Youngs's Representations of the Symmetric Group

Alain Lascoux

CNRS, Institut Gaspard Monge, Université de Marne-la-Vallée

77454 Marne-la-Vallée Cedex, France

Alain.Lascoux@univ-mlv.fr

September 2000

Abstract

We describe the different matrices, due to Young, representing the symmetric group, by reading the same graph with various labellings. Orthogonal idempotents are obtained in the like manner. The only mathematical tools needed for these constructions are comparison of integers and addition of vectors.

Young's purpose was to decompose functions – nowadays one would say "tensors" – into components according to their type of symmetry. In other words, he wanted to generalize the decomposition

$$f(x,y) = \frac{1}{2}\big(f(x,y) + (f(y,x)\big) + \frac{1}{2}\big(f(x,y) - (f(y,x)\big)$$

for functions of two variables to the case of more variables (cf. [Ro]).

He first rewrote the preceding equation into

$$1 = \frac{1}{2}(1+\sigma) + \frac{1}{2}(1-\sigma)\ ,$$

where σ is the transposition of x, y, and thus moved to the study of equations in the group algebra of the symmetric group.

He produced different families of idempotents, and, as a by-product, he obtained various matrices representing the symmetric group (we shall refer to [JK], [Ru] for the traditional approach).

A new impetus to the description of representations was given by the notions of *Gelfand-Tsetlin basis* (cf. [OV]), of maximal commutative subalgebras and of Yang-Baxter equation, which allow to reinterpret the constructions of Young in a simpler manner. As a matter of fact, the only mathematical tools which we require in this text are

- Given two numbers a, b, decide whether $a \geq b$ or not
- compute their difference $a - b$.
- add vectors and test whether the result is a permutation.

Now, introducing various initial elements associated to a given partition, ones generates labelled graphs from which one reads the matrices representing simple transpositions, as well as orthogonal idempotents. One also can directly write other matrices of representation from addition of vectors.

We shall illustrate all these constructions on the partition [3,2], while giving the precise initial data for any partition. There will be some variations, to simplify labellings, in the way a partition is represented graphically by diagrams of boxes in the different quarters of the plane.

When using the term "word" instead of integral vector, we refer to computations in the free algebra, as a proper way of handling such objects as Young tableaux (cf. [LLT2]). For example, $[3,5,1,2,4]$ can be seen as the word $x_3x_5x_1x_2x_4$, as well as the Young tableau $\begin{matrix} 3\,5 \\ 1\,2\,4 \end{matrix}$.

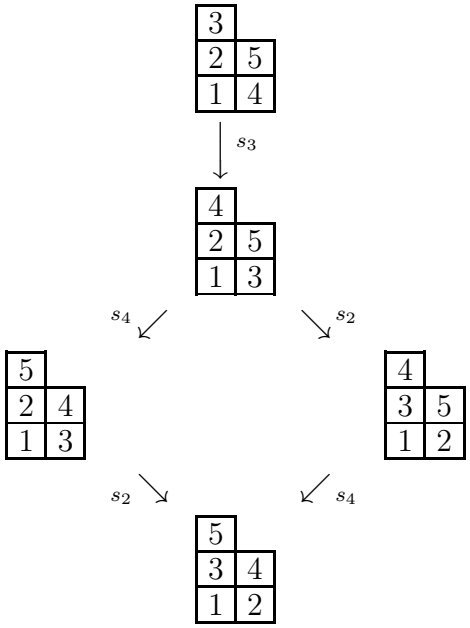
For a start, take the word $[1,0,2,1,0]$, vertical reading of $\begin{matrix} 1\,2 \\ 0\,1 \\ 0 \end{matrix}$, and given any two adjacent letters, a, b, transpose them iff $a > b$.

