ALFRED YOUNG'S CONSTRUCTION OF THE IRREDUCIBLE REPRESENTATIONS OF S_n

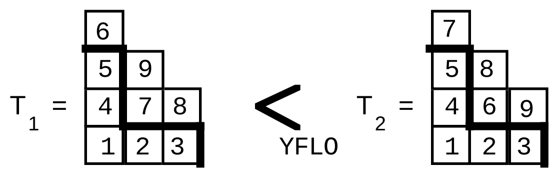
1. Basic Definitions

If λ is a partition of n (in symbols $\lambda \vdash n$), a filling of the Ferrers' diagram of λ by the integers $1, 2, ..., n$ will be referred to as an "injective" tableau of shape λ. The collection of all these tableaux will be denoted by INJ(λ). If $T \in$ INJ(λ) and the entries of T increase from left to right in the rows and from bottom to top on the columns then T is said to be standard. The collection of all these tableaux will be denoted by ST(λ).

Given $T_1, T_2 \in ST(\lambda)$ we shall say that T_1 precedes T_2 in the Young "first letter order" and write "

$$T_1 \;<_{YFLO}\; T_2$$

if the first entry of disagreement between T_1 and T_2 is higher in T_1 than in T_2. For instance we have



Here T_1 and T_2 agree in the positions of $1, 2, 3, 4, 5$. The first letter of disagreement is 6 and it is "higher" in T_1 than in T_2. The positions of the remaining letters do not matter.

Note that if $T_1 <_{YFLO} T_2$ and a is the first letter of disagreement, then there is a letter b such that $a \& b$ are in the same column of T_1 and in the same row of T_2. Note that $b = 4$ in the example given above. In general b is the letter which lies at the intersection of the column of the shape λ in which a lies in T_1 with the row in which a lies in T_2.

Given an injective tableau T with n squares a permutation

$$\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$$

is made to act on T by replacing entry i by σ_i. The resulting tableau is denoted by σT. We say that σ is in the "row group" of T if the rows of T and σT differ only by the order of their entries. The "column group" of T is analogously defined. These two groups will be denoted by $R(T)$ and $C(T)$ respectively. We shall also let

$$P(T) = \sum_{\alpha \in R(T)} \alpha \quad , \quad N(T) = \sum_{\beta \in C(T)} sign(\beta)\beta \qquad 1.1$$

If $T_1, T_2 \in INJ(\lambda)$ by $\sigma_{T_1T_2}$ we denote the permutation that sends T_2 into T_1 that is

$$T_1 \;=\; \sigma_{T_1T_2} T_2$$

This given, it is easy to see that we have

$$\begin{aligned} &a) \quad P(\sigma T) = \sigma P(T)\,\sigma^{-1}\ ;\ N(\sigma T) = \sigma N(T)\sigma^{-1} \qquad \forall\ \sigma \in S_n \\ &b) \quad \sigma_{T_1T_2} P(T_2) \ = \ P(T_1)\sigma_{T_1T_2} \\ &c) \quad \sigma_{T_1T_2} N(T_2) \ = \ N(T_1)\sigma_{T_1T_2} \end{aligned} \tag{1.2}$$

We shall also use the shorthand notations

$$E_{T_1T_2} = N(T_1)\,\sigma_{T_1T_2}\,P(T_2) \quad , \quad E_T = E_{TT} \tag{1.3}$$

Note that from 1.2 a) and b) we derive that

$$E_{T_1T_2} \ = \ E_{T_1}\sigma_{T_1T_2} \ = \ \sigma_{T_1T_2}E_{T_2} \tag{1.4}$$

For a given $\lambda \vdash n$ we let

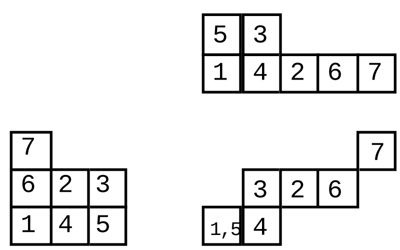
$$S_1^\lambda\ ,\ S_2^\lambda\ ,\ \ldots\ ,\ S_{f_\lambda}^\lambda \tag{1.5}$$

denote the standard tableaux of shape λ in Young's first letter order. It will be convenient to set

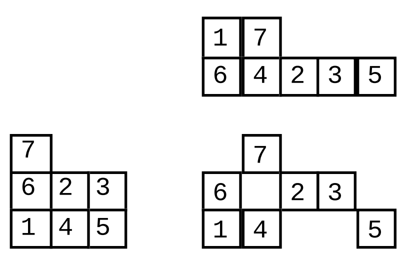
$$E_i^\lambda \ = \ E_{S_i^\lambda} \quad ; \quad E_{ij}^\lambda \ = \ E_{S_i^\lambda S_j^\lambda} \quad ; \quad \sigma_{ij}^\lambda \ = \ \sigma_{S_i^\lambda S_j^\lambda} \quad ; \tag{1.6}$$

The superscript λ may be omitted when dealing with a fixed shape λ.

Given two tableaux T_1, T_2 not necessarily of the same shape we let $T_1 \wedge T_2$ be the diagram obtained by placing in the cell (i,j) the intersection of row i of T_1 with column j of T_2. The diagram below gives an example of this construction



Note that in this case there is a cell which contains more than one entry. When this happens we say that the pair $T_1 \wedge T_2$ is *bad*. Note that in this next example each cell has at most one entry.



When this happens we say that $T_1 \wedge T_2$ is *good*. Note that in the good case if $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ is the shape of T_1 and $\mu = (\mu_1, \mu_2, \ldots, \mu_k)$ is the shape of T_2 then for each i

$$\lambda_1 + \lambda_2 + \cdots + \lambda_i \;\; \leq \;\; \mu_1 + \mu_2 + \cdots + \mu_i \tag{1.7}$$

For the left hand side gives the number of entries in the first i rows of $T_1 \wedge T_2$ and the right hand side gives the number of entries in the diagram obtained by letting the entries of $T_1 \wedge T_2$ drop down their column until the diagram is packed tight. Note that strict inequality can and must hold true only if T_1 and T_2 have different shapes. This means that when T_1 and T_2 have the same shape then $T_1 \wedge T_2$ is good if and only if $T_1 \wedge T_2$ is also a tableau of the same shape and there are permutations $\alpha_1 \in R(T_1)$ and $\beta_2 \in C(T_2)$ such that

$$T_1 \wedge T_2 \;\; = \;\; \alpha_1 T_1 \quad ; \quad T_2 \;\; = \;\; \beta_2 \, T_1 \wedge T_2$$

The following result is easily established

Proposition 1.1

For a pair of tableaux of the same shape the following properties are equivalent

$$\begin{array}{ll} a) & T_1 \wedge T_2 \quad is \;\; good \\ b) & \exists \; \alpha_1 \in R(T_1) \, , \, \exists \; \beta_2 \in C(T_2) \quad such \;\; that \quad T_2 = \beta_2 \alpha_1 \, T_1 \\ c) & \exists \; \alpha_1 \in R(T_1) \, , \, \exists \; \beta_1 \in C(T_1) \quad such \;\; that \quad T_2 = \alpha_1 \beta_1 \, T_1 \\ d) & \exists \; \alpha_2 \in R(T_2) \, , \, \exists \; \beta_2 \in C(T_2) \quad such \;\; that \quad T_2 = \alpha_2 \beta_2 \, T_1 \end{array}$$

Moreover we have

$$sign(\beta_1) \;\; = \;\; sign(\beta_2) \tag{1.8}$$

Proof

We have seen that a) and b) are equivalent. From 1.2 a) we then get that b) implies

$$N(T_1) \;\; = \;\; \alpha_1^{-1} \beta_2^{-1} \, N(T_2) \, \beta_2 \alpha_1 \;\; = \;\; \alpha_1^{-1} \, N(T_2) \, \alpha_1$$

Thus there is a $\beta_1 \in N(T_1)$ such that

$$\beta_1 \;\; = \;\; \alpha_1^{-1} \beta_2 \alpha_1 \tag{1.9}$$

This gives that

$$T_2 \;\; = \;\; \beta_2 \alpha_1 \, T_1 \;\; = \;\; (\alpha_1 \beta_1 \alpha_1^{-1}) \alpha_1 \, T_1$$

which is c). Note that 1.9 implies that $sign(\beta_1) = sign(\beta_2)$.

