Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Advances in Mathematics 225 (2010) 81–120

ADVANCES IN Mathematics

www.elsevier.com/locate/aim

# Explicit combinatorial interpretation of Kerov character polynomials as numbers of permutation factorizations

Maciej Dołęga [a], Valentin Féray [b], Piotr Śniady [a,c,*]

[a] *Institute of Mathematics, University of Wroclaw, pl. Grunwaldzki 2/4, 50-384 Wroclaw, Poland*
[b] *LaBRI, University Bordeaux 1, 351 cours de la Libération, 33400 Talence, France*
[c] *Institute of Mathematics, Polish Academy of Sciences, ul. Śniadeckich 8, 00-956 Warszawa, Poland*

Received 17 October 2008; accepted 9 February 2010

Available online 5 March 2010

Communicated by Dan Voiculescu

**Abstract**

We find an explicit combinatorial interpretation of the coefficients of Kerov character polynomials which express the value of normalized irreducible characters of the symmetric groups $\mathfrak{S}(n)$ in terms of free cumulants $R_2, R_3, \dots$ of the corresponding Young diagram. Our interpretation is based on counting certain factorizations of a given permutation.


*Keywords:* Kerov polynomials; Free cumulants; Representations of symmetric groups; Factorizations of permutations

## 1. Introduction

### 1.1. Generalized Young diagrams

We are interested in the asymptotics of irreducible representations of the symmetric groups $\mathfrak{S}(n)$ for $n \to \infty$ in the scaling of *balanced Young diagrams* which means that we consider a sequence $(\lambda^{(n)})$ of Young diagrams with a property that $\lambda^{(n)}$ has $n$ boxes and $O(\sqrt{n})$ rows and columns. This scaling makes the graphical representations of Young diagrams particularly useful; in this article we will use two conventions for drawing Young diagrams: the French (presented

* Corresponding author at: Institute of Mathematics, University of Wroclaw, pl. Grunwaldzki 2/4, 50-384 Wroclaw, Poland.

*E-mail addresses:* Maciej.Dolega@math.uni.wroc.pl (M. Dołęga), feray@univ-mlv.fr (V. Féray), Piotr.Sniady@math.uni.wroc.pl (P. Śniady).

0001-8708/$ – see front matter 
doi:10.1016/j.aim.2010.02.011

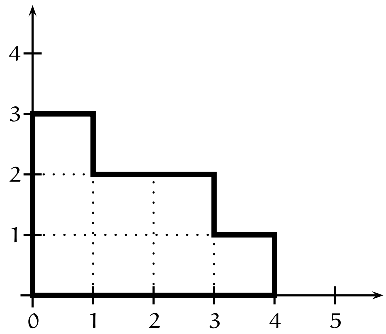


Fig. 1. Young diagram (4, 3, 1) drawn in the French convention.

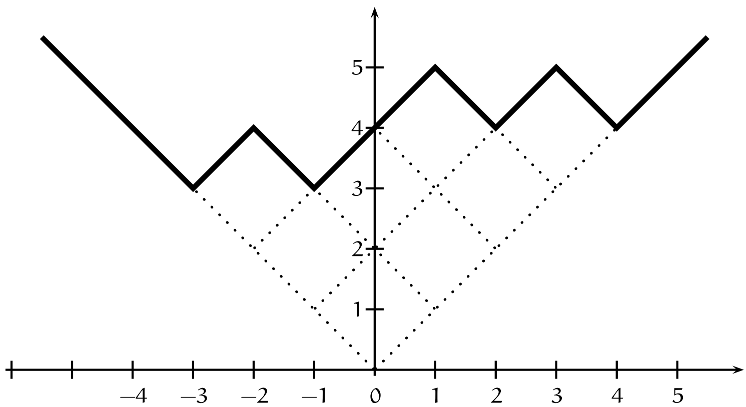


Fig. 2. Young diagram (4, 3, 1) drawn in the Russian convention. The profile of the diagram has been drawn in the solid line.

on Fig. 1) and the Russian one (presented on Fig. 2). Notice that the graphs in the Russian convention are created from the graphs in the French convention by rotating counterclockwise by $\frac{\pi}{4}$ and by scaling by a factor $\sqrt{2}$.

Any Young diagram drawn in the French convention can be identified with its graph which is equal to the set $\{(x, y)\colon\ 0 \leqslant x, 0 \leqslant y \leqslant f(x)\}$ for a suitably chosen function $f : \mathbb{R}_+ \to \mathbb{R}_+$, where $\mathbb{R}_+ = [0, \infty)$. It is therefore natural to define the set of *generalized Young diagrams* $\mathbb{Y}$ *(in the French convention)* as the set of bounded, non-increasing functions $f : \mathbb{R}_+ \to \mathbb{R}_+$ with a compact support; in this way any Young diagram can be regarded as a generalized Young diagram.

We can identify a Young diagram drawn in the Russian convention with its profile, see Fig. 2. It is therefore natural to define the set of *generalized Young diagrams* $\mathbb{Y}$ *(in the Russian convention)* as the set of functions $f : \mathbb{R} \to \mathbb{R}_+$ which fulfill the following two conditions:

- $f$ is a Lipschitz function with constant 1, i.e. $|f(x) - f(y)| \leqslant |x - y|$,
- $f(x) = |x|$ if $|x|$ is large enough.

At the first sight it might seem that we have defined the set $\mathbb{Y}$ of generalized Young diagrams in two different ways, but we prefer to think that these two definitions are just two conventions

(French and Russian) for drawing the same object. This will not lead to confusions since it will be always clear from the context which of the two conventions is being used.

The setup of generalized Young diagrams makes it possible to speak about dilations of Young diagrams. In the geometric language of French and Russian conventions such dilations simply correspond to dilations of the graph. Formally speaking, if $f \in \mathbb{Y}$ is a generalized Young diagram (no matter in which convention) and $s > 0$ is a real number we define the dilated diagram $sf \in \mathbb{Y}$ by the formula

$$(sf)(x) = sf\left(\frac{x}{s}\right).$$

This notion of dilation is very useful in the study of balanced Young diagrams because if $(\lambda^{(n)})_n$ is a sequence of balanced Young diagrams we may for example ask questions about the limit of the sequence $\frac{1}{\sqrt{n}}\lambda^{(n)}$ [17,27].

### 1.2. Normalized characters

Any permutation $\pi \in \mathfrak{S}(k)$ can be also regarded as an element of $\mathfrak{S}(n)$ if $k \leqslant n$ (we just declare that $\pi \in \mathfrak{S}(n)$ has additional $n - k$ fixpoints). For any $\pi \in \mathfrak{S}(k)$ and an irreducible representation $\rho^\lambda$ of the symmetric group $\mathfrak{S}(n)$ corresponding to the Young diagram $\lambda$ we define the normalized character

$$\Sigma_\pi^\lambda = \begin{cases} \underbrace{n(n-1)\cdots(n-k+1)}_{k \text{ factors}} \frac{\operatorname{Tr}\rho^\lambda(\pi)}{\text{dimension of } \rho^\lambda} & \text{if } k \leqslant n, \\ 0 & \text{otherwise.} \end{cases}$$

One of the reasons why such normalized characters are so useful in the asymptotic representation theory is that, as we shall see in Section 4, one can extend the definition of $\Sigma_\pi^\lambda$ to the case when $\lambda \in \mathbb{Y}$ is a generalized Young diagram; furthermore computing their values will turn out to be easy.

Particularly interesting are the values of characters on cycles, therefore we will use the notation

$$\Sigma_k^\lambda = \Sigma_{(1,2,\dots,k)}^\lambda,$$

where we treat the cycle $(1, 2, \dots, k)$ as an element of $\mathfrak{S}(k)$ for any integer $k \geqslant 1$.

### 1.3. Free cumulants

Let $\lambda$ be a (generalized) Young diagram. We define its *free cumulants* $R_2^\lambda, R_3^\lambda, \dots$ by the formula

$$R_k^\lambda = \lim_{s\to\infty} \frac{1}{s^k} \Sigma_{k-1}^{s\lambda}, \tag{1}$$

in other words each free cumulant is asymptotically the dominant term of the character on a cycle of appropriate length in the limit when the Young diagram tends to infinity.

From the above definition it is clear that free cumulants should be interesting for investigations of the asymptotics of characters of symmetric groups, but it is not obvious why the limit should exist and if there is some direct way of calculating it. In Sections 2 and 3 we will review some more conventional definitions of free cumulants and some more direct ways of calculating them.

One of the reasons why free cumulants are so useful in the asymptotic representation theory is that they are homogeneous with respect to dilations of the Young diagrams, namely

$$R_k^{s\lambda} = s^k R_k^{\lambda};$$

in other words the degree of the free cumulant $R_k$ is equal to $k$. This property is an immediate consequence of (1) but it also follows from more conservative definitions of free cumulants.

In fact, the notion of free cumulants origins from the work of Voiculescu [28] where they appeared as coefficients of an $R$-series which turned out to be useful in description of *free convolution* in the context of *free probability theory* [30]. The name of free cumulants was coined by Speicher [23] who found their combinatorial interpretation and their relations with the lattice of non-crossing partitions [22]. Since free probability theory is closely related to the random matrix theory [29] free cumulants quickly became an important tool not only within the framework of free probability but in the random matrix theory as well.

### 1.4. Kerov character polynomials

The following surprising fact is fundamental for this article: it turns out that free cumulants can be used not only to provide asymptotic approximations for the characters of symmetric groups, but also for exact formulas. Kerov during a talk in Institut Henri Poincaré in January 2000 [13] announced the following result (the first published proof was given by Biane [3]): for each permutation $\pi$ there exists a universal polynomial $K_\pi$ with integer coefficients, called *Kerov character polynomial*, with a property that

$$\Sigma_\pi^\lambda = K_\pi\left(R_2^\lambda, R_3^\lambda, \ldots\right) \tag{2}$$

holds true for any (generalized) Young diagram $\lambda$. We say that Kerov polynomial is universal because it does not depend on the choice of $\lambda$. In order to keep the notation simple we make the dependence of the characters and of the free cumulants on $\lambda$ implicit and we write

$$\Sigma_\pi = K_\pi(R_2, R_3, \ldots).$$

As usual, we are mostly concerned with the values of the characters on the cycles, therefore we introduce special notation for such Kerov polynomials

$$\Sigma_k = K_k(R_2, R_3, \ldots).$$

Kerov also found the leading term of the Kerov polynomial:

$$\Sigma_k = R_{k+1} + (\text{terms of degree at most } k-1) \tag{3}$$

which has (1) as an immediate consequence.

The first few Kerov polynomials $K_k$ are as follows [2]:

$$\begin{aligned}
\Sigma_1 &= R_2,\\
\Sigma_2 &= R_3,\\
\Sigma_3 &= R_4 + R_2,\\
\Sigma_4 &= R_5 + 5R_3,\\
\Sigma_5 &= R_6 + 15R_4 + 5R_2^2 + 8R_2,\\
\Sigma_6 &= R_7 + 35R_5 + 35R_3R_2 + 84R_3.
\end{aligned}$$

Based on such numerical evidence Kerov formulated during his talk [13] the following conjecture.

**Conjecture 1.1** *(Kerov). The coefficients of Kerov character polynomials $K_k$ $(k \geqslant 1)$ are* non-negative *integers.*

Biane [3] stated a very interesting conjecture that the underlying reason for positivity of the coefficients of Kerov polynomials is that they are equal to cardinalities of some combinatorial objects. Biane provided also some heuristics what these combinatorial objects could be (we postpone the details until Section 1.11.1).

Since then a number of partial answers were found. Śniady [20] found explicitly the next term (with degree $k-1$) in the expansion (3) (the form of this next term was conjectured by Biane [3]). Goulden and Rattan [9] found an explicit but complicated formula for the coefficients of Kerov polynomials. These results, however, did not shed too much light into possible combinatorial interpretations of Kerov character polynomials.

Some light on the possible combinatorial interpretation of Kerov polynomials was shed by the following result proved by Biane in the aforementioned paper [3] and Stanley [24].

**Theorem 1.2** *(Linear terms of Kerov polynomials). For all integers $l \geqslant 2$ and $k \geqslant 1$ the coefficient of $R_l$ in the Kerov polynomial $K_k$ is equal to the number of pairs $(\sigma_1, \sigma_2)$ of permutations $\sigma_1, \sigma_2 \in \mathfrak{S}(k)$ such that $\sigma_1 \circ \sigma_2 = (1, 2, \dots, k)$ and such that $\sigma_2$ consists of one cycle and $\sigma_1$ consists of $l-1$ cycles.*

For a permutation $\pi$ we denote by $C(\pi)$ the set of cycles of $\pi$. Féray [8] extended the above result to the quadratic terms of Kerov polynomials.

**Theorem 1.3** *(Quadratic terms of Kerov polynomials). For all integers $l_1, l_2 \geqslant 2$ and $k \geqslant 1$ the coefficient of $R_{l_1} R_{l_2}$ in the Kerov polynomial $K_k$ is equal to the number of triples $(\sigma_1, \sigma_2, q)$ with the following properties:*

- *$\sigma_1, \sigma_2$ is a factorization of the cycle; in other words $\sigma_1, \sigma_2 \in \mathfrak{S}(k)$ are such that $\sigma_1 \circ \sigma_2 = (1, 2, \dots, k)$;*
- *$\sigma_2$ consists of two cycles and $\sigma_1$ consists of $l_1 + l_2 - 2$ cycles;*
- *$q : C(\sigma_2) \to \{l_1, l_2\}$ is a surjective map on the two cycles of $\sigma_2$;*
- *for each cycle $c \in C(\sigma_2)$ there are at least $q(c)$ cycles of $\sigma_1$ which intersect non-trivially $c$.*

In fact, Féray [8] managed also to prove positivity of the coefficients of Kerov character polynomials by finding some combinatorial objects with appropriate cardinality, but his proof was so complicated that the resulting combinatorial objects were hardly explicit in more complex cases. We compare this work with our new result in Section 8.

### 1.5. The main result: explicit combinatorial interpretation of the coefficients of Kerov polynomials

The following theorem is the main result of the paper: it gives a satisfactory answer for the Kerov conjecture by providing an explicit combinatorial interpretation of the coefficients of the Kerov polynomials. It was formulated for the first time as a conjecture in June 2008 by Valentin Féray and Piotr Śniady after some computer experiments concerning the coefficient of $R_2^3$ in Kerov polynomials $K_7$ and $K_9$. The original formulation of the conjecture was Theorem 7.1; the form below was pointed out by Philippe Biane in a private communication.

**Theorem 1.4** *(The main result). Let $k \geqslant 1$ and let $s_2, s_3, \dots$ be a sequence of non-negative integers with only finitely many non-zero elements. The coefficient of $R_2^{s_2} R_3^{s_3} \cdots$ in the Kerov polynomial $K_k$ is equal to the number of triples $(\sigma_1, \sigma_2, q)$ with the following properties:*

(a) *$\sigma_1, \sigma_2$ is a factorization of the cycle; in other words $\sigma_1, \sigma_2 \in \mathfrak{S}(k)$ are such that $\sigma_1 \circ \sigma_2 = (1, 2, \dots, k)$;*
(b) *the number of cycles of $\sigma_2$ is equal to the number of factors in the product $R_2^{s_2} R_3^{s_3} \cdots$; in other words $|C(\sigma_2)| = s_2 + s_3 + \cdots$;*
(c) *the total number of cycles of $\sigma_1$ and $\sigma_2$ is equal to the degree of the product $R_2^{s_2} R_3^{s_3} \cdots$; in other words $|C(\sigma_1)| + |C(\sigma_2)| = 2s_2 + 3s_3 + 4s_4 + \cdots$;*
(d) *$q : C(\sigma_2) \to \{2, 3, \dots\}$ is a coloring of the cycles of $\sigma_2$ with a property that each color $i \in \{2, 3, \dots\}$ is used exactly $s_i$ times (informally, we can think that $q$ is a map which to cycles of $C(\sigma_2)$ associates the factors in the product $R_2^{s_2} R_3^{s_3} \cdots$);*
(e) *for every set $A \subset C(\sigma_2)$ which is nontrivial (i.e., $A \neq \emptyset$ and $A \neq C(\sigma_2)$) there are more than $\sum_{i \in A}(q(i) - 1)$ cycles of $\sigma_1$ which intersect $\bigcup A$.*

A careful reader may notice that condition (b) in the above theorem is redundant since it is implied by condition (d); we decided to keep it for the sake of clarity. We postpone presenting interpretations of condition (e) until Sections 1.8 and 1.9.

