Introduction to GL_n representation theory

These notes are from the course Math 668, "Introduction to GL_n representation theory", taught by David E Speyer at the University of Michigan in Fall 2022. They were written by the students, Heitor Anginski Cotosky, Jineon Baek, Joao Pedro Carvalho, Yiwei Fu, Andrew Paul Keisling, Paul Chalakuzhy Mammen, Ryuichi Man,Teo D Miklethun, Scott Neville, Urshita Pal, Amanda Catherine Schwartz, Dawei Shen and Katie Waddle, and edited by David E Speyer.

Contents

August 29: Introduction

For any group G, a representation is a vector space V and a map of groups $\rho : G \to GL(V)$. We are interested in the case where G is also $GL_n(\mathbb{C})$, so we are looking at maps $GL_n(\mathbb{C}) \to GL_N(\mathbb{C})$. I expect this to lead us to

(1) Symmetric functions
(2) Schur polynomials and schur functors
(3) Young tableaux
(4) Jeu de taquin, RSK
(5) Crystals

We might or might not get to

(6) Standard monomial theory
(7) Webs
(8) Representation theory of S_n

There are many other great topics I don't expect to reach:

(9) Infinite dimensional representations
(10) Representation theory of $\mathrm{GL}_n(\mathbb{F}_p)$ over $\mathbb{C}$
(11) Representation theory of $\mathrm{GL}_n(k)$ over k where k has characteristic p
(12) Total positivity
(13) Cluster algebras
(14) Other Lie types

But, the good news is, you will be doing final papers/presentations! So there are opportunities to learn about all these things!

For now, I want to explain what sort of representation we will be talking about, and I want to explain why we will be seeing so many symmetric polynomials in this course.

Write $\mathbb{C}^n$ for the standard n-dimensional representation of $GL_n(\mathbb{C})$.

Some examples of representations we want to study:

(1) The representation $\mathbb{C}^n$, or $(\mathbb{C}^n)^{\oplus r}$ for any $r \geq 0$.
(2) The representations $(\mathbb{C}^n)^{\otimes k}$ and its subrepresentations $\mathrm{Sym}^k(\mathbb{C}^n)$ and $\bigwedge^k(\mathbb{C}^n)$.
(3) The dual representation $(\mathbb{C}^n)^\vee$. In coordinates, this is $g \mapsto g^{-T}$.
(4) We can take $V = \mathbb{C}$ and $\rho(g) = (\det g)^k$ for any integer k.

Here are some homomorphisms $\rho : \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}_N(\mathbb{C})$ we do **not** want to study:

(5) We could map g to $|\det g|^a$ for some real number a.
(6) We could map g to $\overline{g}$, meaning to take the complex conjugate of each entry of g.
(7) We could map g to $\sigma(g)$, where σ is some crazy element of $\mathrm{Gal}(\mathbb{C}/\mathbb{Q})$. (If you believe in the axiom of choice, you believe that there is an automorphism of $\mathbb{C}$ taking $\sqrt{2}$ to $-\sqrt{2}$.)

Let $\rho : \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}_N(\mathbb{C})$ be a group homomorphism. Write g_{ij}, for $1 \leq i, j \leq n$, for the coordinates on $\mathrm{GL}_n(\mathbb{C})$. For $1 \leq p, q \leq N$, the coordinate $\rho(g)_{pq}$ is a function of the g_{ij}. We'll say that ρ is

- a ***polynomial representation*** if each $\rho(g)_{pq}$ is a polynomial in the g_{ij}
- an ***algebraic representation*** if each $\rho(g)_{pq}$ is in $\mathbb{C}[g_{ij}, (\det g)^{-1}]$.

It is easy to check that this is independent of the choice of basis for $\mathbb{C}^N$.

This makes sense from the perspective of algebraic geometry: The polynomial functions $\mathbb{C}[g_{ij}]$ are the regular functions on the n^2-dimensional affine space of $n \times n$ matrices, and the functions in $\mathbb{C}[g_{ij}, (\det g)^{-1}]$ are the regular functions on the locus where $\det g \neq 0$, which is $\mathrm{GL}_n(\mathbb{C})$. In the above list, representations (1) and (2) are polynomial. (In the natural basis, the matrix entries of the action on $(\mathbb{C}^n)^{\otimes k}$ are products of k-entries of g; the matrix entries of the action on $\bigwedge^k(\mathbb{C}^n)$ are $k \times k$ minors of g; the matrix entries of the action on $\mathrm{Sym}^k(\mathbb{C}^n)$ are $k \times k$ permanents of g.) The representation (3) is algebraic but not polynomial. The representation (4) is polynomial if $k \geq 0$, and is algebraic but not polynomial for $k < 0$.

Now, where do the symmetric polynomials come from?

If $\rho : G \to GL(V)$ is an finite dimensional representation, then the character χ of ρ is defined by $\chi(g) = Tr(\rho(g))$. Note that

$$\chi(hgh^{-1}) = Tr(\rho(hgh^{-1})) = Tr(\rho(h)\rho(g)\rho(h)^{-1}) = Tr(\rho(h)^{-1}\rho(h)\rho(g)) = Tr(\rho(g)) = \chi(g)$$

so χ is constant on conjugacy classes. Now, the diagonalizable matrices are dense in $GL_n(\mathbb{C})$, so any continuous function on $GL_n(\mathbb{C})$ is determined by its values on diagonalizable matrices. And "diagonalizable" just means "conjugate to diagonal". So characters of continuous representations are determined by their values on diagonal matrices.

If ρ is polynomial, then $\chi\begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix}$ is a polynomial in $x_1, x_2, \ldots, x_n$. If ρ is rational, then it is a Laurent polynomial.

Moreover, note that

$$\begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix} \begin{bmatrix} x & & \\ & y & \\ & & z \end{bmatrix} \begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix}^{-1} = \begin{bmatrix} y & & \\ & z & \\ & & x \end{bmatrix}$$

and, more generally, conjugating a diagonal matrix by a permutation matrix permutes its entries. So $\chi\left(\begin{bmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{bmatrix}\right)$ is a *symmetric* (Laurent) polynomial in $x_1, x_2, \ldots, x_n$. In this way, we get a symmetric (Laurent) polynomial from each polynomial (rational) representation.

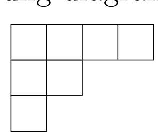
We introduce the notation Λ_n for the ring of symmetric polynomials in $x_1, x_2, \ldots, x_n$, with integer coefficients. We'll write $\Lambda_n^{\pm}$ for symmetric Laurent polynomials. Soon we will also introduce Λ and $\Lambda^{\pm}$, which conceptually are the limits of Λ_n and $\Lambda_n^{\pm}$ as $n \to \infty$.

August 31: The elementary and monomial symmetric functions

Recall that we used the notations Λ_n to denote symmetric polynomials in $x_1, x_2, ..., x_n$, and $\Lambda_n^{\pm}$ to denote symmetric Laurent polynomials in $x_1, x_2, .., x_n$.

Partitions. A ***partition*** is a weakly decreasing sequence of positive integers $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_l \geq 0$.

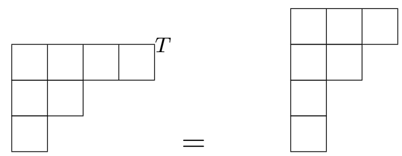
We draw partitions as configurations of boxes called ***Young Diagrams*** - the partition $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_l \geq 0$ is represented as l rows of boxes, with the ith row having λ_i boxes, aligned to the left. For example, the Young diagram for the partition $(4, 2, 1)$ is as follows:



We will often pad our partitions with 0's - for example, $(4, 2, 1), (4, 2, 1, 0), (4, 2, 1, 0, 0)$ all represent the same partition.

For a partition λ, we say the ***size*** of λ is $|\lambda| := \sum_j \lambda_j$ = number of boxes. We say the ***length*** of λ is $\ell(\lambda) := \#\{j : \lambda_j > 0\}$ = number of rows.

The ***transpose*** of λ, denoted λ^T, is obtained by reflecting the Young diagram for λ along its diagonal. For example, $(4, 2, 1)^T = (3, 2, 1, 1)$, because



Another way to describe transposes is to note that ${\lambda^T}_j = \#\{i : \lambda_i \geq j\}$.

We define a partial order, a ***dominance order*** on partitions, as follows: We say $\lambda \succeq \mu$ if

$$\begin{array}{rcl} \lambda_1 & \geq & \mu_1 \\ \lambda_1 + \lambda_2 & \geq & \mu_1 + \mu_2 \\ \lambda_1 + \lambda_2 + \lambda_3 & \geq & \mu_1 + \mu_2 + \mu_3 \\ & & \text{etc.} \end{array}$$

Now we shall look at a few specific types of symmetric polynomials.

Monomial Symmetric Functions. For a partition λ, define m_λ to be the symmetric polynomial defined my

$$m_\lambda(x_1, ..., x_n) = \sum_{(c_1,...,c_n) \in S_n(\lambda_1,...,\lambda_n)} x_1^{c_1} x_2^{c_2} ... x_n^{c_n}.$$

Example. Let $n = 3$. Then $m_{(3,1)}(x_1, x_2, x_3) = x_1^3 x_2 + x_1 x_2^3 + x_1^3 x_3 + x_1 x_3^3 + x_2^3 x_3 + x_2 x_3^3$. Note the padding of $(3, 1)$ here to $(3, 1, 0)$.

Example. Let $n = 3$ again. Then $m_{(1,1)}(x_1, x_2, x_3) = x_1 x_2 + x_2 x_3 + x_3 x_1 \neq 2(x_1 x_2 + x_2 x_3 + x_3 x_1)$. So we do <u>not</u> count the terms with multiplicity. Thus the coefficients of all terms in m_λ are 0 or 1.

Note that the m_λ form a $\mathbb{Z}$-basis for Λ_n, i.e. $\Lambda_n = \bigoplus_{\ell(\lambda) \leq n} \mathbb{Z} \cdot m_\lambda$.

Elementary Symmetric Functions. For a positive integer k let $e_k \in \Lambda_n$ be given by $e_k(x_1, x_2, ..., x_n) = \sum_{i_1 < i_2 < ... < i_k} x_{i_1} x_{i_2} ... x_{i_k} = m_{(1,1,...,1)}$. For a partition $\lambda = (\lambda_1, \lambda_2, ..., \lambda_l)$, we put $e_\lambda = e_{\lambda_1} e_{\lambda_2} ... e_{\lambda_l}$.

Example. We have

$$e_{(3,1)}(w, x, y, z) = e_3 \cdot e_1 = (wxy + xyz + yzw + zwx)(w + x + y + z) =$$
$$(w^2xy + w^2yz + \cdots) + 4wxyz = m_{211} + 4m_{1111}.$$

Note that $(3, 1)^T = (2, 1, 1)$, and that the coefficient of m_{211} in e_{31} is 1. This is an example of the following lemma:

Lemma. e_λ is of the form m_{λ^T} + (a linear combination of m_μ with $\mu \prec \lambda^T$)

Proof. Suppose m_μ occurs with a positive coefficient. Then $x_1^{\mu_1}x_2^{\mu_2}...x_n^{\mu_n}$ occurs in the expansion of $e_{\lambda_1}e_{\lambda_2}...e_{\lambda_l}$. We need to show that $\mu \preceq \lambda$.

Note that each e_{λ_i} contributes a power of atmost 1 for each x_j. Thus $\mu_1 \leq \#\{i : \lambda_i \geq 1\} = (\lambda^T)_1$.

By a similar reasoning, we have $\mu_1 + \mu_2 \leq 2\#\{i : \lambda \leq 2\} + \#\{i : \lambda_i = 1\}$ (each e_{λ_i} with $\lambda_i \geq 2$ can at best have contributed one each of x_1 and x_2, and the rest will have come from e_{λ_i} with $\lambda_i = 1$). Thus $\mu_1 + \mu_2 \leq \#\{i : \lambda_i \geq 2\} + \{i : \lambda_i \geq 1\} = (\lambda^T)_1 + (\lambda^T)_2$.

In general, $\mu_1 + \mu_2 + \cdots + \mu_j \leq j\#\{i : \lambda_i \geq j\} + \sum_{k=1}^{j-1} k\#\{i : \lambda_i = k\} = \sum_{k=1}^{j} \#\{i : \lambda_i \geq k\} = \sum_{k=1}^{j} (\lambda^T)_k$. Thus $\mu \preceq \lambda^T$.

We can trace back this proof to see when equality holds - for equality to hold, the contribution from each e_{λ_i} must precisely be $x_1x_2...x_{\lambda_i}$. So the coefficient of m_{λ^T} is 1. $\square$

Corollary. $\Lambda_n = \bigoplus_{\ell(\lambda^T)\leq n} \mathbb{Z} \cdot e_\lambda = \bigoplus_{\ell(\lambda)\leq n} \mathbb{Z} \cdot e_{\lambda^T}$.

Proof. Fix a degree d. The previous lemma implies that the e_λ with $|\lambda| = d$ are related to m_μ with $|\mu| = d$ by an upper triangular matrix. The diagonal entries of this matrix are all 1's, and thus it is invertible. Inverting this matrix allows us to express the m_μ in terms of e_λ and since we know the m_μ give a basis for Λ_n, the same must be true of the e_{λ_n}. $\square$

Example. Here is the matrix expressing the e's in terms of the m's in degree 3:

$$\begin{array}{ccccccc} e_3 & = & & & & & m_{111} \\ e_{21} & = & & & m_{21} & + & 3m_{111} \\ e_{111} & = & m_3 & + & 3m_{21} & + & 6m_{111}. \end{array}$$

Note that the matrix is "triangular". If we redefined e_λ to be what we are currently calling e_{λ^T}, then it would be upper triangular.

Corollary. $\Lambda_n = \mathbb{Z}[e_1, e_2, ..., e_n]$

Proof. Note that any monomial in $e_1, e_2, ..., e_n$ is precisely equal to e_λ for some λ with all parts of λ being $\leq n$, i.e. $\ell(\lambda^T) \leq n$. The result then follows using the previous Corollary. $\square$

This result is often referred to as the 'Fundamental Theorem of Symmetric Polynomials'.

The ring Λ. We have ring maps $\Lambda_n \to \Lambda_{n-1}$ given by $x_n \mapsto 0, x_k \mapsto x_k$ for $k \leq n-1$. This maps acts on m_λ by: $m_\lambda \mapsto m_\lambda$ if $\ell(\lambda) \leq n-1$, and $m_\lambda \mapsto 0$ otherwise. The map acts on e_λ by : $e_\lambda \mapsto e_\lambda$ if $\ell(\lambda^T) \leq n-1$, and $e_\lambda \mapsto 0$ otherwise.

This allows us to define the ring $\Lambda = \mathbb{Z}[e_1, e_2, ..., e_n, ...] = \bigoplus_\lambda \mathbb{Z} \cdot e_\lambda = \bigoplus_\lambda \mathbb{Z} \cdot m_\lambda$. In terms of category theory, $\Lambda = \lim_{\infty \leftarrow n} \Lambda_n$, the 'graded inverse limit' of Λ_n.

September 2: The ring Λ and the polynomials h_λ

Today we finished discussing the ring Λ and introduced a third basis of the symmetric polynomials h_λ.

When doing manipulations in Λ_n for various n, many identities appear regardless of the particular n. For example, in Λ_n for $n \geq 3$ we have:

$$m_1^2 = m_2 + 2m_{11}$$

$$m_1^2 = m_3 + 3m_{21} + 6m_{111}$$

$$e_{31} = m_{211} + 4m_{1111}.$$

Since there is so much shared structure, it would be nice to have a single ring with all these identities. This motivates Λ, the graded ring of symmetric functions in infinitely many variables, with maps from $\Lambda \to \Lambda_n$ which commute with the natural maps, $\Lambda_n \to \Lambda_{n-1}$ sending $x_n \mapsto 0$ and leaving all else fixed.

The ring Λ. There are two broad approaches to defining something nice, we can give a concrete construction (but then we must work to check that what we've made really does have all the properties) or we can lean on abstract nonsense (in which case we may know very little about what we've made, but it definitely exists and has those properties). We give two concrete constructions, and then an abstract one for those who know a little category theory. Before we begin, we introduce some notation. The degree d piece of R will be denoted R^d. Note that R^d is not a ring, but an abelian group. Thus Λ_n^d denotes the degree d symmetric polynomials in n variables.

First, we can take $\Lambda = \mathbb{Z}[e_1, e_2, \ldots, e_n, \ldots]$, the ring generated by all the e_j. We then define the map $\Lambda \to \Lambda_n$ by sending $e_i \mapsto \begin{cases} e_i & i \leq n \\ 0 & i > n \end{cases}$. Note that these maps are isomorphisms between Λ^d and Λ_n^d for $n \geq d$. It is easy to see that Λ with these maps is compatible with the ring maps $\Lambda_n \to \Lambda_{n-1}$, though it requires some work to find the m_λ's in this construction.

Second, we can first construct an abelian group $\Lambda = \bigoplus_\lambda \mathbb{Z}e_\lambda = \bigoplus_\lambda \mathbb{Z}m_\lambda$, and then define $e_\lambda e_\mu$ (or $m_\lambda m_\mu$) by whatever the expression equals in Λ_N for $N > |\lambda| + |\mu|$ (so that N is large enough that we would see any structure), so that the maps $\Lambda^d \to \Lambda_N^d$ sending $e_\lambda \mapsto e_\lambda$ is an isomorphism between suitably large graded pieces.

Finally, for those who know category theory. The maps

$$\cdots \to \Lambda_3 \to \Lambda_2 \to \Lambda_1$$

form an infinite descending chain in the category of rings. We say the ring $Y = \lim_{\infty \leftarrow n} \Lambda_n$, with maps $\prod_n : Y \to \Lambda_n$ compatible with the chain (ie, applying $\prod_n$ and then moving down the chain is the same as applying $\prod_{n-1}$), is the inverse-limit of this chain if for all Z with compatible maps $\alpha_n : Z \to \Lambda_n$ there is a unique map $f : Z \to Y$ such that $\prod_n \circ f = \alpha_n$ for all n.

Remark. If we restrict our chain to the degree d part, then the chain is eventually constant:

$$\cdots = \Lambda_{d+1}^d = \Lambda_d^d \to \Lambda_{d-1}^d \cdots \to \Lambda_2^d \to \Lambda_1^d.$$

Therefore, in order for our maps $\prod_n$ to be compatible, they must be eventually constant on each graded piece Λ^d.

The polynomials h_λ. The h_λ symmetric polynomials[1] are defined similarly to the e_λ: we first define h_k for $k > 0$, and then set $h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_n}$. The definition of h_k may look familiar:

$$h_k = \sum_{i_1 \leq i_2 \cdots \leq i_k} x_{i_1} \cdots x_{i_k}$$

note that we allow $i_j = i_{j+1}$, unlike in the definition of e_k.

Remark. It will turn out that e_k is the character of $\bigwedge^k(\mathbb{C}^n)$, and h_k is the character of $\mathrm{Sym}^k(\mathbb{C}^n)$. See Homework 1, problem 3.

[1] h stands for "homogeneous", but this is not very helpful. All of the other polynomials we have seen are homogeneous too. Perhaps 'huge', for typically having more terms in h_λ than m_λ or e_λ?

Let us compute a few examples:

$$h_2 = \sum_{i\le j} x_i x_j = \sum_i x_i^2 + \sum_{i<j} x_i x_j = m_2 + m_{1,1}$$

$$h_{1,1} = h_1^2 = (\sum_i x_i)(\sum_j x_j) = m_2 + 2m_{1,1}$$

So we have

$$\begin{bmatrix} h_2 \\ h_{1,1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} m_2 \\ m_{1,1} \end{bmatrix}$$

in particular, the h_λ's are not given by upper triangular combinations of the m_λ's. Still, the matrix above has determinant 1, and so $h_{1,1}, h_2$ are a $\mathbb{Z}$-basis. We will now show they are a basis in general.

Theorem. We have $\Lambda = \mathbb{Z}[h_1, h_2, \ldots]$.

Since the h_i are symmetric polynomials and the e_i form a basis, it is clear that $\mathbb{Z}[h_1, h_2, \ldots]$ is contained in Λ. Our plan is to prove that each e_i is a polynomial in the h_j's, and so $\mathbb{Z}[h_1, h_2, \ldots]$ is exactly Λ. We give the proof after a lemma.

Lemma. For all $k > 0$ we have $\sum_{j=0}^k (-1)^j h_j e_{k-j} = 0$. In particular,

$$e_k = e_k h_0 = \sum_{j=1}^k (-1)^{j+1} h_j e_{k-j}.$$

Proof. Consider the generating functions $e(t) = \sum_j e_j t^j$ and $h(t) = \sum_j h_j t^j$. We can rewrite them as infinite products $e(t) = \prod_{i=1}^\infty (1 + x_i t)$ (the degree j term is gained by picking $x_i t$ from j different i, so we get all combinations of j distinct variables) and

$$h(t) = \prod_{i=1}^\infty (1 + x_i t + (x_i t)^2 + (x_i t)^3 + \cdots) = \prod_{i=1}^\infty \frac{1}{1 - x_i t}$$

(since we now want monomials with arbitrary degree in x_i, but still a total degree of j).

We then see that

$$1 = \prod_{i=1}^\infty \frac{1 + x_i t}{1 + x_i t} = e(t)h(-t) = (\sum_i e_i t^i)(\sum_j (-1)^j h_j t^j)$$

So after expanding the right hand side and collecting the terms with t^k, we find

$$0 = \sum_{j=0}^k (-1)^j h_j e_{k-j}.$$

Solving this equation for e_k finishes the proof. □

Corollary. For $k > 0$, we have e_k is a polynomial in $h_1, \ldots, h_k$.

Proof. We argue by induction on k. If $k = 1$ then $e_1 = h_1$. By Lemma we have

$$e_k = \sum_{j=1}^k (-1)^{j+1} h_j e_{k-j}.$$

Since $k - j < k$, by induction each e_{k-j} is a polynomial in the h_i for $i \le k - j < k$. Therefore e_k is a polynomial in the h_i for $i \le k$. □

Proof of Theorem. Let L be the ring generated by h_i. It is clear that L is contained in Λ, since the e_i generate Λ_n. Corollary implies that $e_k \in L$, thus $L = \Lambda$. Finally, since there are the same number of e_λ and h_λ with $\ell(\lambda) < n$, and e_λ is a basis for Λ_n, the h_λ must be linearly independent. Thus $\Lambda = \mathbb{Z}[h_1, h_2, \ldots]$ as desired. $\square$

Finally, we discuss what happens in Λ_n. The proof above still shows that the h_j generate Λ_n. More precisely, since Λ_n is generated by e_1, e_2, $\ldots$, e_n, the proof above shows that $\Lambda_n = \mathbb{Z}[h_1, h_2, \ldots, h_n]$. In terms of a $\mathbb{Z}$-basis, h_λ is a monomial in $\{h_1, h_2, \ldots, h_n\}$ if and only every every part of λ is $\leq n$ or, in other words, if $\ell(\lambda^T) \leq n$. So we deduce that $\Lambda_n = \mathbb{Z}[h_1, \ldots, h_n] = \bigoplus_{\ell(\lambda^T)\leq n} \mathbb{Z} h_\lambda$.

September 7: The Hall inner product

Motivation. Our objective for this class is to motivate and define a useful bilinear, symmetric, positive definite map $\langle\ ,\ \rangle : \Lambda \times \Lambda \to \mathbb{Z}$. As we'll see, we can also define $\langle\ ,\ \rangle$ on $\Lambda_n \times \Lambda_n$ or on $\Lambda_n^\pm \times \Lambda_n^\pm$.

Definition. Let G be a group and V, W finite dimensional G-representations. We say that a linear map f is ***G-equivariant*** if $f(g * v) = g * f(v)$ forall $v \in V, g \in G$. We denote the set of all such maps as $\mathrm{Hom}^G(V, W)$

The objective of the Hall inner product is to give information about $\mathrm{Hom}^G(V, W)$ when $G = GL_n\mathbb{C}$ in the following manner:
Let V, W be algebraic representations of $\mathrm{GL}_n\mathbb{C}$. As we saw in the first class, the characters χ_V, χ_W are uniquely defined by their values on diagonal matrices. So we can write:

$$\chi_v(x_1, \ldots, x_n) = Tr\rho_V \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_m \end{bmatrix} \qquad \chi_W(x_1, \ldots, x_n) = Tr\rho_W \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_n \end{bmatrix}$$

We know that $\chi_V, \chi_W \in \Lambda_n^\pm$ then we want $\langle\ ,\ \rangle$ to satisfy $\dim_{\mathbb{C}} \mathrm{Hom}^{\mathrm{GL}_n\mathbb{C}}(V, W) = \langle \chi_V, \chi_W \rangle$. In other words, the inner product will allow us to recover $\dim_{\mathbb{C}} \mathrm{Hom}^{\mathrm{GL}_n\mathbb{C}}(V, W)$ only from the characters.

Example. Let $V = W = (\mathbb{C}^n)^{\otimes 2}$, with $\mathrm{GL}_n(\mathbb{C}) = G$ acting on them on the usual manner. Clearly the identity $\mathrm{Id} : (\mathbb{C}^n)^{\otimes 2} \to (\mathbb{C}^n)^{\otimes 2}$ is $\mathrm{GL}_n(\mathbb{C})$ equivariant. Another one will be $\sigma : v_1 \otimes v_2 \mapsto v_2 \otimes v_1$. We have that, in fact, $\mathrm{Hom}^{\mathrm{GL}_n\mathbb{C}}(V, W) = \mathbb{C} \cdot \mathrm{Id} \oplus \mathbb{C} \cdot \sigma$ (for $n \geq 2$). On the other hand, we can use a result from the first problem set to calculate the character:

$$\chi_{\mathbb{C}^n \otimes \mathbb{C}^n}(x_1, \ldots, x_n) = \chi_{\mathbb{C}^n}(x_1, \ldots, x_n)\chi_{\mathbb{C}^n}(x_1, \ldots, x_n) = (x_1 + \cdots + x_n)^2 = h_{1\,1}$$

So, in this case, we would like $\langle h_{1\,1}, h_{1\,1} \rangle = 2$.

We $\dim_{\mathbb{C}} \mathrm{Hom}^{\mathrm{GL}_n\mathbb{C}}(V, W) = \langle \chi_V, \chi_W \rangle$ is certainly positive. It is also "linear" in some senses. First, for $W = W_1 \oplus W_2$, $V = V_1 \oplus V_2$, we know that:

$$\mathrm{Hom}^G(V, W_1 \oplus W_2) \cong \mathrm{Hom}^G(V, W_1) \oplus \mathrm{Hom}^G(V, W_2)$$
$$\mathrm{Hom}^G(V_1 \oplus V_2, W) \cong \mathrm{Hom}^G(V_1, W) \oplus \mathrm{Hom}^G(V_2, W)$$

This fact for linear maps is well known, and the adaptation of the proof for G equivariant maps is straightfoward. Again from the first problem set, we know that $\chi_{V_1 \oplus V_2}(x_1, \ldots, x_n) =$

$\chi_{V_1}(x_1,\ldots,x_n)\oplus\chi_{V_2}(x_1,\ldots,x_n)$, then:

$$\begin{aligned}\dim_{\mathbb{C}}\mathrm{Hom}^{\mathrm{GL}_n\mathbb{C}}(V,W_1\oplus W_2)&=\dim_{\mathbb{C}}\mathrm{Hom}^{\mathrm{GL}_n\mathbb{C}}(V,W_1)+\dim_{\mathbb{C}}\mathrm{Hom}^{\mathrm{GL}_n\mathbb{C}}(V,W_2)\\&=\langle\chi_V,\chi_{W_1}\rangle+\langle\chi_V,\chi_{W_2}\rangle=\langle\chi_V,\chi_{W_1}+\chi_{W_2}\rangle\\&=\langle\chi_V,\chi_{W_1\oplus W_2}\rangle=\langle\chi_V,\chi_W\rangle\end{aligned}$$

Example. Back to $V=(\mathbb{C}^n)^{\otimes 2}$, we have that $\mathbb{C}^n\otimes\mathbb{C}^n=\mathrm{Sym}^2\mathbb{C}\oplus\bigwedge^2\mathbb{C}$ and $\chi_{\mathbb{C}^n\otimes\mathbb{C}^n}=h_{1\,1}=h_1^2=h_2+e_2=\chi_{\mathrm{Sym}^2\mathbb{C}}+\chi_{\bigwedge^2\mathbb{C}}$.
Thus:

$$\begin{aligned}\dim_{\mathbb{C}}\mathrm{Hom}^{\mathrm{GL}_n\mathbb{C}}(V,V)&=\langle h_{1\,1},h_{1\,1}\rangle=\langle h_{1\,1},h_2+e_2\rangle=\langle h_{1\,1},h_2\rangle+\langle h_{1\,1},e_2\rangle\\&=\dim_{\mathbb{C}}\mathrm{Hom}^{\mathrm{GL}_n\mathbb{C}}(V,\mathrm{Sym}^2\mathbb{C})+\dim_{\mathbb{C}}\mathrm{Hom}^{\mathrm{GL}_n\mathbb{C}}(V,\bigwedge^2\mathbb{C})\end{aligned}$$

Now we have motivation for to define $\langle\ ,\ \rangle:\Lambda_n^{\pm}\times\Lambda_n^{\pm}\to\mathbb{Z}$, but why can we expand it to $\Lambda\times\Lambda$? First we'll look at an important property $\langle\ ,\ \rangle$ needs to have:

Proposition. If χ_V and $\chi_W\in\Lambda_n^{\pm}$ are homogeneous of different degree then $\mathrm{Hom}^G(V,W)$.

Therefore, we will want to want to have $\langle p,q\rangle$ if p and q are homogenous of different degrees.

Proof sketch. Let V be a G-representation such that $\chi_V(x_1,\ldots,x_n)$ is a Laurent polynomial of homogeneous degree d in $x_1,\ldots,x_n$, and $z\in\mathbb{C}$. Then we have $\chi_V(z,z,\ldots,z)=cz^d$ for some constant c. So we have $\mathrm{Tr}\rho_v\begin{bmatrix}z&&0\\&\ddots&\\0&&z\end{bmatrix}=cz^d$ for some nonzero constant c. This implies (we are skipping over a nontrivial argument here) that $\rho_V(z\cdot Id)=z^dId_V$.

Now let V,W be representations such that χ_V,χ_W are homogeneous of degree d,e, respectively. So $(z\cdot Id)*v=z^dv$ and $(z\cdot Id)*w=z^ew$ for all $v\in V,w\in W$. Take $f:V\to W$ a GL_n-equivariant map, in particular, f is linear. Thus:

$$z^df(v)=f(z^dv)=f((z\cdot Id)*v)=(z\cdot Id)*f(v)=z^ef(v)$$

Taking z such that $z^d\neq z^e$, we have that $f(v)=0$ for all $v\in V$. So $\mathrm{Hom}^{\mathrm{GL}_n\mathbb{C}}(V,W)=0$ $\square$

Notice that $\langle\ ,\ \rangle:\Lambda_n^d\times\Lambda_n^d\to\mathbb{Z}$ will be an inner product on a finite rank $\mathbb{Z}$-module, so all of your experience from finite dimensional linear algebra will be relevant.

As we've discussed before, the quotient maps $\cdots\to\Lambda_{n+2}^d\to\Lambda_{n+1}^d\to\Lambda_n^d\to\cdots$ are isomorphisms for $n\geq d$. This is not at all obvious, but it will turn out that these isomorphisms preserve $\langle\ ,\ \rangle$. So the inner products $\langle\ ,\ \rangle$ on Λ_n will give rise to a limiting inner product $\langle\ ,\ \rangle$ on Λ.

The definition. Now, we will finally define $\langle\ ,\ \rangle:\Lambda\times\Lambda\to\mathbb{Z}$. Remember that $\Lambda=\oplus_\lambda\mathbb{Z}\cdot m_\lambda=\oplus_\lambda\mathbb{Z}\cdot h_\lambda$. So it is enough to define $\langle h_\lambda,m_\mu\rangle$ and extend linearly. We define:

$$\langle h_\lambda,m_\mu\rangle=\begin{cases}1&\lambda=\mu\\0&\text{otherwise}\end{cases}$$

From this definition, we already know that $\langle\ ,\ \rangle$ is well defined and bilinear, and that $\langle p,q\rangle=0$ if $p,q\in\Lambda$ are homogeneous of different degree, as partitions of different size are different. However there some properties are still unclear like how $\langle\ ,\ \rangle$ is symmetric and positive definite. We will prove that it is positive definite both on the homework and on a future class. For today, we'll look at symmetry.

Define $A_{\lambda\,\mu} \in \mathbb{Z}$ as the coefficients $h_\mu = \sum_\lambda A_{\lambda\,\mu} m_\lambda$. Then

$$\langle h_\lambda, h_\mu \rangle = \langle h_\lambda, \sum_{\lambda'} A_{\lambda'\,\mu} m_{\lambda'} \rangle = \sum_{\lambda'} A_{\lambda'\,\mu} \langle h_\lambda, m_{\lambda'} \rangle = A_{\lambda\,\mu}$$

So it is enough to prove that $A_{\lambda\,\mu} = A_{\mu\,\lambda}$ for any two partitions λ, μ.

Example. Let's zoom in at Λ^3. Let's try to find the coefficients of $h_{2\,1}$ in the m-basis.

$$\begin{aligned} h_{2\,1} = h_2 h_1 &= \left(\sum x_i^2 + \sum_{i<j} x_i x_j \right) \left(\sum x_k \right) \\ &= \sum x_i^3 + \sum_{i \neq k} x_i^2 x_k + 3 \sum_{i<j<k} x_i x_j x_k + \sum_{i \neq j} x_i x_j^2 \\ &= m_3 + 2m_{2\,1} + 3m_{1\,1\,1} \end{aligned}$$

We can do similar computations to get the following identities.

$$\begin{array}{lclclcl} h_3 & = & m_3 & + & m_{2\,1} & + & m_{1\,1\,1} \\ h_{2\,1} & = & m_3 & + & 2m_{2\,1} & + & 3m_{1\,1\,1} \\ h_{1\,1\,1} & = & m_3 & + & 3m_{2\,1} & + & 6m_{1\,1\,1} \end{array}$$

Hence the transition matrix is $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix}$, which is, as expected, symmetric. The reader can check that it is also positive definite.

To prove that $A_{\lambda\,\mu} = A_{\mu\,\lambda}$, we can show that they count the same structure that is symmetric on λ and μ.

Lemma. $A_{\lambda\,\mu} = \#\{\mathbb{Z}_{\geq 0}\text{-matrices with column sum } \mu \text{ and row sum } \lambda\}$.

Notice that if M has column sum μ and row sum λ, then M^T has row sum μ and collumn sum λ. As taking the transposition is a bijective map on $\mathbb{Z}_{\geq 0}$-matrices, we have that this number is the same when we switch μ and λ.

Example. From the previous example, we know that $A_{21;111} = 3$. In fact, these three matrices are

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

They encode a way of getting a factor $x_{i_1} x_{i_2} x_{i_3}$ of $m_{1\,1\,1}$ in the product $h_{2\,1} = h_{\lambda_1\,\lambda_2} = h_{\lambda_1} h_{\lambda_2}$ in the following manner. The k-th row marks all the appearances of the variable x_{i_k} and the j-th collumn marks the factor of h_{λ_j} in which they appear. So, for example, $\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ is for $x_1 x_2 x_3$ as a product of the factor $x_1 x_2$ of $h_2 = \sum_{i \leq j} x_i x_j$ and the factor x_3 of $h_1 = \sum x_i$. We'll do this argument more slowly next time.

September 9: The Hall inner product and generating functions

Recall the computation of last class

$$h_\lambda = \sum_\mu A_{\lambda\mu} m_\mu.$$

Theorem.

$$A_{\lambda\mu} = \#\{\mathbb{Z}_{\geq 0} \text{ matrix } B \text{ with row sum } \lambda \text{ and column sum } \mu\}$$

Corollary. $A_{\lambda\mu} = A_{\mu\lambda}$.

Corollary. $\langle , \rangle$ is symmetric.

Let's do this computation more slowly: Recall the definition of h_λ is all product of $h_{\lambda_i}(x)$'s, where h_{λ_i} is the sum of all monomials of degree λ_i:

$$\begin{aligned} h_\lambda(x) = h_{\lambda_1}(x)h_{\lambda_2}(x)\dots &= \prod_{i=1} h_{\lambda_i}(x) \\ &= \prod_{i=1}\left(\sum_{B_{i1}+B_{i2}+\ldots=\lambda_i} (x_1^{B_{i1}}x_2^{B_{i2}}x_3^{B_{i3}}\cdots)\right) \\ &= \sum_{\substack{B_{ij}\geq 0 \\ B_{i1}+B_{i2}+\ldots=\lambda_i}} \prod_i (x_1^{B_{i1}}x_2^{B_{i2}}x_3^{B_{i3}}\dots) \quad \text{(through distributive law)} \\ &= \sum_{\substack{B_{ij}\geq 0 \\ B_{i1}+B_{i2}+\ldots=\lambda_i}} \left(x_1^{\sum_i B_{i1}}x_2^{\sum_i B_{i2}}x_3^{\sum_i B_{i3}}\cdots\right) \\ &= \sum_{\substack{B_{ij}\geq 0 \\ \text{row sum}(B)=\lambda}} x^{\text{column sum}(B)}. \end{aligned}$$

So the coefficient of x^μ is the number of B with row sum $= \lambda$ and column sum $= \mu$. $\square$

Let's redo this computation using generating functions:

Start with generating function for h_k

$$\sum_{k=0}^{\infty} t^k h_k(y) = \prod_j \frac{1}{1-ty_j}.$$

We can multiply copies of it. Here, t is renamed to x_i:

$$\prod_i \prod_j \frac{1}{1-x_iy_j} = \prod_i \sum_{k_i}^{\infty} x_i^{k_i} h_{k_i}(y)$$

Now using distributive law, we have

$$\prod_i \prod_j \frac{1}{1-x_iy_j} = \sum_{k_1,k_2,\ldots=0}^{\infty} \prod_{i=1}^{\infty} x_i^{k_i} h_{k_i}(y_i) = \sum_{k_1,k_2,\ldots=0} (x_1^{k_1}\dots)h_{k_1}(y_1)h_{k_2}(y_2)\ldots = \sum_{\lambda_1\geq\lambda_2} m_\lambda(x)h_\lambda(y).$$

We can now conclude that

$$\prod_{i,j=1}^{\infty} \frac{1}{1-x_iy_j} = \sum_\lambda m_\lambda(x)h_\lambda(y).$$

Plug in the calculation of h_λ from the beginning we have

$$\prod_{i,j=1}^{\infty} \frac{1}{1-x_iy_j} = \sum_{\lambda,\mu} A_{\lambda\mu} m_\lambda(x) m_\mu(y).$$

This shows that $A_{\lambda\mu}$ is symmetric.

We can recover the "row and column sum" formula from this formula. We have
$$\prod_{i,j=1}^{\infty} \frac{1}{1-x_i y_j} = \prod_{i,j} \sum_{B_{ij}=0}^{\infty} x_i^{B_{ij}} y_j^{B_{ij}}.$$
Expanding the product means taking one term from each factor, which gives a nonnegative matrix B. The row sums of B give the exponents of the x's and the column sums of B give the exponents of the y's.

