Supplement to “Ranking with kernels in Fourier space” (COLT10)

Risi Kondor
Center for the Mathematics of Information
California Institute of Technology
risi@caltech.edu

Marconi Barbosa
NICTA, Australian National University
marconi.barbosa@nicta.com.au

Abstract

This document contains background information and proofs to supplement the paper titled “Ranking with kernels in Fourier space” published at the 23rd International Conference on Learning Theory (COLT), 2010.

1 Definitions from group theory and representation theory

Sets. In the paper and in this supplement for $i, j \in \mathbb{Z}$ with $i \leq j$, $[i, j]$ denotes the set $\{i, i+1, \ldots, j\}$.

Conjugation. We denote the complex conjugate of a number $z \in \mathbb{C}$ as z^* and the Hermitian conjugate of a matrix $M \in \mathbb{C}^{d_1 \times d_2}$ as $M^\dagger \equiv M^{*\top}$.

Groups. A **group** is a set G endowed with an operation $G \times G \to G$ (usually denoted multiplicatively) obeying the following axioms:

G1. for any $x, y \in G$, $xy \in G$ (closure);

G2. for any $x, y, z \in G$, $x(yz) = (xy)z$ (associativity);

G3. there is a unique $e \in G$, called the **identity** of G, such that $ex = xe = x$ for any $x \in G$;

G4. for any $x \in G$, there is a corresponding element $x^{-1} \in G$ called the **inverse** of x, such that $x\,x^{-1} = x^{-1}x = e$.

We *do not* require that the group operation be commutative, i.e., in general, $xy \neq yx$. A subset H of G is called a **subgroup** of G, denoted $H < G$, if H itself forms a group with respect to the same operation as G, i.e., if for any $x, y \in H$, $xy \in H$. If $x \in G$ and $H < G$, then $xH = \{\, xh \mid h \in H \,\} \subset G$ is called a (left–)H–**coset**.

Conjugacy. A pair of group elements $x, y \in G$ are said to be **conjugate** if there is some $z \in G$ such that $x = z^{-1}yz$. Conjugacy is an equivalence relation that partitions G into **conjugacy classes**. A function $f\colon G \to \mathbb{C}$ that is constant on each conjugacy class is called a **class function**.

Generating sets. A subset U of a finite group G is said to be a **generating set** for G if any $x \in G$ can be expressed as a product of elements from U.

Representations. For the purposes of this paper a **representation** of G over $\mathbb{C}$ is a matrix-valued function $\rho\colon G \to \mathbb{C}^{d_\rho \times d_\rho}$ such that $\rho(x)\,\rho(y) = \rho(xy)$ for any $x, y \in G$. We call d_ρ the **order** or the **dimensionality** of ρ. Note that $\rho(e) = I$ for any representation. Two representations ρ_1 and ρ_2 of the same dimensionality d are said to be **equivalent** if for some invertible $T \in \mathbb{C}^{d \times d}$, $\rho_1(x) = T^{-1}\rho_2(x)\,T$ for any $x \in G$. A representation ρ is said to be **reducible** if it decomposes into a direct sum of smaller representations in the form

$$\rho(x) = T^{-1}\left(\rho_1(x) \oplus \rho_2(x)\right) T = T^{-1} \left(\begin{array}{c|c} \rho_1(x) & 0 \\ \hline 0 & \rho_2(x) \end{array}\right) T \qquad \forall\, x \in G$$

for some invertible $T \in \mathbb{C}^{d_\rho \times d_\rho}$.

