LECTURES ON REPRESENTATION THEORY AND INVARIANT THEORY

These are the notes for a lecture course on the symmetric group, the general linear group and invariant theory. The aim of the course was to cover as much of the beautiful classical theory as time allowed, so, for example, I have always restricted to working over the complex numbers. The result is a course which requires no previous knowledge beyond a smattering of rings and modules, character theory, and affine varieties.

These are certainly not the first notes on this topic, but I hope they may still be useful, for [Dieudonné and Carrell] has a number of flaws, and [Weyl], although beautifully written, requires a lot of hard work to read. The only new part of these notes is our treatment of Gordan's Theorem in §13: the only reference I found for this was [Grace and Young] where it was proved using the symbolic method.

The lectures were given at Bielefeld University in the winter semester 1989-90, and this is a more or less faithful copy of the notes I prepared for that course. I have, however, reordered some of the parts, and rewritten the section on semisimple algebras.

The references I found most useful were:

[H. Boerner] "Darstellungen von Gruppen" (1955). English translation "Representations of groups" (North-Holland, 1962,1969).

[J. A. Dieudonné and J. B. Carrell] "Invariant Theory, Old and New," Advances in Mathematics 4 (1970) 1-80. Also published as a book (1971).

[J. H. Grace and A. Young] "The algebra of invariants" (1903, Reprinted by Chelsea).

[H. Weyl] "The classical groups" (Princeton University Press, 1946).

My thanks go to A. J. Wassermann who suggested that I give such a course and explained some of the central ideas to me, and also to my students, both for their patience and for pointing out many inaccuracies in the original version of these notes.

William Crawley-Boevey, April 1990

Fakultät für Mathematik
Universität Bielefeld
Postfach 8640
4800 Bielefeld 1
West Germany

Mathematical Institute
Oxford University
24-29 St. Giles
Oxford OX1 3LB
England

CONTENTS

The Symmetric Group

1. Semisimple algebras 1
2. Young symmetrizers 7
3. Standard Tableaux 13
4. A Character Formula 17
5. The Hook Length Formula 26

The General Linear Group

6. Multilinear and Polynomial Algebra 28
7. Schur-Weyl Duality 33
8. Decomposition of Tensors 36
9. Rational Representations of GL(V) 40
10. Weyl’s Character Formula 44

Invariant Theory

11. Some Examples of Invariants 48
12. The First Fundamental Theorem of Invariant Theory 57
13. Covariants of Bilinear Forms 65

§1. SEMISIMPLE ALGEBRAS

The facts about semisimple algebras which we shall need for the symmetric group should be well-known, and need not be repeated here. For the general linear group, however, we shall need some more delicate results, so some presentation is necessary. Not knowing what to include and what to exclude, we give a very quick development of the whole theory here.

All rings R are associative and have an identity which is denoted by 1 or 1_R. By "module" we always mean left module.

Definition. A $\mathbb{C}$-algebra is ring R which is also a $\mathbb{C}$-vector space with the same addition, satisfying

$$\lambda(rr') = (\lambda r)r' = r(\lambda r') \quad \forall\ r,r'\in R \text{ and } \lambda\in\mathbb{C}.$$

One has the obvious notions of subalgebras and algebra homomorphisms. We shall be particularly interested in the case when R is finite dimensional.

Remarks and Examples.

(1) $\mathbb{C}$ is a $\mathbb{C}$-algebra. $M_n(\mathbb{C})$ is a $\mathbb{C}$-algebra. If G is a group, then the group algebra $\mathbb{C}G$ is the $\mathbb{C}$-algebra with basis the elements of G and multiplication lifted from G.

(2) If R and S are $\mathbb{C}$-algebras, then so is $R\times S$. Here the vector space structure comes from the identification of $R\times S$ with $R\oplus S$.

(3) If 1 = 0 in R then R is the zero ring. Otherwise, for $\lambda,\mu\in\mathbb{C}$ we have $\lambda 1_R = \mu 1_R \Leftrightarrow \lambda = \mu$ so we can identify $\lambda\in\mathbb{C}$ with $\lambda 1_R \in R$. This makes $\mathbb{C}$ a subalgebra of R.

(4) If M is an R-module, then it becomes a $\mathbb{C}$-vector space via

$$\lambda m = (\lambda 1_R)m \quad \forall\ \lambda\in\mathbb{C},\ m\in M.$$

If N is another R-module, then $\mathrm{Hom}_R(M,N)$ is a $\mathbb{C}$-subspace of $\mathrm{Hom}_{\mathbb{C}}(M,N)$. In particular it is a $\mathbb{C}$-vector space. The structure is given by

$$(\lambda f)(m) = \lambda f(m) = f(\lambda m) \text{ for } m\in M,\ \lambda\in\mathbb{C} \text{ and } f \in \mathrm{Hom}_R(M,N).$$

(5) If M is an R-module, then $\mathrm{End}_R(M)$ is a $\mathbb{C}$-algebra, with multiplication given by composition

$$(fg)(m) = f(g(m)) \text{ for } m \in M,\ f,g \in \mathrm{End}_R(M)$$

In particular, if V is a $\mathbb{C}$-vector space, then $\mathrm{End}_{\mathbb{C}}(V)$ is a $\mathbb{C}$-algebra. Of course $\mathrm{End}_{\mathbb{C}}(V) \cong M_{\dim V}(\mathbb{C})$.

(6) If R is a $\mathbb{C}$-algebra and $X \subseteq R$ is a subset, then the centralizer

$$c_R(X) = \{\ r \in R \mid rx=xr\ \forall\ x \in X\ \}$$

of X in R is a $\mathbb{C}$-subalgebra of R. In particular this holds for the centre $c_R(R)$ of R.

(7) If R is a $\mathbb{C}$-algebra and M is an R-module, then the map

$$\alpha : R \longrightarrow \mathrm{End}_{\mathbb{C}}(M), \quad r \longmapsto (m \longmapsto rm)$$

is a $\mathbb{C}$-algebra map. By definition

$$\mathrm{End}_R(M) = c_{\mathrm{End}_{\mathbb{C}}(M)}(\alpha(R)),$$

so that

$$\alpha(R) \subseteq c_{\mathrm{End}_{\mathbb{C}}(M)}(\mathrm{End}_R(M)).$$

(8) If M is an R-module, then it is naturally an $\mathrm{End}_R(M)$-module with the action given by evaluation, and this action commutes with that of R. The inclusion above says that the elements of R act as $\mathrm{End}_R(M)$-module endomorphisms of M. If N is another R-module, then $\mathrm{Hom}_R(N,M)$ is also an $\mathrm{End}_R(M)$-module, with the action given by composition.

Lemma 1. If R is a finite dimensional $\mathbb{C}$-algebra, then there are only finitely many isomorphism classes of simple R-modules, and they are finite dimensional.

PROOF. If S is a simple module, pick $0 \neq s \in S$ and define a map $R \longrightarrow S$ sending r to rs. This is an R-module map, and the image is non-zero, so it is all of S. Thus $\dim_{\mathbb{C}} S \leq \dim_{\mathbb{C}} R < \infty$. Moreover, S must occur in any composition series of R, so by the Jordan-Hölder Theorem there are only finitely many isomorphism classes of simple modules.

Schur's Lemma. Let R be a finite dimensional $\mathbb{C}$-algebra.

(1) If $S \not\cong T$ are non-isomorphic simple modules then $\mathrm{Hom}_R(S,T)=0$.

(2) If S is a simple R-module, then $\mathrm{End}_R(S) \cong \mathbb{C}$.

PROOF. (2) The usual arguments show that $D = \mathrm{End}_R(S)$ is a division ring.

Since S is finite dimensional, so also is D, and therefore if $d \in D$ the elements $1, d, d^2, \ldots$ cannot all be linearly independant, and so there is some non-zero polynomial $p(X)$ over $\mathbb{C}$ with $p(d)=0$. Since $\mathbb{C}$ is algebraically closed this polynomial is a product of linear factors

$$p(X) = c(X-a_1)\ldots(X-a_n), \quad 0 \neq c \in \mathbb{C},\ a_1, \ldots, a_n \in \mathbb{C}$$

so $(d-a_1 1_D)\ldots(d-a_n 1_D)=0$. Now D has no zero-divisors, so one of the terms must be zero. Thus $d=a_i 1_D \in \mathbb{C}1_D$. Since d was arbitrary, $D=\mathbb{C}1_D$.

Definition. An R-module is semisimple if it is a direct sum of simple submodules.

Lemma 2. Submodules, quotients and direct sums of semisimple modules are again semisimple. Every submodule of a semisimple module is a summand.

PROOF. Omitted.

Definition. If R and S are $\mathbb{C}$-algebras, and M and N are an R-module and an S-module, then $M \otimes N$ (the tensor product over $\mathbb{C}$) has the structure of an R-module given by

$$r(m \otimes n) = rm \otimes n \quad \text{for } r \in R,\ m \in M,\ n \in N.$$

and the structure of an S-module given by

$$s(m \otimes n) = m \otimes sn \quad \text{for } s \in S,\ m \in M,\ n \in N.$$

Remarks.

(1) This is completely different to the tensor product of two $\mathbb{C}G$-modules which we shall consider later.

(2) These two actions commute, since

$$r(s(m \otimes n)) = r(m \otimes sn) = rm \otimes sn = s(rm \otimes n) = s(r(m \otimes n)).$$

Thus the images of R and S in $\mathrm{End}_{\mathbb{C}}(M \otimes N)$ commute.

(3) If N has basis $e_1, \ldots, e_m$ then the map

$$M \oplus \ldots \oplus M \longrightarrow M \otimes N,\ (m_1, \ldots, m_m) \longmapsto m_1 \otimes e_1 + \ldots + m_m \otimes e_m$$

is an isomorphism of R-modules. Similarly, if M has basis $f_1, \ldots, f_\ell$ then the map

$$N \oplus \ldots \oplus N \longrightarrow M \otimes N,\ (n_1, \ldots, n_\ell) \longmapsto f_1 \otimes n_1 + \ldots + f_\ell \otimes n_\ell$$

is an isomorphism of S-modules.

Lemma 3. If M is a semisimple R-module, then the evaluation map

$$\bigoplus_S S \otimes \mathrm{Hom}_R(S,M) \longrightarrow M$$

is an isomorphism of R-modules and of $\mathrm{End}_R(M)$-modules. Here S runs over a complete set of non-isomorphic simple R-modules, and we are using the action of R on S and of $\mathrm{End}_R(M)$ on $\mathrm{Hom}_R(S,M)$.

PROOF. The map is indeed an R-module map and an $\mathrm{End}_R(M)$-module map. To see that it is an isomorphism of vector spaces, one can reduce to the case when M is simple, in which case it follows from Schur's Lemma.

Lemma 4. If M is a finite dimensional semisimple R-module, then

$$\mathrm{End}_R(M) \cong \prod_S \mathrm{End}_{\mathbb{C}}(\mathrm{Hom}_R(S,M))$$

where S runs over a complete set of non-isomorphic simple R-modules.

PROOF. The product $E = \prod_S \mathrm{End}_{\mathbb{C}}(\mathrm{Hom}_R(S,M))$ acts naturally on $\oplus_S S\otimes\mathrm{Hom}_R(S,M)$, and since this action commutes with that of R, there is a homomorphism $E\longrightarrow\mathrm{End}_R(M)$ which is in fact injective. To show that it is an isomorphism we count dimensions. By Lemma 3, M is isomorphic to the direct sum of $\dim_{\mathbb{C}}\mathrm{Hom}(S,M)$ copies of each simple module S, and so by Schur's Lemma,

$$\dim_{\mathbb{C}}\mathrm{End}_R(M) \cong \textstyle\sum_S (\dim_{\mathbb{C}}\mathrm{Hom}(S,M))^2,$$

which is also the dimension of E.

Definition. A finite dimensional $\mathbb{C}$-algebra R is semisimple if R is a semisimple R-module.

Remarks.

(1) If R is a semisimple algebra, then any R-module is semisimple, for any module is a quotient of a free module, but these are semisimple.

(2) If G is a finite group then $\mathbb{C}G$ is semisimple by Maschke's Theorem.

Artin-Wedderburn Theorem. Any finite dimensional semisimple $\mathbb{C}$-algebra is isomorphic to a product

$$R = \prod_{i=1}^{h} \mathrm{End}_{\mathbb{C}}(V_i)$$

where the V_i are finite dimensional vector spaces. Conversely, if R has this form it is semisimple, the non-zero V_i form a complete set of non-isomorphic simple R-modules, and as an R-module, R is isomorphic to the direct sum of $\dim_{\mathbb{C}}(V_i)$ copies of each V_i.

PROOF. We prove the assertions about the product first. Clearly we may suppose that all $V_i \neq 0$.

The V_i are naturally R-modules, with the factors other than $\mathrm{End}_{\mathbb{C}}(V_i)$ acting as zero. Now $GL(V_i)$, and hence also R, acts transitively on $V_i \backslash \{0\}$, and it follows that the V_i are simple R-modules.

If the V_i have bases $(e_{i,1},...,e_{i,m_i})$, then the map

$$\begin{array}{ccc} R & \longrightarrow & V_1 \oplus ... \oplus V_1 \oplus ... \oplus V_h \oplus ... \oplus V_h, \\ (f_1,..,f_h) & \longmapsto & (f_1(e_{1,1}),..,f_1(e_{1,m_1}),...,f_h(e_{h,1}),..,f_h(e_{h,m_h})) \end{array}$$

is an R-module map, and is injective, so is an isomorphism by dimensions. Thus R is semisimple.

The V_i are a complete set of simple R-modules by the Jordan-Hölder Theorem, as in Lemma 1, and they are non-isomorphic since if $i \neq j$ then the element $(0,..,0,1,0,..,0) \in R$ (with the 1 in the i-th place) annihilates V_j but not V_i.

If R is any ring then the natural map

$$\alpha : R \longrightarrow \mathrm{End}_{\mathrm{End}_R(R)}(R), \quad r \longmapsto (x \longmapsto rx)$$

is an isomorphism, for it is certainly injective, and if θ lies in the right hand side, then it commutes with the endomorphisms

$$f_r \in \mathrm{End}_R(R),\ f_r(x) = xr.$$

Now if $r \in R$ then

$$\theta(r) = \theta(1r) = \theta(f_r(1)) = f_r(\theta(1)) = \theta(1)r,$$

so θ acts as left multiplication by $\theta(1) \in R$, and hence $\theta = \alpha(\theta(1)) \in \alpha(R)$, so α is surjective.

If now R is a semisimple $\mathbb{C}$-algebra then $\mathrm{End}_R(R)$ is semisimple by Lemma 4 and the proof above, so R is a semisimple $\mathrm{End}_R(R)$-module. A second application of Lemma 4 and the isomorphism α shows that R has the required form.

Lemma 5. If R is semisimple and M is a finite dimensional R-module then the natural map

$$\alpha : R \longrightarrow \mathrm{End}_{\mathrm{End}_R(M)}(M), \quad r \longmapsto (m \longmapsto rm)$$

is surjective.

PROOF. The kernel of this map is the annihilator

$$I = \{r \in R \mid rM = 0\}$$

of M. Now R/I is semisimple, M is an R/I-module and

$$\mathrm{End}_R(M) = \mathrm{End}_{R/I}(M),$$

so we can replace R by R/I and hence we may suppose that M is faithful and that α is injective.

By Lemma 4, $\mathrm{End}_R(M) \cong \prod_S \mathrm{End}_{\mathbb{C}}(\mathrm{Hom}_R(S,M))$, and since M is faithful, all the spaces $\mathrm{Hom}_R(S,M)$ are non-zero, so they are precisely the simple $\mathrm{End}_R(M)$-modules.

By Lemma 3, M is isomorphic as an $\mathrm{End}_R(M)$-module to $\oplus_S\ S \otimes \mathrm{Hom}_R(S,M)$, so it is the direct sum of $\dim_{\mathbb{C}} S$ copies of the simple module $\mathrm{Hom}_R(S,M)$ for each S. As in Lemma 4 this implies that

$$\dim_{\mathbb{C}} \mathrm{End}_{\mathrm{End}_R(M)}(M) = \textstyle\sum_S (\dim_{\mathbb{C}} S)^2,$$

but this is the dimension of R, so α is an isomorphism.

Finally, we note the following fact

Lemma 6. If R is a $\mathbb{C}$-algebra and $h \in R$ is an element with $h^2 = \lambda h$ for some non-zero $\lambda \in \mathbb{C}$, then for any R-module M we have an isomorphism

$$\mathrm{Hom}_R(Rh, M) \cong hM$$

of $\mathrm{End}_R(M)$-modules.

PROOF. Note first that hM is an $\mathrm{End}_R(M)$-submodule of M, since if $\theta \in \mathrm{End}_R(M)$ and $hm \in hM$ then $\theta(hm) = h\theta(m) \in hM$. Replacing h by h/λ we may suppose that h is idempotent. Now we have an $\mathrm{End}_R(M)$-module map

$$\mathrm{Hom}_R(Rh, M) \longrightarrow hM, \ f \longmapsto f(h)$$

with inverse

$$hM \longrightarrow \mathrm{Hom}_R(Rh, M), \ m \longmapsto (r \longmapsto rm),$$

as required.

§2. YOUNG SYMMETRIZERS.

Recall that the representations $\rho: G \longrightarrow GL(V)$ of a group G correspond to $\mathbb{C}G$-modules by setting $gv = \rho(g)(v)$ for $g \in G$ and $v \in V$. The trivial representation is the map $G \longrightarrow \mathbb{C}^{\times}$ sending all $g \in G$ to 1; the corresponding $\mathbb{C}G$-module is denoted by $\mathbb{C}$.

The symmetric group is

$$S_n = \left\{ \text{bijections } \{1,\ldots n\} \longrightarrow \{1,\ldots,n\} \right\}$$

with multiplication given by composition. In this section we compute its representations using certain elements of the group algebra $\mathbb{C}S_n$ called Young Symmetrizers. One representation, the signature

$$\varepsilon : S_n \longrightarrow \{\pm 1\} \quad \text{given by} \quad \varepsilon_\sigma = \prod_{1 \le i < j \le n} \frac{(\sigma i - \sigma j)}{(i-j)},$$

is of course well-known.

For convenience, in this section we set $A = \mathbb{C}S_n$. This is a finite dimensional semisimple $\mathbb{C}$-algebra by Maschke's Theorem.

Definition. A partition of n is a sequence $\lambda = (\lambda_1, \lambda_2, \ldots)$ with $\lambda_i \in \mathbb{N}$, $\lambda_1 \ge \lambda_2 \ge \ldots$ and $\sum_{i=1}^{\infty} \lambda_i = n$. The partitions of n are ordered lexicographically, so that

$$\lambda < \mu \iff \exists\ i \in \mathbb{N} \text{ such that } \lambda_j = \mu_j \text{ for } j < i \text{ and } \lambda_i < \mu_i$$

This is a total ordering on the set of partitions of n.

Example. The partitions of 5 are

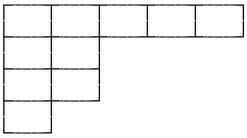
$$(1^5) < (2,1^3) < (2^2,1) < (3,1^2) < (3,2) < (4,1) < (5).$$

Definition. If λ is a partition of n, then the Young frame $[\lambda]$ of λ is the subset

$$\{ (i,j) \mid i \ge 1,\ 1 \le j \le \lambda_i \} \subset \mathbb{N} \times \mathbb{N}$$

We draw a picture for this. For example

$[(5,2^2,1)]$ =



Definition. A Young tableau Σ_λ is a bijection $[\lambda] \longrightarrow \{1,...,n\}$. For example we might have

$$\Sigma_{(5,2^2,1)} \quad = \quad \begin{array}{|c|c|c|c|c|} \hline 9 & 5 & 2 & 7 & 10 \\ \hline 8 & 1 & \multicolumn{3}{c}{} \\ \cline{1-2} 4 & 6 & \multicolumn{3}{c}{} \\ \cline{1-2} 3 & \multicolumn{4}{c}{} \\ \cline{1-1} \end{array}$$

If Σ_λ is a Young tableau and $\sigma \in S_n$ we define a Young tableau $\sigma\Sigma_\lambda$ by $(\sigma\Sigma_\lambda)(x) = \sigma(\Sigma_\lambda(x))$ for $x \in [\lambda]$.

For each partition we need to pick one representative of all the corresponding tableaux, so for definiteness we denote by Σ^0_λ the Young tableau numbered in the order that one reads a book. For example

$$\Sigma^0_{(5,2^2,1)} \quad = \quad \begin{array}{|c|c|c|c|c|} \hline 1 & 2 & 3 & 4 & 5 \\ \hline 6 & 7 & \multicolumn{3}{c}{} \\ \cline{1-2} 8 & 9 & \multicolumn{3}{c}{} \\ \cline{1-2} 10 & \multicolumn{4}{c}{} \\ \cline{1-1} \end{array}$$

We define subgroups $\mathrm{Row}(\Sigma_\lambda)$ and $\mathrm{Col}(\Sigma_\lambda)$ of S_n by

$\sigma \in \mathrm{Row}(\Sigma_\lambda) \Leftrightarrow$ each $i \in \{1,...,n\}$ is in the same row of Σ_λ and $\sigma\Sigma_\lambda$.

$\sigma \in \mathrm{Col}(\Sigma_\lambda) \Leftrightarrow$ each $i \in \{1,...,n\}$ is in the same column of Σ_λ and $\sigma\Sigma_\lambda$.

Definition. If Σ_λ is a Young tableau, then the Young symmetrizer is the element

$$h(\Sigma_\lambda) = \sum_{r \in \mathrm{Row}(\Sigma_\lambda)\ c \in \mathrm{Col}(\Sigma_\lambda)} \varepsilon_c \, rc$$

of A. We also set $h_\lambda = h(\Sigma^0_\lambda)$. The rest of this section is devoted to showing that the left ideals in A of the form Ah_λ with λ running through the partitions of n, are a complete set of non-isomorphic simple A-modules.

Examples.

(1) If $\lambda = (n)$ then $h = h(\Sigma_\lambda) = \sum_{\sigma \in S_n} \sigma$ is the symmetrizer in A. Clearly $gh = h$ for all $g \in S_n$, so $Ah = \mathbb{C}h$ is the trivial representation of S_n.

(2) If $\lambda = (1^n)$ then $h = h(\Sigma_\lambda) = \sum_{\sigma \in S_n} \varepsilon_\sigma \sigma$ is the alternizer in A. Now $gh = \varepsilon_g h$ for all $g \in S_n$, so $Ah = \mathbb{C}h$ is the signature representation of S_n.

Lemma 1. Let λ be a partition of n, Σ_λ a Young tableau and $\sigma \in S_n$

(1) $\mathrm{Row}(\Sigma_\lambda) \cap \mathrm{Col}(\Sigma_\lambda) = \{1\}$.

(2) The coefficient of 1 in $h(\Sigma_\lambda)$ is 1.

(3) $\mathrm{Row}(\sigma\Sigma_\lambda) = \sigma\ \mathrm{Row}(\Sigma_\lambda)\ \sigma^{-1}$ and $\mathrm{Col}(\sigma\Sigma_\lambda) = \sigma\ \mathrm{Col}(\Sigma_\lambda)\ \sigma^{-1}$.

(4) $h(\sigma\Sigma_\lambda) = \sigma\ h(\Sigma_\lambda)\ \sigma^{-1}$.

(5) The A-modules $Ah(\Sigma_\lambda)$ and Ah_λ are isomorphic.

PROOF. (5) $\Sigma_\lambda = \sigma\Sigma^0_\lambda$ for some $\sigma \in S_n$. Postmultiplication by σ defines an isomorphism $Ah(\Sigma_\lambda) \longrightarrow Ah_\lambda$.

Lemma 2. If $\lambda \geq \mu$ are partitions of n and Σ_λ and Σ'_μ are Young tableaux with frames $[\lambda]$ and $[\mu]$, then one of the following is true

(1) there are distinct integers i,j which occur in the same row of Σ_λ and the same column of Σ'_μ.

(2) $\lambda = \mu$ and $\Sigma'_\mu = rc\Sigma_\lambda$ for some $r \in \mathrm{Row}(\Sigma_\lambda)$ and $c \in \mathrm{Col}(\Sigma_\lambda)$.

PROOF. Suppose (1) fails.

If $\lambda_1 \neq \mu_1$ then $\lambda_1 > \mu_1$, so $[\mu]$ has fewer columns than $[\lambda]$, and hence two of the numbers in the first row of Σ_λ are in the same column of Σ'_μ, so (1) holds, contrary to the assumption.

Thus $\lambda_1 = \mu_1$, and since (1) fails some $c_1 \in \mathrm{Col}(\Sigma'_\mu)$ ensures that $c_1\Sigma'_\mu$ has the same elements numbers in the first row as Σ_λ.

Now ignore the first rows of Σ_λ and $c_1\Sigma'_\mu$. By the same argument we find $\lambda_2 = \mu_2$ and can find c_2 such that $c_2c_1\Sigma'_\mu$ have the same numbers in each of the first two rows.

Eventually we find $\lambda = \mu$ and some $c' \in \mathrm{Col}(\Sigma'_\mu)$ such that $\Sigma_\lambda = c'\Sigma'_\mu$ have the same numbers in each row. Then $r\Sigma_\lambda = c'\Sigma'_\mu$ for some $r \in \mathrm{Row}(\Sigma_\lambda)$. Finally $\Sigma'_\mu = rc\Sigma_\lambda$ where

$$c = r^{-1}c'^{-1}r \in r^{-1}c'\mathrm{Col}(\Sigma'_\mu)c'^{-1}r = r^{-1}\mathrm{Col}(c'\Sigma'_\mu)r = \mathrm{Col}(\Sigma_\lambda)$$

since $c' \in \mathrm{Col}(\Sigma'_\mu)$.