Iterating and recording positions by colors $\sigma_1, \sigma_2, \ldots$, one gets a colored graph.

$$\begin{array}{ccccc} & & 1\,0\,2\,1\,0 & & \\ & & \Big\downarrow \sigma_2 & & \\ & & 1\,2\,0\,1\,0 & & \\ & \sigma_1 \swarrow & & \searrow \sigma_3 & \\ 2\,1\,0\,1\,0 & & & & 1\,2\,1\,0\,0 \\ & \sigma_3 \searrow & & \swarrow \sigma_1 & \\ & & 2\,1\,1\,0\,0 & & \end{array} \qquad (\Gamma 1)$$

For a general partition $\beta = [\beta_1, \beta_2, \ldots]$ (written increasingly), the starting word is $[\cdots(\beta_3-1,\ldots,0)(\beta_2-1,\ldots,0)(\beta_1-1,\ldots,0)]$ and can be visualized (as we have done for $[2,3]$) on the north-east diagram with column lengths the partition, each bottom box being occupied by 0, columns being occupied by consecutive numbers.

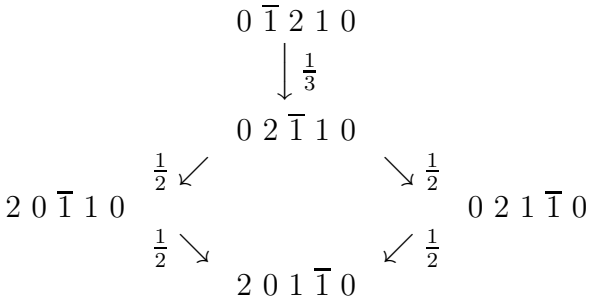
Now, we keep the same graph, but interpret each word as describing the levels occupied by letters $1, 2, \ldots$ in the diagram of [3,2]:



(Γ2)

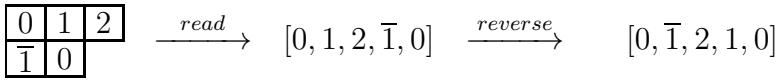
The vertices of the graph are now exactly all the *standard Young tableaux* of a given shape, and the edges are the transpositions of consecutive letters. The labelling of edges has been changed $s_i \leftrightarrow \sigma_{n-i}$ (we could also have read positions from the right).

We did not record values which were exchanged. Label vertices differently once more, and label arrows by the inverse of the **differences** of the values exchanged ($\overline{1}$ stands for -1) :



(Γ3)

The initial vector is the "vector of contents", that is, the diagonal positions of the boxes in the diagram



How to read matrices from the last graph ?

The underlying vector space has a basis coded by the vertices of the graph. To represent any simple transposition s_i, one first erase all edges which are not labelled by s_i. One is left with isolated vertices (corresponding to 1-dimensional representations of $\mathfrak{S}_2$), and pairs of vertices connected by an edge, corresponding to 2-dimensional representations. In this last case, if β is the parameter written on the edge, then the matrix representing s_i on the subspace generated by the two vertices is

$$\begin{bmatrix} -\beta & \sqrt{1-\beta^2} \\ \sqrt{1-\beta^2} & \beta \end{bmatrix} .$$

In the one-dimensional case, if $i, i+1$ is a subword of the vertex, then the restriction of the representation is trivial (i.e. the submatrix is 1), otherwise it is the alternating representation (submatrix $= -1$).

The full matrix M_i representing s_i is composed of these elementary matrices. The reader can check that the elements $\neq \pm 1$ of the following matrices constitute elementary 2×2 matrices with entries $\pm\beta$ or $\sqrt{1-\beta^2}$ given by the graph ($\Gamma 3$).

$$M_1 = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad M_2 = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 & \frac{\sqrt{3}}{2} \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & \frac{-1}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \frac{\sqrt{3}}{2} & 0 & 0 & 0 & \frac{-1}{2} \end{bmatrix}$$

$$M_3 = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{\sqrt{8}}{3} & 0 \\ 0 & 0 & \frac{\sqrt{8}}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad M_4 = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{\sqrt{3}}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

Young's theorem is that these matrices represent the symmetric group, i.e. satisfy the braid relations.

Only the relations $M_iM_{i+1}M_i = M_{i+1}M_iM_{i+1}$ are not immediate, but their proof amounts to check Yang-Baxter equation (see below), and finally reduces to the fact that the "contents" choosen by Young are a "distance" satisfying the additivity:

$$(a-b)+(b-c)=(a-c)$$

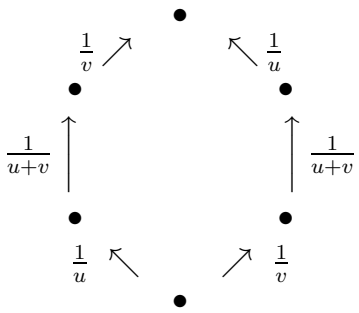
as required by Yang-Baxter's rule of choosing spectral parameters (cf. [Ya], [LLT1]).

