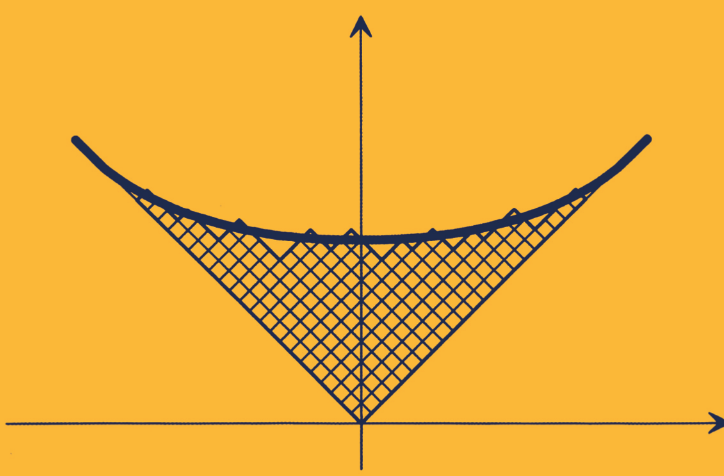
Lecture Notes in Mathematics

1815

Anatoly M. Vershik (Ed.)

Asymptotic Combinatorics with Applications to Mathematical Physics

St. Petersburg 2001



Springer

Lecture Notes in Mathematics 1815

Editors:
J.–M. Morel, Cachan
F. Takens, Groningen
B. Teissier, Paris

Subseries:
European Mathematical Society

Springer
Berlin
Heidelberg
New York
Hong Kong
London
Milan
Paris
Tokyo

Anatoly M. Vershik (Ed.)

Asymptotic Combinatorics with Applications to Mathematical Physics

A European Mathematical
Summer School held at the Euler Institute,
St. Petersburg, Russia
July 9-20, 2001

Springer

Editor

Anatoly M. Vershik

St. Petersburg Department
of the Mathematical Institute
Russian Academy of Sciences
Fontanka 27
St. Petersburg 191011, Russia

e-mail: vershik@pdmi.ras.ru
http://www.pdmi.ras.ru/˜vershik

Technical Editor

Yuri Yakubovich

St. Petersburg Department
of the Mathematical Institute
Russian Academy of Sciences
Fontanka 27
St. Petersburg 191011, Russia

e-mail: yuyakub@mail.ru

Cataloging-in-Publication Data applied for

Bibliographic information published by Die Deutsche Bibliothek

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at http://dnb.ddb.de

Mathematics Subject Classification (2000): 46-XX, 05-XX, 60-XX, 35-XX

ISSN 0075-8434
ISBN 3-540-40312-4 Springer-Verlag Berlin Heidelberg New York

Springer-Verlag Berlin Heidelberg New York a member of BertelsmannSpringer Science + Business Media GmbH

http://www.springer.de

Typesetting: Camera-ready TEX output by the authors

SPIN: 10933565   41/3142/du - 543210 - Printed on acid-free paper

Preface

A European Mathematical Summer School entitled "ASYMPTOTIC COMBINATORICS WITH APPLICATIONS TO MATHEMATICAL PHYSICS" was held at St. Petersburg, Russia, 9–22 July 2001. This meeting was at the same time a NATO Advanced Studies Institute.

It was cosponsored by NATO Science Committee, European Mathematical Society, and Russian Fund for Basic Research.

This volume contains mathematical lectures from the school. Another part of the materials presented at the School, more related to mathematical physics, is already published[1]. Information about the School, its participants, program, etc. can be found at the end of this volume.

The present volume contains lecture courses, as well as several separate lectures, which have mainly mathematical rather than physical orientation. They are aimed mostly at non-specialists and beginners who constituted the majority of the participants of the School. I would like to emphasize that splitting the lectures into "physical" and "mathematical" ones is relative. Moreover, the idea of the School was to unite mathematicians and physicists working essentially on the same problems but following different traditions and notations accepted by their communities.

The last few years were marked by an impressive unification of a number of areas in mathematical physics and mathematics. The Summer School presented some of these major – and until recently mutually unrelated – topics: matrix problems (the study of which was initiated by physicists about 25 years ago), asymptotic representation theory of classical groups (which arose in mathematics approximately at the same time), the theory of random matrices (also initiated by physicists but intensively studied by mathematicians), and, finally, the theory of integrable nonlinear problems in mathematical physics with a wide range of related problems. As a result of the new

[1] *Proceedings of NATO ASI Asymptotic Combinatorics with Application to Mathematical Physics,* V. Malyshev and A. Vershik, Eds., Kluwer Academic Publishers, 2002, 328pp.

interrelations dicovered, all these young theories, which constitute the essential part of modern mathematics and mathematical physics, have become part of one large mathematical area. These interconnections are mainly combinatorial. An illustration of this phenomenon is the perception of the fact that the asymptotic theory of Young tableaux and the theory of spectra of random matrices is essentially the same theory, since the asymptotic microstructure of a random Young diagram with respect to the Plancherel measure coincides with the microstructure of the spectrum of a random matrix in the Gaussian ensemble. The corresponding asymptotic distributions are new for the probability theory. They were originally found by Riemann–Hilbert problem techniques (the Riemann–Hilbert problem arises in calculation of the diagonal asymptotics of orthogonal polynomials). In the present volume this direction is represented by the lectures by P. Deift and A. Borodin. Later these distributions were obtained by another method, that calculates the correlation functions directly using direct relations to integrable problems and hierarchies (A. Borodin, A. Okounkov, G. Olshansky).

The lectures by A. Vershik, G. Olshansky, R. Hora, and partially by A. Borodin and P. Biane are devoted to the asymptotic representation theory. This theory studies the asymptotic behavior of characters of classical groups as the rank of the group grows to infinity. It was started in the beginning of 1970s by works of Vershik–Kerov and Logan–Shepp and one of the first results was the proof of the asymptotic behaviour of the characters of symmetric group and Young diagrams. At that time the similarity and relations to quantum chromodynamics and matrix problems were anticipated but not yet clearly understood. Now these relations are well understood; they have become precise statements rather than vague analogies. These relations are also considered in the lectures by E. Bresin and V. Kasakov which have appeared in the other volume of the School proceedings.

From this point of view, four lectures by A. Okounkov take a particular position. They contain a sketch of the complete proof (obtained jointly with R. Pandharipande) of the Witten–Kontsevich formula relating the generating function of important combinatorial numbers of algebraic geometrical origin and the τ-function of the KdV equation hierarchy. In these lectures, special attention is paid to the role of the theory of symmetric functions and asymptotic representation theory, as well as to relations to random matrices.

The lectures by R. Speicher, M. Nazarov, and by M. Bożejko and R. Szwarc are devoted to more special topics which, however, fit in the same context. At present there are hundreds of journal papers on all these subjects, however the time for accomplished presentations is yet to come. The published lectures of the School should stimulate this process. The reader should keep in mind that the references cited in the lectures are not exhaustive. Of course, the relations between asymptotic combinatorics and mathematical physics extend farther than the topics touched upon during the School. For example, closely related combinatorial problems play a key role in conformal field theory which is now developing fast. I hope that the lectures presented in this volume will be useful for beginners as well as for specialists who want to familiarize themselves with this fascinating area of modern mathematics and mathematical physics and start working in it.

Two prominent mathematicians, Anatoly Izergin (1948–1999) and Sergey Kerov (1946–2000), died a year before the conference which had been planned with their active participation. Their contribution to areas of mathematical physics and mathematics related to the topics of the conference was enormous. One of the sessions of the conference was devoted to their memory.

All the work on preparation of this manuscript was carried out by Yu. Yakubovich, to whom I am very grateful. I would like to express my gratitude to the following organizations which helped greatly in the organization of the School: the European Mathematical Society, the NATO Science Committee, the Euler International Mathematical Institute, the St. Petersburg Department of the Mathematical Institute of the Russian Academy of Sciences and the St. Petersburg Mathematical Society.

Anatoly M. Vershik

Contents

Part III Combinatorics and representation theory

Addendum Information about the school

Asymptotic representation theory and Riemann–Hilbert problem

Alexei Borodin

School of Mathematics
Institute for Advanced Study
Einstein Drive
Princeton NJ 08540
U.S.A.
borodine@math.upenn.edu

Summary. We show how the Riemann–Hilbert problem can be used to compute correlation kernels for determinantal point processes arising in different models of asymptotic combinatorics and representation theory. The Whittaker kernel and the discrete Bessel kernel are computed as examples.

Introduction

A (discrete or continuous) random point process is called *determinantal* if its correlation functions have the form

$$\rho_n(x_1, \ldots, x_n) = \det[K(x_i, x_j)]_{i,j=1}^n,$$

where $K(x, y)$ is a function in two variables called the *correlation kernel*. A major source of such point processes is Random Matrix Theory. All the "unitary" or "$\beta = 2$" ensembles of random matrices lead to determinantal point processes which describe the eigenvalues of these matrices.

Determinantal point processes also arise naturally in problems of asymptotic combinatorics and asymptotic representation theory, see [6]–[9], [5], [15], [21]. Usually, it is not very hard to see that the process that we are interested in is determinantal. A harder problem is to compute the correlation kernel of this process explicitly. The goal of this paper is to give an informal introduction to a new method of obtaining explicit formulas for correlation kernels. It should be emphasized that in representation theoretic models which we consider the kernels cannot be expressed through orthogonal polynomials, as it often happens in random matrix models. That is why we had to invent something different.

The heart of the method is the *Riemann–Hilbert problem* (RHP, for short). This is a classical problem which consists of factorizing a matrix–valued function on a contour in the complex plane into a product of a function which

A. Borodin et al.: LNM 1817, A.M. Vershik (Ed.), pp. 3–19, 2003.

is holomorphic inside the contour and a function which is holomorphic outside the contour. It turns out that the problem of computing the correlation kernels can be reduced to solving a RHP of a rather special form. The input of the RHP (the function to be factorized) is always rather simple and can be read off the representation theoretic quantities such as dimensions of irreducible representations of the corresponding groups. We also employ a discrete analog of RHP described in [2].

The special form of our concrete RHPs allows us to reduce them to certain linear ordinary differential equations (this is the key step), which have classical special functions as their solutions. This immediately leads to explicit formulas for the needed correlation kernels.

The approach also happens to be very effective for the derivation of (nonlinear ordinary differential) Painlevé equations describing the "gap probabilities" in both random matrix and representation theoretic models, see [4], [3]. However, this issue will not be addressed in this paper.

The paper is organized as follows. In Section 1 we explain what a determinantal point process is and give a couple of examples. In Section 2 we argue that in many models correlation kernels give rise to what is called "integrable integral operators". In Section 3 we relate integrable operators to RHP. In Section 4 we derive the Whittaker kernel arising in a problem of harmonic analysis on the infinite symmetric group. In Section 5 we derive the discrete Bessel kernel associated with the poissonized Plancherel measures on symmetric groups.

This paper is an expanded text of lectures the author gave at the NATO Advanced Study Institute "Asymptotic combinatorics with applications to mathematical physics" in July 2001 in St. Petersburg. It is a great pleasure to thank the organizers for the invitation and for the warm hospitality. The author would also like to thank Grigori Olshanski and Percy Deift for helpful discussions.

This research was partially conducted during the period the author served as a Clay Mathematics Institute Long-Term Prize Fellow. This work was also partially supported by the NSF grant DMS-9729992.

1 Determinantal point processes

Definition 1. *Let $\mathfrak{X}$ be a discrete space. A probability measure on $2^{\mathfrak{X}}$ is called a determinantal point process if there exists a function $K : \mathfrak{X} \times \mathfrak{X} \to \mathbb{C}$ such that*

$$\mathrm{Prob}\left\{A \in 2^{\mathfrak{X}} \mid A \supset \{x_1, \dots, x_n\}\right\} = \det[K(x_i, x_j)]_{i,j=1}^{n}$$

for any finite subset $\{x_1, \dots .x_n\}$ of $\mathfrak{X}$. The function K is called the correlation kernel. The functions

$$\rho_n : \{n\text{-point subsets of } \mathfrak{X}\} \to [0, 1]$$
$$\rho_n : \{x_1, \dots, x_n\} \mapsto \mathrm{Prob}\{A \mid A \supset \{x_1, \dots, x_n\}\}$$

are called the correlation functions.

Example 1. Consider a kernel $L : \mathfrak{X} \times \mathfrak{X} \to \mathbb{C}$ such that

- $\det[L(x_i, x_j)]_{i,j=1}^k \geq 0$ for all k-point subsets $\{y_1, \dots, y_k\}$ of $\mathfrak{X}$.
- L defines a trace class operator in $\ell^2(\mathfrak{X})$, for example, $\sum_{x,y\in\mathfrak{X}} |L(x,y)| < \infty$ or L is finite rank. In particular, this condition is empty if $|\mathfrak{X}| < \infty$.

Set

$$\mathrm{Prob}\,\{\{y_1, \dots, y_k\}\} = \frac{1}{\det(1+L)} \cdot \det[L(y_i, y_j)]_{i,j=1}^k \,.$$

This defines a probability measure on $2^{\mathfrak{X}}$ concentrated on finite subsets. Moreover, this defines a determinantal point process. The correlation kernel $K(x,y)$ is equal to the matrix of the operator $K = L(1+L)^{-1}$ acting on $\ell^2(\mathfrak{X})$. See [10], [5], Appendix for details.

Definition 2. *Let $\mathfrak{X}$ be a finite or infinite interval inside $\mathbb{R}$ (e.g., $\mathbb{R}$ itself). A probability measure on locally finite subsets of $\mathfrak{X}$ is called a determinantal point process if there exists a function $K : \mathfrak{X} \times \mathfrak{X} \to \mathbb{C}$ such that*

$$\lim_{\Delta x_1, \dots, \Delta x_n \to 0} \frac{\mathrm{Prob}\{A \in 2^{\mathfrak{X}}_{loc.fin.} \,|\, A \text{ intersects } [x_i, x_i + \Delta x_i] \text{ for all } i = 1, \dots, n\}}{\Delta x_1 \cdots \Delta x_n} = \det[K(x_i, x_j)]_{i,j=1}^n$$

for any finite subset $\{x_1, \dots, x_n\}$ of $\mathfrak{X}$. The function K is called the correlation kernel and the left–hand side of the equality above is called the nth correlation function.

Example 2. Let $w(x)$ be a positive function on $\mathfrak{X}$ such that all the moments $\int_{\mathfrak{X}} x^n w(x) dx$ are finite. Pick a number $N \in \mathbb{N}$ and define a probability measure on N-point subsets of $\mathfrak{X}$ by the formula

$$P_N(dx_1, \dots, dx_N) = c_N \prod_{1 \leq i < j \leq N} (x_i - x_j)^2 \prod_{1 \leq k \leq N} w(x_k) dx_k.$$

Here $c_N > 0$ is a normalizing constant. This is a determinantal point process. The correlation kernel is equal to the Nth Christoffel–Darboux kernel $K_N(x,y)$ associated with $w(x)$, multiplied by $\sqrt{w(x)w(y)}$. That is, let

$$p_0 = 1,\ p_1(x),\ p_2(x), \dots$$

be monic (= leading coefficient 1) orthogonal polynomials on $\mathfrak{X}$ with the weight function $w(x)$:

$$p_m(x) = x^m + \text{ lower order terms },$$

$$\int_{\mathfrak{X}} p_m(x) p_n(x) w(x) dx = h_m \delta_{mn}, \quad m, n = 0, 1, 2, \dots .$$

Then the correlation kernel is equal to

$$K_N(x,y) = \sum_{k=0}^{N} \frac{p_k(x)p_k(y)}{h_k} \sqrt{w(x)w(y)}$$
$$= \frac{1}{h_{N-1}} \frac{p_N(x)p_{N-1}(y) - p_{N-1}(x)p_N(y)}{x-y} \sqrt{w(x)w(y)}.$$

The construction of this example also makes sense in the discrete setting. See [12], [18], [19], [15] for details.

Remark 1. The correlation kernel of a determinantal point process is not defined uniquely! In particular, transformations of the form $K(x,y) \to \frac{f(x)}{f(y)}K(x,y)$ do not change the correlation functions.

2 Correlation kernels as integrable operators

Observe that the kernel $K_N(x,y)$ of Example 2 has the form

$$K_N(x,y) = \frac{\phi(x)\psi(y) - \psi(x)\phi(y)}{x-y}$$

for appropriate ϕ and ψ. Most kernels appearing in “$\beta = 2$ ensembles” of Random Matrix Theory have this form, because they are either kernels of Christoffel–Darboux type as in Example 2 above, or scaling limits of such kernels. However, it is an experimental fact that integral operators with such kernels appear in many different areas of mathematics, see [11].

Definition 3. *An integral operator with kernel of the form*

$$\frac{f_1(x)g_1(y) + f_2(x)g_2(y)}{x-y} \tag{1}$$

is called integrable. Here we assume that $f_1(x)g_1(x) + f_2(x)g_2(x) = 0$ *so that there is no singularity on the diagonal. Diagonal values of the kernel are then defined by continuity.*

The class of integrable operators was singled out in the work of Its, Izergin, Korepin, and Slavnov on quantum inverse scattering method in 1990 [14].

We will also call an operator acting in the ℓ^2*-space on a discrete space integrable if its matrix has the form* (1)*. It is not obvious how to define the diagonal entries of a discrete integrable operator in general. However, in all concrete situations we are aware of, this question has a natural answer.*

Example 3 (poissonized Plancherel measure, cf. [5]). Consider the probability measure on the set of all Young diagrams given by the formula

$$\operatorname{Prob}\{\lambda\} = e^{-\theta}\theta^{|\lambda|}\left(\frac{\dim\lambda}{|\lambda|!}\right)^2. \tag{2}$$

Here $\theta > 0$ is a parameter, $\dim\lambda$ is the number of standard Young tableaux of shape λ or the dimension of the irreducible representation of the symmetric group $S_{|\lambda|}$ corresponding to λ. Denote by $(p_1,\dots,p_d|q_1,\dots,q_d)$ the Frobenius coordinates of λ (see [17], §1 for the definition of Frobenius coordinates). Here d is the number of diagonal boxes in λ. Set $\mathbb{Z}' = \mathbb{Z} + \frac{1}{2} = \{\pm\frac{1}{2}, \pm\frac{3}{2}, \dots\}$.

Let us associate to any Young diagram $\lambda = (p\,|\,q)$ a point configuration $\operatorname{Fr}(\lambda) \subset \mathbb{Z}'$ as follows:

$$\operatorname{Fr}(\lambda) = \left\{p_1 + \tfrac{1}{2}, \dots, p_d + \tfrac{1}{2}, -q_1 - \tfrac{1}{2}, \dots, -q_d - \tfrac{1}{2}\right\}.$$

It turns out that together with (2) this defines a determinantal point process on $\mathbb{Z}'$. Indeed, the well-known hook formula for $\dim\lambda$ easily implies

$$\begin{aligned}\operatorname{Prob}\{\lambda\} &= e^{-\theta}\left(\det\left[\frac{\theta^{\frac{p_i+q_j}{2}}}{(p_i-\frac{1}{2})!(q_j-\frac{1}{2})!(p_i+q_j)}\right]_{i,j=1}^{d}\right)^2\\ &= e^{-\theta}\det[L(y_i,y_j)]_{i,j=1}^{2d}\end{aligned}$$

where $\{y_1,\dots,y_{2d}\} = \operatorname{Fr}(\lambda)$, and $L(x,y)$ is a $\mathbb{Z}'\times\mathbb{Z}'$ matrix defined by

$$L(x,y) = \begin{cases} 0, & \text{if } xy > 0,\\ \dfrac{\theta^{\frac{|x|+|y|}{2}}}{(|x|-\frac{1}{2})!(|y|-\frac{1}{2})!}\,\dfrac{1}{x-y}, & \text{if } xy < 0.\end{cases}$$

In the block form corresponding to the splitting $\mathbb{Z}' = \mathbb{Z}'_+ \sqcup \mathbb{Z}'_-$ it looks as follows

$$L(x,y) = \begin{bmatrix} 0 & \dfrac{\theta^{\frac{x-y}{2}}}{(x-\frac{1}{2})!(-y-\frac{1}{2})!}\,\dfrac{1}{x-y}\\ \dfrac{\theta^{\frac{-x+y}{2}}}{(-x-\frac{1}{2})!(y-\frac{1}{2})!}\,\dfrac{1}{x-y} & 0\end{bmatrix}.$$

The kernel $L(x,y)$ belongs to the class of integrable kernels. Indeed, if we set

$$f_1(x) = g_2(y) = \begin{cases}\dfrac{\theta^{\frac{x}{2}}}{(x-\frac{1}{2})!}, & x > 0,\\ 0, & x < 0,\end{cases} \qquad f_2(x) = g_1(y) = \begin{cases} 0, & x > 0,\\ \dfrac{\theta^{-\frac{x}{2}}}{(-x-\frac{1}{2})!}, & x < 0,\end{cases}$$

then it is immediately verified that $L(x,y) = (f_1(x)g_1(y) + f_2(x)g_2(y))/(x-y)$. Comparing the formulas with Example 1, we also conclude that $e^\theta = \det(1+L)$.[1]

[1] Since $\sum_{x,y\in\mathbb{Z}'}|L(x,y)| < \infty$, the operator L is trace class, and $\det(1+L)$ is well–defined.

What we see in this example is that L is an integrable kernel. We also know, see Example 1, that the correlation kernel K is given by $K = L(1+L)^{-1}$. Is this kernel also integrable? The answer is positive; the general claim in the continuous case was proved in [14], the discrete case was worked out in [2].

Furthermore, it turns out that in many situations there is an algorithm of computing the correlation kernel K if L is an integrable kernel which is "simple enough". The algorithm is based on a classical problem of complex analysis called the *Riemann–Hilbert problem* (RHP, for short).

Let us point out that our algorithm is not applicable to deriving correlation kernels in the "$\beta = 2$" model of Random Matrix Theory. Indeed, the Christoffel–Darboux kernels have norm 1, since they are just projection operators. Thus, it is impossible to define the kernel $L = K(1-K)^{-1}$, because $(1-K)$ is not invertible. In this sense, RMT deals with "degenerate" determinantal point processes.

On the other hand, the orthogonal polynomial method of computing the correlation kernels, which has been so successful in RMT, cannot be applied directly to the representation theoretic models like Example 2.2 above (see, however, [15]). The algorithm explained below may be viewed as a substitute for this method.

3 Riemann–Hilbert problem

Let Σ be an oriented contour in $\mathbb{C}$. We agree that $(+)$-side is on the left of the contour, and $(-)$-side is on the right of the contour. Let v be a 2×2-matrix valued function on Σ.

Definition 4. *We say that a matrix function $m : \mathbb{C}\setminus\Sigma \to \mathrm{Mat}(2,\mathbb{C})$ solves the RHP (Σ, v) if*

(1) m is analytic in $\mathbb{C}\setminus\Sigma$;

(2) $m_+ = m_- v$ on Σ, where $m_\pm(x) = \lim\limits_{\zeta\to x \text{ from } (\pm)\text{-side}} m(\zeta)$.

We say that m solves the normalized RHP (Σ, v) if, in addition, we have

(3) $m(\zeta) \to I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ as $\zeta\to\infty$.

Next we explain what is a *discrete* Riemann–Hilbert problem (DRHP, for short).

Let X be a locally finite subset of $\mathbb{C}$, and let w be a 2×2-matrix valued function on X.

Definition 5. *We say that a matrix function $m : \mathbb{C}\setminus X \to \mathrm{Mat}(2, C)$ solves the DRHP (X, w) if*

(1) m is analytic in $\mathbb{C}\setminus X$;

(2) m has simple poles at the points of X, and

$$\mathrm{Res}_{\zeta=x}\, m(\zeta) = \lim_{\zeta\to x}(m(\zeta)w(x)) \quad \text{for any } x\in X.$$

We say that m solves the normalized DRHP (X, w) if

(3) $m(\zeta) \to I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ *as* $\zeta \to \infty$.

If the set X is infinite, the last relation should hold when the distance from ζ to X is bounded away from zero.

Our next step is to explain how to reduce, for an integrable operator L, the computation of the operator $K = L(1+L)^{-1}$ to a (discrete or continuous) RHP.

3.1 Continuous picture [14]

Let L be an integrable operator on $L^2(\Sigma, |d\zeta|)$, $\Sigma \subset \mathbb{C}$, with the kernel ($x, y \in \Sigma$)

$$L(x,y) = \frac{f_1(x)g_1(y) + f_2(x)g_2(y)}{x - y}, \qquad f_1(x)g_1(x) + f_2(x)g_2(x) \equiv 0.$$

Assume that $(1 + L)$ is invertible.

Theorem 1. *There exists a unique solution of the normalized RHP (Σ, v) with*

$$v = I + 2\pi i \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \begin{bmatrix} g_1 & g_2 \end{bmatrix} = \begin{bmatrix} 1 + 2\pi i f_1 g_1 & 2\pi i f_1 g_2 \\ 2\pi i f_2 g_1 & 1 + 2\pi i f_2 g_2 \end{bmatrix}.$$

For $x \in \Sigma$ set

$$\begin{bmatrix} F_1(x) \\ F_2(x) \end{bmatrix} = \lim_{\zeta \to x} m(\zeta) \begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix},$$

$$\begin{bmatrix} G_1(x) \\ G_2(x) \end{bmatrix} = \lim_{\zeta \to x} m^{-t}(\zeta) \begin{bmatrix} g_1(x) \\ g_2(x) \end{bmatrix}.$$

Then the kernel of the operator $K = L(1+L)^{-1}$ has the form ($x, y \in \Sigma$)

$$K(x,y) = \frac{F_1(x)G_1(y) + F_2(x)G_2(y)}{x - y} \quad \textit{and} \quad F_1(x)G_1(x) + F_2(x)G_2(x) \equiv 0.$$

Example 4. Let Σ be a simple closed curve in $\mathbb{C}$ oriented clockwise (so that the $(+)$-side is outside Σ), and let L be an integrable operator such that the functions f_1, f_2, g_1, g_2 can be extended to analytic functions inside Σ. Then the solution of the normalized RHP (Σ, v) has the form

$$m = \begin{cases} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \text{outside } \Sigma, \\ I - 2\pi i \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \begin{bmatrix} g_1 & g_2 \end{bmatrix} & \text{inside } \Sigma. \end{cases}$$

Then we immediately obtain $F_i = f_i$, $G_i = g_i$, $i = 1, 2$; and $K = L(1+L)^{-1} = L$. On the other hand, this is obvious because $\int_\Sigma L(x,y)L(y,z)dy = 0$ by Cauchy's theorem which means that $L^2 = 0$.

3.2 Discrete picture [2]

Let L be an integrable operator on $\ell^2(X)$, $X \subset \mathbb{C}$, with the kernel

$$L(x,y) = \begin{cases} \dfrac{f_1(x)g_1(y) + f_2(x)g_2(y)}{x-y}, & x \neq y, \\ 0 & x = y, \end{cases}$$

with $f_1(x)g_1(x) + f_2(x)g_2(x) \equiv 0$. Assume that $(1 + L)$ is invertible.

Theorem 2. *There exists a unique solution of the normalized DRHP* (X, w) *with*

$$w = -\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \begin{bmatrix} g_1 & g_2 \end{bmatrix} = \begin{bmatrix} -f_1g_1 & -f_1g_2 \\ -f_2g_1 & -f_2g_2 \end{bmatrix}.$$

For $x \in X$ *set*

$$\begin{bmatrix} F_1(x) \\ F_2(x) \end{bmatrix} = \lim_{\zeta \to x} m(\zeta) \begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix},$$
$$\begin{bmatrix} G_1(x) \\ G_2(x) \end{bmatrix} = \lim_{\zeta \to x} m^{-t}(\zeta) \begin{bmatrix} g_1(x) \\ g_2(x) \end{bmatrix}.$$

Then the kernel of the operator $K = L(1+L)^{-1}$ *has the form* $(x, y \in \Sigma)$

$$K(x,y) = \begin{cases} \dfrac{F_1(x)G_1(y) + F_2(x)G_2(y)}{x-y}, & x \neq y, \\ \begin{bmatrix} G_1(x) & G_2(x) \end{bmatrix} \lim\limits_{\zeta \to x} \left(m'(\zeta) \begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix} \right), & x = y. \end{cases}$$

We also have $F_1(x)G_1(x) + F_2(x)G_2(x) \equiv 0$ *on* X.

Theorem 2 can be extended to the case when $L(x,x) \neq 0$, see [2], Remark 4.2.

Example 5. Let $X = \{a, b\}$ be a two-point subset of $\mathbb{C}$, and

$$L = \begin{bmatrix} 0 & \mu \\ \nu & 0 \end{bmatrix}.$$

Then L is integrable with

$$f_1 = \begin{cases} 0 \\ \nu(b-a) \end{cases}, \quad f_2 = \begin{cases} \mu(a-b) \\ 0 \end{cases}, \quad g_1 = \begin{cases} 1 \\ 0 \end{cases}, \quad g_2 = \begin{cases} 0 \\ 1 \end{cases}.$$

The notation means that, say, $f_1(a) = 0$, $f_1(b) = \nu(b-a)$. Then

$$w(a) = \begin{bmatrix} 0 & 0 \\ \mu(b-a) & 0 \end{bmatrix}, \quad w(b) = \begin{bmatrix} 0 & \nu(a-b) \\ 0 & 0 \end{bmatrix}.$$

Then the matrix $m(\zeta)$ has the form

$$m(\zeta) = I + \frac{1}{\zeta - a}\,\frac{\mu(a-b)}{1-\mu\nu}\begin{bmatrix} \nu & 0 \\ -1 & 0 \end{bmatrix} + \frac{1}{\zeta - b}\,\frac{\nu(b-a)}{1-\mu\nu}\begin{bmatrix} 0 & -1 \\ 0 & \mu \end{bmatrix}.$$

One can check that $\det m \equiv 1$, and

$$m^{-t}(\zeta) = I + \frac{1}{\zeta - a}\,\frac{\mu(a-b)}{1-\mu\nu}\begin{bmatrix} 0 & 1 \\ 0 & \nu \end{bmatrix} + \frac{1}{\zeta - b}\,\frac{\nu(b-a)}{1-\mu\nu}\begin{bmatrix} \mu & 0 \\ 1 & 0 \end{bmatrix}.$$

Further,

$$F_1 = \begin{cases} \frac{\mu\nu(b-a)}{1-\mu\nu} \\ \frac{\nu(b-a)}{1-\mu\nu} \end{cases}, \quad F_2 = \begin{cases} \frac{\mu(a-b)}{1-\mu\nu} \\ \frac{\mu\nu(a-b)}{1-\mu\nu} \end{cases}, \quad G_1 = \begin{cases} \frac{1}{1-\mu\nu} \\ \frac{\mu}{1-\mu\nu} \end{cases}, \quad G_2 = \begin{cases} \frac{\nu}{1-\mu\nu} \\ \frac{1}{1-\mu\nu} \end{cases},$$

$$\lim_{\zeta\to a}\left(m'(\zeta)\begin{bmatrix} f_1(a) \\ f_2(a) \end{bmatrix}\right) = \begin{bmatrix} \frac{-\mu\nu}{1-\mu\nu} \\ \frac{\mu^2\nu}{1-\mu\nu} \end{bmatrix}, \quad \lim_{\zeta\to b}\left(m'(\zeta)\begin{bmatrix} f_1(b) \\ g_1(b) \end{bmatrix}\right) = \begin{bmatrix} \frac{\mu\nu^2}{1-\mu\nu} \\ \frac{-\mu\nu}{1-\mu\nu} \end{bmatrix}.$$

By Theorem 2, this implies that

$$K = \frac{L}{1+L} = \frac{1}{1-\mu\nu}\begin{bmatrix} -\mu\nu & \mu \\ \nu & -\mu\nu \end{bmatrix}$$

which is immediately verified directly. Note that the condition $1 - \mu\nu \neq 0$ is equivalent to the invertibility of $(1 + L)$.

In what follows we will demonstrate how to use Theorems 1 and 2 to compute correlation kernels of determinantal point processes arising in concrete representation theoretic models.

4 Harmonic analysis on $S(\infty)$: Whittaker kernel

As is explained in [7], see also [21], the problem of decomposing generalized regular representations of the infinite symmetric group $S(\infty)$ on irreducible ones reduces to computing correlation kernels of certain determinantal point processes.

Specifically, consider a determinantal point process on $\mathbb{Z}' = \mathbb{Z} + \frac{1}{2}$ constructed using Example 1 with the L-kernel given by

$$L(x,y) = \begin{cases} 0, & xy > 0, \\ \dfrac{|z(z+1)_{x-\frac{1}{2}}(-z+1)_{-y-\frac{1}{2}}|\,\xi^{\frac{x-y}{2}}}{(x-\frac{1}{2})!(-y-\frac{1}{2})!(x-y)}, & x > 0,\ y < 0, \\ \dfrac{|z(-z+1)_{-x-\frac{1}{2}}(z+1)_{y-\frac{1}{2}}|\,\xi^{\frac{-x+y}{2}}}{(-x-\frac{1}{2})!(y-\frac{1}{2})!(x-y)}, & x < 0,\ y > 0. \end{cases}$$

Here $z \in \mathbb{C} \setminus \mathbb{Z}$ and $\xi \in (0,1)$ are parameters. The symbol $(a)_k$ stands for $a(a+1)\cdots(a+k-1) = \Gamma(a+k)/\Gamma(a)$.

Note that as $|z| \to \infty$, $\xi \to 0$, and $|z|^2\xi \to \theta$, this kernel converges to the L-kernel of Example 3.

The problem consists in computing $K = L(1+L)^{-1}$ and taking the scaling limit

$$\mathcal{K}(x,y) = \lim_{\xi\to 1}(1-\xi)^{-1} \cdot K\left([(1-\xi)^{-1}x] + \tfrac{1}{2},\, [(1-\xi)^{-1}y] + \tfrac{1}{2}\right).$$

This problem has been solved in [7]. However, we did not provide a *derivation* of the formula for the kernel K there, we just *verified* the equality $K = L(1+L)^{-1}$.

The goal of this section is to provide a derivation of the kernel $\mathcal{K}(x,y)$ bypassing the computation of $K(x,y)$.

Observe that there exists a limit

$$\mathcal{L}(x,y) = \lim_{\xi\to 1}(1-\xi)^{-1} \cdot L\left([(1-\xi)^{-1}x] + \tfrac{1}{2},\, [(1-\xi)^{-1}y] + \tfrac{1}{2}\right).$$

Indeed, for $a = [(1-\xi)^{-1}x]$, $b = [(1-\xi)^{-1}y]$,

$$\frac{(z+1)_a}{\Gamma(a+1)} = \frac{\Gamma(z+1+a)}{\Gamma(z+1)\Gamma(a+1)} \sim \frac{a^z}{\Gamma(z+1)}, \quad x>0,$$
$$\frac{(-z+1)_b}{\Gamma(b+1)} = \frac{\Gamma(-z+1+|b|)}{\Gamma(-z+1)\Gamma(|b|+1)} \sim \frac{|b|^{-z}}{\Gamma(-z+1)}, \quad y<0,$$
$$\xi^{\frac{a}{2}} \sim (1-(1-\xi))^{\frac{x}{2(1-\xi)}} \sim e^{\frac{x}{2}}, \quad \xi^{-\frac{b}{2}} \sim e^{-\frac{y}{2}},$$

and we get

$$\mathcal{L}(x,y) = \begin{cases} 0, & xy>0,\\ \dfrac{|\sin\pi z|}{\pi}\,\dfrac{(x/|y|)^{\Re z} e^{\frac{-x+y}{2}}}{x-y}, & x>0,\ y<0,\\ \dfrac{|\sin\pi z|}{\pi}\,\dfrac{(y/|x|)^{\Re z} e^{\frac{x-y}{2}}}{x-y}, & x<0,\ y>0. \end{cases}$$

It is natural to assume that $\mathcal{K} = \mathcal{L}(1+\mathcal{L})^{-1}$. It turns out that this relation holds whenever $\mathcal{L}$ defines a bounded operator in $L^2(\mathbb{R}\setminus\{0\})$, which happens when $|\Re z| < \frac{1}{2}$, see [20]. Our goal is to derive $\mathcal{K}$ using the relation $\mathcal{L} = \mathcal{L}(1+\mathcal{L})^{-1}$.

It is easily seen that $\mathcal{L}$ is an integrable operator; we can take

$$f_1(x) = g_2(x) = \begin{cases} \dfrac{|z|^{\frac{1}{2}}}{|\Gamma(z+1)|}\, x^{\Re z} e^{-\frac{x}{2}}, & x>0,\\ 0, & x<0, \end{cases}$$

$$f_2(x) = g_1(x) = \begin{cases} 0, & x>0,\\ \dfrac{|z|^{\frac{1}{2}}}{|\Gamma(-z+1)|}\, |x|^{-\Re z} e^{\frac{x}{2}}, & x<0. \end{cases}$$

Note that $\mathcal{L}(y,x) = -\mathcal{L}(x,y)$ which means that $(1+\mathcal{L})$ is invertible (provided that $\mathcal{L}$ is bounded).

The RHP of Theorem 1 then has the jump matrix

$$v(x) = \begin{cases} \begin{bmatrix} 1 & 2i\,|\sin \pi z\,|\,x^{2\Re z}e^{-x} \\ 0 & 1 \end{bmatrix}, & x > 0 \\ \begin{bmatrix} 1 & 0 \\ 2i\,|\sin \pi z\,|\,|x|^{-2\Re z}e^{x} & 1 \end{bmatrix}, & x < 0. \end{cases}$$

The key property of this RHP which will allow us to solve it, is that it can be reduced to a problem with a piecewise constant jump matrix.

Let m be the solution of the normalized RHP $(\mathbb{R} \setminus \{0\}, v)$. Set

$$\Psi(\zeta) = m(\zeta) \begin{bmatrix} \zeta^{\Re z}e^{-\frac{\zeta}{2}} & 0 \\ 0 & \zeta^{-\Re z}e^{\frac{\zeta}{2}} \end{bmatrix}, \qquad \zeta \notin \mathbb{R}.$$

Then the jump relation $m_+ = m_- v$ takes the form

$$\Psi_+(x) = \Psi_-(x) \begin{bmatrix} x^{-\Re z}e^{\frac{x}{2}} & 0 \\ 0 & x^{\Re z}e^{-\frac{x}{2}} \end{bmatrix}_- v(x) \begin{bmatrix} x^{\Re z}e^{-\frac{x}{2}} & 0 \\ 0 & x^{-\Re z}e^{\frac{x}{2}} \end{bmatrix}_+,$$

and a direct computation shows that the jump matrix for Ψ takes the form

$$\begin{cases} \begin{bmatrix} 1 & 2i\,|\sin \pi z\,| \\ 0 & 1 \end{bmatrix}, & x > 0 \\ \begin{bmatrix} e^{2\pi i\,\Re z} & 0 \\ 2i\,|\sin \pi z\,| & e^{-2\pi i\,\Re z} \end{bmatrix}, & x < 0. \end{cases}$$

Let us first find a solution of this RHP without imposing any asymptotic conditions at infinity. We will denote it by Ψ^0. Set

$$\widetilde{\Psi}^0\zeta = \begin{cases} \Psi^0(\zeta), & \Im\zeta > 0, \\ \Psi^0(\zeta) \begin{bmatrix} 1 & 2i\,|\sin \pi z| \\ 0 & 1 \end{bmatrix}, & \Im\zeta < 0. \end{cases}$$

Then $\widetilde{\Psi}^0$ has no jump across $\mathbb{R}_+$, and the jump matrix $(\widetilde{\Psi}^0_-)^{-1}\widetilde{\Psi}^0_+$ on $\mathbb{R}_-$ has the form

$$\begin{bmatrix} 1 & -2i\,|\sin \pi z| \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e^{2\pi i\,\Re z} & 0 \\ 2i\,|\sin \pi z\,| & e^{-2\pi i\,\Re z} \end{bmatrix}$$
$$= \begin{bmatrix} e^{2\pi i\,\Re z} + 4|\sin \pi z\,|^2 & -2i\,|\sin \pi z|e^{-2\pi i\,\Re z} \\ 2i\,|\sin \pi z\,| & e^{-2\pi i\,\Re z} \end{bmatrix}.$$

The determinant of this matrix is equal to 1, and the trace is equal to $2\cos(2\pi\Re z) + 4|\sin \pi z\,|^2 = e^{\pi\Im z} + e^{-\pi\Im z}$. Thus, if $z \notin \mathbb{R}$, there exists a nondegenerate U such that

$$\begin{bmatrix} e^{2\pi i\,\Re z}+4|\sin\pi z\,|^2 & -2i\,|\sin\pi z|e^{-2\pi i\,\Re z} \\ 2i\,|\sin\pi z\,| & e^{-2\pi i\,\Re z}\end{bmatrix} = U^{-1}\begin{bmatrix} e^{\pi\Im z} & 0 \\ 0 & e^{-\pi\Im z}\end{bmatrix}U.$$

This means that the matrix $\widetilde{\Psi}^0 U^{-1}$ has jump $\begin{bmatrix} e^{\pi\Im z} & 0 \\ 0 & e^{-\pi\Im z}\end{bmatrix}$ across $\mathbb{R}_-$. Note that the matrix $\begin{bmatrix} \zeta^{-i\Im z} & 0 \\ 0 & \zeta^{i\Im z}\end{bmatrix}$ satisfies the same jump relation. Hence,

$$\Psi^0(\zeta) = \begin{cases} \begin{bmatrix} \zeta^{-i\Im z} & 0 \\ 0 & \zeta^{i\Im z}\end{bmatrix} U, & \Im\zeta > 0, \\ \begin{bmatrix} \zeta^{-i\Im z} & 0 \\ 0 & \zeta^{i\Im z}\end{bmatrix}\begin{bmatrix} 1 & -2i\,|\sin\pi z\,| \\ 0 & 1\end{bmatrix} U, & \Im\zeta < 0, \end{cases}$$

is a solution of our RHP for Ψ. It follows that $\Psi(\Psi^0)^{-1}$ has no jump across $\mathbb{R}$ and this implies, modulo some technicalities, that $\Psi = H\Psi^0$ where H is entire.

Now we describe the crucial step. Since the jump matrix for Ψ is piecewise constant, $\Psi' = \frac{d\Psi}{d\zeta}$ satisfies the same jump condition as Ψ, and hence $\Psi'\Psi^{-1}$ is meromorphic in $\mathbb{C}$ with a possible pole at $\zeta = 0$. On the other hand we have

$$\begin{aligned} \Psi'\Psi^{-1} &= H'H^{-1} + H(\Psi^0)'(\Psi^0)^{-1}H^{-1} \\ &= H'H^{-1} - \frac{1}{\zeta}\, i\,\Im z\, H\begin{bmatrix} 1 & 0 \\ 0 & -1\end{bmatrix} H^{-1} = \text{entire function} + \frac{A}{\zeta}, \end{aligned} \tag{3}$$

where A has eigenvalues $\pm i\,\Im z$.

Let us recall now that m solves the *normalized* RHP, which means that $m(\zeta) \sim I$ as $|z| \to \infty$. An additional argument shows that

$$m(\zeta) = I + m^{(1)}\zeta^{-1} + O\left(|\zeta|^{-2}\right), \qquad |\zeta| \to \infty,$$

with a constant matrix $m^{(1)}$. Thus,

$$\Psi(\zeta) = \left(I + m^{(1)}\zeta^{-1} + O\left(|\zeta|^{-2}\right)\right)\begin{bmatrix} \zeta^{\Re z}e^{-\frac{\zeta}{2}} & 0 \\ 0 & \zeta^{-\Re z}e^{\frac{\zeta}{2}}\end{bmatrix}$$

and

$$\Psi'(\zeta)\Psi^{-1}(\zeta) = -\frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1\end{bmatrix} + \frac{1}{\zeta}\begin{bmatrix} \Re z & -m^{(1)}_{12} \\ m^{(1)}_{21} & -\Re z\end{bmatrix} + O(|\zeta|^{-2}).$$

Comparing this relation with (3) we conclude that

$$\Psi'(\zeta) = \left(-\frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1\end{bmatrix} + \frac{1}{\zeta}\begin{bmatrix} \Re z & -m^{(1)}_{12} \\ m^{(1)}_{21} & -\Re z\end{bmatrix}\right)\Psi(\zeta)$$

with $m^{(1)}_{12}m^{(1)}_{21} = (\Re z)^2 + (\Im\zeta)^2 = |z|^2$. This 1st order linear matrix differential equation leads to 2nd order linear differential equations on the matrix elements on Ψ, for example

$$\zeta \Psi_{11}'' + \Psi_{11}' = \left(-\frac{1}{2} - \frac{|z|^2}{\zeta} + \frac{1}{\zeta} \left(\Re z - \frac{\zeta}{2} \right)^2 \right) \Psi_{11}.$$

Using these differential equations and the asymptotics of Ψ at infinity, it is easy to express Ψ in terms of the confluent hypergeometric function or the Whittaker function, see [13], 6.9 for definitions. In terms of the Whittaker function $W_{\kappa,\mu}$ the final formula for Ψ has the form

$$\Psi(\zeta) = \begin{bmatrix} \zeta^{-\frac{1}{2}} W_{\Re z+\frac{1}{2}, i\Im z}(\zeta) & |z|\,(-\zeta)^{-\frac{1}{2}} W_{-\Re z-\frac{1}{2}, i\Im z}(-\zeta) \\ -|z|\,\zeta^{-\frac{1}{2}} W_{\Re z-\frac{1}{2}, i\Im z}(\zeta) & (-\zeta)^{-\frac{1}{2}} W_{-\Re z+\frac{1}{2}, i\Im z}(-\zeta) \end{bmatrix}.$$

It is not hard to show that $\det \Psi \equiv 1$, and

$$\Psi^{-t}(\zeta) = \begin{bmatrix} (-\zeta)^{-\frac{1}{2}} W_{-\Re z+\frac{1}{2}, i\Im z}(-\zeta) & |z|\,\zeta^{-\frac{1}{2}} W_{\Re z-\frac{1}{2}, i\Im z}(\zeta) \\ -|z|\,(-\zeta)^{-\frac{1}{2}} W_{-\Re z-\frac{1}{2}, i\Im z}(-\zeta) & \zeta^{-\frac{1}{2}} W_{\Re z+\frac{1}{2}, i\Im z}(\zeta) \end{bmatrix}.$$

Then Theorem 1 implies

$$F_1(x) = \begin{cases} \frac{|z|^{\frac{1}{2}}}{|\Gamma(z+1)|} \Psi_{11}(x), & x > 0, \\ \frac{|z|^{\frac{1}{2}}}{|\Gamma(-z+1)|} \Psi_{12}(x), & x < 0, \end{cases} \qquad F_2(x) = \begin{cases} \frac{|z|^{\frac{1}{2}}}{|\Gamma(z+1)|} \Psi_{21}(x), & x > 0, \\ \frac{|z|^{\frac{1}{2}}}{|\Gamma(-z+1)|} \Psi_{22}(x), & x < 0, \end{cases}$$

$$G_1(x) = \begin{cases} -\frac{|z|^{\frac{1}{2}}}{|\Gamma(z+1)|} \Psi_{21}(x), & x > 0, \\ \frac{|z|^{\frac{1}{2}}}{|\Gamma(-z+1)|} \Psi_{22}(x), & x < 0, \end{cases} \qquad G_2(x) = \begin{cases} \frac{|z|^{\frac{1}{2}}}{|\Gamma(z+1)|} \Psi_{11}(x), & x > 0, \\ -\frac{|z|^{\frac{1}{2}}}{|\Gamma(-z+1)|} \Psi_{12}(x), & x < 0, \end{cases}$$

and

$$\mathcal{K}(x, y) = \frac{F_1(x)G_1(y) + F_2(x)G_2(y)}{x - y}, \quad x, y \in \mathbb{R} \setminus \{0\}.$$

This kernel is called the *Whittaker kernel*, see [6], [1].

5 Poissonized Plancherel measure: discrete Bessel kernel

We now return to the situation described in Example 3. Our goal is to compute the correlation kernel $K = L(1 + L)^{-1}$. The exposition below follows [2], §7.

According to Theorem 2, we have to find the unique solution of the normalized DRHP $(\mathbb{Z}', w)$ with

$$w(x) = \begin{cases} \begin{bmatrix} 0 & -\frac{\theta^x}{(x-\frac{1}{2})!^2} \\ 0 & 0 \end{bmatrix}, & x \in \mathbb{Z}'_+, \\ \begin{bmatrix} 0 & 0 \\ -\frac{\theta^{-x}}{(-x-\frac{1}{2})!^2} & 0 \end{bmatrix}, & x \in \mathbb{Z}'_-. \end{cases}$$

Note that the kernel L is skew-symmetric, which means that $(1 + L)$ is invertible. If we denote by m the solution of this DRHP then

$$m(\zeta) = I + \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \zeta^{-1} + O(|\zeta^{-2}|), \qquad |\zeta| \to \infty,$$

with constant α, β, γ, δ. The symmetry of the problem with respect to

$$\zeta \leftrightarrow -\zeta, \quad \begin{bmatrix} m_{11}(\zeta) & m_{12}(\zeta) \\ m_{21}(\zeta) & m_{22}(\zeta) \end{bmatrix} \longleftrightarrow \begin{bmatrix} m_{22}(-\zeta) & -m_{21}(-\zeta) \\ -m_{12}(-\zeta) & m_{11}(-\zeta) \end{bmatrix},$$

implies that $\gamma = \beta$ and $\delta = -\alpha$.

Denote $\eta = \sqrt{\theta}$ and set

$$n(\zeta) = m(\zeta) \begin{bmatrix} \eta^{\zeta} & 0 \\ 0 & \eta^{-\zeta} \end{bmatrix}.$$

Then $n(\zeta)$ solves a DRHP with the jump matrix

$$\begin{cases} \begin{bmatrix} 0 & -\frac{1}{(x-\frac{1}{2})!^2} \\ 0 & 0 \end{bmatrix}, & x \in \mathbb{Z}'_+, \\ \begin{bmatrix} 0 & 0 \\ -\frac{1}{(-x-\frac{1}{2})!^2} & 0 \end{bmatrix}, & x \in \mathbb{Z}'_-. \end{cases}$$

Note that this jump matrix does not depend on η. This means that $\frac{\partial n}{\partial \eta}$ has the same jump matrix, and hence the matrix $\frac{\partial n}{\partial \eta}\, n^{-1}$ is entire. Computing the asymptotics as $\zeta \to \infty$, we obtain

$$\frac{\partial n(\zeta)}{\partial \eta}\, n^{-1}(\zeta) = \begin{bmatrix} \zeta & -2\beta \\ 2\beta & -\zeta \end{bmatrix} + O(|z|^{-1}).$$

By Liouville's theorem, the remainder term must vanish, and thus

$$\frac{\partial n(\zeta)}{\partial \eta} = \begin{bmatrix} \zeta & -2\beta \\ 2\beta & -\zeta \end{bmatrix} n(\zeta). \tag{4}$$

This yields 2nd order linear differential equations on the matrix elements of n which involve, however, an unknown function $\beta = \beta(\eta)$.

In order to determine β we need to make one more step. Set

$$p(\zeta) = n(\zeta) \begin{bmatrix} \frac{1}{\Gamma(\zeta+\frac{1}{2})} & 0 \\ 0 & \frac{1}{\Gamma(-\zeta+\frac{1}{2})} \end{bmatrix}.$$

It is immediately verified that the fact that n solves the corresponding DRHP is equivalent to p being entire and satisfying the condition

$$p(x) = (-1)^{x-\frac{1}{2}} p(x) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad x \in \mathbb{Z}'. \tag{5}$$

The key property of this relation is that it depends on x in an insubstantial way. This allows us to do the following trick which should be viewed as a substitute of the differentiation with respect to x. Set

$$\widetilde{p}(\zeta) = \begin{bmatrix} p_{11}(\zeta+1) & -p_{12}(\zeta+1) \\ -p_{21}(\zeta-1) & p_{22}(\zeta-1) \end{bmatrix}.$$

Then $\widetilde{p}$ satisfies the same condition (5) as p does, and hence

$$\widetilde{n}(\zeta) = \widetilde{p}(\zeta) \begin{bmatrix} \Gamma(\zeta+\frac{1}{2}) & 0 \\ 0 & \Gamma(-\zeta+\frac{1}{2}) \end{bmatrix}$$

satisfies the same DRHP as n does. Thus, $\widetilde{n}\, n^{-1}$ is entire. Working out the asymptotics as $\zeta \to \infty$, we obtain

$$\widetilde{n}(\zeta) n^{-1}(\zeta) = \begin{bmatrix} 0 & \frac{\beta}{\eta} \\ -\frac{\beta}{\eta} & 0 \end{bmatrix} + O(|z|^{-1}).$$

Liouville's theorem implies that the remainder vanishes. Hence,

$$\widetilde{p}_{11}(\zeta) = p_{11}(\zeta+1) = \frac{\beta}{\eta}\, p_{21}(\zeta), \quad \widetilde{p}_{21}(\zeta) = -p_{21}(\zeta-1) = -\frac{\beta}{\eta}\, p_{11}(\zeta).$$

This implies that $(\beta/\eta)^2 = 1$. Both cases $\beta = \pm\eta$ lead, via (4), to linear 2nd order differential equations on the matrix elements of n or matrix elements of p. For example, $\beta = -\eta$ yields

$$\left(\frac{\partial^2}{\partial\eta^2} - \frac{\zeta(\zeta-1)}{\eta^2} + 4\right) p_{11}(\zeta) = 0, \quad \left(\frac{\partial^2}{\partial\eta^2} - \frac{\zeta(\zeta+1)}{\eta^2} + 4\right) p_{21}(\zeta) = 0.$$

General solutions of these equations can be written in terms of Bessel functions, and matching the asymptotics at infinity we obtain for $\beta = -\eta$

$$p(\zeta) = \sqrt{\eta} \begin{bmatrix} J_{\zeta-\frac{1}{2}}(2\eta) & J_{-\zeta+\frac{1}{2}}(2\eta) \\ -J_{\zeta+\frac{1}{2}}(2\eta) & J_{-\zeta-\frac{1}{2}}(2\eta) \end{bmatrix}, \tag{6}$$

and for $\beta = \eta$

$$\widehat{p}(\zeta) = \sqrt{\eta} \begin{bmatrix} J_{\zeta-\frac{1}{2}}(2\eta) & -J_{-\zeta+\frac{1}{2}}(2\eta) \\ J_{\zeta+\frac{1}{2}}(2\eta) & J_{-\zeta-\frac{1}{2}}(2\eta) \end{bmatrix}.$$

Here $J_\nu(u)$ is the Bessel function, see [13], 7.2 for the definition.

Using the well–known relation $J_{-n} = (-1)^n J_n$ we immediately see that $p(\zeta)$ given by (6) satisfies (5), while $\widehat{p}(\zeta)$ does not. In fact $\widehat{p}(\zeta)$ satisfies (5) with $(-1)^{x-\frac{1}{2}}$ replaced with $(-1)^{x+\frac{1}{2}}$. This means that

$$m(\zeta) = p(\zeta) \begin{bmatrix} \eta^{-\zeta}\Gamma(\zeta+\frac{1}{2}) & 0 \\ 0 & \eta^{\zeta}\Gamma(-\zeta+\frac{1}{2}) \end{bmatrix}$$

solves the initial DRHP, and by Theorem 2 we obtain

$$F_1(x) = \begin{cases} p_{11}(x), & x > 0, \\ p_{12}(x), & x < 0, \end{cases} \quad F_2(x) = \begin{cases} p_{21}(x), & x > 0, \\ p_{22}(x), & x < 0, \end{cases}$$

$$G_1(x) = \begin{cases} -p_{21}(x), & x > 0, \\ p_{22}(x), & x < 0, \end{cases} \quad G_2(x) = \begin{cases} p_{11}(x), & x > 0, \\ -p_{12}(x), & x < 0, \end{cases}$$

and

$$K(x,y) = \frac{F_1(x)G_1(y) + F_2(x)G_2(y)}{x-y}, \quad x, y \in \mathbb{Z}'.$$

The diagonal values $K(x,x)$ are determined by the L'Hospital rule:

$$K(x,x) = F_1'(x)G_1(x) + F_2'(x)G_2(x).$$

This is the discrete Bessel kernel obtained in [5]. The restriction of $K(x,y)$ to $\mathbb{Z}'_+ \times \mathbb{Z}'_+$ was independently derived in [15].

It is worth noting that the matrix $\widehat{p}$ also has an important meaning. In fact, if we define a kernel $\widehat{K}$ using the formulas above with p replaced by $\widehat{p}$ then $\widehat{K} = L(L-1)^{-1}$.