We will see $\prod_{i,j=1}^{\infty} \frac{1}{1-x_i x_j}$ coming up and up again, and it does not seem to have a name. We'll call it Cauchy's product.

Notice that we have been expanding symmetric things asymmetrically. In general, if we have some identity that can expand the Cauchy product as sum of products of polynomial of x's and polynomial of y's, what can be deduced from this?

Suppose $p_K(x)$ and $q_L(y)$ are two families of homogeneous symmetric polynomials in Λ, indexed by some index sets, and we have an expansion formula:
$$\prod_{i,j=1}^{\infty} \frac{1}{1-x_i y_j} = \sum_{K,L} B_{KL} p_K(x) q_L(y), \quad B_{KL} \in \mathbb{Z}.$$

Proposition. From the above condition we can deduce that p_K $\mathbb{Z}$-span Λ, as do the q_L's. If p_K and q_L are just in $\mathbb{Q}\Lambda$, and $B_{KL} \in \mathbb{Q}$, then p_k must $\mathbb{Q}$-span Λ, as do the q_L's.

Behind the scene we are using this (symmetric polynomials in two sets of variables) to express $\Lambda \otimes_{\mathbb{Z}} \Lambda$.

We use the following lemma to help prove the proposition.

Lemma.
$$\left\langle f(x), \prod_{i,j} \frac{1}{1-x_i y_j} \right\rangle_{\text{in variable } x} = f(y).$$
(In tensor product language, given $\langle f, \rangle : \Lambda \to \mathbb{Z}$, it also induces a map $\langle f, \rangle \otimes \boldsymbol{Id} : \Lambda \otimes_{\mathbb{Z}} \Lambda \to \Lambda$.)

Proof. it is enough to check this for a $\mathbb{Z}$-basis of Λ. We take h_λ.
$$\left\langle h_\lambda(x), \prod \frac{1}{1-x_i y_j} \right\rangle = \left\langle h_\lambda(x), \sum_\lambda m_\lambda(x) h_\lambda(y) \right\rangle = h_\lambda(y). \qquad \square$$

Let's go back to the proposition.

Proof of Proposition. For any $f \in \Lambda$,
$$\begin{aligned} f(y) &= \left\langle f(x), \prod \frac{1}{1-x_i y_j} \right\rangle \\ &= \sum_{K,L} B_{KL} p_K(x) q_L(y) = \sum_{K,L} B_{KL} \left\langle f(x), p_k(\lambda) \right\rangle q_L(y). \qquad \square \end{aligned}$$

Corollary. Suppose p_λ and q_λ are the families of homogeneous symmetric polynomials indexed by partitions with $\deg p_\lambda = \deg q_\lambda = |\lambda|$ and
$$\prod \frac{1}{1-x_i y_j} = \sum B_{\lambda\mu} p_\lambda(x) q_\mu(y), \quad B_{\lambda\mu} \in \mathbb{Z}.$$

Then

$$\Lambda = \bigoplus_\lambda \mathbb{Z}p_\lambda = \bigoplus_\lambda \mathbb{Z}q_\lambda.$$

Remark. Solve this case (Schur polynomials) and get tenure:

$$\prod \frac{1}{1 - x_i y_j z_k} = \sum_{\lambda,\mu,\nu} g_{\lambda\mu\nu} s_\lambda(x) s_\mu(y) s_\nu(z)$$

Suppose now $p_\lambda(X)$ and $q_\mu(y)$ are two families of homogeneous symmetric polynomials in Λ indexed by partition with $\deg p_\lambda = \deg q_\lambda = |\lambda|$ and let

$$\prod_{i,j=1}^{\infty} \frac{1}{1 - x_i y_j} = \sum_{\lambda\mu} B_{\lambda\mu} p_\lambda(x) q_\mu(y), B_{\lambda\mu} \in \mathbb{Z}$$

Proposition. Let $C_{\lambda\mu} = \langle p_\lambda(x), q_\mu(x)\rangle$. Then B and C^T are inverses.

Proof. By Lemma, $\left\langle p_\nu(y), \prod \frac{1}{1-x_i y_j} \right\rangle = p_\nu(x)$. So

$$\sum_{\lambda,\mu} B_{\lambda\mu} p_\lambda(x) \langle p_\nu(y), q_\mu(y)\rangle = \sum_{\lambda,\mu} B_{\lambda\mu} C_{\nu\mu} p_\lambda(x).$$

Both sides are linear combination of p's and since p's form a basis, we can match up coefficients of two sides and get

$$\sum_\mu B_{\lambda\mu} C_{\nu\mu} = \begin{cases} 1 & \lambda = \mu \\ 0 & \lambda \neq \mu \end{cases} \implies BC^T = \boldsymbol{Id}$$ □

In particular, if $\prod \frac{1}{1-x_i y_j} = \sum_\lambda c_\lambda p_\lambda(x) p_\lambda(y)$ the p_λ are orthogonal for $\langle\ ,\ \rangle$. On the homework, you will check that this happens for power sum symmetric functions.

If we have no coefficients in front of $p_\lambda(x)p_\lambda(y)$ i.e.

$$\prod \frac{1}{1 - x_i y_j} = \sum_\lambda s_\lambda(x) s_\lambda(y),$$

then this results in an orthonormal basis. We will see soon that this happens for the Schur polynomials.

September 12: Inner product restricted to Λ_n; Schur polynomials

Recall that we previously defined an inner product $\langle\ ,\ \rangle : \Lambda \times \Lambda \to \mathbb{Z}$ by the formula

$$\langle h_\lambda, m_\mu \rangle = \delta_{\lambda\mu} = \begin{cases} 1 & \lambda = \mu \\ 0 & \text{otherwise} \end{cases}$$

This inner product is

- symmetric (as proved in class);
- positive-definite (HW);
- a perfect pairing when restricted to each degree d: $\Lambda^d \times \Lambda^d \to \mathbb{Z}$. This means that any linear map $\Lambda^d \to \mathbb{Z}$ can be written as $\langle \beta, - \rangle$ for some $\beta \in \Lambda^d$.

What about when we restrict to the case of only n variables?

Hall inner product on Λ_n. We want to define the restriction $\langle\ ,\ \rangle_n$ of $\langle\ ,\ \rangle$ on the quotient space Λ_n of Λ. Recall that $\Lambda_n = \bigoplus_{l(\lambda)\le n} \mathbb{Z}\cdot m_\lambda$. We can still define that on Λ_n,

$$\langle h_\lambda, m_\mu\rangle_n = \delta_{\lambda\mu} \quad \text{for} \quad l(\lambda), l(\mu) \le n.$$

However, while we do know that

$$\Lambda_n = \mathbb{Z}[h_1,\dots,h_n] = \bigoplus_{l(\lambda^T)\le n} \mathbb{Z}\cdot h_\lambda,$$

it's not clear that Λ_n is $\mathbb{Z}$-spanned by h_λ for $l(\lambda)\le n$. We'll gloss past this for now and resolve it eventually.

Example. When $n \ge 2$, we have

$$h_{11} = (\sum x_i)^2 = \sum x_i^2 + 2\sum_{i<j} x_i x_j = m_2 + 2m_{11}.$$

Thus when $n\ge 2$,

$$\langle h_{11}, h_{11}\rangle = \langle h_{11}, m_2 + m_{11}\rangle = 0 + 2 = 2.$$

This coincides with the representation theory perspective: $\mathrm{Hom}^{\mathrm{GL}_n\mathbb{C}}((\mathbb{C}^n)^{\otimes 2}, (\mathbb{C}^n)^{\otimes 2})$ is 2 dimensional, generated by id and the homomorphism $x\otimes y \mapsto y\otimes x$.

On the other hand, when $n=1$, we have $h_{11} = x_1^2 = h_2 = m_2$. Hence

$$\langle h_{11}, h_{11}\rangle = \langle h_2, m_2\rangle = 1,$$

since $\langle h_\lambda, m_\mu\rangle = \delta_{\lambda\mu}$ holds true only for $l(\lambda), l(\mu)\le n = 1$. In this case, the homomorphism $x\otimes y\mapsto y\otimes x$ defined above from $(\mathbb{C})^{\otimes 2} = \mathbb{C}$ to itself coincides with the identity.

By the same proof as last time,

$$\begin{aligned}\prod_{i,j=1}^{n} \frac{1}{1-x_i y_j} &= \sum_{l(\lambda)\le n} h_\lambda(x_1,\dots,x_n) m_\lambda(y_1,\dots,y_n)\\ &= \sum_{l(\lambda),l(\mu)\le n} A_{\lambda\mu} m_\lambda(x_1,\dots,x_n) m_\mu(y_1,\dots,y_n)\end{aligned}$$

for the same coefficients $A_{\lambda\mu}$ as before. Similarly, if $p_\lambda, q_\lambda \in \Lambda_n$ are polynomials indexed by partitions λ with $l(\lambda)\le n$, and

$$\prod_{i,j=1}^{n} \frac{1}{1-x_i y_j} = \sum_{l(\lambda)\le n} p_\lambda(x) q_\lambda(y),$$

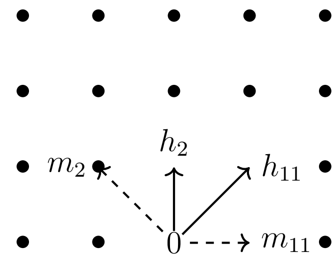
then both the p_λ's and the q_λ's are $\mathbb{Z}$-bases of Λ_n which are dual with respect to the inner product $\langle\ ,\ \rangle_n$. So all discussion in the last lecture holds for $\langle\ ,\ \rangle_n$ as well.

There is a way to think of the definition of Λ_n purely in terms of linear algebra. Let $K = \ker(\Lambda \twoheadrightarrow \Lambda_n) = \bigoplus_{l(\lambda)>n} \mathbb{Z}\cdot m_\lambda$. Then

$$K^\perp := \{f\in\Lambda : \langle f, g\rangle = 0,\ \forall g\in K\} = \bigoplus_{l(\lambda)\le n} \mathbb{Z}\cdot h_\lambda.$$

Since $\langle\ ,\ \rangle$ is positive definite, the map $\mathbb{Q}\otimes K^\perp \to \mathbb{Q}\otimes\Lambda/\mathbb{Q}\otimes K \cong \mathbb{Q}\otimes\Lambda_n$ is an isomorphism which identifies $\langle\ ,\ \rangle|_{K^\perp}$ with $\langle\ ,\ \rangle_n$.

It isn't clear that this works over $\mathbb{Z}$.[2] To see the issue, look at the degree 2 part of Λ. The figure below has been drawn so that $\langle\ ,\ \rangle$ coincides with the standard Euclidean inner product. The h's are labeled with solid arrows and the m's with dashed arrows. The reader can check that the h's and the m's are dual, and can also verify identities like $h_{11} = m_2 + 2m_{11}$.



The quotient map $\Lambda \to \Lambda^2$ is the quotient by $\mathbb{Z}m_{11}$, in other words, projection onto the vertical component. So K is the horizontal axis, and $K^\perp$ is the vertical axis. The claim is that the lattice points of $K^\perp$ map isomorphically to the lattice points of the quotient, which is true.

However, suppose that we instead were quotienting by $\mathbb{Q}m_2$ (the line with slope -1). Then the orthogonal to the kernel would be $\mathbb{Q}h_{11}$ (the line with slope 1). This complement would still map isomorphically to the quotient over $\mathbb{Q}$, but not over $\mathbb{Z}$. So we need to actually use the details of which m's we are setting to 0 in the quotient.

Semi-standard Young Tableaux (SSYT). A ***semi-standard Young Tableau (SSYT)*** is a filling of a Young diagram (corresponding to a partition λ) by positive integers that increase weakly along rows and strictly on columns. For example, below is a SSYT of shape $(4, 2, 1)$:

$$\begin{array}{|c|c|c|c|}\hline 1 & 1 & 2 & 3 \\ \hline 2 & 4 \\ \cline{1-2} 3 \\ \cline{1-1}\end{array}$$

Note that the word "tableau" is singluar, and "tableaux" is its plural form.

For a tableau T, one defines

$$x^T := x_1^{\#(1\text{'s in } T)} x_2^{\#(2\text{'s in } T)} \cdots$$

As an example,

$$x^{\begin{array}{|c|c|c|c|}\hline 1 & 1 & 2 & 3 \\ \hline 2 & 4 \\ \cline{1-2} 3 \\ \cline{1-1}\end{array}} = x_1^2 x_2^2 x_3^2 x_4.$$

Schur polynomials. Now we turn to Schur polynomials s_λ, which corresponds to characters of <u>irreducible</u> representations of $\mathrm{GL}_n\mathbb{C}$. We will prove later that $\langle s_\lambda, s_\mu\rangle = \delta_{\lambda\mu}$, so that the s_λ's form an <u>orthonormal</u> $\mathbb{Z}$-basis of Λ_n.

For a partition λ, define the ***Schur polynomial***

$$s_\lambda := \sum_{T \text{ SSYT of shape } \lambda} x^T.$$

Example. s_1 corresponds to fillings of the diagram $\square$ (by positive integers), so

$$s_1 = x_1 + x_2 + \cdots = e_1 = h_1 = m_1;$$

[2]This material is added by Prof. Speyer, as he understands the issue better now.

For $k \geq 1$, s_k corresponds to filling the diagram with one row and k boxes in that row, thus

$$s_k = \sum_{i_1 \leq \cdots \leq i_k} x_{i_i} \ldots x_{i_k} = h_k;$$

s_{1^k} corresponds to filling the diagram with one column and k boxes in that column, thus

$$s_{1^k} = \sum_{i_1 < \cdots < i_k} x_{i_i} \ldots x_{i_k} = e_k.$$

Example. There are only two degree 2 Schur polynomials, which are all give above:

- $s_2 = h_2$, corresponding to the character of the natural $\mathrm{GL}_n\mathbb{C}$ representation on $\mathrm{Sym}^2\mathbb{C}^n$;
- $s_{11} = e_2$, corresponding to the character of the natural $\mathrm{GL}_n\mathbb{C}$ representation on $\bigwedge^2 \mathbb{C}^n$.

Example. There are three degree 3 Schur polynomials:

- $s_3 = h_3 = m_3 + m_{21} + m_{111}$;
- $s_{111} = e_3 = m_{111}$;
- $$\begin{aligned} s_{21} = \sum_{i \leq j, i<k} x_i x_j x_k &= \sum_{i<k<j} x_i x_j x_k + \sum_{i<j} x_i x_j^2 + \sum_{i \leq j < k} x_i x_j x_k \\ &= \sum_{i<k<j} x_i x_j x_k + \sum_{i<j} x_i x_j^2 + \sum_{i<k} x_i^2 x_k + \sum_{i<j<k} x_i x_j x_k \\ &= m_{21} + 2m_{111}. \end{aligned}$$

Thus one have

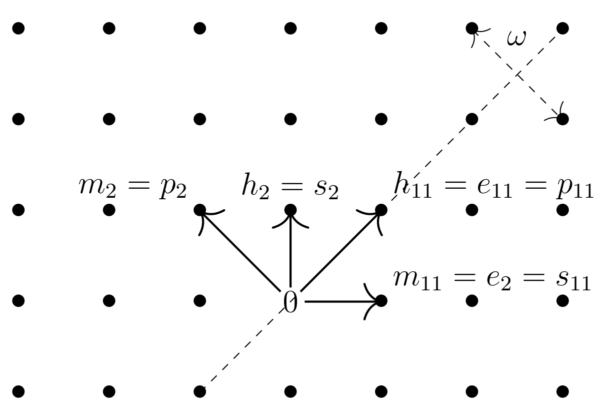
$$\begin{bmatrix} s_3 \\ s_{21} \\ s_{111} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 2 \\ & & 1 \end{bmatrix} \begin{bmatrix} m_3 \\ m_{21} \\ m_{111} \end{bmatrix}.$$

One notices that the s_λ's are related to the m_λ's by an upper-triangular matrix.

As an exercise, one can check that indeed $s_3, s_{21}, s_{1,1,1}$ are indeed orthonormal with respect to $\langle\ ,\ \rangle$.

September 14 – Basic properties of Schur functions

At the start of class, Prof. Speyer drew the following diagram, which he finds to be a useful reference:



The above diagram depicts all of the named symmetric polynomials in degree 2. The Hall inner product has been drawn to coincide with the usual Euclidean inner product on $\mathbb{R}^2$. The involution ω is the reflection over the dashed diagonal line. Some things you can see from this drawing (not all of which we have proved yet):

(1) The s's are orthonormal, the p's are orthogonal but not orthonormal, the h's are dual to the m's.
(2) The involution ω switches e_λ and h_λ, maps s_λ to s_{λ^T}, and maps p_λ to $\pm p_\lambda$.
(3) The bases m_λ, e_{λ^T} and s_λ are upper triangular with respect to each other. The h's are also upper triangular with respect to the s's, in the reverse order.

We now turn to the main topic of today:

Schur polynomials. In the previous lecture, we defined Schur polynomials as a sum over semistandard Young tableaux.

Theorem. For any partition λ, let s_λ be its Schur polynomial. Then,

(1) s_λ is a symmetric function.
(2) s_λ is upper triangular in m_λ, i.e.

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

for some integer $K_{\lambda\mu}$ known as the ***Kotska numbers***.

Proof. To prove the first assertion, we have to show that the coefficient of each monomial $x_1^{\mu_1} x_2^{\mu_2} \cdots x_n^{\mu_n}$ equals to that of $x_1^{\mu_{\sigma(1)}} x_2^{\mu_{\sigma(2)}} \cdots x_n^{\mu_{\sigma(n)}}$ for any permutation $\sigma \in S_n$. It suffices to prove this for a set of generators of S_n, which we shall take as the transpositions $(k \quad k+1)$ for each $1 \le k \le n$. In terms of tableaux, this means that we have to find a bijection between the set of tableaux with x k's and y $(k+1)$'s, and the set of tableaux with x $(k+1)$'s and y k's, which fixes the number of entries of other values.

Observe that in each row, we have a segment that looks like

$\cdots$	k	k	$\cdots$	k	$k+1$	$k+1$	$\cdots$	$k+1$	$\cdots$

Whenever we have k on top of $k+1$, we leave it alone. Then, we are left with disjoint horizontal strips, for each of which we shall interchange the number of k's and $k+1$'s. This gives us an algorithm to convert a tableau with x k's and y $(k+1)$'s to a tableau with x $(k+1)$'s and y k's.

To prove the second assertion, we first express

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

We shall prove that if $K_{\lambda\mu} > 0$, then $\mu \le \lambda$, and that $K_{\lambda\lambda} = 1$.

Suppose that there exists a tableux of shape λ with μ_i i's. Since 1's can only occur in row 1, we have $\mu_1 \le \lambda_1$. Inductively, for any positive integer k, since values at most k can only occur in the first k rows, we have

$$\sum_{i=1}^k \mu_i \le \sum_{i=1}^k \lambda_i.$$

By definition, we have $\mu \leq \lambda$.

If $\mu = \lambda$, then we have

$$\sum_{i=1}^{k} \mu_i = \sum_{i=1}^{k} \lambda_i$$

for any positive integer k. This first implies that the first row of the tableau is entirely filled with 1's, and inductively implies that the ith row of the tableau is entirely filled with i's. This means that the tableau must be of the form

$$\begin{array}{|c|c|c|c|c|c|c|c|c|}
\hline
1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\
\hline
2 & \cdots & 2 & 2 & \cdots & 2 \\
\cline{1-6}
\vdots & \vdots & \vdots \\
\cline{1-3}
n & \cdots & n \\
\cline{1-3}
\end{array}$$

This means that $K_{\lambda\lambda} = 1$. $\square$

To illustrate the algorithm in the proof of the first assertion, we discussed the following example with Young tableau with 3 boxes and values at most 3.

Example. The algorithm exchanges the following pairs:

$$\begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 \\ \cline{1-1}\end{array} \longleftrightarrow \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 \\ \cline{1-1}\end{array}$$

$$\begin{array}{|c|c|}\hline 1 & 1 \\ \hline 3 \\ \cline{1-1}\end{array} \longleftrightarrow \begin{array}{|c|c|}\hline 2 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}$$

$$\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 3 \\ \cline{1-1}\end{array} \longleftrightarrow \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 3 \\ \cline{1-1}\end{array}$$

and maps the following tableaux to itself:

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \quad \text{and} \quad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array}.$$

At the end of the lecture, we also stated the following theorem, whose proof shall be discussed in the next class.

Theorem. $\{s_\lambda\}$ is an orthonormal basis of Λ with respect to Hall's inner product.

September 17 – Cauchy's identity, flawed attempt

Our goal is to show that the Schur polynomials are an orthonormal basis for $\langle\ ,\ \rangle$. Equivalently, this means we want to show

$$\prod_{i,j} \frac{1}{1 - x_i y_j} = \sum_{\lambda} s_\lambda(x) s_\lambda(y).$$

Since $\prod_{i,j} \frac{1}{1-x_i y_j} = \sum_{\lambda,\mu} A_{\lambda\mu} m_\lambda(x) m_\mu(y)$, the coefficient of $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} y_1^{\beta_1} y_2^{\beta_2} \cdots y_n^{\beta_n}$ on the left side is

$$A_{\alpha\beta} = \# \left\{ \mathbb{Z}_{\geq 0} \text{ - matrices} \,\middle|\, \begin{array}{c} \text{row sum} = \alpha \\ \text{column sum} = \beta \end{array} \right\}.$$

The coefficient for this term on the right side (which we wish to prove is equal) is

$$\#\left\{(T,U)\;\middle|\;\begin{array}{c} T,U\in \mathrm{SSYT}(\lambda) \text{ for some } \lambda,\\ \#\text{ of } j\text{'s in } T \text{ is } \alpha_j,\\ \#\text{ of } j\text{'s in } U \text{ is } \beta_j\end{array}\right\}.$$

For instance, with the term $x_1x_2y_1y_2$, this says that

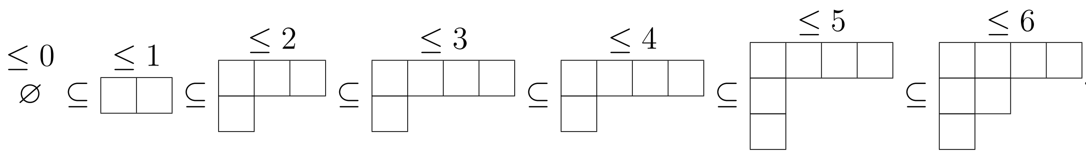
$$\#\left\{\begin{bmatrix}1&0\\0&1\end{bmatrix},\begin{bmatrix}0&1\\1&0\end{bmatrix}\right\}=\#\left\{\left(\,\begin{array}{|c|c|}\hline 1&2\\\hline\end{array},\ \begin{array}{|c|c|}\hline 1&2\\\hline\end{array}\,\right),\left(\begin{array}{|c|}\hline 1\\\hline 2\\\hline\end{array},\ \begin{array}{|c|}\hline 1\\\hline 2\\\hline\end{array}\right)\right\}$$

which is true.

There is a useful way to think of semistandard Young tableaux using chains of partitions, which we now describe. Given an SSYT T, the position of boxes with labels $\leq j$ form a partition. For example, if

$$T=\begin{array}{|c|c|c|c|}\hline 1&1&2&3\\\hline 2&6\\\cline{1-2} 5\\\cline{1-1}\end{array},$$

then we have

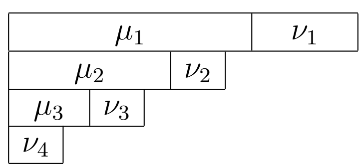


Note that equality is possible when a number is omitted, as 4 is in this example.

Given two partitions μ and ν, we can have μ be $\{\text{boxes}\leq i\}$ and ν be $\{\text{boxes}\leq i+1\}$ if and only if

$$\nu_1\geq\mu_1\geq\nu_2\geq\mu_2\geq\nu_3\geq\cdots$$

We'll say that "ν is μ plus a horizontal strip." We'll write $\mu\uparrow\nu$ and $\nu\downarrow\mu$.



We have that

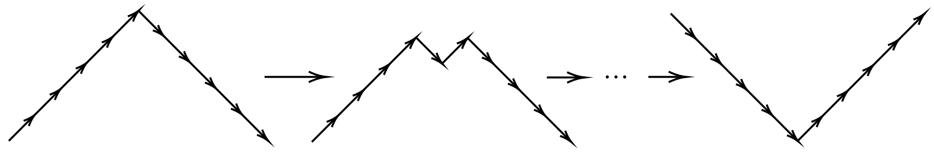
$$\#\left\{(T,U)\;\middle|\;\begin{array}{c} T,U\in \mathrm{SSYT}(\lambda) \text{ for some } \lambda,\\ \#\text{ of } j\text{'s in } T \text{ is } \alpha_j,\\ \#\text{ of } j\text{'s in } U \text{ is } \beta_j\end{array}\right\}$$

$$=\#\left\{(T,U)\;\middle|\;\begin{array}{c} T,U\in \mathrm{SSYT}(\lambda) \text{ for some } \lambda,\\ T \text{ corresponds to } \varnothing=\gamma_0\uparrow\gamma_1\uparrow\gamma_2\uparrow\cdots\uparrow\lambda,\ \forall j:\alpha_j=|\gamma_j|-|\gamma_{j-1}|\\ U \text{ corresponds to } \varnothing=\delta_0\uparrow\delta_1\uparrow\delta_2\uparrow\cdots\uparrow\lambda,\ \forall j:\beta_j=|\delta_j|-|\delta_{j-1}|\end{array}\right\},$$

so

$$\sum_\lambda s_\lambda(x)s_\lambda(y)=\sum_{\varnothing=\gamma_0\uparrow\gamma_1\uparrow\cdots\uparrow\lambda\downarrow\cdots\downarrow\delta_1\downarrow\delta_0=\varnothing} x_1^{|\gamma_1|-|\gamma_0|}x_2^{|\gamma_2|-|\gamma_1|}\cdots y_1^{|\delta_1|-|\delta_0|}y_2^{|\delta_2|-|\delta_1|}\cdots$$

To evaluate the right hand side, we'll essentially flip one by one until $\varnothing=\gamma_0\uparrow\gamma_1\uparrow\cdots\uparrow\lambda\downarrow\cdots\downarrow\delta_1\downarrow\delta_0=\varnothing$ turns into $\varnothing=\gamma_0\downarrow\gamma_1\downarrow\cdots\downarrow\lambda\uparrow\cdots\uparrow\delta_1\uparrow\delta_0=\varnothing$, at which point all γ_j,δ_j will be forced to be $\varnothing$.



Our goal now is to prove the lemma which will do the flipping. This was garbled in class, and we will try to present it correctly now.

Lemma. Let β and γ be two partitions. Then

$$\sum_{\beta\uparrow\delta\downarrow\gamma} x^{|\delta|-|\beta|}y^{|\delta|-|\gamma|} = \frac{1}{1-xy}\sum_{\beta\downarrow\alpha\uparrow\gamma} x^{|\gamma|-|\alpha|}y^{|\beta|-|\alpha|}$$

Example. The smallest example is $\beta = \gamma = \varnothing$. In this case, the lemma says

$$\sum_{\varnothing\uparrow\delta\downarrow\varnothing} x^{|\delta|}y^{|\delta|} = \frac{1}{1-xy}\sum_{\varnothing\downarrow\alpha\uparrow\varnothing} x^{|\varnothing|-|\alpha|}y^{|\varnothing|-|\alpha|}.$$

The left hand side is $\sum_{k\geq 0} x^k y^k = \frac{1}{1-xy}$, while the right hand side is $\frac{1}{1-xy}\cdot 1$, so the lemma holds.

Example. The next smallest example is $\beta = \varnothing, \gamma = (1)$, where the lemma says we should have

$$\sum_{\varnothing\uparrow\delta\downarrow(1)} x^{|\delta|}y^{|\delta|-1} = \frac{1}{1-xy}\sum_{\varnothing\downarrow\alpha\uparrow(1)} x^{1-|\alpha|}y^{0-|\alpha|}.$$

Here the left hand side is $\sum_{k\geq 1} x^k y^{k-1} = \frac{x}{1-xy}$ while the right hand side gives $\frac{1}{1-xy}\cdot x = \frac{x}{1-xy}$, so again the lemma holds.

Start of proof. Let's outline what a proof should look like. The left hand side is a sum over the set

$$\{\delta : \beta\uparrow\delta\downarrow\gamma\}.$$

If we expand $\frac{1}{1-xy}$ on the right hand side as $\sum_{k\geq 0} x^k y^k$, then the right hand side is a sum over

$$\{(\alpha,k) : \beta\downarrow\alpha\uparrow\gamma,\ k\}.$$

So we need a bijection between the set of partitions δ and the set of pairs (α,k).

Moreover, this bijection needs to map $x^{|\delta|-|\beta|}y^{|\delta|-|\gamma|}$ to $x^{|\gamma|-|\alpha|+k}y^{|\beta|-|\alpha|+k}$. In other words, we need

$$\begin{aligned} |\delta|-|\beta| &= |\gamma|-|\alpha|+k \\ |\delta|-|\alpha| &= |\gamma|-|\beta|+k. \end{aligned}$$

Note that these conditions are equivalent to each other, and can be written more symmetrically as

$$|\alpha|+|\delta| = |\beta|+|\gamma|+k. \qquad (*)$$

At this part the proof in class was broken. We will record the wrong approach below and start over next time. $\square$

The broken proof. Let's write down the conditions on δ and α explicitly. The condition that $\beta\uparrow\delta\downarrow\gamma$ means that

$$\delta_1 \geq \beta_1, \gamma_1 \geq \delta_2 \geq \beta_2, \gamma_2 \geq \delta_3 \geq \cdots$$

and the condition that $\beta \downarrow \alpha \uparrow \gamma$ means that

$$\beta_1, \gamma_1 \geq \alpha_1 \geq \beta_2, \gamma_2 \geq \alpha_2 \geq \cdots .$$

Thus, if $\beta \uparrow \delta \downarrow \gamma$, and we define $\delta' = (\delta_2, \delta_3, \ldots)$, then $\beta \downarrow \delta' \uparrow \gamma$.

Professor Speyer thought that we should take $(\alpha, k) = (\delta', \delta_1 - \max(\beta_1, \gamma_1))$. This **is** a bijection between $\{\delta : \beta \uparrow \delta \downarrow \gamma\}$ and $\{(\alpha, k) : \beta \downarrow \alpha \uparrow \gamma,\ k\}$, but it does **not** obey condition $(*)$. We will see the correct bijection next time. "□"

September 19 – Cauchy's identity, Corrected proof

Our goal is to finish the proof of

$$\prod_{i,j} \frac{1}{1 - x_i y_j} = \sum_{\lambda} s_\lambda(x) s_\lambda(y).$$

We will begin by finishing the proof of the Lemma from last class.

Lemma. Let β and γ be two partitions. Then

$$\sum_{\beta\uparrow\delta\downarrow\gamma} x^{|\delta|-|\beta|} y^{|\delta|-|\gamma|} = \frac{1}{1-xy} \sum_{\beta\downarrow\alpha\uparrow\gamma} x^{|\gamma|-|\alpha|} y^{|\beta|-|\alpha|}$$

Proof of Lemma. We can rewrite this as

$$\sum_{\beta\uparrow\delta\downarrow\gamma} x^{|\delta|-|\beta|} y^{|\delta|-|\gamma|} = \sum_{k=0}^{\infty} x^k y^k \sum_{\beta\downarrow\alpha\uparrow\gamma} x^{|\gamma|-|\alpha|} y^{|\beta|-|\alpha|}$$

or, in other words,

$$\sum_{\beta\uparrow\delta\downarrow\gamma} x^{|\delta|-|\beta|} y^{|\delta|-|\gamma|} = \sum_{\beta\downarrow\alpha\uparrow\gamma,\ k\geq 0} x^{|\gamma|-|\alpha|+k} y^{|\beta|-|\alpha|+k}$$

We need is to find a bijection between

$$\{\delta : \beta \uparrow \delta \downarrow \gamma\} \quad \text{and} \quad \{(\alpha, k) : \beta \downarrow \alpha \uparrow \gamma,\ k \geq 0\}$$

$|\delta| - |\beta| = |\gamma| - |\alpha| + k$ and $|\delta| - |\gamma| = |\gamma| - |\beta| + k$. Note that these equations are equivalent to each other, and both can be written more elegantly as

$$|\delta| + |\alpha| - k = |\beta| + |\gamma|.$$

Now, we translate the conditions $\beta \uparrow \delta \downarrow \gamma$ and $\beta \downarrow \alpha \uparrow \gamma$ into inequalities. They are equivalent to:

$$\begin{aligned} \delta_1 \geq \max(\beta_1, \gamma_1) \geq \min(\beta_1, \gamma_1) \geq & \delta_2 \geq \max(\beta_2, \gamma_2) \geq \min(\beta_2, \gamma_2) \geq \delta_3 \geq \cdots \\ \min(\beta_1, \gamma_1) \geq & \alpha_1 \geq \max(\beta_2, \gamma_2) \geq \min(\beta_2, \gamma_2) \geq \alpha_2 \geq \cdots \end{aligned}$$

Given $\delta = (\delta_1, \delta_2 \cdots)$ with $\beta \uparrow \delta \downarrow \gamma$, define (α, k) by

$$\alpha_j = \min(\beta_j, \gamma_j) + \max(\beta_{j+1}, \gamma_{j+1}) - \delta_{j+1} \quad \text{and} \quad k = \delta_1 - \max(\beta_1, \gamma_1).$$

Because of the inequalities, we have $k \geq 0$ and $\beta \downarrow \alpha \uparrow \gamma$. This mapping is also clearly invertible.

We now verify:

$$\begin{aligned}
|\delta| + |\alpha| - k &= \textstyle\sum_{j\geq 1} \delta_j + \sum_{j\geq 1}(\min(\beta_j,\gamma_j) + \max(\beta_{j+1},\gamma_{j+1}) - \delta_{j+1}) - (\delta_1 - \max(\beta_1,\gamma_1)) \\
&= \textstyle\sum_{j\geq 1}(\min(\beta_j,\gamma_j) + \max(\beta_{j+1},\gamma_{j+1}) - \delta_j + \delta_j) + \max(\beta_1,\gamma_1) \\
&= \textstyle\sum_{j\geq 1}(\min(\beta_j,\gamma_j) + \max(\beta_j,\gamma_j)) \\
&= \textstyle\sum_{j\geq 1}(\beta_j + \gamma_j) \\
&= |\beta| + |\gamma|.
\end{aligned}$$

So this bijection satisfies the conditions we needed. □

We are now ready to prove Cauchy's identity:

This lemma is exemplified in the picture below, where the edges are weighted by the variable with exponent the difference of of the size of the partitions from its end to its start. To go from the top path to the bottom path, we have to multiply by the factor in the middle.

$$\begin{tikzcd}
& \delta \arrow[dr, "y"] & \\
\beta \arrow[ur, "x"] \arrow[dr, "y"'] \arrow[rr, phantom, "\frac{1}{1-xy}"] & & \gamma \\
& \alpha \arrow[ur, "x"'] &
\end{tikzcd}$$

Theorem. For every n,

$$\prod_{i,j=1}^{n} \frac{1}{1-x_iy_j} = \sum_{\lambda} s_\lambda(x_1,\cdots,x_n)s_\lambda(y_1,\cdots,y_n).$$

Proof. We can rewrite the right hand side as

$$\sum_{\emptyset=\delta^0\uparrow\delta^1\uparrow\cdots\delta^n\downarrow\delta^{n-1}\downarrow\cdots\downarrow\delta^{2n}=\emptyset} x_1^{|\delta^1|-|\delta^0|}\cdots x_n^{|\delta^n|-|\delta^{n-1}|} y_n^{|\delta^n|-|\delta^{n+1}|}\cdots y_1^{|\delta^{2n-1}|-|\delta^{2n}|}.$$

This is equivalent to a path going from the empty partition to δ and back down, with arrows for each step of the way weighted with the respective x_i or y_i. By the lemma, we can transform this to a path that goes down to some α and then up, multiplying by the appropriate weights. That is, we can instead rewrite the right hand side as

$$\sum_{\emptyset=\alpha^0\downarrow\alpha^1\downarrow\cdots\alpha^n\uparrow\alpha^{n-1}\uparrow\cdots\uparrow\alpha^{2n}=\emptyset} y_n^{|\alpha_0|-|\alpha_1|}\cdots y_1^{|\alpha_{n-1}|-|\alpha_n|} x_1^{|\alpha_{n+1}|-|\alpha_n|}\cdots x_n^{|\alpha_{2n}|-|\alpha_{2n-1}|}\cdot\prod_{i,j=1}^{n}\frac{1}{1-x_iy_j}.$$

But since the only thing we get from the empty partition by going down is the empty partition, and the only partition that goes up to the empty partition is also the empty partition, $\alpha^i = \emptyset$ for every i, so the only possible chain is composed of all empty partitions, and therefore all the exponents are zero, so the sum is 1. This means we can simplify the right hand side to $\prod_{i,j=1}^{n} \frac{1}{1-x_iy_j}$, the result we wanted. □

By taking the limit of these identities as $n \to \infty$, we get the general result $\prod_{i,j=1} \frac{1}{1-x_iy_j} = \sum_\lambda s_\lambda(x)s_\lambda(y)$.

Corollary. The $s_\lambda(x)$ form an orthonormal basis for Λ.

Also, we can see that if $\ell(\lambda) > n$, there are no SSYT of shape λ filled with elements of $\{1, 2, \cdots, n\}$, since the elements in row i of any SSYT have to be $\geq i$. So $\prod_{i,j=1}^{n} \frac{1}{1-x_i y_j} = \sum_{\ell(\lambda)\leq n} s_\lambda(x_1, \cdots, x_n) s_\lambda(y_1, \cdots, y_n)$.

In particular, if we define $\langle\ ,\ \rangle_n$ on Λ_n by $\langle h_\lambda, m_\mu\rangle = \begin{cases} 1 & \lambda = \mu \\ 0 & \lambda \neq \mu \end{cases}$ for $\ell(\lambda), \ell(\mu) \leq n$, we get that $\{s_\lambda | \ell(\lambda) \leq n\}$ are orthonormal.

We can construct a summary of the relations between the basis we have so far. We know that $s_\lambda = \sum_\mu K\lambda\mu m_\mu$, where $K\lambda\mu$ is the number of SSYt of shape λ and content μ. We also know that $h_\pi = \sum_\mu B_{\pi\mu} m_\mu$, where $B_{\pi\mu}$ is the number of $\mathbb{Z}_{\geq 0}-$matrices with row sum π and column sum μ. The following corollary shows how to go directly from h_π to s_λ.

Corollary. We have that $\langle h_\pi, s_\lambda\rangle$ is the number of SSYT of shape λ and content pi, that is, $K_{\lambda\pi}$. Thus $h_\pi = \sum_\lambda K\lambda\pi s_\lambda$.