One can easily see that Theorems 1.2 and 1.3 are special cases of the above result. We decided to postpone the discussion of other applications of this main result until Section 1.12 when more context will be available.

### 1.6. Characters for more complicated conjugacy classes

In order to study characters on more complicated conjugacy classes we will use the following notation. For $k_1, \dots, k_l \geqslant 1$ we define

$$\Sigma^{\lambda}_{k_1, \dots, k_l} = \Sigma^{\lambda}_{\pi},$$

where $\pi \in \mathfrak{S}(k_1 + \cdots + k_l)$ is any permutation with the lengths of the cycles given by $k_1, \dots, k_l$; we may take for example $\pi = (1, 2, \dots, k_1)(k_1 + 1, k_1 + 2, \dots, k_1 + k_2) \cdots$. For simplicity we will often suppress the explicit dependence of $\Sigma_{k_1, \dots, k_l}$ on $\lambda$.

Unfortunately, as it was pointed out by Rattan and Śniady [18], Kerov conjecture is not true for more complicated Kerov polynomials $K_\pi$ for which $\pi$ consists of more than one cycle. However, they conjectured that it would still hold true if the definition (2) of Kerov polynomials was modified as follows.

For $k_1, \ldots, k_l \geqslant 1$ we consider *cumulant* $\kappa^{\mathrm{id}}(\Sigma_{k_1}, \ldots, \Sigma_{k_l})$ *of the conjugacy classes of cycles*. Precise definition of these quantities can be found in [21], for the purpose of this article it is enough to know that their relation to the characters $\Sigma_{k_1,\ldots,k_l}$ is analogous to the relation between classical cumulants of random variables and their moments, as it can be seen on the following examples:

$$\begin{aligned}
\Sigma_r &= \kappa^{\mathrm{id}}(\Sigma_r),\\
\Sigma_{r,s} &= \kappa^{\mathrm{id}}(\Sigma_r, \Sigma_s) + \kappa^{\mathrm{id}}(\Sigma_r)\, \kappa^{\mathrm{id}}(\Sigma_s),\\
\Sigma_{r,s,t} &= \kappa^{\mathrm{id}}(\Sigma_r, \Sigma_s, \Sigma_t) + \kappa^{\mathrm{id}}(\Sigma_r)\kappa^{\mathrm{id}}(\Sigma_s, \Sigma_t) + \kappa^{\mathrm{id}}(\Sigma_s)\kappa^{\mathrm{id}}(\Sigma_r, \Sigma_t)\\
&\quad + \kappa^{\mathrm{id}}(\Sigma_t)\kappa^{\mathrm{id}}(\Sigma_r, \Sigma_s) + \kappa^{\mathrm{id}}(\Sigma_r)\kappa^{\mathrm{id}}(\Sigma_s)\kappa^{\mathrm{id}}(\Sigma_s),
\end{aligned}$$

$$\begin{aligned}
\kappa^{\mathrm{id}}(\Sigma_r) &= \Sigma_r,\\
\kappa^{\mathrm{id}}(\Sigma_r, \Sigma_s) &= \Sigma_{r,s} - \Sigma_r \Sigma_s,\\
\kappa^{\mathrm{id}}(\Sigma_r, \Sigma_s, \Sigma_t) &= \Sigma_{r,s,t} - \Sigma_r \Sigma_{s,t} - \Sigma_s \Sigma_{r,t} - \Sigma_t \Sigma_{r,s} + 2\, \Sigma_r \Sigma_s \Sigma_t.
\end{aligned}$$

As it was pointed out in [21], the above quantities $\kappa^{\mathrm{id}}(\Sigma_r, \Sigma_s, \ldots)$ are very useful in the study of fluctuations of random Young diagrams; in fact they are even more fundamental than the characters $\Sigma_{r,s,\ldots}$ themselves.

**Conjecture 1.5.** *(See Rattan and Śniady [18].) For $k_1, \ldots, k_l \geqslant 1$ there exists a universal polynomial $K_{k_1,\ldots,k_l}$ with nonnegative integer coefficients, called* generalized Kerov polynomial*, such that*

$$(-1)^{l-1}\kappa^{\mathrm{id}}(\Sigma_{k_1}, \ldots, \Sigma_{k_l}) = K_{k_1,\ldots,k_l}(R_2, R_3, \ldots).$$

*The coefficients of this polynomials have some combinatorial interpretation.*

The existence of such a universal polynomial with integer coefficients follows directly from the work of Kerov. The positivity of the coefficients was proved by Féray [8] but his combinatorial interpretation of the coefficients was not very explicit.

In this article will also prove the following generalization of Theorem 1.4 which gives an explicit combinatorial solution to Conjecture 1.5.

**Theorem 1.6.** *Let $k_1, \ldots, k_l \geqslant 1$ and let $s_1, s_2, \ldots$ be a sequence of non-negative integers with only finitely many non-zero elements. The coefficient of $R_2^{s_2} R_3^{s_3} \cdots$ in the generalized Kerov polynomial $K_{k_1,\ldots,k_l}$ is equal to the number of triples $(\sigma_1, \sigma_2, q)$ which fulfill the same conditions as in Theorem* 1.4 *with the following modification: condition* (a) *should be replaced by the following one:*

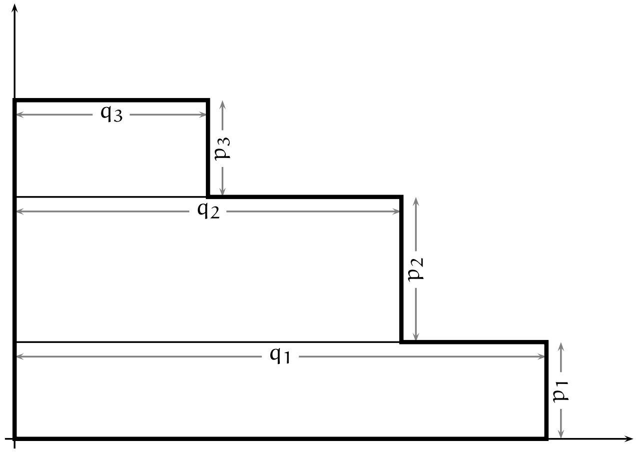


Fig. 3. Generalized Young diagram $\mathbf{p} \times \mathbf{q}$ drawn in the French convention.

(a′) $\sigma_1, \sigma_2 \in \mathfrak{S}(k_1 + \cdots + k_l)$ *are such that*

$$\sigma_1 \circ \sigma_2 = (1, 2, \ldots, k_1)(k_1 + 1, k_1 + 2, \ldots, k_1 + k_2) \cdots$$

*and the group* $\langle \sigma_1, \sigma_2 \rangle$ *acts transitively on the set* $\{1, \ldots, k_1 + \cdots + k_l\}$.

### *1.7. Idea of the proof: Stanley polynomials*

The main idea of the proof of the main result (Theorems 1.4 and 1.6) is to use Stanley polynomials which are defined as follows. For two finite sequences of positive real numbers $\mathbf{p} = (p_1, \ldots, p_m)$ and $\mathbf{q} = (q_1, \ldots, q_m)$ with $q_1 \geqslant \cdots \geqslant q_m$ we consider a multirectangular generalized Young diagram $\mathbf{p} \times \mathbf{q}$, cf. Fig. 3. In the case when $p_1, \ldots, p_m, q_1, \ldots, q_m$ are natural numbers $\mathbf{p} \times \mathbf{q}$ is a partition

$$\mathbf{p} \times \mathbf{q} = (\underbrace{q_1, \ldots, q_1}_{p_1 \text{ times}}, \underbrace{q_2, \ldots, q_2}_{p_2 \text{ times}}, \ldots).$$

If $\mathcal{F} : \mathbb{Y} \to \mathbb{R}$ is a sufficiently nice function on the set of generalized Young diagrams (in this article we use the class of, so called, *polynomial functions*) then $\mathcal{F}(\mathbf{p} \times \mathbf{q})$ turns out to be a polynomial in indeterminates $p_1, p_2, \ldots, q_1, q_2, \ldots$ which will be called *Stanley polynomial*. The Stanley polynomial for the most interesting functions $\mathcal{F}$, namely for the normalized characters $\Sigma_\pi$, is provided by Stanley–Féray character formula (Theorem 4.6) which was conjectured by Stanley [26] and proved by Féray [6], for a more elementary proof we refer to [7].

In the past analysis of some special coefficients of Stanley polynomials resulted in partial results concerning Kerov polynomials [24,25]. In Theorem 4.2 we will show that, in fact, a large class coefficients of Stanley polynomials can be interpreted as coefficients

$$\left. \frac{\partial}{\partial S_{k_1}} \cdots \frac{\partial}{\partial S_{k_l}} \mathcal{F} \right|_{S_2 = S_3 = \cdots = 0}$$

in the Taylor expansion of $\mathcal{F}$ into the *basic functionals* $S_2, S_3, \ldots$ *of shape* of a Young diagram.

These basic functionals $S_2, S_3, \dots$ of shape are not new; they already appeared (possibly with a slightly modified normalization) in the work of Ivanov and Olshanski [10] and implicitly in the work of Kerov [11,12,14].

In this way we are able to express $\Sigma_\pi$ as an explicit polynomial in $S_2, S_3, \dots$. In Proposition 2.2 we will show how to express $S_2, S_3, \dots$ in terms of free cumulants $R_2, R_3, \dots$. Finally, we use some identities fulfilled by Stanley polynomials (Lemma 4.5) in order to express the coefficients of Kerov polynomials in a useful way.

### *1.8. Combinatorial interpretation of condition (e)*

Let $(\sigma_1, \sigma_2, q)$ be a triple which fulfills conditions (a)–(d) of Theorem 1.4. We consider the following polyandrous interpretation of Hall marriage theorem. Each cycle of $\sigma_1$ will be called a *boy* and each cycle of $\sigma_2$ will be called a *girl*. For each girl $j \in C(\sigma_2)$ let $q(j) - 1$ be the *desired number of husbands of* $j$ (notice that condition (c) shows that the number of boys in $C(\sigma_1)$ is right so that if no other restrictions were imposed it would be possible to arrange marriages in such a way that each boy is married to exactly one girl and each girl has the desired number of husbands). We say that a boy $i \in C(\sigma_1)$ is a possible candidate for a husband for a girl $j \in C(\sigma_2)$ if cycles $i$ and $j$ intersect. Hall marriage theorem applied to our setup says that there exists an arrangement of marriages $\mathcal{M} : C(\sigma_1) \to C(\sigma_2)$ which assigns to each boy his wife (so that each girl $j$ has exactly $q(j) - 1$ husbands) if and only if for every set $A \subseteq C(\sigma_2)$ there are at least $\sum_{i \in A} (q(i) - 1)$ cycles of $\sigma_1$ which intersect $\bigcup A$. As one easily see, the above condition is similar but not identical to (e).

**Proposition 1.7.** *Condition* (e) *is equivalent to the following one:*

(e$^2$) *for every nontrivial set of girls $A \subset C(\sigma_2)$ (i.e., $A \neq \emptyset$ and $A \neq C(\sigma_2)$) there exist two ways of arranging marriages $\mathcal{M}_p : C(\sigma_1) \to C(\sigma_2)$, $p \in \{1, 2\}$, for which the corresponding sets of husbands of wives from $A$ are different:*

$$\mathcal{M}_1^{-1}(A) \neq \mathcal{M}_2^{-1}(A).$$

**Proof.** The implication (e$^2$) $\Rightarrow$ (e) is immediate.

For the opposite implication Hall marriage theorem shows existence of $\mathcal{M}_1$. Let us select any boy $i \in \mathcal{M}_1^{-1}(A)$ and let us declare that boy $i$ is not allowed to marry any girl from the set $A$. Applying Hall marriage theorem for the second time shows existence of $\mathcal{M}_2$ with the required properties which finishes the proof of equivalence. □

For permutations $\sigma_1, \sigma_2$ it is convenient to introduce a bipartite graph $\mathcal{V}^{\sigma_1, \sigma_2}$ with the set of vertices $C(\sigma_1) \sqcup C(\sigma_2)$ with edges connecting intersecting cycles [7]. The elements of $C(\sigma_1)$, respectively $C(\sigma_2)$, will be referred to as white, respectively black, vertices. For a bipartite graph with a vertex set $V$ we will denote by $V_\bullet$ the set of black vertices.

The following result gives a strong restriction on the form of the factorizations which contribute to Theorem 1.4 and we hope it will be useful in the future investigations of Kerov polynomials. Notice that this kind of result appears also in the work of Féray [8].

**Proposition 1.8.** *Suppose that $\sigma_1, \sigma_2 \in \mathfrak{S}(k)$ are such that in the graph $\mathcal{V}^{\sigma_1, \sigma_2}$ there exists a disconnecting edge $e$ with a property that each of the two connected components of the resulting* *truncated graph $\mathcal{V}^{\sigma_1,\sigma_2} \setminus \{e\}$ contains at least one vertex from $C(\sigma_2)$. Then triple $(\sigma_1, \sigma_2, q)$ never contributes to the quantities described in Theorems* 1.4 *and* 1.6*, no matter how $q$ and $s_2, s_3, \ldots$ are chosen.*

**Proof.** Before starting the proof notice that the assumptions of Theorems 1.4 and 1.6 show that in order for $(\sigma_1, \sigma_2, q)$ to contribute, graph $\mathcal{V}^{\sigma_1,\sigma_2}$ must be connected.

Let $i \in C(\sigma_1)$, $j \in C(\sigma_2)$ be the endpoints of the edge $e$ and let $V_1, V_2$ be the connected components of $\mathcal{V}^{\sigma_1,\sigma_2} \setminus \{e\}$; we may assume that $i \in V_1$ and $j \in V_2$. We are going to use the condition ($e^2$). Let $A$, respectively $B$, be the set of girls, respectively boys, contained in $V_1$. From the assumption it follows that $V_1 \neq \{i\}$ therefore $A \neq \emptyset$; on the other hand $j \in V_2$ therefore $A \neq C(\sigma_2)$.

If

$$|B| - \sum_{j \in A} \bigl(q(j) - 1\bigr) \tag{4}$$

does not belong to the set {0, 1} then it is not possible to arrange the marriages.

If (4) is equal to zero then any arrangement of marriages $\mathcal{M} : C(\sigma_1) \to C(\sigma_2)$ must fulfill $\mathcal{M}^{-1}(A) = B$; if (4) is equal to 1 then any arrangement of marriages $\mathcal{M} : C(\sigma_1) \to C(\sigma_2)$ must fulfill $\mathcal{M}^{-1}(A) = B \setminus \{i\}$. In both cases, the set of husbands of wives from $A$ is uniquely determined therefore condition ($e^2$) does not hold. □

*1.9. Transportation interpretation of condition (e)*

Let $G$ be a bipartite graph and its set of black vertices be $V_\bullet$. For any set $A \subseteq V_\bullet$ of black vertices we denote by $N_G(A)$ the set of white vertices which have a neighbor in $A$. We can rephrase condition (e) by:

($e^3$) for any nontrivial subset $A$, $|N_{\mathcal{V}^{\sigma_1,\sigma_2}}(A)| \geqslant 1 + \sum_{c \in A}[q(c) - 1]$.

