JOURNAL OF ALGEBRA 69, 287-297 (1981)

A New Construction of Young's Seminormal Representation of the Symmetric Groups

G. E. MURPHY

Department of Mathematics, North East London Polytechnic, Romford Road, Stratford, London E15 4LZ, England

Communicated by Walter Feit

Received January 10, 1980

1. INTRODUCTION

We present a new construction of Young's seminormal representation [1] of the symmetric group over a field of characteristic zero. Let S_K^{μ} [2, p. 90], [3, p. 14] be the Specht module, defined over a field K of characteristic zero, corresponding to the partition μ of n. We consider the effect on S_K^{μ} of the linear operators

$$L_u = \sum_{v=1}^{u-1} (u v), \qquad u = 2, 3, ..., n,$$

where (u v) denotes a transposition. The simultaneous eigenfunctions of the L_u are complete and non-degenerate, and form the basis for S_n^{μ} which gives rise to Young's seminormal representation of the symmetric group \mathfrak{S}_n . The construction is elementary in the sense that it relies only on the standard basis for the Specht module and the Garnir relations [2]. A number of interesting corollaries arise, the principal one being a much simpler derivation of the "branching theorem" for the determinants of the Gram matrices of the Specht modules [4].

The analysis depends critically on the characteristic of the field, and in the more interesting case when K has finite characteristic the crucial Lemma 2.2 does not hold in general, and the eigenfunctions of the L_u are incomplete and degenerate. However, the method does have applications in this case, which will be explored in a later paper.

G. E. MURPHY

2. TABLEAUX AND OPERATORS

We consider the partition $\mu = (\mu_1, \mu_2, ..., \mu_s)$ of *n*, with $\mu_1 \ge \mu_2 \ge \cdots \ge \mu_s$, and the corresponding Young diagram $[\mu]$ with *s* rows of length $\mu_1, \mu_2, ..., \mu_s$. For a μ -tableau *t*, i.e., arrangement of the numbers 1, 2,..., *n* in $[\mu]$, let r_{tu} denote the row occupied by *u* in *t*, and C_t the group of column permutations of *t*. We denote by z_t the monomial in the indeterminates $x_1, x_2, ..., x_n$ defined by

$$z_t = \prod_{u=1}^n x_u^{r_{tu}-1},$$

and by e_i the Specht polynomial

$$e_t = \sum_{\sigma \in C_t} \operatorname{sgn} \sigma \cdot \sigma z_t = \sum_{\sigma \in C_t} \operatorname{sgn} \sigma z_{\sigma t}.$$

We shall require two orderings of the monomials z_t , a total order < and a partial order \lhd [3, p. 10]. If t, t^* are μ -tableaux, then $z_t < z_{t^*}$ if for some $u \le n$, $r_{tu} < r_{t^*u}$, and for each v > u, $r_{tv} = r_{t^*v}$. If $m_{ur}(t)$ denotes the number of entries $v \le u$ in the first r rows of t, then $z_t \le z_{t^*}$ if for all u, r we have $m_{ur}(t) \le m_{ur}(t^*)$, and $z_t < z_{t^*}$ if $z_t \le z_{t^*}$ and for some $u, r, m_{ur}(t) < m_{ur}(t^*)$. By considering the largest u for which this occurs, it is easy to see that $z_t < z_{t^*}$ implies $z_t < z_{t^*}$, although the converse is not true.

The standard μ -tableaux, i.e., those whose rows and columns are strictly increasing, correspond to distinct monomials, and so are totally ordered by <; we denote them by $t_1, t_2, ..., t_d$, with $z_{t_1} < z_{t_2} < \cdots < z_{t_d}$. For brevity we write z_i, e_i for z_{t_i}, e_{t_i} , and i < j for $z_i < z_j$.

If t is column standard, and $t^* = \sigma t \neq t$ for some $\sigma \in C_t$, then clearly $z_{t^*} \triangleleft z_t$. Thus e_t has a leading monomial z_t ; in particular, if t, \bar{t} are column standard, and $z_{\bar{t}}$ occurs in e_t , then $z_{\bar{t}} \triangleleft z_t$.