Lemma 3. If $\sigma \in S_n$ cannot be written as rc for any $r \in \mathrm{Row}(\Sigma_\lambda)$ and $c \in \mathrm{Col}(\Sigma_\lambda)$ then there are transpositions $u \in \mathrm{Row}(\Sigma_\lambda)$ and $v \in \mathrm{Col}(\Sigma_\lambda)$ with $u\sigma = \sigma v$.

PROOF. (2) fails for Σ_λ and $\sigma\Sigma_\lambda$, so there are $i\neq j$ in the same row in Σ_λ and in the same column in $\sigma\Sigma_\lambda$. Let $u = (i\ j)$ and $v = \sigma^{-1}u\sigma$.

Lemma 4. Let Σ_λ be a Young tableau and $a \in A$. The following are equivalent

(1) $rac = \varepsilon_c a$ for all $r \in \mathrm{Row}(\Sigma_\lambda)$ and $c \in \mathrm{Col}(\Sigma_\lambda)$.

(2) $a = \alpha\ h(\Sigma_\lambda)$ for some $\alpha \in \mathbb{C}$.

PROOF.

(2)⇒(1) $rh(\Sigma_\lambda)c = \varepsilon_\sigma h(\Sigma_\lambda)$ since as r' runs through $\mathrm{Row}(\Sigma_\lambda)$ so does rr', and as c' runs through $\mathrm{Col}(\Sigma_\lambda)$ so does $c'c$, with $\varepsilon_{c'c} = \varepsilon_{c'}\varepsilon_c$.

(1)⇒(2) Say $a = \sum_{\sigma\in Sn} a_\sigma\sigma$. If σ is not of the form rc then there are transpositions $u\in\mathrm{Row}(\Sigma_\lambda)$ and $v\in\mathrm{Col}(\Sigma_\lambda)$ with $u\sigma v = \sigma$. By assumption $uav = \varepsilon_v a$, and the coefficient of σ gives $a_{u\sigma v} = \varepsilon_v a_\sigma$, so $a_\sigma = -a_\sigma$, and hence $a_\sigma = 0$. Now the coefficient of rc in (1) gives $a_1 = \varepsilon_c a_{rc}$. Thus

$$a = \textstyle\sum_{r,c} a_{rc}rc = \sum_{r,c} \varepsilon_c a_1\ rc = a_1\ h(\Sigma_\lambda).$$

Lemma 5. If $a \in A$ then $h(\Sigma_\lambda)\ a\ h(\Sigma_\lambda) = \alpha\ h(\Sigma_\lambda)$ for some $\alpha \in \mathbb{C}$.

PROOF. Let $x = h(\Sigma_\lambda)\ a\ h(\Sigma_\lambda)$. This has property (1) above.

Definition. Let $f_\lambda = \dim_{\mathbb{C}}(Ah_\lambda)$.

Lemma 6.

(1) $h(\Sigma_\lambda)^2 = (n!/f_\lambda)\ h(\Sigma_\lambda)$

(2) f_λ divides $n!$

(3) $h(\Sigma_\lambda)\ A\ h(\Sigma_\lambda) = \mathbb{C}\ h(\Sigma_\lambda)$.

In particular $(f_\lambda/n!)\ h(\Sigma_\lambda)$ is an idempotent.

PROOF.

(1) Let $h = h(\Sigma_\lambda)$. We know that $h^2 = \alpha h$ for some $\alpha\in\mathbb{C}$. Right multiplication by h induces a linear map $\hat{h}:A\longrightarrow A$. For $a\in A$ we have $(ah)h = \alpha(ah)$, so $\hat{h}|_{Ah}$ acts as multiplication by α. Take a basis of Ah and extend it to a basis of A. With respect to this $\hat{h}$ has matrix

$$\begin{pmatrix} \alpha I_{f_\lambda} & * \\ 0 & 0 \end{pmatrix}$$

(f_λ since $\dim_{\mathbb{C}} Ah = \dim_{\mathbb{C}} Ah_\lambda$), so $\mathrm{Trace}(\hat{h}) = \alpha f_\lambda$.

With respect to the basis S_n of A, $\hat{h}$ has matrix H with

$$H_{\sigma\tau} = \text{coefficient of } \sigma \text{ in } \tau h.$$

Now $H_{\sigma\sigma} = 1$ so $\mathrm{Trace}(\hat{h}) = n!$.

(2) The coefficient of 1 in $h^2 = \alpha h$ is

$$\alpha = \sum_{r_1,r_2 \in \mathrm{Row}(\Sigma_\lambda)\ c_1,c_2 \in \mathrm{Col}(\Sigma_\lambda)\ r_1c_1r_2c_2=1} \varepsilon_{c_1} \varepsilon_{c_2} \in \mathbb{Z}.$$

(3) By Lemma 5 the only other possibility is $hAh = 0$. But $h^2 \neq 0$.

Lemma 7. $Ah(\Sigma_\lambda)$ is a simple A-module.

PROOF. Let $h = h(\Sigma_\lambda)$. Since Ah is non-zero, and A is semisimple, it suffices to prove that Ah is indecomposable. Say $Ah = U \oplus V$. Then

$$\mathbb{C}h = hAh = hU + hV,$$

so one of hU and hV is non-zero. Without loss of generality $hU \neq 0$, but then $hU = \mathbb{C}h$, so $Ah = AhU \subseteq U$, and hence $U = Ah$ and $V = 0$.

Lemma 8. If $\lambda > \mu$ are partitions and Σ_λ and Σ'_μ are Young tableau, then $h(\Sigma'_\mu)\ A\ h(\Sigma_\lambda) = 0$.

PROOF. Since $\lambda > \mu$, by Lemma 2 there are be two integers in the same row of Σ_λ and in the same column of Σ_μ. The corresponding transposition $\tau \in \mathrm{Row}(\Sigma_\lambda) \cap \mathrm{Col}(\Sigma'_\mu)$. Then

$$h(\Sigma'_\mu)\ h(\Sigma_\lambda) = h(\Sigma'_\mu)\ \tau\ \tau\ h(\Sigma_\lambda) = -\ h(\Sigma'_\mu)\ h(\Sigma_\lambda),$$

so $h(\Sigma'_\mu)\ h(\Sigma_\lambda) = 0$.

Applying this to $\sigma\Sigma_\lambda$ and Σ'_μ for $\sigma \in S_n$ gives

$$0 = h(\Sigma'_\mu)\ h(\sigma\Sigma_\lambda) = h(\Sigma'_\mu)\ \sigma\ h(\Sigma_\lambda)\ \sigma^{-1}$$

so $h(\Sigma'_\mu)\ \sigma\ h(\Sigma_\lambda) = 0$. Thus $h(\Sigma'_\mu)\ A\ h(\Sigma_\lambda) = 0$.

Lemma 9. If $\lambda \neq \mu$ are partitions and Σ_λ and Σ'_μ are Young tableau, then $Ah(\Sigma_\lambda)$ and $Ah(\Sigma'_\mu)$ are not isomorphic.

PROOF. We may assume that $\lambda > \mu$. If there is an isomorphism

$$f : Ah(\Sigma'_\mu) \longrightarrow Ah(\Sigma_\lambda)$$

of A-modules, then

$$f(\mathbb{C}h(\Sigma'_\mu)) = f(h(\Sigma'_\mu)Ah(\Sigma'_\mu)) = h(\Sigma'_\mu)f(Ah(\Sigma'_\mu)) = h(\Sigma'_\mu)Ah(\Sigma_\lambda) = 0,$$

a contradiction.

Remark. The partitions of n correspond to conjugacy classes in S_n, with say $(5,2^2,1)$ corresponding to the permutations in S_{10} of the form

(.....)(..)(..)(.).

Theorem. The left ideals Ah_λ with λ running through the partitions of n are a complete set of non-isomorphic simple A-modules.

PROOF. They are simple, non-isomorphic, and the number of them is equal to the number of conjugacy classes in S_n, which we know from character theory is the number of simple A-modules.

§3. STANDARD TABLEAUX

This section is not really necessary for the main development, but is included because of its cleverness, and because the standard tableaux give an explicit decomposition of tensor space into simple submodules.

Definition. A Young tableau Σ_λ is standard if the numbers increase from left to right in each row and from top to bottom in each column. The standard tableaux with frame $[\lambda]$ are ordered so that Σ_λ is smaller than Σ'_λ if it is smaller in the first place that they differ when you read $[\lambda]$ like a book.

Example. For $\lambda = (3,2)$ the standard tableaux are

$$\begin{matrix}123\\45\end{matrix} < \begin{matrix}124\\35\end{matrix} < \begin{matrix}125\\34\end{matrix} < \begin{matrix}134\\25\end{matrix} < \begin{matrix}135\\24\end{matrix}.$$

We denote by F_λ the number of standard tableaux with frame $[\lambda]$. We shall show that $F_\lambda = f_\lambda$ and as a first step we prove that $\sum F_\lambda^2 = n!$. In the next few lemmas we write λ/μ to mean that λ is a partition of m and μ is a partition of m-1 for some m, and that $[\mu] \subset [\lambda]$.

Lemma 1. If λ is a partition of m then $F_\lambda = \sum_{\mu \text{ st } \lambda/\mu} F_\mu$.

PROOF. If Σ_λ is a standard tableau, then $\Sigma_\lambda^{-1}(\{1,...,m-1\})$ is the frame of a partition μ of m-1, and $\Sigma_\lambda|_{[\mu]}$ is a standard tableau. And conversely.

Lemma 2. If $\lambda \neq \pi$ are partitions of m, then

$$|\{\nu|\nu/\lambda \text{ and } \nu/\pi\}| = |\{\tau|\lambda/\tau \text{ and } \pi/\tau\}| \in \{0,1\}.$$

PROOF. If ν/λ and ν/π then $[\nu] \supseteq [\lambda]\cup[\pi]$, so there must be equality here. Similarly if λ/τ and π/τ then $[\tau] = [\lambda]\cap[\pi]$.

Now $[\lambda]\cup[\pi]$ and $[\lambda]\cap[\pi]$ are always frames of partitions, so

$$\text{there is a } \nu \Leftrightarrow |[\lambda]\cup[\pi]| = m+1 \Leftrightarrow |[\lambda]\cap[\pi]| = m-1 \Leftrightarrow \text{there is a } \tau.$$

Lemma 3. If λ is a partition of m then $(m+1)F_\lambda = \sum_{\nu \text{ st } \nu/\lambda} F_\nu$.

PROOF. This is true for m = 1. We prove it by induction, so suppose it is true for all partitions of m-1. Now

$$\sum_{\nu \text{ st } \nu/\lambda} F_\nu = \sum_{\nu \text{ st } \nu/\lambda} \ \sum_{\pi \text{ st } \nu/\pi} F_\pi \qquad \text{by Lemma 1}$$

$$= |\{\nu | \nu/\lambda\}| \ F_\lambda + \sum_{\nu,\pi \text{ st } \nu/\lambda,\ \nu/\pi,\ \pi\neq\lambda} F_\pi .$$

By inspecting the Young frames one sees that

$$|\{\nu|\nu/\lambda\}| = |\{\tau|\lambda/\tau\}| + 1,$$

and using Lemma 2 we get

$$= \Big(|\{\tau|\lambda/\tau\}|+1\Big) F_\lambda + \sum_{\tau,\pi \text{ st } \lambda/\tau,\ \pi/\tau,\ \pi\neq\lambda} F_\pi$$

$$= F_\lambda + \sum_{\tau,\pi \text{ st } \lambda/\tau,\ \pi/\tau} F_\pi$$

$$= F_\lambda + \sum_{\tau \text{ st } \lambda/\tau} mF_\tau \qquad \text{by the induction}$$

$$= F_\lambda + mF_\lambda = (m+1)F_\lambda \qquad \text{by Lemma 1.}$$

<u>Lemma</u> 4. $\displaystyle\sum_{\lambda \text{ a partition of } m} F_\lambda^2 = m!$

PROOF. It is true for m = 1. We prove it by induction on m. Now

$$\sum_{\lambda \text{ a partition of } m} F_\lambda^2 = \sum_{\lambda \text{ a partition of } m \text{ and } \lambda/\tau} F_\lambda F_\tau \qquad \text{by Lemma 1.}$$

$$= \sum_{\tau \text{ a partition of } m-1} mF_\tau^2 \qquad \text{by Lemma 3.}$$

$$= m! \qquad \text{by the induction.}$$

<u>Lemma</u> 5. If $\Sigma_\lambda > \Sigma'_\lambda$ are standard tableaux then $h(\Sigma_\lambda)\ h(\Sigma'_\lambda) = 0$.

PROOF. It suffices to show that there are two numbers i≠j in the same row in Σ'_λ and in the same column in Σ_λ, for then the transposition t=(i j) is in Row(Σ'_λ) and in Col(Σ_λ). Thus

$$h(\Sigma'_\lambda) = t\ h(\Sigma'_\lambda) \text{ and } h(\Sigma_\lambda) = -\ h(\Sigma_\lambda)\ t,$$

so

$$h(\Sigma_\lambda)\ h(\Sigma'_\lambda) = h(\Sigma_\lambda)\ t\ t\ h(\Sigma'_\lambda) = -\ h(\Sigma_\lambda)\ h(\Sigma'_\lambda).$$

so this product is zero.

Consider where Σ_λ and Σ'_λ first differ. Pictorially we have

$$[\lambda] \quad = \quad \begin{array}{|c|c|c|c|c|c|c|}
\hline
= & = & = & = & = & = & = \\
\hline
= & =z & = & x & & \multicolumn{2}{c}{} \\
\cline{1-5}
 & & & & & \multicolumn{2}{c}{} \\
\cline{1-5}
 & y & & \multicolumn{4}{c}{} \\
\cline{1-3}
 & \multicolumn{6}{c}{} \\
\cline{1-1}
\end{array}$$

where an "=" means that the two tableaux are the same at that box, and x is the first place where they differ. Let $\Sigma'_\lambda(x) = i$ and $y = \Sigma_\lambda^{-1}(i)$. By the assumptions, y must be below and to the left of x; in particular x cannot be in the first column or the last row. Let z be the element of $[\lambda]$ in the same row as x and the same column as y, and let j be the common value of Σ_λ and Σ'_λ at z. Now i and j satisfy the assumptions above.

<u>Theorem</u>. $\mathbb{C}S_n = \oplus\, \mathbb{C}S_n h(\Sigma_\lambda)$ with Σ_λ running over all standard tableaux for all partitions λ of n.

Proof. We show first that the sum is direct, so suppose that there is a non-trivial relation

$$\sum a(\Sigma_\lambda)h(\Sigma_\lambda) = 0 \qquad (*)$$

with $a(\Sigma_\lambda) \in \mathbb{C}S_n$. Pick μ maximal such that some $a(\Sigma_\mu)h(\Sigma_\lambda) \neq 0$, and then pick Σ'_μ minimal with respect to $a(\Sigma'_\mu)h(\Sigma'_\mu) \neq 0$. Multiplying (*) on the right by $h(\Sigma'_\mu)$ we obtain $a(\Sigma'_\mu)\, h(\Sigma'_\mu)^2 = 0$ by Lemma 5 and §2 Lemma 8, so $a(\Sigma'_\mu)\, h(\Sigma'_\mu) = 0$. A contradiction.

Now $\mathbb{C}S_n$ contains $\oplus\, \mathbb{C}S_n h(\Sigma_\lambda)$. By the Artin-Weddurburn Theorem, $\mathbb{C}S_n$ is isomorphic as an $\mathbb{C}S_n$-module to the direct sum of f_λ copies of each $\mathbb{C}S_n h_\lambda$, while in this direct sum there are F_λ copies of $\mathbb{C}S_n h_\lambda$, so by the Jordan Hölder Theorem, $F_\lambda \le f_\lambda$. On the other hand

$$\textstyle\sum_\lambda F_\lambda^2 = n! = \sum_\lambda f_\lambda^2$$

so we must have $F_\lambda = f_\lambda$ for each λ. But this means that the direct sum is equal to $\mathbb{C}S_n$.

<u>Corollary</u>. $f_\lambda = \dim_{\mathbb{C}}(\mathbb{C}S_n h_\lambda)$ is equal to the number F_λ of standard tableaux with frame $[\lambda]$.

<u>Lemma</u> 6. If M is a finite dimensional $\mathbb{C}S_n$-module, then

$$M = \oplus\, h(\Sigma_\lambda)M,$$

where Σ_λ runs over all standard tableaux for all partitions of n.

PROOF. If $\sum m(\Sigma_\lambda) = 0$ is a non-trivial relation with $m(\Sigma_\lambda) \in h(\Sigma_\lambda)M$, choose μ minimal and then Σ'_μ maximal, such that $m(\Sigma'_\mu) \neq 0$. Now premultiply the relation by $h(\Sigma'_\mu)$ to obtain a contradiction by Lemma 5 and §2 Lemma 8. Thus the sum is direct. Now

$$\begin{aligned} \oplus\, h(\Sigma_\lambda)M &\cong \oplus\, \mathrm{Hom}_{\mathbb{C}S_n}(\mathbb{C}S_n h(\Sigma_\lambda), M) \qquad \text{by §1 Lemma 6} \\ &\cong \mathrm{Hom}_{\mathbb{C}S_n}\left(\oplus\, \mathbb{C}S_n h(\Sigma_\lambda), M\right) \cong \mathrm{Hom}_{\mathbb{C}S_n}(\mathbb{C}S_n, M) \cong M, \end{aligned}$$

and all we need is that the dimensions are equal.

<u>Exercise</u>. If $R = M_2(\mathbb{C})$ and h,g are idempotents in R with $R = Rh \oplus Rg$, the argument used in the proof above shows that R is isomorphic to the external direct sum of hR and gR. Show, however, that it is still possible that $R \neq hR+gR$.

§4. A CHARACTER FORMULA

Recall that if M is a finite dimensional $\mathbb{C}G$-module, then the corresponding character is

$$\chi_M(g) = \text{trace of the map } M \longrightarrow M,\ m \longmapsto gm.$$

It is a class function $G \longrightarrow \mathbb{C}$, so if α is a conjugacy class in G we can write $\chi_M(\alpha)$.

If λ is a partition of n, the character of the $\mathbb{C}S_n$-module $\mathbb{C}S_n h_\lambda$ is denoted by χ_λ. In this section we derive a very useful formula which enables one to compute the $\chi_\lambda(\alpha)$. In the present course we shall not use this formula to compute any characters explicitly; instead we use it later to derive Weyl's character formula for the general linear group.

If α is a conjugacy class in S_n, then α consists of all the permutations with a fixed cycle type, which we denote by

$$n^{\alpha n} \dots 2^{\alpha 2} 1^{\alpha 1}$$

meaning that the permutations involve α_n n-cycles, ..., α_2 2-cycles and α_1 1-cycles. The number of permutations in α is denoted by n_α.

<u>Lemma</u> 1. $n_\alpha \;=\; \dfrac{n!}{1^{\alpha 1} 2^{\alpha 2} \dots \alpha 1! \alpha 2! \dots}$.

PROOF. Any permutation in α is one of the n! of the form

$$\underbrace{(*)(*)..(*)}_{\alpha_1}\quad \underbrace{(**)...(**)}_{\alpha_2}\quad \underbrace{(***)}...$$

with the *'s replaced by the numbers 1,...,n. However, each such permutation can be represented in $\alpha_1!\alpha_2!\dots 1^{\alpha 1}2^{\alpha 2}\dots$ ways by permuting the α_i i-cycles in $\alpha_i!$ ways, or rotating an i-cycle in i ways.

<u>Orthogonality</u> <u>relations</u>.

(1) If λ and μ are partitions of n then

$$\sum_{\alpha \text{ conj class in } S_n} n_\alpha\, \chi_\lambda(\alpha)\, \chi_\mu(\alpha) = \begin{cases} n! & (\lambda = \mu) \\ 0 & (\text{else}) \end{cases}$$

(2) If α and β are conjugacy classes in S_n, then

$$\sum_{\lambda \text{ partition of } n} \chi_\lambda(\alpha)\, \chi_\lambda(\beta) = \begin{cases} n!/n_\alpha & (\alpha = \beta) \\ 0 & (\text{else}) \end{cases}$$

PROOF. Every element in S_n is conjugate to its inverse, so

$$\overline{\chi_\lambda(g)} = \chi_\lambda(g^{-1}) = \chi_\lambda(g)$$

for $g \in S_n$. With this observation these relations become the standard orthogonality relations for finite groups.

Notation. Given $x_1,...,x_m \in \mathbb{C}$ and $\ell_1,...,\ell_m \in \mathbb{Z}$ define

$$|x^{\ell_1},...,x^{\ell_m}| = \det(x_i^{\ell_j})_{1\le i,j\le m}.$$

Usually the $\ell_i \ge 0$, in which case it is a homogeneous polynomial of degree $\ell_1+...+\ell_m$ in the x_i.

Example. The Vandermonde $|x^{m-1},...,x,1|$.

Lemma 2. The Vandermonde = $\prod_{i<j}(x_i-x_j)$.

PROOF. Subtracting the second row from the first, the element in position (1,j) is

$$x_1^j - x_2^j = (x_1 - x_2)(x_1^{j-1} + x_1^{j-2}x_2 + ... + x_2^{j-1})$$

so the entire first row is divisible by x_1-x_2. Thus the determinant V is divisible by x_1-x_2 in $\mathbb{C}[x_1,...,x_m]$. Similarly for x_i-x_j with $i<j$. Since polynomial rings are UFDs, V is divisible by the product P. Now V is a polynomial of degree 1+2+...+(m-1), which is the degree of the product, so $V = aP$ for some $a\in\mathbb{C}$. We show by induction on m that a=1. If m=1 then this is clear. In general, if $x_m=0$ then expanding the determinant and using the induction $V = x_1x_2..x_{m-1} \prod_{i<j<m}(x_i-x_j)$, so a=1.

Remark. The same argument shows that if $\ell_i\ge 0$ then $|x^{\ell_1},...,x^{\ell_m}|$ is divisible by the Vandermonde, so that

$$\frac{|x^{\ell_1},...,x^{\ell_m}|}{|x^{m-1},..., 1 |}$$

is a polynomial in $x_1,...,x_m$.

Cauchy's Lemma. If $x_i,y_i \in \mathbb{C}$ $(1\le i\le m)$ and always $x_iy_j\ne 1$ then

$$\det\left(\frac{1}{1-x_iy_j}\right) = |x^{m-1},..,1|.|y^{m-1},..,1|.\prod_{i,j}\left(\frac{1}{1-x_iy_j}\right)$$

PROOF. By induction on m. True for m=1. Now

$$\frac{1}{1-x_iy_j} - \frac{1}{1-x_1y_j} = \frac{x_i-x_1}{1-x_1y_j} \cdot \frac{y_j}{1-x_iy_j}$$

so subtracting the first row from each other row in the determinant one can remove the factor x_i-x_1 from each row $i\neq1$ and $1/(1-x_1y_j)$ from each column, and the determinant equals

$$\prod_{i>1} (x_i-x_1) \cdot \prod_j \left(\frac{1}{1-x_1y_j}\right) \cdot \det \begin{pmatrix} 1 & 1 & \dots \\ y_1/(1-x_2y_1) & y_2/(1-x_2y_2) & \dots \\ y_1/(1-x_3y_1) & y_2/(1-x_3y_2) & \dots \\ \dots & \dots & \dots \end{pmatrix} \qquad (*)$$

Now subtract the first column from each other, and use

$$\frac{y_j}{1-x_iy_j} - \frac{y_1}{1-x_iy_1} = \frac{y_j-y_1}{1-x_iy_1} \cdot \frac{1}{1-x_iy_j}.$$

so the determinant in (*) becomes

$$\prod_{j>1} (y_j-y_1) \cdot \prod_{i>1} \left(\frac{1}{1-x_iy_1}\right) \cdot \det \begin{pmatrix} 1 & 0 & 0 & \dots \\ * & 1/(1-x_2y_2) & 1/(1-x_2y_3) & \dots \\ * & 1/(1-x_3y_2) & 1/(1-x_3y_3) & \dots \\ . & \dots & \dots & \dots \end{pmatrix}$$

and the assertion follows.

Lemma 3. If $x_1,\dots,x_m$ and $y_1,\dots,y_m$ have modulus < 1, then

$$\det\left(\frac{1}{1-x_iy_j}\right) = \sum_{\ell_1>\dots>\ell_m\geq0} |x^{\ell_1},\dots,x^{\ell_m}|.|y^{\ell_1},\dots,y^{\ell_m}|.$$

PROOF. The determinant is

$$\sum_{\pi\in S_m} \varepsilon_\pi \prod_{i=1}^{m} (1 + x_iy_{\pi(i)} + x_i^2y_{\pi(i)}^2 + \dots)$$

and the monomial $x_1^{\ell_1}\dots x_m^{\ell_m}$ (with $\ell_i\in\mathbb{N}$) occurs with coefficient

$$\sum_{\pi\in S_m} \varepsilon_\pi \prod_{i=1}^{m} y_{\pi(i)}^{\ell_i} = |y^{\ell_1},\dots,y^{\ell_m}|.$$

In particular it zero unless the ℓ_i are distinct, so the determinant is

$$\sum_{\ell_1,\dots,\ell_m \text{ distinct}} x_1^{\ell_1}\dots x_m^{\ell_m} |y^{\ell_1},\dots,y^{\ell_m}|$$

$$= \sum_{\ell_1>\dots>\ell_m,\ \pi\in S_m} x_1^{\ell_{\pi(1)}}\dots x_m^{\ell_{\pi(m)}} |y^{\ell_{\pi(1)}},\dots,y^{\ell_{\pi(m)}}|$$

$$= \sum_{\ell_1>\dots>\ell_m,\ \pi\in S_m} x_1^{\ell_{\pi(1)}}\dots x_m^{\ell_{\pi(m)}} \varepsilon_\pi |y^{\ell_1},\dots,y^{\ell_m}|$$

$$= \sum_{\ell_1>\dots>\ell_m} |x^{\ell_1},\dots,x^{\ell_m}|.|y^{\ell_1},\dots,y^{\ell_m}|.$$

Notation. If $\lambda = (\lambda_1,\dots,\lambda_m)\in\mathbb{Z}^m$ (for example if λ is a partition with $\leq m$ parts), we set

$$\ell_i = \lambda_i + m - i.$$

so $\ell_1 = \lambda_1+m-1$, ..., $\ell_m = \lambda_m$.