We leave it to the reader to show that d) implies b) and complete the proof of the proposition.

Definition 1.1

For a pair of tableaux $T_1, T_2 \in INJ(\lambda)$ we shall set

$$C(T_1,T_2) \;=\; \begin{cases} 0 & \text{if } T_1 \wedge T_2 \text{ is bad} \\ sign(\beta_1) = sign(\beta_2) & \text{if } T_1 \wedge T_2 \text{ is good} \end{cases}$$

This given, for a shape λ and a permutation $\sigma \in S_n$ we shall set

$$C^\lambda(\sigma) \;=\; \| C(S_i^\lambda \;,\; \sigma S_j^\lambda) \|_{i'j=1}^{f_\lambda}$$

We have the following basic fact

Proposition 1.2

The matrix $C^\lambda(\epsilon)$ is upper triangular and therefore invertible over the integers.

Proof

We have seen that if $T_1 <_{YFLO} T_2$ then we have two entries a, b that lie in the same column of T_1 and in the same row of T_2 this means that $T_2 \wedge T_1$ is bad. So $T_1 <_{YFLO} T_2$ implies $C(T_2, T_1) = 0$. Thus from the definition of first letter order we derive that

$$C(S_i^\lambda, S_j^\lambda) \;=\; 0 \qquad\qquad \forall\; i > j \;\; .$$

This proves the proposition since $C(S_i^\lambda, S_i^\lambda) = 1$ holds true trivially.

This allows us to set

$$A^\lambda(\sigma) \;=\; C^\lambda(\epsilon)^{-1}\, C^\lambda(\sigma) \qquad 1.11$$

Our goal in the next two sections is to show that the collection of matrix functions $\{\sigma \to A^\lambda(\sigma)\}_{\lambda \vdash n}$ gives a complete set of irreducible representations of S_n.

2. Young's Natural Representation.

We shall start by stating the basic.

Lemma 2.1 (Young's straightening algorithm)

For any $T \in INJ(\lambda)$ we have coefficients $c_1(T), c_2(T), \ldots, c_{f_\lambda}(T)$ such that

$$E_T \;=\; \sum_{i=1}^{f_\lambda} c_i(T)\; E_{S_i^\lambda T} \qquad 2.1$$

We refer the reader to Garsia-Wachs (J.C.T. (A) **50** 47-81) for a proof.

It develops that once 2.1 is known to be possible, the coefficients $c_i(T)$ can be given closed form expressions which facilitate their computation.

The crucial tool is given by the following

Proposition 2.1

For any $T_1, T_2 \in INJ(\lambda)$ and $\sigma \in S_n$

$$E_{T_2 T_1}\,|_\sigma \;=\; C(\sigma T_1, T_2) \qquad 2.2$$

Proof

$$E_{T_2T_1}\,|_\sigma \;=\; \sum_{\beta_2\in C(T_2)}\sum_{\alpha_2\in C(T_2)} \beta_2\alpha_2\;\sigma_{T_1}\,|_\sigma$$

However,

$$\beta_2\alpha_2\;\sigma_{T_2T_1}=\sigma \quad\rightarrow\quad \beta_2\alpha_2=\sigma\;\sigma_{T_1T_2}=\sigma_{\sigma T_1,T_2}$$

or

$$\sigma_{T_2,\sigma T_1} \;=\; \alpha_2^{-1}\beta_2^{-1}$$

which combined with part d) of Proposition 2.1 gives that $\sigma T_1\wedge T_2$ is good. Thus from 1.8 we derive that

$$E_{T_2T_1}\,|_\sigma \;=\; C(\sigma T_1,T_2)$$

as desired.

Q.E.D.

We should note that in particular we must have

$$E_{T_2T_1}\,|_\epsilon \;=\; sign(\beta_2) \;=\; C(T_1,T_2) \qquad 2.3$$

Proposition 2.2

If $c(T)$ denotes the column vector of the coefficients $c_i(T)$ in Young's formula 2.1 and $c(S,T)$ denotes the column vector of coefficients $C(S_j^\lambda,T)$ then

$$c(T) \;=\; C^\lambda(\epsilon)^{-1}\;c(S,T) \qquad 2.4$$

Proof.

Multiplying both sides of 2.1 by σ_{TSj} gives (omitting the λ-superscripts):

$$E_{TS_j} \;=\; \sum_{i=1}^{f_\lambda} E_{S_iS_j}\;\;.$$

Taking coefficients of the identity in both sides, and using 2.3, we get

$$C(S_j,T) \;=\; \sum_{i=1}^{f_\lambda} c_i(T)\;C(S_j,S_i)$$

which, in matrix notation, may be expressed as

$$c(S,T) \;=\; C^\lambda(\epsilon)\;c(T)\;\;,$$

and formula 2.4 follows upon left multiplication of both sides of this identity by $C^\lambda(\epsilon)^{-1}$.

Theorem 2.1

For any tableau $S^* \in INJ(\lambda)$ the group algebra elements

$$E_{S_1S^*}\ ,\ E_{S_2S^*}\ ,\ \ldots\ ,\ E_{S_{f_\lambda}S^*} \tag{2.5}$$

are independent and the action of a $\sigma \in S_n$ on their linear span

$$\mathcal{L}[E_{S_1S^*}\ ,\ E_{S_2S^*}\ ,\ \ldots\ ,\ E_{S_{f_\lambda}S^*}]$$

is given by the matrix $A^\lambda(\sigma)$ defined in 1.11. In other words we have

$$\sigma \,\langle\, E_{S_1S^*}, \ldots, E_{S_{f_\lambda}S^*}\,\rangle \;=\; \langle\, E_{S_1S^*}, \ldots, E_{S_{f_\lambda}S^*}\,\rangle\; A^\lambda(\sigma) \tag{2.6}$$

Proof

Note that the equality

$$c_1\, E_{S_1S^*} + c_2\, E_{S_1S^*} \;+\; \cdots \;+\; c_{f_\lambda}\, E_{S_{f_\lambda}S^*} \;=\; 0$$

hit on the right by $\sigma S^* S_j$ gives

$$c_1\, E_{S_1S_j} + c_2\, E_{S_1S_j} \;+\; \cdots \;+\; c_{f_\lambda}\, E_{S_{f_\lambda}S_j} \;=\; 0\ .$$

Taking the coefficient of the identity permutation and using 2.3 we get

$$c_1\, C(S_j,S_1) + c_2\, C(S_j,S_2) \;+\; \cdots \;+\; c_{f_\lambda}\, C(S_j,S_{f_\lambda}) \;=\; 0 \qquad (\ \forall \quad j=1,2,\ldots,f_\lambda\ )$$

which, in matrix notation, may be rewritten as

$$C^\lambda(\epsilon)\ \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{f_\lambda} \end{pmatrix} \;=\; 0\ .$$

