Recent Progress in Algebraic Combinatorics

Richard P. Stanley[1]
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA 02139
e-mail: rstan@math.mit.edu

version of 19 March 2002

1 Introduction.

Algebraic combinatorics is alive and well at the dawn at the new millenium. Algebraic combinatorics is difficult to define precisely; roughly speaking it involves objects that can be interpreted both combinatorially and algebraically, e.g., as the cardinality of a combinatorially defined set and the dimension of an algebraically defined vector space. Sometimes the combinatorial interpretation is used to obtain an algebraic result, and sometimes *vice versa*. Mathematicians have been engaged in algebraic combinatorics at least since Euler (in particular, his work on partitions), but it wasn't until the 1960's, primarily under the influence of Gian-Carlo Rota, that there was a systematic attempt to establish the foundations of algebraic combinatorics and bring it into the mathematical mainstream. This effort has been highly successful, and algebraic combinatorics has by now become a mature and thriving discipline.

We have chosen three major breakthroughs to highlight recent work in algebraic combinatorics. All three areas have initiated a flurry of further work and suggest many further directions of research to keep practitioners of algebraic combinatorics occupied well into the new century. Our choice of topics was partially influenced by the relative ease in describing the main results to nonexperts in algebraic combinatorics. Much other outstanding work has been done that is not discussed here.

[1]Partially supported by NSF grant #DMS-9988459.

2 The saturation conjecture.

The saturation conjecture concerns certain integers known as *Littlewood-Richardson coefficients*. Given the theme of this paper, it is not surprising that they have both an algebraic and a combinatorial definition. First we discuss the algebraic definition, which is more natural than the combinatorial one.

Let $\mathrm{GL}(n,\mathbb{C})$ denote the group of all invertible transformations from an n-dimensional complex vector space V to itself. After choosing an ordered basis for V we may identify $\mathrm{GL}(n,\mathbb{C})$ with the group of $n \times n$ nonsingular matrices over the complex numbers (with the operation of matrix multiplication). Consider the map $\varphi : \mathrm{GL}(2,\mathbb{C}) \to \mathrm{GL}(3,\mathbb{C})$ defined by

$$\varphi \left[\begin{array}{cc} a & b \\ c & d \end{array} \right] = \left[\begin{array}{ccc} a^2 & 2ab & b^2 \\ ac & ad+bc & bd \\ c^2 & 2cd & d^2 \end{array} \right].$$

This can be checked to be a group homomorphism (and hence a *representation* of $\mathrm{GL}(2,\mathbb{C})$ of degree 3). Moreover, the entries of $\varphi(A)$ are polynomial functions of the entries of A. Hence φ is a *polynomial representation* of $\mathrm{GL}(2,\mathbb{C})$. If $A \in \mathrm{GL}(2,\mathbb{C})$ has eigenvalues x, y, then it can also be checked that $\varphi(A)$ has eigenvalues x^2, xy, y^2. Define the *character* $\operatorname{char}\varphi$ of φ to be the trace of $\varphi(A)$, regarded as a function of the eigenvalues x, y of A. Hence

$$\operatorname{char}\varphi = x^2 + xy + y^2.$$

It was first shown by Schur that the polynomial representations of $\mathrm{GL}(n,\mathbb{C})$ are *completely reducible*, i.e., a direct sum of irreducible representations. The nequivalent irreducible polynomial representations φ_λ of $\mathrm{GL}(n,\mathbb{C})$ are indexed by *partitions* $\lambda = (\lambda_1, \ldots, \lambda_n)$ of length at most n, i.e., $\lambda_i \in \mathbb{Z}$ and $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$. Moreover, $\operatorname{char}\varphi_\lambda$ is a symmetric function $s_\lambda(x_1, \ldots, x_n)$ that had been originally defined by Cauchy and Jacobi and is now known as a *Schur function*. A well-known property of Schur functions is their *stability*:

$$s_\lambda(x_1, \ldots, x_n, 0) = s_\lambda(x_1, \ldots, x_n).$$

For this reason we can let $n \to \infty$ and consider the Schur function s_λ in *infinitely* many variables $x_1, x_2, \ldots$ and specialize to $x_1, \ldots, x_n$ when dealing with $\mathrm{GL}(n,\mathbb{C})$. For more information on symmetric functions and the representation theory of $\mathrm{GL}(n,\mathbb{C})$, see [8][30][37].

If $A : V \to V$ and $B : W \to W$ are linear transformations on finite-dimensional vector spaces, then

$$\mathrm{tr}(A \otimes B) = \mathrm{tr}(A) \cdot \mathrm{tr}(B),$$

where $A \otimes B$ denotes the tensor (or Kronecker) product of A and B, acting on $V \otimes W$. Hence if λ, μ, and ν are partitions and we set

$$c_{\mu\nu}^{\lambda} = \mathrm{mult}(\varphi_\lambda, \varphi_\mu \otimes \varphi_\nu),$$

the multiplicity of φ_λ in the tensor product $\varphi_\mu \otimes \varphi_\nu$ (when written as a direct sum of irreducible representations), then

$$s_\mu s_\nu = \sum_\lambda c_{\mu\nu}^{\lambda} s_\lambda.$$

The nonnegative integers $c_{\mu\nu}^{\lambda}$ are known as *Littlewood-Richardson coefficients*, and the *Littlewood-Richardson rule* [8, Ch. 5][30, §I.9][37, Appendix A1.3] gives a combinatorial interpretation of them (which we will not state here). If m is a positive integer and $\lambda = (\lambda_1, \lambda_2, \ldots)$ a partition, then write $m\lambda = (m\lambda_1, m\lambda_2, \ldots)$.

Saturation conjecture. If $c_{m\mu,m\nu}^{m\lambda} \neq 0$, then $c_{\mu\nu}^{\lambda} \neq 0$.

The saturation conjecture was proved recently by Allen Knutson and Terence Tao [26][27] using a new *honeycomb model* for describing Littlewood-Richardson coefficients. An elegant exposition of the proof was given by Anders Buch [5], and a detailed survey of all the material in this section (and more) was given by William Fulton [9]. A proof of the saturation conjecture based on representations of quivers was later given by Harm Derksen and Jerzy Weyman [7].