Remarks.

(1) $\lambda_1\geq\lambda_2\geq\ldots\geq\lambda_m \Leftrightarrow \ell_1>\ell_2>\ldots>\ell_m$.

(2) If $\sum_i \lambda_i = n$, $\lambda_i\geq 0$, then the polynomial

$$\frac{|x^{\ell_1},\ldots,x^{\ell_m}|}{|x^{m-1},\ldots, 1 |}$$

has degree n in the x_i.

Notation. If $x_1,\ldots,x_m$ and $y_1,\ldots,y_m$ are complex numbers, for $i\in\mathbb{N}$ we set

$$s_i = x_1^i + x_2^i + \ldots + x_m^i \quad \text{and} \quad t_i = y_1^i + y_2^i + \ldots + y_m^i.$$

the power sums of the x_i and the y_i.

Lemma 4.

$$\sum \frac{|x^{\ell_1},\ldots,x^{\ell_m}|}{|x^{m-1},\ldots, 1 |} \cdot \frac{|y^{\ell_1},\ldots,y^{\ell_m}|}{|y^{m-1},\ldots, 1 |} = \frac{1}{n!} \sum_\alpha n_\alpha s_1^{\alpha 1}\ldots s_n^{\alpha n} t_1^{\alpha 1}\ldots t_n^{\alpha n}$$

where the first sum is over the partitions λ of n with $\leq$m parts, and the second sum is over the conjugacy classes α in S_n.

Remark. The quotients on the left are polynomials, so this makes sense even if the x_i or y_i are not distinct.

PROOF. Since both sides are polynomials, we need only prove this when the x_i and y_i have modulus < 1. Now

$$\log \left(\prod_{i,j=1}^m \left(\frac{1}{1-x_iy_j}\right)\right) = \sum_{i,j} \left(\frac{x_iy_j}{1} + \frac{x_i^2y_j^2}{2} + \frac{x_i^3y_j^3}{3} + \ldots\right)$$

$$= \frac{s_1t_1}{1} + \frac{s_2t_2}{2} + \frac{s_3t_3}{3} + \ldots$$

so

$$\prod_{i,j=1}^m \left(\frac{1}{1-x_iy_j}\right) = \exp\left(\frac{s_1t_1}{1} + \frac{s_2t_2}{2} + \frac{s_3t_3}{3} + \ldots\right)$$

$$= \sum \frac{1}{n!}\left(\frac{s_1t_1}{1} + \frac{s_2t_2}{2} + \frac{s_3t_3}{3} + \ldots\right)^n$$

By the multinomial theorem this is

$$= \sum \frac{1}{n!} \frac{n!}{\alpha_1!\alpha_2!\cdots} \left(\frac{s_1t_1}{1}\right)^{\alpha_1} \left(\frac{s_2t_2}{2}\right)^{\alpha_2}\cdots$$

where the sum extends over all sequences $(\alpha_1,\alpha_2,\dots)$ of non-negative integers with only finitely many non-zero terms.

$$= \sum \frac{s_1^{\alpha 1}s_2^{\alpha 2}\cdots\ t_1^{\alpha 1}t_2^{\alpha 2}\cdots}{1^{\alpha 1}2^{\alpha 2}\cdots\ \alpha 1!\alpha 2!\cdots}$$

By Lemma 3 and Cauchy's Lemma,

$$\sum \frac{|x^{\ell_1},\dots,x^{\ell_m}|}{|x^{m-1},\dots,1|}\cdot\frac{|y^{\ell_1},\dots,y^{\ell_m}|}{|y^{m-1},\dots,1|} = \prod_{i,j=1}^{m} \left(\frac{1}{1-x_iy_j}\right)$$

where the sum is over all $\ell_1>\dots>\ell_m\geq 0$. So

$$\sum \frac{|x^{\ell_1},\dots,x^{\ell_m}|}{|x^{m-1},\dots,1|}\cdot\frac{|y^{\ell_1},\dots,y^{\ell_m}|}{|y^{m-1},\dots,1|} = \sum \frac{s_1^{\alpha 1}s_2^{\alpha 2}\cdots\ t_1^{\alpha 1}t_2^{\alpha 2}\cdots}{1^{\alpha 1}2^{\alpha 2}\cdots\ \alpha 1!\alpha 2!\cdots}$$

We can now equate the terms in this which are of degree n in the x_i, getting the required equality.

Definition. For $\lambda = (\lambda_1,\dots,\lambda_m) \in \mathbb{Z}^m$ and α a conjugacy class in S_n, let $\psi_\lambda(\alpha)$ be the coefficient of the monomial $x_1^{\lambda 1}\dots x_m^{\lambda m}$ in $s_1^{\alpha 1}\dots s_n^{\alpha n}$. Thus

$$s_1^{\alpha 1}\dots s_n^{\alpha n} = \sum_{\lambda 1,\dots,\lambda m} \psi_\lambda(\alpha)\ x_1^{\lambda 1}\dots x_m^{\lambda m}.$$

Equivalently we can think of the ψ_λ as class functions $\psi_\lambda : S_n \longrightarrow \mathbb{N}$.

Remarks.

(1) $\psi_\lambda(\alpha) = 0$ if any $\lambda_i<0$ or if $\lambda_1+\dots+\lambda_m\neq n$.

(2) $\psi_\lambda(\alpha)$ is a symmetric function of the λ_i.

Definition. Set $\omega_\lambda(\alpha) = \sum_{\pi\in S_m} \varepsilon_\pi\ \psi_{(\ell_{\pi(1)}+1-m,\dots,\ell_{\pi(m)})}(\alpha)$.

We are eventually going to show that $\omega_\lambda = \chi_\lambda$, but first we need to verify the orthogonality relations for the ω_λ. To do this we need the following lemma, which will eventually be our character formula.

Lemma 5.

$$s_1^{\alpha 1}\dots s_n^{\alpha n}\ |x^{m-1},\dots,1| = \sum_\lambda \omega_\lambda(\alpha)\ |x^{\ell 1},\dots,x^{\ell m}|$$

with summation over the partitions λ of n with $\le m$ parts.

PROOF. $s_1^{\alpha_1}\dots s_n^{\alpha_n}\,|x^{m-1},\dots,1| =$

$$= \sum_{\lambda\in\mathbb{Z}^m,\tau\in S_m} \varepsilon_\tau\, \psi_\lambda(\alpha)\, x_1^{\lambda_1}\dots x_m^{\lambda_m}\, x_{\tau(1)}^{m-1} \cdots x_{\tau(m)}^{0}$$

$$= \sum_{\lambda\in\mathbb{Z}^m,\tau\in S_m} \varepsilon_\tau\, \psi_\lambda(\alpha)\, x_{\tau(1)}^{\lambda_{\tau(1)}+m-1} \cdots x_{\tau(m)}^{\lambda_{\tau(m)}}$$

Let $\ell_i = \lambda_{\tau(i)}+m-i$ instead of the usual convention. Since $\psi_\lambda(\alpha)$ is symmetric in the λ_i, we get

$$= \sum_{\ell_1,..,\ell_m\in\mathbb{Z},\tau\in S_m} \varepsilon_\tau\, \psi_{(\ell_1+1-m,\dots,\ell_m)}(\alpha)\, x_{\tau(1)}^{\ell_1} \cdots x_{\tau(m)}^{\ell_m}$$

$$= \sum_{\ell_1,..,\ell_m\in\mathbb{Z}} \psi_{(\ell_1+1-m,\dots,\ell_m)}(\alpha)\, |x^{\ell_1},\dots, x^{\ell_m}|$$

Since the terms with the ℓ_i not distinct are zero, this becomes

$$= \sum_{\ell_1>\dots>\ell_m,\ \pi\in S_m} \psi_{(\ell_{\pi(1)}+1-m,\dots,\ell_{\pi(m)})}(\alpha)\, \varepsilon_\pi\, |x^{\ell_1},\dots, x^{\ell_m}|.$$

Now setting $\lambda_i = \ell_i+i-m$ as usual, it becomes

$$= \sum_{\lambda,\ \pi\in S_m} \psi_{(\ell_{\pi(1)}+1-m,\dots,\ell_{\pi(m)})}(\alpha)\, \varepsilon_\pi\, |x^{\ell_1},\dots, x^{\ell_m}|$$

where the sum is over all $(\lambda_1,\dots,\lambda_m) \in \mathbb{Z}^m$ with $\lambda_1\ge\dots\ge\lambda_m$. Now the terms for which this is not a partition of n are zero by the remarks above, for if $\lambda_m < 0$ then certainly $\ell_m+\pi^{-1}(m)-m < 0$.

Lemma 6. If λ and λ' are partitions of n with $\le m$ parts then

$$\sum_{\alpha \text{ conj class}} n_\alpha\, \omega_\lambda(\alpha)\, \omega_{\lambda'}(\alpha) = \begin{cases} n! & (\lambda=\lambda') \\ 0 & (\text{else}) \end{cases}$$

PROOF. By Lemma 4, the sum

$$\sum_\lambda |x^{\ell_1},\dots,x^{\ell_m}|\,|y^{\ell_1},\dots,y^{\ell_m}|$$

over the partitions λ of n with $\le m$ parts is equal to

$$\frac{1}{n!} \sum_\alpha n_\alpha\, s_1^{\alpha_1}\dots s_n^{\alpha_n}\, t_1^{\alpha_1}\dots t_n^{\alpha_n}\, |x^{m-1},\dots,1|\, |y^{m-1},\dots,1|.$$

By Lemma 5, this is

$$\frac{1}{n!} \sum_{\alpha,\lambda,\lambda'} n_\alpha\, \omega_\lambda(\alpha)\omega_{\lambda'}(\alpha)\, |x^{\ell_1},\dots,x^{\ell_m}|\, |y^{\ell_1'},\dots,y^{\ell_m'}|$$

with summation over the partitions λ,λ' of n with $\le m$ parts and conjugacy classes α. The assertion follows since as λ and λ' vary, the polynomials

$$|x^{\ell_1},\dots,x^{\ell_m}|\, |y^{\ell_1'},\dots,y^{\ell_m'}|$$

are linearly independent in $\mathbb{C}[x_1,\ldots,x_m,y_1,\ldots,y_m]$.

We now start to relate these ideas with the symmetric group (which has, so far, played no role). The key result is:

Lemma 7. If λ is a partition of n with $\le m$ parts then ψ_λ is the character of the $\mathbb{C}S_n$-module $\mathbb{C}S_n r_\lambda$ where $r_\lambda = \sum_{\sigma\in \mathrm{Row}(\Sigma^0_\lambda)} \sigma$.

Proof. Let θ be the character of $\mathbb{C}S_n r_\lambda$, let $\sigma \in S_n$ be in the conjugacy class α, let $R = \mathrm{Row}(\Sigma^0_\lambda)$, and let

$$S_n = \bigcup_{i=1}^{N} g_i R$$

be a coset decomposition. Then $\mathbb{C}S_n r_\lambda$ has basis $(g_i r_\lambda)_{1\le i\le N}$. We use this basis to compute traces. Now

$$\sigma g_i r_\lambda = g_j r_\lambda \quad \text{if} \quad \sigma g_i \in g_j R.$$

Thus

$$\theta(\alpha) = |\{\, 1\le i\le N \mid g_i^{-1}\sigma g_i \in R \,\}|.$$

Now $g\sigma g^{-1} \in R$ if and only g is in a coset $g_i R$ with $g_i \sigma g_i^{-1} \in R$, and $|R| = \lambda_1!\lambda_2!\ldots$, so

$$\theta(\alpha) = 1/\lambda_1!\lambda_2!\ldots\ |\{\, g\in S_n \mid g\sigma g^{-1} \in R \,\}|.$$

Since $g\sigma g^{-1} = g'\sigma g'^{-1} \Leftrightarrow g'^{-1}g \in c_{S_n}(\sigma)$, each value taken by $g\sigma g^{-1}$ is taken by $|c_{S_n}(\sigma)|$ elements $g\in S_n$. Now

$$|c_{S_n}(\sigma)| = \frac{n!}{n_\alpha} = 1^{\alpha_1}2^{\alpha_2}\ldots\alpha_1!\alpha_2!\ldots$$

so

$$\theta(\alpha) = 1^{\alpha_1}2^{\alpha_2}\ldots\alpha_1!\alpha_2!\ldots/\lambda_1!\lambda_2!\ldots\ |\alpha \cap R|.$$

Now a permutation $\tau \in \alpha \cap R$ restricts to a permutation of the numbers in the i-th row of Σ^0_λ. If this restriction involves say α_{ij} j-cycles, then the α_{ij} satisfy (*):

$$\alpha_{i1} + 2\alpha_{i2} + 3\alpha_{i3} + \ldots = \lambda_i \quad (1\le i\le m)$$
$$\alpha_{1j} + \alpha_{2j} + \alpha_{3j} + \ldots = \alpha_j \quad (1\le j\le n)$$

The number of permutations in R of this type is

$$(\lambda_1!/1^{\alpha_{11}}2^{\alpha_{12}}\ldots\ \alpha_{11}!\alpha_{12}!\ldots)(\lambda_2!/1^{\alpha_{21}}2^{\alpha_{22}}\ldots\ \alpha_{21}!\alpha_{22}!\ldots)\ldots$$

so

$$\theta(\alpha) = \sum \frac{\alpha_1!}{\alpha_{11}!\alpha_{21}!\ldots} \frac{\alpha_2!}{\alpha_{12}!\alpha_{22}!\ldots} \ldots$$

where the summation is over all α_{ij} satisfying (*).

By the multinomial theorem $s_1^{\alpha 1}\ldots s_n^{\alpha n}$ is equal to

$$\sum \left(\frac{\alpha_1!}{\alpha_{11}!\alpha_{21}!\ldots} x_1^{\alpha_{11}} x_2^{\alpha_{21}} \ldots x_m^{\alpha_{m1}}\right)\left(\frac{\alpha_2!}{\alpha_{12}!\alpha_{22}!\ldots} x_1^{2\alpha_{12}} x_2^{2\alpha_{22}} \ldots x_m^{2\alpha_{m2}}\right)\ldots$$

where the sum is over all $\alpha_{ij} \in \mathbb{N}$ ($1 \le i \le m$, $1 \le j \le n$) satisfying

$$\alpha_{1j} + \alpha_{2j} + \ldots + \alpha_{mj} = \alpha_j \quad (1 \le j \le n).$$

Thus $\psi_\lambda(\alpha)$ (the coefficient of the monomial $x_1^{\lambda 1}\ldots x_m^{\lambda m}$) is equal to

$$\sum \frac{\alpha_1!}{\alpha_{11}!\alpha_{21}!\ldots} \frac{\alpha_2!}{\alpha_{12}!\alpha_{22}!\ldots} \ldots$$

where the summation is over all α_{ij} satisfying (*), and hence is equal to $\theta(\alpha)$.

Lemma 8. Let $\mu \le \lambda$ be partitions of n. The simple module $\mathbb{C}S_n h_\mu$ is a submodule of $\mathbb{C}S_n r_\lambda$ if and only if $\mu = \lambda$.

PROOF. If $\mu < \lambda$ and $\sigma \in S_n$, then by §2 Lemma 2 there are two integers in the same row of Σ_λ^0 and in the same column of $\sigma^{-1}\Sigma_\mu^0$, so if τ is their transposition then $\sigma\tau\sigma^{-1} \in \mathrm{Col}(\Sigma_\mu^0)$ and

$$h_\mu \ \sigma \ r_\lambda = h_\mu \ \sigma\tau\sigma^{-1} \ \sigma\tau \ r_\lambda = -h_\mu \ \sigma \ r_\lambda = 0.$$

Thus $0 = h_\mu \mathbb{C}S_n r_\lambda \cong \mathrm{Hom}_{\mathbb{C}Sn}(\mathbb{C}S_n h_\mu, \mathbb{C}S_n r_\lambda)$. Conversely

$$h_\lambda r_\lambda \left(\textstyle\sum_{\sigma \in \mathrm{Col}(\Sigma_\lambda^0)} \varepsilon_\sigma \sigma\right) = h_\lambda^2 \ne 0,$$

so $0 \ne h_\lambda \mathbb{C}S_n r_\lambda \cong \mathrm{Hom}_{\mathbb{C}Sn}(\mathbb{C}S_n h_\lambda, \mathbb{C}S_n r_\lambda)$.

Lemma 9. If λ is a partition of n with $\le$m parts, then $\omega_\lambda = \chi_\lambda$.

PROOF.

(i) If $\pi \in S_n$, let μ_π be the partition with parts

$$\ell_{\pi(1)}+1-m, \ \ldots, \ \ell_{\pi(m)}$$

in the appropriate order, so with parts

$$\lambda_i+\pi^{-1}(i)-i.$$

Since $\psi_\lambda(\alpha)$ is symmetric in the λ_i, we can write

$$\omega_\lambda = \sum_{\pi\in S_m} \varepsilon_\pi \, \psi_{\mu_\pi}$$

If $\pi \neq 1$ then $\mu_\pi > \lambda$, for $\lambda_1+\pi^{-1}(1)-1 \geq \lambda_1$, with equality only if $\pi^{-1}(1)=1$. Then $\lambda_2+\pi^{-1}(2)-2 \geq \lambda_2$, with equality only if $\pi^{-1}(2)=2$, etc.. If $\pi = 1$ then $\mu_\pi = \lambda$. Thus ω_λ is a $\mathbb{Z}$-linear combination of ψ_ν with $\nu\geq\lambda$ and with coefficient of ψ_λ equal to 1.

(ii) By Lemmas 7 and 8, ψ_λ is an $\mathbb{N}$-linear combination of χ_μ's with $\mu\geq\lambda$, and with non-zero coefficient of χ_λ. Thus ω_λ is a $\mathbb{Z}$-linear combination of χ_ν's with $\nu\geq\lambda$ and with positive coefficient of χ_λ. Say

$$\omega_\lambda = \sum_{\nu \text{ partition of } n} k_{\lambda\nu} \, \chi_\nu$$

with the $k_{\lambda\nu} \in \mathbb{Z}$, $k_{\lambda\lambda} > 0$ and $k_{\lambda\nu} = 0$ if $\nu<\lambda$.

(iii) We know that

$$\sum_{\alpha \text{ conj class}} n_\alpha \, \omega_\lambda(\alpha) \, \omega_\mu(\alpha) = \begin{cases} n! & (\lambda=\mu) \\ 0 & (\text{else}) \end{cases}$$

In the case $\lambda = \mu$ the orthogonality of the χ_λ gives $\sum_\nu k_{\lambda\nu}^2 = 1$, so $k_{\lambda\nu} = 0$ if $\lambda\neq\nu$ and $k_{\lambda\lambda} = 1$, as required.

At last our character formula! Recall that $m\in\mathbb{N}$ and $x_1,\ldots,x_m\in\mathbb{C}$ are arbitrary, $\ell_i=\lambda_i+m-i$ and $s_i=x_1^i+\ldots+x_m^i$.

Theorem.

$$s_1^{\alpha 1}\ldots s_n^{\alpha n} \, |x^{m-1},\ldots,1| = \sum_\lambda \chi_\lambda(\alpha) \, |x^{\ell 1},\ldots,x^{\ell m}|$$

with summation over partitions λ of n with $\leq$m parts.

PROOF. Follows from Lemmas 5 and 9.

Remark. In particular, taking $m \geq n$, we can ensure that the right hand side involves all partitions of n.

Remark. If λ is a partition of n with $\leq$m parts, then $\chi_\lambda(\alpha)$ is the coefficient of the monomial $x_1^{\ell 1}\ldots x_m^{\ell m}$ in the expansion of

$$s_1^{\alpha 1}\ldots s_n^{\alpha n} \, |x^{m-1},\ldots,1|.$$

§5. THE HOOK LENGTH FORMULA

We already have one formula for the dimension of the simple $\mathbb{C}S_n$-modules, the number of standard tableaux. In this section we derive two more formulae, one of which is easy to use.

Theorem. If λ is a partition of n with exactly m parts then the degree f_λ of χ_λ is equal to

$$n!\ \prod_{1\le i<j\le m} (\ell_i-\ell_j)\ /\ \ell_1!\dots\ell_m!$$

PROOF. $f_\lambda = \chi_\lambda(1)$, which is the coefficient of $x_1^{\ell_1}\dots x_m^{\ell_m}$ in the expansion of

$$(x_1+\dots+x_m)^n\ |x^{m-1},\dots,1| = \sum_{\tau\in Sm} (x_1+\dots+x_m)^n\ \varepsilon_\tau\ x_1^{\tau(1)-1}\dots x_m^{\tau(m)-1}.$$

By the multinomial theorem this coefficient is equal to

$$\sum_{\tau\in Sm} \varepsilon_\tau\ \frac{n!}{(\ell_1+1-\tau(1))!\dots(\ell_m+1-\tau(m))!}$$

where, by convention $1/x! = 0$ if $x < 0$. Now this is equal to

$$= n!\ |1/(\ell-m+1)!,\ \dots\ 1/(\ell-1)!,1/\ell!|$$

$$= n!/\ell_1!\dots\ell_m!\ |\dots,\ell(\ell-1),\ell,1|$$

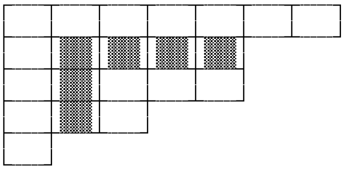
$$= n!/\ell_1!\dots\ell_m!\ |\ell^{m-1},\dots,\ell^2,\ell,1|$$

by adding appropriate columns, and this is what we want since the last determinant is the Vandermonde.

Definition. If λ is a partition of n then the *hook* at $(i,j)\in[\lambda]$ is the set of $(a,b)\in[\lambda]$ with ($a\ge i$ and $b=j$) or ($a=i$ and $b\ge j$). The *hook length* h_{ij} is the number of elements of the hook at (i,j), so that if $[\lambda]$ has column lengths $\mu_1,\mu_2,\dots$ then $h_{ij} = \lambda_i+\mu_j-i-j+1$.

Example. If $\lambda = (7,5^2,3,1)$ then the hook at (2,2) is the shaded part of

$[(7,5^2,3,1)] =$



so $h_{22} = 6$.

Theorem (Hook length formula). $f_\lambda = \dfrac{n!}{\prod_{(i,j)\in[\lambda]} h_{ij}}$.

PROOF. Let λ have m parts. By the previous theorem it suffices to show that

$$\prod_{k=i+1}^{m} (\ell_i-\ell_k) \cdot \prod_{j=1}^{\lambda i} h_{ij} = \ell_i!$$

for each i. Now the product on the left is a product of $\lambda_i+m-i = \ell_i$ terms, so it suffices to show that the terms are precisely $1,2,\dots,\ell_i$ in some order. Now

$$\ell_i-\ell_m > \ell_i-\ell_{m-1} > \ell_i-\ell_{m-2} > \dots$$
$$h_{i1} > h_{i2} > h_{i3} > \dots$$

Since λ has exactly m parts, $\mu_1 = m$ and $h_{i1} = \ell_i$, so each term is $\le\ell_i$. Thus it suffices to show that no h_{ij} is equal to any $\ell_i-\ell_k$. However, if $r = \mu_j$ then $\lambda_r \ge j$ and $\lambda_{r+1} < j$ so

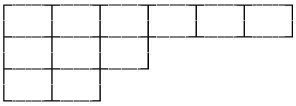
$$h_{ij}-\ell_i+\ell_r = \lambda_i+r-i-j+1-\lambda_i-m+i+\lambda_r+m-r = \lambda_r+1-j > 0,$$

$$h_{ij}-\ell_i+\ell_{r+1} = \lambda_i+r-i-j+1-\lambda_i-m+i+\lambda_{r+1}+m-r-1 = \lambda_{r+1}-j < 0,$$

and hence $\ell_i-\ell_r < h_{ij} < \ell_i-\ell_{r+1}$.

Example. If n = 11 and $\lambda = (6,3,2)$ then

$[\lambda]$ =



so that

$$f_\lambda = \frac{11.10.9.8.7.6.5.4.3.2.1}{8.7.5.3.2.1.4.3.1.2.1} = 990.$$

§6. MULTILINEAR AND POLYNOMIAL ALGEBRA

In this section we recall some rather standard multilinear algebra for finite dimensional $\mathbb{C}G$-modules where G is a group, which may be infinite, or it may be 1, so that we just deal with vector spaces.

Let V,W be finite dimensional $\mathbb{C}G$-modules.

Tensor products. The tensor product $V\otimes W$ (over $\mathbb{C}$) is a $\mathbb{C}G$-module via $g(v\otimes w) = (gv)\otimes(gw)$.

Properties. $V\otimes\mathbb{C} \cong \mathbb{C}$, $V\otimes W \cong W\otimes V$, $(V\otimes W)\otimes Z \cong V\otimes(W\otimes Z)$. If $\theta:V\longrightarrow V'$ and $\phi:W\longrightarrow W'$ are $\mathbb{C}G$-module maps, then so is $\theta\otimes\phi : V\otimes W\longrightarrow V'\otimes W'$.

Hom Spaces. $\mathrm{Hom}_{\mathbb{C}}(V,W)$ is a $\mathbb{C}G$-module via $(gf)(v) = gf(g^{-1}v)$. In particular the *dual* of V is $V^* = \mathrm{Hom}_{\mathbb{C}}(V,\mathbb{C})$, so $(gf)(v)=f(g^{-1}v)$.

