12}+ & B_{11}(a_{11}a_{22} + a_{12}a_{21})+ & B_{02}2a_{21}a_{22} \\ B_{02} \rightarrow & B_{20}a_{12}^2+ & B_{11}a_{12}a_{22}+ & B_{02}a_{22}^2 \end{array}$$

We thus see that the variables B move linearly, with coefficients as homogeneous polynomials of degree 2 in the coefficients of the group elements. In the above case, we know that the *discriminant* $B_{11}^2 - 4B_{20}B_{02}$ is an invariant,

These spaces have been the subject of intense analysis and their study by Gordan, Hilbert and other heralded the beginning of commutative algebra and invariant theory.

2.4 Full Reducibility

Let W be a representation of SL_n and let $Z \subseteq W$ be an invariant subspace. The reducibility question is whether there exists a complement Z' such that $W = Z \oplus Z'$ and Z' is SL_n-invariant as well.

The above result is indeed true although we will not prove it here. There are many proofs known, each with a specific objective in a specific situation, and each extremely instructive.

The simplest is possibly through the Weyl Unitary trick. In this, a suitable **compact** subgroup $U \subseteq SL_n$ is chosen. The theory of compact groups is much like that of finite groups and a complement Z' may easily be found. It is then shown that Z'

The second attack is through showing the fill reducibility of the module $\otimes^d V$, the d-th tensor product representation of V. This goes through the construction of the commutator of the same module regarded as a *right* S_d-module, with the symmetric group permutating the contents of the d positions. The full reducibility then follows from Maschke's theorem and that of the finite group S_d.

The oldest approach was through the construction of a symbolic reynold's operator, which is the **Cayley** Ω-process. Let $\mathbb{C}[W]$ be the ring of polynomial functions on the space W. The operator Ω is a map:

$$\Omega : \mathbb{C}[W] \rightarrow \mathbb{C}[W]^{SL_n}$$

such that if p is an invariant then $\Omega(p \cdot p') = p \cdot \Omega(p')$. The definition of Ω is nothing but **curious**.

3 Invariant Theory

3.1 Algebraic Groups and affine actions

An algebraic group (over $\mathbb{C}$) is an affine variety G equipped with (i) the group product, i.e., a morphism of algebraic varieties $\cdot : G \times G \to G$ which is associative, i.e. $(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$ for all $g_1, g_2, g_3 \in G$, (ii) and algebraic inverse $i : G \to G$, i.e, $g \cdot i(g) = i(g) \cdot g = 1_G$, where 1_G is (iii) a special element $1_G \in G$, which functions as the identity, i.e., $1_G \cdot g = g \cdot 1_G = g$ for all $g \in G$. Let $\mathbb{C}[G]$ be the ring of regular functions on G. The requirements (i)-(iii) above are via morphisms of algebraic varities, and thus the group product and the inverse must be defined algebraically. Thus the product $\cdot : G \times G \to G$ results in a morphism of $\mathbb{C}$-algebras $\cdot^* : \mathbb{C}[G] \to \mathbb{C}[G] \otimes \mathbb{C}[G]$ and the inverse into another morphism $i^* : \mathbb{C}[G] \to \mathbb{C}[G]$.

The essential example is obviously SL_n. Clearly for $G = SL_n$, we have $\mathbb{C}[G] = \mathbb{C}[X]/(det(X) - 1)$, where X is the $n \times n$ indeterminate matrix. The morphism $\cdot^*$ and i^* are clearly:

$$\begin{array}{rcl} \cdot^*(X_{ij}) & = & \sum_k X_{ik} \otimes X_{kj} \\ i^*(X_{ij}) & = & det(M_{ji}) \end{array}$$

where M_{ji} is the corresponding minor of X.

Next, let Z be another affine variety with $\mathbb{C}[Z]$ as its ring of regular functions. We say that Z is a G-**variety** if there is a morphism $\mu : G \times Z \to Z$ which is a group action. Thus, not only must G act on Z, it must do so *algebraically*. This μ induces the map μ^*:

$$\mu^* : \mathbb{C}[Z] \to \mathbb{C}[G] \otimes \mathbb{C}[Z]$$

Thus every function $f \in \mathbb{C}[Z]$ goes to a finite sum:

$$\mu^*(f) = \sum_{i=1}^{k} h_i \otimes f_i$$

where $f_i \in \mathbb{C}[Z]$ and $h_i \in \mathbb{C}[G]$ for all i.

To continue with our example, consider SL_2 and $Z = Sym^d(V^*)$. $\mathbb{C}[Z] = \mathbb{C}[B_{20}, B_{11}, B_{02}]$ and $\mathbb{C}[G] = \mathbb{C}[A_{11}, A_{12}, A_{21}, A_{22}]/(A_{11}A_{22} - A_{21}A_{12} - 1)$. We have for example:

$$\mu^*(B_{20}) = A_{11}^2 \otimes B_{20} + A_{11}A_{21} \otimes B_{11} + A_{21}^2 \otimes B_{02}$$

Let $\mu : G \times Z \to Z$ and $\mu' : G \times Z' \to Z'$ be two G-varities and let $\phi : Z \to Z'$ be a morphism. We say that ϕ is G-**equivariant** if $\phi(\mu(g, z)) = \mu'(g, \phi(z))$ for all $g \in G$ and $z \in Z$. Thus ϕ commutes with the action of G.

3.2 Orbits and Invariants

Every $g \in G$ induces an algebraic bijection on Z by restricting the map $\mu : G \times Z \to Z$ to g. We denote this map by $\mu(g)$ and call it **translation by** g. The map:

$$\mu(g) : Z \to Z$$

induces the isomorphism of $\mathbb{C}$-algebras:

$$\mu(g)^* : \mathbb{C}[Z] \to \mathbb{C}[Z]$$

Given any function $f \in \mathbb{C}[Z]$, $\mu(g)^*(f) \in \mathbb{C}[Z]$ is the *translated* function and denotes the action of G on f. This makes $\mathbb{C}[Z]$ into a G-module. If $\phi : Z \to Z'$ is a G-equivariant map, then the map $\phi^* : \mathbb{C}[Z'] \to \mathbb{C}[Z]$ is also G-equivariant, for the G action on $\mathbb{C}[Z]$ and $\mathbb{C}[Z']$ as above.

We next examine the equation:

$$\mu^*(f) = \sum_{i=1}^{k} h_i \otimes f_i$$

where $f_i \in \mathbb{C}[Z]$ and $h_i \in \mathbb{C}[G]$ for all i. For a fixed g, we see that

$$\mu(g)^*(f) = \sum_{i=1}^{k} h_i(g) \otimes f_i$$

Thus we see that every translate of f lies in the k-dimensional vector space $\mathbb{C}\cdot\{f_1, \dots, f_k\}$. Let

$$M(f) = \mathbb{C} \cdot \{\mu(g)^*(f) | g \in G\} \subseteq \mathbb{C} \cdot \{f_1, \dots, f_k\}$$

Clearly, $M(f)$ is a G-invariant subspace of $\mathbb{C}[Z]$. This may be generalized:

Proposition 16 *Let $S = \{s_1, \dots, s_m\}$ be a finite subset of $\mathbb{C}[Z]$. Then there is a finite-dimensional G-invariant subspace $M(S)$ of $\mathbb{C}[Z]$ containing $s_1, \dots, s_m$.*

Next, let us consider $\mu : G \times Z \to Z$ and fix a $z \in Z$. We get the map $\mu_z : G \to X$. Since G is an affine variety, we see that the image $\mu_z(G)$ is a **constructible set**, whose closure is an affine variety. The image is precisely the orbit $O(z) \subseteq Z$. The closure of $O(z)$ will be denoted by $\Delta(z)$.

If $O(z) = \Delta(z)$, then we have the G-equivariant embedding $i_z : O(z) \to Z$. This gives us the map $i_z^* : \mathbb{C}[Z] \to \mathbb{C}[O(z)]$ which is a surjection. Thus there is an ideal $I(z) = ker(i_z^*)$ such that $\mathbb{C}[O(z)] \cong \mathbb{C}[Z]/I(z)$. Since the map i_z is G-equivariant, $I(z)$ is a G-submodule of $\mathbb{C}[Z]$ and is the ideal of definition of the orbit $O(z)$.

In general, if $I \subseteq \mathbb{C}[Z]$ is an ideal which is G-invariant, then the variety of I is also G-invariant and is the union of orbits.

