Haiman's course UCB

M. Haiman Seminar on combinatorics and algebraic geometry Spring 2000

Notes I: Symmetric functions

1.1. Fundamental definitions.

We denote by Λ_R the algebra of symmetric functions in infinitely many variables $\mathbf{z} = z_1, z_2, \ldots$, with coefficients in a ring R (typically $R = \mathbb{Z}$ or $\mathbb{Q}$).

In this set-up "symmetric functions" are not actually functions, but formal power series in the variables $\mathbf{z}$. By definition a formal power series f is a function assigning to each (finite) monomial in the $\mathbf{z}$'s a *coefficient* in R. We denote the coefficient of a monomial $\mathbf{z}^\nu$ in f by $f|_{\mathbf{z}^\nu}$, and write f formally as the infinite sum

$$f = \sum c_\nu \mathbf{z}^\nu, \tag{1}$$

where $c_\nu = f|_{\mathbf{z}^\nu}$.

The sum of two formal series is given by coefficient-wise addition. The product is given by the usual rule for multiplication of series, which is well-defined since there are only a finite number of contributions to a given term in the product.

More generally, if we have series $\{f_i\}$ such that each monomial $\mathbf{z}^\nu$ occurs with non-zero coefficient in only finitely many of them, the infinite sum $\sum_i f_i$ is well-defined. If, in addition, all but finitely many of the f_i have constant term 1, the infinite product $\prod_i f_i$ is well-defined.

Note that the formal sum expression for f in (1) is a valid infinite sum.

A formal series f as above is *symmetric* if similar monomials have equal coefficient in f. Monomials $\mathbf{z}^\nu$ and $\mathbf{z}^\mu$ are *similar* if they may be obtained from one another by permuting the variables. that is, if they have the same multiset of exponents.

Abusing terminology, we call f a symmetric *polynomial* if only a finite number of similarity classes of monomials have non-zero coefficient in f. Otherwise we will specifically refer to f as a symmetric *infinite series*. We say f is *homogeneous of degree* n if only terms of degree n occur with non-zero coefficient in f. Clearly a symmetric infinite series is a polynomial in our sense if and only if its (non-zero) terms have bounded degree.

One may regard the symmetric functions as series that are invariant for the action of an infinite symmetric group S_∞ permuting the variables. By S_∞ we might mean the group of all (infinite) permutations of the variables, or its subgroup of permutations fixing all but finitely many variables. Which group we choose makes no difference as far as symmetric functions are concerned.

To recover the classical symmetric polynomials in finitely many variables $z_1, \ldots, z_l$ (which are honest polynomials), we may specialize the remaining variables z_i to zero. The specialization of f is simply obtained by deleting all terms of f involving any of the variables we are zeroing.

1.2. Partitions.

A *partition of* n is a multiset of positive integers, or *parts*, whose sum is n. By convention we always write the parts in decreasing order:

$$\lambda = (\lambda_1 \geq \ldots \geq \lambda_l).$$

It is often convenient to admit into the notation some "parts" of zero at the end, so that a sequence as above of length l can stand for any partition with l or fewer parts.

The *diagram* of a partition is an array of *cells* in $\mathbb{N} \times \mathbb{N}$, with λ_i cells in the i-th row, as for example

$$\lambda = (4, 2, 2); \quad D(\lambda) = \begin{array}{llll} \bullet & \bullet & & \\ \bullet & \bullet & & \\ \bullet & \bullet & \bullet & \bullet \end{array} .$$

The *conjugate* partition λ' is defined by setting λ'_i equal to the number of parts $\geq i$ in λ, so that the diagram of λ' is the transpose of the diagram of λ.

The partitions of n are partially ordered by *dominance*, defined by

$$\lambda \leq \mu \quad \text{iff} \quad \lambda_1 + \cdots + \lambda_k \leq \mu_1 + \cdots + \mu_k \quad \text{for all } k.$$

For $i < j$ let $R_{ij}\lambda = (\lambda_1, \ldots, \lambda_i + 1, \ldots, \lambda_j - 1, \ldots, \lambda_l)$. This sequence may or may not be a partition, but in any case it clearly satisfies the defining condition for $R_{ij}\lambda > \lambda$ in dominance order. When $R_{ij}\lambda$ is a partition, its diagram differs from that of λ by the transfer of a cell from row j down to row i. In class we proved the following characterization of dominance.

Proposition: The dominance order on partitions of n is the transitive closure of the relation $\lambda \to \nu$ if $\nu = R_{ij}\lambda$ for some $i < j$.

Corollary: We have $\lambda \leq \mu$ if and only if $\lambda' \geq \mu'$.

1.3. The monomial basis.

Every similarity class of monomials $\mathbf{z}^\nu$ contains a unique monomial $\mathbf{z}^\lambda$ in which the exponent sequence λ is a partition [of $n =$ (degree of $\mathbf{z}^\lambda$)]. The *monomial symmetric functions* are defined by

$$m_\lambda = \sum \mathbf{z}^\nu, \quad \text{where } \mathbf{z}^\nu \text{ ranges over monomials similar to } \mathbf{z}^\lambda.$$

From the definition it is clear that every symmetric polynomial f is a unique R-linear combination of monomial symmetric functions, which we may also state as follows.

Proposition: The monomial symmetric functions m_λ form a basis of Λ_R as a free R-module. More precisely, $\{m_\lambda : |\lambda| = n\}$ is an R-basis of the submodule $\Lambda_R^{(n)}$ of symmetric functions homogeneous of degree n. In particular $\Lambda_R^{(n)}$ is free of rank equal to the number of partitions of n.

In finitely many variables, it is easy to see that we have the following variant, where $l(\lambda)$ denotes the number of parts of λ.

Proposition: The monomial symmetric functions $\{m_\lambda : l(\lambda) \leq l\}$ form an R-basis of $\Lambda_R(z_1, \ldots, z_l)$, while all m_λ with $l(\lambda) > l$ specialize to zero.

1.4. Elementary symmetric functions.

We define the k-th *elementary symmetric function* to be

$$e_k = m_{(1^k)}.$$

This is the sum of all products of k distinct variables z_i.

For a partition λ we set

$$e_\lambda = e_{\lambda_1} \cdots e_{\lambda_k}.$$

Proposition: We have

$$e_\lambda = m_{\lambda'} + \sum_{\nu<\lambda'} c_{\lambda\nu} m_\nu$$

for some integers $c_{\lambda\nu}$.

Proof: The coefficient of m_ν in the expansion by monomials of e_λ is the same as the coefficient $e_\lambda|_{\mathbf{z}^\nu}$ (this principle holds for expansion of any symmetric function by monomials). This is the number of ways to realize $\mathbf{z}^\nu$ as a product of products $\mathbf{z}_{S_i}$, where for a set of indices $S = \{s_1, \ldots, s_k\}$, $\mathbf{z}_S$ denotes $z_{s_1} \cdots z_{s_k}$, and we require $|S_i| = \lambda_i$. Representing the sets S_i by rows of a zero-one matrix, we see that $c_{\lambda\nu}$ is the number of such matrices M with row-sums λ and column-sums ν.

Now for M having fixed row-sums, denoting the column-sums by $\tilde{\nu}_i$, we can simultaneously maximize $\tilde{\nu}_1 + \cdots + \tilde{\nu}_k$ for all k by the "rolling-ball" process: tilt the matrix to the left and roll the ones into the left-most positions in each row. This done, we will have $\tilde{\nu} = \lambda'$. Hence if $c_{\lambda\nu} \neq 0$ we have $\nu \leq \lambda'$, with equality for a unique matrix M, so that $c_{\lambda\lambda'} = 1$.

Corollary: The elementary symmetric functions e_λ form an R-basis of Λ_R. Equivalently, Λ_R is isomorphic to the polynomial ring $R[e_1, e_2, \ldots]$, *i.e.*, the elementary symmetric functions e_k are algebraically independent.

Since λ' has at most l parts iff the parts of λ are at most l, we have the following variant in l variables, known as the fundamental theorem of symmetric functions.

Corollary: $\Lambda_R(z_1, \ldots, z_l)$ is isomorphic to $R[e_1, \ldots, e_l]$.

1.5. Complete homogeneous symmetric functions.

We define the k-th *complete homogeneous symmetric function* to be

$$h_k = \sum_{|\lambda|=k} m_\lambda.$$

This is the sum of all monomials of degree k in the variables $\mathbf{z}$. As with the elementary symmetric functions we define

$$h_\lambda = h_{\lambda_1} \cdots h_{\lambda_k}.$$

We will show presently that these form a basis of Λ_R; this result is not as easy as for the elementary symmetric functions. For now we note that by reasoning similar to that for e_λ we can obtain the following.

Proposition: Let $h_\lambda = \sum_\nu c_{\lambda\nu} m_\nu$. Then $c_{\lambda\nu}$ is the number of non-negative integer matrices with row sums λ and column-sums ν. In particular, $c_{\lambda\nu} = c_{\nu\lambda}$.

1.6. Schur functions.

For the moment, we work with a fixed finite number of variables $z_1, \ldots, z_l$, and take R to be $\mathbb{Z}$. Let α be a partition with *distinct parts*

$$\alpha = (\alpha_1 > \cdots > \alpha_l).$$

(Here we allow $\alpha_l = 0$ so α could have l or $l-1$ parts). The smallest such partition is

$$\delta = (l-1, l-2, \ldots, 1, 0),$$

and we may clearly put

$$\alpha = \lambda + \delta,$$

where λ is an arbitrary partition with at most l parts.

We define

$$a_\alpha(z_1, \dots, z_l) = \det \begin{bmatrix} z_1^{\alpha_1} & \dots & z_1^{\alpha_l} \\ \vdots & & \vdots \\ z_l^{\alpha_1} & \dots & z_l^{\alpha_l} \end{bmatrix} = \sum_{w \in S_l} \epsilon(w) w(\mathbf{z}^\alpha),$$

where $\epsilon(w)$ denotes the sign of the permutation w. We require α to have distinct parts so that a_α will not be zero. Note that a_α is *alternating*, *i.e.*, it changes sign when we exchange two variables z_i and z_j, since this exchanges two rows of the determinant.

Since a_α vanishes upon setting $z_i = z_j$, it must be divisible by

(2) $$\prod_{i<j} (z_i - z_j).$$

In particular, this is true of a_δ. But both a_δ and the above product have degree $\binom{n}{2}$, so a_δ must be a constant multiple of (2). By considering the coefficient of $\mathbf{z}^\delta$ we see that the constant is 1, so we have the *Vandermonde identity*

$$a_\delta = \prod_{i<j} (z_i - z_j).$$

Now for any λ with $l(\lambda) \le l$, both $a_{\lambda+\delta}$ and a_δ are alternating, and hence the "bialternant"

$$s_\lambda(\mathbf{z}) = a_{\lambda+\delta}/a_\delta$$

is a symmetric polynomial in $z_1, \dots, z_l$. These symmetric polynomials are the *Schur functions*. (The definition, which predates Schur, is Jacobi's).