References

1. Borodin, A.: Harmonic analysis on the infinite symmetric group and the Whittaker kernel. St. Petersburg Math. J., **12**, no. 5, (2001)
2. Borodin, A.: Riemann–Hilbert problem and the discrete Bessel kernel. Intern. Math. Research Notices, no. 9, 467–494 (2000), `math/9912093`
3. Borodin, A.: Discrete gap probabilities and discrete Painlevé equations. Preprint (2001)
4. Borodin, A., Deift, P.: Fredholm determinants, Jimbo–Miwa–Ueno tau-functions, and representation theory. Preprint (2001)
5. Borodin, A., Okounkov, A., Olshanski, G.: Asymptotics of Plancherel measures for symmetric groups. J. Amer. Math. Soc., **13**, 491–515 (2000); `math/9905032`
6. Borodin, A., Olshanski, G.: Point processes and the infinite symmetric group. Math. Research Lett., **5**, 799–816 (1998); `math/9810015`
7. Borodin, A., Olshanski, G.: Distributions on partitions, point processes and the hypergeometric kernel. Comm. Math. Phys., **211**, no. 2, 335–358 (2000); `math/9904010`
8. Borodin, A., Olshanski, G.: Z–Measures on partitions, Robinson–Schensted–Knuth correspondence, and $\beta = 2$ random matrix ensembles. Mathematical Sciences Research Institute Publications, **40**, 71–94 (2001); `math/9905189`
9. Borodin, A., Olshanski, G.: Harmonic analysis on the infinite–dimensional unitary group. Preprint (2001); `math/0109194`
10. Daley, D. J., Vere-Jones, D.: An introduction to the theory of point processes. Springer series in statistics, Springer, (1988)

11. Deift, P.: Integrable operators. In: Buslaev, V., Solomyak, M., Yafaev, D. (eds) Differential operators and spectral theory: M. Sh. Birman's 70th anniversary collection. American Mathematical Society Translations, ser. 2, **189**, Providence, R.I., AMS (1999)
12. Dyson, F. J.: Statistical theory of the energy levels of complex systems I, II, III. J. Math. Phys., **3**, 140–156, 157–165, 166–175 (1962)
13. Erdelyi, A. (ed): Higher transcendental functions, Vols. 1, 2. Mc Graw–Hill, (1953)
14. Its, A. R., Izergin, A. G., Korepin, V. E., Slavnov, N. A.: Differential equations for quantum correlation functions. Intern. J. Mod. Phys., **B4**, 1003–1037 (1990)
15. Johansson, K.: Discrete orthogonal polynomial ensembles and the Plancherel measure. Ann. Math. (2), **153**, no. 1, 259–296 (2001); `math/9906120`
16. Kerov, S., Olshanski, G., Vershik, A.: Harmonic analysis on the infinite symmetric group. A deformation of the regular representation. Comptes Rend. Acad. Sci. Paris, Sér. I, **316**, 773–778 (1993); detailed version in preparation
17. Macdonald, I. G.: Symmetric functions and Hall polynomials, 2nd edition. Oxford University Press (1995)
18. Mehta, M. L.: Random matrices, 2nd edition. Academic Press, New York (1991)
19. Nagao, T., Wadati, M.: Correlation functions of random matrix ensembles related to classical orthogonal polynomials. J. Phys. Soc. Japan, **60**, no. 10, 3298–3322 (1991)
20. Olshanski, G.: Point processes and the infinite symmetric group. Part V: Analysis of the matrix Whittaker kernel. Preprint (1998); `math/9810014`
21. Olshanski, G. An introduction to harmonic analysis on the infinite symmetric group. In this volume

Four Lectures on Random Matrix Theory

Percy Deift

Courant Institute
USA
deift@cims.nyu.edu

It is a great pleasure and honor for me to give these lectures in St. Petersburg. I am also very pleased to be here with so many of my friends. My topic is random matrix theory with an emphasis on the relationship to integrable systems.

Note on format

In order to maintain the informal style of the lectures I have not included detailed references. I have simply stated the author(s) of a particular paper or piece of work together with a date. It is then an easy matter using MathSciNet, for example, to trace down the detailed reference. I hope that this procedure is amenable to the reader and captures some of the excitement of the conference in St. Petersburg.

Outline

In this the **first** of four lectures, I am going to begin with some general remarks about integrable systems. Then I will make some general historical remarks about random matrix theory (RMT) and end up with an explanation and formulation of the so-called universality conjecture in RMT. In the **second lecture**, I will introduce a variety of objects and techniques from the modern theory of integrable systems which are needed for the universality conjecture. In the **third** lecture, I will show how to use these techniques to solve the universality conjecture, following the approach of Kriecherbauer, K. McLaughlin, Venakides, X. Zhou and Deift (1997)(1999). In the **fourth and final** lecture, I will show how these techniques can also be used to analyze the so-called Ulam problem in the theory of random permutations.

Acknowledgments

I would like to thank the organizers and particularly Professor Vershik for inviting me to St. Petersburg to give these lectures. The work was supported

A. Borodin et al.: LNM 1817, A.M. Vershik (Ed.), pp. 21–52, 2003.

in part by the NSF Grant No. DMS 0003268 and Grant No. DMS 0296084. Also, I am grateful to the NSF for a travel grant to attend the meeting in St. Petersburg.

Lecture 1

Integrable Systems

The modern theory of integrable systems began in 1967 with the discovery by Gardner, Greene, Kruskal and Miura of a method for integrating the Korteweg de Vries (KdV) equation

$$\begin{aligned} &u_t + u\,u_x + u_{xxx} = 0, \qquad x \in \mathbb{R}, \quad t \geq 0 \\ &u(x, t=0) \to 0 \quad \text{as} \quad |x| \to \infty. \end{aligned}$$

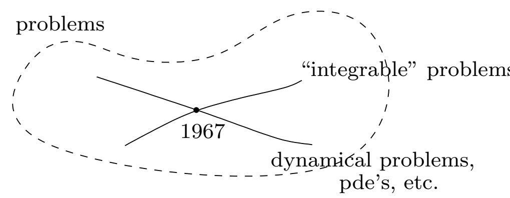
As is well-known, this equation has many applications, and arises, in particular, in the theory of water waves.

Initially the discovery of Gardner et al. was regarded as providing a method of solution for a rather thin set of evolutionary equations, but by the early 1980's it started to become clear that this discovery was just the first glimpse of a far more general integrable method that would eventually have applications across the broad spectrum of problems in pure and applied mathematics. In the narrowest sense, an integrable system is a Hamiltonian dynamical system — finite or infinite dimensional — with "enough" integrals of the motion, all of whose Poisson brackets are zero, to solve the system in some "explicit" form. Now it has been the serendipitous and rather extraordinary experience in the field over the last 30 years, that many systems which are of great mathematical and physical interest, which may not be Hamiltonian, and may not even be dynamical, can be solved "explicitly" using techniques that have some direct link back to the method of solution for the KdV equation discovered by Gardner et al. The kind of developments that I have in mind are for example:

- The resolution of the classical Schottky problem in algebraic geometry in terms of the solution of the Kadomtsev–Petviashvili (KP) equation, by Novikov, Krichever and others
- The introduction of quantum groups
- Integrable statistical models — the connection to Jones polynomials
- Two dimensional quantum gravity and the work of Witten and Kontsevich on the KdV hierarchy
- Nonlinear special function theory — Painlevé theory
- Conformal field theory — the work of Krichever, Dubrovin and others
- Random matrix theory
- Combinatorial problems of Robinson–Schensted–Knuth type.

The thrust of these lectures is to illustrate how the integrable method interacts with two of these topics, viz. random matrix theory and combinatorial problems of Robinson–Schensted–Knuth type.

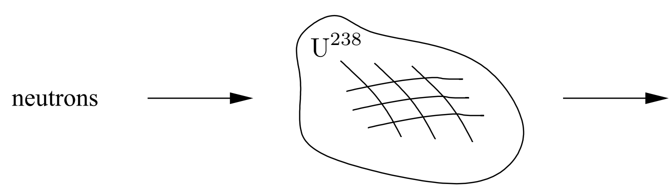
Allow me a very schematic moment! Consider the following picture of the "space" of problems containing in particular the "manifold" of dynamical/pde problems and the "manifold" of "integrable" problems:



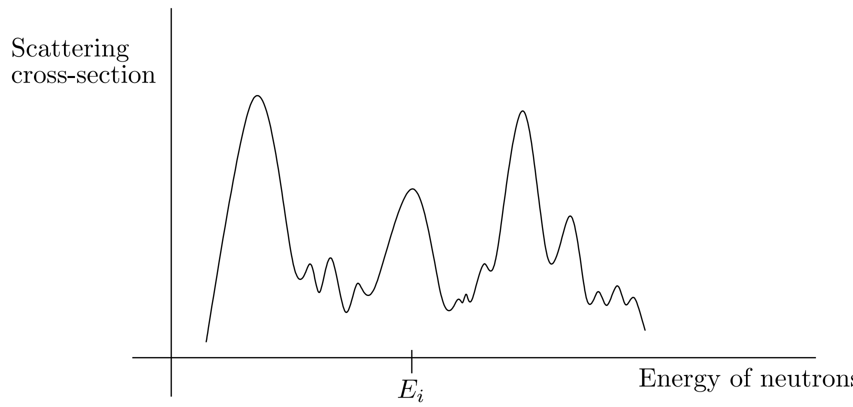
In 1967 these two "manifolds" had a "transverse intersection" at KdV! The initial thought was that to develop the ideas of Gardner et al. one should move along the "pde" manifold. But this turned out to be too limiting: we now know that in order to see the full development of the method one should move in the "transverse" direction. And this is the direction along which we will move in the next three lectures.

Random Matrix Theory

What is random matrix theory? RMT was first studied in the 1930's in the context of mathematical statistics by Hsu, Wihsart and others, but it was Wigner in the early 1950's who introduced RMT to mathematical physics. Wigner was concerned with the scattering of neutrons off large nuclei.



Schematically one sees the following:



The energies E_i for which one obtains large scattering cross sections are called resonances, and for a heavy nucleus like U^{238}, say, there can be hundreds or thousands of them. In theory one could write down some Schrödinger equation for the neutrons + nucleus and try to solve it numerically to compute these E_i. But in Wigner's time (and still in our time, and also into the foreseeable future) this was not a realistic approach, and so people began to think that it was more appropriate to give the resonances a statistical meaning. But what should the statistical model be? At this point Wigner put forward the remarkable hypothesis that the (high) resonances E_i behave like the eigenvalues of a (large) matrix. It is hard to overemphasize what a radical and revolutionary thought this was. We all recall that when we first were learning some physics, we understood that the details of the model were paramount: if you changed the force law in the equations of motion, the behavior of the system would change. But now all of that was out the window. The precise mechanism was no longer important. All that Wigner, and later Dyson, required was that

- The matrices be Hermitian or real symmetric (so that that the eigenvalues were real)
- That the ensemble behave "naturally" under certain physical symmetry groups (GUE, GOE, GSE — more later).

During the late 50's, 60's and early 70's, various researchers (particularly, Bohigas et al.) began to test Wigner's hypothesis against real experimental data in a variety of physical situations and the results were pretty impressive. A classic reference for RMT is Mehta's book *Random Matrices*, and I invite you to look at page 17 for a comparison of 1726 nuclear level spacings against predictions of the GOE ensemble.

In making such comparisons, there is a standard **procedure**. **First**, we must rescale the resonances and the eigenvalues of the random matrices so that number of resonances / unit interval

$$= \text{expected number of eigenvalues / unit interval.}$$

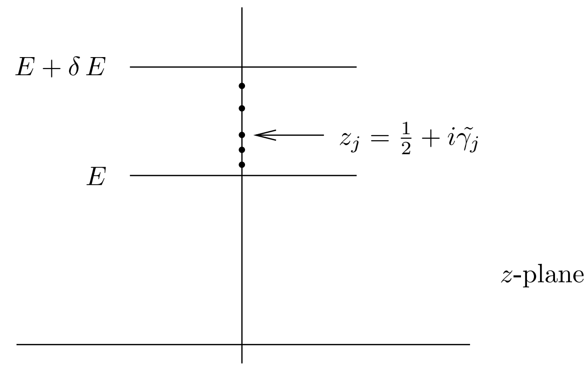
Then we compare the statistics.

More concretely, we typically have a situation where we are interested in resonances in an interval $(E, E + \delta E)$ where $\delta E \ll E$, but δE still contains many resonances. We then rescale the resonances in δE so that their average density is one. We then do a similar scaling for the eigenvalues in some ensemble, and only at that point, after we have adjusted our "microscopes," do we compare statistics. A scientist approaches such problems with two instruments in hand — a microscope and a list of ensembles:

- The dials on the microscope are adjusted to account for the macroscopics of the situation, and vary from system to system
- But once the "slide" is in focus, one sees universal behavior described by one of the entries in the list of ensembles.

Having allowed me a schematic moment, I now ask you for a small philosophical moment! It is always something of a mysterious process when we begin to think of something that is quite deterministic in a statistical way. Nevertheless, it is a process with which we are very familiar. We take a die, for example, which surely obeys the laws of solid mechanics and aerodynamics, but we readily and intuitively understand it as a statistical object. Moreover, we "know" what the stochastic model should be: all six sides have equal probability. Looked at from this point of view, what Wigner had to do was to see the neutron and nucleus system as a "die" and discover a model for its stochasticity. In all such problems there is always some singular/asymptotic process involved and when we cross some "Bayesian" point, phase space opens up, and all bets are on. We are standing, as it were, on the corner of Broadway and 42nd Street and we are watching this little kid play 3 card Monté: if we are fast enough, we can follow all his moves, but then "poof", the cards are out there, and all three cards are equally likely.

Now in the early 1970's a very remarkable thing happened. Montgomery, quite independently of these other goings on, began thinking that the zeros of the Riemann zeta function on the critical line $\operatorname{Re} z = \frac{1}{2}$, should also be viewed statistically. And so, assuming the Riemann Hypothesis and rescaling the imaginary parts of the zeros $\left\{\frac{1}{2} + i\tilde{\gamma}_j\right\}$ he computed the 2-point correlation function ($= \frac{1}{N}\,\#\{\text{ pairs } (i,j), 1 \le i \ne j \le N, \text{ such that } a \le \tilde{\gamma}_i - \tilde{\gamma}_j \le b\}$) for the rescaled zeros in an interval $i\,(a = E, b = E + \delta E) \subset i\mathbb{R}_+$, $E \gg 1$,



and he found a limiting formula

$$R = R(a,b) = \int_a^b \left[1 - \left(\frac{\sin 2\pi u}{2\pi u} \right)^2 \right] du$$

for the 2-point function as $E \to \infty$, $\#\tilde{\gamma}_i \in \delta E \to \infty$. What happened next is very well-known (I even asked Dyson to authenticate this version). Montgomery met Dyson at tea at the Institute in Princeton, and when he told him about his calculations, Dyson immediately wrote down a formula and asked Montgomery, "Did you get this?" Montgomery was astounded, and when he asked him how he knew the answer, Dyson said "Well, if the zeros of the zeta function behaved like the eigenvalues of a random GUE matrix, this would have to be the answer!" Indeed, what Montgomery had obtained for R was precisely the 2-point function for the eigenvalues for a (large) random (GUE) matrix.

At this point, the cat was out of the bag! People began asking whether their favorite list of numbers behaved like the eigenvalues of some random matrix. Through the 80's an extraordinary variety of systems were investigated from this point of view with astounding results: for example, if we take a truncated triangle $\left\{(x,y) \in \mathbb{R}^2 \;:\; 0 \le y \le x \le 1,\; x^2 + y^2 \ge r > 0\right\}$ and look at the eigenvalues $0 \le \lambda_1 \le \lambda_2 \le \cdots$ of the Dirichlet Laplacian in this truncated triangle, they too (after the standard scaling) behave like the eigenvalues of a random GUE matrix (see Mehta).

In the late 1980's, Montgomery's work was taken up numerically by Odlyzko, who then confirmed Montgomery's work to high accuracy and also investigated other statistics for the zeros of the zeta function such as the nearest neighbor spacing. Again there was incredible agreement with random matrix theory. In recent years, the work of Montgomery/Odlyzko has been a wonderful springboard for Sarnak–Rudnick, and then Sarnak–Katz, to prove all kinds of GUE (and GSE) random matrix properties for the zeros of all kinds of automorphic L functions and zeta functions over finite fields.

Until very recently, the physical and mathematical phenomena which were investigated, concerned the eigenvalues of a random matrix in the bulk of the spectrum, but in the last two years or so a very interesting class of problems has started to appear in combinatorics and also in random particle models, which concern the eigenvalues at the top of the spectrum. For example, consider the following version of solitaire called "patience sorting", which is played with a deck of cards labeled $1, 2, \dots N$. As described by Aldous and Diaconis (1999), the game is played as follows.

> The deck is shuffled, cards are turned up one at a time and dealt into piles on the table, according to the following rules:
> A low card may be placed on a higher card (e.g. 2 may be placed on 7), or may be put into a new pile to the right of the existing piles.
> At each stage we see the top card on each pile. If the turned up card is higher than the cards showing, then it *must* be put into a new pile to the right of the others. The object of the game is to finish with as few piles as possible.

The *greedy* strategy is always to place a card on the leftmost possible pile. A simple computation shows that the greedy strategy is optimal. If the shuffled deck is in a "permutation state" π, we let $p_N(\pi)$ denote the number of piles one obtains by playing patience sorting starting from π and using the greedy strategy.

For example, suppose $N = 6$ cards are in the order

$$4 \;\; 1 \;\; 3 \;\; 5 \;\; 6 \;\; 2$$

(corresponding to the permutation $\pi(1) = 4, \pi(2) = 1, \dots$). Then the game proceeds as follows:

$$\begin{array}{cccccc} & 1 & 1 & 1 & 1 & 12 \\ 4 & 4 & 43 & 435 & 4356 & 4356 \end{array}$$

and $p_6(\pi) = 4$.

Question: Putting uniform distribution on the set of shuffles $\{\pi \in S_N\}$, $S_N =$ symmetric group, how does $p_N(\pi)$ behave as $N \to \infty$? Or said differently, how big a table do I need typically to play the game with N cards?

The theorem which we will discuss in the fourth lecture is the following (J. Baik, K. Johannsson, Deift (1999))

Theorem 1. *As $N \to \infty$, p_N, suitably centered and scaled, behaves statistically like the largest eigenvalue of a random GUE matrix.*

More precisely suppose $\lambda_1^{(N)}(M) \geq \lambda_2^{(N)}(M) \geq \cdots \geq \lambda_N^{(N)}(M)$ are the eigenvalues of a $N \times N$ GUE matrix. Then

$$\begin{aligned} \lim_{N\to\infty} \operatorname{Prob}\left(\frac{p_N - 2\sqrt{N}}{N^{1/6}} \leq t \right) &= \lim_{N\to\infty} \operatorname{Prob}\left(\frac{\lambda_1^{(N)}(M) - \sqrt{2N}}{2^{-1/2}N^{-1/6}} \leq t \right) \\ &\equiv F(t) \end{aligned} \tag{1}$$

$F(t)$ is called the Tracy–Widom distribution after its discoverers, and can be expressed explicitly in terms of a solution of the Painlevé II equation.

There are many equivalent formulations of the patience sorting problem — in terms of the longest increasing subsequence of a random permutation (Ulam's problem: see Lecture 4), the height of a nucleating droplet in a supersaturated medium in the so-called Polynuclear Growth (PNG) model, and the number of boxes in the first row of a Young diagram under Plancherel measure (see Deift, *Notices of the AMS*, 2000, and the references therein). The Young diagram version is particularly interesting: it turns out that the number of boxes in the second row behaves statistically like the second largest eigenvalue $\lambda_2^{(N)}(M)$ of a random GUE matrix, and so on.

Now what are these three basic distributions mentioned above that were singled out by Wigner and Dyson (on the basis of the behavior of the system under certain physical symmetry groups)?

I Gaussian Unitary ensemble (GUE) consisting of
  a) $N \times N$ Hermitian matrices $M = (M_{kj}) = (M_{kj}^R + iM_{kj}^I)$
  b) A probability distribution

$$P(M)\,dM = P(M) \prod_{k=1}^{N} dM_{kk} \prod_{k<j} dM_{kj}^R \prod_{k<j} dM_{kj}^I.$$

  which is invariant under unitary conjugations

$$M \mapsto U^{-1}MU = M', \quad U \text{ unitary}$$

  i.e. $P(M')\,dM' = P(M)\,dM$.
  The factor $P(M)$ is introduced to turn Lebesgue measure dM into a probability measure.
  c) $\left\{M_{kk}, M_{kj}^R, M_{kj}^I\right\}$ are independent so

$$P(M) = \prod_{j=1}^{N} f_{jj}^{(0)}(M_{jj}) \prod_{j<k} f_{jk}^{(0)}(M_{jk}^R) \prod_{j<k} f_{jk}^{(1)}(M_{jk}^I)$$

Gaussian Orthogonal ensemble (GOE) consisting of
  a) $N \times N$ real symmetric matrices $(M = M_{kj})$.
  b) A probability distribution

$$P(M)\,dM = P(M) \prod_{k \le j} dM_{kj}$$

  which is invariant under orthogonal conjugation

$$M \mapsto O^{-1}MO = M', \quad O \text{ orthogonal}$$

  i.e. $P(M')dM' = P(M)dM$.

c) $\{M_{kj}\}$, $k \leq j$ are independent so

$$P(M) = \prod_{k \leq j} f_{kj}(M_{kj}).$$

II Gaussian Symplectic ensemble (GSE) consisting of

a) $2N \times 2N$ Hermitian self-dual matrices, i.e. $M = M^*$ and $M = JM^T J^{-1}$ where $J = \begin{pmatrix} \sigma & & & 0 \\ & \sigma & & \\ & & \ddots & \\ 0 & & & \sigma \end{pmatrix}$ where $\sigma = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$.

b) A probability distribution which is invariant under

$$M \mapsto U^{-1} M U = M', \quad U \text{ unitary and symplectic}$$

i.e. $U^* U = I$ and $U^T J U = J$. This invariance is most conveniently expressed by writing M as a $N \times N$ matrix with quaternion entries (see Mehta).

c) The linearly independent components of M are also statistically independent as in Ic, Ic. Again, this is most conveniently expressed by writing M as a $N \times N$ matrix with quaternion entries.

The analysis of GUE is the simplest and I am going to restrict myself to this case throughout these lectures. So, focusing on GUE, the first theorem in the business is that if $P(M)$ satisfies Ia,Ib, Ic, then necessarily

$$P(M) = \text{const}\, e^{-\text{tr}(\alpha M^2 + \beta M + \gamma)}$$

where $\alpha > 0, \beta, \gamma \in \mathbb{R}$. Centering and rescaling we have the **GUE distribution**

$$P(M) = \frac{1}{Z_N} e^{-\text{tr} M^2} dM \tag{2}$$

for some normalization constant Z_N.

Now here comes the problem: whereas conditions Ia and Ib are physically motivated, Ic is just a device and has no physical basis.

If we just assume Ia and Ib we find that

$$P(M) dM = \frac{1}{Z'_N} e^{-\text{tr} V(M)} dM \tag{3}$$

for some real valued function V, $V(x) \to \infty$ as $|x| \to \infty$.

Ensembles of matrices of type (3) are called Unitary Ensembles (UE's). Now physicists turn this all on its head and say: "As there is no physical way to distinguish between different choices of V, then whatever answer to a physical question that we compute, the answer must be independent of V." This is a rough form of what is meant by universality in random matrix theory. In the coming lectures we will focus on a particular basic statistic, the probability $P^{(V)}_{(a,b)}$ that a (large) GUE matrix has no eigenvalues in an interval

(a, b), and we will show that as $N \to \infty$ the $P^{(V)}_{(a,b)}$ indeed has a universal form independent of V. In proving this, we will employ techniques that have arisen in integrable theory over the years and can be traced back in some form to the original method of solution for the KdV equation discovered in 1967 by Gardner, Greene, Kruskal and Miura.

To summarize, what has emerged, somewhat mysteriously and to everyone's surprise is a very powerful heuristic: all kinds of physical and mathematical objects, behave, in some limit, like random matrices. Next time you have a system that you want to model in such a way with a large matrix, go for it. You have a good chance of being right!

Lecture 2

As indicated in Lecture 1, in this lecture I will introduce various methods and ideas from the theory of Integrable Systems that are needed in the solution of the Universality Conjecture. So recall that a Unitary Ensemble (UE) is the ensemble of $N \times N$ Hermitian matrices $M = M^* = (M_{ij})$ with probability distribution

$$\begin{aligned} P(M)\, dM &= \frac{1}{Z_N} e^{-\mathrm{tr} V(M)} dM \qquad (4) \\ &= \frac{1}{Z_N} e^{-\mathrm{tr} V(M)} \prod_{i=1}^{N} dM_{ii} \prod_{i<j} d\, M_{ij}^R \prod_{i<j} d\, M_{ij}^I \end{aligned}$$

where $V : \mathbb{R} \to \mathbb{R},\ V(x) \to \infty$ sufficiently rapidly as $|x| \to \infty$ and Z_N is the normalization constant. Of course the probability distribution (4) turns the eigenvalues

$$\lambda_1(M) \geq \lambda_2(M) \geq \cdots \geq \lambda_N(M)$$

of a matrix M in the ensemble into random variables.

Now fix an energy E and look at the eigenvalues in a window $[E-x, E+x]$ about E, and scale the x axis $x \mapsto y = \gamma_N(x-E)$ so that the expected number of eigenvalues/y-interval equals 1.

Now for

$$P_N(\theta; E) \equiv \mathrm{Prob}\,\{M : M \text{ has no eigenvalues in}\,(E-\theta, E+\theta)\}$$

compute

$$\lim_{N \to \infty} P_N\left(\frac{y}{\gamma_N}; E\right).$$

In the case of GUE, $P(M) = \frac{1}{Z_N} e^{-\mathrm{tr} M^2} dM$, a calculation of Gaudin and Mehta using classical special function theory (more later) showed that

$$\gamma_N \sim N^{\frac{1}{2}} \quad \left(= N^{1-\frac{1}{2m}},\ \text{for } V = x^{2m}\right) \qquad (5)$$

and that

$$\lim_{N\to\infty} P_N\left(\frac{y}{\gamma_N}; E\right) = \det(1 - S_y) \tag{6}$$

where S_y denotes the trace class operator with kernel

$$S_y(\xi, \eta) = \frac{\sin \pi(\xi - \eta)}{\pi(\xi - \eta)} \tag{7}$$

acting on $L^2(-y, y)$. The **universality conjecture** is the claim that (6) is true with the same right hand side for all (suitable) V. The only thing that depends on V is the "setting" for the microscope of Lecture 1, viz. γ_N.

This conjecture was first considered in the physics literature by Brezin and Zee (about 1992), and in the mathematical literature in 1997 by Pastur and Scherbina, and in a special case, $V(x) = x^4 + tx^2$, by Its and Bleher (1999). We will follow the method of Deift, Kriecherbauer, K. McLaughlin, Venakides and X. Zhou (DKMVZ (1999)). A pedagogic presentation of the method is given in Deift, Courant Lecture Notes #3, AMS 2000.

In DKMVZ the authors prove the universality conjecture for two kinds of potentials V

- $V(x) = t_{2m}x^{2m} + \cdots + t_0, \quad t_{2m} > 0$ and
- $V(x) = NQ(x)$ where
  1. $Q(x)$ is real analytic in a neighborhood of $\mathbb{R}$
  2. $\frac{Q(x)}{|\log(x)|} \to +\infty$ as $|x| \to \infty$

In these lectures, for simplicity, I will only consider the special case

$$\begin{aligned} V(x) &= x^{2m} \quad \text{for some positive integer } m \\ \text{and} \quad E &= 0 \end{aligned}$$

This case contains most (but not all) of the difficulties — more later.

The proof of (6) proceeds in steps:

Step 1 (Weyl integration formula)

Every Hermitian matrix M has a spectral representation

$$M = U\Lambda U^*$$

where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$, $\lambda_1 \geq \cdots \geq \lambda_N$ are the eigenvalues of M and U is the matrix of orthonormal eigenvectors. If the λ_i are distinct (which is true on an open dense set of matrices of full measure), then U is uniquely determined as an element of $U(N)/T(N)$ where $U(N)$ is the unitary group and $T(N)$ is the N-torus, and we consider the change of variables

$$M = U\Lambda U^* \mapsto (\Lambda, U) \in (\mathbb{R}^N)^{+\downarrow} \times U(N)/T(N)$$

where $(\mathbb{R}^N)^{+\downarrow}$ is the set of ordered points $\lambda_1 > \cdots > \lambda_N$ in $\mathbb{R}^N$.

The critical fact is that under this change of variables the λ_i and the U_{ij} become statistically independent

$$\frac{1}{Z_N} e^{-\mathrm{tr} V(M)\, dM} = \frac{1}{Z_N} e^{-\sum V(\lambda_i)} \prod_{i<j} (\lambda_i - \lambda_j)^2 \; d^N\lambda \; K(p) \; dp_1 \dots dp_{N(N-1)}$$

where the variables p_i describe the U_{ij} and $K = K(p)$ depends only on the p_i's.

Thus if we are interested in computing the expectation of functions F, say, which are invariant under conjugation

$$F(M) = F(UMU^*)$$

so that such functions only depend on (symmetric functions of) the eigenvalues of M, we have the Weyl integration formula

$$\begin{aligned}\mathbb{E}(F) &= \frac{1}{Z_N} \int_{\lambda_1 > \cdots > \lambda_N} F(\lambda_1, \dots, \lambda_N) e^{-\sum_i V(\lambda_i)} \prod_{i<j} (\lambda_i - \lambda_j)^2 \, d^N\lambda \, K(p) \, d^{N(N-1)} p \\ &= \frac{1}{Z'_N} \int_{\lambda_1 > \cdots > \lambda_N} F(\lambda_1, \dots, \lambda_N) e^{-\sum_i V(\lambda_i)} \prod_{i<j} (\lambda_i - \lambda_j)^2 \, d^N\lambda\end{aligned}$$

where the $dp_i \cong dU_{ij}$ have been integrated out, and Z'_N is the new norming constant. Thus we are lead to the distribution on the eigenvalues

$$P_N(\lambda) d^N\lambda = \frac{1}{Z'_N} e^{-\sum_i V(\lambda_i)} \prod_{i<j} (\lambda_i - \lambda_j)^2 \chi_{\lambda_1 > \cdots > \lambda_N} d^N\lambda \tag{8}$$

(For GOE, $\cdots \prod_{i<j} |\lambda_i - \lambda_j|$, for GSE, $\cdots \prod_{i<j} |\lambda_i - \lambda_j|^4$. Thus, in the language of thermodynamics, physicists speak of the inverse temperature $\beta = 1, 2$, or 4 — more later.) From (8) we see that the probability that eigenvalues are close together is small, and one speaks of "eigenvalue repulsion." This is a fundamental property of such ensembles which implies, in particular, that the eigenvalue spacings are not Poisson.

Step 2 (Enter orthogonal polynomials: OP's)

Recall that if $d\mu(x)$ is a measure on $\mathbb{R}$ with finite moments

$$\int |x|^q d\mu(x) < \infty, \quad q = 0, 1, 2, \dots$$

then $d\mu$ generates via the Gramm–Schmidt procedure a unique set of orthonormal polynomials

$$p_n(x) = c_n \pi_n(x) = c_n(x^n + \dots), \; c_n > 0, \; n = 0, 1, 2, \dots$$

satisfying

$$\int p_n(x)\, p_m(x)\, d\mu(x) = \delta_{n,m}, \quad n, m \geq 0 \tag{9}$$

Let $\{p_n(x) = p_n(x, V)\}_{n \geq 0}$ denote the orthonormal polynomials generated by $d\mu(x) = e^{-V(x)}\, dx$. The standard fact in the business is that the p_n's satisfy a three term recurrence relation

$$L\, p = \begin{pmatrix} a_0 & b_0 & 0 & \cdots \\ b_0 & a_1 & b_1 & \ddots \\ 0 & b_1 & a_2 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ \vdots \end{pmatrix} = x \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ \vdots \end{pmatrix} = xp.$$

It is a curious, helpful and somewhat mysterious fact that these polynomials have a very interesting "functorial interrelationship" with RMT, viz., given V, there are 2 natural constructions

$$\begin{tikzcd}
 & V \arrow[dl] \arrow[dr] & \\
e^{-\operatorname{tr} V(M)}\, dM \arrow[d] & & e^{-V(x)}\, dx \arrow[d] \\
\text{mean field theory} & & \{p_n\}_{n \geq 0} \arrow[d] \\
 & & L = \begin{pmatrix} a_0 & b_0 & 0 \\ b_0 & a_1 & b_1 \\ 0 & \ddots & \ddots \end{pmatrix}, \text{ Jacobi operator}
\end{tikzcd}$$

The "random matrix version" of e^{-V} plays the role of a sort of "$\frac{1}{2}$-quantization" of the "orthogonal polynomial version" of e^{-V}. In a very concrete way mean field theory results imply critical properties of orthogonal polynomials and, on the other hand, results from orthogonal polynomials are crucial in understanding the $N \to \infty$ limit in RMT.

For example, we have the following basic result of Gaudin and Mehta. Let $\theta > 0$, and consider the quantity $P_N(\theta) \equiv P_N(\theta; E = 0)$ introduced earlier. Then it turns out that $P_N(\theta)$ can be expressed explicitly in terms of orthogonal polynomials. Indeed

$$P_N(\theta) = \det(I - K_N) \tag{10}$$

where K_N is the finite rank (and hence trace class operator) acing on $L^2(-\theta, \theta)$ with kernel

$$K_N(x, y) = \sum_{j=0}^{N-1} \phi_j(x) \phi_j(y) \tag{11}$$

where $\phi_j(x) = e^{-\frac{1}{2}V(x)} p_j(x)$. Note that

$$\int \phi_j(x)\, \phi_l(x)\, dx = \delta_{jl}, \quad j, l \geq 0 \tag{12}$$

The essential technical step in deriving (10) is a so-called "integrating out" lemma of Gaudin.

Of course, what we are interested in is $\lim_{N\to\infty} P_N\left(\frac{y}{\gamma_N}; E = 0\right)$ and so we see that the question of universality reduces to a question of the asymptotics of orthogonal polynomials. More precisely one needs, in the case $V(x) = x^{2m}$, the asymptotics of

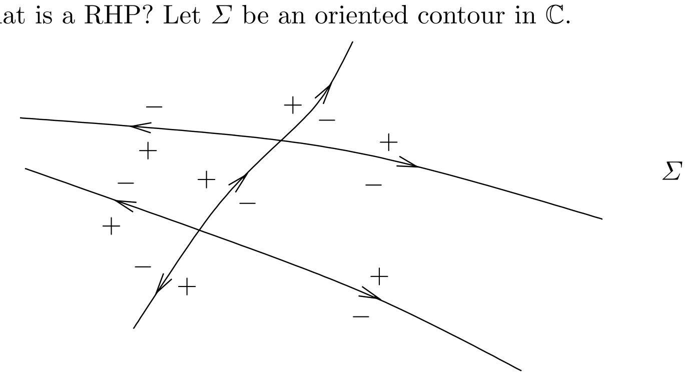
$$p_N(N^{\frac{1}{2m}} x), \quad p_{N-1}(N^{\frac{1}{2m}} x) \quad \text{as} \quad N \to \infty. \tag{13}$$

Such asymptotics is of "double-scaling type" and was known in the classical literature only for the Hermite polynomials corresponding to the weight $e^{-x^2} dx$ (Plancherel–Rotach, 1929). This is precisely the weight that arises in the Gaussian UE, $e^{-\mathrm{tr} M^2} dM$, and this is the reason that Gaudin and Mehta were able to compute (6) for GUE. Apart from the special case considered by Its and Bleher $V(x) = x^4 + tx^2$, the papers DKMVZ were the first to prove Plancherel–Rotach-type asymptotics for general orthogonal polynomials, yielding the universality conjecture. It is here that various techniques from integral systems begin to play an explicit role.

Step 3 (Enter Riemann–Hilbert problems RHP)

The "double-scaling" analysis in DKMVZ proceeds by expressing orthogonal polynomials in terms of the solution of a RHP, initially introduced by Fokas, Its and Kitaev (1992) and then analyzing this RHP asymptotically using the non-commutative steepest descent method for RHP's introduced by X. Zhou and Deift in 1993. I will say much more about this method in the next lecture.

So what is a RHP? Let Σ be an oriented contour in $\mathbb{C}$.



Note that (by convention) the (+)-side (resp. (-)-side) of Σ lies to the left (resp. right) as one moves in the direction of the orientation.

Suppose in addition that for some $k \in \mathbb{N}$ we have a map

$$v : \Sigma \to GL(k, \mathbb{C})$$

such that $v, v^{-1} \in L^\infty(\Sigma)$. The RHP (Σ, v) consists of the following: find an $l \times k$ matrix function $m = m(z)$ such that

- m is analytic in $\mathbb{C} \setminus \Sigma$
- (jump condition) $m_+(z) = m_-(z)\, v(z)$ for a.e. $z \in \Sigma$ where

$$m_\pm(z) = \lim_{\substack{z' \to z \\ z' \in (\pm)\,\text{side}}} m(z').$$

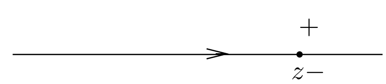
If in addition $l = k$ and

- $m(z) \to I \quad \text{as} \quad z \to \infty$

we say that the solution of the RHP is normalized at ∞.

It is a remarkable fact that the solution of an extraordinary number of systems of great mathematical and physical interest can be expressed in terms of the solution of some associated RHP. This includes all the classical integrable systems and provides a gateway for integrable methods into a wide variety of a priori unrelated scientific problems.

The RHP that Fokas, Its and Kitaev found for orthogonal polynomials is the following. Recall that $p_k = c_k \pi_k$ where π_k is monic, $\pi_k = x^k + \cdots$. Let $\Sigma = \mathbb{R}$, oriented from left to right.



Let $v(z) = \left(\begin{smallmatrix} 1 & e^{-V(z)} \\ 0 & 1 \end{smallmatrix}\right)$, $z \in \mathbb{R}$. Fix a positive integer k and let $Y = Y^{(k)} = (Y_{ij})_{1 \le i,j \le 2}$ be the (unique) 2×2 solution of the RHP (Σ, v).

$$\begin{cases} \bullet\ Y(z) \ \text{ analytic in } \mathbb{C} \setminus \mathbb{R} \\ \bullet\ Y_+(z) = Y_-(z) v(z), \quad z \in \mathbb{R} \\ \bullet\ Y(z) \begin{pmatrix} z^{-k} & 0 \\ 0 & z^k \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{as} \quad z \to \infty. \end{cases} \tag{14}$$

Then

$$\pi_k(x) = Y_{11}(x) \tag{15}$$

and

$$c_{k-1} = \sqrt{\frac{(Y_{21})_1}{-2\pi i}} \tag{16}$$

where

$$Y_{21}(x) = (Y_{21})_1\, x^{k-1} + \cdots \tag{17}$$

Now I haven't proved anything so far in these lectures, so let me prove (15). As you'll see it's easy and it's fun. I will also prove uniqueness because the method of proof will be important in the third lecture.

Proof. Consider the first row of the jump relation $Y_+ = Y_- v$,

$$\begin{pmatrix} Y_{11} & Y_{12} \end{pmatrix}_+ = \begin{pmatrix} Y_{11} & Y_{12} \end{pmatrix}_- \begin{pmatrix} 1 & e^{-V(z)} \\ 0 & 1 \end{pmatrix}, \qquad z \in \mathbb{R}. \tag{18}$$

Hence $(Y_{11})_+(z) = (Y_{11})_-(z)$ which implies that Y_{11} is entire. But

$$\begin{pmatrix} Y_{11} & Y_{12} \end{pmatrix} \begin{pmatrix} z^{-k} & 0 \\ 0 & z^k \end{pmatrix} = \begin{pmatrix} Y_{11} z^{-k} & Y_{12} z^k \end{pmatrix} \to \begin{pmatrix} 1 & 0 \end{pmatrix} \tag{19}$$

as $z \to \infty$. Therefore, Y_{11} is a monomial of order k, by Liouville's Theorem: $Y_{11}(z) = z^k + \cdots$. On the other hand, from (18), we see that

$$(Y_{12})_+ = (Y_{12})_- + Y_{11} e^{-V(z)}$$

and hence by the Plemelj formula (by (19), $Y_{12}(z) \to 0$ as $z \to \infty$)

$$Y_{12}(z) = \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{Y_{11} e^{-V(s)}}{s - z} \, ds.$$

Expanding in powers of $\frac{1}{z}$,

$$Y_{12}(z) = -\frac{1}{2\pi i} \int_{\mathbb{R}} Y_{11}(s) e^{-V(s)} \Big(\frac{1}{z} + \frac{s}{z^2} + \cdots \frac{s^{k-1}}{z^k} + \frac{s^k}{z^{k+1}} + \cdots \Big) \, ds.$$

But from (19), $Y_{12} = o(\frac{1}{z^k})$ as $z \to \infty$. Hence

$$\int_{\mathbb{R}} Y_{11} e^{-V(s)} s^j ds = 0, \qquad 0 \le j \le k-1.$$

By the construction of the orthogonal polynomials it follows that

$$Y_{11}(s) = \pi_k(s).$$

Now we prove uniqueness. Note that on $\mathbb{R}$

$$(\det Y)_+ = (\det Y)_- \det v(z) = (\det Y)_-$$

and hence $\det Y(z)$ is entire. But as $z \to \infty$

$$\det Y(z) = \det \left(Y(z) \begin{pmatrix} z^{-k} & 0 \\ 0 & z^k \end{pmatrix} \right) \to 1 \ \text{ as } \ z \to \infty.$$

Hence, again by Liouville's Theorem, $\det Y(z) \equiv 1$. In particular $Y^{-1}(z)$ is non-singular. Now let $\widetilde{Y}$ be a second solution of the RHP and set $H(z) = \widetilde{Y}(z) Y(z)^{-1}$. Then on $\mathbb{R}$

$$H_+(z) = \widetilde{Y}_+ Y_+^{-1} = (\widetilde{Y}_- v)(Y_- v)^{-1} = \widetilde{Y}_- v v^{-1} Y_-^{-1} = H_-(z).$$

Hence $H(z)$ is entire, and as

$$H(z) = \left(\widetilde{Y}(z)\begin{pmatrix} z^{-k} & 0 \\ 0 & z^k \end{pmatrix}\right)\left(Y(z)\begin{pmatrix} z^{-k} & 0 \\ 0 & z^k \end{pmatrix}\right)^{-1} \to I \times I = I \quad \text{as} \quad z \to \infty$$

it follows that $H(z) \equiv I$. Therefore, $\widetilde{Y}(z) = Y(z)$ and hence the solution is unique.

Fokas, Its and Kitaev arrived at their RHP through their investigation of two dimensional quantum gravity, which is an integrable system. A special case of this RHP also appeared in work of Deift, Kamvissis, Kriecherbauer and X. Zhou (1996) on the Toda rarefaction problem.

In the third lecture we will show how to solve the RHP asymptotically as $N = k \to \infty$, leading to the required Plancherel–Rotach asymptotics.

Lecture 3

Recall from Lecture 2 that the universality conjecture reduces to the verification of Plancherel–Rotach asymptotics for orthogonal polynomials with respect to the measure $e^{-V(x)}dx$. For this lecture, as is in Lecture 2, we restrict ourselves to the special case $V(x) = x^{2m}$, $m = 1, 2, \dots$ (but more later). Recall that $\pi_N(z)$ is the 11-entry of the solution Y of the Fokas–Its–Kitaev RHP, as in (15). Now it turns out that it is useful to make a preliminary rescaling of the problem.

$$z \to N^{\frac{1}{2m}} z$$

Or more precisely, set

$$S(z) = S^{(N)}(z) \equiv \begin{pmatrix} N^{-\frac{1}{2m}} & 0 \\ 0 & N^{\frac{1}{2m}} \end{pmatrix} Y(N^{\frac{1}{2m}} z) \tag{20}$$

Then $S(z) = S^{(N)}(z)$ solves (note $V(zN^{\frac{1}{2m}}) = NV(z)$) the equivalent RHP

$$\begin{cases} \bullet \ S(z) \text{ analytic in } \mathbb{C} \setminus \mathbb{R} \\ \bullet \ S_+(z) = S_-(z)\begin{pmatrix} 1 & e^{-NV(z)} \\ 0 & 1 \end{pmatrix} \text{ for } z \in \mathbb{R} \\ \bullet \ S(z)\begin{pmatrix} z^{-N} & 0 \\ 0 & z^N \end{pmatrix} \to I \quad \text{as} \quad z \to \infty \end{cases} \tag{21}$$

This scaling can be motivated as follows: we have the eigenvalue distribution

$$\frac{1}{Z'_N} e^{-\sum V(x_i)} \prod_{i<j} (x_i - x_j)^2 \, d^N x = \frac{1}{Z'_N} e^{-\left[\sum_{i \neq j} \log|x_i - x_j|^{-1} + \sum V(x_i)\right]} d^N x$$

After scaling $x \to N^{\frac{1}{2m}} x$, the distribution takes the form

$$\frac{1}{Z_N''} \, e^{-\left[\sum_{i\neq j} \log|x_i-x_j|^{-1} + N\sum V(x_i)\right]} \, d^N x$$

where Z_N'' is the new normalization constant. For any $x = (x_1, \ldots, x_N)$ let

$$h_x = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}$$

denote the normalized counting measure for the x_i's. Then the above distribution takes the form

$$\frac{1}{Z_N''} \, e^{-N^2 \Phi^V(h_x)} \, d^N x$$

where

$$\Phi^V(h_x) = \int\int \log|s-t|^{-1} \, dh_x(s) \, dh_x(t) + \int V(s) \, dh_x(s)$$

in which the repulsive "electrostatic forces" between the "charges" act on the same scale by the external potential. We expect the main contribution to come from a charge distribution where Φ^V is minimal. More precisely we are lead to consider the following auxiliary variational problem

$$\begin{aligned} E^V &= \inf_{\mu\in\mathbb{P}} \Phi^V(\mu) \\ &= \inf_{\mu\in\mathbb{P}} \left[\int\int \log|s-t|^{-1} \, d\mu(s) \, d\mu(t) + \int V(s) \, d\mu(s) \right] \end{aligned} \tag{22}$$

where $\mathbb{P}$ is the space of probability distributions on $\mathbb{R}$.

It turns out (see Saff–Totik (1997)) that the infimum is indeed achieved at a unique $\mu = \mu^V$, the **equilibrium measure**,

$$E^V = \Phi^V(\mu^V).$$

Moreover, μ^V has compact support.

The equilibrium measure $d\mu^V$ can not only be thought of as describing the equilibrium distribution of repelling charges as above, but it also turns out to be connected to a rather extraordinary collection of problems in classical analysis. We describe two of these connections which are relevant to the problem at hand:

- In the scaled unitary ensemble

  $$\frac{1}{Z_N} e^{-N \operatorname{tr} V(M)} dM \qquad (V(x) = x^{2m})$$

  we have

  $$\lim_{N\to\infty} \frac{1}{N} \mathbb{E}\,(\#\,\text{eigenvalues in } B) = \mu^V(B) \tag{23}$$

  for any Borel set B. It is this relation that teaches us how to adjust our "microscopes" to see the universal behavior. Together with the scaling in (20), (23) implies that $\gamma_N \sim N \cdot N^{-\frac{1}{2m}} = N^{1-\frac{1}{2m}}$ (cf. (5) above).

- For the scaled measure $e^{-NV(x)}dx$, let $x_1^{(N)} \geq \cdots \geq x_N^{(N)}$ denote the zeros of the Nth orthogonal polynomial $p_N(s)$. Define the normalized counting measure

$$\delta^{(N)} = \frac{1}{N}\sum_{j=1}^{N}\delta_{x_i^N}.$$

Then as $N \to \infty$

$$\delta^N \rightharpoonup \mu^V \tag{24}$$

$$\text{i.e.} \quad \lim_{N\to\infty}\int f(s)\,d\delta^N(s) = \int f(s)\,d\mu^V(s)$$

for all continuous functions f decaying at $s = \pm\infty$.

We will return to the relation (24) later on. The Euler–Lagrange relations for the equilibrium measure $d\mu^V$ have the following form: there exists a unique $l \in \mathbb{R}$ (the Lagrange multiplier) so that

(i) $2\int \log|x-y|^{-1}\,d\mu^V(y) + V(x) \geq l \quad \forall\, x \in \mathbb{R}$
(ii) $2\int \log|x-y|^{-1}\,d\mu^V(y) + V(x) = l \quad \forall\, x \in \operatorname{supp} d\mu^V$

Conversely, if $d\tilde{\mu}$ is any probability measure with compact support satisfying (i) (ii) for some real number $\tilde{l}$, then necessarily $\tilde{\mu} = \mu^V$ and $\tilde{l} = l$.

In the special case $V(x) = x^{2m}$ these variational conditions can be solved explicitly, and one finds

$$d\mu^V(z) = \psi(z)\,dz \tag{25}$$

where $\psi(z)$ is supported on a finite interval $(-a, a)$

$$\psi(z) = \frac{m}{\pi}\,(a^2 - z^2)^{\frac{1}{2}}\,h(z)\,\chi_{(-a,a)}(z) \tag{26}$$

where

$$h(z) = z^{2m-2} + \sum_{j=0}^{m-1}\beta_j z^{2j}, \qquad \beta_j > 0 \tag{27}$$

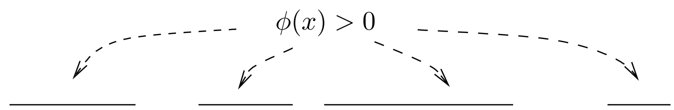
and

$$a = \left(m\prod_{k=1}^{m}\frac{2k-1}{2k}\right)^{-\frac{1}{2m}}. \tag{28}$$

Aside: For more general weights $e^{-NV(x)}\,dx$ where

- $V(x)$ is analytic in a neighborhood of $\mathbb{R}$
- $\frac{V(x)}{\log|x|} \to \infty \;$ as $\; |x| \to \infty$

one can show using a mean field theory calculation going back essentially to Itzykson that $d\mu^V$ is absolutely continuous with respect to Lebesgue measure $d\mu^V(x) = \psi(x)\,dx$ and is supported on a finite union of intervals.



This turns out to be critical information in analyzing the asymptotics of orthogonal polynomials in the general case.

We now return to the RHP (21). The calculations that follow are motivated by work which Zhou and I did together with Venakides (1994) on the zero-dispersion problem for KdV where genuinely non-linear oscillations ($\mathrm{sn}(\alpha x+\beta)$ as opposed to $\sin(\alpha x + \beta)$, where sn denotes the Jacobi function) develop.

Our first task is to turn the RHP into a normalized RHP. To this end set

$$T(z) = T^{(N)}(z) = \begin{pmatrix} e^{N\frac{\tilde{l}}{2}} & 0 \\ 0 & e^{-N\frac{\tilde{l}}{2}} \end{pmatrix} S(z) \begin{pmatrix} e^{-Ng(z)} & 0 \\ 0 & e^{Ng(z)} \end{pmatrix} \begin{pmatrix} e^{-N\frac{\tilde{l}}{2}} & 0 \\ 0 & e^{N\frac{\tilde{l}}{2}} \end{pmatrix} \quad (29)$$

Here $\tilde{l}$ is a constant, to be determined below, and $g(z)$ is an as yet undetermined function with the following properties

(i) $g(z)$ is analytic in $\mathbb{C} \setminus \mathbb{R}$
(ii) $g(z) \sim \log z$ as $z \to \infty$.

A simple, direct calculation shows that $T(z)$ solves the following RHP:

$$\begin{cases} \bullet \ \ T(z) \text{ analytic in } \mathbb{C} \setminus \mathbb{R} \\ \bullet \ \ T_+(z) = T_-(z)\, v_T(z), \quad z \in \mathbb{R}, \\ \text{where } \ v_T(z) = \begin{pmatrix} e^{N(g_-(z)-g_+(z))} & e^{N(g_-(z)+g_+(z)-V+\tilde{l})} \\ 0 & e^{N(g_+(z)-g_-(z))} \end{pmatrix} \\ \bullet \ \ T(z) \to I \ \text{ as } \ z \to \infty \end{cases} \quad (30)$$

The normalization condition $T(z) \to I$ follows as

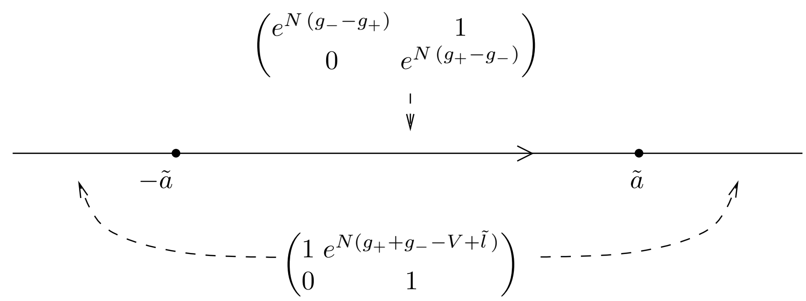
$$\begin{pmatrix} e^{-Ng(z)} & 0 \\ 0 & e^{Ng(z)} \end{pmatrix} \sim \begin{pmatrix} z^{-N} & 0 \\ 0 & z^N \end{pmatrix} \quad \text{as } \ z \to \infty.$$

Note that the RHP (30) is equivalent to the original RHP (14) in the sense that the solution of one of them immediately gives the solution of the other.

We now list the additional properties that $g(z)$ should have. The reason for these properties will become clear a little further on.

(a) Suppose that on some finite interval $(-\tilde{a}, \tilde{a})$, $g_+(z)+g_-(z)-V(z)+\tilde{l} = 0$
(b) $g_+(z) + g_-(z) - V(z) + \tilde{l} < 0$ for $|z| > \tilde{a}$
(c) $g_+(z) - g_-(z) \in i\mathbb{R}$ and $i\frac{d}{dz}(g_+(z) - g_-(z)) > 0$ for $z \in (-\tilde{a}, \tilde{a})$
(d) $e^{N(g_+(z)-g_-(z))} = 1$ for $|z| > \tilde{a}$

Now observe that if g satisfies (a) and (d) $v_T(z)$ takes the form



Set

$$G(z) = g_+(z) - g_-(z) \quad \text{for } z \in (-\tilde{a}, \tilde{a})$$

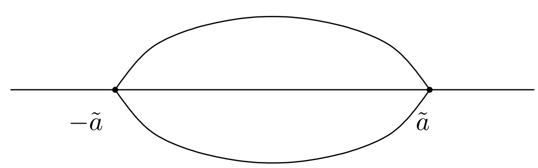
and observe that

$$\begin{pmatrix} e^{-NG} & 1 \\ 0 & e^{NG} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ e^{NG} & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ e^{-NG} & 1 \end{pmatrix}$$

Also as

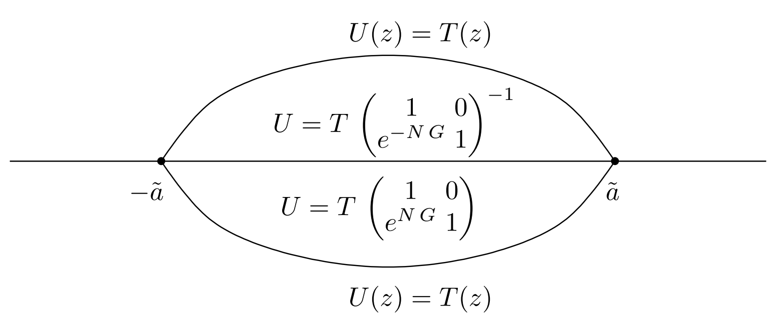
$$G(z) = g_+ - (V - \tilde{l} - g_+) = 2g_+ - V + \tilde{l} = V - \tilde{l} - g_- - g_- = V - \tilde{l} - 2g_-$$

G has an analytic continuation to the lips

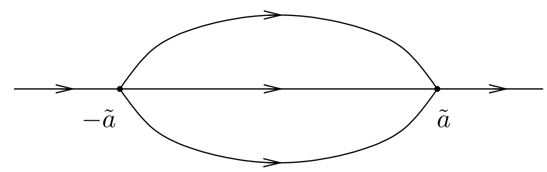


around $(-\tilde{a}, \tilde{a})$.