Why is the proof of the saturation conjecture an important breakthrough? The answer is that it is related in a surprising way to a number of other topics. The first concerns the eigenvalues of hermitian matrices. Let A, B, C be $n \times n$ hermitian matrices. Hence their eigenvalues are real. Denote the eigenvalues of A as

$$\alpha \; : \; \alpha_1 \geq \cdots \geq \alpha_n,$$

and similarly β and γ for B and C. Considerable attention has been given to the following problem.

Problem. Characterize those triples (α, β, γ) for which there exist hermitian matrices $A + B = C$ with eigenvalues α, β, and γ.

By taking traces we see that

$$\sum \gamma_i = \sum \alpha_i + \sum \beta_i. \tag{1}$$

After much work by a number of researchers, A. Horn conjectured a complete characterization of triples (α, β, γ), consisting of (1) together with linear inequalities of the form

$$\sum_{k \in K} \gamma_k \leq \sum_{i \in I} \alpha_i + \sum_{j \in J} \beta_j, \tag{2}$$

for certain sets

$$I, J, K \subset \{1, \ldots, n\}, \quad |I| = |J| = |K|.$$

For instance, when $n = 2$ Horn's inequalities (which are easy to show that together with (1) characterize (α, β, γ) in this case) become

$$\begin{aligned} \gamma_1 &\leq \alpha_1 + \beta_1 \\ \gamma_2 &\leq \alpha_2 + \beta_1 \\ \gamma_2 &\leq \alpha_1 + \beta_2. \end{aligned}$$

For $n = 3$ there are twelve inequalities, as follows:

$$\begin{aligned} \gamma_1 &\leq \alpha_1 + \beta_1 \\ \gamma_2 &\leq \min(\alpha_1 + \beta_2, \alpha_2 + \beta_1) \\ \gamma_3 &\leq \min(\alpha_1 + \beta_3, \alpha_2 + \beta_2, \alpha_3 + \beta_1) \\ \gamma_1 + \gamma_2 &\leq \alpha_1 + \alpha_2 + \beta_1 + \beta_2 \\ \gamma_1 + \gamma_3 &\leq \min(\alpha_1 + \alpha_2 + \beta_1 + \beta_3, \alpha_1 + \alpha_3 + \beta_1 + \beta_2) \\ \gamma_2 + \gamma_3 &\leq \min(\alpha_1 + \alpha_2 + \beta_2 + \beta_3, \alpha_1 + \alpha_3 + \beta_1 + \beta_3, \alpha_2 + \alpha_3 + \beta_1 + \beta_2). \end{aligned}$$

The connection between the Saturation Conjecture and Horn's conjecture was given by Alexander Klyachko [24].

Theorem. *The Saturation Conjecture implies Horn's conjecture.*

A more precise connection between Littlewood-Richardson coefficients and eigenvalues of hermitian matrices is provided by the following result, implicit in the work of Heckman [22] and more explicit in Klyachko [24].

Theorem. *Let $\alpha, \beta,$ and γ be partitions of length at most n. The Saturation Conjecture implies that the following two conditions are equivalent:*

- $c_{\alpha\beta}^{\gamma} \neq 0$.
- *There exist $n \times n$ hermitian matrices $A + B = C$ with eigenvalues $\alpha, \beta,$ and γ.*

Since equation (2) consists of *linear* inequalities, the two theorems above show that the nonvanishing of $c_{\alpha\beta}^{\gamma}$ depends on (explicit) linear inequalities among the coordinates of α, β, γ. Thus for fixed n the points $(\alpha, \beta, \gamma) \in \mathbb{R}^{3n}$ for which $c_{\alpha\beta}^{\gamma} \neq 0$ are the integer points in a certain convex cone. Hence the subject of polyhedral combinatorics is closely associated with the theory of Littlewood-Richardson coefficients. For further information on this point of view, see [41].

The theorems stated above involve hermitian matrices. It is known [9, Thm. 3] that exactly the same results hold for the class of real symmetric matrices.

There are a number of other situations in which Littlewood-Richardson coefficients play a surprising role. These situations are thoroughly discussed in [9]. We mention one of them here. Given a partition $\lambda = (\lambda_1, \lambda_2, \ldots)$ and a prime p, let G be a (finite) abelian p-group of type λ, i.e.,

$$G \cong \left(\mathbb{Z}/p^{\lambda_1}\mathbb{Z}\right) \times \left(\mathbb{Z}/p^{\lambda_2}\mathbb{Z}\right) \times \cdots.$$

Given further partitions μ and ν, let $g_{\mu\nu}^{\lambda}(p)$ denote the number of subgroups H of G of type μ such that the quotient group G/H has type ν.

Theorem. (a) $g_{\mu\nu}^{\lambda}(p)$ *is a polynomial function of p with integer coefficients.*

(b) *For any prime p we have that $g_{\mu\nu}^{\lambda}(p) \neq 0$ if and only if $c_{\mu\nu}^{\lambda} \neq 0$.*

The polynomial $g_{\mu\nu}^{\lambda}(t)$ is called a *Hall polynomial* after the pioneering work of Philip Hall [20]. Hall established the above theorem, except that in part (b) he only showed that $g_{\mu\nu}^{\lambda}(t)$ vanishes identically (as a polynomial in t) if and only if $c_{\mu\nu}^{\lambda} = 0$. Subsequently Miller Maley [31] showed that the polynomial $g_{\mu\nu}^{\lambda}(t+1)$ has nonnegative coefficients, from which (b) follows. For an exposition of the basic properties of Hall polynomials, see [30, Chs. II and III.2]. The theory of Hall polynomials holds in the more general context of the ring of integers (i.e., the unique maximal order) of a division algebra of finite rank over a $\mathfrak{p}$-adic field [30, Remark 3, p. 179] or even more generally for q-primary lattices [38, Thm. 4.81].