Lemma: For $l(\lambda) \le l$ we have

$$a_{\lambda+\delta_{l+1}}(z_1, \dots, z_l, 0) = (z_1 \cdots z_l) a_{\lambda+\delta_l}(z_1, \dots, z_l).$$

Proof: Write out the two determinants and check that the result follows by elementary properties of determinants.

Corollary: For $l(\lambda) \le l$ we have $s_\lambda(z_1, \dots, z_l, 0) = s_\lambda(z_1, \dots, z_l)$, where the first Schur function is as defined in $l+1$ variables and the second is as defined in l variables.

Corollary: Let $s_\lambda(\mathbf{z}) = \sum_\mu K_{\lambda\mu} m_\mu(\mathbf{z})$. Then the integers $K_{\lambda\mu}$ do not depend on the number of variables l, provided only that we take it large enough for s_λ to be defined.

In infinitely many variables, and over any coefficient ring R, we now define the Schur functions s_λ by the formula

$$s_\lambda = \sum_\mu K_{\lambda\mu} m_\mu,$$

where the $K_{\lambda\mu}$ are as in the corollary. Then we automatically have that the specialization of s_λ to l variables, for $l \ge l(\lambda)$, agrees with the bialternant formula. The specialization of s_λ to fewer than $l(\lambda)$ variables, let us note, is zero. To see this, observe that in the bialternant formula for s_λ in $l = l(\lambda)$ variables, we have $\alpha_i > 0$ for all i, so if we specialize any further variable to zero, the determinant in the the numerator will have a zero row and will vanish.

There is a simple procedure for expressing any symmetric polynomial f in terms of Schur functions. Namely, specialize f to l variables, where l is large enough (l at least the degree of f will do), multiply f by a_δ, and extract the coefficient of $\mathbf{z}^{\lambda+\delta}$, which is also the coefficient of $a_{\lambda+\delta}$ in $a_\delta f$ and hence the coefficient of s_λ in f. Note that, at least over $\mathbb{Z}$ or any ring R in which 2 is not a zero-divisor, the $a_{\lambda+\delta}$'s form a "monomial" basis of the alternating polynomials just as the m_λ's do for the symmetric polynomials.

Over $\mathbb{Z}$, this shows in particular that every m_λ is a $\mathbb{Z}$-linear combination of Schur functions (of the same degree), and therefore, for each n, the square integer matrix $[K_{\lambda\mu}]$ with rows and columns indexed by partitions λ, μ of n is invertible. This implies the following result, for any coefficient ring R.

Proposition: The Schur functions s_λ form an R-basis of Λ_R.

1.7. Power-sums.

We define the k-th *power-sum symmetric function* to be

$$p_k = m_{(k)} = \sum_i z_i^k.$$

As with the elementary and complete homogeneous symmetric functions, we define

$$p_\lambda = p_{\lambda_1} \cdots p_{\lambda_l}.$$

Proposition: We have

$$p_\lambda = \sum_{\nu \geq \lambda} u_{\lambda\nu} m_\lambda,$$

with $u_{\lambda\lambda} = \prod_i r_i!$, where $\lambda = (1^{r_1}, 2^{r_2}, \dots)$.

Proof: By considering the monomials that could appear in the product $p_{\lambda_1} \cdots p_{\lambda_k}$ we see immediately that $u_{\lambda\nu} = 0$ unless λ *refines* ν, that is, the parts of λ can be obtained by further partitioning the parts of ν. It is not hard to see that if λ refines ν then also $\lambda \leq \nu$ in the dominance order.

The formula for $u_{\lambda\lambda}$, the coefficient of $\mathbf{z}^\lambda$ in p_λ, is straightforward.

We now assume that our ground ring R contains $\mathbb{Q}$, so that we have division by non-zero integers, and the triangular matrix $[u_{\lambda\nu}]$ is invertible.

Corollary: Over R containing $\mathbb{Q}$, the power-sums p_λ form a basis of Λ_R, and we have $\Lambda_R \cong R[p_1, p_2, \dots]$.

1.8. Plethystic substitution.

Assume $\mathbb{Q} \subseteq R$. Then by Corollary 1.7, if $\mathbb{F}$ is an R-algebra, and $f_1, f_2, \dots \in \mathbb{F}$ are arbitrary, we may define a unique R-algebra homomorphism $\Lambda_R \to \mathbb{F}$ mapping the k-th power-sum p_k to f_k, for all k.

Specifically, we take $\mathbb{F}$ to be the algebra of formal Laurent series with coefficients in R, in an alphabet of variables, or *letters*, $a_1, a_2, \dots$. (In practice the letters will be denoted x_i, y_i, q, t, etc. rather than a_i.) Now given $A \in \mathbb{F}$, we define

$$\mathrm{ev}_A \colon \Lambda_{\mathbb{F}} \to \mathbb{F}$$

by the rule

$$\mathrm{ev}_A(p_k) = A|_{a_i \mapsto a_i^k},$$

that is, the image of p_k is the result of replacing each letter by its k-th power in A. Note that we can perfectly well define this on $\Lambda_{\mathbb{F}}$, not just on Λ_R, which will be useful later on. Here the symmetric function variables $\mathbf{z}$ are necessarily *not* part of the alphabet of letters, otherwise it would make no sense to speak of $\Lambda_{\mathbb{F}}$.

Definition: The *plethystic substitution* of A into f is $\mathrm{ev}_A(f)$, henceforth denoted $f[A]$.

Here is the simplest, but most important, example of this construction. Let $X = x_1 + x_2 + \cdots$ where $x_1, x_2 \ldots$ are letters. Then by definition, $p_k[X] = X|_{x_i \mapsto x_i^k} = p_k(x_1, x_2, \ldots) = p_k(\mathbf{x})$. Since plethystic substitution is an algebra homomorphism it immediately follows that for all $f \in \Lambda$ we have

$$f[X] = f(\mathbf{x}), \quad \text{when } X = x_1 + x_2 + \cdots.$$

Here X may be a finite or infinite sum of letters. This means we are now free to forget about our original variables $\mathbf{z}$: we may view Λ_R purely formally, and recover symmetric functions in actual variables $\mathbf{x}$ through the plethystic substitution of their formal sum X.

For clarity, let me repeat the set-up here. We first fix a ground ring R and an alphabet of letters, denoting by $\mathbb{F}$ the algebra of formal Laurent series in the letters. We form the algebra of symmetric functions $\Lambda_{\mathbb{F}}$ in formal variables $\mathbf{z}$ which are not part of the alphabet of letters. Immediately and forever, however, we forget about the variables $\mathbf{z}$, viewing $\Lambda_{\mathbb{F}}$ purely formally as $\mathbb{F}[p_1, p_2, \ldots]$. To recover an actual symmetric function $f(\mathbf{x})$ in *letters* $\mathbf{x}$, we take the plethystic evaluation $f[X]$, where here and throughout, X will always stand for the formal sum $x_1 + x_2 + \cdots$.

More generally, if $A = t_1 + t_2 + \cdots$ is a sum of monomials in the letters, each with coefficient 1, then $f[A] = f(t_1, t_2, \ldots)$. Note that if A is a series with positive integer coefficients, we can treat it as a sum of monomials with coefficient 1 by repeating terms as many times as their coefficients indicate, *e.g.*, $f[(x+y)^2] = f[x^2 + 2xy + y^2] = f(x^2, xy, xy, y^2)$.

One caution must be observed with plethystic substitution: since the letters have a distinguished role, it is generally not permissible to substitute for them inside the plethystic bracket. Thus if f is homogeneous of degree n it is true that $f[tX] = t^n f[X]$, where t is a letter, but NOT true that $f[-X] = (-1)^n f[X]$. This happens because the evaluation map ev_A is defined in terms of replacing letters by their k-th powers, but it makes no sense to "replace -1 by its k-th power." However the substitution of a monomial in the letters for a letter in a plethystic identity IS permissible, since this operation is compatible with the definition of ev_A.

Because we defined plethystic substitution using power-sums, we have assumed to this point that our coefficient ring R contains $\mathbb{Q}$. For series A with *integer coefficients* , however, possible to define the plethystic substitution $f[A]$ for $f \in \Lambda_{\mathbb{F}}$ over an arbitrary ground ring R, as we shall see below.

1.9. Plethystic identities.

If we take $X = x_1 + x_2 + \cdots$, where $\mathbf{x}$ is infinite, then the $p_k[X] = p_k(\mathbf{x})$ are algebraically independent. It follows that any plethystic equation in X which holds for $X = x_1 + x_2 + \cdots$, holds identically. The same principle may be applied with more than one formal alphabet $X, Y, \ldots$.

Example: From the definition of the elementary symmetric functions it is easy to see that

$$e_n(x_1, x_2, \ldots, y_1, y_2, \ldots) = \sum_{k+l=n} e_k(x_1, x_2, \ldots)e_l(y_1, y_2, \ldots),$$

where here and elsewhere we adopt the convention that $e_0 = 1$. This proves the plethystic identity

(3) $$e_n[X+Y] = \sum_{k+l=n} e_k[X]e_l[Y],$$

for arbitrary X, Y. By similar reasoning we have the identity

(4) $$h_n[X+Y] = \sum_{k+l=n} h_k[X]h_l[Y].$$

1.10. The "Cauchy kernel" Ω.

It is useful for many purposes to define the symmetric infinite series

$$\Omega = \sum_{n \geq 0} h_n,$$

with the convention that $h_0 = 1$.

The identity (4) implies the following identity.

Proposition: $\Omega[X+Y] = \Omega[X]\Omega[Y]$.

This immediately leads to some interesting formulas. To begin with, for a single letter x, we have $h_n[x] = h_n(x) = x^n$, so

$$\Omega[x] = \frac{1}{1-x}.$$

Then by the multiplicative property, with $X = x_1 + x_2 + \cdots$ as usual, we have

$$\Omega[X] = \sum_n h_n(\mathbf{x}) = \prod_i \frac{1}{1-x_i}.$$

Now using $\Omega[X]\Omega[-X] = 1$, we have

$$\Omega[-X] = \prod_i (1-x_i) = \sum_n (-1)^n e_n(\mathbf{x}).$$

Here the last sum is easily seen to be equal to the product, by inspection. Comparing homogeneous components in each degree n we obtain the plethystic identity

(5) $$h_n[-X] = (-1)^n e_n[X],$$

and hence also

(6) $$e_n[-X] = (-1)^n h_n[X].$$

Note that X is not a letter—it is *arbitrary*, so it is correct to substitute $-X$ for X in (5) to get (6).