Let a coloring $q : V_\bullet \to \{2, 3, \ldots\}$ of the black vertices of a bipartite graph $G$ be given. We say that $G$ is $q$-admissible if for every set $A \subseteq V_\bullet$ of black vertices $|N_A| \geqslant \sum_{c \in A}[q(c) - 1]$ and furthermore the equality holds if and only if $V_\bullet$ is equal to the set of all black vertices in a union of some connected components of $G$.

Notice that if $G$ is connected then it is $q$-admissible if and only if it satisfies condition ($e^3$).

**Proposition 1.9.** *Condition* (e) *is equivalent to the following one:*

($e^4$) *there exists a strictly positive solution to the following system of equations:*

**Set of variables:**

$\{x_{i,j}$: *white vertex $i$ is connected to black vertex $j$*$\}$.

**Equations:** $\begin{cases} \forall i, \sum_j x_{i,j} = 1, \\ \forall j, \sum_i x_{i,j} = q(j) - 1. \end{cases}$

*More generally, graph $G$ is $q$-admissible if and only if condition* ($e^4$) *is fulfilled.*

Before starting the proof note that the possibility of arranging marriages (see Section 1.8) can be rephrased as existence of a solution to the above system of equations with a requirement that $x_{i,j} \in \{0, 1\}$.

The system of equations in condition ($e^4$) can be interpreted as a transportation problem where each white vertex is interpreted as a factory which produces a unit of some ware and each black vertex $j$ is interpreted as a consumer with a demand equal to $q(j) - 1$. The value of $x_{i,j}$ is interpreted as amount of ware transported from factory $i$ to the consumer $j$.

**Proof of Proposition 1.9.** Suppose that the above system has a positive solution. For any $A \subseteq V_\bullet$ we have

$$\sum_{j \in A} \bigl(q(j) - 1\bigr) = \sum_{j \in A} \sum_{\substack{i: \\ (i,j) \text{ is an edge}}} x_{i,j} = \sum_{i \in N_G(A)} \sum_{\substack{j \in A: \\ (i,j) \text{ is an edge}}} x_{i,j} \leqslant \bigl|N_G(A)\bigr|.$$

Furthermore, if $|N_G(A)| = \sum_{j \in A} (q(j) - 1)$ then the above inequality is an equality which means that for each $i \in N_G(A)$ one has

$$\sum_{\substack{j \in A: \\ (i,j) \text{ is an edge}}} x_{i,j} = 1.$$

As $\sum_j x_{i,j} = 1$ and $x_{i,j} > 0$ if $(i, j)$ is an edge, this implies that there is no edge $(i, j)$ with $i \in N_G(A)$ and $j \notin A$. In this way we have proved that $A$ is the set of black vertices of a union of some disjoint components, therefore $G$ is $q$-admissible.

The opposite implication is easy: we consider the mean of all solutions of the system with the condition $x_{i,j} \in \{0, 1\}$. This gives us a strictly positive solution because the $q$-admissibility ensures that if we force some variable $x_{i,j}$ to be equal to 1 we can find a solution to the system. □

### *1.10. Open problems*

#### *1.10.1. C-expansion*

In analogy to (1) we define for $k \geqslant 2$

$$C_k^\lambda = \frac{24}{k(k+1)(k+2)} \lim_{s \to \infty} \frac{1}{s^k} \bigl(\Sigma_{k+1}^{s\lambda} - R_{k+2}^{s\lambda}\bigr) \tag{5}$$

which (up to the unusual numerical factor in front) gives the leading terms of the deviation from the first-order approximation $\Sigma_{k+1}^\lambda \approx R_{k+2}^\lambda$. The explicit form of $C_k$

$$C_k = \sum_{\substack{j_2, j_3, \ldots \geqslant 0, \\ 2j_2 + 3j_3 + \cdots = k}} \frac{(j_2 + j_3 + \cdots)!}{j_2! j_3! \cdots} \prod_{i \geqslant 2} \bigl((i-1) R_i\bigr)^{j_i}$$

as a polynomial in free cumulants $R_2, R_3, \ldots$ was conjectured by Biane [3] and was proved by Śniady [20]. Goulden and Rattan [9] proved that for each $k \geqslant 1$ there exists a universal polynomial $L_k$ called *Goulden–Rattan polynomial* with rational coefficients such that

$$\Sigma_k - R_{k+1} = L_k(C_2, C_3, \ldots) \tag{6}$$

and they found an explicit but complicated formula for $L_k$. A simpler proof and some more related results can be found in the work of Biane [4].

**Conjecture 1.10.** *(See Goulden and Rattan [9].) The coefficients of $L_k$ are* non-negative *rational numbers with relatively small denominators.*

It is natural to conjecture that the underlying reason for positivity of the coefficients is that they have (after some additional rescaling) a combinatorial interpretation.

In some sense the free cumulants $(R_k)$ are analogous to the above quantities $(C_k)$: both have natural interpretations as leading (respectively, subleading) terms in the asymptotics of characters, cf. (1), respectively (5). Also, Conjecture 1.1 is analogous to Conjecture 1.10: both conjectures state that there are exact formulas which express the characters $\Sigma_k$ (respectively, the subdominant terms of the characters $\Sigma_k - R_{k+1}$) as polynomials in free cumulants (respectively, $(C_k)_{k\geqslant 2}$) with non-negative integer coefficients (respectively, non-negative rational coefficients with small denominators) which have a combinatorial interpretation.

The advantage of the quantities $(C_k)$ over free cumulants $(R_k)$ is that the Goulden–Rattan polynomials $L_k$ seem to have a simpler form than Kerov polynomials $K_k$ while the numerical evidence for Conjecture 1.10 suggests that their coefficients should have a rich and beautiful structure. Also, Kerov's conjecture (Conjecture 1.1) would be an immediate corollary from Conjecture 1.10. For these reasons we tend to believe that the quantities $(C_k)$ are even better suitable for the asymptotic representation theory then the free cumulants $(R_k)$ and Conjecture 1.10 deserves serious interest.

*1.10.2. $\mathcal{R}$-expansion*

Another interesting direction of research was pointed out by Lassalle [16] who presented quite explicit conjectures on the form of the coefficients of Kerov polynomials.

*1.10.3. Arithmetic properties of Kerov polynomials*

**Proposition 1.11.** *If $p$ is an odd prime number then $\frac{\Sigma_p - R_{p+1} + 2R_2}{p}$ and $\frac{\Sigma_{p-1} - R_p}{p}$ are polynomials in free cumulants $R_2, R_3, \ldots$ with nonnegative integer coefficients.*

**Proof.** In order to prove that the coefficients of $\frac{\Sigma_p - R_{p+1} + 2R_2}{p}$ are integer we consider the action of the group $\mathbb{Z}/p\mathbb{Z}$ on the set of triples $(\sigma_1, \sigma_2, q)$ which contribute to Theorem 1.4 defined by conjugation

$$\psi(i)(\sigma_1, \sigma_2, q) = \left(c^i \sigma_1 c^{-i}, c^i \sigma_2 c^{-i}, q'\right),$$

where $c = (1, 2, \ldots, k)$ is the cycle; we leave the details how to define $q'$ as a simple exercise. All orbits of this action consist of $p$ elements except for the fixpoints of this action which are of the form $\sigma_1 = c^a$, $\sigma_2 = c^{1-a}$. These fixpoints contribute to the monomial $R_{k+1}$ (with multiplicity 1) and to the monomial $R_2$ (with multiplicity $p-2$).

In order to prove that the coefficients of $\frac{\Sigma_{p-1} - R_p}{p}$ are integer we express $R_p$ as a linear combination of the conjugacy classes $\Sigma_\pi$. A formula for such an expansion presented in the paper [20] involves summation over all partitions of the set $\{1, \ldots, p\}$. The group $\mathbb{Z}/p\mathbb{Z}$ acts on such partitions; all orbits in this action consist of $p$ elements except for the fixpoints of this action: the minimal partition (which gives $\Sigma_{p-1}$) and the maximal partition (which turns out not to contribute). We express all summands (except for the summand corresponding to $\Sigma_{p-1}$) as polynomials in free cumulants, which finishes the proof. $\square$

The following conjecture was formulated by Światosław Gal (private communication) based on numerical calculations.

**Conjecture 1.12.** *If $p$ is an odd prime number then and $\frac{\Sigma_{p+1}-R_{p+2}+R_3}{p}$ is a polynomial in free cumulants $R_2, R_3, \ldots$ with nonnegative integer coefficients.*

We hope that the above claims will shed some light on more precise structure of Kerov polynomials and on the form of $C$-expansion and $\mathcal{R}$-expansion described above; they suggest that (maybe up to some small error terms) $\Sigma_n - R_{n+1}$ should in some sense be divisible by $(n-1)n(n+1)$ which supports the conjectures of Lassalle [16].

*1.10.4. Discrete version of the functionals $S_2, S_3, \ldots$*

One of the fundamental ideas in this paper is the use of the fundamental functionals $S_2, S_3, \ldots$ of the shape of a Young diagram defined as integrals over the area of a Young diagram of the powers of the contents:

$$S_n^\lambda = (n-1) \iint_{\square \in \lambda} (\text{contents}_\square)^{n-2}\, d\square \tag{7}$$

(we postpone the precise definition until Section 3.2).

It would be interesting to investigate properties of analogous quantities

$$T_n^\lambda = (n-1) \sum_{\square \in \lambda} (\text{contents}_\square)^{n-2} \tag{8}$$

in which the integral over the Young diagram was replaced by a sum over its boxes. Notice that unlike the integrals (7) which are well defined for generalized Young diagrams, the sum (8) makes sense only if $\lambda$ is a conventional Young diagram but since the resulting object is a polynomial function on the set of Young diagrams it can be extended to generalized Young diagrams. This type of quantities have been investigated by Corteel, Goupil and Schaeffer [5].

The reason why we find the functional $T_n$ so interesting is that via non-commutative Fourier transform it corresponds to a central element of the symmetric group algebra $\mathbb{C}[\mathfrak{S}(k)]$ given by the following very simple formula

$$T_n = (n-1) \sum_{2 \leqslant i \leqslant n} X_i^{n-2},$$

where

$$X_i = (1,i) + (2,i) + \cdots + (i-1,i) \in \mathbb{C}\big[\mathfrak{S}(k)\big]$$

are the Jucys–Murphy elements.

The hidden underlying idea behind the current paper is the differential calculus on the (polynomial) functions on the set of generalized Young diagrams $\mathbb{Y}$ in which we study derivatives corresponding to infinitesimal changes of the shape of a Young diagram, as it can be seen in the proof of Theorem 4.2. It is possible to develop the formalism of such a differential calculus and to express the results of this paper in such a language instead of the language of Stanley polynomials (and, in fact, the initial version of this article was formulated in this way), nevertheless if the main goal is to prove the Kerov conjecture then this would lead to unnecessary complication of the paper.

On the other hand, just like the usual differential and integral calculus has an interesting discrete difference and sum analogue, the above described differential calculus on generalized Young diagrams has a discrete difference analogue in which we study the change of the function on the set of Young diagrams corresponding to addition or removal of a single box. We expect that just like functionals $(S_n)$ are so useful in the framework of differential calculus on the set of generalized Young diagrams, functionals $(T_n)$ will be useful in the framework of the difference calculus on Young diagrams.

It would be very interesting to develop such a difference calculus and to verify if free cumulants $(R_n)$ have some interesting discrete version which nicely fits into this setup.

*1.10.5. Characterization of Stanley polynomials*

Lemma 4.5 contains some identities fulfilled by Stanley polynomials. It would be interesting to find some more such identities. In particular we state the following problem here.

**Problem 1.13.** Find (minimal set of) conditions which fully characterize the class of Stanley polynomials $\mathcal{F}(\mathbf{p} \times \mathbf{q})$ where $\mathcal{F}$ is a polynomial function on the set of Young diagrams.

It seems plausible that the answer for this problem is best formulated in the language of the differential calculus of function on the set of generalized Young diagrams about which we mentioned in Section 1.10.4.

*1.10.6. Various open problems*

Is there some analogue of Kerov character polynomials for the representation theory of semisimple Lie groups, in particular for the unitary groups $U(d)$? Does existence of Kerov polynomials for characters of symmetric groups $\mathfrak{S}(n)$ tell us something (for example via Schur–Weyl duality) about representations of the unitary groups $U(d)$? Is there some analogue of Kerov character polynomials in the random matrix theory? Is it possible to study Kerov polynomials in such a scaling that phenomena of universality of random matrices occur?

*1.11. Exotic interpretations of Kerov polynomials*

Theorem 1.4 gives some interpretation of the coefficients of Kerov polynomials but clearly it does not mean that there are no other interpretations.

*1.11.1. Biane's decomposition*

The original conjecture of Biane [3] suggested that the coefficients of Kerov polynomials are equal to multiplicities in some unspecified decomposition of the Cayley graph of the symmetric group into a signed sum of non-crossing partitions. This result was proved by Féray [8] but the details of his construction were quite implicit. In Section 8 we shall revisit the conjecture of Biane in the light of our new combinatorial interpretation of Kerov polynomials. Unfortunately, our understanding of this interpretation of the coefficients of Kerov polynomials is still not satisfactory and remains as an open problem.

*1.11.2. Multirectangular random matrices*

For a given Young diagram $\lambda$ we consider a Gaussian random matrix $(A^{\lambda}_{ij})$ with the shape of $\lambda$. Formally speaking, the entries of $(A^{\lambda}_{ij})$ are independent with $A^{\lambda}_{ij}=0$ if box $(i,j)$ does not belong to $\lambda$; otherwise $\Re A^{\lambda}_{ij}, \Im A^{\lambda}_{ij}$ are independent Gaussian random variables with mean zero and variance $\frac{1}{2}$. One can think that either $(A^{\lambda}_{ij})$ is an infinite matrix or it is a square (or rectangular) matrix of sufficiently big size.

**Theorem 1.14.** *Kerov polynomials express the moments of the random matrix $A^{\lambda}$ in terms of the genus-zero terms in the genus expansion (up to the sign). More precisely,*

$$\mathbb{E}\big[\operatorname{Tr}\big(A^{\lambda}\big(A^{\lambda}\big)^{\star}\big)^{n}\big] = -K_n(-R_2,-R_3,-R_4,\dots),$$

*where $R_i$ is defined as the genus zero term in the expansion for*

$$\mathbb{E}\big[\operatorname{Tr}\big(A^{\lambda}\big(A^{\lambda}\big)^{\star}\big)^{i-1}\big],$$

*or, precisely speaking,*

$$R_i = \lim_{s\to\infty}\frac{1}{s^i}\mathbb{E}\big[\operatorname{Tr}\big(A^{s\lambda}\big(A^{s\lambda}\big)^{\star}\big)^{i-1}\big].$$

This is an immediate consequence of the results from [7].

*1.11.3. Dimensions of (co)homologies*

In analogy to Kazhdan–Lusztig polynomials it is tempting to ask if the coefficients of Kerov polynomials might have a topological interpretation, for example as dimensions of (co)homologies of some interesting geometric objects, maybe related to Schubert varieties, as suggested by Biane (private communication). This would be supported by the Biane's decomposition from Section 1.11.1 which maybe is related to Bruhat order and Schubert cells. In this context it is interesting to ask if the conditions from Theorem 1.4 can be interpreted as geometric conditions on intersections of some geometric objects. Another approach towards establishing link between Kerov polynomials and Schubert calculus would be to relate Kerov polynomials and Schur symmetric polynomials.

*1.11.4. Schur polynomials*

Each Schur polynomial can be written as quotient of two determinants. Exactly the same quotient of determinants appears in the Harish-Chandra–Itzykson–Zuber integral

$$\int_{U(d)} e^{AUBU^{\star}}\,dU$$

if $A$ and $B$ are hermitian matrices with suitably chosen eigenvalues (say $(x_i)$ for $A$ and $(\log\lambda_i)$ for $B$).