Properties. $(V\otimes W)^* \cong V^*\otimes W^*$. If V is one-dimensional then $V^*\otimes V\cong\mathbb{C}$. If $\theta:V\longrightarrow W$ is a $\mathbb{C}G$-module map, then so is $\theta^*:W^*\longrightarrow V^*$. The map $V^*\otimes W\longrightarrow\mathrm{Hom}_{\mathbb{C}}(V,W)$ taking $f\otimes w$ to the map $v\longmapsto f(v)w$ is a $\mathbb{C}G$-module isomorphism.

Tensor powers. The n-th *tensor power* of V is

$$T^nV = V\otimes\ldots\otimes V \text{ (n copies, if } n>0\text{)},\ T^0V = \mathbb{C}.$$

Properties.

(1) If V has basis $e_1,\ldots,e_m$ then T^nV has basis $e_{i1}\otimes\ldots\otimes e_{in}$ so it has dimension m^n.

(2) T^nV is a $\mathbb{C}S_n$-module via

$$\sigma(v_1\otimes\ldots\otimes v_f) = v_{\sigma^{-1}(1)}\otimes\ldots\otimes v_{\sigma^{-1}(f)}\quad(\sigma\in S_f).$$

and the actions of $\mathbb{C}S_n$ and $\mathbb{C}G$ commute: $g\sigma x = \sigma gx$ for $g\in G$, $\sigma\in S_n$, $x\in T^nV$.

Definition. The n-th *exterior power* of V is $\Lambda^nV = T^nV / X$ where

$$X = \mathbb{C}\langle x - \varepsilon_\sigma\sigma x \mid x\in T^nV,\ \sigma\in S_n\rangle.$$

The image of $v_1\otimes\ldots\otimes v_n$ in Λ^nV is denoted by $v_1\wedge\ldots\wedge v_n$. We also define

$$T^nV_{\mathrm{anti}} = \{\ x \in T^nV \mid \sigma x = \varepsilon_\sigma x\ \forall\ \sigma \in S_n\ \},$$

the set of *antisymmetric tensors*.

Properties.

(1) $v_1\wedge\ldots\wedge v_n = \varepsilon_\sigma\, v_{\sigma^{-1}(1)} \wedge \ldots \wedge v_{\sigma^{-1}(n)}$ for $\sigma \in S_n$. In particular, considering a transposition, $v_1\wedge\ldots\wedge v_n = 0$ whenever two of the v_i are equal.

(2) T^nV_{anti} and Λ^nV are $\mathbb{C}G$-modules.

(3) Λ^nV has basis $e_{i1}\wedge\ldots\wedge e_{in}$ with $i_1<\ldots<i_n$, so it has dimension $\binom{m}{n}$. In particular Λ^mV is one-dimensional and $\Lambda^{m+1}V = \Lambda^{m+2}V = \ldots = 0$.

Remark. For a vector space V over an arbitrary field k one should define the exterior powers by $\Lambda^nV = T^nV/X$ where X is spanned by the tensors of form $v_1\otimes v_2\otimes\ldots\otimes v_n$ with two of the v_i equal. If k has characteristic $\neq 2$ this reduces to the given definition.

Lemma 1. $T^nV_{anti} = aT^nV$ where $a = \sum_{\sigma\in Sn} \varepsilon_\sigma\sigma$ is the alternizer. The natural map $T^nV_{anti} \longrightarrow \Lambda^nV$ is an isomorphism of $\mathbb{C}G$-modules.

PROOF. If x is antisymmetric, then $ax = (n!)x$, so $x\in aT^nV$. Conversely since $\sigma a=\varepsilon_\sigma a$, any element of aT^nV is antisymmetric. The map

$$aT^nV \longrightarrow T^nV \longrightarrow \Lambda^nV$$

is a $\mathbb{C}G$-module map with kernel $X \cap aT^nV \leq aX$ since $x = 1/n!\ ax$ for x antisymmetric. However $aX = 0$ since for $y\in T^nV$ and $\sigma\in S_n$ we have $a(y-\varepsilon_\sigma\sigma y)=0$. The map is surjective since if $x\in\Lambda^nV$ is the image of $y\in T^nV$, then x is also the image of $1/n!\ ay$.

Lemma 2. $\Lambda^n(V^*) \cong (\Lambda^nV)^*$.

PROOF. The natural map $T^nV \twoheadrightarrow \Lambda^nV$ gives an inclusion

$$(\Lambda^nV)^* \hookrightarrow (T^nV)^* \cong T^n(V^*).$$

By the universal property of Λ^nV - that any alternating multilinear map $V\times\ldots\times V \longrightarrow \mathbb{C}$ factors through Λ^nV - the image of this map is $T^n(V^*)_{anti}$, which is isomorphic to $\Lambda^n(V^*)$.

Definition. The n-th symmetric power of V is $S^nV = T^nV / Y$ where

$$Y = \mathbb{C}\langle x - \sigma x \mid x\in T^nV,\ \sigma\in S_n\rangle.$$

The image of $v_1 \otimes \ldots \otimes v_n$ in $S^n V$ is denoted by $v_1 \vee \ldots \vee v_n$. We also define

$$T^n V_{symm} = \{ x \in T^n V \mid \sigma x = x \ \forall \ \sigma \in S_n \},$$

the set of symmetric tensors.

Properties.

(1) For $\sigma \in S_n$ one has $v_1 \vee \ldots \vee v_n = v_{\sigma^{-1}(1)} \vee \ldots \vee v_{\sigma^{-1}(n)}$.

(2) $T^n V_{symm}$ and $S^n V$ are a $\mathbb{C}G$-modules.

(3) $S^n V$ has basis $e_{i1} \vee \ldots \vee e_{in}$ with $i_1 \leq \ldots \leq i_n$, so has dimension $\binom{m+n-1}{n}$. To see this, note that

$$(1-X_1)^{-1} \ldots (1-X_m)^{-1} = \sum_{i1,\ldots,im} X_1^{i1} \ldots X_m^{im}$$

so the number of terms with total degree n is the coefficient of X^n in $(1-X)^{-m}$, which is $(-1)^n \binom{-m}{n} = \binom{m+n-1}{n}$.

As in the case of exterior powers one has

Lemma 3. $T^n V_{symm} = sT^n V$ where $s = \sum_{\sigma \in Sn} \sigma$ is the symmetrizer. The natural map $T^n V_{symm} \longrightarrow S^n V$ is an isomorphism of $\mathbb{C}G$-modules.

Lemma 4. $S^n(V^*) \cong (S^n V)^*$.

Next we consider polynomial maps between vector spaces. These generalize the usual notion of linear maps.

Definition. Let V and W be finite dimensional $\mathbb{C}$-vector spaces. A function $\phi : V \longrightarrow W$ is a polynomial (resp. homogeneous n-ic) map provided that V and W have bases $e_1, \ldots, e_m$ and $f_1, \ldots, f_h$ such that for all $X_1, \ldots, X_m \in \mathbb{C}$ we have

$$\phi(X_1 e_1 + \ldots + X_m e_m) = \phi_1(X_1, \ldots, X_m) f_1 + \ldots + \phi_h(X_1, \ldots, X_m) f_h$$

where the $\phi_i(X_1, \ldots, X_m)$ are polynomials (resp. homogeneous polynomials of degree n).

Lemma 5. If there are such functions ϕ_i with respect to some bases, then there are such functions with respect to any bases.

Proof. Suppose that $e'_i = \sum_j p_{ji} e_j$ and $f_i = \sum_j q_{ji} f'_j$. Then

$$\begin{aligned}\phi(\textstyle\sum_i x_i e'_i) &= \phi(\textstyle\sum_{i,j} x_i p_{ji} e_j)\\ &= \textstyle\sum_r \phi_r(\sum_{i_1} x_{i_1} p_{1i_1},\ldots, \sum_{i_m} x_{i_m} p_{mi_m})\ f_r\\ &= \textstyle\sum_{r,s} \phi_r(\sum_{i_1} x_{i_1} p_{1i_1},\ldots,\sum_{i_m} x_{i_m} p_{mi_m})\ q_{sr}\ f'_s\end{aligned}$$

and the functions

$$\textstyle\sum_r \phi_r(\sum_{i_1} x_{i_1} p_{1i_1},\ldots,\sum_{i_m} x_{i_m} p_{mi_m})\ q_{sr}$$

are polynomials or homogeneous polynomials of degree n like the ϕ_i.

Notation. We denote by $\mathrm{Poly}_{\mathbb{C}}(V,W)$ and $\mathrm{Hom}_{\mathbb{C},n}(V,W)$ the spaces of such maps. Clearly these are vector spaces.

Lemma 6. A composition of polynomial maps $X\longrightarrow W$ and $W\longrightarrow Z$ is a polynomial map. The composition of a homogeneous n-ic and a homogeneous n′-ic map is a homogeneous nn′-ic map.

PROOF. $(x^n)^{n'} = x^{nn'}$.

Examples.

(1) $\mathrm{Hom}_{\mathbb{C},0}(V,W) = W$ and $\mathrm{Hom}_{\mathbb{C},1}(V,W) = \mathrm{Hom}_{\mathbb{C}}(V,W)$,

(2) The map $\Delta : v \longmapsto v\vee\ldots\vee v$ lies in $\mathrm{Hom}_{\mathbb{C},n}(V,S^nV)$, since

$$\begin{aligned}\Delta(\textstyle\sum x_i e_i) &= \textstyle\sum_{i_1,\ldots,i_n} x_{i_1}\ldots x_{i_n}\ e_{i_1}\vee\ldots\vee e_{i_n}\\ &= \textstyle\sum_{i_1\le\ldots\le i_n} c_{i_1,\ldots,i_n}\ x_{i_1}\ldots x_{i_n}\ e_{i_1}\vee\ldots\vee e_{i_n}\end{aligned}$$

for suitable constants $c_{i_1,\ldots,i_n}$.

Theorem. If V and W are vector spaces, then $\psi\longmapsto\psi\circ\Delta$ induces an isomorphism $\mathrm{Hom}_{\mathbb{C}}(S^nV,W) \longrightarrow \mathrm{Hom}_{\mathbb{C},n}(V,W)$.

PROOF. To show that the map is injective, suppose that $\psi\circ\Delta = 0$. We show by descending induction on i that $\psi(v_1\vee\ldots\vee v_n) = 0$ whenever i of the terms are equal. The case i=n is true by assumption, and the case i=1 is what we want. Suppose true for i+1, then for $\alpha\in\mathbb{C}$,

$$\begin{aligned}0 &= \psi((x+\alpha y)\vee\ldots\vee(x+\alpha y)\vee v_{i+2}\vee\ldots\vee v_n)\\ &= \textstyle\sum_{j=0}^{i+1} \alpha^j \binom{i+1}{j}\ \psi(\underbrace{x\vee\ldots\vee x}_{i+1-j} \vee \underbrace{y\vee\ldots\vee y}_{j} \vee v_{i+2}\vee\ldots\vee v_n)\end{aligned}$$

Since this is zero for each $\alpha \in \mathbb{C}$, each term is zero. In particular

$$\binom{i+1}{1}\ \psi(\underbrace{x\vee\dots\vee x}_{i}\vee y\vee v_{i+2}\vee\dots\vee v_n) = 0,$$

as required.

Now if $\dim_{\mathbb{C}}V = m$, $\dim_{\mathbb{C}}W = h$ then

$$\dim_{\mathbb{C}}\mathrm{Hom}_{\mathbb{C}}(S^nV,W) = h\ \binom{m+n-1}{n} = \dim_{\mathbb{C}}\mathrm{Hom}_{\mathbb{C},n}(V,W)$$

so the map is an isomorphism.

Lemma 7. The elements of the form $v\vee\dots\vee v$ with $v\in V$ span S^nV.

PROOF. Take $W = \mathbb{C}$. If these elements do not span S^nV then there is a non-zero linear map $S^nV\longrightarrow\mathbb{C}$ whose composition with Δ is zero.

Remark. We can construct an inverse explicitly. Let $\phi \in \mathrm{Hom}_{\mathbb{C},n}(V,W)$, so

$$\phi(X_1e_1+\dots+X_me_m) = \phi_1(X_1,\dots,X_m)f_1 + \dots + \phi_h(X_1,\dots,X_m)f_h$$

with ϕ_i a homogeneous polynomial of degree n. We define the total polarization $P\phi \in \mathrm{Hom}_{\mathbb{C}}(S^nV,W)$ of ϕ by

$$(P\phi)(e_{i_1}\vee\dots\vee e_{i_n}) = \textstyle\sum_{j=1}^{h} \dfrac{\partial^n\phi_j}{\partial X_{i1}\dots\partial X_{in}}\ f_j$$

This makes sense since the right hand side is symmetric in $i_1,\dots,i_n$. Note that the partial derivative is a complex number since ϕ_j has degree n. Now for $v\in V$ we have $(P\phi)\circ\Delta = n!\phi$. Namely,

$$\begin{aligned}(P\phi)\Delta(X_1e_1+\dots+X_me_m) &= \textstyle\sum_{i1,\dots,in} X_{i1}\dots X_{in}\ (P\phi)(e_{i1}\vee\dots\vee e_{in})\\ &= \textstyle\sum_j \sum_{i1,\dots,in} X_{i1}\dots X_{in}\ \dfrac{\partial^n\phi_j}{\partial X_{i1}\dots\partial X_{in}}\ f_j.\end{aligned}$$

By iteration of Euler's Theorem, that if F is homogeneous of degree r in variables X_i then $\sum_i X_i\partial F/\partial X_i = rF$, we obtain

$$= \textstyle\sum_j\ n!\ \phi_j(X_1,\dots,X_n)\ f_j \quad = \quad n!\ \phi(X_1e_1+\dots+X_me_m).$$

Example. If $\phi:V\longrightarrow\mathbb{C}$ is a quadratic form, so

$$\phi(X_1e_1+\dots+X_me_m) = \textstyle\sum_{i,j}\ a_{ij}X_iX_j$$

with $a_{ij} = a_{ji}$, then

$$(P\phi)\Big[(\textstyle\sum_i X_ie_i)\vee(\sum_j Y_je_j)\Big] = 2\ \sum_{i,j}\ a_{ij}\ X_i\ Y_j$$

is (2×) the corresponding symmetric bilinear form.

§7. SCHUR-WEYL DUALITY

Let V be a vector space of dimension m and let $n \in \mathbb{N}$. We know that T^nV is a $\mathbb{C}S_n$-module, so we have a map

$$\mathbb{C}S_n \longrightarrow \mathrm{End}_{\mathbb{C}}(T^nV) \quad \text{sending } \sigma \in S_n \text{ to } (x \longmapsto \sigma x)$$

Also, regarding V as a representation of $GL(V)$ in the natural way, T^nV becomes a $\mathbb{C}GL(V)$-module, and we have a corresponding map

$$\mathbb{C}GL(V) \longrightarrow \mathrm{End}_{\mathbb{C}}(T^nV) \quad \text{sending } \phi \in GL(V) \text{ to } T^n\phi = \phi\otimes\ldots\otimes\phi.$$

Remark. In this section we prove Schur-Weyl duality, that the images of $\mathbb{C}S_n$ and $\mathbb{C}GL(V)$ in $\mathrm{End}_{\mathbb{C}}(T^nV)$ are each others centralizers. Despite its innocuous appearance this result is absolutely fundamental. For example it is precisely this fact which explains why the symmetric group and the general linear group are related.

Definition. The algebra $A^n(V)$ of bisymmetric transformations is the subalgebra of $\mathrm{End}_{\mathbb{C}}(T^nV)$ consisting of the endomorphisms which commute with the image of $\mathbb{C}S_n$. Thus

$$A^n(V) = \mathrm{End}_{\mathbb{C}S_n}(T^nV).$$

Since $\mathbb{C}S_n$ is semisimple and T^nV is a finite dimensional $\mathbb{C}S_n$-module, $A^n(V)$ is a semisimple $\mathbb{C}$-algebra by §1 Lemma 4.

We set $W = \mathrm{End}_{\mathbb{C}}(V)$, which is a $\mathbb{C}GL(V)$-module by conjugation.

Lemma 1. There is an isomorphism

$$\alpha : T^nW \longrightarrow \mathrm{End}_{\mathbb{C}}(T^nV)$$

sending $f_1\otimes\ldots\otimes f_n$ to the map

$$v_1 \otimes \ldots \otimes v_n \longmapsto f_1(v_1) \otimes \ldots \otimes f_n(v_n).$$

This is an isomorphism of $\mathbb{C}GL(V)$-modules, and of $\mathbb{C}S_n$-modules.

PROOF. $T^nW = W\otimes\ldots\otimes W \cong (V\otimes V^*) \otimes \ldots \otimes (V\otimes V^*)$

$$\cong (V\otimes\ldots\otimes V) \otimes (V^*\otimes\ldots\otimes V^*)$$
$$\cong T^nV \otimes (T^nV)^* \cong \mathrm{End}_{\mathbb{C}}(T^nV).$$

Now T^nW has a natural structure of $\mathbb{C}S_n$-module, and $\mathrm{End}_{\mathbb{C}}(T^nV)$ inherits its structure from T^nV (as conjugation). One can check that α is an $\mathbb{C}S_n$-module map (exercise).

Lemma 2. $A^n(V) = \alpha(T^nW_{symm})$.

PROOF. $A^n(V)$ is the set of $x \in \mathrm{End}_{\mathbb{C}}(T^nV)$ fixed under the action of S_n, and T^nW_{symm} is the set of $y \in T^nW$ fixed under the action of S_n.

Lemma 3. Affine n-space $\mathbb{A}^n$ is irreducible, that is, if

$$\mathbb{A}^n = X \cup Y$$

and X and Y are Zariski-closed subsets, then $X = \mathbb{A}^n$ or $Y = \mathbb{A}^n$.

PROOF. The ring of regular functions on $\mathbb{A}^n$ is $R = \mathbb{C}[X_1,...,X_n]$. If X and Y are the zero sets of ideals I,J in R, then the assumption is that any maximal ideal contains either I or J. If I and J are both non-zero then we can pick $0\neq i\in I$ and $0\neq j\in J$. Now any maximal ideal contains ij, so

$$(ij)(a_1,...,a_n) = 0$$

for all $a_1,...,a_n\in\mathbb{C}$. Thus by Hilbert's Nullstellensatz $ij \in \sqrt{\{0\}} = \{0\}$, which contradicts the fact that R is an integral domain.

Lemma 4. If Y is a subspace of $\mathbb{C}^d$, then identifying $\mathbb{C}^d = \mathbb{A}^d$, Y is Zariski-closed.

PROOF. Choose a basis $f_1,...,f_h$ of $\mathrm{Hom}_{\mathbb{C}}(\mathbb{C}^d/Y,\mathbb{C})$, and regard these as maps $\mathbb{C}^d\longrightarrow\mathbb{C}$. Then Y is the zero set of the f_i.

Lemma 5. T^nW_{symm} is spanned by the $\phi\otimes...\otimes\phi$ with $\phi \in GL(V)$.

PROOF. Let X be the subspace of T^nW_{symm} spanned by the $\phi\otimes...\otimes\phi$ with $\phi \in GL(V)$. Now the map

$$W \xrightarrow{\alpha} T^nW, \quad \phi\longmapsto\phi\otimes...\otimes\phi$$

is a regular map between the affine spaces

$$W \cong \mathbb{A}^{m^2} \text{ and } T^nW \cong \mathbb{A}^{m^{2n}}.$$

Since X is a subspace, it is Zariski-closed by Lemma 4, and hence $\alpha^{-1}(X)$ is Zariski-closed. Thus

$$W = \alpha^{-1}(X) \cup \{\text{the endomorphisms with determinant zero}\}$$

is a union of Zariski-closed subsets. But $\mathbb{A}^{m^2}$ is irreducible, so $\alpha^{-1}(X)=W$. Thus X contains all maps of the form $\phi\otimes...\otimes\phi$ with $\phi \in W$. But these span

T^nW_{symm}, since the $\phi\vee\ldots\vee\phi$ span S^nW by §6 Lemma 7.

Restating this, we have

Theorem. $A^n(V)$ is spanned by the $T^n\phi$ with $\phi \in GL(V)$.

Finally, we have Schur-Weyl duality

Theorem. The images of $\mathbb{C}S_n$ and $\mathbb{C}GL(V)$ in $End_{\mathbb{C}}(T^nV)$ are each others centralizers.

PROOF. The statement that the image of $\mathbb{C}GL(V)$ is the centralizer of the image of $\mathbb{C}S_n$ is just a reformulation of the assertion that $A^n(V)$ is spanned by the $T^n\phi$ with $\phi\in GL(V)$, which was the last theorem.

Recall that $A^n(V)$ is a semisimple $\mathbb{C}$-algebra. By §1 Lemma 5 we know that $\mathbb{C}S_n$ maps onto $End_{A^n(V)}(T^nV)$, and since the image of $GL(V)$ spans $A^n(V)$ it follows that

$$End_{A^n(V)}(T^nV) = End_{\mathbb{C}GL(V)}(T^nV).$$

Thus $\mathbb{C}S_n$ maps onto $End_{\mathbb{C}GL(V)}(T^nV)$, or in other words, the image of $\mathbb{C}S_n$ in $End_{\mathbb{C}}(T^nV)$ is the centralizer of the image of $\mathbb{C}GL(V)$.

§8. DECOMPOSITION OF TENSORS

Still V is a vector space of dimension m.

One learns in school physics that any rank two tensor, ie any element of $V\otimes V$, can be written in a unique way as a sum of a symmetric and an antisymmetric tensor. The Young symmetrizers enable one to generalize this to higher rank tensors, namely by §3 Lemma 6 we have

$$(*) \qquad T^nV = \bigoplus_{\lambda \text{ a partition of } n \text{ and } \Sigma_\lambda \text{ standard}} h(\Sigma_\lambda)T^nV$$

Example. $h_{(1^n)}T^nV = T^nV_{anti} \cong \Lambda^nV$ and $h_{(n)}T^nV = T^nV_{symm} \cong S^nV$, so taking n=2 this decomposition becomes

$$T^2V = T^2V_{anti} \oplus T^2V_{symm} \cong \Lambda^2V \oplus S^2V.$$

Since the actions of S_n and GL(V) on T^nV commute, if λ is a partition of n and Σ_λ is a Young tableau with frame $[\lambda]$, then $h(\Sigma_\lambda)T^nV$ is a $\mathbb{C}GL(V)$-submodule of T^nV. Note that $h(\Sigma_\lambda)T^nV \cong h_\lambda T^nV$ as $\mathbb{C}GL(V)$-modules, for

$$h_\lambda = \sigma\ h(\Sigma_\lambda)\ \sigma^{-1}$$

for some $\sigma\in S_n$, so premultiplication by σ induces an isomorphism

$$h(\Sigma_\lambda)T^nV \longrightarrow h_\lambda T^nV.$$

Lemma 1. The non-zero modules $h_\lambda T^nV$ with λ a partition of n, are non-isomorphic simple $\mathbb{C}GL(V)$-modules. If M is an $A^n(V)$-module, and M is regarded as a $\mathbb{C}GL(V)$-module by restriction via the natural map $\mathbb{C}GL(V)\longrightarrow A^n(V)$, then M is isomorphic to a direct sum of copies of the $h_\lambda T^nV$.

Proof. Recall that

$$A^n(V) = \mathrm{End}_{\mathbb{C}S_n}(T^nV) \quad \text{and} \quad h_\lambda T^nV \cong \mathrm{Hom}_{\mathbb{C}S_n}(\mathbb{C}S_nh_\lambda,\ T^nV)$$

by §1 Lemma 6. By the Artin-Wedderburn Theorem and §1 Lemma 4, the non-zero spaces $h_\lambda T^nV$ are a complete set of non-isomorphic simple $A^n(V)$-modules. Note also that $A^n(V)$ is semisimple, so the lemma follows from the next two assertions, which both follow immediately from the fact proved in §7 that the map $\mathbb{C}GL(V)\longrightarrow A^n(V)$ is onto.

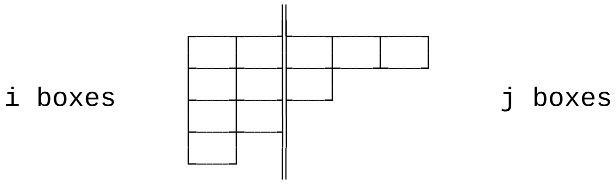
(1) If M is an $A^n(V)$-module and N is a $\mathbb{C}GL(V)$-submodule of M, then N is

an $A^n(V)$-submodule, and

(2) If M and N are $A^n(V)$-modules and $\theta:M\longrightarrow N$ is a $\mathbb{C}GL(V)$-module map, then θ is an $A^n(V)$-module map.

Remark. Thus (*) is a decomposition of T^nV into $\mathbb{C}GL(V)$-submodules which are either zero or simple. Obviously it is important to know which of these submodules are non-zero, and that is what the rest of this section is devoted to. First we have a rather technical lemma.

Let λ be a partition and suppose that $[\lambda]$ is partitioned into two non-empty parts, say of i and j = n-i boxes, by a vertical bar.



Let Σ_λ be a tableau whose numbers in the left hand part are {1,...,i} and in the right hand part are {i+1,...,n}. Let μ be the partition of i corresponding to the left hand part, and let Σ_μ be the restriction of Σ_λ to $[\mu]$. Let ν be the partition of j corresponding to the right hand part, and let Σ_ν be the corresponding tableau. This is a map from $[\nu]$ to $\{1',\ldots,j'\}$ if we set $1'=i+1$, $2'=i+2$, ..., $j'=n$.