Combining (6) with (3), we have

$$(7) \qquad e_n[X-Y] = \sum_{k+l=n} (-1)^l e_k[X] h_l[Y].$$

This formula allows us to evaluate $e_n[A]$ for any A with integer coefficients, without reference to the definition of plethysm in terms of power sums. We simply write A as $X - Y$ where X and Y have positive integer coefficients, repeat terms as necessary to obtain sums of monomials with coefficient 1, and evaluate using (7).

Since $\Lambda_R \cong R[e_1, e_2, \ldots]$, we can define $f[A]$ for A with integer coefficients, over any coefficient ring R.

Another application of Ω is to express the complete homogeneous symmetric functions h_n in terms of the power-sums. For this, take $X = x_1 + x_2 + \cdots$ as always, so

$$\begin{aligned}
\Omega[X] &= \prod_i \frac{1}{1-x_i} \\
&= \exp(\sum_i \log \frac{1}{1-x_i}) \\
&= \exp(\sum_i \sum_{k>0} x_i^k/k) \\
&= \exp(\sum_{k>0} p_k[X]/k) \\
&= \prod_{k>0} \exp(p_k[X]/k) \\
&= \prod_{k>0} \sum_{r\geq 0} p_k[X]^r/r!k^r \\
&= \sum_\lambda p_\lambda[X]/z_\lambda,
\end{aligned}$$

where if $\lambda = (1^{r_1}, 2^{r_2}, \ldots)$, we define $z_\lambda = \prod_k r_k! k^{r_k}$. Hence in particular we have

$$h_n = \sum_{|\lambda|=n} p_\lambda/z_\lambda.$$

The quantity z_λ has an interesting combinatorial interpretation: it is the number of elements in the symmetric group S_n commuting with any given permutation w whose cycle type (the partition given by the lengths of its disjoint cyles) is λ. The significance of this point will appear later on.

1.11. The involution ω.

Since $\Lambda_R \cong R[e_1, e_2, \ldots]$, there is a unique endomorphism

$$\omega\colon \Lambda_R \to \Lambda_R \quad \text{defined by} \quad \omega e_k = h_k.$$

(Caution: ω is defined on the algebra of symmetric functions in infinitely many variables but is not compatible with specialization to finitely many variables, since the latter makes

$e_k = 0$ for $k > l$, but does not make $h_k = 0$.) By virtue of (6), for symmetric functions homogeneous of degree n, we have

$$\omega f[X] = (-1)^n f[-X].$$

Thus ω is essentially plethystic substitution of $-X$ for X, up to a sign depending on the degree. In particular we immediately see that

$$\omega^2 = 1,$$

i.e., ω is an involution. In particular, ω is an isomorphism, which yields the following result.

Proposition: The complete homogeneous symmetric functions h_λ form an R-basis of Λ_R.

1.12. First Cauchy formula and Hall inner product.

As usual, let $X = x_1 + x_2 + \cdots$, $Y = y_1 + y_2 + \cdots$. Then we have the *first Cauchy identity*

$$\begin{aligned}\Omega[XY] &= \prod_{i,j} \frac{1}{1 - x_i y_j} \\ &= \prod_i \Omega[x_i Y] \\ &= \prod_i \sum_n x_i^n h_n[Y] \\ &= \sum_{\nu_1, \nu_2, \ldots} \mathbf{x}^\nu h_\nu[Y] \\ &= \sum_\lambda m_\lambda[X] h_\lambda[Y].\end{aligned}$$

By symmetry we also have

$$\Omega[XY] = \sum_\lambda h_\lambda[X] m_\lambda[Y],$$

which is another version of the fact we observed earlier, that the matrix giving the expansion of the h_λ by the μ_ν is symmetric.

Definition: The *Hall inner product* $\langle -, - \rangle$ is the R-bilinear form on Λ_R uniquely defined by requiring

$$\langle h_\lambda, m_\nu \rangle = \delta_{\lambda\mu},$$

i.e., we make the monomial symmetric functions and the complete homogeneous symmetric functions into dual bases.

According to this definition, the coefficient of m_λ in the monomial expansion of any f is given by $\langle h_\lambda, f \rangle$. In particular,

$$\langle h_\lambda, h_\mu \rangle = c_{\lambda\mu}, \quad \text{where} \quad h_\lambda = \sum_\mu c_{\lambda\mu} m_\mu.$$

Since $c_{\lambda\mu} = c_{\mu\lambda}$, the inner product $\langle -, - \rangle$ is *symmetric*.

Proposition: Taking $X = x_1 + x_2 + \cdots$, and considering the Hall inner product with respect to symmetric functions in $\mathbf{x}$, we have the identity

$$\langle \Omega[AX], f(\mathbf{x})\rangle = f[A].$$

Proof. Both sides are linear in f so it suffices to verify the identity for $f = m_\lambda$. Then the left hand side is the coefficient of $h_\lambda[X]$ in $\Omega[AX]$, which is $m_\lambda[A]$ by the Cauchy identity. Indeed, the Proposition can be viewed as a basis-free formulation of the Cauchy identity, as is made more precise by the following important corollary.

Corollary: Let $\{u_\lambda\}$ be a homogeneous R-basis of Λ_R. Then $\{v_\lambda\}$ is the Hall-dual basis if and only if we have the Cauchy identity

$$\Omega[XY] = \sum_\lambda u_\lambda[X]v_\lambda[Y].$$

Proof: The R-basis $\{u_\lambda\}$ has a unique dual basis $\{v_\lambda\}$, since the Hall inner product is non-degenerate on each homogeneous component $\Lambda_R^{(n)}$, which is a finitely generated free R-module. There is also a unique expansion of $\Omega[XY]$ in terms of the u_λ with some coefficients $w_\lambda[Y]$. Thus the "if" and "only if" are both equivalent to the identity $v_\lambda = w_\lambda$. By the previous Proposition, $v_\lambda[Y] = \langle \Omega[XY], v_\lambda[X]\rangle_{\mathbf{x}} = w_\lambda[Y]$.

Corollary: We have the identity

$$\langle g[X]\Omega[AX], f[X]\rangle_{\mathbf{x}} = \langle g[X], f[A+X]\rangle_{\mathbf{x}}.$$

In other words the operator of multiplication by $\Omega[AX]$ is Hall-adjoint to the plethystic shift operator sending $f[X]$ to $f[A+X]$.

Proof. Here we use another handy plethystic trick. Since both sides are linear in $g[X]$, it is sufficient to verify the identity with $g[X] = \Omega[BX]$, where B is arbitrary, since by the Cauchy formula, $\Omega[BX]$ is a linear combination of the symmetric function basis elements $m_\lambda[X]$ with coefficients $h_\lambda[B]$ that are themselves linearly independent for general B (say, $B = b_1 + b_2 + \cdots$).

Now by Proposition 1.12 we have

$$\langle \Omega[BX]\Omega[AX], f[X]\rangle = \langle \Omega[(A+B)X], f[X]\rangle = f[A+B] = \langle \Omega[BX], f[A+X]\rangle.$$

1.13. Raising operators for Schur functions.

In l variables $x_1, \ldots, x_l$, using the Vandermonde identity, we may write the bialternant formula for the Schur function s_λ as

$$s_\lambda(\mathbf{x}) = \frac{\sum_{w\in S_l} \epsilon(w)w(\mathbf{x}^{\lambda+\delta})}{\prod_{i<j}(x_i - x_j)}.$$

The denominator is alternating, so we can account for the sign $\epsilon(w)$ by bringing it inside the sum to get

$$s_\lambda(\mathbf{x}) = \sum_{w\in S_l} w\left(\frac{\mathbf{x}^{\lambda+\delta}}{\prod_{i<j}(x_i - x_j)}\right) = \sum_{w\in S_l} w\left(\frac{\mathbf{x}^{\lambda}}{\prod_{i<j}(1 - x_j/x_i)}\right). \tag{8}$$

Here we have divided numerator and denominator by $\mathbf{x}^\delta$ to obtain the last formula.

Now let S_{l-1} be the subgroup of S_l consisting of permutations that fix 1. We can isolate the role of the first part λ_1 of λ by organizing the terms in (8) according to cosets of S_{l-1}

containing w. Note that the elements v of a given right coset are those with a given value of $v(1) = k$. The sum with w ranging only over S_{l-1} is equal to

$$\frac{x_1^{\lambda_1}}{\prod_{j\neq 1}(1 - x_j/x_1)} s_\nu(x_2, \dots, x_l),$$

where $\nu = (\lambda_2, \lambda_3, \dots)$. Operating on this with a coset representative v such that $v(1) = k$, and summing over cosets, that is, over k, gives the total sum in the form

$$s_\lambda(\mathbf{x}) = \sum_k \frac{x_k^{\lambda_1}}{\prod_{j\neq k}(1 - x_j/x_k)} s_\nu[X - x_k]. \tag{9}$$

Here we have conveniently used the plethystic notation in terms of $X = x_1 + \cdots + x_l$ for the Schur function s_ν in the variables $\mathbf{x}$ with x_k omitted.

The interesting aspect of the above formula appears when we compare it with the following partial fraction expansion for $\Omega[zX]$, where $X = x_1 + \cdots + x_l$ as above, and we treat $\Omega[zX]$ as a function of z:

$$\Omega[zX] = \prod_{i=1}^{l} \frac{1}{1 - zx_i} = \sum_{k=1}^{l} \frac{1}{1 - zx_k} \prod_{j\neq k} \frac{1}{1 - x_j/x_k}. \tag{10}$$

To fully exploit the resemblance between the terms above with those of (9), consider the following simple identity, valid for any polynomial or formal power series $f(z)$ in z whose terms involve only non-negative powers z^n.

Lemma:

$$\frac{1}{1 - zu} f(z^{-1})|_{z^0} = \sum_{n\geq 0} z^n u^n f(z^{-1})|_{z^0} = \sum_{n\geq 0} u^n f(z^{-1})|_{z^{-n}} = f(u).$$

Here the vertical bar denotes the taking of a coefficient, and the statement of the Lemma contains its proof, since the last equation is essentially the *definition* of $f(u)$. More generally, we have for $r \geq 0$

$$\frac{1}{1 - zu} f(z^{-1})|_{z^r} = u^r f(u),$$

as can easily be seen by replacing $f(z)$ with $z^r f(z)$ in the Lemma. Now summing this over $u = x_1, \dots, x_l$ and using (10) yields the plethystic identity

$$f[X - z^{-1}]\Omega[zX]|_{z^r} = \sum_{k=1}^{l} \frac{x_k^r}{\prod_{j\neq k}(1 - x_j/x_k)} f[X - x_k],$$

valid when $X = x_1 + \cdots + x_l$. Note that the substitution of x_k for z^{-1} inside the plethystic bracket is permissible, since both are letters.

