Schur–Weyl duality
Hjalmar Rosengren, 11 October 2015

The Schur–Weyl duality is a useful relation between representations of two seemingly very different groups: the finite group S_n and the infinite group $\mathrm{GL}(V)$, with V a finite-dimensional complex vector space. It provides an appropriate end for the course as it connects two of our main topics: tensor products of vector spaces and representations of finite groups. It also gives a glimpse into the representation theory of Lie groups, a subject with many applications in mathematics and physics. A Lie group is a group which is also a smooth manifold, such that the group operations are smooth functions. (We will not define smooth manifold but mention that any open subset in $\mathbb{R}^N$ is an example.) In particular, $\mathrm{GL}(V)$ is a Lie group since if $\dim V = m$ its elements can be identified with invertible $m \times m$ matrices, which form an open subset of $\mathbb{C}^{m^2} \simeq \mathbb{R}^{2m^2}$.

Let us first consider a tensor product $V \otimes V$. Clearly, S_2 acts on this space by $\mathrm{id}(u \otimes v) = u \otimes v$, $(12)(u \otimes v) = v \otimes u$. Since S_2 has only one-dimensional representations, it's easy to decompose $V \otimes V$ under this group action. Namely, $V \otimes V \simeq V_+ \oplus V_-$, where $V_\pm$ is the subspace consisting of elements x with $(12)x = \pm x$. Then, V_+ is a multiple of the trivial representation and V_- a multiple of the alternating representation. Moreover, the decomposition of $u \otimes v$ with respect to this splitting is

$$u \otimes v = \frac{u \otimes v + v \otimes u}{2} + \frac{u \otimes v - v \otimes u}{2}. \tag{1}$$

Note that $V_\pm \simeq (V \otimes V)/V_\mp$. In particular, V_- can be identified with the quotient of $V \otimes V$ by the subspace generated by $u \otimes v + v \otimes u$. That's exactly our definition of $V \wedge V$. Similarly, $V_+ \simeq (V \otimes V)/[u \otimes v - v \otimes u]$. This is the *symmetric tensor product*; common notations are $S^2 V$ and $V \odot V$. In summary, the decomposition of $V \otimes V$ under S_2 is

$$V \otimes V = (V \wedge W) \oplus (V \odot V).$$

When $n \geq 3$, the situation is more complicated. Consider the action of S_n on $\bigotimes^n V$ given by

$$\sigma(v_1 \otimes \cdots \otimes v_n) = v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(n)}$$

(we need inverses to get a left group action). The elements $x \in \bigotimes V^n$ satisfying $gx = x$ for all $g \in S_n$ form an invariant subspace denoted $\bigodot^n V$ and the elements satisfying $gx = \operatorname{sgn}(g)x$ an invariant subspace isomorphic to $\bigwedge^n V$. However, there are further components corresponding to the other representations of S_n. To give an idea how to find these, consider again (1). In this case, the irreducible representations are labelled by the partitions (2) and $(1,1)$ and the corresponding Young symmetrizers are

$$c_{(2)} = \mathrm{id} + (12), \qquad c_{(1,1)} = \mathrm{id} - (12).$$

We see that, up to a multiplicative constant, the action of the Young symmetrizer c_λ gives the projection from $V \otimes V$ to the corresponding invariant subspace. Thus, it's natural to look for invariant subspaces of the form $V^\lambda = c_\lambda \bigotimes^n V$.

We also need to incorporate the group $\mathrm{GL}(V)$. The map $A \mapsto \bigotimes^n A$ defines an action of $\mathrm{GL}(V)$ on $\bigotimes^n V$, that is, $\mathrm{GL}(V)$ acts by

$$A(v_1 \otimes \cdots \otimes v_n) = Av_1 \otimes \cdots \otimes Av_n.$$

Clearly, this commutes with the action of S_n. It follows that V^λ is a representation of $\mathrm{GL}(V)$. We can then formulate Schur–Weyl duality as follows.

Theorem 1. *As representations of $S_n \times \mathrm{GL}(V)$,*

$$\bigotimes^n V \simeq \bigoplus_\lambda V_\lambda \otimes V^\lambda, \tag{2}$$

where the sum runs over partitions of n. Moreover, the spaces V^λ are either zero or irreducible.