We now deform the RHP (30) for T by setting

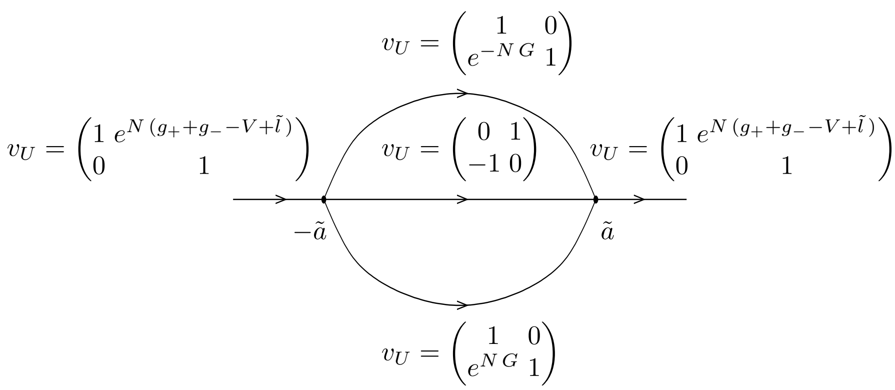


A simple calculation shows that U solves the equivalent RHP on the extended contour Σ_U

$\Sigma_U =$ 

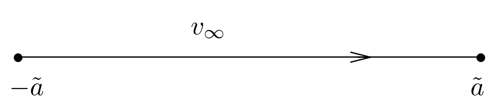
$$\begin{cases} \bullet \;\; U(z) \text{ analytic in } \mathbb{C} \setminus \Sigma_U \\ \bullet \;\; U_+(z) = U_-(z)\, v_U(z) \text{ for } z \;\in \Sigma_U \\ \bullet \;\; U(z) \to I \;\; \text{ as } \;\; z \to \infty \end{cases} \tag{31}$$

where



Note finally that by (c) and the Cauchy–Riemann conditions, Re $G > 0$ on the upper lip and Re $G < 0$ on the lower lip. And also by (b), $g_+ + g_- - V + \tilde{l} < 0$ for $|z| > \tilde{a}$.

At this stage, morally, we are done! Indeed we see that the choices (a)-(d) were made precisely so that the jump matrix v_U converges to the limiting form $v_\infty = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ on $(-\tilde{a}, \tilde{a})$.



and we expect that

$$U \to U^\infty$$

where U^∞ solves the RHP

$$\begin{cases} \bullet \;\; U^\infty \text{ analytic in } \mathbb{C} \setminus [-\tilde{a}, \tilde{a}] \\ \bullet \;\; U^\infty_+(z) = U^\infty_-(z) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \text{ for } z \in (-\tilde{a}, \tilde{a}) \\ \bullet \;\; U^\infty(z) \to I \;\; \text{ as } \;\; z \to \infty \end{cases}$$

This RHP is easily solved explicitly and hence gives, after we unravel all our transformations, the leading asymptotics for the orthogonal polynomials as desired.

Now how do we know that a $g(z)$ satisfying (a)–(d) exists? This is where one of the truly lucky things in the whole business happens. Suppose μ is a probability measure and set

$$g(z) \equiv \int \log(z-s)\, d\mu(s) \tag{32}$$

Then a direct calculation shows that $g(z)$ solves (a)-(d) $\iff$ $d\mu$ satisfies the Euler–Lagrange equations (i), (ii), i.e. $d\mu = d\mu^V$, the equilibrium measure!

So once we have computed the equilibrium measure, we are done, at least morally.

Now why only morally? It turns out that to show that the solution of a RHP converges to the solution of some other RHP, we need to know that the jump matrices converge in the $L^\infty(\Sigma)$ norm. But although $v_U(z) \to v_\infty(z)$ as $N \to \infty$ pointwise on Σ_U, the rate of convergence is slower and slower near the endpoints $\pm\tilde{a}$, and so the convergence is not in L^∞. This is an absolutely major issue and most of the analytical work in the method goes into overcoming this difficulty. We proceed by constructing a "local solution" to the RHP and patching it together to obtain a parametrix for the full problem.

The problem we face here has similarities to the homogenization problem in elliptic PDE theory, e.g. suppose we have

$$H_\epsilon = -\nabla \cdot a\left(\frac{x}{\epsilon}\right)\nabla$$

so H_ϵ converges weakly to some operator H_0 as $\epsilon \downarrow 0$. What can we say about $\lim_{\epsilon\downarrow 0} f(H_\epsilon)$ for various functions f, e.g. does $(H_\epsilon+1)^{-1}$ converge (weakly) to $(H_0+1)^{-1}$? In general the answer is no and oscillations play a crucial role in determining the limit, just as they do in the RHP.

Finally, we note that there is another way to see that $\begin{pmatrix} e^{-Ng(z)} & 0 \\ 0 & e^{Ng(z)} \end{pmatrix}$ is the "right" normalizer for the RHP. Recall that

$$Y_{11} = \pi_N(z) = \prod_{i=1}^N (z - x_i^{(N)}) = e^{\sum_i \log(z-x_i^N)} = e^{N\int \log(z-s)\, d\delta^N(s)}$$

where again $\delta^N = \frac{1}{N}\sum_i \delta_{x_i^N}$ is the normalized counting measure for the zeros of π_N. But by (24),

$$\delta^N \rightharpoonup d\mu^V.$$

Thus we expect

$$Y_{11} \sim e^{N\int \log(z-s)\, d\mu^V(s)} = e^{N\,g(z)}$$

and we anticipate that

$$Y\begin{pmatrix} e^{-N\,g} & 0 \\ 0 & e^{N\,g} \end{pmatrix} = \begin{pmatrix} Y_{11}e^{-N\,g} & * \\ * & * \end{pmatrix}$$

is well behaved!

Lecture 4

In this the final lecture I will consider Ulam's problem in combinatorics and show how integrable methods lead to its solution. I will describe joint work with J. Baik and K. Johansson (1999).

Let S_N be the group of permutations $\{\pi\}$ of the numbers $1, 2, \ldots, N$. For $1 \leq i_1 < i_2 < \cdots < i_k \leq N$ we say that $\pi(i_i), \ldots, \pi(i_k)$ is an ***increasing subsequence*** of length k in π if $\pi(i_1) < \cdots < \pi(i_k)$. Let $l_N(\pi)$ be the length of the longest increasing subsequence, e.g. if $N = 5$ and π is the permutation $(1\,2\,3\,4\,5) \mapsto (5\,1\,3\,2\,4)$, then 1 3 4 and 1 2 4 are both increasing subsequences of length 3 and are the longest increasing subsequences.

Equip S_N with the uniform distribution so that

$$q_{n,M}^{(1)} \equiv \operatorname{Prob}\{l_N(\pi) \leq n\} = \frac{f_{n,N}}{N!}$$

where $f_{n,N} = \#$ of π's with $l_N(\pi) \leq n$.

Our aim is the following: Determine the asymptotics of $q_{n,N}$ as $N \to \infty$.

Now why is l_N of interest?

There are connections to

- Representation theory of S_N (Young tableaux — more later).
- Ulam's metric on S_N, $d(\pi, \sigma) = N - l_N(\pi\sigma^{-1})$ which is useful for a variety of statistical questions.
- A variety of statistical mechanical particle models — directed first passage percolation questions, growth models, exclusion process models, "vicious walker" models, super-conductivity models
- Random topologies on surfaces (Okounkov).
- Patience sorting as described in Lecture 1. Indeed $p_N(\pi) = l_N(\pi)$!

Some History: The beginning of the business

(1935) Erdös and Szekeres proved a Ramsey theory type result: every permutation of N numbers has either an increasing or decreasing subsequence of length at least $\sqrt{N-1}$. Therefore, $\mathbb{E}_N(l_N) \geq \frac{1}{2}\sqrt{N-1}$.

(1961) Working on one of the early computers, Ulam did some Monté Carlo simulations which indicated

$$\mathbb{E}_N(l_N) \sim 1.7\, N^{\frac{1}{2}}$$

Here $N \leq 10$, but still he believed this was a strong hint for the case $N \to \infty$, and he conjectured that

$$\lim_{N\to\infty} \frac{\mathbb{E}_N(l_N)}{N^{\frac{1}{2}}} = c \tag{33}$$

exists. This became known as Ulam's problem: prove the limit exists and compute c.

(1968) Numerical work of Baer and Brock indicated that $c = 2$, if it exists. The critical analytical development came in

(1972) Hammersley showed that the above limit exists, but he did not compute the value of c. He used Kingman's subadditive ergodic theory ideas for some associated stochastic process. Then in

(1977) Logan and Shepp showed $c \geq 2$, and independently Vershik and Kerov showed $c = 2$.
Independent proofs of (33) with $c = 2$ have now been given by Aldous and Diaconis (1995), Seppäläinen (1996) and Johansson (1997), amongst others.

Over the years many conjectures have been made about other statistics for l_N. In particular there were various conjectures for the variance $\mathrm{Var}_N(l_N)$ of the form

$$\mathrm{Var}_N(l_N) \sim cN^{\alpha}$$

for different values of α. In particular, a calculation of Kesten based on statistical mechanical considerations indicated that $\alpha = \frac{1}{3}$.

Then in 1993, together with Eric Rains, Andrew Odlyzko began a series of large scale ($N \sim 10^9$) Monté Carlo simulations on $\mathrm{Var}_N(l_N)$. They found that

$$\lim_{N\to\infty} \frac{\mathrm{Var}_N(l_N)}{N^{\frac{1}{3}}} = c_0 \sim 0.813,$$

indicating that $\alpha = \frac{1}{3}$ as in Kesten's calculations, and also

$$\lim_{N\to\infty} \frac{\mathbb{E}_N(l_N) - 2\sqrt{N}}{N^{\frac{1}{6}}} = c_1 \sim -1.779.$$

Apart from various large deviation results of Deuschel and Zeitouni (1995) (1997), J.-H. Kim (1996) and Seppälainen (1998), that is where the matter stood.

In order to state our results on the problem I need to introduce the Tracy–Widom distribution. Let $u(x)$ be the (unique) solution of the Painlevé II (PII) equation

$$u_{xx} = 2u^3 + xu, \ u \sim -\mathrm{Ai}(x) \quad \text{as} \quad x \to +\infty$$

where $\mathrm{Ai}(x)$ is the Airy function.

We have (and I will say more about this later)

$$\begin{aligned} u(x) &= -\mathrm{Ai}(x) + O\left(\frac{e^{-\frac{4}{3}x^{\frac{3}{2}}}}{x^{\frac{1}{4}}}\right), \ x \to +\infty \\ &= -\sqrt{\frac{-x}{2}}\left(1 + O\left(\frac{1}{x^2}\right)\right) \quad \text{as} \quad x \to -\infty \end{aligned}$$

The existence of a solution of PII with this asymptotics is due to Hastings and McLeod (1980). Define the (first) Tracy–Widom distribution

$$
\begin{aligned}
F^{(1)}(t) &= e^{-\int_t^\infty (x-t)\, u^2(x)\, dx} \\
&\sim e^{ct^3} \quad \text{as} \quad t \to -\infty \\
&\sim 1 - e^{-c\, t^{\frac{3}{2}}} \quad \text{as} \quad t \to +\infty
\end{aligned}
$$

Thus $\frac{dF^{(1)}}{dt}(t) > 0$, $F^{(1)}(t) \to 0$ as $t \to -\infty$, $F^{(1)}(t) \to 1$ as $t \to \infty$, so $F^{(1)}(t)$ is indeed a distribution function.

Theorem 2. *Let S_N be the group of permutations of $1, \ldots, N$ with uniform distribution, and let $l_N(\pi)$ be the length of the longest increasing subsequence for $\pi \in S_N$. Let $\chi^{(1)}$ be a random variable whose distribution is $F^{(1)}$. Then as $N \to \infty$*

$$
\chi_N = \frac{l_N - 2\sqrt{N}}{N^{\frac{1}{6}}} \to \chi
$$

in distribution. That is,

$$
\lim_{N\to\infty} \operatorname{Prob}\left(\chi_N \le t\right) = F^{(1)}(t), \quad t \in \mathbb{R}.
$$

We also have convergence of the moments.

Theorem 3. *For any $m = 1, 2, 3, \ldots$ we have*

$$
\lim_{N\to\infty} \mathbb{E}_N\ (\chi_N^m) = \mathbb{E}\,(\chi^m)
$$

where $\mathbb{E}$ denotes expectation with respect to $F^{(1)}$. In particular for $m = 2$,

$$
\lim_{N\to\infty} \frac{\operatorname{Var}(l_N)}{N^{\frac{1}{3}}} = \int t^2\, dF^{(1)}(t) - \left(\int t\, dF^{(1)}(t)\right)^2 \tag{34}
$$

and for $m = 1$

$$
\lim_{N\to\infty} \frac{\mathbb{E}_N(l_N) - 2\sqrt{N}}{N^{\frac{1}{6}}} = \int t\, dF^{(1)}(t) \tag{35}
$$

If one solves PII numerically and computes the right hand side of (34), (35) resp., one finds 0.8132 and -1.771 which agrees pretty well with the Monté-Carlo values c_0 and c_1 above.

Now it turns out that there is a very interesting connection between the above results and random matrix theory, in particular with GUE. The fact is this (Tracy–Widom (1994)): as $N \to \infty$ the distribution of the largest eigenvalue, suitably centered and scaled $\left(\lambda_1 \mapsto \left(\lambda_1 - \sqrt{2N}\right) N^{\frac{1}{6}} 2^{\frac{1}{2}}\right)$, of a matrix M converges precisely to the function $F^{(1)}(t)$! That is,

$$
\operatorname{Prob}\left\{\frac{\lambda_1 - \sqrt{2N}}{N^{-\frac{1}{6}} 2^{-\frac{1}{2}}} \le t\right\} \to F^{(1)}(t) \quad \text{as} \quad N \to \infty.
$$

So our theorem says that the length of the longest increasing subsequence of a permutation behaves like the largest eigenvalue of a random matrix. Tracy and Widom also computed the distribution functions for 2nd, 3rd, . . . largest eigenvalues.

The natural question is this: Is there anything in the random permutation picture that behaves like the 2nd (or 3rd, 4th, ...) largest eigenvalue of a random GUE matrix? It can not be the 2nd largest increasing subsequence which is clearly distributed in the same way as the longest. To see what to do we must introduce some standard ideas from combinatorics/representation theory.

Let $\mu = (\mu_1, \ldots, \mu_l)$, $\mu_1 \geq \cdots \geq \mu_l \geq 1$ be a partition of N, $\sum_{i=1}^{l} \mu_i = N$. Associate to μ a frame (alternatively Ferrer's diagram, Young diagram). For example, for $N = 10 = 4 + 3 + 2 + 1$

$$\ydiagram{4,3,2,1} \qquad \begin{array}{l} \mu_1 = 4 \\ \mu_2 = 3 \\ \mu_3 = 2 \\ \mu_4 = 1 \end{array}$$

If we insert the numbers $1, \ldots, N$ bijectively into the boxes of the frame we obtain $N!$ Young tableaux. For example, for $N = 6 = 3 + 2 + 1$

$$\begin{ytableau} 1 & 4 & 5 \\ 6 & 3 \\ 2 \end{ytableau} \qquad N = 6$$

If we ensure that the rows and columns are increasing, for example,

$$\begin{ytableau} 1 & 4 & 5 \\ 6 & 3 \\ 2 \end{ytableau} \qquad\qquad \begin{ytableau} 1 & 3 & 5 \\ 2 & 4 \\ 6 \end{ytableau} \qquad\qquad N = 6$$

we obtain two *Standard Young Tableaux* with frame $\mu = (3, 2, 1)$.

Now there is a remarkable theorem of Robinson (1938) and Schensted (1961) which says that there is a bijection from S_N onto pairs of standard Young tableaux with the same frame

$$S_N \ni \pi \mapsto (P(\pi), Q(\pi))$$

where $\mathrm{frame}(P(\pi)) = \mathrm{frame}(Q(\pi)) = (\mu_1(\pi), \ldots, \mu_l(\pi))$. Furthermore (Schensted 1961)

$$l_N(\pi) = \mu_1(\pi)$$

Thus in this language, the width of a Standard Young Tableau behaves statistically (with Plancherel measure $\equiv$ push forward of uniform distribution on S_N) like the largest eigenvalue of a random GUE matrix.

What about $\mu_2(\pi)$, the number of boxes in the second row of frame$(P(\pi)) =$ frame$(Q(\pi))$? Well if we go back to the simulations of Odlyzko and Rains, we find that the computations indicate

$$\lim_{N\to\infty} \frac{\mathrm{Var}_N(\mu_2(\pi))}{N^{\frac{1}{3}}} = 0.545$$

$$\lim_{N\to\infty} \mathbb{E}_N \left(\frac{\mu_2(\pi) - 2\sqrt{N}}{N^{\frac{1}{6}}} \right) = -3.618$$

Well, these values agree, once again, with the variation and mean of the 2nd largest eigenvalue (suitably centered and scaled) of a random GUE matrix as computed by Tracy and Widom! In Baik, Deift, Johansson (1999), the authors conjectured that the length of the first k rows of a Standard Young Tableau, suitably centered and scaled, should behave statistically as $N \to \infty$ like the first k eigenvalues of a random GUE matrix. In a subsequent paper they proved the following result for the second row.

Let $u(x;s)$ solve PII $u'' = 2u^3 + xu$ with $u(x;s) \sim -\sqrt{s}\mathrm{Ai}(x)$ as $x \to +\infty$ $(0 < s \leq 1)$. Set $F(t;s) = e^{-\int_t^\infty (x-t)\, u^2(x;s)\, dx}$. Then $F^{(2)}(t) \equiv F^{(1)}(t) - \frac{\partial}{\partial s}\Big|_{s=1} F(t;s)$ is the Tracy–Widom distribution function for the second largest eigenvalue of a GUE matrix. Set $\chi_N^{(2)} = \frac{\mu_2 - 2\sqrt{N}}{N^{\frac{1}{6}}}$ where $\mu_2 = \mu_2(\pi)$ is the number of boxes in the second row of $P(\pi)$ (or $Q(\pi)$). Let $\chi^{(2)}$ be a random variable with distribution $F^{(2)}$. We have the following results (second paper of Baik, Deift, Johansson (1999))

Theorem 4. *As $N \to \infty$, $\chi_N^{(2)}$ converges to $\chi^{(2)}$ in distribution, i.e.* $\lim_{N\to\infty} \mathrm{Prob}\left(\chi_N^{(2)} \leq t\right) = F^{(2)}(t), \quad t \in \mathbb{R}$

Theorem 5. *For $m = 1, 2, 3, \ldots$, $\mathbb{E}_N\left(\left(\chi_N^{(2)}\right)^m\right) \to \mathbb{E}\left(\left(\chi^{(2)}\right)^m\right)$ as $N \to \infty$.*

The general conjecture for the k rows was then proved (in three independent ways!) by Okounkov (1999), Borodin, Olshanski, Okounkov (1999) and Johansson (1999). These authors proved convergence in distribution. Convergence of the moments for the k rows was proved subsequently by Baik, Deift and Rains (2001).

The results in Theorems 2 and 3 have generated a lot of activity in a variety of areas including the representation theory of large groups, polynuclear growth models, percolation models, random topologies on surfaces, digital boiling, amongst many others (see Adler, Baik, Borodin, Diaconis, Forrester, Johansson, van Moerbeke, Okounkov, Olshanski, Prähoffer, Spohn, Tracy, Widom, ..., 1999—). Early indications of the relationship between Standard Young Tableaux and Random Matrix Theory can be found, for example, in the work of Regev (1981).

How do we prove theorems 2, 3, 4, and 5? The key analytic fact is that the problems at hand can be rephrased as Riemann–Hilbert problems with (large) external parameters. This reformulation helps for the following reason: in the early 90's Xin Zhou and I introduced a steepest-descent type method for oscillating RH problems. This work was developed by a number of people and eventually (1997) placed in a very general form by Zhou, Venakides and myself. The method is a non-commutative, non-linear analog of the classical steepest descent method for scalar integrals and, I'll say more about this later. It is now recognized that there are many problems in many different areas that reduce to the asymptotic evaluation of some oscillatory RH problem.

Step 1 (Poissonization)

Set

$$\phi_n^{(1)}(\lambda) \equiv \sum_{N=0}^{\infty} \frac{e^{-\lambda}\lambda^N}{N!} q_{n,N}^{(1)}$$

By a rather elementary, but fortunate, de-Poissonization lemma due to Johansson (1998), the asymptotics of $q_{n,N}$ as $N \to \infty$ can be inferred from the asymptotics of $\phi_n^{(1)}(\lambda)$ for $\lambda \sim N$. So we must investigate the double scaling limit for $\phi_n^{(1)}(\lambda)$ for $1 \leq n \leq N \sim \lambda$. The critical region is around $n \sim 2\sqrt{\lambda} \sim 2\sqrt{N}$. Poissonization helps because of the following wonderful fact: there is an exact formula for $\phi_n^{(1)}(\lambda)$

$$\phi_n^{(1)}(\lambda) = e^{-\lambda} D_{n-1}\left(e^{2\sqrt{\lambda}\cos\theta}\right) = e^{-\lambda} D_{n-1}(\lambda)$$

D_{n-1} is the $n \times n$ Toeplitz determinant with weight function $f(e^{i\theta}) = e^{2\sqrt{\lambda}\cos\theta}$,

$$D_{n-1}(f) = \det\left(c_{kj} = c_{k-j} = \int_{-\pi}^{\pi} e^{-i(k-j)\theta} f(e^{i\theta}) \frac{d\theta}{2\pi}\right)_{0\leq k,j\leq n-1}$$

This formula was first found by Gessel (1990), but has since been discovered independently by many authors: Johansson, Diaconis–Shahshahani, Gessel–Weinstein–Wilf, Odylyzsko, Poonen, Widom, Wilf, Rains, Baik, Deift and Johansson (1999) also give a new proof.

Step 2 (Orthogonal polynomials)

Set

$$\kappa_n^2(\lambda) = \frac{D_{n-1}(\lambda)}{D_n(\lambda)}$$

The quantity $\kappa_n^2(\lambda)$ is the normalization coefficient for the nth orthonormal polynomial $p_n(z) = k_n(\lambda)(z^n + \cdots),\ k_n > 0$, with respect to $f(e^{i\theta})\frac{d\theta}{2\pi}$ where $f = e^{2\sqrt{\lambda}\cos\theta}$, that is

$$\int_{-\pi}^{\pi} \overline{p_n(e^{i\theta})} p_m(e^{i\theta}) f(e^{i\theta}) \frac{d\theta}{2\pi} = \delta_{n,m}$$

Using the Strong–Szegö limit theorem we have

$$\log \phi_n(\lambda) = \sum_{k=n}^{\infty} \log \kappa_k^2(\lambda)$$

This is the main formula: to control $\phi_n(\lambda)$ we must control $\kappa_k(\lambda)$ for $k \geq n$ in the case where $n, \lambda \to \infty$.

Step 3 (Riemann–Hilbert Problem)

A critical component in our approach is the following extension of the 1991 result of Fokas, Its, Kitaev (compare Lecture 2): Let Σ denote the unit circle oriented counterclockwise. Let $Y(z;\, k+1,\, \lambda) = (Y_{ij}(z;\, k+1,\, \lambda))_{1 \leq i,j \leq 2}$ be the 2×2 matrix function satisfying the following RHP

$$\begin{cases} \bullet\ Y(z; k+1, \lambda) \text{ is analytic for } z \in \mathbb{C} \setminus \Sigma \\ \bullet\ Y_+(z; k+1, \lambda) = Y_-(z; k+1, \lambda) \begin{pmatrix} 1 & \frac{1}{z^{k+1}} e^{\sqrt{\lambda}(z+\frac{1}{z})} \\ 0 & 1 \end{pmatrix} \quad \text{for} \quad z \in \Sigma \\ \bullet\ Y(z; k+1, \lambda) \begin{pmatrix} z^{-(k+1)} & 0 \\ 0 & z^{k+1} \end{pmatrix} = I + O(\frac{1}{z}) \quad \text{as} \quad z \to \infty \end{cases}$$

Then Y is unique and

$$\kappa_k^2(\lambda) = -Y_{21}(z=0; k+1, \lambda).$$

Also

$$\pi_{k+1} = Y_{11}$$

where $\pi_k(z) = \frac{p_k(z)}{\kappa_k} = z^k + \dots$. So to evaluate $\kappa_k^2(\lambda)$, $k \geq n$, and hence $\phi_n^{(1)}(\lambda)$, we must be able to control the solution Y of the above RHP in the limit when the two parameters $k, \sqrt{\lambda}$ in $z^{-(k+1)} e^{\sqrt{\lambda}(z+\frac{1}{z})}$ are large.

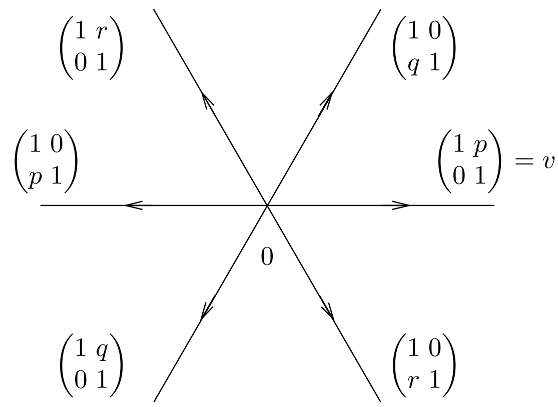
This is precisely the situation that can be analyzed by the non-linear steepest descent method. The calculations are very similar to those that arose in the work of Deift, McLaughlin, Kriechaber, Venakides, and Zhou described in the earlier lectures on the proof of universality for various statistical quantities in random matrix theory.

Step 4 (Painlevé Theory)

Where does PII come into the picture? Now there is a wonderful way (going back essentially to Ablowitz and Segur, then Flaschka–Newell, and Jimbo–Miwa–Ueno) to solve PII

$$u'' = 2u^3 + xu.$$

Let p, q, r be three numbers satisfying $p+q+r+pqr = 0$. Consider the oriented contour Σ consisting of six rays with v attached as follows



For fixed x let Ψ solve the following RHP:

$$\begin{cases} \bullet \ \Psi \text{ analytic in } \mathbb{C} \setminus \Sigma \\ \bullet \ \Psi_+ = \Psi_-(z) v \text{ on } \Sigma \\ \bullet \ m \equiv \Psi \begin{pmatrix} e^{i(\frac{4z^3}{3}+xz)} & 0 \\ 0 & e^{-i(\frac{4z^3}{3}+xz)} \end{pmatrix} \to I \quad \text{as} \quad z \to \infty \end{cases}$$

Then if we expand

$$m = I + \frac{m_1(x)}{z} + O\left(\frac{1}{z^2}\right)$$

one can show that

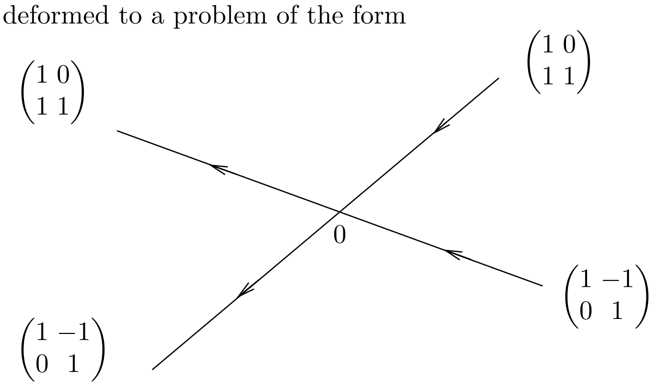
$$u(x) \equiv 2i\,(m_1(x))_{12}$$

solves PII.

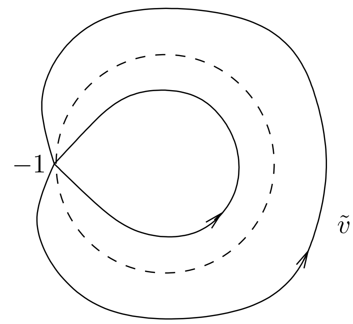
Of particular interest is the case where

$$p = -q = 1, \ r = 0$$

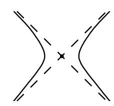
which can be deformed to a problem of the form



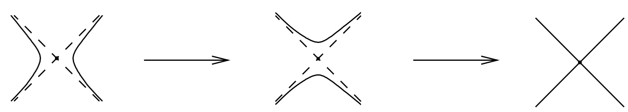
Now it turns out, at least for $\frac{2\sqrt{\lambda}}{k+1} \leq 1$ that the RHP for Ψ is equivalent to a RHP on a contour of the following form



Away from $z = -1$, for $\frac{\sqrt{\lambda}}{k+1} \sim 1$, $\tilde{v} \to I$. The leading contribution then comes from $z = -1$



Rotating by $90°$ we get



which turns out to be precisely the RHP with the right jump matrices for PII ($p = -q = 1$, $r = 0$) as above!

This is the way in which PII comes into the picture.

Free Probability Theory and Random Matrices

R. Speicher*

Department of Mathematics and Statistics
Queen's University, Kingston
ON K7L 3N6, Canada
speicher@mast.queensu.ca

Summary. Free probability theory originated in the context of operator algebras, however, one of the main features of that theory is its connection with random matrices. Indeed, free probability can be considered as *the* theory providing concepts and notations, without relying on random matrices, for dealing with the limit $N \to \infty$ of $N \times N$-random matrices.

One of the basic approaches to free probability, on which I will concentrate in this lecture, is of a combinatorial nature and centers around so-called free cumulants. In the spirit of the above these arise as the combinatorics (in leading order) of $N \times N$-random matrices in the limit $N = \infty$. These free cumulants are multi-linear functionals which are defined in combinatorial terms by a formula involving non-crossing partitions.

I will present the basic definitions and properties of non-crossing partitions and free cumulants and outline its relations with freeness and random matrices. As examples, I will consider the problems of calculating the eigenvalue distribution of the sum of randomly rotated matrices and of the compression (upper left corner) of a randomly rotated matrix.

1 Random matrices and freeness

Free probability theory, due to Voiculescu, originated in the context of operator algebras, however, one of the main features of that theory is its connection with random matrices. Indeed, free probability can be considered as *the* theory providing concepts and notations, without relying on random matrices, for dealing with the limit $N \to \infty$ of $N \times N$-random matrices.

Let us consider a sequence $(A_N)_{N\in\mathbb{N}}$ of selfadjoint $N \times N$-random matrices A_N. In which sense can we talk about the limit of these matrices? Of course, such a limit does not exist as a $\infty \times \infty$-matrix and there is no convergence in the usual topologies connected to operators. What converges and survives in the limit are the moments of the random matrices.

* Research supported by a grant of NSERC, Canada

A. Borodin et al.: LNM 1817, A.M. Vershik (Ed.), pp. 53–73, 2003.

To talk about moments we need in addition to the random matrices also a state. This is given in a canonical way by the averaged trace: Let tr_N be the normalized trace on $N \times N$-matrices, i.e. for $A = (a_{ij})_{i,j=1}^N$ we have

$$\mathrm{tr}_N(A) := \frac{1}{N} \sum_{i=1}^{N} a_{ii}.$$

In the same way, we get the averaged trace $\mathrm{tr}_N \otimes \mathbb{E}$ for $N \times N$-random matrices, i.e. for $A = (a_{ij}(\omega))_{i,j=1}^N$ (where the entries a_{ij} are random variables on some probability space Ω equipped with a probability measure P) we have

$$\mathrm{tr}_N \otimes \mathbb{E}(A) := \frac{1}{N} \sum_{i=1}^{N} \int_{\Omega} a_{ii}(\omega) dP(\omega).$$

Given these states $\mathrm{tr}_N \otimes E$, we can now talk about the k-th moment $\mathrm{tr}_N \otimes \mathbb{E}(A_N^k)$ of our random matrix A_N, and it is well known that for nice random matrix ensembles these moments converge for $N \to \infty$. So let us denote by α_k the limit of the k-th moment,

$$\lim_{N\to\infty} \mathrm{tr}_N \otimes E(A_N^k) =: \alpha_k.$$

Thus we can say that the limit $N = \infty$ consists exactly of the collection of all these moments α_k. But instead of talking about a collection of numbers α_k we prefer to identify these numbers as moments of some variable A. Abstractly it is no problem to find such an A, we just take a free algebra $\mathcal{A}$ with generator A and define a state φ on $\mathcal{A}$ by setting

$$\varphi(A^k) := \alpha_k.$$

Of course, nothing deep has happened, this is just a shift in language, but it provides us with a more conceptual way of looking at the limit of our random matrices. Now we can say that our random matrices A_N converge to the variable A in distribution (which just means that the moments of A_N converge to the moments of A). We will denote this by $A_N \to A$. Note that the nature of the limit $N = \infty$ is quite different from the case of finite N. In the latter case the A_N live in classical probability spaces of $N \times N$-random matrices, whereas the $N = \infty$ limit object A is not of a classical nature any more, but lives in a 'non-classical probability space' given by some algebra $\mathcal{A}$ and a state φ.

1.1 Remark

One should note that for a selfadjoint operator $A = A^*$, the collection of moments (or, equivalently, the state φ corresponding to these moments) corresponds also to a probability measure μ_A on the real line, determined by

$$\varphi(A^k) = \int_{\mathbb{R}} t^k d\mu_A(t).$$

(We can ignore the problem of non-uniqueness of this moment problem, because usually our operators A are bounded, which ensures uniqueness.) In particular, for a selfadjoint $N \times N$-matrix $A = A^*$ this measure is given by the eigenvalue distribution of A, i.e. it puts mass $1/N$ on each of the eigenvalues of A (counted with multiplicity):

$$\mu_A = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i},$$

where $\lambda_1, \dots, \lambda_N$ are the eigenvalues of A. In the same way, for a random matrix A, μ_A is given by the averaged eigenvalue distribution of A. Thus, moments of random matrices with respect to the averaged trace $\mathrm{tr}_N \otimes \mathbb{E}$ contain exactly that type of information in which one is usually interested when dealing with random matrices.

1.2 Example

Let us consider the basic example of random matrix theory, expressed in our new language. Let G_N be the usual selfadjoint Gaussian $N \times N$-random matrices (i.e., entries above the diagonal are independently and normally distributed). Then the famous theorem of Wigner can be stated in our language in the form that

$$G_N \to s, \qquad \text{where } s \text{ is a semi-circular variable,}$$

where semi-circular just means that the measure μ_s is given by the semi-circular distribution (or, equivalently, the even moments of the even variable s are given by the Catalan numbers).

Up to now, nothing crucial has happened, we have just shifted a bit the usual way of looking on things. A new and crucial concept, however, appears if we go over from the case of one variable to the case of more variables. Of course, again joint moments are the surviving quantities in multi-matrix models (even if it is now not so clear any more how to prove this convergence in concrete models) and we can adapt our way of looking on things to this situation by making the following definition.

1.3 Definition

Consider $N \times N$-random matrices $A_N^{(1)}, \dots, A_N^{(m)}$ and variables $A_1, \dots, A_m$ (living in some abstract algebra $\mathcal{A}$ equipped with a state φ). We say that

$$(A_N^{(1)}, \dots, A_N^{(m)}) \to (A_1, \dots, A_m) \qquad \text{in distribution,}$$

if

$$\lim_{N\to\infty} \mathrm{tr}_N \otimes \mathbb{E}[A_N^{(i_1)} \cdots A_N^{(i_k)}] = \varphi(A_{i_1} \cdots A_{i_k})$$

for all choices of k, $1 \le i_1, \dots, i_k \le m$.

1.4 Remark

The $A_1, \dots, A_m$ arising in the limit of random matrices are a priori abstract elements in some algebra $\mathcal{A}$, but it is good to know that in many cases they can also be concretely realized by some kind of creation and annihilation operators on a full Fock space. Indeed, free probability theory was introduced by Voiculescu for investigating the structure of special operator algebras generated by these type of operators. In the beginning, free probability had nothing to do with random matrices.

1.5 Example

Let us now consider the example of two independent Gaussian random matrices $G_N^{(1)}, G_N^{(2)}$ (i.e., each of them is a selfadjoint Gaussian random matrix and all entries of $G_N^{(1)}$ are independent from all entries of $G_N^{(2)}$). Then one knows that all joint moments converge, and we can say that $(G_N^{(1)}, G_N^{(2)}) \to (s_1, s_2)$, where Wigner tells us that both s_1 and s_2 are semi-circular. The question is: What is the relation between s_1 and s_2? Does the independence between $G_N^{(1)}$ and $G_N^{(2)}$ survive in some form also in the limit? The answer is yes and is provided by a basic theorem of Voiculescu which says that s_1 and s_2 are **free** in the following sense.

1.6 Definition

Let $\mathcal{A}$ be a unital algebra and $\varphi : \mathcal{A} \to \mathbb{C}$ a linear functional on $\mathcal{A}$, which is unital, i.e., $\varphi(1) = 1$. Then $a_1, \dots, a_m \in \mathcal{A}$ are called *free* (with respect to φ) if

$$\varphi[p_1(a_{i(1)}) \cdots p_k(a_{i(k)})] = 0$$

whenever

- $p_1, \dots, p_k$ are polynomials in one variable;
- $i(1) \ne i(2) \ne i(3) \ne \cdots \ne i(k)$ (only neigbouring elements are required to be distinct);
- $\varphi[p_j(a_{i(j)})] = 0$ for all $j = 1, \dots, k$.

1.7 Remark

Note that the definition of freeness can be considered as a way of organizing the information about all joint moments of free variables in a systematic and conceptual way. Indeed, the above definition allows to calculate mixed moments of free variables in terms of moments of the single variables. For example, if a, b are free, then the definiton of freeness requires that

$$\varphi[(a - \varphi(a) \cdot 1)(b - \varphi(b) \cdot 1)] = 0,$$

which implies that

$$\varphi(ab) = \varphi(a) \cdot \varphi(b) \qquad \text{if } a, b \text{ are free.}$$

In the same way,

$$\varphi[(a - \varphi(a) \cdot 1)(b - \varphi(b) \cdot 1)(a - \varphi(a) \cdot 1)(b - \varphi(b) \cdot 1)] = 0$$

leads finally to

$$\varphi(abab) = \varphi(aa) \cdot \varphi(b) \cdot \varphi(b) + \varphi(a) \cdot \varphi(a) \cdot \varphi(bb) - \varphi(a) \cdot \varphi(b) \cdot \varphi(a) \cdot \varphi(b).$$

Analogously, all mixed moments can (at least in principle) be calculated by reducing them to alternating products of centered variables as in the definition of freeness.

Thus the statement 's_1, s_2 are free and each of them is semicircular' determines all joint moments in s_1 and s_2.

Formulating our knowledge about the joint moments of s_1 and s_2 in this peculiar way might look not very illuminating on first sight, but it will turn out that recognizing this notion of freeness as the organizing principle for the collection of moments adds a new perspective on the limit of random matrices.

In particular, we are now in the context of non-commutative probability theory which consists mainly of the doctrine that one should use notations and ideas from classical probability theory in order to understand problems about non-commutative algebras.

Free probability theory can be described as that part of non-commutative probability theory where the notion of 'freeness' plays an essential role. Furthermore, according to the basic philosophy of Voiculescu this notion of freeness should be considered (and investigated) in analogy with the classical notion of 'independence'—both freeness and independence prescribe special relations between joint moments of some variables. (Of course, both cases correspond to very special, but also very fundamental situations.)

One of the most interesting features of freeness is that this concept appears in at least two totally different mathematical situations. Originally it was introduced by Voiculescu in the context of operator algebras, later it turned out that there is also some relation, as described above, with random matrices. This gives a non-trivial connection between these two different fields. For example, modelling operator algebras by random matrices has led to some of the most impressive results about operator algebras in the last years.

Furthermore, apart from the concrete manifestations of freeness via random matrices or operator algebras, there exist also an abstract probabilistic theory of freeness, which shows the depth of this concept and which I want to address in the following.

2 Combinatorial approach to free probability: non-crossing partitions and free cumulants

'Freeness' of random variables is defined in terms of mixed moments; namely the defining property is that very special moments (alternating and centered ones) have to vanish. This requirement is not easy to handle in concrete calculations. Thus we will present here another approach to freeness, more combinatorial in nature, which puts the main emphasis on so called 'free cumulants'. These are some polynomials in the moments which behave much better with respect to freeness than the moments. The nomenclature comes from classical probability theory where corresponding objects are also well known and are usually called 'cumulants' or 'semi-invariants'. There exists a combinatorial description of these classical cumulants, which depends on partitions of sets. In the same way, free cumulants can also be described combinatorially, the only difference to the classical case is that one has to replace all partitions by so called 'non-crossing partitions'.

This combinatorial description of freeness is due to me [8, 9] (see also [3]); in a series of joint papers with A. Nica [4, 5, 6] it was pursued very far and yielded a lot of new results in free probability theory. For more information on other aspects of freeness, in particular the original analytical approach of Voiculescu, one should consult the papers [10, 11, 13], the collection of various articles [13], or the monographs [14, 1]

2.1 Definitions

A *partition* of the set $S := \{1, \dots, n\}$ is a decomposition

$$\pi = \{V_1, \dots, V_r\}$$

of S into disjoint and non-empty sets V_i, i.e.

$$V_i \cap V_j = \emptyset \quad (i, j = 1, \dots, r; \quad i \neq j) \qquad \text{and} \qquad S = \bigcup_{i=1}^{r} V_i .$$

We call the V_i the *blocks* of π.
For $1 \leq p, q \leq n$ we write

$$p \sim_\pi q \qquad \text{if } p \text{ and } q \text{ belong to the same block of } \pi.$$

A partition π is called *non-crossing* if the following does not occur: There exist $1 \leq p_1 < q_1 < p_2 < q_2 \leq n$ with

$$p_1 \sim_\pi p_2 \not\sim_\pi q_1 \sim_\pi q_2.$$

The set of all non-crossing partitions of $\{1, \ldots, n\}$ is denoted by *NC(n)*.

Non-crossing partitions were introduced by Kreweras [2] in a purely combinatorial context without any reference to probability theory.

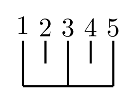
2.2 Examples

We will also use a graphical notation for our partitions; the term 'non-crossing' will become evident in such a notation. Let

$$S = \{1, 2, 3, 4, 5\}.$$

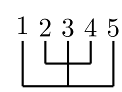
Then the partition

$$\pi = \{(1, 3, 5), (2), (4)\} \qquad \hat{=}$$



is non-crossing, whereas

$$\pi = \{(1, 3, 5), (2, 4)\} \qquad \hat{=}$$



is crossing.

2.3 Remarks

1) In an analogous way, non-crossing partitions $NC(S)$ can be defined for any linearly ordered set S; of course, we have

$$NC(S_1) \cong NC(S_2) \qquad \text{if} \qquad \#S_1 = \#S_2.$$

2) In most cases the following recursive description of non-crossing partitions is of great use: a partition π ist non-crossing if and only if at least one block $V \in \pi$ is an interval and $\pi \backslash V$ is non-crossing; i.e. $V \in \pi$ has the form

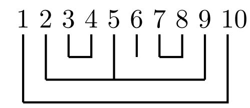
$$V = (k, k+1, \ldots, k+p) \qquad \text{for some } 1 \leq k \leq n \text{ and } p \geq 0,\ k+p \leq n$$

and we have

$$\pi \backslash V \in NC(1, \ldots, k-1, k+p+1, \ldots, n) \cong NC(n-(p+1)).$$

Example: The partition

$$\{(1,10),(2,5,9),(3,4),(6),(7,8)\} \quad \hat{=}$$



can, by successive removal of intervals, be reduced to

$$\{(1,10),(2,5,9)\}\hat{=}\{(1,5),(2,3,4)\}$$

and finally to

$$\{(1,5)\}\hat{=}\{(1,2)\}.$$

3) By writing a partition π in the form $\pi = \{V_1, \dots, V_r\}$ we will always assume that the elements within each block V_i are ordered in increasing order.

2.4 Definition

Let $(\mathcal{A}, \varphi)$ be a probability space, i.e. $\mathcal{A}$ is a unital algebra and $\varphi : \mathcal{A} \to \mathbb{C}$ is a unital linear functional. We define the *(free or non-crossing) cumulants*

$$k_n : \mathcal{A}^n \to \mathbb{C} \qquad (n \in \mathbb{N})$$

(indirectly) by the following system of equations:

$$\varphi(a_1 \dots a_n) = \sum_{\pi \in NC(n)} k_\pi[a_1, \dots, a_n] \qquad (a_1, \dots, a_n \in \mathcal{A}),$$

where k_π denotes a product of cumulants according to the block structure of π:

$$k_\pi[a_1, \dots, a_n] := k_{V_1}[a_1, \dots, a_n] \dots k_{V_r}[a_1, \dots, a_n] \qquad \text{for } \pi = \{V_1, \dots, V_r\} \in NC(n)$$

and

$$k_V[a_1, \dots, a_n] := k_{\#V}(a_{v_1}, \dots, a_{v_l}) \qquad \text{for} \qquad V = (v_1, \dots, v_l).$$

2.5 Remarks and Examples

1) Note: the above equations have the form

$$\varphi(a_1 \dots a_n) = k_n(a_1, \dots, a_n) + \text{smaller order terms}$$

and thus they can be resolved for the $k_n(a_1, \dots, a_n)$ in a unique way.
2) Examples:

- $n = 1$
$$\varphi(a_1) = k_{|}[a_1] = k_1(a_1),$$
thus
$$k_1(a_1) = \varphi(a_1).$$
- $n = 2$
$$\begin{aligned}\varphi(a_1a_2) &= k_{\sqcup}[a_1, a_2] + k_{|\,|}[a_1, a_2]\\ &= k_2(a_1, a_2) + k_1(a_1)k_1(a_2),\end{aligned}$$
thus
$$k_2(a_1, a_2) = \varphi(a_1a_2) - \varphi(a_1)\varphi(a_2).$$
- $n = 3$
$$\begin{aligned}\varphi(a_1a_2a_3) &= k_{\sqcup\!\sqcup}[a_1, a_2, a_3] + k_{|\,\sqcup}[a_1, a_2, a_3] + k_{\sqcup\,|}[a_1, a_2, a_3]\\ &\quad + k_{\sqcup\!\!\!\!\!\;\shortmid\;}[a_1, a_2, a_3] + k_{|\,|\,|}[a_1, a_2, a_3]\\ &= k_3(a_1, a_2, a_3) + k_1(a_1)k_2(a_2, a_3) + k_2(a_1, a_2)k_1(a_3)\\ &\quad + k_2(a_1, a_3)k_1(a_2) + k_1(a_1)k_1(a_2)k_1(a_3),\end{aligned}$$
and thus
$$\begin{aligned}k_3(a_1, a_2, a_3) &= \varphi(a_1a_2a_3) - \varphi(a_1)\varphi(a_2a_3) - \varphi(a_1a_3)\varphi(a_2)\\ &\quad - \varphi(a_1a_2)\varphi(a_3) + 2\varphi(a_1)\varphi(a_2)\varphi(a_3).\end{aligned}$$

3) For $n = 4$ we consider the special case where all $\varphi(a_i) = 0$. Then we have
$$k_4(a_1, a_2, a_3, a_4) = \varphi(a_1a_2a_3a_4) - \varphi(a_1a_2)\varphi(a_3a_4) - \varphi(a_1a_4)\varphi(a_2a_3).$$

4) The k_n are multi-linear functionals in their n arguments.

The meaning of the concept ‘cumulants’ for freeness is shown by the following theorem.

2.6 Theorem

Consider $a_1, \dots, a_m \in \mathcal{A}$. Then the following two statements are equivalent:

i) $a_1, \dots, a_m$ are free.
ii) mixed cumulants vanish, i.e.: We have for all $n \geq 2$ and for all $1 \leq i(1), \dots, i(n) \leq m$:
$$k_n(a_{i(1)}, \dots, a_{i(n)}) = 0,$$
whenever there exist $1 \leq p, q \leq n$ with $i(p) \neq i(q)$.

2.7 Remarks

1) An example of the vanishing of mixed cumulants is that for a, b free we have $k_3(a, a, b) = 0$, which, by the definition of k_3 just means that

$$\varphi(aab) - \varphi(a)\varphi(ab) - \varphi(aa)\varphi(b) - \varphi(ab)\varphi(a) + 2\varphi(a)\varphi(a)\varphi(b) = 0.$$

This vanishing of mixed cumulants in free variables is of course just a reorganization of the information about joint moments of free variables – but in a form which is much more useful for many applications.
2) The above characterization of freeness in terms of cumulants is the translation of the definition of freeness in terms of moments – by using the relation between moments and cumulants from Definition 2.4. One should note that in contrast to the characterization in terms of moments we do not require that $i(1) \neq i(2) \neq \cdots \neq i(n)$ or $\varphi(a_i) = 0$. (That's exactly the main part of the proof of that theorem: to show that on the level of cumulants the assumption 'centered' is not needed and that 'alternating' can be weakened to 'mixed'.) Hence the characterization of freeness in terms of cumulants is much easier to use in concrete calculations.

3 Addition of free variables

3.1 Notation

For a random variable $a \in \mathcal{A}$ we put

$$k_n^a := k_n(a, \ldots, a)$$

and call $(k_n^a)_{n \geq 1}$ the *(free) cumulants of* a.

Our main theorem on the vanishing of mixed cumulants in free variables specialises in this one-dimensional case to the linearity of the cumulants.

3.2 Proposition

Let a and b be free. Then we have

$$k_n^{a+b} = k_n^a + k_n^b \qquad \text{for all } n \geq 1.$$

Proof. We have

$$\begin{aligned} k_n^{a+b} &= k_n(a+b, \ldots, a+b) \\ &= k_n(a, \ldots, a) + k_n(b, \ldots, b) \\ &= k_n^a + k_n^b, \end{aligned}$$

because cumulants which have both a and b as arguments vanish by Theorem 2.6.

Thus, the addition of free random variables is easy to describe on the level of cumulants; the cumulants are additive under this operation. It remains to make the connection between moments and cumulants as explicit as possible. On a combinatorial level, our definition specializes in the one-dimensional case to the following relation.

3.3 Proposition

Let $(m_n)_{n\geq 1}$ and $(k_n)_{n\geq 1}$ be the moments and free cumulants, respectively, of some random variable. The connection between these two sequences of numbers is given by

$$m_n = \sum_{\pi \in NC(n)} k_\pi,$$

where

$$k_\pi := k_{\#V_1} \cdots k_{\#V_r} \qquad \text{for} \qquad \pi = \{V_1, \dots, V_r\}.$$

Example. For $n = 3$ we have

$$\begin{aligned} m_3 &= k_{\sqcup\!\sqcup} + k_{|\,\sqcup} + k_{\sqcup\,|} + k_{\overset{\,\cdot}{\sqcup}} + k_{|\,|\,|} \\ &= k_3 + 3k_1k_2 + k_1^3. \end{aligned}$$

For concrete calculations, however, one would prefer to have a more analytical description of the relation between moments and cumulants. This can be achieved by translating the above relation to corresponding formal power series.

3.4 Theorem

Let $(m_n)_{n\geq 1}$ and $(k_n)_{n\geq 1}$ be two sequences of complex numbers and consider the corresponding formal power series

$$\begin{aligned} M(z) &:= 1 + \sum_{n=1}^{\infty} m_n z^n, \\ C(z) &:= 1 + \sum_{n=1}^{\infty} k_n z^n. \end{aligned}$$

Then the following three statements are equivalent:

(i) We have for all $n \in \mathbb{N}$

$$m_n = \sum_{\pi \in NC(n)} k_\pi = \sum_{\pi = \{V_1, \dots, V_r\} \in NC(n)} k_{\#V_1} \dots k_{\#V_r}.$$

(ii) We have for all $n \in \mathbb{N}$ (where we put $m_0 := 1$)

$$m_n = \sum_{s=1}^{n} \sum_{\substack{i_1,\dots,i_s \in \{0,1,\dots,n-s\} \\ i_1+\dots+i_s=n-s}} k_s m_{i_1} \dots m_{i_s}.$$

(iii) We have

$$C[zM(z)] = M(z).$$

Proof. We rewrite the sum

$$m_n = \sum_{\pi \in NC(n)} k_\pi$$

in the way that we fix the first block V_1 of π (i.e. that block which contains the element 1) and sum over all possibilities for the other blocks; in the end we sum over V_1:

$$m_n = \sum_{s=1}^{n} \sum_{V_1 \text{ with } \#V_1 = s} \sum_{\substack{\pi \in NC(n) \\ \text{where } \pi = \{V_1, \dots\}}} k_\pi.$$

If

$$V_1 = (v_1 = 1, v_2, \dots, v_s),$$

then $\pi = \{V_1, \dots\} \in NC(n)$ can only connect elements lying between some v_k and v_{k+1}, i.e. $\pi = \{V_1, V_2, \dots, V_r\}$ such that we have for all $j = 2, \dots, r$: there exists a k with $v_k < V_j < v_{k+1}$. There we put

$$v_{s+1} := n + 1.$$

Hence such a π decomposes as

$$\pi = V_1 \cup \tilde{\pi}_1 \cup \dots \cup \tilde{\pi}_s,$$

where

$$\tilde{\pi}_j \text{ is a non-crossing partition of } \{v_j + 1, v_j + 2, \dots, v_{j+1} - 1\}.$$

For such π we have

$$k_\pi = k_{\#V_1} k_{\tilde{\pi}_1} \dots k_{\tilde{\pi}_s} = k_s k_{\tilde{\pi}_1} \dots k_{\tilde{\pi}_s},$$

and thus we obtain

$$
\begin{aligned}
m_n &= \sum_{s=1}^{n} \sum_{1=v_1<v_2<\cdots<v_s\le n} \sum_{\substack{\pi=V_1\cup\tilde\pi_1\cup\cdots\cup\tilde\pi_s\\ \tilde\pi_j\in NC(v_j+1,\dots,v_{j+1}-1)}} k_s k_{\tilde\pi_1}\dots k_{\tilde\pi_s}\\
&= \sum_{s=1}^{n} k_s \sum_{1=v_1<v_2<\cdots<v_s\le n} \Big(\sum_{\tilde\pi_1\in NC(v_1+1,\dots,v_2-1)} k_{\tilde\pi_1}\Big)\dots\Big(\sum_{\tilde\pi_s\in NC(v_s+1,\dots,n)} k_{\tilde\pi_s}\Big)\\
&= \sum_{s=1}^{n} k_s \sum_{1=v_1<v_2<\cdots<v_s\le n} m_{v_2-v_1-1}\dots m_{n-v_s}\\
&= \sum_{s=1}^{n} \sum_{\substack{i_1,\dots,i_s\in\{0,1,\dots,n-s\}\\ i_1+\cdots+i_s+s=n}} k_s m_{i_1}\dots m_{i_s} \qquad (i_k := v_{k+1}-v_k-1).
\end{aligned}
$$

This yields the implication (i) $\Longrightarrow$ (ii).
We can now rewrite (ii) in terms of the corresponding formal power series in the following way (where we put $m_0 := k_0 := 1$):

$$
\begin{aligned}
M(z) &= 1+\sum_{n=1}^{\infty} z^n m_n\\
&= 1+\sum_{n=1}^{\infty}\sum_{s=1}^{n}\sum_{\substack{i_1,\dots,i_s\in\{0,1,\dots,n-s\}\\ i_1+\cdots+i_s=n-s}} k_s z^s m_{i_1} z^{i_1}\dots m_{i_s} z^{i_s}\\
&= 1+\sum_{s=1}^{\infty} k_s z^s \Big(\sum_{i=0}^{\infty} m_i z^i\Big)^s\\
&= C[zM(z)].
\end{aligned}
$$

This yields (iii).
Since (iii) describes uniquely a fixed relation between the numbers $(k_n)_{n\ge1}$ and the numbers $(m_n)_{n\ge1}$, this has to be the relation (i).

If we rewrite the above relation between the formal power series in terms of the Cauchy transform

$$G(z) := \sum_{n=0}^{\infty} \frac{m_n}{z^{n+1}}$$

and the R-transform

$$R(z) := \sum_{n=0}^{\infty} k_{n+1} z^n,$$

then we obtain Voiculescu's formula.

3.5 Corollary

The relation between the Cauchy transform $G(z)$ and the R-transform $R(z)$ of a random variable is given by

$$G[R(z) + \frac{1}{z}] = z.$$

Proof. We just have to note that the formal power series $M(z)$ and $C(z)$ from Theorem 3.4 and $G(z)$, $R(z)$, and $K(z) = R(z) + \frac{1}{z}$ are related by:

$$G(z) = \frac{1}{z} M(\frac{1}{z})$$

and

$$C(z) = 1 + zR(z) = zK(z), \qquad \text{thus} \qquad K(z) = \frac{C(z)}{z}.$$

This gives

$$K[G(z)] = \frac{1}{G(z)} C[G(z)] = \frac{1}{G(z)} C[\frac{1}{z} M(\frac{1}{z})] = \frac{1}{G(z)} M(\frac{1}{z}) = z,$$

thus $K[G(z)] = z$ and hence also

$$G[R(z) + \frac{1}{z}] = G[K(z)] = z.$$

3.6 Free convolution

The above results give us a quite effective tool for calculating the distribution of the sum $a+b$ of free variables from the distribution of a and the distribution of b. In analogy with the usual convolution (which corresponds to the sum of independent random variables) we introduce the notion $\boxplus$ of *free convolution* as operation on probability measures by

$$\mu_{a+b} = \mu_a \boxplus \mu_b \qquad \text{if } a, b \text{ are free.}$$

Then we know that free cumulants and the R-transform linearize this free convolution.