Lemma 2. There is a $\mathbb{C}GL(V)$-module surjection

$$\left[h(\Sigma_\mu)T^iV\right] \otimes \left[h(\Sigma_\nu)T^jV\right] \quad \twoheadrightarrow \quad \left[h(\Sigma_\lambda)T^nV\right].$$

Proof. $S_i = \mathrm{Aut}\{1,\ldots,i\}$ and $S_j = \mathrm{Aut}\{1',\ldots,n'\}$ are embedded in S_n, so we can regard $\mathbb{C}S_i$ and $\mathbb{C}S_j$ as subsets of $\mathbb{C}S_n$ which commute. Now $\mathrm{Col}(\Sigma_\lambda) = \mathrm{Col}(\Sigma_\mu)\times\mathrm{Col}(\Sigma_\nu)$ and $H = \mathrm{Row}(\Sigma_\mu)\times\mathrm{Row}(\Sigma_\nu)$ is the subgroup of $\mathrm{Row}(\Sigma_\lambda)$ on the permutations which keep each number on the same side of the bar. Let $\mathrm{Row}(\Sigma_\lambda) = \bigcup_i r_i H$ be a coset decomposition.

$$\begin{aligned} h(\Sigma_\lambda) &= \textstyle\sum_{r\in\mathrm{Row}(\Sigma_\lambda)} \sum_{c\in\mathrm{Col}(\Sigma_\lambda)} \varepsilon_c\, rc \\ &= \textstyle\sum_i \sum_{r'\in\mathrm{Row}(\Sigma_\mu)} \sum_{r''\in\mathrm{Row}(\Sigma_\nu)} \sum_{c'\in\mathrm{Col}(\Sigma_\mu)} \sum_{c''\in\mathrm{Col}(\Sigma_\nu)} \varepsilon_{c'}\varepsilon_{c''} r_i r' r'' c' c'' \\ &= \textstyle\sum_i r_i h(\Sigma_\mu) h(\Sigma_\nu). \end{aligned}$$

Thus $h(\Sigma_\lambda)h(\Sigma_\mu)h(\Sigma_\nu) = \sum_i r_i h(\Sigma_\mu)^2 h(\Sigma_\nu)^2 = \alpha\ h(\Sigma_\lambda)$ where $\alpha = i!j!/f_\mu f_\nu$.

We have a $\mathbb{C}GL(V)$-module map $T^iV \otimes T^jV \cong T^nV \twoheadrightarrow h(\Sigma_\lambda)T^nV$ given by premultiplying by $h(\Sigma_\lambda)$. The restriction of this map to

$$\left[h(\Sigma_\mu)T^iV\right] \otimes \left[h(\Sigma_\nu)T^jV\right]$$

is onto, since

$$h(\Sigma_\lambda)\ (x \otimes y) = 1/\alpha\ h(\Sigma_\lambda)\left(\ h(\Sigma_\mu)x \ \otimes\ h(\Sigma_\nu)y\ \right).$$

Lemma 3.

(1) If $\lambda_{m+1} = 0$ then $h_\lambda T^nV \neq 0$.

(2) If $n>0$ and $\lambda_m = 0$ then $\dim_{\mathbb{C}} h_\lambda T^nV \geq 2$.

Proof.

(1) Let i_j = row in which j occurs in Σ_λ^0, and $x = e_{i_1}\otimes\ldots\otimes e_{i_n}$. Then for $\sigma\in S_n$

$$\begin{aligned} \sigma x = x \ &\Leftrightarrow\ i_j = i_{\sigma^{-1}(j)} \text{ for } 1\leq j\leq n \\ &\Leftrightarrow\ j \text{ and } \sigma^{-1}(j) \text{ occur in the same row} \\ &\Leftrightarrow\ \sigma \in \mathrm{Row}(\Sigma_\lambda^0). \end{aligned}$$

Thus the coefficient of x in the decomposition of $h_\lambda x$ wrt the standard basis of T^nV is $|\mathrm{Row}(\Sigma_\lambda^0)| \neq 0$, so $h_\lambda x \neq 0$.

(2) If $y = e_{1+i_1}\otimes\ldots\otimes e_{1+i_n}$ then the argument above shows that $h_\lambda x$ and $h_\lambda y$ are linearly independent.

Lemma 4. If $\lambda_{m+1} = 0$ and $\lambda_m > 0$ then

$$h_\lambda T^nV \cong \Lambda^m(V) \otimes h_{(\lambda_1-1,\ldots,\lambda_m-1)}T^{n-m}V.$$

Proof. Divide $[\lambda]$ into the first column and the rest. Let Σ_λ be a tableau whose first column consists of the numbers $\{1,\ldots,m\}$. By Lemma 2 there is a surjection

$$\Lambda^m(V) \otimes h(\Sigma_\nu)T^{n-m}V \twoheadrightarrow h(\Sigma_\lambda)T^nV,$$

where $\nu = (\lambda_1-1,\lambda_2-1,\ldots,\lambda_m-1)$. Using the usual isomorphisms this gives a map

$$\Lambda^m(V) \otimes h_\nu T^{n-m}V \twoheadrightarrow h_\lambda T^nV.$$

Now both $h_\nu T^{n-m}V$ and $h_\lambda T^nV$ are non-zero, and hence are simple

$\mathbb{C}GL(V)$-modules by Lemma 1. Since $\Lambda^m(V)$ is one-dimensional, both sides are simple modules and the map must be an isomorphism.

Theorem. If λ is a partition of n and $m = \dim_{\mathbb{C}} V$, then

$$\dim_{\mathbb{C}} h_\lambda T^n V = \begin{cases} 0 & (\lambda_{m+1} \neq 0) \\ 1 & (\lambda_1 = \lambda_2 = \ldots = \lambda_m,\ \lambda_{m+1} = 0) \\ \geq 2 & (\text{else}) \end{cases}$$

PROOF. If $\lambda_{m+1} \neq 0$ then $[\lambda]$ has $i>m$ rows and as in the previous lemma there is a surjection $\Lambda^i(V) \otimes h_\nu T^{n-i}V \twoheadrightarrow h_\lambda T^n V$. But $\Lambda^i(V) = 0$.

On the other hand, if $\lambda_{m+1} = 0$, then by iterating Lemma 4 we have

$$\dim_{\mathbb{C}} h_\lambda T^n V = \dim_{\mathbb{C}} h_{(\lambda_1-\lambda_m,\ldots,\lambda_{m-1}-\lambda_m)} T^{n-m\lambda_m} V$$

which is one if $\lambda_1 = \ldots = \lambda_m$, and otherwise is ≥ 2 by Lemma 3.

§9. RATIONAL REPRESENTATIONS OF GL(V)

Throughout, V is a vector space with basis $e_1,\dots,e_m$.

Definition. A finite dimensional $\mathbb{C}GL(V)$-module W with basis $w_1,\dots,w_h$ is said to be rational (resp. polynomial, resp. homogeneous n-ic) provided that there are rational functions (resp. polynomials, resp. homogeneous polynomials of degree n) $f_{ij}(X_{rs})$ $(1\le i,j\le h)$ in the m^2 variables X_{rs} $(1\le r,s\le m)$, such that the map

$$\begin{tikzcd} GL_m(\mathbb{C}) \arrow[r, "\text{basis } e_i"] & GL(V) \arrow[r, "\text{representation}"] & \mathrm{End}_{\mathbb{C}}(W) \arrow[r, "\text{basis } w_i"] & M_h(\mathbb{C}) \end{tikzcd}$$

sends a matrix $(A_{rs})_{rs} \in GL_m(\mathbb{C})$ to the matrix $(f_{ij}(A_{rs}))_{ij}$.

Lemma 1. These notions do not depend on the bases $e_1,\dots,e_m$ and $w_1,\dots,w_h$.

Remark. W is a rational $\mathbb{C}GL(V)$-module if and only if the map $GL(V)\longrightarrow GL(W)$ is a regular map of affine varieties. Recall that a rational map of affine varieties is not everywhere defined: we definitely don't want that.

Examples.

(1) T^nV is a homogeneous n-ic $\mathbb{C}GL(V)$-module.

(2) $\mathbb{C} \oplus V$ is polynomial, but not homogeneous.

(3) V^* is rational, but not polynomial.

Lemma 2.

(1) Submodules, quotient modules and direct sums of rational (resp. polynomial, resp. homogeneous n-ic) modules are of the same type.

(2) If U is rational, then so is U^*.

(3) If U and W are rational (resp. polynomial, resp homogeneous n-ic and n'-ic) then $U\otimes W$ is rational (resp. polynomial, resp. homogeneous $n+n'$-ic).

(4) If U is homogeneous n-ic and homogeneous n'-ic, then $U=0$.

PROOF. For (2) note that the entries of A^{-1} are rational functions of the entries of A.

Definition. If $\lambda_1\geq\lambda_2\geq\dots\geq\lambda_m\geq 0$ and $\sum_i \lambda_i = n$, we set

$$D_{\lambda_1,..,\lambda_m}(V) = h_{(\lambda_1,...,\lambda_m)}\, T^nV.$$

Theorem. Every homogeneous n-ic $\mathbb{C}GL(V)$-module is a direct sum of simple submodules. The modules $D_{\lambda_1,..,\lambda_m}(V)$ with $\lambda_1\geq\dots\geq\lambda_m\geq 0$ and $\sum_i \lambda_i = n$ are a complete set of non-isomorphic simple homogeneous n-ic $\mathbb{C}GL(V)$-modules.

PROOF. In view of §8 Lemma 1, it suffices to show that any homogeneous n-ic $\mathbb{C}GL(V)$-module is obtained from some $A^n(V)$-module by restriction.

Let U be a homogeneous n-ic $\mathbb{C}GL(V)$-module, so U corresponds to a map $\rho:GL(V)\longrightarrow \mathrm{End}_{\mathbb{C}}(U)$. In the following diagram, the maps across the top are the natural maps, and their composite $\gamma:GL(V)\longrightarrow A^n(V)$ is in fact the natural map we use for restricting $A^n(V)$-modules to $\mathbb{C}GL(V)$-modules. We shall show that there are maps ρ_i making the diagram commute.

$$\begin{tikzcd}
GL(V) \arrow[r, hook] \arrow[d, "\rho"'] & \mathrm{End}_{\mathbb{C}}(V) \arrow[r] \arrow[d, dashed, "\rho_1"'] & S^n(\mathrm{End}_{\mathbb{C}}(V)) \arrow[r, phantom, "\cong"] \arrow[d, dashed, "\rho_2"'] & T^n(\mathrm{End}_{\mathbb{C}}(V))_{\mathrm{symm}} \arrow[r, phantom, "\cong"] \arrow[d, dashed, "\rho_3"'] & A^n(V) \arrow[d, dashed, "\rho_4"'] \\
\mathrm{End}_{\mathbb{C}}(U) \arrow[r, equal] & \mathrm{End}_{\mathbb{C}}(U) \arrow[r, equal] & \mathrm{End}_{\mathbb{C}}(U) \arrow[r, equal] & \mathrm{End}_{\mathbb{C}}(U) \arrow[r, equal] & \mathrm{End}_{\mathbb{C}}(U)
\end{tikzcd}$$

Since U is homogeneous n-ic, we can extend the domain of definition of ρ to obtain a homogeneous n-ic map ρ_1. By the property of symmetric powers there is a linear map ρ_2. Since the remaining maps across the top are isomorphisms there are certainly linear maps ρ_3 and ρ_4, as required.

Now $\rho_4(1) = \rho_4(\gamma(1)) = \rho(1) = 1$ and

$$\rho_4(\gamma(gg')) = \rho(gg') = \rho(g)\rho(g') = \rho_4(\gamma(g))\rho_4(\gamma(g'))$$

for $g,g'\in GL(V)$. Since $\gamma(GL(V))$ spans $A^n(V)$ it follows that ρ_4 is a $\mathbb{C}$-algebra map. This turns U into an $A^n(V)$-module, and the restriction to $\mathbb{C}GL(V)$ is the module we started with, as required.

Lemma 3. Every polynomial module for $GL(\mathbb{C}) = \mathbb{C}^\times$ decomposes as a direct sum of submodules on which $g \in \mathbb{C}^\times$ acts as multiplication by g^n (some n).

PROOF. Here is a silly proof. If $\rho : \mathbb{C}^\times \longrightarrow GL(U) \cong GL_h(\mathbb{C})$ is a polynomial representation, then each $\rho(g)_{ij}$ is a polynomial in g, and we can choose $N\in\mathbb{N}$ such that each $\rho(g)_{ij}$ has degree strictly less than N. By restriction, U becomes a $\mathbb{C}G$-module where

$$G = \{\exp 2\pi ij/N \mid 0\leq j<N\} \subset \mathbb{C}^\times$$

is cyclic of order N. Now

$$U = U_1 \oplus U_2 \oplus \ldots \oplus U_h$$

as a $\mathbb{C}G$-module, with each U_i one-dimensional and $g \in G$ acting as multiplication by g^{n_i} on U_i $(0 \leq n_i < N)$ (since these are the possible simple $\mathbb{C}G$-modules). Choosing non-zero elements of the U_i gives a basis of U, and if $(\rho(g)_{ij})$ now denotes the matrix of $\rho(g)$ with respect to this basis then

$$\rho(g)_{ij} = \begin{cases} g^{n_i} & (i = j) \\ 0 & (i \neq j) \end{cases}$$

for $g = \exp\{2\pi ij/N\}$ with $0 \leq j < N$, and hence for all $g \in \mathbb{C}^\times$ since the $\rho(g)_{ij}$ are polynomials of degree < N in g.

Lemma 4. Every polynomial $\mathbb{C}GL(V)$-module decomposes as a direct sum of homogeneous n-ic modules.

PROOF. Say $\rho: GL(V) \longrightarrow GL(U)$ is a polynomial representation of GL(V). The inclusion $\mathbb{C}^\times \longrightarrow GL(V)$ enables us to regard U as a polynomial representation of $\mathbb{C}^\times$, so by Lemma 3,

$$U = U_0 \oplus \ldots \oplus U_N$$

with $\alpha 1 \in GL(V)$ acting as multiplication by α^n on U_n. If $u \in U_n$ and $g \in GL(V)$, let

$$gu = u_0 + \ldots + u_N$$

with $u_i \in U_i$. Now $(\alpha 1)gu = g(\alpha 1)u$ for $\alpha \in \mathbb{C}^\times$, so

$$u_0 + \alpha u_1 + \alpha^2 u_2 + \ldots + \alpha^N u_N = \alpha^i u_0 + \alpha^i u_1 + \alpha^i u_2 + \ldots + \alpha^i u_N$$

and hence $u_i = 0$ for $i \neq n$. Thus $gu \in U_n$ and the spaces U_n are $\mathbb{C}GL(V)$-submodules of U. Since $\alpha 1$ acts as multiplication by α^n on U_n it follows that U_n is a homogeneous n-ic $\mathbb{C}GL(V)$-module.

Theorem. Every polynomial $\mathbb{C}GL(V)$-module is a direct sum of simple submodules. The modules $D_{\lambda_1,..,\lambda_m}(V)$ with $\lambda_1 \geq \ldots \geq \lambda_m \geq 0$ are a complete set of non-isomorphic simple polynomial $\mathbb{C}GL(V)$-modules.

Definition. If $n \in \mathbb{Z}$ then the one-dimensional $\mathbb{C}GL(V)$-module corresponding to the representation

$$GL(V) \longrightarrow \mathbb{C}^\times, \; g \longmapsto [\det(g)]^n$$

is denoted by $\det^n$. Thus $\det^1 \cong \Lambda^m(V)$, $\det^n \cong T^n(\det^1)$ if $n \geq 0$, and $\det^n \cong (\det^{-n})^*$ if $n \leq 0$.

<u>Definition</u>. If $\lambda_1 \geq \ldots \geq \lambda_m$ but $\lambda_m < 0$, we define

$$D_{\lambda_1,\ldots,\lambda_m}(V) = D_{\lambda_1-\lambda_m,\ldots,\lambda_{m-1}-\lambda_m,0}(V) \otimes \det^{\lambda m}$$

<u>Remark</u>. If $\lambda_1 \geq \ldots \geq \lambda_m > 0$ then we have already seen that

$$D_{\lambda_1,\ldots,\lambda_m}(V) \cong D_{\lambda_1-\lambda_m,\ldots,\lambda_{m-1}-\lambda_m,0}(V) \otimes \det^{\lambda m}$$

<u>Theorem</u>. Every rational $\mathbb{C}GL(V)$-module is a direct sum of simple submodules. The modules $D_{\lambda 1,\ldots,\lambda m}(V)$ with $\lambda_1 \geq \ldots \geq \lambda_m$ are a complete set of non-isomorphic simple rational $\mathbb{C}GL(V)$-modules.

PROOF. The rational functions $f : GL(V) \longrightarrow \mathbb{C}$ are all of the form $f = p/\det^i$ with p a polynomial function. Thus if U is a rational $\mathbb{C}GL(V)$-module, then $W = U \otimes \det^N$ is a polynomial $\mathbb{C}GL(V)$-module for some N. Since W decomposes as a direct sum of simples, so does U. If U is simple, then so is W, and thus $W \cong D_{\mu 1,\ldots,\mu m}(V)$ for some $\mu_1 \geq \ldots \geq \mu_m$. Finally $U \cong D_{\mu 1-N,\ldots,\mu m-N}(V)$, using the remark above.

<u>Theorem</u>. The one-dimensional rational $\mathbb{C}GL(V)$-modules are precisely the $\det^n$ with $n \in \mathbb{Z}$.

PROOF. After passing, as above, to polynomial modules, this follows from the theorem in §8.

§10. WEYL'S CHARACTER FORMULA

Notation. V is a vector space of dimension m. If $\lambda=(\lambda_1,\dots,\lambda_m) \in \mathbb{Z}^m$ and $\lambda_1\geq\dots\geq\lambda_m$, then the character of the $\mathbb{C}GL(V)$-module

$$D_{\lambda 1,\dots,\lambda m}(V) \quad (= h_\lambda T^n V \text{ if } \lambda_m\geq 0)$$

is denoted by ϕ_λ.

Lemma 1. If ξ is an endomorphism of V then $\phi_\lambda(\xi)$ is a symmetric rational function of the eigenvalues of ξ.

PROOF. The function $P(x_1,\dots,x_m) = \phi_\lambda(\mathrm{diag}(x_1,\dots,x_m))$ is a rational function of $x_1,\dots,x_m$, and it is symmetric since $\mathrm{diag}(x_1,\dots,x_m)$ is conjugate to $\mathrm{diag}(x_{\tau(1)},\dots,x_{\tau(m)})$ for $\tau \in S_m$.

Now choose a basis of V so that the matrix A_1 of ξ is in Jordan Normal Form, and for $t\in\mathbb{C}$, let A_t be the matrix obtained from A_1 by changing the 1's on the upper diagonal into t's. Let ξ_t be the endomorphism corresponding to A_t. For $t\neq 0$, A_t is conjugate to A_1, so $\phi_\lambda(\xi_t) = \phi_\lambda(\xi)$. Since ϕ_λ is a rational function it is continuous (where defined), so

$$\phi_\lambda(\xi) = \lim_{t\to 0} \phi_\lambda(\xi_t) = \phi_\lambda\left(\lim_{t\to 0} \xi_t \right) = \phi_\lambda(\xi_0) = P(x_1,\dots,x_m).$$

Exercise. Phrase this using the Zariski topology, by means of the discriminant of the characteristic polynomial.

Lemma 2. Let α be a conjugacy class in S_n with cycle type $n^{\alpha n}\dots 1^{\alpha 1}$ and let ξ be an endomorphism of V with eigenvalues $x_1,\dots,x_m$. If $s_i=x_1^i+\dots+x_m^i$, then

$$s_1^{\alpha 1}\dots s_n^{\alpha n} = \sum_\lambda \chi_\lambda(\alpha)\, \phi_\lambda(\xi)$$

with summation over partitions λ of n with $\leq$m parts

PROOF. Let $g \in \alpha$. We may suppose that ξ has matrix $\mathrm{diag}(x_1,\dots,x_m)$ with respect to the standard basis $e_1,\dots,e_m$ of V. Consider the endomorphism of $T^n V$ sending x to $g\xi x = \xi g x$. We compute its trace in two ways.

Considering $T^n V$ as a $\mathbb{C}S_n$-module, by §1 Lemma 3, we have

$$T^nV = \bigoplus_\lambda \mathbb{C}S_n h_\lambda \otimes \mathrm{Hom}_{\mathbb{C}S_n}(\mathbb{C}S_n h_\lambda, T^nV)$$

and then by §1 Lemma 6 this becomes

$$T^nV \cong \bigoplus_\lambda (\mathbb{C}S_n h_\lambda) \otimes (h_\lambda T^nV).$$

Now this is an isomorphism as both a $\mathbb{C}S_n$-module and an $\mathrm{End}_{\mathbb{C}S_n}(T^nV)$-module, and since the action of GL(V) on T^nV commutes with that of S_n, the corresponding action of $g\xi$ on the right hand side is given by the action of g on $\mathbb{C}S_n h_\lambda$ and of ξ on $h_\lambda T^nV$, so the trace of this action is $\sum_\lambda \chi_\lambda(\alpha)\ \phi_\lambda(\xi)$.

On the other hand we can compute the trace directly:

$$g\xi(e_{i_1} \otimes \ldots \otimes e_{i_n}) = x_{i_1} \ldots x_{i_n}\ \ e_{i_{g^{-1}(1)}} \otimes \ldots \otimes e_{i_{g^{-1}(n)}}$$

so the trace is $\sum x_{i1} \ldots x_{in}$ summed over

$$\left\{ (i_1, \ldots, i_m) \;\middle|\; 1 \le i_1 \le m,\ \ldots,\ 1 \le i_n \le m \text{ and } i_{g^{-1}(j)} = i_j \text{ for each } j \right\}$$

Now the condition that $i_{g^{-1}(j)} = i_j$ for each j is equivalent to requiring that the function $j \longmapsto i_j$ be constant on the cycles involved in g. It follows that the trace is equal to $s_1^{\alpha 1} \ldots s_n^{\alpha n}$.

Equating the two calculations of the trace gives the required equation.

Theorem. Let λ be a partition of n with $\le$ m parts and let ξ be an endomorphism of V with eigenvalues $x_1, \ldots, x_m$. If $s_i = x_1^i + \ldots + x_m^i$, then

$$\phi_\lambda(\xi) = \sum_{\alpha \text{ conj class}} \frac{\chi_\lambda(\alpha)}{\alpha_1! \ldots \alpha_n!} \left(\frac{s_1}{1}\right)^{\alpha_1} \ldots \left(\frac{s_n}{n}\right)^{\alpha_n}.$$

PROOF. Take the formula in Lemma 2, multiply by $n_\alpha\ \chi_\mu(\alpha)$, and sum over α, using the orthogonality of the χ_λ.

Theorem. (Weyl's Character Formula for the general linear group).

If $\xi \in$ GL(V) has eigenvalues $x_1, \ldots, x_m$ then

$$\phi_\lambda(\xi) = \frac{|x^{\ell_1}, \ldots, x^{\ell_m}|}{|x^{m-1}, \ldots, 1|}$$

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

PROOF. First suppose that the $\lambda_i \geq 0$, so that λ is a partition of n with $\leq m$ parts. By Lemma 2 and the character formula for the symmetric group we know that $s_1^{\alpha 1} \dots s_n^{\alpha n}$ is equal to both

$$\sum \chi_\lambda(\alpha)\ |x^{\ell_1},\dots,x^{\ell_m}|\ /\ |x^{m-1},\dots,1| \quad \text{and} \quad \sum \chi_\lambda(\alpha)\ \phi_\lambda(\xi)$$

with summation over partitions λ of n with $\leq m$ parts. The orthogonality of the χ_λ enables us to equate coefficients.

For general λ, since

$$D_\lambda(V) \cong D_{\lambda 1-\lambda m,\dots,\lambda m-\lambda m}(V) \otimes \det^{\lambda m}$$

it follows that

$$\phi_\lambda(\xi) = \frac{|x^{\ell 1-\lambda m},\dots,x^{\ell m-\lambda m}|}{|x^{m-1},\dots,\ 1\ |} \cdot (x_1 \dots x_m)^{\lambda m} = \frac{|x^{\ell 1},\dots,x^{\ell m}|}{|x^{m-1},\dots,\ 1\ |}.$$

Remark. Our proof of Weyl's character formula looks quite short, but this is because most of the proof, the character formula for the symmetric group, is in §4. There is, however, another approach to these formulae which derives Weyl's character formula first, and then deduces the character formula for S_n. The idea is to use integration to compute Weyl's character formula for the compact subgroup U_m of unitary matrices in $GL_m(\mathbb{C})$, and then to translate that to $GL_m(\mathbb{C})$. Finally one can use Lemma 2 to pass to characters of the symmetric group. See the details and the discussion in [Weyl].

Theorem. The degree of ϕ_λ, the dimension of $D_{\lambda 1,\dots,\lambda m}(V)$, is

$$\prod_{1\leq i<j\leq m} (\ell_i-\ell_j)\ /\ \prod_{1\leq i<j\leq m} (j-i).$$

PROOF. For $t \in \mathbb{C}$ set

$$x_m = 1,\ x_{m-1} = e^t,\ x_{m-2} = e^{2t},\ \dots,\ x_1 = e^{(m-1)t}$$

Then

$$|x^{\ell_1},\dots,x^{\ell_m}| = \prod_{i<j} (e^{\ell_i t}-e^{\ell_j t})$$

since it is the transpose of a Vandermonde matrix. The term of lowest degree in t is $\prod_{i<j} [(\ell_i-\ell_j)t]$. Also

$$|x^{m-1},\dots,1| = \prod_{i<j} (e^{(m-i)t}-e^{(m-j)t})$$

and the term of lowest degree in t is $\prod_{i<j} [(j-i)t]$.