Proof. We have

$$\begin{aligned} \langle h_\pi, s_\lambda\rangle &= \langle h_\pi, \sum_\mu K_{\lambda\mu} m_\mu\rangle \\ &= \sum_\mu K_{\lambda\mu}\langle h_\pi, s_\mu\rangle \\ &= K_{\lambda\pi}. \end{aligned}$$

Since the s_λ are orthonormal, the inner product gives the coefficients of the decomposition of h this basis. □

In the diagram below, we have labeled the combinatorial objects that are counted by each transition coefficient. The labels $\pi \preceq \lambda$ and $\lambda \succeq \mu$ are meant to remind the reader that, if the coefficient is nonzero, then this inequality holds.

$$\begin{tikzcd}
h_\pi \arrow[r, "\text{SSYT, shape } \lambda \text{, content } \pi", "\pi \preceq \lambda"'] \arrow[dr, "\mathbb{Z}_{\geq 0}\text{-matrices with row sum } \pi \text{ and column sum } \mu"'] & s_\lambda \arrow[d, "\text{SSYT, shape } \lambda \text{, content } \mu"', "\mu \preceq \lambda"] \\
 & m_\mu
\end{tikzcd}$$

September 21: The Jacobi-Trudi Identity

Last time, we ended with the following diagram showing how the h, s, and m bases of symmetric functions relate to each other:

$$
\begin{tikzcd}
h_\pi \arrow[rr, "\text{SSYT, shape } \lambda\text{, content } \pi", "\pi \preceq \lambda"'] \arrow[drr, "\begin{array}{l}\mathbb{Z}_{\geq 0}\text{-matrices with}\\ \text{row sum } \pi \text{ and}\\ \text{column sum } \mu\end{array}"'] & & s_\lambda \arrow[d, "\begin{array}{l}\text{SSYT,}\\ \text{shape } \lambda\text{,}\\ \text{content } \mu\end{array}"', "\mu \preceq \lambda"] \\
& & m_\mu
\end{tikzcd}
$$

We denote the number of SSYT with shape λ and content π (with $\pi \preceq \lambda$) by $K_{\lambda\pi}$, and the number of $\mathbb{Z}_0$-matrices with row sum π and column sum μ by $A_{\pi\mu}$. Then the above diagram can be interpreted as saying that

$$A_{\pi\mu} = \sum_\lambda K_{\lambda\pi} K_{\lambda\mu}.$$

That is, we have a bijection between the set of $\mathbb{Z}_{\geq 0}$ matrices with row sum π, column sum μ, and the set of pairs (T, U) of SSYT of the same shape with $\text{content}(T) = \pi$ and $\text{content}(U) = \mu$. This is the Robinson-Schensted-Knuth (RSK) correspondence. The above equation looks like a matrix equation: $A = K^T K$. Indeed, here are the transition matrices for degree 3 symmetric polynomials and their relation:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Our goal today is to begin to prove the following identity, which allows us to write the s_λ in terms of the h_π, and hence gives the coefficients of the matrix K^{-1}.

Theorem (Jacobi-Trudi)**.** If $\lambda = (\lambda_1, \dots, \lambda_\ell)$, then

$$s_\lambda = \det \begin{bmatrix} h_{\lambda_1} & h_{\lambda_1+1} & \cdots & h_{\lambda_1+(\ell-1)} \\ h_{\lambda_2-1} & h_{\lambda_2} & \cdots & h_{\lambda_2+(\ell-2)} \\ \vdots & & \ddots & \vdots \\ h_{\lambda_\ell-(\ell-1)} & & \cdots & h_{\lambda_\ell} \end{bmatrix},$$

where we adopt the conventions $h_0 = 1$, $h_k = 0$ for $k < 0$.

Example. If $\lambda = (2, 1)$, we have

$$s_{21} = \det \begin{bmatrix} h_2 & h_3 \\ 1 & h_1 \end{bmatrix} = h_{21} - h_3 = (m_3 + 2m_{21} + 3m_{111}) - (m_3 + m_{21} + m_{111}) = m_{21} + 2m_{111},$$

which is the decomposition of s_{21} into monomial symmetric functions that we have seen before.

Example. If $\lambda = (5,2,1)$, we have

$$s_{521} = \det \begin{bmatrix} h_5 & h_6 & h_7 \\ h_1 & h_2 & h_3 \\ 0 & 1 & h_1 \end{bmatrix}.$$

Example. If $\lambda = 1^k = (1,\dots,1)$, we have

$$s_{1^k} = \det \begin{bmatrix} h_1 & h_2 & \cdots & h_k \\ 1 & h_1 & \cdots & h_{k-1} \\ 0 & 1 & \cdots & h_{k-2} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & h_1 \end{bmatrix},$$

which is equal to e_k by homework 1, problem 5. This gives the expected result that $s_{1^k} = e_k$.

Before we begin to prove the identity, we will compute explicitly what it gives for the coefficients of K^{-1}. We have

$$\det \begin{bmatrix} h_{\lambda_1} & h_{\lambda_1+1} & \cdots & h_{\lambda_1+(\ell-1)} \\ h_{\lambda_2-1} & h_{\lambda_2} & \cdots & h_{\lambda_2+(\ell-2)} \\ \vdots & & \ddots & \vdots \\ h_{\lambda_\ell-(\ell-1)} & & \cdots & h_{\lambda_\ell} \end{bmatrix} = \sum_{w\in S_\ell} (-1)^w h_{\lambda_1+w(1)-1}\cdots h_{\lambda_\ell+w(\ell)-\ell}.$$

The coefficient for h_π is then

$$(K^{-1})_{\pi\lambda} = \sum_{\substack{w\in S_\ell \\ \mathrm{sort}(\lambda_1+w(1)-1,\dots,\lambda_\ell+w(\ell)-\ell)=(\pi_1,\dots,\pi_\ell)}} (-1)^w.$$

The way we will prove Jacobi-Trudi next time is by viewing both sides as counting paths in directed graphs. This interpretation of s_λ comes from the bijection constructed in homework 3, problem 4, between SSYTs of shape λ with entries from 1 to n and a set of collections of vertex disjoint directed paths connecting certain points in $\mathbb{Z}^2$.

For the remainder of class, we build our intuition toward relating the determinant in the Jacobi-Trudi identity with counting paths. We first note that $h_k(x_1,\dots,x_n)$ is the generating function for the weighted paths connecting the opposite corners in the following directed graph (with $n-1$ columns and $k+1$ rows):

$$\begin{tikzcd}
\bullet \arrow[r, "1"] & \bullet \arrow[r, "1"] & \cdots \arrow[r, "1"] & \bullet \\
\bullet \arrow[u, "x_1"] \arrow[r, "1"] & \bullet \arrow[u, "x_2"] \arrow[r, "1"] & \cdots \arrow[r, "1"] & \bullet \arrow[u, "x_n"] \\
\vdots \arrow[u, "x_1"] & \vdots \arrow[u, "x_2"] & & \vdots \arrow[u, "x_n"] \\
\bullet \arrow[u, "x_1"] \arrow[r, "1"] & \bullet \arrow[u, "x_2"] \arrow[r, "1"] & \cdots \arrow[r, "1"] & \bullet \arrow[u, "x_n"]
\end{tikzcd}$$

For example, when $n=3$ and $k=2$, the sum of the weights of all possible paths from the bottom left to the top right corner are

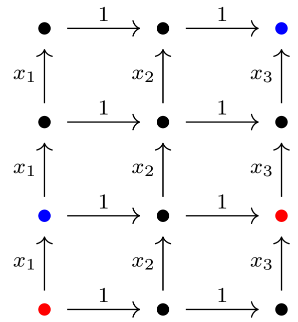
$$h_2(x_1,x_2,x_3) = x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_2x_3 + x_1x_3,$$

where the weight of a path is given by the product of the weights of its edges.

Next time we will put these into the context of the determinant in Jacobi-Trudi via the Lindström-Gessel-Viennot theorem. As a sneak peak, consider the expression.

$$\det \begin{bmatrix} h_2 & h_3 \\ h_0 & h_1 \end{bmatrix} = h_2h_1 - h_3h_0$$

When $n = 3$, each of the h_k counts weighted paths in



Namely, h_2 counts paths (${\color{blue}\bullet} \to {\color{blue}\bullet}$), h_1 counts paths (${\color{red}\bullet} \to {\color{red}\bullet}$), h_3 counts paths (${\color{red}\bullet} \to {\color{blue}\bullet}$), and h_0 counts paths (${\color{blue}\bullet} \to {\color{red}\bullet}$). The expression given by the determinant then counts all pairs of a path (${\color{blue}\bullet} \to {\color{blue}\bullet}$) and a path (${\color{red}\bullet} \to {\color{red}\bullet}$), then subtracts off those pairs of paths that intersect each other.

September 23 – Jacobi-Trudi identity, proof conclusion

Today's goal is to finish the proof of the Jacobi-Trudi Identity, which states that

$$s_{\lambda_1}s_{\lambda_2}\dots s_{\lambda_\ell} = \det \begin{bmatrix} h_{\lambda_1} & h_{\lambda_1+1} & \cdots & h_{\lambda_1+(\ell-1)} \\ h_{\lambda_2-1} & h_{\lambda_2} & \cdots & h_{\lambda_2+(\ell-2)} \\ \vdots & & \ddots & \vdots \\ h_{\lambda_\ell-(\ell-1)} & & \cdots & h_{\lambda_\ell} \end{bmatrix},$$

where we put $h_0 = 1$ and $h_k = 0$ for $k < 0$.

Example.

$$s_{(5,2,1)} = \det \begin{bmatrix} h_5 & h_6 & h_7 \\ h_1 & h_2 & h_3 \\ 0 & 1 & h_1 \end{bmatrix}$$

Let Γ be a finite, acyclic directed graph with weights $x(e)$ on each edge e. The weight of a path γ in Γ of the form

$$v_0 \xrightarrow{e_1} v_1 \xrightarrow{e_2} v_2 \xrightarrow{e_3} \dots \xrightarrow{e_\ell} v_\ell$$

is defined to be $\mathtt{wt}(\gamma) = \prod_{i=1}^{\ell} x(e_i)$. For two vertices v, w of Γ, we define

$$P(v \to w) = \sum_{v \xrightarrow{\gamma} w} \mathtt{wt}(\gamma)$$

Therefore, if $v_1, v_2, \ldots, v_\ell, w_1, w_2, \ldots, w_\ell$ are vertices of Γ then

$$\begin{aligned}\det[P(v_i \to w_j)] &= \sum_{\sigma \in S_\ell} (-1)^\sigma \prod_{i=1}^{\ell} P(v_i \to w_{\sigma(i)}) \\ &= \sum_{\sigma \in S_\ell} (-1)^\sigma \sum_{\substack{v_1 \xrightarrow{\gamma_1} w_{(\sigma(1)} \\ v_2 \xrightarrow{\gamma_2} w_{(\sigma(2)} \\ \cdots \\ v_\ell \xrightarrow{\gamma_\ell} w_{(\sigma(\ell)}}} \mathtt{wt}(\gamma_1)\mathtt{wt}(\gamma_2)\ldots\mathtt{wt}(\gamma_\ell)\end{aligned}$$

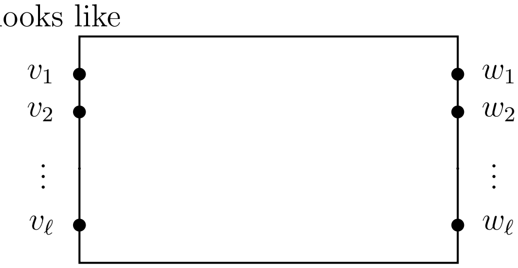
We claim then that

$$\begin{aligned}\det[P(v_i \to w_j)] &= \sum_{\sigma \in S_\ell} (-1)^\sigma \prod_{i=1}^{\ell} P(v_i \to w_{\sigma(i)}) \\ &= \sum_{\sigma \in S_\ell} (-1)^\sigma \sum_{\substack{v_1 \xrightarrow{\gamma_1} w_{\sigma(1)} \\ v_2 \xrightarrow{\gamma_2} w_{\sigma(2)} \\ \cdots \\ v_\ell \xrightarrow{\gamma_\ell} w_{\sigma(\ell)}}} \mathtt{wt}(\gamma_1)\mathtt{wt}(\gamma_2)\ldots\mathtt{wt}(\gamma_\ell) \\ &= \sum_{\sigma \in S_\ell} (-1)^\sigma \sum_{\substack{v_1 \xrightarrow{\gamma_1} w_{\sigma(1)} \\ v_2 \xrightarrow{\gamma_2} w_{\sigma(2)} \\ \cdots \\ v_\ell \xrightarrow{\gamma_\ell} w_{\sigma(\ell)} \\ \gamma_i \cap \gamma_j = \emptyset\ \forall i \neq j}} \mathtt{wt}(\gamma_1)\mathtt{wt}(\gamma_2)\ldots\mathtt{wt}(\gamma_\ell)\end{aligned}$$

since all terms which involve intersecting paths cancel.

Proof. Choose a total order on the vertices of Γ with $i < j$ if $i \to j$. If $\gamma_1, \ldots, \gamma_\ell$ is a collection of paths which are not disjoint, choose x to be the first vertex which is in two of the paths. Then choose the first i and j such that x is on γ_i ad γ_j. Then swap the portions of γ_i and γ_j which come before the paths reach x. This procedure is a sign reversing involutions. That is, the new paths have the same weight, same point x, and opposite sign, and when we apply the map again it undos itself. Thus, everything cancels except when the paths are disjoint. □

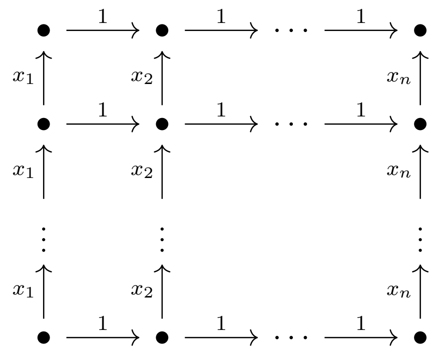
In the particular case where Γ is embedded in a disc in $\mathbb{R}^2$ and $v_1, \ldots, v_\ell, w_1, \ldots, w_\ell$ are vertices in Γ such that Γ looks like



then the only way to have disjoint paths between these vertices is to have the paths be of the form $v_i \to w_i$. In this case, we then have

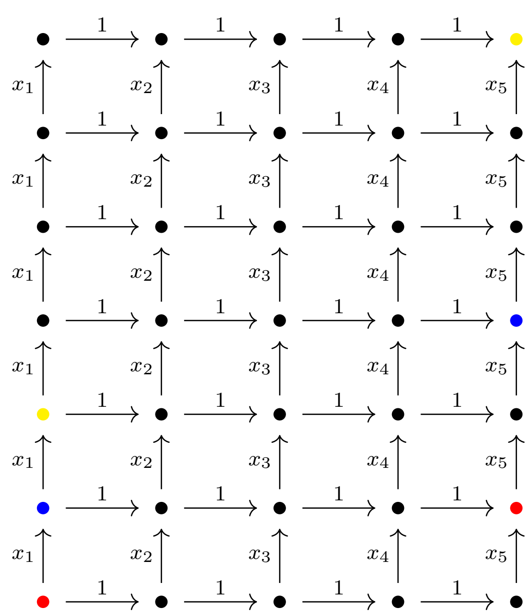
$$\det[P(v_i \to w_j)] = \sum_{\substack{v_1 \xrightarrow{\gamma_1} w_1 \\ \cdots \\ v_\ell \xrightarrow{\gamma_\ell} w_\ell \\ \gamma_i \cap \gamma_j = \emptyset}} \mathtt{wt}(\gamma_1) \dots \mathtt{wt}(\gamma_\ell)$$

We recall from last class that $h_k(x_1, \dots, x_n)$ is the generating function for directed paths in the following graph which connect the bottom left point to the top right point



where each column has k or vertical arrows. To compute $s_{\lambda_1} s_{\lambda_2} \dots s_{\lambda_\ell}$ using our above formulas, we use a graph as above with $\lambda_1 + (\ell - 1)$ vertical steps in each column. We place v_ℓ in the bottom left corner at height 0, and the remaining $v_2, \dots, v_\ell$ in above v_ℓ along the leftmost "avenue" at height $\ell - i$. We place the w_j for $j = 1, \dots, \ell$ along the rightmost avenue at height $\lambda_j + (\ell - j) = \lambda_j + v_j$.

Example. As as example, consider again the function $s_{(5,2,1)}$. To this function we associate the following grid

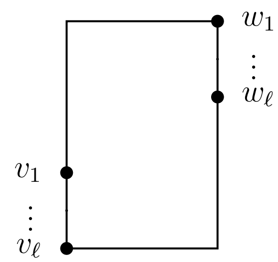


where v_1 and w_1 are the yellow vertices, v_2 and w_2 are the blue vertices, and v_3 and w_3 are the red vertices. We see that the statement of the Jacobi-Trudi identity in this case

$$s_{(5,2,1)} = \det \begin{bmatrix} h_5 & h_6 & h_7 \\ h_1 & h_2 & h_3 \\ 0 & 1 & h_1 \end{bmatrix}$$

makes sense with the paths we see in the graph. For example, the $(3,1)$ entry in the matrix, corresponding to v_1 and w_3, is zero, and there are no paths in the graph from v_1 and w_3.

In general, we note that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell$ implies in our construction that the height of w_i is strictly greater than the height of w_{i+1} for all $i = 1, \ldots, \ell - 1$. Therefore, we do get a planar picture



so the only collection of disjoint paths connecting the v_i to the w_j are the paths $v_1 \to w_1$, $v_2 \to w_2$, ... , $v_\ell \to w_\ell$. Therefore, the formula

$$\det[P(v_i \to w_j)] = \sum_{\substack{v_1 \xrightarrow{\gamma_1} w_1 \\ \cdots \\ v_\ell \xrightarrow{\gamma_\ell} w_\ell \\ \gamma_i \cap \gamma_j = \emptyset}} \mathtt{wt}(\gamma_1) \ldots \mathtt{wt}(\gamma_\ell)$$

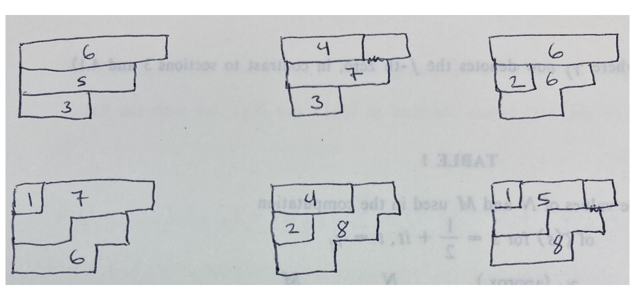
applies and yields the Jacobi-Trudi formula.

September 28 – The ratio of alternants formula

Remark. In problem 4 on the homework the goal is to have a graphical way to draw the Jacobi-Trudi identity. For example

$$s_{(6,5,3)} = \det \begin{bmatrix} h_6 & h_7 & h_8 \\ h_4 & h_5 & h_6 \\ h_1 & h_2 & h_3 \end{bmatrix} = h_{(6,5,3)} - h_{(6,6,2)} - h_{(7,4,3)} - h_{(7,6,1)} + h_{(8,4,2)} - h_{(8,5,1)}.$$

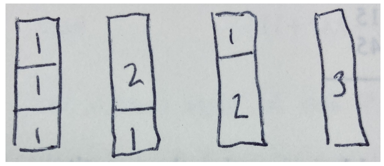
We can draw these like this:



We run into an issue the way the problem was written with the example

$$s_{(1,1,1)} = \det \begin{bmatrix} h_1 & h_2 & h_3 \\ h_0 & h_1 & h_2 \\ 0 & h_0 & h_1 \end{bmatrix} = h_{(1,1,1)} - h_{(2,1,0)} - h_{(2,1,0)} + h_{(3)}$$

We'd like this to give the pictures: This problem will be rewritten for next week.



We now turn to our main topic, the Ratio of Alternants Formula.

Definition. A function $f(x_1, \dots, x_n) \in \mathbb{Z}[x_1, \dots, x_n]$ is called **alternating** if

$$f(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}) = (-1)^\sigma f(x_1, \dots, x_n)$$

for all $\sigma \in S_n$.

Example. The Vandermonde determinant

$$\Delta(x_1, \dots, x_n) := \prod_{i<j} (x_i - x_j)$$

is alternating.

Lemma. (From Homework) The alternating polynomials are exactly $\Delta \cdot$(symmetric polynomial), i.e. elements of $\Delta \cdot \Lambda_n$.

We are looking for an analogous basis to the symmetric polynomial bases, but for the alternating polynomials. For $\alpha_1 > \alpha_2 > \cdots > \alpha_n$, set

$$A_\alpha = \det \begin{bmatrix} x_1^{\alpha_1} & x_1^{\alpha_2} \dots x_1^{\alpha_n} & & \\ x_2^{\alpha_1} & x_2^{\alpha_2} \dots x_n^{\alpha_n} & & \\ \vdots & \vdots & \dots & \vdots \\ x_n^{\alpha_1} & \dots & \dots & x_n^{\alpha_n} \end{bmatrix} = \sum_{\sigma \in S_n} (-1)^\sigma x_1^{\alpha_{\sigma(1)}} x_2^{\alpha_{\sigma(2)}} \cdots x_n^{\alpha_{\sigma(n)}}.$$

Note we want $\alpha_1 > \alpha_2 > \cdots > \alpha_n$, if $\alpha_i = \alpha_{i+1}$ this is 0.

Then the alternating polynomials are $\bigoplus_{\alpha_1 > \alpha_2 > \cdots > \alpha_n} \mathbb{Z} A_\alpha$. So

$$\Delta \cdot \Lambda_n = \bigoplus_{\alpha_1 > \cdots > \alpha_n} \mathbb{Z} \cdot A_\alpha$$

$$\Lambda_n = \bigoplus_{\alpha_1 > \cdots > \alpha_n} \mathbb{Z} \cdot \frac{A_\alpha}{\Delta}$$

Theorem.

$$s_{(\lambda_1, \lambda_2, \dots, \lambda_n)}(x_1, \dots, x_n) = \frac{A_{\lambda_1 + n - 1, \lambda_2 + n - 2, \dots, \lambda_n}}{\Delta}$$

Remark. The theorem above could have been the definition of s_λ, see Friday for an alternate order of the definitions we've seen so far.

A special case of this theorem states that

$$\begin{aligned}1 = S_{\emptyset} =& \frac{A_{(n-1,n-2,\ldots,2,1,0)}}{\Delta}\\ =& \frac{\det|x_i^{n-j}|_{1\le i,j\le n}}{\prod_{i<j}(x_i - x_j)}.\end{aligned}$$

In other words, $\det(x_i^{n-j}) = \prod_{i<j}(x_i - x_j)$, which is called Vandermonde's identity.

Proof. (Proof of special case of theorem.) The degree of $\frac{A_{(n-1,n-2,\ldots,2,1,0)}}{\Delta}$ is

$$0 + 1 + \cdots + (n-1) - \binom{n}{2},$$

which is 0, so this quantity is an integer. We can look at the coefficient of $x_1^{n-1}x_2^{n-2}\ldots x_{n-1}$ and see that it is 1.

□

We will define

$$\rho = (n-1, n-2, \ldots, 2, 1, 0), \text{ with } \rho_j = n - j.$$

We just showed $A_\rho = \prod_{i<j}(x_i - x_j)$. Our next goal is to show that

$$\frac{A_{\lambda+\rho}}{A_\rho} = s_\lambda$$

or in other words that

$$A_{\lambda+\rho} = s_\lambda \cdot A_\rho.$$

We want to show that

$$\sum_{\tau\in S_n}(-1)^\tau x_1^{\lambda_{\tau(1)}+\rho_{\tau(1)}} x_2^{\lambda_{\tau(2)}+\rho_{\tau(2)}} \cdots x_n^{\lambda_{\tau(n)}+\rho_{\tau(n)}} = \left(\sum_{\sigma\in S_n}(-1)^\sigma x_1^{\rho_{\sigma(1)}} x_2^{\rho_{\sigma(2)}} \cdots x_n^{\rho_{\sigma(n)}}\right)\left(\sum_\mu K_{\lambda\mu} m_\mu\right)$$

or we can write the far right term on the right hand side as

$$\left(\sum_{\mu_1,\ldots,\mu_n\in\mathbb{Z}^n_{\ge 0}} K_{\lambda,\mathrm{sort}(\mu)} x_1^{\mu_1} x_2^{\mu_2} \cdots x_n^{\mu_n}.\right)$$

On the right hand side the coefficient of $x_1^{\nu_1+\rho_1} x_2^{\nu_2+\rho_2} \cdots x_n^{\nu_n+\rho_n}$ is

$$\sum \left((-1)^\sigma K_{\lambda,\mathrm{sort}(\mu)} \,\middle|\, \nu_j + \rho_j = \rho_{\sigma(j)} + \mu_j\right) = \sum_{\mu_1\ge\mu_2\ge\cdots\ge\mu_n} K_{\lambda\mu} \sum_\sigma \left((-1)^\sigma \,\middle|\, (\nu_j + \rho_j - \rho_{\sigma(j)}) \sim (\mu)\right)$$

where $\sim$ means "if the left thing is a permutation of the right thing".

Set

$$L_{\mu\nu} := \sum_\sigma \left((-1)^\sigma \,\middle|\, (\nu_j + \rho_j - \rho_{\sigma(j)}) \sim (\mu)\right)$$

then the coefficient above is equal to $\sum K_{\lambda\mu} L_{\mu\nu}$. Thus we want to know where $K \cdot L = Id$. That is where

$$L_{\mu\nu} = \sum_\sigma \left((-1)^\delta \middle| v_j + \rho_j - \rho_{\sigma(j)} \sim \mu\right).$$

We will compare to K^{-1} from the Jacobi-Trudi identity.

Jacobi-Trudi says

$$\begin{aligned} s_\lambda &= \det \begin{bmatrix} h_{\lambda_1} & h_{\lambda_1+1} & \dots & \dots \\ \vdots & \vdots & \dots & \vdots \\ \dots & \dots & \dots & h_{\lambda_n} \end{bmatrix} \\ &= \sum_\sigma (-1)^\sigma h_{\lambda_j+j-\sigma(j)} \\ &= \sum_\sigma (-1)^\sigma h_{\lambda-\rho+\sigma(\rho)}. \end{aligned}$$

The coefficient of h_μ is

$$\sum_\sigma \left((-1)^\sigma \,\middle|\, \lambda-\rho+\sigma(\rho)\sim\mu\right) = L_{\mu\lambda}.$$

By Jacobi-Trudi,

$$\sum_\mu K_{\nu\mu}L_{\mu\lambda} = \begin{cases} 1 & \lambda=\nu \\ 0 & \lambda\neq\nu \end{cases}$$

So we get

$$L_{\mu\nu} = \sum_\sigma \left((-1)^\delta \middle| v_j+\rho_j-\rho_{\sigma(j)}\sim\mu\right)$$

as desired.

September 30 – Starting from the ratio of alternants

The goal for today is to follow Cauchy's definition of the Schur polynomials to derive Jacobi-Trudi identity and Cauchy's identity.

As before, let $\rho_n = (n-1, n-2, \dots, 2, 1, 0)$. Let $A_{\alpha_1\alpha_2\dots\alpha_n} = \det\left[x_i^{\alpha_j}\right]_{1\leq i,j\neq n}$. What happens if we define $s_\lambda(x_1, x_2, \cdots, x_n)$ to be $\dfrac{A_{\lambda+\rho_n}(x_1,x_2,\cdots,x_n)}{A_{\rho_n}(x_1,x_2,\cdots,x_n)}$? The ratio is in Λ_n. It is not clear that this ratio stabilizes as n grows, but we will know that it is true once we prove the Jacobi-Trudi identity.

Consider the following n by ∞ matrix $M(\underline{x})$.

$$M(\underline{x}) = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^i & \cdots & \\ 1 & x_2 & x_2^2 & \cdots & x_2^i & \cdots & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \\ 1 & x_n & x_n^2 & \cdots & x_n^i & \cdots & \end{bmatrix}$$

Then $A_{\lambda+\rho}$ is the n by n minor of $M(\underline{x})$ corresponding to columns indexed by λ_i+n-i.

For an invertible n by n matrix g, the n by n minors of $gM(\underline{x})$ are $\det(g)A_{\lambda+\rho}$. So ratio of minors in M is the same as ratio of minors in gM. We prove Jacobi-Trudi identity by finding an invertible g such that ratio of minors in gM gives us the desired expression in the $h's$.

Suppose that x_1, x_2,..., x_n are distinct. Then the rows of $M(\underline{x})$ span all solutions $[a_0, a_1, \cdots]$ to the recursion

$$a_{k+n} - e_1a_{k+n-1} + e_2a_{k+n-2}\cdots + (-1)^n a_k.$$

We can also express this in terms of generating functions: The solutions to this recursion are the coefficients of generating functions $\sum a_k t^k$ of the form $\frac{\text{polynomial of degree} \leq n-1}{\prod(x-t_j)}$. The rows of $M(\underline{x})$, specifically, are the rational functions $\frac{1}{1-tx_j}$.

Now, another basis for the vector space of rational functions of the form $\frac{\text{polynomial of degree} \leq n-1}{\prod(1-tx_j)}$ is the rational functions $\frac{x^k}{\prod(1-tx_j)}$, for $0 \leq k \leq n-1$. We have $\frac{x^k}{\prod(1-tx_j)} = t^k + h_1 t^{k+1} + h_2 t^{k+2} + \cdots$. So the row span of $M(\underline{x})$ is the same as the row span of the matrix

$$H := \begin{bmatrix} 1 & h_1 & h_2 & h_3 & \cdots & \cdots \\ 0 & 1 & h_1 & h_2 & \cdots & \cdots \\ 0 & 0 & 1 & h_1 & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \cdots \\ 0 & 0 & 0 & \cdots\cdots & 1 & h_1 \cdots \end{bmatrix}.$$

By our previous argument, the quotient of corresponding n by n minors of M is equal to the quotient of corresponding n by n minors of H. And this is the statement of Jacobi-Trudi identity because the first n by n minor of H is upper triangular with all diagonal entries equal to 1, and thus has determinant 1.

Now, we provide Cauchy's proof of Cauchy's identity. First, we write

$$\left[\frac{1}{1-x_i y_j}\right]_{1 \leq i,j \leq n} = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & \\ 1 & x_2 & x_2^2 & \cdots & \\ \vdots & \vdots & \vdots & \ddots & \\ 1 & x_n & x_n^2 & \cdots & \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ y_1 & y_2 & y_3 & \cdots & y_n \\ y_1^2 & y_2^2 & y_3^2 & \cdots & y_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{bmatrix}$$

So $\det \left[\frac{1}{1-x_i y_j}\right] = \sum_{\alpha_1 > \alpha_2 > \cdots > \alpha_n} \det M(\underline{x})_\alpha \det M(\underline{y})_\alpha^T$, where $M(\underline{x})_\alpha$ denotes the minor corresponding to columns $\alpha_1 > \alpha_2 > \cdots > \alpha_n$. On the right hand side, going through strictly decreasing n-tuples $\alpha_1 > \alpha_2 > \cdots > \alpha_n$ is the same as going through non-increasing n-tuples $\alpha_1 - 1 \geq \alpha_2 - 2 \geq \cdots \geq \alpha_n - n$, which corresponds to a parition λ. So the right hand side is equal to $\sum_\lambda A_{\lambda+\rho}(\underline{x}) A_{\lambda+\rho}(\underline{y})$.

What we need to show is that

$$\left[\frac{1}{1-x_i y_j}\right]_{1 \leq i,j \leq n} = \frac{A_\rho(x) A_\rho(y)}{\prod(1-x_i y_j)}.$$

This follows from a determinantal identity proved by Cauchy in another source, 26 years later:

$$\left[\frac{1}{u_i + v_j}\right]_{1 \leq i,j \leq n} = \frac{\prod_{i<j}(u_i - u_j) \prod_{i<j}(v_i - v_j)}{\prod(u_i + v_j)}.$$

To prove this, note that the left hand side must be of the form $\frac{f(u_1,\ldots,u_n,v_1,\ldots,v_n)}{\prod(u_i+v_j)}$ for some polynomial f. The polynomial f must vanish whenever $u_i = u_j$ and whenever $v_i = v_j$, so it is divisible by $\prod_{i<j}(u_i - u_j) \prod_{i<j}(v_i - v_j)$. But also, computing degrees, the polynomial f must be homogeneous of degree $n(n-1)$, so f is a scalar multiple of $\prod_{i<j}(u_i - u_j) \prod_{i<j}(v_i - v_j)$, and a little work checks that the scalar is 1.

October 3-October 7 Introduction to Representation Theory of Finite Groups (IBL)

We solved problems about representation theory. Here they are!

Let G be a group and k a field. A ***representation*** of G is a k-vector space V and a group homomorphism $\rho : G \to \mathrm{GL}(V)$. In other words, a representation is an action of G on V by linear maps. We will often denote the representation as "V", and we will sometimes write ρ_V for ρ.

Given two representations (V_1, ρ_1) and (V_2, ρ_2), the direct sum representation is the action of G on $V_1 \oplus V_2$ where g acts by $\left[\begin{smallmatrix} \rho_1(g) & 0 \\ 0 & \rho_2(g) \end{smallmatrix}\right]$. A representation V is called ***indecomposable*** if $V \neq 0$ and we cannot write $V = V_1 \oplus V_2$ for V_1, $V_2 \neq 0$. It is easy to see that any finite dimensional representation is a direct sum of indecomposable representations.

Given a representation V, a ***subrepresentation*** of V is a vector subspace U of V such that G maps U to U. A representation V is called ***simple*** if $V \neq 0$ and the only subrepresentations of V are V and $\{0\}$.

Given two G-representations U and V, a ***morphism of G-representations*** is a linear map $\phi : U \to V$ obeying

$$\phi(g * u) = g * \phi(u) \quad \text{for } g \in G,\ u \in U.$$

Problem 1. Let U and V be G-representations and let $\phi : U \to V$ be a morphism of G-representations. We write $\mathrm{Hom}_G(U, V)$ for the space of morphism of G-representations from U to V. Show that $\mathrm{Ker}(\phi)$ and $\mathrm{Im}(\phi)$ are subrepresentations of U and V respectively.

Our first goal is to prove:

Theorem (Maschke's Theorem, first version)**.** Let k have characteristic zero and let G be a finite group. Then V is simple if and only if it is indecomposable..

Problem 2. Show that, if V is not indecomposable then V is not simple.

Problem 3. Let k have characteristic zero, let G be a finite group and let V be a finite dimensional nonzero representation of G. Suppose that V is not simple, so that $U \subset V$ is a nontrivial subrepresentation. Let $\eta : V \to U$ be a linear map with $\eta(u) = u$ for $u \in U$. Define

$$\pi(v) = \tfrac{1}{|G|} \sum_{g \in G} \left(\rho(g) \circ \eta \circ \rho(g)^{-1}\right)(v).$$

(1) Show that π is a morphism of G-representations (even though η may not be).
(2) Show that π has image U and that $\pi(u) = u$ for $u \in U$.
(3) Show that $V = U \oplus \mathrm{Ker}(\pi)$. Deduce that V is not indecomposable.

Thus we conclude:

Theorem (Maschke's Theorem, second version)**.** Let k have characteristic zero, let G be a finite group and let V be a finite dimensional representation of G. Then V is a direct sum of simple representations.

We next will consider the question of whether the decomposition of V as a direct sum of simple representations is unique.

Problem 4 (Schur's Lemma, first version)**.** Let U and V be simple representations of G and let $\phi : U \to V$ be a morphism of G-representations. Show that either $\phi = 0$ or else ϕ is an isomorphism.

Problem 5. Let $U_1, U_2, \ldots, U_k$ be a collection of nonisomorphic simple G-representations and let $V = \bigoplus_{j=1}^k U_j^{\oplus c_j}$ for some nonnegative integers c_j.

(1) Show that $\dim_k \mathrm{Hom}_G(U_j, V) = c_j \dim_k \mathrm{Hom}_G(U_j, U_j)$.
(2) Deduce that, if $\bigoplus_{j=1}^k U_j^{\oplus c_j} \cong \bigoplus_{j=1}^k U_j^{\oplus d_j}$, then $(c_1, c_2, \ldots, c_k) = (d_1, d_2, \ldots, d_k)$. (Nitpicker alert: Did you use that $U_j \neq 0$?)

We thus deduce:

Theorem. Let G be a group and let V be a finite dimensional representation of G which can be written as $V = \bigoplus_{j=1}^k U_j^{\oplus c_j}$ for a collection of nonisomorphic simple representations $U_1, U_2, \ldots, U_k$ of G. Then the U_j (up to isomorphism) and the multiplicities c_j are uniquely determined by V.

This theorem doesn't require that G is finite or that k has characteristic zero, but without those hypotheses, Maschke's theorem does not apply, so there may be many representations V which cannot be written in this way.

We now see how things are better if k is algebraically closed.

Problem 6. Let k be an algebraically closed field and let U be a simple G-representation. Let $\phi : U \to U$ be a morphism of G-representations. Show that ϕ is a scalar multiple of the identity. (Hint: Let λ be an eigenvalue of ϕ, and apply Problem 4 to the linear map $\phi - \lambda \mathrm{Id}$.)

Problem 7. Let k be an algebraically closed field. Let $U_1, U_2, \ldots, U_k$ be a collection of nonisomorphic simple G-representations and let $V = \bigoplus_{j=1}^k U_j^{\oplus c_j}$ for some nonnegative integers c_j. Show that $\dim_k \mathrm{Hom}_G(U_j, V) = c_j$.

We now introduce character theory into the story. Given a representation V, the ***character*** of V is the function $G \to k$ defined by

$$\chi_V(g) = \mathrm{Tr}(\rho_V(g)).$$

Problem 8. Suppose that g_1 and g_2 are conjugate elements of G, meaning that $g_2 = hg_1h^{-1}$ for some $h \in G$. For any character χ of G, show that $\chi(g_1) = \chi(g_2)$.

Problem 9. Check that $\chi_{V_1 \oplus V_2}(g) = \chi_{V_1}(g) + \chi_{V_2}(g)$.

Problem 10. Let k have characteristic zero, let G be a finite group, and let V be a finite-dimensional representation of G. Let

$$V^G = \{v \in V : g(v) = v \ \forall g \in G\}.$$

Define $\pi : V \to V$ by

$$\pi(v) = \tfrac{1}{|G|} \sum_{g \in G} g * v.$$

(1) Show that $\mathrm{Im}(\pi) = V^G$ and $V = V^G \oplus \mathrm{Ker}(\pi)$.
(2) Show that $\mathrm{Tr}(\pi) = \dim V^G$.
(3) Deduce that

$$\dim V^G = \tfrac{1}{|G|} \sum_{g \in G} \chi_V(g).$$

Problem 11. Let k have characteristic zero, let G be a finite group, and let U and V be finite-dimensional representations of G. Let $\mathrm{Hom}(U, V)$ be the vector space of k-linear maps $U \to V$ (all maps, not just the morphisms of G-representations). Define an action of G on $\mathrm{Hom}(U, V)$ by

$$(g * \phi)(u) = \left(\rho_V(g) \circ \phi \circ \rho_U(g^{-1})\right)(u).$$

(1) Show that $\mathrm{Hom}_G(U, V) = \mathrm{Hom}(U, V)^G$ for this G-action.
(2) Show that

$$\dim_k \mathrm{Hom}_G(U, V) = \tfrac{1}{|G|} \sum_{g \in G} \chi_U(g^{-1}) \chi_V(g).$$

We can improve the formula above in the case that $k = \mathbb{C}$.

