THE WEYL CHARACTER FORMULA FOR $U(n)$, AND GEL'FAND-CETLIN PATTERNS NOTES FOR MATH 261, FALL 2001

ALLEN KNUTSON

1. The Weyl character formula

Let $\Delta = \{x_i - x_j\}_{i \neq j}$ denote the **roots of** $GL_n(\mathbb{C})$, which we defined as the nonzero weights of the **adjoint representation** of $U(n)$ on its complexified Lie algebra, $\mathfrak{gl}_n(\mathbb{C})$. Split this into $\Delta_+ = \{x_i - x_j\}_{i<j}$ and $\Delta_- = -\Delta_+$.

Define the **Weyl denominator** as

$$Den(t) = \prod_{\alpha \in \Delta_-} (1 - t^\alpha) = \prod_{i>j} (1 - t_i t_j^{-1}), \quad t \in T$$

a polynomial function on the torus T. Calling it "denominator" suggests that we'll want to invert it; expanding each $(1 - t^\alpha)^{-1}$ as a power series $1 + t^\alpha + t^{2\alpha} + \ldots$, the coefficient of t^μ in $1/Den$ is the Kostant partition function of $-\mu$.

Careful: expanding in a power series is not unique, e.g. $\frac{1}{1-z} = \frac{-z^{-1}}{1-z^{-1}} = -z^{-1} - z^{-2} - \ldots$, and this can lead to all sorts of problems if not done carefully. When we say "expand Den in a power series" in this episode we will always mean in the way above.

Since every element of $U(n)$ is conjugate to an element of T (usually $n!$ of them), we should be able to reduce integrals of class functions on $U(n)$ to integrals over T, by weighting the points in T by the volume of the conjugacy class in $U(n)$. The conjugacy classes with repeated eigenvalues are lower dimensional than the generic dimension, so this weighting should vanish along any $t_i = t_j$. That makes Den sounds like a good candidate. But the weighting should also be S_n-invariant, and Den isn't, but $Den\overline{Den}$ is.

Theorem (Weyl integration formula)**.** *Let $f : U(n) \to \mathbb{C}$ be a smooth class function on $U(n)$, and $d\mu_{U(n)}, d\mu_T$ invariant measures on $U(n)$, T of total mass 1. Then*

$$\int_{U(n)} f \, d\mu_{U(n)} = \frac{1}{n!} \int_T f|_T \, Den \, \overline{Den} \, d\mu_T.$$

Proof. Let $\Phi : (U(n)/T) \times T \to U(n)$ take $(gT, t) \mapsto gtg^{-1}$. For each element of $U(n)$ with distinct eigenvalues, which is the generic case, this map is $n!$-to-one. Therefore

$$\int_{U(n)} f \, d\mu_{U(n)} = \frac{1}{n!} \int_{(U(n)/T) \times T} \Phi^*(f) \Phi^*(\, d\mu_{U(n)})$$

The first term, $\Phi^*(f)$, is easy to calculate because f is a class function – $\Phi^*(f)(gT, t) = f(\Phi(gT, t)) = f(gtg^{-1}) = f(t)$. To calculate the second, we need to compute the determinant of the derivative of Φ. (The derivative is easy, the determinant is trickier.)

To calculate the derivative, we wiggle the input a little bit, using ϵ with $\epsilon^2 = 0$.

$$\Phi(g(1+\epsilon X)T, t) = g(1+\epsilon X)t(1-\epsilon X)g^{-1} = g(1+\epsilon(X - tXt^{-1}))tg^{-1}$$
$$\Phi(gT, t(1+\epsilon Y)) = gt(1+\epsilon Y)g^{-1}$$

Now we need to be specific about how we identify the tangent spaces to the manifolds. Identify

- the tangent space to $gT \in U(n)/T$ with zero-diagonal skew-Hermitian matrices via $X \mapsto g(1+\epsilon X)T$
- the tangent space to $t \in T$ with diagonal imaginaries via $Y \mapsto t(1+\epsilon Y)$
- the tangent space to $g \in U(n)$ with skew-Hermitians via $X \mapsto g(1+\epsilon X)$

In particular, since the skew-Hermitians are the sum of the diagonal imaginaries and the zero-diagonal skew-Hermitians, this lets us think of the above derivative as a square matrix,

$$d\Phi = \begin{bmatrix} X \mapsto X - tXt^{-1} & \\ & Y \mapsto Y \end{bmatrix}$$