Thus the invertibility of $C^\lambda(\epsilon)$ forces

$$c_1 = c_2 = \cdots = c_{f_\lambda} \;=\; 0\ .$$

This proves the independence. To complete the proof we simply note that Young's formula 2.1 with $T = \sigma S_j$ gives

$$E_{\sigma S_j} \;=\; \sum_{i=1}^{f_\lambda} c_i(\sigma S_j)\, E_{S_i \sigma S_j}\ . \tag{2.7}$$

But

$$E_{\sigma S_j} \;=\; N(\sigma S_j) P(\sigma S_j) \;=\; \sigma\, N(S_j) P(S_j) \sigma^{-1} \;=\; \sigma\, E_{S_j}\, \sigma^{-1}$$

Thus

$$\sigma\, E_{S_j} \;=\; E_{\sigma S_j}\, \sigma$$

and we may rewrite 2.7 in the form

$$\sigma\, E_{S_j} \;=\; \sum_{i=1}^{f_\lambda} c_i(\sigma S_j)\, E_{S_i \sigma S_j}\, \sigma \;=\; \sum_{i=1}^{f_\lambda} E_{S_i S_j}$$

Thus for any $S^* \in INJ(\lambda)$ we have

$$\sigma\, E_{S_j S^*} \;=\; \sum_{i=1}^{f_\lambda} c_i(\sigma S_j)\, E_{S_i S^*} \qquad 2.8$$

and we see that the coefficients here do not depend on the particular tableau S^*. To find out what they are, we simply equate the coefficients of ϵ in both sides of 2.8 with $S^* = S_k$ and get

$$E_{S_j S_k}\,|_{\sigma^{-1}} \;=\; \sum_{i=1}^{f_\lambda} c_i(\sigma S_J)\, E_{S_i S_k}\,|_\epsilon \;\;.$$

Using 2.2 and 2.3, this may be rewritten as

$$C(\sigma^{-1} S_k, S_j) \;=\; \sum_{i=1}^{f_\lambda} c_i(\sigma S_J)\, C(S_k, S_i) \;\;.$$

Since for any pair $T_1, T_2 \in INJ(\lambda)$ and any $\sigma \in S_n$ we have $C(\sigma T_1, \sigma T_2) = C(T_1, T_2)$ we can rewrite this last relation in matrix form as

$$C^\lambda(\sigma) \;=\; C^\lambda(\epsilon)\, \| c_i(\sigma S_j) \|_{i,i=1}^{f_\lambda} \;\;. \qquad 2.9$$

Thus we must have

$$\| c_i(\sigma S_j) \|_{i,i=1}^{f_\lambda} \;=\; C^\lambda(\epsilon)^{-1}\, C^\lambda(\sigma) \;=\; A^\lambda(\sigma) \;\;.$$

This completes the proof.

We have seen then that the map $\sigma \to A^\lambda(\sigma)$ is indeed a representation of S_n. In the next section we show that it is irreducible.

3. Characters of left ideals in a Group Algebra

Let G be a finite group and let $\mathcal{A}(G)$. denote its group algebra. Given an element $\phi \in \mathcal{A}(G)$ the subspace spanned by the elements $\gamma\phi$ as γ varies in G will be denoted here by $\mathcal{A}(G)\phi$ and referred to as *left ideal* generated by ϕ in $\mathcal{A}(G)$. In symbols

$$\mathcal{A}(G)\phi \;=\; \mathcal{L}[\gamma\phi \;:\; \gamma \in G\,]$$

It clear from the definition that $\forall \tau \in G$ we have

$$\psi \in \mathcal{A}(G)\phi \;\longrightarrow\; \tau\psi \in \mathcal{A}(G)\phi \;\;.$$

In other words $\mathcal{A}(G)$ is invariant under the action of G induced by left multiplication. Our goal here is to derive a formula for the character of the representation of G corresponding to this action. To state our result we need to introduce some notation and recall some definitions. Let Λ be an index set for the irreducible representations of G and let $\{A^\lambda\}_{\lambda\in\Lambda}$ be a complete set of irreducible unitary representations of G. For a given $\phi \in \mathcal{A}(G)$ we set

$$A^\lambda(\phi) \;=\; \sum_{\gamma\in G} \phi(\gamma)\overline{A}^\lambda(\gamma) \; . \tag{3.1}$$

We may call this matrix the *Fourier coefficient* of ϕ corresponding to λ. We sall also denote by $\mathcal{F}\,(\phi)$ the block diagonal matrix with blocks $\{\overline{A}^\lambda(\phi)\}_{\lambda\in\Lambda}$. In symbols

$$\mathcal{F}\,(\phi) \;=\; \bigoplus_{\lambda\in\Lambda} \overline{A}^\lambda(\phi) \; . \tag{3.2}$$

We may refer to $\mathcal{F}\,\phi$ as the *Fourier transform* of ϕ. Now we know that if n_λ is the dimension of A^λ then the map

$$\phi \quad\longrightarrow\quad \mathcal{F}\,(\phi)$$

is an isomorphism of the group algebra $\mathcal{A}(G)$ onto the direct sum of matrix algebras

$$\mathcal{M} \;=\; \bigoplus_{\lambda\in\Lambda} \mathcal{M}_{n_\lambda\times n_\lambda} \tag{3.3}$$

where $\mathcal{M}_{n_\lambda\times n_\lambda}$ here denotes the algebra of $n_\lambda \times n_\lambda$ matrices. This given, we may use the Fourier transform to translate any algebraic question about $\mathcal{A}(G)$ into a question about $\mathcal{M}$. To answer such questions we need to familiarize ourselves with algebraic properties of $\mathcal{M}$. For a given $\lambda \in \Lambda$ let E^λ_{ij} denote the $n_\lambda \times n_\lambda$ matrix with i,j entry equal to 1 and all other entries equal to zero. We see that left multiplication of E^λ_{ij} by an arbitrary matrix $A = \|a_{ij}\|_{i,j=1}^{n_\lambda}$ gives

$$A\, E^\lambda_{ij} \;=\; a_{1i}E^\lambda_{1j} + a_{2i}E^\lambda_{2j} + \cdots + a_{n_\lambda i}E_{n_\lambda j} \; . \tag{3.4}$$

This simple identity has several fundamental implications. To begin with we see that the linear span

$$\mathcal{E}^\lambda_j \;=\; \mathcal{L}[E^\lambda_{1j}, E^\lambda_{2j}, \ldots, E^\lambda_{n_\lambda j}]$$

is an irreducible invariant subspace of $\mathcal{M}_{n_\lambda\times n_\lambda}$ under left multiplication. This observation immediately translates into the following beautiful result.

Proposition 3.1

For a $\lambda \in \Lambda$ let e^λ_{ij} be the element of $\mathcal{A}(G)$ with Fourier coefficients

$$\overline{A}^\mu(e^\lambda_{ij}) \;=\; \begin{cases} 0 & \textit{if} \quad \mu \neq \lambda \\ E^\lambda_{ij} & \textit{if} \quad \mu = \lambda \end{cases} \qquad\qquad \mu\in\Lambda \tag{3.5}$$

then for any $j \in [1, n_\lambda]$ the (left multiplication) action of G on the linear span

$$\mathbf{V}^\lambda_j \;=\; \mathcal{L}[e^\lambda_{1j}, e_{2j}, \ldots, e^\lambda_{n_\lambda j}]$$

induces the irreducible representation A^λ.