Definition: The *raising operator* B_r is the coefficient of z^r in the plethystic operator $B(z)$ defined by

$$B(z)f = f[X - z^{-1}]\Omega[zX].$$

Later we will see q-analogs of these raising operators in connection with Hall-Littlewood polynomials—the q versions are known as Jing's vertex operators. For now, our calculations above prove the raising operator formulas

$$s_\lambda[X] = B_{\lambda_1} s_{(\lambda_2,\lambda_3,\dots)}[X],$$

and hence

$$s_\lambda[X] = B_{\lambda_1} B_{\lambda_2} \cdots B_{\lambda_l}(1).$$

Since we derived the formulas for a finite but arbirtrarily number of variables $X = x_1 + \cdots + x_l$, they are actually valid as plethystic identities; in particular they are valid in infinitely many variables.

The operators B_r do not commute, but there is a commutation formula, which is of interest to derive. Introducing separate dummy variables u, v we have

$$\begin{aligned} B(u)B(v)f &= B(u)f[X - v^{-1}]\Omega[vX] \\ &= f[X - u^{-1} - v^{-1}]\Omega[v(X - u^{-1})]\Omega[uX] \\ &= \Omega[-v/u]f[X - u^{-1} - v^{-1}]\Omega[(u+v)X]. \end{aligned}$$

Since $\Omega[-v/u] = (1 - v/u)$, as u and v are letters, we have

$$uB(u)B(v)f = (u - v)f[X - u^{-1} - v^{-1}]\Omega[(u+v)X].$$

Now the right-hand side changes sign if we exchange u and v, giving the identity

$$uB(u)B(v) = -vB(v)B(u).$$

To extract information about the operators B_r, we may compare coeffecients on each side. Doing this with the coefficient of $u^{r+1}v^{s+1}$ yields the commutation relation

$$B_r B_{s+1} = -B_s B_{r+1}. \tag{11}$$

In particular, taking $r = s$ we have

$$B_r B_{r+1} = 0. \tag{12}$$

The commutation relations may be used to express any product of the operators B_r as plus-or-minus such a product with the indices decreasing, or zero. In particular this gives a definition for s_ν when ν is an integer sequence that is not a partition, as plus-or-minus a Schur function, or zero. You can easily verify that this definition agrees with what the bialternant formula would suggest for s_ν, as it should, since we did not actually use the fact that λ was a partition in deriving the raising operator formula from the bialternant formula.

1.14. Cauchy formula.

Let $Y = y_1 + \cdots + y_l$. We may write $s_\lambda[X] = B_{\lambda_1} \cdots B_{\lambda_l}(1)$ as

$$s_\lambda[X] = B(y_1) \cdots B(y_l) 1|_{\mathbf{y}^\lambda}.$$

Now it is an easy exercise to evaluate the above operator product, much as we did for $B(u)B(v)$ in 1.13, to obtain

$$\begin{aligned} s_\lambda[X] &= \Omega[XY]\Omega[-\sum_{i<j} y_j/y_i]|_{\mathbf{y}^\lambda} \\ &= \Omega[XY]\prod_{i<j}(1 - y_j/y_i)|_{\mathbf{y}^\lambda} \\ &= \Omega[XY]\prod_{i<j}(y_i - y_j)|_{\mathbf{y}^{\lambda+\delta}}. \end{aligned}$$

Now recall that the coefficient of $s_\lambda[Y]$ in any symmetric function $f[Y]$ is found by multiplying $f[Y]$ by $a_\delta(\mathbf{y})$ and extracting the coefficient of $\mathbf{y}^{\lambda+\delta}$. In view of this the above identity reads

$$\Omega[XY] = \sum_\lambda s_\lambda[X]s_\lambda[Y].$$

Of course if we obtain this with $Y = y_1 + \cdots + y_l$, for a fixed l, the sum only ranges over terms with $l(\lambda) \leq l$, all others having $s_\lambda[Y] = 0$. However, l was arbitrary, so the identity is valid as a plethystic idenitity, with the sum ranging over all λ. This is the *Cauchy formula* for Schur functions. By 1.12, it has the following consequence.

Proposition: The Schur functions form an orthonormal basis of Λ_R with respect to the Hall inner product $\langle -, - \rangle$.

1.15. Pieri formulas.

We may derive expressions for elementary and complete homogeneous symmetric functions in terms of Schur functions from formulas known as *Pieri formulas* for the multiplication of a Schur function by e_k or h_k. Viewing $\Omega[\pm uX]$ as a generating function for elementary or homogeneous symmetric functions, we can obtain the Pieri formulas from commutation formulas for the raising operator generating function $B(z)$ with the operator of multiplication by $\Omega[\pm uX]$. We begin with $\Omega[-uX]$, which is the easier case.

We have

$$\begin{aligned} B(z)\Omega[-uX]f &= \Omega[-u(X - z^{-1})]f[X - z^{-1}]\Omega[zX] \\ &= \Omega[u/z]\Omega[-uX]B(z)f, \end{aligned}$$

or

$$\Omega[-uX]B(z) = (1 - u/z)B(z)\Omega[-uX]$$

as an operator identity, where $\Omega[-uX]$ stands for the operator of multiplication by itself. Taking the coefficient of $(-u)^k z^r$ yields the operator identity

$$e_k B_r = (B_r e_k + B_{r+1}e_{k-1}). \qquad (13)$$

We'll need an auxilliary result to properly handle B_0:

$$B_0\Omega[-uX] = \Omega[u/z]\Omega[-uX]\Omega[zX]|_{z^0} = \frac{1}{1 - uz^{-1}}\Omega[uX]\Omega[-zX]|_{z^0}.$$

Now by Lemma 1.13, the last expression is equal to $\Omega[uX]\Omega[-uX] = 1$, so we have

$$B_0(e_k) = \begin{cases} 1 & k = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Combining this with (13) for $r = 0$, $f = 1$ we see by induction on k that

$$B_1^k(1) = e_k.$$

Now let us use (13) to compute $e_k s_\lambda[X]$. Applying (13) repeatedly to move the e_k to the right we obtain

$$e_k B_{\lambda_1} \cdots B_{\lambda_l}(1) = \sum_\nu B_{\nu_1} \cdots B_{\nu_l}(e_{k-j}) = \sum_\nu B_{\nu_1} \cdots B_{\nu_l} B_1^{k-j}(1),$$

where ν ranges over all integer sequences obtained from λ by increasing some (or none, or all) of the parts by 1, and j is the number of parts so increased.

By (12), any terms in the above sum for which $(\nu, 1^{k-j})$ is not a partition will vanish. This leaves one term for each partition $\mu = (\nu, 1^{k-j})$ obtainable from λ by increasing some parts by 1 and adding some parts of 1 at the end (effectively increasing extraneous parts of zero to 1). Such a partition is said to differ from λ by the addition of a *vertical k-strip* (k being the size), a terminology whose significance in terms of partition diagrams is apparent. Denoting the condition that μ is λ plus a vertical k-strip by $\mu/\lambda \in V_k$, we may express our first Pieri formula as follows.

Proposition:

$$e_k s_\lambda = \sum_{\mu/\lambda \in V_k} s_\mu.$$

For the second Pieri formula we rearrange our previous operator identity to read

$$\Omega[uX]B(z) = \frac{1}{1 - u/z} B(z)\Omega[uX],$$

obtaining

$$h_k B_r = \sum_{j=0}^{k} B_{r+j} h_j.$$

Note also that $h_k = B_k(1)$, directly from the definition of B_k. Hence

$$h_k s_\lambda[X] = h_k B_{\lambda_1} \cdots B_{\lambda_l}(1) = \sum_{j_1 + \cdots + j_{l+1} = k} B_{\lambda_1 + j_1} \cdots B_{\lambda_l + j_l} B_{j_{l+1}}(1).$$

As we have seen, each term on the right is plus-or-minus a Schur function, or zero. I claim there is perfect cancellation of all terms with some index i such that $\lambda_i + j_i > \lambda_{i-1}$. Consider such a term and let i be the greatest such index. Compare

$$B_{\lambda_1 + j_1} \cdots B_{\lambda_{i-1} + j_{i-1}} B_{\lambda_i + j_i} \cdots 1;$$
$$B_{\lambda_1 + j_1} \cdots B_{\lambda_i + j_i - 1} B_{\lambda_{i-1} + j_{i-1} + 1} \cdots 1.$$

The second term has $\lambda_i + j_i - 1 \geq \lambda_{i-1}$ and $\lambda_{i-1} + j_{i-1} + 1 > \lambda_i$ so $\lambda_i + j_i - 1 = \lambda_{i-1} + j'_{i-1}$, say, and $\lambda_{i-1} + j_{i-1} + 1 = \lambda_i + j'_i$. Written this way the second term also satisfies $\lambda_i + j'_i > \lambda_{i-1}$, with i the greatest such index, and the same construction applied to it returns us to the first term. Thus terms of this type pair off with opposite sign, except when $\lambda_i + j_i = \lambda_{i-1} + 1$, in which case the term is zero.

The surviving terms have $\lambda + \mathbf{j} = \mu$, where μ is λ plus a *horizontal* k-strip (conjugate concept to vertical k-strip). This gives the second Pieri formula.

Proposition:

$$h_k s_\lambda = \sum_{\mu/\lambda \in H_k} s_\mu.$$

Note that either Pieri formula completely determines every s_λ, by providing an inductive rule for expressing the e_λ's (respectively the h_λ's) in terms of Schur functions. From the conjugate symmetry between the two formulas we obtain the following corollary.

Corollary: We have $\omega s_\lambda = s_{\lambda'}$, or $s_\lambda[-X] = (-1)^{|\lambda|} s_{\lambda'}[X]$.

Another corollary, this one to the second Pieri rule, is the combinatorial formula for the *Kostka coefficients* $K_{\lambda\mu}$ giving Schur functions in terms of monomials. Recall the adjunction formula $\langle f[X+A], g\rangle = \langle f, \Omega[AX]g\rangle$ from 1.10. In particular, taking $A = u$ a single letter, and comparing coefficients of u^k, we have

$$\langle f[X+u], g\rangle|_{u^k} = \langle f, h_k g\rangle.$$

(This is a plethystically disguised version of the dual-basis relationship between the m_λ's and h_λ's.) By the second Pieri formula,

$$\langle s_\nu[X+u], s_\lambda\rangle|_{u^k} = \langle s_\nu, h_k s_\lambda\rangle$$

is 1, if ν/λ is a horizontal k-strip, and zero otherwise. In other words,

$$s_\nu[X+u] = \sum_k u^k \sum_{\lambda:\, \nu/\lambda \in H_k} s_\lambda.$$

Iterating this, we see that the coefficient of $\mathbf{u}^\kappa$ in $s_\nu[u_1 + \cdots + u_l]$ is equal to the number of sequences

$$\emptyset = \lambda^{(0)} \subseteq \lambda^{(1)} \subseteq \cdots \subseteq \lambda^{(l)} = \nu \quad \text{such that } \lambda^{(i)}/\lambda^{(i-1)} \in H_{\kappa_i} \text{ for all } i = 1, \ldots, l.$$

Filling in the difference diagram between $\lambda^{(i-1)}$ and $\lambda^{(i)}$ with i's, we get a filling of the diagram of ν with κ_1 1's, κ_2 2's, and so on, with non-decreasing rows and columns, in which the horizontal strip condition on the positions of each i means that *columns* increase *strictly*. Such a filling is a *column strict Young tableau of shape* ν and *content* κ.