2.1 LEMMA. Suppose u > v, and v occurs in a column to the right of u in t_i . Then there are numbers a_{ii} such that

$$(u\,v)\,e_i=\sum_{j\triangleleft i}a_{ij}e_j.$$

Proof. The e_i form a basis for S_k^{μ} [2], so that there are certainly numbers a_{ii} such that

$$(u v) e_i = \sum_{j=1}^d a_{ij} e_j.$$

We must show that $a_{ii} = 0$ unless $j \triangleleft i$.

Suppose *u* is the r_1 th element of a column $u_1, u_2, ..., u_k$ of t_i , and *v* the r_2 th element of a later column $v_1, v_2, ..., v_l$, so that $r_1 > r_2$. Let t be the column standard tableau obtained by ordering the columns of $(u v) t_i$. To obtain t from t_i we interchange *u* and *v* and allow each to "float" to its correct position, i.e., we replace each v_r with $v \le v_r \le u$ by $\min(u, v_{r+1})$, or by *u* if r = l, and each u_r with $v \le u_r \le u$ by $\max(v, u_{r-1})$. Now let us compare the set of elements in the first *r* rows of t_i with the corresponding set of t_i , ignoring individual positions. For $r < r_2$ or $r \ge r_1$ these are identical; for $r_2 \le r < r_1$ they differ only in that an element $x_r = \max(v, u_r)$ has been removed, and replaced by $y_r = \min(u, v_{r+1})$, or simply $y_r = u$ if r = l. Because t_i is standard, $v_{r+1} > u_r$, and $y_r > x_r$. Therefore

$$\begin{split} m_{wr}(\bar{t}) &= m_{wr}(t_i) - 1 \qquad \text{if} \quad x_r \leqslant w < y_r, \ r_2 \leqslant r < r_1, \\ &= m_{wr}(t_i) \qquad \text{otherwise,} \end{split}$$

so that $\overline{t} \triangleleft t_i$.

Since \overline{t} is obtained by a column permutation from $(u v) t_i$, $(u v) e_i = \pm e_{\overline{t}}$. Suppose there is some $j \triangleleft i$ such that $a_{ij} \neq 0$; choose the largest such j. z_j occurs in this term and no other on the right side of the expansion, and so must occur in e_i . But this means that $z_j \leq z_{\overline{t}} < z_i$, a contradiction, and so $a_{ij} = 0$ unless j < i.

Let us define the class of the (k, l) node of $[\mu]$ to be the difference l - k. Nodes in the same class therefore occur in diagonal lines, and the classes of the removable nodes of $[\mu]$, i.e., those nodes at the end of both a row and a column, are all distinct. Let α_{ui} be the class of the node occupied by u in t_i . By considering the successive removal of nodes n, n - 1,..., 2 from t_i , we see that the sequence $(\alpha_{ui}), u = 1, 2,..., n$, determines t_i completely. Also for fixed i, j, the sequence (α_{ui}) is simply a permutation of (α_{uj}) . Together these give us a simple but important lemma.

2.2 LEMMA. (a) If
$$\alpha_{vi} = \alpha_{vj}$$
 for all $v \le n$, then $t_i = t_j$.
(b) If $\alpha_{vi} = \alpha_{vj}$ for all $v \le n$ except possibly $v = u, u - 1$, then $t_i = t_j$ or $t_i = (u, u - 1)t_j$.

We shall make extensive use of the following commutation relations for the L_{u} , which are easily verified.

$$L_{u}(u, u-1) = (u, u-1)L_{u-1} + 1;$$
(2.3)

$$L_{u-1}(u, u-1) = (u, u-1) L_u - 1;$$
(2.4)

$$L_{u}(v, v-1) = (v, v-1) L_{u}, \qquad v \neq u, u-1;$$
(2.5)

$$L_u L_v = L_v L_u \qquad \text{for all} \quad u, v. \tag{2.6}$$

The distinct z_t form a K-basis for the space $\sum_t K \cdot z_t$, where t runs over all

 μ -tableaux. Let \langle , \rangle be the bilinear form [3, p. 14] defined on this space, such that

$$\langle z_t, z_{t^*} \rangle = 1$$
 if $z_t = z_{t^*}$
= 0 otherwise.