3 The $n!$ and $(n+1)^{n-1}$ conjectures.

The $n!$ and $(n+1)^{n-1}$ conjectures concern the action of the symmetric group $\mathfrak{S}_n$ on two sets $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ of n variables. In order to appreciate these conjectures, knowledge of the situation for one set of n variables is of value. We therefore first review this theory (for which the proofs are *much* easier). $\mathfrak{S}_n$ acts on the polynomial ring $A = \mathbb{C}[x_1, \ldots, x_n]$ by permuting variables, i.e., for $w \in \mathfrak{S}_n$ let $w \cdot x_i = x_{w(i)}$ and extend to all of A in the obvious way. Let

$$A^{\mathfrak{S}_n} = \{f \in A \,:\, w \cdot f = f \quad \forall w \in \mathfrak{S}_n\},$$

the *ring of invariants* of the action of $\mathfrak{S}_n$ on A. The invariant polynomials $f \in A^{\mathfrak{S}_n}$ are the *symmetric polynomials* in the variables $x_1, \ldots, x_n$ (over $\mathbb{C}$). The "fundamental theorem of symmetric functions" asserts that

$$A^{\mathfrak{S}_n} = \mathbb{C}[e_1, \ldots, e_n],$$

a polynomial ring in the algebraically independent elementary symmetric functions

$$e_k = \sum_{1 \leq i_1 < \cdots < i_k \leq n} x_{i_1} \cdots x_{i_k}.$$

Regard n as fixed and define the ring

$$R = A/(e_1, \ldots, e_n).$$

The ring R inherits the usual grading from A, i.e.,

$$R = R_0 \oplus R_1 \oplus \cdots,$$

where R_i is spanned by (the images of) all homogeneous polynomials of degree i in the variables $x_1, \ldots, x_n$. Because the generators $e_1, \ldots, e_n$ of $R^{\mathfrak{S}_n}$ are algebraically independent of degrees $1, 2, \ldots, n$, it is easy to see that

$$\dim_{\mathbb{C}} R = n!,$$

and more generally,

$$\sum_i \dim_{\mathbb{C}}(R_i)\, q^i = (1+q)(1+q+q^2)\cdots(1+q+\cdots+q^{n-1}), \qquad (3)$$

the standard "q-analogue" of $n!$.

Since the ideal $(e_1, \ldots, e_n)$ of R is $\mathfrak{S}_n$-invariant, $\mathfrak{S}_n$ acts on R. Moreover, this action respects the grading of R, i.e., $w \cdot R_i = R_i$ for all $w \in \mathfrak{S}_n$. Thus R is in fact a *graded* $\mathfrak{S}_n$*-module*, and we can ask, as a refinement of (3), for the multiplicity of each irreducible representation of $\mathfrak{S}_n$ in R_i. For the action on R as a whole the situation is simple to describe (and not difficult to prove): R affords the *regular representation* of $\mathfrak{S}_n$, i.e., the multiplicity of each irreducible representation is its degree (or dimension).

To describe the $\mathfrak{S}_n$-module structure of R_i, we need some understanding of the (inequivalent) irreducible representations of $\mathfrak{S}_n$. They are indexed by partitions λ of n (denoted $\lambda \vdash n$), i.e, $\lambda = (\lambda_1, \ldots, \lambda_\ell) \in \mathbb{Z}^\ell$ where $\lambda_1 \geq \cdots \geq \lambda_\ell > 0$ and $\sum \lambda_i = n$. The dimension of the irreducible $\mathfrak{S}_n$-module M_λ indexed by $\lambda \vdash n$ is denoted by f^λ and is equal to the number of *standard Young tableaux* (SYT) of shape λ, i.e., the number of ways to insert the numbers $1, 2, \ldots, n$ (without repetition) into an array of shape λ (i.e., left-justified with λ_i entries in row i) so that every row and column is increasing. For instance $f^{(3,2)} = 5$, as shown by the five SYT

$$\begin{array}{ccccc} 1\,2\,3 & 1\,2\,4 & 1\,2\,5 & 1\,3\,4 & 1\,3\,5 \\ 4\,5 & 3\,5 & 3\,4 & 2\,5 & 2\,4 \end{array}.$$

There is also a simple explicit formula (e.g., [30, Exam. I.5.2][37, Cor. 7.21.6]), known as the *hook-length formula*, for f^λ.

Since R affords the regular representation of $\mathfrak{S}_n$, the multiplicity of M_λ in R is equal to f^λ. Thus we would like to describe the multiplicity of M_λ in R_i as the number of SYT T of shape λ with some additional property depending on i. This property is the value of the *major index* of T, denoted $\mathrm{MAJ}(T)$. It is defined by

$$\mathrm{MAJ}(T) = \sum_{i+1 \text{ below } i \text{ in } T} i,$$

where the sum ranges over all entries i of T such that $i+1$ appears in a lower row than i. For instance, the SYT of shape $(3,2,2)$ shown below has $\mathrm{MAJ}(T) = 2+3+6 = 11$.

$$T = \begin{array}{ll} 1 & \mathbf{2} \; \mathbf{6} \\ \mathbf{3} & 5 \\ 4 & 7 \end{array} .$$

The following result is due independently to Lusztig (unpublished) and Stanley [36, Prop. 4.11].

Theorem. *Let $\lambda \vdash n$. Then*

$$\mathrm{mult}(M_\lambda, R_i) = \#\{\text{SYT } T : \mathrm{shape}(T) = \lambda,\ \mathrm{MAJ}(T) = i\}.$$

For example, let $n = 5$. There are three SYT with five entries and major index 3, namely,

$$\begin{array}{lll} 1\,2\,\mathbf{3}\,5 & 1\,2\,\mathbf{3} & \mathbf{1}\,4\,5 \\ 4 & 4\,5 & \mathbf{2} \\ & & 3 \end{array} .$$

It follows that

$$R_3 \cong M_{41} \oplus M_{32} \oplus M_{311}.$$

There is another description of R which leads to a different generalization to two sets of n variables. Given any polynomial $P(x_1, \ldots, x_n)$ over $\mathbb{C}$, define ∂P to be the complex vector space spanned by P and all its partial derivatives of all orders. For instance $\partial(x+y)^2$ has dimension three, one basis being $\{(x+y)^2, x+y, 1\}$. Let

$$V_n = \prod_{1 \le i < j \le n} (x_i - x_j). \tag{4}$$

It is easy to see that

$$R \cong \partial V_n$$

as graded $\mathfrak{S}_n$-modules. In particular, $\dim(\partial V_n) = n!$ and ∂V_n affords the regular representation of $\mathfrak{S}_n$.