It would be interesting to verify if Kerov polynomials can be used to express the exact values of Schur polynomials by some limit value of Harish-Chandra–Itzykson–Zuber integral when the size of the matrix tends to infinity and each variable $x_i$ occurs with a multiplicity which tends to infinity; also the shape of the Young diagram $\lambda$ should tend to infinity, probably in the "balanced Young diagram" way.

### *1.11.5. Analytic maps*

We conjecture that Kerov polynomials are related to moduli space of analytic maps on Riemann surfaces or ramified coverings of a sphere.

### *1.11.6. Integrable hierarchy*

Jonathan Novak (private communication) conjectured that Kerov polynomials might be algebraic solutions to some integrable hierarchy (maybe Toda?) and their coefficients are related to the tau function of the hierarchy.

## *1.12. Applications of the main result*

### *1.12.1. Positivity conjectures and precise information on Kerov polynomials*

The advantage of the approach to characters of symmetric groups presented in this article over some other methods is that the formula for the coefficients given by Theorem 1.4 does not involve summation of terms of positive and negative sign unlike most formulas for characters such as Murnaghan–Nakayama rule or Stanley–Féray formula (Theorem 4.6). In this way we avoid dealing with complicated cancellations. For this reason the main result of the current paper seems to be a perfect tool for proving stronger results, such as Conjecture 1.10 of Goulden and Rattan or the conjectures of Lassalle [16].

### *1.12.2. Genus expansion*

One of the important methods in the random matrix theory and in the representation theory is to express the quantity we are interested in (for example: moment of a random matrix or character of a representation) as a sum indexed by some combinatorial objects (for example: partitions of an ordered set or maps) to which one can associate canonically a two-dimensional surface [15]. Usually the asymptotic contribution of such a summand depends on the topology of the surface with planar objects being asymptotically dominant. This method is called genus expansion since exponent describing the rate of decay of a given term usually linearly depends on the genus.

The main result of this article fits perfectly into this philosophy since to any pair of permutations $\sigma_1, \sigma_2$ which contributes to Theorem 1.4 or Theorem 1.6 we may associate a canonical graph on a surface, called a *map*. It is not difficult to show that also in this situation the degree of the terms $R_2^{s_2} R_3^{s_3} \cdots$ to which such a pair of permutations contributes decreases as the genus increases.

It is natural therefore to ask about the structure of factorizations $\sigma_1 \circ \sigma_2 = (1, 2, \ldots, k)$ with a prescribed genus. As we already pointed out in Proposition 1.8, condition (e) of Theorem 1.4 gives strong limitations on the shape of the resulting bipartite graph $\mathcal{V}^{\sigma_1,\sigma_2}$ which translate to limitations on the shape of the corresponding map. Very analogous situation was analyzed in the paper [20] where it was proved that by combining a restriction on the genus and a condition analogous to the one from Proposition 1.8 ("*evercrossing partitions*") one gets only a finite number of allowed patterns for the geometric object concerned.

Similar analysis should be possible for the formulas for Kerov polynomials presented in the current paper which should shed some light on Conjecture 1.10 of Goulden and Rattan and the conjectures of Lassalle [16].

*1.12.3. Upper bounds on characters*

It seems plausible that the main result of this article, Theorems 1.4 and 1.6, can be used to prove new upper bounds on the characters of symmetric groups

$$\chi^{\lambda}(\pi) = \frac{\operatorname{Tr} \rho^{\lambda}(\pi)}{\text{dimension of } \rho^{\lambda}} \tag{9}$$

for balanced Young diagram $\lambda$ in the scaling when the length of the permutation $\pi$ is large compared to the number of boxes of $\lambda$.

The advantage of such approach to estimates on characters over other methods, such as via Frobenius formula as in the work of Rattan and Śniady [18] or via Stanley–Féray formula [7], becomes particularly visible in the case when the shape of the Young diagram becomes close to the limit curve for the Plancherel measure [17,27] for which all free cumulants (except for $R_2$) are close to zero. Indeed, for $\lambda$ in the neighborhood of this limit curve one should expect much tighter bounds on the characters (9) because such Young diagrams maximize the dimension of the representation which is the denominator of the fraction, while the numerator can be estimated by Murnaghan–Nakayama rule and some combinatorial tricks [19].

*1.13. Overview of the paper*

In Section 2 we recall some basic facts about free cumulants $R_1, R_2, \dots$ and quantities $S_1, S_2, \dots$ for probability measures on the real line and their relations with each other. The main result of this section is formula (14) which allows to express functionals $S_1, S_2, \dots$ in terms of free cumulants $R_1, R_2, \dots$.

In Section 3 we define the fundamental functionals $S_2, S_3, \dots$ for generalized Young diagrams and study their geometric interpretation.

In Section 4 we study Stanley polynomials and their relations to the fundamental functionals $S_2, S_3, \dots$ of shape of a Young diagram.

Section 5 is devoted to a toy example: we shall prove Theorem 1.4 in the simplest nontrivial case of coefficients of the quadratic terms, which is exactly the case in Theorem 1.3. In this way the reader can see all essential steps of the proof in a simplified situation when it is possible to avoid technical difficulties.

In Section 6 we prove some auxiliary combinatorial results.

In Section 7 we present the proof of the main result: Theorems 1.4 and 1.6.

Finally, in Section 8 we revisit the paper [8] and we show how rather implicit constructions of Féray become much more concrete once one knows the formulation of the main result of the current paper, Theorem 1.4. In fact, Section 8 provides a short proof of Theorem 1.4 based on results from of Féray. This simplicity is however slightly misleading since the original paper [8] is not easy.

## 2. Functionals of measures

In this section we present relations between moments $M_1, M_2, \dots$ of a given probability measure, its free cumulants $R_1, R_2, \dots$ and its functionals $S_1, S_2, \dots$. The only result of this section which will be used in the remaining part of the article is equality (14), nevertheless we find functionals $S_1, S_2, \dots$ so important that we collected in this section also some other formulas involving them.

Assume that $\nu$ is a compactly supported measure on $\mathbb{R}$. For integer $n \geqslant 0$ we consider moments of $\nu$

$$M_n^\nu = \int z^n \, d\nu(z)$$

and its Cauchy transform

$$G^\nu(z) = \int \frac{1}{z-x} \, d\nu(x) = \sum_{n \geqslant 0} \frac{M_n^\nu}{z^{1+n}};$$

the integral and the series make sense in a neighborhood of infinity.

From the following on we assume that $\nu$ is a compactly supported probability measure on $\mathbb{R}$. We define a sequence $(S_n^\nu)_{n \geqslant 1}$ of the coefficients of the expansion

$$S^\nu(z) = \log z G^\nu(z) = \sum_{n \geqslant 1} \frac{S_n^\nu}{z^n}$$

in a neighborhood of infinity and a sequence $(R_n^\nu)_{n \geqslant 1}$ of free cumulants as the coefficients of the expansion

$$R^\nu(z) = \left(G^\nu\right)^{\langle -1 \rangle}(z) - \frac{1}{z} = \sum_{n \geqslant 1} R_n^\nu z^{n-1} \tag{10}$$

in a neighborhood of 0, where $(G^\nu)^{\langle -1 \rangle}$ is the right inverse of $G^\nu$ with respect to the composition of functions [28]. When it does not lead to confusions we shall omit the superscript in the expressions $M_n^\nu$, $G^\nu$, $S^\nu$, $R^\nu$, $S_n^\nu$, $R_n^\nu$.

The relation between the moments and the free cumulants is given by the following combinatorial formula which, in fact, can be regarded as an alternative definition of free cumulants [23]:

$$M_n = \sum_{\Pi \in \mathrm{NC}_n} R_\Pi, \tag{11}$$

where the summation is carried over all non-crossing partitions of $n$-element set and where $R_\Pi$ is defined as the multiplicative extension of $(R_k)$:

$$R_\Pi = \prod_{b \in \Pi} R_{|b|},$$

where the product is taken over all blocks $b$ of the partition $\Pi$ and $|b|$ denotes the number of the elements in $b$ [23].

Information about the measure $\nu$ can be described in various ways; in this article descriptions in terms of the sequences $(S_n)$ and $(R_n)$ play eminent role and we need to be able to relate each of these sequences to the other. We shall do it in the following.

**Lemma 2.1.** *For any integer $k \geqslant 1$*

$$\frac{\partial G(R_1, R_2, \ldots)}{\partial R_k}(z) = -\frac{1}{k}\big(\big[G(z)\big]^k\big)' = -G^{k-1}(z)G'(z),$$

*where both sides of the above equality are regarded as formal power series in powers of $\frac{1}{z}$ with the coefficients being polynomials in $R_1, R_2, \ldots$.*

**Proof.** Eq. (10) is equivalent to

$$G\left(R(z) + \frac{1}{z}\right) = z. \tag{12}$$

We denote

$$t = R(z) + \frac{1}{z}.$$

Let us keep all free cumulants fixed except for $R_k$, we shall treat $G$ as a function of free cumulants. By taking the derivatives of both sides of (12) it follows that

$$\begin{aligned} 0 &= \frac{\partial}{\partial R_k}\left[G\left(R(z) + \frac{1}{z}\right)\right] = \frac{\partial G}{\partial R_k}(t) + G'(t)\frac{\partial}{\partial R_k}\left(R(z) + \frac{1}{z}\right) \\ &= \frac{\partial G}{\partial R_k}(t) + G'(t)z^{k-1} = \frac{\partial G}{\partial R_k}(t) + G'(t) \cdot G^{k-1}(t) \end{aligned}$$

which finishes the proof. □

**Proposition 2.2.** *For any integer $n \geqslant 1$*

$$M_n = \sum_{l \geqslant 1} \frac{1}{l!}(n)_{l-1} \sum_{\substack{k_1,\ldots,k_l \geqslant 1 \\ k_1+\cdots+k_l=n}} R_{k_1} \cdots R_{k_l}, \tag{13}$$

$$S_n = \sum_{l \geqslant 1} \frac{1}{l!}(n-1)_{l-1} \sum_{\substack{k_1,\ldots,k_l \geqslant 1 \\ k_1+\cdots+k_l=n}} R_{k_1} \cdots R_{k_l}, \tag{14}$$

$$R_n = \sum_{l \geqslant 1} \frac{1}{l!}(-n+1)^{l-1} \sum_{\substack{k_1,\ldots,k_l \geqslant 1 \\ k_1+\cdots+k_l=n}} S_{k_1} \cdots S_{k_l}, \tag{15}$$

*where*

$$(a)_b = \underbrace{a(a-1)\cdots(a-b+1)}_{b \text{ factors}}$$

*denotes the falling factorial.*

**Proof.** Lemma 2.1 shows that

$$\frac{\partial^2 G(R_1, R_2, \ldots)}{\partial R_k \partial R_l}(z) = \frac{1}{k+l-1}\left(\left[G(z)\right]^{k+l-1}\right)''$$

therefore if $k + l = k' + l'$ then

$$\frac{\partial^2 M_n(R_1, R_2, \ldots)}{\partial R_k \partial R_l} = \frac{\partial^2 M_n(R_1, R_2, \ldots)}{\partial R_{k'} \partial R_{l'}}.$$

It follows by induction that

$$\frac{\partial^l M_n(R_1, R_2, \ldots)}{\partial R_{k_1} \cdots \partial R_{k_l}} = \frac{\partial^l M_n(R_1, R_2, \ldots)}{(\partial R_1)^{l-1} \partial R_{k_1+\cdots+k_l-(l-1)}}.$$

From the moment-cumulant formula (11) it follows that for $R_1 = R_2 = \cdots = 0$ the right-hand side of the above equation is equal to the number of non-crossing partitions with an ordering of blocks, such that the numbers of elements in consecutive blocks are as follows:

$$\underbrace{1, \ldots, 1}_{l-1 \text{ times}}, k_1 + \cdots + k_l - (l-1).$$

Such non-crossing partitions have a particularly simple structure therefore it is very easy to find their cardinality. Therefore

$$\left.\frac{\partial^l M_n(R_1, R_2, \ldots)}{(\partial R_1)^{l-1} \partial R_{k_1+\cdots+k_l-(l-1)}}\right|_{R_1=R_2=\cdots=0} = \begin{cases} (n)_{l-1} & \text{if } n = k_1 + \cdots + k_l, \\ 0 & \text{otherwise,} \end{cases} \tag{16}$$

which finishes the proof of (13).

Lemma 2.1 shows that for $k \geqslant 2$

$$\frac{\partial S(R_1, R_2, \ldots)}{\partial R_k}(z) = \frac{\partial \log[zG(z)]}{\partial R_k} = -G^{k-2}G' = \frac{\partial G(R_1, R_2, \ldots)}{\partial R_{k-1}}$$

therefore

$$\frac{\partial S_n(R_1, R_2, \ldots)}{\partial R_k} = \frac{\partial M_{n-1}(R_1, R_2, \ldots)}{\partial R_{k-1}}.$$

Assume that $k_l \geqslant 2$; then

$$\frac{\partial^l S_n(R_1, R_2, \ldots)}{\partial R_{k_1} \cdots \partial R_{k_l}} = \frac{\partial^l M_{n-1}(R_1, R_2, \ldots)}{\partial R_{k_1} \cdots \partial R_{k_l-1}}$$

which is calculated in Eq. (16). In this way we proved that if $(k_1, \dots, k_l) \neq (1, 1, \dots, 1)$ then

$$\left.\frac{\partial^l S_n(R_1, R_2, \dots)}{\partial R_{k_1} \cdots \partial R_{k_l}}\right|_{R_1=R_2=\dots=0} = \begin{cases} (n-1)_{l-1} & \text{if } n = k_1 + \dots + k_l, \\ 0 & \text{otherwise.} \end{cases}$$

In order to prove the case $k_1 = \dots = k_l = 1$ it is enough to consider the Dirac point measure $\nu = \delta_a$ for which $G(z) = \frac{1}{z-a}$, $R_1 = a$, $R_2 = R_3 = \dots = 0$ and $S(z) = -\log(1 - \frac{a}{z})$, $S_n = \frac{a^n}{n}$. In this way the proof of (14) is finished.

Lagrange inversion formula shows that

$$\begin{aligned} R_{n+1} &= -\frac{1}{n}\left[\frac{1}{z}\right]\left(\frac{1}{G(z)}\right)^n = -\frac{1}{n}\left[\frac{1}{z^{n+1}}\right]\exp\bigl[-nS(z)\bigr] \\ &= \sum_{l \geqslant 1} \frac{1}{l!}(-n)^{l-1} \sum_{\substack{k_1,\dots,k_l \geqslant 1 \\ k_1+\dots+k_l=n+1}} S_{k_1} \cdots S_{k_l} \end{aligned}$$

which finishes the proof of (15). □

## 3. Generalized Young diagrams

The main result of this section is formula (17) which relates the fundamental functionals $S_2, S_3, \dots$ to the geometric shape of the Young diagram.

In the following we base on the notations introduced in Section 1.1.

### 3.1. *Measure on a diagram and contents of a box*

Notice that each unit box of a Young diagram drawn in the French convention becomes in the Russian notation a square of side $\sqrt{2}$. For this reason, when drawing a Young diagram according to the French convention we will use the plane equipped with the usual measure (i.e. the area of a unit square is equal to 1) and when drawing a Young diagram according to the Russian notation we will use the plane equipped with the usual measure divided by 2 (i.e. the area of a unit square is equal to $\frac{1}{2}$). In this way a (generalized) Young diagram has the same area when drawn in the French and in the Russian convention.

Speaking very informally, the setup of generalized Young diagrams corresponds to looking at a Young diagram from very far away so that individual boxes become very small. Therefore by the term *box of a Young diagram* $\lambda$ we will understand simply any point $\square$ which belongs to $\lambda$. In the case of the Russian convention this means that $\square = (x, y)$ fulfills

$$|x| < y < \lambda(x).$$

We define the contents of the box $\square = (x, y)$ in the Russian convention by $\text{contents}_\square = x$.