The second construction that we make is that of the **quotient** Z/G. Since $\mathbb{C}[Z]$ is a G-module, we examine $\mathbb{C}[Z]^G$, the subring of G-invariant functions in $\mathbb{C}[Z]$. We define Z/G as the spectrum $Spec(\mathbb{C}[Z]^G)$. The inclusion $\mathbb{C}[Z]^G \to \mathbb{C}[Z]$ gives us the quotient map:

$$\pi : Z \to Z/G$$

Example 17 *Let us consider $Z = Sym^2(V^*) \cong \mathbb{C}^3$ where V is the standard representation of $G = SL_2$. As we have seen, $\mathbb{C}[Z] = \mathbb{C}[B_{20}, B_{11}, B_{02}]$. There is only one invariant*

$\delta = B_{11}^2 - 4B_{02}B_{20}$. Thus $\mathbb{C}[Z]^G = \mathbb{C}[\delta]$. Thus Z/G is precisely $Spec(\mathbb{C}[\delta]) = \mathbb{C}$, the complex plane. The map π is executed as follows: given a form $aX_1^2 + bX_1X_2 + cX_2^2 \equiv (a,b,c) \in Z$, we eveluate the invariant δ at the point (a,b,c). Thus $\pi : \mathbb{C}^3 \to \mathbb{C}$ is given by:

$$\pi(a,b,c) = b^2 - 4ac$$

Clearly, if $f, f' \in Z$ such that $f = g \cdot f'$ for $g \in SL_2$ then $\delta(f) = \delta(f')$. We look at the converse: if f, f' are such that $\delta(f) = \delta(f')$ then is it that $f \in O(f')$? We begin with $f = aX_1^2 + bX_1X_2 + cX_2^2$. We assume for the moment that $a \neq 0$. If that is the case, we make the substitution $X_1 \to X_1 - \alpha X_2$ and $X_2 \to X_2$. Note that this transformation is unimodular for all α. This transforms f to:

$$a(X_1 - \alpha X_2)^2 + b(X_1 - \alpha X_2)X_2 + cX_2^2$$

The coefficient of X_1X_2 is $-2a\alpha + b$. Thus by choosing α as $b/2a$ we see that f is transformed into $a''X_1^2 - c''X_2^2$ for some a'', c''. By a similar token, even if $a = 0$ one may do a similar transform. Thus in general, if f is not the zero form, there is a point in $O(f)$ which is of the form $aX_1^2 - cX_2^2$. Thus, we may assume that both f and f' are in this form. We can simplify the form further by a diagonal element of SL_2 to put both f and f' as $X_1^2 - cX_2^2$ and $X_1^2 - c'X_2^2$. It is now clear that $\delta(f) = \delta(f')$ implies that $c = c'$. Thus the general answer is that if $f, f' \neq 0$ and $\delta(f) = \delta(f')$ then $f \in O(f')$.

Next, let us examine the form 0. we see that $\delta(0) = 0$. Thus we see that (i) for any point $d \in \mathbb{C}$, if $d \neq 0$, then $\pi^{-1}(d)$ consists of a single orbit, (ii) $\pi^{-1}(0)$ consists of two orbits, $O(0)$ and $O(X_1^2)$, the perfect square, (iii) the orbits $O(f)$ are closed when $\delta(f) \neq 0$, (iv) $O(X_1^2)$ is not closed. Its closure includes the closed orbit 0.

The above example illustrates the utility of contructing the quotient Z/G as a variety parametrizing closed orbits albeit with some deviant points. In the example above, the discrepancy was at the point 0, wherein the pre-image is closed but decomposes into two orbits.

We state the all-important theorem linking a space and its quotient in the restricted case when G is finite and Z is a finite-dimensional G-module. Thus $\mathbb{C}[Z]$ is a polynomial ring and the action of G is homogeneous.

Theorem 18 *Let G be a finite group and act on the space Z. Let $R = \mathbb{C}[Z]$ and $R^G = \mathbb{C}[Z]^G$ be the ring of invariants. Let $\pi : Z \to Z/G$ be the quotient map. Then*

(i) For any ideal $J \subseteq R^G$, we have $(J \cdot R) \cap R^G = J$.

(ii) The map π is surjective. Further, for any $x \in Z/G$, $\pi^{-1}(x)$ is a single orbit in Z.

Proof: Let $f_1, \ldots, f_k$ be elements of R^G and $f \in R^G$ be such that:

$$f = r_1 f_1 + \ldots + r_k f_k$$

where r_i are elements of R. Since G is finite, we apply the Reynolds operator $p \to p^G$.

In other words, we have:

$$\begin{aligned} f &= \frac{1}{|G|}\sum_{g\in G} f \\ &= \frac{1}{|G|}\sum_{g\in G}\sum_i r_i f_i \\ &= \frac{1}{|G|}\sum_i f_i \sum_{g\in G} r_i \\ &= \sum_i f_i \frac{1}{|G|}\sum_{g\in G} r_i \\ &= \sum_i f_i r_i^G \end{aligned}$$

Note that we have used the fact that if $h \in R^G$ and $p \in R$ then $(h\cdot p)^G = h\cdot p^G$. Thus we see that any element f of R^G which is expressible as an R-linear combination of elements f_i's of R^G is already expressible as an R^G-linear combination of the same elements. This proves (i) above.

Now we prove (ii). Firstly, let J be a maximal ideal of R^G. By part (i) above, $J \cdot R$ is a proper ideal of R and thus π is surjective. Let $x \in Z/G$ and $J_x \subseteq R^G$ be the maximal ideal for the point x. Let z be such that $\pi(z) = x$ and let $I_{O(x)} \subseteq R$ be the ideal of all functions in R vanishing at all points of the orbit $O(z)$ of z. We show that $rad(J_x \cdot R) = I_{O(z)}$ which proves (ii). Towards that, it is clear that (a) $J_x R \subseteq I_{O(z)}$ and (b) the variety of $J_x R$ is G-invariant. Let $O(z')$ is another orbit in the variety of $J_x \cdot R$. If $O(z') \neq O(z)$ then we already have that there is a $p \in R^G$ such that $p(O(z)) = 0$ and $p(O(z')) = 1$. Since this $p \in J_x \cdot R$, the variety of $J_x \cdot R$ excludes the orbit $O(z')$ proving our claim. $\square$

Theorem 19 *Let Z be a finite-dimesional G-module for a finite group G. Then $\mathbb{C}[Z]^G$, the ring of G-invariants, is finitely generated as a $\mathbb{C}$-algebra.*

Proof: Since the action of G is homogeneous, every invariant $f \in \mathbb{C}[Z]^G$ is the sum of homogeneous invariants. Let $I_G \subseteq \mathbb{C}[Z]$ be the ideal generated by the **positive** degree invariants. By the Hilbert Basis theorem, $I_G = (f_1, \ldots, f_k) \cdot \mathbb{C}[Z]$, where each f_i is itself an invariant. we claim that $\mathbb{C}[Z]^G = \mathbb{C}[f_1, \ldots, f_k] \subseteq \mathbb{C}[Z]$, or in other words, every invariant is a polynomial over $\mathbb{C}$ in the terms $f_1, \ldots, f_k$. To this end, let f be a homogeneous invariant. We prove this by induction over the degree d of f. Since $f \in I_G$, we have

$$f = \sum_i r_i f_i \ \text{ where for all } i,\, r_i \in \mathbb{C}[Z]$$

By applying the reynolds operator, we have:

$$f = \sum_i h_i f_i \ \text{ where for all } i,\, h_i \in \mathbb{C}[Z]^G$$

Now since each h_i has degree less than d, by our induction hypothesis, each is an element of $\mathbb{C}[f_1, \ldots, f_k]$. This then shows that f too is an element of $\mathbb{C}[f_1, \ldots, f_k]$. $\square$

3.3 The Nagata Hypothesis

We now generalize the above two theorems to the case of more general groups. This generlization is possible if the group G satisfies what we call the *Nagata Hypothesis.*

Definition 20 *Let G be an algebraic group. We say that G satisfies the Nagata hypothesis if for every finite-dimensional module M over $\mathbb{C}$ and a G-invariant hyperplane $N \subseteq M$, there is a decomposition $M = H \oplus P$ as G-modules, where H is a hyperplane and P is an invariant (i.e., P is the trivial representation.*

The group SL_n satisfies the hypothesis, and so do the so-called reductive groups over $\mathbb{C}$.

Theorem 21 *Let G satisfy the Nagata hypothesis and let Z be an affine G-variety. The ring $\mathbb{C}[Z]^G$ is finitely generated as a $\mathbb{C}$-algebra.*

The proof will go through several steps.

Let $f \in \mathbb{C}[Z]$ be an arbitrary element. We define two modules $M(f)$ and $N(f)$.

$$\begin{aligned} M(f) &= \mathbb{C} \cdot \{s \cdot f | s \in G\} \\ N(f) &= \mathbb{C} \cdot \{s \cdot f - t \cdot f | s, t \in G\} \end{aligned}$$

Thus $M(f) \supseteq N(f)$ are finite-dimensional submodules of $\mathbb{C}[Z]$.

