AN EXACT FORMULA FOR INTEGRATING POLYNOMIALS OVER $U(d)$

IAN WEINER

Abstract. This report is primarily an exposition of a result from the 2004 paper of Collins and Śniady[1], *Integration with Respect to the Haar Measure on Unitary, Orthogonal and Symplectic Group.* Beginning with the classical Schur-Weyl duality result, we derive the expression for the moments of unitary group $U(d)$ in terms of the Weingarten function. Also included is a description of a basic MATLAB implementation allowing the computation of arbitrary moments of $U(d)$ of order ≤ 10 for any dimension d.

1. Introduction

Let $U(d)$ be the group of unitary $d \times d$ matrices over $\mathbb{C}$. As a compact Lie group it comes equipped with a unique probability measure which is invariant under left and right group multiplication. This is the Haar measure, which we denote dU. We are interested in a formula for computing the *moments* of $U(d)$, by which we mean expressions of the form

$$\int_{U(d)} U_{i_1 j_1} \cdots U_{i_n j_n} \overline{U}_{i'_1 j'_1} \cdots \overline{U}_{i'_m j'_m} dU$$

where i, j, i', j' are arbitrary multi-indices. Of course, armed with a formula for the moments, any polynomial expression in the entries of U may be integrated with respect to Haar measure. We will see that the symmetry of the Haar measure with respect to group multiplication allows us to compute the moments with an exact algebraic expression.

Before treating the general case we will examine a few special cases. First up: What is $\int U_{ij} dU$ for some indices i, j? The symmetries of Haar measure require this to be zero. One neat way to see this is to form a $d \times d$ matrix M whose i, j^{th} entry is the above, for $1 \leq i, j \leq d$. Writing the entirety in matrix form, we have, for *any* fixed $U_0 \in U(d)$, $M = \int U\, dU = \int U_0 U\, dU = U_0 \int U\, dU = AM$, where we used the fact that dU is invariant under the change of variables $U \mapsto U_0 U$. But $M = U_0 M$ for all unitary U_0 can only hold if $M = 0$. In fact, using $U_0 = e^{i\theta} I$ we can deal directly with the moments themselves: $\int U_{ij} dU = \int e^{i\theta} U_{ij} dU = e^{i\theta} \int U_{ij} dU \implies \int U_{ij} dU = 0$. This also works for the more general moments, giving an overall phase factor unless the number of $U'_{ij}s$ and the number of $\overline{U}'_{i'j'}s$ are equal, so $\int_{U(d)} U_{i_1 j_1} \cdots U_{i_n j_n} \overline{U}_{i'_1 j'_1} \cdots \overline{U}_{i'_m j'_m} dU = 0$ whenever $n \neq m$.

In fact, multiplying on the left and right by diagonal unitary matrices U_0 and U_1 with different phase factors $e^{i\theta_k^{(0)}}, e^{i\theta_k^{(1)}}$ for each diagonal entry shows that, for the moment to be non-zero, each index appearing in the multi-index i must appear *the same number of times* in the multi-index i', and similarly for the indices of j and j'. Combining the last two observations, we have:

Proposition. *$\int_{U(d)} U_{i_1j_1}\cdots U_{i_nj_n}\overline{U}_{i'_1j'_1}\cdots\overline{U}_{i'_mj'_m}dU = 0$ whenever $m \neq n$. Furthermore, when $m = n$, for the moment to be non-zero it is necessary that there exist permutations $\sigma, \tau \in S_n$ such that $\sigma(i) = i'$ and $\tau(j) = j'$.*

Finally, let us consider an evidently non-zero integral: $\int U_{ij}\overline{U}_{ij}dU = \int |U_{ij}|^2 \, dU$. To compute this, first note that for any unitary matrix, $\Sigma_j |U_{ij}|^2 = 1$, so $\Sigma_j \int |U_{1j}|^2 \, dU = 1$. But since we can permute the column indices of U by right-multiplication with a unitary permutation matrix, which leaves dU invariant, all the terms in this sum must be equal, so $\int |U_{ij}|^2 \, dU = 1/d$ for each $1 \leq i, j \leq d$.

When $n > 1$ we cannot use such elementary arguments. However, the interaction between the unitary group and permutations acting on indices that we've seen hinted at above will be crucial. The classical result of Schur described below categorizes this interaction in an elegant way, and is key to the main result.