In the case of the French convention $\square = (x, y)$ belongs to a diagram $\lambda$ if

$$x > 0 \quad \text{and} \quad 0 < y < \lambda(x)$$

and the contents of the box $\square = (x, y)$ is defined by $\text{contents}_\square = x - y$.

### 3.2. Functionals of Young diagrams

The above definitions of the measure on the plane and of the contents in the case of French and Russian conventions are compatible with each other, therefore it is possible to define some quantities in a convention-independent way. In particular, we define the *fundamental functionals of shape* of a generalized Young diagram

$$S_n^\lambda = (n-1) \iint_{\square\in\lambda} (\text{contents}_\square)^{n-2}\, d\square \tag{17}$$

for integer $n \geqslant 2$. Clearly, each functional $S_n$ is a homogeneous function of the Young diagram with degree $n$.

Let a generalized Young diagram $\lambda : \mathbb{R} \to \mathbb{R}_+$ drawn in the Russian convention be fixed. We associate to it a function

$$\tau^\lambda(x) = \frac{\lambda(x) - |x|}{2}$$

which gives the distribution of the contents of the boxes of $\lambda$. When it does not lead to confusions we will write for simplicity $\tau$ instead of $\tau^\lambda$. In the following we shall view $\tau$ as a measure on $\mathbb{R}$. Its Cauchy transform can be written as

$$G^\tau(z) = \iint_{\square\in\lambda} \frac{1}{z - \text{contents}_\square}\, d\square.$$

With these notations we have that

$$S_n^\lambda = (n-1)\int x^{n-2}\tau(x)\, dx = -\int x^{n-1}\tau'(x)\, dx$$

are (rescaled) moments of the measure $\tau$ or, alternatively, (shifted) moments of the Schwartz distribution $-\tau'$.

We define

$$S^\lambda(z) = \sum_{n\geqslant 2} \frac{S_n^\lambda}{z^n} = \iint_{\square\in\lambda} \frac{1}{(z - \text{contents}_\square)^2}\, d\square$$

where the second equality follows by expanding right-hand side into a power series and (17). It follows that

$$\begin{aligned} S^\lambda(z) &= -\frac{d}{dz} G^\tau(z) = G^{-\tau'}(z) = -\int \frac{1}{z-x}\tau'(x)\, dx \\ &= -\int \log(z-x)\tau''(x)\, dx, \end{aligned}$$

in particular $S^\lambda(z)$ coincides with the Cauchy transform of a Schwartz distribution $-\tau'$. The above formulas show that $S^\lambda(z)$ and $S_n^\lambda(z)$ coincide (up to small modifications) with the quantities considered by Kerov [12,14], Ivanov and Olshanski [10].

### 3.3. *Kerov transition measure*

The corresponding Cauchy transform

$$G^\lambda(z)=\frac{1}{z}\exp S^\lambda(z) \tag{18}$$

is a Cauchy transform of a probability measure $\mu_\lambda$ on the real line, called *Kerov transition measure* of λ [12,14]. Probably it would be more correct to write $G^{\mu_\lambda}$ instead of $G^\lambda$ and to write $S^{\mu_\lambda}$ instead of $S^\lambda$, but this would lead to unnecessary complexity of the notation.

One of the reasons why Kerov's transition measure was so successful in the asymptotic representation theory of symmetric groups is that it can be defined in several equivalent ways, related either to the shape of λ or to representation theory or to moments of Jucys–Murphy elements or to certain matrices. For a review of these approaches we refer to [1].

### 3.4. *Free cumulants of a Young diagram*

In order to keep the introduction as non-technical as possible, we introduced free cumulants of a Young diagram by formula (1). The conventional way of defining them is to use (10) for the Cauchy transform given by (18). Therefore, one should make sure that these two definitions are equivalent. This can be done thanks to Frobenius formula

$$\Sigma_{k-1}^\lambda=-\frac{1}{k-1}\left[\frac{1}{z}\right]\frac{1}{G^\lambda(z-1)G^\lambda(z-2)\cdots G^\lambda(z-(k-1))}$$

which shows that

$$\frac{1}{s^k}\Sigma_{k-1}^{s\lambda}=-\frac{1}{k-1}\left[\frac{1}{z}\right]\frac{1}{G^\lambda(z-\frac{1}{s})G^\lambda(z-\frac{2}{s})\cdots G^\lambda(z-\frac{k-1}{s})};$$

therefore definition (1) would give

$$R_k^\lambda=-\frac{1}{k-1}\left[\frac{1}{z}\right]\left(\frac{1}{G^\lambda(z)}\right)^{k-1}$$

which coincides with the value given by the Lagrange inversion formula applied to (10).

### 3.5. *Polynomial functions on the set of Young diagrams*

For simplicity we shall often drop the explicit dependence of the functionals of Young diagrams from λ. Since the transition measure $\mu^\lambda$ is always centered it follows that $M_1=R_1=S_1=0$.

Existence of Kerov polynomials allows us define formally the normalized characters $\Sigma_\pi^\lambda$ even if λ is a generalized Young diagram.

We will say that a function on the set of generalized Young diagrams $\mathbb{Y}$ is a *polynomial function* if one of the following equivalent conditions hold [10]:

- it is a polynomial in $M_2, M_3, \ldots$;

- it is a polynomial in $S_2, S_3, \ldots$;
- it is a polynomial in $R_2, R_3, \ldots$;
- it is a polynomial in $(\Sigma_\pi)_\pi$.

## 4. Stanley polynomials and Stanley–Féray character formula

### 4.1. Stanley polynomials

**Proposition 4.1.** *Let $\mathcal{F} : \mathbb{Y} \to \mathbb{R}$ be a polynomial function on the set of generalized Young diagrams. Then $(\mathbf{p}, \mathbf{q}) \mapsto \mathcal{F}(\mathbf{p} \times \mathbf{q})$ for $\mathbf{p} = (p_1, \ldots, p_m)$, $\mathbf{q} = (q_1, \ldots, q_m)$ is a polynomial in indeterminates $p_1, \ldots, p_m, q_1, \ldots, q_m$, called* Stanley polynomial.

**Proof.** It is enough to prove this proposition for some family of generators of the algebra of polynomial functions on $\mathbb{Y}$ for example for functionals $S_2, S_3, \ldots$. We leave it as an exercise. □

**Theorem 4.2.** *Let $\mathcal{F} : \mathbb{Y} \to \mathbb{R}$ be a polynomial function on the set of generalized Young diagrams, we shall view it as a polynomial in $S_2, S_3, \ldots$. Then for any $k_1, \ldots, k_l \geqslant 2$*

$$\frac{\partial}{\partial S_{k_1}} \cdots \frac{\partial}{\partial S_{k_l}} \mathcal{F} \bigg|_{S_2 = S_3 = \cdots = 0} = \left[ p_1 q_1^{k_1 - 1} \cdots p_l q_l^{k_l - 1} \right] \mathcal{F}(\mathbf{p} \times \mathbf{q}). \tag{19}$$

**Proof.** Let $\mathbf{p} = (p_1, \ldots, p_m)$, $\mathbf{q} = (q_1, \ldots, q_m)$. For a given index $i$ we consider a trajectory in the set of generalized Young diagrams $q_i \mapsto (\mathbf{p} \times \mathbf{q})$, where all other parameters $(p_j)$ and $(q_j)_{j \neq i}$ are treated as constants. In the Russian convention we have

$$\left( \frac{\partial}{\partial q_i} (\mathbf{p} \times \mathbf{q}) \right)(x) = \begin{cases} 2 & \text{if } q_i - p_1 - \cdots - p_i < x < q_i - p_1 - \cdots - p_{i-1}, \\ 0 & \text{otherwise,} \end{cases}$$

which shows the change of the contents distribution. From (17) it follows therefore

$$\frac{\partial}{\partial q_i} S_n^{\mathbf{p} \times \mathbf{q}} = \int_{q_i - p_1 - \cdots - p_i}^{q_i - p_1 - \cdots - p_{i-1}} (n-1) x^{n-2} \, dx$$

and

$$\frac{\partial}{\partial q_i} \mathcal{F}(\mathbf{p} \times \mathbf{q}) = \sum_{n \geqslant 2} \int_{q_i - p_1 - \cdots - p_i}^{q_i - p_1 - \cdots - p_{i-1}} \frac{\partial \mathcal{F}}{\partial S_n} (n-1) x^{n-2} \, dx.$$

By iterating the above argument we show that

$$\begin{aligned} \frac{\partial}{\partial q_1} \cdots \frac{\partial}{\partial q_l} \mathcal{F}(\mathbf{p} \times \mathbf{q}) = & \sum_{n_1, \ldots, n_l \geqslant 2} \frac{\partial}{\partial S_{n_1}} \cdots \frac{\partial}{\partial S_{n_l}} \mathcal{F}(\mathbf{p} \times \mathbf{q}) \\ & \times \int_{q_1 - p_1}^{q_1} (n_1 - 1) x_1^{n_1 - 2} \, dx_1 \cdots \int_{q_l - p_1 - \cdots - p_l}^{q_l - p_1 - \cdots - p_{l-1}} (n_l - 1) x_l^{n_l - 2} \, dx_l. \end{aligned}$$

We shall treat both sides of the above equality as polynomials in $\mathbf{p}$ and we will treat $\mathbf{q}$ as constants. We are going to compute the coefficient of $p_1 \cdots p_m$ of both sides; we do this by computing the dominant term of the right-hand side in the limit $\mathbf{p} \to 0$. It follows that

$$[p_1 \cdots p_m] \frac{\partial}{\partial q_1} \cdots \frac{\partial}{\partial q_l} \mathcal{F}(\mathbf{p} \times \mathbf{q}) = \sum_{n_1,\ldots,n_l \geqslant 2} \frac{\partial}{\partial S_{n_1}} \cdots \frac{\partial}{\partial S_{n_l}} \mathcal{F}(\mathbf{p} \times \mathbf{q}) \bigg|_{p_1=\cdots=p_l=0} \times (n_1-1) q_1^{n_1-2} \cdots (n_l-1) q_l^{n_l-2}$$

which finishes the proof. □

**Corollary 4.3.** *If $k_1, \ldots, k_l \geqslant 2$ then*

$$\left[p_1 q_1^{k_1-1} \cdots p_l q_l^{k_l-1}\right] \mathcal{F}(\mathbf{p} \times \mathbf{q})$$

*does not depend on the order of the elements of the sequence $(k_1, \ldots, k_l)$.*

### 4.2. Identities fulfilled by coefficients of Stanley polynomials

The coefficients of Stanley polynomials of the form $[p_1 q_1^{k_1-1} \cdots p_l q_l^{k_l-1}] \mathcal{F}(\mathbf{p} \times \mathbf{q})$ with $k_1, \ldots, k_l \geqslant 2$ have a relatively simple structure, as it can be seen for example in Corollary 4.3. In the following we will study the properties of such coefficients if some of the numbers $q_1, \ldots, q_{k_l}$ are equal to 1.

Let $\mathcal{F} : \mathbb{Y} \to \mathbb{R}$ be a fixed polynomial function. For a sequence $(a_1, b_1), \ldots, (a_m, b_m)$ of ordered pairs, where $a_1, \ldots, a_m \geqslant 2$ and $b_1, \ldots, b_m \geqslant 1$ are integers we define an auxiliary quantity

$$Q^{\mathcal{F}}_{(a_1,b_1),\ldots,(a_m,b_m)} = \left( \prod_r (-1)^{b_r-1} \, (a_r-1)_{(b_r-1)} \right) \left[p_1 q_1^{a_1-1} \cdots p_m q_m^{a_m-1}\right] \mathcal{F}(\mathbf{p} \times \mathbf{q}),$$

which thanks to Corollary 4.3 does not depend on the order of the elements in the tuple $(a_1, b_1), \ldots, (a_m, b_m)$.

**Corollary 4.4.** *For any polynomial function $\mathcal{F}$ on the set of generalized Young diagrams and $k_1, \ldots, k_l \geqslant 2$*

$$\frac{\partial}{\partial R_{k_1}} \cdots \frac{\partial}{\partial R_{k_l}} \mathcal{F} \bigg|_{R_2=R_3=\cdots=0} = \sum_{\Pi \in P(1,2,\ldots,l)} (-1)^{l-|\Pi|} Q^{\mathcal{F}}_{((\sum_{i \in b} k_i, |b|):\, b \in \Pi)},$$

*where the sum runs over all partitions of $\{1, \ldots, l\}$.*

**Proof.** It is enough to use Theorem 4.2 and Eq. (14). □

**Lemma 4.5.** *For any polynomial function $\mathcal{F} : \mathbb{Y} \to \mathbb{R}$ and any sequence of integers $k_1, \ldots, k_m \geqslant 1$*

$$\left[p_1 q_1^{k_1-1} \cdots p_m q_m^{k_m-1}\right] \mathcal{F}(\mathbf{p} \times \mathbf{q}) = \sum_{\Pi} Q^{\mathcal{F}}_{((\sum_{i \in b} k_i, |b|):\, b \in \Pi)},$$

*where the sum runs over all partitions $\Pi$ of the set $\{1,\dots,m\}$ with a property that if $(a_1,\dots,a_l)$ with $a_1<\dots<a_l$ is a block of $\Pi$ then $k_{a_1}=\dots=k_{a_{l-1}}=1$ and $k_{a_l}\geqslant 2$ or, in other words, the set of rightmost legs of the blocks of $\Pi$ coincides with the set of indices $i$ such that $k_i\geqslant 2$.*

**Proof.** We shall treat $\mathcal{F}(\mathbf{p}\times\mathbf{q})$ as a polynomial in $\mathbf{p}$ and we shall treat $\mathbf{q}$ as constants. Our goal is to understand the coefficient $[p_1\cdots p_m]\mathcal{F}(\mathbf{p}\times\mathbf{q})$. Since $\mathcal{F}$ is a polynomial in $S_2,S_3,\dots$ we are also going to investigate analogous coefficients for $\mathcal{F}=S_n$.

For the purpose of the following calculation we shall use the French notation:

$$S_n(\lambda)=(n-1)\iint(\mathrm{contents}_\square)^{n-2}\,d\square$$
$$=(n-2)!\sum_{1\leqslant r\leqslant n-1}(-1)^{r-1}\iint_{(x,y)\in\lambda}\frac{x^{n-1-r}}{(n-1-r)!}\frac{y^{r-1}}{(r-1)!}\,dx\,dy.$$

Since the integral

$$\iint_{(x,y)\in\lambda}\frac{x^{n-1-r}}{(n-1-r)!}\frac{y^{r-1}}{(r-1)!}\,dx\,dy$$

can be interpreted as the volume of the set

$$\bigl\{(x_1,\dots,x_{n-r},y_1,\dots,y_r)\colon\ 0<x_1<\dots<x_{n-r}\text{ and }0<y_1<\dots<y_r\text{ and }(x_{n-r},y_r)\in\lambda\bigr\}$$

therefore for any $i_1<\dots<i_r$

$$[p_{i_1}\cdots p_{i_r}]S_n(\mathbf{p}\times\mathbf{q})=(-1)^{r-1}(n-1)_{r-1}q_{i_r}^{n-r}. \tag{20}$$

We express $\mathcal{F}$ as a polynomial in $S_2,S_3,\dots$. Notice that the monomial $p_1\cdots p_m$ can arise in $\mathcal{F}(\mathbf{p}\times\mathbf{q})$ only in the following way: we cluster the factors $p_1\cdots p_m$ in all possible ways or, in other words, we consider all partitions $\Pi$ of the set $\{1,\dots,m\}$. Each block of such a partition corresponds to one factor $S_n$ for some value of $n$. Thanks to Eq. (20) we can compare the factors $q_1,\dots,q_m$ which appear with a non-zero exponent and see that only partitions $\Pi$ which contribute are as prescribed in the formulation of the lemma; furthermore we can find the correct value of $n$ for each block of $\Pi$.