Lemma 22 *Let $f \in \mathbb{C}[Z]$ be an arbitrary element. Then there exists an $f^* \in \mathbb{C}[Z]^G \cap M(f)$ such that $f - f^* \in N(f)$.*

Proof: We prove this by induction over $dim(M(f))$. Note that $s \cdot f = f + (s \cdot f - 1 \cdot f)$ and thus $M(f) = N(f) \oplus \mathbb{C} \cdot f$ as vector spaces. Thus $N(f)$ is a G-invariant hyperplane of $M(f)$. By the Nagata hypothesis, $M(f) = f' \oplus H$, where f' is an invariant. If $f' \notin N(f)$ then $M(f) = f' \oplus N(f)$ and the lemma follows. However, if $f' \in N(f)$, then $f = f' + h$ where $h \in H$. Since H is a G-module, we see that $M(h) \subseteq H$ which is of dimension less than that of $M(f)$. Thus there is an invariant h^* such that $h - h^* \in N(h)$. Note that $f - f' = h$ with f' invariant, implies that $s \cdot h - t \cdot h = s \cdot f - t \cdot f$. Thus $N(h) \subseteq N(f)$. We take $f^* = h^*$. Examining $f - f^*$, we see that

$$f - h^* = f' + h - h^* \in N(f)$$

This proves the lemma. □

Lemma 23 *Let $f_1, \ldots, f_r \in \mathbb{C}[Z]^G$. Then $\mathbb{C}[Z]^G \cap (\sum_i \mathbb{C}[Z] \cdot f_i) = \sum_i \mathbb{C}[Z]^G \cdot f_i$. Thus if an invariant f is a $\mathbb{C}[Z]$linear combination of invariants, then it is already a $\mathbb{C}[Z]^G$ linear combination.*

Proof: This is proved by induction on r. Say $f = \sum_{i=1}^r h_i f_i$ with $h_i \in \mathbb{C}[Z]$. Applying the above lemma to h_r, there is an $h'' \in \mathbb{C}[Z]^G$ and an $h' \in N(h_k)$ such that $h_r = h' + h''$. We tackle h' as follows. Since f is an invariant, we have for all $s, t \in G$:

$$\sum_{i=1}^{r} (s \cdot h_i - t \cdot h_i) f_i = s \cdot f - t \cdot f = 0$$

Hence:

$$(s \cdot h_r - t \cdot h_r)f_r = \sum_{i=1}^{r-1}(s \cdot h_i - t \cdot h_i)f_i$$

It follows from this that $h' f_r = \sum_{i=1}^{r-1} h'_i f_i$ for some $h'_i \in \mathbb{C}[Z]$. Substituting this in the expression

$$f = \sum_{i=1}^{r-1} h_i f_i + (h' + h'')f_r$$

we get:

$$f - h'' f_r = \sum_{i=1}^{r-1}(h_i + h'_i)f_i$$

Thus the invariant $f - h'' f_r$ is $\mathbb{C}[Z]$-linear combination of $r-1$ invariants, and thus the induction hypothesis applies. This then results in an expression for f as a $\mathbb{C}[Z]^G$-linear combination for f. $\square$

The above lemma proves part (i) of Theorem 18 for groups G for which the Nagata hypothesis holds. The route to Theorem 19 is now a straight-forward adaptation of its proof for finite groups. In the case when Z is a G-module, $\mathbb{C}[Z]^G$ would then be homogeneous and the proof of Theorem 19 holds. For general Z here is a trick which converts it to the homogeneous case:

Proposition 24 *Let Z be an affine G-variety. Then there is a G-module W and an equivariant embedding $\phi : Z \to W$.*

Proof: Since $\mathbb{C}[Z]$ is a finitely generated $\mathbb{C}$-algebra, $\mathbb{C}[Z] = \mathbb{C}[f_1, \ldots, f_k]$ where $f_1, \ldots, f_k$ are some specific elements of $\mathbb{C}[Z]$. By expanding the list of generators, we may assume that the vector space $\mathbb{C} \cdot \{f_1, \ldots, f_k\} \subseteq \mathbb{C}[Z]$ is a finite-dimensional G-module. Let us construct W as an isomorphic copy of this module with the dual basis $W_1, \ldots, W_k$. We construct the map $\phi^* : \mathbb{C}[W] \to \mathbb{C}[Z]$ by defining $\phi(W_i) = f_i$. Since $\mathbb{C}[W]$ is a free algebra, we indeed have the surjection:

$$\phi^* : \mathbb{C}[W] \to \mathbb{C}[Z]$$

This proves the proposition. $\square$

We are now ready to prove:

Theorem 25 *Let G satisfy the Nagata hypothesis. Let Z be an affine G-variety with coordinate ring $\mathbb{C}[Z]$. Then $\mathbb{C}[Z]^G$ is finitely generated as a $\mathbb{C}$-algebra.*

Proof: We construct the equivariant surjection $\phi^* : \mathbb{C}[W] \to \mathbb{C}[Z]$. Note that $\mathbb{C}[W]^G$ is already finitely generated over $\mathbb{C}$, say $\mathbb{C}[W]^G = \mathbb{C}[h_1, \ldots, h_k]$. We claim that $\phi^*(h_1), \ldots, \phi^*(h_k)$ generate $\mathbb{C}[Z]^G$.

Now, let $f \in \mathbb{C}[Z]^G$. By the surjectivity of ϕ^*, there is an $h \in \mathbb{C}[W]$ such that $\phi^*(h) = f$. Consider the space $N(h)$. A typical generator of $N(h)$ is $s \cdot h - t \cdot h$. Applying ϕ^* to this, we see that:

$$\begin{aligned} \phi^*(s \cdot h - t \cdot h) &= s \cdot \phi^*(h) - t \cdot \phi^*(h) \\ &= f - f = 0 \end{aligned}$$

By an earlier lemma there is an invariant h^* such that $h - h^* \in N(h)$. thus applying ϕ^* we see that $\phi^*(h^*) = \phi(h) = f$. Thus there is an invariant h^* such that $\phi^*(h^*) = f$. Now $h^* \in \mathbb{C}[h_1, \ldots, h_k]$ implies that $f = \phi^*(h^*) \in \mathbb{C}[\phi^*(h_1), \ldots, \phi^*(h_k)]$. □

4 Orbit-closures

In this subsection we will analyse the validity of Theorem 18 for general groups G with the Nagata hypothesis. The objective is to analyse the map $\pi : Z \to Z/G$.

Proposition 26 *Let G satisfy the Nagata hypothesis and act on the affine variety Z. Then the map $\pi : Z \to Z/G$ is surjective.*

Proof: Let $J \subseteq \mathbb{C}[Z]^G$ be a maximal ideal. By lemma 23 $J \cdot \mathbb{C}[Z]$ is a proper ideal of $\mathbb{C}[Z]$. This implies that π is surjective. □

Theorem 27 *Let G satisfy the Nagata hypothesis and act on the affine variety Z. Let W_1 and W_2 be G-invariant (Zariski-)closed subsets of Z such that $W_1 \cap W_2$ is empty. Then there is an invariant $f \in \mathbb{C}[Z]^G$ such that $f(W_1) = 0$ and $f(W_2) = 1$.*

Proof: Let I_1 and I_2 be the ideals of W_1 and W_2 in $\mathbb{C}[Z]$. Since their intersection is empty, by Hilbert Nullstellensatz, we have $I_1 + I_2 = 1$. Whence there are functions $f_1 \in I_1$ and $f_2 \in I_2$ such that $f_1 + f_2 = 1$. For arbitrary $s, t \in G$ we have:

$$(s \cdot f_1 - t \cdot f_1) + (s \cdot f_2 - t \cdot f_2) = 1 - 1 = 0$$

Note that $M(f_i)$ and $N(f_i)$ are submodules of the G-invariant ideal I_i. Now applying lemma 22 to f_1, we see that there are elements $f_i' \in N(f_i)$ such that $f_1 + f_1' \in \mathbb{C}[Z]^G$. Whence we see that:

$$(f_1 + f_1') + (f_2 + f_2') = 1$$

where $f = f_1 + f_1'$ is an invariant. Since $f_1 + f_1' \in I_1$ we see that $f(W_1) = 0$. On the other hand, $f = 1 - (f_2 + f_2') \in 1 + I_2$ and thus $f(W_2) = 1$. □

Recall that $\Delta(z)$ denotes the closure of the orbit $O(z)$ in Z.

Theorem 28 *Let G satisfy the Nagata hypothesis and act on an affine variety Z. We define the relation $\approx$ on Z as follows: $z_1 \approx z_2$ if and only if $\Delta(z_1) \cap \Delta(z_2)$ is non-empty. Then*

(i) $\approx$ is an equivalence relation.