2. Representations and Schur-Weyl Duality

Both $U(d)$ and S_n act on $\mathbb{C}^d \otimes \cdots \otimes \mathbb{C}^d = \left(\mathbb{C}^d\right)^{\otimes n}$ in a natural way. If $v_k \in \mathbb{C}^d, U \in U(d)$, and $\sigma \in S_n$, we define:

$$\begin{aligned} U^{\otimes n} \cdot (v_1 \otimes \cdots \otimes v_n) &= (Uv_1) \otimes \cdots \otimes (Uv_n) \\ \sigma \cdot (v_1 \otimes \cdots \otimes v_n) &= v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(n)} \end{aligned}$$

These actions are linear over $\mathbb{C}$, so they induce *representations* of $U(d)$ and S_n as groups of complex matrices, formally written $\rho^n_{U(d)} : U(d) \to End\left(\mathbb{C}^d\right)^{\otimes n}$ and $\rho^d_{S_n} : S_n \to End\left(\mathbb{C}^d\right)^{\otimes n}$. The actions of U and σ commute, so they give a natural combined representation $\rho_{S_n \times U(d)}$.

To describe this representation, first we need to know about the irreducible representations of $U(d)$ and S_n, respectively, which are well-studied and characterized[2, 3]. For brevity we only review what is needed. For notation, we define a *partition* λ to be a non-increasing sequence of non-negative integers. Its length $l(\lambda)$ is the number of positive integers in the sequence, and we write $\lambda \vdash n$ if $|\lambda| \equiv \Sigma_i \lambda_i = n$.

Fact. *To each partition λ with $l(\lambda) \leq d$ there is a unique irreducible complex representation $\rho^\lambda_{U(d)} : U(d) \to V^\lambda$. The dimension $d^\lambda_{U(d)} \equiv \dim_{\mathbb{C}} V^\lambda$ is given by*

$$d^\lambda_{U(d)} = \prod_{1 \leq i < j \leq d} \frac{\lambda_i - \lambda_j + j - i}{j - i}$$

To each partition $\lambda \vdash n$ there is a unique irreducible complex representation $\rho^\lambda_{S_n} : S_n \to W^\lambda$. The corresponding characters are denoted χ^λ and their character tables may be computed algorithmically via the Murnaghan-Nakayama rule[4]. *As always, $d^\lambda_{S_n} \equiv \dim_{\mathbb{C}} W^\lambda = \chi^\lambda(e)$. The group algebra $\mathbb{C}[S_n]$ decomposes as the direct sum*

$$\mathbb{C}[S_n] \cong \bigoplus_{\lambda \vdash n} End\, W^\lambda$$

Now we can state Schur-Weyl duality[5]:

Fact. *Let $P_{n,d} = \{\lambda : \lambda \vdash n$ and $l(\lambda) \leq d\}$. The action of $S_n \times U(d)$ on $\left(\mathbb{C}^d\right)^{\otimes n}$ can be expressed by the following decomposition:*

$$\left(\mathbb{C}^d\right)^{\otimes n} \cong \bigoplus_{\lambda \in P_{n,d}} \left(V^\lambda \otimes W^\lambda\right)$$

with $\rho_{S_n \times U(d)}$ acting on the $\left(V^\lambda \otimes W^\lambda\right)$ subspace as $\rho^\lambda_{S_n} \otimes \rho^\lambda_{U(d)}$.

Recall that the group algebra $\mathbb{C}[S_n]$ is defined to be the algebra of complex-valued functions on S_n, with multiplication given by convolution. It is isomorphic to the algebra of formal linear combinations of elements inS_n with coefficients in $\mathbb{C}$, with multiplication defined by extending group multiplication to be linear over $\mathbb{C}$. We define $\mathbb{C}_d[S_n]$ to be the subalgebra obtained by only considering λ with $l(\lambda) \leq d$:

$$\mathbb{C}_d[S_n] = \bigoplus_{\lambda \in P_{n,d}} End\, W^\lambda$$

From Schur-Weyl duality, we can naturally interpret $\mathbb{C}_d[S_n]$ in yet another way– as a subalgebra of $End\ \left(\mathbb{C}^d\right)^{\otimes n}$, that is, complex matrices acting on vectors of the form $v_1 \otimes \cdots \otimes v_n$. Moving freely between these interpretations will be useful in the next section.