Eq. (19) finishes the proof. $\square$

### *4.3. Stanley–Féray character formula*

The following result was conjectured by Stanley [26] and proved by Féray [6] and therefore we refer to it as Stanley–Féray character formula. For a more elementary proof we refer to [7].

**Theorem 4.6.** *The value of the normalized character on $\pi\in\mathfrak{S}(n)$ for a multirectangular Young diagram $\mathbf{p}\times\mathbf{q}$ for $\mathbf{p}=(p_1,\dots,p_r)$, $\mathbf{q}=(q_1,\dots,q_r)$ is given by*

$$\Sigma_\pi^{\mathbf{p}\times\mathbf{q}} = \sum_{\substack{\sigma_1,\sigma_2\in\mathfrak{S}(n)\\ \sigma_1\circ\sigma_2=\pi}} \sum_{\phi_2: C(\sigma_2)\to\{1,\dots,r\}} (-1)^{\sigma_1} \left[\prod_{b\in C(\sigma_1)} q_{\phi_1(b)} \prod_{c\in C(\sigma_2)} p_{\phi_2(c)}\right], \tag{21}$$

*where $\phi_1 : C(\sigma_1) \to \{1, \dots, r\}$ is defined by*

$$\phi_1(c) = \max_{\substack{b\in C(\sigma_2),\\ b \text{ and } c \text{ intersect}}} \phi_2(b).$$

Interestingly, the above theorem shows that some partial information about the family of graphs $(\mathcal{V}^{\sigma_1,\sigma_2})_{\sigma_1,\sigma_2}$ can be extracted from the coefficients of Stanley polynomial $\Sigma_\pi^{\mathbf{p}\times\mathbf{q}}$. This observation will be essential for the proof of the main result.

The following result is a simple corollary from Theorem 4.6 and it was proved by Féray [8].

**Theorem 4.7.** *For any integers $k_1, \dots, k_l \geqslant 1$ the value of the cumulant $\kappa^{\mathrm{id}}(\Sigma_{k_1}, \dots, \Sigma_{k_l})$ evaluated at the Young diagram $\mathbf{p}\times\mathbf{q}$ is given by*

$$\kappa^{\mathrm{id}\ \mathbf{p}\times\mathbf{q}}(\Sigma_{k_1}, \dots, \Sigma_{k_l}) = \sum_{\substack{\sigma_1,\sigma_2\in\mathfrak{S}(n)\\ \sigma_1\circ\sigma_2=\pi\\ \langle\sigma,\pi\rangle \text{ transitive}}} \sum_{\phi_2: C(\sigma_2)\to\{1,\dots,r\}} (-1)^{\sigma_1} \left[\prod_{b\in C(\sigma_1)} q_{\phi_1(b)} \prod_{c\in C(\sigma_2)} p_{\phi_2(c)}\right],$$

*where $n = k_1 + \cdots + k_l$ and $\pi$ is a fixed permutation with the cycle structure $k_1, \dots, k_l$, for example $\pi = (1, 2, \dots, k_1)(k_1 + 1, k_1 + 2, \dots, k_1 + k_2)\cdots$, and where $\phi_1$ is as in Theorem* 1.4.

## 5. Toy example: Quadratic terms of Kerov polynomials

We are on the way towards the proof of Theorem 1.4 which, unfortunately, is a bit technically involved. Before dealing with the complexity of the general case we shall present in this section a proof of Theorem 1.3 which concerns a simplified situation in which we are interested in quadratic terms of Kerov polynomials. This case is sufficiently complex to show the essential elements of the complete proof of Theorem 1.4 but simple enough not to overwhelm the reader with unnecessary difficulties.

We shall prove Theorem 1.3 in the following equivalent form:

**Theorem 5.1.** *For all integers $l_1, l_2 \geqslant 2$ and $k \geqslant 1$ the derivative*

$$\left.\frac{\partial^2}{\partial R_{l_1}\partial R_{l_2}} K_k\right|_{R_2=R_3=\cdots=0}$$

*is equal to the number of triples $(\sigma_1, \sigma_2, q)$ with the following properties:*

(a) *$\sigma_1, \sigma_2$ is a factorization of the cycle; in other words $\sigma_1, \sigma_2 \in \mathfrak{S}(k)$ are such that $\sigma_1 \circ \sigma_2 = (1, 2, \dots, k)$;*
(b) *$\sigma_2$ consists of two cycles and $\sigma_1$ consists of $l_1 + l_2 - 2$ cycles;*

(c) $\ell : C(\sigma_2) \to \{1,2\}$ *is a bijective labeling of the two cycles of* $\sigma_2$*;*
(d) *for each cycle* $c \in C(\sigma_2)$ *there are at least* $l_{\ell(c)}$ *cycles of* $\sigma_1$ *which intersect non-trivially* $c$.

**Proof.** Eq. (14) shows that for any polynomial function $\mathcal{F}$ on the set of generalized Young diagrams

$$\frac{\partial^2}{\partial R_{l_1} \partial R_{l_2}} \mathcal{F} = \frac{\partial^2}{\partial S_{l_1} \partial S_{l_2}} \mathcal{F} + (l_1 + l_2 - 1) \frac{\partial}{\partial S_{l_1+l_2}} \mathcal{F},$$

where all derivatives are taken at $R_2 = R_3 = \cdots = S_2 = S_3 = \cdots = 0$. Theorem 4.2 shows that the right-hand side is equal to

$$\left[p_1 p_2 q_1^{l_1-1} q_2^{l_2-1}\right] \mathcal{F}(\mathbf{p} \times \mathbf{q}) + (l_1 + l_2 - 1)\left[p_1 q_1^{l_1+l_2-1}\right] \mathcal{F}(\mathbf{p} \times \mathbf{q}).$$

Lemma 4.5 applied to the second summand shows therefore that

$$\frac{\partial^2}{\partial R_{l_1} \partial R_{l_2}} \mathcal{F} = \left[p_1 p_2 q_1^{l_1-1} q_2^{l_2-1}\right] \mathcal{F}(\mathbf{p} \times \mathbf{q}) - \left[p_1 p_2 q_2^{l_1+l_2-2}\right] \mathcal{F}(\mathbf{p} \times \mathbf{q}). \tag{22}$$

In fact, the above equality is a direct application of Corollary 4.4, nevertheless for pedagogical reasons we decided to present the above expanded derivation. In the following we shall use the above identity for $\mathcal{F} = \Sigma_k$.

On the other hand, let us compute the number of the triples $(\sigma_1, \sigma_2, \ell)$ which contribute to the quantity presented in Theorem 5.1. By inclusion–exclusion principle it is equal to

$$\begin{aligned}
&\big(\text{number of triples which fulfill conditions (a)–(c)}\big) \\
&\quad + (-1)\big(\text{number of triples for which the cycle } \ell^{-1}(1) \text{ intersects at most } l_1 - 1 \text{ cycles of } \sigma_1\big) \\
&\quad + (-1)\big(\text{number of triples for which the cycle } \ell^{-1}(2) \text{ intersects at most } l_2 - 1 \text{ cycles of } \sigma_1\big).
\end{aligned} \tag{23}$$

At first sight it might seem that the above formula is not complete since we should also add the number of triples for which the cycle $\ell^{-1}(1)$ intersects at most $l_1 - 1$ cycles of $\sigma_1$ and the cycle $\ell^{-1}(2)$ intersects at most $l_2 - 1$ cycles of $\sigma_1$, however this situation is not possible since $\sigma_1$ consists of $l_1 + l_2 - 2$ cycles and $\langle \sigma_1, \sigma_2 \rangle$ acts transitively.

By Stanley–Féray character formula (21) the first summand of (23) is equal to

$$(-1) \sum_{\substack{i+j=l_1+l_2-2, \\ 1 \leqslant j}} \left[p_1 p_2 q_1^i q_2^j\right] \Sigma_k^{\mathbf{p} \times \mathbf{q}}, \tag{24}$$

the second summand of (23) is equal to

$$\sum_{\substack{i+j=l_1+l_2-2, \\ 1 \leqslant i \leqslant l_1-1}} \left[p_1 p_2 q_1^j q_2^i\right] \Sigma_k^{\mathbf{p} \times \mathbf{q}}, \tag{25}$$

and the third summand of (23) is equal to

$$\sum_{\substack{i+j=l_1+l_2-2,\\ 1\leqslant j\leqslant l_2-1}} \left[p_1 p_2 q_1^i q_2^j\right] \Sigma_k^{\mathbf{p}\times\mathbf{q}}. \tag{26}$$

We can apply Corollary 4.3 to the summands of (25); it follows that (25) is equal to

$$\sum_{\substack{i+j=l_1+l_2-2,\\ 1\leqslant i\leqslant l_1-1}} \left[p_1 p_2 q_1^i q_2^j\right] \Sigma_k^{\mathbf{p}\times\mathbf{q}}. \tag{27}$$

It remains now to count how many times a pair $(i, j)$ contributes to the sum of (24), (25), (27). It is not difficult to see that the only pairs which contribute are $(0, l_1 + l_2 - 2)$ and $(l_1 - 1, l_2 - 1)$, therefore the number of triples described in the formulation of the theorem is equal to the right-hand of (22) which finishes the proof. □

## 6. Combinatorial lemmas

Our strategy of proving the main result of this paper will be to start with the number of factorizations described in Theorem 1.4 and to interpret it as certain linear combination of coefficients of Stanley polynomials for $\Sigma_k$. The first step in this direction is promising: Stanley–Féray character formula (Theorem 4.6) shows that indeed Stanley polynomial for $\Sigma_k$ encodes certain information about the geometry of the bipartite graphs $\mathcal{V}^{\sigma_1,\sigma_2}$ for all factorizations. Unfortunately, condition (e) is quite complicated and at first sight it is not clear how to extract the information about the factorizations fulfilling it from the coefficients of Stanley polynomials.

In this section we will prove three combinatorial lemmas: Corollaries 6.2, 6.4 and 6.5 which solve this difficulty.

### *6.1. Euler characteristic*

Let $\mathcal{I}$ be a family of some subsets of a given finite set $\mathcal{X}$. We define

$$\chi(\mathcal{I}) = \sum_{l\geqslant 1} \sum_{\substack{C=(C_1\subsetneq\cdots\subsetneq C_l),\\ C_1,\ldots,C_l\in\mathcal{I}}} (-1)^{l-1},$$

where the sum runs over all non-empty chains $C = (C_1 \subsetneq \cdots \subsetneq C_l)$ contained in $\mathcal{I}$. In such a situation we will also say that $C$ is $l$-chain and $|C| = l$. Notice that family $\mathcal{I}$ gives rise to a simplicial complex $\mathcal{K}$ with $l-1$-simplices corresponding to $l$-chains contained in $\mathcal{I}$ and the above quantity $\chi(\mathcal{I})$ is just the Euler characteristic of $\mathcal{K}$.

The following lemma shows that under certain assumptions this Euler characteristic is equal to 1; we leave it as an exercise to adapt the proof to show a stronger statement that under the same assumptions $\mathcal{K}$ is in fact contractible (we will not use this stronger result in this article).

**Lemma 6.1.** *Let $\mathcal{I}$ be a non-empty family with a property that*

$$A \cap B \in \mathcal{I} \quad \textit{or} \quad A \cup B \in \mathcal{I} \quad \textit{holds for all } A, B \in \mathcal{I}. \tag{28}$$

*Then*

$$\chi(\mathcal{I}) = 1.$$

**Proof.** Let $\mathcal{X} = \{x_1, \dots, x_n\}$. We define

$$\mathcal{I}_k = \bigl\{A \cup \{x_1, \dots, x_k\}\colon\ A \in I\bigr\}.$$

Clearly $\mathcal{I}_0 = \mathcal{I}$ and $\mathcal{I}_n = \{\mathcal{X}\}$ therefore $\chi(\mathcal{I}_n) = 1$. It remains to prove that $\chi(\mathcal{I}_{k-1}) = \chi(\mathcal{I}_k)$ holds for all $1 \leqslant k \leqslant n$ and we shall do it in the following.

Let us fix $k$. For an $l$-chain $C = (C_1 \subsetneq \cdots \subsetneq C_l)$ contained in $\mathcal{I}_{k-1}$ we define

$$\iota_k(C) = \bigl(C_1 \cup \{x_k\} \subseteq \cdots \subseteq C_l \cup \{x_k\}\bigr)$$

which is a chain contained in $\mathcal{I}_k$. Notice that $\iota_k(C)$ is either an $l-1$-chain (if $C_{i+1} = C_i \cup \{x_k\}$ for some $i$) or $l$-chain (otherwise). With these notations we have

$$\begin{aligned}\chi(\mathcal{I}_{k-1}) &= \sum_{C\colon \text{ non-empty chain in } \mathcal{I}_{k-1}} (-1)^{|C|-1} \\ &= \sum_{D\colon \text{ non-empty chain in } \mathcal{I}_k} \ \sum_{\substack{C\colon \text{ non-empty chain in } \mathcal{I}_{k-1},\\ \iota_k(C)=D}} (-1)^{|C|-1}.\end{aligned}$$

In order to prove $\chi(\mathcal{I}_{k-1}) = \chi(\mathcal{I}_k)$ it is enough now to show that for any non-empty chain $D = (D_1 \subsetneq \cdots \subsetneq D_l)$ contained in $\mathcal{I}_k$

$$(-1)^{|D|-1} = \sum_{\substack{C\colon \text{ non-empty chain in } \mathcal{I}_{k-1},\\ \iota_k(C)=D}} (-1)^{|C|-1}. \tag{29}$$

Let $1 \leqslant p \leqslant l$ be the maximal index with a property that $D_p \notin \mathcal{I}_{k-1}$; if no such index exists we set $p = 0$. In the remaining part of this paragraph we will show that $D_i \setminus \{x_k\} \in \mathcal{I}_{k-1}$ holds for all $1 \leqslant i \leqslant p$. Clearly, in the cases when $p = 0$ or $i = p$ there is nothing to prove. Assume that $i < p$ and $D_i \setminus \{x_k\} \notin \mathcal{I}_{k-1}$. Since $D_i \in \mathcal{I}_k$ it follows that $x_k \in D_i \in \mathcal{I}_{k-1}$. It is easy to check that an analogue of (28) holds true for the family $\mathcal{I}_{k-1}$. We apply this property for $A = D_i$ and $B = D_p \setminus \{x_k\}$ which results in a contradiction since $A \cap B = D_i \setminus \{x_k\} \notin \mathcal{I}_{k-1}$ and $A \cup B = D_p \notin \mathcal{I}_{k-1}$.

Let $1 \leqslant q \leqslant l$ be the minimal index with a property that $x_k \in D_q$; if no such index exists we set $q = n + 1$. Similarly as above we show that $D_i \in \mathcal{I}_{k-1}$ holds for all $q \leqslant i \leqslant l$.