ŕ

Since L_u is a sum of transpositions and \langle , \rangle is invariant under the action of \mathfrak{S}_n , we have for any $\xi, \eta \in S_k^u$

$$\langle \xi, L_u \eta \rangle = \langle L_u \xi, \eta \rangle.$$

Moreover, if E is any polynomial function of the L_u , we may use (2.6) to give

$$\langle \xi, E\eta \rangle = \langle E\xi, \eta \rangle.$$

3. THE ORTHOGONAL BASIS

We now consider the effect of operating with L_u on e_i . Suppose C is the set of elements in some column to the left of u in t_i . By the simplest of the Garnir relations [2, p. 92]

$$\sum_{v \in C} (u v) e_i = e_i. \tag{3.1}$$

If v is in the same column as u,

$$(u v) e_i = -e_i. \tag{3.2}$$

Let

 $Y = \{ y \mid y < u \text{ and } y \text{ is to the right of } u \text{ in } t_i \},$ $Z = \{ z \mid z > u \text{ and } z \text{ is to the left of } u \text{ in } t_i \}.$

Then if u is in the kth row and lth column of t_i , t_i has the form

290

in an obvious notation. Using (3.1) and (3.2)

$$L_{u}e_{i} = \sum_{v=1}^{u-1} (u v) e_{i}$$

$$= (l-1) e_{i} - (k-1) e_{i} - \sum_{z \in Z} (z u) e_{i} + \sum_{y \in Y} (y u) e_{i}.$$
(3.3)

The first two terms on the right are just $a_{ui}e_i$, and all the remaining terms satisfy the conditions of Lemma 2.1, so that there are numbers a_{ii} such that

$$(L_u - \alpha_{ui}) e_i = \sum_{j < i} a_{ij} e_j.$$
(3.4)

3.5 LEMMA. Let $u_1, u_2, ..., u_i \leq n$ be any sequence of integers; then

$$\prod_{k=1}^{i} (L_{u_k} - \alpha_{u_k k}) e_i = 0.$$

Proof. $(L_{u_i} - \alpha_{u_i i}) e_j$ is a sum of terms e_k with k < j; since the factors commute the lemma follows by a simple induction on *i*.

Define

$$E_i = \prod_{c=-n+1}^{n-1} \prod_{\{u \mid \alpha_{ui} \neq c, u \leq n\}} \frac{c - L_u}{c - \alpha_{ui}},$$

$$f_i = E_i^d e_i.$$

Notice that the range of c includes all the distinct classes of $[\mu]$, and i runs over all the indices of the standard μ -tableaux.

3.6 LEMMA. $\{f_1, f_2, ..., f_d\}$ is a set of mutually orthogonal simultaneous eigenfunctions of the L_u , $(2 \le u \le n)$.

Proof. If $i \neq j$, Lemma 2.2 assures us that there is a u such that $\alpha_{ui} \neq \alpha_{uj}$. Thus for any j < i there is a sequence $u_1, u_2, ..., u_j$ such that E_i^d has a factor $(L_{u_k} - \alpha_{u_kk})$ for any $k \leq j$; the exponent d ensures that repeated terms occur sufficiently often. Thus by Lemma 3.5,

$$E_i^d e_j = 0 \qquad \text{for} \quad j < i, \tag{3.7}$$

and applying E_i^d to (3.4),

$$(L_u - a_{ui}) f_i = 0. (3.8)$$

Consequently $(c - L_u)f_i = (c - \alpha_{ui})f_i$ for all $c \in K$. Using the definition of E_i and the first line of the proof

$$E_i f_i = f_i,$$

$$E_i f_j = 0 \quad \text{if} \quad j \neq i. \tag{3.9}$$

Therefore if $j \neq i$,

$$\langle f_i, f_j \rangle = \langle E_i f_i, f_j \rangle = \langle f_i, E_i f_j \rangle = 0,$$

and the f_i are mutually orthogonal.

3.10 THEOREM. The f_i form an orthogonal basis for S_k^u , obtained from the standard basis by a unimodular linear transformation, and

$$e_i = \sum_{j \leq i} a_{ij} f_j,$$

where $a_{ii} = 1$, $a_{ij} = \langle e_i, f_j \rangle / \langle f_j, f_j \rangle$.