Problem 12. Let G be a finite group, and let V be a finite dimensional representation of G over the field $\mathbb{C}$. Let $g \in G$.

(1) Show that all the eigenvalues of $\rho_V(g)$ are roots of unity.
(2) Show that $\chi_V(g^{-1}) = \overline{\chi_V(g)}$, where $\overline{z}$ is the complex conjugate of z.

Thus, let G be a finite group and let p and q be $\mathbb{C}$-valued functions on g. Define

$$\langle p, q \rangle := \tfrac{1}{|G|} \sum_{g \in G} \overline{p(g)} q(g).$$

This is a positive definite Hermitian inner product. Then our results above specialize to say:

Theorem. With the above assumptions and notations, let U and V be two representations of G over $\mathbb{C}$. Then

$$\dim \mathrm{Hom}_G(U, V) = \langle \chi_U, \ \chi_V \rangle.$$

In particular, if U is a simple representation, then $\langle \chi_U, \ \chi_V \rangle$ is the multiplicity of the summand U in V.

October 10 - Haar measure and Representation Theory of Compact Groups

Given a compact Hausdorff topological group G we recall the Haar measure which gives us the following map:

$$\int_G : \{\text{continuous Real (or Complex) valued functions on } G\} \to \mathbb{R}(\text{or}\mathbb{C})$$

. This map has the following properties.

(1) $\int_G f$ is linear in f

(2) If $f_n \to f$ uniformly then $\int_G f_n \to \int_G f$

(3) $\int_G f(g)dg = \int_G f(hg)dg = \int_G f(gh)dg$

(4) $\int_G 1 = 1$

(5) If $f > 0$ then $\int_G f > 0$

If G is a compact Lie group we have $\int_G f = \int_G f\omega$ where ω is a left and right invariant volume form.

For finite G we have already seen such a map where $\int_G f = \frac{\sum_{g \in G} f(g)}{|G|}$. The Haar measure gives us a way to average in a way similar to what we did for finite groups. In particular Character Theory works the same way as before.

Theorem. If V is a continuous finite dimensional representation of G and U is a subrepresentation, then there is a subrepresentation T such that $V = U \oplus T$

Proof. Let $\eta : V \to U$ be a projection. Define $\pi(v) = \int_G g \star \eta(g^{-1} \star v)$. Because of the invariance of the haar measure, as in the finite case, π is a G representation homomorphism, $\pi^2 = \pi$ and thus $V = U \oplus \ker(\pi)$. □

Thus we have that V is simple if and only if it is indecomposible. In analogy to the finite group case we also have.

Theorem. $\dim(\mathrm{Hom}_G(U, V)) = \int_G \overline{\chi_U(g)}\chi_V(g) = \int_G \chi_U(g^{-1})\chi_V(g)$

Proof. The following facts follow similar to the finite group case

(1) $\dim(W^G) = \int_G \chi_W(g)$
(2) $\mathrm{Hom}_G(U, V) = \mathrm{Hom}(U, V)^G$
(3) $\chi_{Hom_G(U,V)}(g) = \chi_U(g^{-1})\chi_V(g)$

To finish the proof we need the fact that all the eigenvalues of $\rho(g)$ have norm 1. In the finite case we used the fact that $|G|$ has finite order but now we don't have any such restriction. But the fact still holds and we prove the following lemma

Lemma. Let $\rho : G \to GL(V)$ be a continuous finite dimensional representation. Then there exists a positive definite symmetric/hermitian form for which every $\rho(g)$ is orthogonal or unitary.

Proof. If $(,)$ is any positive definite symmetric/hermitian form we can define.

$$\langle u, v\rangle = \int_G (gu, gv)dg$$

By the properties of the Haar measure this bilinear form is positive definite and it is still symmetric/Hermitian. In addition this bilinear form is also G invariant and thus $g \in G$ acts by an orthogonal/unitary operator. $\square$

In particular all $\rho(g)$ are diagonalizable and have eigenvalues of norm 1. Thus $\chi_U(g^{-1}) = \overline{\chi_U(g)}$ and the Theorem holds. $\square$

We now look at the following basic example of $GL(1) = \{[z] : z \in \mathbb{C}^*\}$. $GL(1)$ has a compact subgroup $U(1) = \{[z] : |z| = 1\}$.
For each integer k we have the one dimensional representation $\rho_k : GL_1 \to GL_1$ given by $[z] \to [z^k]$. We have the following remarkable decomposition of algebraic representations.

Theorem. Let V be an algebraic representation of GL_1 of dimension n. Then $V = \oplus_{k\in\mathbb{Z}} V_k$ where all but finitely many V_k are zero and $z \in GL_1$ acts by z^k on V_k.

Proof. Let $\rho(z) = \sum_{k\in\mathbb{Z}} p_k z^k$ where $p_k \in Hom(V, V)$ and finitely many p_k are non zero.
The condition $\rho(1) = I_n$ translates into the equation

$$\sum_k p_k = I_n.$$

The condition $\rho(X)\rho(Y) = \rho(XY)$ translates into the equation $\sum_{i,j} p_i p_j X^i Y^j = \sum_k p_k X^k Y^k$. Taking coefficients of $X^I Y^j$ on both sides, we deduce:

$$p_i p_j = \begin{cases} p_i & i = j \\ 0 & \text{otherwise.} \end{cases}$$

In other words, the p_i are mutually orthogonal idempotents. Thus, by HW6, Problem 3, we know that $V = \oplus_k V_k$ where p_k projects V to V_k. We claim that $\oplus V_k$ is the desired decomposition because if $v \in V_k$, then we have $\rho(z)v = \sum p_j(z^j v) = z^k v$. $\square$

An analogous result holds for U_1.

Theorem. Let V be a continuous representation of U_1 of dimension n. Then $V = \oplus_{k\in\mathbb{Z}} V_k$ where all but finitely many V_k are zero and $z \in U_1$ acts by z^k on V_k.

Proof. Using the Fourier decomposition of $\rho(z)$, let $\rho(e^{i\theta}) = \sum_{k\in\mathbb{Z}} p_k e^{ik\theta}$ where $p_k \in Hom(V, V)$.

$$\rho(1) = I_n \implies \sum_k p_k = I_n$$

and

$$\rho(e^{iX})\rho(e^{iY}) = \rho(e^{i(X+Y)}) \implies \sum_{a,b} p_a p_b e^{iaX} e^{ibY} = \sum_k p_c e^{ic(X+Y)}$$

Thus $p_a p_b = p_a$ if $a = b$ and 0 otherwise. In order to apply HW6, Problem 3 we need to know that only finitely p_c are non zero and that follows because $p_c^2 = p_c$ and so all eigenvalues of p_c are either 0 or 1. Since $\sum_k p_c = I_n$ $\sum_k tr(p_c) = n$ and so only finitely many p_c are non-zero.
Thus we know that $V = \oplus_k V_k$ where p_k projects V to V_k. We claim that $\oplus V_k$ is the desired decomposition because if $v \in V_k$ $\rho(z)v = \sum p_j(z^j v) = z^k v$. $\square$

Note: We can use the idea of taking the trace to conclude that the p_k are finite and replace the algebraic representations in the GL_1 case with holomorphic representations.

In the case of finite dimensional finite group representations over G we have the following result that we will prove later.

Theorem. $|G| = \sum_{V_i \text{represents an isomorphism class of simples}} (\dim(V_i))^2$, infact something stronger holds $\mathbb{C}G = \bigoplus_i \operatorname{Hom}(V_i, V_i)$ as $G \times G$ representations where $\mathbb{C}G$ is the vector space with basis elements corresponding to elements in G.

For compact groups, we want to show

$$\text{“}O(G)\text{”} = \text{“}\bigoplus\text{”}\operatorname{Hom}(V_i, V_i).$$

So we want nice functions on G which we will call $O(G)$ and want to be able to talk about directs sums. Fourier Analysis gives us multiple answers. We will look at one such way in the next lecture.

October 12 – Matrix coefficients

We started with a short discussion about Problem 2 from Problem Set 6. In that problem, we are assuming that the set X on which G acts is finite.

In general, if $G \curvearrowright X$, this can be viewed as a representation, as follows:

In terms of matrices, we have $G \to S_X \hookrightarrow \mathrm{GL}_X$. The matrix for g is the permutation matrix given by g. For example, for the action $S_3 \curvearrowright \{1, 2, 3\}$, the matrix for the cycle (123) is given by $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.

Another way to define this representation is as follows: Let V be the free vector space on X. Then $g \cdot e_x = e_{g \cdot x}$.

When $|X| < \infty$, then $V \cong V^\vee$ (the dual space), and we can also have the following description: $V^\vee$ is the set of functions $X \to k$, with $(g * f)(x) = f(g^{-1}x)$.

Now we come back to the topic for today's class.

If G is a finite group, then $G \times G$ acts on G as follows:

$$(\sigma, \tau) * g = \sigma g \tau^{-1}$$

We have the following theorem from the representation theory of finite groups:

Theorem.

$$\begin{aligned} \mathbb{C}G = &\text{ Functions } (G \to \mathbb{C}) \\ \cong &\bigoplus_{V \text{ an irrep}} V \otimes V^\vee \\ \cong &\bigoplus \operatorname{Hom}(V, V) \\ \cong &\bigoplus \operatorname{Hom}(V, V)^\vee \end{aligned}$$

Our goal is to prove an analogue of this when G is a compact Hausdorff topological group. That is, we will show that: $\{$ nice functions $G \to CC\} \cong \oplus_V \mathrm{Hom}(V,V)^\vee$ as $(G \times G)$-representations.

Remark. $\mathrm{Hom}(V,V) \cong \mathrm{Hom}(V,V)^\vee$ as G-representations. An isomorphism is given by $A \mapsto [M \mapsto \mathrm{Tr}(AM)]$. We will stick to writing the dual for aesthetic reasons.

An intermediate goal, which is today's goal, is to make precise what we mean by 'nice functions'.

Fix a topological group G (we don't need compactness yet). Let $\rho : G \to \mathrm{GL}(V)$ be a continuous finite-dimensional representation. Let $\lambda \in \mathrm{Hom}(V,V)^\vee$.

So we have:

$$G \xrightarrow{\rho} \mathrm{GL}(V) \subset \mathrm{Hom}(V,V) \xrightarrow{\lambda} \mathbb{C}.$$

We'll define $f : G \to \mathbb{C}$ to be a ***matrix coefficient*** if $f = \lambda \circ \rho$ for some (V, ρ, λ) as above.

In other words, f is a linear combination of the functions $(\rho(g))_{ij}$.

Example. For a finite group G, $G \curvearrowright G$ by $g * h = gh$.

Then $\rho(g)_{h_1 h_2} = \begin{cases} 1 & h_1 = gh_2 \\ 0 & \text{otherwise.} \end{cases}$ In particular, $\rho(g)_{he} = \begin{cases} 1 & g = h \\ 0 & \text{otherwise.} \end{cases}$. So all functions $G \to \mathbb{C}$ are matrix coefficients.

Example. $G = S^1 = \mathbb{R}/2\pi\mathbb{Z}$.

Last time, we saw that any finite dimensional representation of G is a direct sum of 1-dimensional irreducibles ρ_k, given by: $\rho_k(\theta) = [e^{ik\theta}]$

So Matrix Coefficients are the trigonometric polynomials $\sum_{k=-N}^{N} a_k e^{ik\theta}$

Definition. Let $C_0(G) = \{\text{continuous functions } G \to \mathbb{C}\}$

Then $G \times G \curvearrowright C_0(G)$, with action given by $[(h_1, h_2) * f](g) = f(h_1^{-1} g h_2)$.

Theorem. Suppose $f \in C_0(G)$. Then the following are equivalent:

(1) f is a matrix coefficient, i.e. $\exists (V, \rho, \lambda)$ with $f = \lambda \circ \rho$.

$(2)_{LR}$ The $G \times G$ orbit of f spans a finite dimensional subspace of $C_0(G)$.

$(2)_L$ The $G \times \{1\}$ orbit of f spans a finite dimensional subspace of $C_0(G)$.

$(2)_R$ The $\{1\} \times G$ orbit of f spans a finite dimensional subspace of $C_0(G)$.

Proof. The implications $(2)_{LR} \implies (2)_L$ and $(2)_{LR} \implies (2)_R$ are easy.

Let's prove $(1) \implies (2)_{LR}$. We have some (V, ρ, λ) so that $f = \lambda(\rho(g))$. Elements in the $G \times G$ orbit of f look like:

$$\begin{aligned} [(h_1, h_2) * f](g) &= f(h_1^{-1} g h_2) \\ &= \lambda(\rho(h_1^{-1} g h_2)) \\ &= \lambda(\rho(h_1)^{-1} \rho(g) \rho(h_2)) \end{aligned}$$

For $x \in \mathrm{Hom}(V,V)$, the map $x \mapsto \lambda(\rho(h_1)^{-1} x \rho(h_2))$ is a linear functional, i.e. this map is in $\mathrm{Hom}(V,V)^\vee$.

So every function $(h_1, h_2) * f$ is of the form $\lambda' \circ \rho$ for some $\lambda' \in \mathrm{Hom}(V, V)^\vee$. Now the result follows from the fact that $\dim(\mathrm{Hom}(V, V)^\vee) < \infty$.

Now we'll prove $(2)_R \implies (2)_{LR}$. The proof of $(2)_L \implies (2)_{LR}$ is analogous.

So we know that $\mathrm{Span}[(1, h) * f] < \infty$. We want to find (V, ρ, λ) so that $f = \lambda \circ \rho$. Set $V = \mathrm{Span}[(1, h) * f]$. We can define $\rho : G \to \mathrm{GL}(V)$ given by $g \mapsto [(1, h) * f \mapsto (1, gh) * f]$. In other words, ρ sends $g \mapsto [p \mapsto (1, g) * p]$. We need to check that ρ is continuous.

Since V is finite dimensional, we can pick a basis $(1, \sigma_1) * f, (1, \sigma_2) * f, \ldots, (1, \sigma_k) * f$ for V, and we can also assume $\sigma_1 = 1$.

For $\tau \in G$, let $\epsilon_\tau : V \to \mathbb{C}$ be the linear functional $f \mapsto f(\tau)$. Since V is a vector space of functions on G, the dual $V^\vee$ is spanned by the ϵ_τ. Let $\epsilon_{\tau_1}, \epsilon_{\tau_2}, \ldots, \epsilon_{\tau_k}$ be a basis of $V^\vee$, and assume that $\tau_1 = 1$.

We have

$$\begin{aligned} \epsilon_{\tau_j}(\rho(g)[(1, \sigma_i) * f]) &= \epsilon_{\tau_j}[(1, g) * (1, \sigma_i) * f] \\ &= [(1, g) * (1, \sigma_i) * f](\tau_j) \\ &= f(\tau_j \sigma_i g) \end{aligned}$$

Set $S = [f(\sigma_i \tau_j)]_{ij}$. Then $\rho(g)$ written in the $(1, \sigma_i) * f$ basis is $S^{-1}[f(\tau_j \sigma_i g)]$, which is continuous in g.

We also need λ so that $\lambda(\rho(g)) = f(g)$. Note that:

$$\begin{aligned} f(g) &= [(1, g) * f](1) \\ &= \epsilon_{\tau_1}((1, g) * f) \\ &= \epsilon_{\tau_1}((1, g) * (1, \sigma_1) * f) \end{aligned}$$

So f is $g \mapsto f(\tau_j \sigma_i g)$ for $i = j = 1$, i.e. f extracts the top left matrix entry of the matrix $[f(\tau_j \sigma_i g)] = S \cdot \rho(g)$. Thus f is a linear combination of the matrix entries of $\rho(g)$, which is what we wanted. $\square$

October 14: The (weak) Peter-Weyl theorem

Suppose G is a topological group (no compactness assumption), we have defined

$$\mathcal{O}(G) = \{\text{matrix coefficient of } G\},$$

which is the set of nice $\mathbb{C}$-valued functions $f : G \to \mathbb{C}$.

One way to define nice functions is that

$$\left\{ f : G \to \mathbb{C} \;\middle|\; \begin{array}{c} f \text{ is of the form } \lambda \circ \rho_V \text{ where} \\ \rho_V : G \to \mathrm{GL}(V) \text{ is a continuous map} \\ \text{and } \lambda : \mathrm{End}(V) \to \mathbb{C} \text{ is linear} \end{array} \right\}$$

Now we suppose that G is compact and Hausdorff. So G has Haar measure. Let V_i be a collection of representatives for isomorphic class of G-irreducible representations (which is continuous and finite dimensional).

For each irreducible representation V_i, we have the linear map

$$\mathrm{End}(V_i)^\vee \to \mathcal{O}(G).$$

Consequently, we can construct the Peter-Weyl map through direct sum

$$\boldsymbol{PW} : \bigoplus \mathrm{End}(V_i)^\vee \to \mathcal{O}(G).$$

Theorem (Weak Peter-Weyl)**.** The Peter-Weyl map

$$\boldsymbol{PW} : \bigoplus \mathrm{End}(V_i)^\vee \to \mathcal{O}(G)$$

is an isomorphism of $(G \times G)$-representations.

To show this, we need one more lemma:

Lemma. Let G and H be compact Hausdorff topological group and U and V be their simple G-representation and H-representation respectively.

Let $G \times H \curvearrowright \mathrm{Hom}(U, V)$ by

$$[(g, h) * \phi](\mathbf{u}) = h * \phi(g^{-1} * \mathbf{u}).$$

Then (U, V) is a simple $(G \times H)$-representation.

This is also true for groups and $\mathbb{C}$-representations. But to use the proof below we need the restrictions.

Slick character proof. The Haar measure on $G \times H$ is the product of Haar measures on G and H. We need to show that

$$\left\langle \chi_{\mathrm{Hom}(U,V)}, \chi_{\mathrm{Hom}(U,V)} \right\rangle = 1.$$

We have

$$\begin{aligned}
\left\langle \chi_{\mathrm{Hom}(U,V)}, \chi_{\mathrm{Hom}(U,V)} \right\rangle &= \int_{G\times H} \overline{\chi_{\mathrm{Hom}(U,V)}(g, h)} \chi_{\mathrm{Hom}(U,V)}(g, h) \\
&= \int_{G\times H} \overline{\overline{\chi_V(g)} \chi_V(h)} \overline{\chi_U(g)} \chi_V(h) \\
&= \int_G \overline{\chi_U(g)} \chi_U(g) \int_H \overline{\chi_V(h)} \chi_V(h) \\
&= 1 \cdot 1 = 1
\end{aligned}$$

$\square$

Now we can go on to prove weak Peter-Weyl theorem:

Proof. So the Peter-Weyl map

$$\boldsymbol{PW} : \bigoplus \mathrm{End}(V_i)^\vee \to \mathcal{O}(G)$$

Injective: This is clearly a map of $(G \times G)$-representations. By Lemma, $\mathrm{End}(V_i)^\vee \cong \mathrm{End}(V_i) \cong \mathrm{Hom}(V_i, V_i)$ is simple as $(G \times G)$-representations. And $\mathrm{End}(V_i) \ncong \mathrm{End}(V_j)$ for $i \neq j$ as $(G \times G)$-representations.

As $(\{1\} \times G)$-representations, $\mathrm{End}(V_i) \cong V_i^{\bigoplus \dim V_i}$. So $\ker(\boldsymbol{PW})$ is $\bigoplus_{i \in S} \mathrm{End}(V_i)^\vee$ for some set S of simple representations. But $\mathrm{End}(V_i)^\vee \to \mathcal{O}(G)$ is not 0 map. (Otherwise $\rho_{V_i}(g)$ is the 0 matrix for all G, which obviously contradicts with the fact that $\rho_{V_i}(g) \in \mathrm{GL}(V_i)$)

Hence the set S has to be empty set, which shows that $\boldsymbol{PW}$ is injective.

Surjective: Let $f = \lambda \circ \rho_W$ be a matrix coefficient for some representation $\rho_W : G \to \mathrm{GL}(W)$. Then $W = \bigoplus V_i^{\oplus c_i}$. We can write $\rho_W(g)$ in block fashion:

$$\rho_W(g) = \begin{bmatrix} \rho_{V_1}(g) & & & 0 \\ & \ddots & & \\ & & \rho_{V_2}(g) & \\ 0 & & & \ddots \end{bmatrix}.$$

So any linear combination of matrix entries of $\rho_W(g)$ is a linear combination of $\rho_{V_i}(g)$'s. We can consequently write

$$\lambda \circ \rho_W \in \boldsymbol{PW}\left(\bigoplus_{V_i \text{ in } W} \mathrm{End}(V_i)\right)$$

and conclude that $\boldsymbol{PW}$ is surjective. □

Remark. For G finite we have

$$\mathbb{C}G \cong \bigoplus \mathrm{End}(V_i) \text{ as } (G \times G)\text{-representation.}$$

Remark. $\mathcal{O}(G)$ is a subring of $\mathcal{C}_0(G)$ by pointwise addition and multiplication. So the isomorphism

$$\underbrace{\mathcal{O}(G)}_{\text{rng}} \cong \bigoplus \underbrace{\mathrm{End}(V_i)}_{\text{ring}}$$

gives convolution in Haar measure.

Remark. We have $G \hookrightarrow G \times G$ by $g \mapsto (g, g)$. Hence

$$\begin{aligned} \mathcal{O}(G)^G &\cong \bigoplus \mathrm{End}(V_i)^G \\ &\cong \bigoplus \left(\mathrm{End}(V_i)^\vee\right)^G. \end{aligned}$$

Recall that $G \curvearrowright G$ by conjugation: $g * h = ghg^{-1}$. $G \curvearrowright \mathcal{O}(G)$ by $g * f = f(ghg^{-1})$. So $\mathcal{O}(G)^G$ is the conjugacy invariant matrix coefficients of G.

Meanwhile,

$$\begin{aligned} \mathrm{End}(V_i)^G &= \mathrm{Hom}(V_i, V_i)^G \\ &= \mathrm{Hom}_G(V_i, V_i) \\ &= \mathbb{C} \cdot \mathrm{Id}_{V_i} \\ \left(\mathrm{End}(V_i)^\vee\right)^G &= \mathbb{C} \cdot \mathrm{Trace}\,. \end{aligned}$$

Trace is linear and conjugacy invariant. In particular $\mathrm{Trace} \circ \rho_{V_i} = \chi_{\rho_{V_i}}$. We can conclude that

$$\left\{ \begin{array}{c} \text{conjugacy-invariant} \\ \text{matrix coefficients} \end{array} \right\} \cong \bigoplus \mathbb{C} \cdot \chi_{V_i}.$$

In particular, when G is finite, this translates to

$$\#\{\text{conjugacy classes}\} = \#\{\text{irreducible representations}\}.$$

Remark. $\mathcal{O}(G)$ has a Hermitian form

$$\langle p, q\rangle = \int_G \overline{p(g)}q(g).$$

The isomorphism $\mathcal{O}(G) \cong \bigoplus \mathrm{End}(V_i)$ is orthogonal for this form $\langle , \rangle$.

If we choose coordinates on $\mathrm{End}(V_k)$ such that ρ_{V_k} is unitary then $\sqrt{\dim V_k}(\rho_{V_k})_{ij}$ are orthonormal basis of $\mathcal{O}(G)$.

Two points if you are interested in analysis:

(1) $\mathcal{C}_0(G)$ is the completion of $\mathcal{O}(G)$ for uniform convergence.
(2) $L^2(G)$ is the L^2 completion of $\mathcal{O}(G)$.

October 19 – The unitary trick

Review of previous class: As we saw last class, the space of matrix coeficients of G, denoted $\mathcal{O}(G)$ is isomorphic to $\bigoplus_{V \text{ is isom. class of irred. rep.}} \mathrm{End}(V) \cong \bigoplus \mathrm{End}(V)^\vee = \bigoplus V^\vee \times V$. Indeed, the Peter-Weil map $\mathtt{PW} : \bigoplus \mathrm{End}(V)^\vee \to \mathcal{O}(G)$ is an isomorphism of $G \times G$-representations.

Other things $\mathcal{O}(G)$ is:

- $G \times G-$representation
- Commutative ring (subring of $\mathcal{C}^0(G)$)
- Rng by convolution ($f_1 * f_2 = \int_G f_1(h)f_2(h^{-1}g)$)
- Looking at the space of functions as an algebra with inner product $\langle f_1, f_2\rangle = \int_G \overline{f_1}(g)f_2(g)$, we get that $\mathcal{O}(G)$ is orthogonal.

Looking at the map and action $G \hookrightarrow G\times G \to G$ given by $g \mapsto (g, g)$ and $(g, g)*h = ghg^{-1}$ we also get that $\mathcal{O}(G)^G$ is the set of conjugation invariant matrix coefficients of G. But it also is $\bigoplus \mathrm{End}(V)^G \cong \mathrm{Hom}(V, V)^G \cong \mathrm{Hom}_G(V, V) \cong \mathbb{C}\mathrm{Id}_V$. Looking at the dual map, we also get it is $\bigoplus [\mathrm{End}(V)^\vee]^G = \bigoplus \mathbb{C}[(Tr) \circ V] = \bigoplus \mathbb{C}\chi_V$. In particular, this means that the traces uniquely define the representations, that is, if $\chi_V = \chi_W$ then $V \cong W$.

The unitary trick. Now, we turn our attention to **the unitary trick**, a way of showing that classifying representations of $\mathrm{GL}_n(\mathbb{C})$ is the same doing it for $U(n)$ (the unitary group).

Lemma. Let $\Omega \subset \mathbb{C}^n$ be an open neighborhood of 0. Let $f : \Omega \to \mathbb{C}$ be holomorphic and $f|_{\Omega\cap\mathbb{R}^n} = 0$. Then $f = 0$ on an open neighborhood of 0.

Proof. We proceed by induction on n. The base case $n = 0$ is trivial, as $\mathbb{R}^0 = \mathbb{C}^0$. Now suppose the result is true for $k < n$. Write $f(z_1, \cdots, z_n)$ as a power series. If $f \neq 0$, let d be the smallest exponent of $z_n$with non-zero coefficients. So $f(z_1, \cdots, z_n) = z_n^d(g(z_1, \cdots, z_{n-1})+ z_n h(z_1, \cdots, z_n))$. So $z_n^{-d}f(z_1, \cdots, z_n) = g(z_1, \cdots, z_{n-1}) + z_n h(z_1, \cdots, z_n)$ is holomorphic. If $f|_{\Omega\cap\mathbb{R}^n} = 0$, then so is $g|_{\Omega\cap\mathbb{R}^n}$. And since g will be holomorphic, by the induction step, $g(z_1, \cdots, z_{n-1}) = 0$ on an open neighborhood of 0. Therefore, d is not the smallest as we had supposed, and this a contradiction. So $f = 0$ around an open neighborhood of 0.

□

Corollary. Let Z be a complex $n-$fold and let $\sigma : Z \to Z$ be a complex antilinear involution. Let $X \subset Z$ be the set of fixed points of σ and let $x \in X$. If f is a holomorphic function on Z defined near x and $f|_X = 0$, then $f = 0$ near x on Z.

So let $G = \mathrm{GL}_n(\mathbb{C})$ and $\sigma : G \to G$ be the map $g \mapsto (\bar{g}^{-1})^T$. So the set of fixed points of σ is $K = U(n)$. Moreover the unitary group is compact. So we have G as a complex, connected Lie group and K as a compact real Lie group that is the set of fixed points of an anti-linear holomorphic involution in G. So holomorphic function in G are 0 if and only if they are so in K. But as K is compact, we can use the results we have proved so far for compact, Hausdorff, topological groups.

Theorem. Let V and W be complex $G-$representations with $\rho_V : G \to \mathrm{GL}(V)$ and $\rho_W : G \to \mathrm{GL}(W)$ holomorphic. Then $\mathrm{Hom}_G(V, W) = \mathrm{Hom}_K(V, W)$.

Proof. Let $\phi : V \to W$ be $\mathbb{C}-$linear. So we need to show that $\phi(gv) = g\phi(v)$ for all $g \in G$ if and only in the same is true for all $g \in K$. The forward direction is clear, as $K \subset G$. For the other direction, note that $\phi(gv) - g\phi(v)$ is a holomorphic function on g and it vanishes for every $g \in K$. Therefore, it vanishes on all of G by the unitary trick.

□

Corollary. If V, W are as above. Then $V \cong W$ as $G-$representations if and only is they are so as $K-$representations.

Proof. This follows from the fact this isomorphism would be an invariant map in $\mathrm{Hom}_G(V, W) = \mathrm{Hom}_K(V, W)$. □

October 21 – Applications of the unitary trick

We did this day as IBL. Here are the problems:

Let G be a connected complex Lie group, let $\sigma : G \to G$ be an antiholomorphic involution and let K be the fixed points of σ. Then K is a real Lie-subgroup of G. We proved, and you may assume:

Theorem. Any holomorphic function f on G which restricts to 0 on K is 0 on G.

For a complex vector space V and a representation $\rho : G \to \mathrm{GL}(V)$, we say that V is a ***holomorphic representation*** of G if the matrix entries of ρ are holomorphic functions.

The first two problems were done in the previous class, but please talk them over. Most of these problems are very short:

Problem 13. Let V and W be holomorphic representations of G. Show that $\mathrm{Hom}_K(V, W) = \mathrm{Hom}_G(V, W)$. (I.e. a linear map $\phi : V \to W$ is a map of G-representations if and only if it is a map of K-representations.)

Problem 14. Let V and W be holomorphic representations of G. Show that $V \cong W$ as G-representations if and only if $V \cong W$ as K-representations.

Problem 15. Let W be a holomorphic G-representation and let V be a vector subspace of W. Show that V is a sub-G-representation if and only if it is a sub-K-representation.

Problem 16. Let W be a holomorphic representation of G. Show that W is simple as a G-representation if and only if W is simple as a K-representation.

Problem 17. Let W be a holomorphic representation of G. Show that W is indecomposable as a G-representation if and only if W is indecomposable as a K-representation.

Now, suppose that K is compact.

Problem 18. Let W be a holomorphic representation of G. Show that W is indecomposable as a G-representation if and only if W is simple as a G-representation.

Afterwards, we proved the following results:

Lemma. Let V be a K-representation and let Y be a G-representation such that V embeds K-equivariantly into Y. Then the K-action on V extends to a G-action.

Proof. We may as well think of V as a subspace of Y. Then use Problem 15. $\square$

Lemma. Let K be compact, let V be a K-representation, and let X and Y be G-representations such that $\chi_V + \chi_X|_K = \chi_Y|_K$. Then the K-action on V extends to a G-action.

Proof. Since K is compact, the equality of characters implies that $V \oplus X|_K \cong Y|_K$ as K-representations. In particular, the inclusion of $V \oplus 0$ into Y is a K-equivariant map $V \to Y$. Now use the previous lemma. $\square$

October 24 - Extending $U(n)$ representations to $GL_n(\mathbb{C})$

On the IBL section of last class, we proved several similar results. Let G be a connected $\mathbb{C}$-Lie group, $\sigma : G \to G$ an anti-holomorphic involution and $K = G^\sigma$. Given V, W and G-representations, then, if some properties hold for V, W as K-representations, they will also be true for V, W as G-representations:

(1) $V|_K \cong W|_K$ if and only if $V \cong W$;
(2) $\mathrm{Hom}_K(V|_K, W|_K) = \mathrm{Hom}(V, W)$;
(3) Given $U \subset V$ a subspace, then U is a K-subrepresentation if, and only if it is a G-subrepresentation;
(4) $\chi_V|_K = \chi_W|_K \iff \chi_V = \chi_W$;

Furthermore, if K is also compact, we have:

(5) V is simple $\iff$ V is indecomposable;
(6) $\chi_V = \chi_W \iff V \cong W$.

All these results take a G-representation and look at what we can say about it as a K-representation. Now, what if we start with a K-representation V? Is it possible to extend it to a G?
First of all notice that if such an extension exists, it will be unique, by result (1) of the list above. Our problem is existence.

Suppose that we have a representation $\rho : K \to \mathrm{GL}_N$; let ρ_{ij} be the matrix entries of this map. If the ρ_{ij} have analytic continuations to G, then the analytic continuations will give a map $\tilde{\rho} : G \to \mathrm{GL}_N$, and this map will also be a representation. (The condition $\tilde{\rho}(gh) = \tilde{\rho}(g)\tilde{\rho}(h)$ will hold by analytic continuation from the same relation on K.) So the question is basically whether all matrix coefficients of K have analytic extensions to G. Here are three examples of things going wrong:

Example. Take $G = \mathbb{C}^*$, $\sigma : \mathbb{C}^* \to \mathbb{C}^*$ the usual conjugation. Then $K = \mathbb{R}^*$. Then $t \mapsto \left[\, t^{1/3} \,\right]$ is a well defined continuous representation $K \to \mathrm{GL}_1(\mathbb{C})$, but it branches if we try to continue it to G.

Example. Let E be an elliptic curve defined over $\mathbb{R}$ with one real component, for example, $y^2 = x^3 - 1$. As a topological group, $E(\mathbb{R}) \cong S^1$, so we have a group homomorphism $E(\mathbb{R}) \to \mathrm{GL}_1(\mathbb{C})$ sending $\theta \in S^1$ to $e^{i\theta}$. However, this branches on $E(\mathbb{C})$: We have $E(\mathbb{C}) \cong (\mathbb{C}^*)/q^{\mathbb{Z}}$ for some $q < 1$, and the matrix coefficient is the coordinate on the cover $\mathbb{C}^*$.

Example. Take $G = \mathbb{PSL}_3(\mathbb{C}) = \mathrm{SL}_3(\mathbb{C})/\langle \zeta \mathrm{Id}_3 \rangle$, where ζ is a primitive cube root of unity. Then we can take $K = \mathbb{PSL}_3(\mathbb{R}) \cong \mathrm{SL}_3(\mathbb{R})$; the isomorphism is because the group $\langle \zeta \mathrm{Id}_3 \rangle$ which we quotient by to form $\mathbb{PSL}_3$ has trivial intersection with the real points.

Then we have a representation $\rho : \mathbb{PSL}_3(\mathbb{R}) \to \mathrm{GL}_3$ by the isomorphism $\mathbb{PSL}_3(\mathbb{R}) \cong \mathrm{SL}_3(\mathbb{R})$. If we try to analytically continue it to $\mathbb{PSL}_3(\mathbb{C})$, we get branching; the analytic continuation is only well defined on $\mathrm{SL}_3(\mathbb{C})$.

For our purposes, the most important example of a pair $K \subset G$ for when the extension exists is $U(n) \subset \mathrm{GL}_n(\mathbb{C})$. We will spend the rest of the class treating this example. Let's start with notation for some subgroups of $\mathrm{GL}_n(\mathbb{C})$:

$$\begin{array}{ccc} G = \mathrm{GL}_n(\mathbb{C}) & \supset & T = \begin{bmatrix} z_1 & & 0 \\ & \ddots & \\ 0 & & z_n \end{bmatrix} \\ \cup & & \cup \\ K = U(n) & \supset & S = \begin{bmatrix} e^{i\theta_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\theta_n} \end{bmatrix} \end{array}$$

Notice that, as a topological group $S \cong S^1 \times \cdots \times S^1 = T^n$, the n-torus. Because of this, sometimes we call either S or T a ***torus***.

Given a representation $\rho : K \to \mathrm{GL}_N(\mathbb{C}) = \mathrm{GL}(V)$, we want to extend it to G. Let $\chi(g) = Tr(\rho(g))$, be the character of ρ, and $\chi|_S$ its restriction to S and denote $\chi|_S \left(\begin{bmatrix} e^{i\theta_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\theta_n} \end{bmatrix} \right) = \chi|_S(e^{i\theta_1}, \dots, e^{i\theta_n})$. In particular, as any permutation matrix is unitary and χ is conjugacy invariant, for any $\sigma \in S^n$, we have that:

$$\begin{aligned} \chi|_S(e^{i\theta_{\sigma(1)}}, \dots, e^{i\theta_{\sigma(n)}}) &= \chi|_S \left(\begin{bmatrix} e^{i\theta_{\sigma(1)}} & & 0 \\ & \ddots & \\ 0 & & e^{i\theta_{\sigma(n)}} \end{bmatrix} \right) = \chi \left(\sigma^{-1} \begin{bmatrix} e^{i\theta_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\theta_n} \end{bmatrix} \sigma \right) \\ &= \chi \left(\begin{bmatrix} e^{i\theta_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\theta_n} \end{bmatrix} \right) = \chi|_S(e^{i\theta_1}, \dots, e^{i\theta_n}) \end{aligned}$$

Hence $\chi|_S$ is a symmetric function on $e^{i\theta_1}, \dots, e^{i\theta_n}$. Now as $S \cong S^1 \times \cdots \times S^1$, by our classification of continuous S^1 representations, we have that $V = \oplus_{\mu \in \mathbb{Z}^n} V_\mu$ as an S-representation. Where all but finitely many $V_\mu = 0$ and $\begin{bmatrix} e^{i\theta_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\theta_n} \end{bmatrix} \in S$ acts on V_μ by multiplication by $e^{i(\mu_1\theta_1 + \cdots + \mu_n\theta_n)}$. Then $(\chi|_S)_{V_\mu}(e^{i\theta_1}, \dots, e^{i\theta_n}) = Tr(e^{i(\mu_1\theta_1 + \cdots + \mu_n\theta_n)} Id) = (\dim V_\mu) e^{i(\mu_1\theta_1 + \cdots + \mu_n\theta_n)}$. Then as $\chi_{V \oplus W} = \chi_V + \chi_W$, as in the first homework, we have that:

$$\chi|_S(e^{i\theta_1}, \dots, e^{i\theta_n}) = \sum_{\mu \in \mathbb{Z}^n} (\dim V_\mu) e^{i(\mu_1\theta_1 + \cdots + \mu_n\theta_n)}$$

In particular, $\chi|_S$ is a symmetric Laurent polynomial on $e^{i\theta_1}, \dots, e^{i\theta_n}$, with non-negative integer coefficients. As every unitary matrix is diagonalizable and S is the set of diagonalizable unitary matrices, as we know $\chi|_S$, we also know χ.