Proof

We only need to observe that using the map $\phi \longrightarrow \mathcal{F}(\phi)$ we can translate the identities in 3.4 into the matrix identities

$$\langle\, \gamma e^{\lambda}_{1j}, \gamma e^{\lambda}_{2j}, \ldots, \gamma e^{\lambda}_{n_\lambda j} \,\rangle \;=\; \langle\, e^{\lambda}_{1j}, e^{\lambda}_{2j}, \ldots, e^{\lambda}_{n_\lambda j} \,\rangle \;\; A^{\lambda}(\gamma) \qquad \forall \quad j \in [1, n_\lambda] \tag{3.6}$$

Remark 3.1

Let us recall that from Schur's lemma it follows that the entries a^{λ}_{ij} in a complete set of irreducible unitary representations of a finite group G satisfy the orthogonality relations

$$\sum_{\gamma\in G} a^{\lambda}_{ij}(\gamma)\, \overline{a}^{\mu}_{rs}(\gamma) \;=\; \begin{cases} 1/n_\lambda & \text{if } \lambda=\mu,\ i=r \text{ and } j=s \\ 0 & \text{otherwise} \end{cases}$$

This implies that every element $f \in \mathcal{A}(G)$ may be recovered from its Fourier coefficients $\overline{A}^{\lambda}(f) = \|\overline{a}^{\lambda}_{ij}(f)\|^{n_\lambda}_{i,j=1}$ by means of the formula

$$f \;=\; \sum_{\lambda\in\Lambda} n_\lambda \sum_{i,j=1}^{n_\lambda} \overline{a}^{\lambda}_{ij}(f)\, a^{\lambda}_{ij} \;\; ,$$

which may be viewed as giving the inverse of the Fourier transform. In particular we see that the elements e^{λ}_{ij} defined in 3.5 are none other than multiples of the corresponding matrix entries. More precisely this formula gives

$$e^{\lambda}_{ij} \;=\; n_\lambda\, a^{\lambda}_{ij}$$

Remark 3.2

It is also important to note at this point that the elements e^{λ}_{ij} are related by the following identities

$$e^{\lambda}_{ij}\;\; e^{\mu}_{rs} \;=\; \begin{cases} e^{\lambda}_{is} & \text{if } \mu=\lambda \text{ and } j=s \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

To see this we need only verify that their Fourier transforms are related by the same identities. For these reasons a collection of elements of a group algebra satisfying these relations are sometimes referred to as *matrix units*. Remarkably, A. Young, was able to write down explicitely a complete set of matrix units for any symmetric group S_n. We shall see in the next sections how Young units may be expressed in terms of the elements $E_{T_1T_2}$ introduced in the last section.

To proceed we need to recall two basic results from linear algebra.

Proposition 3.2

If A is any $n\times n$ matrix of rank $r \le n$ and $s = n-r$ then we can find a pair of invertible matrices P and Q such that

$$P\; A\; Q \;=\; \begin{pmatrix} I_{rr} & O_{rs} \\ O_{sr} & O_{ss} \end{pmatrix} \tag{3.8}$$

where I_{rr} is the $r\times r$ identity matrix and O_{rs} , O_{sr} , O_{ss} are respectively$r\times s$, $s\times r$ and $s\times s$ matrices with all entries equal to zero.

Proof

By right multiplication by a suitable permutation matrix π we can permute the columns of A so that the first r columns of the resulting matrix $A' = A\pi$ span the column space of A'. This done we can express each of the remaining columns of A' as a linear combination of the first r columns. Using these expressions we can put together a non singular matrix M such that $B = A'M$ is of the form

$$B \;=\; (\, A'_{nr} \quad O_{ns}\,)$$

where A'_{nr} is the matrix consisting of the first r columns of A' and O_{ns} is an $n \times s$ matrix of zeros. This gives

$$A\pi M \;=\; (\, A'_{nr} \quad O_{ns}\,)$$

Working on B^T (*) in the same way as we did with A we can construct a permutation matrix τ and a nonsigular matrix N such that $C = B\tau N$ is of the form

$$C \;=\; \begin{pmatrix} C_{rr} & O_{rs} \\ O_{sr} & O_{ss} \end{pmatrix}$$

where C_{rr} is a non singular $r \times r$ matrix. This gives

$$N^T\, \tau^T\, A\, \pi\, M \;=\; \begin{pmatrix} C_{rr}^T & O_{rs} \\ O_{sr} & O_{ss} \end{pmatrix}$$

Note that if D_{rr} denotes the inverse of the matrix C_{rr}^T then

$$\begin{pmatrix} D_{rr} & O_{rs} \\ O_{sr} & O_{ss} \end{pmatrix} \begin{pmatrix} C_{rr}^T & O_{rs} \\ O_{sr} & O_{ss} \end{pmatrix} \;=\; \begin{pmatrix} I_{rr} & O_{rs} \\ O_{sr} & O_{ss} \end{pmatrix} \;.$$

This gives

$$\begin{pmatrix} D_{rr} & O_{rs} \\ O_{sr} & O_{ss} \end{pmatrix} N^T\, \tau^T\, A\, \pi\, M \;=\; \begin{pmatrix} I_{rr} & O_{rs} \\ O_{sr} & O_{ss} \end{pmatrix} \;,$$

which is 3.8 with

$$P \;=\; \begin{pmatrix} D_{rr} & O_{rs} \\ O_{sr} & O_{ss} \end{pmatrix} N^T\, \tau^T \qquad \text{and} \qquad Q \;=\; \pi\, M \;.$$

Proposition 3.3

A matrix is diagonable if and only if it satisfies a polynomial equation with simple roots

Proof

Even though this fact is an immediate consequence of the Jordan normal form result, we should point out that it has a simple direct proof. To this end note that, if $c_1, c_2, \ldots, c_k$ and $u_1, u_2, \ldots, u_k$ are complex numbers, and $u_1, u_2, \ldots, u_k$ are distinct then the unique polynomial $Q(x)$ of degree $\leq k$ that satisfies

$$Q(u_i) \;=\; c_i \qquad\qquad (\, i = 1, 2, \ldots, k\,)$$

(*) The superscript T denotes transposition is given by the *Lagrange interpolation formula*:

$$Q(x) \;=\; \sum_{i=1}^{k} c_i \, \frac{\prod_{\substack{j=1 \\ j\neq i}}^{k} (x-u_j)}{\prod_{\substack{j=1 \\ j\neq i}}^{k} (u_i-u_j)} \tag{3.9}$$

In particular we must have

$$1 \;=\; \sum_{i=1}^{k} \frac{\prod_{\substack{j=1 \\ j\neq i}}^{k} (x-u_j)}{\prod_{\substack{j=1 \\ j\neq i}}^{k} (u_i-u_j)}$$

This gives that for any $n\times n$ matrix we have

$$I_{nn} \;=\; \sum_{i=1}^{k} \frac{\prod_{\substack{j=1 \\ j\neq i}}^{k} (A-u_j)}{\prod_{\substack{j=1 \\ j\neq i}}^{k} (u_i-u_j)}$$

where I_{nn} denotes the $n\times n$ identity matrix. In particular, for any n-dimensional vector v we have

$$\mathbf{v} \;=\; \sum_{i=1}^{k} \mathbf{v}_i \tag{3.10}$$

with

$$\mathbf{v}_i \;=\; \frac{\prod_{\substack{j=1 \\ j\neq i}}^{k} (A-u_j)}{\prod_{\substack{j=1 \\ j\neq i}}^{k} (u_i-u_j)} \; \mathbf{v} \qquad (\; for \quad i=1,2,\ldots,k \;)$$

Suppose now that A satisfies the polynomial equation $P(A)=0$ with

$$P(x) \;=\; (x-u_1)(x-u_2)\cdots(x-u_k)$$

We must then have

$$A \prod_{\substack{j=1 \\ j\neq i}}^{k} (A-u_j) \;=\; u_i \prod_{\substack{j=1 \\ j\neq i}}^{k} (A-u_j)$$

in particular we get that

$$A\;\mathbf{v}_i \;=\; u_i\;\mathbf{v}_i\;.$$

In other words, 3.10 shows that any n-vector $\mathbf{v}$ has an expansion in terms of eigenvectors of A. The latter result is easily seen to be equivalent to the diagonability of A. This proves the sufficiency of the condition. To complete the proof we need to show necessity. However this is easily verified.