In particular, we also have the free convolution powers

$$\mu^{\boxplus r} := \mu \boxplus \cdots \boxplus \mu \qquad (r\text{-times})$$

of μ, which are in terms of cumulants characterized by

$$k_n(\mu^{\boxplus r}) = r \cdot k_n(\mu).$$

If we are given free variables a and b and we want to calculate the distribution of $a + b$, then we calculate the R-transforms R_a and R_b and get thus by the linearization property the R-transform of $a + b$,

$$R_{a+b} = R_a + R_b.$$

It remains to extract the distribution out of this. From the R-transform we can get the Cauchy transform G_{a+b} by Corollary 3.5, and then we use the classical Stietljes inversion formula for extracting the distribution from this. In general, the relation between R-transform and Cauchy transform might lead to equations which have no analytic solutions, however, in many concrete case these equations can be solved. For example, if we put $\mu = \frac{1}{2}(\delta_0 + \delta_1)$, then the above machinery shows that the distribution $\mu \boxplus \mu$ is given by the arcsine law.

3.7 Application: random sum of matrices [7]

Fix two (compactly supported) probability measures μ and ν on the real line and consider deterministic (e.g., diagonal) $N \times N$-matrices A_N and C_N, whose eigenvalue distribution converges, for $N \to \infty$, towards the given measures. To put it in the language introduced in Section 1, we assume that

$$A_N \to a \qquad \text{with} \qquad \mu_a = \mu$$

and

$$C_N \to c \qquad \text{with} \qquad \mu_c = \nu.$$

Now we rotate C_N against A_N randomly by replacing C_N by

$$B_N := U_N C_N U_N^*,$$

where U_N is a random Haar unitary matrix from the ensemble of unitary $N \times N$-matrices equipped with the Haar measure. Of course, the eigenvalue distribution of B_N is the same as the one of C_N, however, any definite relation between the eigenspaces of A_N and the eigenspaces of C_N has now been destroyed. A_N and B_N are in the limit $N \to \infty$ generic realizations of the given eigenvalue distributions μ and ν. The question which we want to address is: What is the eigenvalue distribution of the sum $A_N + B_N$ in the limit $N \to \infty$, i.e. what can we say about

$$A_N + B_N \to ?$$

A version of the theorem of Voiculescu about the connection between random matrices and freeness tells us that A_N and B_N become free in the limit $N \to \infty$, i.e. it yields that

$$(A_N, B_N) \to (a, b) \qquad \text{with } \mu_a = \mu,\ \mu_b = \nu, \text{ and } a, b \text{ free.}$$

Thus we know that the eigenvalue distribution of $A_N + B_N$ converges towards the distribution of $a + b$ where a and b are free. But the distribution of $a + b$ can be calculated with our tools from free probability in a very effective and systematic way by using the R-transform machinery. For example, if we take the generic sum of two projections of trace 1/2, (i.e., $\mu = \nu = \frac{1}{2}(\delta_0 + \delta_1)$, then our example from above shows us that the distribution of

$$\begin{pmatrix} 1 &&&&& \\ & 0 &&&& \\ && 1 &&& \\ &&& 0 && \\ &&&& 1 & \\ &&&&& \ddots \end{pmatrix} + U_N \begin{pmatrix} 1 &&&&& \\ & 0 &&&& \\ && 1 &&& \\ &&& 0 && \\ &&&& 1 & \\ &&&&& \ddots \end{pmatrix} U_N^*,$$

is, in the limit $N \to \infty$, given by the arcsine law.

4 Multiplication of free variables

Finally, to show that our description of freeness in terms of cumulants has also a significance apart from dealing with additive free convolution, we will apply it to the problem of the product of free random variables: Consider $a_1, \dots, a_n, b_1, \dots, b_n$ such that $\{a_1, \dots, a_n\}$ and $\{b_1, \dots, b_n\}$ are free. We want to express the distribution of the random variables $a_1 b_1, \dots, a_n b_n$ in terms of the distribution of the a's and of the b's.

4.1 Notation

1) Analogously to k_π we define for

$$\pi = \{V_1, \dots, V_r\} \in NC(n)$$

the expression

$$\varphi_\pi[a_1 \dots, a_n] := \varphi_{V_1}[a_1, \dots, a_n] \dots \varphi_{V_r}[a_1, \dots, a_n],$$

where

$$\varphi_V[a_1, \dots, a_n] := \varphi(a_{v_1} \dots a_{v_l}) \qquad \text{for} \qquad V = (v_1, \dots, v_l).$$

Examples:

$$\begin{aligned} \varphi_{\sqcup\!\sqcup}[a_1, a_2, a_3] &= \varphi(a_1 a_2 a_3) \\ \varphi_{|\,\sqcup}[a_1, a_2, a_3] &= \varphi(a_1)\varphi(a_2 a_3) \\ \varphi_{\sqcup\,|}[a_1, a_2, a_3] &= \varphi(a_1 a_2)\varphi(a_3) \\ \varphi_{\sqcup\!\!\!\!\sqcup\!\!\!'}[a_1, a_2, a_3] &= \varphi(a_1 a_3)\varphi(a_2) \\ \varphi_{|\,|\,|}[a_1, a_2, a_3] &= \varphi(a_1)\varphi(a_2)\varphi(a_3) \end{aligned}$$

2) Let $\sigma, \pi \in NC(n)$. Then we write

$$\sigma \leq \pi$$

if each block of σ is contained as a whole in some block of π, i.e. σ can be obtained out of π by refinement of the block structure.
Example:

$$\{(1),(2,4),(3),(5,6)\} \leq \{(1,5,6),(2,3,4)\}$$

With these notations we can generalize the relation

$$\varphi(a_1 \dots a_n) = \sum_{\pi \in NC(n)} k_\pi[a_1,\dots,a_n]$$

in the following way.

$$\varphi_\sigma[a_1,\dots,a_n] = \sum_{\substack{\pi \in NC(n) \\ \pi \leq \sigma}} k_\pi[a_1,\dots,a_n].$$

Consider now

$$\{a_1,\dots,a_n\},\{b_1,\dots,b_n\} \qquad \text{free.}$$

We want to express alternating moments in a and b in terms of moments of a and moments of b. We have

$$\varphi(a_1 b_1 a_2 b_2 \dots a_n b_n) = \sum_{\pi \in NC(2n)} k_\pi[a_1,b_1,a_2,b_2,\dots,a_n,b_n].$$

Since the a's are free from the b's, Theorem 2.6 tells us that only such π contribute to the sum whose blocks do not connect a's with b's. But this means that such a π has to decompose as

$$\pi = \pi_1 \cup \pi_2 \qquad \text{where} \quad \begin{aligned} &\pi_1 \in NC(1,3,5,\dots,2n-1) \\ &\pi_2 \in NC(2,4,6,\dots,2n). \end{aligned}$$

Thus we have

$$\begin{aligned}
&\varphi(a_1 b_1 a_2 b_2 \dots a_n b_n) \\
&\quad = \sum_{\substack{\pi_1 \in NC(odd), \pi_2 \in NC(even) \\ \pi_1 \cup \pi_2 \in NC(2n)}} k_{\pi_1}[a_1,a_2,\dots,a_n] \cdot k_{\pi_2}[b_1,b_2,\dots,b_n] \\
&\quad = \sum_{\pi_1 \in NC(odd)} \Big(k_{\pi_1}[a_1,a_2,\dots,a_n] \cdot \sum_{\substack{\pi_2 \in NC(even) \\ \pi_1 \cup \pi_2 \in NC(2n)}} k_{\pi_2}[b_1,b_2,\dots,b_n] \Big).
\end{aligned}$$

Note now that for a fixed π_1 there exists a maximal element σ with the property $\pi_1 \cup \sigma \in NC(2n)$ and that the second sum is running over all $\pi_2 \leq \sigma$.

4.2 Definition

Let $\pi \in NC(n)$ be a non-crossing partition of the numbers $1, \ldots, n$. Introduce additional numbers $\bar{1}, \ldots, \bar{n}$, with alternating order between the old and the new ones, i.e. we order them in the way

$$1\bar{1}2\bar{2}\ldots n\bar{n}.$$

We define the *complement* $K(\pi)$ of π as the maximal $\sigma \in NC(\bar{1}, \ldots, \bar{n})$ with the property

$$\pi \cup \sigma \in NC(1, \bar{1}, \ldots, n, \bar{n}).$$

If we present the partition π graphically by connecting the blocks in $1, \ldots, n$, then σ is given by connecting as much as possible the numbers $\bar{1}, \ldots, \bar{n}$ without getting crossings among themselves and with π.

(This natural notation of the complement of a non-crossing partition is also due to Kreweras [2]. Note that there is no analogue of this for the case of all partitions.)

With this definition we can continue our above calculation as follows:

$$\begin{aligned}
&\varphi(a_1 b_1 a_2 b_2 \ldots a_n b_n) \\
&\qquad = \sum_{\pi_1 \in NC(n)} \Big(k_{\pi_1}[a_1, a_2, \ldots, a_n] \cdot \sum_{\substack{\pi_2 \in NC(n) \\ \pi_2 \leq K(\pi_1)}} k_{\pi_2}[b_1, b_2, \ldots, b_n] \Big) \\
&\qquad = \sum_{\pi_1 \in NC(n)} k_{\pi_1}[a_1, a_2, \ldots, a_n] \cdot \varphi_{K(\pi_1)}[b_1, b_2, \ldots, b_n].
\end{aligned}$$

Thus we have proved the following result.

4.3 Theorem

Consider

$$\{a_1, \ldots, a_n\}, \{b_1, \ldots, b_n\} \qquad \text{free.}$$

Then we have

$$\varphi(a_1 b_1 a_2 b_2 \ldots a_n b_n) = \sum_{\pi \in NC(n)} k_\pi[a_1, a_2, \ldots, a_n] \cdot \varphi_{K(\pi)}[b_1, b_2, \ldots, b_n].$$

Similar to the additive case one can translate this combinatorial description of the product of free variables in an analytic way in terms of the so-called S-transfrom. However, this is more complicated as in the case of the R-transform and we will not address this problem here. Instead, we want to show that the above combinatorial description of the product of free variables can lead to quite explicit (and unexpected) results without running through an analytic reformulation. Such a result is given in our final application to the problem of the compression of a random matrix.

4.4 Application: Compression of random matrix

Consider, as in Section 3.7, a sequence of deterministic $N \times N$-matrices C_N with prescribed eigenvalue distribution μ in the limit $N \to \infty$ and consider the randomly rotated version $A_N := U_N C_N U_N^*$ of this matrix. The question we want to address is the following: Can we calculate the eigenvalue distribution of upper left corners of the matrices A_N. Formally, we get these corners by compressing A_N with projections of the form

$$P_N := \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

where $\mathrm{tr}_N(P_N) \to \alpha$ for some fixed α with $0 < \alpha \leq 1$. Thus we ask for the eigenvalue distribution of $P_N A_N P_N$ in the limit $N \to \infty$. (However, we have to calculate this in the compressed space, throwing away the bunch of trivial zero eigenvalues outside of the non-trivial corner of $P_N A_N P_N$, i.e., we consider $P_N A_N P_N$ not as $N \times N$-matrix, but as $\alpha N \times \alpha N$-matrix.)

Now, again by Voiculescu's theorem about asymptotic freeness of random matrices, we know that

$$(A_N, P_N) \to (a, p),$$

where a has the prescribed distribution μ, p is a projection of trace α and a and p are free. Thus, the answer for our question on the distribution of corners in randomly rotated matrices is provided by calculating the distribution of pap in the compressed space, i.e. by calculating the renormalized moments

$$\frac{1}{\alpha}\varphi[(pap)^n],$$

which is, by the trace property of φ and the projection property $p^2 = p$, the same as

$$\frac{1}{\alpha}\varphi[(ap)^n].$$

This fits now exactly in the above frame of calculating the moments of products of free variables, in the special case where the second variable is a projection of trace α. Using $p^k = p$ for all $k \geq 1$ and $\varphi(p) = \alpha$ gives

$$\varphi_{K(\pi)}[p, p, \dots, p] = \varphi(p \dots p)\varphi(p \dots p) \cdots = \alpha^{|K(\pi)|},$$

where $|K(\pi)|$ denotes the number of blocks of $K(\pi)$. We can express this number of blocks also in terms of π, since we always have the relation

$$|\pi| + |K(\pi)| = n + 1.$$

Thus we can continue our calculation of Theorem 4.3 in this case as

$$\begin{aligned}\frac{1}{\alpha}\varphi[(ap)^n] &= \frac{1}{\alpha}\sum_{\pi \in NC(n)} k_\pi[a, \dots, a]\alpha^{n+1-|\pi|} \\ &= \sum_{\pi \in NC(n)} \frac{1}{\alpha^{|\pi|}} k_\pi[\alpha a, \dots, \alpha a],\end{aligned}$$

which shows that

$$k_n(pap, \dots, pap) = \frac{1}{\alpha} k_n(\alpha a, \dots, \alpha a)$$

for all n. By our remarks on the additive free convolution, this gives the surprising result that the renormalized distribution of *pap* is given by

$$\mu_{pap} = \mu_{\alpha a}^{\boxplus 1/\alpha}.$$

In particular, for $\alpha = 1/2$, we have

$$\mu_{pap} = \mu_{1/2a}^{\boxplus 2} = \mu_{1/2a} \boxplus \mu_{1/2a}.$$

This means that the distribution of the upper left corner of size 1/2 of a randomly rotated matrix is, apart from rescaling with the factor 1/2, the same as the distribution of the sum of the considered matrix and another randomly rotated copy of itself. E.g., if we take the example $\mu = \frac{1}{2}(\delta_0 + \delta_1)$, then the corner of size 1/2 of such a randomly rotated projection has as eigenvalue distribution the arcsine law.

References

1. Hiai, F., Petz, D.: The semicircle law, free random variables and entropy. Mathematical Surveys and Monographs, Vol. 77, AMS (2000)
2. Kreweras, G.: Sur les partitions non-croisees d'un cycle. Discrete Math., **1**, 333–350 (1972)

3. Nica, A.: R-transforms of free joint distributions, and non-crossing partitions. J. Funct. Anal., **135**, 271–296 (1996)
4. Nica, A., Speicher, R.: On the multiplication of free n-tuples of non-commutative random variables (with an appendix by D. Voiculescu). Amer. J. Math., **118**. 799–837 (1996)
5. Nica, A., Speicher, R.: R-diagonal pairs—a common approach to Haar unitaries and circular elements. In: Voiculescu, D.-V. (ed) Free Probability Theory, AMS, 149–188 (1997)
6. Nica, A., Speicher, R.: Commutators of free random variables. Duke Math. J., **92**, 553–592 (1998)
7. Speicher, R.: Free convolution and the random sum of matrices. RIMS, **29**, 731–744 (1993)
8. Speicher, R.: Multiplicative functions on the lattice of non-crossing partitions and free convolution. Math. Ann., **298**, 611–628 (1994)
9. Speicher, R.: Combinatorial theory of the free product with amalgamation and operator-valued free probability theory. Memoirs of the AMS, **627** (1998)
10. Voiculescu, D.: Addition of certain non-commuting random variables. J. Funct. Anal., **66**, 323–346 (1986)
11. Voiculescu, D.: Limit laws for random matrices and free products. Invent. math., **104**, 201–220 (1991)
12. Voiculescu, D.: Free probability theory: random matrices and von Neumann algebras. Proceedings of the ICM 1994, Birkhäuser, 227–241 (1995)
13. Voiculescu, D. (ed): Free Probability Theory. Fields Institute Communications, vol. 12, AMS (1997)
14. Voiculescu, D. V., Dykema, K. J., Nica, A.: Free Random Variables. CRM Monograph Series, vol. 1, AMS (1993)

A Noncommutative Version of Kerov's Gaussian Limit for the Plancherel Measure of the Symmetric Group

Akihito Hora

Faculty of Environmental Science and Technology
Okayama University
Okayama 700-8530, Japan
hora@ems.okayama-u.ac.jp

Summary. We give a noncommutative extension of Kerov's central limit theorem for irreducible characters of the symmetric group with respect to the Plancherel measure [S.Kerov: C. R. Acad. Sci. Paris 316 (1993)] in the framework of algebraic probability theory. For adjacency operators associated with the cycle classes, we consider their decomposition according to the length function on the Cayley graph of the symmetric group. We develop a certain noncommutative central limit theorem for them, in which the limit picture is described by creation and annihilation operators on an analogue of the Fock space equipped with an orthonormal basis labelled by Young diagrams. The limit Gaussian measure in Kerov's theorem appears as the spectral distribution of the field operators in our setting.

1 Introduction

Let $S(n)$ be the symmetric group of degree n and $\hat{S}(n)$ the set of the equivalence classes of its irreducible representations. χ_ρ^λ denotes the value of the irreducible character χ^λ $(\lambda \in \hat{S}(n))$ on each element of conjugacy class C_ρ in $S(n)$ corresponding to partition ρ of n. The Plancherel measure of $S(n)$ is the probability on $\hat{S}(n)$ defined by $M_n(\lambda) = \dim^2 \lambda / n!$ where $\dim \lambda$ is the dimension of $\lambda \in \hat{S}(n)$. Considering the value $\chi_{(k)}^\lambda$ on a cycle of length k for each λ, we set

$$\phi_k(\lambda) = n^{k/2} \chi_{(k)}^\lambda / \dim \lambda \qquad (\lambda \in \hat{S}(n),\ 2 \le k \le n).$$

In [13], Kerov showed the following asymptotic behaviour of joint distributions of ϕ_k's with respect to M_n.

A. Borodin et al.: LNM 1817, A.M. Vershik (Ed.), pp. 77–88, 2003.

Kerov's Theorem. *For* $\forall m \geq 2$ *and* $\forall x_2, \ldots, x_m \in \mathbf{R}$*, we have*

$$\lim_{n\to\infty} M_n(\{\lambda \in \hat{S}(n) | \phi_k(\lambda) \leq x_k \ for\ 2 \leq k \leq m\}) = \prod_{k=2}^{m} \frac{1}{\sqrt{2\pi k}} \int_{-\infty}^{x_k} e^{-y^2/2k} dy \ . \quad (1)$$

In other words, ϕ_k's corresponding to cycles of distinct length are asymptotically independent random variables with respect to the Plancherel measure whose limit distributions are Gaussian.

The aim of this note is to give a noncommutative extension (the meaning will be clarified later) of the above theorem by using the framework of algebraic probability theory. We will develop a certain central limit theorem for creation and annihilation operators acting on an analogous object to the Fock space and derive Kerov's theorem as its corollary.

We begin with rephrasing (1) in terms of adjacency operators on the symmetric group. Let L denote the left regular representation of a finite group G with identity element e. For conjugacy class C of G, the operator on $\ell^2(G)$ defined by $A_C = \sum_{x\in C} L(x)$ is called an adjacency operator. The Plancherel measure of G is defined by $M(\alpha) = \dim^2 \alpha / \sharp G$ $(\alpha \in \hat{G})$. δ_x denotes the delta function supported by $x \in G$. For arbitrary conjugacy classes $C_1, \ldots, C_p$ in G, we have

$$\langle \delta_e, A_{C_1} \cdots A_{C_p} \delta_e \rangle_{\ell^2(G)} = \frac{1}{\sharp G} tr(A_{C_1} \cdots A_{C_p}) = \int_{\hat{G}} \frac{(\sharp C_1)\chi^{\alpha}_{C_1}}{\dim \alpha} \cdots \frac{(\sharp C_p)\chi^{\alpha}_{C_p}}{\dim \alpha} M(d\alpha) \ .$$

In the case of $G = S(n)$, since the conjugacy class of the cycles of length k has cardinality

$$\sharp C_{(k)} = n(n-1)\cdots(n-k+1)/k \sim n^k/k \quad (\text{as } n \to \infty),$$

(1) is equivalent to

$$\lim_{n\to\infty} \langle \delta_e, \Big(\frac{A_{(2)}}{\sqrt{\sharp C_{(2)}}}\Big)^{p_2} \cdots \Big(\frac{A_{(m)}}{\sqrt{\sharp C_{(m)}}}\Big)^{p_m} \delta_e \rangle_{\ell^2(S(n))} = \prod_{k=2}^{m} \int_{-\infty}^{\infty} x^{p_k} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \quad (2)$$

$(\forall m \geq 2,\ \forall p_2, \ldots, p_m \in \mathbf{N})$.

In [9], we obtained an extension of this result to arbitrary conjugacy classes in the symmetric group. Let $\mathcal{Y}$ denote the set of the Young diagrams which does not contain a row consisting of a single box. $|\rho|$ denotes the number of boxes contained in $\rho \in \mathcal{Y}$. The nontrivial conjugacy classes in the infinite symmetric group $S(\infty)$ [resp. $S(n)$] are parametrised by $\mathcal{Y}$ [resp. $\{\rho \in \mathcal{Y} | \, |\rho| \leq n\}$]. $k_j(\rho)$ denotes the number of the rows of length j in $\rho \in \mathcal{Y}$ and then we use the expression $\rho = (2^{k_2(\rho)} \cdots j^{k_j(\rho)} \cdots)$. Let $H_k(x)$ be the monic Hermite polynomial obeying the recurrence formula

$$xH_k(x) = H_{k+1}(x) + kH_{k-1}(x) \ (k \geq 1), \quad H_0(x) = 1, \quad H_1(x) = x \ .$$

Theorem 0 ([9]). *For $\forall m \in \mathbf{N}$, $\forall \rho_1, \ldots, \rho_m \in \mathcal{Y}$ and $\forall r_1, \ldots, r_m \in \mathbf{N}$, we have*

$$\lim_{n\to\infty} \langle \delta_e, \Big(\frac{A_{\rho_1}}{\sqrt{\sharp C_{\rho_1}}} \Big)^{r_1} \cdots \Big(\frac{A_{\rho_m}}{\sqrt{\sharp C_{\rho_m}}} \Big)^{r_m} \delta_e \rangle_{\ell^2(S(n))}$$
$$= \prod_{j\geq 2} \int_{-\infty}^{\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} \Big(\frac{H_{k_j(\rho_1)}(x)}{\sqrt{k_j(\rho_1)!}} \Big)^{r_1} \cdots \Big(\frac{H_{k_j(\rho_m)}(x)}{\sqrt{k_j(\rho_m)!}} \Big)^{r_m} dx \ . \quad (3)$$

(*Note that $k_j(\rho_1) = \cdots = k_j(\rho_m) = 0$ for $j > \max(|\rho_1|, \ldots, |\rho_m|)$.*)

For $\rho \in \mathcal{Y}$ and $n \geq |\rho|$, adjacency operator A_ρ has mean $\langle \delta_e, A_\rho \delta_e \rangle_{\ell^2(S(n))} = 0$ and variance $\langle \delta_e, A_\rho^2 \delta_e \rangle_{\ell^2(S(n))} = \sharp C_\rho$ with respect to the vacuum state. The scaling limit (3) is hence referred to as a sort of central limit theorem. Such a central limit theorem involving adjacency operators on discrete groups or regular graphs enjoys considerable literature. Even rectricted to the ones on the symmetric group and directly related to this note, we refer to [13], [3], [9] etc. A lot of studies have been done around the free group and the concept of freeness due to Voiculescu, though we omit here concrete citations. Besides Cayley graphs of discrete groups, we can mention e.g. [10], [8] and [7] dealing with distance-regular graphs. As Kerov already suggested in [13], such a kind of central limit theorem has much to do with the limit shape of Young diagrams ([14], [15]) and their fluctuation around it. In [12], Ivanov and Olshanski have done a detailed study succeeding [13]. It includes a different proof of (3) from ours in [9]. They give an insight also into the asymptotic formula due to Biane ([4]) for the irreducible characters of the symmetric group in terms of free cumulants.

In this note we will lift the convergence of distributions expressed by (2) to the level of the convergence of operators. Roughly speaking, we will construct limit operators on a Hilbert space in a natural way and show the convergence of mixed matrix elements of normalised adjacency operators to those of limit operators. Our procedure goes on as follows. See §2 for precise definitions.

(I) According to a length function on the Cayley graph of $S(n)$, each adjacency operator is canonically decomposed: $A_{(j)} = A_{(j)}^+ + A_{(j)}^-$, $(A_{(j)}^-)^* = A_{(j)}^+$, $(A_{(j)}^+)^* = A_{(j)}^-$.

(II) Introduce an analogue of the Fock space containing an orthonormal basis labelled by the Young diagrams in $\mathcal{Y}$, creation operators B_j^+ and annihilation operators B_j^- acting on it.

(III) Show that any mixed product of $A_{(j)}^+$'s and $A_{(j)}^-$'s converges to the same type product of B_j^+'s and B_j^-'s in the meaning of convergence of all matrix elements.

(IV) Observe that $(B_j^+, B_j^-)_{j=2,3,\dots}$ are independent with respect to the Fock vacuum state and that the spectral distribution of $B_j^+ + B_j^-$ is Gaussian.

Decomposing an operator under imaging creation and annihilation (of particles) and considering their scaling limit are popular methods often used in quantum probability. In [6], [8] and [7], an approach to central limit theorems through such a "quantum decomposition" has been studied on the basis of a useful connection found by Accardi and Bożejko ([1]) between the so-called interacting Fock spaces and orthogonal polynomials. In these articles, we see several concrete limit distributions besides Gaussian ones. As a merit of this approach, we can mention that taking limits *after* decomposing an operator makes the involved combinatorial argument much more transparent. We will see that the limit operators $B_j^\pm$ are easier to handle than $A_{(j)}^\pm$ from a combinatorial viewpoint and that the spectral structure of $B_j^\pm$ is clear. In fact, our argument in this note relies on elementary combinatorics without using representation theory of the symmetric group or theory of symmetric functions. Central limit theorems as convergence of matrix elements enjoy an application to some limit theorems with respect to non-vacuum states treated e.g. in [5] and [11].

In §2 we give a precise description of the procedures (I)–(IV) and state our main results. Their proofs are given in §3.

Acknowledgement I would like to express my deep appreciation to Professor A. M. Vershik and the organizers for giving me a nice opportunity to contribute the School and the Proceedings. Also many thanks are due to Professor N. Obata for stimulating discussions, which are certainly reflected in this note. The present research activity is partially supported by Grant-in-Aid for Scientific Research (C), JSPS (No. 13640175).

2 Main Results

2.1 Young diagrams and length function

Let us consider the Cayley graph of $S(n)$ with all the transpositions as its generators. $[x]$ denoting the number of the cycles (including trivial ones) appearing in the cycle decomposition of $x \in S(n)$, the length of a geodesic connecting x with e is equal to $n - [x]$. For a given diagram $\rho \in \mathcal{Y}$, we take n such that $|\rho| \leq n$ and consider $x \in S(n)$ yielding the cycle decomposition of type ρ. Set $l(\rho) = |\rho| - \sharp(\text{rows in } \rho)$. Then, $l(\rho) = n - [x]$. We thus refer to $l : \mathcal{Y} \longrightarrow \mathbf{N}$ as the length function on diagrams. This length function leads us to a stratified arrangement of the diagrams in $\mathcal{Y}$. Namely, we arrange ρ's having the common value $l(\rho)$ in the same stratum. In the usual Young graph, each stratum consists of the diagrams with a common number of boxes (including 1-box rows), where each diagram corresponds to an irreducible representation of the symmetric group of the same degree with the number of boxes. Comparing these two, we see that our arrangement of the diagrams in $\mathcal{Y}$ is made from the Young graph by the following operation; for every diagram in the Young graph, add a copy of the first (longest) column to the left side of the diagram. (The edges of the original Young graph are now irrelevant.)

Let us define quantum decomposition of adjacency operator $A_{(j)}$ associated with conjugacy class $C_{(j)}$ of the j-cycles in $S(n)$. The degree n is fixed until the end of this subsection. For $s \in C_{(j)}$ ($j \in \{2, \ldots, n\}$), we define operators s^+ and s^- acting on $\ell^2(S(n))$ by

$$s^+\delta_x = \begin{cases} \delta_{sx} & \text{if } [sx] < [x] \\ (1/2)\delta_{sx} & \text{if } [sx] = [x] \\ 0 & \text{if } [sx] > [x], \end{cases} \qquad s^-\delta_x = \begin{cases} \delta_{sx} & \text{if } [sx] > [x] \\ (1/2)\delta_{sx} & \text{if } [sx] = [x] \\ 0 & \text{if } [sx] < [x]. \end{cases}$$

Clearly $L(s) = s^+ + s^-$. Then we set

$$A^+_{(j)} = \sum_{s \in C_{(j)}} s^+ , \qquad A^-_{(j)} = \sum_{s \in C_{(j)}} s^- . \tag{4}$$

Clearly $A_{(j)} = A^+_{(j)} + A^-_{(j)}$, $(A^+_{(j)})^* = A^-_{(j)}$ and $(A^-_{(j)})^* = A^+_{(j)}$. For $\rho \in \mathcal{Y}$ such that $|\rho| \leq n$, we set

$$\xi_\rho = \sum_{x \in C_\rho} \delta_x , \qquad \Phi(\rho) = \frac{1}{\sqrt{\sharp C_\rho}} \xi_\rho . \tag{5}$$

For convenience we set $\Phi(\emptyset) = \delta_e$ also for empty diagram $\emptyset$. $\{\Phi(\rho) | \rho \in \mathcal{Y} \cup \{\emptyset\}, |\rho| \leq n\}$ is an orthonormal system in $\ell^2(S(n))$. In what follows,

$$\Gamma(S(n)) = \Phi(\emptyset) \oplus \bigoplus_{\rho \in \mathcal{Y}, |\rho| \leq n} \Phi(\rho)$$

plays a role of a finite-dimensional Fock space.

2.2 Noncommutative central limit theorem

In order to describe the limit of normalised adjacency operators, we introduce an analogue of the Fock space with creation and annihilation operators as follows. A unit vector $\Psi(\rho)$ is assigned to every $\rho \in \mathcal{Y}$. Also to empty diagram $\emptyset$, for which $l(\emptyset) = 0$ by definition, is assigned a unit vector $\Psi(\emptyset)$. Set Hilbert space Γ as the completion of

$$\Gamma_0 = \Psi(\emptyset) \oplus \bigoplus_{\rho \in \mathcal{Y}} \Psi(\rho) = \bigoplus_{j=0}^{\infty} \bigoplus_{\rho : l(\rho) = j} \Psi(\rho) ,$$

in which $\Psi(\rho)$'s form an orthonormal basis. $\rho \sqcup (j) \in \mathcal{Y}$ [resp. $\rho \setminus (j) \in \mathcal{Y} \cup \{\emptyset\}$] denotes the diagram we get by adding [resp. removing] a j-row to [resp. from]

$\rho \in \mathcal{Y}$. If ρ does not have a j-row (i.e. $k_j(\rho) = 0$), we set $\Psi(\rho \setminus (j)) = 0$. For $j \in \{2, 3, \ldots\}$, we introduce operators B_j^+ (creation) and B_j^- (annihilation) densely defined on Γ by

$$B_j^+ \Psi(\rho) = \sqrt{k_j(\rho) + 1}\, \Psi(\rho \sqcup (j)) \ , \quad B_j^- \Psi(\rho) = \sqrt{k_j(\rho)}\, \Psi(\rho \setminus (j)) \tag{6}$$

($\rho \in \mathcal{Y} \cup \{\emptyset\}$). These are mutually adjoint: $(B_j^+)^* = B_j^-$, $(B_j^-)^* = B_j^+$ and satisfy the commutation relations on Γ_0:

$$\begin{aligned} &[B_i^-, B_j^-] = [B_i^+, B_j^+] = [B_i^-, B_j^+] = 0 \quad (i \neq j) \\ &[B_j^-, B_j^+] = I \ . \end{aligned}$$

These commutation relations coincide with those for creation and annihilation operators on the Boson Fock space.

Theorem 1. *For* $\forall \tau, \rho \in \mathcal{Y} \cup \{\emptyset\}$, $\forall m \in \mathbf{N}$, $\forall \epsilon_1, \ldots, \epsilon_m \in \{+, -\}$ *and* $\forall j_1, \ldots, j_m \in \{2, 3, \ldots\}$, *we have*

$$\lim_{n \to \infty} \Big\langle \Phi(\tau), \frac{A_{(j_1)}^{\epsilon_1}}{\sqrt{\sharp C_{(j_1)}}} \cdots \frac{A_{(j_m)}^{\epsilon_m}}{\sqrt{\sharp C_{(j_m)}}} \Phi(\rho) \Big\rangle_{\ell^2(S(n))} = \Big\langle \Psi(\tau), B_{j_1}^{\epsilon_1} \cdots B_{j_m}^{\epsilon_m} \Psi(\rho) \Big\rangle_{\Gamma} \ . \tag{7}$$

Theorem 2. *For* $\forall m \in \mathbf{N}$, $\forall Q_2, \ldots, Q_m$*: noncommutative polynomials in two variables, and* $\forall \tau, \rho \in \mathcal{Y} \cup \{\emptyset\}$ *having no* j*-rows for* j *such that* Q_j *is nontrivial (i.e. not a constant), we have*

$$\begin{aligned} &\langle \Psi(\tau), Q_2(B_2^+, B_2^-) \cdots Q_m(B_m^+, B_m^-) \Psi(\rho) \rangle_{\Gamma} \\ &= \langle \Psi(\tau), Q_2(B_2^+, B_2^-) \Psi(\rho) \rangle_{\Gamma} \cdots \langle \Psi(\tau), Q_m(B_m^+, B_m^-) \Psi(\rho) \rangle_{\Gamma} \ . \end{aligned} \tag{8}$$

In particular, applying Theorem 1 and Theorem 2 to the case of $\tau = \rho = \emptyset$, we get the following limit of normalised mixed moments for $A_{(j)} = A_{(j)}^+ + A_{(j)}^-$.

Corollary. *For* $\forall m \in \mathbf{N}$ *and* $\forall p_2, \ldots, p_m \in \mathbf{N}$, *we have*

$$\begin{aligned} &\lim_{n \to \infty} \Big\langle \delta_e, \Big(\frac{A_{(2)}}{\sqrt{\sharp C_{(2)}}} \Big)^{p_2} \cdots \Big(\frac{A_{(m)}}{\sqrt{\sharp C_{(m)}}} \Big)^{p_m} \delta_e \Big\rangle_{\ell^2(S(n))} \\ &= \langle \Psi(\emptyset), (B_2^+ + B_2^-)^{p_2} \Psi(\emptyset) \rangle_{\Gamma} \cdots \langle \Psi(\emptyset), (B_m^+ + B_m^-)^{p_m} \Psi(\emptyset) \rangle_{\Gamma} \ . \end{aligned} \tag{9}$$

Definition (6) implies that, restricted on the subspace $\Psi(\emptyset) \oplus \overline{\bigoplus_{n=1}^{\infty}} \Psi((j^n))$ of Γ, B_j^+ and B_j^- agree with creation and annihilation operators respectively on the 1-mode Boson Fock space. Hence a field operator $B_j^+ + B_j^-$ obeys the standard Gaussian distribution with respect to the vacuum state $\langle \Psi(\emptyset), \ \cdot \ \Psi(\emptyset) \rangle_{\Gamma}$. Namely, (9) is equivalent to (2). We thus reconstructed (1) in Kerov's theorem.

Remark. It is useful to describe a limit picture by creation and annihilation operators through quantum decomposition of an adjacency operator. In the final step to read out the distribution of a field operator, we need in general not only the Boson Fock space but also so-called interacting Fock spaces. The key is an isometry due to Accardi and Bożejko from a 1-mode interacting Fock space $\mathcal{F}$ to $L^2(\mathbf{R}, \mu)$. It is based on the correspondence between creation and annihilation operators on $\mathcal{F}$ and a recurrence formula for the orthogonal polynomials with respect to probability measure μ (having all the moments). See [1] for details. In [7], we used the interacting Fock spaces associated with the Laguerre and Meixner polynomials as well as the usual Boson and free Fock spaces (which are associated with the Hermite and Chebychev polynomials respectively) to describe scaling limits of adjacency operators on certain regular graphs.

3 Proof of Main Results

3.1 Action of $A^{\pm}_{(j)}$ on $\Gamma(S(n))$

In this subsection we fix $n \in \mathbf{N}$ and treat Young diagrams $\rho \in \mathcal{Y} \cup \{\emptyset\}$ such that $|\rho| \leq n$. Adjacency operator A_ρ is expressed by the matrix $[\tilde{A}_\rho]_{x,y \in S(n)}$:

$$(\tilde{A}_\rho)_{x,y} = \begin{cases} 1 & \text{if } \ x^{-1}y \in C_\rho \\ 0 & \text{if } \ x^{-1}y \notin C_\rho \end{cases}$$

with respect to the canonical basis $\{\delta_x | x \in S(n)\}$ in $\ell^2(S(n))$. Set

$$p^{\sigma}_{\tau\rho} = \sharp\{z \in S(n) | x^{-1}z \in C_\tau, \ z^{-1}y \in C_\rho\}$$

for $x^{-1}y \in C_\sigma$. Note that $p^{\sigma}_{\tau\rho}$ does not depend on the choice of x, y whenever $x^{-1}y \in C_\sigma$. $p^{\sigma}_{\tau\rho}$ is called an intersection number of the group association scheme of $S(n)$ (see [2]) and satisfies

$$A_\tau A_\rho = \sum_{\sigma} p^{\sigma}_{\tau\rho} A_\sigma \ .$$

$A^{\pm}_{(j)}$ act on the vectors in (5) as follows.

Lemma 1. *For $j \in \{2, \ldots, n\}$ and $|\rho| \leq n$, we have*

$$A^{\pm}_{(j)}\xi_\rho = \sum_{i=1}^{j-1} \sum_{\sigma : l(\sigma) = l(\rho) \pm i, |\sigma| \leq n} p^{\sigma}_{(j)\rho}\xi_\sigma + \frac{1}{2} \sum_{\sigma : l(\sigma) = l(\rho), |\sigma| \leq n} p^{\sigma}_{(j)\rho}\xi_\sigma \ .$$

(In the '$-$' case, i actually runs over $1, \ldots, (j-1) \wedge l(\rho)$.)

Proof. We have only to show the '+' case because the situation is the same. From (4) and (5) we have

$$\begin{aligned} A^+_{(j)}\xi_\rho &= \sum_{x\in C_\rho}\sum_{s\in C_{(j)}} s^+\delta_x \\ &= \sum_{x\in C_\rho}\sum_{s\in C_{(j)}:[sx]<[x]} \delta_{sx} + \frac{1}{2}\sum_{x\in C_\rho}\sum_{s\in C_{(j)}:[sx]=[x]} \delta_{sx} \end{aligned}$$

Here the first term is equal to

$$\begin{aligned} &\sum_{y\in S(n)} \sharp\{(s,x)|x\in C_\rho, s\in C_{(j)}, [sx]<[x], y=sx\}\delta_y \\ &= \sum_{i=1}^{\infty}\sum_{\sigma:|\sigma|\le n}\sum_{y\in C_\sigma} \sharp\{(s,x)|x\in C_\rho, s\in C_{(j)}, l(\sigma)=l(\rho)+i, y=sx\}\delta_y \\ &= \sum_{i=1}^{j-1}\sum_{\sigma:l(\sigma)=l(\rho)+i,|\sigma|\le n} p^\sigma_{(j)\rho}\xi_\sigma \ . \end{aligned}$$

Dealing similarly with the second term, we get the desired expression.

We rewrite Lemma 1 as

$$\begin{aligned} \frac{A^\pm_{(j)}}{\sqrt{\sharp C_{(j)}}}\Phi(\rho) &= \sum_{i=1}^{j-1}\sum_{\sigma:l(\sigma)=l(\rho)\pm i,|\sigma|\le n} p^\sigma_{(j)\rho}\sqrt{\frac{\sharp C_\sigma}{\sharp C_{(j)}\sharp C_\rho}}\Phi(\sigma) \\ &\quad + \frac{1}{2}\sum_{\sigma:l(\sigma)=l(\rho),|\sigma|\le n} p^\sigma_{(j)\rho}\sqrt{\frac{\sharp C_\sigma}{\sharp C_{(j)}\sharp C_\rho}}\Phi(\sigma) \end{aligned} \tag{10}$$

(i runs over $1,\ldots,(j-1)\wedge l(\rho)$ in the '$-$' case).

3.2 Estimate for $p^\sigma_{(j)\rho}$ as $n\to\infty$

This subsection is devoted to an estimate for the coefficients of (10) as $n\to\infty$.

Let $j\in\{2,3,\ldots\}$, $\sigma,\rho\in\mathcal{Y}\cup\{\emptyset\}$ and $|\sigma|\le n, |\rho|\le n$. Take $x\in C_\rho$ in $S(n)$. Note $x\in C_\rho$ in $S(m)$ for $\forall m\ge n$. The definition of an intersection number yields

$$p^\sigma_{(j)\rho} = \sharp\{s\in C_{(j)}|s^{-1}x\in C_\rho\} \qquad (x\in C_\rho \text{ is fixed}).$$

In order to estimate $p^\sigma_{(j)\rho}$, we thus have to count possible expressions of the form

$$s^{-1}x = (a_j\,a_{j-1}\,\ldots\,a_1)x = (a_j\,a_{j-1})\cdots(a_3\,a_2)(a_2\,a_1)x\in C_\rho \ . \tag{11}$$

Let us prepare some terms for this argument. When transposition $(a\,b)$ and $y\in C_\tau$ are multiplied to be $(a\,b)y\in C_{\tau'}$, we have $l(\tau')=l(\tau)+1$ or $l(\tau)-1$.

Then we say that $(a\,b)$ is a $+$ or $-$ length action for y respectively. Moreover, we have $|\tau'| \in \{|\tau| \pm i | i = 0, 1, 2\}$ and say that $(a\,b)$ is a $+i$ or $-i$ size action for y. At every step of multiplying $(a_i\,a_{i-1})$ succesively $(j-1)$ times in (11), we give colours — R(red), Y(yellow), G(green) — to $a_1, a_2, \ldots, a_j$. First Y is put on a_1 and G is put on a_2. After performing $(a_2\,a_1)x$, we put R, Y, G on a_1, a_2, a_3 respectively. Next, after performing $(a_3\,a_2)(a_2\,a_1)x$, we put R, R, Y, G on a_1, a_2, a_3, a_4 respectively, ... In other words, in the succesive multiplying actions in (11), each colour indicates the role of a letter in $\mathrm{supp}s = \{a_1, \ldots, a_j\}$ at the moment as follows:

- R : letter which cannot move any more in the succeeding actions
- Y : letter which moves next once
- G : letter which moves next twice.

Increase [resp. decrease] of the size ($= \sharp$boxes) caused by a $+$ [resp. $-$] length action is 2, 1, or 0. Just before the step $y \mapsto (a\,b)y$ in the multiplying actions in (11), a has colour G and b has colour Y. At the moment, this step possibly contributes to the growth of $p^{\sigma}_{(j)\rho}$ with order n if $a \notin \mathrm{supp}y$ and with order n^2 if $y = x, a \notin \mathrm{supp}x, b \notin \mathrm{supp}x$. Otherwise the contribution remains finite (i.e. with order n^0). We refer to this power of n as the possible increase for $p^{\sigma}_{(j)\rho}$ of each multiplying action in (11).

Lemma 2. *For $j \in \{2, 3, \ldots\}$ and $\sigma, \rho \in \mathcal{Y} \cup \{\emptyset\}$ such that neither $\sigma = \rho \sqcup (j)$ nor $\sigma = \rho \setminus (j)$, we have*

$$p^{\sigma}_{(j)\rho} \sqrt{\frac{\sharp C_\sigma}{\sharp C_{(j)} \sharp C_\rho}} = O(1/\sqrt{n}) \quad (\text{ as } n \to \infty). \tag{12}$$

Proof. [Step 1] The left hand side of (12) is majorised by a constant multiple of

$$\frac{n^{\text{sum of the possible increase for } p^{\sigma}_{(j)\rho}}}{n^{\text{sum of the increase of size}}} \times \sqrt{\frac{n^{\text{sum of the increase and the decrease of size}}}{n^j}}\ , \tag{13}$$

where the sums are taken over the $(j-1)$ times multiplying actions in $x \in C_\sigma \mapsto s^{-1}x \in C_\rho$ in (11). If neither $+2$ nor -2 size action appears in (11), then (13) $= O(1/\sqrt{n})$.

[Step 2] If a $+2$ size action once appears, then neither $+2$ nor -2 size action appears afterwards. For, after a $+2$ size action, we always see a letter of colour R as the left neighbour of a letter of colour Y. Then it is impossible that a $+2$ or -2 size action occurs. On the other hand, if a -2 size action once appears, then, except for possible appearance of a $+2$ size action at the next step, we do not see afterwards $+2$ or -2 size actions. For, after the next action of a -2 size action, we always see R as the left neighbour of Y.

[Step 3] If a -2 size action appears on the way, then, since a 1-cycle with colour Y is produced, the next is a + length action; if it is a $+1$ [resp. $+2$] size action, then, since it contributes to $p^{\sigma}_{(j)\rho}$ with 0th [resp. 1st] order and Step 2 assures 'sum of the increase and the decrease of size' in (13) $\leq j$ [resp. $j+1$], we see (13) $= O(1/n)$ [resp. $O(1/\sqrt{n})$]. If the last (i.e. $(j-1)$ th) action is a -2 size one, there is a 2-cycle consisting of a Y-letter and a G-letter. The 2-cycle comes from a cycle of the form $\cdots$ G$*$Y$\cdots$ ($*$ indicates any box) just before. If the '$\cdots$' part actually exists in the left or right, the action of cutting out G$*$ (which turns into YG) does not reduce the size. Then, (13) $= O(1/\sqrt{n})$. Otherwise, G$*$Y (= YG$*$) comes form $\cdots$ G$**$Y$\cdots$ just before. If one '$\cdots$' part exists at least, a similar argument assures (13) $= O(1/\sqrt{n})$. Tracing back in this way, we see that (13) $= O(1/\sqrt{n})$ unless $s^{-1}x$ is made by removing a j-cycle from x.

[Step 4] Assume that a $+2$ size action appears. The case where a -2 size action occurred just before is already discussed in Step 3. Otherwise, 'sum of the increase and the decrease of size' in (13) $\leq j$. One of the two added letters has colour Y and either (i)is a_1 or (ii)came from a cycle of length ≥ 2 before. In case (ii), since the possible increase for $p^{\sigma}_{(j)\rho}$ of this $+2$ size action is 1, (13) $= O(1/n)$. In case (i), the $+2$ size action is the first one and produces a cycle of YR. Then, unless letters are added to this cycle one by one succesively, there is a + length action of connecting YR$\cdots$R with a cycle of length ≥ 2, which does not cause increase of the size. Then (13) $= O(1/\sqrt{n})$. Hence, unless $s^{-1}x$ is made by adding a j-cycle to x, (13) $= O(1/\sqrt{n})$.

We thus showed (12) with the indicated exceptions.

3.3 Proof of theorems

Proof of Theorem 1. [Step 1] If C_ρ is a conjugacy class in $S(n)$, then

$$\sharp C_\rho = n(n-1)\cdots(n-|\rho|+1)/\prod_{j\geq 2} j^{k_j(\rho)} k_j(\rho)!\ .$$

For $\sigma = \rho \sqcup (j)$, we have $k_j(\sigma) = k_j(\rho)+1$ and, as $n \to \infty$,

$$p^{\sigma}_{(j)\rho}\sqrt{\frac{\sharp C_\sigma}{\sharp C_{(j)}\sharp C_\rho}} \sim k_j(\sigma)\sqrt{\frac{j\cdot j^{k_j(\rho)}k_j(\rho)!}{j^{k_j(\sigma)}k_j(\sigma)!}} = \sqrt{k_j(\rho)+1}\ .$$

For $\sigma = \rho \setminus (j)$, we have $k_j(\sigma) = k_j(\rho)-1$ and, as $n \to \infty$,

$$p^{\sigma}_{(j)\rho}\sqrt{\frac{\sharp C_\sigma}{\sharp C_{(j)}\sharp C_\rho}} \sim \frac{1}{j}\sqrt{\frac{j\cdot j^{k_j(\rho)}k_j(\rho)!}{j^{k_j(\sigma)}k_j(\sigma)!}} = \sqrt{k_j(\rho)}\ .$$

If σ is related to ρ in the other manners, Lemma 2 assures

$$p^{\sigma}_{(j)\rho}\sqrt{\frac{\sharp C_\sigma}{\sharp C_{(j)}\sharp C_\rho}} \longrightarrow 0 \quad \text{as } n \to \infty.$$

[Step 2] We show Theorem 1 by induction on m. The case of $m = 0$ is trivial. Assume that (7) holds for m. Using (10), Step 1 and the assumption of induction, we get, for $\forall \epsilon_1, \ldots, \epsilon_{m+1} \in \{+, -\}$, $\forall j_1, \ldots, j_{m+1} \in \{2, 3, \ldots\}$ and $\forall \tau, \rho \in \mathcal{Y} \cup \{\emptyset\}$,

$$
\begin{aligned}
&\langle \Phi(\tau), \frac{A_{(j_1)}^{\epsilon_1}}{\sqrt{\sharp C_{(j_1)}}} \cdots \frac{A_{(j_m)}^{\epsilon_m}}{\sqrt{\sharp C_{(j_m)}}} \frac{A_{(j_{m+1})}^{+}}{\sqrt{\sharp C_{(j_{m+1})}}} \Phi(\rho) \rangle_{\ell^2(S(n))} \\
&= \sum_{i=1}^{j_{m+1}-1} \sum_{\sigma : l(\sigma) = l(\rho) + i, |\sigma| \le n} p_{(j_{m+1})\rho}^{\sigma} \sqrt{\frac{\sharp C_\sigma}{\sharp C_{(j_{m+1})} \sharp C_\rho}} \\
&\qquad\qquad \times \langle \Phi(\tau), \frac{A_{(j_1)}^{\epsilon_1}}{\sqrt{\sharp C_{(j_1)}}} \cdots \frac{A_{(j_m)}^{\epsilon_m}}{\sqrt{\sharp C_{(j_m)}}} \Phi(\sigma) \rangle_{\ell^2(S(n))} \\
&\qquad + \frac{1}{2} \sum_{\sigma : l(\sigma) = l(\rho), |\sigma| \le n} (\text{the same summand}) \\
&\longrightarrow \sqrt{k_{j_{m+1}}(\rho) + 1} \langle \Psi(\tau), B_{j_1}^{\epsilon_1} \cdots B_{j_m}^{\epsilon_m} \Psi(\rho \sqcup (j_{m+1})) \rangle_\Gamma \quad (\text{as } n \to \infty) \\
&= \langle \Psi(\tau), B_{j_1}^{\epsilon_1} \cdots B_{j_m}^{\epsilon_m} B_{j_{m+1}}^{+} \Psi(\rho) \rangle_\Gamma \ .
\end{aligned}
$$

The same computation is valid for '$-$' instead of '$+$'. Hence our induction proceeds.

Proof of Theorem 2. By virtue of linearity it suffices to show (8) for monomials $Q_2, \ldots, Q_m$. Using the commutation relation, we rewrite a part of non-normal order (if it exists) as $B_j^- B_j^+ = B_j^+ B_j^- + I$. Then, again by linearity, only the case where $Q_j(B_j^+, B_j^-) = (B_j^+)^{\alpha_j} (B_j^-)^{\beta_j}$ remains to be checked. However, the assumption for ρ and τ implies that, for such Q_j's, the both sides of (8) vanish unless $\alpha_j = \beta_j = 0$. This completes the proof.

References

1. Accardi, L., Bożejko, M.: Interacting Fock spaces and Gaussianization of probability measures. Infin. Dimen. Anal. Quantum Probab. Relat. Top., **1**, 663–670 (1998)
2. Bannai, E., Ito, T.: Algebraic combinatorics I, association schemes. Menlo Park, California, Benjamin / Cummings (1984)
3. Biane, P.: Permutation model for semi-circular systems and quantum random walks. Pacific J. Math., **171**, 373–387 (1995)
4. Biane, P.: Representations of symmetric groups and free probability. Advances in Math., **138**, 126–181 (1998)
5. Hashimoto, Y.: Deformations of the semicircular law derived from random walks on free groups. Probab. Math. Stat., **18**, 399–410 (1998)
6. Hashimoto, Y.: Quantum decomposition in discrete groups and interacting Fock spaces. Infin. Dimen. Anal. Quantum Probab. Relat. Top., **4**, 277–287 (2001)

7. Hashimto, Y., Hora, A., Obata, N.: Central limit theorems for large graphs: a method of quantum decomposition. Preprint (2001)
8. Hashimoto, Y., Obata, N., Tabei, N.: A quantum aspect of asymptotic spectral analysis of large Hamming graphs. In: Hida, T., Saitô, K. (eds) Quantum Information III, Singapore, World Scientific, 45–57 (2001)
9. Hora, A.: Central limit theorem for the adjacency operators on the infinite symmetric group. Commun. Math. Phys., **195**, 405–416 (1998)
10. Hora, A.: Central limit theorems and asymptotic spectral analysis on large graphs. Infin. Dimen. Anal. Quantum Probab. Relat. Top., **1**, 221–246 (1998)
11. Hora, A.: Gibbs state on a distance-regular graph and its application to a scaling limit of the spectral distributions of discrete Laplacians. Probab. Theory Relat. Fields, **118**, 115–130 (2000)
12. Ivanov, V., Olshanski, G.: Kerov's central limit theorem for the Plancherel measure on Young diagrams. Preprint, (2001)
13. Kerov, S.: Gaussian limit for the Plancherel measure of the symmetric group. C. R. Acad. Sci. Paris, **316**, Série I, 303–308 (1993)
14. Logan, B. F., Shepp, L. A.: A variational problem for random Young tableaux. Advances in Math., **26**, 206–222 (1977)
15. Vershik, A. M., Kerov, S. V.: Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux. Doklady AN SSSR, **233**, 1024–1027 (1977); English translation: Soviet Mathematics Doklady, **18**, 527–531 (1977)

Random trees and moduli of curves

Andrei Okounkov

Department of Mathematics
University of California at Berkeley
Evans Hall #3840, Berkeley, CA 94720–3840
USA
okounkov@math.berkeley.edu

Summary. This is an expository account of the proof of Kontsevich's combinatorial formula for intersections on moduli spaces of curves following the paper [14]. It is based on the lectures I gave on the subject in St. Petersburg in July of 2001.

These are notes from the lectures I gave in St. Petersburg in July of 2001. Our goal here is to give an informal introduction to intersection theory on the moduli spaces of curves and its relation to random matrices and combinatorics. More specifically, we want to explain the proof of Kontsevich's formula given in [14] and how it is connected to other topics discussed at this summer school such as, for example, the combinatorics of increasing subsequences in a random permutations.

These lectures were intended for an audience of mostly analysts and combinatorialists interested in asymptotic representation theory and random matrices. This is very much reflected in both the selection of the material and its presentation. Since absolutely no background in geometry was assumed, there is a long and very basic discussion of what moduli spaces of curves and intersection theory on them are about. We hope that a reader trained in analysis or combinatorics will get some feeling for moduli of curves (without worrying too much about the finer points of the theory, all but a few of which were swept under the rug).

Conversely, in the second, asymptotic, part of the text, I allowed myself to operate more freely because the majority of the audience was experienced in asymptotic analysis. Also, since many fundamental ideas such as e.g. the KdV equations for the double scaling limit of the Hermitian 1-matrix model were at length discussed in other lectures of the school, their discussion here is much more brief that it would have been in a completely self-contained course. A much more detailed treatment of both geometry and asymptotics can be found in the paper [14], on which my lectures were based.

It is needless to say that, since this is an expository text based on my joint work with Rahul Pandharipande, all the credit should be divided while any

A. Borodin et al.: LNM 1817, A.M. Vershik (Ed.), pp. 89–126, 2003.

blame is solely my responsibility. Many people contributed to the success of the St. Petersburg summer school, but I want to especially thank A. Vershik for organizing the school and for the invitation to participate in it. I am grateful to NSF (grant DMS-0096246), Sloan foundation, and Packard foundation for partial financial support.

1 Introduction to moduli of curves

1.1

Let me begin with an analogy. In the ideal world, the moduli spaces of curves would be quite similar to the Grassmann varieties

$$Gr_{k,n} = \{L \subset \mathbb{C}^n, \dim L = k\}$$

of k-dimensional linear subspaces L of an n-dimensional space. While any such subspace L is geometrically just a k-dimensional vector space $\mathbb{C}^k$, nontrivial things can happen if L is allowed to vary in families, and this nontriviality is captured by the geometry of $Gr_{k,n}$.

A convenient formalization of the notion of a family of linear spaces parameterized by points of some base space B is a (locally trivial) vector bundle over B. There is a natural *tautological* vector bundle over the Grassmannian $Gr_{k,n}$ itself, namely the space

$$\mathcal{L} = \{(L, v), v \in L \subset \mathbb{C}^n\}$$

formed by pairs (L, v), where L is a k-dimensional subspace of $\mathbb{C}^n$ and v is a vector in L. Forgetting the vector v gives a map $\mathcal{L} \to Gr_{k,n}$ whose fiber over $L \in Gr_{k,n}$ is naturally identified with L itself.