Now suppose ρ does extend to $\rho' : G \to \mathrm{GL}_N(\mathbb{C})$, and χ' is the character of ρ'. Then χ' is uniquely defined by $\chi'|_T$, as T is dense in G and χ' is continuous. Furthermore $\chi'|_T \left(T = \begin{bmatrix} z_1 & & 0 \\ & \ddots & \\ 0 & & z_n \end{bmatrix} \right) = \chi'|_T \in \Lambda_n^{\pm}$, that is, it is a symmetrical Laurent polynomial on

$z_1, \ldots, z_n$, and $\chi'|_T(e^{i\theta_1}, \ldots, e^{i\theta_n}) = \chi|_S(e^{i\theta_1}, \ldots, e^{i\theta_n})$. So, if there's an extension ρ', its character will be $\chi \in \Lambda_n^{\pm}$.
As K is compact, any K-representation is uniquely defined by its character, so if there's $\rho' : G \to \mathrm{GL}_n(\mathbb{C})$ with character χ, then $\rho = \rho'|_K$. Hence, what we need to do now is to prove that we can build a representation ρ' with character f, for any $f \in \Lambda_n^{\pm}$.
Let's list some facts we saw on previous classes/homeworks:

- The character of the obvious representation of $\bigwedge^k \mathbb{C}^n$ is $e_k(z_1, \ldots, z_n)$. So we get representations whose character is e_k;
- If V, W are G-representations, then we can induce representations of $V \oplus W$, $V \otimes W$ such that $\chi_{V \oplus W} = \chi_V \oplus \chi_W$ and $\chi_{V \otimes W} = \chi_V \chi_W$. Therefore, we have representations with e_λ as their characters, for any partition λ and any linear combination of them with non-negative integers;
- $g \to \det(g)^{-1} \in \mathrm{GL}_1(\mathbb{C}) = \mathrm{GL}(D)$ is a representation with character $\chi(z_1, \ldots, z_n) = \frac{1}{z_1 \cdots z_n}$. Now, if $g \in \Lambda_n^{\pm}$, there is $M \geq 0$ such that $(z_1 \cdots z_n)^M g \in \Lambda_n$. Then if we find a G-representation V with character $\chi_V = (z_1 \cdots z_n)^M g$, we have $g = \chi_{U^{\oplus d} \otimes V}$.

Resuming what we have up until now. Suppose we qare given $\rho : K \to \mathrm{GL}(V)$, we have $f = \chi_V \in \Lambda_n^{\pm}$ and we want to find a G-representation with character f as well. By the last bullet point, we can assume $f \in \Lambda_n$. Then as $e_\lambda' s$ form a basis for Λ_n, we can write $f = \sum_\lambda c_\lambda e_\lambda$. If all the c_λ's are non-negative, we can build Y a G-rep whose character is f by tensors and direct products of $\bigwedge^k \mathbb{C}^N$. The last piece of the puzzle is when there is some $c_\lambda < 0$. For that, we use the criterion from last class.
Write $f = \sum b_\lambda e_\lambda - \sum a_\lambda e_\lambda$, so that $b_\lambda, a_\lambda \geq 0$ for all λ. Then $f + \sum a_\lambda e_\lambda = \sum b_\lambda e_\lambda$. As above, we can build X, Y G-reps whose character is $\sum a_\lambda e_\lambda, \sum b_\lambda e_\lambda$, respectively. Then, as $Y|_K$ $V \oplus X|_K$ are K-representations with same character, $Y|_K \cong V \oplus X|_K$. Then $V \hookrightarrow V \oplus X|_K$ is a K-invariant embedding, and by the criterion, we can extend the action of K on V until G. This finishes our proof.
This was the big result of this class. Our next object is to prove that the characters of simple representaions are precisely the Schur Polynomials. We'll do that next class. Here is some useful foreshadowing for now. First, from the list of facts, we proved last class, we have that indecomposable G-representations are simple. Then G breaks down as a direct sum of simple representations, which is unique up to iso. Form the main proof of this class, we can build a G-rep with any character $g \in \Lambda_n^{\pm}$. Thus, the characters of simple G-representations, form a $\mathbb{Z}$-basis of $\Lambda_n^{\pm}$. The main claim is that these characters are of the form $(z_1 \cdots z_n)^{-k} s_\lambda(z_1, \ldots, z_n)$.

October 26 – Matrix coefficients of $\mathrm{GL}_n(\mathbb{C})$ and its Peter-Weyl decomposition

Recall that

$$G = \mathrm{GL}_n(\mathbb{C}) = \left\{ g = \begin{bmatrix} z_{11} & \cdots & z_{1n} \\ \vdots & \ddots & \vdots \\ z_{nn} & \cdots & z_{nn} \end{bmatrix} : \det(g) \neq 0 \right\}$$

has the compact subgroup

$$K = U(n) = \{g : g\overline{g}^T = \mathrm{Id}_n\}.$$

We learned the unitary trick that lifts K-representations to G-representations in the last IBL class. And in last class, we lifted any arbitrary finite-dimensional, continuous K-representation to G-representations especially for $K = U(n)$ and $G = \mathrm{GL}_n(\mathbb{C})$.

Let $g = (z_{ij})$ be an arbitrary element in $\mathrm{GL}_n(\mathbb{C})$. Recall that a *polynomial* representation ρ of $\mathrm{GL}_n(\mathbb{C})$ is a representation with the matrix coefficients of $\rho(g)$ being in $\mathbb{C}[z_{ij}]$. Likewise, the *algebraic* representations of $\mathrm{GL}_n(\mathbb{C})$ are exactly the representations ρ such that every coefficients of the matrix $\rho(g)$ is in $\mathbb{C}[z_{ij}, \det(g)^{-1}]$. Thus, we can talk about the rings of polynomial matrix coefficients, $\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{poly}}$ and algebraic matrix coefficients, $\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{alg}}$.

We will prove the following.

Theorem. The characters of irreducible *polynomial* representations of $\mathrm{GL}_n(\mathbb{C})$ are the Schur polynomials $s_\lambda(x_1, \cdots, x_n)$.

The characters of irreducible *algebraic* representations of $\mathrm{GL}_n(\mathbb{C})$ are the Laurent polynomials $(x_1 x_2 \cdots x_n)^{-M} s_\lambda(x_1, \cdots, x_n)$ for suitable partition λ and $M \in \mathbb{Z}$.

We first compute the matrix coefficients of $\mathrm{GL}_n(\mathbb{C})$.

Theorem. The rings of polynomial, algebraic and holomorphic matrix coefficients are related as in the diagram below:

$$\begin{tikzcd}
\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{poly}} \arrow[rr, equal] \arrow[d, hook] & & \mathbb{C}[z_{ij}] \arrow[dd, hook] \\
\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{alg}} \arrow[d, equal] & & \\
\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{hol}} \arrow[r, equal] & \mathcal{O}(U(n)) \arrow[r, equal] & \mathbb{C}[z_{ij}, \det(g)]
\end{tikzcd}$$

Proof. (Sketch) First check $\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{hol}} = \mathcal{O}(U(n))$. Restriction gives the map $\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{hol}} \to \mathcal{O}(U(n))$. The restriction map is injective by the unitary trick. It is also surjective by the fact that we can lift any K-representation to G-representation in the last class.

Next, check $\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{alg}} = \mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{hol}}$. The containment $\subseteq$ is immediate. Note that in the last class, any lift of a continuous $U(n)$-representation to a $\mathrm{GL}_n(\mathbb{C})$-representation happened to be algebraic as well. This shows the containment $\supseteq$.

We show $\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{alg}} = \mathbb{C}[z_{ij}, (\det g)^{-1}]$. The containment $\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{alg}} \subseteq \mathbb{C}[z_{ij}, (\det g)^{-1}]$ holds by definition. We use the following lemma.

Lemma. $\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{adj}}$ for any adj = poly, alg, hol is a subring of the ring of continuous functions $C_0(\mathrm{GL}_n(\mathbb{C}))$

Proof. (of Lemma) For any $f_1, f_2 \in \mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{adj}}$, we have $f_1 = \lambda_1 \circ \rho_{V_1}$ and $f_2 = \lambda_2 \circ \rho_{V_2}$ for functionals λ_1, λ_2 and $\mathrm{GL}_n(\mathbb{C})$-representations ρ_{V_1}, ρ_{V_2} of V_1, V_2 that satisfy the adj = poly, alg, hol. Check the followings

$$f_1 + f_2 = (\lambda_1 \oplus \lambda_2) \circ \rho_{V_1 \oplus V_2}$$
$$f_1 f_2 = (\lambda_1 \otimes \lambda_2) \circ \rho_{V_1 \otimes V_2}$$

so that $f_1 + f_2, f_1 f_2 \in \mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{adj}}$ as well. $\square$

So $\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\mathrm{alg}}$ is a $\mathbb{C}$-subalgebra of $\mathbb{C}[z_{ij}, (\det g)^{-1}]$. It remains to check that $z_{ij}, (\det g)^{-1} \in \mathcal{O}_{\mathrm{alg}}$ to show the equality $\mathcal{O}_{\mathrm{alg}} = \mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\mathrm{alg}} = \mathbb{C}[z_{ij}, (\det g)^{-1}]$. From the standard representation $\mathbb{C}^n$ of $\mathrm{GL}_n(C)$, we get $z_{ij} \in \mathcal{O}_{\mathrm{alg}}$. The representation $\rho : \mathrm{GL}_n(\mathbb{C}) \to \mathbb{C}$ defined as $g \mapsto (x \mapsto (\det g)^{-1}x)$ gives $(\det g)^{-1} \in \mathcal{O}_{\mathrm{alg}}$, and we are done.

To check $\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\mathrm{poly}} = \mathbb{C}[z_{ij}]$, note that the containment $\subseteq$ holds by definition, and $z_{ij} \in \mathcal{O}_{\mathrm{poly}}$ by the fact that the standard representation $\mathbb{C}^n$ of $\mathrm{GL}_n(\mathbb{C})$ is polynomial. $\square$

So we have

$$\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\mathrm{poly}} = \mathbb{C}[z_{ij}] \simeq \bigoplus_{V \text{poly irred.}} V \otimes V^\vee$$

and

$$\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\mathrm{alg}} = \mathbb{C}[z_{ij}, (\det g)^{-1}] \simeq \bigoplus_{V \text{alg irred.}} V \otimes V^\vee$$

as $\mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C})$ representations.

Break $\mathbb{C}[z_{ij}] = \bigoplus_{d=0}^{\infty} \mathbb{C}[z_{ij}]_d$ into degree d homogeneous polynomials. so that each part $\mathbb{C}[z_{ij}]_d$ is a finite $\mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C})$ representation. Let's see how $\mathbb{C}[z_{ij}] \simeq \bigoplus_{V \text{poly irred.}} V \otimes V^\vee$ restricts to such degree d counterparts.

For $d = 0$,

$$\mathbb{C}[z_{ij}]_0 \simeq \mathbb{C} \times 1 \simeq (\text{trivial}) \otimes (\text{trivial})^\vee.$$

For $d = 1$,

$$\mathbb{C}[z_{ij}]_1 \simeq \{n \times n \text{ matrices of coefficients}\} \simeq \mathbb{C}^n \otimes (\mathbb{C}^n)^\vee$$

where the $\mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C})$ action is $(g, h) * M = gMh^{-1}$.

For $d = 2$,

$$\begin{aligned} \mathbb{C}[z_{ij}]_2 &= \mathrm{Span}\left(\det \begin{bmatrix} z_{pr} & z_{ps} \\ z_{qr} & z_{qs} \end{bmatrix}\right)_{\substack{1 \le p < q \le n \\ 1 \le r < s \le n}} \oplus \mathrm{Span}\left(\mathrm{perm} \begin{bmatrix} z_{pr} & z_{ps} \\ z_{qr} & z_{qs} \end{bmatrix}\right)_{\substack{1 \le p \le q \le n \\ 1 \le r \le s \le n}} \\ &= \left(\vee^2 \mathbb{C}^n \otimes \left(\vee^2 \mathbb{C}^n\right)^\vee\right) \oplus \left(\mathrm{Sym}^2 \mathbb{C}^n \otimes \left(\mathrm{Sym}^2 \mathbb{C}^n\right)^\vee\right) \end{aligned}$$

where the first component has dimension $\binom{n}{2}^2$ and the second component has dimension $\binom{n+1}{2}^2$. Here, $\mathrm{perm} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is defined as $ad + bc$ which is bilinear.

Recall the Cauchy identity.

$$\prod_{i,j=1}^{n} \frac{1}{1 - x_i y_j} = \sum_{l(\lambda) \le n} s_\lambda(x) s_\lambda(y)$$

Expanding the right-hand side for the partitions , (1), (1, 1) and (2), we have

$$\sum_{l(\lambda)\leq n} s_\lambda(x)s_\lambda(y) = 1 + \left(\sum_{i=1}^n x_i\right)\left(\sum_{j=1}^n y_j\right) + \left(\sum_{1\leq i_1<i_2\leq n} x_{i_1}x_{i_2}\right)\left(\sum_{1\leq j_1<j_2\leq n} y_{j_1}y_{j_2}\right) + \left(\sum_{1\leq i_1\leq i_2\leq n} x_{i_1}x_{i_2}\right)\left(\sum_{1\leq j_1\leq j_2\leq n} y_{j_1}y_{j_2}\right) + \cdots$$

and note that the first four summands of Schur polynomials somewhat corresponds to the factors $V \otimes V^\vee$ of $\mathbb{C}[z_{ij}]_d$ we found for $d = 0, 1, 2$.

To clarify this further, let's restrict $\mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C})$ to $T \times T$ where T is the set of diagonal matrices. Write the coordinates of $T \times T$ as

$$\left(\begin{bmatrix} x_1 & & \\ & \ddots & \\ & & x_n\end{bmatrix}, \begin{bmatrix} y_1^{-1} & & \\ & \ddots & \\ & & y_n^{-1}\end{bmatrix}\right) \in T \times T$$

so that its action on z_{ij} is the multiplication by x_iy_j.

Then the $T\times T$ character of $\mathbb{C}[z_{ij}]$ is $\prod_{i,j=1}^n \frac{1}{1-x_iy_j}$, the left-hand side of the Cauchy identity. Meanwhile, the $T \times T$ character on $V \otimes V^\vee$ is $\chi_V(x)\chi_V(y)$. Comparing the character of $\mathbb{C}[z_{ij}] \simeq \bigoplus_{V\text{poly irred.}} V \otimes V^\vee$, we can observe that $\chi_V(x)$'s obey Cauchy's identity. That is, χ_V's form an orthonormal basis of Λ_n with the inner product of symmetric polynomials. This means that the collection of χ_V is the collection of $\pm s_\lambda$'s. The sign should be positive, because assigning $x_i = 1$ makes $\chi_V(x)$ equal to $\dim V$ which is positive, and the monomial coefficients of s_λ is positive. In conclusion, the characters of polynomial irreducible representations of $\mathrm{GL}_n(\mathbb{C})$ are the Schur polynomials s_λ. We will expand on this in later classes.

October 28 – Constructing the irreps of GL_n, first time

Today's first goal is to show that characters of polynomial irreducible represemntations of $\mathrm{GL}_n(\mathbb{C})$ are the schur polynomials $s_\lambda(x_1, \cdots, x_n)$.

Recall from last class that the matrix coefficients $\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{poly reps}} = \mathbb{C}[z_{ij}]_{1\leq i,j\leq n}$.

By the Peter-Weyl theorem, we have

$$\mathcal{O}(\mathrm{GL}_n(\mathbb{C}))_{\text{poly reps}} = \bigoplus_{V\text{poly irrep}} \mathrm{End}(V^\vee) = \bigoplus_{V\text{poly irrep}} V \otimes V^\vee$$

.

Consider characters of the action of $\begin{bmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_n\end{bmatrix}$ and $\begin{bmatrix} y_1^{-1} & & & \\ & y_2^{-1} & & \\ & & \ddots & \\ & & & y_n^{-1}\end{bmatrix}$ on both sides of the above representations.

For the left hand side, these two diagonal matrices act on an n by n matrix $[z_{ij}]$ by multiplying on the left and multiplying inverse on the right, sending $z_{ij} \mapsto x_iy_jz_{ij}$. Thus, by acting on monomials in z_{ij}, we get all powers of x_iy_j. By adding up all combinations of $(i,j) \in [n]^2$, we get that the character equals $\prod_{i,j=1}^n \frac{1}{1-x_iy_j}$. Furthermore, by Cauchy's identity, $\prod_{i,j=1}^n \frac{1}{1-x_iy_j} = \sum_\lambda s_\lambda(x_1,\cdots,x_n)s_\lambda(y_1,\cdots,y_n)$

For the right hand side, we have a direct sum of tensor product of two representations V and $V^\vee$. Use the fact that

(1) Character of direct sum equals sum of characters.
(2) Character of tensor product equals product of characters.
(3) Character of dual representation equals character evaluated on the inverse element

We have that the character on the right hand side equals $\sum_{V \text{ polyn irrep}} \chi_V(x_1,\cdots,x_n)\chi_V(y_1,\cdots,y_n)$.

Thus, we arrive at the following equality of polynomials

$$\sum_\lambda s_\lambda(x_1,\cdots,x_n)s_\lambda(y_1,\cdots,y_n) = \sum_{V \text{ polyn irrep}} \chi_V(x_1,\cdots,x_n)\chi_V(y_1,\cdots,y_n)$$

We claim that $\chi_V \begin{bmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{bmatrix}$ is in Λ_n. Proof of claim:

(1) It is a polynomial because V is a polynomial representation.
(2) It is symmetric because any permutation σ acts by conjugation as
$$\sigma \begin{bmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{bmatrix} \sigma^{-1} = \begin{bmatrix} x_{\sigma(1)} & & \\ & \ddots & \\ & & x_{\sigma(n)} \end{bmatrix}$$
and character is constant on conjugacy classes.
(3) The coefficients are in $\mathbb{Z}_{\geq 0}$ because V restricted to the torus is a direct sum of 1-dimensional representations. Thus, by adding up the characters of 1-dim representations, we get nonnegative integer coefficients.

Now, assume that $\chi_V\left(\begin{bmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{bmatrix}\right) = \sum c_{V\lambda}s_\lambda(\underline{x})$ (the Schur polynomials form a $\mathbb{Z}$-basis). We extract the coefficient of $s_\lambda(\underline{x})s_\lambda(\underline{y})$ from the previously established equality $\sum_\lambda s_\lambda(x_1,\cdots,x_n)s_\lambda(y_1,\cdots,y_n) = \sum_{V \text{ polyn irrep}} \chi_V(x_1,\cdots,x_n)\chi_V(y_1,\cdots,y_n)$ to get that $\sum_V c_{V\lambda}^2 = 1$. Thus, for any fixed λ, exactly one $c_{V\lambda}$ equals ± 1 and the rest are 0. But notice that χ_V and s_λ both have positive integer coefficients in $x_1,\cdots,x_n$, the sign must be a plus sign.

By now, we showed that for each λ, exactly one $c_{V\lambda}$ equals 1 and the other equals 0. And if $c_{V\lambda} = c_{V\mu} = 1$ for $\lambda \neq \mu$, we can look at the coefficient of $s_\lambda(\underline{x})s_\mu(\underline{y})$ to see that $0 = \sum c_{V\lambda}c_{V\mu}$. But notice that each C is either 0 or 1. And if $c_{V\lambda} = c_{V\mu} = 1$, then thet sum on the right hand side is at least 1. Contradiction. Thus, we have a one-to-one correspondence between the V's and the λ's such that $c_{V\lambda} = 1$. This shows that the characters of polynomial irreducible representations are exactly the Schur polynomials.

Note that the Hall inner product makes s_λ orthonormal. And the Hall inner product of characters is given by $\langle \chi_V, \chi_W \rangle = \dim \operatorname{Hom}(V,W) = \int_K \overline{\chi_V(g)}\chi_W(g)$ for a compact subgroup K. It makes χ_V's orthonormal, where V's are the irreducible representations. So the two Hall inner products are the same.

Next, we attempt the answer the following question: What is the irreducible representation V_λ with character s_λ? We find a clue in the upper-triangularity results for s_λ's, e_λ's, and h_λ's.

Recall that we have

$$h_\lambda = s_\lambda + \sum_{\mu \succ \lambda} K_{\mu\lambda} s_\mu$$

$$e_{\lambda^T} = s_\lambda + \sum_{\mu \prec \lambda} K_{\lambda\mu} s_\mu$$

Thus, for any λ, s_λ is the only term that appears in both h_λ and e_{λ^T}. So when we take the Hall inner product, only one terms survives. We have $\langle e_{\lambda^T}, h_\lambda \rangle = 1$. By our previous discussion of equivalence between the two Hall inner products, we have that

$$\dim \operatorname{Hom}(\bigotimes \bigwedge^{\lambda_k^T} \mathbb{C}^n, \bigotimes \operatorname{Sym}^{\lambda_k} \mathbb{C}^n) = 1$$

This is because the anti-symmetric tensor has character $e_{\lambda_1^T} \cdots e_{\lambda_m^T} = e_{\lambda^T}$ and the symmetric tensor has character $h_{\lambda_1} \cdots h_{\lambda_n} = h_\lambda$. (Character of tensor product is product of characters.) So we have a unique homomorphism $\bigotimes \bigwedge^{\lambda_k^T} \mathbb{C}^n \to \bigotimes \operatorname{Sym}^{\lambda_k} \mathbb{C}^n$ up to a scalar, which factors through V_λ. Then the image of such a (nonzero) GL_n-equivariant map is V_λ. (This can be seen by decomposing the two representations into irreducible subrepresentations. And by Schur's lemma and its corollary (see problem 4 and 5 on the October 3-7 worksheet), maps between representations are the direct sum of maps between the irreducible components. And maps between different components are trivial, and the domain and target only have one common component, which is V_λ, the component corresponding to s_λ.)

Here is an example. We have

$$\begin{array}{ccccccc} h_3 & = & s_3 & & & & \\ h_{21} & = & s_3 & + & s_{21} & & \\ h_{111} & = & s_3 & + & 2s_{21} & + & s_{111} \end{array} \quad \text{and} \quad \begin{array}{ccccccc} e_{3^T} & = & s_3 & + & 2s_{21} & + & s_{111} \\ e_{21^T} & = & & & s_{21} & + & s_{111} \\ e_{111^T} & = & & & & & s_{111}. \end{array}$$

To obtain V_{21}, we consider the image of the unique (up to scalar multiple) GL_3-equivariant map $\bigwedge^2 \mathbb{C}^n \otimes \mathbb{C}^n \to \operatorname{Sym}^2 \mathbb{C}^n \otimes \mathbb{C}^n$.

To find such a map, we notice that $\bigwedge^2 \mathbb{C}^n$ is a subspace of $\mathbb{C}^n \otimes \mathbb{C}^n$ via $u \wedge v \mapsto u \otimes v - v \otimes u$. And $\operatorname{Sym}^2 \mathbb{C}^n$ is a quotient of $\mathbb{C}^n \otimes \mathbb{C}^n$ via $u \otimes v \mapsto (uv)$. Thus, we want to find a nonzero map $\bigwedge^2 \mathbb{C}^n \otimes \mathbb{C}^n \to \mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n \to \mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n \to \operatorname{Sym}^2 \mathbb{C}^n \otimes \mathbb{C}^n$, where the first map is given by the above inclusion, the third map is given by the above quotient, and the middle map is to permute the entries by a permutation in S_3 (We showed this in homework). Taking the middle map to be (23) will give us a nonzero composition (whereas identity will result in a zero composition). The composition is $(u \wedge v) \otimes w \mapsto (uw) \otimes v - (vw) \otimes u$. And the image is V_{21}.

October 31 – Constructing the irreps of $\mathrm{GL}_n(\mathbb{C})$, try two

This lecture was confused. November 2nd was much better.

Before beginning with today's spooky material, here is a quick summary of where we are right now:

(1) We showed using compact representation theory that continuous representations of $U(n)$ are determined by their characters, and that these characters are symmetric Laurent polynomials in n-variables (when evaluated on the torus).

(2) We showed using the unitary trick that every algebraic representation of GL_n is determined by its restriction to $U(n)$. Moreover, the continuous representations of $U(n)$ can always be extended uniquely to algebraic representations of GL_n. Hence,

the algebraic representations of GL_n are also determined by their characters, which are also symmetric Laurent polynomials in n-variables.

(3) We showed using the Peter-Weyl theorem and Cauchy's identity that the irreducible polynomial representations of GL_n correspond exactly with the Schur polynomials.

We are now interested in explicitly constructing the irreducible representation V_λ corresponding with the Schur polynomial s_λ. To do this, we will continue our method from the previous class.

Given a partition λ, we set

$$E_\lambda = \bigotimes_k \bigwedge^{\lambda_k^T} \mathbb{C}^n,$$

which has character e_{λ^T}, and

$$H_\lambda = \bigotimes_\ell \mathrm{Sym}^{\lambda_\ell} \mathbb{C}^n,$$

which has character h_λ. Then

$$\dim \mathrm{Hom}_{\mathrm{GL}_n}(E_\lambda, H_\lambda) = \langle e_{\lambda^T}, h_\lambda \rangle = 1.$$

The unique non-zero homomorphism (up to scaling) will factor through V_λ because s_λ is the only common Schur polynomial between e_{λ^T} and h_λ when both written in terms of the s-basis. Hence, the image of this homomorphism will be isomorphic to V_λ.

We now construct such a non-zero map. We use $(\mathbb{C}^n)^{\otimes \lambda}$ to denote the tensor product indexed by the boxes of λ. Let $S_{\mathrm{col},\lambda}$ denote the group of permutations of the boxes of λ that only permute within the columns. We have an inclusion $\alpha : E_\lambda \to (\mathbb{C}^n)^{\otimes \lambda}$ given by

$$\alpha\left(\bigotimes_k v_{1,k} \wedge \cdots \wedge v_{\lambda_k^T,k}\right) = \sum_{\sigma \in S_{\mathrm{col},\lambda}} (-1)^\sigma \bigotimes_{\ell,k} v_{\sigma(\ell,k)},$$

where (ℓ, k) denotes the box of λ in row ℓ and column k. After wrapping one's head around the indices, this is just the usual inclusion of an alternating product into a tensor product.

On the other hand, we have a projection $\pi : (\mathbb{C}^n)^{\otimes \lambda} \to H_\lambda$ given by

$$\pi\left(\bigotimes_{\ell,k} v_{\ell,k}\right) = \bigotimes_\ell v_{\ell,1} v_{\ell,2} \cdots v_{\ell,\lambda_\ell},$$

which is just the usual projection from a tensor product to a symmetric product.

All that remains is to show that $\pi \circ \alpha$ is non-zero. This is non-trivial, and we will leave it for next time. A key ingredient is that we symmetrize $(\mathbb{C}^n)^{\otimes \lambda}$ along the *rows*, while the inherited alternating behavior only resides within the *columns*.

We consider an example. If $\lambda = \begin{array}{|c|c|}\hline & \\ \hline \\ \cline{1-1}\end{array}$, then

$$\alpha : (u \wedge v) \otimes w \mapsto \begin{array}{|c|c|}\hline u & w \\ \hline v \\ \cline{1-1}\end{array} - \begin{array}{|c|c|}\hline v & w \\ \hline u \\ \cline{1-1}\end{array},$$

and

$$\pi : \begin{array}{|c|c|}\hline p & q \\ \hline r \\ \cline{1-1}\end{array} \mapsto (pq) \otimes r.$$

Thus, the map $\pi \circ \alpha$ for this choice of λ is given by

$$(u \wedge v) \otimes w \mapsto (uw) \otimes v - (vw) \otimes u.$$

November 2 – Constructing the irreps of GL_n, getting it right

Let $n \geq 1$ be a positive integer, l be a partition with $\ell(\lambda) \leq n$. Denote $d = |\lambda|$. Denote

$$E_\lambda := \bigotimes_k \wedge^{\lambda_k^T}(\mathbb{C}^n) \quad \text{and} \quad H_\lambda := \bigotimes_k \mathrm{Sym}^{\lambda_k^T}(\mathbb{C}^n).$$

Let $S_{\lambda,C}$ (resp $S_{\lambda,R}$) be the set of permutations within columns (resp. rows) of boxes in Young tableau of shape λ. Then, we have

$$\prod_k S_{\lambda_k^T} \cong S_{\lambda,C} \subset S_{\text{boxes}} \cong S_d \quad \text{and} \quad \prod_k S_{\lambda_k} \cong S_{\lambda,R} \subset S_{\text{boxes}} \cong S_d,$$

where S_{boxes} denotes the set of permutations of all the boxes in the Young tableau of shape λ.

Let $V^{\otimes \text{boxes}}$ be the quotient of $\mathbb{C}[S_{\text{boxes}}]$ by multi-linearity. Define

$$\alpha_\lambda := \sum_{\sigma \in S_{\lambda,C}} \mathrm{sgn}(\sigma) \cdot \sigma \quad \text{and} \quad \beta_\lambda := \sum_{\sigma \in S_{\lambda,R}} \sigma$$

as maps from $V^{\otimes \text{boxes}}$ to itself.

Example. Let $\lambda = (2,1)$ and label the boxes as $\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$. Then, we have

$$\alpha_{(2,1)} = \mathrm{id} - (1 \quad 3) \quad \text{and} \quad \beta_{(2,1)} = \mathrm{id} + (1 \quad 2).$$

Then, we have the following commutative diagram.

$$\begin{tikzcd}
V^{\otimes \text{boxes}} \arrow[rr, "\alpha_\lambda"] \arrow[dr, two heads] & & V^{\otimes \text{boxes}} \arrow[rr, "\beta_\lambda"] \arrow[dr, two heads] & & V^{\otimes \text{boxes}} \\
& E_\lambda \cong \mathrm{im}(\alpha_\lambda) \arrow[ur, hook] \arrow[rr, "\Phi"] \arrow[dr, two heads] & & \mathrm{im}(\beta_\lambda) \cong H_\lambda \arrow[ur, hook] & \\
& & \mathrm{im}(\beta_\lambda \circ \alpha_\lambda) \cong \mathrm{im}(\Phi) \cong V_\lambda \arrow[ur, hook] & &
\end{tikzcd}$$

Example. In our running example with $\lambda = (2,1)$, the above diagram becomes

$$\begin{tikzcd}
\begin{array}{|c|c|}\hline u & v \\ \hline w & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \arrow[rr, mapsto, "\alpha_\lambda"] \arrow[dr, mapsto] & & \begin{array}{|c|c|}\hline u & v \\ \hline w & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} - \begin{array}{|c|c|}\hline w & v \\ \hline u & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \arrow[rr, mapsto, "\beta_\lambda"] \arrow[dr, mapsto] & & \begin{array}{|c|c|}\hline u & v \\ \hline w & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} + \begin{array}{|c|c|}\hline v & u \\ \hline w & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} - \begin{array}{|c|c|}\hline w & v \\ \hline u & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} - \begin{array}{|c|c|}\hline v & w \\ \hline u & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \\
& (u \wedge w) \otimes v \arrow[ur, mapsto] \arrow[rr, mapsto, "\Phi"] \arrow[dr, mapsto] & & (uv) \otimes w - (vw) \otimes u \arrow[ur, mapsto] & \\
& & (uv) \otimes w - (vw) \otimes u \arrow[ur, mapsto] & &
\end{tikzcd}$$

Lemma. $\beta_\lambda \circ \alpha_\lambda \neq 0$

Proof. Let $\{e_1, \cdots, e_n\}$ be a basis for $\mathbb{C}^n$. Note that it suffices to check that $\beta_\lambda \circ \alpha_\lambda(T_0)$ is nonzero, where T_0 denotes the Young tableau of shape λ with e_1's on the first row, e_2's on the second row and so on. Note that the coefficient of T_0 in $\beta(U)$ is given by

$$\prod_k \lambda_k! \cdot (\text{coefficient of } T_0 \text{ in } U)$$

for any U. Since we do not get enough e_j's in the jth column to perform any non-trivial column permutation, the result follows. $\square$

Remark. $\beta_\lambda \circ \alpha_\lambda \in \mathbb{C}[S_n]$ is known as the ***Young symmetrizer***. It plays an important role in the representation theory of S_n; we sketch some of the key points. As you will check on the homework, $\beta_\lambda \mathbb{C}[S_n] \alpha_\lambda$ is one-dimensional, and is spanned by the Young symmetrizer. In the decomposition $\mathbb{C}[S_n] = \bigoplus_{W_\lambda:\text{irreps of } S_n} \mathrm{End}(W_\lambda)$, the Young symmetrizer lies in one summand. In that summand, it is rank one, and we have $(\beta_\lambda \circ \alpha_\lambda)^2 = c_\lambda(\beta_\lambda \circ \alpha_\lambda)$ where c_λ is a certain positive integer. (Specifically, c_λ is the product of the "hook lengths" of λ.) If we think of $\mathbb{C}[S_n]$ as block matrices indexed by the representations of S_n, then $\beta_\lambda \circ \alpha_\lambda$ is 0 in all but one block and, on that block, can be thought of as having a c_λ in the upper left and 0's everywhere else.

The representation W_λ of $\mathbb{C}[S_n]$ can be described as $\beta_\lambda \alpha_\lambda \mathbb{C}[S_n]$, acted on from the right, or as $\mathbb{C}[S_n]\beta_\lambda\alpha_\lambda$, acted on from the left.

Remark. What happens when k is a field of prime characteristic p? In this case, we have the commutative diagram

$$\begin{tikzcd}
V^{\otimes d} \arrow[r, "\alpha_\lambda"] \arrow[d, two heads] & V^{\otimes d} \arrow[rr, "\beta_\lambda"] \arrow[d, two heads] & & V^{\otimes d} \\
\wedge^d V \cong \mathrm{im}(\alpha_\lambda) \cong E_\lambda \arrow[ur, hook] \arrow[dr, two heads] & \mathrm{Sym}^d V \arrow[r, two heads] & \mathrm{im}(\beta) \arrow[r, hook] & (V^{\otimes d})^{S_d} \arrow[u, hook] \\
& \mathcal{S}_\lambda(V) \arrow[ur, hook] & &
\end{tikzcd}$$

Definition. The ***Schur functor***, denoted $\mathcal{S}_\lambda$, is a functor from the category of vector spaces over k to itself defined by

$$\mathcal{S}_\lambda(V) := \mathrm{im}\left(\beta_\lambda \circ \alpha_\lambda : V^{\otimes d} \to V^{\otimes d}\right)$$

for any vector space V over field k, and

$$\mathcal{S}_\lambda(f) := f^{\otimes d}|_{\mathcal{S}_\lambda(V)} : \mathcal{S}_\lambda(V) \to \mathcal{S}_\lambda(W)$$

for any linear map $f : V \to W$.

Example. For instance, we have

$$\begin{aligned}
s_{(k)}(V) &= \mathrm{Sym}^k(V) \\
s_{1^k}(V) &= \wedge^k V \\
s_{(2,1)}(V) &= \mathrm{im}\left(\beta_{(2,1)} \circ \alpha_{(2,1)} : \wedge^2 V \otimes V \to \mathrm{Sym}^2 V \otimes V\right)
\end{aligned}$$

November 4 – Constructing the irreps of GL_n with matrix minors

Today was an impromptu worksheet day! David was sick, so he sent in problems to solve,

We found (or nearly found) a representation with character s_λ. We are specifically concerned with the right action of GL_n on polynomials on n^2 variables z_{ij}, where we think of these as entries in an $n \times n$ matrix Z. So A acts by replacing Z with ZA^{-1}.

Note that the span of each row of Z is an n-dimensional sub-representation, since multiplication on the right sends z_{i*}, ie row i, to $z_{i*}A^{-1}$. In other words, A acts on each row seperately. Because this is a subrep, we can compute the character by setting A to the diagonal matrix with x_i^{-1} in the i,ith entry. This sends z_{ij} to $x_j z_{ij}$, so the character is $\sum x_j$.

Problem 19. Fix $1 \leq i \leq n$ and a nonnegative integer d. Show that the vector space of degree d homogeneous polynomials in $\{z_{i1}, z_{i2}, \ldots, z_{in}\}$ is a GL_n-representation and, as such, is isomorphic to $\mathrm{Sym}^d\mathbb{C}^n$.

As discussed before, the right action sends row i to row i linearly. We act on degree d polynomials by replacing each variable by a linear combination of the others, which is again a degree d polynomial. The torus (our diagonal matrix of x_i^{-1}) acts by rescaling each variable. The monomials are a basis, and we send each degree d monomial $z_{ij_1} \cdots z_{ij_d}$ to $x_{j_1} \cdots x_{j_d} z_{ij_1} \cdots z_{ij_d}$. So the character is

$$\sum_{j_1 \leq j_2 \leq \cdots \leq j_d} x_{j_1} \cdots x_{j_d} = h_d(x).$$

Since characters determine our representations, our representation must be isomorphic to Sym^d, whose character is h_d.

We can also see the isomorphism directly by identifying the product $z_{ij_1} \cdots z_{ij_d}$ with $e_{j_1} \cdots e_{j_d} \in \mathrm{Sym}^d$. Either way, this concludes the Problem.

Let $J = \{j_1, \ldots, j_k\}$ be a k-element subset of $\{1, \ldots, n\}$. Define the $k \times k$ minor Δ_J as the determinant of the first k rows and J columns of Z.

Problem 20. For $1 \leq k \leq n$, check that the subspace of R spanned by the Δ_J (where $|J| = k$) is a GL_n-subrepresentation, and that this subrepresentation is isomorphic to $\bigwedge^k \mathbb{C}^n$ as a GL_n-representation.

Let Z^j denote the jth column of Z, and $Z^j|_k$ the restriction of the jth column to the first k rows. Recall that $(ZA)^j = \sum_i Z^i A_{ij}$. So $\Delta_J = \det(Z^{j_1}|_k, \ldots, Z^{j_k}|_k)$ and acting by A gives us:

$$\begin{aligned}
\det((AZ)^{j_1}|_k, \ldots, (AZ)^{j_k}|_k) &= \sum_{i_1=1}^{n} A_{i_1 j_1} \det(Z^{i_1}|_k, (AZ)^{j_2}|_k, \ldots, (AZ)^{j_k}) \\
&= \sum_{I=\{i_1, i_2, \ldots, i_k\}} (\prod_{m=1}^{k} A_{i_m j_m}) \Delta_I
\end{aligned}$$

which is in the span of Δ_J. So this is a subrep. One may also decompose A in various ways to get nicer formulas, or notice some cancelation in the sum above.

The character of this representation is $\sum_{i_1 < \ldots < i_k} x_{i_1} \cdots x_{i_k} = e_k(x)$, the same as the character for $\bigwedge^k \mathbb{C}^n$. So they must be isomorphic. We can find a particular isomorphism by identifying Δ_J with $e_{j_1} \wedge e_{j_2} \wedge \ldots \wedge e_{j_k}$. This concludes the Problem.

For λ a partition, define U_λ to be the subspace of R spanned by $\Delta_{J_1} \Delta_{J_2} \cdots \Delta_{J_m}$ where $|J_i| = \lambda_i^T$. Recall the representations E_λ and H_λ, with characters e_λ and h_λ, from the previous classes.

Problem 21. Define a surjection $E_\lambda \to U_\lambda$.

Let $J_i = \{j_{1i}, j_{2i} \cdots\}$. Send

$$e_{j_{11}} \wedge e_{j_{21}} \wedge \cdots \wedge e_{j_{\lambda_1^T 1}} \otimes e_{j_{12}} \wedge e_{j_{22}} \wedge \cdots \wedge e_{j_{\lambda_2^T 2}} \otimes \cdots$$

to $\Delta_{J_1}\Delta_{J_2}\cdots$. Essentially the multiplication map after we take our identification with $\bigwedge^k \mathbb{C}^n$ from the previous problem. This is surjective, we can get any particular product of Δ_J's.

Problem 22. Define an injection $U_\lambda \to H_\lambda$.