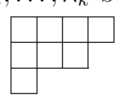
Irreps. A maximal set of pairwise inequivalent, irreducible representations we call a system of **irreps**. It is possible to show that if $\mathcal{R}_1$ and $\mathcal{R}_2$ are two different systems of irreps of the same finite group G, then $\mathcal{R}_1$ and $\mathcal{R}_2$ have the same (finite) cardinality, and there is a bijection $\phi\colon \mathcal{R}_1 \to \mathcal{R}_2$ such that if $\phi(\rho_1) = \rho_2$, then ρ_1 and ρ_2 are equivalent. The theorem of total reducibility asserts that given a system of irreps $\{\rho_1, \rho_2, \ldots, \rho_k\}$, any representation ρ can be reduced into a direct sum of the ρ_i's in the sense that there is an invertible matrix T and sequence of mutiplicities $m_1, m_2, \ldots, m_k$ such that $\rho(x) = T^{-1}\big[\bigoplus_{i=1}^{k} \bigoplus_{j=1}^{m_k} \rho_i(x)\big]\, T$. For more information on representation theory, see, e.g. (J.–P. Serre: Linear Representations of Finite Groups, Springer–Verlag, 1977).

Unitarity irreps. A representation ρ is said to be **unitary** if $\rho(x)^{-1} = \rho(x^{-1}) = \rho(x)^{\dagger}$ for all $x \in G$. A finite group always has at least one system of irreps which is unitary. In particular, Young's orthogonal reprsentation (see below) is such a system. Throughout the paper by "irrep" we will tacitly always mean "unitary irrep".

Permutations. A **permutation** of $[1, n]$ is a bijective mapping $\sigma\colon [1, n] \to [1, n]$. The product of two such permutations is defined by composition, $(\sigma_2\sigma_1)(i) = \sigma_2(\sigma_1(i))$ for all $i \in [1, n]$. With respect to this operation the set of all $n!$ possible permutations of $[1, n]$ form a group called the **symmetric group** of degree n, which we denote $\mathbb{S}_n$. For $m < n$ we identify $\mathbb{S}_m$ with the subgroup of permutations that only permute $[1, m]$, i.e., $\mathbb{S}_m = \{\, \tau \in \mathbb{S}_n \mid \tau(m+1) = m+1,\ \tau(m+2) = m+2,\ \ldots,\ \tau(n) = n \,\}$.

Cycle notation. A **cycle** in a permutation $\sigma \in \mathbb{S}_n$ is a sequence $(c_1, c_2, \ldots, c_k)$ such that for $i = 1, 2, \ldots, k-1$, $\sigma(c_i) = c_{i+1}$ and $\sigma(c_k) = c_1$. Any permutation can be expressed as a product of disjoint cycles. Some special permutations that we are interested in are the **transpositions** (i, j), the **adjacent transpositions** $\tau_i = (i, i+1)$, and the **contiguous cycles** $[\![i, j]\!] = (i, i+1, \ldots, j)$.

Partitions and Young diagrams. A sequence of positive integers $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ is said to be an **integer partition** of n (denoted $\lambda \vdash n$) if $\sum_{i=1}^{k} \lambda_i = n$ and $\lambda_i \geq \lambda_{i+1}$ for $i = 1, 2, \ldots, k-1$. The **Young diagram** of λ consists of $\lambda_1, \lambda_2, \ldots, \lambda_k$ boxes laid down in consecutive rows, as in



for $\lambda = (4, 3, 1)$. The **transpose** of λ is the partition λ' that we get by exchanging the rows with the columns in the diagram of λ. For example, $(4, 3, 1)' = (3, 2, 2, 1)$.

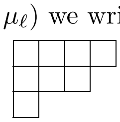
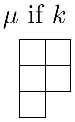
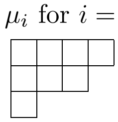
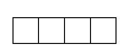
Young tableaux. A Young diagram with numbers in its cells is called a **Young tableau**. A **standard Young tableau (SYT)** is a Young tableau in which each of the numbers $1, 2, \ldots, n$ is featured exactly once, and in such a way that in each row the numbers increase from left to right and in each column they increase from top to bottom. For example,

1	2	5	8
3	4	7	
6			

(1)

is a standard tableau of shape $\lambda = (4, 3, 1)$. An alternative way to specify a standard tableau t is by its **Yamanouchi symbol** $t = [r_n r_{n-1} \ldots r_1]$, where r_i is the row index of the number i in t. We use the shorthand $[r_n r_{n-1} \ldots r_{n-k} \ldots]_n$ for the Yamanouchi symbol beginning with $r_n r_{n-1} \ldots r_{n-k}$ and padded with the appropriate number of 1's so as to result in a tableau of n boxes. For example, $[123122\ldots]_8$ is equivalent to (1), and $[\ldots]_n$ is the unique standard tableau of shape $\lambda = (n)$.