Adriano Garsia and Mark Haiman had the idea of generalizing the above constructions of R and ∂V_n to two sets $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ of n variables. For the first generalization, let $\mathfrak{S}_n$ act *diagonally* on $B = \mathbb{C}[x, y]$, i.e.,

$$w \cdot x_i = x_{w(i)}, \quad w \cdot y_i = y_{w(i)}.$$

Let

$$B^{\mathfrak{S}_n} = \{f \in B \,:\, w \cdot f = f \quad \forall w \in \mathfrak{S}_n\},$$

the ring of invariants of the action of $\mathfrak{S}_n$ on B. It is no longer the case that $B^{\mathfrak{S}_n}$ is generated by algebraic independent elements. (For general information about rings of invariants of finite groups, see for instance [35][36].) However, we can still define

$$R^{(2)} = S/I,$$

where I is the ideal of B generated by elements of $B^{\mathfrak{S}_n}$ with zero constant term. The $(n+1)^{n-1}$ conjecture of Garsia and Haiman [12][13] was recently proved by Haiman [19], based on techniques he developed to prove the $n!$ conjecture discussed below, together with a theorem of Bridgeland, King, and Reid on the McKay correspondence.

Theorem ($(n+1)^{n-1}$ conjecture). $\dim_{\mathbb{C}} R^{(2)} = (n+1)^{n-1}$

Just as R had the additional structure of a graded $\mathfrak{S}_n$-module, similarly $R^{(2)}$ is a *bigraded* $\mathfrak{S}_n$-module. In other words,

$$R^{(2)} = \bigoplus_{i,j} R^{(2)}_{ij} \quad \text{(vector space direct sum)},$$

where $R^{(2)}_{ij}$ is the subspace of $R^{(2)}$ spanned by (the images of) polynomials that are homogeneous of degree i in the x variables and degree j in the y variables, and moreover $R^{(2)}_{ij}$ is invariant under the action of $\mathfrak{S}_n$ on $R^{(2)}$. For instance, when $n = 4$ it can be computed that

$$R^{(2)}_{2,1} \cong 2M_{211} \oplus M_{22} \oplus M_{31}.$$

In particular,

$$\dim_{\mathbb{C}} R^{(2)}_{2,1} = 2f^{211} + f^{22} + f^{31} = 2 \cdot 3 + 2 + 3 = 12.$$

Garsia and Haiman stated in [11] (see also [17, Conj. 7.5]) a complicated conjectured formula for $\mathrm{mult}(M_\lambda, R^{(2)}_{ij})$. Haiman's proof of the $(n+1)^{n-1}$ conjecture mentioned above actually establishes this stronger conjecture of Garsia and Haiman. A consequence of Haiman's result asserts the following [11][17, p. 246]. Let Γ be the anti-invariant subspace of $R^{(2)}$, i.e.,

$$\Gamma = \{f \in R^{(2)} \,:\, w \cdot f = \mathrm{sgn}(w) f \quad \forall f \in \mathfrak{S}_n\},$$

where $\mathrm{sgn}(w)$ denotes the sign of the permutation w. Then

$$\dim_{\mathbb{C}} \Gamma = \frac{1}{n+1}\binom{2n}{n},$$

a *Catalan number*. James Haglund [16] conjectured and Garsia and Haglund [10] proved a combinatorial interpretation of the Γ bigrading, i.e., a combinatorial interpretation of the numbers $\dim_{\mathbb{C}} \Gamma_{ij}$. For some information on the ubiquitious appearance of Catalan (and related) numbers throughout mathematics, see [37, Exer. 6.19–6.38] and the addendum at www-math.mit.edu/∼rstan/ec.html.

The number $\dim_{\mathbb{C}} R^{(2)} = (n+1)^{n-1}$ has a number of combinatorial interpretations, e.g., it is the number of forests of rooted trees on n vertices [37, Prop. 5.3.2] or the number of parking functions of length n [37, Exer. 5.49]. It is natural to ask whether one can give a combinatorial interpretation of $\dim_{\mathbb{C}} R^{(2)}_{ij}$ that refines some known interpretation of $(n+1)^{n-1}$. At present this question is open.

We turn to the second generalization of R due to to Garsia and Haiman. First we need to define a generalization of the Vandermonde product (4) to two sets of variables. Let $\mu \vdash n$. Coordinatize the squares of the diagram of μ by letting $(i-1, j-1)$ be the coordinate of the square in the ith row and jth column. For instance, the coordinates of the squares of the diagram of $\mu = (3, 2)$ are given by

0,0	0,1	0,2
1,0	1,1	

Let $(i_1, j_1), \ldots, (i_n, j_n)$ be the coordinates of the squares of the diagram of μ (in some order), and define the $n \times n$ determinant

$$D_\mu = \left| x_r^{i_s} y_r^{j_s} \right|_{r,s=1,\ldots,n}.$$

For instance,

$$D_{32} = \begin{vmatrix} 1 & y_1 & y_1^2 & x_1 & x_1 y_1 \\ 1 & y_2 & y_2^2 & x_2 & x_2 y_2 \\ 1 & y_3 & y_3^2 & x_3 & x_3 y_3 \\ 1 & y_4 & y_4^2 & x_4 & x_4 y_4 \\ 1 & y_5 & y_5^2 & x_5 & x_5 y_5 \end{vmatrix}.$$

Note that if μ consists of a single row (i.e., μ consists of the single part n) then $D_\mu = V_n(y)$, while if μ consists of a single column then $D_\mu = V_n(x)$.

The $n!$ conjecture of Garsia and Haiman [12][13], later proved by Haiman [18], is the following assertion.

Theorem ($n!$ conjecture). *For any $\mu \vdash n$, we have*

$$\dim_{\mathbb{C}} \partial D_\mu = n!.$$

The space ∂D_μ, just as $R^{(2)}$, is a bigraded $\mathfrak{S}_n$-module. For each $i, j \geq 0$ and $\lambda \vdash n$, we can ask for a "description" of the integer $\mathrm{mult}\Big(M_\lambda, (D_\mu)_{ij}\Big)$. Garsia and Haiman [12][13] gave such a description, and Haiman [17, Thm. 5.4] showed that it actually followed from the $n!$ conjecture. The Garsia-Haiman description involves the theory of *Macdonald symmetric functions*, a generalization of Schur functions due to I. G. Macdonald [29][30, Ch. VI] and currently of great interest in several different areas, such as the representation theory of quantum groups, affine Hecke algebras, and the Calegero-Sutherland model in particle physics (see [18] for references). We won't

define Macdonald symmetric functions here but will give a brief indication of Haiman's result.