In particular,

$$\bigotimes^n V \simeq \bigoplus_\lambda \dim(V^\lambda) V_\lambda \tag{3}$$

as a representation of S_n and

$$\bigotimes^n V \simeq \bigoplus_\lambda \dim(V_\lambda) V^\lambda \tag{4}$$

as a representation of $\mathrm{GL}(V)$. The notion of Schur–Weyl duality refers to the symmetric role played by the groups S_n and $\mathrm{GL}(V)$.

To prove Theorem 1, we will need the following result. We remark that Lemma 2 is conversely an immediate consequence of Theorem 1, as the span of the elements $v \otimes \cdots \otimes v$ is clearly invariant under $\mathrm{GL}(V)$.

Lemma 2. *For any vector space V, the space $V^{(n)} = \bigodot^n V$ is spanned by $v \otimes \cdots \otimes v$, $v \in V$.*

Proof. Let $U \subseteq \bigodot^n V$ be the span of all elements $v \otimes \cdots \otimes v$. If $(e_j)_{j=1}^m$ is a basis for V and $\alpha_j \in \mathbb{C}$, then

$$\begin{aligned} u &= (\alpha_1 e_1 + \cdots + \alpha_m e_m) \otimes \cdots \otimes (\alpha_1 e_1 + \cdots + \alpha_m e_m) \\ &= \sum_{k_1,\dots,k_n=1}^{m} \alpha_{k_1} \cdots \alpha_{k_n} e_{k_1} \otimes \cdots \otimes e_{k_n} \in U. \end{aligned}$$

If we consider u as a function of α_1, then $h^{-1}(u(\alpha_1 + h) - u(\alpha_1)) \in U$. By continuity $\partial u/\partial \alpha_1 \in U$. Repeating this argument,

$$\frac{\partial^{l_1} \cdots \partial^{l_m} u}{\partial \alpha_1^{l_1} \cdots \partial \alpha_m^{l_m}} \in U.$$

If $l_1 + \cdots + l_m = n$, this gives

$$\sum_{\sigma \in S_n} \sigma(e_1 \otimes \cdots \otimes e_1 \otimes \cdots \otimes e_n \otimes \cdots \otimes e_n) \in U,$$

where each e_j appears l_j times. These vectors clearly span $\bigodot^n V$. $\square$

We will also need the following Lemma. It probably seems familiar, as we essentially derived it twice last time as part of the proof of Theorem 2 and Lemma 10.

Lemma 3. *If R is a ring, e and idempotent in R and M an R-module, then*

$$\mathrm{Hom}_R(Re, M) \simeq eM,$$

where $x \in eM$ corresponds to the homomorphism $y \mapsto yx$, $y \in Re$.

Proof. Let $\phi \in \mathrm{Hom}_R(Re, M)$ and let $x = \phi(e)$. Then $x = \phi(e^2) = e\phi(e) = ex$, so $x \in eM$. Moreover, for any $y \in Re$ we have $\phi(y) = \phi(ye) = y\phi(e) = yx$. Conversely, if $x \in eM$, $\phi(y) = yx$ is clearly an R-module homomorphism. $\square$

We apply Lemma 3 when $R = \mathbb{C}[S_n]$, e is proportional to c_λ and $M = \bigotimes^n V$. It follows that $\mathrm{Hom}_{S_n}(V_\lambda, \bigotimes^n V) \simeq V^\lambda$. This is equivalent to (3). The equation (2) means that, in addition, the map $V_\lambda \otimes V^\lambda \to \bigotimes^n V$ given by $x \otimes y \mapsto xy$, commutes with the action of $\mathrm{GL}(V)$. But that's just a restatement of the fact that the actions of S_n and $\mathrm{GL}(V)$ commute.

It remains to prove that V^λ is either zero or irreducible under the action of $\mathrm{GL}(V)$. Let us first consider the action of the algebra $B = \mathrm{End}_{\mathbb{C}[S_n]}(\bigotimes^n V)$. It follows from (2) and the fact that V_λ is irreducible that

$$B = \bigoplus_{\lambda,\mu} \mathrm{Hom}_{\mathbb{C}[S_n]}(V_\lambda \otimes V^\lambda, V_\mu \otimes V^\mu) \simeq \bigoplus_{\lambda} \mathrm{End}_{\mathbb{C}}(V^\lambda).$$

Thus, when restricted to V^λ, B acts as the full matrix algebra. Clearly, V^λ has no non-trivial invariant subspaces under that action, since any non-zero vector is mapped to any other non-zero vector by some linear transformation. Thus, as a representation of B, V^λ is either zero or irreducible.