Proof. Each factor of E_i is of the form $1 - (L_u - a_{ui})/(c - a_{ui})$. Considering the effect of this on e_i , and on any e_j with $j \triangleleft i$, we see that $E_i^d e_i$ is the sum of e_i and a set of terms e_j with $j \triangleleft i$, ie for some b_{ij} ,

$$f_i = \sum_{j \leq i} b_{ij} e_j$$

with $b_{ii} = 1$. The transformation matrix is lower triangular, with 1's on the diagonal, and so is unimodular; the inverse transformation is of the same form, i.e.,

$$e_i = \sum_{j \leq i} a_{ij} f_j,$$

and since the f_i are orthogonal, $a_{ij} = \langle e_i, f_j \rangle / \langle f_j, f_j \rangle$. Operating with E_i , and using (3.9), we have immediately

3.11 COROLLARY. $f_i = E_i e_i$.

We now construct the matrix representing the transposition (u - 1, u) with respect to the basis $f_1, f_2, ..., f_d$.

3.12 THEOREM (Young's seminormal representation) [1, p. 217; reprint p. 453]. Let $t_i = (u - 1, u) t_i$ if this is standard. Then

$$(u-1, u)f_i = \rho_1 f_i + \rho_2 f_j,$$

292

where

$$\rho_1^{-1} = \alpha_{ui} - \alpha_{u-1,i}$$

and

$$\begin{split} \rho_2 &= 0 & \text{if } (u-1,u) \ t_i \text{ is not standard,} \\ &= 1 & \text{if } j > i, \\ &= 1 - \rho_1^2 & \text{if } j < i. \end{split}$$

Proof. The denominators of E_i, E_j are in fact independent of the particular tableaux, depending only on the partition, and are identical. The numerators have common factors

$$c - L_v, \qquad v \neq u, u - 1 \tag{3.13}$$

and common pairs of factors

$$(c - L_u)(c - L_{u-1}), \qquad c \neq \alpha_{ui}, \, \alpha_{u-1,i},$$
 (3.14)

and differ only in that E_i has factors $(\alpha_{ui} - L_{u-1})(\alpha_{u-1,i} - L_u)$, where E_j has $(\alpha_{ui} - L_u)(\alpha_{u-1,i} - L_{u-1})$, since $\alpha_{vi} = \alpha_{vj}$, $v \neq u, u-1$, and $\alpha_{ui} = \alpha_{u-1,j}$, $\alpha_{u-1,i} = \alpha_{uj}$. It is easy to verify, using (2.3) to (2.6), that (u-1, u) commutes with (3.13) and (3.14), while

$$(u-1, u)(\alpha_{ui} - L_{u-1})(\alpha_{u-1,i} - L_u) = (\alpha_{ui} - L_u)(\alpha_{u-1,i} - L_{u-1})(u, u-1) + (\alpha_{u-1,i} - \alpha_{ui}).$$
(3.15)

Consequently, if we write

$$E_{i} = \frac{(\alpha_{ui} - L_{u-1})(\alpha_{u-1,i} - L_{u})}{(\alpha_{ui} - \alpha_{u-1,i})(\alpha_{u-1,i} - \alpha_{ui})} F,$$

$$E_{j} = \frac{(\alpha_{ui} - L_{u})(\alpha_{u-1,i} - L_{u-1})}{(\alpha_{ui} - \alpha_{u-1,i})(\alpha_{u-1,i} - \alpha_{ui})} F,$$
(3.16)

(u-1, u) commutes with F, and

$$(u-1, u) E_i = E_j(u-1, u) + (\alpha_{ui} - \alpha_{u-1,i})^{-1} F.$$
(3.17)

If $(u-1, u) t_i$ is not standard, t_j and E_j are not defined; however, we may take (3.16) as the definition of E_j in this case. Comparing (3.16) and (3.9), we see that $Ff_i = f_i$, $Ff_j = f_j$. If $k \neq i, j$, then by Lemma 2.2(b) F has a factor $(\alpha_{vk} - L_v)$ for some v, so that $Ff_k = 0$. Therefore, unless i > j, $Fe_i = f_i$ by Theorem 3.10.