Note that the variables z_{1j}, those in the first row of Z, appear with degree 1 in every Δ_J (provided J is not empty). Likewise the z_{ij} appear with degree 1 in every Δ_J with $|J| \geq i$. Thus in every monomial of $\Delta_{J_1}\Delta_{J_2}\cdots\Delta_{J_m} \in U_\lambda$ can be written as a product of $\lambda_i = \#\{k|\lambda_k^T \geq i\}$ of the z_{kj}s, ie grouping the variables from the same rows. We've already identified monomials of degree d with elements of $\mathrm{Sym}^d\mathbb{C}^n$ in the first problem of today. For example, we would send the monomial:

$$z_{1a}z_{2b}z_{3c}z_{1d}z_{2e}z_{1f}z_{1g} \mapsto e_ae_de_fe_g \otimes e_be_e \otimes e_c$$

Since this map is an injection on monomials, it is an injection on U_λ.

And now we are done - we have found a nonzero map from $E_\lambda \to H_\lambda$. Since the only common subrep of E_λ and H_λ is V_λ, we conclude that U_λ must have character s_λ, and is thus an irrep.

November 7 – Some examples of GL_n representations

We first review the material from the previous IBL class:

$$R = \mathbb{C}\begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & xnn \end{bmatrix},$$

with $\mathrm{GL}_n(\mathbb{C})$ acting via right multiplication. For $J = \{j_1, j_2, \cdots, j_k\} \subseteq \{1, \cdots, n\}$, define

$$\Delta_J = \det \begin{bmatrix} x_{1j_1} & \cdots & x_{1j_k} \\ \vdots & \ddots & \vdots \\ x_{kj_1} & \cdots & xkj_k \end{bmatrix}.$$

Then we showed $\underset{|J|=k}{\mathrm{span}}\Delta_J \cong \bigwedge^k \mathbb{C}^n$ as a $\mathrm{GL}_n(\mathbb{C})$-rep. If $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_n)$ is a partition then

$$\underset{|J_k|=\lambda_k^T}{\mathrm{span}}\ (\Delta_{J_1}\Delta_{J_2}\cdots\Delta_{J_n})$$

is an irreducible representation of $\mathrm{GL}_n(\mathbb{C})$ with character s_λ. For example, take $n = 3$ and $\lambda = (2, 1)$ so that

$$s_\lambda(x_1, x_2, x_3) = x_1^2x_2 + x_1x_2^2 + x_1^2x_3 + 2x_1x_2x_3 + x_2^2x_3 + x_1x_3^2 + x_2x_3^2.$$

Let

$$V_\lambda := \underset{|J_k|=\lambda_k^T}{\mathrm{span}}\ (\Delta_{J_1}\Delta_{J_2}\cdots\Delta_{J_n}).$$

We have

$$\dim(V_\lambda) = \chi_{V_\lambda}(\mathrm{Id}_3) = s_\lambda(1, 1, 1) = 8,$$

despite V_λ being the span of 9 polynomials:

$$V_\lambda = \operatorname{span}\begin{pmatrix} \Delta_{12}\Delta_1 & \Delta_{12}\Delta_2 & \Delta_{12}\Delta_3 \\ \Delta_{13}\Delta_1 & \Delta_{13}\Delta_2 & \Delta_{13}\Delta_3 \\ \Delta_{23}\Delta_1 & \Delta_{23}\Delta_2 & \Delta_{23}\Delta_3 \end{pmatrix}.$$

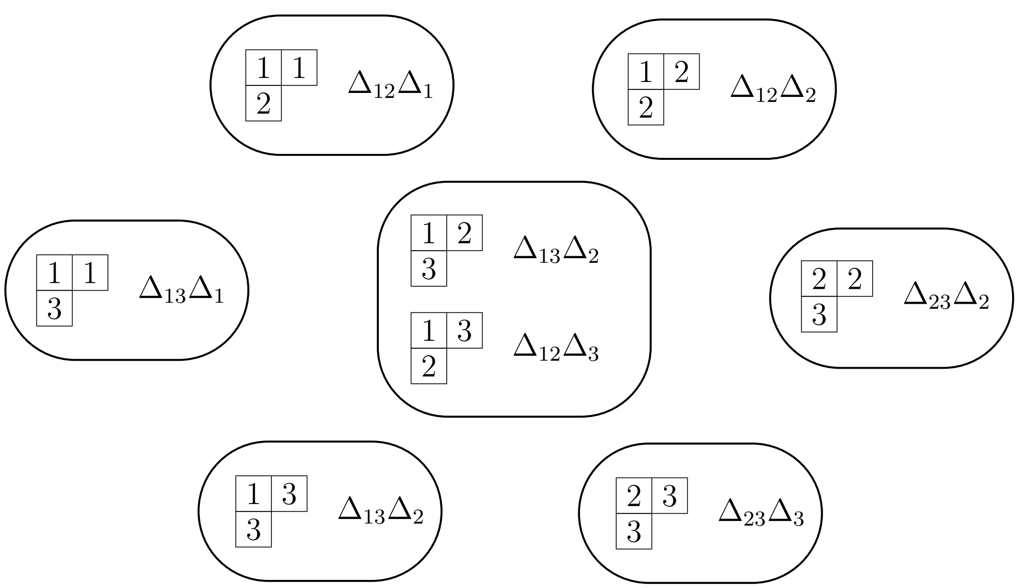
There must therefore be a linear relationship between these polynomials. Indeed $\Delta_{12}\Delta 3 - \Delta_{12}\Delta_2 + \Delta_{23}\Delta_1 = 0$. One way to find this relationship is that

$$\begin{aligned}\Delta_{12}\Delta_3 - \Delta_{12}\Delta_2 + \Delta_{23}\Delta_1 &= \begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} x_{13} - \begin{vmatrix} x_{11} & x_{13} \\ x_{21} & x_{23} \end{vmatrix} x_{12} + \begin{vmatrix} x_{12} & x_{13} \\ x_{22} & x_{23} \end{vmatrix} x_{11} \\ &= \det\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix} = 0.\end{aligned}$$

Another way to look at this is to decompose V_λ into T-eigenspaces, where $T \subseteq \mathrm{GL}_3(\mathbb{C})$ is the torus. The matrix $\begin{bmatrix} x_1 & & \\ & x_2 & \\ & & x_3 \end{bmatrix} \in T$ acts on $\operatorname{span}(\Delta_{ab}\Delta_c)$ through multiplication by $x_a x_b x_c$. This is because it takes z_{ij} to $x_j z_{ij}$, so it maps

$$\Delta_{ab}\Delta_c = (z_{1a}z_{2b} - z_{2a}z_{1b})z_{1c} \mapsto ((x_a z_{1a})(x_b z_{2b}) - (x_a z_{2a})(x_b z_{1b}))x_c z_{1c} = (x_a x_b x_c)\Delta_{ab}\Delta_c.$$

Therefore this matrix will act on $\operatorname{span}(\Delta_{23}\Delta_1, \Delta_{13}\Delta_2, \Delta_{12}\Delta_3)$ via multiplication by $x_1x_2x_3$. Since the coefficient of the $x_1x_2x_3$ term in $s_\lambda(x_1, x_2, x_3)$ (the character of V_λ) is 2, the dimension of $\operatorname{span}(\Delta_{23}\Delta_1, \Delta_{13}\Delta_2, \Delta_{12}\Delta_3)$ must be 2 or less, meaning the linear dependence must be among these three terms. In other words, if we associate the polynomial $\Delta_{ab}\Delta_c$ with the young tableau $\begin{array}{|c|c|}\hline a & c \\ \hline b \\ \cline{1-1}\end{array}$, then the coefficient for $x_1x_2x_3$ in $s_\lambda(x_1, x_2, x_3)$ is already covered by $\Delta_{13}\Delta_2$ and $\Delta_{12}\Delta_3$ since they're associated with a semi standard young tableau, whereas $\Delta_{23}\Delta_1 \leftrightarrow \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 \\ \cline{1-1}\end{array}$ is not. Therefore we can safely discard $\Delta_{23}\Delta_1$ without impacting the span. The T-eigenspaces of this representation can therefore be drawn as follows:



Recall from previous lectures the decomposition of algebraic representations of $C^* = \mathrm{GL}_1$. Namely, if C^* acts on V then V can be broken up as the direct sum

$$V = \bigoplus_{k=-\infty}^{\infty} V_k,$$

where $[t]$ acts on V_k by t^k. On homework, we showed that algebraic representations of $\mathbb{C}^+ \cong \left\{ \begin{bmatrix} 1 & u \\ 0 & 1 \end{bmatrix} : u \in \mathbb{C} \right\}$ are $\exp(uN) = \sum_{k=0}^{\infty} \frac{u^k N^k}{k!}$ for some nilpotent N. Also shown on homework was that if $G = \left\{ \begin{bmatrix} t & u \\ 0 & 1 \end{bmatrix} : u \in \mathbb{C}, t \in \mathbb{C}^* \right\}$ then the nilpotent N takes V_k to V_{k+1} where $\left[\begin{smallmatrix} t & 0 \\ 0 & 1 \end{smallmatrix}\right]$ acts on V_k by multiplication by t^k, i.e.

$$\begin{tikzcd}[column sep=small] V = \cdots & V_{-2} \arrow[r, bend left, "N"] & \oplus V_{-1} \arrow[r, bend left, "N"] & \oplus V_0 \arrow[r, bend left, "N"] & \oplus V_1 \arrow[r, bend left, "N"] & \oplus V_2 \cdots \end{tikzcd}$$

The key relationship here is that

$$\begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & u \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & tu \\ 0 & 1 \end{bmatrix} \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix}.$$

If $v \in V_k$, this means that

$$\rho\left(\begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix}\right)\left(v + uNv + \frac{u^2N^2v}{2} + \cdots\right) = \left(I + tuN + \frac{t^2u^2N^2}{2} + \cdots\right) t^k v.$$

The linear term says that

$$\rho\left(\begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix}\right)(Nv) = t^{k+1}Nv,$$

so $Nv \in V_{k+1}$. More generally, if GL_n acts on V then V can be written as a direct sum

$$V = \bigoplus_{\mu} V_\mu,$$

where $\left[\begin{smallmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{smallmatrix}\right] \in T \subseteq \mathrm{GL}_n$ acts on V_μ by multiplication by $t_1^{\mu_1} \cdots t_n^{\mu_n}$. Rather than just looking at the subgroup T of diagonal matrices, we can put a nonzero entry in the ith row and jth column, e.g.

$$\begin{array}{c c} & \begin{array}{cccc} & & j & \end{array} \\ \begin{array}{c} \\ i \\ \\ \\ \end{array} & \begin{pmatrix} * & & & \\ & * & * & \\ & & * & \\ & & & * \end{pmatrix} \end{array} \qquad \cong T \ltimes \mathbb{C}^+$$

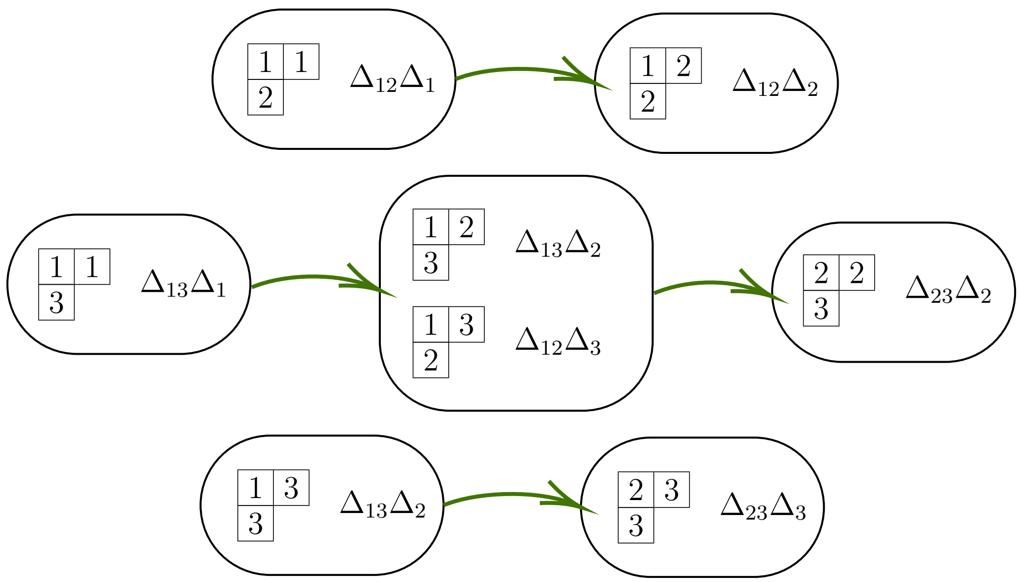
We'll have $\begin{array}{c c} & \begin{array}{cccc} & & j & \end{array} \\ \begin{array}{c} \\ i \\ \\ \\ \end{array} & \begin{pmatrix} 1 & & & \\ & 1 & u & \\ & & 1 & \\ & & & 1 \end{pmatrix} \end{array}$ acting by $\exp(uN_{ij})$ where $N_{ij} : V_\mu \to V_{\mu+e_i-e_j}$. For example, with $\lambda = (2,1)$, the matrix $\left[\begin{smallmatrix} 1 & & \\ u & 1 & \\ & & 1 \end{smallmatrix}\right]$ will take $\Delta_{12} \mapsto \Delta_{12}$ since a column operation won't change the determinant of the minor. It will map $\Delta_1 \mapsto \Delta_1 + u\Delta_2$. Therefore

$$\sum_{k=0}^{\infty} \frac{u^k N_{21}^k}{k!} \Delta_{12}\Delta_1 = \exp(uN_{21})(\Delta_{12}\Delta_1) = \Delta_{12}\Delta_1 + u\Delta_{12}\Delta_2.$$

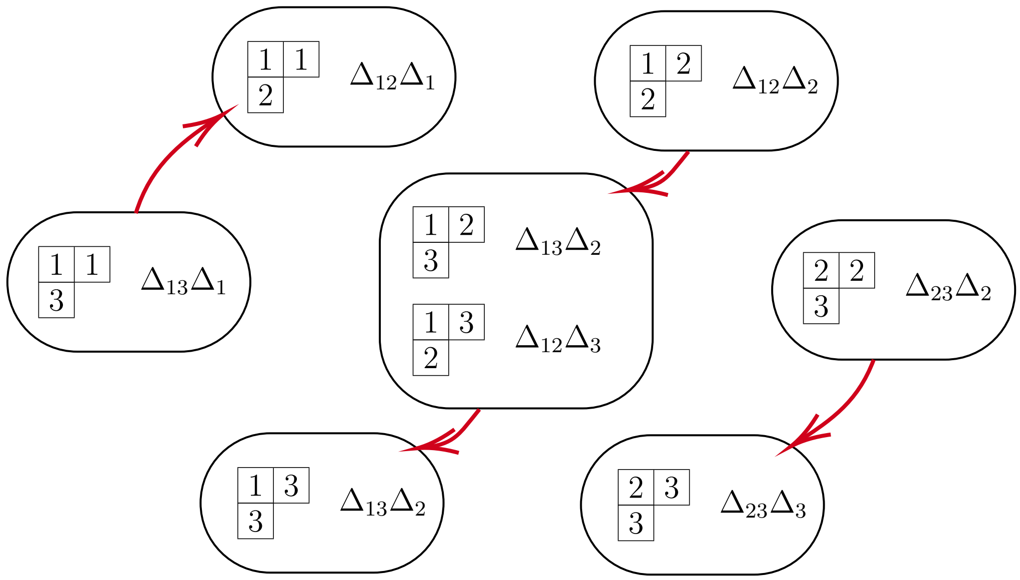
Looking at the linear term tells us that N_{21} will take elements in $\mathrm{span}(\Delta_{12}\Delta_1)$ to $\mathrm{span}(\Delta_{12}\Delta_1)$. Similarly the matrix takes $\Delta_{13} \mapsto \Delta_{13} + u\Delta_{23}$ and $\Delta_2 \mapsto \Delta_2$ so N_{21} takes $\Delta_{13}\Delta_2$ to $\Delta_{23}\Delta_2$. As a final example

$$\Delta_{13}\Delta_1 \mapsto (\Delta_{13} + u\Delta_{23})(\Delta_1 + u\Delta_2) = \Delta_{13}\Delta_1 + u(\Delta_{13}\Delta_2 + \Delta_{23}\Delta_1) + u^2\Delta_{23}\Delta 2.$$

Again looking at the linear term, we find that N_{21} takes $\Delta_{13}\Delta_1$ into the $\mathrm{span}(\Delta_{13}\Delta_2, \Delta_{23}\Delta_1) = \mathrm{span}(\Delta_{13}\Delta_2, \Delta_{12}\Delta_3)$. Completing this process for every basis element gives us the following arrows for N_{21}



The transpose matrix $\begin{bmatrix} 1 & u & \\ & 1 & \\ & & 1 \end{bmatrix}$ will have corresponding nilpotent N_{12} which reverses the direction of these arrows. The matrix $\begin{bmatrix} 1 & & \\ & 1 & \\ & u & 1 \end{bmatrix}$ will have corresponding nilpotent N_{32} which follows different arrows:



Again N_{23} will just reverse the direction of these arrows.

November 9 – Weight spaces and the maps between them

We begin by discussing the language of weight spaces. Consider the torus $T \cong (\mathbb{C}^*)^n \subset \mathrm{GL}_n$ which consists of matrices of the form $\begin{bmatrix} t_1 & & \\ & \ddots & \\ 0 & & t_n \end{bmatrix}$. If V is an algebraic representation of T, then $V = \oplus_{\alpha \in \mathbb{Z}^n} V_\alpha$ where

$$V_\alpha = \left\{ v \in V \mid \begin{bmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{bmatrix} \vec{v} = t_1^{\alpha_1} \dots t_n^{\alpha_n} \vec{v} \right\}.$$

We call $(\alpha_1, \dots, \alpha_n)$ a ***weight***. The space V_α is the α-weight space, and a vector $v \in V$ is an α-weight vector.

We now discuss how to prove the decomposition of V given above. When $\rho : T \to \mathrm{GL}_n$ is algebraic, this implies the entries of $\big[\rho_{ij}(t_1, \dots, t_n)\big]$ are Laurent polynomials in the t_i. That is, they are of the form $\sum P_{ij,\alpha} t^\alpha \in \mathbb{C}[t_1^\pm, \dots, t_n^\pm]$. Thus $\rho \begin{bmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{bmatrix} = \sum_{\alpha \in \mathbb{Z}^n} P_\alpha t_1^{\alpha_1} \dots t_n^{\alpha_n}$. Since ρ is a homomorphism,

$$\rho\left(\begin{bmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{bmatrix}\begin{bmatrix} u_1 & & \\ & \ddots & \\ & & u_n \end{bmatrix}\right) = \rho\left(\begin{bmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{bmatrix}\right)\rho\left(\begin{bmatrix} u_1 & & \\ & \ddots & \\ & & u_n \end{bmatrix}\right)$$

Therefore, $\sum_\gamma P_\gamma t^\gamma u^\gamma = \big(\sum_\alpha P_\alpha t^\alpha\big)\big(\sum_\beta P_\beta u^\beta\big) = \sum_{\alpha,\beta} P_\alpha P_\beta t^\alpha u^\beta$. Matching coefficients on both sides yields

$$P_\alpha P_\beta = \begin{cases} P_\alpha & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases}$$

Further, $\rho(\mathrm{Id}_n) = \mathrm{Id}_n$, so $\sum_\alpha P_\alpha = \mathrm{Id}_n$. It follows that V breaks up as described above.

When $n = 2$, we see that

$$V_{\square\square} = \mathrm{Sym}^2\mathbb{C}^2 = \mathbb{C}e_1e_1 \oplus \mathbb{C}e_1e_2 \oplus \mathbb{C}e_2e_2$$

and

$$V_{\ydiagram{1,1}} = \bigwedge^2 \mathbb{C}^2 = \mathbb{C}(e_1 \wedge e_2)$$

where V_λ has character s_λ. Since $\begin{bmatrix} t_1 & 0 \\ 0 & t_2 \end{bmatrix}$ acts on $\mathbb{C}e_1e_1$ by t_1^2, the weight is $(2,0)$. Similarly, the weight on $\mathbb{C}e_1e_2$ and on $\mathbb{C}(e_1 \wedge e_2)$ is $(1,1)$, and the weight on $\mathbb{C}e_2e_2$ is $(0,2)$.

Example. In the homework, we particularly thought about the case where $|\lambda| = n$ and we look at the weight $(1,1,\dots,1)$. We called the corresponding space W_λ.

Here are the values for $n = 2$ and $n = 3$:

$$W_{\ydiagram{2}} = \mathbb{C}e_1e_2$$

and

$$W_{\ydiagram{1,1}} = \mathbb{C}(e_1 \wedge e_2)$$

and S_2 acts trivially on $W_{\ydiagram{2}}$ and by sign on $W_{\ydiagram{1,1}}$. In the $n = 3$ case we have

$$\begin{aligned}
V_{\ydiagram{3}} &= \mathrm{Sym}^3\mathbb{C}^3 = \mathbb{C}e_1^3 \oplus \cdots \oplus \mathbb{C}e_1e_2e_3 \oplus \cdots \oplus \mathbb{C}e_3^3 \\
V_{\ydiagram{2,1}} &= \mathbb{C}(\Delta_{12}\Delta_1) \oplus \cdots \oplus \mathbb{C}(\Delta_{23}\Delta_3) \oplus \mathbb{C}\{\Delta_{12}\Delta_3, \Delta_{13}\Delta_2\} \\
V_{\ydiagram{1,1,1}} &= \mathbb{C}(e_1 \wedge e_2 \wedge e_3) \\
W_{\ydiagram{3}} &= \mathbb{C}(e_1e_2e_3) \\
W_{\ydiagram{2,1}} &= \mathbb{C} = (\Delta_{12}\Delta_3) \oplus \mathbb{C}(\Delta_{13}\Delta_2) = \mathrm{Span}(\Delta_{12}\Delta_3, \Delta_{13}\Delta_2, \Delta_{23}\Delta_1) \\
W_{\ydiagram{1,1,1}} &= \mathbb{C}(e_1 \wedge e_2 \wedge e_3)
\end{aligned}$$

The W's are called Specht modules, and they are irreps of S_n.

Let $1 \leq i \neq j \leq n$. Let E_{ij} be the matrix with 1 in the (i,j) entry and 0 elsewhere. Then $Id + uE_{ij} \cong \mathbb{C}$ forms a copy of the additive group and acts on V by $\sum \frac{u^k}{k!} N_{ij}^k = \exp(uN_{ij})$. Matrices of the form $\begin{bmatrix} t_1 & & \\ & \ddots & u \\ & & t_n \end{bmatrix}$ form a semidirect product, and N_{ij} maps V_α to $V_{\alpha+e_i-e_j}$.

We now begin a discussion on Lie algebras. Let G be a smooth Lie group and $\mathfrak{g} = T_{\mathrm{Id}}G$. We will think of having $G = \mathrm{GL}_n\mathbb{C}$ and $\mathfrak{g} = \mathrm{Mat}_{n\times n}\mathbb{C}$.

Let $\rho : G \to \mathrm{GL}_n$ be a representation and $g \in \mathfrak{g}$ be an $n \times n$ matrix. Then $\rho(\mathrm{Id} + \epsilon g)$ is $\rho(\mathrm{Id}) + \epsilon r(g)$ plus higher terms in ϵ where $r : \mathfrak{g} \to \mathrm{Mat}_{n\times n}\mathbb{C}$ is $r = D\rho$, a map of tangent spaces.

Proposition. Let $G = \mathrm{GL}_n$. Then r determines ρ.

Proof. Suppose we have two representations $\rho_1, \rho_2 : G \to \mathrm{GL}_n$ with the same derivative r. For any $g \in \mathfrak{g}$ with $M >> 0$, we have $\rho_1((1 + \frac{g}{M})^M) = (\rho_1(1 + \frac{g}{M}))^M = (\mathrm{Id} + \frac{r(g)}{M} + O(\frac{1}{M^2}))^M$.

Then $\exp(g) = \sum_{k=0}^{\infty} \frac{g^k}{k!} = \lim_{n\to\infty}(1+\frac{g}{n})^n$, so $\rho(\exp(g)) = \exp(r(g))$. Thus, $\rho_1 = \rho_2$ on $\{\exp(g) \mid g \in \mathfrak{g}\}$.

There is an open set Ω containing Id_n such that every matrix in Ω is an exponential. (Intuition: log series converegs near the identity.) Let $X = \{g \in \mathrm{GL}_n \mid \rho_1(g) = \rho_2(g)\}$. Clearly, X is closed in G. If $g \in X$, $u \in \Omega$, then $\rho_1(gu) = \rho_1(g)\rho_1(u) = \rho_2(g)\rho_2(u) = \rho_2(gu)$, which implies $gu \in X$. So, for any $g \in G$, $g\Omega$ is an open neighborhood of g in X. It follows that X is both open and closed. Since it is in a connected space, we conclude $X = G$. $\square$

If

$$\rho\left(\begin{bmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{bmatrix}\right) = \begin{bmatrix} t_1^{\alpha^1} & & \\ & \ddots & \\ & & t_n^{\alpha^N} \end{bmatrix}$$

then

$$r\left(\begin{bmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{bmatrix}\right) = \begin{bmatrix} \alpha^1 \cdot s & & \\ & \ddots & \\ & & \alpha^N \cdot s \end{bmatrix}$$

Consider $T \subset \mathrm{GL}_n$, $t \subset \mathfrak{g}$ where $\mathfrak{g}$ is the set of $n \times n$ matrices and t is the set of diagonal matrices. Weight spaces are also eigenspaces of $r\left(\begin{bmatrix} * & & 0 \\ & \ddots & \\ 0 & & * \end{bmatrix}\right)$ and $r\left(\begin{bmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{bmatrix}\right)$ acts by $\alpha_1 \cdot s_1 + \cdots + \alpha_n \cdot s_n$ on V_α. We also have

$$\rho\left(\begin{bmatrix} 1 & & \\ & \ddots & u \\ & & 1 \end{bmatrix}\right) = \exp(uN)$$

Taking the derivative at $u = 0$, we see

$$r\left(\begin{bmatrix} 0 & & \\ & \ddots & 1 \\ & & 0 \end{bmatrix}\right) = N$$

November 11 – The Lie Algebra

Let G be a smooth Lie group and let $\mathfrak{g}$ be T_eG, the tangent space to G at the identity. Let $\rho : G \to GL_n$ be a representation, . Then $r : \mathfrak{g} \to \mathrm{Mat}_{N\times N}$ is the derivative $D_e(\rho)$. Concretely, if G is GL_n, then $\mathfrak{g}$ is $\mathrm{Mat}_{n\times n}$, and we have

$$\rho(\mathrm{Id} + \epsilon g) = \mathrm{Id}_N + \epsilon r(g) + O(\epsilon^2).$$

For G connected, the representation ρ is determined by r:

$$\rho(\exp(g)) = \exp(r(g)).$$

Here is a sketch of a proof for $G = \mathrm{GL}_n$: We have $\exp(g) = \lim_{M\to\infty}(1+g/M)^M$. So

$$\begin{aligned}\rho(\exp(g)) = \rho\left(\lim_{M\to\infty}(1+g/M)^M\right) &= \lim_{M\to\infty}\rho((1+g/M)^M) = \lim_{M\to\infty}\rho(1+g/M)^M \\ &= \lim_{M\to\infty}(1 + r(g)/M + O(1/M^2))^M = \exp(r(g)).\end{aligned}$$

Next we discussed in the GL_n case what happens with commutators. Let $g, h \in \mathrm{Mat}_{n\times n}$, and $\delta, \epsilon \in \mathbb{R}_{>0}$. Then we have

$$\begin{aligned}(1+g\delta)(1+h\epsilon)(1+g\delta)^{-1}(1+h\epsilon)^{-1} =&(1+g\delta)(1+h\epsilon)(1-g\delta+g^2\delta^2-\dots)(1-h\epsilon+h^2\epsilon^2-\dots)\\ =&1+(gh-gh-hg+gh)\delta\epsilon+O(\delta+\epsilon)^3\\ =&1+[g,h]\delta\epsilon+O(\delta+\epsilon)^3\end{aligned}$$

where $[g,h] := gh - hg$.

In general we get

$$[,] : \mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$$

such that if $\alpha(t)$ and $\beta(u)$ are paths, then

$$\begin{aligned}\alpha(0) =&\mathrm{Id}\\ \alpha'(0) =&g\\ \beta(0) =&\mathrm{Id}\\ \beta'(0) =&h\end{aligned}$$

and

$$\alpha(t)\beta(u)\alpha(t)^{-1}\beta(u)^{-1} = \mathrm{Id} + (tu)[g,h] + \dots.$$

So

$$(1+r(g)\delta+\dots)(1+r(h)\epsilon+\dots)(1+r(g)\delta+\dots)^{-1}(1+r(h)\epsilon+\dots)^{-1} = 1+r([g,h])\delta\epsilon+\dots$$

which implies

$$[r(g), r(h)] = r([g,h]). \tag{1}$$

A map $r : \mathfrak{g} \to \mathrm{Mat}_{N\times N}$ obeying $[r(g), r(h)] = r([g,h])$ is called a ***Lie algebra homomorphism***.

Example. What happens in GL_2?

We can split

$$\mathfrak{g}l_2 = \{\left[\begin{smallmatrix} * & * \\ * & * \end{smallmatrix}\right]\} = \mathbb{C}\underbrace{\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]}_{\mathrm{E}} \oplus \underbrace{\left[\begin{smallmatrix} * & 0 \\ 0 & * \end{smallmatrix}\right]}_{\mathfrak{t}} \oplus \mathbb{C}\underbrace{\left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]}_{\mathrm{F}}.$$

Then we can compute

$$\begin{gathered}\left[\left[\begin{smallmatrix} t_1 & 0 \\ 0 & t_2 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]\right] = \left[\begin{smallmatrix} 0 & t_1 \\ 0 & 0 \end{smallmatrix}\right] - \left[\begin{smallmatrix} 0 & t_2 \\ 0 & 0 \end{smallmatrix}\right] = (t_1-t_2)E\\ \left[\left[\begin{smallmatrix} t_1 & 0 \\ 0 & t_2 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]\right] = \left[\begin{smallmatrix} 0 & 0 \\ t_2 & 0 \end{smallmatrix}\right] - \left[\begin{smallmatrix} 0 & 0 \\ t_1 & 0 \end{smallmatrix}\right] = (t_2-t_1)F\\ \left[\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]\right] = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] - \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right].\end{gathered}$$

Continuing with this example we have

$$\rho : GL_2 \to GL(V)$$

with $V = \oplus_{(a_1,a_2)\in\mathbb{Z}^2} V_{a_1a_2}$. We can view ρ as

$$\rho\left(\left[\begin{smallmatrix} T_1 & 0 \\ 0 & T_2 \end{smallmatrix}\right]\right) : V_{a_1a_2} \xrightarrow{T_1^{a_1}T_2^{a_2}} V_{a_1a_2}$$

then

$$r\left(\left[\begin{smallmatrix} t_1 & 0 \\ 0 & t_2 \end{smallmatrix}\right]\right) : V_{a_1a_2} \xrightarrow{a_1t_1+a_2t_2} V_{a_1a_2}.$$

Note that

$$r(E) : V_{a_1a_2} \to V_{(a_1+1)(a_2-1)}$$

and

$$(1+\epsilon t_1)^{a_1}(1+\epsilon t_2)^{a_2} = 1+\epsilon(a_1t_1+a_2t_2)+\dots.$$

We see that

$$\left[r\left[\begin{smallmatrix} t_1 & 0 \\ 0 & t_2\end{smallmatrix}\right], r(E)\right] = (t_1-t_2)r(E)$$

so

$$r\left[\begin{smallmatrix} t_1 & 0 \\ 0 & t_2\end{smallmatrix}\right] r(E) = -r(E)r\left[\begin{smallmatrix} t_1 & 0 \\ 0 & t_2\end{smallmatrix}\right] = (t_1-t_2)r(E).$$

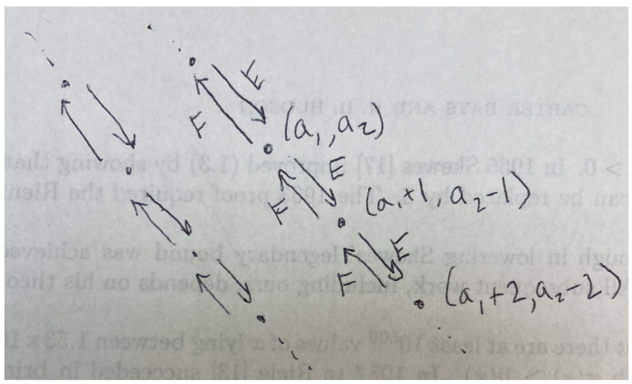
If $v \in V_{a_1a_2}$ we get

$$r\left[\begin{smallmatrix} t_1 & 0 \\ 0 & t_2\end{smallmatrix}\right] r(E)v - r(E)(t_1a_1+t_2a_2)V = (t_1-t_2)r(E)v$$

or

$$r\left[\begin{smallmatrix} t_1 & 0 \\ 0 & t_2\end{smallmatrix}\right] r(E)v = (t_1(a_1+1)+t_2(a_2-1))r(E)v.$$

So we get this diagram (possibly with more diagonals):



On v_{a_1,a_2}, we have the relation $EF-FE=(a_1-a_2)\mathrm{Id}$, since $[E,F]=\left[\begin{smallmatrix} 1 & 0 \\ 0 & -1\end{smallmatrix}\right]$, and we know that $\left[\begin{smallmatrix} 1 & 0 \\ 0 & -1\end{smallmatrix}\right]$ acts on the (a_1,a_2) weight space by a_1-a_2.

Let's see how this works for the irreps of GL_2.. We know GL_2 irreps correspond to $(\lambda_1,\lambda_2)\in\mathbb{Z}^2$ with $\lambda_1\geq\lambda_2$, so we can write $(k+\ell,\ell)$ with $k\geq 0$. They are of the form $(\bigwedge^2\mathbb{C}^2)^{\otimes l}\otimes \mathrm{Sym}^k\mathbb{C}^2$. So the character is

$$(x_1x_2)^l(x_1^k+x_1^{k-1}x_2+\dots+x_2^k).$$

The weight diagram looks like

$$\begin{tikzcd}[column sep=small, row sep=small]
(\ell,k+\ell) \arrow[dr, shift left, "E"] & & & & \\
& (\ell+1,k+\ell-1) \arrow[ul, shift left, "F"] \arrow[dr, shift left, "E"] & & & \\
& & \cdots \arrow[ul, shift left, "F"] \arrow[dr, shift left, "E"] & & \\
& & & (k+\ell-1,\ell+1) \arrow[ul, shift left, "F"] \arrow[dr, shift left, "E"] & \\
& & & & (k+\ell,\ell) \arrow[ul, shift left, "F"]
\end{tikzcd}$$

Let's compute the actual maps. They are maps between one dimensional vector spaces, but we need to choose a particular vector in each vector space in order to write down specific maps. Let e_1, e_2 be a basis for $\mathbb{C}^2$. Then, for $0 \le a \le k$, the vector $(e_1 \wedge e_2)^\ell e_1^a e_2^{k-a}$ is in $(a+\ell, k+\ell-a)$ weight space of $(\bigwedge^2 \mathbb{C}^2)^{\otimes l} \otimes \mathrm{Sym}^k \mathbb{C}^2$.

The subgroup $\left[\begin{smallmatrix} 1 & u \\ 0 & 1 \end{smallmatrix}\right]$ preserves $e_1 \wedge e_2$. The action on $e_1^a e_2^{k-a}$ is

$$\rho\left(\left[\begin{smallmatrix} 1 & u \\ 0 & 1 \end{smallmatrix}\right]\right)\left(e_1^a e_2^{k-a}\right) = e_1^a (e_2 + u e_1)^{k-a} = e_1^a e_2^{k-a} + u(k-a) e_1^{a+1} e_2^{k-a-1} + \cdots .$$

So

$$r\left(\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]\right) e_1^a e_2^{k-a} = (k-a) e_1^{a+1} e_2^{k-a-1}$$

and the maps are

$$\begin{tikzcd}
(\ell, k+\ell) \arrow[dr, shift left, "k"] & & & & \\
& (\ell+1, k+\ell-1) \arrow[ul, shift left, "1"] \arrow[dr, shift left, "k-1"] & & & \\
& & \cdots \arrow[ul, shift left, "2"] \arrow[dr, shift left, "k-2"] & & \\
& & & (k+\ell-1, \ell+1) \arrow[ul, shift left, "k-1"] \arrow[dr, shift left, "2"] & \\
& & & & (k+\ell, \ell) \arrow[ul, shift left, "k"]
\end{tikzcd}$$

To repeat, every algebraic GL_2 rep breaks up into strings like this, which we call GL_2-strings.