We are finally in a position to state and prove the basic result of this section

Theorem 3.1

Let χ^ϕ denote the character of the representation resulting from the left multiplication action of G on the left ideal $\mathcal{A}(G)\phi$, and let χ^λ denote the character of A^λ. Then in full generality we have

$$\chi^\phi \;=\; \sum_{\lambda\in\Lambda} r_\lambda(\phi)\,\chi^\lambda \tag{3.11}$$

where we have set

$$r_\lambda(\phi) \;=\; rank\;\; \overline{A}^\lambda(\phi)\;\; . \tag{3.12}$$

In particular, if ϕ is an idempotent element of $\mathcal{A}(G)$ this identity can be rewritten in the following beautiful form due to A. Young:

$$\chi^\phi \;=\; \sum_{\gamma\in G} \gamma\,\phi\,\gamma^{-1} \tag{3.13}$$

Proof

Note that from Proposition 3.2 we immediately deduce that we can find invertible elements p and q of the group algebra $\mathcal{A}(G)$ giving

$$p\,\phi\,q \;=\; \sum_\lambda \sum_{j=1}^{r_\lambda} e^\lambda_{jj}\;\; . \qquad\qquad (\, r_\lambda = r_\lambda(\phi)\,) \tag{3.14}$$

Observe now that formula 3.4 yields also the identity

$$A(E^\lambda_{11}+E^\lambda_{22}+\cdots+E^\lambda_{r_\lambda r_\lambda}) \;=\; \langle\, E^\lambda_{11}, E^\lambda_{21},\ldots,E^\lambda_{n_\lambda 1}\,\rangle \begin{pmatrix} a_{11}\\ a_{21}\\ \vdots \\ a_{n1}\end{pmatrix} \;+$$

$$+\;\; \langle\, E^\lambda_{12}, E^\lambda_{22},\ldots,E^\lambda_{n_\lambda 2}\,\rangle \begin{pmatrix} a_{12}\\ a_{22}\\ \vdots \\ a_{n2}\end{pmatrix} \;+\;\cdots\;+$$

$$+\;\; \langle\, E^\lambda_{1n_\lambda}, E^\lambda_{2n_\lambda},\ldots,E^\lambda_{n_\lambda n_\lambda}\,\rangle \begin{pmatrix} a_{1n_\lambda}\\ a_{2n_\lambda}\\ \vdots \\ a_{n_\lambda n_\lambda}\end{pmatrix}\;\; .$$

Note that using the Fourier transform we can construct an element $f\in\mathcal{A}(G)$ all of whose Fourier coefficients $A^\mu(f)$ for $\mu\neq\lambda$ vanish and $A^\lambda(f)$ equal to an arbitrary $n_\lambda\times n_\lambda$ matrix $A=\|a_{ij}\|$. For such an element we shall have

$$
\begin{aligned}
f\,(e^\lambda_{11}+e^\lambda_{22}+\cdots+e^\lambda_{n_\lambda n_\lambda}) \;&=\; \langle\, e^\lambda_{11}, e^\lambda_{21},\ldots,e^\lambda_{n_\lambda 1}\,\rangle \begin{pmatrix} a_{11}\\ a_{21}\\ \vdots \\ a_{n1}\end{pmatrix} \;+\\
&+\;\langle\, e^\lambda_{12}, e^\lambda_{22},\ldots,e^\lambda_{n_\lambda 2}\,\rangle \begin{pmatrix} a_{12}\\ a_{22}\\ \vdots \\ a_{n2}\end{pmatrix} \;+\;\cdots\;+\\
&+\;\langle\, e^\lambda_{1n_\lambda}, e^\lambda_{2n_\lambda},\ldots,e^\lambda_{n_\lambda n_\lambda}\,\rangle \begin{pmatrix} a_{1n_\lambda}\\ a_{2n_\lambda}\\ \vdots \\ a_{n_\lambda n_\lambda}\end{pmatrix}\;.
\end{aligned}
\qquad 3.14
$$

This shows that the left ideal, in the group algebra $\mathcal{A}(G)$, generated by the element

$$e^\lambda_{11}+e^\lambda_{22}+\cdots+e^\lambda_{n_\lambda r_\lambda}$$

decomposes into the direct sum of the linear spans

$$\mathbf{V}^\lambda_j \;=\; \mathcal{L}[\,e^\lambda_{1j}, e^\lambda_{2j},\ldots,e^\lambda_{n_\lambda j}\,] \qquad (\,i=1,2,\ldots,r_\lambda\,)\;. \qquad 3.15$$

Clearly we can do this for every $\lambda\in\Lambda$, and since the elements e^λ_{ij} are independent, these observations yield us the direct sum decomposition

$$\mathcal{A}(G)\,p\,\phi\,q \;=\; \bigoplus_{\lambda\in\Lambda}\bigoplus_{j=1}^{r_\lambda(\phi)} \mathbf{V}^\lambda_j$$

But we know (by Proposition 3.1) that the action of G on $\mathbf{V}^\lambda_j$, when expressed in terms of the basis $\langle e^\lambda_{1j}, e^\lambda_{2j},\ldots,e^\lambda_{n_\lambda j}\rangle$ is given precisely by the irreducible A^λ. Thus 3.15 gives that the action of G on $\mathcal{A}(G)\,p\,\phi\,q$ breaks up into $r_\lambda(\phi)$ occurrences of A^λ. In other words, the character of this representation is given precisely by the right hand side of 3.11. However, we see that we can write (*)

$$\mathcal{A}(G)\,p\,\phi\,q \;=\; \mathcal{A}(G)\,\phi\;q \;=\; q^{-1}\left[\mathcal{A}(G)\,\phi\right]q$$

Thus the representations corresponding to $\phi\mathcal{A}(G)$ and $p\phi q\mathcal{A}(G)$ are similar and formula 3.11 must hold true as asserted.