Given any space B and a map

$$\phi : B \to Gr_{k,n}$$

we can form the pull-back of $\mathcal{L}$

$$\phi^*\mathcal{L} = \{(b, v), b \in B, v \in C^n, v \in \phi(b)\}$$

which is a rank k vector bundle over B. For a compact base B, in the $n \to \infty$ limit this becomes a bijection between (homotopy classes) of maps $B \to Gr_{k,\infty}$ and (isomorphism classes) of rank k vector bundles on B.

In particular, one can associate to a vector bundle $\phi^*\mathcal{L}$ its characteristic cohomology classes obtained by pulling back the elements of $H^*(Gr_{k,n})$ via the map ϕ. Intersections of these classes describe the enumerative geometry of the bundle $\phi^*\mathcal{L}$. It is thus especially important to understand the intersection theory on the space $Gr_{k,n}$ itself — and this leads to a very beautiful classical combinatorics, in particular, Schur functions play a central role (see for example [5], Chapter 14).

1.2

One would like to have a similar theory with families of linear spaces replaced by families of curves of some genus g. That is, given a family F of, say, smooth genus g algebraic curves parameterized by some base B we want to have a natural map $\phi : B \to \mathcal{M}_g$ that captures the essential information about the family F. Here $\mathcal{M}_g$ is the *moduli space* of smooth curves of genus g, that is, the space of isomorphism classes of smooth genus g curves. At this point it may be useful to be a little naive about what we mean by a family of curves etc., in this way we should be able to understand the basic issues without too many technicalities getting in our way. Basically, a family F of curves with base B is a "nice" morphism

$$\pi : F \to B$$

of algebraic varieties whose fibers $\pi^{-1}(b)$, $b \in B$, are smooth complete genus g curves. We want the moduli space $\mathcal{M}_g$ and the induced map $\phi : B \to \mathcal{M}_g$ to be also algebraic.

We will see that the first difficulty with the above program is that, in general, the family F will not be a pull-back of any universal family over $\mathcal{M}_g$. To get a sense of why this is the case we can cheat a little and consider the (normally forbidden) case $g = 0$. Up to isomorphism, there is only one curve of genus 0, namely the projective line $\mathbb{P}^1$. Hence the map ϕ in this case can only be the trivial map to a point. There exist, however, highly nontrivial families with fibers isomorphic to $\mathbb{P}^1$ or even $\mathbb{C}^1$ as we, in fact, already saw above in the example of the tautological rank 1 bundle over $Gr_{1,n} \cong \mathbb{P}^{n-1}$.

The reason why there exist locally trivial yet globally nontrivial families with fiber $\mathbb{P}^1$ is that $\mathbb{P}^1$ has a large automorphism group which one can use to glue trivial pieces in a nontrivial way. Basically, the automorphisms are the principal issue behind the nonexistence of a universal family of curves over $\mathcal{M}_g$. The situation becomes manageable, if not entirely perfect, once one can get the automorphism group to be finite (which is automatic for smooth curves of genus $g > 1$). A standard way to achieve this is to consider curves with sufficiently many marked points on them (≥ 3 marked points for $g = 0$ and ≥ 1 for $g = 1$). Since curves with marked points arise very naturally in many other geometric situations, the moduli spaces $\mathcal{M}_{g,n}$ of smooth genus g curves with n distinct marked points should be considered on equal footing with the moduli spaces $\mathcal{M}_g$ of plain curves.

1.3

As the first example, let us consider the space $\mathcal{M}_{1,1}$ of genus $g = 1$ curves C with one parked point $p \in C$. By Riemann–Roch the space of $H^0(C, \mathcal{O}(2p))$ of meromorphic functions on C with at most double pole at p has dimension 2. Hence, in addition to constants, there exists a (unique up to linear combinations with constants) nonconstant meromorphic function

$$f : C \to \mathbb{P}^1 ,$$

with a double pole at p and no other poles (this is essentially the Weierstraß function $\wp$). Thus, f defines at a 2-fold branched covering of $\mathbb{P}^1$ doubly ramified over $\infty \in \mathbb{P}^1$. For topological reasons, it has three additional ramification points which, after normalization, we can take to be $0, 1$ and some $\lambda \in \mathbb{P}^1 \setminus \{0, 1, \infty\}$.

Now it is easy to show that such a curve C must be isomorphic to the curve

$$C \cong \{y^2 = x(x-1)(x-\lambda)\} \in \mathbb{P}^2 \tag{1}$$

in such a way that the point p becomes the unique point at infinity and the function f becomes the coordinate function x. It follows that every smooth pointed $g = 1$ curves occurs in the following family of curves

$$F = \{(x, y, \lambda), y^2 = x(x-1)(x-\lambda)\} \tag{2}$$

with the base

$$B = \{\lambda\} = \mathbb{P}^1 \setminus \{0, 1, \infty\} ,$$

where the marked point is the point p at infinity. For example, the curve (1) corresponding to $\lambda = \frac{3}{2}$ is plotted in Figure 1.

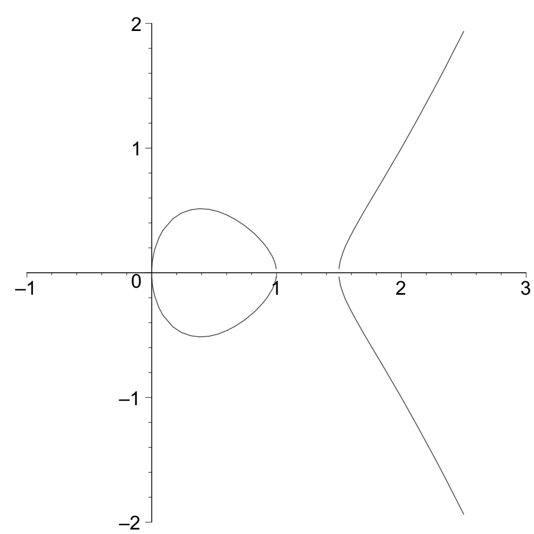


Fig. 1. The $\lambda = \frac{3}{2}$ member of the family (2).

However, a given curve C occurs in the family (2) more than once. Indeed, we made arbitrary choices when we normalized the 3 critical values of f to be 0, 1, and λ, respectively. At this stage, we can choose any of $6 = 3!$ possible assignments which makes the symmetric group $S(3)$ act on the base

B preserving the isomorphism class of the fiber. Concretely, this group is generated by involutions

$$\lambda \mapsto 1 - \lambda\,,$$

which interchanges the roles of 0 and 1, and by

$$\lambda \mapsto 1/\lambda\,,$$

exchanging the roles of 1 and λ. It can be shown that two members of the family F are isomorphic if and only if they belong to the same $S(3)$ orbit. Thus, the structure map $\phi : B \to \mathcal{M}_{1,1}$ should be just the quotient map

$$\phi : B \to B/S(3) = \operatorname{Spec} \mathbb{C}[B]^{S(3)}\,.$$

Here $\mathbb{C}[B]^{S(3)}$ is the algebra of $S(3)$-invariant regular functions on B. This algebra is a polynomial algebra with one generator, the traditional choice for which is the following

$$j(\lambda) = 256\,\frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(\lambda - 1)^2}\,.$$

Thus, $\mathcal{M}_{1,1}$ is simply a line

$$\mathcal{M}_{1,1} = \operatorname{Spec} \mathbb{C}[j] \cong \mathbb{C}\,.$$

It is now time to point out that the family F is not a pull-back ϕ^* of some universal family over $\mathcal{M}_{1,1}$. The simplest way to see this is to observe that the group $S(3)$ fails to act on F. Indeed, let us try to lift the involution $\lambda \to 1-\lambda$ from B to F. There are two ways to do this, namely

$$(x, y, \lambda) \mapsto (1 - x, \pm iy, 1 - \lambda)\,,$$

neither of which is satisfactory because the square of either map

$$(x, y, \lambda) \mapsto (x, -y, \lambda)$$

yields, instead of identity, a nontrivial automorphism of every curve in the family F. One should also observe that both choices act by a nontrivial automorphism on the fiber over the fixed point $\lambda = \frac{1}{2} \in B$. In fact, the fibers of F over fixed points of a transposition and a 3-cycle in $S(3)$, respectively, (with $j(\lambda) = 1728$ and $j(\lambda) = 0$, resp.) are precisely the curves with extra large automorphism groups (of order 4 and 6, resp.).

The existence of an nontrivial automorphism of every pointed genus 1 curve leads to the somewhat unpleasant necessity to consider every point of $\mathcal{M}_{1,1}$ as a "half-point" in some suitable sense in order to get correct enumerative predictions. Again, automorphisms make the real world not quite ideal.

While it is important to be aware of these automorphism issues (for example, to understand how intersection numbers on moduli spaces can be rational numbers), there is no need to be pessimistic about them. In fact, by allowing spaces more general than algebraic varieties (called *stacks*) one can live a life in which $\mathcal{M}_{g,n}$ is smooth and with a universal family over it. This is, however, quite technical and will remain completely outside the scope of these lectures.

1.4

Clearly, the space $\mathcal{M}_{1,1} \cong \mathbb{C}$ is not compact. The j-invariant of the curve (1) goes to ∞ as the parameter λ approaches the three excluded points $\{0, 1, \infty\}$. As λ approaches 0 or 1, the curve C acquires a nodal singularity; for example, for $\lambda = 1$ we get the curve

$$y^2 = x(x-1)^2$$

plotted in Figure 2. It is natural to complete the family (2) by adding the

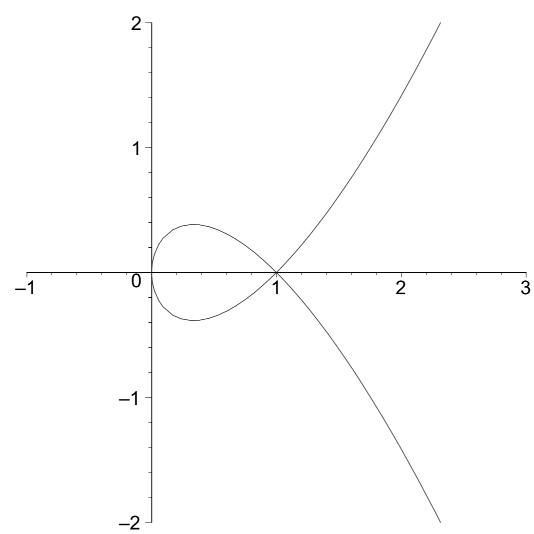


Fig. 2. The nodal $\lambda = 1$ curve in the family (2).

corresponding nodal cubics for $\lambda \in \{0, 1\}$. All plane cubic with a node being isomorphic, the function j extends to a map

$$j : \mathbb{C} \to \mathbb{P}^1/S(3) = \overline{\mathcal{M}}_{1,1} \cong \mathbb{P}^1$$

to the moduli space of $\overline{\mathcal{M}}_{1,1}$ of curves of arithmetic genus 1 with at most one node and a smooth marked point.[1]

In general, it is very desirable to have a nice compactification for the noncompact spaces $\mathcal{M}_{g,n}$. First of all, interesting families of curves over a complete base B are typically forced to have singular fibers over some points in the base (as in the example above). Fortunately, as we will see below, it often happens that precisely these special fibers contain key information about the geometry of the family. Also, since eventually we will be interested in intersection theory on the moduli spaces of curves, having a complete space can be a significant advantage.

A particularly remarkable compactification $\overline{\mathcal{M}}_{g,n}$ of $\mathcal{M}_{g,n}$ was constructed by Deligne and Mumford. The space $\overline{\mathcal{M}}_{g,n}$ is the moduli space of *stable* curves C of arithmetic genus g with n distinct marked points. Stability, by definition, means that the curve C is complete and connected with at worst nodal singularities, all marked points are smooth, and that C, together with marked points, admits only finitely many automorphisms. In practice, the last condition means that every rational component of the normalization of C should have at least 3 special (that is, lying over marked or singular points of C) points. Observe that, in particular, the curve C is allowed to be reducible. A typical stable curve can be seen in Figure 3.

1.5

Those who have not seen this before are probably left wondering how it is possible for $\overline{\mathcal{M}}_{g,n}$ to be compact. What if, for example, one marked point p_1 on some fixed curve C approaches another marked point $p_2 \in C$? We should be able to assign some meaningful stable limit to such a 1-parametric family of curves, but it is somewhat nontrivial to guess what it should be.

A family of curves with a 1-dimensional base B is a surface F together with a map $\pi : F \to B$ whose fibers are the curves of the family. Marking n points on the curves means giving n sections of the map π, that is, n maps

$$p_1, \dots, p_n : B \to F \,,$$

such that

$$\pi(p_k(b)) = b \,, \quad b \in B \,, \quad k = 1, \dots, n \,.$$

[1] For future reference we point out that something rather different happens in the family (2) as $\lambda \to \infty$. Indeed, the equation

$$\lambda^{-1} y^2 = x(x-1)(\lambda^{-1}x - 1)$$

becomes $x(x-1) = 0$ in the $\lambda \to \infty$ limit, which means that we get three lines (one of which is the line at infinity), all three of them intersecting in the marked point at infinity. In other words, the fiber of the family (2) at $\lambda = \infty$ is very much not the kind of curve by which we want to compactify $\mathcal{M}_{1,1}$. This problem can be cured, but in a not completely trivial way, see below.

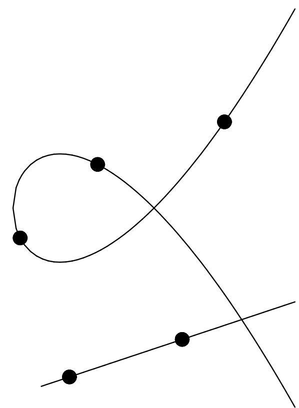


Fig. 3. A boundary element of $\overline{\mathcal{M}}_{1,5}$

We will denote by

$$S_i = p_i(B)$$

the trajectories of the marked points on F; they are curves on the surface F.

Now suppose we have a 1-dimensional family of 2-pointed curves such that at some bad point b_0 of the base B we have $p_1(b_0) = p_2(b_0)$, that is, over this point two marked points hit each other, see Figure 4, and therefore the fiber $\pi^{-1}(b_0)$ is not a stable 2-pointed curve. It is quite easy to repair this family:

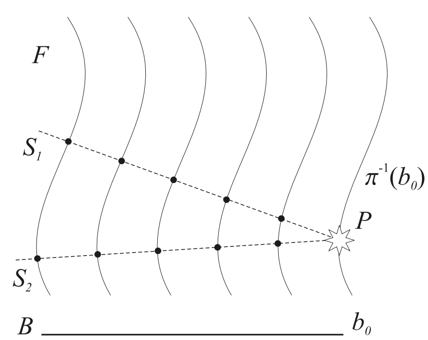


Fig. 4. A family with colliding marked points

just blow up the offending (but smooth) point $P = p_1(b_0) = p_2(b_0)$ on the surface F. Let

$$\sigma : \widetilde{F} \to F$$

be the blow-up at P. Then

$$\widetilde{\pi} = \pi \circ \sigma : \widetilde{F} \to B$$

is new family of curves with base B. Outside b_0 this the same family as before, whereas the fiber $\widetilde{\pi}^{-1}(b_0)$ is the old fiber $\pi^{-1}(b_0)$ plus the exceptional divisor $E = \sigma^{-1}(P) \cong \mathbb{P}^1$ of the blow-up, see Figure 5. Assuming the sections S_1

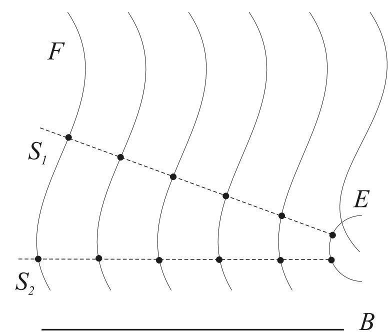


Fig. 5. Same family, except with collision point blown up

and S_2 met each other and the fiber $\pi^{-1}(b_0)$ transversally at P, as in Figure 4, the marked points on $\widetilde{\pi}^{-1}(b_0)$ are two distinct point on the exceptional divisor E. Therefore, $\widetilde{\pi}^{-1}(b_0)$ is a stable 2-pointed curve which is the stable limit of the curves $\pi^{-1}(b)$ as $b \to b_0$.

To summarize, if one marked point on a curve C approaches another then C bubbles off a projective line $\mathbb{P}^1$ with these two points on it as in Figure 6.

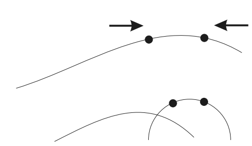


Fig. 6. Bubbling off a projective line

More generally, if F is any family of curves with a smooth 1-dimensional base B that are stable except over one offending point $b_0 \in B$ then after a sequence of blow-ups and blow-downs and, possibly, after passing to a branched covering of B, one can always arrive at a new family with all fibers stable. Moreover, the fiber over b_0 in this family is determined uniquely. This process is called the *stable reduction* and how it works is explained, for example, in [8], Chapter 3C.

In particular, there exists a stable reduction of the family (2) which, as we saw in Section 1.4, fails to have a stable fiber over the point $\lambda = \infty$ in the base. This is an example where only blow-ups and blow-downs will not suffice, that is, a base change is necessary.

1.6

The topic of this lectures is intersection theory on the Deligne–Mumford spaces $\overline{\mathcal{M}}_{g,n}$ and, specifically, intersections of certain divisors ψ_i which will be defined presently. It was conjectured by Witten [19] that a suitable generating function for these intersections is a τ-function for the Korteweg–de Vries hierarchy of differential equations. This conjecture was motivated by an analogy with matrix models of quantum gravity, where the same KdV hierarchy appears (this was already discussed in other lectures at this school). The KdV equations were deduced by Kontsevich in [9] from an explicit combinatorial formula for the intersections of the ψ-classes (see also, for example, [2] for more information about the connection to the KdV equations). The main goal of these lectures is to explain a proof of this combinatorial formula of Kontsevich following the paper [14].

The definition of the divisors ψ_i is the following. A point in $\overline{\mathcal{M}}_{g,n}$ is a stable curve C with marked points $p_1, \dots, p_n$. By definition, all points p_i are smooth points of C, hence the tangent space $T_{p_i}C$ to C at p_i is a line. Similarly, we have the cotangent lines $T^*_{p_i}C$, $i = 1, \dots, n$. As the point $(C, p_1, \dots, p_n) \in \overline{\mathcal{M}}_{g,n}$ varies, these cotangent lines $T^*_{p_i}C$ form n line bundles over $\overline{\mathcal{M}}_{g,n}$. By definition, ψ_i is the first Chern class of the line bundle $T^*_{p_i}C$. In other words, it is the divisor of any nonzero section of the line bundle $T^*_{p_i}C$.

1.7

To get a better feeling for these classes let us intersect ψ_i with a curve in $\overline{\mathcal{M}}_{g,n}$. The answer to this question should be a number. Let B be a curve. A map $B \to \overline{\mathcal{M}}_{g,n}$ is morally equivalent to a 1-dimensional family of curves with base B (in reality we may have to pass to a suitable branched covering of B to get an honest family.[2]) So, let us consider a family $\pi : F \to B$ of stable pointed curves with base B and the induced map $\phi : B \to \overline{\mathcal{M}}_{g,n}$. As usual, the marked points $p_1, \dots, p_n$ are sections of π and

$$S_i = p_i(B)\,, \quad i = 1, \dots, n\,,$$

are disjoint curves on the surface F. A section s of $\phi^*(T_{p_i})$ is a vector field on the curve S_i which is tangent to fibers of π and, hence, s is a section of the normal bundle to $S_i \subset F$, see Figure 7. The degree of this normal bundle is the self-intersection of the curve S_i on the surface F, that is,

$$\deg(s) = (S_i, S_i)_F\,,$$

where (s) is the divisor of s. In other words,

[2] We already saw an example of this in Section 1.3. Indeed, $\overline{\mathcal{M}}_{1,1}$ is itself a line $\mathbb{P}^1$. However, in order to get an actual family over it we have to go to a branched covering. As always, the automorphisms are to blame.

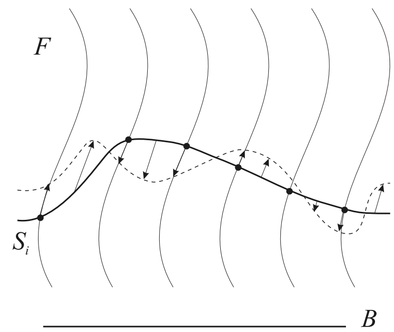


Fig. 7. The self-intersection of $S_i = p_i(B)$ on F.

$$\int_B c_1\left(\phi^*(T_{p_i})\right) = (S_i, S_i)_F\,,$$

where c_1 denotes the 1st Chern class. Dually, we have

$$\int_{\phi(B)} \psi_i = -(S_i, S_i)_F\,.$$

We will now use this formula to compute the intersections of ψ_i with $\overline{\mathcal{M}}_{g,n}$ in the cases when the space $\overline{\mathcal{M}}_{g,n}$ is itself 1-dimensional. Since

$$\dim \overline{\mathcal{M}}_{g,n} = 3g - 3 + n\,, \tag{3}$$

this happens for $\overline{\mathcal{M}}_{0,4}$ and $\overline{\mathcal{M}}_{1,1}$.

1.8

The space $\overline{\mathcal{M}}_{0,4}$ is easy to understand. After all, there is only one smooth curve of genus 0, namely $\mathbb{P}^1$. Moreover, any 3 distinct points of $\mathbb{P}^1$ can be taken to the points $\{0, 1, \infty\}$ by an automorphism of $\mathbb{P}^1$ (in particular, this means that $\overline{\mathcal{M}}_{0,3}$ is a point). After we identified the first three marked points with $\{0, 1, \infty\}$, we can take any point $x \in \mathbb{P}^1 \setminus \{0, 1, \infty\}$ as the fourth marked point. Thus the locus of smooth curves in $\overline{\mathcal{M}}_{0,4}$ is isomorphic to $\mathbb{P}^1 \setminus \{0, 1, \infty\}$. Singular curves are obtained as we let x approach the 3 excluded points $\{0, 1, \infty\}$, which, by the process described in Section 1.5, bubbles off a new $\mathbb{P}^1$ with two marked points on it. This completes the description of $\overline{\mathcal{M}}_{0,4} \cong \mathbb{P}^1$.

In addition, this gives a description of the universal family over $\overline{\mathcal{M}}_{0,4}$ (oh yes, in genus 0 it does exist !). Take $\mathbb{P}^1 \times \mathbb{P}^1$ with coordinates (x, y). The map $(x, y) \to x$ with 4 sections

$$p_1(x) = (x, 0)\,, \quad p_2(x) = (x, 1)\,, \quad p_3(x) = (x, \infty)\,, \quad p_4(x) = (x, x)\,,$$

defines a family of 4-pointed smooth genus 0 curves for $x \in \mathbb{P}^1 \setminus \{0, 1, \infty\}$. The section p_4 collides with the other three at the points $(0, 0)$, $(1, 1)$, and

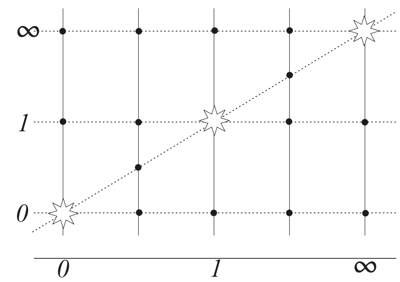


Fig. 8. The trivial family $\mathbb{P}^1 \times \mathbb{P}^1$ with sections $p_1, \dots, p_4$.

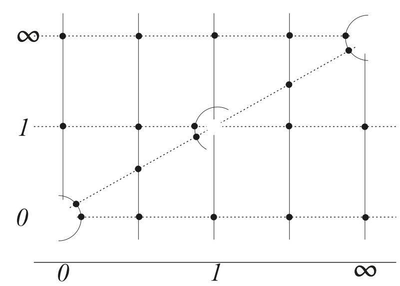


Fig. 9. The universal family over $\overline{\mathcal{M}}_{0,4}$.

(∞, ∞), see Figure 8. To extend this family over all of $\mathbb{P}^1$, we blow up these collision points as in Section 1.5 and get the surface F shown in Figure 9. The closures of the curves $p_1(x), \dots, p_4(x)$ give 4 sections which are now disjoint everywhere.

Incidentally, this surface F can be naturally identified with $\overline{\mathcal{M}}_{0,5}$ and, more generally, for any n there exists a natural map $\overline{\mathcal{M}}_{0,n+1} \to \overline{\mathcal{M}}_{0,n}$ giving the universal family over $\overline{\mathcal{M}}_{0,n}$. This map forgets the $(n+1)$st marked point and, if the curve becomes unstable after that, blows down all unstable components.

Now let us compute $\int_{\overline{\mathcal{M}}_{0,4}} \psi_1$ using the recipe given in Section 1.7. Recall that $S_1 \subset F$ denotes the closure of the curve $\{(x, 0)\}$, $x \neq 0$ in F (a.k.a. the proper transform of the corresponding curve in $\mathbb{P}^1 \times \mathbb{P}^1$). Let E denote the preimage of $(0, 0)$ under the blow-up, that is, let E be the exceptional divisor. The self-intersection of S_1 with any curve $\{y = c\}$, $c \neq 0$, on $\mathbb{P}^1 \times \mathbb{P}^1$ is clearly zero. Letting $c \to 0$ we get

$$(S_1, E + S_1) = 0\,.$$

Since, obviously, $(S_1, E) = 1$ we conclude that

$$\int_{\overline{\mathcal{M}}_{0,4}} \psi_1 = -(S_1, S_1) = -(-1) = 1\,.$$

1.9

Now let us analyze the integral $\int_{\overline{\mathcal{M}}_{1,1}} \psi_1$. In the absence of a universal family, we have to look for another suitable family to compute this integral. A particularly convenient family can be obtained in the following way. Consider the projective plane $\mathbb{P}^2$ with affine coordinates (x, y). Pick two generic cubic polynomials $f(x, y)$ and $g(x, y)$ and consider the family of cubic curves

$$F = \{(x, y, t), f(x, y) - t\, g(x, y) = 0\} \subset \mathbb{P}^2 \times \mathbb{P}^1 , \tag{4}$$

with base $B = \mathbb{P}^1$ parameterized by t. The cubic curves $f(x, y) = 0$ and $g(x, y) = 0$ intersect in 9 points $p_1, \ldots, p_9$ and we can choose any of those points as the marked point in our family. An example of such family of plane cubics is plotted in Figure 10.

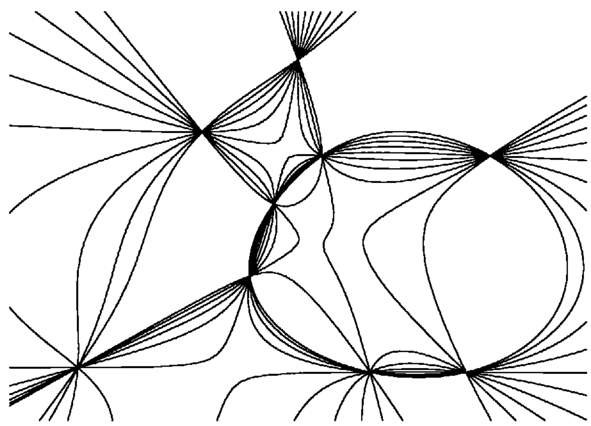


Fig. 10. A pencil of plane cubics.

Our first observation is that the surface F is the blow-up of $\mathbb{P}^2$ at the points $p_1, \ldots, p_9$ (which are distinct for generic f and g). Indeed, we have a rational map

$$\mathbb{P}^2 \ni (x, y) \to \left(x, y, \frac{g}{f}\right) \in F ,$$

which is regular away from $p_1, \ldots, p_9$. Each of the points p_i is a transverse intersection of $f = 0$ and $g = 0$, which is another way of saying that at those points the differentials df and dg are linearly independent. Thus, this map identifies F with the blow-up of $\mathbb{P}^2$ at $p_1, \ldots, p_9$. The graph of the function $\frac{g(x,y)}{f(x,y)}$ is shown in Figure 11. This graph goes vertically over the points $p_1, \ldots, p_9$ that are blown up.

Since the section S_1 is exactly one of the exceptional divisors of this blow-up, arguing as in Section 1.8 above we find that

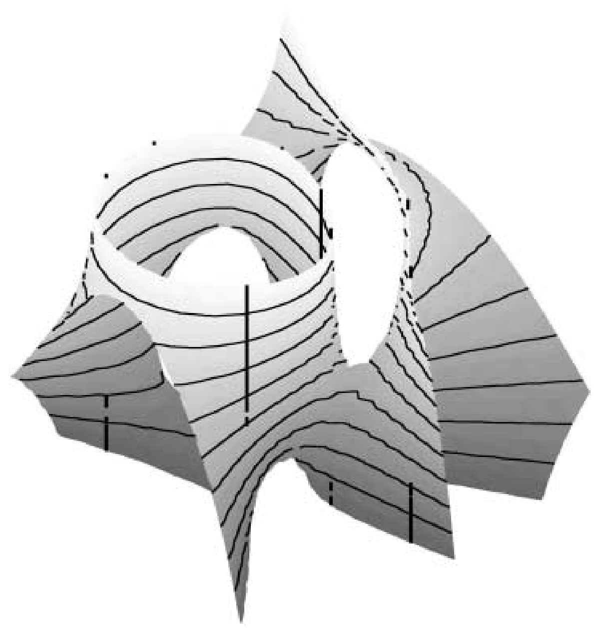


Fig. 11. A fragment of the surface (4)

$$(S_1, S_1) = -1 .$$

It does not mean, however, that we are done with the computation of the integral, because the induced map $\phi : B \to \overline{\mathcal{M}}_{1,1}$ is very far from being one-to-one. In fact, set-theoretically, the degree of the map ϕ is 12 as we shall now see.

To compute the degree of ϕ we need to know how many times a fixed generic elliptic curve appears in the family F. This is a classical computation. First, one can show that the singular cubic is generic enough. Then we claim that, as t varies, there will be precisely 12 values of t that produce a singular curve. There are various ways to see this. For example, the singularity of the curve is detected by vanishing of the discriminant. The discriminant of a cubic polynomial is a polynomial of degree 12 in its coefficients, hence a polynomial of degree 12 in t.

A alternative way to obtain this number 12 is to compute the Euler characteristic of the surface F in two different ways. On the one hand, viewing F as a blow-up, we get

$$\chi(F) = \chi(\mathbb{P}^2) + 9(\chi(\mathbb{P}^1) - 1) = 12 .$$

On the other hand, F is fibered over B and the generic fiber is a smooth elliptic curve whose Euler characteristic is 0. The special fibers are the nodal elliptic curves with Euler characteristic equal to 1. Hence, there are 12 special fibers.

However, as remarked in Section 1.3, each point of $\overline{\mathcal{M}}_{1,1}$ is really a half-point because of automorphism of order 2 of any pointed genus 1 curve. Therefore, the $24 = 2 \cdot 12$ is the true degree of the map ϕ. By the push-pull formula we thus obtain

$$\int_{\overline{\mathcal{M}}_{1,1}} \psi_1 = \frac{1}{\deg \phi} \int_B \phi^* \psi_1 = \frac{1}{\deg \phi} (-(S_1, S_1)) = \frac{1}{24}\,.$$

An interesting corollary of this computation is that if $F \to B$ is a smooth family of 1-pointed genus 1 stable curves over a smooth complete curve B then the set-theoretic degree of the induced map $B \to \overline{\mathcal{M}}_{1,1}$ has to be divisible by 12.

1.10

It is difficult to imagine being able to compute many intersections of the ψ-classes in the above manner. To begin with, it is essentially impossible to write down a sufficiently explicit family of general high genus curves, see the discussion in Chapter 6F of [8]. It is therefore amazing that there exist several complete and beautiful descriptions of the all possible intersection numbers of the form

$$\langle \tau_{k_1} \dots \tau_{k_n} \rangle \overset{\text{def}}{=} \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{k_1} \cdots \psi_n^{k_n}\,, \quad k_1 + \dots + k_n = 3g - 3 + n\,. \tag{5}$$

The most striking description was conjectured by Witten [19] and says the exponential of the following generating function for the numbers (5)

$$F(t_1, t_2, \dots) = \sum_n \frac{1}{n!} \sum_{k_1,\dots,k_n} \langle \tau_{k_1} \dots \tau_{k_n} \rangle \, t_{k_1} \cdots t_{k_n} \tag{6}$$

is a τ-function for the KdV hierarchy of differential equations. This conjecture was motivated by the (physical) analogy with the random matrix models of quantum gravity and, in fact, the τ-function thus obtained is the same as the one that arises in the double scaling of the 1-matrix model (and discussed in other lectures of this school). The KdV equation and the string equation satisfied by the τ-function uniquely determine all numbers (5). Alternatively, the numbers (5) are uniquely determined by the associated Virasoro constraints. Further discussion can be found, for example, in [2].

1.11

Kontsevich in [9] obtained the KdV equations for (5) from a combinatorial formula for the following (somewhat nonstandard) generating function

$$K_{g,n}(z_1, \dots, z_n) = \sum_{k_1+\dots+k_n=3g-3+n} \langle \tau_{k_1} \dots \tau_{k_n} \rangle \prod \frac{(2k_i - 1)!!}{z_i^{2k_i+1}}\,, \tag{7}$$

for the numbers (5) with fixed g and n.

The main ingredient in Kontsevich's combinatorial formula is a 3-valent graph G embedded in the a topological surface Σ_g. A further condition on this graph G is that the complement $\Sigma_g \setminus G$ is a union of n topological disks (in particular, this forces G to be connected). These disks, called *cells*, have to (bijectively) numbered by $1, \dots, n$. Two such graphs G and G' are identified if there exist an orientation preserving homeomorphism of Σ_g that takes G to G' and preserves the labels of the cells. In particular, every graph G has an automorphism group $\operatorname{Aut} G$, which is finite and only seldom nontrivial. Let $\mathsf{G}^3_{g,n}$ denote the set of distinct such graphs G with given values of g and n; this is a finite set. An example of an element of $\mathsf{G}^3_{2,3}$ is shown in Figure 12.

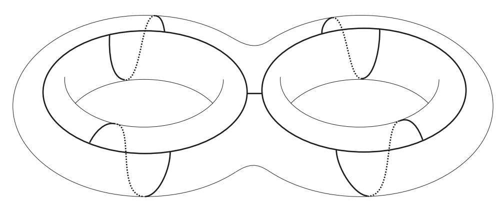


Fig. 12. A trivalent map on a genus 2 surface

Another name for a graph $G \subset \Sigma_g$ such that $\Sigma_g \setminus G$ is a union of cells is a *map* on Σ_g. One can imagine that the cells are the countries in which the graph G divides the surface Σ_g.

Kontsevich's combinatorial formula for the function (7) is the following:

$$K_{g,n}(z_1, \dots, z_n) = \\ 2^{2g-2+n} \sum_{G \in \mathsf{G}^3_{g,n}} \frac{1}{|\operatorname{Aut} G|} \prod_{\text{edges } e \text{ of } G} \frac{1}{z_{\text{one side of } e} + z_{\text{other side of } e}}, \quad (8)$$

where the meaning of the term

$$z_{\text{one side of } e} + z_{\text{other side of } e}$$

is the following. Each edge e of G separates two cells (which may be identical). These cell carry some labels, say, i and j. Then $(z_i + z_j)^{-1}$ is the factor in (8) corresponding to the edge e.

To get a better feeling for how this works let us look at the cases $(g, n) = (0, 3), (1, 1)$ that we understand well. The space $\overline{\mathcal{M}}_{0,3}$ is a point and the only nontrivial integral over it is

$$\langle \tau_0 \, \tau_0 \, \tau_0 \rangle = \int_{\overline{\mathcal{M}}_{0,3}} 1 = 1 \, .$$

Thus,

$$K_{0,3} = \frac{1}{z_1 z_2 z_3} \,.$$

The combinatorial side of Kontsevich's formula, however, is not quite trivial. The set $\mathsf{G}^3_{0,3}$ consists of 4 elements. Two of them are shown in Figure 13; the

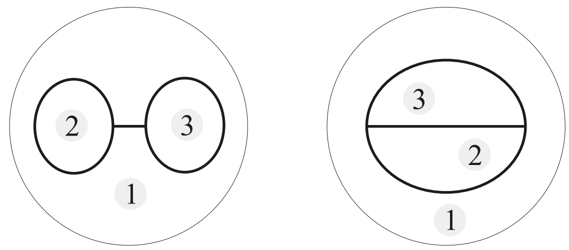


Fig. 13. Elements of $\mathsf{G}^3_{0,3}$

other two are obtained by permuting the cell labels of the graph of the left. All these graphs have only trivial automorphisms. Hence, we get

$$K_{0,3} = 2\left(\frac{1}{2z_1(z_1+z_2)(z_1+z_3)} + \text{permutations} \right. \\ \left. + \frac{1}{(z_1+z_2)(z_1+z_3)(z_2+z_3)}\right) ,$$

and, indeed, this simplifies to $(z_1 z_2 z_3)^{-1}$. What is apparent in this example is that it is rather mysterious how (8), which a priori is only a rational function of the z_i's, turns out to be a polynomial in the variables z_i^{-1}.

Perhaps this example created a somewhat wrong impression because in this case (8) was much more complicated than (7). So, let us consider the case $(g, n) = (1, 1)$, where the computation of the unique integral

$$\langle \tau_1 \rangle = \frac{1}{24}$$

already does require some work. The unique element of $\mathsf{G}^3_{1,1}$ is shown in Figure 14. This graph can be obtained by gluing the opposite sides of a hexagon, which also explains why the automorphism group of this graph is the cyclic group of order 6 (acting by rotations of the hexagon). Thus, (8) specializes in this case to

$$2\,\frac{1}{6}\,\frac{1}{(z_1+z_1)^3} = \frac{1}{24}\,\frac{1}{z_1^3} \,,$$

as it should.

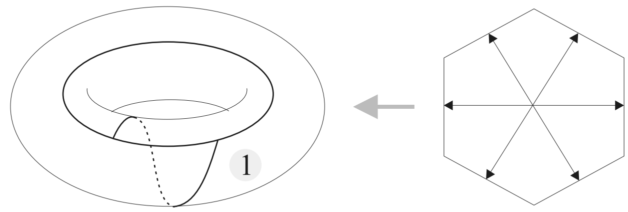


Fig. 14. The unique graph in $\mathsf{G}^3_{1,1}$

1.12

Kontsevich was led to the formula (8) by considering a cellular decomposition of $\mathcal{M}_{g,n}$ coming from Strebel differentials. In these lectures we shall explain, following [14], different approach to the formula (8) via the asymptotics in the Hurwitz problem of enumerating branched covering of $\mathbb{P}^1$. This approach is based on the relation between the Hurwitz problem and intersection theory on $\overline{\mathcal{M}}_{g,n}$ discovered in [3, 4] and on the asymptotic analysis developed in [10]. It has several advantages over the approach based on Strebel differentials.

2 Hurwitz problem

2.1

Intersection theory on $\overline{\mathcal{M}}_{g,n}$ is about enumerative geometry of families of stable n-pointed curves of genus g. The significance of the space $\overline{\mathcal{M}}_{g,n}$ is that its geometry captures some essential information about all possible families of curves. Through the space $\overline{\mathcal{M}}_{g,n}$, one can learn something about curves in general from any specific enumerative problem. If the specific enumerative problem is sufficiently rich, one can gather a lot of information about intersection theory on $\overline{\mathcal{M}}_{g,n}$ from it. Potentially, one can get a complete understanding of the whole intersection theory, which then can be applied to any other enumerative problem.

Our strategy will be to study such a particular yet representative enumerative problem. This specific problem will be the Hurwitz problem about branched covering of $\mathbb{P}^1$. That there exists a direct connection between Hurwitz problem and the intersection theory on $\overline{\mathcal{M}}_{g,n}$ was first realized in [3, 4]. The beautiful formula of [3] for the Hurwitz numbers will be the basis for our computations.

In fact, we will see that the (exact) knowledge of the numbers (5) is equivalent to the *asymptotics* in the Hurwitz problem. This is, in some sense, very fortunate because asymptotic enumeration problems often tend to be more structured and accessible than exact enumeration.

2.2

It is a century-old theme in combinatorics to enumerate branched coverings of a Riemann surface by another Riemann surface (an example of which is shown schematically in Figure 16). Given degree d, positions of ramification points downstairs, and their types (that is, given the conjugacy class in $S(d)$ of the monodromy around each one of them), there exist only finitely many possible coverings and the natural question is: how many? This very basic enumerative problem arises all over mathematics, from complex analysis to ergodic theory. These numbers of branched coverings are directly connected to other fundamental objects in combinatorics, namely to the class algebra of the symmetric group and — via the representation theory of finite groups — to the characters of symmetric groups.

We also mention that there is a general, and explicit, correspondence between enumeration of branched covering of a curve and the the Gromov–Witten theory of the same curve, see [15]. From this point of view, the computation of the numbers (5), that is, the Gromov–Witten theory of a point, arises as a limit in the Gromov–Witten theory of $\mathbb{P}^1$ as the degree goes to infinity. This is parallel to how the free energy (6) equation arises as the limit in the 1-matrix model.

2.3

The particular branched covering enumeration problem that we will be concerned with can be stated as follows. The data in the problem are a partition μ and genus g. Let

$$f : C \to \mathbb{P}^1$$

be a map of degree

$$d = |\mu| = \sum \mu_i \,,$$

where C is smooth connected complex curve of genus g. We require that $\infty \in \mathbb{P}^1$ is a critical value of the map f and the corresponding monodromy has cycle type μ. Equivalently, this can be phrased as the requirement that divisor $f^{-1}(\infty)$ has the form

$$f^{-1}(\infty) = \sum_{i=1}^{n} \mu_i \, [p_i] \,,$$

where $n = \ell(\mu)$ is the length of the partition μ and $p_1, \ldots, p_n \in C$ are the points lying over $\infty \in \mathbb{P}^1$. We further require that all other critical values of f are distinct and nondegenerate. In other words, the map f^{-1} has only square-root branch points in $\mathbb{P}^1 \setminus \{\infty\}$. The number r of such square-root branch points is given by the Riemann–Hurwitz formula

$$r = 2g - 2 + |\mu| + \ell(\mu) \,. \tag{9}$$

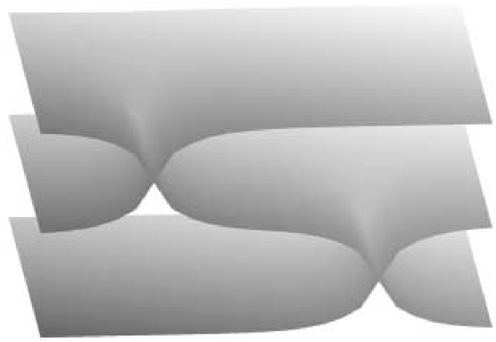


Fig. 15. A Hurwitz covering with $\mu = (3)$

An example of such a covering can be seen in Figure 15 where $\mu = (3)$ and $r = 2$, hence $d = 3$, $n = 1$, and $g = 0$.

We will call a covering satisfying the above conditions a *Hurwitz covering*. Once the positions of the r simple branchings are fixed, there are only finitely many Hurwitz coverings provided we identify two coverings

$$f : C \to \mathbb{P}^1 , \quad f' : C' \to \mathbb{P}^1$$

for which there exists an isomorphism $h : C \to C'$ such that $f = f' \circ h$. Similarly, we define automorphisms of f as an automorphisms $h : C \to C$ such that $f = f \circ h$. We will see that, with a very rare exception, Hurwitz coverings have only trivial automophisms.

By definition, the *Hurwitz number* $\mathrm{Hur}_g(\mu)$ is the number of isomorphism classes of Hurwitz coverings with given positions of branch points. In the special case when such a covering has a nontrivial automorphism, it should be counted with multiplicity $\frac{1}{2}$.

2.4

The Hurwitz problem can be restated as a problem about factoring permutations into transpositions. This goes as follows.

Let us pick a point $x \in \mathbb{P}^1$ which is not a ramification point. Then, by basic topology, all information about the covering is encoded in the homomorphism

$$\pi_1(\mathbb{P}^1 \setminus \{\text{ramification points}\}, x) \to \mathrm{Aut}\, f^{-1}(x) \cong S(d)\,.$$

The identification of $\mathrm{Aut}\, f^{-1}(x)$ with $S(d)$ here is not canonical, but it is convenient to pick any one of the $d!$ possible identifications. Then, by construction, the loop around ∞ goes to a permutation $s \in S(d)$ with cycle type μ and loops around finite ramification points correspond to some transpositions $t_1, \dots, t_r$ in $S(d)$.

The unique relation between those loops in π_1 becomes the equation

$$t_1 \cdots t_r = s\,. \tag{10}$$

This establishes the equivalence of the Hurwitz problem with the problem of factoring general permutations into transpositions (up to conjugation, since we picked an arbitrary identification of $\operatorname{Aut} f^{-1}(x)$ with $S(d)$). More precisely, the Hurwitz number $\mathrm{Hur}_g(\mu)$ is the number (up to conjugacy, and possibly with an automorphism factor) of factorization of the form (10) that correspond to a connected branched covering. A branched covering is connected when we can get from any point of $f^{-1}(x)$ to any other point by the action of the monodromy group. Thus, the transpositions $t_1, \ldots, t_n$ have to generate a transitive subgroup of $S(d)$, which is then automatically forced to be the whole of $S(d)$.

The fact that $t_1, \ldots, t_n$ generate $S(d)$ greatly constraints the possible automorphisms of f. Indeed, the action of any nontrivial automorphism on $f^{-1}(x)$ has to commute with $t_1, \ldots, t_n$, and hence with $S(d)$, which is only possible if $d = 2$.

By the usual inclusion–exclusion principle, it is clear that one can go back and forth between enumeration of connected and possibly disconnected coverings. Thus, the Hurwitz problem is essentially equivalent to decomposing the powers of one single element of the class algebra of the symmetric group, namely of the conjugacy class of a transposition

$$\sum_{1 \le i < j \le d} (ij) \tag{11}$$

in the standard conjugacy class basis. There is a classical formula, going back to Frobenius, for all such expansion coefficients in terms of irreducible characters. The character sums that one thus obtains can be viewed as finite analogs of Hermitian matrix integrals, with the dimension of a representation λ playing the role of the Vandermonde determinant and the central character of (11) in the representation λ playing the role of the Gaussian density, see, for example [11, 12] for a further discussion of properties of such sums.

2.5

For us, the crucial property of the Hurwitz problem is its connection with the intersection theory on the Deligne–Mumford spaces $\overline{\mathcal{M}}_{g,n}$. This connection was discovered, independently, in [4] and [3], the latter paper containing the following general formula

$$\mathrm{Hur}_g(\mu) = \frac{r!}{|\operatorname{Aut}\mu|} \prod_{i=1}^{n} \frac{\mu_i^{\mu_i}}{\mu_i!} \int_{\overline{\mathcal{M}}_{g,n}} \frac{1 - \lambda_1 + \cdots \pm \lambda_g}{\prod (1 - \mu_i \psi_i)}\,, \tag{12}$$

where r is number of branch points given by (9), $n = \ell(\mu)$ is the length of the partition μ, $\operatorname{Aut}\mu$ is the stabilizer of the vector μ in $S(n)$,

$$\lambda_i \in H^{2i}\left(\overline{\mathcal{M}}_{g,n}\right), \quad i = 1, \dots, g,$$

are the Chern classes of the Hodge bundle over $\overline{\mathcal{M}}_{g,n}$ (it is not important for what follows to know what this is), and finally, the denominators are supposed to be expanded into a geometric series

$$\frac{1}{1-\mu_i\psi_i} = 1 + \mu_i\psi_i + \mu_i^2\psi_i^2 + \dots, \tag{13}$$

which terminates because $\psi_i \in H^2\left(\overline{\mathcal{M}}_{g,n}\right)$ is nilpotent.

In particular, the integral in the ELSV formula (12) is a polynomial in the μ_i's. The monomials in this polynomial are obtained by picking a term in the expansion (13) for each $i = 1, \dots, n$ and then adding a suitable λ-class to bring the total degree to the dimension of $\overline{\mathcal{M}}_{g,n}$. It is, therefore, clear that the top degree term of this polynomial involves only intersections of the ψ-classes and no λ-classes. That is,

$$\int_{\overline{\mathcal{M}}_{g,n}} = \sum_{k_1+\dots+k_n=3g-3+n} \prod \mu_i^{k_i} \langle \tau_{k_1} \dots \tau_{k_n} \rangle + \text{lower degree}. \tag{14}$$

These top degree terms are precisely the numbers (5) that we want to understand.

2.6

A natural way to infer something about the top degree part of a polynomial is to let its arguments go to infinity. The behavior of the prefactors in (12) is given by the Stirling formula

$$\frac{m^m}{m!} \sim \frac{e^m}{\sqrt{2\pi m}}, \quad m \to \infty.$$

Let N be a large parameter and let μ_i depend on N in such a way that

$$\frac{\mu_i}{N} \to x_i, \quad i = 1, \dots, n, \quad N \to \infty,$$

where $x_1, \dots, x_n$ are finite. We will also additionally assume that all μ_i's are distinct and hence $|\operatorname{Aut}\mu| = 1$. Then by (14) and the Stirling formula, we have the following asymptotics of the Hurwitz numbers:

$$\frac{1}{N^{3g-3+n/2}} \frac{\operatorname{Hur}_g(\mu)}{e^{|\mu|} r!} \to \\ \frac{1}{(2\pi)^{n/2}} \sum_{k_1+\dots+k_n=3g-3+n} \prod \mu_i^{k_i - \frac{1}{2}} \langle \tau_{k_1} \dots \tau_{k_n} \rangle =: H_g(x). \tag{15}$$

It is convenient to Laplace transform the asymptotics $H_g(x)$. Since

$$\int_0^\infty e^{-sx}\, x^{k-1/2}\, dx = \frac{\Gamma(k+1/2)}{s^{k+1/2}} = \sqrt{\pi}\,\frac{(2k-1)!!}{2^k\, s^{k+1/2}}\,,$$

we get

$$\int_{\mathbb{R}^n_{>0}} e^{-s\cdot x}\, H_g(x)\, dx = \sum_{k_1+\cdots+k_n=3g-3+n} \langle \tau_{k_1} \dots \tau_{k_n} \rangle \prod \frac{(2k_i-1)!!}{(2s_i)^{k_i+1/2}}\,, \tag{16}$$

which up to the following change of variables

$$z_i = \sqrt{2s_i}\,, \quad i = 1, \dots, n\,,$$

is precisely the Kontsevich generating function (7) for the numbers (5).

Thus, we find ourselves in situation which looks rather comfortable: the generating function that we seek to compute is not only related to a specific enumerative problem but, in fact, it is the Laplace transform of the asymptotics in that enumerative problem. People who do enumeration know that asymptotics tends to be simpler than exact enumeration and, usually, the Laplace (or Fourier) transform of the asymptotics is the most natural thing to compute.

This general philosophy is, of course, only good if we can find a handle on the Hurwitz problem. In the following subsection, we will discuss a restatement of the Hurwitz problem in terms of enumeration of certain graphs on genus g surfaces that we call branching graphs. This description will turn out to be particularly suitable for our purposes (which may not be a huge surprise because, after all, Kontsevich's formula (8) is stated in terms of graphs on surfaces).

2.7

A very classical way to study branched coverings is to cut the base into simply-connected pieces. Over each of the resulting regions the covering becomes trivial, that is, consisting of d disjoint copies of the region downstairs, where d is the degree of the covering. The structure of the covering is then encoded in the information on how those pieces are patched together upstairs. Typically, this gluing data is presented in the form of a graph, usually with some additional labels etc.

There is, obviously, a considerable flexibility in this approach and some choices may lead to much more convenient graph enumeration problems than the others. For the Hurwitz problem, we will follow the strategy from [1], which goes as follows.

Let

$$f : C \to \mathbb{P}^1$$

be a Hurwitz covering with partition μ and genus g. In particular, the number r of finite ramification points of f is given by the formula (9). Without loss of generality, we can assume these ramification points to be rth roots of unity in $\mathbb{C}$. Let us cut the base $\mathbb{P}^1$ along the unit circle $S = \{|z| = 1\}$, that is, let us write

$$\mathbb{P}^1 = D_- \sqcup S \sqcup D_+$$

where

$$D_\pm = \{|z| \lesseqgtr 1\}$$

are the Southern and Northern hemisphere in Figure 16, respectively.

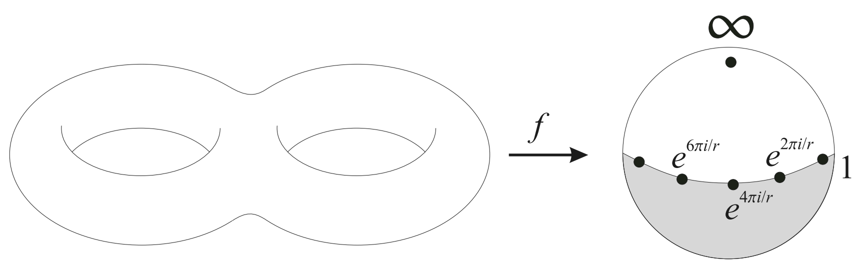


Fig. 16. A Hurwitz covering $f : \Sigma_2 \to \mathbb{P}^1$

Since the map f is unramified over D_-, its preimage $f^{-1}(D_-)$ consists of d disjoint disks. Their closures, however, are not disjoint: they come together precisely at the critical points of f. By construction, critical points of f are in bijection with its critical values, that is, with the rth roots of unity in $\mathbb{P}^1$. Thus, the the set $f^{-1}(\overline{D_-}) \subset C$ looks like the structure in Figure 17. This structure is, in fact, a graph Γ embedded in a genus g surface. Its vertices are the components of $f^{-1}(D_-)$ and its edges are the critical points of f that join those components together. In addition, the edges of Γ (there are r of them) are labeled by the roots of unity.

This edge-labeled graph $\Gamma \subset \Sigma_g$ is subject to some additional constraints. First, the cyclic order of labels around any vertex should be in agreement with the cyclic order of roots of unity. Next, the complement of Γ consists of n topological disks, where n is the length of the partition μ. Indeed, the complement of Γ corresponds to $f^{-1}(D_+)$ and $z = \infty$ is the only ramification point in D_+. The connected components of $f^{-1}(D_+)$ thus correspond to parts of μ.

The partition μ can be reconstructed from the edge labels of Γ as follows. Pick a cell U_i in $f^{-1}(D_+)$. The length of the corresponding part μ_i of μ is precisely the number of times the map f wraps the boundary ∂U_i around the circle S. As we follow the boundary ∂U_i, we see the edge labels appear in a certain sequence. As we complete a full circle around ∂U_i, the edge labels will make exactly μ_i turns around S. It is natural to call this number μ_i the *perimeter* of the cell U_i. This perimeter is $(2\pi)^{-1}$ times the sum of *angles* between pairs of the adjacent edges on ∂U_i, where the angle is the usual angle in $(0, 2\pi)$ between the corresponding roots of unity, see Figure 17.

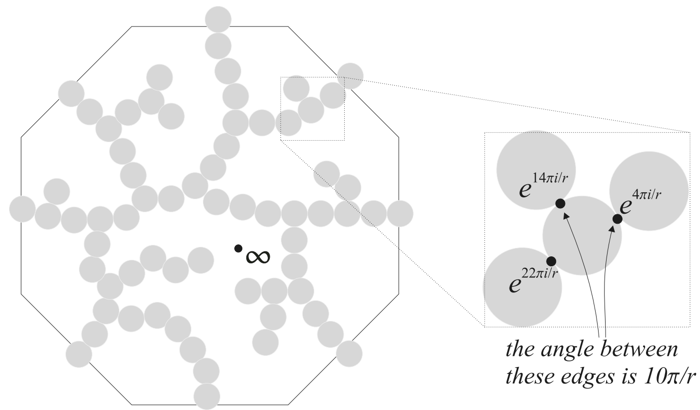


Fig. 17. Preimage $f^{-1}(D_-)$ on Σ_2

We call a edge-labeled embedded graph Γ as above a *branching graph*. By the above correspondence, the number $\mathrm{Hur}_g(\mu)$ is the number of genus g branching graphs with n cells of perimeter $\mu_1, \dots, \mu_n$. As usual, in the trivial $d = 2$ case, those graphs have to be counted with automorphism factors.

It is this definition of Hurwitz numbers that we will use for the asymptotic analysis in the next lecture.

2.8

It may be instructive to consider an example of how this correspondence between coverings and graphs works. Consider the covering corresponding to factorization

$$(12)\,(13)\,(24)\,(14)\,(13) = (1243)$$

of the form (10). The degree of this covering is $d = 4$, it has $r = 5$ ramification points, and the monodromy $\mu = (4)$ around infinity. It follows that its genus is $g = 1$. Let us denote the five finite ramification points by

$$\{a, b, c, d, e\} = \{1, e^{2\pi i/5}, \dots, e^{8\pi i/5}\} .$$

The preimage of D_- on the torus Σ_1 consists of 4 disks and the momodromies tell us which disk is connected to which at which critical point: for example, at the critical point lying over a, the 1st disk is connected to the 2nd disk. This is illustrated in Figure 18 where, among the 3 preimages of any critical value, the one which is a critical point is typeset in boldface. Clearly, any disk in $f^{-1}(D_-)$ has the alphabet $\{a, b, c, d, e\}$ going counterclockwise around its boundary and, in particular, the cyclic order of the critical values on its boundary is in agreement with the orientation on Σ_1.