3. The Main Result

In this section we prove the main result:

Theorem.

$$\int\limits_{U(d)} U_{i_1 j_1} \cdots U_{i_n j_n} \overline{U}_{i'_1 j'_1} \cdots \overline{U}_{i'_n j'_n} dU = \sum_{\sigma : \sigma(i) = i'} \sum_{\tau : \tau(j) = j'} Wg(\tau \sigma^{-1})$$

where Wg is the Weingarten function, defined as

$$Wg(\sigma) \equiv \frac{1}{(n!)^2} \sum_{\lambda \in P_{n,d}} \frac{\left(d^\lambda_{S_n}\right)^2}{d^\lambda_{U(d)}} \chi^\lambda(\sigma)$$

We begin with some more general definitions to illuminate the algebraic structure to be exploited. First, note that for any $A \in End\ \left(\mathbb{C}^d\right)^{\otimes n}$ and any $U \in U(d)$ there is a conjugation action $A \mapsto U^{\otimes n} A \overline{U}^{\otimes n}$. This produces an operator acting on (linear combinations of) vectors of the form $v_1 \otimes \cdots \otimes v_n$ by first multiplying by $\overline{U}$ on each factor of the tensor product, then applying A, and finally multiplying all resulting factors by U. Going one step further by integrating the result over all U, we define the *conditional expectation* of A by

$$\mathbb{E}(A) \equiv \int\limits_{U(d)} U^{\otimes n} A \left(U^*\right)^{\otimes n} dU$$

The result is also in $End\left(\mathbb{C}^d\right)^{\otimes n}$ but it has been symmetrized in the following sense: For any fixed $U_0 \in U(d)$,

$$\begin{aligned} U_0^{\otimes n}\mathbb{E}(A) &= \int_{U(d)} U_0^{\otimes n} U^{\otimes n} A \left(U^*\right)^{\otimes n} dU \\ &= \left(\int_{U(d)} (U_0 U)^{\otimes n} A \left((U_0 U)^*\right)^{\otimes n} dU\right) U_0^{\otimes n} = \mathbb{E}(A) U_0^{\otimes n} \end{aligned}$$

where in the last equality we have used the invariance of Haar measure with respect to the coordinate transformation $U \mapsto U_0 U$. Since $\mathbb{E}(A)$ commutes with the action of every unitary matrix, the Schur-Weyl theorem tells us that its $U(d)$ piece is trivial; it is a matrix representation arising entirely from the algebra $\mathbb{C}_d[S_n]$. (This explains the terminology "conditional expectation"– we have integrated out the $U(d)$ component but left a dependance on the S_n piece).

Using a similar argument based on invariance of Haar measure, one can show a second important property: for all A, $Tr(\mathbb{E}(A)) = Tr(A)$.

The final property that we will note for future use is that for any $\sigma \in S_n$, $\mathbb{E}\left(A\rho^d_{S_n}(\sigma)\right) = \mathbb{E}(A)\rho^d_{S_n}(\sigma)$, which is evident from the fact that $\overline{U}^{\otimes n}$ commutes with the action of permutations σ.

Now we'll show how the conditional expectation will be reduced to the expression defining a moment of $U(d)$: Let multi-indices i, j, i', j' be given. Define$A, B \in End\left(\mathbb{C}^d\right)^{\otimes n}$ in the standard basis by

$$\begin{aligned} A\left(e_{i_1} \otimes \cdots \otimes e_{i_n}\right) &= e_{i'_1} \otimes \cdots \otimes e_{i'_n}, \quad \text{zero for other basis vectors of } \left(\mathbb{C}^d\right)^{\otimes n} \\ B\left(e_{j'_1} \otimes \cdots \otimes e_{j'_n}\right) &= e_{j_1} \otimes \cdots \otimes e_{j_n}, \quad \text{zero for other basis vectors of } \left(\mathbb{C}^d\right)^{\otimes n} \end{aligned}$$

(Note the different location of the primes). Then, denoting the columns of U by U_k, we have

$$\begin{aligned} Tr\left(A\mathbb{E}(B)\right) &= Tr\left(A \int_{U(d)} U_{j_1}\left(U_{j'_1}\right)^* \otimes \cdots \otimes U_{j_n}\left(U_{j'_n}\right)^*\right) \\ &= \int_{U(d)} U_{i_1 j_1}\overline{U}_{i'_1 j'_1} \cdots U_{i_n j_n}\overline{U}_{i'_n j'_n} dU \end{aligned}$$