To complete the proof we must derive formula 3.13 from 3.11. To this end let us recall that the natural scalar product in the group algebra $\mathcal{A}(G)$ of a finite group G is obtained by setting for $f,g\in\mathcal{A}(G)$

$$\langle f,g\rangle \;=\; \frac{1}{|G|}\sum_{\gamma\in G} f(\gamma)\bar{g}(\gamma) \qquad 3.16$$

(*) Because of the invertibility of p and q

We also know that for any element $\gamma \in G$ we have

$$\langle \gamma f, g \rangle \;=\; \langle f, \gamma^{-1} g \rangle \qquad \text{and} \qquad \langle f\gamma, g \rangle \;=\; \langle f, g\gamma^{-1} \rangle \qquad 3.17$$

Note next that if ϕ satisfies $\phi^2 = \phi$ then for each $\lambda \in \Lambda$ we have as well that $A^\lambda(\phi)^2 = A^\lambda(\phi)$. Proposition 3.3 gives then that $A^\lambda(\phi)$ is diagonable. Since all its non-zero eigenvalues must be equal to 1, we immediately deduce that $rank\; A^\lambda(\phi) = trace\; A^\lambda(\phi)$. Thus when ϕ is an idempotent, formula 3.11 may be rewritten as

$$\chi^\phi \;=\; \sum_{\lambda \in \Lambda} \chi^\lambda \, trace\, \overline{A}^\lambda(\phi) \; .$$

Now the orthonormality of the characters with respect to the scalar product in 3.16 yields that this formula is equivalent to the identities

$$\langle \chi^\phi, \chi^\lambda \rangle \;=\; trace\, \overline{A}^\lambda(\phi) \; . \qquad\qquad (\,\forall \quad \lambda \in \Lambda\,) \qquad 3.18$$

Thus to prove 3.13 we need only verify that the right hand side of 3.12 satisfies the same identities. However, we see that the equalities in 3.16 (and the fact that the characters χ^λ are central elements of $\mathcal{A}(G)$) give

$$\begin{aligned} \langle (\frac{1}{|G|} \sum_{\gamma \in G} \gamma\phi\gamma^{-1}\,), \chi^\lambda \rangle \;&=\; \sum_{\gamma \in G} \langle \gamma\phi\gamma^{-1}\,, \chi^\lambda \rangle \;=\; \sum_{\gamma \in G} \langle \phi, \gamma^{-1}\, \chi^\lambda\, \gamma \rangle \\ &=\; \sum_{\gamma \in G} \langle \phi, \chi^\lambda \rangle \;=\; |G|\; \langle \phi, \chi^\lambda \rangle \\ &=\; |G|\, \frac{1}{|G|} \; \sum_{\gamma \in G} \phi(\gamma)\, \bar{\chi}^\lambda(\gamma) \;= \\ &=\; \sum_{\gamma \in G} \phi(\gamma)\, trace\, \overline{A}^\lambda(\gamma) \;=\; trace\, \overline{A}^\lambda(\phi) \; . \end{aligned}$$

This completes our proof.

4. Irreducibility

In this section we shall show that Young's representation $A^\lambda(\sigma)$ (defined in 1.11) is in fact irreducible. Let us start by observing that the equalities

$$\sigma_{S_r^\lambda S_s^\lambda}\, E_{S_s^\lambda S_j^\lambda} \;=\; E_{S_r^\lambda S_j^\lambda} \qquad 4.1$$

which hold for any given pair $r, s = 1, .., f_\lambda$ imply that

$$\mathcal{L}[E_{S_1^\lambda S_j^\lambda}, E_{S_2^\lambda S_j^\lambda}, \ldots, E_{S_{f_\lambda}^\lambda S_j^\lambda}] \;=\; \mathcal{A}(S_n) E_{S_j^\lambda S_j^\lambda} \; . \qquad 4.2$$

The reason for this is that 4.1 shows the containments

$$E_{S_r^\lambda S_j^\lambda} \in \mathcal{A}(S_n) E_{S_j^\lambda S_j^\lambda} \qquad\qquad (\forall \quad r = 1, .., f_\lambda\,)$$

which give that the left hand side of 4.2 is contained in the right hand side. However the reverse containment must also hold true because formula 2.8 (with $S^* = S_j$) shows that left multiplication of $E_{S_j^\lambda S_j^\lambda}$ by any $\sigma \in S_n$ always yields an element in the linear span of $E_{S_1^\lambda S_j^\lambda}, E_{S_2^\lambda S_j^\lambda}, \ldots, E_{S_{f_\lambda}^\lambda S_j^\lambda}$.

This places us in a position to compute the character of Young's representation by means of Theorem 3.1. More precisely we can show that each $E_{S_j^\lambda S_j^\lambda}$ is essentially an idempotent and thereby obtain the character of A^λ via formula 3.13. The crucial ingredient in this last step is given by the following beautiful identity.

Proposition 4.1 (Von Neuman Sandwich Lemma)

For any element $f \in \mathcal{A}(S_n)$ and any $T \in INJ(\lambda)$ we have

$$N(T)\; f\; P(T) \;=\; c_T(f) \;\; N(T)P(T) \;, \tag{4.3}$$

with

$$c_T(f) \;=\; f\; P(T)N(T)\,|_\epsilon \;. \tag{4.4}$$

Proof

Expanding f in the left hand side of 4.3 we get

$$N(T)\; f\; P(T) \;=\; \sum_{\sigma \in S_n} f(\sigma)\; N(T)\; \sigma\; P(T) \;. \tag{4.5}$$

Now note that using 1.2 a) we can write

$$N(T)\; \sigma\; P(T) \;=\; N(T)\; P(\sigma T)\; \sigma \;.$$

That means that the only terms that survive in 4.5 are those for which $N(T)\; P(\sigma T) \neq 0$. However, this can happen only if $\sigma T \wedge T$ is good. But part d) of Proposition 1.1 gives that we must have

$$T \;=\; \alpha\, \beta\, \sigma\, T \tag{4.6}$$

with $\alpha \in R(T)$ and $\beta \in C(T)$. In other words the only terms that survive in 4.5 are those coming from permutations σ of the form

$$\sigma \;=\; \beta^{-1}\alpha^{-1} \qquad\qquad (\text{with} \quad \alpha \in R(T) \;\;,\;\; \beta \in C(T)\;)$$

Since any permutation that can be so expressed has a unique expression of this form we can rewrite 4.5 as

$$N(T)\; f\; P(T) \;=\; \sum_{\alpha \in R(T)} \sum_{\beta \in C(T)} f(\beta^{-1}\alpha^{-1})\; N(T)\; \beta^{-1}\alpha^{-1}\; P(T) \;. \tag{4.7}$$

But then the simple identity

$$N(T)\; \beta^{-1}\alpha^{-1}\; P(T) \;=\; sign(\beta)\; N(T)P(T)$$

reduces 4.7 to

$$N(T)\; f\; P(T) \;=\; \left(\sum_{\alpha \in R(T)} \sum_{\beta \in C(T)} sign(\beta) f(\beta^{-1}\alpha^{-1}) \right)\; N(T)P(T) \;,$$

which is 4.3 with

$$c_T(f) \;=\; \sum_{\alpha\in R(T)}\sum_{\beta\in C(T)} sign(\beta) f(\beta^{-1}\alpha^{-1})$$

However the latter is but another way of writing 4.4.

Q.E.D.

An immediate consequence of Proposition 4.1 is the following fundamental fact

Theorem 4.1

For each $\lambda\vdash n$ we have a constant h_λ depending only on λ such that for all tableaux $T\in INJ(\lambda$ we have

$$E_{TT}^2 \;=\; h_\lambda \; E_{TT} \tag{4.8}$$

Proof

We simply use 4.3 with $f=P(T)N(T)$ and get

$$N(T)\;P(T)N(T)\;P(T) \;=\; h_\lambda \;\; N(T)P(T) \;\;,$$

with

$$h_\lambda \;=\; P(T)N(T)\;P(T)N(T)\,|_\epsilon \;=\; E_{TT}^2\,|_\epsilon \;\;.$$

Next we need to show that h_λ is the same for all injective tableaux of shape λ. However this is a simple consequence of the fact that if $T_1=\sigma T$ then $E_{T_1T_1}^2=\sigma E_{TT}^2\sigma^{-1}$ and thus

$$E_{T_1T_1}^2\,|_\epsilon \;=\; E_{TT}^2\,|_\epsilon \;\;.$$

Finally we must show that $h_\lambda\neq 0$. But if that were so the element E_{TT} would be nilpotent. This in turn would imply that $E_{TT}\,|_\epsilon=0$ (!!! why?). Now this is absurd since as we can easily see we have $E_{TT}\,|_\epsilon=1$. This completes our proof.