Corollary: The Kostka coefficient $K_{\lambda\mu}$ given by the expansion

$$s_\lambda = \sum_\mu K_{\lambda\mu} m_\mu$$

is equal to the number of column-strict Young tableaux of shape λ and content μ.

It also follows that the number of column-strict Young tableaux is symmetric with respect to permutations of the content vector κ. *Exercise:* prove this directly and combinatorially by considering what happens for adjacent transpositions in κ.

1.16. Jacobi-Trudi formula.

Consider the problem of expanding Schur functions in terms of the complete homogeneous symmetric functions h_μ. By Hall duality, the coefficient of h_μ in s_λ is given by

$$\langle s_\lambda, m_\mu \rangle,$$

which is also the coefficient of s_λ in m_μ. We find the latter by the usual device for Schur function expansions to be

$$\begin{aligned} m_\mu a_\delta\big|_{\mathbf{x}^{\lambda+\delta}} &= \sum_{w \in S_l} \epsilon(w) m_\mu\big|_{\mathbf{x}^{\lambda+\delta-w(\delta)}} \\ &= \sum_{w \in S_l} \epsilon(w) \begin{cases} 1 & \text{if } \mathbf{x}^{\lambda+\delta-w(\delta)} \approx \mathbf{x}^\mu, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

This implies the *Jacobi-Trudi identity*, given as follows.

Proposition:

$$s_\lambda = \sum_{w \in S_l} \epsilon(w) h_{\lambda+\delta-w(\delta)} = \det \begin{bmatrix} h_{\lambda_1} & h_{\lambda_1+1} & h_{\lambda_1+2} & \cdots \\ h_{\lambda_2-1} & h_{\lambda_2} & h_{\lambda_2+1} & \cdots \\ & & \vdots & \\ \cdots & & h_{\lambda_l-1} & h_{\lambda_l} \end{bmatrix},$$

with the convention that $h_0 = 1$ and $h_{-k} = 0$ for $-k$ negative.

The format of the Jacobi-Trudi matrix on the right is: $h_{\lambda_1}, \ldots, h_{\lambda_l}$ on the diagonal, indices changing by 1 from column to column within each row. Here l can be any integer greater than or equal to the number of parts of λ. Applying ω gives the analog for elementary symmetric functions.

Corollary:

$$s_{\lambda'} = \det \begin{bmatrix} e_{\lambda_1} & e_{\lambda_1+1} & e_{\lambda_1+2} & \cdots \\ e_{\lambda_2-1} & e_{\lambda_2} & e_{\lambda_2+1} & \cdots \\ & & \vdots & \\ \cdots & & e_{\lambda_l-1} & e_{\lambda_l} \end{bmatrix}.$$

Notes II: Representation theory.

2.1. Group representation basics.

A (complex) *matrix representation* of a group G is a homomorphism

$$\rho\colon G \to GL_n(\mathbb{C})$$

from G to the group of $n \times n$ invertible complex matrices. A (complex) *linear representation* of G is a homomorphism

$$\rho\colon G \to \mathrm{End}(V)$$

where V is a finite-dimensional vector space over $\mathbb{C}$. Of course every matrix representation induces a linear representation on $V = \mathbb{C}^n$, and conversely, choosing a basis in a linear representation on V induces a matrix representation.

Matrix representations ρ, ρ' are *similar* if there is a matrix A such that $\rho'(g) = A^{-1}\rho(g)A$ for all $g \in G$. Similar matrix representation induce isomorphic linear representations; conversely, isomorphic linear representations with arbitrary bases, or the same linear representation with two choices of basis, induce similar matrix representations.

The space V of a linear representation, with its G-action given by ρ, is also called a G-module. When ρ is understood from context we often write gv instead of $\rho(g)v$, for $g \in G$ and $v \in V$.

A *submodule* of V is a subspace $W \subseteq V$ such that $gW \subseteq W$ for all $g \in G$. We say that V is *irreducible* if its only submodules are 0 and V itself. We say that V is *completely reducible* if every submodule of V is a direct summand of V, which implies V is a direct sum of irreducible modules.

When the group G is $GL_n(\mathbb{C})$ or a subgroup thereof, we will chiefly be concerned with *polynomial representations*, in which the entries of the representing matrix $\rho(g)$ are polynomials in the entries of the matrix $g \in GL_n$. As an algebraic variety, GL_n is by definition the open set in the affine n^2 dimensional space of $n \times n$ matrices defined by the non-vanishing of the polynomial $\det g$. Therefore the globally defined *regular functions* on GL_n are generated by the polynomials in the entries of g and the multiplicative inverse of the determinant $(\det g)^{-1}$. If the entries of $\rho(g)$ are regular functions of the entries of g, we say that ρ is a *rational representation.*

Note that $\det\colon g \mapsto \det(g)$ is itself a 1-dimensional matrix representation of GL_n, and it follows that if ρ is a rational representation so is $(\det)^k \otimes \rho$, where the latter is defined by $g \mapsto (\det g)^k \rho(g)$, for any integer k. Since the regular functions on g have denominator a power of $\det g$ it follows that for every rational representation ρ, the representation $(\det)^k \otimes \rho$ is polynomial, for k sufficiently large. In other words, every rational representation has the form $(\det)^k \otimes \rho$ for some polynomial representation ρ and integer k. Thus the theories of polynomial representations and of rational representations are essentially interchangeable.

Example: The symmetric group S_n acts on $V = \mathbb{C}^n$ by permuting the basis vectors $e_1, \ldots, e_n$. The vector $v = e_1 + \cdots + e_n$ is invariant, so its span $W = \mathbb{C}v$ is a 1-dimensional submodule of V, on which S_n acts by the *trivial* representation. If we set $W' = \{\sum_i a_i e_i : a_1 + \cdots + a_n = 0\}$, then W' is also a submodule, spanned by the vectors $e_i - e_{i+1}$, and

$V = W \oplus W'$. This exhibits V as a direct sum of two irreducible submodules. *Exercise:* prove W' is irreducible.

Example: Let $B \subseteq GL_2$ be the subgroup of upper triangular 2×2 matrics. We have the *defining representation* of B on $V = \mathbb{C}^2$, in which each matrix in B is represented by itself. Since

$$\begin{bmatrix} x_1 & y \\ 0 & x_2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix},$$

the subspace W spanned by the basis vector $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is a submodule, on which B acts through the homomorphism sending $\begin{bmatrix} x_1 & y \\ 0 & x_2 \end{bmatrix}$ to $[x_1]$. There is a quotient module V/W, on which B acts via the homomorphism sending $\begin{bmatrix} x_1 & y \\ 0 & x_2 \end{bmatrix}$ to $[x_2]$, but there is no submodule W' such that $V = W \oplus W'$, so V is not completely reducible.

Theorem: Every complex representation of a finite group G is completely reducible.

This theorem also holds over any field whose characteristic does not divide the order of G. *Exercise:* find an example of a non-completely reducible representation of a finite group G over a field of prime characteristic dividing $|G|$.

Theorem: Every rational representation of $G = GL_n$ is completely reducible.

This theorem holds more generally for rational representations of semi-simple algebraic groups and for continuous representations of semi-simple Lie groups. For more on this, consult a textbook on Lie groups and Lie algebras.

Example: The defining representation of GL_n on $V = \mathbb{C}^n$ is obviously a polynomial representation. It is irreducible since for any non-zero vector v, we can find group elements g_i such that the vectors $g_i v$ span V.

Example: As we have seen, the homomorphism sending $g \in GL_n$ to $[\det g]$ defines a 1-dimensional matrix representation of GL_n, the *determinant* representation. Clearly it is a polynomial representation, and irreducible because 1-dimensional.

Example: $V \otimes V$ is a polynomial representation. A basis of $V \otimes V$ is given by the elements $e_i \otimes e_j$, where $e_1, \ldots, e_n$ is the usual basis of V. Then $g(e_i \otimes e_j) = ge_i \otimes ge_j = \sum_{i',j'} g_{i'i} g_{j'j} e_{i'} \otimes e_{j'}$, which shows that the entries of the matrix representing g on $V \otimes V$ are products $g_{i'i} g_{j'j}$, so this is a polynomial representation. *Exercise:* show that $V \otimes V = \bigwedge^2 V \oplus S^2 V$ is a decomposition of $V \oplus V$ into irreducibles, where $\bigwedge^2 V$ and $S^2 V$ are the second exterior and symmetric powers of V respectively (this will be easy later on).

Example: GL_n acts on the dual space V^* via the homorphism sending g to $(g^{-1})^T$. This is a rational but not a polynomial representation. The representation $(\det V) \otimes V^*$ is polynomial, since the entries of $(\det g)(g^{-1})^T$ are polynomials in the entries of g.

2.2. Characters.

If ρ and ρ' are similar matrix representations of G, $\rho'(g) = A^{-1}\rho(g)A$, for some A, then clearly we have $\operatorname{tr} \rho'(g) = \operatorname{tr} \rho(g)$. In particular, given a linear representation $\rho \colon G \to GL(V)$, the trace of $\rho(g)$ in a corresponding matrix representation is independent of choice of basis, and equal to its value in any isomorphic representation W.

Definition: The function $\chi^V : G \to \mathbb{C}$ sending g to the trace of its representing matrix is the *character* of the representation V.

Proposition: Characters are *class functions*, that is, they are constant on conjugacy classes of G.

Proof: If $g' = h^{-1}gh$ then $\rho(g') = \rho(h)^{-1}\rho(g)\rho(h)$, so the matrices $\rho(g')$ and $\rho(g)$ have the same trace.

The following two properties of characters are easy to verify by analyzing the structure of the matrix of a direct sum or tensor product of two linear maps.