Recall our observation that $\mathbb{E}(B) \in \mathbb{C}_d[S_n]$ via the injective representation induced by the Schur-Weyl theorem. Thus, to finish the theorem, we will study the function $\Phi : End\left(\mathbb{C}^d\right)^{\otimes n} \to \mathbb{C}_d[S_n]$ given by

$$\Phi(A) \equiv \sum_{\sigma \in S_n} \left[Tr\left(A\rho^d_{S_n}\left(\sigma^{-1}\right)\right)\right]\sigma$$

(Here we are treating $\mathbb{C}_d[S_n]$ as complex-linear combinations of the elements of S_n). We see that Φ is compatible with the operations of $\mathbb{C}_d[S_n]$ since for any $\tau \in S_n$,

$$\begin{aligned} \Phi\left(A\rho^d_{S_n}(\tau)\right) &= \sum_{\sigma \in S_n} \left[Tr\left(A\rho^d_{S_n}\left(\tau\sigma^{-1}\right)\right)\right]\sigma \\ &= \sum_{\pi \in S_n} \left[Tr\left(A\rho^d_{S_n}\left(\pi^{-1}\right)\right)\right]\pi\tau = \Phi\left(A\right)\tau \end{aligned}$$

where $\pi \equiv \sigma\tau^{-1}$. Similarly we can show that $\Phi\left(\rho_{s_n}^d(\tau)A\right) = \tau\Phi(A)$.

We saw above that $Tr(B) = Tr\left(\mathbb{E}(B)\right)$ and $\mathbb{E}\left(B\rho_{S_n}^d(\sigma)\right) = \mathbb{E}(B)\rho_{S_n}^d(\sigma)$, so we also have $Tr\left(B\rho_{S_n}^d\left(\sigma^{-1}\right)\right) = Tr\left(\mathbb{E}\left(B\rho_{S_n}^d\left(\sigma^{-1}\right)\right)\right) = Tr\left(\mathbb{E}(B)\rho_{S_n}^d\left(\sigma^{-1}\right)\right)$ which, along with the above compatibility property of Φ, gives

$$\Phi(B) = \Phi\left(\mathbb{E}(B)\right) = \mathbb{E}(B)\Phi\left(id\right)$$

If we assume $\Phi(id)$ is invertible in $\mathbb{C}_d[S_n]$ (as we show below), this in turn allows us to write

$$\Phi\left(A\mathbb{E}\left(\mathbb{B}\right)\right) = \Phi(A)\mathbb{E}(B) = \Phi(A)\Phi(B)\Phi(id)^{-1}$$

We saw above that when A and B are defined properly in terms of the multi-indices i, j, i', j', the coefficient in front of the identity permutation e in this equation is the corresponding moment on the left hand side. On the right hand side, it is obtained by a convolution, summing over coefficients of the three factors whose corresponding elements of S_n multiply to e. By the definition of A, $Tr\left(A\rho_{S_n}^d\left(\sigma^{-1}\right)\right) = 1$ whenever $\sigma(i) = i'$and zero otherwise. Also, $Tr\left(B\rho_{s_n}^d\left(\tau^{-1}\right)\right) = 1$ whenever $\tau(j') = j$ and zero otherwise, which is equivalent to being 1 when $\tau^{-1}\left(j\right) = j'$ and zero otherwise. Thus the convolution can be written as

$$\begin{aligned}\left[\Phi(A)\Phi(B)\Phi(id)^{-1}\right](e) &= \sum_{\sigma\in S_n}\sum_{\tau\in S_n}\left[\Phi(A)\right](\sigma)\left[\Phi(B)\right]\left(\tau^{-1}\right)\left[\Phi(id)^{-1}\right]\left(\tau\sigma^{-1}\right)\\ &= \sum_{\sigma:\sigma(i)=i'}\ \sum_{\tau:\tau(j)=j'}\left[\Phi(id)^{-1}\right]\left(\tau\sigma^{-1}\right)\end{aligned}$$

where we are now treating $\Phi(id)^{-1} \in \mathbb{C}_d[S_n]$ as a function $S_n \to \mathbb{C}$.