To see that Young's matrices can be interpreted as equations in the group algebra of $\mathfrak{S}_n$, one needs now another reading of the graph. The vertices, whatever equivalent labelling is choosen (tableaux, contents, etc.), will be interpreted as element of the group algebra of $\mathfrak{S}_n$, and an oriented edge of color i and parameter β will be

"multiplication by $(s_i+\beta)\frac{1}{\sqrt{1-\beta^2}}$**"**

Non-physicists usually require $\beta \neq \frac{1}{0}, 1$.

The braid relation $M_iM_{i+1}M_i = M_{i+1}M_iM_{i+1}$ is now seen graphically as



that is (*Yang-Baxter equation*)

$$(s+\frac{1}{u})\,(s'+\frac{1}{u+v})\,(s+\frac{1}{v}) = (s'+\frac{1}{v})\,(s+\frac{1}{u+v})\,(s'+\frac{1}{u}) \qquad (\star)$$

which is checked directly in the group algebra of $\mathfrak{S}_3$.

Instead of using the basis $1, s, s', ss', s's, ss's$ of the group algebra of $\mathfrak{S}_3$, one prefers the Yang-Baxter basis $\quad 1,\ (s+\frac{1}{u}),\ (s'+\frac{1}{v}),\ (s+\frac{1}{u})\,(s'+\frac{1}{u+v}),$ $(s'+\frac{1}{v})\,(s+\frac{1}{u+v}),\ (s+\frac{1}{u})\,(s'+\frac{1}{u+v})\,(s+\frac{1}{v})$, the change of basis being triangular.

We shall keep the same edges, but instead of starting from a bottom element equal to 1, we shall use a more sophisticated element to get irreducible representations.

Let us first notice that there are 2 distinguished elements in $\mathbb{C}[\mathfrak{S}_n]$, which correspond to the two 1-dimensional representations:

$$\begin{cases} \clubsuit & = & \sum_{\sigma\in\mathfrak{S}} \sigma \\ \nabla & = & \sum_{\sigma\in\mathfrak{S}} (-1)^{\ell(\sigma)}\, \sigma \end{cases}$$

One has more general elements by taking sums over a Young subgroup (that is, a direct product of symmetric groups diagonally embedded).

Thus for any composition J of n (denoting the order of the successive symmetric group composing a Young subgroup), one has two elements

$$\clubsuit_J \ , \ \nabla_J \ .$$

Studying the product (that he denoted $P(t)N(t)$) of the sum of all permutations preserving the rows of a given tableau, multiplied by the alternating sum of the permutations fixing the columns of the same tableau, Young ([Y], QS1) found 1-dimensional spaces from which it is easy to write idempotents. This construction can be rewritten as follows.

Lemma(Young). *For two conjugate compositions I, J, then the spaces $\clubsuit_J\, \mathbb{C}[\mathfrak{S}_n]\, \nabla_I$ and $\nabla_I\, \mathbb{C}[\mathfrak{S}_n]\, \clubsuit_J$ are 1-dimensional.*

Take J to be a partition λ, and let μ be the conjugate partition. Then

$$\clubsuit_\lambda\ \mathbb{C}[\mathfrak{S}_n]\ \nabla_\mu\ \mathbb{C}[\mathfrak{S}_n]\ \clubsuit_\lambda$$

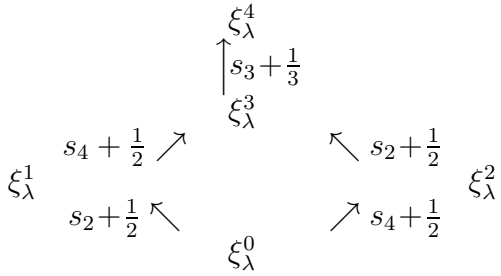
is still 1-dimensional. It contains a unique idempotent ξ_λ. Similarly

$$\nabla_\mu \clubsuit_\lambda\ \mathbb{C}[\mathfrak{S}_n]\ \clubsuit_\lambda\ \mathbb{C}[\mathfrak{S}_n]\ \nabla_\mu$$

contains another unique idempotent ξ_λ^{top}.