Now let us apply (3.17) to e_i . If $(u-1, u) t_i$ is standard, and j > i, then

$$(u-1, u)f_i = E_j e_j + (\alpha_{ui} - \alpha_{u-1,i})^{-1} F e_i$$

= $f_j + \rho_1 f_i.$ (3.18)

If $(u-1, u) t_i$ is not standard, then F is a projection operator for f_i , so that E_j annihilates the whole of the orthogonal basis, and therefore every element of S_k^u , in particular $E_j(u-1, u) e_i = 0$, so that

$$(u-1, u)f_i = (\alpha_{ui} - \alpha_{u-1,i})^{-1} Fe_i = \rho_1 f_i.$$
(3.19)

If $(u-1, u) t_i$ is standard, and j < i, we may reverse i, j and the sign of ρ_1 in (3.18) to obtain

$$(u-1,u)f_j=f_i-\rho_1f_j,$$

so that operating again with (u-1, u) and collecting terms,

$$(u-1, u)f_i = f_j + \rho_1(u-1, u)f_j = (1-\rho_1^2)f_j + \rho_1f_i.$$
(3.20)

3.21 COROLLARY. If i > j then $\langle f_i, f_i \rangle = \rho_2 \langle f_j, f_j \rangle$.

Proof. Taking the inner product of (3.20) with itself, and using the orthogonality of f_i , f_j ,

$$\langle f_i, f_i \rangle = (1 - \rho_1^2)^2 \langle f_j, f_j \rangle + \rho_1^2 \langle f_i, f_i \rangle,$$

so that $\langle f_i, f_i \rangle = \rho_2 \langle f_j, f_j \rangle$.

4. DIAGONALIZATION OF THE GRAM MATRIX

We may use Corollary 3.21 to calculate $\langle f_i, f_i \rangle$. Let us number the removable nodes of $[\mu]$ as $r_1, r_2, ...,$ starting at the top, e.g.,

The (i, j) hook of $[\mu]$ consists of the node (i, j) together with the nodes to the right of it in the *i*th row and below it in the *j*th column, and has length h_{ij} ;

294

here the (2, 2) hook is illustrated, and $h_{22} = 6$. If the *i*th removable node is in the (k, l) position, we define the "hook quotient" q_i by

$$q_i = \prod_{j=1}^{k-1} \frac{h_{jl}}{h_{jl}-1}.$$

Simply, if $[\mu(i)]$ is the diagram obtained by removing this node, q_i is the ratio of the product of the hooks in the *l*th column of $[\mu]$ to that in $[\mu(i)]$.

The removal of all nodes larger than u from t_i leaves a tableau t_i^u for some partition of u. u occupies a removable node of t_i^u , and we denote the corresponding hook quotient by γ_{ui} .

The partition conjugate to μ is μ' , where $[\mu']$ is the transpose of μ . We denote the transpose of t_i by t'_i , and similarly the corresponding quantities e'_i , γ'_{ui} , f'_i . Note that transposition reverses the standard ordering, so that $t'_d < t'_{d-1} < \cdots < t'_1$.

4.1. THEOREM. $\langle f_i, f_i \rangle = \prod_{u=1}^n \gamma_{ui}$.

Proof. We proceed by induction. If μ' has s' parts, then $[\mu]$ has columns of length $\mu'_1, \mu'_2, ..., \mu'_{s'}$. Now $e_1 = f_1$, and it is simple to verify that

$$\langle e_1, e_1 \rangle = \prod_{l=1}^{s'} \mu_l'!.$$

u is the last element in t_1^u , and so γ_{u1} is simply the length of the last column. Consequently the product of the γ_{u1} for all *u* in the *l*th column of t_1 is just $\mu'_l!$, which gives

$$\prod_{u=1}^n \gamma_{u1} = \prod_{l=1}^{s'} \mu_l'! = \langle f_1, f_1 \rangle.$$