We can diagonalize this matrix, if we complexify from zero-diagonal skew-Hermitians to zero-diagonal arbitrary matrices, and use the basis of root spaces $\mathbb{C}e_{ij}$. (Careful: i and j run from 1 through n, but this matrix is much bigger – its rows are indexed by pairs (i,j).) So

$$\det d\Phi = \prod_{i\neq j}(1 - t_i t_j^{-1}) \prod_{i=j} 1$$

and this is $\mathrm{Den}\overline{\mathrm{Den}}$. □

We implicitly used the rational function t^λ/Den in the N-orbits episode; we proved that the weight multiplicities in the irrep V_λ were bounded by the coefficients of the corresponding monomials in the power series t^λ/Den.

Since S_n acts on T, it acts on rational functions on T. Define the rational function

$$WCF_\lambda = \sum_{w\in S_n} w \cdot \frac{t^\lambda}{\mathrm{Den}} = \sum_{w \in S_n} \frac{t^{w\cdot\lambda}}{\prod_{i>j}(1 - t_{w(i)}t_{w(j)}^{-1})}$$

This version of the function is obviously S_n-invariant.

Proposition. *Let $\rho = (n-1, n-2, \ldots, 1, 0)$, and $(-1)^w$ be 1 for even permutations, -1 for odd.*

1. $$\mathrm{Den} \cdot WCF_\lambda = \sum_{w\in S_n} (-1)^w \, t^{w\cdot(\lambda+\rho)-\rho}.$$
2. *The rational function WCF_λ is actually a Laurent polynomial...*
3. *...and it has integer coefficients.*

Proof. We need to rewrite our denominators:

$$\prod_{i>j}(1 - t_{w(i)}t_{w(j)}^{-1}) = \prod_{i>j, w(i)>w(j)} (1 - t_{w(i)}t_{w(j)}^{-1}) \prod_{i>j, w(i)<w(j)} (1 - t_{w(i)}t_{w(j)}^{-1})$$

$$= \prod_{i>j, w(i)>w(j)} (1 - t_{w(i)}t_{w(j)}^{-1}) \prod_{i>j, w(i)<w(j)} (-t_{w(i)}t_{w(j)}^{-1})(1 - t_{w(j)}t_{w(i)}^{-1}) = \prod_{i>j, w(i)<w(j)} (-t_{w(i)}t_{w(j)}^{-1}) \prod_{i>j}(1 - t_i t_j^{-1})$$

the last equality basically because w permutes the set $\binom{n}{2}$. Then

$$\mathrm{Den} \cdot WCF_\lambda = \sum_{w \in S_n} t^{w \cdot \lambda} \frac{\prod_{i>j}(1 - t_i t_j^{-1})}{\prod_{i>j}(1 - t_{w(i)} t_{w(j)}^{-1})} = \sum_{w \in S_n} t^{w \cdot \lambda} \prod_{i>j, w(i)<w(j)} (-t_{w(i)}^{-1} t_{w(j)})$$

so it remains to understand this factor. Obviously it kicks out the right sign, $(-1)^w$. What remains is

$$\prod_{i>j, w(i)<w(j)} t_{w(i)}^{-1} t_{w(j)} = \prod_{k<l, w^{-1}(k)>w^{-1}(l)} t_k^{-1} t_l$$

so the exponent on t_m in this product is

$$\begin{aligned}
&\#\{i < m : w(i) > w(m)\} - \#\{i > m : w(i) < w(m)\} \\
&= \#\{i \leq m : w(i) > w(m)\} - \#\{i \geq m : w(i) < w(m)\} \\
&= m - \#\{i \leq m : w(i) \leq w(m)\} - \#\{i \geq m : w(i) < w(m)\} \\
&= m - w(m)
\end{aligned}$$

which is indeed the mth entry in $w \cdot \rho - \rho$.

For part 2, we just need to check that where the function seems to blow up (on $(\mathbb{C}^\times)^n$) it doesn't actually. This is where some t_i, t_j approach one another. It's enough to consider the case $t_1 \sim t_2$ (all others distinct) because of the S_n-invariance.

$$\lim_{t_1 - t_2 \to 0} WCF_\lambda \sim (1 - t_2 t_1^{-1})^{-1} \sum_w (-1)^w t^{w \cdot (\lambda+\rho) - \rho} \quad \text{(i.e. up to a factor that isn't blowing up)}$$

$$= \sum_{w, w(1)<w(2)} (-1)^w \frac{t^{w \cdot (\lambda+\rho)-\rho} - t^{ws_1 \cdot (\lambda+\rho)-\rho}}{1 - t_2 t_1^{-1}}$$

Each of these terms looks like $(t_1^a t_2^b - t_1^b t_2^a)/(1 - t_2 t_1^{-1})$, and that's regular at $t_1 = t_2$.