To have a precise expression of these two idempotents (i.e. to give their decomposition in the basis of permutations), one needs to handle double cosets. Non reduced expressions are easier to write : Young took, up to scalars, $P(t)N(t)P(t)$ and $N(t')P(t')N(t')$, with t and t' the bottom and top tableaux respectively (in the second graph). One can also use a coding of elements of the group algebra as polynomials to obtain compact expressions (cf. [La?]).

Now we are ready to describe the orthogonal idempotents. Take the graph with vertices labelled by the contents of Young tableaux of shape λ, put $\xi_\lambda^0 := \xi_\lambda$ in the bottom place and play the Yang-Baxter game : a path in the graph means a product of factors $(s_i + \beta)/\sqrt{1-\beta^2}$, as indicated by the labelling of edges.



(Γ4)

For any pair of vertices t, t', let $\phi(t, t')$ be the evaluation of a path from t to t' in the Yang-Baxter graph (with proper normalization, i.e. adding the factors $1/\sqrt{1-\beta^2}$ that we have not written).

Define

$$e_{t,t'} := \phi(t, t^0)\, \xi_\lambda\, \phi(t^0, t')$$

Then the $e_{t,t'}$ are **Young's orthogonal units**, the $e_{t,t}$ are **Young's orthogonal idempotents** ([Y],QS4); they are a complete system of mutually orthogonal idempotents. Thus, the orthogonal idempotents (written ξ_λ^i instead of $e_{t,t}$) are obtained from the bottom one by following any path (even non reduced), and multiplying left and right, at each step, by the factor $(s_i + \beta)/\sqrt{1-\beta^2}$ or $(s_i - \beta)/\sqrt{1-\beta^2}$ according to the orientation of the edge.

One could also have started from the top of the graph, using the idempotent ξ_λ^{top} instead. As we have said, both these idempotents were explicited by Young. He obtained the others by a not too easy orthonormalization process that we shall bypass ([Y], QS3). Notice that expanding the idempotents in the basis of permutations gives the entries of Young's matrices.

Modern proof is very easy, because it was realized by Jucys [Ju] that Young had found in fact a *Gelfand-Tsetlin basis* of the group algebra. The $e_{t,t'}$ are simultaneous two-sided eigenvectors of the *Jucys-Murphy* elements which constitute a maximal commutative subalgebra of $\mathbb{C}[\mathfrak{S}_n]$ (Jucys-Murphy elements are just sums of transpositions: $J_k := \sum_{i<k}(i,k)$).

In fact, Thrall gave an induction rule for the $e_{t,t}$ which renders their characterization immediate. Denote by u the tableau obtained by erasing the last letter n of t. Then one has (up to normalization; [Th])

$$e_{t,t} = e_{u,u}\, P(t) N(t)\, e_{u,u} \qquad (\star\star)$$

This is checked by having the Jucis-Murphy elements act on both sides of $(\star\star)$. For $J_1, \ldots, J_{n-1}$, it is known by induction on n that $e_{u,u}$, and thus $e_{t,t}$, is a two-sided eigenvector. To add the action of last element J_n, one can as well take $J_1 + \cdots + J_n$, but it is a central element, and thus easy to handle.

I shall finish by a last description, more in spirit of Young who, as I have already said, was motivated by the classification of polynomials according to their type of symmetry.

Young ([Y], QS3) described what was later called "Specht modules" [Sp], [Ga], by having the symmetric group act on products of Vandermonde determinants associated to tableaux. One can simplify his construction by having recourse to the quotient ring $\mathbb{C}[x_1, \dots, x_n]/\mathfrak{S}ym_+$ ($\mathfrak{S}ym_+$ = ideal generated by symmetric functions without constant term). The *Specht polynomials* are in this interpretation replaced by mere monomials.