(ii) $z_1 \approx z_2$ *iff* $f(z_1) = f(z_2)$ *for all* $f \in \mathbb{C}[Z]^G$.

(iii) Within each $\Delta(z)$ *there is a unique closed orbit, and this is of minimum dimension among all orbits in* $\Delta(z)$.

Proof: It is clear that (ii) proves (i). Towards (ii), if $\Delta(z_1) \cap \Delta(z_2)$ is empty, then by Theorem 27, there is an invariant separating the two points. Thus it remains to show that if $\Delta(z_1 0 \cap \Delta(z_2)$ is *non-empty*, and f is any invariant, then $f(z_1) = f(z_2)$. So let z be an element of the intersection. Since f is an invariant, we have $f(O(z_1))$ is a constant, say α. Since $O(z_1) \subseteq \Delta(z_1)$ is dense, and f is continuous, we have $f(z) = \alpha$. Thus $f(z_1) = f(z_2) = f(z) = \alpha$.

Now (iii) is easy. Clearly $\Delta(z)$ cannot have *two* closed distinct orbits, for otherwise they woulbe separated by an invariant. But this must take the same value on all points of $\Delta(z)$. That it is of minimum dimension follows algebraic arguments. □

Definition 29 *Let* Z *be a* G*-variety and* $z \in Z$. *We say that* z *is* **stable** *if the orbit* $O(z) \subseteq Z$ *is closed in* Z.

By the above theorem, every point x of $Spec(C[Z]^G)$ corresponds to exactly one stable point: the point whose orbit is of minimum dimension in $\pi^{-1}(x)$.

Example 30 *Consider the action of* $G = SL_n$*on* $\mathcal{M}$, *the space of* $n \times n$*-matrices by conjugation. Thus, given* $A \in SL_n$ *and* $M \in \mathcal{M}$, *we have:*

$$A \cdot M = AMA^{-1}$$

Let $R = \mathbb{C}[\mathcal{M}] = \mathbb{C}[X_{11}, \ldots, X_{nn}]$ *be the ring of functions on* $\mathcal{M}$. *The invariants* R^G *is generated as a* $\mathbb{C}$*-algebra by the forms* $e_i(X) = Tr(X^i)$, *for* $i = 1, \ldots, n$. *The forms* $\{e_i | 1 \leq i \leq n\}$ *are algebraically independent and thus* R^G *is the polynomial ring* $\mathbb{C}[e_1, \ldots, e_n]$. *Clearly then* $Spec(R^G) \cong \mathbb{C}^n$ *and we have:*

$$\pi : \mathcal{M} \cong \mathbb{C}^{n^2} \to \mathbb{C}^n$$

Given a matrix M *with eigenvalues* $\{\lambda_1, \ldots, \lambda_n\}$, *we have:*

$$e_i(M) = \lambda_1^i + \ldots + \lambda_n^i$$

Thus, by the fundamental theorem of algebra, the image $\pi(M)$ *determines the set* $\{\lambda_1, \ldots, \lambda_n\}$ *(with multiplicities). On the other hand, given any tuple* $\mu = (\mu_1, \ldots, \mu_n)$ *there is a unique set* $\lambda_\mu = \{\lambda_1, \ldots, \lambda_n\}$ *such that* $\sum_r \lambda_r^i = \mu_i$. *Clearly, for the diagonal matrix* $D(\lambda_\mu) = diag(\lambda_1, \ldots, \lambda_n)$, *we have that* $\pi(D(\lambda)) = \mu$. *This verifies that* π *is surjective.*

For a given μ, *the set* $\pi^{-1}(\mu)$ *are all matrices* M *with* $Spec(M) = \lambda = \lambda_\mu$. *By the Jordan canonical form (JCF), this set may be stratified by the various Jordan canonical blocks of spectrum* λ. *If* λ *has no multiplicities then* $\pi^{-1}(\mu)$ *consists of just one orbit: matrices* M *such that* $JCF(M) = D(\lambda)$. *For a general* λ, *the orbit of* M *with* $JCF(M) = D(\lambda)$ *is the unique closed orbit of minimum dimension. All other orbits contain this orbit in its closure. Thus stable points* $M \in \mathcal{M}$ *are the diagonalizable matrices.*

As an example, consider the case when $n = 2$ and the matrix:

$$N = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

Consider the family $A(t) = diag(t, t^{-1}) \in SL_2$. We see that:

$$N(t) = A(t)NA(t)^{-1} = \begin{bmatrix} \lambda & t^2 \\ 0 & \lambda \end{bmatrix}$$

Thus $\lim_{t \to 0} N(t) = diag(\lambda, \lambda)$, the diagonal matrix.

Thus, we see that the invariant ring $\mathbb{C}[Z]^G$ puts a different equivalence relation $\approx$ on points in Z which is coarser than $\cong$, the orbit equivalence relation. The relation $\approx$ is more 'topological' than group-theoretic and correctly classifies orbits by their separability by invariants. The special case of Theorem 28 when Z is a representation was analysed by Hilbert in 1893. The point $0 \in Z$ is then the smallest closed orbit, and the equivalence class $[0]_{\approx}$ is termed as the **null-cone** of Z. We see that the null-cone consists of all points $z \in Z$ such that 0 lies in the orbit-closure $\Delta(z)$ of z. It was Hilbert who discovered that if $0 \in \Delta(z)$ then 0 lies in the orbit-closure of a 1-parameter diagonal subgroup of SL_n. To understand the intricay of Hilbert's constructions, it is essential that we understand diagonal subgroups of SL_n.

5 Tori in SL_n

Let $\mathbb{C}^*$ denote the multiplicative group of non-zero complex numbers. A torus is the abstract group $(\mathbb{C}^*)^m$ for some m. Note that $\mathbb{C}^*$ is an abelian algebraic group with $\mathbb{C}[G] = \mathbb{C}[T, T^{-1}]$. Furthermore, $\mathbb{C}^*$ has a compact subroup $S^1 = \{z \in \mathbb{C}, |z| = 1\}$, the unit circle.

Next, let us look at representations of tori. For $\mathbb{C}^*$, the simplest representations are indexed by integers $k \in \mathbb{Z}$. So let $k \in \mathbb{Z}$. The representation $\mathbb{C}_{[k]}$ corresponds to the 1-dimensional vector space $\mathbb{C}$ with the action:

$$t \cdot z = t^k z$$

Thus a non-zero $t \in \mathbb{C}^*$ acts on z by multiplication of the k-th power. Next, for $(\mathbb{C}^*)^m$, let $\chi = (\chi[1], \ldots, \chi[m])$ be a sequence of integers. For such a χ, we define the representation $\mathbb{C}_\chi$ as follows: Let $\overline{t} = (t_1, \ldots, t_m) \in (\mathbb{C}^*)^m$ be a general element and $z \in \mathbb{C}$. The action is given by:

$$\overline{t} \cdot z = t_1^{\chi[1]} \ldots t_m^{\chi[m]} z$$

Such a χ is called a **character** of $(\mathbb{C}^*)^m$.

These 1-dimensional representations of tori are crucial in the analysis of algebraic group actions.

Let us begin by understanding the structure of algebraic homomorphisms from $\mathbb{C}^*$ to $SL_n(\mathbb{C})$. So let

$$\lambda : \mathbb{C}^* \to SL_n(\mathbb{C})$$

be such a map such that $\lambda(t) = [a_{ij}(t)]$ where $a_{ij}(T) \in \mathbb{C}[T, T^{-1}]$. An important substitution is for $t = e^{i\theta}$, and we obtain a 2π-periodic map

$$\overline{\lambda} : \mathbb{C} \to SL_n$$

We see that $\overline{\lambda}(0) = I$. Let the derivative at 0 for $\overline{\lambda}$ be X.

We have the following general lemma:

Lemma 31 *Let $f : \mathbb{C} \to SL_n$ be a smooth map such that $f(0) = I$ and $f'(0) = X$, where X is an $n \times n$-matrix. Then for $\theta \in \mathbb{C}$,*

$$\lim_{k \to \infty} \left[f\left(\frac{\theta}{k} \right) \right]^k = e^{\theta X}$$

The proof follows from the local diffeomorphism of the exponential map in the neighborhood of the identity matrix.