Let $\lambda, \mu \vdash n$. The coefficient of $x^\mu = x_1^{\mu_1} x_2^{\mu_2} \cdots$ in the Schur function s_λ is known as a *Kostka number*, denoted $K_{\lambda\mu}$, and has a simple combinatorial interpretation in terms of semistandard Young tableaux [30, (5.13)][37, §7.10]. In the theory of Macdonald polynomials there arises naturally a two-parameter generalization $K_{\lambda\mu}(q,t)$ of the Kostka number $K_{\lambda\mu} = K_{\lambda\mu}(0,1)$. *A priori* $K_{\lambda\mu}(q,t)$ is only a rational function of q and t, but Macdonald conjectured that it was a polynomial with nonnegative integer coefficients. In 1996–98 several independent proofs were given that $K_{\lambda\mu}(q,t)$ was indeed a polynomial with integer coefficients, but nonnegativity remained open. Haiman showed the remarkable fact that $K_{\lambda\mu}(q,t)$ is essentially the bigraded Hilbert series for the λ-isotypic component of D_μ. More precisely,

$$t^{b(\mu)} K_{\lambda\mu}(q, 1/t) = \sum_{r,s \geq 0} \text{mult}\left(M_\lambda, (D_\mu)_{r,s}\right)\, t^r q^s,$$

where $b(\mu) = \sum (i-1)\mu_i$. This formula establishes the nonnegativity of the coefficients of $K_{\lambda\mu}(q,t)$, though a combinatorial interpretation of these coefficients remains open.

Hamian's proof is based on the geometry of the Hilbert scheme $\text{Hilb}^n(\mathbb{C}^2)$ of n points in the plane. (Claudio Procesi suggested to Haiman the possible relevance of the Hilbert scheme.) Let X and Y be indeterminates. We can define $\text{Hilb}^n(\mathbb{C}^2)$ as a set by

$$\text{Hilb}^n(\mathbb{C}^2) = \{I \subseteq \mathbb{C}[X,Y] \,:\, \dim_{\mathbb{C}} \mathbb{C}[X,Y]/I = n\},$$

i.e., all ideals I of $\mathbb{C}[X,Y]$ such that the quotient ring $\mathbb{C}[X,Y]/I$ is an n-dimensional vector space. Suppose that $\mathcal{Z} = \{z_1, \ldots, z_n\}$ is a set of n *distinct* points in $\mathbb{C}^2$. Let

$$I_{\mathcal{Z}} = \{f \in \mathbb{C}[X,Y] \,:\, f(z_1) = \cdots = f(z_n) = 0\}.$$

Then $I_{\mathcal{Z}}$ is an ideal of $\mathbb{C}[X,Y]$ such that $\mathbb{C}[X,Y]/I_{\mathcal{Z}}$ can be identified with the space of all functions $f : \mathcal{Z} \to \mathbb{C}$, so $I_{\mathcal{Z}} \in \text{Hilb}^n(\mathbb{C}^2)$. This explains why $\text{Hilb}^n(\mathbb{C}^2)$ is called the Hilbert scheme of n points in the plane — it is a closure of the space of all n-element subsets of $\mathbb{C}^2$. In fact, $\text{Hilb}^n(\mathbb{C}^2)$ has the structure of a smooth irreducible algebraic variety, of dimension $2n$.

The remarkable connections between $\mathrm{Hilb}^n(\mathbb{C}^2)$ and the $n!$ and $(n+1)^{n-1}$ conjectures are too technical to discuss here, but let us give a vague hint or two. Write $H^n = \mathrm{Hilb}^n(\mathbb{C}^2)$. Given a partition $\mu \vdash n$, let U_μ be the set of all ideals $I \in H^n$ such that a basis for $\mathbb{C}[x,y]/I$ consists of the (images of the) monomials $x^h y^k$, where the (h,k)'s are the coordinates for the squares of the diagram of μ. Then the sets U_μ are open, affine, and cover H^n, suggesting the possible relevance of H^n to the $n!$ conjecture. Moreover, for each $I \in H^n$ there is a natural way to associate an n-element *multiset* $\pi(I) \subset \mathbb{C}^2$. The n-element multisets contained in $\mathbb{C}^2$ form an affine variety $\mathrm{Sym}^n(\mathbb{C}^2)$, viz.,

$$\mathrm{Sym}^n(\mathbb{C}^2) = (\mathbb{C}^2)^n/\mathfrak{S}_n = \mathrm{Spec}\ \mathbb{C}[x_1,\ldots,x_n,y_1,\ldots,y_n]^{\mathfrak{S}_n},$$

suggesting the possible relevance of H^n to the $(n+1)^{n-1}$ conjecture. See the papers [17] and [18] for details.

It is natural to ask about generalizing the work of Garsia and Haiman to more than two sets of variables. However, all obvious conjectures turn out to be false. One difficult is that the Hilbert scheme $\mathrm{Hilb}^n(\mathbb{C}^k)$ is no longer smooth for $k > 2$.

The $(n+1)^{n-1}$ and $n!$ conjectures are just the beginning of an amazing edifice of conjectures due to Garsia, Haiman, and their collaborators. For instance, we defined a determinant D_λ when λ is a partition of n, regarded as a certain subset of $\mathbb{N} \times \mathbb{N}$ (where $\mathbb{N} = \{0,1,2,\ldots\}$). In exactly the same way we can define D_X for *any* n-element subset X of $\mathbb{N} \times \mathbb{N}$. Bergeron, Garsia, and Tesler [3] then conjecture (and prove in some special cases) for several classes of subsets X that $\dim_{\mathbb{C}}(\partial D_X) = k_X n!$ for some positive integer k_X; and in fact ∂D_X, regarded as an $\mathfrak{S}_n$-module, affords k_X copies of the regular representation.