Partial orders on partitons and SYT's. There is a natural partial order on partitions induced by the inclusion order of their respective diagrams. In particular, for $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ and $\mu = (\mu_1, \mu_2, \ldots, \mu_\ell)$ we write $\lambda \geq \mu$ if $k \geq \ell$, and $\lambda_i \geq \mu_i$ for $i = 1, 2, \ldots, \ell$. For example,

 $\geq$  and  $\geq$ .

If $\lambda \geq \mu$, then we say that λ is descended from μ or that μ is an ancestor of λ. Similarly, for a pair of standard Young tableaux t and t' (or the corresponding Yamanouchi symbols) we write $t \geq t'$ (and say that t is a descendant of its ancestor t') if t' is a subtableau of t. For example,

1	2	5	8
3	4	7	
6			

$\geq$

1	2	5
3	4	
6		

Naturally, if t is of shape λ and t' is of shape μ, then $t \geq t'$ imples $\lambda \geq \mu$, but not the other way round.

Indexing the irreps. The significance of integer partitions and standard Young tableaux is that given a system of irreps $\mathcal{R}$ of $\mathbb{S}_n$, the former are in bijection with the individual irreps $\rho \in \mathcal{R}$, while the latter are in bijection with the rows and columns of the actual representation matrices. Exploiting these bijections, we label the irreps by the partitions $\{\lambda \vdash n\}$, and we label the individual rows and columns of $\rho_\lambda(\sigma)$ by the standard tableaux of shape λ (or, alternatively, the Yamanouchi symbols of shape λ). The dimensionalities $d_\lambda \equiv d_{\rho_\lambda}$ are determined by how many SYT there are of shape λ. A useful formula to answer this question is the so-called **hook rule**

$$d_\lambda = \frac{n!}{\prod_i \ell(i)}, \tag{2}$$

where i ranges over all the cells of λ, and $\ell(i)$ are the lengths of the corresponding "hooks", i.e., the number of cells to the right of cell i plus the number of cells below i plus one. For example, it is easy to check that $d_{(1)} = 1$, $d_{(n-1,1)} = n-1$, $d_{(n-2,2)} = n(n-3)/2$, and $d_{(n-2,1,1)} = (n-1)(n-2)/2$.

Young's Orthogonal Representation. The specific system of irreps that we use in this paper is called **Young's orthogonal representation**, or just **YOR**. A special feature of YOR is that its irreps are not only unitary, but also real-valued, hence the $\rho_\lambda(\sigma)$ matrices are orthogonal. YOR is defined by explicitly specifying the representation matrices corresponding to adjacent transpositions. For any standard tableau t, letting $\tau_i(t)$ be the tableau that we get from t by exchanging the numbers i and $i+1$ in its diagram, the rule defining $\rho_\lambda(\tau_i)$ in YOR is the following: if $\tau_i(t)$ is *not* a standard tableau, then the column of $\rho_\lambda(\tau_i)$ indexed by t is zero, except for the diagonal element $[\rho_\lambda(\tau_i)]_{t,t} = 1/d_t(i,i+1)$; if $\tau_i(t)$ is a standard tableau, then in addition to this diagonal element, we also have a single non-zero off-diagonal element, $[\rho_\lambda(\tau_i)]_{\tau_k(t),t} = (1 - 1/d_t(i,i+1)^2)^{1/2}$. All other matrix entries of $\rho_\lambda(\tau_i)$ are zero. In the above $d_t(i,i+1) = c_t(i+1) - c_t(i)$, where $c(j)$ is the column index minus the row index of the cell where j is located in t. Note that the trivial representation $\rho_{\text{triv}}(\sigma) \equiv 1$ is the irrep indexed by $\lambda = (n)$. Also note that for any $\lambda \vdash n$, each row/column of $\rho_\lambda(\tau_i)$ has at most two non-zero entries.