The above analysis shows that a chain $C$ contained in $\mathcal{I}_{k-1}$ such that $\iota_k(C) = D$ must have one of the following two forms:

(1) if $C = (C_1, \dots, C_l)$ is a $l$-chain then there exists a number $r$ $(p \leqslant r < q)$ such that

$$C_i = \begin{cases} D_i \setminus \{x_k\} & \text{for } 1 \leqslant i \leqslant r, \\ D_i & \text{for } r < i \leqslant l, \end{cases}$$

(2) if $C = (C_1, \ldots, C_{l+1})$ is a $l+1$-chain then there exists a number $r$ $(p < r < q)$ such that

$$C_i = \begin{cases} D_i \setminus \{x_k\} & \text{for } 1 \leqslant i \leqslant r, \\ D_{i-1} & \text{for } r < i \leqslant l+1. \end{cases}$$

There are $q - p$ choices for the first case and there are $q - p - 1$ choices for the second case and (29) follows. □

*6.2. Applications of Euler characteristic*

**Corollary 6.2.** *Let $\sigma_1, \sigma_2 \in \mathfrak{S}(k)$ be permutations such that $\langle \sigma_1, \sigma_2 \rangle$ acts transitively and let $q : C(\sigma_2) \to \{2, 3, \ldots\}$ be a coloring with a property that*

$$\sum_{i \in C(\sigma_2)} q(i) = |C(\sigma_1)| + |C(\sigma_2)|.$$

*We define $\mathcal{I}$ to be a family of the sets $A \subset C(\sigma_2)$ with the following two properties:*

- *$A \neq \emptyset$ and $A \neq C(\sigma_2)$,*
- *there are at most $\sum_{i \in A}(q(i) - 1)$ cycles of $\sigma_1$ which intersect $\bigcup A$.*

*Then*

$$\sum_{l \geqslant 0} \sum_{\substack{C = (C_1 \subsetneq \cdots \subsetneq C_l), \\ C_1, \ldots, C_l \in \mathcal{I}}} (-1)^l = \begin{cases} 1 & \text{if } \mathcal{I} = \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{30}$$

**Proof.** It is enough to prove that the family $\mathcal{I}$ fulfills the assumption of Lemma 6.1; we shall do it in the following. For $A \subseteq C(\sigma_2)$ we define

$$f(A) = \Big(\text{number of cycles of } \sigma_1 \text{ which intersect } \bigcup A\Big) - \sum_{i \in A} \big(q(i) - 1\big).$$

In this way $A \in \mathcal{I}$ iff $A \neq \emptyset$, $A \neq C(\sigma_2)$ and $f(A) \leqslant 0$.

It is easy to check that for any $A, B \subseteq C(\sigma_2)$

$$f(A) + f(B) \geqslant f(A \cup B) + f(A \cap B) \tag{31}$$

therefore if $A, B \in \mathcal{I}$ then $f(A \cup B) \leqslant 0$ or $f(A \cap B) \leqslant 0$. If $A \cap B \neq \emptyset$ and $A \cup B \neq C(\sigma_2)$ this finishes the proof. Since $f(\emptyset) = f(C(\sigma_2)) = 0$ also the case when either $A \cap B = \emptyset$ or $A \cup B = C(\sigma_2)$ follows immediately.

It follows that if $A, B \in \mathcal{I}$ and $A \cap B, A \cup B \notin \mathcal{I}$ then $A, B \neq \emptyset$, $A \cap B = \emptyset$, $A \cup B = C(\sigma_2)$, $f(A) = f(B) = 0$. The latter equality shows that

$$\Big(\text{number of cycles of } \sigma_1 \text{ which intersect } \bigcup A\Big) + \Big(\text{number of cycles of } \sigma_1 \text{ which intersect } \bigcup B\Big) = |C(\sigma_1)|$$

therefore each cycle of $\sigma_1$ intersects either $\bigcup A$ or $\bigcup B$ which contradicts transitivity. □

**Lemma 6.3.** *For any $n \geqslant 1$*

$$\sum_k (-1)^k \left\{ {n \atop k} \right\} k! = (-1)^n, \tag{32}$$

*where $\left\{ {n \atop k} \right\}$ denotes the Stirling symbol of the first kind, namely the number of ways of partitioning $n$-element set into $k$ non-empty classes.*

**Proof.** A simple inductive proof follows from the recurrence relation

$$\left\{ {n \atop k} \right\} = k \left\{ {n-1 \atop k} \right\} + \left\{ {n-1 \atop k-1} \right\}. \qquad \square$$

**Corollary 6.4.** *Let $r \geqslant 1$ and let $k_1, \dots, k_r$ and $n_1, \dots, n_r$ be numbers such that $k_1 + \dots + k_r = n_1 + \dots + n_r$. We define $\mathcal{I}$ to be a family of the sets $A \subset \{1, \dots, r\}$ with the following properties: $A \neq \emptyset$ and $A \neq \{1, \dots, r\}$ and*

$$\sum_{i \in A} k_i \leqslant \sum_{i \in A} n_i.$$

*Then*

$$\sum_{l \geqslant 0} \sum_{\substack{C=(C_1 \subsetneq \cdots \subsetneq C_l), \\ C_1, \dots, C_l \in \mathcal{I}}} (-1)^l = \begin{cases} (-1)^{r-1} & \text{if } (k_1, \dots, k_r) = (n_1, \dots, n_r), \\ 0 & \text{otherwise.} \end{cases} \tag{33}$$

**Proof.** If $(k_1, \dots, k_r) = (n_1, \dots, n_r)$ then $\mathcal{I}$ consists of all subsets of $\{1, \dots, r\}$ with the exception of $\emptyset$ and $\{1, \dots, r\}$. Therefore there is a bijective correspondence between the chains $C = (C_1 \subsetneq \dots \subsetneq C_l)$ which contribute to the left-hand side of (33) and sequences $(D_1, \dots, D_{l+1})$ of non-empty and disjoint sets such that $D_1 \cup \dots \cup D_{l+1} = \{1, 2, \dots, r\}$; this correspondence is defined by requirement that

$$C_i = D_1 \cup \dots \cup D_i.$$

It follows that the left-hand side of (33) is equal to

$$\sum_l (-1)^l \left\{ {n \atop l+1} \right\} (l+1)!$$

which can be evaluated thanks to (32).

We consider the case when $(k_1, \dots, k_r) \neq (n_1, \dots, n_r)$; for simplicity we assume that $k_1 \neq n_1$. We define

$$f(A) = \sum_{i \in A} (k_i - n_i)$$

which fulfills (31) and similarly as in the proof of Corollary 6.2 we conclude that condition (28) is fulfilled under additional assumption that $A \cap B \neq \emptyset$ or $A \cup B \neq \{1, 2, \dots, r\}$; this means that Lemma 6.1 cannot be applied directly and we must analyze the details of its proof. We select the sequence $x_1, x_2, \dots$ used in the proof of Lemma 6.1 in such a way that $x_1 = 1$. A careful inspection shows that the proof of the equality $\chi(\mathcal{I}_0) = \chi(\mathcal{I}_1)$ presented above is still valid. Since families $\mathcal{I}_1, \mathcal{I}_2, \dots$ fulfill condition (28) therefore $\chi(\mathcal{I}_1) = \chi(\mathcal{I}_2) = \cdots = 0$. □

**Corollary 6.5.** *Let $r \geqslant 1$, let $\Pi \in P(1, 2, \dots, r)$ be a partition, let $n_1, \dots, n_r$ be numbers and let $\phi : \Pi \to \mathbb{R}$ be a function on the set of blocks of the partition $\Pi$ with a property that*

$$\sum_{b \in \Pi} \phi(b) = n_1 + \cdots + n_r$$

*and $\phi(b) \geqslant |b|$ holds for each block $b \in \Pi$. We define $\mathcal{I}$ to be a family of the sets $A \subset \{1, \dots, r\}$ with the following properties: $A \neq \emptyset$ and $A \neq \{1, \dots, r\}$ and*

$$\sum_{\substack{b \in \Pi, \\ b \cap A \neq \emptyset}} \big(\phi(b) - |b \setminus A|\big) \leqslant \sum_{i \in A} n_i .$$

*Then*

$$\sum_{l \geqslant 0} \sum_{\substack{C = (C_1 \subsetneq \cdots \subsetneq C_l), \\ C_1, \dots, C_l \in \mathcal{I}}} (-1)^l = \begin{cases} (-1)^{|\Pi|-1} & \text{if } \phi(b) = \sum_{i \in b} n_i \text{ holds for each block } b \in \Pi, \\ 0 & \text{otherwise.} \end{cases} \tag{34}$$

**Proof.** We define

$$f(A) = |A| + \sum_{\substack{b \in \Pi, \\ b \cap A \neq \emptyset}} \big(\phi(b) - |b|\big)$$

which fulfills (31). The remaining part of the proof is analogous to Corollary 6.4. □

## 7. Proof of the main result

We will prove Theorem 1.4 in the following equivalent form.

**Theorem 7.1** *(The main result, reformulated). Let $k \geqslant 1$ and let $n_1, \dots, n_r \geqslant 2$ be a sequence of integers. The derivative of Kerov polynomial*

$$\frac{\partial}{\partial R_{n_1}} \cdots \frac{\partial}{\partial R_{n_r}} K_k \bigg|_{R_2 = R_3 = \cdots = 0}$$

*is equal to the number of triples $(\sigma_1, \sigma_2, \ell)$ with the following properties:*

(a) *$\sigma_1, \sigma_2$ is a factorization of the cycle; in other words $\sigma_1, \sigma_2 \in \mathfrak{S}(k)$ are such that $\sigma_1 \circ \sigma_2 = (1, 2, \dots, k)$;*

(b) $|C(\sigma_2)| = r$;
(c) $|C(\sigma_1)| + |C(\sigma_2)| = n_1 + \cdots + n_r$;
(d) $\ell : C(\sigma_2) \to \{1, \dots, r\}$ *is a bijection;*
(e) *for every set* $A \subset C(\sigma_2)$ *which is nontrivial (i.e.,* $A \neq \emptyset$ *and* $A \neq C(\sigma_2)$*) we require that there are more than* $\sum_{i \in A}(n_{\ell(i)} - 1)$ *cycles of* $\sigma_1$ *which intersect* $\bigcup A$.

**Proof.** Let us sum both sides of (30) over all triples $(\sigma_1, \sigma_2, \ell)$ for which conditions (a)–(d) are fulfilled; for such triples we define the coloring $q : C(\sigma_2) \to \{2, 3, \dots\}$ by $q(i) = n_{\ell(i)}$. It follows that the number of triples which fulfill all conditions from the formulation of the theorem is equal to

$$\sum_{l \geqslant 0} \sum_{\substack{C=(C_1,\dots,C_l),\\ \emptyset \subsetneq C_1 \subsetneq \cdots \subsetneq C_l \subsetneq \{1,2,\dots,r\}}} (-1)^l \operatorname{Bad}_C, \tag{35}$$

where $\operatorname{Bad}_C$ for $C = (C_1, \dots, C_l)$ denotes the number of triples $(\sigma_1, \sigma_2, \ell)$ which fulfill (a)–(d) and such that for each $1 \leqslant j \leqslant l$ there are at most $\sum_{i \in C_j}(n_i - 1)$ cycles of $\sigma_1$ which intersect $\bigcup_{i \in C_j} \ell^{-1}(i)$.

Theorem 4.6 shows that

$$\operatorname{Bad}_C = (-1)^{r-1} \sum_{k_1,\dots,k_r} \left[p_1 \cdots p_r q_1^{k_1-1} \cdots q_r^{k_r-1}\right] \Sigma_k^{\mathbf{p}\times\mathbf{q}}, \tag{36}$$

where the above sum is taken over all integers $k_1, \dots, k_r \geqslant 1$ such that

$$k_1 + \cdots + k_r = n_1 + \cdots + n_r \tag{37}$$

and

$$\underbrace{k_{r+1-|C_j|} + \cdots + k_r}_{|C_j| \text{ summands}} \leqslant \sum_{i \in C_j} n_i \quad \text{holds for each } 1 \leqslant j \leqslant l. \tag{38}$$

We apply Lemma 4.5 to the right-hand side of (36). Therefore

$$\operatorname{Bad}_C = (-1)^{r-1} \sum_{\Pi \in P(1,2,\dots,r)} \sum_{k_1,\dots,k_r} Q^{\Sigma_k}_{((\sum_{i \in b} k_i, |b|):\ b \in \Pi)}, \tag{39}$$

where the first sum runs over all partitions $\Pi$ of the set $\{1, 2, \dots, r\}$ and the second sum runs over the tuples $k_1, \dots, k_r$ which fulfill conditions (37), (38) and such that the set of indices $i$ such that $k_i \geqslant 2$ coincides with the set of rightmost legs of the blocks of $\Pi$.

For simplicity, before dealing with the general case, we shall analyze first the contribution of the trivial partition which consists only of singletons. We define $\operatorname{Bad}_C^{\text{trivial}}$ to be the expression (39) with the sum over partitions replaced by only one summand for $\Pi = \{\{1\}, \{2\}, \dots, \{r\}\}$, i.e.

$$\operatorname{Bad}_C^{\text{trivial}} = (-1)^{r-1} \sum_{k_1,\dots,k_r} Q^{\Sigma_k}_{(k_1,1),\dots,(k_r,1)}, \tag{40}$$

where the sum runs over the same set as in Eq. (36) with an additional restriction $k_1, \dots, k_r \geqslant 2$.

Corollary 4.3 shows that we may change the order of the elements in the sequence $(k_1, \dots, k_r)$ hence (40) holds true also if the sum on the right-hand side runs over all integers $k_1, \dots, k_r \geqslant 2$ such that $k_1 + \dots + k_r = n_1 + \dots + n_r$ and such that for each $1 \leqslant j \leqslant l$

$$\sum_{i \in C_j} k_i \leqslant \sum_{i \in C_j} n_i.$$

Therefore for an analogue of the sum (35) Corollary 6.4 shows that

$$\sum_{l \geqslant 0} \sum_{\substack{C=(C_1,\dots,C_l), \\ \emptyset \subsetneq C_1 \subsetneq \dots \subsetneq C_l \subsetneq \{1,2,\dots,r\}}} (-1)^l \operatorname{Bad}_C^{\text{trivial}} = (-1)^{r-1} Q^{\Sigma_k}_{(n_1,1),\dots,(n_r,1)}.$$

Notice the right-hand side is the summand appearing in Corollary 4.4 for the trivial partition $\Pi$ which is quite encouraging.

Having in mind the above simplified case let us tackle the general partitions $\Pi$. Corollary 4.3 shows that we may shuffle the blocks of partition $\Pi$ hence from (39) it follows that

$$\operatorname{Bad}_C = (-1)^{r-1} \sum_{\Pi \in P(1,2,\dots,r)} \sum_{\phi} Q^{\Sigma_k}_{((\phi(b),|b|):\ b \in \Pi)}, \tag{41}$$

where the second sum runs over all functions $\phi$ which assign integer numbers to blocks of $\Pi$ and such that:

- $\phi(b) \geqslant |b| + 1$ holds for every block $b \in \Pi$;
- $\sum_{b \in \Pi} \phi(b) = n_1 + \dots + n_r$;
- $$\sum_{\substack{b \in \Pi, \\ b \cap C_j \neq \emptyset}} \big(\phi(b) - |b \setminus C_j|\big) \leqslant \sum_{i \in C_j} n_i$$

  holds for each $1 \leqslant j \leqslant l$.

Therefore (35) is equal to

$$\sum_{\Pi \in P(1,2,\dots,r)} \sum_{\phi} Q^{\Sigma_k}_{((\phi(b),|b|):\ b \in \Pi)} \left[ \sum_{l \geqslant 0} \sum_{\substack{C=(C_1,\dots,C_l), \\ \emptyset \subsetneq C_1 \subsetneq \dots \subsetneq C_l \subsetneq \{1,2,\dots,r\}}} (-1)^{l+r-1} \right].$$

Corollary 6.5 can be used to calculate the expression in the bracket hence the above sum is equal to

$$\sum_{\Pi \in P(1,2,\dots,r)} Q^{\Sigma_k}_{((\sum_{i \in b} n_i, |b|):\ b \in \Pi)} (-1)^{r-|\Pi|}.$$

Corollary 4.4 finishes the proof. $\square$

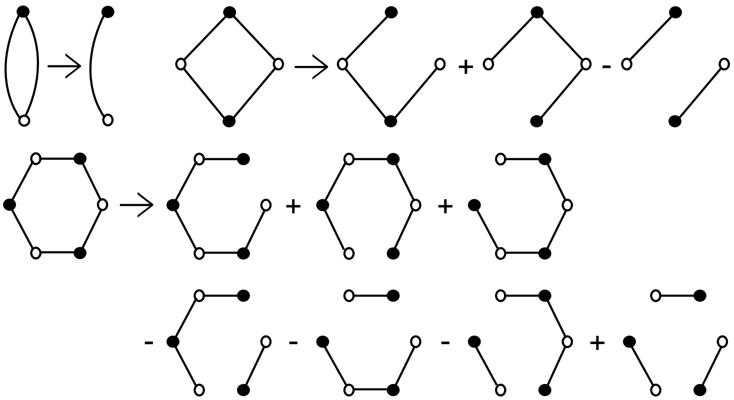

Fig. 4. Local transformations on graphs.