To complete the proof we only need to show that $\Phi(id)^{-1}$ is the Weingarten function defined in the theorem. To see this, first note that for any $\sigma \in S_n$ we have

$$\left[\Phi(id)\right](\sigma) = Tr\left(\rho_{Sn}^d\left(\sigma^{-1}\right)\right) = Tr\left(\rho_{S_n}^d\left(\sigma\right)\right) = \chi_{S_n}^d(\sigma) = \sum_{\lambda\in P_{n,d}} d_{U(d)}^{\lambda}\chi^{\lambda}\left(\sigma\right)$$

where $\chi_{S_n}^d(\sigma)$ denotes the character of the $\rho_{S_n}^d$ representation. The second equality follows from the fact that $\rho_{S_n}^d(\sigma)$ is a permutation matrix, hence its trace is the same as the trace of its inverse. The final equality is simply the apparent form of the character given the characterization of the Schur-Weyl duality theorem. Now multiplying in $\mathbb{C}_d[S_n]$ we have, by the Schur orthogonality relations for irreducible characters,

$$\Phi(id)Wg = \frac{1}{(n!)^2}\sum_{\lambda\in P_{n,d}}\left(d_{S_n}^{\lambda}\right)^2\left(\chi^{\lambda}\right)^2 = \frac{1}{n!}\sum_{\lambda\in P_{n,d}}\left(d_{S_n}^{\lambda}\right)^2 e = e$$

which completes the proof.

4. MATLAB Implementation

Included with this report is a MATLAB implementation of the above result. Due to time constraints, the full algorithmic Monaghan-Nakayama rule was not implemented; instead, the character tables for S_n were hard-coded for $n \leq 5$, permitting computation of all moments of order ≤ 10 for any d. The functions included are:

- sdim(L,d) : computes $d^{\lambda}_{U(d)}$ for λ specified by the row vector L.
- charSn(L,M) : computes $\chi^{\lambda}(\mu)$ where $\lambda$and μ are specified by row vectors L and M, respectively. μ is a conjugacy class of a permutation, specified as a partition $\mu \vdash n$ representing the sizes of the disjoint cycles of the permutation. Currently requires $n \leq 5$.
- getLs(n,d,m) : returns cell array representing $P_{n,d}$ with optional parameter m constraining the maximum entry of the partitions.
- Wg(M,d) : computes the Weingarten function for given d evaluated on permutation conjugacy class specified in partition form by row vector M.
- getConjClass(sigma,tau) : Computes the conjugacy class of $\tau\sigma^{-1}$. Input permutations are specified by row vectors $[\sigma(1), \sigma(2), \dots, \sigma(n)]$, etc. The output is a row vector representing the partition of the conjugacy class corresponding to disjoint cycles.
- uMoment(d,I,J,Ic,Jc) : Given dimension d and multi-indices i, j, i', j' specified by row vectors I, J, Ic, Jc respectively, computes the corresponding moment of $U(d)$ using the formula of the main theorem.

Example. Compute $Wg(\sigma)$ for $d = 3$ and σ containing two length-2 cycles and one fixed point.

```
>> Wg([2 2 1],3)
ans =
  -1.3558e-004
```

Example. Compute $\int_{U(6)} U_{1,2}U_{1,2}U_{3,1}U_{2,1}U_{2,3}\overline{U}_{3,2}\overline{U}_{1,1}\overline{U}_{1,1}\overline{U}_{2,2}\overline{U}_{2,3}dU$.

```
>> uMoment(6,[1,1,3,2,2],[2,2,1,1,3],[3,1,1,2,2],[2,1,1,2,3])
ans =
  6.6138e-006
```

References

[1] B. Collins, P. Śniady, *Integration With Respect to the Haar Measure on Unitary, Orthogonal and Symplectic Group*, Commun. Math. Phys. 264 (2006) 773{795. arXiv:math-ph/0402073

[2] W. Fulton, J. Harris, *Representation theory: A First Course*, Vol. 129 of Graduate Texts in Mathematics, Springer Verlag, 1991

[3] B. Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms, and symmetric Functions*, Vol. 203, Springer Verlag, 2001.

[4] D. Bernstein, *The Computational Complexity of Rules for the Character Table of S_n*, Journal of Symbolic Computation 37 (6) (2004) 727{748.

[5] Hermann Weyl. *The Classical Groups. Their Invariants and Representations.* Princeton University Press, Princeton, N.J., 1939.