The quotient ring, as a representation of $\mathfrak{S}_n$, is isomorphic to the regular representation, i.e., to $\mathbb{C}[\mathfrak{S}_n]$. Thus all the constructions given above can be interpreted in terms of polynomials. Not only it is easier to manipulate polynomials in commutative indeterminates, but moreover, one has many tools on polynomials which can be put into use. I shall refrain from introducing *divided differences*, though they are perfectly fit to describe representations and shall only use a canonical scalar product (with $\Delta := \prod_{i<j}(x_i - x_j)$) :

$$(f, g) := \sum (-1)^{\ell(\sigma)} (fg)^\sigma \frac{1}{\Delta} |_{x_1=0=x_2=\cdots}$$

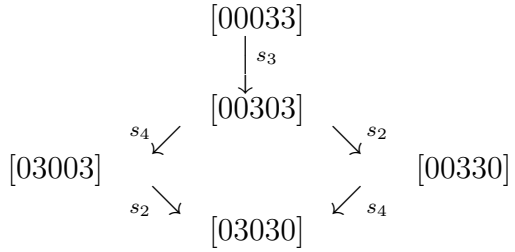
(more generally, one does not need to specialize variables to 0, one should work in the ring of polynomials as a free module over the ring of symmetric polynomials, with quadratic form $\sum (-1)^{\ell(\sigma)} (fg)^\sigma / \Delta$).

Explicitely, the scalar product boils down to

$$(x^u, x^v) = (x^{u+v}, 1) = \begin{cases} (-1)^{\ell(\sigma)} & if\ u+v = [\dots 210]^\sigma \\ 0 & \text{otherwise} \end{cases}$$

that is, one has to test whether the vector $u + v$ is a permutation of $[n-1, \dots, 1, 0]$ or not, and if so, keep the sign of the permutation.

The reader will not be surprised that we have to change the labelling of vertices of our graph : the top vertex is now [00033] (in general, it would be $[0^{\lambda_1}\, \lambda_1^{\lambda_2}\, (\lambda_1 + \lambda_2)^{\lambda_3}\, (\lambda_1 + \lambda_2 + \lambda_3)^{\lambda_4} \cdots]$ for a decreasing partition λ) :



(Γ5)

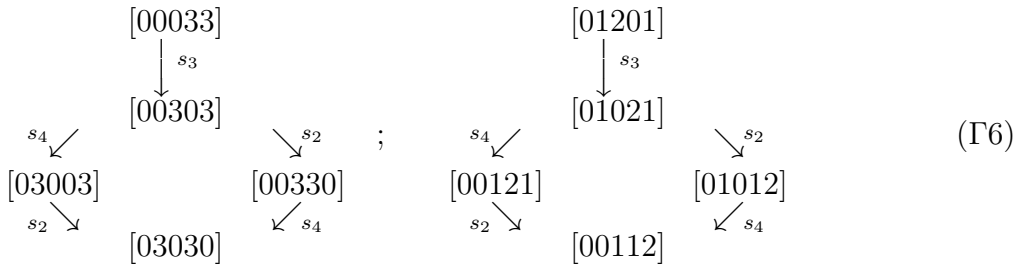
Interpret the vertices as **exponents** of monomials. Then one has the following theorem, for which we refer to [CLL] (a priori one should expect

a "permutation representation", but evaluating modulo $\mathfrak{S}ym_+$ entails irreducibility – in characteristics 0, of course).

Theorem. *For a given partition λ, then the set of monomials x^u read from the graph associated to λ constitute a basis of a copy of the irreducible representation of $\mathfrak{S}_n$ of index λ, the symmetric group acting by permutations of the indeterminates x_i's.*

Matrices of representations, or decompositions of elements in the standard basis (*straightening*, [DKR], [Ga]) reduce to computing scalar products.

One can in fact describe the adjoint basis, which generates the "Specht representation". As a first approximation, it is sufficient to take the "complementary monomials", applying the same transpositions to the top vector $[(0,\dots,\lambda_1-1),(0,\dots,\lambda_2-1),\dots]$ (or, equivalently, replacing repeated letters by $0,1,2,\dots$, from left to right)



(Γ6)

It is not difficult to check that the matrix of scalar products between these two bases is triangular, with ± 1 in the diagonal. One has choosen the second top vector to be such that it sums with the first one is $[0123\dots]$, and acting by the same permutations on both vectors generate the diagonal of the matrix. The proof of the above theorem mostly relies on this elementary fact [CLL]. In fact, this matrix was given by Rutherford [Ru], [DKR], with another interpretation.