As a corollary we obtain Young's formulas for the character of his representation

Theorem 4.2

$$trace\;\;A^\lambda \;=\; \frac{1}{h_\lambda}\sum_{T\in INJ(\lambda)} N(T)P(T) \;=\; \frac{1}{h_\lambda}\sum_{T\in INJ(\lambda)} P(T)N(T) \tag{4.9}$$

Proof

The equality in 4.2 with $j=1$ gives that

$$\mathcal{L}[E_{S_1^\lambda S_1^\lambda},E_{S_2^\lambda S_1^\lambda},\ldots,E_{S_{f_\lambda}^\lambda S_1^\lambda}] \;=\; \mathcal{A}(S_n)E_{S_1^\lambda S_1^\lambda} \;=\; \mathcal{A}(S_n)E_{S_1^\lambda S_1^\lambda}/h_\lambda \;\;.$$

Since Theorem 4.1 with $T=S_1^\lambda$ gives that $E_{S_1^\lambda S_1^\lambda}/h_\lambda$ is an idempotent. Using Theorem 2.1 with $S^*=S_1^\lambda$ and formula 3.13 with $\phi=E_{S_1^\lambda S_1^\lambda}/h_\lambda$ we get

$$trace\;A^\lambda \;=\; \sum_{\sigma\in S_n}\sigma\,\frac{E_{S_1^\lambda S_1^\lambda}}{h_\lambda}\,\sigma^{-1} \;\;.$$

The identity in 4.9 now follows immediately since $\sigma\, E_{S_1^\lambda S_1^\lambda}\, \sigma^{-1} = E_{\sigma S_1^\lambda \sigma S_1^\lambda}$ and the tableau σS_1^λ, as σ varies in S_n, describes the injective tableaux T of shape λ each once and only once.

To prove the second equality in 4.9 we only need to observe that for any two elements f, g of a group algebra $\mathcal{A}(G)$ we have

$$\sum_{\gamma\in G} \gamma\, f\, g\, \gamma^{-1} \;=\; \sum_{\gamma\in G} \gamma\, g\, f\, \gamma^{-1} \;.$$

The identities in 4.8 and 4.9 have a most important consequence for us here.

Proposition 4.2

The idempotents E_{TT}/h_λ are of total rank 1 and consequently each of the characters

$$\chi^\lambda = trace\ A^\lambda$$

is irreducible.

Proof

For a moment let $\{\xi^\mu\}_{\mu\vdash n}$ denote a complete set of irreducible characters of S_n. (*) Then the idempotency of E_{TT}/h_λ and Theorem 3.1 give that

$$\chi^\lambda \;=\; \frac{1}{h_\lambda}\sum_{\sigma\in S_n} \sigma\, N(T)P(T)\, \sigma^{-1} \;=\; \sum_{\mu\vdash n} r_\mu(E_{TT}/h_\lambda)\, \xi^\mu \;. \qquad 4.10$$

This given, the orthonormality of the ξ^μ gives that

$$\langle\chi^\lambda,\chi^\lambda\rangle \;=\; \sum_{\mu\vdash n} (r_\mu(E_{TT}/h_\lambda))^2 \;.$$

Thus to show that χ^λ must be equal to one of the ξ^μ we need only verify that

$$\langle\chi^\lambda,\chi^\lambda\rangle \;=\; 1 \;. \qquad 4.11$$

To this end we use the first equality in 4.10 and get

$$\begin{aligned}\langle\, \chi^\lambda\,,\, \chi^\lambda\,\rangle \;&=\; \sum_{\sigma\in S_n} \langle \sigma\, N(T)P(T)\, \sigma^{-1}\,,\, \chi^\lambda\,\rangle \;=\; \sum_{\sigma\in S_n} \langle N(T)P(T)\,,\, \sigma^{-1}\chi^\lambda\, \sigma\,\rangle \\ &=\; \frac{1}{h_\lambda}\sum_{\sigma\in S_n} \langle\, N(T)P(T)\,,\, \chi^\lambda\,\rangle \;=\; \frac{n!}{h_\lambda}\,\langle\, N(T)P(T)\,,\, \chi^\lambda\,\rangle \end{aligned} \;.$$

Next we use the second equality in 4.9 and make the substitution

$$\chi^\lambda \;=\; \frac{1}{h_\lambda}\sum_{\sigma\in S_n} P(\sigma T)N(\sigma T) \;.$$

(*) Of course our final goal is to show that we may take $\xi^\lambda = \chi^\lambda$, but at this point we do not know

This gives

$$\begin{aligned}\langle\, \chi^\lambda \,,\, \chi^\lambda \,\rangle \;&=\; \frac{n!}{h_\lambda^2}\sum_{\sigma\in S_n} \langle\, N(T)P(T)\,,\, P(\sigma T)N(\sigma T)\,\rangle \\ &=\; \frac{n!}{h_\lambda^2}\sum_{\sigma\in S_n} \langle\, N(T)P(T)N(\sigma T)P(\sigma T)\,,\, \epsilon\rangle\end{aligned} \qquad 4.12$$

But we know that $P(T)N(\sigma T)\neq 0$ if and only if $\sigma=\alpha\beta$ with $\alpha\in R(T)$ and $\beta\in C(T)$. This gives that

$$\begin{aligned}N(T)P(T)N(\sigma T)P(\sigma T) &= N(T)P(T)\alpha\beta N(T)P(T)\beta^{-1}\alpha^{-1}\\ &= N(T)P(T)N(T)P(T)sign(\beta)\beta^{-1}\alpha^{-1} = h_\lambda\, N(T)P(T)sign(\beta)\beta^{-1}\alpha^{-1}\;\;.\end{aligned}$$

Putting this back into 4.12 we finally obtain

$$\begin{aligned}\langle\, \chi^\lambda \,,\, \chi^\lambda \,\rangle &= \frac{n!}{h_\lambda}\sum_{\sigma\in S_n} \langle\, N(T)P(T)sign(\beta)\beta^{-1}\alpha^{-1}\,,\, \epsilon\,\rangle = \frac{n!}{h_\lambda}\langle\, N(T)P(T)N(T)P(T)\,,\, \epsilon\,\rangle\\ &= n!\,\langle N(T)P(T)\,,\, \epsilon\,\rangle = N(T)P(T)\,\big|_\epsilon \;=\; 1\end{aligned}$$

Q.E.D.

We are finally in a position to show that

Theorem 4.3

The group algebra elements $\{\chi^\lambda\}_{\lambda\vdash n}$ form a complete set of irreducible irreducible characters of S_n.

Proof

We have seen that each χ^λ is irreducible. Since we have as many of them as conjucacy classes, we need only show that thay are mutually orthogonal, or equivalently that

$$\chi^\lambda\,\chi^\mu \;=\; 0 \qquad\qquad \text{when} \quad \lambda\neq\mu \qquad 4.13$$

But that is immediate, since $\chi^\lambda\,\chi^\mu\neq 0$ (by Young's formulas) implies that

$$\sum_{T_1\in INJ(\lambda)}\sum_{T_2\in INJ(\mu)} P(T_1)N(T_1)P(T_2)N(T_2) \;\neq\; 0$$

Now, the non-vanishing of a single term in this double sum implies that $T_2\wedge T_1$ is good and in particular (by 1.7) we get that $\lambda\le_D\mu$. However, the centrality of χ^λ and χ^μ gives that $\chi^\lambda\chi^\mu = \chi^\mu\chi^\lambda$ and this forces $\mu\le_D\lambda$ as well yielding $\lambda=\mu$. Thus 4.13 must hold true as asserted.