Observe that the preimage $f^{-1}(D_+)$ is one cell whose boundary is a 4-fold covering of the equator. In particular, the alphabet $\{a, b, c, d, e\}$ is repeated 4

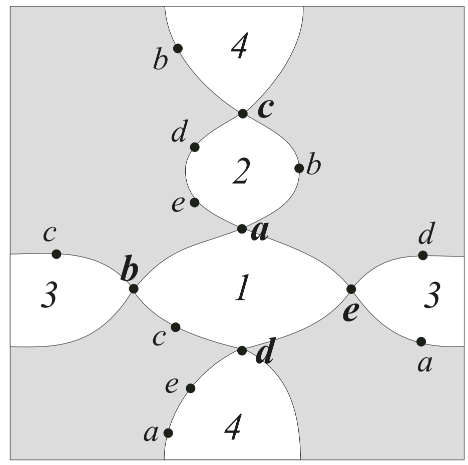


Fig. 18. Preimage of $f^{-1}(D_-)$ on the torus Σ_1

times around the boundary of $f^{-1}(D_+)$. Finally, Figure 19 shows the branching graph translation of Figure 18.

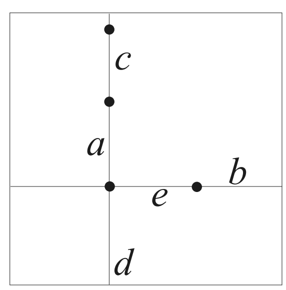


Fig. 19. Branching graph corresponding to Figure 18

2.9

Finally, a few remarks about how one can prove a formula like (12). This will necessary be a very sketchy account; the actual details of the proof can be found in [7, 14], as well as in the original paper [3].

As mentioned before, the numbers like $\mathrm{Hur}_g(\mu)$ a special case of in the integrals in the Gromov–Witten theory of $\mathbb{P}^1$, that is, certain intersections on the Kontsevich moduli space $\overline{\mathcal{M}}_{g,d,n}(\mathbb{P}^1)$ of stable degree d maps

$$f : C \to \mathbb{P}^1$$

from a varying n-pointed genus g domain curve C to the fixed target curve $\mathbb{P}^1$.

Since such a map can be composed with any automorphism of $\mathbb{P}^1$, we have a $\mathbb{C}^\times$-action on $\overline{\mathcal{M}}_{g,d,n}(\mathbb{P}^1)$. A theory due to Graber and Pandharipande [6] explains how to localize the integrals in Gromov–Witten theory to the fixed points of the action of the torus $\mathbb{C}^\times$. These fixed point loci in $\overline{\mathcal{M}}_{g,d,n}(\mathbb{P}^1)$ are, essentially, products of Deligne–Mumford spaces $\overline{\mathcal{M}}_{g_i,n_i}$ for some g_i's and n_i's. Indeed, only very few maps are fixed by the action of the torus. Namely, for the standard $\mathbb{C}^\times$-action on $\mathbb{P}^1$ and an irreducible domain curve C the only choices are the degree 0 constant maps to $\{0,\infty\} = \left(\mathbb{P}^1\right)^{\mathbb{C}^\times}$ or the degree d map

$$\mathbb{P}^1 \ni z \mapsto z^d \in \mathbb{P}^1 .$$

In general, the domain curve is allowed to be reducible, but still any torus-invariant map has to be of the above form on each component C_i of C. Once all discrete invariants of the curve C are fixed (that is, the combinatorics of its irreducible components, their genera and numbers of marked points on them) the remaining moduli parameters are only a choice of a bunch of curves to collapse plus a choice of where to attach the non-collapsed $\mathbb{P}^1$'s to them. That is, the torus-fixed loci are products of Deligne–Mumford spaces, modulo possible automorphisms of the combinatorial structure.

In this way integrals in the Gromov–Witten theory of $\mathbb{P}^1$ can be reduced, at least in principle, to computing intersections on $\overline{\mathcal{M}}_{g,n}$. An elegant localization analysis leading to the ELSV formula is presented in [7], see also [14].

3 Asymptotics in Hurwitz problem

3.1

Our goal now is to see how the Laplace transform (16) of the asymptotics (15) in the Hurwitz problem turns into Kontsevich's combinatorial formula (8). The formulation of the Hurwitz problem in terms of branching graphs, see Section 2.7, looks promising. Indeed, a branching graph Γ is by definition embedded in a topological genus g surface Σ_g and and it cuts Σ_g into n cells. Here the numbers g and n are the same as the indices in $\overline{\mathcal{M}}_{g,n}$, on the intersection theory on which we are trying to understand. Similarly, in Kontsevich's formula we have a graph G embedded into Σ_g and cutting it into n cells. This graph G, however, is a more modest object: it does not have any edge labels and it is allowed to have only 3-valent vertices.

Recall that we denote by $\mathsf{G}^3_{g,n}$ the set of all possible 3-valent graphs as in Kontsevich's formula (8). Let us introduce two larger sets

$$\mathsf{G}^3_{g,n} \subset \mathsf{G}^{\geq 3}_{g,n} \subset \mathsf{G}_{g,n} ,$$

on which, by definition, the 3-valence condition is weakened to allow vertices of valence 3 or more, and dropped altogether, respectively. The elements of

$\mathsf{G}_{g,n}^{\geq 3}$ can be obtained from elements of $\mathsf{G}_{g,n}^{3}$ by contracting some edges. In particular, the set $\mathsf{G}_{g,n}^{\geq 3}$ is still a finite set. Similarly, denote by $\mathsf{H}_{g,\mu}$ the set of all branching graphs with given genus g and perimeter partition μ. Our first order of business is to construct a map

$$\mathsf{H}_{g,\mu} \to \mathsf{G}_{g,n}^{\geq 3}\,,$$

which we call the *homotopy type* map. This map is the composition of the map

$$\mathsf{H}_{g,\mu} \to \mathsf{G}_{g,n}$$

which simply forgets the edge labels with the map

$$\mathsf{G}_{g,n} \to \mathsf{G}_{g,n}^{\geq 3}\,,$$

which does the following. First, we remove all univalent vertices together with the incident edge. After that, we remove the all remaining 2-valent vertices joining their two incident edges. What is left, by construction, has only vertices of valence 3 and higher and still cuts Σ_g into n cells.

We remark that in the two exceptional cases $(g, n) = (0, 1), (0, 2)$, which correspond to unstable moduli spaces, what we get in the end (a point and a circle, respectively) is not really an element of $\mathsf{G}_{g,n}^{\geq 3}$. In all other cases, however, we do get an honest element of $\mathsf{G}_{g,n}^{\geq 3}$. Figure 20 illustrates this procedure applied to the branching graph from Figure 17.

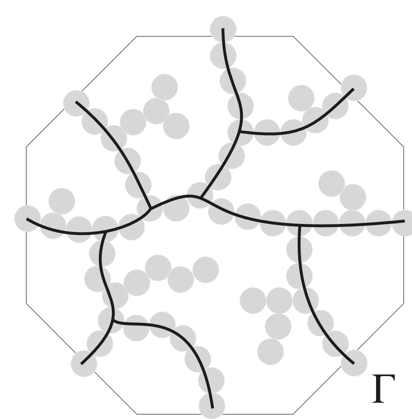


Fig. 20. The homotopy type of the branching graph from Figure 17

3.2

Now let us make the following simple but important observation. Since the set $\mathsf{G}_{g,n}^{\geq 3}$ is *finite* and we are interested in the asymptotics of $\mathrm{Hur}_g(\mu)$ as $\mu \to \infty$

while keeping g and n fixed, we can just do the asymptotics separately for each homotopy type and then sum over all possible homotopy types. The Laplace transform (16) will then be also expressed as a sum over all corresponding homotopy types G in $\mathsf{G}_{g,n}^{\geq 3}$.

We now claim that not only Kontsevich's combinatorial formula (8) is the Laplace transformed asymptotics (16) but, in fact, the summation over $G \in \mathsf{G}_{g,n}^{3}$ in Kontsevich's formula corresponds precisely to summation over possible homotopy types. Since there are non-trivalent homotopy types, implicit in this claim is the statement that *non-trivalent homotopy types do not contribute to asymptotics*.

3.3

What do we need to do to get the asymptotics of the number of branching graphs of a given homotopy type G? What would suffice is to have a simple way to enumerate all such branching graphs. To enumerate all branching graphs with given homotopy type G, we need to retrace the steps of the homotopy type map. Imagine that the homotopy type graph G is a fossil from which we want to reconstruct some prehistoric branching graph $\mathit{\Gamma}$. What are the all possible ways to do it?

The answer to all these rhetoric questions is quite simple. It is easy to see that the preimage of any edge in G is some subtree in the original branching graph $\mathit{\Gamma}$. In addition, all these trees carry edge-labels which were erased by the homotopy type map. Thus, for any edge e of G, we need to take a tree T_e whose edges are labeled by roots of unity. In particular, there is a canonical way to make this tree planar, that is, embed it in the plane in such a way that the cyclic order of edges around each vertex agrees with the order of their labels. In particular, each such tree is a *branching tree*, that is, it satisfies the $(g,n) = (0,1)$ case of our definition of a branching graph[3].

Next, these trees are to be glued into the graph $\mathit{\Gamma}$ by identifying some of their vertices, as in Figure 21. This means that each of these branching trees carries two special vertices, which we call its *root* and *top*. These special vertices of T_e mark the places where T_e is attached to the other trees in $\mathit{\Gamma}$. We will call such a branching tree with two marked vertices an *edge tree*.

3.4

Now we have a procedure which from a homotopy type G and a collection of edge trees $\{T_e\}$ with distinct labels assembles a branching graph $\mathit{\Gamma}$. This procedure, which we will call *assembly*, does have some imperfections. Those imperfections will be discussed momentarily, but first we want to make the following important observation.

[3] A small and inessential detail is that the labels T_e are taken from a larger set of roots of unity

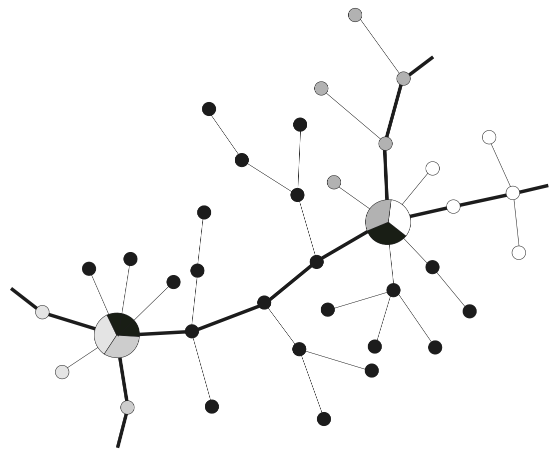


Fig. 21. Assembling a branching graph from edge trees

Since the homotopy type graph G is something fixed and finite, *the whole asymptotics of the branching graphs lies in the edge trees*. For a large random branching graph Γ, those edge trees will be large random trees. This is how the theory of random trees enters the scene. Fortunately, a large random tree is a very well studied and a very nicely behaved object, see for example [16] for a particularly enjoyable introduction. It turns out that all the information we need about random trees is either already classical or can be easily deduced from known results.

In fact, all required knowledge about random trees can be quite easily deduced (as was done in [14]) from the first principles, which in this case, is the following formula going back to Cayley [17]. Consider all possible trees T with the vertex set $\{1, \dots, m\}$. For any such tree T, we have a function $\mathrm{val}_T(i)$ which takes the vertex $i = 1, \dots, m$ to its valence in T. The information about all vertex valences in all possible trees T is encoded in the following generating function

$$\sum_T z_1^{\mathrm{val}_T(1)} \cdots z_m^{\mathrm{val}_T(m)} = z_1 \cdots z_m (z_1 + \cdots + z_m)^{m-2} . \tag{17}$$

A probabilistic restatement of this result is the following. The valence $\mathrm{val}_T(i)$ is the number of edges of T incident to the vertex i. Let us cut all edges in half; since there were $m-1$ edges of T, we get $2m-2$ half edges. The formula (17) says that the same distribution of half-edges can be obtained as follows: give every vertex a half-edge and the remaining $m-2$ edges just throw at the vertices randomly like darts.

What is then the valence of a given vertex in a random tree T? It is 1 for the half edge allowance that it always gets plus its share in the random distribution of $m-2$ darts among m targets. As $m \to \infty$, this share goes to a Poisson random variable with mean 1. In other words, as $m \to \infty$ we have

$$\mathrm{Prob}\{\mathrm{val}_T(i) = v\} \to \frac{e^{-1}}{(v-1)!}\,, \quad v = 1, 2, \dots\,. \tag{18}$$

For different vertices, their valences become independent in the $m \to \infty$ limit.

Also, setting all variables in (17) to 1 we find that the total number of trees with vertex set $\{1, \dots, m\}$ is m^{m-2}.

3.5

Now it is time to talk about how the assembly map differs from being one-to-one (it clear that it is onto).

First, it may happen that the cyclic order of edge labels is violated at one of the vertices of G where we patch together different edge trees. If this is the case, we simply declare the assembly to be a failure and do nothing. The probability of such an assembly failure in the large graph limit can be computed as follows. Suppose that we need to glue together three vertices with valences v_1, v_2, and v_3. From (18), the chance of seeing these particular valences is

$$\frac{e^{-3}}{(v_1-1)!(v_2-1)!(v_3-1)!}\,.$$

On the other hand, the conditional probability that the edge labels in the resulting graph are cyclically ordered, given that they were cyclically ordered before gluing is easily seen to be

$$\frac{(v_1-1)!(v_2-1)!(v_3-1)!}{(v_1+v_2+v_3-1)!}\,.$$

Hence the success rate of the assembly at a particular trivalent vertex is

$$e^{-3} \sum_{v_1, v_2, v_3 \geq 1} \frac{1}{(v_1+v_2+v_3-1)!} = \frac{e^{-2}}{2}\,.$$

Assembly failures at distinct vertices being asymptotically independent events, this goes into an overall factor and, eventually in the prefactor in (8).

At this point it should be clear that there in no need to consider nontrivalent vertices. Indeed, a homotopy type graph with a vertex of valence ≥ 4 can be obtained from a trivalent graph by contracting some edges, hence corresponds to the case when some of the edge graphs are trivial. It is obvious that the chances that a large random tree came out empty are negligible. Hence, nontrivalent graphs make indeed no contribution to the asymptotics and can safely be ignored.

3.6

The second (minor) issue with the assembly map is that we can get the same, that is, isomorphic branching graphs starting from different collections of the edge trees. This happens if the homotopy type graph G has nontrivial automorphisms. It is clear that the group $\mathrm{Aut}(G)$ acts on edges of G and, hence, acts by permutations on collections of edge trees preserving the isomorphism class of the assembly output. It is also clear that the chance for a large edge tree to be isomorphic to another edge tree (or to itself with root and top permuted) is, asymptotically, zero. Hence almost surely this $\mathrm{Aut}(G)$ action is free and hence there is an overcounting of branching graphs by exactly a factor of $|\mathrm{Aut}(G)|$. This explains the division by $|\mathrm{Aut}(G)|$ in (8).

3.7

Now, after explaining the summation over trivalent graphs and the automorphism factor in (8), we get to the heart of Kontsevich's formula — the product over the edges.

It is at this point that the convenience (promised in Section 2.6) of working with the Laplace transform (16) rather then the asymptotics (15) itself can be appreciated. We will see shortly that, asymptotically, the cell perimeters of a branching graph Γ assembled from a 3-valent graph G and bunch of random edge trees $\{T_e\}$ is a sum of independent contributions from each edge of G. This makes the Laplace transform (16) factor over the edges of G as in (8). To justify the above claim, we need to take a closer look at a large typical edge tree.

3.8

Let T be an edge tree. It has two marked vertices, root and top; let us call the path joining them the *trunk* of T. The tree T naturally splits into 3 parts: the root component, the top component, and the trunk component, according to their closest trunk point. This is illustrated in Figure 22.

Figure 22 may give a wrong idea of the relative size of these components for a typical large edge tree. Let $M \to \infty$ be the size (e.g. the number of vertices) in T. It is known and can be without difficulty deduced from (17) (see for example [14]) that the size M distributes itself among the three components of T in the $M \to \infty$ limit as follows.

First, the size of the root and the top component stays finite in the $M \to \infty$ limit. In fact, it goes to the *Borel distribution*, given by the following formula

$$\mathrm{Prob}(k) = \frac{k^{k-1}\, e^{-k}}{k!}\,, \quad k = 1, 2, \dots\,.$$

Second, the typical size of the trunk is of order $\sqrt{M}$. More precisely, scaled by $\sqrt{M}$, the trunk size distribution goes to the Rayleigh distribution with density

$$x\, e^{-x^2/2}\, dx\,, \quad x \in (0, \infty)\,.$$

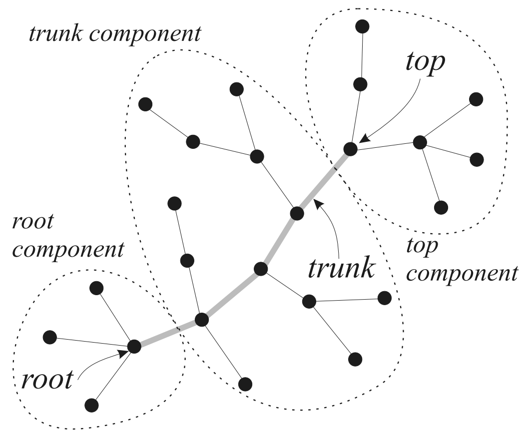


Fig. 22. The components of an edge tree

For our purposes, however, it only matters that the size of all these parts is $o(M)$ as $M \to \infty$.

The overwhelming majority of vertices lie, therefore, somewhere in the branches of the trunk component of T. What is very important is that, after assembly, any such vertex will find itself completely surrounded by a unique cell. As a result, it will contribute exactly 1 to that cell's perimeter. What this analysis shows it that, asymptotically, the cell perimeters are determined simply by the the number of such interior trunk vertices ending up in a given cell, all other contributions to perimeters being $o(M)$. It should be clear that such contributions of distinct edges of G are indeed independent, leading to the factorization in (8).

3.9

What remains is to determine is what the edge factors are, that is, to determine the actual contribution of an edge tree T to the perimeters of the adjacent cells.

All we need to know for this is to know how vertices in the trunk component distribute themselves between the two sides of the trunk, as in Figure 22. One shows, see [14] and below, that the fraction of the vertices that land on a given side of the trunk is, asymptotically, *uniformly distributed* on $[0, 1]$. This reduces the computation of the edge factor to computing one single integral. That computation will be presented in a moment, after we review the knowledge that we have accumulated so far.

3.10

Let G be a 3-valent map with n cells. It has

$$|E(G)| = 6g - 6 + 3n$$

edges and

$$|V(G)| = 4g - 4 + 2n$$

vertices, which follows from the Euler characteristic equation

$$|V(G)| - |E(G)| + n = 2 - 2g$$

combined with the 3-valence condition $3|V(G)| = 2|E(G)|$.

Let $e \in E(G)$ be an edge of G and let T_e be the corresponding edge tree. Let d_e be the number of vertices of T_e. Ignoring the few vertices on the trunk itself, the vertices of T_e distribute themselves between the two sides of the trunk of T_e. Let's say that p_e vertices are on the one side and define the number q_e by

$$p_e + q_e = d_e \,. \tag{19}$$

It is clear that q_e is the approximate number of vertices on the other side of the trunk. We call the numbers p_e and q_e the *semiperimeters* of the tree T_e.

The basic question, which we now can answer in the large graph limit, is how many branching graphs Γ have given semiperimeters $\{(p_e, q_e)\}_{e\in E(G)}$. This distribution can be computed asymptotically as follows.

3.11

First, there are some overall factors that come from automorphisms of G and the assembly success rate. Recall that in Section 3.5 we saw that the assembly success rate is $e^{-2|V(G)|}2^{-|V(G)|}$.

Second, for every edge $e \in E(G)$ we need to pick an edge tree T_e with d_e vertices. As we already learned from (17), the number of vertex-labeled trees with d_e vertices is $d_e^{d_e-2}$. Vertex labels can be traded for edge labels at the expense of the factor $d_e!/(d_e-1)! = d_e$, hence there are $d_e^{d_e-3}$ edge labeled trees with d_e vertices. The choice of the root vertex brings in additional factor of d_e choices. Once the root is fixed, the condition (19) dictates the position of the top, so there is no additional freedom in choosing it. To summarize, there are $\sim d_e^{d_e-2}$ edge trees with given semiperimeters p_e and q_e.

Third, the edge labels of Γ is a shuffle of edge labels of the trees T_e. Let

$$r = \sum_e (d_e - 1)$$

be the total number of edges in Γ (and, hence, also the total number of simple branch points in the Hurwitz covering corresponding to Γ). Obviously, there are

$$\frac{r!}{\prod_{e\in E(G)}(d_e-1)!} \tag{20}$$

ways to shuffle edge labels of $\{T_e\}$ into edge labels of Γ.

Putting it all together, we obtain the following approximate expression for the number of branching graphs with given semiperimeters $\{(p_e, q_e)\}_{e\in E(G)}$

$$\frac{r!\, e^{-2|V(G)|}\, 2^{-|V(G)|}}{|\operatorname{Aut}(G)|} \prod_{e\in E(G)} \frac{d_e^{d_e-2}}{(d_e-1)!} \sim \frac{r!\, e^{d}\, 2^{-|V(G)|}}{|\operatorname{Aut}(G)|} \prod_{e\in E(G)} \frac{1}{\sqrt{2\pi}\, d_e^{3/2}}\,, \tag{21}$$

where

$$d = |V(\Gamma)| = \sum_e d_e - 2|V(G)|$$

is the degree of the corresponding Hurwitz covering and the RHS of (21) is obtained from the LHS by the Stirling formula.

Note that the factor $r!\, e^d$ precisely cancels with prefactor in (15).

3.12

Since the cell perimeters of Γ are the sums of edge tree semiperimeters along the boundaries of the cells, the computation of the Laplace transform (16) indeed boils down to the computation of a singe edge factor

$$\begin{aligned}\frac{1}{\sqrt{2\pi}} \iint_{p,q>0} \frac{e^{-ps_1-qs_2}}{(p+q)^{3/2}}\, dp\, dq &= \frac{1}{\sqrt{2\pi}}\, \frac{1}{s_1-s_2} \int_{x>0} \left(e^{-s_1x} - e^{-s_2x}\right) \frac{dx}{x^{3/2}} \\ &= \frac{1}{\sqrt{2\pi}}\, \Gamma\left(-\tfrac{1}{2}\right) \frac{\sqrt{s_1}-\sqrt{s_2}}{s_1-s_2} \\ &= \frac{\sqrt{2}}{\sqrt{s_1}+\sqrt{s_2}}\,,\end{aligned}$$

where we set $x = p+q$.

Recall that the relation between the Laplace transform variables s_i in (16) and the variables z_i in Kontsevich's generation function 7 is

$$z_i = \sqrt{2s_i}\,, \quad i = 1, \dots, n\,.$$

Thus we get indeed the LHS of (8), including the correct exponent of 2, which is

$$|E(G)| - |V(G)| = 2g - 2 + n\,.$$

This completes the proof of Kontsevich's formula (8).

4 Remarks

4.1

Since random matrices are the common thread of many talks at this school, let us point out various connections between moduli of curves and random matrices. As we already discussed, the original KdV conjecture of Witten was based on physical parallelism between intersection theory on $\overline{\mathcal{M}}_{g,n}$ and the double scaling limit of the Hermitian 1-matrix model. Despite many spectacular achievements by physicists as well as mathematicians, this double scaling seems to remain a source of serious mathematical challenges, in particular, it appears that no direct mathematical connection between it and moduli of curves is known. On the other hand, there is a very direct connection between what we did and another, much simpler, matrix model, namely, the edge scaling of the standard GUE model. This connection goes as follows.

Recall that by Wick formula the coefficients of the $1/N$ expansion of the following $N \times N$ Hermitian matrix integral

$$\int e^{-\operatorname{tr} H^2} \prod_1^m \operatorname{tr} H^{k_i} \, dH \tag{22}$$

are the numbers of ways to glue a surface of given topology from m polygons with perimeters $k_1, \dots, k_n$. The double scaling limit of the 1-matrix model is concerned with gluing a given surface out of a very large number of small pieces. An opposite asymptotic regime is when the number m of pieces stays fixed while their sizes k_i go to infinity. Since for large k the trace $\operatorname{tr} H^k$ picks out the maximal eigenvalues of H, this asymptotic regime is about largest eigenvalues of a Hermitian random matrix. In the large N limit, the distribution of largest eigenvalues of H is well known to be the Airy ensemble. This edge scaling random matrix ensemble is very rich, yet susceptible to a very detailed mathematical analysis. In particular, the individual distributions of eigenvalues were found by Tracy and Widom in [18]. They are given in terms of certain solutions of the Painlevé II equation.

The connection between GUE edge scaling and what we were doing is the following. If one takes a branching graph as in Figure 17 and strips it off its edge labels, one gets a map on genus Σ_g with a few cells of very large perimeter, that is, an object of precisely the kind enumerated by (22) in the edge scaling regime. We argued that almost all vertices of a large branching Γ graph are completely surrounded by a unique cell, hence contribute exactly 1 to that cell's perimeter regardless of the edge labels. This shows that edge labels play no essential role in the asymptotics, thus establishing a direct connection between Hurwitz numbers asymptotics and GUE edge scaling. A similar direct connection can be established in other situations, for example, between GUE edge scaling and distribution of long increasing subsequences in a random permutation, see [10]. Since a great deal is known about GUE edge scaling, one can profit very easily from having a direct connection to it. In particular, one can give closed error-function-type formulas for a natural generating functions (known as n-point functions) for the numbers (5), see [13].

There exists another matrix model, namely the Kontsevich's matrix model [9], specifically designed to reproduce the graph summation in (8) as its diagrammatic expansion. Once the combinatorial formula (8) is established, this Kontsevich's model can be used to analyze it, in particular, to prove the KdV equations, see [9] and also [2].

Alternatively, the KdV equations can be pulled back from the GUE edge scaling (where they have been studied in depth by Adler, Shiota, and van Moerbeke) via the above described connection, see the exposition in [13].

4.2

In our approach, the intersections (5), the combinatorial formula (8), the KdV equations etc. appear through the asymptotic analysis of the Hurwitz problem. The ELSV formula (12), which is the bridge between enumeration of branched coverings and the intersection theory of $\overline{\mathcal{M}}_{g,n}$ is, on the hand, an exact formula. It is, therefore, natural to ask for more exact bridges between intersection theory, combinatorics, and integrable systems.

After the moduli space $\overline{\mathcal{M}}_{g,n}$ of stable curves, a natural next step is the Gromov–Witten theory of $\mathbb{P}^1$, that is, the intersection theory on the moduli space $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^1, d)$ of stable degree d maps

$$C \to \mathbb{P}^1$$

from an n-pointed genus g curve C to the projective line $\mathbb{P}^1$. More generally, one can replace $\mathbb{P}^1$ by some higher genus target curve X. It turns out, see [15], that there is a simple dictionary, which we call the Gromov–Witten–Hurwitz correspondence, between enumeration of branched coverings of $\mathbb{P}^1$ and Gromov–Witten theory of $\mathbb{P}^1$. This correspondence naturally connects with some very beautiful combinatorics and integrable systems, the role of random matrices now being played by random partitions. The connection with the integrable systems is seen best in the equivariant Gromov–Witten theory of $\mathbb{P}^1$, where the 2-Toda lattice hierarchy of Ueno and Takasaki plays the role that KdV played for $\overline{\mathcal{M}}_{g,n}$.

References

1. Arnold, V. I.: Topological classification of complex trigonometric polynomials and the combinatorics of graphs with an identical number of vertices and edges. Funct. Anal. Appl., **30**, no. 1, 1–14 (1996)
2. Di Francesco, P.: 2-D quantum and topological gravities, matrix models, and integrable differential systems. The Painlevé Property, Springer–Verlag, 229–285 (1999)
3. Ekedahl, T., Lando, S., Shapiro, M., Vainshtein, A.: Hurwitz numbers and intersections on moduli spaces of curves. Invent. Math., **146**, no. 2, 297–327 (2001)
4. Fantechi, B., Pandharipande, R. Stable maps and branch divisors. `math.AG/9905104`
5. Fulton, W.: Intersection theory. Springer–Verlag (1998)

6. Graber, T., Pandharipande, R.: Localization of virtual classes. Invent. Math., **135**, 487–518 (1999)
7. Graber, T., Vakil, R.: Hodge integrals and Hurwitz numbers via virtual localization. `math.AG/0003028`
8. Harris, J., Morrison, I.: Moduli of Curves. Springer–Verlag (1998)
9. Kontsevich, M.: Intersection theory on the moduli space of curves and the matrix Airy function. Commun. Math. Phys., **147**, 1–23 (1992)
10. Okounkov, A.: Random matrices and random permutations. IMRN, no. 20, 1043–1095 (2000)
11. Okounkov, A.: Infinite wedge and random partitions. Selecta Math., New Ser., **7**, 1–25 (2001)
12. Okounkov, A.: Toda equations for Hurwitz numbers. Math. Res. Lett., **7**, no. 4, 447–453 (2000)
13. Okounkov, A.: Generating functions for intersection numbers on moduli spaces of curves. IMRN, no. 18, 933–957 (2002)
14. Okounkov, A., Pandharipande, R.: Gromov–Witten theory, Hurwitz numbers, and matrix models, I. `math.AG/0101147`
15. Okounkov, A., Pandharipande, R.: in preparation
16. Pitman, J.: Enumeration of trees and forests related to branching processes and random walks. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, **41** (1998)
17. Stanley, R.: Enumerative combinatorics, Vol. 2. Cambridge Studies in Advanced Mathematics, **62**, Cambridge University Press, Cambridge (1999)
18. Tracy, C. A., Widom, H.: Level-spacing distributions and the Airy kernel. Commun. Math. Phys., **159**, 151–174 (1994)
19. Witten, E.: Two-dimensional gravity and intersection theory on moduli space. Surveys in Diff. Geom., **1**, 243–310 (1991)

An introduction to harmonic analysis on the infinite symmetric group

Grigori Olshanski

Dobrushin Mathematics Laboratory
Institute for Information Transmission Problems
Bolshoy Karetny 19
Moscow 101447, GSP–4
Russia
olsh@iitp.ru, olsh@online.ru

Introduction

The aim of the present survey paper is to provide an accessible introduction to a new chapter of representation theory — harmonic analysis for *noncommutative groups with infinite–dimensional dual space.*

I omitted detailed proofs but tried to explain the main ideas of the theory and its connections with other fields. The fact that irreducible representations of the groups in question depend on infinitely many parameters leads to a number of new effects which never occurred in conventional noncommutative harmonic analysis. A link with stochastic point processes is especially emphasized.

The exposition focuses on a single group, the infinite symmetric group $S(\infty)$. The reason is that presently this particular example is worked out the most. Furthermore, $S(\infty)$ can serve as a very good model for more complicated groups like the infinite–dimensional unitary group $U(\infty)$.

The paper is organized as follows. In §1, I explain what is the problem of harmonic analysis for $S(\infty)$. §§2–5 contain the necessary preparatory material. In §6, the main result is stated. It was obtained in a cycle of papers by Alexei Borodin and myself. In §7, the scheme of the proof is outlined. The final §8 contains additional comments and detailed references.

This paper is an expanded version of lectures I gave at the Euler Institute, St.–Petersburg, during the NATO ASI Program “Asymptotic combinatorics with applications to mathematical physics”. I also partly used the material of my lectures at the Weizmann Institute of Science, Rehovot. I am grateful to Anatoly Vershik, Amitai Regev, and Anthony Joseph for warm hospitality in St. Petersburg and Rehovot, and to Vladimir Berkovich for taking notes of my lectures at the Weizmann. Finally, I would like to thank Alexei Borodin for cooperation and help.

A. Borodin et al.: LNM 1817, A.M. Vershik (Ed.), pp. 127–160, 2003.

1 Virtual permutations and generalized regular representations

1.1 The Peter–Weyl theorem

Let $\mathcal{K}$ be a compact group, μ be the normalized Haar measure on $\mathcal{K}$ (i.e., $\mu(\mathcal{K})=1$), and H be the Hilbert space $L^2(K,\mu)$. The group $\mathcal{G}=\mathcal{K}\times\mathcal{K}$ acts on $\mathcal{K}$ on the right as follows: if $g=(g_1,g_2)\in G$ and $x\in\mathcal{K}$, then $x\cdot g=g_2^{-1}xg_1$. This action gives rise to a unitary representation T of $\mathcal{G}$ on H:

$$(T(g)f)(x)=f(x\cdot g), \qquad f\in H, \quad g\in G.$$

It is called the *biregular representation.* Let $\widehat{\mathcal{K}}$ be the set of equivalence classes of irreducible representations of $\mathcal{K}$. Recall that all of them are finite dimensional and unitarizable. For $\pi\in\widehat{\mathcal{K}}$, let $\overline{\pi}$ denote the dual representation. Since π is unitary, $\overline{\pi}$ is obtained from π by the conjugation automorphism of the base field $\mathbb{C}$.

Peter–Weyl's theorem. *The biregular representation T is equivalent to the direct sum of the irreducible representations of G of the form $\pi\otimes\overline{\pi}$,*

$$T\sim\bigoplus_{\pi\in\widehat{\mathcal{K}}}(\pi\otimes\overline{\pi}).$$

This is one of the first results of *noncommutative harmonic analysis.* The aim of noncommutative harmonic analysis can be stated as *decomposing natural representations into irreducible ones.* The biregular representation can be called a natural representation because it is fabricated from the group itself in a very natural way. The Peter–Weyl theorem serves as a guiding example for more involved theories of noncommutative harmonic analysis.

1.2 The infinite symmetric group

Let $S(n)$ be the symmetric group of degree n, i.e., the group of permutations of the set $\{1,\dots,n\}$. By the very definition, $S(n)$ acts on $\{1,\dots,n\}$. The stabilizer of n is canonically isomorphic to $S(n-1)$, which makes it possible to define, for any $n=2,3,\dots$, an embedding $S(n-1)\to S(n)$. Let $S(\infty)$ be the inductive limit of the groups $S(n)$ taken with respect to these embeddings. We call $S(\infty)$ the *infinite symmetric group.*

Clearly, $S(\infty)$ is a countable, locally finite group. It can be realized as the group of all *finite* permutations of the set $\{1,2,\dots\}$.

1.3 The biregular representation for $S(\infty)$

The definition of a biregular representation given in §1.1 evidently makes sense for the group $S(\infty)$. Namely, set $\mathcal{K}=S(\infty)$, $\mathcal{G}=S(\infty)\times S(\infty)$, and take as

μ the counting measure on $S(\infty)$. Then the unitary representation T of the group $S(\infty) \times S(\infty)$ in the Hilbert space $L^2(S(\infty), \mu)$ is defined by exactly the same formula as in §1.1.

Proposition. *The biregular representation T of the group $S(\infty) \times S(\infty)$ is irreducible.*

Sketch of proof. Let $\mathrm{diag}(S(\infty))$ be the image of $S(\infty)$ under the diagonal embedding $S(\infty) \to S(\infty) \times S(\infty)$. The Dirac function δ_e is a unique (up to a constant factor) $\mathrm{diag}(S(\infty))$–invariant vector in the space of T. This follows from the fact that all conjugacy classes in $S(\infty)$, except for $\{e\}$, are infinite. On the other hand, δ_e is a cyclic vector, i.e., it generates under the action of $S(\infty) \times S(\infty)$ a dense subspace in $L^2(S(\infty))$. It follows that there is no proper closed $S(\infty) \times S(\infty)$–invariant subspace.

Thus, in the case of the group $S(\infty)$, the naive analog of the biregular representation is of no interest for harmonic analysis. We will explain how to modify the construction in order to get interesting representations.

From now on we are using the notation

$$G = S(\infty) \times S(\infty), \qquad K = \mathrm{diag}(S(\infty)).$$

We call G the *infinite bisymmetric group.*

1.4 Virtual permutations

Note that in the construction of §1.1, the group $\mathcal{K}$ plays two different roles: it is the carrier of a Hilbert space of functions and it acts (by left and right shifts) in this space. The idea is to separate these two roles. As the carrier of a Hilbert space we will use a remarkable compactification $\mathfrak{S}$ of $S(\infty)$. It is not a group but still a G–space, which is sufficient for our purposes.

For any $n \geq 2$, we define a projection $p_n : S(n) \to S(n-1)$ as removing the element n from the cycle containing it. That is, given a permutation $\sigma \in S(n)$, if n is fixed under σ then $p_n(\sigma) = \sigma$, and if n enters a nontrivial cycle $(\cdots \to i \to n \to j \to \cdots)$ then we simply replace this cycle by $(\cdots \to i \to j \to \cdots)$. We call p_n the *canonical projection*.

Proposition. *The canonical projection $p_n : S(n) \to S(n-1)$ commutes with the left and right shifts by the elements of $S(n-1)$. Moreover, for $n \geq 5$ it is the only map $S(n) \to S(n-1)$ with such a property.*

Let $\mathfrak{S}$ be the projective limit of the finite sets $S(n)$ taken with respect to the canonical projections. Any point $x \in \mathfrak{S}$ is a collection $(x_n)_{n \geq 1}$ such that $x_n \in S(n)$ and $p_n(x_n) = x_{n-1}$. For any m, we identify $S(m)$ with the subset of those points $x = (x_n)$ for which $x_n \in S(m)$ for all $n \geq m$. This allows us to embed $S(\infty)$ into $\mathfrak{S}$.

We equip $\mathfrak{S}$ with the projective limit topology. In this way we get a totally disconnected compact topological space. We call it the *space of virtual permutations.*

The image of $S(\infty)$ is dense in $\mathfrak{S}$. Hence, $\mathfrak{S}$ is a *compactification* of the discrete space $S(\infty)$.

There exists an action of the group G on the space $\mathfrak{S}$ by homeomorphisms extending the action of G on $S(\infty)$. Such an action is unique.

There are several different realizations of the space $\mathfrak{S}$. One of them looks as follows. Set $I_n = \{0, \dots, n-1\}$. There exists a bijection

$$\mathfrak{S} \to I := I_1 \times I_2 \times \dots, \qquad x = (x_1, x_2, \dots) \mapsto (i_1, i_2, \dots)$$

such that $i_n = 0$, if $x_n(n) = n$, and $i_n = j$, if $x_n(n) = j < n$. This bijection is a homeomorphism (here we equip the product space I with the product topology). It gives rise, for every $n \geq 1$, to a bijection $S(n) \to I_1 \times \dots \times I_n$. In this realization, the canonical projection $p_n : S(n) \to S(n-1)$ turns into the natural projection $I_1 \times \dots \times I_n \to I_1 \times \dots \times I_{n-1}$.

1.5 Ewens' measures on $\mathfrak{S}$

Let $\mu_1^{(n)}$ be the normalized Haar measure on $S(n)$. Its pushforward under the canonical projection p_n coincides with the measure $\mu_1^{(n-1)}$, because p_n commutes with the left (and right) shifts by elements of $S(n-1)$. Thus, the measures $\mu_1^{(n)}$ are pairwise *consistent* with respect to the canonical projections. Hence, we can define their projective limit, $\mu_1 = \varprojlim \mu_1^{(n)}$, which is a probability measure on $\mathfrak{S}$.

The measure μ_1 is invariant under the action of G, and it is the only probability measure on $\mathfrak{S}$ with this property. Thus, viewing $\mathfrak{S}$ as a substitute of the group space, we may view μ_1 as a substitute of the normalized Haar measure.

Now we define a one–parameter family of probability measures containing the measure μ_1 as a particular case.

For $t \geq 0$, let $\mu_t^{(n)}$ be the following measure on $S(n)$:

$$\mu_t^{(n)}(x) = \frac{t^{[x]-1}}{(t+1)(t+2) \cdot \dots \cdot (t+n-1)},$$

where $[x] = [x]_n$ is the number of cycles of x in $S(n)$. If $t = 1$ then this reduces to above definition of the measure $\mu_1^{(n)}$.

Proposition. (i) $\mu_t^{(n)}$ *is a probability measure on* $S(n)$, *i.e.,*

$$\sum_{x \in S(n)} t^{[x]} = t(t+1) \cdot \dots \cdot (t+n-1).$$

(ii) *The measures* $\mu_t^{(n)}$ *are pairwise consistent with respect to the canonical projections.*

(iii) *The pushforward of* $\mu_t^{(n)}$ *under the bijective map* $S(n) \to I_1 \times \cdots \times I_n$ *of §1.4 is the product measure* $\nu_t^{(1)} \times \cdots \times \nu_t^{(n)}$*, where, for any* m*,* $\nu_t^{(m)}$ *is the following probability measure on* I_m*:*

$$\nu_t^{(m)}(i) = \begin{cases} \frac{t}{t+m-1}, & i = 0 \\ \frac{1}{t+m-1}, & i = 1, \dots, m-1. \end{cases}$$

Proof. (i) Induction on n. Assume that the equality in question holds for $n-1$. Notice that $[p_n(x)]_{n-1}$ is equal to $[x]_n$ when $x \notin S(n-1)$, and to $[x]_n - 1$ when $x \in S(n-1)$. We have

$$\begin{aligned} \sum_{x \in S(n)} t^{[x]} &= \sum_{y \in S(n-1)} \sum_{p_n(x)=y} t^{[x]} = \sum_{y \in S(n-1)} \left(t \cdot t^{[y]} + (n-1) t^{[y]} \right) \\ &= t \cdot t(t+1) \cdot \dots \cdot (t+n-2) + (n-1)t(t+1) \cdot \dots \cdot (t+n-2) \\ &= t(t+1) \cdot \dots \cdot (t+n-1). \end{aligned}$$

(ii) We have to verify that for every $y \in S(n-1)$

$$\frac{t^{[y]-1}(t+1) \cdot \dots \cdot (t+n-2)}{=} \sum_{p_n(x)=y} \frac{t^{[x]-1}(t+1) \cdot \dots \cdot (t+n-1)}{}.$$

It is precisely what is done in the proof of (i).

(iii) This follows from the fact that, under the bijection $x \mapsto (i_1, \dots, i_n)$ between $S(n)$ and $I_1 \times \cdots \times I_n$, the number of zeros in $(i_1, \dots, i_n)$ equals $[x]$.

The consistency property makes it possible to define, for any $t \geq 0$, a probability measure $\mu_t = \varprojlim \mu_t^{(n)}$ on $\mathfrak{S}$. This measure is invariant under the diagonal subgroup K but is not G–invariant (except the case $t = 1$). As $t \to \infty$, μ_t tends to the Dirac measure at $e \in S(\infty) \subset \mathfrak{S}$. Let us denote this limit measure by μ_∞.

Following S. V. Kerov, we call the measures μ_t the *Ewens measures*. The next claim gives a characterization of the family $\{\mu_t\}$.

Proposition. *The measures* μ_t*, where* $0 \leq t \leq \infty$*, are exactly those probability measures on* $\mathfrak{S}$ *that are* K*–invariant and correspond to product measures on* $I_1 \times I_2 \times \dots$ *.*

1.6 Transformation properties of the Ewens measures

Recall that $[\sigma]_n$ denotes the number of cycles of a permutation $\sigma \in S(n)$.

Proposition. (i) *For any* $x = (x_n) \in \mathfrak{S}$ *and* $g \in G$*, the quantity* $[x_n \cdot g]_n - [x_n]_n$ *does not depend on* n *provided that* n *is large enough.*

(ii) *Denote by* $c(x,g)$ *the stable value of this quantity. The function* $c(x,g)$ *is an additive cocycle with values in* $\mathbb{Z}$*, that is,*

$$c(x,gh) = c(x\cdot g, h) + c(x,g), \qquad x \in \mathfrak{S}, \quad g,h \in G.$$

Recall that a measure is called *quasi–invariant* under a group of transformations if, under the shift by an arbitrary element of the group, the measure is transformed to an equivalent measure.

Proposition. *Assume* $t \in (0,+\infty)$.

(i) *The measure* μ_t *is quasi–invariant under the action of the group* G.

(ii) *We have*

$$\frac{\mu_t(d(x\cdot g))}{\mu_t(dx)} = t^{c(x,g)},$$

where the left–hand side is the Radon–Nikodym derivative.

Note that $c(x,g)=1$ whenever $g \in K$. This agrees with the fact that the measures are K–invariant.

1.7 The representations T_z

We start with a general construction of unitary representations related to group actions on measure spaces with cocycles.

Assume we are given a space $\mathcal{S}$ equipped with a Borel structure (i.e., a distinguished sigma–algebra of sets), a discrete group $\mathcal{G}$ acting on $\mathcal{S}$ on the right and preserving the Borel structure, and a Borel measure μ, which is quasi–invariant under $\mathcal{G}$. A complex valued function $\tau(x,g)$ on $\mathcal{S}\times\mathcal{G}$ is called a *multiplicative cocycle* if

$$\tau(x,gh) = \tau(x\cdot g,h)\tau(x,g), \qquad x\in\mathcal{S}, \quad g,h\in\mathcal{G}.$$

Next, assume we are given a multiplicative cocycle $\tau(x,g)$ which is a Borel function in x and which satisfies the relation

$$|\tau(x,g)|^2 = \frac{\mu(d(x\cdot g))}{\mu(dx)}\,.$$

Then these data allow us to define a unitary representation $T = T_\tau$ of the group $\mathcal{G}$ acting in the Hilbert space $L^2(\mathcal{S},\mu)$ according to the formula

$$(T(g)f)(x) = \tau(x,g)f(x\cdot g), \qquad f\in L^2(\mathcal{S},\mu), \quad x\in\mathcal{S}, \quad g\in\mathcal{G}.$$

Let $z\in\mathbb{C}$ be a nonzero complex number. We apply this general construction for the space $\mathcal{S}=\mathfrak{S}$, the group $\mathcal{G}=G$, the measure $\mu=\mu_t$ (where $t=|z|^2$), and the cocycle $\tau(x,g)=z^{c(x,g)}$. All the assumptions above are satisfied, so that we get a unitary representation $T=T_z$ of the group G.

Using a continuity argument it is possible to extend the definition of the representations T_z to the limit values $z=0$ and $z=\infty$ of the parameter z. It turns out that the representation T_∞ is equivalent to the biregular representation of §1.3. Thus, the family $\{T_z\}$ can be viewed as a deformation of the biregular representation.

We call the T_z's the *generalized regular representations.* These representations are reducible (with the only exception of T_∞). Now we can state the main problem that we address in this paper.

Problem of harmonic analysis on $S(\infty)$**.** *Describe the decomposition of the generalized regular representations* T_z *into irreducibles ones.*

2 Spherical representations and characters

2.1 Spherical representations

By a *spherical representation* of the pair (G,K) we mean a pair (T,ξ), where T is a unitary representation of G and ξ is a unit vector in the Hilbert space $H(T)$ such that:

(i) ξ is K–invariant and

(ii) ξ is cyclic, i.e., the span of the vectors of the form $T(g)\xi$, where $g\in G$, is dense in $H(T)$.

We call ξ the *spherical vector.* We call two spherical representations (T_1,ξ_1) and (T_2,ξ_2) *equivalent* if there exists an isometric isomorphism between their Hilbert spaces which commutes with the action of G and preserves the spherical vectors. Such an isomorphism is unique within multiplication by a scalar. The equivalence $(T_1,\xi_1)\sim(T_2,\xi_2)$ implies the equivalence $T_1\sim T_2$ but the converse is not true in general.

The matrix coefficient $(T(g)\xi,\xi)$, where $g\in G$, is called the *spherical function.* Two spherical representations are equivalent if and only if their spherical functions coincide.

We aim to give an independent characterization of spherical functions for (G,K).

2.2 Positive definite functions

Recall that a complex–valued function f on a group $\mathcal{G}$ is called *positive definite* if:

(i) $f(g^{-1})=\overline{f(g)}$ for any $g\in\mathcal{G}$ and

(ii) for any finite collection $g_1,\dots,g_n$ of elements of $\mathcal{G}$, the $n\times n$ Hermitian matrix $[f(g_j^{-1}g_i)]$ is nonnegative.

Positive definite functions on $\mathcal{G}$ are exactly diagonal matrix coefficients of unitary representations of $\mathcal{G}$.

Now return to our pair (G,K). The spherical functions for (G,K) can be characterized as the *positive definite,* K*–biinvariant functions on* G*, normalized at* $e\in G$.

2.3 Characters

Recall that the character of an irreducible representation π of a compact group $\mathcal{K}$ is the function $g \mapsto \chi^\pi(g) = \mathrm{Tr}(\pi(g))$. If $\mathcal{K}$ is noncompact, an irreducible representation π of $\mathcal{K}$ is not necessarily finite dimensional, and so the function $g \mapsto \mathrm{Tr}(\pi(g))$ does not make sense in general. But it turns out that in certain cases the ratio

$$\widetilde{\chi}^\pi(g) = \frac{\chi^\pi(g)}{\chi^\pi(e)}$$

does make sense.

Let $\mathcal{K}$ be an arbitrary group. A function on $\mathcal{K}$ is said to be *central* if it is constant on conjugacy classes. Denote by $\mathcal{X}(K)$ the set of central, positive definite, normalized functions on $\mathcal{K}$ (if $\mathcal{K}$ is a topological group then we additionally require the functions to be continuous). If $\varphi, \psi \in \mathcal{X}(\mathcal{K})$, then for every $t \in [0,1]$ the function $(1-t)\varphi + t\psi$ is also an element of $\mathcal{X}(\mathcal{K})$, i.e., $\mathcal{X}(\mathcal{K})$ is a convex set.

Recall that a point of a convex set is called *extreme* if it is not contained in the interior of an interval entirely contained in the set. Let $\mathrm{Ex}(\mathcal{X}(\mathcal{K}))$ denote the subset of extreme points of $\mathcal{X}(\mathcal{K})$.

If the group $\mathcal{K}$ is compact then the functions from $\mathrm{Ex}(\mathcal{X}(\mathcal{K}))$ are exactly the *normalized* irreducible characters $\widetilde{\chi}^\pi(g)$, where $\pi \in \widehat{\mathcal{K}}$. As for general elements of $\mathcal{X}(\mathcal{K})$, they are (possibly infinite) convex linear combinations of these functions.

In particular, if $\mathcal{K}$ is finite then $\mathcal{X}(\mathcal{K})$ is a finite–dimensional simplex.

We will call the elements of $\mathcal{X}(\mathcal{K})$ the *characters* of $\mathcal{K}$. The elements of $\mathrm{Ex}(\mathcal{X}(\mathcal{K}))$ will be called the *extreme characters*. Notice that this terminology does not agree with the conventional terminology of representation theory. However, in the case of the group $S(\infty)$ this will not lead to a confusion.

2.4 Correspondence between spherical representations of (G, K) and characters of $S(\infty)$

There is a natural 1–1 correspondence between spherical functions for (G, K) and characters of $S(\infty)$. Specifically, given a function f on the group $G = S(\infty) \times S(\infty)$, let χ be the function on $S(\infty)$ obtained by restricting f to the first copy of $S(\infty)$. Then $f \mapsto \chi$ establishes a 1–1 correspondence between K–biinvariant functions on G and central functions on $S(\infty)$. Moreover, this correspondence preserves the positive definiteness property. This implies that the equivalence classes of spherical representations of (G, K) are parametrized by the characters of $S(\infty)$.

Proposition. *Let T be a unitary representation of G and $H(T)^K$ be the subspace of K–invariant vectors in the Hilbert space $H(T)$ of T.*

If T is irreducible then $H(T)^K$ has dimension 0 or 1. Conversely, if the subspace $H(T)^K$ has dimension 1 then its cyclic span is an irreducible subrepresentation of T.

Corollary. *For an irreducible spherical representation of (G, K), the spherical vector ξ is defined uniquely, within a scalar multiple, which does not affect the spherical function.*

A spherical function corresponds to an irreducible representation if and only if the corresponding character is extreme. Thus, the (equivalence classes of) irreducible spherical representations of (G, K) are parametrized by extreme characters of $S(\infty)$.

2.5 Spectral decomposition

Proposition. (i) *For any character $\psi \in \mathcal{X}(S(\infty))$, there exists a probability measure P on the set $\mathrm{Ex}(\mathcal{X}(S(\infty)))$ of extreme characters such that*

$$\psi(\sigma) = \int_{\chi \in \mathrm{Ex}(\mathcal{X}(S(\infty)))} \chi(\sigma) P(d\chi), \qquad \sigma \in S(\infty).$$

(ii) *Such a measure is unique.*

(iii) *Conversely, for any probability measure P on the set of extreme characters, the function ψ defined by the above formula is a character of $S(\infty)$.*

We call this integral representation the *spectral decomposition* of a character. The measure P will be called the *spectral measure* of ψ. If ψ is extreme then its spectral measure reduces to the Dirac mass at ψ.

Let (T, ξ) be a spherical representation of (G, K), ψ be the corresponding character, and P be its spectral measure. If ξ is replaced by another spherical vector in the same representation then the character ψ is changed, hence the measure P is changed, too. However, P is transformed to an equivalent measure. Thus, the equivalence class of P is an invariant of T as a unitary representation.

The spectral decomposition of ψ determines a decomposition of the representation T into a continual integral of irreducible spherical representations.

3 Thoma's theorem and spectral decomposition of the representations T_z with $z \in \mathbb{Z}$

3.1 First example of extreme characters

Let $\alpha = (\alpha_1 \geq \cdots \geq \alpha_p \geq 0)$ and $\beta = (\beta_1 \geq \cdots \geq \beta_q \geq 0)$ be two collections of numbers such that

$$\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j = 1.$$

Here one of the numbers p, q may be zero (then the corresponding collection α or β disappears). To these data we will assign an extreme character $\chi^{(\alpha,\beta)}$ of $S(\infty)$, as follows.

Let

$$p_k(\alpha,\beta)=\sum_{i=1}^{p}\alpha_i^k+(-1)^{k-1}\sum_{j=1}^{q}\beta_j^k.$$

Note that

$$p_1(\alpha,\beta)\equiv 1.$$

Given $\sigma\in S(\infty)$, we denote by $m_k(\sigma)$ the number of k–cycles in σ. Since σ is a finite permutation, we have

$$m_1(\sigma)=\infty,\qquad m_k(\sigma)<\infty \text{ for } k\ge 2,\qquad m_k(\sigma)=0 \text{ for } k \text{ large enough.}$$

In this notation, we set

$$\chi^{(\alpha,\beta)}(\sigma)=\prod_{k=1}^{\infty}(p_k(\alpha,\beta))^{m_k(\sigma)}=\prod_{k=2}^{\infty}(p_k(\alpha,\beta))^{m_k(\sigma)},\qquad \sigma\in S(\infty),$$

where we agree that $1^\infty=1$ and $0^0=1$.

Proposition. *Each function $\chi^{(\alpha,\beta)}$ defined by the above formula is an extreme character of $S(\infty)$.*

If $p=1$ and $q=0$ (i.e., $\alpha_1=1$ and all other parameters disappear) then we get the trivial character, which equals 1 identically. If $p=0$ and $q=1$ then we get the alternate character $\mathrm{sgn}(\sigma)=\pm 1$, where the plus–minus sign is chosen according to the parity of the permutation. More generally, we have

$$\chi^{(\alpha,\beta)}\cdot\mathrm{sgn}=\chi^{(\beta,\alpha)}.$$

3.2 Thoma's set

Let $\mathbb{R}^\infty$ denote the direct product of countably many copies of $\mathbb{R}$. We equip $\mathbb{R}^\infty$ with the product topology. Let Ω be the subset of $\mathbb{R}^\infty\times\mathbb{R}^\infty$ formed by couples $\alpha\in\mathbb{R}^\infty$, $\beta\in\mathbb{R}^\infty$ such that

$$\alpha=(\alpha_1\ge\alpha_2\ge\cdots\ge 0),\quad \beta=(\beta_1\ge\beta_2\ge\cdots\ge 0),\quad \sum_{i=1}^{\infty}\alpha_i+\sum_{j=1}^{\infty}\beta_j\le 1.$$

We call Ω the *Thoma set.* We equip it with topology induced from that of the space $\mathbb{R}^\infty\times\mathbb{R}^\infty$. It is readily seen that Ω is a compact space.

The couples (α,β) that we dealt with in §3.1 can be viewed as elements of Ω. The subset of such couples (with given p,q) will be denoted by Ω_{pq}.

Note that each Ω_{pq} is isomorphic to a simplex of dimension $p+q-1$. As affine coordinates of the simplex one can take the numbers

$$\alpha_1-\alpha_2,\ \ldots,\ \alpha_{p-1}-\alpha_p,\ \alpha_p,\ \beta_1-\beta_2,\ \ldots,\ \beta_{q-1}-\beta_q,\ \beta_q$$

but we will not use these coordinates.