November 14 - $\mathfrak{gl}(2)$ Strings

Recall from last time that every $\mathfrak{gl}(2)$ representation breaks up into irreducible representations as follows: There is some ordered pair (a, b) of integers, with $a \le b$, such that the weights of the representation are (a, b), $(a+1, b-1)$, $(a+2, b-2)$, $\dots$, (b, a). The actions of the Lie algebra elements $E = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ and $F = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$ are as shown below:

$$\begin{tikzcd}
(a,b)\ \bullet \arrow[dr, bend left, "b-a"] & & & & \\
& \bullet \arrow[ul, bend left, "1"] \arrow[dr, bend left, "b-a-1"] & & & \\
& & \bullet \arrow[ul, bend left, "2"] & & \\
& & & \bullet \arrow[ul, phantom, "\ddots"] \arrow[dr, bend left, "1"] & \\
& & & & \bullet\ (b,a) \arrow[ul, bend left, "b-a"]
\end{tikzcd}$$

Let's look in detail at the representation $V_{(2,1)}$ of $\mathfrak{gl}(3)$. It is 8 dimensional and spanned by matrix minors. We have a copy of gl(2) sitting inside generated by $\left[\begin{smallmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$ We can obtain the action of these elements on this representation by looking at the coffeicient of u in the action of $\left[\begin{smallmatrix} 1 & 0 & 0 \\ u & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 & u & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right]$

$$\rho\left(\left[\begin{smallmatrix} 1 & 0 & 0 \\ u & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right]\right)(\Delta_{12}\Delta_1) = \Delta_{12}(\Delta_1 + u\Delta_2) = \Delta_{12}\Delta_1 + u\Delta_{12}\Delta_2$$

Thus we have the following string

$$\begin{tikzcd} \Delta_{12}\Delta_2 \arrow[r, bend right] & \Delta_{12}\Delta_1 \arrow[l, bend right] \end{tikzcd}$$

Similarly we get the following string

$$\begin{tikzcd} \Delta_{13}\Delta_3 \arrow[r, bend right] & \Delta_{12}\Delta_1 \arrow[l, bend right] \end{tikzcd}$$

$$\rho(\left[\begin{smallmatrix} 1 & 0 & 0 \\ u & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right])(\Delta_{13}\Delta_1) = (\Delta_{13} + u\Delta_{23})(\Delta_1 + u\Delta_2) = \Delta_{13}\Delta_1 + u(\Delta_{13}\Delta_2 + \Delta_{23}\Delta_1) + u^2\Delta_{23}\Delta_2$$

$$\rho(\left[\begin{smallmatrix} 1 & 0 & 0 \\ u & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right])(\Delta_{13}\Delta_2+\Delta_{23}\Delta_1) = (\Delta_{13}+u\Delta_{23})(\Delta_2)+\Delta_{23}(\Delta_1+u\Delta_2) = \Delta_{13}\Delta_2+\Delta_{23}\Delta_1+u(2\Delta_{23}\Delta_2$$

So we can normalize our basis vectors and get the following string

$$\begin{tikzcd} \Delta_{23}\Delta_2 \arrow[r, bend right, "2"'] & \frac{\Delta_{23}\Delta_1+\Delta_{13}\Delta_2}{2} \arrow[l, bend right] \arrow[r, bend right] & \Delta_{13}\Delta_1 \arrow[l, bend right, "2"'] \\ & \Delta_{13}\Delta_2 - \Delta_{23}\Delta_1 & \end{tikzcd}$$

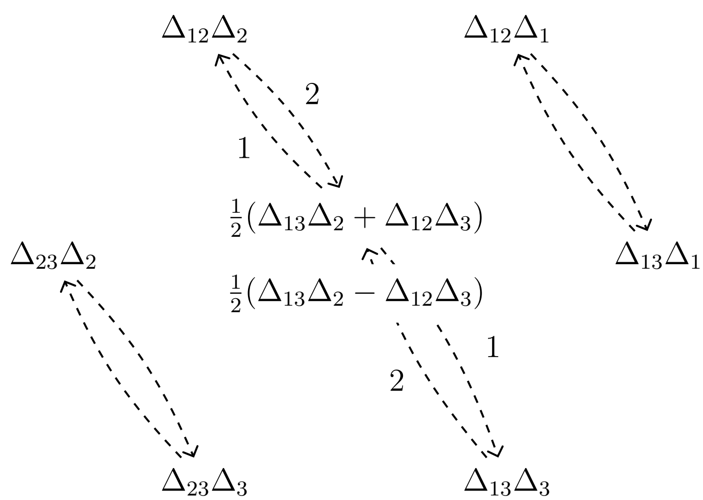
Putting it all together we get

$$\begin{tikzcd} \Delta_{12}\Delta_2 \arrow[r, bend left] & \Delta_{12}\Delta_1 \arrow[l, bend left] \end{tikzcd}$$

$$\begin{tikzcd} & \frac{1}{2}(\Delta_{23}\Delta_1 + \Delta_{13}\Delta_2) \arrow[dl, bend right, "1"] \arrow[dr, bend left, "1"] & \\ \Delta_{23}\Delta_2 \arrow[ur, bend left, "2"] & & \Delta_{13}\Delta_1 \arrow[ul, bend left, "2"] \\ & \frac{1}{2}(\Delta_{13}\Delta_2 - \Delta_{23}\Delta_1) & \end{tikzcd}$$

$$\begin{tikzcd} \Delta_{23}\Delta_3 \arrow[r, bend left] & \Delta_{13}\Delta_3 \arrow[l, bend left] \end{tikzcd}$$

We have a similar diagram for the action of $\left[\begin{smallmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$



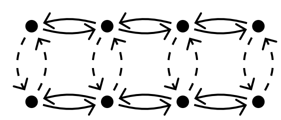
We now consider the general $\mathfrak{gl}(n)$. We define e_k as the matrix with 1 in the position $(k, k+1)$ and zero everywhere else. Similarly we define f_k is defined as the matrix with 1 in $(k+1, k)$. These are called the **Chevally Generators** and generate a copy of $\mathfrak{gl}(2)$ inside $\mathfrak{gl}(n)$.

Proposition. If $|i - j| > 1$, then e_j, f_j map (e_i, f_i) strings to strings of the same length or to 0

Proof. This follows from the fact that

$$[e_i, f_j] = [f_i, e_j] = [e_i, e_j] = [f_i, f_j] = 0$$

Thus we have the following diagram where the the solid arrows denote (e_i, f_i) strings and the dotted lines denote the action of e_j and f_j



Now if a terminal vertex is non-zero but it maps above or below to 0 on going right we have a composite map which is zero. But if we first go right and then go up or down we get a composite map which is non zero hence contradiction.

□

For $i, i+1$ we have the following relation called ***Serre's Relation***

$$[e_i, [e_i, e_{i+1}]] = [e_{i+1}, [e_i, e_{i+1}]] = [f_i, [f_i, f_{i+1}]] = [f_{i+1}, [f_i, f_{i+1}]] = 0.$$

We also still have

$$[e_i, f_{i+1}] = [e_{i+1}, f_i] = 0.$$

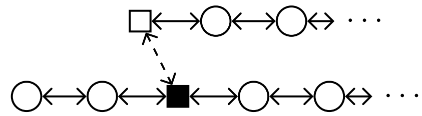
The corresponding result for strings is

Proposition. e_{i+1}, f_{i+1} map length ℓ (e_i, f_i) strings to length $\ell \pm 1$ strings.

Proof. This follows from translating Serre's relation into the commutator for the relation. We work it out for e_i, e_{i+1}

$$r(e_i)^2 r(e_{i+1}) - 2r(e_i)r(e_{i+1})r(e_i) + r(e_{i+1})r(e_i)^2 = 0$$

Indeed, suppose that we had an (e_i, f_i)-string of length ℓ mapping to an (e_i, f_i) string of length $\geq \ell + 2$. We'll draw e_i mapping rightward with solid arrows and e_{i+1} slanting down and left with dashed arrows. Then, at the left hand end of the shorter string, we would have



In the above picture, start at the white square node. We see that $r(e_i)^2 r(e_{i+1})$ is non zero but $r(e_i)r(e_{i+1})r(e_i)$ and $r(e_{i+1})r(e_i)^2$ map to 0 and thus we have a contradiction to Serre's relation. Thus the result follows. $\square$

Remark. This was the argument presented in class, but it now seems easier to start at the black square node and use the relation $[e_i, f_{i+1}] = 0$.

November 16 – High weight vectors

Our goal for this class is to study High Weight Vectors (HWVs) for a representation V of GL_n. From here on we have $\rho : GL_n \to GL(V)$ assigns matrices to elements of the group and $r : \mathfrak{gl}_n \to End(V)$ is the corresponding derivative map for the Lie Algebra. We also have that $V = \bigoplus_{\alpha \in \mathbb{Z}^n} V_\alpha$ is a decomposition of V into weight spaces.

For any element of the torus $t = \begin{bmatrix} t_1 & & & \\ & t_2 & & \\ & & \ddots & \\ & & & t_n \end{bmatrix} \in T$ we have that t acts on V_α by $t_1^{\alpha_1} t_2^{\alpha_2} \cdots t_n^{\alpha_n}$. And for any $i \neq j$, we have $E_{ij} = \begin{bmatrix} & & \cdot & \\ \cdot & \cdot & 1 & \cdot \\ & & \cdot & \\ & & \cdot & \end{bmatrix}$ is the matrix with a single non-zero entry in row i and column j), such that $r(E_{ij}) : V_\alpha \to V_{\alpha + e_i - e_j}$. Note that, by applying $r(E_{ij})$ with $i < j$, the weight of the result keeps increasing in lexicographical order so that eventually we get to a point where $V_\alpha \neq 0$, but $r(E_{ij})(V_\alpha) = 0$.

Definition. We say $v \in V$ is a *high weight vector* (HWV) if $r(E_{ij})(v) = 0$ for all $1 \leq i < j \leq n$.

Examples for GL_3: For each of this examples, $\rightarrow$ represents the E_{12} action, $\leftarrow$ the E_{21} one, $\nearrow$ the E_{23} one, and $\swarrow$ the E_{32} one. The high weight vector is in green.

(1) Representation on $\mathbb{C}^3$.

$$\begin{tikzcd} e_1 & & e_2 \arrow[ll] \\ & e_3 \arrow[ur] & \end{tikzcd}$$

(2) Representation on $\bigwedge^2 \mathbb{C}^3$.

$$\begin{tikzcd} & e_1 \wedge e_2 & \\ e_1 \wedge e_3 \arrow[ur] & & e_2 \wedge e_3 \arrow[ll] \end{tikzcd}$$

(3) Representation defined by $n = 3$, $\lambda = \ydiagram{2,1}$.

$$\begin{tikzcd}[column sep=small] & 2\Delta_{12}\Delta_1 & & 2\Delta_{12}\Delta_2 \arrow[ll] & \\ \Delta_{13}\Delta_1 \arrow[ur] & & (\Delta_{13}\Delta_2 + \Delta_{23}\Delta_1) \arrow[ll] & & \Delta_{23}\Delta_2 \arrow[ll] \\ & & (\Delta_{12}\Delta_3 - \Delta_{13}\Delta_2) \arrow[uur] & & \\ & \Delta_{13}\Delta_3 \arrow[ur] & & \Delta_{23}\Delta_3 \arrow[ll] \arrow[uur] & \end{tikzcd}$$

Remark. The following are equivalent:

(1) v is a high weight vector;

(2) v is fixed by all $\begin{bmatrix} 1 & & & & \\ & 1 & & u & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}$;

(3) v is fixed by $N_+ = \{ \begin{bmatrix} 1 & & * \\ & \ddots & \\ & & 1 \end{bmatrix} \}$ (upper triangular matrices with 1s in the diagonal).

Theorem. Let V_λ be a GL_n-irrep. Then the vector space of high weight vectors for V_λ has dimension 1.

This proof was presented badly. Read the presentation in the following class instead.

November 18 – High weight vectors, and the easy part of the PBW theorem

We'll start with redoing the proof from last class, this time also giving context from general Lie algebras. Here's our setting: we had a GL_n-representation V. Inside GL_n, we have the subgroups N_+ of upper triangular matrices with diagonal entries 1, N_- of lower diagonal matrices with diagonal entries all 1, and the torus subgroup T of (invertible) diagonal matrices. The Lie algebras corresponding to these matrices are: $\mathfrak{gl}$ consisting of all matrices, $\mathfrak{n}_+$ with strictly upper triangular matrices, $\mathfrak{n}_-$ with strictly lower triangular matrices, and $\mathfrak{t}$ of diagonal matrices. We have a representation $\rho : GL_n \to GL(V)$ and a corresponding Lie algebra representation $r : \mathfrak{g} \to End(V)$. V decomposes into a direct sum $\bigoplus_\alpha V_\alpha$, where each V_α is an eigenspace for T and $\mathfrak{t}$. For $e_{ij} \in \mathfrak{n}_+$ with $i < j$, $r(e_{ij})$ maps between weight spaces, with $r(e_{ij}) : V_\alpha \to V_{\alpha+e_i-e_j}$, with $\alpha + e_i - e_j >_{lex} \alpha$. (Thus we called these e_{ij} the 'raising operators'). So there exists $v \in V, v \neq 0$, such that $\mathfrak{n}_+ v = 0$. Equivalently, $\rho(u)v = v$ for all $u \in N_+$.

Definition. $V_{\text{high weight}} = \{v \in V | ev = 0 \text{ for all } e \in \mathfrak{n}_+\}$. This is also $\{v \in V | u \cdot v = v \text{ for all } u \in N_+$.

Our claim was that $V_{\text{high weight}}$ is one dimensional.

Here's general context from Lie algebras: Suppose $\mathfrak{g}$ is a finite dimensional Lie algebra, with $g_1, g_2, \dots, g_m$ forming a basis for $\mathfrak{g}$. Let $r : \mathfrak{g} \to End(V)$ be a Lie algebra representation. This means $r([g, h]) = [r(g), r(h)]$ for all g, h.

Let $v \in V$. The following subspaces of V are equal:

(1) $\text{Span}(r(g_{i_1})r(g_{i_2}) \dots r(g_{i_l})v)$, for $l \geq 0$ and $1 \leq i_1, i_2, \dots, i_l \leq m$.
(2) $\text{Span}(r(g_{i_1})r(g_{i_2}) \dots r(g_{i_l})v)$, for $l \geq 0$ and $1 \leq i_1 \leq i_2 \leq \dots \leq i_l \leq m$.

Proof. The inclusion of the latter subspace into the former is clear. We need the opposite inclusion. We'll show by induction on k, that any $r(g_{i_1})r(g_{i_2}) \dots r(g_{i_l})v$ with $l \leq k$ is in the span of $r(g_{j_i})r(g_{j_2}) \dots r(g_{j_{l'}})v$ for $l' < k$ and $1 \leq j_1 \leq j_2 \dots \leq j_{l'} \leq m$. First note that we can use our induction hypothesis and write $r(g_{i_2})r(g_{i_3}) \dots r(g_{i_l})v$ in the span of $r(g_{j_1})r(g_{j_2}) \dots r(g_{l'})v$ with $j_2 \leq j_3 \dots \leq j_{l'}, l' \leq k$. So it's okay to deal with vectors of the form $r(g_{i_1})r(g_{j_2})r(g_{j_3}) \dots r(g_{j_{l'}})v$ with $j_2 \leq j_3 \leq \dots \leq j_{l'}$. Suppose $j_2 \leq j_3 \leq \dots \leq j_a \leq i_1 \leq j_{a+1} \dots \leq j_{l'}$.

We can write

$$\begin{aligned} r(g_{i_1})r(g_{j_2}) \dots r(g_{j_{l'}})v &= [r(g_{i_1}), r(g_{j_2})]r(g_{j_3}) \dots r(g_{j_{l'}})v \\ &+ r(g_{j_2})[r(g_{i_1}), r(g_{j_3})]r(g_{j_4}) \dots r(g_{j_{l'}})v \\ &+ \dots \\ &+ r(g_{j_2})r(g_{j_3}) \dots [r(g_{i_1}), r(g_{j_a})] \dots v \\ &+ r(g_{j_2})r(g_{j_3}) \dots r(g_{j_a})r(g_{i_1})r(g_{j_{a+1}}) \dots v \end{aligned}$$

Note that $[r(g_{i_1}), r(g_{j_b})] = r([g_{i_1}, g_{j_b}])$. Here, all but the last term is gotten by letting atmost $k - 1$ elements on v, and so by induction these should be in the span of the vector space we're looking at. The last term is one of our spanning elements, since $i_2 \leq i_3 \leq \dots \leq j_a \leq i_1 \leq j_{a+1} \dots$. □

Now we return to $\mathfrak{g} = \mathfrak{gl}_n$. This is spanned by n^2 basis elements $g_1, g_2, \dots g_{n^2}$, and we can pick them in such a way that the first few of them form a basis for $\mathfrak{n}_-$, the next few for $\mathfrak{t}$, and the rest for $\mathfrak{n}_+$. If v is a high weight vector, then $\mathfrak{n}_+ v = 0$, and $\mathfrak{t}v \in \text{Span}(v)$.

By our theorem, if we want to look at the subrepresentation generated by v, we can look at the span of $r(g_{j_1})r(g_{j_2}) \dots r(g_{j_l})v$, with the first few g_{j_k} in $\mathfrak{n}_-$, the next few in $\mathfrak{t}$ and the last few in $\mathfrak{n}_+$. If there are a non-zero number of g_{j_k}'s from $\mathfrak{n}_+$, then the vector becomes 0, and so we can disregard these. The action of the elements from $\mathfrak{t}$ gives us a scalar multiple of v. So we are left with the span of $r(g_{j_1})r(g_{j_2}) \dots r(g_{j_l})v$ with all $j_k \in \mathfrak{n}_-$. So the subspace of V spanned by $r(g_{j_1})r(g_{j_2}) \dots r(g_{j_a})v$, for $j_b \in \mathfrak{n}_-$ forms a $\mathfrak{gl}_n$-subrepresentation of V. So if V

is an irreducible representation, and v a high weight vector, then the $\mathfrak{n}_-$-subrepresentation generated by v is V itself. Let $v \in V_\alpha$. Then, if $V_\beta \neq 0$, we must have $\beta <_{\text{lex}} \alpha$ and $V_\alpha = \mathbb{C} \cdot v$.

Corollary. The high weight space is 1-dimensional

Proof. Suppose $v \in V_\alpha$ and $w \in V_\beta$ are both linearly independent high weight vectors. WLOG assume $\alpha \leq_{\text{lex}} \beta$. $V_\beta \neq 0$ implies $\beta \leq_{\text{lex}} \alpha$ and so we must have $\alpha = \beta$. But then $w \in V_\beta = V_\alpha = \mathbb{C} \cdot v$. This gives a contradiction. $\square$

In terms of constructing V_λ as products of matrix minors, $\Delta_{(\lambda^T)_1} \Delta_{(\lambda^T)_2} \dots \Delta_{(\lambda^T)_l}$ is a high weight vector. Its associated SSYT is: $\begin{array}{|c|c|c|c|} \hline 1 & 1 & \dots & 1 \\ \hline 2 & \dots & 2 \\ \cline{1-3} 3 & \dots \\ \cline{1-2} \end{array} \quad \vdots$

Abstractly, V_λ is characterised as the unique GL_n-irrep whose high weight vector space is in weight λ. Here, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

Recall that $R = \mathbb{C}\begin{bmatrix} x_{11} & \dots & x_{1n} \vdots x_{n1} & \dots & x_{nn} \end{bmatrix} = \bigoplus V_\lambda^\vee \otimes V_\lambda$. We have ${}^{N_+}R = \bigoplus_\lambda {}^{N_+}V_\lambda^\vee \otimes V_\lambda = \bigoplus(\text{ high weight }) \otimes V_\lambda \cong \bigoplus V_\lambda$ as GL_n-reps. So what we are learning is that $R = \mathbb{C}[\Delta_I]$.

Theorem. The ring of functions on $n \times n$ matrices invariant for left N_+ action is $\bigoplus V_\lambda$ as a representation for the right GL_n action, and is generated by the minors Δ_I.

$V_\lambda \otimes V_\lambda^\vee$ is the GL_n subrep generated by $\Delta\left(\begin{array}{|c|c|c|c|} \hline 1 & 1 & \dots & 1 \\ \hline 2 & \dots & 2 \\ \cline{1-3} 3 & \dots \\ \cline{1-2} \end{array}\right)$.

Remark. *Question from the floor: Given the high weight space, is it clear that we could construct the irrep?* If v is a high weight vector, then $V_\lambda = \text{Span}(r(g_{j_1}) \dots r(g_{j_l})v)$ for $g_1, g_2, \dots, g_l \in \mathfrak{n}_-$, $j_1 \leq j_2 \leq \cdots \leq j_l$. We can multiply with an $r(h)$ using the equation:

$$r(h)r(g_{j_1}) \dots r(g_l)v = \sum r(g_{j_1}) \dots [r(g_{j_a}), r(h)] r(g_{j_{a+1}}) \dots \qquad + r(g_{j_1}) \dots r(g_{j_l}) r(h) v.$$

Abstractly, we can obtain V_λ as a quotient of an infinite dimensional vector space, called the Verma module. The λ-Verma module is an infinite dimensional vector space with $\mathfrak{g}$-action, whose basis we think of as $(g_{j_1}, \dots, g_{j_l})$, $j_1 \leq \cdots \leq j_l$ with $j_a \in \mathfrak{n}_-$. V_λ turns out to be its unique finite dimensional $\mathfrak{g}$-invariant quotient.

Remark. So, what is this about the PBW theorem? For any Lie algebra $\mathfrak{g}$, there is a non-commutative ring $U(\mathfrak{g})$, called the universal enveloping algebra of $\mathfrak{g}$, such that every $\mathfrak{g}$-representation is an $U(\mathfrak{g})$-module. In this language, $\text{Span} r(g_{i_1}) r(g_{i_2}) \dots r(g_{i_l})v)$ is the $U(\mathfrak{g})$-submodule generated by v, and each individual product $r(g_{i_1}) r(g_{i_2}) \dots r(g_{i_l})v$ is the product of v with an element $g_{i_1} g_{i_2} \cdots g_{i_l}$ in $U(\mathfrak{g})$. The theorem we have proved is equivalent to showing that the monomials $g_{i_1} g_{i_2} \cdots g_{i_l}$ with $1 \leq i_1 \leq i_2 \leq \cdots \leq i_l \leq m$ span $U(\mathfrak{g})$. The full PBW theorem states that these monomials are a basis of $U(\mathfrak{g})$. The linear independence is significantly harder than the spanning; fortunately, we don't need it.

November 28 – Crystals

The goal of this section is to attempt to combinatorialize GL_n reps as best we can. We will do this in the form of an axiomatization.

Axiomatization of GL_n-action on basis of GL_n-rep. We will require the following:

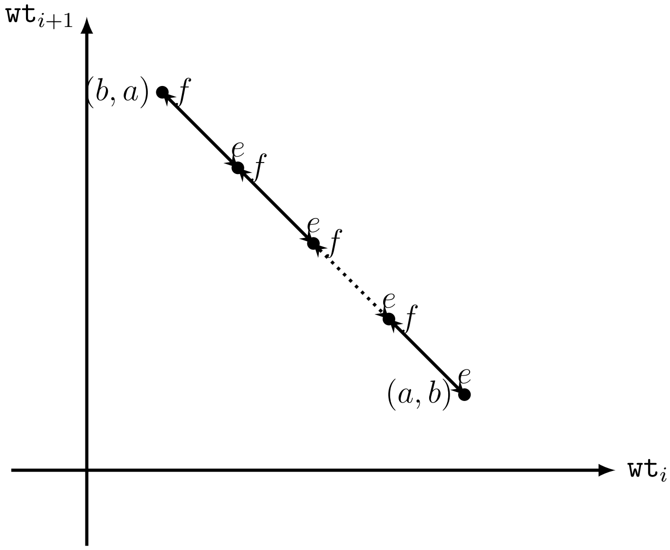
- a finite set set B (think of this set as indexing a basis of a GL_n rep)
- a weight function $\texttt{wt} : B \to \mathbb{Z}^n$.

These two axioms set up the weight spaces.

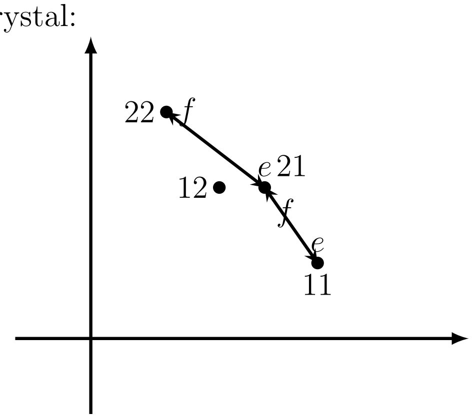
We will further require:

- $\widetilde{B} = B \sqcup \{0\}$
- functions
$$e_1, e_2, \ldots, e_{n-1} : \widetilde{B} \to \widetilde{B}$$
and
$$f_1, \ldots, f_{n-1} : \widetilde{B} \to \widetilde{B}$$
- $e_i(0), f_i(0) = 0$
- If $b \in B$ and $e_i(b) \neq 0$ then
$$\texttt{wt}(e_i(b)) = \texttt{wt}(b) + (0, \ldots, 0, \underset{i}{1}, \underset{i+1}{-1}, 0, \ldots, 0).$$
- If $f_i(b) \neq 0$ then
$$\texttt{wt}(f_i(b)) = \texttt{wt}(b) - (0, \ldots, 0, 1, -1, 0, \ldots, 0).$$

If we just consider e_i, f_i, then we are just looking at the i and $i+1$ component of $\texttt{wt}()$. Then B breaks up into disjoint sets where $(\texttt{wt}(), e_i, f_i)$ looks like a GL_2 string.



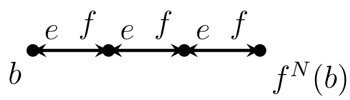
An example is the GL_2 crystal:



Define $s_i = (i\ i+1) \in S_n$. To ensure we get the desired behavior we need to also require the following two axioms:

- (Local) If $b \in B$, $e_i(b) \neq 0$, then $f_i(e_i(b)) = b$ and if $b \in B$, $f_i(b) \neq 0$, then $e_i(f_i(b)) = b$.
- (Global) Let $b \in B$, $e_i(b) = 0$, $f_i^N(b) \neq 0$, $f_i^{N+1}(b) = 0$, then $\mathtt{wt}(f_i^N(b)) = s_i(\mathtt{wt}(b))$.

Note that the Global statement with e_i and f_i switched is equivalent to the one written above. Note also that the local axiom says we must have strings, and the global one ensures they look like



This concludes the axioms of a crystal.

We now conclude with many examples and some non-examples.

Example. Let $B = [n]$, and for $j \in B$, let $\mathtt{wt}(j) = (0, \ldots, \underset{j}{1}, 0, \ldots, 0)$. Then

$$e_i(j) = \begin{cases} j+1 & \text{if } j = i \\ 0 & \text{else} \end{cases} \qquad \text{and} \qquad e_i(j) = \begin{cases} j-1 & \text{if } j = i+1 \\ 0 & \text{else} \end{cases}$$

We get

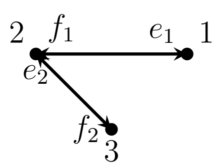


Figure 1. Modeling how GL_3 acts on $\mathbb{C}^3$

This is in fact the only crystal we can put on these weights that satisfies the axioms outlined above. A crystal will be unique whenever the weight function is injective.

Example. Let $B = \{\cdot\}$, and $\mathtt{wt}(\cdot) = (k, k, \ldots, k)$ for some $k \in \mathbb{Z}$. Then $e_i(\cdot) = f_i(\cdot) = 0$ for all i. This models $(\det)^k$.

Example. Let $B = \binom{[n]}{k}$, the set of k-element subsets of $[n]$. For S, a k-element subset of $[n]$, let $\mathtt{wt}(S) = (1, 0, \ldots, 0, 1, 1, \ldots, 0)$ (where we have 1's in the slots for the elements of S). Then

$$e_i(S) = \begin{cases} S - \{i\} \cup \{i+1\} & i \in S, i+1 \notin S \\ 0 & \text{else} \end{cases}$$

and

$$f_i(S) = \begin{cases} S - \{i+1\} \cup \{i\} & i+1 \in S, i \notin S \\ 0 & \text{else} \end{cases}$$

We get

Example. Let $B = \left(\binom{[n]}{k}\right) = \{k - \text{element multi-subsets of } [n]\}$. An example would be

Example. When modeling $(\mathbb{C}^2)^{\otimes 2}$ we start to run into trouble. We now can make two different pictures, which are both acceptable according to the axioms.

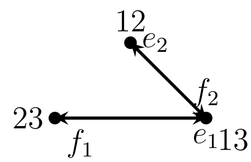


FIGURE 2. Modeling how GL_n acts on $\bigwedge^k \mathbb{C}^n$

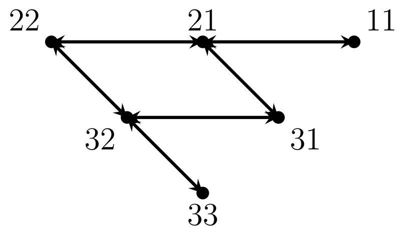
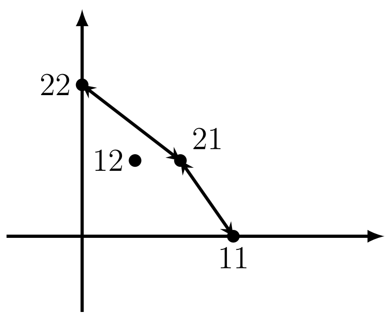
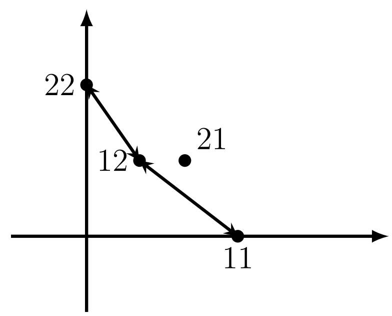


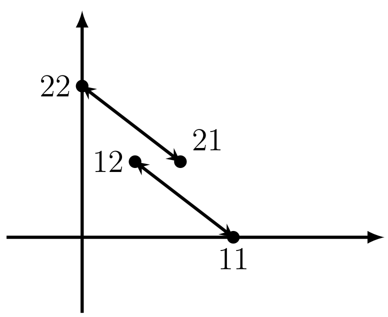
FIGURE 3. Modeling how GL_n acts on $\mathrm{Sym}^k\mathbb{C}^n$



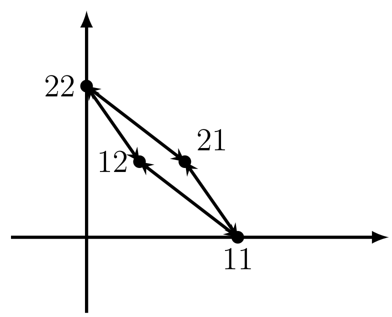
or



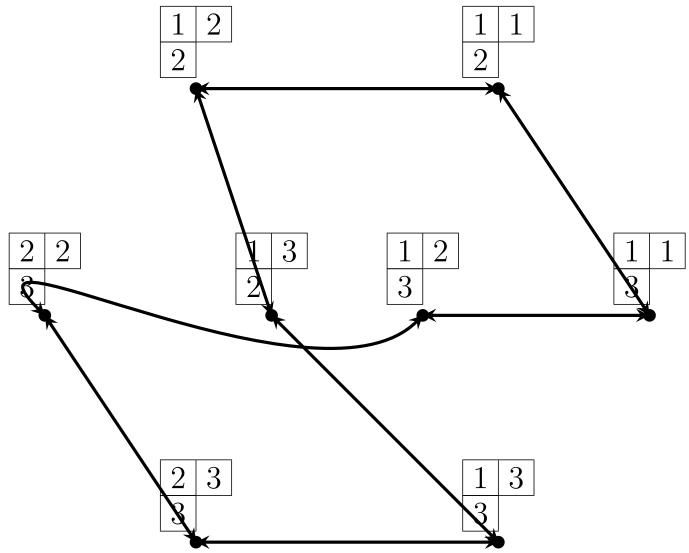
Example. There are a few options that are NOT allowed by the axioms. These violate the global axiom.



or

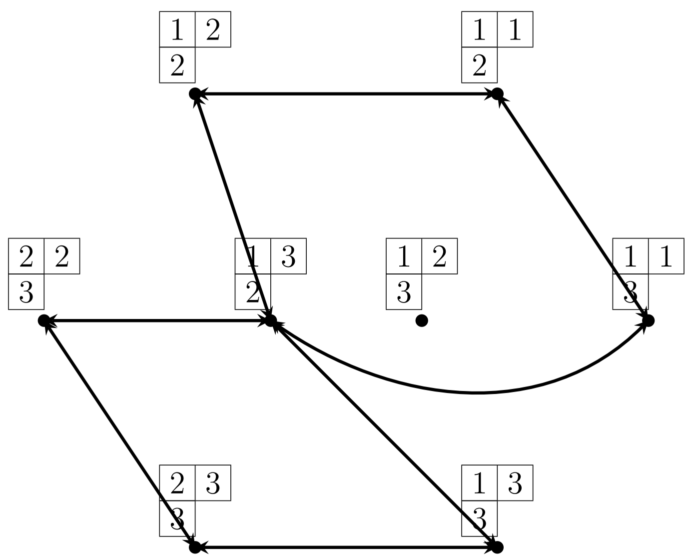


Example. To model $s_{\ydiagram{2,1}}\mathbb{C}^3$ one possibility is



(this picture is in the plane where coordinates add to 3)

Another picture would be the following, but it doesn't match the representation theory.

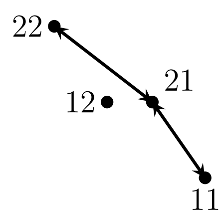


An Overview of What's Coming Next. We are going to build a model of $(\mathbb{C}^n)^{\otimes d}$ called the "word" crystal. Let

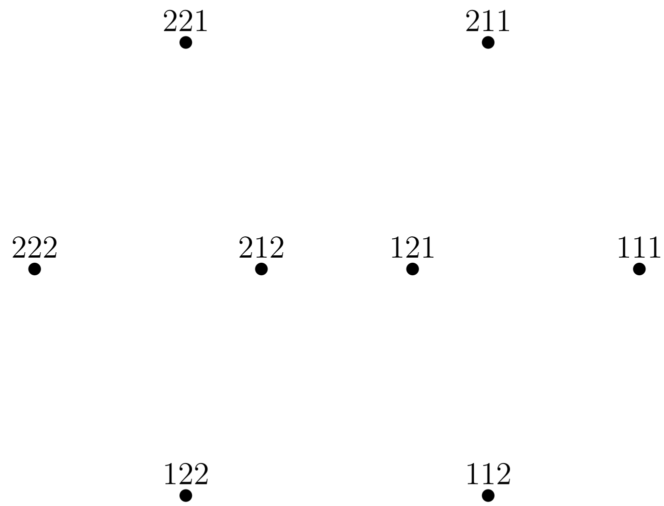
$$B = \{w_1 w_2 \cdots w_d \mid w_1, \ldots, w_d \in [n]\}.$$

Then each e_i will change one w_j from i to $i+1$ (or be 0). Similarly each f_i will change one w_j from $i+1$ to i (or be 0).

An example is below.



For next time, think about where to put arrows in the picture below



November 30 – The word crystal on $[2]^d$

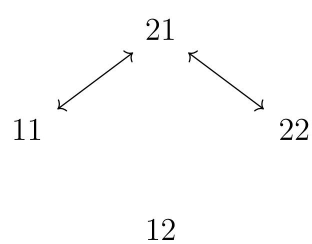
Last time, we defined the axioms for GL_n crystals, and saw several examples. Through these examples, we saw that the string axioms do not give us uniqueness. In particular, given a GL_n-representation V, there are crystals on the basis of V (given by the weights of the representation) that are not the one induced by the representation.

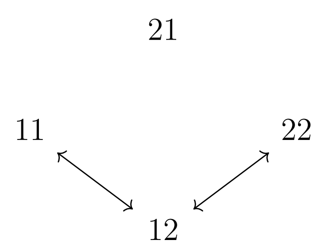
In the next two classe, we will consider what crystal structure we should put on $[n]^d$ in order to mimic $(\mathbb{C}^n)^{\otimes d}$. Our objective today is to study the word crystal on $[n]^d$ when $n = 2$. That is the crystals defined by:

- The set $B = [2]^d$;
- The weight function $wt : B \to \mathbb{Z}^2$ defined by $wt(w) = (\#1\text{'s in } w, \#2\text{'s in } w)$;
- The rising function $e : \tilde{B} \to \tilde{B}$ that changes some occurrence of 2 to 1 in w;
- The lowering function $f : \tilde{B} \to \tilde{B}$ that changes some occurrence of 1 to 2 in w;

Under these definition, there's only one word crystal for $d = 1$: $1 \longleftrightarrow 2$.

For $d = 2$, we have 2 choices:





Where every e, f that not shown in this diagram is just 0, e moves always to the left and f always to the right. For $d \geq 3$, we have many options and it is not clear which we should follow. To narrow the field, let's go back to our motivation, the GL_n-representation of $(\mathbb{C}^n)^{\otimes d}$, and try to find an additional condition from there.

Tensor products of Lie algebra and Lie group representations. Let G be a Lie group and $\rho_1 : G \to GL(V_1), \rho_2 : G \to GL(V_2)$ two G-representations. Ever since problem set 1, we know what the tensor product $\rho_\otimes : G \to GL(V_1 \otimes V_2)$ of ρ_1, ρ_2 looks like: $\rho_\otimes(g)(v_1 \otimes v_2) = (\rho_1(g)v_1) \otimes (\rho(g)v_2)$.

Now, given $\mathfrak{g}$ a Lie algebra and $r_1 : \mathfrak{g} \to \mathfrak{gl}(V_1), r_2 : \mathfrak{g} \to \mathfrak{gl}(V_2)$ representations, what does the tensor product $r_\otimes : \mathfrak{g} \to \mathfrak{gl}(V_1 \otimes V_2)$ looks like? If G is the corresponding Lie group, we know that r_1, r_2 give us G-representations $\rho_j : G \to GL(V_j)$ by: $\rho_j(Id+\varepsilon g) = Id+\varepsilon r_j(g)+\cdots$, where $\cdots$ means higher (≥ 2) order terms in ε. Then we can compute $r_\otimes$ from our earlier definition of $\rho_\otimes$:

$$\begin{aligned}\rho_\otimes(Id + \varepsilon g)(v_1 \otimes v_2) &= (\rho_1(Id + \varepsilon g))(v_1) \otimes (\rho_2(Id + \varepsilon g))(v_2)\\ v_1 \otimes v_2 + \varepsilon r_\otimes(g)(v_1 \otimes v_2) + \cdots &= (v_1 + \varepsilon r_1(g)(v_1) + \cdots) \otimes (v_2 + \varepsilon r_2(g)(v_2) + \cdots)\\ v_1 \otimes v_2 + \varepsilon r_\otimes(g)(v_1 \otimes v_2) + \cdots &= v_1 \otimes v_2 + \varepsilon(r_1(g)(v_1) \otimes v_2 + v_1 \otimes r_2(g)(v_2)) + \cdots\end{aligned}$$

So we deduce that

$$r_\otimes(g)(v_1 \otimes v_2) = r_1(g)(v_1) \otimes v_2 + v_1 \otimes r_2(g)(v_2).$$

We can think of this as a "product rule" for how a Lie algebra acts on a tensor product. We now want to find a variant of this rule for crystals.

The tensor product of two crystals. Given B_1, B_2 two crystals, how should we define their tensor product? There is a standard answer to this, which we'll include at the end of the section but, for now, we are trying to lead up to it gradually.

First off, the underlying set will be $B_\otimes = B_1 \times B_2$, as if B_1, B_2 are bases for V_1, V_2, then $\{b_1 \otimes b_2; b_1 \in B_1, b_2 \in B_2\}$ is a basis for $V_1 \otimes V_2$. From here, we will denote (b_1, b_2) as $b_1 \otimes b_2$. Also from basic linear algebra, we have that $b_1 \otimes 0$ and $0 \otimes b_2$ should both be 0.

The `wt` function comes from the weights of the action of the Torus on $V_1 \otimes V_2$, so $wt(b_1 \otimes b_2) = wt(b_1) + wt(b_2)$.

Finally, we need to define the e and f operators. These come from the Lie algebra action. In the Lie algebra case, we would have $e_j(b_1 \otimes b_2) = e_j(b_1) \otimes b_2 + b_1 \otimes e_j(b_2)$. In the crystal setting, there is no additive structure, so we ask for:

$$e_j(b_1 \otimes b_2) = e_j(b_1) \otimes b_2 \text{ or } b_1 \otimes e_j(b_2) \qquad (*)$$

We impose the same condition on f_j.

The Word Crystal on $B = [2]^d$. Just as $(\mathbb{C}^2)^{\otimes d} \cong (\mathbb{C}^2)^{\otimes j} \otimes (\mathbb{C}^2)^{\otimes d-j}$ for any $1 \leq j \leq d$, i.e. the tensor product of representations is associative, we would like the crystal $[2]^d$ to be $[2]^j \otimes [2]^{d-j}$, that is to say, to obey condition $(*)$ for all $1 \leq j \leq d-1$. This turns out to be extremely restrictive.