Proposition: We have

$$\chi^{V\oplus W} = \chi^V + \chi^W$$

(more generally, $\chi^V = \chi^{V/W} + \chi^W$ for any submodule $W \subseteq V$, even if not a direct summand), and

$$\chi^{V\otimes W} = \chi^V \chi^W.$$

For a finite group G, the *group algebra* $\mathbb{C}G$ is the vector space of formal linear combinations of elements of G (with basis G), and multiplication defined by extending the multiplication in G linearly. Then G acts on $\mathbb{C}G$ by left multiplication, yielding a representation of G, called the *regular representation.* The basic facts about finite group representations are summarized in the following theorem.

Theorem: Let G be a finite group. Then every irreducible representation V_α of G occurs as a submodule of the regular representation $\mathbb{C}G$, and in fact $\mathbb{C}G$ is the direct sum over all α of $\chi^\alpha(1) = \dim V_\alpha$ copies of V_α. In particular there are only finitely many isomorphism classes of irreducible G modules. Moreover their characters form a basis of the space of class functions, so the number of non-isomorphic irreducibles is equal to the number of conjugacy classes in G.

The character of a representation of an infinite group G is well-defined, in particular, we can speak of the character of a polynomial representation of GL_n. Note that the character is itself a polynomial in the entries of g. The diagonalizable elements of GL_n form a dense set, so χ^V is determined by its values on diagonalizable elements g, and hence, since it is a class function, by its values on the diagonal elements. Denoting by $\tau(\mathbf{x})$ the diagonal matrix with diagonal entries $x_1, \ldots, x_n$, $\chi^V(\tau(\mathbf{x}))$ is a polynomial in $\mathbf{x}$. We will denote it simply by $\chi^V(\mathbf{x})$ and refer to this polynomial as the character of V.

Note that for any permutation $w \in S_n$, $\tau(x_{w(1)}, \ldots, x_{w(n)})$ is conjugate to $\tau(\mathbf{x})$ by a permutation matrix in GL_n. This implies that $\chi^V(\mathbf{x})$ is a *symmetric* polynomial in $\mathbf{x}$.

Note also that the character of $\det g$ is $x_1 \cdots x_n = e_n(\mathbf{x})$. The character of the rational representation $(\det g)^k$ is $(x_1 \cdots x_n)^k$, where k may be any integer. Thus the characters of the rational representations of GL_n are rational functions of the form $(x_1 \cdots x_n)^k f(\mathbf{x})$, where f is a symmetric polynomial. In other words they are symmetric *Laurent polynomials* in $\mathbf{x}$. These characters make sense literally, not just formally, since the x_i's are the eigenvalues of g, and these are non-zero as g is invertible.

2.3. Schur's Lemma.

Lemma: If V and W are irreducible G-modules, and $\phi\colon V \to W$ is a G-module homorphism (a linear map that commutes with the G actions), then either $\phi = 0$ or ϕ is an isomorphism.

Proof: Assume $\phi \neq 0$. Then the kernel of ϕ is a submodule of V, not equal to V, and since V is irreducible, $\ker(\phi) = 0$. Similarly the image of ϕ is a submodule of W, not equal to 0, and since W is irreducible, $\operatorname{im}(\phi) = W$. This shows ϕ is injective and surjective, hence an isomorphism.

Corollary: If V is irreducible then the space of G-module homomorphisms $\operatorname{Hom}_G(V, V)$ is 1-dimensional, *i.e.*, every $\phi\colon V \to V$ is a scalar multiple of the identity map.

Proof: Given $\phi\colon V \to V$ let α be an eigenvalue of ϕ. Then $\phi - \alpha 1_V$ is singular, hence not an isomorphism, hence zero by Schur's Lemma. Note the corollary uses the fact that $\mathbb{C}$ is algebraically closed, so that ϕ must have an eigenvalue $\alpha \in \mathbb{C}$.

2.4. Polynomial representations of GL_n.

For the moment we fix $G = GL_n$. We denote by B the *Borel subgroup* of upper triangular matrices, by N the *unipotent radical* of B consisting of upper triangular matrices with 1's on the diagonal, and by T the *torus* of diagonal matrices in G. Note that $T \cong (\mathbb{C}^*)^n$ is indeed an algebraic torus group. We also write B^- and N^- for the lower-triangular analogs of B and N, the *opposite* Borel subgroup and its unipotent radical.

The elements of T will be denoted $\tau(x_1, \dots, x_n)$, as in 2.2. Thus our convention for representing a character of g as a function $\chi(x_1, \dots, x_n)$ amounts to evaluating χ on the torus T.

The torus group T is itself semi-simple (it's a product of copies of GL_1), so any G module V is completely reducible as a T module. One can show that the polynomial representations of T are simply the one-dimensional representations $\mathbb{C}_\lambda$, in which $\tau(\mathbf{x})$ acts as multiplication by $\mathbf{x}^\lambda$. Here λ is a sequence of non-negative integers $(\lambda_1, \dots, \lambda_n)$. The rational representation are the same except that the λ_i are any integers, possibly negative. Complete reducibility as a T module then means that V decomposes as a direct sum of *weight spaces* V_λ, whose elements $v \in V_\lambda$ are simultaneous eigenvectors for all $\tau(\mathbf{x}) \in T$, with eigenvalue $\mathbf{x}^\lambda$. Weights for which λ is a partition are called *dominant.*

Proposition: The character $\chi^V(\mathbf{x})$ is equal to

$$\chi^V(\mathbf{x}) = \sum_\lambda (\dim V_\lambda)\mathbf{x}^\lambda.$$

In particular the character of any representation of GL_n has non-negative integer coefficients.

Proof: The trace of $\tau(\mathbf{x})$ on V is the sum of its eigenvalues, counted with multiplicities. Since the weight spaces are by definition the eigenspaces, χ^V contains the term $\mathbf{x}^\lambda$ with coefficient equal to the dimension of the corresponding weight space V_λ.

Note that the Proposition implies that when $\lambda' = w(\lambda)$ is a permutation of λ, we have $\dim V_{\lambda'} = \dim V_\lambda$. This can also be seen directly, as the permutation matrix of w carries V_λ into $V_{\lambda'}$.

Since GL_n is an open subset of the affine space of all $n \times n$ matrices $M_n(\mathbb{C})$, its tangent space at 1 is just that of $M_n(\mathbb{C})$, and can itself be identified with $M_n(\mathbb{C})$. In this context it is usually denoted $\mathfrak{gl}_n$ rather than $M_n(\mathbb{C})$. The group structure of GL_n is reflected in a *Lie algebra* structure on $\mathfrak{gl}_n$, which will not concern us for now.

Given a polynomial representation

$$\rho\colon G \to GL(V),$$

since $\rho(1) = 1_V$, the differential of ρ is a linear map of tangent spaces

$$d\rho\colon \mathfrak{gl}_n \to \mathfrak{gl}(V).$$

We can compute it explicitly from its definition

$$d\rho(a) = \lim_{\epsilon\to 0} \frac{1}{\epsilon}(\rho(1+\epsilon a) - \rho(1)),$$

or from either of the equivalent and more convenient formulas

$$d\rho(a) = \frac{d}{dt}\rho(1+ta)|_{t=0} = \frac{d}{dt}\rho\exp(ta)|_{t=0}.$$

Lemma (A):

$$\exp(t\; d\rho(a)) = \rho(\exp(ta)).$$

Proof: Let $g_t = \exp(ta)$. Note that $g_t g_{t'} = g_{t+t'}$. Therefore,

$$\frac{d}{dt}\rho(g_t)|_{t=t_0} = \frac{d}{dt}\rho(g_{t_0})\rho(g_{t-t_0})|_{t=t_0} = \rho(g_{t_0})\frac{d}{dt}\rho(g_t)|_{t=0} = \rho(g_{t_0})d\rho(a).$$

This shows that $\rho(g_t)$ satisfies the first-order differential equation and initial condition

$$\frac{d}{dt}\rho(g_t) = \rho(g_t)d\rho(a); \quad \rho(g_0) = 1,$$

which characterize $\exp(t\; d\rho(a))$.

Lemma (B): If $v \in V_\lambda$ is a weight vector, then $d\rho(E_{ij})v \in V_{R_{ij}\lambda}$ is a weight vector of weight $R_{ij}\lambda$, where E_{ij} is the unit matrix with a 1 in row i and column j, and zeroes elsewhere.

Proof: We have

$$\begin{aligned}\rho(\tau(\mathbf{x}))d\rho(E_{ij})v &= \frac{d}{dt}\rho(\tau(\mathbf{x})\exp(tE_{ij}))v|_{t=0}\\ &= \frac{d}{dt}\rho(\exp(t\frac{x_i}{x_j}E_{ij})\tau(\mathbf{x}))v|_{t=0},\end{aligned}$$

since $\tau(\mathbf{x})E_{ij}\tau(\mathbf{x})^{-1} = (x_i/x_j)E_{ij}$. Since v has weight λ, the above is equal to

$$\frac{d}{dt}\rho(\exp(t\frac{x_i}{x_j}E_{ij}))\mathbf{x}^\lambda v|_{t=0} = \mathbf{x}^{R_{ij}\lambda}\frac{d}{dt}\rho(\exp(tE_{ij}))v|_{t=0} = \mathbf{x}^{R_{ij}\lambda}d\rho(E_{ij})v.$$

Corollary: If λ is a partition, maximal in dominance order among the weights of V, then every weight vector $v \in V_\lambda$ is N-invariant, *i.e.*, $V_\lambda \subseteq V^N$.

Proof: Matrices of the form $\exp(tE_{ij})$ generate N. By Lemma 2.4 (A), we have $\rho(\exp(tE_{ij}))v = \exp(td\rho(E_{ij}))v = v$, since $d\rho(E_{ij})v = 0$, as the weight space $V_{R_{ij}\lambda}$ is zero by maximality.

Let us remark that since T normalizes N, the space of N invariants V^N is a T submodule of V, and hence is itself the direct sum of its weight spaces V_λ^N.

Corollary: Every non-zero representation of G has $V^N \neq 0$. Hence $\dim V^N = 1$ implies V is irreducible.

Proof: The first part is clear, since there is some maximal non-zero weight space. If $V = W \oplus W'$ were a direct sum of non-zero submodules, then each would contribute to V^N, giving $\dim V^N > 1$.

2.5. The flag variety.

Definition: A (complete) *flag* F in $V = \mathbb{C}^n$ is a sequence of subspaces $0 \subseteq F_1 \subseteq F_2 \subseteq \cdots \subseteq F_{n-1} \subseteq V$, with $\dim F_i = i$.

Recall that the i-dimensional subspaces $W \subseteq V$ form a projective variety, the *Grassmann variety* $G_i(V)$. (Sometimes we will denote it instead by $G^{n-i}(V)$, when we want to view it as parametrizing the $n-i$ dimensional quotient spaces V/W of V.)