4 Longest increasing subsequences.

Let $w = a_1 a_2 \cdots a_n \in \mathfrak{S}_n$. An *increasing subsequence* of w is a subsequence $a_{i_1} a_{i_2} \cdots a_{i_k}$ of w for which $a_{i_1} < a_{i_2} < \cdots < a_{i_k}$. Let $\mathrm{is}_n(w)$ denote the length of the longest increasing subsequence of $w \in \mathfrak{S}_n$. For instance, if $w = 274163958 \in \mathfrak{S}_9$ then $\mathrm{is}_9(w) = 4$, exemplified by the increasing subsequences

2469 and 1358. There has been much recent interest in the behavior of the function $\mathrm{is}_n(w)$. A survey of much of this work has been given by Percy Deift [6].

The first question of interest is the expected value $E(n)$ of $\mathrm{is}_n(w)$, where w ranges uniformly over $\mathfrak{S}_n$. Thus

$$E(n) = \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} \mathrm{is}_n(w).$$

Elementary arguments show that

$$\frac{1}{2}\sqrt{n} \leq E(n) \leq e\sqrt{n},$$

and Hammersley [21, Thm. 4] showed in 1972, using subadditive ergodic theory, that the limit

$$c = \lim_{n \to \infty} \frac{E(n)}{\sqrt{n}}$$

exists. Vershik and Kerov [40] (with the difficult direction $c \geq 2$ shown independently by Logan and Shepp [28]) showed in 1977 that $c = 2$.

The proof of Vershik-Kerov and Logan-Shepp is based on the identity

$$E(n) = \frac{1}{n!} \sum_{\lambda \vdash n} \lambda_1 \left(f^\lambda\right)^2, \tag{5}$$

where $\lambda = (\lambda_1, \lambda_2, \ldots)$ and f^λ denotes the number of SYT of shape λ as in Section 3. Equation (5) is due to Craige Schensted [34] and is an immediate consequence of the Robinson-Schensted-Knuth algorithm; see also [37, Exer. 7.109(a)].

The work of Vershik-Kerov and Logan-Shepp only determines the asymptotic behavior of the expectation of $\mathrm{is}_n(w)$. What about stronger results? A major breakthrough was made by Jinho Baik, Percy Deift, and Kurt Johansson [1], and has inspired much further work. To describe their results, let $\mathrm{Ai}(x)$ denote the *Airy function*, viz., the unique solution to the second-order differential equation

$$\mathrm{Ai}''(x) = x\,\mathrm{Ai}(x),$$

subject to the condition

$$\mathrm{Ai}(x) \sim \frac{e^{-\frac{2}{3}x^{3/2}}}{2\sqrt{\pi}x^{1/4}} \text{ as } x \to \infty.$$

Let $u(x)$ denote the unique solution to the nonlinear third order equation

$$u''(x) = 2u(x)^3 + xu(x), \tag{6}$$

subject to the condition

$$u(x) \sim -\mathrm{Ai}(x), \text{ as } x \to \infty.$$

Equation (6) is known as the *Painlevé II equation*, after Paul Painlevé (1863–1933)[2]. Painlevé completely classified differential equations (from a certain class of second order equations) whose "bad" singularities (branch points and essential singularities) were independent of the initial conditions. Most of the equations in this class were already known, but a few were new, including equation (6).

Now define the *Tracy-Widom distribution* to be the probability distribution on $\mathbb{R}$ given by

$$F(t) = \exp\left(-\int_t^\infty (x-t)u(x)^2\,dx\right). \tag{7}$$

It is easily seen that $F(t)$ is indeed a probability distribution, i.e., $F(t) \geq 0$ and $\int_{-\infty}^{\infty} F(t)dt = 1$. Let χ be a random variable with distribution F, and let χ_n be the random variable on $\mathfrak{S}_n$ defined by

$$\chi_n(w) = \frac{\mathrm{is}_n(w) - 2\sqrt{n}}{n^{1/6}}.$$

We can now state the remarkable results of Baik, Deift, and Johansson.

Theorem. *As $n \to \infty$, we have*

$$\chi_n \to \chi \quad \text{in distribution,}$$

[2]In addition to being a distinguished mathematician, in 1908 Painlevé was the first passenger of Wilbur Wright, during which they set a flight duration record of 70 minutes, and in 1917 and 1925 he held a position equivalent to Prime Minister of France.

i.e., for all $t \in \mathbb{R}$,

$$\lim_{n\to\infty} \mathrm{Prob}(\chi_n \le t) = F(t).$$

Theorem. *For any* $m = 0, 1, 2, \ldots$,

$$\lim_{n\to\infty} E(\chi_n^m) = E(\chi^m).$$

Corollary. *We have*

$$\begin{aligned} \lim_{n\to\infty} \frac{\mathrm{Var}(\mathrm{is}_n)}{n^{1/3}} &= \int t^2\, dF(t) - \left(\int t\, dF(t)\right)^2 \\ &= 0.8132\cdots, \end{aligned}$$

where Var *denotes variance, and*

$$\begin{aligned} \lim_{n\to\infty} \frac{E(\mathrm{is}_n) - 2\sqrt{n}}{n^{1/6}} &= \int t\, dF(t) \\ &= -1.7711\cdots. \end{aligned} \tag{8}$$

The above theorems are a vast refinement of the Vershik-Kerov and Logan-Shepp results concerning $E(n)$, the expectation of $\mathrm{is}_n(w)$. The first theorem gives the entire limiting distribution (as $n \to \infty$) of $\mathrm{is}_n(w)$, while the second theorem gives an asymptotic formula for the mth moment. Note that equation (8) may be rewritten

$$E(n) = 2\sqrt{n} + \alpha n^{1/6} + o\left(n^{1/6}\right),$$

where $\alpha = \int t\, dF(t)$, thereby giving the second term in the asymptotic behavior of $E(n)$.

We will say only a brief word on the proof of the above results, explaining how combinatorics enters into the picture. Some kind of analytic expression is needed for the distribution of $\mathrm{is}_n(w)$. Such an expression is provided by the following result of Ira Gessel [14], later proved in other ways by various persons.