The Fourier transform. The **Fourier transform** of a function $f\colon \mathbb{S}_n \to \mathbb{C}$ is defined as the collection of matrices

$$\widehat{f}(\lambda) = \sum_{\sigma \in \mathbb{S}_n} f(\sigma)\, \rho_\lambda(\sigma) \qquad \lambda \vdash n. \tag{3}$$

As always, we assume that the ρ_λ representations are given in YOR. Since YOR is a system of real valued representations, if f is a real valued function, then the $\widehat{f}(\lambda)$ Fourier components are real valued matrices. Non-commutative Fourier transforms such as (3) enjoy many of the same properties as the usual Fourier transforms on the real line and the unit circle. In particular (3) is an invertible, unitary mapping $\mathbb{C}^{\mathbb{S}_n} \to \bigoplus_{\lambda \vdash n} \mathbb{C}^{d_\lambda \times d_\lambda}$. The **inverse Fourier transform** is

$$f(\sigma) = \frac{1}{n!} \sum_{\lambda \vdash n} d_\lambda \operatorname{tr}\big[\widehat{f}(\lambda)\, \rho_\lambda(\sigma)^{-1}\big].$$

Convolution and correlation. The **convolution** of $f\colon \mathbb{S}_n \to \mathbb{C}$ with $g\colon \mathbb{S}_n \to \mathbb{C}$ is defined $(f * g)(\sigma) = \sum_{\tau \in \mathbb{S}_n} f(\sigma\tau^{-1})\, g(\tau)$. The convolution theorem states that $\widehat{f * g}(\lambda) = \widehat{f}(\lambda)\, \widehat{g}(\lambda)$ for each $\lambda \vdash n$. The **correlation** of f with g is defined $(f \star g)(\sigma) = \sum_{\tau \in \mathbb{S}_n} f(\sigma\tau) g(\tau)^*$ and the correlation theorem states that $\widehat{f \star g}(\lambda) = \widehat{f}(\lambda)\, \widehat{g}(\lambda)^\dagger$.

2 Proofs

Proof of Proposition 1. K is positive definite if and only if

$$Q = \sum_{\sigma \in \mathbb{S}_n} \sum_{\tau \in \mathbb{S}_n} a(\sigma)\, K(\sigma, \tau)\, a(\tau) = \sum_{\sigma \in \mathbb{S}_n} \sum_{\tau \in \mathbb{S}_n} a(\sigma)\, \kappa(\sigma\tau^{-1})\, a(\tau) > 0$$

for any non-zero $a\colon \mathbb{S}_n \to \mathbb{C}$. Now by the Fourier inversion formula

$$Q = \sum_{\sigma \in \mathbb{S}_n} \sum_{\tau \in \mathbb{S}_n} a(\sigma)\, a(\tau)\, \frac{1}{n!} \sum_{\lambda \vdash n} d_\lambda \operatorname{tr}\left[\widehat{\kappa}(\lambda)\, \rho_\lambda(\tau\sigma^{-1})\right],$$

which, by $\rho_\lambda(\tau\sigma^{-1}) = \rho_\lambda(\tau)\, \rho_\lambda(\sigma^{-1})$ and $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, we can also write as

$$Q = \frac{1}{n!} \sum_{\lambda \vdash n} d_\lambda \operatorname{tr}\left[\left(\sum_{\tau \in \mathbb{S}_n} a(\tau) \rho_\lambda(\tau)\right) \widehat{\kappa}(\lambda) \left(\sum_{\sigma \in \mathbb{S}_n} a(\sigma) \rho_\lambda(\sigma^{-1})\right)\right].$$