For shape [3,2], here are the scalar products between the two monomial bases:

$$Q := \begin{bmatrix} 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & -1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot \\ -1 & \cdot & \cdot & \cdot & -1 \end{bmatrix}$$

Because there are some entries outside the diagonal, one needs to invert the preceding sparse matrix to get the adjoint basis. The matrix representing

a permutation σ will simply be obtained by multiplying the matrix of scalar products $((x^u)^\sigma\ ,\ x^v)$ by the inverse of Q.

For example, the matrix of scalar products with the image of the basis under $\sigma = [54321]$, and the matrix representing this permutation are

$$\begin{bmatrix} 1 & -1 & \cdot & 1 & \cdot \\ \cdot & \cdot & 1 & \cdot & -1 \\ \cdot & \cdot & \cdot & -1 & 1 \\ \cdot & 1 & -1 & -1 & 1 \\ -1 & \cdot & \cdot & \cdot & -1 \end{bmatrix} \quad , \quad \begin{bmatrix} 1 & -1 & \cdot & 1 & \cdot \\ \cdot & \cdot & -1 & \cdot & 1 \\ \cdot & \cdot & \cdot & -1 & 1 \\ \cdot & 1 & -1 & -1 & 1 \\ \cdot & 1 & \cdot & -1 & 1 \end{bmatrix}$$

One has needed only addition of vectors to compute the entries of the matrices of scalar products !

References

[BR] H. BARCELLO, A. RAM, Combinatorial representation theory, New perspectives in algebraic combinatorics, Math. Sci. Res. Inst. Publ., **38**, Cambridge Univ. Press, Cambridge (1999) 23–90.

[CLL] C. CARRE, A. LASCOUX, B. LECLERC, Turbo straightening, *Int J. of Algebra and Comp.* (1992) 275-290.

[DKR] J. DESARMENIEN, J. KUNG, G.C . ROTA, Invariant Theory, Young bitableaux and combinatorics, Adv. in M. **27** (1978) 63–92.

[JK] JAMES, A. KERBER, The representation theory of the symmetric group, Encyclopedia of Maths, vol **16** Addison-Wesley (1981).

[Ju] A.A. JUCYS, On the Young operators on the symmetric groups, Lietuvos fizikos rinkinys VI:2 (1966) 163-180; Factorization of Young projection operators for the symmetric group, ibidem, X1:1 (1971) 5–10.

[Ga] H. GARNIR,Théorie de la représentation des groupes symétriques, Mémoires de la Soc. Royale des Sc. de Liège **10** (1950).

[La] A. LASCOUX, *Young's natural idempotents as polynomials*, Annals of Combinatorics **1** (1997) 91–98.

[LLT1] A. LASCOUX, B. LECLERC, J-Y THIBON, *Flag Varieties and the Yang-Baxter Equation* , Letters in Math. Phys. **40** (1997) 75–90.

[LLT2] A. LASCOUX, B. LECLERC, J-Y THIBON, The Plactic Monoid, in "Combinatorics On Words", M.Lothaire, Acad. Press (to appear).

[0V] A. OKOUNKOV, A. VERSHIK, A new approach to representation theory of symmetric groups, Selecta Math. **2** (1996) 581–605.

[Ro] G. C. ROTA, Combinatorial Theory and Invariant Theory, Notes by L.Guibas, Bowdoin College (1971).

[Ru] D. RUTHERFORD, Substitutional Analysis, Edinburgh University Press (1948).

[Sp] W. SPECHT, Die irreduziblen darstellungen der symmetrischen gruppe, Math. Z. **39** (1935) 606–711.
[Th] R. M. THRALL, Young's semi normal representation of the symmetric group, Duke M. **8** (1941) 611–624.
[Ya] C.N. YANG, Some exact results for the many-body problem in one dimension with repulsive delta-function interaction, Phys. Rev. Lett. **19** (1967), 1312-1315.
[Y] A. YOUNG, The Collected Papers of Alfred Young, University of Toronto Press (1977). Representations are described in the eight articles *On Quantitative Substitutional Analysis*, published in the Proc. London Math. Soc. We refer more specially to QS1, **33** (1900) 97–146; QS3, **28** (1928) 255–292; QS4, **31** (1930) 253–272.