Lemma 4. *The algebra* $B = \mathrm{End}_{\mathbb{C}[S_n]}(\bigotimes^n V)$ *is spanned algebraically by* $\mathrm{End}(V)$ *(acting diagonally on* $\bigotimes^n V$*) and topologically by* $\mathrm{GL}(V)$*.*

Proof. We have $\mathrm{End}_{\mathbb{C}[S_n]}(\bigotimes^n V) = \bigodot^n \mathrm{End}(V)$. By Lemma 2, this space is indeed spanned by the elements $A \otimes A \otimes \cdots \otimes A$, where $A \in \mathrm{End}(V)$. The second fact follows since $\mathrm{GL}(V)$ is dense in $\mathrm{End}(V)$ (the determinant is a continuous function of the matrix elements, so if $\det(A) = 0$ there is an arbitrarily small perturbation of A with non-zero determinant). $\square$

If a subspace U of V^λ is closed under $\mathrm{GL}(V)$, it follows from Lemma 4 that U is closed under B and from the preceeding discussion that $U = \{0\}$ or $U = V^\lambda$. Thus. V^λ is either zero or irreducible. This completes the proof of Theorem 1.

There is a converse of Lemma 4, saying that $\mathrm{End}_{\mathrm{End}(V)}(\bigotimes^n V)$ is given by the action of $\mathbb{C}[S_n]$. Thus, when acting on $\bigotimes^n V$, each of the algebras $\mathrm{End}(V)$ and $\mathbb{C}[S_n]$ is the commutant of the other. The "double commutant theorem" is a far-raching generalization of Theorem 1 within the general theory of semisimple algebras, which we do not go into here.

We will also compute the character of V^λ. To this end, we first note the following fact.

Lemma 5. *Let* $A \in \mathrm{GL}(V)$ *and let* $\sigma \in S_n$ *have cycle structure* $(c_1, \dots, c_n)$ *(that is, it contains exactly* k *cycles of length* k*). Then,*

$$\mathrm{Tr}_{\bigotimes^n V}(A\sigma) = \prod_{k=1}^{n} \mathrm{Tr}_V(A^k)^{c_k}.$$

Proof. Choosing a basis $(e_j)_{j=1}^m$ in V, we write $Ae_j = \sum_k A_j^k e_k$. Then,

$$(A\sigma)(e_{j_1} \otimes \cdots \otimes e_{j_n}) = \sum_{k_1,\ldots,k_n=1}^{m} A^{k_1}_{j_{\sigma^{-1}(1)}} \ldots A^{k_n}_{j_{\sigma^{-1}(n)}} e_{k_1} \otimes \cdots \otimes e_{k_n}.$$

It follows that

$$\mathrm{Tr}_{\bigotimes^n V}(A\sigma) = \sum_{k_1,\ldots,k_n=1}^{m} A^{k_{\sigma(1)}}_{k_1} \cdots A^{k_{\sigma(n)}}_{k_n}.$$

Since

$$\mathrm{Tr}(A^j) = \sum_{k_1,\ldots,k_j} A_{k_1k_2} A_{k_2k_3} \cdots A_{k_{j-1}k_j} A_{k_jk_1},$$

this can be written as indicated. For instance, when $n = 5$ and $\sigma = (12)(345)$, the sum is

$$\sum_{k_1,\ldots,k_5} A^{k_2}_{k_1} A^{k_1}_{k_2} A^{k_4}_{k_3} A^{k_5}_{k_4} A^{k_3}_{k_5} = \mathrm{Tr}(A^2)\,\mathrm{Tr}(A^3).$$

□

Suppose now that A has eigenvalues $x_1, \ldots, x_m$. Then,

$$\mathrm{Tr}_{\bigotimes^n V}(A\sigma) = \prod_{k=1}^{n} (x_1^k + \cdots + x_m^k)^{c_k} = \sum_{\lambda} \chi_\lambda(\sigma)\,\mathrm{Tr}_{V^\lambda}(A), \tag{5}$$

where the first equality follows from Lemma 5 and the second from Theorem 1.