Theorem. *Let*

$$u_k(n) = \#\{w \in \mathfrak{S}_n : \mathrm{is}_n(w) \le k\}$$

$$\begin{aligned} U_k(x) &= \sum_{n\geq 0} u_k(n)\frac{x^{2n}}{n!^2} \\ B_i(x) &= \sum_{n\geq 0} \frac{x^{2n+i}}{n!\,(n+i)!}. \end{aligned}$$

Then

$$U_k(x) = \det\left(B_{|i-j|}(x)\right)_{i,j=1}^{k}.$$

Example. We have

$$\begin{aligned} U_2(x) &= \begin{vmatrix} B_0(x) & B_1(x) \\ B_1(x) & B_0(x) \end{vmatrix} \\ &= B_0(x)^2 - B_1(x)^2. \end{aligned}$$

From this it is easy to deduce that

$$u_2(n) = \frac{1}{n+1}\binom{2n}{n},$$

a Catalan number. This result was first stated by John Michael Hammersley in 1972, with the first published proofs by Knuth [25, §5.1.4] and Rotem [33]. There is a more complicated expression for $u_3(n)$ due to Gessel [14, §7][37, Exer. 7.16(e)], namely,

$$u_3(n) = \frac{1}{(n+1)^2(n+2)}\sum_{j=0}^{n}\binom{2j}{j}\binom{n+1}{j+1}\binom{n+2}{j+2},$$

while no "nice" formula for $u_k(n)$ is known for fixed $k > 3$.

Gessel's theorem reduces the theorems of Baik, Deift, and Johansson to "just" analysis, viz., the Riemann-Hilbert problem in the theory of integrable systems, followed by the method of steepest descent to analyze the asymptotic behavior of integrable systems. For further information see the survey [6] of Deift mentioned above.

The asymptotic behavior of $\mathrm{is}_n(w)$ (suitably scaled) turned out to be identical to the Tracy-Widom distribution $F(t)$ of equation (7). It is natural to ask how the Tracy-Widom distribution arose in the first place. It seems

surprising that such an "unnatural" looking function as $F(t)$ could have arisen independently in two different contexts. Originally the Tracy-Widom distribution arose in connection with the *Gaussian Unitary Ensemble* (GUE). GUE is a certain natural probability distribution on the space of all $n \times n$ hermitian matrices $M = (M_{ij})$, namely,

$$Z_n^{-1} e^{-\mathrm{tr}(M^2)} dM,$$

where Z_n is a normalization constant and

$$dM = \prod_i dM_{ii} \cdot \prod_{i<j} d(\mathrm{Re}\, M_{ij}) d(\mathrm{Im}\, M_{ij}).$$

Let the eigenvalues of M be $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$. The following result marked the eponymous appearance [39] of the Tracy-Widom distribution:

$$\lim_{n\to\infty} \mathrm{Prob}\left(\left(\alpha_1 - \sqrt{2n}\right)\sqrt{2}n^{1/6} \leq t\right) = F(t). \tag{9}$$

Thus as $n \to \infty$, $\mathrm{is}_n(w)$ and α_1 have the same distribution (after scaling).

It is natural to ask, firstly, whether there is a result analogous to equation (9) for the other eigenvalues α_k of the GUE matrix M, and, secondly, whether there is some connection between such a result and the behavior of increasing subsequences of random permutations. A generalization of (9) was given by Tracy and Widom [39] (expressed in terms of the Painlevé II function $u(x)$). The connection with increasing subsequences was conjectured in [1] and proved independently by Borodin-Okounkov-Olshanski [4], Johannson [23], and Okounkov [32]. Given $w \in \mathfrak{S}_n$, define integers $\lambda_1, \lambda_2, \ldots$ by letting $\lambda_1 + \cdots + \lambda_k$ be the largest number of elements in the union of k increasing subsequences of w. For instance, let $w = 247951368$. The longest increasing subsequence is 24568, so $\lambda_1 = 5$. The largest union of two increasing subsequences is 24791368 (the union of 2479 and 1368), so $\lambda_1 + \lambda_2 = 8$. (Note that it is impossible to find a union of length 8 of two increasing subsequences that contains an increasing subsequence of length $\lambda_1 = 5$.) Finally w itself is the union of the three increasing subsequences 2479, 1368, and 5, so $\lambda_1 + \lambda_2 + \lambda_3 = 9$. Hence $(\lambda_1, \lambda_2, \lambda_3) = (5, 3, 1)$ (and $\lambda_i = 0$ for $i > 3$). Readers familiar with the theory of the Robinson-Schensted-Knuth algorithm will recognize the sequence $(\lambda_1, \lambda_2, \ldots)$ as the *shape* of the two standard Young tableaux obtained by applying this algorithm to w, a well-known result of

Curtis Greene [15][37, Thm. A1.1.1]. (In particular, $\lambda_1 \geq \lambda_2 \geq \cdots$, a fact which is by no means obvious.) The result of [4][23][32] asserts that as as $n \to \infty$, λ_k and α_k are equidistributed, up to scaling.

The Tracy-Widom distribution arose completely independently in the behaviour of $\mathrm{is}_n(w)$ and GUE matrices. Is this connection just a coincidence? The work of Okounkov [32] provides a connection, *via* the theory of random topologies on surfaces.

References

[1] J. Baik, P. Deift, and K. Johansson, On the distribution of the length of the longest increasing subsequence of random permutations, *J. Amer. Math. Soc.* **12** (1999), 1119–1178, `math.CO/9810105`[3].

[2] A. Berenstein and A. Zelevinsky, Triple multiplicities for $sl(r+1)$ and the spectrum of the exterior algebra of the adjoint representation, *J. Algebraic Combinatorics* **1** (1992), 7–22.

[3] F. Bergeron, A. Garsia, and G. Tesler, Multiple left regular representations generated by alternants, *J. Combinatorial Theory (A)* **91** (2000), 49–83.

[4] A. Borodin, A. Okounkov, and G. Olshanski, Asymptotics of Plancherel measures for symmetric groups, *J. Amer. Math. Soc.* **13** (2000), 481–515, `math.CO/9905032`.

[5] A. Buch, The saturation conjecture (after A. Knutson and T. Tao), *Enseign. Math.* **46** (2000), 43–60, `math.CO/9810180`.

[6] P. Deift, Integrable systems and combinatorial theory, *Notices Amer. Math. Soc.* **47** (2000), 631–640.