Suppose the corollary is true for $f_1, f_2, ..., f_{i-1}$. If i > 1, there is a u such that u - 1 is to the right of u in t_i , so that $t_j = (u - 1, u) t_i$ is standard, with i > j. For any v, γ_{vi} depends only on v and the shape of t_i^v , so that $\gamma_{vi} = \gamma_{vj}$ for all $v \neq u, u - 1$, and $\gamma_{u-1,i} = \gamma_{uj}$. Let h be the hook joining u and u - 1 in t_i^u , so that $h = \alpha_{u-1,i} - \alpha_{ui} + 1 = 1 - \rho_1^{-1}$. Then γ_{ui} differs from $\gamma_{u-1,j}$ only in that the former has a factor h/(h-1), the latter (h-1)/(h-2), and

$$\gamma_{ui} = \frac{h(h-2)}{(h-1)^2} \gamma_{u-1,j} = (1-\rho_1^2) \gamma_{u-1,j}.$$

Consequently, by Corollary 3.21 and the inductive hypothesis,

$$\prod_{u=1}^{n} \gamma_{ui} = \rho_2 \prod_{u=1}^{n} \gamma_{uj} = \rho_2 \langle f_j, f_j \rangle = \langle f_i, f_i \rangle.$$

4.2 COROLLARY. $\langle f_i, f_i \rangle \langle f'_i, f'_i \rangle$ is the product of all the hooks of $[\mu]$.

Proof. The corollary is equivalent to the statement that the hook product is

$$\prod_{u=1}^n \gamma_{ui} \gamma'_{ui}.$$

We prove this by induction on *n*. It is trivially true for n = 1, suppose it is true for 1, 2, ..., n - 1. Let $[\lambda]$ be the diagram for t_i^{n-1} ; then by the inductive hypothesis the product of the hooks in $[\lambda]$ is

$$\prod_{u=1}^{n-1} \gamma_{ui} \gamma'_{ui}.$$

Now if *n* is in the (k, l) position of $[\mu]$, $\gamma_{ni}\gamma'_{ni}$ is just the product of the hooks in the *k*th row and the *l*th column of $[\mu]$ divided by the product for the same row and column of $[\lambda]$. But the former are the $[\mu]$ hooks which are not also $[\lambda]$ hooks, and conversely for the latter; the corollary follows.

Let G^{μ} be the Gram matrix [3, p3] of S^{μ} with elements $\langle e_i, e_j \rangle$, i, j = 1, 2, ..., d. By transforming to the basis $\{f_i\}$ we apply a unimodular transformation to G^{μ} which reduces it to diagonal form with diagonal elements $\langle f_i, f_i \rangle$, i = 1, 2, ..., d. If again $\mu(i)$ is obtained from μ by removing the *i*th removable node, we have the "branching theorem" for determinants [4, p. 225].

4.3 COROLLARY. det $G^{\mu} = \prod_{i} q_i^{d_i} \det G^{\mu(i)}$, where $d_i = \dim(S^{\mu(i)})$.

Proof. From Corollary 4.2 we have

det
$$G^{\mu} = \prod_{i=1}^{d} \langle f_i, f_i \rangle = \prod_{i=1}^{d} \prod_{u=1}^{n} \gamma_{ui}.$$

Now the tableaux with n in the *i*th removable node are simply the $\mu(i)$ tableaux with this node added; there are d_i of them, and they contribute a factor

$$q_i^{d_i}$$
 det $G^{\mu(i)}$,

and det G^{μ} is the product of such factors.

ACKNOWLEDGMENTS

The author is indebted to M. H. Peel and G. D. James for their considerable advice and assistance during the preparation of this paper, and to the referee for his sympathetic and helpful comments on style and form.

References

- A. YOUNG, On quantitative substitutional analysis VI, Proc. London Math. Soc. (2) 31 (1931), 253-289. Reprinted in "The Collected Papers of Alfred Young, 1873-1940," Mathematical Expositions 21, Univ. of Toronto Press, Toronto and Buffalo, 1977.
- 2. M. H. PEEL, Specht modules and the symmetric groups, J. Algebra 36 (1975), 88-97.
- 3. G. D. JAMES, "The Representation Theory of the Symmetric Groups," Lecture Notes in Mathematics, Vol. 682, Springer-Verlag, Berlin/New York, 1978.
- 4. G. D. JAMES AND G. E. MURPHY, The determinant of the Gram matrix for a Specht module, J. Algebra 59 (1979), 222-235.