Finally, to see the integer coefficients just notice that in the power series expansion of Den we get integer coefficients. Then WCF_λ is that power series times a polynomial with ± 1 coefficients (as we saw above). So it's a power series (actually a polynomial) with integer coefficients. (It's totally nonobvious from here that it's got *natural* coefficients, i.e. nonnegative.) ☐

Theorem (Weyl character formula). *If V_λ is an irrep with high weight λ, then* $\mathrm{Tr}\,(t|_{V_\lambda}) = WCF_\lambda(t)$.

Proof. So far we know that $WCF_\lambda(t)$ is an S_n-invariant Laurent polynomial with integer coefficients. We already showed that the characters of irreps give a $\mathbb{Z}$-basis for the space of all such. Therefore $WCF_\lambda = \sum_\mu m_\mu \mathrm{Tr}(\cdot|_{V_\mu})$, for some integers m_μ (only finitely many terms used).

Even better, they were an *orthonormal* basis, with respect to the pairing $\int_{U(n)} \phi\bar{\psi}\, d\mu$. So $|WCF_\lambda|^2 = \sum_\mu m_\mu^2$.

Let's compute that, using the Weyl integration formula:

$$\int_{U(n)} WCF_\lambda \overline{WCF_\lambda}\, d\mu_{U(n)} = \frac{1}{n!} \int_T WCF_\lambda \overline{WCF_\lambda}\, \mathrm{Den}\, \overline{\mathrm{Den}}\, d\mu_T$$

$$= \frac{1}{n!} \int_T \Big(\sum_{w \in S_n} (-1)^w t^{w \cdot (\lambda+\rho)-\rho} \Big) \Big(\sum_{v \in S_n} (-1)^v \overline{t^{v \cdot (\lambda+\rho)-\rho}} \Big)\, d\mu_T$$

$$= \frac{1}{n!} \sum_{w,v \in S_n} (-1)^{wv} \int_T t^{w\cdot(\lambda+\rho)} \overline{t^{v\cdot(\lambda+\rho)}}\, d\mu_T = \frac{1}{n!} \sum_{w,v \in S_n} (-1)^{wv} \delta_{w\cdot(\lambda+\rho), v\cdot(\lambda+\rho)}$$

this last because the characters t^β of irreps of T are orthonormal (you may be more familiar with this from thinking about Fourier series). But since λ weakly decreasing and ρ strictly decreasing, $\lambda + \rho$ is strictly decreasing, and two different permutations of it are automatically different.

$$= \frac{1}{n!} \sum_{w,v \in S_n} (-1)^{wv} \delta_{w,v} = \frac{1}{n!} \sum_{w \in S_n} 1 = 1.$$

So $|WCF_\lambda|^2 = 1$, so $\sum_\mu m_\mu^2 = 1$, which can only happen if $WCF_\lambda = \pm \mathrm{Tr}\, (\cdot|_{V_\nu})$ for a particular ν. Since we have the description of WCF_λ as a power series strongly dominated by 0 times a polynomial strongly dominated by λ, we know WCF_λ is strongly dominated by λ, and the coefficient of t^λ is 1 (i.e. positive), so this $\nu = \lambda$. □

2. Simple corollaries

We get a couple immediate corollaries, using our formula for $\mathrm{Den} \cdot WCF_\lambda$.

Theorem (Kostant multiplicity formula)**.** *Let λ be a dominant weight, μ a weight. Then the multiplicity of the weight μ in the irrep V_λ dominated by λ is*

$$\sum_{w \in S_n} (-1)^w K(w \cdot (\lambda + \rho) - \rho - \mu).$$

Proof. This is because the Kostant partition function $K(-\cdot)$ computes the coefficients in the power series $1/\mathrm{Den}$. □

We knew already abstractly that given the character of a representation V, one can figure out isomorphism type, or equivalently, the decomposition into irreducibles. We can now give a formula for this:

Theorem. *Let V be a representation of $U(n)$, $\chi : T \to \mathbb{C}$ its character. Then the multiplicity of the irrep V_λ in V (λ a dominant weight) is the coefficient of t^λ in $\mathrm{Den} \cdot \chi$.*

Proof. It's enough to check for irreducibles, because both the question and answer are linear in χ. For irreducibles it follows from our formula for $\mathrm{Den} \cdot WCF_\lambda$, because that Laurent polynomial has only one term with a dominant exponent, namely $1 \cdot t^\lambda$. □

Theorem (Steinberg tensor product formula)**.** *Let λ, μ, ν be dominant weights. Then when we decompose $V_\lambda \otimes V_\mu$ into irreducibles, the number of components isomorphic to V_ν is*

$$\sum_{w \in S_n} (-1)^w (\textit{the multiplicity of the weight } \nu - w \cdot (\lambda + \rho) + \rho \textit{ in } V_\mu)$$

Proof. We're computing the t^ν term in $\mathrm{Den} \cdot WCF_\lambda \cdot WCF_\mu$, which is

$$\mathrm{Den} \cdot WCF_\lambda \cdot WCF_\mu = \sum_{w \in S_n} (-1)^w\, t^{w\cdot(\lambda+\rho)-\rho}\, WCF_\mu.$$

So for each w, that's the coefficient on the t^ν term in $t^{w\cdot(\lambda+\rho)-\rho} WCF_\mu$, which is the same as the coefficient on the $t^{\nu - w\cdot(\lambda+\rho)+\rho}$ term in WCF_μ, and that's the advertised weight multiplicity. □

Note that for certain λ and μ, many of these terms may be automatically zero. For example, if each $\lambda_i - \lambda_{i+1}$ is greater than $\mu_1 - \mu_n$, then they are all zero except for $w = \mathbf{Id}$, and this question reduces to a weight multiplicity.

There is a picturesque way of dropping these automatically zero terms. Recenter the V_μ weight diagram at λ. It may flop out of the dominant cone $\gamma_1 \geq \gamma_2 \geq \ldots \gamma_n$. So fold it in and subtract. Each time you do that, it may flop out a different side (though less far), so you may need to fold it again. When you're done you've computed the sum above. (The folding is really noticing the contribution of a term $w \neq \mathbf{Id}$.)

3. Gel'fand-Cetlin patterns

First we give what is called a *branching rule* from $GL_n(\mathbb{C})$ to $GL_{n-1}(\mathbb{C})$ (the block diagonal subgroup).

Theorem. *Let V_λ be an irrep of $GL_n(\mathbb{C})$ with high weight $\lambda = (\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n)$. This is a (usually reducible) representation of $GL_{n-1}(\mathbb{C})$. The multiplicity of V_μ in V_λ, $\mu = (\mu_1 \geq \ldots \geq \mu_{n-1})$, is 1 if each $\mu_i \in [\lambda_{i+1}, \lambda_i]$, 0 if not.*

This is just like the Kostant multiplicity formula, which was really a branching rule from $GL_n(\mathbb{C})$ to T. That rule wasn't *manifestly positive* like this one, though.

Proof. Let Den_n denote the Weyl denominator for $GL_n(\mathbb{C})$, Den_{n-1} the one for $GL_{n-1}(\mathbb{C})$. Then we're looking at the coefficient of t^μ (with $t_n = 1$) in

$$\sum_{w \in S_n} (-1)^w t^{w\cdot(\lambda+\rho)-\rho} \frac{Den_{n-1}}{Den_n} = \sum_{w \in S_n} (-1)^w t^{w\cdot(\lambda+\rho)-\rho} \prod_{j<n} \frac{1}{(1 - t_j^{-1} t_n)}$$

Expanding this denominator as a power series $\prod_{j<n}(1 + t_j^{-1} + t_j^{-2} + \ldots)$, we get something very simple: $\sum_{(d_i) \in \mathbb{N}^{n-1}} \prod_{i=1}^{n-1} t_i^{-d_i}$, with all coefficients 1.

So the coefficient on the t^μ term is

$$\sum_{w \in S_n} (-1)^w \prod_{i=1}^{n-1} \left[\mu_i \leq (w \cdot (\lambda + \rho) - \rho)_i\right]$$

$$= \sum_{w \in S_n} (-1)^w \prod_{i=1}^{n-1} \left[\mu_i + (n - i) \leq \lambda_{w(i)} + (n - w(i))\right]$$

using the computer science notation $[P] = 1$ if the statement P is true, $[P] = 0$ if false.