Applying this lemma to $\overline{\lambda}$ we see that $e^{\theta X}$ is in the image of $\overline{\lambda}$ for all θ. Now, since $\overline{\lambda}$ is 2π-periodic, we must have $e^{(n2\pi+\theta)X} = e^{\theta X}$. This forces (i) X to be diagonalizable, and (ii) with integer eigenvalues. This proves:

Proposition 32 *Let $\lambda : \mathbb{C}^* \to SL(V)$ be an algebraic homomorphism. Then the image of λ is closed and $V \cong \mathbb{C}_{[m_1]} \oplus \ldots \oplus \mathbb{C}_{[m_n]}$, for some integers $m_1, \ldots, m_n$, where $n = dim_{\mathbb{C}}(V)$.*

Based on this, we have the generalization:

Proposition 33 *Let $\lambda : (\mathbb{C}^*)^r \to SL(V)$ be an algebraic homomorphism. Then the image of λ is closed and $V \cong \mathbb{C}_{\chi_1} \oplus \ldots \oplus \mathbb{C}_{\chi_n}$, for some integers $\chi_1, \ldots, \chi_n$, where $n = dim_{\mathbb{C}}(V)$.*

Thus, in effect, for every homomorphism $\lambda : (\mathbb{C}^*)^r \to SL_n$, there is a fixed invertible matrix A such that for all $\overline{t} \in (\mathbb{C}^*)^r$, the conjugate $A\lambda(\overline{t})A^{-1}$ is diagonal.

A torus in SL_n is defined as an abstract subgroup H of SL_n which is isomorphic to $(\mathbb{C}^*)^r$ for some r. The **standard maximal torus** D of SL_n is the diagonal matrices $diag(t_1, \ldots, t_n)$ where $t_i \in \mathbb{C}^*$ and $t_1 t_2 \ldots t_n = 1$.

This clears the way for the important theorem:

Theorem 34 *(i) Every torus is contained in a maximal torus. All maximal tori in SL_n are isomorphic to $(\mathbb{C}^*)^{n-1}$.*

(ii) If T and T' are two maximal tori then there is an $A \in SL_n$ such that $T' = ATA^{-1}$. Thus all maximal tori are conjugate to D above.

(iii) Let $N(D)$ be the normalizer of D and $N(D)^o$ be the connected component of $N(D)$. Then $N(D)^o = D$ and $N(D)/D$ is the **Weyl group** *W, isomorphic to the symmetric group S_n.*

Definition 35 *Let G be an algebraic group. $\Gamma(G)$ will denote the collection of all* **1-parameter subgroups** *of G, i.e., morphisms $\lambda : \mathbb{C}^* \to G$. $X(G)$ will denote the collection of all* **characters** *of G, i.e., homomorphisms $\chi : G \to \mathbb{C}^*$.*

We consider the case when $G = (\mathbb{C}^*)^r$. Clearly, for a given $\lambda : \mathbb{C}^* \to G$, there are integers $m_1, \ldots, m_r \in \mathbb{Z}$ such that:

$$\lambda(t) = (t^{m_1}, \ldots, t^{m_r})$$

In the same vein, for the character $\chi : G \to \mathbb{C}^*$, we have integers $a_1, \ldots, a_r$ such that

$$\chi(t_1, \ldots, t_r) = \prod_i t_i^{a_i}$$

We also have the composition $\lambda \circ \chi : \mathbb{C}^* \to \mathbb{C}^*$, where by:

$$\lambda \circ \chi(t) = t^{m_1 a_1 + \ldots + m_r a_r}$$

Consolidating all this, we have:

Theorem 36 *Let $G = (\mathbb{C}^*)^r$. Then $\Gamma(G) \cong \mathbb{Z}^r$ and $X(G) \cong \mathbb{Z}^r$. Furthermore, there is the pairing*

$$<,>: \Gamma(G) \times X(G) \to \mathbb{Z}$$

which is a unimodular pairing on lattices.

Example 37 *Let $G = (\mathbb{C}^*)^3$ and λ and χ be as follows:*

$$\begin{aligned} \lambda(t) &= (t^3, t^{-1}, t^2) \\ \chi(t_1, t_2, t_3) &= t_1^{-1} t_2 t_3^2 \end{aligned}$$

Then, $\lambda \cong [3, -1, 2]$ and $\chi \cong [-1, 1, 2]$. We evaluate the pairing:

$$< \lambda, \chi >= 3 \cdot -1 + (-1) \cdot 1 + 2 \cdot 2 = 0$$

We now turn to the special case of $D \subseteq SL_n$, the maximal torus which is isomorphic to $(\mathbb{C}^*)^{n-1}$. By the above theorem, $\Gamma(D), X(D) \cong \mathbb{Z}^{n-1}$. However, it will more convenient to identify this space as a subset of $\mathbb{Z}^n$. So let:

$$\mathbb{Y}^n = \{[m_1, \ldots, m_n] \in \mathbb{Z}^n | m_1 + \ldots + m_n = 0\}$$

It is easy to see that $\mathbb{Y}^n \cong \mathbb{Z}^{n-1}$. In fact, we will set up a special bijection $\theta : \mathbb{Y}^n \to \mathbb{Z}^{n-1}$ defined as:

$$\theta([m_1, m_2, \ldots, m_n]) = [m_1, m_1 + m_2, \ldots, m_1 + \ldots + m_{n-1}]$$

The inverse θ^{-1} is also easily computed:

$$\theta^{-1}[a_1, \ldots, a_{n-1}] = [a_1, a_2 - a_1, a_3 - a_2, \ldots, a_{n-1} - a_{n-2}, -a_{n-1}]$$

This θ corresponds to the $\mathbb{Z}$-basis of $\mathbb{Y}^n$ consisting of the vectors $e_1 - e_2, \ldots, e_{n-1} - e_n$ where e_i is the standard basis of $\mathbb{Z}^n$. This is also equivalent to the embedding $\theta^* : (\mathbb{C}^*)^{n-1} \to D$ as follows:

$$(t_1, \ldots, t_{n-1}) \to \begin{bmatrix} t_1 & 0 & \ldots & & 0 \\ 0 & t_1^{-1}t_2 & 0 & \ldots & 0 \\ & & \vdots & & \\ 0 & \ldots & 0 & t_{n-2}^{-1}t_{n-1} & 0 \\ 0 & & \ldots & 0 & t_{n-1}^{-1} \end{bmatrix}$$

A useful computation is to consider the inclusion $D \subseteq D^*$, where $D^* \subseteq GL_n$ is subgroup of *all* diagonal matrices. Clearly $\Gamma(D) \subseteq \Gamma(D^*)$, however there is a surjection $X(D^*) \to X(D)$. It will be useful to work out this surjection explicitly via θ and θ^*. If $[m_1, \ldots .m_n] \in \mathbb{Z}^n \cong X(D^*)$, then it maps to $[m_1 - m_2, \ldots, m_{n-1} - m_n] \in \mathbb{Z}^{n-1} \cong X((\mathbb{C}^*)^{n-1})$ via θ^*. If we push this back into $\mathbb{Y}^n$ via θ^{-1}, we get:

$$[m_1, \ldots, m_n] \to [m_1 - m_2, 2m_2 - m_1 - m_3, , \ldots, 2m_{n-1} - m_{n-2} - m_n, m_n - m_{n-1}]$$

We are now ready to define the **weight spaces** of an SL_n-module W. So let W be such a module. By restricting this module to $D \subseteq G$, via Proposition 33, we see that W is a direct sum $W = \mathbb{C}_{\chi_1} \oplus \ldots \oplus \mathbb{C}_{\chi_N}$, where $N = dim_{\mathbb{C}}(W)$. Collecting identical characters, we see that:

$$W = \oplus_{\chi \in X(D)} \mathbb{C}_\chi^{m_\chi}$$

Thus W is a sum of m_χ copies of the module $\mathbb{C}_\chi$. Clearly $m_\chi = 0$ for all but a finite number, and is called the **multiplicity** of χ. For a given module W, computing m_χ is an intricate combinatorial exercise and is given by the celebrated **Weyl Character Formula**.