Our treatment would not be complete if we did not identify our mysterious constant h_λ. The following result which is also due to Alfred Young may be shown here (thanks to Von Neumann) in a much simpler way than in the original proof.

Theorem 4.3

$$h_\lambda \;=\; \frac{n!}{f_\lambda}$$

Proof

There is another way to compute the character χ^λ. That is to compute it as the trace of A^λ directly from the definition of A^λ itself. In fact, using 1.11 and dropping some of the λ superscripts we can write

$$\chi^\lambda \;=\; \sum_{\sigma\in S_n} \sigma\, trace\, C(\epsilon)^{-1}C(\sigma) \;=\; trace\, C(\epsilon)^{-1}\, F$$

where $F = \|F_{ij}\|_{i=1}^{f_\lambda}$ is a matrix with group algebra entries

$$F_{ij} \;=\; \sum_{\sigma\in S_n} \sigma\, C(S_i, \sigma S_j)\ \ .$$

Now Proposition 2.1 gives that

$$C(S_i,\sigma S_j) \;=\; C(\sigma^{-1}S_i, S_j) \;=\; E_{S_jS_i}\,|_{\sigma^{-1}}$$

Thus (*)

$$\begin{aligned} F_{ij} \;&=\; \sum_{\sigma\in S_n} \sigma\, E_{S_jS_i}\,|_{\sigma^{-1}} \\ &=\; \sum_{\sigma\in S_n} \sigma^{-1}\, E_{S_jS_i}\,|_{\sigma} \\ &=\; P(S_i)\sigma_{ij}N(S_j)\ \ . \end{aligned}$$

Note further that since χ^λ is central we must have as well that

$$\chi^\lambda \;=\; \frac{1}{n!}\sum_{\sigma\in S_n} \sigma\chi^\lambda\sigma^{-1}$$

We can thus also write

$$\chi^\lambda \;=\; \frac{1}{n!}\sum_{\sigma\in S_n} \sigma\left(trace\, C(\epsilon)^{-1}F \right)\sigma^{-1} \;=\; \frac{1}{n!} trace\, C(\epsilon)^{-1}\sum_{\sigma\in S_n} \sigma\, F\sigma^{-1}\ \ .$$

But we see that

$$\sum_{\sigma\in S_n} \sigma\, F_{ij}\sigma^{-1} \;=\; \sum_{\sigma\in S_n} \sigma\, P(S_i)\sigma_{ij}N(S_j)\sigma^{-1} \;=\; \sum_{\sigma\in S_n} \sigma\, N(S_j)P(S_i)\sigma_{ij}\sigma^{-1} \;=\; 0$$

at least when $i > j$. This means that the matrix

$$\sum_{\sigma\in S_n} \sigma\, F\sigma^{-1}$$

is lower triangular and since $C(\epsilon)^{-1}$ is lower triangular with unit diagonal elements we must conclude that

$$\chi^\lambda = \frac{1}{n!}\, trace \sum_{\sigma\in S_n} \sigma\, F\sigma^{-1} = \frac{1}{n!}\sum_{i=1}^{f_\lambda}\sum_{\sigma\in S_n} \sigma\, P(S_i)N(S_i)\sigma^{-1} = \frac{f_\lambda}{n!}\sum_{T\in INJ(\lambda)} P(T)N(T)$$

(*) For convenience we are setting $\sigma_{S_iS_j} = \sigma_{ij}$

comparing with formula 4.9 we see that we must have $h_\lambda = n!/f_\lambda$ precisely as asserted.

We terminate with Young's formulas giving a complete sets of matrix units for $\mathcal{A}(S_n)$. To this end it is convenient to set $\sigma_{S_i^\lambda, S_j^\lambda} = \sigma_{ij}^\lambda$ and

$$\gamma_i^\lambda \quad = \quad \frac{1}{h_\lambda}\, N(S_i^\lambda)P(S_i^\lambda) \qquad 4.14$$

Theorem 4.4

The group algebra elements

$$e_{ij}^\lambda \quad = \quad \gamma_i^\lambda\, \sigma_{ij}^\lambda\, (1-\gamma_{j+1}^\lambda)(1-\gamma_{j+2}^\lambda)\cdots(1-\gamma_{f_\lambda}^\lambda)$$

satisfy the identities

$$e_{ij}^\lambda\; e_{rs}^\mu \quad = \quad \begin{cases} e_{is}^\lambda & \text{if} \quad \lambda=\mu \quad \text{and} \quad j=r \\ 0 & \text{otherwise} \end{cases} \qquad 4.15$$

Proof

Note that if T_1 and T_2 are injective tableaux of different shapes λ and μ we necessarily have

$$N(T_1)P(T_1)N(T_2)P(T_2) \quad = \quad 0\ .$$

The reason for this is that $P(T_1)N(T_2) \neq 0$ immediately gives (by 1.7) $\lambda \leq_D \mu$. Moreover the non vanishing of the whole product $N(T_1)P(T_1)N(T_2)P(T_2)$ yields that for at least one permutation σ we must have $N(T_1)\sigma P(T_2) \neq 0$. But since $N(T_1)\sigma P(T_2) = N(T_1)P(\sigma T_2)\sigma$ we are also forced to conclude that $N(T_1)P(\sigma T_2) \neq 0$ as well, which gives the reverse inequality $\mu \leq_D \lambda$. This given, we need only establish the identities in 4.15 when $\lambda = \mu$. To simplify our notation we shall hereafter drop the superscript λ in all our elements. Now note that we can write

$$e_{ij}\; e_{rs} \quad = \quad \sigma_{ij}\gamma_j(1-\gamma_{j+1})\cdots(1-\gamma_{f_\lambda})\gamma_r\sigma_{rs}(1-\gamma_{s+1})\cdots(1-\gamma_{f_\lambda})\ .$$

Note further that

$$j > i \quad \longrightarrow \quad \gamma_j\gamma_i = 0\ . \qquad 4.16$$

This is because we have seen that for $j > i$ we have two elements a, b that are in the same column of S_i and the same row of S_j forcing $P(S_j)N(S_i) = 0$. Thus, when $j < r$ we get

$$(1-\gamma_{j+1})\cdots(1-\gamma_{f_\lambda})\gamma_r \quad = \quad 0\ ,$$

and when $j > r$ we must have

$$\gamma_j(1-\gamma_{j+1})\cdots(1-\gamma_{f_\lambda})\gamma_r \quad = \quad \gamma_j\gamma_r \quad = \quad 0$$

as well. Finally for $j = r$

$$\begin{aligned} e_{ij}e_{js} \quad &= \quad \sigma_{ij}\gamma_j(1-\gamma_{j+1})\cdots(1-\gamma_{f_\lambda})\gamma_j\sigma_{js}(1-\gamma_{s+1})\cdots(1-\gamma_{f_\lambda}) \\ &= \quad \sigma_{ij}\gamma_j^2\sigma_{js}(1-\gamma_{s+1})\cdots(1-\gamma_{f_\lambda}) \\ &= \quad \sigma_{ij}\sigma_{js}\gamma_s(1-\gamma_{s+1})\cdots(1-\gamma_{f_\lambda}) \quad = \quad e_{is} \end{aligned}$$

Q.E.D.