Proposition. *The union of the simplices Ω_{pq} is dense in Ω.*

For instance, the point $(\underline{0},\underline{0}) = (\alpha \equiv 0, \beta \equiv 0) \in \Omega$ can be approximated by points of the simplices Ω_{p0} as $p \to \infty$,

$$(\underline{0},\underline{0}) = \lim_{p\to\infty}((\underbrace{1/p,\dots,1/p}_{p}),\underline{0}).$$

3.3 Description of extreme characters

Now we extend by continuity the definition of §3.1. For any $k = 2, 3, \dots$ we define the function p_k on Ω as follows. If $\omega = (\alpha, \beta) \in \Omega$ then

$$p_k(\omega) = p_k(\alpha,\beta) = \sum_{i=1}^{\infty} \alpha_i^k + (-1)^{k-1} \sum_{j=1}^{\infty} \beta_j^k.$$

Note that p_k is a continuous function on Ω. It should be emphasized that the condition $k \geq 2$ is necessary here: the similar expression with $k = 1$ (that is, the sum of all coordinates) *is not* continuous.

Next, for any $\omega = (\alpha, \beta) \in \Omega$ we set

$$\chi^{(\omega)}(\sigma) = \chi^{(\alpha,\beta)}(\sigma) = \prod_{k=2}^{\infty} (p_k(\alpha,\beta))^{m_k(\sigma)}, \qquad \sigma \in S(\infty),$$

Thoma's theorem. (i) *For any $\omega \in \Omega$ the function $\chi^{(\omega)}$ defined above is an extreme character of $S(\infty)$.*

(ii) *Each extreme character is obtained in this way.*

(iii) *Different points $\omega \in \Omega$ define different characters.*

In particular, the character $\chi^{(\underline{0},\underline{0})}$ is the delta function at $e \in S(\infty)$. It corresponds to the biregular representation defined in §1.3.

Note that the topology of Ω agrees with the topology of pointwise convergence of characters on $S(\infty)$. This implies, in particular, that the characters of §3.1 are dense in the whole set of extreme characters with respect to the topology of pointwise convergence.

Corollary. *For any character ψ of $S(\infty)$, its spectral measure P can be viewed as a probability measure on the compact space Ω, and the integral representation of §2.5 can be rewritten in the following form*

$$\psi(\sigma) = \int_{\Omega} \chi^{(\omega)} P(d\omega), \qquad \sigma \in S(\infty).$$

3.4 Spectral decomposition for integral values of z

Consider the generalized regular representations T_z of the group G introduced in §1.7.

Theorem. *Assume z is an integer, $z = k \in \mathbb{Z}$.*

(i) *The representation T_k possesses K–invariant cyclic vectors, i.e., it can be made a spherical representation.*

(ii) *Let ξ be any such vector, ψ be the corresponding character, and P be its spectral measure on Ω. Then P is supported by the subset*

$$\bigcup_{p,q\ge 0,\,(p,q)\ne(0,0),\,p-q=k} \Omega_{pq}$$

and for any Ω_{pq} entering this subset, the restriction of P to Ω_{pq} is equivalent to Lebesgue measure on the simplex Ω_{pq}.

When $k \ne 0$, the restriction $(p,q) \ne (0,0)$ is redundant because it follows from the condition $p - q = k$.

The condition $p - q = k$ also implies that the spectral measures corresponding to different integral values of the parameter z are mutually singular. This, in turn, implies that the corresponding representations are *disjoint*, i.e., they do not have equivalent subrepresentations.

4 The characters χ_z

4.1 Definition of χ_z and its explicit expression

Let T_z be a generalized regular representation of G. Assume first $z \ne 0$. Recall that T_z is realized in the Hilbert space $L^2(\mathfrak{S}, \mu_t)$, where $t = |z|^2$. Let $\mathbf{1}$ denote the function on $\mathfrak{S}$ identically equal to 1. It can be viewed as a vector of $L^2(\mathfrak{S}, \mu_t)$. Since μ_t is K–invariant and the cocycle $z^{c(x,g)}$ entering the construction of T_z is trivial on K, the vector $\mathbf{1}$ is a K–invariant vector. Consider the corresponding matrix coefficient and pass to the corresponding character (see §2.4), which we denote by χ_z. Thus,

$$\chi_z(\sigma) = (T_z(\sigma, e)\mathbf{1}, \mathbf{1}), \qquad \sigma \in S(\infty).$$

We aim to give a formula for χ_z. To do this we will describe the expansion of $\chi_z\mid_{S(n)}$ in irreducible characters of $S(n)$ for any n. Recall that the irreducible representations of $S(n)$ are parametrized by Young diagrams with n boxes. Let $\mathbb{Y}_n$ be the set of these diagrams. For $\lambda \in \mathbb{Y}_n$ we denote by χ^λ the corresponding irreducible character (the trace of the irreducible representation of $S(n)$ indexed by λ). Let $\dim\lambda = \chi^\lambda(e)$ be the dimension of this representation. In combinatorial terms, $\dim\lambda$ is the number of standard Young tableaux of shape λ. Note that for this number there exist closed expressions. Below the notation $(i,j) \in \lambda$ means that the box on the intersection of the ith row and the jth column belongs to λ.

Theorem. *For any* $n = 1, 2, \dots,$

$$\chi_z \mid_{S(n)} = \sum_{\lambda \in \mathbb{Y}_n} \left(\frac{\prod\limits_{(i,j)\in\lambda} |z+j-i|^2}{|z|^2 \,(|z|^2+1)\dots(|z|^2+n-1)} \, \frac{\dim\lambda}{n!} \right) \chi^{\lambda}.$$

Note that this formula also makes sense for $z = 0$.

The next claim is a direct consequence of the formula.

Proposition. *The function* $\mathbf{1}$ *is a cyclic vector for* T_z *if and only if* $z \notin \mathbb{Z}$.

Thus, for nonintegral z, the couple $(T_z, \mathbf{1})$ is a spherical representation and the character χ_z entirely determines T_z.

Note that for $z = k \in \mathbb{Z}$, the cyclic span of $\mathbf{1}$ is a proper subrepresentation that "corresponds" to a particular simplex Ω_{pq} (see §3.4). Specifically,

$$(p, q) = \begin{cases} (k, 0), & \text{if } k > 0 \\ (0, |k|), & \text{if } k < 0 \\ (1, 1), & \text{if } k = 0. \end{cases}$$

4.2 The symmetry $z \leftrightarrow \bar{z}$

Proposition. *For any* z, *the representations* T_z *and* $T_{\bar{z}}$ *are equivalent.*

Proof. Indeed, if $z \in \mathbb{R}$ then there is nothing to prove. If $z \notin \mathbb{R}$ then $\mathbf{1}$ is cyclic, so that the claim follows from the fact that $\chi_z = \chi_{\bar{z}}$, which in turn is evident from Theorem of §4.1.

Note that this is by no means evident from the construction of the representations T_z.

4.3 Disjointness

Let P_z be the spectral measure of the character χ_z, see §3.3. When z is integral, the measure P_z lives on a simplex Ω_{pq}, see §4.1. Now we focus on the measures P_z with $z \notin \mathbb{Z}$.

Theorem. (i) *Let* $z \notin \mathbb{Z}$. *Then all simplices* Ω_{pq} *are null sets with respect to the measure* P_z.

(ii) *Let* z_1 *and* z_2 *be two complex number, both nonintegral,* $z_1 \neq z_2$, *and* $z_1 \neq \bar{z}_2$. *Then the measures* P_{z_1} *and* P_{z_2} *are mutually singular.*

It follows that the generalized regular representations T_z are mutually disjoint, with the exception of the equivalence $T_z \sim T_{\bar{z}}$.

4.4 A nondegeneracy property

Proposition. *All measures P_z, $z \in \mathbb{C}$, are supported by the subset*

$$\Omega_0 := \left\{(\alpha, \beta) \mid \sum \alpha_i + \sum \beta_j = 1\right\}.$$

On the contrary, the measure P_∞ that corresponds to the biregular representation T_∞ is the Dirac measure at the point $(\underline{0}, \underline{0})$, which is outside Ω_0. This does not contradicts the fact that the family $\{T_z\}$ is a deformation of T_∞, because Ω_0 is dense in Ω.

5 Determinantal point processes

5.1 Point configurations

Let $\mathfrak{X}$ be a locally compact separable topological space. By a *point configuration* in $\mathfrak{X}$ we mean a locally finite collection C of points of the space $\mathfrak{X}$. These points will also be called *particles.* Here "locally finite" means that the intersection of C with any relatively compact subset is finite. Thus, C is either finite or countably infinite. Multiple particles in C are, in principle, permitted but all multiplicities must of course be finite. However, we will not really deal with configurations containing multiple particles. Let us emphasize that the particles in C are unordered.

The set of all point configurations in $\mathfrak{X}$ will be denoted by $\mathrm{Conf}(\mathfrak{X})$.

5.2 Definition of a point process

A relatively compact Borel subset $A \subset \mathfrak{X}$ will be called a *window.* Given a window A and $C \in \mathrm{Conf}(\mathfrak{X})$, let $\mathcal{N}_A(C)$ be the cardinality of the intersection $A \cap C$ (with multiplicities counted). Thus, $\mathcal{N}_A$ is a function on $\mathrm{Conf}(\mathfrak{X})$ taking values in $\mathbb{Z}_+$. We equip $\mathrm{Conf}(\mathfrak{X})$ with the Borel structure generated by the functions of the form $\mathcal{N}_A$.

By a measure on $\mathrm{Conf}(\mathfrak{X})$ we will mean a Borel measure with respect to this Borel structure.

By definition, a *point process* on $\mathfrak{X}$ is a probability measure $\mathcal{P}$ on the space $\mathrm{Conf}(\mathfrak{X})$.

In practice, point processes often arise as follows. Assume we are given a Borel space Y and a map $\phi : Y \to \mathrm{Conf}(\mathfrak{X})$. The map ϕ must be a Borel map. i.e., for any window A, the superposition $\mathcal{N}_A \circ \phi$ must be a Borel function on Y. Further, assume we are given a probability Borel measure P on Y. Then its pushforward $\mathcal{P}$ under ϕ is well defined and it is a point process.

Given a point process, we can speak about *random* point configurations C. Any reasonable (that is, Borel) function of C becomes a random variable. For instance, $\mathcal{N}_A$ is a random variable for any window A, and we may consider the probability that $\mathcal{N}_A$ takes any prescribed value.

5.3 Example: Poisson process

Let ρ be a measure on $\mathfrak{X}$. It may be infinite but must take finite values on any window. The *Poisson process with density* ρ is characterized by the following properties:

(i) For any window A, the random variable $\mathcal{N}_A$ has the Poisson distribution with parameter $\rho(A)$, i.e.,

$$\operatorname{Prob}\{\mathcal{N}_A = n\} = \frac{\rho(A)^n}{n!} e^{-\rho(A)}, \quad n \in \mathbb{Z}_+ .$$

(ii) For any pairwise disjoint windows $A_1, \dots, A_k$, the corresponding random variables are independent.

In particular, if $\mathfrak{X} = \mathbb{R}$ and ρ is the Lebesgue measure then this is the classical Poisson process.

5.4 Correlation measures and correlation functions

Let $\mathcal{P}$ be a point process on $\mathfrak{X}$. One can assign to $\mathcal{P}$ a sequence $\rho_1, \rho_2, \dots$ of measures, where, for any n, ρ_n is a symmetric measure on the n–fold product $\mathfrak{X}^n = \mathfrak{X} \times \dots \times \mathfrak{X}$, called the *$n$–particle correlation measure*. Under mild assumptions on $\mathcal{P}$ the correlation measures exist and determine $\mathcal{P}$ uniquely. They are defined as follows.

Given n and a compactly supported bounded Borel function f on $\mathfrak{X}^n$, let $\widetilde{f}$ be the function on $\operatorname{Conf}(\mathfrak{X})$ defined by

$$\widetilde{f}(C) = \sum_{i_1, \dots, i_n} f(x_{i_1}, \dots, x_{i_n}), \qquad C = \{x_1, x_2, \dots\} \in \operatorname{Conf}(\mathfrak{X}),$$

summed over all n–tuples of pairwise distinct indices. Here we have used an enumeration of the particles in C but the result does not depend on it.

Then the measure ρ_n is characterized by the equality

$$\int_{\mathfrak{X}^n} f \rho_n = \int_{C \in \operatorname{Conf}(\mathfrak{X})} \widetilde{f}(C)\, \mathcal{P}(dC),$$

where f is an arbitrary compactly supported bounded Borel function on $\mathfrak{X}^n$.

Examples. (i) If $\mathcal{P}$ is a Poisson process then $\rho_n = \rho^{\otimes n}$, where ρ is the density of $\mathcal{P}$.

(ii) Assume that $\mathfrak{X}$ is discrete and $\mathcal{P}$ lives on multiplicity free configurations. Then the correlation measures say what is the probability that the random configuration contains an arbitrary given finite set of points.

Often there is a natural measure ν on $\mathfrak{X}$ (a *reference measure*) such that each ρ_n has a density with respect to $\nu^{\otimes n}$. This density is called the nth *correlation function*. For instance, if $\mathfrak{X}$ is a domain of an Euclidean space and ν is the Lebesgue measure then, informally, the nth correlation function equals the density of the probability that the random configuration has particles in given n infinitisemal regions $dx_1, \dots, dx_n$.

5.5 Determinantal point processes

Let $\mathcal{P}$ be a point process on $\mathfrak{X}$. Assume that $\mathfrak{X}$ is equipped with a reference measure ν such that the correlation functions (taken with respect to ν) exist. Let us denote these functions by $\rho_n(x_1,\dots,x_n)$. The process $\mathcal{P}$ is said to be *determinantal* if there exists a function $K(x,y)$ on $\mathfrak{X}\times\mathfrak{X}$ such that

$$\rho_n(x_1,\dots,x_n)=\det[K(x_i,x_j)]_{i,j=1}^n, \qquad n=1,2,\dots.$$

Then $K(x,y)$ is called a *correlation kernel* of $\mathcal{P}$.

If $K(x,y)$ exists it is not unique since for any nonvanishing function $\phi(x)$ on $\mathfrak{X}$, the kernel $\phi(x)K(x,y)(\phi(y))^{-1}$ leads to the same result.

If we replace the reference measure by an equivalent one then we always can appropriately change the kernel. Specifically, if ν is multiplied by a positive function $f(x)$ then $K(x,y)$ can be replaced, say, by $K(x,y)(f(x)f(y))^{-1/2}$.

Examples. (i) Let $\mathfrak{X}=\mathbb{R}$, ν be the Lebesgue measure, and $K(x,y)=\overline{K(y,x)}$ be the kernel of an Hermitian integral operator K in $L^2(\mathbb{R})$. Then $K(x,y)$ is a correlation kernel of a determinantal point process if and only if $0\le K\le 1$ and the restriction of the kernel to any bounded interval determines a trace class operator.

(ii) The above conditions are satisfied by the *sine kernel*

$$K(x,y)=\frac{\sin(\pi(x-y))}{\pi(x-y)}, \qquad x,y\in\mathbb{R}.$$

The sine kernel arises in random matrix theory. It determines a translation invariant point process on $\mathbb{R}$, which is a fundamental and probably the best known example of a determinantal point process.

6 The point processes $\mathcal{P}_z$ and $\widetilde{\mathcal{P}}_z$. The main result

6.1 From spectral measures to point processes

Let $I=[-1,1]\subset\mathbb{R}$ and $I^*=[-1,1]\setminus\{0\}$. Let us take I^* as the space $\mathfrak{X}$. We define an embedding $\Omega\to\operatorname{Conf}(I^*)$ as follows

$$\omega=(\alpha,\beta)\mapsto C=\{\alpha_i\ne0\}\cup\{-\beta_j\ne0\}.$$

That is, we remove the possible zero coordinates, change the sign of the β–coordinates, and forget the ordering. In this way we convert ω to a point configuration C in the punctured segment I^*. In particular, the empty configuration $C=\varnothing$ corresponds to $\omega=(\underline{0},\underline{0})$.

Given a probability measure P on Ω, its pushforward under this embedding is a probability measure $\mathcal{P}$ on $\operatorname{Conf}(I^*)$, i.e., a point process on the space I^*, see §5.2. Applying this procedure to the spectral measures P_z (§4.3) we get point processes $\mathcal{P}_z$ on I^*.

6.2 Lifting

We aim to define a modification of the point processes $\mathcal{P}_z$. Fix $z \in \mathbb{C} \setminus \{0\}$ and set as usual $t = |z|^2$. Let $s > 0$ be a random variable whose distribution has the form

$$\frac{1}{\Gamma(t)} s^{t-1} e^{-s} ds$$

(the gamma distribution on $\mathbb{R}_+$ with parameter t.) We assume that s is independent of $\mathcal{P}_z$. Given the random configuration C of the process $\mathcal{P}_z$, we multiply the coordinates of all particles of C by the random factor s. The result is a random point configuration $\widetilde{C}$ on $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$.

We call this procedure the *lifting*. Under the lifting the point process $\mathcal{P}_z$ is transformed to a point process on $\mathbb{R}^*$ which we denote by $\widetilde{\mathcal{P}}_z$.

The lifting is in principle reversible. Indeed, due to Proposition of §4.4, we can recover C from $\widetilde{C}$ by dividing all the coordinates in $\widetilde{C}$ by the sum of their absolute values.

It turns out that the lifting leads to a simplification of the initial point process.

6.3 Transformation of the correlation functions under the lifting

Fix the parameter z. Let $\rho_n(x_1, \dots, x_n)$ and $\widetilde{\rho}_n(x_1, \dots, x_n)$ be the correlation functions of the processes $\mathcal{P}_z$ and $\widetilde{\mathcal{P}}_z$, respectively (see §5.4). Here we take the Lebesgue measure as the reference measure.

The definition of the lifting implies that

$$\widetilde{\rho}_n(x_1, \dots, x_n) = \int_0^\infty \frac{s^{t-1}e^{-s}}{\Gamma(t)} \rho_n\left(\frac{x_1}{s}, \dots, \frac{x_n}{s}\right) \frac{ds}{s^n},$$

where we agree that the function ρ_n vanishes on $(\mathbb{R}^*)^n \setminus (I^*)^n$. Thus, the action of the lifting on the correlation functions is expressed by a ray integral transform.

This ray transform can be readily reduced to the Laplace transform. It follows that it is injective, which agrees with the fact that lifting is reversible.

6.4 The main result

To state the result we need some notation.

Let $W_{\kappa,\mu}(x)$ denote the *Whittaker function* with parameters $\kappa, \mu \in \mathbb{C}$. It is a unique solution of the differential equation

$$W'' - \left(\frac{1}{4} - \frac{\kappa}{x} + \frac{\mu^2 - \frac{1}{4}}{x^2}\right) W = 0$$

with the condition $W(x) \sim x^\kappa e^{-\frac{x}{2}}$ as $x \to +\infty$.

This function is initially defined for real positive x and then can be extended to a holomorphic function on $\mathbb{C} \setminus (-\infty, 0]$.

Next, we write $z = a + ib$ with real a, b and set

$$P_{\pm}(x) = \frac{t^{\frac{1}{2}}}{|\Gamma(1 \pm z)|} W_{\pm a + \frac{1}{2}, ib}(x), \qquad Q_{\pm}(x) = \frac{t^{\frac{3}{2}} x^{-\frac{1}{2}}}{|\Gamma(1 \pm z)|} W_{\pm a - \frac{1}{2}, ib}(x) .$$

Main Theorem. *For any $z \in \mathbb{C} \backslash \{0\}$, the point process $\widetilde{\mathcal{P}}_z$ is a determinantal process whose correlation kernel can be written as*

$$K(x,y) = \begin{cases} \dfrac{P_+(x)Q_+(y) - Q_+(x)P_+(y)}{x-y}, & x > 0\,, y > 0 \\ \dfrac{P_+(x)P_-(-y) + Q_+(x)Q_-(-y)}{x-y}, & x > 0\,, y < 0 \\ \dfrac{P_+(x)P_+(y) + Q_-(-x)Q_+(y)}{x-y}, & x < 0\,, y > 0 \\ -\dfrac{P_-(-x)Q_-(-y) - Q_-(-x)P_-(-y)}{x-y}, & x < 0\,, y < 0 \end{cases}$$

where $x, y \in \mathbb{R}^*$ and the indeterminacy arising for $x = y$ is resolved via the L'Hospital rule.

We call the kernel $K(x, y)$ the *Whittaker kernel.*

Note that $K(x, y)$ is real valued. It is not symmetric but satisfies the symmetry property

$$K(x, y) = \operatorname{sgn}(x) \operatorname{sgn}(y) K(y, x),$$

where $\operatorname{sgn}(x)$ equals ± 1 according to the sign of x. This property can be called *J–symmetry*, it means that the kernel is symmetric with respect to an indefinite inner product.

6.5 The L–operator

Split the Hilbert space $L^2(\mathbb{R}^*)$ into the direct sum $L^2(\mathbb{R}_+) \oplus L^2(\mathbb{R}_-)$, where all L^2 spaces are taken with respect to the Lebesgue measure. According to this splitting we will write operators in $L^2(\mathbb{R}^*)$ in block form, as 2×2 operator matrices.

Let

$$L = \begin{bmatrix} 0 & A \\ -A^t & 0 \end{bmatrix},$$

where A is the integral operator with the kernel

$$A(x, y) = \left| \frac{\sin(\pi z)}{\pi} \right|^2 \cdot \frac{\left(\frac{x}{|y|}\right)^{\operatorname{Re} z} e^{-\frac{x-y}{2}}}{x - y}, \quad x > 0, \quad y < 0.$$

By A^t we denote the conjugate operator $L^2(\mathbb{R}_+) \to L^2(\mathbb{R}_-)$.

Theorem. *Assume that* $-\frac{1}{2} < \Re z < \frac{1}{2}$, $z \neq 0$. *Then* A *is a bounded operator* $L^2(\mathbb{R}_-) \to L^2(\mathbb{R}_+)$ *and the correlation kernel* $K(x,y)$ *is the kernel of the operator* $L(1+L)^{-1}$.

Note that, in contrast to K, the kernel of L does not involve special functions.

6.6 An application

Fix $z \in \mathbb{C} \setminus \mathbb{Z}$ and consider the probability space (Ω, P_z). For any $k = 1, 2, \dots$ the coordinates α_k and β_k are functions in $\omega \in \Omega$, hence we may view them as random variables. The next result provides an information about the rate of their decay as $i, j \to \infty$.

Theorem. *With probability 1, there exist limits*

$$\lim_{k\to\infty} (\alpha_k)^{\frac{1}{k}} = \lim_{k\to\infty} (\beta_k)^{\frac{1}{k}} = q(z) \in (0,1),$$

where

$$q(z) = \exp\left(\pi\,\frac{\operatorname{ctg}\pi z - \operatorname{ctg}\pi\bar{z}}{z-\bar{z}}\right) = \exp\left(-\sum_{n\in\mathbb{Z}} \frac{1}{|z-n|^2}\right)$$

7 Scheme of the proof of the Main Theorem

7.1 The z–measures

Recall that by $\mathbb{Y}_n$ we denote the finite set of Young diagrams with n boxes. Set

$$P_z^{(n)}(\lambda) = \frac{\prod\limits_{(i,j)\in\lambda} |z+j-i|^2}{|z|^2\,(|z|^2+1)\dots(|z|^2+n-1)}\,\frac{(\dim\lambda)^2}{n!}\,, \qquad \lambda \in \mathbb{Y}_n\,.$$

Comparing this with the expression of $\chi_z\mid_{S(n)}$ (§4.1) we see that the quantities $P_z^{(n)}(\lambda)$ are the coefficients in the expansion of χ_z in the *normalized* irreducible characters $\chi^\lambda/\dim\lambda$. It follows that

$$\sum_{\lambda\in\mathbb{Y}_n} P_z^{(n)}(\lambda) = 1.$$

Thus, for any fixed $n = 1, 2, \dots$, the quantities $P_z^{(n)}(\lambda)$ determine a probability measure on $\mathbb{Y}_n$. We will denote it by $P_z^{(n)}$ and call it the *z–measure* on Y_n.

7.2 Frobenius coordinates and the embedding $\mathbb{Y}_n \hookrightarrow \Omega$

Given $\lambda \in \mathbb{Y}_n$, let λ' be the transposed diagram and d be the number of diagonal boxes in λ. We define the *modified Frobenius coordinates* of λ as

$$a_i = \lambda_i - i + \frac{1}{2}, \qquad b_i = \lambda'_i - i + \frac{1}{2}, \qquad i = 1, \dots, d.$$

Note that

$$a_1 > \dots > a_d > 0, \qquad b_1 > \dots > b_d > 0, \qquad \sum_{i=1}^{d}(a_i + b_i) = n.$$

For any $n = 1, 2, \dots$ we embed $\mathbb{Y}_n$ into Ω by making use of the map

$$\lambda \mapsto \omega_\lambda = (\alpha, \beta),$$
$$\alpha = \left(\frac{a_1}{n}, \dots, \frac{a_d}{n}, 0, 0, \dots\right), \qquad \beta = \left(\frac{b_1}{n}, \dots, \frac{b_d}{n}, 0, 0, \dots\right).$$

As $n \to \infty$, the points ω_λ coming from the diagrams $\lambda \in \mathbb{Y}_n$ fill out the space Ω more and more densely. Thus, for large n, the image of $\mathbb{Y}_n$ in Ω can be viewed as a discrete approximation of Ω.

7.3 Approximation of P_z by z–measures

Let $\underline{P}_z^{(n)}$ be the pushforward of the measure $P_z^{(n)}$ under the embedding $\mathbb{Y}_n \hookrightarrow \Omega$. This is a probability measure on Ω.

Approximation Theorem. *As $n \to \infty$, the measures $\underline{P}_z^{(n)}$ weakly converge to the measure P_z.*

This fact is the starting point for explicit computations related to the measures P_z.

7.4 The mixed z–measures

Let $\mathbb{Y} = \mathbb{Y}_0 \cup \mathbb{Y}_1 \cup \mathbb{Y}_2 \cup \dots$ be the set of all Young diagrams. We agree that $\mathbb{Y}_0$ consists of a single element – the empty diagram $\varnothing$. Fix $z \in \mathbb{C} \setminus \{0\}$ and $\xi \in (0, 1)$. We define a measure $\widetilde{P}_{z,\xi}$ on $\mathbb{Y}$ as follows:

$$\widetilde{P}_{z,\xi}(\lambda) = P_z^{(n)}(\lambda) \cdot (1 - \xi)^{|z|^2} \, \frac{|z|^2\,(|z|^2 + 1) \dots (|z|^2 + n - 1)}{n!} \, \xi^n, \qquad \lambda \in \mathbb{Y},$$

where n is the number of boxes in λ and $P_z^{(0)}(\varnothing) := 1$.

In other words, $\widetilde{P}_{z,\xi}$ is obtained by mixing together all the z–measures $P_z^{(0)}, P_z^{(1)}, \dots$, where the weight of the nth component is equal to

$$\pi_{t,\xi}(n) = (1-\xi)^t \, \frac{t(t+1)\dots(t+n-1)}{n!} \, \xi^n, \qquad t = |z|^2.$$

Note that

$$\sum_{n=0}^{\infty} \pi_{t,\xi}(n) = 1.$$

It follows that $\widetilde{P}_{z,\xi}$ is a probability measure. Let us call it the *mixed z–measure.*

Note that, as $z \to 0$, the measure $\widetilde{P}_{z,\xi}$ tends to the Dirac mass at $\{\varnothing\}$ for any fixed ξ.

7.5 The lattice process $\widetilde{\mathcal{P}}_{z,\xi}$

Set

$$\mathbb{Z}' = \mathbb{Z} + \frac{1}{2} = \left\{ \dots, -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, \dots \right\}.$$

Using the notation of §7.2 we assign to an arbitrary Young diagram a point configuration $C \in \operatorname{Conf}(\mathbb{Z}')$, as follows

$$\lambda \mapsto C = \{-b_1, \dots, -b_d, a_d, \dots, a_1\}.$$

The correspondence $\lambda \mapsto C$ defines an embedding $\mathbb{Y} \hookrightarrow \operatorname{Conf}(\mathbb{Z}')$. Take the pushforward of the measure $P_{z,\xi}$ under this embedding. It is a probability measure on $\operatorname{Conf}(\mathbb{Z}')$, hence a point process on $\mathbb{Z}'$. Let us denote it by $\widetilde{\mathcal{P}}_{z,\xi}$.

Theorem. *The process $\widetilde{\mathcal{P}}_{z,\xi}$ on the lattice $\mathbb{Z}'$ is determinantal. Its correlation kernel can be explicitly computed: it has the form quite similar to that of the kernel $K(x,y)$ from §6.4, where the corresponding functions $P_\pm$ and $Q_\pm$ are now expressed through the Gauss hypergeometric function.*

7.6 Idea of proof of the Main Theorem

Given $\xi \in (0,1)$, we embed the lattice $\mathbb{Z}'$ into $\mathbb{R}^*$ as follows

$$\mathbb{Z}' \ni x \mapsto (1-\xi)x \in \mathbb{R}^*.$$

Let $\underline{\widetilde{\mathcal{P}}}_{z,\xi}$ be the pushforward of $\widetilde{\mathcal{P}}_{z,\xi}$ under this embedding. We can view $\underline{\widetilde{\mathcal{P}}}_{z,\xi}$ as a point process on $\mathbb{R}^*$.

Remark that the probability distribution $\pi_{t,\xi}$ on $\mathbb{Z}_+$ introduced in §7.4 approximates in an appropriate scaling limit as $\xi \nearrow 1$ the gamma distribution on $\mathbb{R}^*$ with parameter t. Specifically, the scaling has the form $n \mapsto (1-\xi)n$. Recall that we have used the gamma distribution in the definition of the lifting, see §6.2.

Combining this fact with the Approximation Theorem of §7.3 we conclude that the process $\underline{\widetilde{\mathcal{P}}}_{z,\xi}$ must converge to the process $\widetilde{\mathcal{P}}_z$ as $\xi \nearrow 1$ in a certain sense. More precisely, we prove that the correlation measures of the former process converge to the respective correlation measures of the latter process.

On the other hand, we can explicitly compute the scaled limit of the lattice correlation measures using the explicit expression of the lattice kernel from §7.5. It turns out that then the Gauss hypergeometric function degenerates to the Whittaker function and we get the formulas of §6.4.

8 Notes and references

8.1 Section 1

The main reference to this section is the paper Kerov–Olshanski–Vershik [41].

§1.1. Peter–Weyl's theorem is included in many textbooks on representation theory. See, e.g., Naimark [55], §32.

§1.2. From the purely algebraic point of view, there is no single infinite analog of the permutation groups $S(n)$ but a number of different versions. The group $S(\infty) = \varinjlim S(n)$ formed by *finite* permutations of the set $\{1, 2, \dots\}$ and the group of *all* permutations of this set may be viewed as the minimal and the maximal versions. There is also a huge family of intermediate groups. The choice of an appropriate version may vary depending on the applications we have in mind. Certain *topological* groups connected with $S(\infty)$ are discussed in Olshanski [61], Okounkov [56], [57].

§1.3. The result of the Proposition is closely related to von Neumann's classical construction of II_1 factors. See Murray–von Neumann [51], ch. 5, and Naimark [55], ch. VII, §38.5.

§1.4. The G–space $\mathfrak{S}$ of virtual permutations was introduced in Kerov–Olshanski–Vershik [41]. Notice that the canonical projection p_n emerged earlier, see Aldous [1], p. 92. A closely related construction, which also appeared earlier, is the so–called *Chinese restaurant process,* see, e.g., Arratia–Barbour–Tavaré [2], §2 and references therein. Projective limit constructions for classical groups and symmetric spaces are considered in Pickrell [71], Neretin [53], Olshanski [64]. Earlier papers: Hida–Nomoto [29], Yamasaki [90], [91], Shimomura [73].

§1.5. The definition of the Ewens measures μ_t on the space $\mathfrak{S}$ was proposed in [41], see also Kerov–Tsilevich [42], Kerov [37] (the latter paper deals with a generalization of these measures). The definition of [41] was inspired by the fundamental concept of the *Ewens sampling formula*, which was derived in 1972 by Ewens [26] in the context of population genetics. There is a large literature concerning Ewens' sampling formula (or Ewens' partition structure). See, e.g., the papers Watterson [89], Kingman [44], [45], [47], Arratia–Barbour–Tavaré [2],[3], Ewens [27], which contain many other references.

§1.6. The results were established in [41]. For projective limits of classical groups and symmetric spaces, there also exist distinguished families of measures with good transformation properties, see Pickrell [71], Neretin [53], Olshanski [64].

§1.7. The representations T_z were introduced in [41]. A parallel construction exists for infinite–dimensional classical groups and symmetric spaces, see the pioneer paper Pickrell [71] and also Neretin [53], Olshanski [64].

8.2 Section 2

§2.1. The concept of spherical representations is usually employed for *Gelfand pairs* $(\mathcal{G},\mathcal{K})$. According to the conventional definition, $(\mathcal{G},\mathcal{K})$ is said to be a Gelfand pair if the subalgebra of $\mathcal{K}$–biinvariant functions in the group algebra $L^1(\mathcal{G})$ is commutative. This works for locally compact $\mathcal{G}$ and compact $\mathcal{K}$. There exists, however, a reformulation which makes sense for arbitrary groups, see Olshanski [62]. Our pair (G,K) is a Gelfand pair, see Olshanski [63].

§2.2. For general facts concerning positive definite functions on groups, see, e.g., Naimark [55].

§2.3. There exist at least two different ways to define characters for infinite–dimensional representations. The most known recipe (Gelfand, Harish–Chandra) is to view characters not as ordinary functions but as distributions on the group. This idea works perfectly for a large class of Lie groups and p–adic groups but not for groups like $S(\infty)$. The definition employed here follows another approach, which goes back to von Neumann. Extreme characters of a group $\mathcal{K}$ are related to finite factor representations of $\mathcal{K}$ in the sense of von Neumann. See Thoma [76], [77], Stratila–Voiculescu [75], Voiculescu [87].

§2.4. The correspondence between extreme characters and irreducible spherical representations was pointed out in Olshanski [61], [62]. The Proposition follows from the fact that our pair (G,K) is a Gelfand pair, see Olshanski [63].

The irreducible spherical representations of (G,K) form a subfamily of a larger family of representations called *admissible representations,* see Olshanski [61], [62], [63]. On the other hand, aside from *finite* factor representations of $S(\infty)$ that correspond to extreme characters, there exist interesting examples of factor representations of quite different nature, see Stratila–Voiculescu [75].

Explicit realizations of finite factor representations of $S(\infty)$ and irreducible spherical representations of (G,K) are given in Vershik–Kerov [82], Wassermann [88], Olshanski [63].

§2.5. There are various methods to establish the existence and uniqueness of the spectral decomposition. See, e.g., Diaconis–Freedman [23], Voiculescu [87], Olshanski [64]. One more approach, which is specially adapted to the group $S(\infty)$ and provides an explicit description of $\mathrm{Ex}(\mathcal{X}(S(\infty)))$, is proposed in Kerov–Okounkov–Olshanski [39].

8.3 Section 3

§3.1. The expressions $p_k(\alpha,\beta)$ are *supersymmetric* analogs of power sums. About the role of supersymmetric functions in the theory of characters of $S(\infty)$ see Vershik–Kerov [83], Olshanski–Regev–Vershik [65].

§3.2. The Thoma set Ω can be viewed as an infinite–dimensional simplex. The subsets Ω_{pq} are exactly its finite–dimensional faces.

§3.3. Thoma's paper [76] was the first work about characters of $S(\infty)$. It contains the classification of extreme characters (Thoma's theorem), which was obtained using complex–analytic tools. Thoma's theorem is equivalent to another classification problem — that of *one–sided totally positive sequences.* Much earlier, that problem was raised by Schoenberg and solved by Edrei [25]. The equivalence of both problems was implicit in Thoma's paper [76] but Thoma apparently was not aware of the works on total positivity.

The next step was made by Vershik and Kerov [83]. Following a general principle earlier suggested in Vershik [79], Vershik and Kerov found a new proof of Thoma's theorem. Their approach is based on studying the limit transition from characters of $S(n)$ to characters of $S(\infty)$. This provides a very natural interpretation of Thoma's parameters α_i, β_j.

Developing further the asymptotic approach of [83], Kerov–Okounkov–Olshanski [39] obtained a generalization of Thoma's theorem. An even more general claim was conjectured by Kerov in [35].

One of the fruitful ideas contained in Vershik–Kerov's paper [83] concerns the combinatorics of irreducible characters χ^λ of the finite symmetric groups. Assume that $\lambda\in\mathbb{Y}_n$ and ρ is a partition of m, where $m\le n$. Let χ^λ_ρ denote the value of χ^λ at the conjugacy class in $S(n)$ indexed by the partition $\rho\cup 1^{n-m}$ of n. The idea was to consider χ^λ_ρ as a function in λ with ρ viewed as a parameter. Vershik and Kerov discovered that the function $\lambda\mapsto\chi^\lambda_\rho$, after a simple normalization, becomes a supersymmetric function in the modified Frobenius coordinates of λ. This function is inhomogeneous and its top degree homogeneous term is the supersymmetric (product) power sum indexed by ρ. Further results in this directions: Kerov–Olshanski [40], Okounkov–Olshanski [60], Olshanski–Regev–Vershik [65]. Even in the simplest case when ρ consists of a single part ($\rho=(m)$) the function $\lambda\mapsto\chi^\lambda_\rho=\chi^\lambda_{(m)}$ is rather nontrivial. See Wassermann [88], Kerov [36], Biane [5], Ivanov–Olshanski [32].

§3.4. The spectral decomposition of T_z's for integral values of z was obtained in Kerov–Olshanski–Vershik [41].

8.4 Section 4

§4.1. The results were obtained in Kerov–Olshanski–Vershik [41]. Similar results for other groups: Pickrell [71], Olshanski [64].

§4.2. One can define intertwining operators for the representations T_z and $T_{\bar z}$. These operators have interesting properties. See Kerov–Olshanski–Vershik [41].

§4.3. The result was obtained in Kerov–Olshanski–Vershik [41]. Note that the Theorem of §6.6 implies a weaker result: the spectral measures P_{z_1} and P_{z_2} are mutually singular for any $z_1, z_2 \in \mathbb{C} \setminus \mathbb{Z}$ such that $q(z_1) \neq q(z_2)$.

§4.4. The result was announced in Kerov–Olshanski–Vershik [41]. It can be proved in different ways, see Olshanski [66], Borodin [67].

8.5 Section 5

§§5.1 – 5.4. The material is standard. See Daley and Vere-Jones [21], Lenard [48], Kingman [47]. Point processes are also called *random point fields.*

§5.5. The class of determinantal point process was first singled out by Macchi [49], [50] under the name of *fermion processes.* The motivation comes from a connection with the fermionic Fock space. The term "determinantal" was suggested in Borodin–Olshanski [15]. We found it more appropriate, because in our concrete situation, point configurations may be viewed as consisting of particles of *two opposite charges.* A number of important examples of determinantal point processes emerged in random matrix theory, see, e.g., Dyson [24], Mehta [52], Nagao–Wadati [54], Tracy–Widom [78], and the references therein. However, to our knowledge, up to the recent survey paper by Soshnikov [74], the experts in this field did not pay attention to general properties of determinantal processes and did not introduce any general name for them.

The result stated in Example (i) is due to Soshnikov [74].

8.6 Section 6

§6.1. The spectral measures P_z with nonintegral parameter z originally looked mysterious: it was unclear how to handle them.

The idea of converting the measures P_z into point processes $\mathcal{P}_z$ and computing the correlation functions was motivated by the following observation. It turns out that the coefficients of the expansion of §4.1 can be interpreted as moments of certain auxiliary measures (we called them the *controlling measures*). The controlling measures are determined by these moments uniquely. On the other hand, the correlation functions can be expressed through the controlling measures. It follows that evaluating the correlation functions can be reduced to solving certain (rather complicated) multidimensional moment problems.

We followed first this way (see the preprints [66]–[70]; part of results was published in Borodin [7], [8]; a summary is given in Borodin–Olshanski [14]). A general description of the method and the evaluation of the first correlation function are given in Olshanski [66]. In Borodin [7] the moment problem in question is studied in detail. This leads (Borodin [67]) to some formulas for the higher correlation functions: a multidimensional integral representation and an explicit expression through a multivariate Laurichella hypergeometric series of type B. Both are rather involved.

§§6.2–6.3. The idea of lifting (Borodin [69]) turned out to be extremely successful, because it leads to a drastic simplification of the correlation functions. What is even more important is that due to this procedure we finally hit a nice class of point processes, the determinantal ones.

§6.4. The derivation of the Whittaker kernel by the first method is given in Borodin [69], [8]. It should be noted that the Whittaker kernel belongs to the class of *integrable kernels.* This class was singled out by Its–Izergin–Korepin–Slavnov [31], see also Deift [22], Borodin [11].

§6.5. The claim concerning the L–operator and some related facts are contained in Olshanski [70]. A conclusion is that (at least when $|\Re z| < 1/2$) the whole information about the spectral measure P_z is encoded in a very simple kernel $L(x, y)$.

§6.6. The result is obtained in Borodin–Olshanski [68]. It can be viewed as a strong law of large numbers. Roughly speaking, the coordinates α_k, β_k decay like the terms of the geometric progression $\{q(z)^k\}$. A similar result holds for point processes of quite different type (Poisson–Dirichlet distributions), see Vershik–Shmidt [86].

Notice that the preprints [66]–[70] contain a number of other results, some of them remain still unpublished.

8.7 Section 7

The main reference for this section is the paper Borodin–Olshanski [15], which gives an alternate way of proving the Main Theorem. The method of [15] is simpler than the previous approach based on a moment problem. Furthermore, our second approach explains the origin of the lifting. However, the correlation functions for the initial process $\mathcal{P}_z$ are not directly obtained in this way.

§7.1. The z–measures $P_z^{(n)}$ with fixed parameter z and varying index n satisfy the *coherency relation*

$$P_z^{(n)}(\mu) = \sum_{\lambda \in \mathbb{Y}_{n+1}:\, \lambda \supset \mu} \frac{\dim \mu}{\dim \lambda} P_z^{(n+1)}(\lambda), \qquad n = 1, 2, \dots, \quad \mu \in \mathbb{Y}_n\,.$$

It expresses the fact that the function $\chi_z \mid_{S(n+1)}$ is an extension of the function $\chi_z \mid_{S(n)}$. The coherency relation is not evident from the explicit expression for the z–measures.

As $|z| \to \infty$, the measures $P_z^{(n)}$ converge to the *Plancherel measure* on $\mathbb{Y}_n$,

$$P_\infty^{(n)}(\lambda) = \frac{(\dim \lambda)^2}{n!}\,.$$

Note that the expression for $P_z^{(n)}(\lambda)$ looks as a product over the boxes of λ times $P_\infty^{(n)}(\lambda)$. This property together with the coherency relation can be used for a combinatorial characterization of the z–measures, see Rozhkovskaya [72].

Actually, the term "z–measures" has a somewhat wider meaning: the family $\{P_z^{(n)}\}$ forms the "principal series" while the whole family of the z–measures also includes a "complementary series" and a "degenerate series" of measures which are given by similar expressions.

A much larger family of *Schur measures* was introduced by Okounkov [59]. In general, the Schur measures do not obey the coherency relation and hence do not correspond to characters of $S(\infty)$. However, they also give rise to determinantal point processes. It would be interesting to know whether the z–measures exhaust all Schur measures satisfying the coherency relation.

Kerov [38] introduced analogs of z–measures satisfying a certain one–parameter *deformation* of the coherency relation (the coherency relation written above is closely related to the Schur functions, while Kerov's more general form of the coherency relation is related to the Jack symmetric functions, see also Kerov–Okounkov–Olshanski [39]). For another approach, see Borodin–Olshanski [16]. Study of the point processes corresponding to these more general z–measures was started in Borodin–Olshanski [20].

An analog of z–measures corresponding to *projective* characters of $S(\infty)$ was found in Borodin [6]. See also Borodin–Olshanski [16].

The paper Borodin–Olshanski [17] presents a survey of connections between z–measures and a number of models arising in combinatorics, tiling, directed percolation and random matrix theory.

§7.2. The idea of embedding $\mathbb{Y}$ into Ω is due to Vershik and Kerov [83]. In a more general context it is used in Kerov–Okounkov–Olshanski [39].

§7.3. The Approximation Theorem actually holds for spectral measures corresponding to arbitrary characters of $S(\infty)$. See Kerov–Okounkov–Olshanski [39].

§7.4. What we called "mixing" is a well–known trick. Under different names it is used in various asymptotic problems of combinatorics and statistical physics. See, e.g., Vershik [81]. The general idea is to replace a large n limit, where the index n enumerates different probabilistic ensembles, by a limit transition of another kind (we are dealing with a unifying ensemble depending on a parameter and let the parameter tend to a limit). In many situations the two limit transitions lead to the same result. For instance, this usually happens for the *poissonization procedure,* when the mixing distribution on $\mathbb{Z}_+$ is a Poisson distribution. (About the poissonized Plancherel measure, see Baik–Deift–Johansson [4], Borodin–Okounkov–Olshanski [13], Johansson [33].) A key property of the Poisson distribution is that as its parameter goes to infinity, the standard deviation grows more slowly than the mean. In our situation, instead of Poisson we have to deal with the distribution $\pi_{t,\xi}$, a particular case of the *negative binomial distribution.* As $\xi \nearrow 1$, the standard deviation and the mean of $\pi_{t,\xi}$ have the same order of growth, which results in a nontrivial transformation of the large n limit (the lifting).

§7.5. The fact that the lattice process $\widetilde{\mathcal{P}}_z$ is determinantal is checked rather easily. The difficult part of the Theorem is the calculation of the correlation kernel. This can be done in different ways, see Borodin–Olshanski [15], Okounkov [58], [59]. Borodin [9], [11] describes a rather general procedure of computing correlation kernels via a Riemann–Hilbert problem.

§7.6. For more details see Borodin–Olshanski [15].

8.8 Other problems of harmonic analysis leading to point processes

A parallel but more complicated theory holds for the infinite–dimensional unitary group $U(\infty)=\varinjlim U(N)$. For this group, there exists a completion $\mathfrak{U}$ of the group space $U(\infty)$, which plays the role of the space $\mathfrak{S}$ of virtual permutations. On $\mathfrak{U}$, there exists a family of measures with good transformation properties which give rise to certain unitary representations of $U(\infty)\times U(\infty)$ — analogs of the representations T_z. See Neretin [53], Olshanski [64]. The problem of harmonic analysis for these representations is studied in Borodin–Olshanski [19]. It leads to determinantal point processes on the space $\mathbb{R}\setminus\{\pm\frac{1}{2}\}$. Their correlation kernels were found in [19]: these are integrable kernels expressed through the Gauss hypergeometric function.

There exists a similarity between decomposition of unitary representations into irreducible ones and decomposition of invariant measures on ergodic components. Both problems often can be interpreted in terms of barycentric decomposition on extreme points in a convex set. Below we briefly discuss two problems of "harmonic analysis for invariant measures" that lead to point processes.

The first problem concerns invariant probability measures for the action of the diagonal group $K\subset G$ on the space $\mathfrak{S}$. Recall that K is isomorphic to $S(\infty)$. Such measures are in 1–1 correspondence with *partition structures* in the sense of Kingman [43]. The set of all K–invariant probability measures on $\mathfrak{S}$ (or partition structures) is a convex set. Its extreme points correspond to *ergodic invariant measures* whose complete classification is due to Kingman [45], [46], see also Kerov [34]. Kingman's result is similar to Thoma's theorem. The decomposition of Ewens' measures μ_t on ergodic components leads to a remarkable one–parameter family of point processes on $(0,1]$ known as *Poisson–Dirichlet distributions.* There is a large literature on Poisson–Dirichlet distributions, we cite only a few works: Watterson [89], Griffiths [28], Vershik–Shmidt [86], Ignatov [30], Kingman [43], [44], [47], Vershik [80], Arratia–Barbour–Tavaré [3]. One can show that the lifting of the Poisson–Dirichlet distribution with parameter $t>0$ is the Poisson process on $(0,+\infty)$ with density $\frac{t}{x}e^{-x}dx$.

In the second problem, one deals with $(U(\infty),\mathfrak{U})$ instead of $(S(\infty),\mathfrak{S})$. Here we again have a distinguished family of invariant measures, see Borodin–Olshanski [18], Olshanski [64]. Their decomposition on ergodic components is described in terms of certain determinantal point processes on $\mathbb{R}^*$. The corresponding correlation kernels are integrable and are expressed through another solution of Whittaker's differential equation (§6.4), see [18]. This subject is closely connected with Dyson's *unitary circular ensemble*, see [18], [64].

For the point processes mentioned above, a very interesting quantity is the position of the rightmost particle in the random point configuration. In the Poisson–Dirichlet case, the distribution of this random variable is given by a curious piece–wise analytic function satisfying a linear difference–differential equation: see Vershik–Shmidt [86], Watterson [89]. For the (discrete and continuous) determinantal point processes arising in harmonic analysis, the distribution of the rightmost particle can be expressed through solutions of certain nonlinear (difference or differential) Painlevé equations: see Borodin [10], Borodin–Deift [12].