Now the clever bit, if I do say so myself. Make an $n \times n$ matrix where the (i, j) entry is $[\mu_i + i \leq \lambda_j + j]$ for $i \leq n - 1$, and 1 across the last row $i = n$. Then the above expression is just the determinant of this matrix!

What can such a matrix look like? It's all 0s and 1s. Since λ is weakly decreasing, this matrix is decreasing across rows (all 1s, then all 0s). Since μ is weakly decreasing, this matrix is increasing down columns (all 0s, then all 1s). So each row starts with a bunch of 1s, at least as many as in the last row.

If any row has the same number of 1s as the previous row, then the matrix has a repeated row – determinant zero. Or if a row has two (or more) more 1s than the previous, it's got a repeated column – determinant zero again. Otherwise, it's a lower unipotent matrix, so determinant one. This happens exactly if

$$\lambda_i + (n - i) \geq \mu_i + (n - i) > \lambda_{i+1} + (n - (i + 1))$$

i.e. $\mu_i \in [\lambda_{i+1}, \lambda_i]$ for each $i = 1, \ldots, n - 1$. □

Define a **Gel′fand-Cetlin pattern** as a triangle of numbers, bottom row labeled $\lambda_1, \lambda_2, \ldots, \lambda_n$, where each entry is weakly in between the two numbers to its southwest and southeast. Define the **weight of a GC pattern** as the differences in the row sums. (Consider the top row to be the empty row before the first entry, so that the first difference one considers is the very top entry.)

Corollary. *The irrep V_λ has a basis of weight vectors indexed by GC patterns, canonical up to individual rescaling, where the T-weight is the weight of the GC pattern.*

Proof. The basic idea is very simple – decompose the irrep under $GL_{n-1}(\mathbb{C})$ and go down from there. Each time, it breaks up into irreps, and finally they're irreps of the trivial group $GL_0(\mathbb{C})$, so they're one-dimensional (and have canonical bases, up to individual rescaling). All that's left is to keep track of the weights.

We state this as an inductive proof, where it is obvious for $GL_1(\mathbb{C})$ (decomposing a 1-dimensional irrep under the trivial group), the GC pattern has a single entry, and it is the weight of the pattern.

Decompose V_λ under $GL_{n-1}(\mathbb{C}) \times GL_1(\mathbb{C})$,

$$V_\lambda \cong \sum_{(\mu,n)} m_{\mu,k} V_\mu \otimes \mathbb{C}_k$$

where μ is a dominant weight for $GL_{n-1}(\mathbb{C})$, $k \in \mathbb{Z}$ is a dominant[1] weight for $GL_1(\mathbb{C})$, and $V_\mu \otimes \mathbb{C}_k$ is the corresponding irrep for the product.

This decomposition is greatly constrained by the (a priori coarser) decomposition we know under $GL_{n-1}(\mathbb{C})$, because each irrep $V_\mu \otimes \mathbb{C}_k$ of $GL_{n-1}(\mathbb{C}) \times GL_1(\mathbb{C})$ "breaks up as" just V_μ under the subgroup $GL_{n-1}(\mathbb{C})$. So $m_{\mu,k}$ is nonzero only for those μ interspersing between λ, and for each such μ, only one k appears.

To figure out this k, consider the action of the scalars $\{z \cdot \mathbf{1}\}, z \cdot \mathbf{1} \mapsto \prod_i z^{\lambda_i}$ on V_λ, i.e. it acts with weight $\sum_{i=1}^n \lambda_i$. The scalars are a subgroup of $GL_{n-1}(\mathbb{C}) \times GL_1(\mathbb{C})$ too. On $V_\mu \otimes \mathbb{C}_k$, it acts with weight $\sum_{i=1}^{n-1} \mu_i + k$. So $k = \sum_i^n \lambda_i - \sum_i^{n-1} \mu_i$, the difference of the last two rows' sums.

That computed the action of the lower right $GL_1(\mathbb{C})$ on the components of the breakup $V_\lambda = \sum V_\mu$; inducting from there we compute the action of each other $GL_1(\mathbb{C})$ along the diagonal in the finer and finer breakups. □

Now we have a manifestly positive branching rule from $GL_n(\mathbb{C})$ to T!

Corollary. *The μ-weight multiplicity in V_λ is the number of GC patterns with bottom λ and weight μ.*

[1] A one-element list is automatically weakly decreasing!