For example, let's fix

$$\begin{tikzcd}
& 12 \arrow[dl, <->] \arrow[dr, <->] & \\
11 & & 22 \\
& 21 &
\end{tikzcd}$$

as $[2]^2$. **Warning: The other convention was used in class, but it is going to mesh horribly with RSK, so these notes have been fixed to use the good convention.**

Let's try to find $[2]^3$. Consider $f(111)$. Thinking of $[2]^3$ as $[2]^2 \otimes [2]$, we must have $f(111) \in \{(12)1, (11)2\}$. Thinking of $[2]^3$ as $[2] \otimes [2]^2$, we must have $f(111) \in \{1(12), 2(11)\}$. So we must have $f(111) = 112$.

Continuing in this way, the arrows which are immediately forced are

$$\begin{tikzcd}
& 112 \arrow[r, <->] & 122 & \\
111 \arrow[ur] & 121 \arrow[dr] & 212 \arrow[dl] & 222 \arrow[ul] \\
& 211 & 221 &
\end{tikzcd}$$

But then the reversal of these arrows is also forced and, at that point, we already have enough arrows to fill up the required strings. So the crystal must be

$$\begin{tikzcd}
& 112 \arrow[r, <->] \arrow[dl, <->] & 122 \arrow[dr, <->] & \\
111 & 121 \arrow[dr, <->] & 212 \arrow[dl, <->] & 222 \\
& 211 & 221 &
\end{tikzcd}$$

We will now describe what the final answer will be. We will check that it obeys condition $(*)$. Then, at the end, we will sketch a proof that it is the only thing that obeys $(*)$.

Use the following rule to convert each word into a mountain range. Each 1 will be an upward slope, and each 2 will be a downward slope. That way, the word 11211221211222 will be converted to the image in Figure 4.

Continuing this analogy, we can think both sides of the mountain range are illuminated by a source of light, the sun. This makes some slopes sunlit, and others are shadowed, as seen in the image. Then, we define the maps e, f in the following manner: e changes the leftmost sunlit 2, here marked with green, to a 1 and f turns the rightmost 2, in red, into a 1. If there's no sunlit 1, then f takes the mountain range to 0 and similarly, if there's no sunlit 2, e takes the mountain range to 0.

Notice that if a slope in a word w is shadowed, the same slope in $e(w), f(w)$ is still shadowed, since the change either doesn't affect the valley it is at, or makes the border of the valley even taller. Therefore e and f will only affect the sunlit parts of our mountain range. Notice that we cannot have a sunlit upward slope after a sunlit downward one. Hence

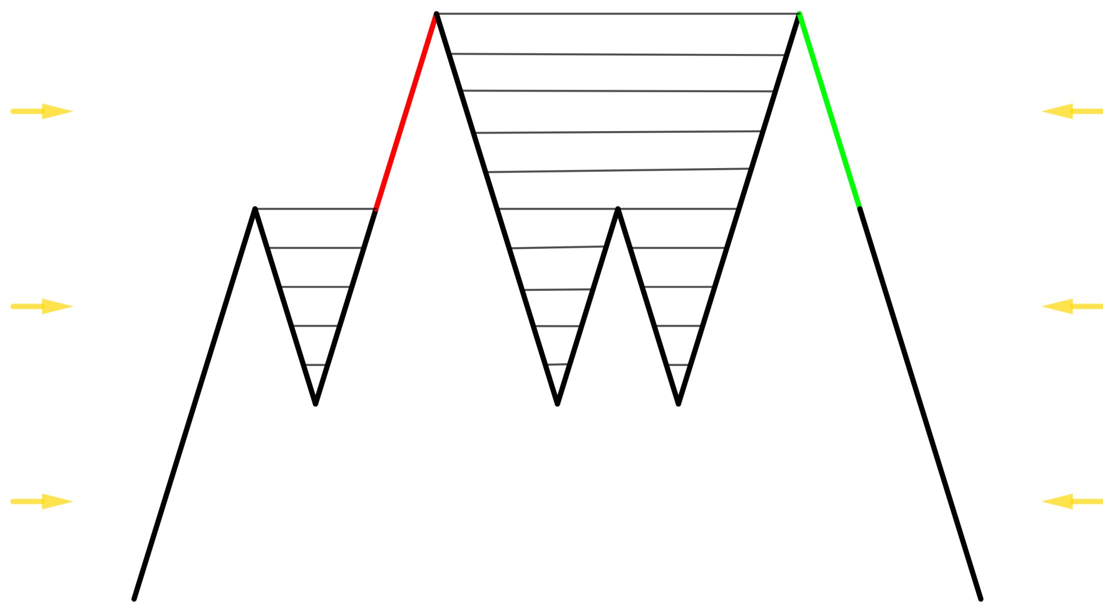


FIGURE 4. Mountain range corresponding to 11211221211222

the sub-word of sunlit slopes is of the form 1^j2^{k-j}. It is also easy to see that the action on that sunlit subword is of the form:

$$1^d \longleftrightarrow 1^{d-1}2 \longleftrightarrow \cdots \longleftrightarrow 12^{d-1} \longleftrightarrow 2^d$$

This gives us the local string property of a crystal. For the global string property, we just need to notice that the sub-word of shadowed slopes has the same amount of 1's and 2's. If $f(w) = 0$ then the sunlit subword is 2^j, and if $e(w') = 0$, the sunlit word is 1^j. So if w and w' are the ends of a chain, then they have the same shadowed area, say, with k 1's and k 2's. Then $\mathtt{wt}(w) = (k, k + j)$ and $\mathtt{wt}(w') = (k + j, k)$. This is the global property.

Finally, let's prove $(*)$. Suppose, w is obtained concatenating w_1 and w_2, that is , $w = w_1w_2$. For simplicity, call the leftmost sunlit 1 of a word v as $\lambda(v)$ and similarly $\rho(v)$ for the rightmost sunlit 2 in v. Then if $\lambda(w_2)$ is taller than $\lambda(w_1)$, it will be $\lambda(w)$. Otherwise, $\lambda(w)$ is $\lambda(w_1)$. Therefore, we either have $f(w) = f(w_1)w_2$ or $f(w) = w_1f(w_2)$. An analogous argument applies for e.

The uniqueness of the mountain rule.

Proposition. Fix $d \geq 3$ and one of the word crystals for $[2]^2$ given above. The mountain range crystal is the unique crystal for $[2]^d$ that satisfies $(*)$ for every split $[2]^j \otimes [2]^{d-j}$ of $[2]^d$.

We have already proved that the mountain range rule has this property. Assume by induction that the mountain range rule is the only rule which works for all $d' < d$.

For each word w of length d, we can carry out the following procedure: For each $1 \leq j \leq d-1$, split w into $w'w''$ where w' is the first j characters and w'' is the remaining $d-j$. Compute $f(w')w''$ and $w'f(w'')$ using the mountain range rule. So, for each j, we get two possible words. There is at least one word which is in every pair, namely $f(w)$ as computed with mountain ranges. If the other word in the pair varies with j, then the only way to define $f(w)$ consistent with $(*)$ is by the mountain range rule. (Of course, we can do the same thing with e.)

We now look closely at the remaining ambiguous cases. We'll use the symbol V to denote a word in which every segment is shadowed (and, thus, the two sides of V are at the same height). The ambiguous cases fall into three classes.

The cases where there are two 1's that could turn into 2's: These look like $1V1$. The mountain range rule says to change the second 1, but $(*)$ is also consistent with changing the first 1.

The cases there the mountain range rule gives 0, but there is also a consistent choice of which 1 to change: These look like $2V1$. The mountain range rule says 0, but changing the final 1 is also consistent with $(*)$.

The cases where the mountain range rule changes a 1 to a 2, but 0 is also consistent: These look like $V12^b$. The mountain range rule says to change the 1 to a 2, but the segment of the cut not containing that 1 always maps to 0.

In particular, note that $2V1$ is **not** ambiguous: We know that $f(2V1) = 2V2$ and $e(2V1) = 1V1$. Thus, by the local string rule, we must have $f(1V1) = 1V2$ and $e(2V2) = 1V2$. This resolves the issue of what to do with the first cases.

Then we can't have $f(2V1) = 2V2$, since we have just shown that $e(2V2) = 1V2$. So we must have $f(2V1) = 0$ instead. In the same way, we must have $e(2V1) = 0$. This resolves the second case.

Finally, note that $V2^{b+1}$ is not ambiguous, we have $e(V2^{b+1}) = V12^b$. So $f(V12^b) = V2^{b+1}$, resolving the third case.

The definition of the tensor product of crystals. As promised above, there is a standard definition of the tensor product of crystals, which we give now. (This was not covered in class, but is natural to put here.)

Let B be a crystal, and let $b \in B$. We define $\epsilon_j(b)$ to be the exponent such that $e_j^{\epsilon_j(b)}(b) \neq 0$ and $e_j^{\epsilon_j(b)+1}(b) = 0$. In other words, ϵ_j is the distance from b to the end of the (e_j, f_j) chain through b. Similarly, let $\phi_j(b)$ be the exponent such that $f_j^{\phi_j(b)}(b) \neq 0$ and $f_j^{\phi_j(b)+1}(b) = 0$. In terms of the mountain range story, $\epsilon(b)$ is the number of sunlit downslopes and $\phi(b)$ is the number of sunlit upslopes.

Let B_1 and B_2 be two crystals. The standard definition of the tensor product of crystals is the crystal structure on $B_1 \times B_2$ given by

$$e(b_1 \otimes b_2) = \begin{cases} e(b_1) \otimes b_2 & \epsilon(b_1) > \phi(b_2) \\ b_1 \otimes e(b_2) & \epsilon(b_1) \leq \phi(b_2) \end{cases} \qquad f(b_1 \otimes b_2) = \begin{cases} f(b_1) \otimes b_2 & \epsilon(b_1) \geq \phi(b_2) \\ b_1 \otimes f(b_2) & \epsilon(b_1) < \phi(b_2) \end{cases}.$$

(Here we have omitted the subscript j on e, f, ϵ and ϕ for readability.)

We can understand this rule in terms of mountain ranges. The sunlit subword portion of the b_1-range is $2^{\phi(b_1)}1^{\epsilon(b_1)}$, and the sunlit portion of the b_2 range is $2^{\phi(b_2)}1^{\epsilon(b_2)}$. The inequality between $\epsilon(b_1)$ and $\phi(b_2)$ controls which of the two peaks is higher when they are shoved together.

So another, to me much less clear, way to motivate the mountain rule is to say that it is $[2]^{\otimes n}$ computed by the above formula.

December 2 – The word crystal for $n > 2$

Define the word crystal $[n]^d$ (sometimes written $[n]^{\otimes d}$), in analogy to the word crystals $[2]^d$ from last time. Recall our vertices are words, our weight function is $\mathtt{wt}(w) = (\#1's \text{ in } w, \#2's \text{ in } w, \ldots, \#n's \text{ in } w)$, and our e_i, f_i action is given by restricting our word to the $(i, i+1)$ characters only (ignoring the rest), and acting as e, f would as though the entries were 1, 2 respectively replacing an $i+1$ with an i or an i with an $i+1$.

An example. The word $w = 4311134$ in $[4]^7$ has $wt(w) = (3, 0, 2, 2)$, the lowering operators give:

- $f_1(w) = 4311234$ as the substring with 1s and 2s is 111,
- $f_2(w) = 0$ as there are no 2's in w,
- $f_3(w) = 4311144$.

Lemma. If $w_i \leq w_{i+1}$ then the same is true for $e_j(w)$ and $f_j(w)$.

Proof. The operators only change a value by 1, so if $w_i < w_{i+1}$ then the claim is immediate.

Suppose $w_i = w_{i+1} = j$. Regardless of the rest of w, e_{j-1} will never decrease w_{i+1} (it is either shadowed, or there is a higher choice in w_i), and likewise f_j will never increase w_i. □

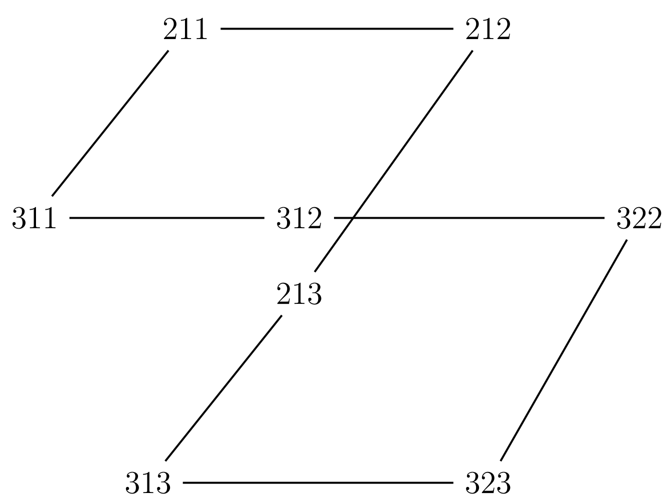
Lemma. If $w_i > w_{i+1}$ then the same is true for $e_j(w)$ and $f_j(w)$.

Proof. If $w_i > w_{i+1} + 1$, then no operator can change either value enough to change the inequality.

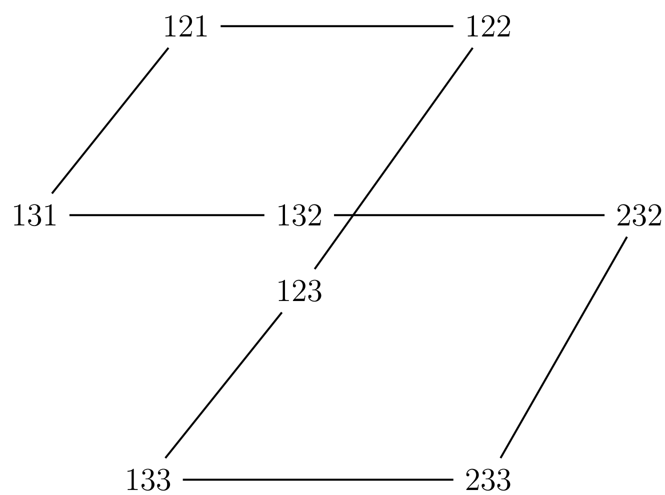
Suppose $w_i = j + 1 > j = w_{i+1}$. Regardless of the rest of w, both w_i and w_{i+1} are in 'shadow' - they are paired with each other, and no operator will change them. □

In other words, the relative order (properly understood) of a word is invariant to e_i, f_i. In the example above $w = 4311134$ has order $>>\leq\leq\leq\leq$, and so do all of its (nonzero) neighbors.

Every word lattice has a component from the $>>> \cdots$ order, corresponding to the $\bigwedge^d \mathbb{C}^n$ representation. Every word lattice has a component from the $\leq\leq\leq \cdots$ order, corresponding to the $\mathrm{Sym}^d\mathbb{C}^n$ representation. The other components may seem a bit familiar. We decompose the word lattice $[3]^3$ as $\bigwedge^3 \mathbb{C}^3, \mathrm{Sym}^3\mathbb{C}^3$, and two more components depicted below:



associated to $(>, \leq)$, and another associated to $(\leq>)$:



These look like the irrep associated to the shape $(2,1)$. Recall that $\mathcal{S}_\lambda(\mathbb{C}^n)$ is the irreducible representation of GL_n associated to the partition λ. Then, indeed, we have:

$$
\begin{aligned}
(\mathbb{C}^n)^{\otimes 3} &\cong \mathrm{Sym}^3\mathbb{C}^n \oplus \bigwedge^3 \mathbb{C}^n \oplus W_{2,1} \otimes \mathbb{S}_{2,1}\mathbb{C}^n \\
&\cong \mathrm{Sym}^3\mathbb{C}^n \oplus \bigwedge^3 \mathbb{C}^n \oplus \mathbb{S}_{2,1}\mathbb{C}^n \oplus \mathbb{S}_{2,1}\mathbb{C}^n,
\end{aligned}
$$

as $W_{2,1} \cong \mathbb{C}^2$.

Suppose $\ell(\lambda) \leq n$ and $|\lambda| = d$. We have a map from $SSYT(\lambda) \to [n]^d$ by reading the entries of the tableau from left to right, starting at the bottom row and going up. Formally, given $T \in SSYT(\lambda)$, define the *reading word* w of T to be $w_1 \cdots w_{\lambda_n} = T_n, w_{\lambda_n+1} \cdots w_{\lambda_n+\lambda_{n-1}} = T_{n-1}$, and so on until $w_{d-\lambda_1+1} \cdots w_d = T_1$, where T_i denotes the ith row of the tableau. For example, the reading word of

1	1	1	3	4
3				
4				

is 4.3.11134 (with periods added to indicate the end of the tableau rows).

Theorem. Let λ be a partition as above. Let $B_\lambda \subset [n]^d$ be the set of reading words of $T \in SSYT(\lambda, [n])$. Then B_λ is a connected component of $[n]^d$.

Proof. We proceed in three parts. First, B_λ is closed under e_i, f_i. We prove the case of e_i, the others are symmetric. Fix a reading word w_T from a tableau T, assume that $e_i(w_T) \neq 0$, and look at the positions of all the i and $j = i + 1$ entries of T, for example:

				i	i	j
		i	i	j	j	
i	j					
j						

observe that we get a sequence of horizontal strips, each of the form $\boxed{i}\boxed{i}\boxed{\cdots}\boxed{i}\boxed{j}\boxed{j}\boxed{\cdots}\boxed{j}$. If we ignore other characters, these are substrings of the reading word of T, which we place from bottom to top. In w_T, these give 'peaks' $i^a(i+1)^b$. Since e_i changes the leftmost unmatched $i+1$, we must change the $i+1$ right after the string of i's in some row. This gives a new tableau word (just by changing this entry from $i+1$ to i in T) unless the entry immediately above it is an i. But every such entry is clearly matched, they pair with the strip of i's immediately above them. So we must get another reading word, as claimed.

Observations about highest weight vectors Now we show that B_λ is a connected component. Exactly as with weight vectors for representations, a highest weight word w is one such that $e_i(w) = 0$ for all i. In analogy with highest weight vector spaces for representaitons, there must be a (unique) highest weight word in each connected component of $[n]^d$. We claim that comes from the tableau:

$$T_{high} = \begin{array}{|c|c|c|c|c|c|c|c|}
\hline
1 & 1 & 1 & 1 & \cdots & 1 & 1 & 1 \\
\hline
2 & 2 & 2 & \cdots & 2 \\
\cline{1-5}
3 & 3 & \cdots & 3 \\
\cline{1-4}
\vdots & & \\
\cline{1-3}
n & \cdots & n \\
\cline{1-3}
\end{array}$$

ie, the tableau with all 1s for the first row, all 2's for the second, and so on. Fix some tableau other than T_{high}, and say that it first differs from T_{high} in the i,jth coordinate (in lexicographic order, so minimize the row then column). Since T_{high} had the minimum possible value in row i, we must have a larger entry k in T. Apply e_{k-1} to the reading word of T, we see that T has at least one unmatched k (the first difference, note the rows above this entry are at most $k-2$), so the result is not 0.

Since there must be some highest weight word in B_λ, and we have eliminated all the others, T_{high} must have a highest weight reading word. $\square$

December 7 – Combinatorial consequences of crystals

Recall that we have defined the $GL_n(\mathbb{C})$-crystal structure on $[n]^d$ and $\mathcal{B}_\lambda \subset [n]^d$ to model the $GL_n(\mathbb{C})$-representations $(\mathbb{C}^n)^{\otimes d}$ and $\mathcal{S}_\lambda(\mathbb{C}^n)$ respectively. We have also shown in the previous lecture that each $\mathcal{B}_\lambda$ is connected with a unique highest weight vector of weight λ. Then, we introduce the notion of the *regular* crystal as follows.

Definition. A crystal $\mathcal{B}$ is ***regular*** if every connected component of $\mathcal{B}$ is isomorphic to $\mathcal{B}_\lambda$ for some weight λ. In this case, we have

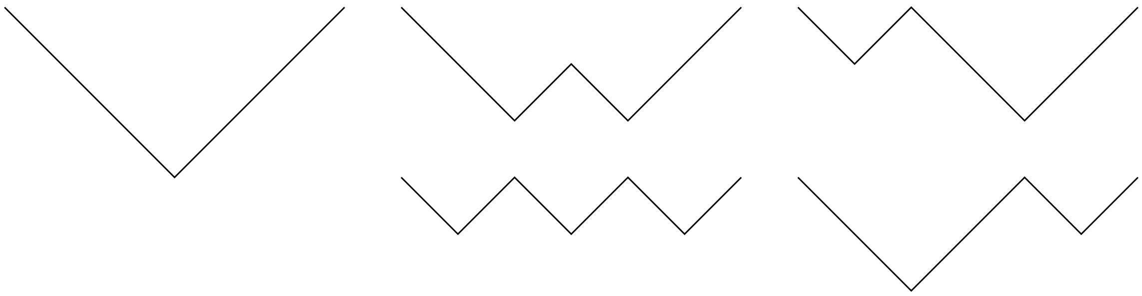
$$\sum_{b\in\mathcal{B}} x^{\mathtt{wt}(b)} = \sum_{\substack{b\in\mathcal{B} \\ b:\text{highest weight}}} s_{\mathtt{wt}(b)}(x).$$

We shall show in the next class that $[n]^d$ is regular. For the purpose of this lecture, let us first take this fact for granted and see what we can get out of it.

The first question for today is to determine whether a word $w = w_1 \cdots w_d$ is of the highest weight. By definition, this is equivalent to requiring $e_{i+1}(w) = 0$ for any $1 \le i \le n-1$. In the mountain range model, this means that all $i+1$ in the $(i,i+1)$-word must be shaded, which is then equivalent to requiring at least as many i's as $(i+1)$'s in every segment $w_j \cdots w_d$. Hence, we make the following definition.

Definition. A word $w = w_1 \cdots w_d$ is ***Yamanouchi*** if for each $1 \le i \le n-1$ and any final segment $w_j \cdots w_d$ of w, there are at least as many i's as $(i+1)$'s.

Example. The multiplicity of $(\det)^{\frac{n}{2}}$ in $(\mathbb{C}^2)^{\otimes n}$ is given by the number of Yamanouchi words with $\frac{n}{2}$ 1's and $\frac{n}{2}$ 2's. For instance, when $n = 6$, there are 5 such words, namely 222111, 221211, 212211, 212121 and 221121. In the mountain range model, these words can be visualized as



Note that these mountain ranges are simply Dyck paths from $(0, 0)$ to $(\frac{n}{2}, \frac{n}{2})$ tilted by 45 degrees. Since it is well-known that the number of Dyck paths are given by the Catalan's numbers, the multiplicity of $(\det)^{\frac{n}{2}}$ in $(\mathbb{C}^2)^{\otimes n}$ is given by

$$C_{\frac{n}{2}} = \frac{2}{n+2}\binom{n}{\frac{n}{2}}.$$

Example. The multiplicity of $(\det)^{\frac{n}{3}}$ in $(\mathbb{C}^3)^{\otimes n}$ is given by the number of Yamanouchi words with $\frac{n}{3}$ 1's, $\frac{n}{3}$ 2's and $\frac{n}{3}$ 3's. By encoding each appearance of 1, 2 and 3 in the word by a unit step in each of the three coordinate directions in a three-dimensional space, we can conclude that the multiplicity of $(\det)^{\frac{n}{3}}$ in $(\mathbb{C}^3)^{\otimes n}$ equals the number of three-dimensional Dyck paths.

Then, we proceed to derive a version of the Littlewood-Richardson rule using the concept of the height weight vectors in a crystal. For any pair of partitions λ and μ, the product of $s_\lambda(x)$ and $s_\mu(x)$ is again a symmetric function, so there exists coefficients $c^\nu_{\lambda\mu}$ such that

$$s_\lambda(x) \cdot s_\mu(x) = \sum_\nu c^\nu_{\lambda\mu} \cdot s_\nu(x).$$

These coefficients $c^\nu_{\lambda\mu}$ are called the ***Littlewood-Richardson coefficients***, and any combinatorial rules to determine them is known as a ***Littlewood-Richardson Rule***. We also define $\lambda * \mu$ to be the Young diagram formed by attaching the Young diagrams of λ and μ on the bottom and the right of a rectangular grid of dimension $\ell(\lambda^T)$ by $\ell(\mu)$. It is clear that $\mathrm{SSYT}(\lambda * \mu)$ corresponds bijectively to $\mathrm{SSYT}(\lambda) \times \mathrm{SSYT}(\mu)$.

Now, let us look for the condition for which an SSYT of shape $\lambda * \mu$ is of highest weight. Note that a word in $[n]^{|\lambda|+|\mu|}$ is Yamanouchi if and only if the corresponding SSYT U of shape μ is super-standard (i.e. having all i's on the i-th row) and in every final segment of the reading word of U, we have

$$(\text{number of } j\text{'s in } T) + \mu_j \geq (\text{number of } (j+1)\text{'s in } T) + \mu_{j+1}.$$

If $T \in \mathrm{SSYT}(\lambda)$ satisfies this condition, it is said to be ***μ-dominated***. Then, we have

$$c^\nu_{\lambda\mu} = |\{\mu\text{-dominated SSYT of shape } \lambda \text{ and content } \nu - \mu\}|.$$

Then, we shall obtain another version of the Littlewood-Richardson rule by assuming that the equality

$$s_\nu(x_1, \cdots, x_k, y_1, \cdots, y_{n-k}) = \sum_{\lambda,\mu} c^\nu_{\lambda\mu} \cdot s_\lambda(x) s_\mu(y).$$

holds. Note that for other Lie types in general, the coefficients in the above expression do not match with the Littlewood-Richardson coefficients.

To show this, we take the crystal $\mathcal{B}_\nu$ and consider the $GL_k \times GL_{n-k}$-crystal structure on $\mathrm{SSYT}(\nu, [n])$ defined by the actions of $(e_1, \cdots, e_{k-1}, e_{k+1}, \cdots, e_n)$ and $(f_1, \cdots, f_{k-1}, f_{k+1}, \cdots, f_n)$. Note that the (e, f)-actions for GL_k commute with those actions for GL_{n-k} and do not change the weights on GL_{n-k}, and vice versa. Thus, we have the isomorphism

$$\mathcal{B}_\nu([n]) \cong \bigsqcup_\lambda \mathcal{B}_\lambda([k]) \times \mathcal{B}_{\nu/\lambda}([k+1, n]).$$

An SSYT of shape ν corresponds to a highest weight vector for $GL_k \times GL_{n-k}$ if and only if the sub-tableau of shape λ is super-standard, and the remainder of skew shape ν/λ has Yamanouchi reading word. Hence, we have

$$c^\nu_{\lambda\mu} = |\{\text{SSYT of shape } \nu/\lambda \text{ and content } \mu\}|\,.$$

Finally, recall the bijection as sets

$$M_{d\times n}(\mathbb{Z}_{\geq 0}) \longleftrightarrow \bigsqcup_\lambda \mathrm{SSYT}(\lambda, [n]) \times \mathrm{SSYT}(\lambda, [d])$$

with the weight function $\mathtt{wt}(A) := (\text{row sums of } A, \text{column sums of } A)$ corresponding to the weight function $\mathtt{wt}(T, U) := (\text{entries of } T, \text{entries of } U)$. Our goal is to build commuting $GL_d \times GL_n$-crystal structures on $M_{d\times n}(\mathbb{Z}_{\geq 0})$ to get an isomorphism as crystals.

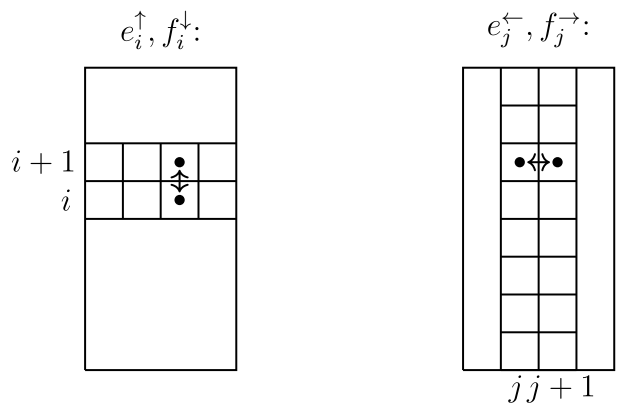
We shall see how this is done in the next class, but let us observe that the implications if this is done are

(1) every $GL_d \times GL_n$ connected component is isomorphic to $\mathcal{B}_1 \times \mathcal{B}_2$ for some GL_d-crystal $\mathcal{B}_1$ and GL_n-crystal $\mathcal{B}_2$, and
(2) every $GL_d \times GL_n$ connected component contains a $GL_d \times GL_n$ highest weight vector.

December 9 – Crystals and RSK

Today we will discuss how we can see RSK as an isomorphism of crystals. This was discovered independently by Danilov and Koshevoy, "Arrays and the combinatorics of Young tableaux" and van Leeuwen, "Double crystals of binary and integral matrices". Today's presentation blends both perspectives.

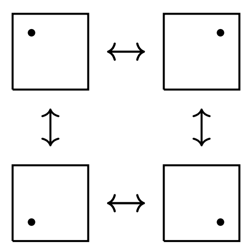
We are going to make a $\mathrm{GL}_d \times \mathrm{GL}_n$ crystal on the set of $\mathbb{Z}_{\geq 0}$-valued $d \times n$ matrices which mimics the action of $\mathrm{GL}_d \times \mathrm{GL}_n$ on $\mathbb{C}[x_{ij}]_{1\leq i\leq d, 1\leq j\leq n}$. Given a matrix, the $\mathbb{Z}^d$ weight is the row sum, and the $\mathbb{Z}^n$ weight is the column sum. We will define operators $e_i^\uparrow$ and $f_i^\downarrow$ for GL_d and operators $e_j^\leftarrow$ and $f_j^\rightarrow$ for GL_n. In order to do this, we will think about the matrices as piles of chips on a rectangular $d \times n$ board where the number of chips on a particular spot is given by the number in the corresponding entry of the matrix.



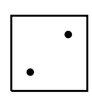
Each operator moves one chip. In order to decide which chip to move, we embed into the word crystal. For $e_j^{\leftarrow}$ and $f_j^{\rightarrow}$, we read like English, left to right across the first row, then left to right across the second row, etc. For $e_i^{\uparrow}$ and $f_i^{\downarrow}$, we read first down the first (leftmost) column, then down the second column, etc.

It is clear that this gives a GL_d crystal and a GL_n crystal, that the crystal operators slide within rows and columns, and that $e^{\leftarrow}$ and $f^{\rightarrow}$ preserve row sums, while $e^{\uparrow}$ and $f^{\downarrow}$ preserve column sums. What is not clear is that $(e_j^{\leftarrow}, f_j^{\rightarrow})$ commutes with $(e_i^{\uparrow}, f_i^{\downarrow})$. We will postpone this proof. For now, we consider some examples.

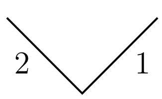
When we have just one chip, the crystal is isomorphic to $B_{\square}(d)\times B_{\square}(n)$. When $d=n=2$, we have the following picture.



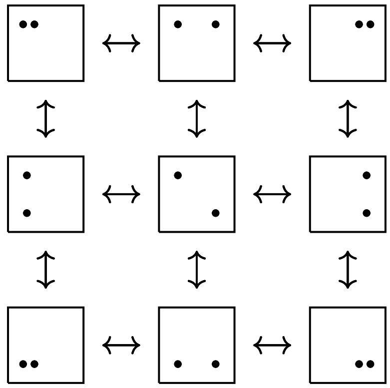
When $d=n=2$ and there are two chips, there is one connected component consisting of just one matrix which looks like



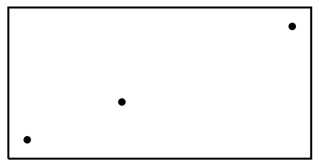
This component is isomorphic to $B_{\substack{\square\\\square}} \times B_{\substack{\square\\\square}}(2)$ The row column weight are both $2,1$ and both correspond to valleys



when considering the corresponding mountain range, but the operators change peaks. The other component is isomorphic to $B_{\square\square}(2) \times B_{\square\square}(2)$ and looks like

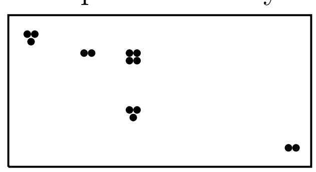


If we consider the set of configurations with k chips on a strictly NE-SW line, i.e. configurations that look like



then this forms a connected component isomorphic to $B_{\begin{smallmatrix}\square\\\square\\\square\end{smallmatrix}}(d) \times B_{\begin{smallmatrix}\square\\\square\\\square\end{smallmatrix}}(n)$ (with k boxes in the vertical column).

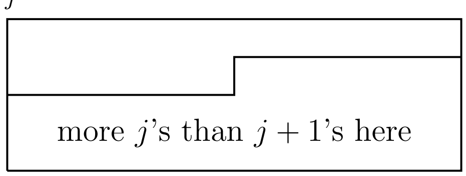
If we instead consider the set of m chips on a weakly NW-SE line, i.e. those that look like



we get a connected component that is isomorphic to $B_{\square\square\square}(d) \times B_{\square\square\square}(n)$ (with k boxes in the horizontal row).

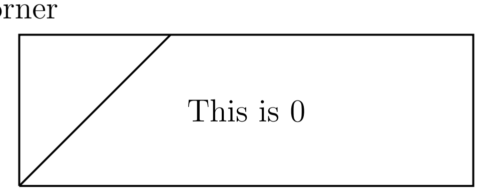
For now, let us assume that the left and right operators $\leftrightarrow$ commute with the up and down operators $\updownarrow$. Then every $(\leftrightarrow, \updownarrow)$ connected component is $C_{\updownarrow} \times C_{\leftrightarrow}$ for some $\updownarrow$ crystal and some $\leftrightarrow$ crystal. Every $(\leftrightarrow, \updownarrow)$ connected component of the $\mathrm{GL}_d \times \mathrm{GL}_n$ crystal has an element which is sent to 0 by $f_i^{\downarrow}$ and $e_j^{\leftarrow}$.

If an element is killed by $e_j^{\leftarrow}$ means that

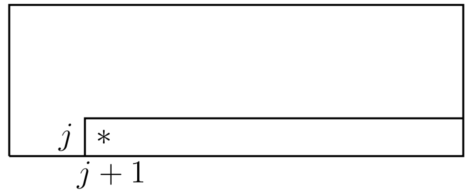


$j \;\; j+1$

A consequence of this is that everything must be zero to the left of the main diagonal that begins in the bottom left corner

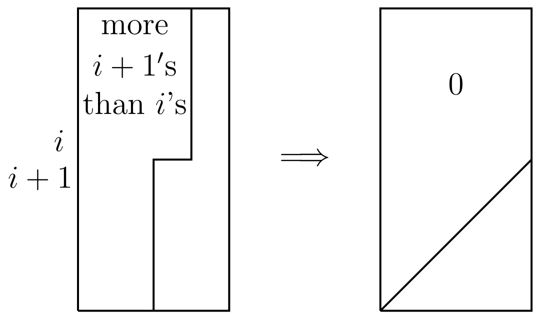


This can be proven by induction. Beginning in the bottom row, if we had

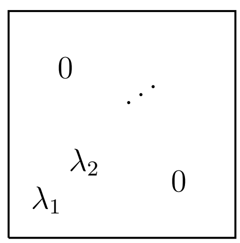


Then there could be no mater j's to the right of that last entry, so the bottom row must be 0 besides potentially the leftmost entry.

Similarly, if an element is killed by $f_i^{\downarrow}$, that means that

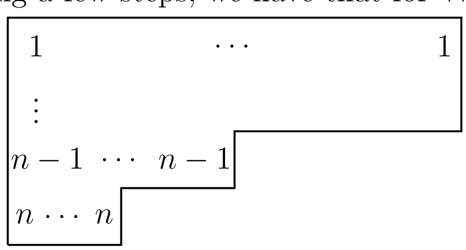


That is, everything above the main diagonal beginning at the bottom left corner is 0. Therefore, elements which are killed both by $e_j^{\leftarrow}$ and $f_i^{\downarrow}$ are 0 except on this diagonal and has the form



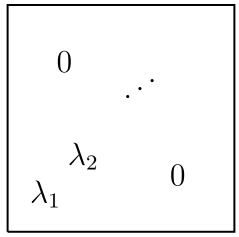
The horizontal word for $\leftrightarrow$ is then $k^{\lambda_k}(k-1)^{\lambda_{k-1}}\dots 1^{\lambda_1}$ and the vertical word for $\updownarrow$ is $d^{\lambda_1}(d-1)^{\lambda_2}\dots 1^{\lambda_k}$ where $k = \min(d, n)$.

We want to show that an element is killed by $e_i^{\leftarrow}$ and $f_j^{\downarrow}$ if and only if given $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$ from the matrix as above is a partition. We will get one connected component per partition. Skipping a few steps, we have that for $\leftrightarrow$ we get



which is $B_\lambda(n) \times B_\lambda(d)$ (where the above diagram has shape λ).

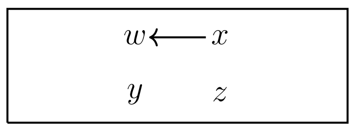
A consequence is that $[n]^d$ is regular as a GL_n-crystal. We can embed this is matrices with one chip per row where the chip in the i-th row occurs in the column j where j is the i-th entry in the corresponding element of $[n]^d$. We must have that this set is isomorphic to a horizontal component of



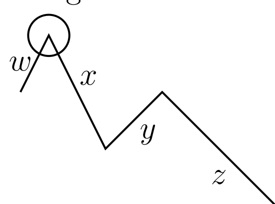
When we do a $\updownarrow$ crystal move in rows i and $i+1$ we look just at those rows. Some chips are sunlit, shown in yellow. Others are shadowed, shown in red, and are scattered around.

The vertical operators either take the rightmost sunlit chip of the top row and slide it down or the leftmost sunlit chip in the bottom row and slide it up.

We claim that the horizontal operators don't change which chips are sunlit. Let us consider what happens when we use the $\leftrightarrow$ operators in these rows. If we have

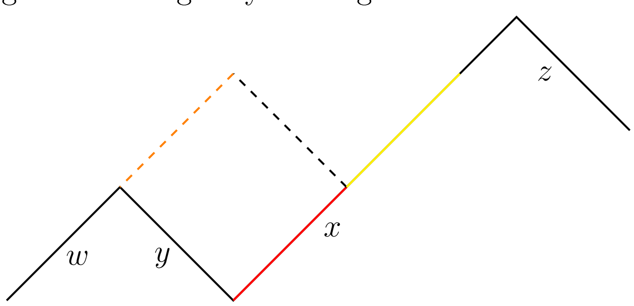


then the corresponding $\leftrightarrow$ mountain range looks like



where the operator acts on the circled peak. We see then that we must have $x > y$.

The $\updownarrow$ mountain range which we get by reading down columns looks like



where the solid lines represent the mountain range before applying $\leftarrow$ and the dotted lines represent what the range looks like after. We see that there must have been a red portion in shadow. When the operator is applied, a corresponding new segment of the w portion of the range is made and is sunlit. This is the dotted orange portion in the above diagram. The same amount of a previously sunlit portion of the x part of the mountain range, colored in yellow, is made shadowed by this. Therefore, the total amount of parts of the x and w portion of the mountain range that are sunlit is preserved, and we get commutation.