Example 38 *Let us look at SL_3 and the weight-spaces for some modules of SL_3. All modules that we discuss will also be GL_3-modules and thus D^* modules. The formula for converting D^*-modules to D-modules will be useful. This map is $\mathbb{Z}^3 \to \mathbb{Y}^3$ and is given by:*

$$[m_1, m_2, m_3] \to [m_1 - m_2, 2m_2 - m_1 - m_3, m_3 - m_2]$$

The simplest SL_n module is $\mathbb{C}^3$ with the basis $\{X_1, X_2, X_3\}$ with D^* weights $[1, 0, 0], [0, 1, 0]$ *and* $[0, 0, 1]$. *This converted to D-weights give us $\{[1, -1, 0], [-1, 2, -1], [0, -1, 1]\}$, with $\mathbb{C}_{[1,-1,0]} \cong \mathbb{C} \cdot X_1$ and so on.*

The next module is $Sym^2(\mathbb{C}^3)$ with the basis X_i^2 and X_iX_j. The six D^* and D-weights with the weight-spaces are given below:

D^*-wieghts	*D-weights*	*weight-space*
$[2, 0, 0]$	$[2, -2, 0]$	X_1^2
$[0, 2, 0]$	$[-2, 4, -2]$	X_2^2
$[0, 0, 2]$	$[0, -2, 2]$	X_3^2
$[0, 1, 1]$	$[-1, 1, 0]$	X_2X_3
$[1, 0, 1]$	$[1, -2, 1]$	X_1X_3
$[1, 1, 0]$	$[0, 1, -1]$	X_1X_2

The final example is the space of 3×3*-matrices* $\mathcal{M}$ *acted upon by conjugation. We see at once that* $\mathcal{M} = \mathcal{M}_0 \oplus \mathbb{C}\cdot I$ *where* $\mathcal{M}_0$ *is the* 8*-dimensional space of trace-zero matrices, and* $\mathbb{C}\cdot I$ *is* 1*-dimensional space of multiples of the idenity matrix. Weight vectors are* E_{ij}*, with* $1 \leq i,j \leq 3$*. The* D^* *weights are* $[1,-1,0],[1,0,-1],[0,1,-1],[-1,0,1],[0,-1,1],[-1,1,0]$ *and* $[0,0,0]$*. The multiplicity of* $[0,0,0]$ *in* $\mathcal{M}$ *is* 3 *and in* $\mathcal{M}_0$ *is* 2*. Note that* $E_{ii} \notin \mathcal{M}_0$*. The* D*-weights are* $[2,-3,1],[1,0,-1],[-1,3,-2]$ *and its negatives, and obviously* $[0,0,0]$*.*

The normalizer $N(D)$ gives us an action of $N(D)$ on the weight spaces. If w is a weight-vector of weight χ, $t \in D$ and $g \in N(D)$, then $g \cdot w$ is also a weight vector. Afterall $t \cdot (g \cdot w) = g \cdot t' \cdot w$ where $t' = g^{-1}tg$. Thus

$$t \cdot (g \cdot w) = \chi(t')(g \cdot w)$$

whence $g \cdot w$ must also be a weight-vector with some weight χ'. This χ' is easily computed via the action of D^*. Here the action of $N(D^*)$ is clear: if $\chi = [m_1, \ldots, m_n]$, then $\chi' = [m_{\sigma(1)}, \ldots, m_{\sigma(n)}]$ for some permutation $\sigma \in S_n$ determined by the component of $N(D^*)$ containing g. Thus the map χ to χ' for D-weights in the case of SL_3 is as follows:

$$[m_1 - m_2, 2m_2 - m_1 - m_3, m_3 - m_2] \to [m_{\sigma(1)} - m_{\sigma(2)}, 2m_{\sigma(2)} - m_{\sigma(1)} - m_{\sigma(3)}, m_{\sigma(3)} - m_{\sigma(2)}]$$

Caution: Note that though $\mathbb{Y}^3 \subseteq \mathbb{Z}^3$ is an S_3-invariant subset, the action of S_3 on $\chi \in \mathbb{Y}^3$ is **different**. Note that, e.g., in the last example above, $[2,-3,1]$ is a weight but not the 'permuted' vector $[-3,2,1]$. This is because of our peculiar embedding of $\mathbb{Z}^{n-1} \to \mathbb{Y}^n$.

6 The Null-cone and the Destabilizing flag

The fundamental result of **Hilbert** states:

Theorem 39 *Let* W *be an* SL_n*-module, and let* $w \in W$ *be an element of the null-cone. Then there is a* 1*-parameter subgroup* $\lambda : \mathbb{C}^* \to SL_n$ *such that*

$$\lim_{t \to 0} \lambda(t) \cdot w = 0_W$$

In other words, if the zero-vector 0_W lies in the orbit-closure of w, then there is a 1-parameter subgroup taking it there, in the limit. We will not prove this statement here. Our objective for this section is to interpret the geometric content of the theorem. We will show that there is a *standard form* for an element of the null-cone. For well-known representations, this standard form is easily identified by geometric concepts.

6.1 Characters and the half-space criterion

To begin, let D be the fixed maximal torus. For any $w \in W$, we may express:

$$w = w_1 + w_2 + \ldots + w_r$$

where $w_i \in W_{\chi_i}$, the weight-space for character χ_i. Note the the above expression is unique if we insist that each w_i be non-zero. The set of characters $\{\chi_1, \ldots, \chi_r\}$ will be called the **support** of w and denoted as $supp(w)$. Let $\lambda : \mathbb{C}^* \to SL_n$ be such that $Im(\lambda) \subseteq D$. In this case, the action of $t \in \mathbb{C}^*$ via λ is easily described:

$$t \cdot w = t^{(\lambda,\chi_1)} w_1 + \ldots + t^{(\lambda,\chi_r)} w_r$$

Thus, if $\lim_{t\to 0} t \cdot w$ exists (and is 0_W), then for all $\chi \in supp(w)$, we have $(\lambda, \chi) \geq 0$ (and further $(\lambda, \chi) > 0$).

Note that (λ, χ) is implemented as a linear functional on $\mathbb{Y}^n$. Thus, if $\lim_{t\to 0} t \cdot w$ exists (and is $)_W$) then there is a **hyperplane** in $\mathbb{Y}^n$ such that the support of w is on one side of the hyperplane (**strictly** on one side of the hyperplane). The normal to this hyperplane is given by the conversion of λ into $\mathbb{Y}^n$ notation.

On the other hand if the support $supp(w)$ enjoys the geometric/combinatorial property, then by the approximability of reals by rationals, we see that there is a λ such that $\lim_{t\to 0} t \cdot w$ exists (and is zero).

Thus for 1-parameter subgroups of D, Hilbert's theorem translates into a combinatorial statement on the lattice subset $supp(w) \subset \mathbb{Y}^n$. We call this the (**strict**) **half-space** property. In the general case, we know that given any $\lambda : \mathbb{C}^* \to SL_n$, there is a maximal torus T containing $Im(\lambda)$. By the conjugacy result on maximal tori, we know that $T = ADA^{-1}$ for some $A \in SL_n$. Thus, we may say that w is in the null-cone iff there is a translate $A \cdot w$ such that $supp(A \cdot w)$ satisfies the strict half-space property.

Example 40 *Let us consider SL_3 acting of the space of forms of degree 2. For the standard torus D, the weight-spaces are $\mathbb{C} \cdot X_i^2$ and $\mathbb{C} \cdot X_i X_j$. Consider the form $f = (X_1 + X_2 + X_3)^2$. We see that $supp(f)$ is set of all characters of $Sym^2(\mathbb{C}^3)$ and* **does not** *satisfy the combinatorial property. However, under a basis change A:*

$$\begin{array}{lcl} X_1 & \to & X_1 + X_2 + X_3 \\ X_2 & \to & X_2 \\ X_3 & \to & X_3 \end{array}$$

we see that $A \cdot f = X_1^2$. Thus $A \cdot f$ **does** *satisfy the strict half-space property. Indeed consider the λ*

$$\lambda(t) = \begin{bmatrix} t & 0 & 0 \\ 0 & t^{-1} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

We see that

$$\lim_{t\to 0} t \cdot (A \cdot f) = t^2 X_1^2 = 0$$

Thus we see that every form in the null-cone has a standard form with a very limited sets of possible supports.

Let us look at the module $\mathcal{M}$ of 3×3-matrices under conjugation. Let us fix a λ:

$$\lambda(t) = \begin{bmatrix} t^{n_1} & 0 & 0 \\ 0 & t^{n_2} & 0 \\ 0 & 0 & t^{n_3} \end{bmatrix}$$

such that $n_1 + n_2 + n_3 = 0$. We may assume that $n_1 \geq n_2 \geq n_3$. Looking at the action of $\lambda(t)$ on a general matrix X, we see that:

$$t \cdot X = (t^{n_i - n_j} x_{ij})$$

Thus if $\lim_{t \to 0} t \cdot X$ is to be 0 then $x_{ij} = 0$ for all $i > j$. In other words, X is strictly upper-triangular. Considering the general 1-parameter group tells us that X is in the null-cone iff there is an A such that AXA^{-1} is strictly upper-triangular. In other words, X is **nilpotent**. *The 1-parameter subgroup identifies this transformation and thus the flag of nilpotency.*

6.2 The destabilizing flag

In this section we do a more refined analysis of elements of the null-cone. The basic motivation is to identify a unique set of 1-parameter subgroups which drive a null-point to zero. The simplest example is given by $X_1^2 \in Sym^2(\mathbb{C}^3)$. Let λ, λ' and λ'' be as below:

$$\lambda(t) = \begin{bmatrix} t & 0 & 0 \\ 0 & t^{-1} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \lambda'(t) = \begin{bmatrix} t & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & t^{-1} \end{bmatrix} \quad \lambda''(t) = \begin{bmatrix} t & 0 & 0 \\ 0 & 0 & -1 \\ 0 & t^{-1} & 0 \end{bmatrix}$$

We see that all the three λ, λ' and λ'' drive X_1^2 to zero. The question is whether these are related, and to classify such 1-parameter subgroups. Alternately, one may view this to a more refined classification of points in the null-cone, such as the stratification of the nilpotent matrices by their Jordan canonical form.