For $k < l$, the locus in $G_k \times G_l$ consisting of pairs (W_k, W_l) such that $W_k \subseteq W_l$ is a closed subvariety, *i.e.*, containment is a *closed condition.* It follows that the complete flags in V can be represented as a closed subvariety X of the product of Grassmann varieties,

$$X \subseteq G_1(V) \times \cdots \times G_{n-1}(V).$$

Now $G = GL(V)$ acts (algebraically) on X, with $g \in G$ mapping the flag F to the flag gF defined by $(gF)_i = g(F_i)$. Consider the *standard flag* E:

$$E_i = \{e_1, \ldots, e_i\},$$

where $e_1, \ldots, e_n$ is the usual basis of unit coordinate vectors in $V = \mathbb{C}^n$. We have $gE_i = E_i$ if and only if the first i columns of the matrix g have non-zero entries only in the first i rows, that is, iff g is block upper-triangular, with diagonal blocks of sizes $i \times i$ and $(n-i) \times (n-i)$. We have $gE = E$ iff this holds for all i, that is, the matrix g is upper-triangular. This shows that the stabilizer of E is B.

Given any flag F, we can choose a compatible basis $f_1, \ldots, f_n$ such that $\{f_1, \ldots, f_i\}$ spans F_i for all i, by simply taking f_i to be any element of F_i not in F_{i-1}. Then the element $g \in G$ carring e_i to f_i for all i has $gE = F$. Thus the G orbit of the standard flag E is the whole flag variety X. In particular, X is irreducible, since it is the image of the irreducible variety G by the morphism mapping g to gE.

Since G acts transitively on X and B is the stabilizer of the flag $E \in X$, we can identify X with the space of cosets G/B, where a coset gB corresponds to the flag gE, a correspondence that does not depend on the choice of coset representative. In this way G/B is given the structure of a projective variety.

(For any closed subgroup B of an algebraic group G, there is a canonical way to make G/B into an algebraic variety, but usually it will not be a projective variety. Note that G itself here is an affine, not a projective variety.)

We will have something to say later about the construction of homogeneous coordinates for G/B. For now, we work out local coordinates in a neighborhood of the standard flag E, or the coset $1B$. Recall the following result of matrix algebra.

Proposition: Every $n \times n$ matrix can be written in the form $LPDU$, where $U \in N$ is upper uni-triangular, $L \in N^-$ is lower uni-triangular, P is a permutation matrix, and D is a diagonal matrix.

In the language of G/B this says that every coset gB is contained in N^-wB for some permutation w, and thus every flag belongs to N^-wE for some w. Consider the morphism

$$\phi\colon N^- \to G/B; \quad \phi(g) = gE.$$

This morphism is obviously equivariant with respect to left multiplication by elements of N^-; in particular its fibers over all points of its image are isomorphic. We have $\phi^{-1}(\{E\}) = N^- \cap B = \{1\}$, so ϕ is injective, and therefore, since G/B is non-singular, ϕ is an isomorphism onto its image N^-E. Thus N^-E is an affine cell of dimension $\dim N^- = \binom{n}{2}$.

To make this explicit, observe that for $n \in N^-$, nE is the flag F with F_i spanned by the vectors $ne_1, \dots, ne_i$, that is, by the first i columns of the matrix

$$n = \begin{bmatrix} 1 & & \dots & 0 \\ a_{21} & 1 & & \vdots \\ a_{31} & a_{32} & 1 & \\ & & \dots & \\ a_{n1} & a_{n2} & \dots & 1 \end{bmatrix}.$$

The entries of this matrix thus serve as local coordinates on N^-E. Using this description, moreover, one can show (exercise) that a flag F belongs to N^-E if and only if, for all i, with $W_i = \mathbb{C}\{e_{i+1}, \dots, e_n\}$, we have $F_i + W_i = \mathbb{C}^n$, or equivalently, $F_i \cap W_i = 0$. Now for any fixed space W, the condition $F_i \cap W \neq 0$ is a closed condition on the Grassmannian $G_i(\mathbb{C}^n)$, so our conditions on F are open conditions, and N^-E is an *open* affine cell in the flag variety.

By symmetry, for any $g \in G$, the set gN^-E is an open affine neighborhood of gE, so this provides us with a covering of the flag variety by open affines, which are actually affine cells.

2.6. The Borel-Weil construction.

Let W be a rational representation of the Borel subgroup $B \subseteq G = GL_n$. We are going to describe a canonical procedure for constructing an algebraic vector bundle $G \times_B W$ on G/B, with an equivariant G action, whose fiber at the standard flag E (corresponding to the coset $1B \in G/B$) is W.

First we review terminology. A *vector bundle* of rank r over a space X is a space E, together with a map $\pi\colon E \to X$, and a structure of r-dimensional vector space on each fiber of π. In particular, $\mathbb{C}^r \times X$, with π the projection on X, is a vector bundle, called the *trivial* bundle over X. In general we require every vector bundle to be *locally trivial.* This means that every point of X has an open neighborhood U such that $\pi\colon \pi^{-1}(U) \to U$ is isomorphic to a trivial bundle. Here *isomorphism* has the meaning you would expect: $\pi\colon E \to X$ and $\pi'\colon E' \to X$ are isomorphic if there is a map $\alpha\colon E \to E'$ such that $\pi' \circ \alpha = \pi$, and α induces a linear isomorphism on each fiber.

The meaning of the term "vector bundle" depends on the category in which we work. Thus if we require that E should be a topological space, and that the map π and the trivializing isomorphisms should be continuous, we have defined *topological* vector bundles. For our purposes we will work with *algebraic* vector bundles. Thus X should be a complex algebraic variety, as should E, and we regard $\mathbb{C}^r$ as affine r space. The map π and the trivializing isomorphisms should be *morphisms*, that is, they have to be given by regular functions in local coordinates, and of course the trivializing isomorphisms should be isomorphisms of algebraic varieties.

I should point out that the concept of algebraic vector bundle is distinct from and more rigid than that of topological vector bundle. There exist topological vector bundles which cannot be given an algebraic structure, and topological isomorphisms between algebraically non-isomorphic vector bundles.

Definition: Given a B-module W, the space $G \times_B W$ is the orbit space $(G \times W)/B$, where B acts on $G \times W$ by $b(g, w) = (gb^{-1}, bw)$. There is a canonical map $\pi: G \times_B W \to G/B$ given by $(g, w) \mapsto gB$.

Proposition: For any $g \in G$, the map $\phi_g: W \to G \times_B W$ sending w to $B(g, w)$ is a bijection from W onto the fiber of π over the point $gB \in G/B$. Moreover, for any other $g' = gb \in gB$, the map $\phi_g^{-1} \circ \phi_{g'}: W \to W$ is $\rho(b)$, where ρ is the given representation of B on W.

Proof: Since $\pi(g, w) = gB$, ϕ_g does map W to the fiber over gB, and indeed surjectively onto it, since w is arbitrary. Now let $g' = gb$. I claim that for all w, $w' = b^{-1}w$ is the unique element such that $\phi_{g'}(w') = \phi_g(w)$. This establishes both that ϕ_g is injective (take $b = 1$) and the "moreover" part of the proposition.

For the claim, suppose $B(g', w') = B(g, w)$. The element b such that $g' = gb$ is unique, so the only representation of (g', w') as element of $B(g, w)$ is as $(gb, b^{-1}w)$. This says exactly that $w' = b^{-1}w$.

Now we assign the fibers of $\pi: G \times_B W \to G/B$ vector space structures via their bijections ϕ_g with W. By the Proposition, since $\phi_g^{-1} \circ \phi_{g'}$ is linear, this assignment is independent of the choice of g.

We want to verify this structure makes $G \times_B W$ into an algebraic vector bundle, that is, we have algebraic local trivializations. Since everything is G-equivariant it is enough to do this on the open set $N^- E \in G/B$. The space $\pi^{-1}(N^- E)$ consists of the orbits $B(n, w)$, where $n \in N^-$ and $w \in W$. Since ϕ is injective, the pairs (n, w) are a system of orbit representatives, *i.e.*, their orbits are distinct. This gives an isomorphism of $\phi^{-1}(N^- E)$ with $N^- \times W \cong N^- E \times W$, which preserves the linear structure we assigned to the factors. This shows that $G \times_B W$ is an algebraic vector bundle, and that it trivializes over $N^- E$ (and hence also over $gN^- E$ for every $g \in G$).

The Borel-Weil construction actually sets up an equivalence between B-modules and G-equivariant vector bundles over G/B, as given by the following proposition, whose proof I leave as an exercise.

Proposition: Let A be a G-equivariant algebraic vector bundle over G/B, that is, a vector bundle with an algebraic action of G that commutes with the map $\pi: A \to G/B$, and such that each g carries the fiber over a flag F linearly onto the fiber over gF. Then $A \cong G \times_B W$, where W is the fiber of A over the B-fixed point $E = 1B$, viewed as a B-module.

Exercise: Show that if W is originally a G-module, with the G action restricted to B, then $G \times_B W$ is isomorphic to the trivial bundle $(G/B) \times W$, with the obvious equivariant G action on the latter.

To avoid confusion we'll use a different notation for the algebraic vector bundle $G \times_B W$ itself (viewed as an algebraic variety), and its sheaf of sections, a locally free sheaf of $\mathcal{O}$-modules on G/B. The latter we denote by $E(W)$.

Proposition: The space of global sections $H^0(G/B, E(W))$ is a rational representation of G.

Proof: It's finite-dimensional because G/B is projective. Since $E(W)$ is an algebraic vector bundle, G acts algebraically on its space of sections, so it's a rational representation.

Let us remark that the higher sheaf cohomology groups of $E(W)$ are also rational G-modules, and that H^0 can perfectly well be zero, even if W is not.

2.7. Line bundles.

A rank-1 vector bundle is called a *line bundle*. By the correspondence in 2.7, the equivariant line bundles on G/B correspond to 1-dimensional representations of B.

Proposition: Every 1-dimensional rational representation of B is of the form $\mathbb{C}_\lambda$, where λ is a weight, and B acts on $\mathbb{C}_\lambda$ through the homomorphism $B \to T$ with kernel N, mapping each upper-triangular matrix to its diagonal part, and T acts by $\tau(\mathbf{x})v = \mathbf{x}^\lambda v$.

Proof: If W is 1-dimensional then it has only one weight space $W = W_\lambda$. Hence N acts trivially on W, by the corollary to Lemma 2.4(B), while T acts according to the weight, by definition. As B is the semi-direct product of T and N, the result follows.