[7] H. Derksen and J. Weyman, Semi-invariants of quivers and saturation for Littlewood-Richardson coefficients, *J. Amer. Math. Soc.* **13** (2000), 467–479.

[8] W. Fulton, *Young Tableaux*, London Mathematical Society Student Texts **35**, Cambridge University Press, Cambridge, 1997.

[9] W. Fulton, Eigenvalues, invariant factors, highest weights, and Schubert calculus, *Bull. Amer. Math. Soc.* **37** (2000), 209–249, `math.AG/9908012`.

[10] A. M. Garsia and J. Haglund, A proof of the q,t-Catalan positivity conjecture, *Discrete Math.*, to appear, `www.math.upenn.edu/~jhaglund`.

[11] A. M. Garsia and M. Haiman, A remarkable q,t-Catalan sequence and q-Lagrange inversion, *J. Algebraic Combin.* **5** (1996), 191–244.

[3] math.AG, math.CO, math.RT and hep-th refer to sections of the LANL preprint archive `xxx.lanl.gov`.

[12] A. M. Garsia and M. Haiman, A graded representation model for Macdonald's polynomials, *Proc. Nat. Acad. Sci. U.S.A.* **90** (1993), 36-7–3610.

[13] A. M. Garsia and M. Haiman, Some natural bigraded S_n-modules and q,t-Kostka coefficients, *Electron. J. Combin.* **3** (1996), RP24.

[14] I. Gessel, Symmetric functions and P-recursiveness, *J. Combinatorial Theory (A)* **53** (1990), 257–285.

[15] C. Greene, An extension of Schensted's theorem, *Advances in Math.* **14** (1974), 254-265.

[16] J. Haglund, Conjectured statistics for the q,t-Catalan numbers, *Advances in Math.*, to appear, `www.math.upenn.edu/~jhaglund`.

[17] M. Haiman, Macdonald polynomials and geometry, in *New perspectives in algebraic combinatorics (Berkeley, CA, 1996–97)* (L. J. Billera, *et al.*, eds.), MSRI Publ. **38**, Cambridge Univ. Press, Cambridge, 1999, pp. 207-254.

[18] M. Haiman, Hilbert schemes, polygraphs, and the Macdonald positivity conjecture, *J. Amer. Math. Soc.* **14** (2001), 941–1006, `www/math.berkeley.edu/~mhaiman`.

[19] M. Haiman, Vanishing theorems and character formulas for the Hilbert scheme of points in the plane, preliminary draft, `www/math.berkeley.edu/~mhaiman`; abbreviated version in *Physics and Combinatorics* (A. N. Kirillov and N. Liskova, eds.), World Scientific, London, 2001, pp. 1–21.

[20] P. Hall, The algebra of partitions, in *Proc. 4th Canadian Math. Congress (Banff)*, 1959, pp. 147–159.

[21] J. M. Hammersley, A few seedlings of research, in *Proc. Sixth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, University of California Press, Berkeley/Los Angeles, 1972, pp. 345–394.

[22] G. J. Heckman, Projections of orbits and asymptotic behavior of multiplicities for compact connected Lie groups, *Invent. Math.* **67** (1982), 333–356.

[23] K. Johansson, Discrete orthogonal polynomial ensembles and the Plancerel measure, *Ann. Math.* **153** (2001), 259–296, `math.CO/9906120`.

[24] A. A. Klyachko, Stable bundles, representation theory and Hermitian operators, *Selecta Math.* **4** (1998), 419–445.

[25] D. E. Knuth, *The Art of Computer Programming*, vol. 3, *Sorting and Searching*, Addison-Wesley, Reading, Massachusetts, 1973; second edition, 1998.

[26] A. Knutson and T. Tao, The honeycomb model of $\mathrm{GL}_n(\mathbb{C})$ tensor products I: proof of the saturation conjecture, *J. Amer. Math. Soc.* **12** (1999), 1055–1090, `math.RT/9807160`.

[27] A. Knutson and T. Tao, Honeycombs and sums of Hermitian matrices, *Notices Amer. Math. Soc.* **48** (2001), 175–186, `math.RT/0009048`.

[28] B. F. Logan and L. A. Shepp, A variational problem for random Young tableaux, *Advances in Math.* **26** (1977), 206–222.

[29] I. G. Macdonald, A new class of symmetric functions, *Actes 20^e Séminaire Lotharingien*, Publ. I.R.M.A., Strasbourg, 1992, pp. 5–39.

[30] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, second ed., Oxford University Press, Oxford, 1995.

[31] F. M. Maley, The Hall polynomial revisited, *J. Algebra* **184** (1996), 363–371.

[32] A. Okounkov, Random matrices and random permutations, *Internat. Math. Res. Notices* **2000**, 1043–1095, `math.CO/9903176`.

[33] D. Rotem, On a correspondence between binary trees and a certain type of permutation, *Inf. Proc. Letters* **4** (1975/76), 58–61.

[34] C. E. Schensted, Longest increasing and decreasing subsequences, *Canad. J. Math.* **13** (1961), 179–191.

[35] L. Smith, *Polynomial Invariants of Finite Groups*, A K Peters, Wellesley, Massachusetts, 1995.

[36] R. Stanley, Invariants of finite groups and their applications to combinatorics, *Bull. Amer. Math. Soc.* (new series) **1** (1979), 475–511.

[37] R. Stanley, *Enumerative Combinatorics*, vol. 2, Cambridge University Press, Cambridge, 1999.

[38] G. Tesler, Semi-primary lattices and tableaux algorithms, Ph.D. thesis, M.I.T., 1995.

[39] C. A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel, *Comm. Math. Phys.* **159** (1994), 151–174, `hep-th/9211141`.

[40] A. M. Vershik and S. V. Kerov, Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux, *Dokl. Akad. Nauk SSSR* **233** (1977), 1024–1027. English translation in *Soviet Math. Dokl.* **18** (1977), 527–531.

[41] A. Zelevinsky, Littlewood-Richardson semigroups, in *New perspectives in algebraic combinatorics (Berkeley, CA, 1996–97)* (L. J. Billera, *et al.*, eds.), MSRI Publ. **38**, Cambridge Univ. Press, Cambridge, 1999, pp. 337–345, `math.CO/9704228`.