There are two aspects to this analysis. Firstly, to identify a metric by which to choose the 'best' 1-parameter subgroup driving a null-point to zero. Next, to show that there is a unique equivalence class of such 'best' subgroups.

Towards the first objective, let $\lambda : \mathbb{C}^* \to SL_n$ be a 1-parameter subgroup. Without loss of generality, we may assume that $Im(\lambda) \subseteq D$. If w is a null-point then we have:

$$t \cdot w = t^{n_1} w_1 + \ldots + t^{n_k} w_k$$

where $n_i > 0$ for all i. Clearly, a measure of how fast λ drives w to zero is $m(\lambda) = \min\{n_1, \ldots, n_k\}$. Verify that this really does not depend on the choice of the maximal torus at all, and thus is well-defined.

Next, we see that for a λ as above, we consider $\lambda^2 : \mathbb{C}^* \to SL_n$ such that $\lambda^2(t) = \lambda(t^2)$. It is easy to see that $m(\lambda^2) = 2 \cdot m(\lambda)$. Clearly, λ and λ^2 are intrinsically identical and we would like to have a measure invariant under such scaling. This comes about by associating a length to each λ. Let λ be as above and let $Im(\lambda) \subseteq D$. Then, there are integers $a_1, \ldots, a_n$ such that

$$\lambda(t) = \begin{bmatrix} t^{a_1} & 0 & 0 & 0 \\ 0 & t^{a_2} & 0 & 0 \\ & & \vdots & \\ 0 & 0 & 0 & t^{a_n} \end{bmatrix}$$

We define $\|\lambda\|$ as

$$\|\lambda\| = \sqrt{a_1^2 + \ldots + a_n^2}$$

We must now show that this does not depend on the choice of the maximal torus D. Let $\mathcal{T}(SL_n)$ denotes the collection of all maximal tori of SL_n as abstract subgroups. For every $A \in SL_n$, we may define the map $\phi_A : \mathcal{T} \to \mathcal{T}$ defined by $T \to ATA^{-1}$. The stabilizer of a torus T for this action of SL_n is clearly $N(T)$, the normalizer of T. Also recall that $N(T)/T = W$ is the (discrete) weyl group. Let $Im(\lambda) \subseteq D$ and $Im(\lambda) \subseteq D'$ for some two maximal tori D and D'. Since there is an A such that $AD'A^{-1} = D$, it is clear that $\|\lambda\| = \|A\lambda A^{-1}\|$. Thus, we are left to check if $\|\lambda'\| = \|\lambda\|$ when (i) $Im(\lambda), Im(\lambda') \subseteq D$, and (ii) $\lambda' = A\lambda A^{-1}$ for some $A \in SL_n$. This throws the question to invariance of $\|\lambda\|$ under $N(D)$, or in other words, symmetry under the weyl group. Since $W \cong S_n$, the symmetric group, and since $\sqrt{a_1^2 + \ldots + a_n^2}$ is a symmetric function on $a_1, \ldots, a_n$, we have that $\|\lambda\|$ is well-defined.

We now define the *efficiency* of λ on a null-point w to be

$$e(\lambda) = \frac{m(\lambda)}{\|\lambda\|}$$

We see immediately that $e(\lambda) = e(\lambda^2)$.

Lemma 41 *Let W be a representation of SL_n and let $w \in W$ be a null-point. Let $\mathcal{N}(w, D)$ be the collection of all $\lambda : \mathbb{C}^* \to D$ such that $\lim_{t\to 0} t \cdot w = 0_W$. If $\mathcal{N}(w, D)$ is non-empty then there is a* unique *$\lambda' \in \mathcal{N}(w, D)$ which drives w to zero and maximizes the efficiency, i.e., $e(\lambda') > e(\lambda)$ for all $\lambda \in \mathcal{N}(w, D)$ and $\lambda \neq (\lambda')^k$ for any $k \in \mathbb{Z}$. This 1-parameter subgroup will be denoted by $\lambda(w, D)$.*

Proof: Suppose that $\mathcal{N}(w, D)$ is non-empty. Then in the weight-space expansion of w for the maximal torus D, we see that $supp(w)$ satisfies the half-space property for some $\lambda \in \mathbb{Y}^n$. Note that the $\lambda \in \mathcal{N}(w, D)$ are parametrized by lattice points $\lambda \in \mathbb{Y}^n$ such that $(\lambda, \chi) > 0$ for all $\chi \in supp(w)$. Let $Cone(w)$ be the conical combination (over $\mathbb{R}$) of all $\chi \in supp(w)$ and $Cone(w)^o$ its **polar**. Thus, in other words, $\mathcal{N}(w, D)$ is precisely the collection of lattice points in the cone $Cone(w)^o$. Next, we see that $e(\lambda)$ is a convex function on $Cone(w)^o$ which is contant over rays $\mathbb{R}^+ \cdot \lambda$ for all $\lambda \in Cone(w)^o$. By a routine analysis, the maximum of such a function must be a unique ray with rational entries. This proves the lemma. $\square$

This covers one important part in our task of identifying the 'best' 1-parameter subgroup driving a null-point to zero. The next part is to relate D to other maximal tori.

Let $\lambda : \mathbb{C}^* \to SL_n$ and let $P(\lambda)$ be defined as follows:

$$P(\lambda) = \{A \in SL_n | \lim_{t\to 0} \lambda(t) A \lambda(t^{-1}) = I \in SL_n\}$$

Having fixed a maximal torus D containing $Im(\lambda)$, we easily identify $P(\lambda)$ as a **parabolic** subgroup, i.e., block upper-triangular. Indeed, let

$$\lambda(t) = \begin{bmatrix} t^{a_1} & 0 & 0 & 0 \\ 0 & t^{a_2} & 0 & 0 \\ & & \vdots & \\ 0 & 0 & 0 & t^{a_n} \end{bmatrix}$$

with $a_1 \geq a_2 \geq \ldots \geq a_n$ (obviously with $a_1 + \ldots + a_n = 0$). Then

$$P(\lambda) = \{(x_{ij}) | x_{ij} = 0 \text{ for all } i, j \text{ such that } a_i < a_j\}$$

The **unipotent radical** $U(\lambda)$ is a normal subgroup of $P(\lambda)$ defined as:

$$U(\lambda) = (x_{ij}) \text{ where } x_{ij} = \begin{cases} x_{ij} = 0 & \text{if } a_i < a_j \\ x_{ij} = \delta_{ij} & \text{if } a_i = a_j \end{cases}$$

Lemma 42 *Let $\lambda \in \mathcal{N}(w, D)$ and let $g \in P(\lambda)$, then (i)$g\lambda g^{-1} \subseteq P(\lambda)$ and $P(g\lambda g^{-1}) = P(\lambda)$, (ii) $g\lambda g^{-1} \in \mathcal{N}(w, gDg^{-1})$, and (iii) $e(\lambda) = e(g\lambda g^{-1})$.*

This actually follows from the construction of the explicit SL_n-modules and is left to the reader. We now come to the unique object that we will define for every $w \in W$ in the null-cone. This is the parabolic subgroup $P(\lambda)$ for any 'best' λ. We have already seen above that if λ' is a $P(\lambda)$-conjugate of a best λ then λ' is 'equally best' and $P(\lambda') = P(\lambda)$.

We now relate two general equally best λ and λ'. For this we need a preliminary definition and a lemma:

Definition 43 *Let V be a vector space over $\mathbb{C}$. A* **flag** *$\mathcal{F}$ of V is a sequence $(V_0, \ldots, V_r)$ of nested subspaces $0 = V_0 \subset V_1 \subset \ldots \subset V_r = V$.*

Lemma 44 *Let $dim_{\mathbb{C}}(V) = r$ and let $\mathcal{F} = (V_0, \ldots, V_r)$ and $\mathcal{F}' = (V_0', \ldots, V_r')$ be two (complete) flags for V. Then there is a basis $b_1, \ldots, b_r$ of V and a permutation $\sigma \in S_n$ such that $V_i = \overline{\{b_1, \ldots, b_i\}}$ and $V_i' = \overline{\{b_{\sigma(1)}, \ldots, b_{\sigma(i)}\}}$ for all i.*

This is proved by induction on r.