Assuming, without loss of generality that the ρ_λ are all unitary representations,

$$Q = \frac{1}{n!} \sum_{\lambda \vdash n} d_\lambda \operatorname{tr}\left[\widehat{a}(\lambda)^\dagger \widehat{\kappa}(\lambda) \widehat{a}(\lambda)\right]. \tag{4}$$

Clearly, if each $\widehat{\kappa}(\lambda)$ is positive definite, then $Q > 0$. To prove the converse, note that since $\{\bigoplus_{\lambda \vdash n} \rho_\lambda(\sigma)\}_{\sigma \in \mathbb{S}_n}$ form a basis for $\bigoplus_{\lambda \vdash n} \mathbb{R}^{d_\lambda \times d_\lambda}$, the $\widehat{a}(\lambda)$ are general matrices (not all zero). Hence if $Q > 0$ for any a, then each $\widehat{\kappa}(\lambda)$ must be positive definite. To prove the positive semi–definite case, just replace $Q > 0$ with $Q \geq 0$. ■

Proof of Proposition 2. The regular representation of $\mathbb{S}_n$ is an $n!$-dimensional representation ρ_{reg} in which the rows/columns of the representation matrices are indexed by the $\sigma \in \mathbb{S}_n$ group elements themselves, and

$$[\rho_{\text{reg}}(\tau)]_{\sigma',\sigma} = \begin{cases} 1 & \text{if } \sigma' = \tau\sigma \\ 0 & \text{otherwise.} \end{cases}$$

By the theorem of total reducibility, we know that the regular representation can be decomposed as

$$\rho(\tau) = T^{-1} \Big[\bigoplus_{\lambda \vdash n} \bigoplus_{i=1}^{m_\lambda} \rho_\lambda(\tau) \Big] T$$

for some matrix T and multiplicities $(m_\lambda)_{\lambda \vdash n}$. In particular, if a matrix $Q \in \mathbb{C}^{\mathbb{S}_n \times \mathbb{S}_n}$ has the property that $Q_{\sigma',\sigma} = q(\sigma'\sigma^{-1})$ for some function q (in which case it is called an $\mathbb{S}_n$–circulant), then it can be written as

$$Q = \sum_{\tau \in \mathbb{S}_n} q(\tau)\, \rho_{\text{reg}}(\tau) = T^{-1} \Big[\bigoplus_{\lambda \vdash n} \bigoplus_{i=1}^{m_\lambda} \sum_{\tau \in \mathbb{S}_n} q(\tau) \rho_\lambda(\tau) \Big] T = T^{-1} \Big[\bigoplus_{\lambda \vdash n} \bigoplus_{i=1}^{m_\lambda} \widehat{q}(\lambda) \Big] T.$$

Conversely, any matrix of the form $R = T^{-1}\big[\bigoplus_{\lambda \vdash n} \bigoplus_{i=1}^{m_\lambda} \widehat{r}(\lambda)\big]\, T$ is necessarily an $\mathbb{S}_n$–circulant with $R_{\sigma',\sigma} = r(\sigma'\sigma^{-1})$. Now the graph Laplacian $\Delta_{\sigma',\sigma} = q(\sigma'\sigma^{-1})$ is an $\mathbb{S}_n$–circulant, and for any matrix M and invertible T, $\exp(T^{-1}MT) = T^{-1}\exp(M)\, T$, therefore

$$K = \exp(\beta\Delta) = \exp\Big(T^{-1} \Big[\bigoplus_{\lambda \vdash n} \bigoplus_{i=1}^{m_\lambda} \beta\, \widehat{q}(\lambda) \Big] T \Big) = T^{-1} \Big[\bigoplus_{\lambda \vdash n} \bigoplus_{i=1}^{m_\lambda} \exp\left(\beta\, \widehat{q}(\lambda)\right) \Big] T,$$

from which, letting $\widehat{\kappa}(\lambda) = \exp(\beta\, \widehat{q}(\lambda))$, we read off that K is an $\mathbb{S}_n$–circulant with $K_{\sigma',\sigma} = \kappa(\sigma'\sigma^{-1})$, i.e., $K(\sigma',\sigma) = K_{\sigma',\sigma}$ is right–invariant and $K(\sigma',\sigma) = \kappa(\sigma'\sigma^{-1})$. ■