If $\xi_t = \mathrm{diag}(x_1,...,x_m)$, then the degree of ϕ_λ is

$$\phi_\lambda(1) \;=\; \lim_{t\to 0} \phi_\lambda(\xi_t) \;=\; \prod_{1\le i<j\le m} (\ell_i-\ell_j) \;/\; \prod_{1\le i<j\le m} (j-i)$$

by Weyl's character formula.

Theorem. If two rational $\mathbb{C}GL(V)$-modules have the same characters, then they are isomorphic.

Proof. As λ varies, the rational functions in Weyl's character formula are linearly independent elements of $\mathbb{C}(x_1,...,x_m)$.

Lemma 3. $D_{\lambda 1,...,\lambda m}(V)^* \cong D_{-\lambda m,...,-\lambda 1}(V)$.

Proof. the character ψ of the left hand module is given by

$$\psi(\xi) \;=\; \phi_\lambda(\xi^{-1}) \;=\; \frac{|x^{-\ell_1},...,x^{-\ell_m}|}{|x^{-(m-1)},...,1\,|} \;=\; \frac{|x^{-\ell_m},...,x^{-\ell_1}|}{|1,...,x^{-(m-1)}|}$$

$$=\; \frac{|x^{m-1-\ell_m},...,x^{m-1-\ell_1}|}{|x^{m-1},\, ...,\quad 1\,|} \;=\; \phi_\mu(\xi)$$

where $\mu = (-\lambda_m,...,-\lambda_1)$. Thus the dual is isomorphic to $D_\mu(V)$.

In the same vein one has the following result, which is left as an exercise. We shall make extensive use of this formula later.

Clebsch-Gordan formula. If $V = \mathbb{C}^2$ and $p,q \in \mathbb{N}$ then

$$D_{p,0}(V) \otimes D_{q,0}(V) \cong \bigoplus_{r=0}^{\min(p,q)} D_{p+q-r,r}(V).$$

Remark. We list some important rational GL(V)-modules.

$\mathbb{C} = D_{0,0,...,0}(V)$

$\det^i = D_{i,i,...,i}(V)$

$V = D_{1,0,...,0}(V)$

$S^n(V) = D_{n,0,...,0}(V)$

$\Lambda^n(V) = D_{1,..,1,0,..,0}(V)$ (n≤m. With n 1's and m-n 0's)

$V^* = D_{0,...,0,-1}(V)$

$S^n(V)^* = D_{0,...,0,-n}(V)$

$\Lambda^n(V)^* = D_{0,..,0,-1,...,-1}(V)$ (n≤m. With m-n 0's and n -1's)

§11. SOME EXAMPLES OF INVARIANTS

The notion of an invariant embraces many classical constructions: the discriminant, the determinant, the Hessian, etc..Presumably because of the importance of these examples, the idea of classifying all invariants arose. In this section we shall examine in detail some of the important examples of invariants, and a few simple cases in which all invariants can be classified. In the next two sections we shall investigate the general problem of classifying all invariants.

Definition. If U is a finite dimensional vector space, then $\mathbb{C}[U]$ denotes the set of all polynomial maps $U \longrightarrow \mathbb{C}$. This is an (infinite dimensional if $U \neq 0$) commutative $\mathbb{C}$-algebra via

$$(\lambda f)(u) = \lambda f(u) \qquad (f+f')(u) = f(u) + f'(u)$$
$$(ff')(u) = f(u)f'(u) \qquad 1_{\mathbb{C}[U]}(u) = 1$$

for $\lambda \in \mathbb{C}$, $f,f' \in \mathbb{C}[U]$ and $u \in U$.

Remarks.

(1) $\mathbb{C}[U]$ is the ring of regular functions of the affine variety $U \cong \mathbb{A}^{\dim U}$.

(2) We have

$$\mathbb{C}[U] \cong \oplus_{n=0}^{\infty} \mathrm{Hom}_{\mathbb{C},n}(U,\mathbb{C}) \cong \oplus_{n=0}^{\infty} S^n(U^*)$$

by polarization. This is the symmetric algebra on U^*.

(3) If U^* has basis $\xi_1,\dots,\xi_h$, then the map

$$\mathbb{C}[X_1,\dots,X_h] \longrightarrow \mathbb{C}[U], \quad X_i \longmapsto \xi_i$$

is an isomorphism of $\mathbb{C}$-algebras.

(4) If U is a $\mathbb{C}G$-module then $\mathbb{C}[U]$ is a $\mathbb{C}G$-module via

$$(gf)(u) = f(g^{-1}u) \quad \forall\ g\in G,\ f\in\mathbb{C}[U],\ u\in U.$$

If $g\in G$, then the map $\mathbb{C}[U] \longrightarrow \mathbb{C}[U]$, $f \longmapsto gf$, is a $\mathbb{C}$-algebra automorphism:

$$[g(ff')](u) = (ff')(g^{-1}u) = f(g^{-1}u)f'(g^{-1}u) =$$
$$= [(gf)(u)][(gf')(u)] = [(gf)(gf')](u)$$
$$(g1_{\mathbb{C}[U]})(u) = 1_{\mathbb{C}[U]}(g^{-1}u) = 1 = 1_{\mathbb{C}[U]}(u).$$

Definition. A function $f:U \longrightarrow W$ between $\mathbb{C}G$-modules is a concomitant if $f(gu) = gf(u)\ \forall\ u \in U$ and $g \in G$.

Examples.

(1) A linear concomitant is precisely a $\mathbb{C}G$-module map.

(2) $\Delta:U \longrightarrow S^nU,\ u \longmapsto u\vee\ldots\vee u$ is a homogeneous n-ic concomitant.

Definition. An invariant for a $\mathbb{C}G$-module U is a concomitant $f:U \longrightarrow \mathbb{C}$, so

$$f(gu) = f(u)\ \forall\ u \in U \text{ and } g \in G.$$

Lemma 1. The set

$$\mathbb{C}[U]^G = \{\ f \in \mathbb{C}[U] \mid gf = f\ \}.$$

of polynomial invariants for U is a $\mathbb{C}$-subalgebra of $\mathbb{C}[U]$.

PROOF. Trivial.

Remark. The main problem of invariant theory can now be formulated as computing $\mathbb{C}[U]^G$. Some important general results which I shall not cover are:

(1) If G is finite then $\mathbb{C}[U]^G$ is a finitely generated $\mathbb{C}$-algebra (E. Noether).

(2) If U is a rational $\mathbb{C}GL(V)$-module then $\mathbb{C}[U]^{GL(V)}$ and $\mathbb{C}[U]^{SL(V)}$ are finitely generated $\mathbb{C}$-algebras (Hilbert) and Cohen-Macaulay rings (Hochster and Roberts [Adv.Math. 13(1974) 115-175]). Moreover $\mathbb{C}[U]^{SL(V)}$ is a UFD, so a Gorenstein ring.

(3) One can compute an explicit bound on the number of generators needed for $\mathbb{C}[U]^{SL(V)}$ (V. Popov [Asterisque vol 87/88]).

Instead we consider in this section some simple examples.

Example (Symmetric polynomials).

A vector space U with basis $f_1,\ldots,f_n$ becomes a $\mathbb{C}S_n$-module via $\sigma f_i = f_{\sigma(i)}$. If $\xi_1,\ldots,\xi_n$ is the dual basis of U^*, the isomorphism

$$\mathbb{C}[X_1,\ldots,X_n] \longrightarrow \mathbb{C}[U],\ X_i \longmapsto \xi_i.$$

enables us to identify $\mathbb{C}[U]$ with $\mathbb{C}[X_1,...,X_n]$. The action of S_n on this is given by $\sigma X_i = X_{\sigma(i)}$. The set of polynomial invariants of the $\mathbb{C}S_n$-module U is thus

$$\mathbb{C}[X_1,...,X_n]^{S_n} = \{ f \in \mathbb{C}[X_1,...,X_n] \mid f \text{ is symmetric in the } X_i \}.$$

Recall that the elementary symmetric polynomials

$$E_i(X_1,...,X_n) \in \mathbb{C}[X_1,...,X_n]^{S_n}$$

are defined by

$$(t+X_1)...(t+X_n) = t^n + E_1 t^{n-1} + E_2 t^{n-2} + ... + E_n$$

so

$$E_1(X_1,...,X_n) = X_1 + ... + X_n,$$

$$E_i(X_1,...,X_n) = \textstyle\sum_{1\le j_1<j_2<...<j_i\le n} X_{j_1}X_{j_2}...X_{j_i}.$$

$$E_n(X_1,...,X_n) = X_1X_2...X_n.$$

The Fundamental Theorem of Symmetric Functions computes the polynomial invariants for U since it states that the $\mathbb{C}$-algebra map

$$\mathbb{C}[Y_1,...,Y_n] \longrightarrow \mathbb{C}[X_1,...,X_n]^{S_n},\ Y_i \longmapsto E_i(X_1,...,X_n)$$

is an isomorphism.

Before moving on to other examples, recall that if $f(t) \in \mathbb{C}[t]$ is a monic polynomial of degree n

$$f(t) = t^n + a_1t^{n-1} + ... + a_{n-1}t + a_n = (t+\lambda_1)...(t+\lambda_n),$$

then its <u>discriminant</u> is

$$\mathrm{disc}(f) = \textstyle\prod_{i<j} (\lambda_i-\lambda_j)^2.$$

Since the polynomial

$$|X^{n-1},...,1|^2 = \textstyle\prod_{i<j} (X_i-X_j)^2$$

is symmetric in the X_i, it can be expressed as a polynomial in the elementary symmetric polynomials $E_1,...,E_n$, say

$$\textstyle\prod_{i<j} (X_i-X_j)^2 = D(E_1(X_i),...,E_n(X_i)).$$

Since also $E_i(\lambda_1,...,\lambda_n) = a_i$, it follows that disc(f) is a polynomial in $a_1,...,a_n$. For example

$$\mathrm{disc}(t^2+bt+c) = b^2 - 4c\ ,\ \mathrm{disc}(t^3+bt+c) = -4b^3 - 27c^2$$

<u>Example</u> (The alternating group).

In the previous example, restriction enables us to consider U as a representation of the alternating group A_n (here we suppose $n\ge2$). The set of polynomial invariants is then $\mathbb{C}[X_1,...,X_n]^{A_n}$.

In this case there is another polynomial invariant, the Vandermonde

$$|X^{n-1},...,1| = \prod_{i<j} (X_i-X_j).$$

Now $|X^{n-1},...,1|^2$ is an S_n-invariant, so is a polynomial $D(E_1,...,E_n)$ in the elementary symmetric polynomials.

The $\mathbb{C}$-algebra map

$$\theta:\mathbb{C}[Y_1,..,Y_n,Z]\longrightarrow\mathbb{C}[X_1,..,X_n]^{A_n}, \quad \theta(Y_i)=E_i, \quad \theta(Z)=|X^{n-1},...,1|$$

is surjective, and the kernel is the ideal $(Z^2-D(Y_1,...,Y_n))$.

PROOF.

(1) θ is surjective. Let $\tau\in S_n$ be a transposition. (Here we use that $n\geq 2$!) If $f \in \mathbb{C}[X_1,...,X_n]^{A_n}$, then $f_s = f+\tau f$ is a symmetric polynomial since

$$\sigma f_s = \sigma f + \tau(\tau^{-1}\sigma\tau)f = f + \tau f = f_s$$

for $f \in A_n$, while

$$\tau f_s = \tau f+\tau^2 f = \tau f + f = f_s.$$

Similarly $f_a = f-\tau f$ is an alternating polynomial.

Now $f_a(X_1,...,X_n) = 0$ whenever $X_1 = X_2$, so by Hilbert's Nullstellensatz

$$f_a \in \sqrt{(X_1-X_2)} = (X_1-X_2).$$

Thus f_a is divisible by X_1-X_2. Similarly f_a is divisible by X_i-X_j $(i<j)$, and hence

$$f_a = |X^{n-1},...,1|.h$$

for some polynomial h. Clearly h is a symmetric polynomial. Thus

$$f = \tfrac{1}{2}(f_s + |X^{n-1},...,1|.h).$$

Since f_s and h are symmetric they are in the image of θ, and hence so is f. Thus θ is onto.

(2) $(Z^2-D(Y_i)) \subseteq \mathrm{Ker}(\theta)$. This is clear since

$$\theta(Z^2-D(Y_i)) = |X^{n-1},...,1|^2 - D(E_1,...,E_n) = 0.$$

(3) $\mathrm{Ker}(\theta) \subseteq (Z^2-D(Y_i))$. By polynomial division, any polynomial $P(Y_1,...,Y_n,Z)$ is of the form

$$P(Y_i,Z) = Q(Y_i,Z)\Big[Z^2-D(Y_i)\Big] + \Big[A(Y_i) + B(Y_i)Z\Big]$$

To prove the assertion it thus suffices to show that if $P \in \mathrm{Ker}(\theta)$ has form $A(Y_i)+B(Y_i)Z$, then $P=0$.

If $a_1,\ldots,a_n \in \mathbb{C}$, let

$$f(t) = t^n+a_1t^{n-1}+\ldots+a_n = (t+\lambda_1)\ldots(t+\lambda_n).$$

Then

$$P(a_1,\ldots,a_n,|\lambda^{n-1},\ldots,1|) = \theta(P)(\lambda_1,\ldots,\lambda_n) = 0$$

since $P \in \mathrm{Ker}(\theta)$. Exchanging λ_1 and λ_2 changes the sign of the Vandermonde, so

$$P(a_1,\ldots,a_n,\pm\delta) = 0 \quad \text{where} \quad \delta = [\mathrm{disc}(t^n+a_1t^{n-1}+\ldots+a_n)]^{1/2}.$$

Using the special form of P this becomes

$$A(a_1,\ldots,a_n) \pm B(a_1,\ldots,a_n)\delta = 0$$

and hence $A = B = 0$ on the Zariski-dense open subset

$$\{(a_1,\ldots,a_n) \in \mathbb{C}^n \mid \mathrm{disc}(t^n+a_1t^{n-1}+\ldots+a_n) \neq 0\}$$

of $\mathbb{C}^n$. A Zariski topology argument now shows that $A = B = 0$ everywhere on $\mathbb{C}^n$. Thus $P = 0$, as required.

Example (Characteristic polynomial).

Let V be a vector space of dimension m. Recall that the natural action of $GL(V)$ on $U = \mathrm{End}_{\mathbb{C}}(V)$ is given by

$$(g\bullet\theta)(v) = g\theta(g^{-1}v)$$

so $g\bullet\theta = g\theta g^{-1}$. If

$$\chi_\theta(t) = \det(t1_V + \theta)$$

denotes the characteristic polynomial of θ (more or less), and $c_n(\theta)$ is the coefficient of t^{m-n} in $\chi_\theta(t)$, then

$$\chi_{g\bullet\theta}(t) = \det(t1_V + g\theta g^{-1}) = \det g(t1_V + \theta)g^{-1} = \chi_\theta(t)$$

and so $c_n : U \longrightarrow \mathbb{C}$ is an homogeneous n-ic invariant. Note that if θ has eigenvalues $\lambda_1,\ldots,\lambda_m$, then putting θ into Jordan Normal Form we have

$$\chi_\theta(t) = \prod_{j=1}^m (t + \lambda_j)$$

so $c_n(\theta) = E_n(\lambda_1,\ldots,\lambda_m)$, so

$$c_1(\theta) = \lambda_1+\ldots+\lambda_m = \mathrm{tr}(\theta), \quad c_m(\theta) = \lambda_1\ldots\lambda_m = \det(\theta).$$

The $\mathbb{C}$-algebra map

$$c : \mathbb{C}[Y_1,\ldots,Y_m] \longrightarrow \mathbb{C}[U]^{GL(V)}, \; Y_n \longmapsto c_n$$

is an isomorphism.

Proof. It suffices to observe that if f is a polynomial invariant for U, then $f(\theta)$ is a symmetric polynomial function of the eigenvalues of θ. The proof of this is the same as the proof that the characters of rational $\mathbb{C}GL(V)$-modules are symmetric rational functions in the eigenvalues of $g\in GL(V)$, §10 Lemma 1.

Example (Discriminant of a quadratic form).

Let $V = \mathbb{C}^m$, and let $U = \mathrm{Hom}_{\mathbb{C},2}(V,\mathbb{C})$ be the set of quadratic forms on V. By polarization U can be identified with the set of symmetric $m\times m$ matrices A with

$$f(x) = x^TAx \quad (x \text{ a column vector in } \mathbb{C}^m).$$

We define the discriminant disc(f) of f by

$$\mathrm{disc}(f) = \det(A).$$

Now U is a $\mathbb{C}GL_m(\mathbb{C})$-module via

$$(gf)(x) = f(g^{-1}x) = x^T(g^{-T}Ag^{-1})x$$

for $x \in \mathbb{C}^m$, $g \in GL_m(\mathbb{C})$, so

$$\mathrm{disc}(gf) = \det(g^{-T}Ag^{-1}) = \det(g)^{-2}\mathrm{disc}(f).$$

Thus $\mathrm{disc}:U\longrightarrow\mathbb{C}$ is an $SL_m(\mathbb{C})$-invariant (but not an $GL_m(\mathbb{C})$-invariant).

The $\mathbb{C}$-algebra map $\mathbb{C}[X]\longrightarrow\mathbb{C}[U]^{SL_2(\mathbb{C})}$, $X\longmapsto\mathrm{disc}$ is an isomorphism.

PROOF. The map is injective since if there is a polynomial P with $P(\mathrm{disc}(f)) = 0$ for all $f\in U$, then $P(\lambda) = 0$ for all $\lambda\in\mathbb{C}$, since the quadratic form

$$f_\lambda(X_1,\dots,X_m) = X_1^2+X_2^2+\dots+\lambda X_m^2$$

has discriminant λ. Thus $P = 0$.

Now let $\theta:U\longrightarrow\mathbb{C}$ be a polynomial $SL_m(\mathbb{C})$-invariant, and define

$$F : \mathbb{C} \longrightarrow \mathbb{C}, \quad F(\lambda) = \theta(f_\lambda).$$

This is a polynomial function since θ is a polynomial map. We want to show that $\theta(f) = F(\mathrm{disc}(f))$.

In fact we need only prove this for f with $\mathrm{disc}(f)\neq 0$, for if this case is known, then

$$U = \{f \mid \mathrm{disc}(f) = 0\} \cup \{f \mid \theta(f) = F(\mathrm{disc}(f))\}$$

is a union of two Zariski-closed subsets, so by the irreducibility of U we deduce that $U = \{f \mid \theta(f) = F(\mathrm{disc}(f))\}$.

Recall that any matrix is congruent (over $\mathbb{C}$) to a matrix

$$\mathrm{diag}(1,\ldots,1,0,\ldots,0).$$

Thus if $f \in U$ corresponds to matrix A and

$$\lambda = \mathrm{disc}(f) = \det(A) \neq 0,$$

then there is some $B \in GL_m(\mathbb{C})$ with $B^TAB = I$. If now

$$C = \mathrm{diag}(1,\ldots,1,\det(B)^{-1})$$

then $BC \in SL_m(\mathbb{C})$ and

$$(BC)^TABC = C^TB^TABC = \mathrm{diag}(1,\ldots,1,\det(B)^{-2}) = \mathrm{diag}(1,\ldots,1,\lambda)$$

so $\theta(f) = \theta(f_\lambda) = F(\lambda)$ as required.

Example (Discriminant of a binary form)

Let $U = \mathrm{Hom}_{\mathbb{C},n}(\mathbb{C}^2,\mathbb{C})$ be the set of homogeneous polynomials of degree $n\geq 1$ in two variables X_1,X_2. If $f \in U$, say

$$f = a_0X_1^n + a_1X_1^{n-1}X_2 + \ldots + a_nX_2^n = b\,(\lambda_1X_1+\mu_1X_2)\ \ldots\ (\lambda_nX_1+\mu_nX_2)$$

define

$$\mathrm{disc}(f) = b^{2n-2} \prod_{i<j} (\lambda_i\mu_j-\lambda_j\mu_i)^2.$$

This is well-defined since it is unchanged if two terms are exchanged or if one term is enlarged and another reduced by the same factor.

For example, when n=3 one can check that

$$\mathrm{disc}(f) = -27\ a_0^2a_3^2+18\ a_0a_1a_2a_3-4\ a_0a_2^3-4\ a_1^3a_3+a_1^2a_2^2.$$

The map $\mathrm{disc}:U\longrightarrow\mathbb{C}$ is a homogeneous (2n-2)-ic $SL_2(\mathbb{C})$-invariant.

PROOF.

(1) $\mathrm{disc}:U\longrightarrow\mathbb{C}$ is an $SL_2(\mathbb{C})$-invariant. If $g\in GL_2(\mathbb{C})$ and $g^{-1}= \begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$ then

$$(gf)(X_1,X_2) = f(\alpha X_1+\beta X_2,\gamma X_1+\delta X_2) = b\prod_{i=1}^{n} \Big(\lambda_i(\alpha X_1+\beta X_2) + \mu_i(\gamma X_1+\delta X_2)\Big)$$

so

$$\begin{aligned}\mathrm{disc}(gf) &= b^{2n-2} \prod_{i<j}\Big((\lambda_i\alpha+\mu_i\gamma)(\lambda_j\beta+\mu_j\delta) - (\lambda_j\alpha+\mu_j\gamma)(\lambda_i\beta+\mu_i\delta)\Big)^2\\ &= b^{2n-2} \prod_{i<j} (\lambda_i\mu_j - \lambda_j\mu_i)^2(\alpha\delta - \beta\gamma)^2\\ &= \mathrm{disc}(f)\ \det(g)^{-n(n-1)}.\end{aligned}$$

(2) There is a homogeneous polynomial $Q(Z_0,\ldots,Z_n)$ of degree 2n-2 with $\mathrm{disc}(f) = Q(a_0,\ldots,a_n)$ when $a_0 \neq 0$. Let $D(Y_1,\ldots,Y_n)$ be the polynomial with

$D(E_1(X_i),...,E_n(X_i)) = \prod_{i<j} (X_i-X_j)^2$. If $a_0 \neq 0$, then taking $\lambda_1=...=\lambda_n=1$ and $b=a_0$, we see that

$$\mathrm{disc}(f) = a_0^{2n-2} D(a_1/a_0,...,a_n/a_0) = a_0^{-s} Q(a_0,...,a_n)$$

where $Q(Z_0,...,Z_n)$ is a polynomial and we arrange things so that s is non-negative, but otherwise is as small as possible. Dually, if $a_n \neq 0$ then

$$\mathrm{disc}(f) = a_n^{2n-2} D(a_{n-1}/a_n,...,a_0/a_n) = a_n^{-t} Q'(a_0,...,a_n).$$

Since

$$a_0^s Q'(a_0,...a_n) = a_n^t Q(a_0,...,a_n)$$

on the Zariski-dense open subset of U defined by $a_0a_n \neq 0$, we have

$$Z_0^s Q'(Z_0,...,Z_n) = Z_n^t Q(Z_0,...,Z_n).$$

Thus Z_0^s divides Q, and so s = 0 by minimality. Finally observe that

$$Q(\alpha Z_0,...,\alpha Z_n) = (\alpha Z_0)^{2n-2} D(\alpha Z_1/\alpha Z_0,...) = \alpha^{2n-2} Q(Z_0,...,Z_n),$$

so that Q is homogeneous of degree 2n-2.

(3) The map $Q:U \longrightarrow \mathbb{C}$, $f \longmapsto Q(a_0,...,a_n)$ is an $SL_2(\mathbb{C})$-invariant. If $g \in SL_2(\mathbb{C})$ then on the Zariski-dense open subset

$$\{f \in U \mid a_0 \neq 0 \text{ and gf has non-zero coefficient of } X_1^n\}$$

of U we have Q(f) = disc(f) = disc(gf) = Q(gf). Thus Q(f) = Q(gf) on U.

(4) disc(f) = Q(f). If f = 0 this is clear, so suppose that f ≠ 0. There is $g \in SL_2(\mathbb{C})$ such that gf has non-zero coefficient of X_1^n. Then

$$\mathrm{disc}(f) = \mathrm{disc}(gf) = Q(gf) = Q(f).$$

For a $\mathbb{C}SL(V)$-module, as well as the usual invariants, one wants to consider another construction:

Definition. A covariant for a $\mathbb{C}SL(V)$-module U is a polynomial invariant $U \oplus V \longrightarrow \mathbb{C}$.

Examples.

(1) Every invariant θ for U gives a covariant

$$U \oplus V \longrightarrow \mathbb{C}, \ (u,x) \longmapsto \theta(u).$$

(2) If $U = \mathrm{Hom}_{\mathbb{C},n}(V,\mathbb{C})$ then there is a trivial "evaluation" covariant

defined by

$$ev : U \oplus V \longrightarrow \mathbb{C},\ (f,v) \longmapsto f(v).$$

<u>Example</u> (Hessian).

If f is a function of variables $X_1,\dots,X_m$, then the <u>Hessian</u>

$$H(f) = \det\left(\frac{\partial^2 f}{\partial X_i \partial X_j}\right)$$

is another function of $X_1,\dots,X_m$.

Let $U = \mathrm{Hom}_{\mathbb{C},n}(\mathbb{C}^m,\mathbb{C})$ be the set of homogeneous n-ic polynomials in $X_1,\dots,X_m$. The Hessian defines a polynomial map

$$U \oplus \mathbb{C}^m \longrightarrow \mathbb{C},\quad (f,x) \longmapsto H(f)(x)$$

Now U is naturally a $\mathbb{C}GL_m(\mathbb{C})$-module via

$$(gf)(x) = f(g^{-1}x) \quad (f \in U,\ g \in GL_m(\mathbb{C}),\ x \in \mathbb{C}^m).$$

By the chain rule for differentiation

$$H(gf)(gx) = \det(g)^{-2}H(f)(x) \quad (f \in U,\ g \in GL_m(\mathbb{C}),\ x \in \mathbb{C}^m),$$

so that H is an $SL_m(\mathbb{C})$-covariant.