Corollary 45 *Let λ and λ' be two 1-parameter subgroups and $P(\lambda)$ and $P(\lambda')$ be their corresponding parabolic subgroups. Then there is a maximal torus T of SL_n such that $T \in P(\lambda) \cap P(\lambda')$.*

Proof: It is clear that there is a correspondence between parabolic subgroups of SL_n and flags. We refine the flags associated to the parabolic subgroups $P(\lambda)$ and $P(\lambda')$ to complete flags and apply the above lemma. □

We are now prepared to prove Kempf's theorem:

Theorem 46 *Let W be a representation of SL_n and $w \in W$ a null-point. Then there is a 1-parameter subgroup $\lambda \in \Gamma(SL_n)$, such that (i) for all $\lambda' \in \Gamma(SL_n)$, we have $e(\lambda) \geq e(\lambda')$, and (ii) for all λ' such that $e(\lambda) = e(\lambda')$ we have $P(\lambda) = P(\lambda')$ and that there is a $g \in P(\lambda)$ such that $\lambda' = g\lambda g^{-1}$.*

Proof: Let $\mathcal{N}(w)$ be all elements of $\Gamma(SL_n)$ which drive w to zero. Let $\Xi(W)$ be the (finite) collection of D-characters appearing in the representation W. For every $\lambda(w, T)$ such that $Im(\lambda) \subseteq T$, we may consider an $A \in SL_n$ such that $A\lambda A^{-1} \in \mathcal{N}(A \cdot w, D)$ and $e(\lambda) = e(A\lambda A^{-1})$. Since the 'best' element of $\mathcal{N}(A \cdot w, D)$ is determined by $supp(A.w) \subseteq \Xi$, we see that there are only finitely many possibilities for $e(A \cdot w, A\lambda A^{-1})$ and therefore for $e(\lambda)$ for 'best' λ driving w to zero.

Thus the length k of any sequence $\lambda(w,T_1),\ldots,\lambda(w,T_k)$ such that $e(\lambda(w,T_1)) < \ldots < e(\lambda(w,T_k))$ must be bounded by the number $2^{|\Xi|}$. This proves (i).

Next, let $\lambda_1 = \lambda(w,T_1)$ and $\lambda_2 = \lambda(w,T_2)$ be two 'best' elements of $\mathcal{N}(w,T_1)$ and $\mathcal{N}(w,T_2)$ respectively. By lemma 45, we have a torus, say D, and $P(\lambda(w,T_i))$-conjugates λ_i such that (i) $e(\lambda_i) = e(\lambda(w,T_i))$ and (ii) $Im(\lambda_i) \subseteq D$. By lemma 41, we have $\lambda_1 = \lambda_2$ and thus $P(\lambda_1) = P(\lambda_2)$. On the other hand, $P(\lambda(w,T_i)) = P(\lambda_i)$ and this proves (ii). $\Box$

Thus Theorem 46 associates a unique parabolic subgroup $P(w)$ to every point in the null-cone. This subgroup is called the **destabilizing flag** of w. Clearly, if w is in the null-cone then so is $A \cdot w$, where $A \in SL_n$. Furthermore, it is clear that $P(A \cdot w) = AP(w)A^{-1}$.

Corollary 47 *Let $w \in W$ be in the null-cone and let $G_w \subseteq SL_n$ stabilize w. Then $G_w \subseteq P(w)$.*

Proof: Let $g \in G_w$. Since $g \cdot w = w$, we see that $gP(w)g^{-1} = P(w)$, and that g normalizes $P(w)$. Since the normalizer of any parabolic subgroup of SL_n is itself, we see that $g \in P(w)$. $\Box$

7 Stability

Recall that $z \in W$ is stable iff its orbit $O(z)$ is closed in W. In the last section, we tackled the points in the null-cone, i.e., points in the set $[0_W]_{\approx}$, or in other words, points which close onto the stable point 0_W. A similar analysis may be done for arbitrary stable points.

Following Kempf, let $S \subseteq W$ be a closed SL_n-invariant subset. Let $z \in W$ be arbitrary. If the orbit-closure $\Delta(z)$ intersects S, then we will associate a unique parabolic subgroup $P_{z,S} \subseteq SL_n$ as a witness to this fact. The construction of this parabolic subgroup is in several steps.

As the first step, we construct a representation X of G and a closed SL_n-equivariant embedding $\phi : W \to X$ such that $\phi^{-1}(0_X) = S$ scheme-theoretically. This may be done as follows: since S is a closed sub-variety of W, there is an ideal $I_S = (f_1,\ldots,f_k)$ of definition for S. We may further assume that the vector space $\overline{\{f_1,\ldots,f_k\}}$ is itself an SL_n-module, say X. We assume that X is k-dimensional.

We now construct that map $\phi : W \to X$ as follows:

$$\phi(w) = (f_1(x),\ldots,f_k(x))$$

Note that $\phi(S) = 0_X$ and that $I_S = (f_1,\ldots,f_k)$ ensures that the requirements on our ϕ do hold.

Next, there is an adaptation of Hilbert's Theorem 39 which we do not prove:

Theorem 48 *Let W be an SL_n-module, and let $y \in W$ be a stable point. Let $z \in [y]_{\approx}$ be an element which closes onto y. Then there is a 1-parameter subgroup $\lambda : \mathbb{C}^* \to SL_n$ such that*

$$\lim_{t\to 0} \lambda(t) \cdot w \in O(y)$$

Thus the limit exists and lies in the closed orbit of y.

Now suppose that $\Delta(z) \cap S$ is non-empty. Then there must be a stable $y \in S$ such that $y \in \Delta(z)$. We apply the above theorem to y and obtain the λ as above. This shows that there is indeed a 1-parameter subgroup driving z into S. Next, it is easy to see that

$$\lim_{t \to 0} [\lambda(t) \cdot \phi(z)] = 0_X$$

Thus, $\phi(z)$ actually lies in the null-cone of X. We may now be tempted to employ the techniques of the previous section to come up with the 'best' λ and its parabolic, now called $P(z, S)$. This is almost the technique that is to be adopted, except that this 'best' λ drives $\phi(z)$ into 0_X, but $\lim_{t \to 0}[\lambda(t) \cdot z]$ (which is supposed to be in S) may not exist! This is because we are using the unproved (and untrue) converse of the assertion that 1-parameter subgroups which drive z into S drive $\phi(z)$ into 0_X.

This above argument is rectified by limiting the domain of allowed 1-parameter subgroups to (i) $Cone(supp(\phi(z))^o$ as before, and (ii) those λ such that $\lim_{t \to 0}[\lambda(t) \cdot z]$ exists. This second condition is also a 'convex' condition and then a 'best' λ does exist. This completes the construction of $P(z, S)$.

As before, if $G_z \subseteq SL_n$ stabilizes z then it normalizes $P(z, S)$ thus must be contained in it:

Proposition 49 *If G_z stabilizes z then $G_z \subseteq P(z, S)$.*

Let us now consider the **permanent** and the **determinant**. Let $\mathcal{M}$ be the n^2-dimensional space of all $n \times n$-matrices. Since *det* and *perm* are homogeneous n-forms on $\mathcal{M}$, we consider the $SL(\mathcal{M})$-module $W = Sym^n(\mathcal{M}^*)$. We recall now certain stabilizing groups of the *det* and the *perm*. We will need the definition of a certain group L'. This is defined as the group generated by the **permutation matrices** and the diagonal matrices in GL_n. In other words, L' is the normalizer of the complete standard torus $D^* \subseteq GL_n$. L is defined as that subgroup of L' which is contained in SL_n.

Proposition 50

(A) *Consider the group $K = SL_n \times SL_n$. We define the action μ_K of a typical element $(A, B) \in K$ on $X \in \mathcal{M}$ as given by:*

$$X \to AXB^{-1}$$

Then (i) $\mathcal{M}$ is an irreducible representation of K and $Im(K) \subseteq SL(\mathcal{M})$, and (ii) K stabilizes the determinant.

(B) *Consider the group $H = L \times L$. We define the action μ_H of a typical element $(A, B) \in H$ on $X \in \mathcal{M}$ as given by:*

$$X \to AXB^{-1}$$

Then (i) $\mathcal{M}$ is an irreducible representation of H and $Im(H) \subseteq SL(\mathcal{M})$, and (ii) H stabilizes the permanent.

We are now ready to show:

Theorem 51 *The points det and perm in the $SL(\mathcal{M})$-module $W = Sym^n(\mathcal{M}^*)$ are stable.*

Proof: Lets look at det. If it were not stable, then there would be a closed $SL(\mathcal{M})$-invariant subset $S \subset W$ such that the $det \notin S$ but closes onto S: just take S to be the unique closed orbit in $[det]_{\approx}$. Whence there is a parabolic $P(det, S)$ which, by proposition 49, would contain K. This would mean that there is a K-invariant flag in $\mathcal{M}$ corresponding to $P(det, S)$. This contradicts the irreducibility of $\mathcal{M}$ as a K-module. $\square$