Proof of Proposition 3. Since there are $\binom{n}{2}$ transpositions in total, q can be expressed as

$$q(\sigma) = u(\sigma) - \tbinom{n}{2}\delta_e(\sigma),$$

where $u(\sigma) = 1$ if σ is a transposition and 0 otherwise, and $\delta_e(e) = 1$ and 0 otherwise. By linearity $\widehat{q}(\lambda) = \widehat{u}(\lambda) - \binom{n}{2}\widehat{\delta}_e(\lambda) = \widehat{u}(\lambda) - \binom{n}{2} I$. Using the general result that each Fourier component of a class function is a multiple of the identity,

$$\widehat{u}(\lambda) = d_\lambda^{-1} \operatorname{tr}\left[\widehat{u}(\lambda)\right] I = \frac{1}{d_\lambda} \sum_{i<j} \operatorname{tr}\left(\rho_\lambda((i,j))\right)\, I = \binom{n}{2} \frac{\chi_\lambda((2,1,1,\ldots))}{d_\lambda}\, I.$$

By the formula on page 40 of [Diaconis, 1988], $\chi_\lambda((2,1,1,\ldots))/d_\lambda = \binom{n}{2}^{-1} \sum_i \binom{\lambda_i}{2} - \binom{\lambda'}{2}$, hence

$$\widehat{q}(\lambda) = \left(\left(\sum_i \binom{\lambda_i}{2} - \binom{\lambda'}{2} \right) - \binom{n}{2} \right) I$$

The result follows by Proposition 2. ■

Proof of Lemma 4. By (main text, eq. 7), $w(\sigma) = \sum_{\nu \in \mathbb{S}_n} u(\sigma\nu^{-1})\, v(\nu)$, so w is exactly the convolution $u * v$. The result follows by the convolution theorem. ■

Proof of Lemma 5. Let $u^-(\sigma) = u(\sigma^{-1})^*$. By the unitarity of the irreps $\widehat{u^-}(\lambda) = \widehat{u}(\lambda)^\dagger$. Now

$$(u^- * v)(e) = \sum_{\tau \in \mathbb{S}_n} u^-(\tau^{-1}) v(\tau) = \sum_{\tau \in \mathbb{S}_n} u(\tau)^* v(\tau) = \langle \boldsymbol{u}, \boldsymbol{v} \rangle,$$

and the result follows by plugging the convolution theorem into the inverse Fourier transform. ■

Proof of Proposition 7. Follows immediately from Proposition 9, since all tableaux t with $t \geq [\ldots]_{n-k}$ must have at least $n-k$ boxes in their first row. Also see the argument at the beginning of Section 5.1. ■

Proof of Proposition 8. See main text. Since in our computational model multiplying by scalars is free, the $\sum_{\lambda \in \Lambda^n_{n-\min(k,k')}} 2\, d_\lambda^3$ operation count is just the cost of computing the $\widehat{\kappa}(\lambda) \cdot \widehat{R}(\lambda)$ and $\widehat{R}'(\lambda) \cdot (\widehat{\kappa}(\lambda)\widehat{R}(\lambda))$ matrix products. ■

Proof of Proposition 9. For $\tau \in \mathbb{S}_{n-k}$, by the transitivity of $\leq$, recursively applying (main text, eq. 18) gives

$$[\rho_\lambda(\tau)]_{t,t'} = \begin{cases} [\rho_{\lambda^-}(\tau)]_{u,u'} & \text{if } u \leq t,\ u' \leq t', \text{ and } u \text{ and } u' \text{ are both of shape } \lambda^- \vdash n-k, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