§12. THE FIRST FUNDAMENTAL THEOREM OF INVARIANT THEORY

The First Fundamental Theorem of Invariant Theory (for GL(V) or SL(V)) gives generators for the set of polynomial invariants in the important special case when the module is a direct sum of copies of V and V^*. This is important because in principle one is supposed to be able to use the FFT to compute the invariants for an arbitrary rational module. In fact history has shown that such a transition is not possible, but the idea will be demonstrated with an example in the next section.

Let V be a vector space with basis $e_1,\dots,e_m$.

Theorem (Multilinear First Fundamental Theorem). If $n,r \in \mathbb{N}$ then

(1) $\mathrm{Hom}_{\mathbb{C}GL(V)}(T^n(V^*) \otimes T^rV, \mathbb{C}) = 0$ if $n \neq r$.

(2) $\mathrm{Hom}_{\mathbb{C}GL(V)}(T^n(V^*) \otimes T^nV, \mathbb{C})$ is spanned by the maps μ_σ ($\sigma \in S_n$) defined by

$$\mu_\sigma(\phi_1\otimes\dots\otimes\phi_n\otimes v_1\otimes\dots\otimes v_n) = \phi_{\sigma(1)}(v_1)\dots\phi_{\sigma(n)}(v_n).$$

Proof. If X and Y are $\mathbb{C}G$-modules, then

$$\mathrm{Hom}_{\mathbb{C}}(X^*\otimes Y,\mathbb{C}) = (X^*\otimes Y)^* \cong X^{**}\otimes Y^* \cong X\otimes Y^* \cong \mathrm{Hom}_{\mathbb{C}}(Y,X)$$

and taking the G-fixed points, we obtain

$$\mathrm{Hom}_{\mathbb{C}G}(X^*\otimes Y,\mathbb{C}) \cong \mathrm{Hom}_{\mathbb{C}G}(Y,X)$$

Using this and the isomorphism $T^n(V^*) \cong (T^nV)^*$ we have an isomorphism

$$\pi\colon \mathrm{Hom}_{\mathbb{C}GL(V)}(T^rV,T^nV) \longrightarrow \mathrm{Hom}_{\mathbb{C}GL(V)}(T^n(V^*) \otimes T^rV, \mathbb{C}).$$

Explicitly, this sends a homomorphism f to the map

$$\phi_1\otimes\dots\otimes\phi_n \otimes x \longmapsto (\phi_1\otimes\dots\otimes\phi_n)(f(x)) \quad (\phi_i \in V^* \text{ and } x \in T^rV).$$

(1) $\mathrm{Hom}_{\mathbb{C}GL(V)}(T^rV,T^nV) = 0$, since $\mathrm{Hom}_{\mathbb{C}GL(V)}(S',S) = 0$ if S' (resp. S) is a simple homogeneous r-ic (resp. n-ic) $\mathbb{C}GL(V)$-module.

(2) This is a restatement of Schur-Weyl duality. Recall that $\mathrm{End}_{\mathbb{C}GL(V)}(T^nV)$ is spanned by the maps λ_σ ($\sigma \in S_n$) where

$$\lambda_\sigma(v_1\otimes\dots\otimes v_n) = v_{\sigma^{-1}(1)}\otimes\dots\otimes v_{\sigma^{-1}(n)}.$$

Now π sends λ_σ to μ_σ.

Notation. If U and W are $\mathbb{C}G$-modules, we denote by $\mathrm{Hom}_{\mathbb{C}G,n}(U,W)$ the vector space of all homogeneous n-ic concomitants $U \longrightarrow W$. Note that $\mathrm{Hom}_{\mathbb{C},n}(U,W)$ is a $\mathbb{C}G$-module by conjugation:

$$(gf)(u) = gf(g^{-1}u) \quad (g \in G,\ u \in U,\ f \in \mathrm{Hom}_{\mathbb{C},n}(U,W)),$$

and $\mathrm{Hom}_{\mathbb{C}G,n}(U,W) = \mathrm{Hom}_{\mathbb{C},n}(U,W)^G$.

Lemma 1. If U is a $\mathbb{C}G$-module, then $\mathbb{C}[U]^G \cong \oplus_{n=0}^{\infty} \mathrm{Hom}_{\mathbb{C}G,n}(U,\mathbb{C})$.

Proof. $\mathbb{C}[U] \cong \oplus_{n=0}^{\infty} \mathrm{Hom}_{\mathbb{C},n}(U,\mathbb{C})$. This is an isomorphism of $\mathbb{C}G$-modules. Now take G-fixed points.

Lemma 2. If U and W are $\mathbb{C}G$-modules, then there are inverse isomorphisms of vector spaces

$$\begin{tikzcd}
\mathrm{Hom}_{\mathbb{C}G}(S^nU,W) \arrow[r, shift left, "\psi \longmapsto \psi\circ\Delta"] & \mathrm{Hom}_{\mathbb{C}G,n}(U,W) \arrow[l, shift left, "f \longmapsto \frac{1}{n!}Pf"]
\end{tikzcd}$$

where $\Delta : U \longrightarrow S^nU$, $u \longmapsto u\vee\ldots\vee u$ and Pf is the total polarization of f.

Proof. These maps induce inverse isomorphisms between $\mathrm{Hom}_{\mathbb{C}}(S^nU,W)$ and $\mathrm{Hom}_{\mathbb{C},n}(U,W)$, so it suffices to prove

(a) if ψ is a concomitant, then so is $\psi\circ\Delta$, and

(b) if f is a concomitant, then so is $\frac{1}{n!}Pf$.

Now (a) is obvious, since Δ is a concomitant. There are 3 ways to prove (b).

1st way. Use the formula for Pf. This is very long.

2nd way. If $f \in \mathrm{Hom}_{\mathbb{C}G}(S^nU,W)$, then $f = (\frac{1}{n!}Pf)\circ\Delta$, so

$$\begin{aligned}(\tfrac{1}{n!}Pf)(g(u\vee\ldots\vee u)) &= (\tfrac{1}{n!}Pf)(g\Delta(u)) = (\tfrac{1}{n!}Pf)\Delta(gu) = f(gu)\\ &= gf(u) = g(\tfrac{1}{n!}Pf)(u\vee\ldots\vee u)\end{aligned}$$

for all $u \in U$ and $g \in G$. Now the elements of the form $u\vee\ldots\vee u$ span S^nU.

3rd way. $\mathrm{Hom}_{\mathbb{C}}(S^nU,W)$ and $\mathrm{Hom}_{\mathbb{C},n}(U,W)$ are $\mathbb{C}G$-modules with G acting by conjugation, and $\psi \longmapsto \psi\circ\Delta$ is a $\mathbb{C}G$-module map and an isomorphism. Thus its inverse $f \longmapsto \frac{1}{n!}Pf$ is also a $\mathbb{C}G$-module map. Now these maps restrict to isomorphisms between the sets of G-fixed points.

Remark. The map Δ factors as $U \xrightarrow{\delta} T^nU \xrightarrow{\text{nat}} S^nU$ where $\delta(u)=u\otimes\ldots\otimes u$, so composition with δ induces a surjection

$$\mathrm{Hom}_{\mathbb{C}G}(T^nU,W) \twoheadrightarrow \mathrm{Hom}_{\mathbb{C}G,n}(U,W).$$

It is in this form that we shall use Lemma 2.

Theorem (First Fundamental Theorem for GL_m). If

$$U = \underbrace{V\oplus\ldots\oplus V}_{p} \oplus \underbrace{V^*\oplus\ldots\oplus V^*}_{q},$$

then $\mathbb{C}[U]^{GL(V)}$ is generated as a $\mathbb{C}$-algebra by elements

$$\{\ \rho_{ij}\ |\ 1\le i\le p,\ 1\le j\le q\ \}$$

defined by

$$\rho_{ij}(v_1,v_2,\ldots,v_p,\phi_1,\ldots,\phi_q) = \phi_j(v_i).$$

PROOF. Note first that the ρ_{ij} are polynomial invariants:

$$(g\phi_j)(gv_i) = \phi_j(g^{-1}gv_i) = \phi_j(v_i).$$

By Lemma 1 it suffices to prove that any homogeneous n-ic invariant f is a linear combination of products of ρ_{ij}. By Lemma 2 we have a surjection

$$\mathrm{Hom}_{\mathbb{C}GL(V)}(T^nU,\mathbb{C}) \twoheadrightarrow \mathrm{Hom}_{\mathbb{C}GL(V),n}(U,\mathbb{C}).$$

Let

$$V_1 = V,\ \ldots,\ V_p = V,\ V_{p+1} = V^*,\ \ldots,\ V_{p+q} = V^*,$$

so that

$$U = \oplus_{i=1}^{p+q} V_i.$$

This decomposition of U gives a decomposition of T^nU:

$$T^nU \;=\; \bigoplus_{i_1=1}^{p+q}\ \bigoplus_{i_2=1}^{p+q}\ \ldots \bigoplus_{i_n=1}^{p+q}\ V_{i_1}\otimes\ldots\otimes V_{i_n}$$

so

$$\mathrm{Hom}_{\mathbb{C}GL(V)}(T^nU,\mathbb{C}) = \bigoplus_{i_1,\ldots,i_n} \mathrm{Hom}_{\mathbb{C}GL(V)}(V_{i_1}\otimes\ldots\otimes V_{i_n},\mathbb{C}).$$

We consider one of the summands. By the multilinear version of the FFT, either $\mathrm{Hom}_{\mathbb{C}GL(V)}(V_{i1}\otimes\ldots\otimes V_{in},\mathbb{C}) = 0$ or

$$\begin{cases} n = 2k, \\ k \text{ of the } i_j \text{ are } \le p, \text{ say for } j = \alpha_1,\ldots,\alpha_k; \\ k \text{ of the } i_j \text{ are } >p, \text{ say for } j = \beta_1,\ldots,\beta_k; \end{cases}$$

and in this case $\mathrm{Hom}_{\mathbb{C}GL(V)}(V_{i1}\otimes....\otimes V_{in},\mathbb{C})$ is spanned by the k! maps τ_σ $(\sigma \in S_k)$ given by

$$\tau_\sigma(v_1\otimes...\otimes v_n) \;=\; v_{\beta_{\sigma(1)}}(v_{\alpha_1}) \;...\; v_{\beta_{\sigma(k)}}(v_{\alpha_k}).$$

Tracking back, the homogeneous n-ic invariant of U corresponding to τ_σ is

$$\rho_{i_{\alpha_1},i_{\beta_{\sigma(1)}}-p} \cdots \rho_{i_{\alpha_k},i_{\beta_{\sigma(k)}}-p},$$

and the assertion follows.

<u>Theorem</u>. If $r \in \mathbb{N}$, and

$$U = \mathrm{End}_{\mathbb{C}}(V) \oplus \;...\; \oplus \mathrm{End}_{\mathbb{C}}(V) \quad \text{(r copies)},$$

then $\mathbb{C}[U]^{GL(V)}$ is generated as a $\mathbb{C}$-algebra by the invariants

$$t_{i_1,...,i_k} : U \longrightarrow \mathbb{C},\ (\theta_1,...,\theta_r) \longmapsto \mathrm{Tr}(\theta_{i_1}\theta_{i_2}...\theta_{i_k}),$$

where $k \geq 1$ and $1 \leq i_1,...,i_k \leq r$.

<u>Remark</u>. In fact $\mathbb{C}[U]^{GL(V)}$ is generated by the $t_{i1,...,ik}$ with $k \leq 2^m-1$. See [Procesi, The invariant theory of n×n matrices, Adv. Math. 19 (1976), 306-381].

<u>Remark</u>. Regarding Schur-Weyl duality as hard (since it fails in characteristic p>0), Procesi shows that this theorem is equivalent to the FFT. Thus invariant theory is about representations of quivers: the quivers with one vertex and n loops.

PROOF. By polarization (Lemmas 1 and 2) we need to compute

$$\mathrm{Hom}_{\mathbb{C}GL(V)}(T^n(\mathrm{End}_{\mathbb{C}}(V)),\mathbb{C}).$$

Since $\mathrm{End}_{\mathbb{C}}(V) \cong V^*\otimes V$ this is isomorphic to

$$\mathrm{Hom}_{\mathbb{C}GL(V)}(T^n(V^*)\otimes T^nV,\mathbb{C})$$

which, by the multilinear FFT, is spanned by the maps μ_σ $(\sigma \in S_n)$

$$\mu_\sigma(\phi_1\otimes...\otimes\phi_n\otimes v_1\otimes...\otimes v_n) = \phi_{\sigma(1)}(v_1)...\phi_{\sigma(n)}(v_n).$$

We compute the corresponding map

$$\nu_\sigma : T^n(\mathrm{End}_{\mathbb{C}}(V)) \longrightarrow \mathbb{C}.$$

Let V have basis $e_1,...,e_m$ and V^* dual basis $\eta_1,...,\eta_m$. If $\theta \in \mathrm{End}_{\mathbb{C}}(V)$ has matrix A_{ij} with respect to this basis, then

$$\theta(v) = \textstyle\sum_{ij} A_{ij}\ \eta_j(v)\ e_i$$

so the corresponding element of $V^* \otimes V$ is

$$\textstyle\sum_{ij} A_{ij}\, \eta_j \otimes e_i .$$

Now if θ_k has matrix A^k_{ij}, then $\theta_1 \otimes \ldots \otimes \theta_n$ corresponds to

$$\sum_{\substack{a_1,..,a_n \\ b_1,..,b_n}} A^1_{a_1 b_1} A^2_{a_2 b_2} \ldots\ \eta_{b_1} \otimes \eta_{b_2} \otimes .. \otimes e_{a_1} \otimes e_{a_2} \otimes \ldots \in T^n(V^*) \otimes T^n V$$

so we have

$$\begin{aligned}
\nu_\sigma(\theta_1 \otimes \ldots \otimes \theta_n) &= \sum_{\substack{a_1,..,a_n \\ b_1,..,b_n}} A^1_{a_1 b_1} A^2_{a_2 b_2} \ldots\ \eta_{b_{\sigma(1)}}(e_{a_1})\ \eta_{b_{\sigma(2)}}(e_{a_2}) \ldots \\
&= \sum_{b_1,..,b_n} A^1_{b_{\sigma(1)}, b_1}\ A^2_{b_{\sigma(2)}, b_2} \ldots
\end{aligned}$$

If $\sigma = (i_1 i_2 \ldots i_k)(j_1 j_2 \ldots j_\ell) \ldots$ we can reorder this to get

$$\begin{aligned}
&= \sum_{b_1,..,b_n} A^{i_1}_{b_{i_2}, b_{i_1}}\ A^{i_2}_{b_{i_3}, b_{i_2}} \ldots A^{i_k}_{b_{i_1}, b_{i_k}}\ A^{j_1}_{b_{j_2}, b_{j_1}} \ldots \\
&= \operatorname{Tr}(\theta_{i_k} \ldots \theta_{i_1})\ \operatorname{Tr}(\theta_{j_l} \ldots \theta_{j_1}) \ldots
\end{aligned}$$

The assertion follows.

Before we move on to the FFT for SL_m we need to know a little about how rational $\mathbb{C}GL(V)$-modules behave when regarded as $\mathbb{C}SL(V)$-modules by restriction. First we make a non-standard definition

Definition. If $r \in \mathbb{Z}$ and U is a f.d rational $\mathbb{C}GL(V)$-module, we shall say that U has rational degree r provided that $(\lambda 1_V)u = \lambda^r u$ for all $u \in U$, $\lambda \in \mathbb{C}^\times$, where $\lambda 1_V \in GL(V)$.

Exercises.

(1) If U is homogeneous r-ic, then it has rational degree r.

(2) If U has rational degree r, then U^* has rational degree $-r$.

(3) If U and W have rational degrees r and n, then $U \otimes W$ has rational degree $n+r$.

(4) $D_{\lambda_1,\ldots,\lambda_m}(V)$ has rational degree $\sum_{i=1}^m \lambda_i$.

Lemma 3. If U and W are rational $\mathbb{C}GL(V)$-modules, with rational degrees r and n respectively, then

(1) If $m \nmid r-n$ then $\operatorname{Hom}_{\mathbb{C}SL(V)}(U,W) = 0$

(2) If $r-n = mk$ and $0 \neq d \in \det^{-k}$, then the map

$$\pi : \operatorname{Hom}_{\mathbb{C}GL(V)}(\det^{-k} \otimes U, W) \longrightarrow \operatorname{Hom}_{\mathbb{C}SL(V)}(U,W),\ \pi(f)(u) = f(d \otimes u)$$

is an isomorphism of vector spaces.

Proof. Any element $g \in GL(V)$ can be written as $g = \lambda s$ where λ is an m^{th} root of $\det(g)$ and $s \in SL(V)$. Now

$$(\lambda 1_V)u = \lambda^r u, \quad (\lambda 1_V)w = \lambda^n w \qquad (u \in U,\ w \in W)$$

If $\theta \in \mathrm{Hom}_{\mathbb{C}SL(V)}(U,W)$, then

$$\theta(gu) = \theta((\lambda s)u) = s\,\theta((\lambda 1_V)u) = s\,\theta(\lambda^r u) = s\,\lambda^r\,\theta(u)$$
$$= \lambda^{r-n}\, s(\lambda 1_V)\,\theta(u) = \lambda^{r-n}\, g\,\theta(u).$$

If $\theta(u) \neq 0$ then $m \mid r-n$, otherwise choosing a different m^{th} root of $\det(g)$ would give a different answer. Hence (1).

Now let $r-n = mk$. The map π is defined since $gd = d$ for $g \in SL(V)$, and π is injective, so we only need to prove that it is surjective. Let $\theta \in \mathrm{Hom}_{\mathbb{C}SL(V)}(U,W)$, and define

$$f : \det^{-k} \otimes U \longrightarrow W, \ d\otimes u \longmapsto \theta(u).$$

If f is a $\mathbb{C}GL(V)$-module map, then (2) is proved since $\theta = \pi(f)$. Now

$$\lambda^{r-n} = \lambda^{mk} = \det(\lambda 1_V)^k = (\det g)^k,$$

so

$$f(gd\otimes gu) = f((\det g)^{-k} d\otimes gu) = (\det g)^{-k}\theta(gu) = g\theta(u) = gf(d\otimes u).$$

Exercises.

(1) $D_{\lambda 1,\dots,\lambda m}(V)$ is simple as a $\mathbb{C}SL(V)$-module.

(2) $D_{\lambda 1,\dots,\lambda m}(V) \cong D_{\mu 1,\dots,\mu m}(V)$ as $\mathbb{C}SL(V)$-modules if and only if

$$\lambda_1 - \mu_1 = \lambda_2 - \mu_2 = \dots = \lambda_m - \mu_m.$$

Remark. Although we have not done so, one can develop a theory of rational $\mathbb{C}SL(V)$-modules, and can prove that every rational $\mathbb{C}SL(V)$-module is the restriction of a rational $\mathbb{C}GL(V)$-module. Thus the $D_{\lambda 1,\dots,\lambda m}(V)$ with $\lambda_m=0$ are a complete set of non-isomorphic simple rational $\mathbb{C}SL(V)$-modules.

Theorem (First Fundamental Theorem for SL_m).

Let V be a vector space with basis $e_1,\dots,e_m$, and let V^* have dual basis $\eta_1,\dots,\eta_m$. If

$$U = \underbrace{V\oplus\dots\oplus V}_{p} \oplus \underbrace{V^*\oplus\dots\oplus V^*}_{q},$$

then $\mathbb{C}[U]^{SL(V)}$ is generated as a $\mathbb{C}$-algebra by the polynomial invariants which send $(v_1,v_2,\dots,v_p,\phi_1,\dots,\phi_q) \in U$ to

$$\begin{array}{ll} \phi_j(v_i) & (1\le i\le p,\ 1\le j\le q), \\ [v_{i1},\dots,v_{im}] & (1\le i_1<\dots<i_m\le p), \\ [\phi_{j1},\dots,\phi_{jm}] & (1\le j_1<\dots<j_m\le q), \end{array}$$

where $[v_{i1},\dots,v_{im}]$ is the determinant of the matrix whose n^{th} column is the coordinates of v_{in} with respect to $e_1,\dots,e_m$, and $[\phi_{j1},\dots,\phi_{jm}]$ is the determinant of the matrix whose n^{th} column is the coordinates of ϕ_{jn} with respect to $\eta_1,\dots,\eta_m$.

Proof. Clearly the indicated functions are SL(V)-invariants. Moreover, the restrictions on i_k and j_k can be replaced by

$$1\le i_1,\dots,i_m\le p \quad\text{and}\quad 1\le j_1,\dots,j_m\le q.$$

By polarization (Lemmas 1 and 2) we reduce to having to compute $\mathrm{Hom}_{\mathbb{C}SL(V)}(T,\mathbb{C})$ where

$$T = T^n(V^*) \otimes T^rV.$$

Now T has rational degree r-n, so by Lemma 3, $\mathrm{Hom}_{\mathbb{C}SL(V)}(T,\mathbb{C}) = 0$ if $m \nmid r-n$. Thus we may suppose that $r-n = mk$. If $0 \ne d \in \det^{-k}$, then by Lemma 3 the map

$$\mathrm{Hom}_{\mathbb{C}GL(V)}(\det^{-k}\otimes T,\mathbb{C}) \longrightarrow \mathrm{Hom}_{\mathbb{C}SL(V)}(T,\mathbb{C}),\ f \longmapsto (t \longmapsto f(d\otimes t))$$

is an isomorphism.

We shall consider the case $k\ge 0$, the case $k<0$ is similar.

Identifying $\det^{-1}$ with the summand $T^m(V^*)_{anti}$ of $T^m(V^*)$, we can identify $\det^{-k} \otimes T$ with a summand of

$$T^k[T^m(V^*)] \otimes T \cong T^r(V^*) \otimes T^rV.$$

Thus we have a restriction map

$$\mathrm{res} : \mathrm{Hom}_{\mathbb{C}GL(V)}(T^r(V^*) \otimes T^rV,\ \mathbb{C}) \twoheadrightarrow \mathrm{Hom}_{\mathbb{C}GL(V)}(\det^{-k} \otimes T,\ \mathbb{C})$$

which is surjective since $\det^{-k} \otimes T$ is a summand. By the multilinear FFT the left hand Hom space is spanned by the maps μ_σ ($\sigma \in S_r$) defined by

$$\mu_\sigma(\phi_1\otimes\dots\otimes\phi_r\otimes v_1\otimes\dots\otimes v_r) = \phi_1(v_{\sigma^{-1}(1)})\dots\phi_n(v_{\sigma^{-1}(n)}).$$

Setting

$$\delta = \textstyle\sum_{\tau\in Sm} \varepsilon_\tau\ \eta_{\tau(1)}\otimes\dots\otimes\eta_{\tau(m)} \in T^m(V^*)_{anti} = \det^{-1},$$

$$d = \delta \otimes \dots \otimes \delta \in \det^{-k}$$

and if $t = \phi_1\otimes\dots\otimes\phi_n\otimes v_1\otimes\dots\otimes v_r \in T$, then we have

$$\mathrm{res}(\mu_\sigma)(d\otimes t) = [v_{\sigma^{-1}(1)}\ldots v_{\sigma^{-1}(m)}][v_{\sigma^{-1}(m+1)}\ldots v_{\sigma^{-1}(2m)}]\ldots$$
$$\times\ \phi_1(v_{\sigma^{-1}(km+1)})\ \phi_2(v_{\sigma^{-1}(km+2)})\ldots$$

The corresponding polynomial SL(V)-invariant is thus a product of $\phi_j(v_i)$ and brackets $[v_{i1}\ldots v_{im}]$.

In the similar case when $k<0$ one obtains a product of $\phi_j(v_i)$ and brackets $[\phi_{j1}\ldots\phi_{jm}]$, and the assertion of the theorem follows.

§13. COVARIANTS OF BINARY FORMS

Notation. Throughout this section the following notation will be fixed: $V = \mathbb{C}^2$, $G = SL(V) = SL_2(\mathbb{C})$, and

$$C_n = \mathrm{Hom}_{\mathbb{C},n}(V,\mathbb{C}) \cong (S^nV)^* \cong S^n(V^*),$$

which can be identified with the set of homogeneous polynomials of degree n in variables X_1, X_2.

Remarks.

(1) The C_n $(n\geq 0)$ are non-isomorphic simple $\mathbb{C}G$-modules.

(2) $D_{i,j}(V) \cong C_{i-j}$ as $\mathbb{C}G$-modules, so $C_n \cong S^nV$. In particular $V \cong V^*$.

(3) The Clebsch-Gordan Formula becomes

$$C_p \otimes C_q \cong \bigoplus_{r=0}^{\min(p,q)} C_{p+q-2r}.$$

Fix $n\in\mathbb{N}$. Recall that a covariant for C_n is a polynomial $\mathbb{C}SL(V)$-invariant $C_n\oplus V \longrightarrow \mathbb{C}$. We have already met some of these covariants:

$\mathrm{ev}(f,v) = f(v)$ the evaluation map,
$\mathrm{disc}(f,v) = \mathrm{disc}(f)$ the discriminant,
$H(f,v) = H(f)(v)$ the Hessian.

Our aim is to compute generators for $\mathbb{C}[C_n\oplus V]^G$, or, stated more grandly, to

compute all covariants of binary forms of degree n.

In general this problem is not solved, but it is answered for small n. In this section we prove a useful theorem due to Gordan, and then solve the cases n=3 and n=4.

Example. Classically this problem was tackled with the symbolic method: using polarization to reduce it to the FFT. For example we shall compute $\mathrm{Hom}_{\mathbb{C}G,2}(C_2\oplus V,\mathbb{C})$.