On the other hand, Frobenius character formula (Theorem 3 last time) says that $\chi_\lambda(g)$ is the coefficient of $x^{\lambda+\rho}$ in the generating function

$$\prod_{1\le i<j\le m} (x_i - x_j) \prod_{k=1}^{n} (x_1^k + \cdots + x_m^k)^{c_k}, \tag{6}$$

where $\rho = (m-1, \ldots, 1, 0)$. The following simple Lemma will now be useful.

Lemma 6. *Let P be an anti-symmetric polynomial in m variables, that is, $P(x_{\sigma(1)}, \ldots, x_{\sigma(m)}) = P(x_1, \ldots, x_m)$. Write*

$$P(x_1, \ldots, x_m) = \sum_{k_1,\ldots,k_m \ge 0} C_{k_1,\ldots,k_m} x_1^{k_1} \cdots x_m^{k_m}.$$

Then,

$$\frac{P(x_1,\ldots,x_m)}{\prod_{1\le i<j\le m}(x_i-x_j)} = \sum_{\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_m\ge 0} C_{\lambda+\rho}\, S_\lambda(x_1,\ldots,x_m),$$

where S_λ is the Schur polynomial

$$S_\lambda(x_1,\ldots,x_m) = \frac{\det_{1\le i,j\le m}(x_i^{\lambda_j+m-j})}{\prod_{1\le i<j\le m}(x_i-x_j)}. \tag{7}$$

Proof. That P is anti-symmetric means that the coefficients $C_{k_1,\ldots,k_m}$ are anti-symmetric in $(k_1,\ldots,k_m)$. In particular, $C_{k_1,\ldots,k_m}=0$ when $k_i=k_j$ for some $i\neq j$. If that is not the case, we can write $k=\sigma(\lambda+\rho)$ for a unique sequence $\lambda_1\ge\cdots\ge\lambda_m\ge 0$ and a unique $\sigma\in S_m$. This gives

$$P(x_1,\ldots,x_m) = \sum_{\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_m\ge 0} C_{\lambda+\rho}\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)x_{\sigma(1)}^{\lambda_1+\rho_1}\cdots x_{\sigma(m)}^{\lambda_m+\rho_m}.$$

Writing the inner sum as a determinant completes the proof. $\square$

Note that S_λ is really a polynomial, since when $x_i=x_j$ for some $i\neq j$, two rows in the determinant are equal and hence the numerator in (7) vanishes.

Frobenius character formula says that when P is given by (6) and $\lambda_m>0$, then $C_{\lambda+\rho}=\chi_\lambda(g)$. If we want to compute $C_{\lambda+\rho}$ when $\lambda_k>0$ and $\lambda_{k+1}=0$ then it follows from the case $m\mapsto k$ of Frobenius character formula that $C_{\lambda+\rho}=\chi_{[\lambda]}(g)$, where $[\lambda]$ is obtained from λ by deleting 0s at the end. Comparing this with (5), we can draw the following conclusion.

Proposition 7. *Let $\dim(V)=m$, let λ be a partition with k rows and $A\in\mathrm{GL}(V)$ have eigenvalues $x_1,\ldots,x_m$. Then, if $k\le m$,*

$$\mathrm{Tr}_{V^\lambda}(A) = S_\lambda(x_1,\ldots,x_m),$$

where on the right-hand side we complete λ to an m-dimensional vector by defining

$$(\lambda_1,\ldots,\lambda_m) = (\lambda_1,\ldots,\lambda_k,0,\ldots,0). \tag{8}$$

Moreover, if $k>m$, then $V^\lambda=\{0\}$.

We can easily use Proposition 7 to compute the dimension of V^λ. Using the Vandermonde determinant, it follows that

$$S_\lambda(1, t, \dots, t^{m-1}) = \prod_{1 \le i < j \le m} \frac{t^{\lambda_i + m - i} - t^{\lambda_j + m - j}}{t^{m-i} - t^{m-j}}.$$

As $t \to 1$, this gives

$$\dim(V^\lambda) = S_\lambda(1, 1, \dots, 1) = \prod_{1 \le i < j \le m} \frac{\lambda_i - \lambda_j + j - i}{j - i},$$

where we are still using (8).