In particular, if $f \colon \mathbb{S}_n \to \mathbb{C}$ is a function with support restricted to $\mathbb{S}_{n-k}$, and its restriction $f{\downarrow}_{\mathbb{S}_{n-m}} \colon \mathbb{S}_{n-k} \to \mathbb{C}$ is defined $f{\downarrow}_{\mathbb{S}_{n-k}}(\tau) = f(\tau)$, then

$$[\widehat{f}(\lambda)]_{t,t'} = \Big[\sum_{\tau \in \mathbb{S}_{n-m}} f(\tau)\, \rho_\lambda(\tau) \Big]_{t,t'} =$$
$$\begin{cases} \big[\widehat{f{\downarrow}_{\mathbb{S}_{n-k}}}(\lambda^-)\big]_{u,u'} & \text{if } u \leq t,\ u' \leq t', \text{ and } u \text{ and } u' \text{ are both of shape } \lambda^- \vdash n-k, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

The inverse Fourier transform for a function $g \colon \mathbb{S}_{n-k} \to \mathbb{C}$ is

$$g(\tau) = \frac{1}{(n-k)!} \sum_{\lambda^- \vdash n-k} d_{\lambda^-} \operatorname{tr}\big[\widehat{g}(\lambda^-)\, \rho_{\lambda^-}(\tau^{-1})\big].$$

If g is the identity function $g(\tau) = 1$, considering that $\rho_{(n-k)}$ is the trivial representation $\rho_{(n-k)}(\tau) \equiv 1$, we can read off that $\widehat{g}(\lambda^-) = 0$ for all λ^-, except for $\lambda^- = (n-k)$, in which case $\widehat{g}((n-k)) = (n-k)!$.

By defintion, $\mathbf{S}_{n-k}$, regarded as a function on $\mathbb{S}_n$, is a function with support restriced to $\mathbb{S}_{n-k}$, and its restriction to $\mathbb{S}_{n-k}$ is exactly g, so by plugging into (6) and considering that the only SYT of shape $(n-k)$ is $[\ldots]_{n-k}$ we get the desired result. ■

Proof of Proposition 10. Viewed as a map $\boldsymbol{v} \colon \boldsymbol{u} \mapsto \boldsymbol{v}\boldsymbol{u}$, the group algebra vector $\boldsymbol{v}$ can be regarded as a linear operator on $\mathbb{C}[\mathbb{S}_n]$. The existence of the adjoint then follows from general linear algebra. Now by Lemmas 4 and 5,

$$\langle \boldsymbol{u}, \boldsymbol{v}\boldsymbol{w} \rangle = \frac{1}{n!} \sum_{\lambda \vdash n} d_\lambda \operatorname{tr}[\widehat{u}(\lambda)^\dagger \widehat{v}(\lambda)\, \widehat{u}(\lambda)] \qquad \text{and}$$
$$\langle \boldsymbol{v}^\dagger \boldsymbol{u}, \boldsymbol{w} \rangle = \frac{1}{n!} \sum_{\lambda \vdash n} d_\lambda \operatorname{tr}\big[\big(\widehat{v}^\dagger(\lambda)\, \widehat{u}(\lambda)\big)^\dagger\, \widehat{u}(\lambda) \big] = \frac{1}{n!} \sum_{\lambda \vdash n} d_\lambda \operatorname{tr}\big[\widehat{u}(\lambda)^\dagger (\widehat{v^\dagger}(\lambda))^\dagger\, \widehat{u}(\lambda) \big],$$

so for these two expressions to be equal for any $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{w}$, we must have $\widehat{v^\dagger}(\lambda) = (\widehat{v}(\lambda))^\dagger$. Note that in these expressions † is overloaded to serve two different roles: to signify the group algebra adjoint and the Hermitian conjugate. Our result shows that the two roles, in fact, match. ■