Identifying C_2 with $T^2(V^*)_{\mathrm{symm}}$ we have maps

$$\mathrm{Hom}_{\mathbb{C}G,2}(C_2\oplus V,\mathbb{C}) \twoheadleftarrow \mathrm{Hom}_{\mathbb{C}G,2}(T^2V^*\oplus V,\mathbb{C}) \twoheadleftarrow \mathrm{Hom}_{\mathbb{C}G}(T^2(T^2V^*\oplus V),\mathbb{C})$$

$$\cong \mathrm{Hom}_{\mathbb{C}G}(T^2(T^2V^*) \oplus T^2V^*\otimes V \oplus V\otimes T^2V^* \oplus T^2V,\ \mathbb{C})$$

Now

(1) $\mathrm{Hom}_{\mathbb{C}G}(T^2V,\mathbb{C})$ is spanned by the map $v_1\otimes v_2 \longmapsto [v_1,v_2]$, but the corresponding covariant is $(f,v)\longmapsto[v,v]=0$.

(2) $\mathrm{Hom}_{\mathbb{C}G}(T^2V^*\otimes V,\mathbb{C}) = \mathrm{Hom}_{\mathbb{C}G}(V\otimes T^2V^*,\mathbb{C}) = 0$.

(3) $\mathrm{Hom}_{\mathbb{C}G}(T^2(T^2V^*),\mathbb{C})$ is spanned by the maps which send a tensor $(\phi_1\otimes\phi_2)\otimes(\phi_3\otimes\phi_4)$ to

$$[\phi_1,\phi_2][\phi_3,\phi_4], \qquad [\phi_1,\phi_3][\phi_2,\phi_4], \qquad [\phi_1,\phi_4][\phi_2,\phi_3]$$

The corresponding covariants send (f,v) to

$$0 \qquad\qquad -\tfrac{1}{2}\,\mathrm{disc}(f) \qquad\qquad \tfrac{1}{2}\,\mathrm{disc}(f)$$

Thus $\mathrm{Hom}_{\mathbb{C}G,2}(C_2\oplus V,\mathbb{C})$ is spanned by the discriminant. More generally one can show that $\mathbb{C}[C_2\oplus V]^G$ is generated as a $\mathbb{C}$-algebra by ev and disc.

We shall not use the symbolic method, however, since we have not found it necessary.

Definition. If f and g are functions of two variables X_1,X_2 and $r\in\mathbb{N}$, then the r-th transvectant (Überschiebung) of f and g, denoted by $\tau_r(f,g)$, is defined by

$$\tau_r(f,g) = \sum_{i=0}^{r} \frac{(-1)^i}{i!(r-i)!}\,\frac{\partial^r f}{\partial X_1^{r-i}\partial X_2^{i}}\,\frac{\partial^r g}{\partial X_1^{i}\partial X_2^{r-i}}$$

$$= \frac{1}{r!}\left(\frac{\partial}{\partial X_1}\frac{\partial}{\partial Y_2} - \frac{\partial}{\partial X_2}\frac{\partial}{\partial Y_1}\right)^r \left(f(X_1,X_2)\,g(Y_1,Y_2)\right)\Bigg|_{Y_1=X_1,Y_2=X_2}$$

Examples.

(0) $\tau_0(f,g) = fg$;

(1) $\tau_1(f,g) = \partial(f,g)/\partial(X_1,X_2)$ is the Jacobian of f and g;

(2) $\tau_2(f,f) = H(f)$ is the Hessian of f.

(r) $\tau_r(f,g) = (-1)^r\,\tau_r(g,f)$, so $\tau_r(f,f) = 0$ if r is odd.

Remark. If f and g are homogeneous polynomials of degrees p, q, then $\tau_r(f,g) = 0$ unless $r \le \min(p,q)$, in which case it is a homogeneous polynomial of degree $p+q-2r$. The normalization used for $\tau_r(f,g)$ is my own; usually the definition is

$$(f,g)^r = \frac{r!(p-r)!(q-r)!}{p!q!} \tau_r(f,g).$$

Our first lemma shows how transvectants are related to the Clebsch-Gordan formula.

Lemma 1.

(1) If $r \le \min(p,q)$ then the map

$$\tau_r : C_p \otimes C_q \longrightarrow C_{p+q-2r}, \quad f\otimes g \longmapsto \tau_r(f,g)$$

is non-zero map of $\mathbb{C}G$-modules.

(2) Any $\mathbb{C}G$-module map $C_p \otimes C_q \longrightarrow C_{p+q-2r}$ is a multiple of τ_r.

(3) The τ_r combine to give an isomorphism of $\mathbb{C}G$-modules.

$$C_p \otimes C_q \xrightarrow{(\tau_r)} \oplus_{r=0}^{\min(p,q)} C_{p+q-2r}.$$

Proof.

(1) The map is certainly a vector space map, so we need to show that it commutes with the action of $s\in G$. Let $s^{-1} = \begin{pmatrix}\alpha & \beta \\ \gamma & \delta\end{pmatrix}$ and $x\in\mathbb{C}^2$. Introduce new variables

$$X_1' = \alpha X_1+\beta X_2, \; X_2' = \gamma X_1+\delta X_2,$$

ie $(X_1',X_2') = s^{-1}(X_1,X_2)$. Thus $(sf)(X_1,X_2) = f(X_1',X_2')$. Now

$$\partial/\partial X_1 = \partial X_1'/\partial X_1.\partial/\partial X_1' + \partial X_2'/\partial X_1.\partial/\partial X_2' = \alpha.\partial/\partial X_1' + \gamma.\partial/\partial X_2'$$
$$\partial/\partial X_2 = \beta.\partial/\partial X_1' + \delta.\partial/\partial X_2'$$

Also, introducing $Y_1' = \alpha Y_1+\beta Y_2$, $Y_2' = \gamma Y_1+\delta Y_2$, we have similar formulae for the $\partial/\partial Y_i$ and

$$\frac{\partial}{\partial X_1}\frac{\partial}{\partial Y_2} - \frac{\partial}{\partial X_2}\frac{\partial}{\partial Y_1} = (\alpha\delta-\beta\gamma)\left(\frac{\partial}{\partial X_1'}\frac{\partial}{\partial Y_2'} - \frac{\partial}{\partial X_2'}\frac{\partial}{\partial Y_1'}\right),$$

and $\alpha\delta-\beta\gamma = 1$. Thus $\tau_r(sf,sg)(x)$ is equal to

$$\frac{1}{r!}\left(\frac{\partial}{\partial X_1}\frac{\partial}{\partial Y_2} - \frac{\partial}{\partial X_2}\frac{\partial}{\partial Y_1}\right)^r \left(f(X_1',X_2')\; g(Y_1',Y_2')\right) \Bigg|_{(X_1,X_2)=(Y_1,Y_2)=x}$$

which is the same as

$$\frac{1}{r!}\left(\frac{\partial}{\partial X_1'}\frac{\partial}{\partial Y_2'} - \frac{\partial}{\partial X_2'}\frac{\partial}{\partial Y_1'}\right)^r \left(f(X_1',X_2')\; g(Y_1',Y_2')\right) \Bigg|_{(X_1',X_2')=(Y_1',Y_2')=s^{-1}x}$$

and hence equal to $\tau_r(f,g)(s^{-1}x) = (s\tau_r(f,g))(x)$, as required.

The map τ_r is non-zero since if

$$f(X_1,X_2) = X_1^p \quad\text{and}\quad g(X_1,X_2) = X_2^q$$

then

$$\partial^r f/\partial X_1{}^r = p(p-1)..(p-r+1)\ X_1{}^{p-r} \quad\text{and}$$
$$\partial^r g/\partial X_2{}^r = q(q-1)..(q-r+1)\ X_2{}^{q-r},$$

so that

$$\tau_r(f,g) = 1/r!\ p(p-1)..(p-r+1)q(q-1)..(q-r+1)\ X_1^{p-r}\ X_2^{q-r} \neq 0.$$

(2) By Clebsch-Gordan,

$$C_p \otimes C_q \cong \oplus_{r=0}^{\min(p,q)}\ C_{p+q-2r}.$$

Since the summands on the RHS are simple and non-isomorphic,

$$\dim_{\mathbb{C}} \operatorname{Hom}_{\mathbb{C}G}(C_p \otimes C_q,\ C_{p+q-2r}) = 1,$$

so any map is a multiple of τ_r.

(3) Combine (2) with Clebsch-Gordan.

We shall need the following technical lemma later.

Lemma 2. If $p,q,n,r \in \mathbb{N}$ and $r \leq \min(q,n)$ then setting $N=\max(0,r-p)$, there are a_N, .., $a_r \in \mathbb{C}$, with $a_r \neq 0$, such that

$$f\ \tau_r(g,h) = \textstyle\sum_{k=N}^{r} a_k\ \tau_k(\tau_{r-k}(f,g),h)$$

for all $f \in C_p$, $g \in C_q$, $h \in C_n$.

Remark. Realizing that the left hand side can be rewritten as $\tau_0(f,\tau_r(g,h))$ this lemma expresses a sort of associativity for the τ_i. In fact there are many formulae of this nature; see the chapter on Gordan's Series in [Grace and Young]. In particular one may also find expressions for the coefficients a_k there.

PROOF. We have an isomorphism of $\mathbb{C}G$-modules

$$C_p \otimes C_q \otimes C_n \xrightarrow{(\tau_i\otimes 1)} \bigoplus_{i=0}^{\min(p,q)} C_{p+q-2i} \otimes C_n$$
$$\xrightarrow{(\tau_k)} \bigoplus_{i=0}^{\min(p,q)} \bigoplus_{k=0}^{\min(n,p+q-2i)} C_{p+q+n-2i-2k}$$

By Schur's Lemma, any $\mathbb{C}G$-module map

$$C_p \otimes C_q \otimes C_n \longrightarrow C_{p+q+n-2r}$$

is a linear combination of the maps $\tau_k \circ (\tau_i \otimes 1)$ with

$$i+k = r,\ 0 \le i \le \min(p,q) \text{ and } 0 \le k \le \min(n,p+q-2i).$$

With our assumptions this condition is just

$$N \le k \le r \text{ and } i = r-k.$$

In particular this holds for the map

$$C_p \otimes C_q \otimes C_n \longrightarrow C_{p+q+n-2r}, \quad f\otimes g\otimes h \longmapsto f\tau_r(g,h),$$

so there are $a_k \in \mathbb{C}$ as required.

Now suppose that $a_r = 0$. Set

$$f = X_1^p \qquad g = X_1^q \qquad h = X_2^r.$$

Then

$$\tau_i(f,g) = \begin{cases} X_1^{p+q} & (i=0) \\ 0 & (\text{else}) \end{cases}$$

so the right hand side is zero. On the other hand

$$\tau_r(g,h) = \frac{1}{r!}\,\frac{\partial^r g}{\partial X_1{}^r}\,\frac{\partial^r h}{\partial X_2{}^r} = \frac{1}{r!}q(q-1)\ldots(q-r+1)X_1^{q-r}n(n-1)\ldots(n-q+1)X_2^{n-r}$$

so $f\tau_r(g,h) \ne 0$, a contradiction.

Notation. Fix $n\in\mathbb{N}$.

(1) Let $R = \mathbb{C}[C_n\oplus V]$. Thus R^G is the set of covariants of C_n.

(2) If $\phi \in R$ and $f \in C_n$, define $\phi(f) \in \mathbb{C}[V]$ by $\phi(f)(x) = \phi(f,x)$.

(3) If $r\in\mathbb{N}$ and $\phi,\psi \in R$, then $\tau_r(\phi,\psi)$ denotes the map

$$C_n\oplus V\longrightarrow\mathbb{C}, \qquad (f,x)\longmapsto\tau_r(\phi(f),\psi(f))(x).$$

(4) If $d,i\in\mathbb{N}$ let R_{di} be the set of $\phi \in R$ which are homogeneous, of degree d in C_n and degree i in V. Thus the elements of R_{d0} are invariants and $\mathrm{ev} \in R_{1n}$.

Lemma 3. If $r\in\mathbb{N}$ and $\phi,\psi : C_n\oplus V\longrightarrow\mathbb{C}$ are covariants, then so is

$$\tau_r(\phi,\psi) : C_n \oplus V \longrightarrow \mathbb{C}, \quad (f,x) \longmapsto \tau_r(\phi(f),\psi(g))(x).$$

Proof. Follows from Lemma 1.

Lemma 4.

(1) The R_{di} are $\mathbb{C}G$-submodules of R, and $R = \bigoplus_{d,i=0}^{\infty} R_{di}$, so $R^G = \bigoplus_{d,i=0}^{\infty} R_{di}^G$.

(2) $R_{di}R_{ej} \subseteq R_{d+e,i+j}$ and $\tau_r(R_{di},R_{ej}) \subseteq R_{d+e,i+j-2r}$ for $r\le\min(i,j)$.

(3) The assignment $\phi \longmapsto (f\longmapsto\phi(f))$ induces an isomorphism

$$R^G_{di} \cong \mathrm{Hom}_{\mathbb{C}G,d}(C_n, C_i).$$

In particular

$$R^G_{0i} = \begin{cases} \mathbb{C}.1_R & (i=0) \\ 0 & (i\neq 0) \end{cases} \quad \text{and} \quad R^G_{1i} = \begin{cases} \mathbb{C}.\mathrm{ev} & (i=n) \\ 0 & (i\neq n) \end{cases}.$$

PROOF.

(1),(2) Clear.

(3) The assignment $\phi \longmapsto (f\longmapsto\phi(f))$ induces an isomorphism $R_{di}\longrightarrow\mathrm{Hom}_{\mathbb{C},d}(C_n,C_i)$ of $\mathbb{C}G$-modules, and the assertion follows after taking fixed points. Now

$$R^G_{0i} \cong C^G_i \cong \mathrm{Hom}_{\mathbb{C}G}(\mathbb{C},C_i) = \mathrm{Hom}_{\mathbb{C}G}(C_0,C_i) \quad \text{and} \quad R^G_{1i} \cong \mathrm{Hom}_{\mathbb{C}G}(C_n,C_i)$$

and the dimensions of these are known since the C_i are non-isomorphic simple $\mathbb{C}G$-modules.

<u>Lemma</u> 5. Any covariant $\theta\in R^G_{di}$ $(d\geq 1)$ can be expressed as a linear combination

$$\theta = \sum_{r=0}^{\min(n,i)} \tau_{n-r}(\phi_r,\mathrm{ev})$$

with $\phi_r \in R^G_{d-1,n+i-2r}$.

PROOF. Using the correspondence in Lemma 4, we need to prove the surjectivity of the map

$$\alpha : \bigoplus_{r=0}^{\min(n,i)} \mathrm{Hom}_{\mathbb{C}G,d-1}(C_n,C_{n+i-2r}) \longrightarrow \mathrm{Hom}_{\mathbb{C}G,d}(C_n,C_i)$$

which sends $\phi \in \mathrm{Hom}_{\mathbb{C}G,d-1}(C_n,C_{n+i-2r})$ to the map

$$f \longmapsto \tau_{n-r}(\phi(f),f).$$

If $\delta_{d-1} : C_n\longrightarrow T^{d-1}C_n$, $f \longmapsto f\otimes\ldots\otimes f$ denotes the diagonal map, then composition with δ_{d-1} gives a homomorphism

$$\beta : \bigoplus_{r=0}^{\min(n,i)} \mathrm{Hom}_{\mathbb{C}G}(T^{d-1}C_n,C_{n+i-2r}) \longrightarrow \bigoplus_{r=0}^{\min(n,i)} \mathrm{Hom}_{\mathbb{C}G,d-1}(C_n,C_{n+i-2r})$$

and it suffices to prove that $\alpha\circ\beta$ is surjective.

For $r\leq i$ we have non-zero $\mathbb{C}G$-module maps

$$C_{n+i-2r} \longrightarrow \mathrm{Hom}_{\mathbb{C}}(C_n,C_i),\ f \longmapsto (g\longmapsto\tau_{n-r}(f,g))$$

and since by Clebsch-Gordan we have an isomorphism

$$\mathrm{Hom}_{\mathbb{C}}(C_n,C_i) \cong C_n^*\otimes C_i \cong C_n\otimes C_i \cong \bigoplus_{r=0}^{\min(n,i)} C_{n+i-2r}.$$

it follows that these maps combine to give an isomorphism

$$\bigoplus_{r=0}^{\min(n,i)} C_{n+i-2r} \longrightarrow \mathrm{Hom}_{\mathbb{C}}(C_n,C_i).$$

Applying $\mathrm{Hom}_{\mathbb{C}}(T^{d-1}C_n,-)$ to this and using the isomorphism

$$\mathrm{Hom}_{\mathbb{C}}(T^{d-1}C_n,\ \mathrm{Hom}(C_n,C_i)) \cong \mathrm{Hom}_{\mathbb{C}}(T^dC_n,\ C_i)$$

gives an isomorphism

$$\bigoplus_{r=0}^{\min(n,i)} \mathrm{Hom}_{\mathbb{C}}(T^{d-1}C_n,C_{n+i-2r}) \longrightarrow \mathrm{Hom}_{\mathbb{C}}(T^dC_n,C_i).$$

which sends $\psi \in \mathrm{Hom}_{\mathbb{C}}(T^{d-1}C_n,C_{n+i-2r})$ to the map

$$f_1\otimes\ldots\otimes f_d \longmapsto \tau_{n-r}(\psi(f_1\otimes\ldots\otimes f_{d-1}),f_d).$$

Taking G-fixed points now gives an isomorphism

$$\gamma : \bigoplus_{r=0}^{\min(n,i)} \mathrm{Hom}_{\mathbb{C}G}(T^{d-1}C_n,C_{n+i-2r}) \longrightarrow \mathrm{Hom}_{\mathbb{C}G}(T^dC_n,C_i).$$

By polarization, composition with δ_d induces a surjection

$$\zeta : \mathrm{Hom}_{\mathbb{C}G}(T^dC_n,C_i) \twoheadrightarrow \mathrm{Hom}_{\mathbb{C}G,d}(C_n,C_i).$$

Now $\alpha\circ\beta = \zeta\circ\gamma$ is surjective, as required.

Theorem (Gordan. A weak form of Gordan’s famous theorem).

If T is a $\mathbb{C}$-subalgebra of R^G with the property that $\tau_r(\phi,ev) \in T$ whenever $r \in \mathbb{N}$ and $\phi \in T$, then $T = R^G$.

Proof. By definition $1_R \in T$, so each $R^G_{0i} \subseteq T$. If $R^G_{d-1,i} \subseteq T$ for all i, then by Lemma 5, $R^G_{di} \subseteq T$ for all i. Thus $R^G \subseteq T$.

Remark. We shall use this to find a set of generators of R^G in case n=3 and n=4, but we need to be more precise, and we need a preliminary lemma.

Lemma 6. If $\phi \in R^G_{dp}$, $\psi \in R^G_{eq}$ and $r \leq \min(q,n)$, then

$$\tau_r(\phi\psi,ev) = \alpha\phi\tau_r(\psi,ev) + \textstyle\sum_{k=N}^{r-1} \tau_k(\theta_k,ev)$$

where $N = \max(0,r-p)$, for some $\alpha \in \mathbb{C}$ and $\theta_k \in R^G_{d+e,p+q+2k-2r}$.

PROOF. By Lemma 2, $\phi\tau_r(\psi,ev)$ is a linear combination of the covariants $\tau_k(\tau_{r-k}(\phi,\psi),ev)$ with $N \le k \le r$, and the coefficient of the term with $k=r$ is non-zero. This term is $\tau_r(\phi\psi,ev)$, so we can turn the equation around, and write

$$\tau_r(\phi\psi,ev) = \alpha\phi\tau_r(\psi,ev) + \textstyle\sum_{k=N}^{r-1} \alpha_k\tau_k(\tau_{r-k}(\phi,\psi),ev).$$

Setting

$$\theta_k = \alpha_k\tau_{r-k}(\phi,\psi) \in R_{d+e,p+q+2k-2r'}$$

the assertion follows.

Example (Covariants of cubic forms). If $n=3$, then R^G is generated by the covariants

$ev \in R^G_{13}$

$H = \tau_2(ev,ev)$, the Hessian, in R^G_{22}.

$t = \tau_1(H,ev) \in R^G_{33}$

$D = \tau_3(t,ev)$, the discriminant ($\times$48), in R^G_{40}.

PROOF. Let T be the $\mathbb{C}$-subalgebra of R^G generated by ev, H, t and D. We must show that $T = R^G$. As before, $R^G_{0i} \subseteq T$ for all i. Suppose by induction that

$$R^G_{d',i} \subseteq T \text{ for all } i \text{ and all } d'<d.$$

We have to show that $R^G_{di} \subseteq T$ for all i. By Lemma 5, it suffices to prove for all $0 \le r \le n$ the property

$$(P_r) \qquad \tau_r(\phi,ev) \in T \text{ for all } \phi \in R^G_{d-1,j} \text{ with } j \ge r.$$

(P_0) is trivial: by the induction on d we have $\phi \in T$. Thus

$$\tau_0(\phi,ev) = \phi.ev \in T.$$

Now suppose that $0 < r \le n$ and that $(P_{r'})$ is true for all $r'<r$. We have to prove (P_r). Again $\phi \in T$, so ϕ is a linear combination of monomials

$$ev^x\ H^y\ t^z\ D^w \in R^G_{x+2y+3z+4w,3x+2y+3w'}$$

and it suffices to prove (P_r) when ϕ is a monomial. There are two cases.

Case 1. If the monomial decomposes as a product

$$\phi = \phi_1\phi_2 \text{ where } \phi_2 \in R^G_{eq} \text{ with } q \ge r \text{ and } e < d-1.$$

By Lemma 6 we have

$$\tau_r(\phi,ev) = \alpha\, \phi_1\tau_r(\phi_2,ev) + \textstyle\sum_{k=N}^{r-1} \tau_k(\theta_k,ev) \quad (\theta_k \in R^G_{d-1,j+2k-2r}).$$

Now $\tau_k(\theta_k,ev) \in T$ by property (P_k), and $\tau_r(\phi_2,ev) \in R^G_{e+1,q+n-2r} \subseteq T$ by the induction on d. Thus $\tau_r(\phi,ev) \in T$.

Case 2. If ϕ does not decompose, there are only the following cases

(a) ϕ = ev, r = 1,2,3.

(b) ϕ = H, r = 1,2.

(c) ϕ = t, r = 1,2,3.

(d) $\phi = H^2$, r = 3.

Namely, suppose that ϕ is not one of these. Since $r \le n=3$, if ev, t or H^2 occurs in ϕ this factor can be removed. Thus $\phi = D^w$ or $\phi = H.D^w$. In the first case r=0, but this has been dealt with; in the second case $r \le 2$, so this decomposes unless w=0.

Now

(a1) $\tau_1(ev,ev) = 0$ since 1 is odd.

(a2) $\tau_2(ev,ev) = H \in T$ by definition.

(a3) $\tau_3(ev,ev) = 0$ since 3 is odd.

(b1) $\tau_1(H,ev) = t \in T$ by definition.

(b2) $\tau_2(H,ev) = 0$ by calculation.

(c1) $\tau_1(t,ev) = -3/2\, H^2 \in T$ by calculation.

(c2) $\tau_2(t,ev) = 0$ by calculation.

(c3) $\tau_3(t,ev) = D \in T$ by definition.

(d3) $\tau_3(H^2,ev) = 0$ by calculation.

The calculations are tedious, but not "difficult". For example, if

$$f = a_0X_1^3 + a_1X_1^2X_2 + a_2X_1X_2^2 + a_3X_2^3,$$

then

$$H(f) = \begin{vmatrix} 6a_0X_1+2a_1X_2 & 2a_1X_1+2a_2X_2 \\ 2a_1X_1+2a_2X_2 & 6a_3X_2+2a_2X_1 \end{vmatrix}$$

$$= (12a_0a_2-4a_1^2)X_1^2 + (36a_0a_3-4a_1a_2)X_1X_2 + (12a_1a_3-4a_2^2)X_2^2.$$

so

$$\tau_2(H,ev)(f) = \tau_2(H(f),f)$$

$$= \frac{1}{2}\frac{\partial^2 H}{\partial X_1^2}\frac{\partial^2 f}{\partial X_2^2} - \frac{\partial^2 H}{\partial X_1 \partial X_2}\frac{\partial^2 f}{\partial X_1 \partial X_2} + \frac{1}{2}\frac{\partial^2 H}{\partial X_2^2}\frac{\partial^2 f}{\partial X_1^2}$$

$$= (12a_0a_2-4a_1)(2a_2^2X_1+6a_3X_2) - (36a_0a_3-4a_1a_2)(2a_1X_1+2a_2X_2)$$

$$+ (12a_1a_3-4a_2^2)(6a_0X_1+2a_1X_2)$$

$$= 0.$$

Remark. Associated with the symbolic method there is a symbolic notation which makes the calculations easier, but still non-trivial. See [Grace and Young], or indeed any old text on invariant theory.

Exercise (Covariants of quartic forms). Take n=4, and consider the covariants

$$ev,\ H = \tau_2(ev,ev),\ i = \tau_4(ev,ev),\ t = \tau_1(H,ev),\ j = \tau_4(H,ev).$$

Which R_{di} do they lie in? Show that they generate R^G, using the calculations

$\tau_2(H,ev)$ is a multiple of i.ev.

$\tau_3(H,ev) = 0$.

$\tau_1(t,ev)$ is a linear combination of H^2 and $i.ev^2$.

$\tau_2(t,ev) = 0$.

$\tau_3(t,ev)$ is a linear combination of i.H and j.ev.

$\tau_4(t,ev) = 0$.

Remark. In Sylvester’s Collected Works one can find tables of details about higher degree forms. For example one has

Degree of binary form	0	1	2	3	4	5	6	7	8	9	10
Number of generators needed for covariants	0	1	2	4	5	23	26	124	69	415	475