Proof of Theorem 1.6 is analogous (the reference to Theorem 4.6 should be replaced by Theorem 4.7) and we skip it.

## 8. Graph decomposition

In this section we will compare our main result with the previous complicated combinatorial description of the coefficients of Kerov's polynomials proposed by Féray in [8]. This will lead us to a new proof of the main result of this paper, Theorems 1.4 and 1.6.

### *8.1. Reformulation of the previous result*

Let us consider the formal sum of the collection of graphs $(\mathcal{V}^{\sigma_1,\sigma_2})_{\sigma_1,\sigma_2}$ over all factorizations $\sigma_1 \cdot \sigma_2 = (1, 2, \dots, k)$ (these graphs were defined in Section 1.8 but in order to be compatible with the notation of the paper [8] it might be more convenient to allow multiple edges connecting two cycles with the multiplicity equal to the number of the elements in the common support). Let us apply the local transformations presented on Fig. 4 (the reader can easily imagine the generalizations of the drawn transformation to bigger loops: for a given oriented loop of length $2k$ we remove in $2^k - 1$ ways all non-empty subsets of the set of edges oriented from a black vertex to a white vertex with the plus or minus sign depending if the number of removed edges is odd or even) to each of the summands and let us interate this procedure until we obtain a formal linear combination of forests. Of course, the final result $S$ may depend on the choice of the loops used for the transformations, so in order to have a uniquely determined result we have to choose the loops in some special way, for example as described in paper [8], the details of which will not be important for this article.

Then we have the following result:

**Theorem 8.1.** *(See Féray [8].) The coefficient of* $R_2^{s_2} R_3^{s_3} \cdots$ *in* $K_k$ *is equal to* $(-1)^{1+s_2+s_3+\cdots}$ *times the total sum of coefficients of all forests in* $S$ *which consist of* $s_i$ *trees with one black and* $i-1$ *white vertices (*$i$ *runs over* $\{2, 3, \dots\}$*).*

We will reformulate this result in a form closer to Theorem 1.4. For this purpose, if $(\sigma_1, \sigma_2, q)$ is a triple verifying conditions (a)–(d) and $F$ is a subforest of $\mathcal{V}^{\sigma_1,\sigma_2}$ with the same set of vertices, we will say that $F$ is a $q$-forest if the following two conditions are fulfilled:

- all cycles of $\sigma_2$ (black vertices) are in different connected components,
- each cycle $c$ of $\sigma_2$ is the neighbor of exactly $q(c) - 1$ cycles of $\sigma_1$ (white vertices).

**Theorem 8.2.** *Let $k \geqslant 1$ and let $s_2, s_3, \ldots$ be a sequence of non-negative integers with only finitely many non-zero elements. The coefficient of the monomial $R_2^{s_2} R_3^{s_3} \cdots$ in the Kerov polynomial $K_k$ is equal to the number of triples $(\sigma_1, \sigma_2, q)$ which fulfill conditions* (a)–(d) *of Theorem* 1.4 *and such that*

(e$^5$) *when we apply the transformations from Fig.* 4 *as prescribed in* [8, Section 3]*, in the resulting linear combination of forests there is (exactly one) $q$-forest.*

It is easy to see that this theorem is a reformulation of Theorem 8.1. A priori, it might seem that in the theorem above we should count each triplet $(\sigma_1, \sigma_2, q)$ with multiplicity equal to the number of $q$-forests appearing in the result, but we will prove in Corollary 8.4 that it is always equal to 0 or 1.

Comparing Theorem 8.2 with Theorem 1.4 we may wonder if conditions (e) and (e$^5$) are equivalent. We will prove their equivalence in the following section.

### *8.2. Equivalence of conditions (e) and (e$^5$)*

In Section 1.9 we introduced the notion of $q$-admissibility of a graph. Recall that if a graph $G$ is connected then it is $q$-admissible if and only if it satisfies condition (e$^3$) which is a reformulation of (e). Notice also that if $G$ contains no loops then it is $q$-admissible if and only if it is a $q$-forest.

**Lemma 8.3.** *The sum of coefficients of $q$-admissible graphs $G$ multiplied by $(-1)^{(\text{number of connected components of } G)}$ in a formal linear combination of bipartite graphs with a given set of vertices and labeling $q : V_\bullet \to \{2, 3, \ldots\}$ does not change after performing any transformation of the form presented on Fig.* 4.

**Proof.** Let us choose some oriented loop $L$ in graph $G$ and let us denote by $E$ the set of edges which can be erased in the corresponding local transformation from Fig. 4; in other words $E$ consists of every second edge in the loop $L$.

Consider the convex polyhedron $P$ (without boundary) which is the set of all positive solutions $(x_e)_{e \text{ is an edge of } G}$ to the system of equations from condition (e$^4$).

If $f$ is a real function on the set of edges of $G$ and $v$ is a vertex of $G$ we define $(\Phi(f))(v)$ to be the sum of values of $f$ on edges adjacent to $v$. If $P$ is non-empty then its dimension is equal to the dimension of $\ker \Phi$. It is a simple exercise to show that $\operatorname{Im} \Phi$ consists of all functions on vertices of $G$ with a property that for each connected component of $G$ the sum of values on black vertices is equal to the sum of values on white vertices hence

$$\dim \operatorname{Im} \Phi = (\text{number of vertices of } G) - (\text{number of components of } G).$$

It follows from rank-nullity theorem that

$$\begin{aligned}\dim P &= \dim\ker\Phi \\ &= (\text{number of edges of } G) - (\text{number of vertices of } G) \\ &\quad + (\text{number of connected components of } G). \end{aligned} \tag{42}$$

For a positive solution $(x_e)$ of our system of equations and a real number $t$ we define

$$y_e = \begin{cases} x_e & \text{if } e \notin L, \\ x_e + t & \text{if } e \in (L \setminus E), \\ x_e - t & \text{if } e \in E. \end{cases}$$

which is also a solution. Let $t$ be the minimal positive number for which $(y_e)$ is not positive. In this way we define a map $\Pi : (x_e) \mapsto (y_e)$.

For any non-empty $A \subseteq E$ we define $P_A$ to be the set of positive solutions with a property that

$$\forall_{e\in E} \quad e \in A \quad \iff \quad x_e = \min_{f\in E} x_f .$$

Since the defining condition for $P_A$ can be written in terms of some equations and inequalities it follows that $P_A$ is a convex polyhedron. It is easy to check that

$$\Pi(P_A) = \bigl\{(x_e)\colon \text{ non-negative solution such that } \forall_{e\colon \text{ edge of } G}\ (x_e = 0) \iff (e \in A)\bigr\}.$$

The latter set can be identified with the set of positive solutions for our system of equations corresponding to the graph $G' = G \setminus A$. It follows that

$$\begin{aligned}\dim P_A &= 1 + \dim \Pi(P_A) \\ &= 1 + \bigl(\text{number of edges of } (G \setminus A)\bigr) - (\text{number of vertices of } G) \\ &\quad + \bigl(\text{number of connected components of } (G \setminus A)\bigr), \end{aligned} \tag{43}$$

where the last equality is just (42) applied to $G' = G \setminus A$.

It is easy to see that $P = \bigsqcup_{A\neq\emptyset} P_A$ is a disjoint union therefore we have the equality between the Euler characteristics:

$$\chi(P) = \sum_{A\neq\emptyset} \chi(P_A)$$

which thanks to (42) and (43) shows that

$$\begin{aligned}&(-1)^{(\text{number of connected components of } G)}\ [P \text{ is non-empty}] \\ &\quad = \sum_{A\neq\emptyset} (-1)^{|A|-1} (-1)^{(\text{number of components of } G\setminus A)}\ [P_A \text{ is non-empty}], \end{aligned} \tag{44}$$

where we use the convention that

$$\big[\text{(condition)}\big] = \begin{cases} 1 & \text{if (condition) is true,} \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 1.9 shows that $P$ (respectively, $P_A$) is non-empty if and only if $G$ (respectively, $G \setminus A$) is $q$-admissible therefore (44) is equivalent to

$$\begin{aligned} &(-1)^{\text{(number of components of } G)} \, [G \text{ is } q\text{-admissible}] \\ &\quad = \sum_{A \neq \emptyset} (-1)^{|A|-1} (-1)^{\text{(number of components of } G\setminus A)} \, \big[(G \setminus A) \text{ is } q\text{-admissible}\big], \end{aligned}$$

which is the desired equality. □

**Corollary 8.4.** *Suppose that* $(\sigma_1, \sigma_2, q)$ *is a triple verifying the conditions* (a)–(d) *of Theorem* 1.4. *If we iterate local transformations from Fig.* 4 *on* $\mathcal{V}^{\sigma_1,\sigma_2}$ *until we obtain a formal linear combination of forests (not necessarily choosing the loops as prescribed in* [8]*) then the sum of coefficients of* $q$*-forests in the result is equal to*

$$\begin{cases} (-1)^{1+s_2+s_3+\cdots} & \textit{if condition (e) is fulfilled;} \\ 0 & \textit{otherwise.} \end{cases}$$

*In the case when we perform the transformations as prescribed in* [8, Section 3]*, the sign property of this decomposition* [8, Proposition 3.3.1] *implies that there is exactly one* $q$*-forest (with the appropriate sign) in the resulting sum if condition* (e) *is fulfilled and there are no* $q$*-forests otherwise; in other words condition* (e) *is equivalent to* (e$^5$).

The above corollary together with Theorem 8.2 give another proof of the main result of the paper, Theorem 1.4. Analogous results can be stated for the situation presented in Theorem 1.6.

## Acknowledgments

Research of P.Ś. is supported by the MNiSW research grant P03A 013 30, by the EU Research Training Network "QP-Applications", contract HPRN-CT-2002-00279 and by the EC Marie Curie Host Fellowship for the Transfer of Knowledge "Harmonic Analysis, Nonlinear Analysis and Probability", contract MTKD-CT-2004-013389.

P.Ś. thanks Marek Bożejko, Philippe Biane, Akihito Hora, Jonathan Novak, Światosław Gal and Jan Dymara for several stimulating discussions during various stages of this research project.

## References

[1] Philippe Biane, Representations of symmetric groups and free probability, Adv. Math. 138 (1) (1998) 126–181.

[2] Philippe Biane, Free cumulants and representations of large symmetric groups, in: XIIIth International Congress on Mathematical Physics, London, 2000, Int. Press, Boston, MA, 2001, pp. 321–326.

[3] Philippe Biane, Characters of symmetric groups and free cumulants, in: Asymptotic Combinatorics with Applications to Mathematical Physics, St. Petersburg, 2001, in: Lecture Notes in Math., vol. 1815, Springer, Berlin, 2003, pp. 185–200.

[4] Philippe Biane, On the formula of Goulden and Rattan for Kerov polynomials, Sém. Lothar. Combin., 55, Art. B55d, 5 pp. (electronic), 2005/07.
[5] Sylvie Corteel, Alain Goupil, Gilles Schaeffer, Content evaluation and class symmetric functions, Adv. Math. 188 (2) (2004) 315–336.
[6] Valentin Féray, Stanley's formula for characters of the symmetric group, Ann. Comb. 13 (2010) 453–461.
[7] Valentin Féray, Piotr Śniady, Asymptotics of characters of symmetric groups related to Stanley–Féray character formula, preprint available at arXiv:math/0701051, 2007.
[8] Valentin Féray, Combinatorial interpretation and positivity of Kerov's character polynomials, J. Algebraic Combin. 29 (4) (2009) 473–507.
[9] I.P. Goulden, A. Rattan, An explicit form for Kerov's character polynomials, Trans. Amer. Math. Soc. 359 (8) (2007) 3669–3685 (electronic).
[10] Vladimir Ivanov, Grigori Olshanski, Kerov's central limit theorem for the Plancherel measure on Young diagrams, in: Symmetric Functions 2001: Surveys of Developments and Perspectives, in: NATO Sci. Ser. II Math. Phys. Chem., vol. 74, Kluwer Acad. Publ., Dordrecht, 2002, pp. 93–151.
[11] Sergei Kerov, Interlacing measures, in: Kirillov's Seminar on Representation Theory, in: Amer. Math. Soc. Transl. Ser. 2, vol. 181, Amer. Math. Soc., Providence, RI, 1998, pp. 35–83.
[12] S. Kerov, A differential model for the growth of Young diagrams, in: Proceedings of the St. Petersburg Mathematical Society, vol. IV, in: Amer. Math. Soc. Transl. Ser. 2, vol. 188, Amer. Math. Soc., Providence, RI, 1999, pp. 111–130.
[13] S. Kerov, Talk in Institute Henri Poincaré, Paris, January 2000.
[14] S.V. Kerov, Asymptotic representation theory of the symmetric group and its applications in analysis, Transl. Math. Monogr., vol. 219, American Mathematical Society, Providence, RI, 2003, translated from the Russian manuscript by N.V. Tsilevich, With a foreword by A. Vershik and comments by G. Olshanski.
[15] Sergei K. Lando, Alexander K. Zvonkin, Graphs on surfaces and their applications, in: Encyclopaedia Math. Sci., vol. 141, Springer-Verlag, Berlin, 2004, With an appendix by Don B. Zagier, Low-Dimensional Topology, II.
[16] Michel Lassalle, Two positivity conjectures for Kerov polynomials, Adv. in Appl. Math. 41 (3) (2008) 407–422.
[17] B.F. Logan, L.A. Shepp, A variational problem for random Young tableaux, Adv. Math. 26 (2) (1977) 206–222.
[18] Amarpreet Rattan, Piotr Śniady, Upper bound on the characters of the symmetric groups for balanced Young diagrams and a generalized Frobenius formula, Adv. Math. 218 (3) (2008) 673–695.
[19] Yuval Roichman, Upper bound on the characters of the symmetric groups, Invent. Math. 125 (3) (1996) 451–485.
[20] Piotr Śniady, Asymptotics of characters of symmetric groups, genus expansion and free probability, Discrete Math. 306 (7) (2006) 624–665.
[21] Piotr Śniady, Gaussian fluctuations of characters of symmetric groups and of Young diagrams, Probab. Theory Related Fields 136 (2) (2006) 263–297.
[22] Roland Speicher, The lattice of admissible partitions, in: Quantum Probability & Related Topics, QP-PQ, vol. VIII, World Sci. Publ., River Edge, NJ, 1993, pp. 347–352.
[23] Roland Speicher, Combinatorial theory of the free product with amalgamation and operator-valued free probability theory, Mem. Amer. Math. Soc. 132 (627) (1998) x+88.
[24] Richard P. Stanley, Kerov's character polynomial and irreducible symmetric group characters of rectangular shape. Transparencies from a Conference in Québec City, June 2002.
[25] Richard P. Stanley, Irreducible symmetric group characters of rectangular shape, Sém. Lothar. Combin., 50, Art. B50d, 11 pp. (electronic), 2003/04.
[26] Richard P. Stanley, A conjectured combinatorial interpretation of the normalized irreducible character values of the symmetric group, preprint available at arXiv:math.CO/0606467, 2006.
[27] A.M. Veršik, 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 (6) (1977) 1024–1027.
[28] Dan Voiculescu, Addition of certain noncommuting random variables, J. Funct. Anal. 66 (3) (1986) 323–346.
[29] Dan Voiculescu, Limit laws for random matrices and free products, Invent. Math. 104 (1) (1991) 201–220.
[30] D.V. Voiculescu, K.J. Dykema, A. Nica, A noncommutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups, in: Free Random Variables, in: CRM Monogr. Ser., vol. 1, American Mathematical Society, Providence, RI, 1992.