For reasons that will appear shortly, we denote by L_λ the line bundle

$$L_\lambda = G \times_B \mathbb{C}_{w_0 \lambda},$$

where $w_0 \in S_n$ is the permutation reversing the indices: $w_0(i) = n + 1 - i$. The sheaf of sections of L_λ is the locally free sheaf

$$E(\mathbb{C}_{w_0 \lambda}).$$

Now let us recall how the Grassmann variety $G_k(V)$ is embedded in projective space. A point of $G_k(V)$ is a k-dimensional subspace $W \subseteq V$, and the k-th exterior power of the latter, $\bigwedge^k W$, is a 1-dimensional subspace of $\bigwedge^k V$. By definition the projective space $\mathbb{P}(\bigwedge^k V)$ is the variety $G_1(\bigwedge^k V)$ of such 1-dimensional subspaces, so the map $W \mapsto \bigwedge^k W$ defines a map $G_k(V) \mapsto \mathbb{P}(\bigwedge^k V)$. This is the *Plücker embedding*.

In concrete terms the Plücker embedding assigns W the homogeneous coordinates given by the $k \times k$ minors of any $k \times n$ matrix with row space W (the ratios of these minors being independent of the choice of basis for W). To see this note that, if $e_1, \ldots, e_n$ is a basis of V, and $v_1, \ldots, v_k$ is a basis of W, then the coordinates of $v_1 \wedge \cdots \wedge v_k$ in terms of the basis $\{e_{i_1} \wedge \cdots \wedge e_{i_k} : i_1 < \cdots < i_k\}$ are given precisely by the minors of the matrix with rows v_i.

Recall that the standard ample sheaf $\mathcal{O}(1)$ on projective space $\mathbb{P}(U)$ is *dual* to the sheaf of sections of the tautological line bundle L, whose fiber over a 1-dimensional subspace $S \in \mathbb{P}(U)$, $S \subseteq U$, is S itself. Each linear coordinate function x on U gives rise by restriction to a linear functional on each fiber S, and so to a global section of the dual bundle to L, or a global section of $\mathcal{O}(1)$. By multiplication, polynomials homogeneous of degree d in the coordinates on U give rise to global sections of the d-th tensor power $\mathcal{O}(d)$ of $\mathcal{O}(1)$. This gives the ring homomorphism, which is actually an isomorphism, identifying the coordinate ring of U, which is also the homogeneous coordinate ring of $\mathbb{P}(U)$, with the graded ring of global sections

$$\bigoplus_{d\geq 0} H^0(\mathbb{P}(U), \mathcal{O}(d)).$$

In particular, the sheaves $\mathcal{O}(d)$ have plenty of non-zero global sections—enough that their vanishing loci can distinguish between any distinct closed subschemes of $\mathbb{P}(U)$, whence the term "ample."

Restricting $\mathcal{O}(1)$ from $\mathbb{P}(\bigwedge^k(V))$ to $G_k(V)$, as embedded by the Plücker embedding, it becomes the dual to the sheaf of sections of the highest exterior power $\bigwedge^k W$ of the tautological bundle W on $G_k(V)$, whose fiber at a point W is W itself. We can summarize this as follows:

Proposition: The ample bundle $\mathcal{O}(1)$ induced on $G_k(V)$ by the Plücker embedding is the dual of the k-th exterior power of the tautological bundle.

For our purposes, it will be more natural to formulate this in terms of $G^d(V) = G_k(V)$, where $d = n-k$. The relevant bundle on $G^d(V)$ is the tautological *quotient bundle* with fiber V/W at W, a rank d vector bundle. Exterior multiplication $\bigwedge^k W \otimes \bigwedge^d(V/W) \to \bigwedge^n V \cong \mathbb{C}$ identifies $\bigwedge^d(V/W)$ with the dual space of $\bigwedge^k W$. This identification is canonical except for the choice of a non-zero basis vector in $\bigwedge^n V$, which does not depend on W. Thus $\bigwedge^d(V/W)$ is isomorphic to the dual bundle of $\bigwedge^k W$.

Corollary: The ample bundle $\mathcal{O}(1)$ induced on $G^d(V)$ by the Plücker embedding is isomorphic to the d-th exterior power of the tautological quotient bundle.

Corollary: Let Q_d be the tautological quotient bundle on the flag variety G/B whose fiber at a flag F is V/F_{n-d}. Then its highest exterior power $M_d = \bigwedge^d Q_d$ is isomorphic to $\mathcal{O}(1)$, pulled back from the Grassmann variety G^d, through the embedding of G/B in the product of Grassmann varieties. In particular, M_d has non-zero global sections, and any tensor product $M_1^{e_1} \otimes \cdots \otimes M_{n-1}^{e_{n-1}}$ with all $e_d > 0$ is very ample on G/B, *i.e.*, it is the restriction of $\mathcal{O}(1)$ for some projective embedding of G/B.

Note that we could have also constructed M_d as the highest exterior power of the dual of the tautological subspace bundle S_{n-d} with fiber F_{n-d} at F. This gives the same line bundle with a different G action, because G does not act trivially on $\bigwedge^n V$. *Exercise:* show that $\bigwedge^{n-d}(S_{n-d})^*$ is equivariantly isomorphic to $(\det)^{-1} \otimes M_d$, where det is the trivial line bundle $G/B \times \mathbb{C}$, with G acting by the determinant representation on $\mathbb{C}$, or $\det = L_{(-1,-1,\ldots,-1)}$ in our notation for equivariant line bundles.

As Q_d is clearly G-equivariant, it must be $L_\lambda = G \times_B \mathbb{C}_{w_0\lambda}$ for some weight λ. To find this weight, we need only inspect the B action (or just the T action) on the fiber of Q_d at the standard flag, corresponding to the coset $1B$. This fiber is $\bigwedge^d(V/E_{n-d})$. It is spanned by the single vector $e_{n-d+1} \wedge \cdots \wedge e_n$, on which T acts with weight $w_0\lambda = (0, \ldots, 0, 1, \ldots, 1)$, with d ones and $n-d$ zeroes. Therefore we have the following result.

Proposition: The G-equivariant line bundle $\bigwedge^d Q_d$ on G/B is isomorphic to L_{λ_d} where $\lambda_d = \varepsilon_1 + \cdots + \varepsilon_d = (1, \ldots, 1, 0, \ldots, 0)$ corresponds to the partition (1^d).

We remark that in a homogeneous polynomial representation of degree d of G, the scalar matrices tI must act as scalars $t^d I$. Equivalently, every weight space V_λ has $|\lambda| = d$. Similar considerations apply to B and T. The defining representation V of G is clearly homogeneous of degree 1. The equivariant bundle Q_d corresponds to the B-module V/E_{n-d}—note that E_{n-d} is a B-submodule of V, so this makes sense. This is again homogeneous of degree 1, and its d-th exterior power is homogeneous of degree d. Thus we should expect, as we have just seen, that λ for the corresponding line bundle L_λ should be a partition of d.

Corollary: For every partition λ, the space of global sections $H^0(G/B, L_\lambda)$ is non-zero.
Proof: Every such L_λ is a tensor product $L_{\lambda_1}^{\mu_1} \otimes \cdots \otimes L_{\lambda_n}^{\mu_n}$, where in fact μ is the conjugate partition λ'. Since $L_{\lambda_d} = M_d$, it has a non-zero global section.

We have seen that $L_{(1,\ldots,1)}$ is the determinant representation of G on the trivial bundle. By tensoring with $L_{(1,\ldots,1)}^k$, for any integer k, we see that the Corollary also applies for λ of the form partition plus $(k, \ldots, k)$, that is, for λ any non-increasing integer sequence. Shifting λ by $(k, \ldots, k)$ has the effect of tensoring $H^0(G/B, L_\lambda)$ by the k-th power of the determinant representation.

2.8. Irreducibile representations.

Theorem: The representation of G on $H^0(G/B, L_\lambda)$ is irreducible whenever it is non-zero.
Proof: By Corollary 2.4, there is an N-invariant section in $H^0(G/B, L_\lambda)$. Equally well, by symmetry, there is an N^--invariant section σ. Then σ is determined on all of N^-E by its value in the fiber over the standard flag E. Since N^-E is open, it is dense, and thus σ is determined completely by its value in the fiber over E. In particular, since this fiber is 1-dimensional, any two N^--invariant sections in $H^0(G/B, L_\lambda)$ are scalar multiples of each other. Again by Corollary 2.4 (with N^- in place of N) this implies that $H^0(G/B, L_\lambda)$ is irreducible.

Now let V be an arbitrary irreducible rational representation of G and let λ be a maximal weight, that is, λ is a partition plus $(k, \ldots, k)$ for some integer k, and the partition in question is maximal in dominance order. Then as we have seen, V contains a non-zero N-invariant weight vector $v \in V_\lambda$. Let $W = \mathbb{C}v$ be the span of v. Then W is a 1-dimensional B-submodle of V, since N fixes v, and v is a weight vector for T. Hence we can construct the line bundle $L_{w_0\lambda}$ on G/B as $G \times_B W$.
There is a morphism ϕ of algebraic varieties from $G \times_B W$ to V given by

$$B(g, tv) \mapsto tgv.$$

This is well-defined since a general element of $B(g, v)$ has the form (gb^{-1}, bv). Its image is the union of $\{0\}$ and the G orbit of v in V.
The map ϕ is linear in the scalar t. Hence if $\alpha \in V^*$ is a linear functional on V, then the composite $\alpha \circ \phi$ is linear on each fiber of the line bundle $G \times_B W$, that is, it represents a global section of the dual bundle $L_{-w_0\lambda}$.
In particular, by taking a basis of weight vectors including v in V, and its dual basis (also of weight vectors) in V^*, we can choose α of weight $-\lambda$, with $\alpha(v) \neq 0$. Then obviously the section $\sigma = \alpha \circ \phi \in H^0(G/B, L_{-w_0\lambda})$ is non-zero. This shows that the map from V^* to $H^0(G/B, L_{-w_0\lambda})$ sending α to $\alpha \circ \phi$ is non-zero, and it is clearly linear and G-equivariant. Since V^* and $H^0(G/B, L_{-w_0\lambda})$ are both irreducible, this map is an isomorphism, by Schur's

Lemma. Furthermore, $-w_0\lambda$ is a maximal weight (indeed, the unique maximal weight, by irreducibility and Corollary 2.4) of V^*. Of course, since V was arbitrary, so is V^*, so we have proved the following fundamental theorem of Borel and Weil.

Theorem: Every irreducible rational representation V of GL_n is isomorphic to $H^0(G/B, L_\lambda)$, where λ is the unique maximal weight of V.

Corollary: The irreducible representations V^λ with highest weight λ are non-isomorphic for distinct λ.

Proof: Immediate from the uniqueness of λ.

Corollary: A rational G module V is irreducible if and only if $\dim V^N = 1$.

Proof: We have already seen the "if". For the only if, use the Borel-Weil theorem and the fact shown in the proof of the previous theorem, that $H^0(G/B, L_\lambda)$ has a unique N-invariant vector, up to a scalar multiple.