References

1. Aldous, D. J.: Exchangeability and related topics. In: Springer Lecture Notes in Math., **1117**, 2–199 (1985)
2. Arratia, R., Barbour, A. D., Tavaré, S.: Poisson processes approximations for the Ewens sampling formula. Ann. Appl. Probab., **2**, 519–535 (1992)
3. Arratia, R., Barbour, A. D., Tavaré, S.: Random combinatorial structures and prime factorizations. Notices Amer. Math. Soc., **44**, no. 8, 903–910 (1997)
4. Baik, J., Deift, P., Johansson, K.: On the distribution of the length of the longest increasing subsequence of random permutations. J. Amer. Math. Soc., **12**, no. 4, 1119–1178 (1999)
5. Biane, Ph.: Representations of symmetric groups and free probability. Advances in Math., **138**, 126–181 (1998)
6. Borodin, A.: Multiplicative central measures on the Schur graph. In: Vershik, A. M. (ed) Representation theory, dynamical systems, combinatorial and algorithmic methods II. Zapiski Nauchnykh Seminarov POMI, **240**, 44–52 (1997) (Russian); English translation: J. Math. Sci., **96**, no. 5, 3472–3477 (1999)
7. Borodin, A.: Characters of symmetric groups and correlation functions of point processes. Funktsional. Anal. Prilozhen., **34**, no. 1, 12–28 (2000) (Russian); English translation: Funct. Anal. Appl., **34**, no. 1, 10–23 (2000)
8. Borodin, A.: Harmonic analysis on the infinite symmetric group and the Whittaker kernel. Algebra Anal. **12**, no. 5, 28–63 (2001) (Russian); English translation: St. Petersburg Math. J., **12**, no. 5, 733–759 (2001)
9. Borodin, A.: Riemann–Hilbert problem and the discrete Bessel kernel. Intern. Math. Research Notices, no. 9, 467–494 (2000), `math/9912093`
10. Borodin, A.: Discrete gap probabilities and discrete Painlevé equations. Duke Math. J., to appear; `math-ph/0111008`
11. Borodin, A.: Asymptotic representation theory and Riemann–Hilbert problem. In this volume; `math/0110318`
12. Borodin, A., Deift, P.: Fredholm determinants, Jimbo–Miwa–Ueno tau–functions, and representation theory. Comm. Pure Appl. Math., **55**, no. 9, 1160–1230 (2002); `math-ph/0111007`
13. Borodin, A., Okounkov, A., Olshanski, G.: Asymptotics of Plancherel measures for symmetric groups. J. Amer. Math. Soc., **13**, 491–515 (2000); `math/9905032`
14. Borodin, A., Olshanski, G.: Point processes and the infinite symmetric group. Math. Research Lett., **5**, 799–816 (1998); `math/9810015`

15. Borodin, A., Olshanski, G.: Distributions on partitions, point processes and the hypergeometric kernel. Comm. Math. Phys., **211**, no. 2, 335–358 (2000); math/9904010
16. Borodin, A., Olshanski, G.: Harmonic functions on multiplicative graphs and interpolation polynomials. Electronic J. Comb., **7** (2000), paper #R28; math/9912124
17. Borodin, A., Olshanski, G.: Z–Measures on partitions, Robinson–Schensted–Knuth correspondence, and $\beta = 2$ random matrix ensembles. In: Bleher, P. M., Its, A. R. (eds) Random matrix models and their applications. Mathematical Sciences Research Institute Publications **40**, Cambridge Univ. Press, 71–94 (2001); math/9905189
18. Borodin, A., Olshanski, G.: Infinite random matrices and ergodic measures. Comm. Math. Phys., **223**, 87–123 (2001); math-ph/0010015
19. Borodin, A., Olshanski, G.: Harmonic analysis on the infinite–dimensional unitary group and determinantal point processes. Ann. Math., to appear; math/0109194
20. Borodin, A., Olshanski, G.: Z-measures on partitions and their scaling limits, math-ph/0210148 (2002)
21. Daley, D. J., Vere-Jones, D.: An introduction to the theory of point processes. Springer series in statistics, Springer, (1988)
22. Deift, P.: Integrable operators. In: Buslaev, V., Solomyak, M., Yafaev, D. (eds) Differential operators and spectral theory: M. Sh. Birman's 70th anniversary collection. American Mathematical Society Translations, ser. 2, **189**, Providence, R.I., AMS, (1999)
23. Diaconis, P., Freedman, D.: Partial exchangeability and sufficiency. In: Statistics: Applications and New Directions (Calcutta, 1981), Indian Statist. Inst., Calcutta, 205–236 (1984)
24. Dyson, F. J.: Statistical theory of the energy levels of complex systems I, II, III. J. Math. Phys., **3**, 140–156, 157–165, 166–175 (1962)
25. Edrei, A.: On the generating functions of totally positive sequences II. J. Analyse Math., **2**, 104–109 (1952)
26. Ewens, W. J.: The sampling theory of selectively neutral alleles. Theoret. Population Biology, **3**, 87–112 (1972)
27. Ewens, W. J.: Population Genetics Theory – the Past and the Future. In: Lessard, S. (ed) Mathematical and Statistical Developments of Evolutionary Theory. Proc. NATO ASI Symp., Kluwer, Dordrecht, 117–228 (1990)
28. Griffiths, R. C.: On the distribution of points in a Poisson Dirichlet process. J. Appl. Probab., **25**, 336–345 (1988)
29. Hida, T., Nomoto, H.: Gaussian measure on the projective limit space of spheres. Proc. Japan Academy, **40**, 31–34 (1964)
30. Ignatov, Ts.: On a constant arising in the asymptotic theory of symmetric groups and on Poisson–Dirichlet measures. Teor. Veroyatnost. Primenen., **27**, no. 1, 129–140 (1982) (Russian); English translation: Theory Probab. Appl., **27**, 136–147 (1982)
31. Its, A. R., Izergin, A. G., Korepin, V. E., Slavnov, N. A.: Differential equations for quantum correlation functions. Intern. J. Mod. Phys., **B4**, 1003–1037 (1990)
32. Ivanov, V., Olshanski, G.: Kerov's central limit theorem for the Plancherel measure on Young diagrams. In: Fomin, S. (ed) Symmetric Functions 2001: Surveys of Developments and Perspectives. NATO Science Series II. Mathematics, Physics and Chemistry, vol. 74, Kluwer, 93–151 (2001)

33. Johansson, K.: Discrete orthogonal polynomial ensembles and the Plancherel measure. Ann. Math. (2), **153**, no. 1, 259–296 (2001); `math/9906120`
34. Kerov, S. V.: Combinatorial examples in the theory of AF-algebras. In: Differential geometry, Lie groups and mechanics X, Zapiski Nauchnykh Seminarov LOMI, **172**, 55–67 (1989) (Russian); English translation: J. Soviet Math., **59**, no. 5, 1063–1071 (1992)
35. Kerov, S. V.: Generalized Hall–Littlewood symmetric functions and orthogonal polynomials. In: Vershik, A. M. (ed) Representation Theory and Dynamical Systems. Advances in Soviet Math., Vol. 9, Amer. Math. Soc., Providence, R.I., 67–94 (1992)
36. Kerov, S. V.: Gaussian limit for the Plancherel measure of the symmetric group. Comptes Rendus Acad. Sci. Paris, Série I, **316**, 303–308 (1993)
37. Kerov, S. V.: Subordinators and the actions of permutations with quasi–invariant measure. In: Zapiski Nauchnyh Seminarov POMI, **223**, 181–218 (1995) (Russian); English translation: J. Math. Sci. (New York), **87**, no. 6, 4094–4117 (1997)
38. Kerov, S. V.: Anisotropic Young diagrams and Jack symmetric functions. Funktsional. Anal. Prilozhen., **34**, no. 1, 51–64 (2000) (Russian); English translation: Funct. Anal. Appl., **34**, no. 1, 41–51 (2000)
39. Kerov, S., Okounkov, A., Olshanski, G.: The boundary of Young graph with Jack edge multiplicities. Intern. Math. Res. Notices, **1998:4**, 173–199 (1998); `q-alg/9703037`.
40. Kerov, S., Olshanski, G.: Polynomial functions on the set of Young diagrams. Comptes Rendus Acad. Sci. Paris Sér. I, **319**, 121–126 (1994)
41. Kerov, S., Olshanski, G., Vershik, A.: Harmonic analysis on the infinite symmetric group. A deformation of the regular representation. Comptes Rend. Acad. Sci. Paris, Sér. I, **316**, 773–778 (1993); detailed version in preparation
42. Kerov, S. V., Tsilevich, N. V.: Stick breaking process generates virtual permutations with Ewens distribution In: Zapiski Nauchnyh Seminarov POMI, **223**, 162–180 (1995) (Russian); English translation: J. Math. Sci. (New York), **87**, no. 6, 4082–4093 (1997)
43. Kingman, J. F. C.: Random discrete distributions. J. Royal Statist. Soc. B, **37**, 1–22 (1975)
44. Kingman, J. F. C.: The population structure associated with the Ewens sampling formula. Theoret. Population Biology, **11**, 274–283 (1977)
45. Kingman, J. F. C.: Random partitions in population genetics. Proc. Roy. Soc. London A., **361**, 1–20 (1978)
46. Kingman, J. F. C.: The representation of partition structures. J. London Math. Soc. (2), **18**, 374–380 (1978)
47. Kingman, J. F. C.: Poisson processes. Oxford University Press (1993)
48. Lenard, A.: Correlation functions and the uniqueness of the state in classical statistical mechanics. Comm. Math. Phys, **30**, 35–44 (1973)
49. Macchi, O.: The coincidence approach to stochastic point processes. Adv. Appl. Prob., **7**, 83–122 (1975)
50. Macchi, O.: The fermion process — a model of stochastic point process with repulsive points. In: Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes and of the Eighth European Meeting of Statisticians (Tech. Univ. Prague, Prague, 1974), Vol. A, Reidel, Dordrecht, 391–398 (1977)

51. Murray, F. J., von Neumann, J.: On rings of operators IV. Ann. Math. **44**, 716–808 (1943)
52. Mehta, M. L.: Random matrices, 2nd edition. Academic Press, New York (1991)
53. Neretin, Yu. A.: Hua type integrals over unitary groups and over projective limits of unitary groups. Duke Math. J., **114**, 239–266 (2002); `math-ph/0010014`
54. Nagao, T., Wadati, M.: Correlation functions of random matrix ensembles related to classical orthogonal polynomials. J. Phys. Soc. Japan, **60**, no. 10, 3298–3322 (1991)
55. Naimark, M. A.: Normed rings. Nauka, Moscow (1962) (Russian); English translation: Normed algebras. Wolters–Noordhoff, Groningen, (1972)
56. Okounkov, A.: Thoma's theorem and representations of infinite bisymmetric group. Funktsion. Anal. Prilozhen., **28**, no. 2, 31–40 (1994) (Russian); English translation: Funct. Anal. Appl., **28**, no. 2, 101–107 (1994)
57. Okounkov, A. Yu.: On representations of the infinite symmetric group. In: Vershik, A. M. (ed) Representation Theory, Dynamical Systems, Combinatorial and Algorithmic Methods II, Zap. Nauchn. Semin. POMI, **240**, 167–229 (1997) (Russian); English translation: J. Math. Sci. (New York), **96**, no. 5, 3550–3589 (1999)
58. Okounkov, A.: $SL(2)$ and z–measures. In: Bleher, P. M., Its, A. R. (eds) Random matrix models and their applications. Mathematical Sciences Research Institute Publications, **40**, Cambridge Univ. Press, 407–420 (2001); `math/0002136`
59. Okounkov, A.: Infinite wedge and measures on partitions. Selecta Math. (New Ser.), **7**, 57–81 (2001); `math/9907127`
60. Okounkov, A., Olshanski, G.: Shifted Schur functions. Algebra i Analiz, **9**, no. 2, 73–146 (1997) (Russian); English translation: St. Petersburg Math. J., **9**, no. 2, 239–300 (1998)
61. Olshanski, G.: Unitary representations of infinite-dimensional pairs (G, K) and the formalism of R. Howe. Doklady AN SSSR, **269**, 33–36 (1983) (Russian); English translation: Soviet Math. Doklady, **27**, no. 2, 290–294 (1983)
62. Olshanski, G.: Unitary representations of infinite–dimensional pairs (G, K) and the formalism of R. Howe. In: Vershik, A., Zhelobenko, D. (eds) Representation of Lie Groups and Related Topics. Advanced Studies in Contemporary Math., **7**, Gordon and Breach Science Publishers, New York etc., 269–463 (1990)
63. Olshanski, G.: Unitary representations of (G, K)-pairs connected with the infinite symmetric group $S(\infty)$. Algebra i Analiz, **1**, no. 4, 178–209 (1989) (Russian); English translation: Leningrad Math. J., **1**, 983–1014 (1990)
64. Olshanski, G.: The problem of harmonic analysis on the infinite–dimensional unitary group. J. Funct. Anal., to appear; `math/0109193`
65. Olshanski, G., Regev, A., Vershik, A.: Frobenius–Schur functions. In: Joseph, A., Melnikov, A., Rentschler, R. (eds) Studies in Memory of Issai Schur. Birkhäuser, to appear; `math/0110077`.
66. Olshanski, G.: Point processes and the infinite symmetric group. Part I: The general formalism and the density function. `math/9804086` (1998)
67. Borodin, A.: Point processes and the infinite symmetric group. Part II: Higher correlation functions. `math/9804087` (1998)
68. Borodin, A., Olshanski, G.: Point processes and the infinite symmetric group. Part III: Fermion point processes. `math/9804088` (1998)
69. Borodin, A.: Point processes and the infinite symmetric group. Part IV: Matrix Whittaker kernel. `math/9810013` (1998)
70. Olshanski, G.: Point processes and the infinite symmetric group. Part V: Analysis of the matrix Whittaker kernel. `math/9810014` (1998)

71. Pickrell, D.: Measures on infinite dimensional Grassmann manifold. J. Funct. Anal., **70**, 323–356 (1987)
72. Rozhkovskaya, N. A.: Multiplicative distributions on Young graph. In: Vershik, A. M. (ed) Representation Theory, Dynamical Systems, Combinatorial and Algorithmic Methods II. Zapiski Nauchnykh Seminarov POMI,**240**, Nauka, St. Petersburg, 246–257 (1997) (Russian); English translation: J. Math. Sci. (New York), **96**, no. 5, 3600–3608 (1999)
73. Shimomura, H.: On the construction of invariant measure over the orthogonal group on the Hilbert space by the method of Cayley transformation. Publ. RIMS Kyoto Univ., **10**, 413–424 (1974/75)
74. Soshnikov, A.: Determinantal random point fields. Uspekhi Mat. Nauk, **55**, no. 5, 107–160 (2000) (Russian); English translation: Russian Math. Surveys, **55**, no. 5, 923–975 (2000); `math/0002099`
75. Stratila, S., Voiculescu, D.: Representations of AF–algebras and of the group $U(\infty)$. Springer Lecture Notes, **486** (1975)
76. Thoma, E.: Die unzerlegbaren, positive–definiten Klassenfunktionen der abzählbar unendlichen, symmetrischen Gruppe. Math. Zeitschr., **85**, 40–61 (1964)
77. Thoma, E.: Characters of infinite groups. In: Arsene, Gr., Strătilă, S., Verona, A., Voiculescu, D. (eds) Operator Algebras and Group Representations, Vol. 2, Pitman, 23–32 (1984)
78. Tracy, C. A., Widom, H.: Universality of distribution functions of random matrix theory II. In: Harnad, J., Sabidussi, G., Winternitz, P. (eds) Integrable Systems: From Classical to Quantum. CRM Proceedings & Lecture Notes, **26**, Amer. Math. Soc., Providence, 251–264 (2000)
79. Vershik, A. M.: Description of invariant measures for the actions of some infinite-dimensional groups. Doklady AN SSSR, **218**, 749–752 (1974) (Russian); English translation: Soviet Math. Doklady, **15**, 1396–1400 (1974)
80. Vershik, A. M.: Asymptotic distribution of decompositions of natural numbers into prime divisors. Dokl. Akad. Nauk SSSR, **289**, no. 2, 269–272 (1986) (Russian); English translation: Soviet Math. Doklady, **34**, 57–61 (1987)
81. Vershik, A. M.: Statistical mechanics of combinatorial partitions, and their limit shapes. Funktsional. Anal. Prilozhen., **30**, no. 2, 19–39 (1996) (Russian); English translation: Funct. Anal. Appl., **30**, 90–105 (1996)
82. Vershik, A. M., Kerov, S. V.: Characters and factor representations of the infinite symmetric group. Doklady AN SSSR, **257**, 1037–1040 (1981) (Russian); English translation: Soviet Math. Doklady, **23**, 389–392 (1981)
83. Vershik, A. M., Kerov, S. V.: Asymptotic theory of characters of the symmetric group. Funktsion. Anal. Prilozhen., **15**, no. 4, 15–27 (1981) (Russian); English translation: Funct. Anal. Appl., **15**, no. 4, 246–255 (1981)
84. Vershik, A. M., Kerov, S. V.: Characters and factor representations of the infinite unitary group. Doklady AN SSSR, **267**, no. 2, 272–276 (1982) (Russian); English translation: Soviet Math. Doklady, **26**, 570–574 (1982)
85. Vershik, A. M., Kerov, S. V.: Locally semisimple algebras. Combinatorial theory and the K_0 functor. In: Itogi Nauki, Sovr. Probl. Mat., Noveish. Dostizh., VINITI, **26**, 3–56 (1985) (Russian); English translation: J. Soviet Math., **38**, 1701–1733 (1987)
86. Schmidt, A., Vershik, A.: Limit measures that arise in the asymptotic theory of symmetric groups. I, II. Teor. Verojatn. i Prim., **22**, no. 1, 72–88 (1977); **23**,

no. 1, 42–54 (1978) (Russian); English translation: Theory of Probab. Appl., **22**, 70–85 (1977); **23**, 36–49 (1978)
87. Voiculescu, D.: Représentations factorielles de type II_1 de $U(\infty)$. J. Math. Pures Appl., **55**, 1–20 (1976)
88. Wassermann, A. J.: Automorphic actions of compact groups on operator algebras. Thesis, University of Pennsylvania (1981)
89. Watterson, G. A.: The stationary distribution of the infinitely many–alleles diffusion model. J. Appl. Probab., **13**, 639–651 (1976)
90. Yamasaki, Y.: Projective limit of Haar measures on $O(n)$. Publ. RIMS, Kyoto Univ., **8**, 141–149 (1972/73)
91. Yamasaki, Y.: Kolmogorov's extension theorem for infinite measure. Publ. RIMS, Kyoto Univ., **10**, 381–411 (1974/75)

Two lectures on the asymptotic representation theory and statistics of Young diagrams

A. Vershik

St. Petersburg Department
Steklov Mathematical Institute
Fontanka, 27
St. Petersburg 197011
Russia
vershik@pdmi.ras.ru

The first lecture: The main problems and short history. Statistics on Young diagrams and partitions.

Introduction: What is asymptotical representation theory

Investigation of classical groups of high ranks leads to two kinds of problems. Questions of the first kind deal with asymptotical properties of groups, their representations, characters and other attributes as group rank grows to infinity. Another kind of questions (in the spirit of infinite dimensional analysis) deal with properties of infinite dimensional analogues of classical groups. Let us discuss, for instanse, the most simple nontrivial example of classical group series, that is the series of symmetrical groups S_N. Typical question of the first kind is what is the structure of the symmetric group of high rank and its representations? A question of the second kind is what can be told about infinite symmetric group, that is a group of finite permutations of natural numbers? Both kinds of questions are closely connected, but it is appropriate mention here that while questions of the first kind seem to be more natural and their importance was emphasized as early as 1940s by H. Weyl [33] and J. von Neumann [14], nevertheless the functional analysis evolved mainly into investigation of infinite dimensional groups, which is certainly caused by its applications to physics. Questions of both kinds are parts of asymptotical representation and group theory, but proper asymptotical problems were, strangely enough, investigated much less and thus they make up the main part of the theory. In a wide context their study was started in the 1970s, but a lot of separate problems were considered earlier. It is important to emphasize from the very beginning that the questions considered deal with the structure of groups and their representations as a whole rather than with investigation of particular functionals on groups or its representations. For example, the

A. Borodin et al.: LNM 1817, A.M. Vershik (Ed.), pp. 161–182, 2003.

question about the distribution of maximal cycle length in typical permutation was considered by V. L. Goncharov in 1940s [4] while the question about the general structure of typical conjugacy class in symmetric group was studied only in the middle of 1970s by A. Schmidt and myself [19, 20]. The same can be said about representations of the symmetric group.

It turned out that the questions under consideration are related to a wide range of mathematical and physical problems which at the first sight have nothing common with the original issues. To take some liberties, we can think of these problems as of problems from statistical physics of algebraic (and, particularly, group) structures. At first, asymptotical theory naturally leads investigators to probabilistic statements of questions. Indeed, in order to make the limiting transition in an ensemble of representations, conjugacy classes etc. of a group we have to choose some measure on this ensemble. The measure differs depending on the problem, for example, it is natural to consider the Haar measure to study symmetric groups, and the Plancherel measure to study their representation, and there are a lot of other useful measures as will be explained below. When the problem becomes probabilistic one we can state traditional questions like a law of large numbers, limit theorems, large deviation theorems etc. Answering these questions we immediately see that proofs use different methods than that of classical probability theory and the answers, that is limit distributions, measures etc., are different from usual ones. Nevertheless, the common principles are preserved, for example, the effect similar to the law of large numbers take place, although in a peculiar form; for representations of the symmetric group it can be stated in the following way: almost all (in the sense of the Plancherel measure) irreducible representations of the symmetric group S_N as $N \to \infty$ come close to each other in the natural sense (the corresponding characters, matrix elements, etc. have close values after a necessary renormalization). These probabilistic problems can not be reduced to the schemes with sequences of independent or weakly dependent random variables unlike the traditional ones, but the general principles (for instance, subadditive ergodic theorem) remain useful.

One important circumstance should be mentioned from the very beginning: in many cases such as symmetric group, general linear group and others the parametrization of representations, conjugacy classes etc. have visual geometric sense, namely the parameters are partitions, Young diagrams and so on. Thus the problems considered get a combinatorial and geometrical meaning and become questions about random partitions, random Young tableaux and diagrams and so on. Direct utilization of these parametrizations allow us to include our questions in a wide range of problems concerning so called "limit shape", which arise in the theory of random growth understood in the very wide sense. Later we will give several examples and now emphasize that the problems of asymptotical group and representation theory often include some classical questions as particular cases; they can be solved by new methods while they remained unsolved before. One of the striking examples is the Ulam problem about asymptotical formula of the maximal length of monotone subsequence of a sequence of independent random variables. Its first full solution was found as a corollary of the proof of already mentioned statement about typical representation of the symmetric group or, in the other words, about typical Young diagrams with respect to the Plancherel measure (Vershik and Kerov [7], see also Logan and Shepp [12]).

Recent years a number of problems about fluctuations of typical diagram and on the microstructure of random Yound diagrams have been solved, it turned out that the answers are suitable not only for this problem but after minor generalizations give solutions of the problems about the spectrum of random Gaussian matrices and of some physical issues. The general description of some of these new methods is given in the lectures of P. Deift in the present volume (I mean the Riemann–Hilbert problem and asymptotics of orthogonal polynomials) and these methods lead to integrable problems (Painlevé equations). Other, more straightforward approach to investigation of asymptotical microstructure of Young diagrams based so called point determinant processes was developed by Okounkov, Borodin and Olshansky. They found direct links between formulas used for calculation of correlation functions for these point processes and hierarchies of integrable problems. Among mathematicians who took part in this impressive progress I mention only principal ones: Deift–Baik–Johansson, Okounkov–Borodin–Olshansky, Widom–Tracy, whose works in its turn were based on previous results (Izergin–Its–Korepin, Krein–Gohberg, Vershik–Kerov, see references in lectures in this volume).

The whole range of questions of asymptotial representation theory is closely connected to the probability theory and the measure theory and this connection leads to discovery of new classes of infinite dimensional distributions and new properties of known distributions. To illustrate this I can name the Airy distribution and the Airy process which appeared in problems concerning the fluctuation of Young diagrams, and the class of point determinant processes, interest in which has quickened because correlation functions of these proesses coinsides with kernels of integrable operators in integrable problems theory. Another example, the Poisson–Dirichlet measure has been known and was used in combinatorics and mathematical genetics, but a number of its new properties and symmetries were discovered only after it had appeared as the limit measure for conjugacy classes of symmetric groups.

These and other applications, possible unexpected ones, are good confirmations for a thesis that fundamental and natural asymptotical problems of purely combinatorial or algebraic nature often have deep applications to physics, analysis etc.

Even more paradoxical conclusion from asymptotical representation theory experience is its reverse influence on the classical representation theory of finite and finite-dimensional groups. Here we should keep in mind that asymptotical approach requires more standartized point of view to the theory of prelimit groups and their representation, in particular, inductive construction of the theory. This “look from infinity” on the finite theory leads to serious simplifications and clarifications of the foundations of classical representation theory. This can be illustrated by the revision of the symmetric group representation theory undertaken by A. Vershik and A. Okounkov [27, 16] and later of other Coxeter groups by A. Vershik and I. Pushkarev [18]. This revision used the Gelfand–Tsetlin and Murphy–Jucys bases. Particularly, the work by A. Vershik and A. Okounkov clarify the appearance of the Young tables in this theory as a spectrum of the Gelfand–Tsetlin algebra in the Murphy–Jucys basis and leads directly to realization of the presentation in Young's orthogonal form. This at last takes away the question about somehow synthetic and nontrival combinatorial constructions which are present in all expositions of classical representation theory starting from the pioneer works of Frobenius, Schur and Young with further refinements of Weyl, von Neumann and others.

Now we will touch the second kind of problems, that is problems about limit groups for classical series. First of all we should mention that for each given group series, for instanse for symmetric groups, there are numerous limit groups: we can consider the group of finite permutations of a countable set, the group of all permutations, or intermediate groups. Obviously, the choice should be made depending on problems and applications. Moreover this choice is not evident as we will see by the example of groups of matrices over finite field. On the other hand these infinite groups are usually "wild", that is their irreducible unitary representations do not allow good ("smooth") classification. So the investigation can be fruitful only after applying a restriction on the category of representations. Only one category (or class) of representations is doubtless. It is the category of representations with a trace which generalizes the catogory of representations of finite groups in the most direct way. Its study can be significantly reduced to an investigation of finite and semifinite character on the group. Another approach developed by G. Olshansky [17] is a theory of symmetric pairs which investigates only representations whose restriction to a certain subgroup belongs to a fixed simple class of representations.

Infinite dimensional (or, more generally, "large", that is similar to infinite dimensional) group can have few representations having finite or semifinite character, or can, on the contrary, have them abundantly and unclassified. Remarkable property of limit groups of classical group series is that, from the one hand, the quantity of their representations of stated type is large enough to construct substantial harmonic analysis and, on the other hand, these representations can be effectively described and smoothly parametrized. The characters of infinite symmetric group were found in a pioneer work of E. Thoma [22]; later new proof of his theorem based on ergodic method with interpretation of parameters of characters in terms of the concrete representations appeared in [8], and one more proof was made by Okounkov [15] in terms of spectral theory. The description and investigation of charaters of classical group limits is another kind of questions which belongs to asymptotical representation theory. Symmetric group plays here a role of main example as well, and all other groups (infinite unitary, symplectic etc.) can be to a great extent reduced to symmetric one.

In the first lecture I will mention only several questions about limit shapes of random diagrams and partitions; they mainly belong to the first kind of problems. In the second lecture I will tell about the asymptotical theory of infinite matrices over finite field; it is closer to the second kind of questions. These results were obtained jointly with the late Sergey Vasilievich Kerov.

Other problems of asymptotical representation theory concern a construction of implicit realizations of representations with characters and similar ones, investigation of limits of Grottendick groups of classical series and K-theory of locally semisimple algebras (particularly, group algebras of locally finite groups), investigation of continual limits and theory generalizations (continual diagrams, relations to problems of moments and so on). I will not stop on these topics now. In conclusion, note that asymptotical theory in the above sense began with works by me and my students in the early 1970s (A. Vershik and A. Schmidt [19, 20], A. Vershik and S. Kerov, later by N. Tsilevich [24], Yu. Yakubovich [32]) and was continued by G. Olshansky and his students (A. Okounkov, A. Borodin, V. Ivanov), see references in their lectures in this volume. Invaluable contribution on the development of this theory was made by S. Kerov (1946–2000); his list of works can be found in [31] and his dotoral thesis will be soon published in one of AMS series. Recent progress in the theory was mentioned above, more about it can be found in the current volume. In the nearest future it is supposed to prepare a number of surveys of asymptotical representation theory with a list of open problems.

The work on these lectures was supported in part by FRBF grant No. 02-01-00093.

Asymptotical properties of statistics on partitions and Young diagrams

Without further detailed explanations we start with description of questions and results arising in asymptotical representation theory and close areas concerning statistical diagram theory. Essentially, very wide range of problems can be reduced to this theory, including problems of representation theory of symmetric and other groups as a very special case. Remind that the Young diagrams with n cells parametrize in a natural way all irreducible representations of the symmetric group S_n, as well as conjugacy classes of this group. Denote the set of all diagrams with n cells (or in the other words all partitions of the number n) by $\mathcal{P}_n$ and their union over n by $\mathcal{P}$. Natural measure on diagrams as conjugacy classes is the Haar measure, and the probability of the class (or the diagram or the partition) is proportional to the number of elements in conjugacy class:

$$\mathrm{Prob}_n(g) = \frac{1}{n! \prod_k r_k(g)! k^{r_k(g)}}, \qquad r_k(g) = \#\{\text{cycles in } g \in S_n \text{ of length } k\}$$

If one regards a diagram as an irreducible representation of the group S_n then in view of the Bernside formula (sum of the squares of dimensions of representations is equal to the order of the group) it is natural to consider the Plansherel measure: the probability is proportional to a square of the dimension of the irreducible representation. Thus it is given by Frobenius formula or more handy by the hook formula (FRT=Frame–Robinson–Threll):

$$\mathrm{Prob}_n(\lambda) = \frac{\dim^2(\lambda)}{n!} = \frac{n!}{\prod_{c\in\lambda} h^2(c)}$$

where $\lambda \in \mathcal{P}_n$ is the partition (the Young diagram), c runs over all cells of λ and $h(c)$ is hook length of cell c.

Asymptotical properties of these statistics as n grows to infinity are absolutely different.

Theorem (see [19]). *Regard a conjugacy class of the substition (that is an element of the group S_n) as nonincreasing finite numeric sequence of cycle lengthes of the permutation normed by n. Then the sequence of the Haar measures as measures on infinite dimensional symplex of nonnegative monotone nonincreasing sequences with a sum not exceeding 1 is weakly converging (in the symplex topology) to a measure which is called the Poisson–Dirichlet measure ($PD(1)$).*

Note that this measure appeared in numerous situations and was investigated in detail (see [25] and references therein).

This result has a lot of corollaries and in fact gives the full information about limit behaviour of cylindric (i.e. depending only on finite number of cycles) functionals on the symmetric group. Here is a typical example:

> For n large enough, more then $0.99n!$ permutations in symmetric group S_n have $0.99n$ permuting elements in 11 cycles [20, 26].

Now turn to the Plancherel measure.

Theorem (see [7, 9]). *Regard the Young diagram as a configuration on the positive part of the two dimensional lattice $\mathbf{Z}_+^2$ and norm it by $\sqrt{n}$. Then the sequence of images of the Plancherel measures (see above) weakly converges (in the weak topology on the quadrant $\mathbf{R}_+^2$) to the δ-measure concentrated on the curve Ω having the following equation in the coordinate system rotated by $45°$:*

$$\Omega(u) = \begin{cases} \frac{2}{\pi}\left(u \arcsin \frac{u}{2} + \sqrt{4-u^2}\right), & |u| \le 2; \\ |u|, & |u| > 2. \end{cases}$$

Fundamental difference is seen at the first sight: in the second case the limit is a δ-measure (this situation is called *ergodic*) while in the first case the limit was a nondegenrated measure on a space of numeric sequences (*nonergodic* case). We will see that this difference is not determined by the distinction of normalization.

The proofs of these theorems are also differ a lot. It is relatively simple to ascertain the existance of a limit for the Haar measures and thus the investigation of a limit measure went far enough; limit joint distributions, functional equation for density, asymptotics of typical sequence, central limit theorem etc. were found.

For the Plancherel measure the limit curve is more comlicated to find, but now there are several proofs of this fact and also of central limit theorem in a direction transversal to the curve (see the lecture of R. Hora in the current volume). As for the limit microstructure of the diagram and, particularly, for a question about fluctuation of lengthes of the first and subsequent rows, its solution has been late for 20 years (see above and other lectures in this volume) but now we have a full information about both microstruture of bulk of the diargam and its border. The connections to integrable problems, the Riemann–Hilbert problem, point determenant processes became clear during this investigation.

Evident formulas for prelimit probabilities make both cases easier. Even while formulas are very cumbersome in both problems the question about limit transition meets only technical difficulties. However the probability is often given implicitly in many natural (in particular physical and applied) problems.

There are wide range of applications of the stated above result about limit curve. One of more bright examples is the solution of the Ulam problem about maximal length of monotone increasing subsequence of random variables distributed uniformly on the unit interval (or the same sequence in the permutation). The answer (see [7, 9]) is that the ratio ot the length of this subsequence and $2\sqrt{n}$ tends to 1 with the probability 1. Because of the RSK algorithm (Robinson–Shenstedt–Knuth) this length coincides with the length of the first row of the random Young diagram with respect to the Plancherel measure, see also the lectures of P. Deift. Lower estimation of this problem can be easily obtained from the theorem of the limit shape of the Young diagram due to [7, 12]. Upper estimation was obtained in [7], detailed proof can be found in [9]. Another example of the same corollary from the representation theory is the following: the restriction of typical irreducible representation of S_n to its subgroup $S_{\sqrt{n}}$ contains an invariant vector, and this estimate is exact.

Let us look a bit wider on these topics and extend the class of measures on diagrams and partitions. We skip most examples of these measures (see [28] but concentrate on a certain important cases and general principles. By anology with statistical physics sometimes it is useful to pass from measures on $\mathcal{P}_n$ ("microcanonical ensemble") to measures on $\mathcal{P}$ ("grand canonical ensemble"), to find asymptotics for the latter over continuous parameter and then proof that both ensembles are equivalent, or in the other words that the limits coinside. This techniques is very usefull for so called multiplicative measures (see [28, 29, 32]) but sometimes it is unsuitable.

Suppose that the probability measure μ_n on the set of diagrams $\mathcal{P}_n$ is given, then the following three questions arise:

a) What normaization of diagrams is good in the sense that there exists a weak limit of measures μ_n in a space of curves (i.e. continuous diagrams) or in a space of sequences (of lengthes of diagram rows or columns)?
   Of course, there always exist trival normalizations such that the limit is degenerate but we are speaking about the unique nontrivial normalization. For instanse, in the above examples the normalizations are as follows. For the Haar measure only one coordinate (row lengthes) is normalized by n while the second coordinate is fixed. For the Plancherel measures the normalization of both axes is equal to $\sqrt{n}$. If these normalizations are altered then the limit becomes degenerate if it exists.

b) Is this limit measure a δ-measure or dispersed measure (is it ergodic or nonergodic case)? In the nonergodic case the difference between trivial and nontricial normalization is evident because the limit measure is non-degenerate, but in the ergodic case this difference is

c_1) In the nonergodic case find properties of the limit measure, in particular describe its support, give its characterization in a class of all measures.
   The practice shows that limit measures arising in asymptotical algebraic problems usually have natural charaterizing properties, for instanse, the Poisson–Dirichlet measure which is the limit in the problem about conjugacy classes of the symmetric group and has numerous characterizing properties. The same measure appears in the question about the distribution of prime factors in typical natural numbers and others.

c_2) In the ergodic case find the distribution of fluctuations in different scales, find rate function (formulate variational principle) and prove large deviation principle.

Note that in the last case the distinctions between scales in different "directions" of fluctuations can be essential. For example, the central limit theorem for the Plancherel measure (Kerov; other version by Hora and Olshansky, see the references in their lectures in this volume) is much easier then limit theorems in "tangent" scales. For instance, the correct scale for fluctuations of first row length ($n^{1/6}$) is different from the natural scale in diagonal direction.

The following natural example belongs to A. Borodin. It is the problem about asymptotical distribution of sizes of the Jordan blocks of upper triangular nilpotent matrices over finite field as matrix order grows to infinity. This case is ergodic but the limit measure is a δ-measure on the sequences (not on functions as for the Plancherel measure).

Theorem (see [1]). *The limit distribution of decreasing normalized by n sizes $l_1(n) \geq l_2(n) \geq \dots$ of the Jordan blocks of upper triangular nilpotent matrices of the field with q elements (with uniform distribution) is the δ-measure concentrated on the sequence $((1-t),(1-t)t,(1-t)t^2,\dots)$ where $t = q^{-1}$. In the other words, for each $k \in \mathbf{N}$ and $\epsilon > 0$ consider the set of upper* *triangular nilpotent matrices of order n such that normalized first (ordered decreasing) k sizes of the Jordan blocks differs from the first k numbers of this sequence less than by ϵ, then the uniform measure of this set tends to 1 as n grows to infinity.*

This means that asymptotically the sizes of the Jordan blocks decrease as the geometric progression with ratio $t = q^{-1}$. Central limit theorem for this case and the estimates for the rate of (exponential) convergence can be found in the same paper. It is interesting that if we considers all nilpotent matrices (not only upper triangular ones) then will find out that almost all of them asymtotically have only one Jordan block, more rigorously almost all matrices have a one block sized $n(1 - o(1))$ (Fulman, [3]). However, for our considerations the upper triangular matrices are of the most interest (see Lecture 2).

So this example lies between the Haar and the Plancherel measures. The proof is rather simple: it is sufficient to calculate the expectation of the size of the Jordan block and to make certain that it grows linearly as n grows, and at the same time the dispersion vanishes. This calculation is done with the help of one useful and general technique, that is by a construction of the Markov chain for transition from matrices of the order $n-1$ to matrices of the order n which plays a role of the dynamics of the Jordan blocks. The same technique was used earlier for symmetric groups.

Thus, in the ergodic case the limit δ-measure can be concentrated both on a sequence and on a curve. The natural example of nonergodic behaviour such that the support of the limit measure is a continual set of curves is unknown yet.

The problems considered above belong to a class of so called "limit shape problems" which often arise in biology (colony growth), medicine (infection spreading), ecology (fire expansion), and also in physics, because one of the most popular problems in quantum field theory and statistical physics (DLA=diffusion limit of aggregation) exactly fits in this class; moreover, one specialization of this problem is the problem about the Plancherel measure, and in the full generality this problem adjoins the SLE (Stochastic Loewner equation) problems. And, surely, limit shape problems appears in geometry (dynamics of polyhedrons, curves etc.), probability theory (nontraditional random walks, heap growth) and so on. Interpretation of the problems under consideration affects diverse mathematical areas.

The second lecture: Introduction to the theory of the characters of the group of infinite matrices over finite fields.

1 Definitions and the problem

This lecture will be devoted to the topics which we had studied jointly with S. V. Kerov (1946–2000), and now it is on the way to the detailed publication. The frist short announcement had been published in [15]. Here I will explain some details and the first constructions, mainly concerned to the subjects of that paper and also I will retell some classical facts about $GL(n, F_q)$ in a new form which is convinient for the further work. In fact I will discuss the general approach to the problem.

This theory is another part of the asymptotic representation theory, namely about the represetations and characters of the group of infinite matrices over finite field. This is also the next step after the theory of infinite symmetric group toward the more general algebraic theory. A different approach consists in the consideration of the asymptotic theory of the general Coxeter groups, or continious groups like $U(\infty)$, or other infinite dimensional Lie groups and so on. But it turned out that the group of infinite matrices over finite field plays the special role, – its theory is more complicated than examples mentioned above and revealed some new effects with comparison to symmetric group more deeper than in the case of other Coxeter or classical Lie groups.

On the other side, we know from the theory of representations of $GL(n, F_q)$ that many formulas with parameter q turn out to the formulas for symmetric group S_n if we set $q = 1$, symbolically:

$$\text{“}S_n = GL(n, F_1)\text{”}.$$

Although this equality is metaphorical one because F_1 is nonsense (there are no such a field in which $0 = 1$), there are many reasons to say like this and we can consider these groups like “q-deformation” of the symmetric groups. What we intend to do first of all is to give right “q-deformation” of infinite symmetric group. We will see that even this step is not evident.

Let us consider the group $GL(\infty, F_q)$ of all infinite invertible matrices over finite field F_q with q elements of the form $Id+g$, where g is a finite matrix and Id infinite identity matrix. This is locally finite countable group which can be considered as inductive limit of the groups $GL(n, F_q)$ with natural imbedding: $GL(n, F_q) \subset GL(n + 1, F_q)$. Now we can put a question: Is it true (even in methaphorcal sense) that

$$\text{“}S_\infty = GL(\infty, F_1)\text{”?}$$

In the paper which followed to the paper with formulas for the characters of infinite symmetric group E. Thoma [23] got calculation of the characters of the group $GL(\infty, F_q)$. The final solution was done by his collaborator R. Skudlarek [21]; but surprisingly the answer was drastically differ from the case of infinite syummetric group and very poor: it happened that there are only countably many of the indecomposable normalized by 1 at unity characters of the group $GL(\infty, F_q)$ (see [23, 21]):

$$\chi_n(e) = 1, \quad \chi_n(g) = q^{-n(\mathrm{rk}(I-g))}, \qquad n = 1, \dots, \infty; \quad g \neq e; \quad \chi_0 \equiv 1.$$

These characters as well as the corresponding representations are not so interesting and the answer *does not give a good base for the serious harmonic analysis* on the group as it was in the case of infinite symmetric group. The reason was that the problem had been posed in too straightfoward way and in a sense had ignored the classical representation theory of the groups $GL(n, F_q)$. As we will see even the choice of the group which plays the role of the analog of S_∞ in this case is not so evident (see below).

It worth mentioning that the representation theory of the groups $GL(n, F_q)$ was developed by several authors starting from G. Frobenius, I. Schur and at nowadays by R. Steinberg and especially by J. Green who had suggested a systematic approach to the theory and found the list of characters in his paper [5]. Later D. K. Faddeev [2] revised Green's approach using the Mackey theory of inducing representations, it gave the serious simplification of whole picture. A. Zelevinsky in his book [34] suggested an axiomatis version of theory which was based of the notion of the Hopf algbebras; his idea was important for our work first of all because he had shown the crucial role of the Hopf algrbra generated by the symmerics group as an "atom" of general type of those Hopf algerbas and secondly, because the definition of the group GLB (see below) was suggested together with him and S. Kerov. This group appears inevitably if we want to study right infinite analog of the representations theory of groups $GL(n, F_q)$ and the harmonic analysis on groups of infinite matrices. We will explain why the character theory on that group is more natural than on the group $GL(\infty, F_q)$.

2 Group GLB and algebra $\mathcal{A}$.

The right approach to this question is the following: we must consider not only *normalized characters* but also generalized characters which could take an infinite values on the elements of the group but must be finite on sufficiently large subalgebra of the group algebra. As in the theory of semisimple Lie groups (f.e. $SL(2, r)$) we need to consider *infinitesimal* characters which are infdinite on the group itself. We already had described such characters (locally finite) in the paper ([8]) for the group S_∞, but in that case even the set of finite characters was sufficiently rich. So first of all we must consider semifinite characters instead of finite ones. But for $GL(\infty, F_q)$ the situation is more complicated than for S_∞, because it is not clear how to choose a subalgebra of the group algebra such that the character must be finite on it.

The reasonable point of view came from classical representation theory of the group $GL(n,F_q)$, namely from so called parabolic phylosophy. The representation theory of $GL(n,F_q)$ was created mainly by J. Green and continued in the latest important papers by D. Faddeev, A. Zelevinsky mentioned above. The operation of induction in this theory used so called *parabolic imbedding* of the group $GL(n,F_q)$ (and consequently its group algebra) to the group algebra of the group $GL(n+1,F_q)$:

$$i_n : \mathbf{C}(Gl(n,F_q)) \to \mathbf{C}(GL(n+1,F_q)), \qquad n = 1,2,\dots$$

let $g \in GL(n,F_q)$ then i_n is defined by the following values on the group and by linear extension from the group to the group algebra $(GL(n,F_q)$:

$$i_n(g) = \frac{1}{q^n}\Sigma_{h\in F_q^n} g^h$$

(the summation is in the group algebra of the group $GL(n+1,F_q)$) where g^h is a matrix of order $n+1$

$$g^h = \begin{pmatrix} g & h \\ 0 & 1 \end{pmatrix}$$

with g an $n\times n$ block, $h \in F_q^n$ is the last column and 0 is zero row of length n. It is clear that i_n is injective homomorphism of algebras but it does not preserve the unity $(i_n(1_n) \neq 1_{n+1})$.

So we have nonunital algebra which is an inductive limit of algebras

$$\mathcal{A}_q = \mathcal{A} = \lim ind(\mathbf{C}(GL(n,F_q)), i_n) \tag{1}$$

(we omit q if there is no misunderstanding.)

Now we define the group $GLB_q = GLB$ of *infinite almost triangular matrix* over F_q — this is the group of all infinite martices $a = \{a_{i,j}\}$ which have zero entries $a_{i,j}$ for $i > j$ with sufficiently large $i = i(a)$

Theorem 1 (Group GLB and Algebra $\mathcal{A}$).

1. *The group GLB is locally compact with compact Borel subgroup B of upper triangular matrices, unimodular, amenable group. The unique two-sided invariant Haar measure m is normalized by 1 on the subgroup B.*
2. *The algebra $\mathcal{A}$ could be imbedded to the group algebra $L^1(GLB,m)$ as Bruhat–Schwartz subalgebra of cylindric functions in the natural sense. The algebra $\mathcal{A}$ is locally-semisimple $\mathbf{Z}$-graded nonunital algebra; inductive limit of the finite dimensional semisimple algebras.*

Proof. 1) Remark that the group $GL(\infty,F_q)$ is a countable subgroup of the group GLB_q. The Borel subgroup as a set is an infinite product of the finite sets, so it is compact in the weak topology, and the group GLB is a countable union over g of the compact subgroups of type $\{gBg^{-1}, g \in GL(\infty,F_q)\}$. We equip GLB with inductive topology of this union, now we have locally compact σ-compact nondiscrete topology on the group GLB and the subgroup $GL(\infty, F_q)$ is dense subgroup in this topology. Haar measure is defined as a normalized product-measure on the compact subgroup B and naturally extends to all group; it is two-sided invariant measure on B and consquently on GLB. Amenability follows from the fact that GLB is a union of compact subgroups. We can define a group algebra as a vector space $L^1(GLB, m)$ over $\mathbf{C}$ of all integrable functions by Haar measure m on the group GLB with convolution as multiplication and with usual involution.

2) For any element $g \in GL(\infty)$ define the set $\hat{g}$ of all matrices in GLB which has as NW-corner the matrix g. This is the compact subset in GLB, and its characteristic function a_g is the element of the space $L^1(GLB, m)$ (cylindric function). The algebra (under convolution) generated by all functions a_g over $g \in GL(\infty)$ is evidently isomorphic to $\mathcal{A}$ and we will call it the Bruhat–Schwartz subalgebra. It plays a role of smooth functions in the theory of Lie groups. □

It is interesting that the group GLB looks like an infinite dimensional group but has properties of locally compact groups. Thus, continuing of our methaphorical parallels we can say that the right q-deforamtion of the group S_∞ is the group $GLB(F_q)$, at least more natural than $GL(\infty, F_q)$.

Now our program is the following: we want to list semifinite characters of the group GLB (which are by definition the semifinite traces on the group algebra $L^1(GLB, m)$) and which are *finite on the dense subalgebra* $\mathcal{A}$ of group algebra. Remark that the group $GL(\infty, F_q)$ is the dense subgroup of GLB so we indeed had generalized the problem about the characters on that group in the right way.

We will see that the technique uses very essentially theory of representations of the finite groups $GL(n, F_q)$ and the answer is very interesting. It will be clear also the deep analogy with the previous theory of characters of infinite symmetric group.

We will assume that the readers know some general facts about the connections between representations of the locally compact groups and representations of its group algerbas of the various types (see for example the book by Hewitt and Ross [6]). Namely there are natural bijections between the continuous $*$-representatons of dense subalgebras of the group algebra of the locally compact group (in our case subalgebra $\mathcal{A}$ of the group algebra $L^1(GLB, m)$) and continuous unitary representations of the group. In particular, the representations of $\mathcal{A}$ with finite (infinite) indecomposable trace correspond (one-to-one) to the representations of type II_1(or II_∞) of the group with finite or infinitesimal character, in our case, continuous *-representation of the algebra $\mathcal{A}$ could be extended uniquely on the group GLB with the same weak closure of the image.

3 Structure of the algebra $\mathcal{A}$

3.1 First of all we describe the structure of algebra $\mathcal{A}$ as a locally semisimple algebra. We alreay had described it as an inductive limit of the algebras $\mathbf{C}(GL(n, F_q))$ (see (1)), and let us fix this structure further. Now we want to clear up the structure of ideals of this algebra in order to study its characters, representations and so on. More exactly, we want to describe the *branching rule* of the representations of the series of algebras $\mathbf{C}(GL(n, F_q))$ of that inductive family.

As well as each locally semisimple algebra the algebra $\mathcal{A}$ can be described in terms of $\mathbf{Z}_+$ graded locally finite (multi)graph which is called *the Bratteli diagram* (see [11]). The vertices of such (multi)graphs correspond to the simple modules of all finite dimensional semisimple algebras which include to the inductive family (in our case (1)). In particular, the vertices of the degree n correspond to simple modules of finite dimensional algebra of the level n (in our case $\mathbf{C}(GL(n, F_q))$). The edges between two vertices of the consequtive degree n and $n+1$, say π_n and π_{n+1} are defined according to a branching rule: the muliplicity of the edge between that pair of vertices is equal to the multiplicity of the simple module π_n (more carefully: corresponding to the vertex π_n) as a submodule of the simple module π_{n+1} considered as a module of algebra of the level n. If module π_n is not a submodule of the module π_{n+1}, then there is no edge. So the Bratteli diagram can be a multigraph if there occurs multiplicity more than one.

It means that we must repeat a classical task to decompose simple modules over algebra $\mathbf{C}(GL(n+1, F_q)$ as a sum of simple modules over subalgerba $\mathbf{C}(GL(n, F_q)$ which is *parabolically* imbedded to the previous. This will be a branching rule and definition of the required Bratteli diagram. It is important to emphasize that the answer hardly depends on the imbeddding: for exapmle if we imbed group $GL(n, F_q)$ into the group $GL(n+1, F_q)$ in ordinary way then we have a multiplicities in branching but in our case of *parabolic imbedding* there is not and as we will see we obtain a true graph as Brattlei diagram.

3.2 The general answer is given with the following theorem which we will be clarified later:

Theorem 2 (Decomposion of the algebra $\mathcal{A}$).
Algebra $\mathcal{A}$ is the sum of countably many of two-sided ideals each of which is isomorphic to the group algerba of the infinite symmetric group:

$$\mathcal{A} = \bigoplus_{i \in \mathcal{I}} A(i)$$

where i runs over parametrization set I. The Bratteli diagram of algebra $\mathcal{A}$ is the union of the countable many disconnected graphs each of which is isomorphic to the Young graph.

The parametrization set, graded structure and branching rule for algebra $\mathcal{A}$ as well as structure of Bratteli diagram is defined below.

3.3 Consider the lattice $(\mathbf{Z}_+)^2$ with ordinary partial order. Young diagram is a finite ideal of this lattice. Consider the partial ordered set of all finite ideals (including the empty ideal–empty Young diagram) which is ordered by inclusion; a number of elements define the grading of it. So we define the *Young graph* $\mathbf{Y}$ as $\mathbf{Z}_+$-graded graph, (which is the Hasse diagram of this partial ordered set). The empty Young diagram has degree 0, there are $p(n)$ (Euler function or the number of partitions of naturals) Young diagrams of the level n. We will denote as $|\lambda|$ the number of the cells (elements) of Young diagram λ. The classical fact is that the Young graph as a locally finite $\mathbf{Z}_+$-graded graph is the Bratteli diagram of the group algebras over $\mathbf{C}$ of the infinite symmetric group. (see for instance [11]). Below we will use also the other graded structure on it.

3.4 Let $C(q) \equiv C = \bigcup_{d=1}^{\infty} C_d$ be the set of all irreducible polynomials over the finite field F_q where C_d is a subset of the polynomials of degree d, particularly $C_1 = \{\mathbf{1}, \dots, q-1\}$ is a set of all irreducible polynomials of degree 1, or polynomials $t + a$ where $a \neq 0$. The element $\mathbf{1} \in C_1$ (polynomial $t+1$) will play a special role in our construction.

Let $\mathcal{R}$ be the set of all *finite* functions on C with values in the set of all Young diagrams $\mathbf{Y}$; finiteness means that images of only finite numbers of the elements of C are nonempty diagram. There is "empty function" or "empty vertex" $\emptyset$ which values at all pointrs is an empty diagram.

We can represent also the set $\mathcal{R}$ as infinite direct sum of the Young graphs numerated by the elements c of the set C:

$$\mathcal{R} = \oplus \mathbf{Y}_c,$$

where $\mathbf{Y}_c$ for all c are isomorphic to $\mathbf{Y}$. We will call the values of the function $\phi \in \mathcal{R}$ at the elements $c \in C$ as the coordinates of ϕ. The c-th coordinate of the element $\phi \in \mathcal{R}$ is the value $\phi(c)$. The graded structure in each $\mathbf{Y}_c$ depends on c accordingly to the definition of graduation.

Define the graduation $\|\cdot\|$ on the set $\mathcal{R}$ as follows: $\|\phi\| \equiv \Sigma_{c \in C} |\phi(p)| d(c)$ where c a polynomial of degree $d(c)$, and $|\lambda|$ is number of the cells in the diagram $|\lambda|$. Denote by $\mathcal{R}_n$ the set of functions from $\mathcal{R}$ of degree n.

3.5 Let $\widehat{GL(n, F_q)}$ be the set of all simple modules over $\mathbf{C}$ (that is irreducible complex representations) of all algerbas $\mathbf{C}(GL(n, F_q))$ graded with the rank of the group n.

Lemma 1 (The list of the vertices). *There is a one-to-one canonical correpondense between the graded set $\hat{G}L(n, F_q)$ and the graded set $\mathcal{R}_n$.*

This is a classical fact, remember that the same natural parametrizations has the set of conjugacy classes of the group $GL(n, F_q)$ (see [5]).

Example. The function $\phi_{\mathbf{1}}$ which has one-cell diagram value at $\mathbf{1}$ and empty diagrams as all other values corresponds to the identity representation of

$GL(1, F_q)$. The function which corresponds to the characters of $GL(1, F_q)$ (that is to one-dimensional representations) has one-cell diagram value at one element of $C_1 = \hat{F}^*$ and empty diagram at another $c \in C$.

Corollary. *The set of vertices of the Bratteli diagrams of the algebra $\mathcal{A}$ is the set $\mathcal{R}$. The n-th level of the diagram is the set $\mathcal{R}_n$, $n = 0, \ldots$.*

Remark. The set $\mathcal{R}$ is the union of all irreducible representations of the groups $GL(n, F_q)$, $n = 1, \ldots$; we got all of them toghether for further purpose – to define graph structure. Upto now we did not speak about representations of the group GLB or algebra $\mathcal{A}$. Nevertheless we already can interprete empty functions as the representations: they correspond to zero representations of algebra $\mathcal{A}$ or to identity representation of the group GLB. (Recall that $\mathcal{A}$ is nonunital algebra.)

Sometimes we will call the functions from $\mathcal{R}$ as the verticies, and the values of functions as the coordinates of the vertices.

The structure of the locally finite graph on the set $\mathcal{R}$ is defined by the following partial order:

Lemma 2 (Definition of the edges and branching rule). *The function $\phi_1 \in \mathcal{R}$ are less that function $\phi_2 \in \mathcal{R}$ if the values at all points $c \neq \mathbf{1}$ are the same and diagram $\phi_1(\mathbf{1})$ is a subdiagram of the diagram $\phi_2(\mathbf{1})$. In another words, the edge between two vertices (=functions) exists iff all values exept the value at $\mathbf{1}$ coinside and diagrams at $\mathbf{1}$ differ on one cell. This partial order and the structure of the graph corresponds to the branching rule of the representations of $Gl(n, F_q)$, $n = 1, \ldots$*

So we have defined the $\mathbf{Z}_+$-graded locally finite graph $\mathcal{R}$ which is Bratteli diagram of algebra $\mathcal{A}$.

Lemma 2 is equivalent to the classical fact about branching of the representations of $GL(n, F_q)$, and we repeat once more that the crucial role in the definition of branching rule plays parabolic imbedding. We should have a different answer for branching rule if we use instead of parabolic an ordinary imbeddins of the groups. [1]

Corollary. *Bratteli diagram is disconnected sum of the subgraphs each of which is isomormorphic to Young graph, the grading of the components are various accordingly to the graded structure of graph. Each component corresponds to the minimal two-sided ideal of algerba $\mathcal{A}$.*

[1] For example, an initial vertex corresponds to the module of the algebra $\mathbf{C}(GL(n+1, F_q))$ which has no nonzero submodules over subalgebra $\mathbf{C}(GL(n, F_q))$. (Recall that zero representation of nonunital algebra corresponds to identity representation of the group.) If instead of the parabolic imbedding of the last algebra to the previous one we use ordinary imbedding $\mathbf{C}(GL(n, F_q))$ to the group $\mathbf{C}(GL(n+1, F_q))$ the branching rule will be completely different.

Remark. The two-sided ideal corresponded to the vertex $\mathbf{1}$ is *the Hecke subalgebra* of the algebra $\mathcal{A}$, namely the subalgebra of the functions on the group GLB which are constant on the left and right classes over Borel subgroup. (See “principle series” below).

3.6 Denote by $\mathcal{I} \in \mathcal{R}$ the set of all vertices which has nonempty preceding vertices in the sense of order defined above together with empty vertex $\emptyset$ (function with identically empty diagram as values). Call these vertices *initial vertices.* Each initial vertex is the origin of one of the components of our graph. It is clear that we have the following decompostion:

$$\mathcal{R} = \mathbf{Y_1} \oplus \bigoplus_{c \in C \smallsetminus \{\mathbf{1}\}} \mathbf{Y}_c \equiv \mathbf{Y_1} \oplus \mathcal{I}$$

Indeed, each vertex could be written as a pair of initial vertex and some vertex from Young graph $\mathbf{Y_1}$. Now we can say that the graph $\mathcal{R}$ is “left-side direct product” of the Young graph Y and set $\mathcal{I}$.[2] So the parametric set in the formulation of Theorem 2, which parametrizes the ideals of the algebra $\mathcal{A}$ can be identified with the set $\mathcal{I}$. So each initial vertex is an initial vertex for some Young subgraph in $\mathcal{R}$.

3.7 Using the structure of the set of initial vertices we can give the further classificaion of the initial vertices and representations of the groups $GL(n, F_q)$.

Let us call a vertex (function) from $\mathcal{I}$ a *caspidal* if there is only one nonemtpy diagram as the values of it at some point $c \in C$ and this diagram has one cell; denote this element as ϕ_c. It is clear that this function can be identified with the element $c \in C$ in which the value is nonempty diagram. We will preserve denotation $\mathcal{C}$ for the set of all caspidal vertices (functions). For example $\phi_{\mathbf{1}}$ is the caspidal vertex. The degree of caspidal vertex at point c is $\deg(\phi_c) = \deg(c)$ where the degree at the right side is degree of c as of the irreducible polynomial. So for given n the set of caspidal vertices of degree n is the same as the finite set of the irreducible polynomials of degree n.

The *primary vertex* of $\mathcal{R}$ is a vertex ϕ for which there exists only one element $c \in C$ whose image is nonempty diagram; the denotation is $\mathcal{P}$. Each primary vertex is associated with some caspidal vertex which has one-cell diagram at the same c. The primary vertex which belongs to $\hat{G}L(n, F_q)$ (or has dergee n) can be associated with element c if $d(c) \mid n$, and is represented by the Young diagram with $n/d(c)$ cells or, in another words, degree of the primary vertex which is the Young diagram with m cells (as unigue nonempty coordinate) is equal to $m \cdot d(c)$. As was mentioned the set of primary vertices

[2] Let Γ be a graph and M be a set, left-sided direct product is a graph which has the direct product $\Gamma \times M$ as the set of vertices and the set of all pairs of type $((\gamma, m), (\gamma_1, m_1))$ where $m = m_1$ and the pair $(\gamma\gamma_1)$ is an edge of graph Γ as the set of edges. I do not know if such definition is popular in the graph theory.

associated with given caspidal vertex is a set of vertices of the Young graph which starts from that caspidal vertex (we denoted its before as $\mathbf{Y}_c$).

The *principle vertex* is the element of $\mathcal{R}$ whose values is nonempty diagram only at $\mathbf{1}$, denote the set of princile vertices by $\mathcal{M}$. Each principle vertex by definition is primary one. They are vertices of the Young graph $Y_{\mathbf{1}}$ which is *true* subgraph of our graph $\mathcal{R}$.

We have inclusions:

$$\mathcal{C} \subset \mathcal{P} \subset \mathcal{I} \subset \mathcal{R}, \mathcal{M} \subset \mathcal{P}.$$

The terminology above coinsides with classical one: vertices (functions or elements of $\mathcal{R}$) which we have called caspidal (primary, principle), corresponds to the caspidal, (correspondingly to primary, of principle series) representations of $Gl(n, F_q)$. We will see that detale analysis of the representations is reduced in precise sense to the studying of the principle series.

Remark. As we saw above, the graph $\mathcal{R}$ is the union of the Young subgraphs of the graph $\mathcal{R}$ and each of them starts at some initial vertex. From the other side the set of primary vertices associated with a given caspidal vertex can be also considered as a set of vertices of the union of Young diagrams (as functions on C with values in the Young graph), we denoted it as $\mathbf{Y}_c$. But the edges of **those** Young graphs are **not** the edges of the graph $\mathcal{R}$! Let us call the set $\mathbf{Y}_c$ of primary vertices associated with caspidal vertex c as a *virtual Young subgraph*; it is important not to confuse $\mathbf{Y}_c$ with the true Young subgraphs of $\mathcal{R}$! Nevertheless we will use this virtual "Young structure" of the set vertices.

4 Parabolic product, and *K*-functor

4.1 Consider the free Abelian group over the set of vertices of Young graph $\mathbf{Z}(\mathbf{Y})$; this is the group of integer linear combinations of the vertices. As it was proved in [10] the Grothendick group K_0 of the infinite symmetric group, (or more exactly Grothendick group of its group algebra $K_0(\mathbf{C}(S_\infty)))$ is a quotient of $\mathbf{Z}(\mathbf{Y})$ over a suitable subgroup (see below). Moreover it was proved in [10] that there is a canonical structure of commutative ring in this group which is defined by means of natural multiplication; the multiplication came from the operation of the induction of the representations: $ind_{S_n \times S_m}^{S_{n+m}}(\pi_n \bigotimes \pi_m)$ where π_n, π_m are the representations of the groups S_n and S_m correspondingly. This multiplication make sense only when we consider all symmetric groups together. The resulting operation can be described directly: the space $\mathbf{Z}(\mathbf{Y})$ as a ring is the ring Λ of all symmetric functions of infinite number of variables, and K_0 is a quotient of $\mathbf{Z}(\mathbf{Y}) = \Lambda$ over ideal generated by function $s_1 - \mathbf{1}$ where s_1 is the first symmetric function (sum of all varibales). This factorization of the space of symmetric functions corresponds to the identification of the group S_n as a subgroup of S_{n+1} for all, and consequently identification of the identity representations of each group S_n; all of them are the same element of $K_0(\mathbf{C}(S_\infty$ – the initial vertex of the graph.

Recall that *Schur functions* (s_λ) correspond to the character of the representation of S_n with diagram λ) (see [13]). In terms of Schur basis of the space of symmetric functions the branching rule of the representations of symmetric group and the structure of Young graph has the following explanation: let λ the diagram of the level n then for Schur function we have

$$s_\lambda \cdot s_1 = \sum s_\Lambda$$

where Λ runs over all Young diagrams which follow to diagram λ on the $(n+1)$-th level. This gives a branching rule in the terms of multiplication formula.

4.2 Let $\mathbf{Z}(\mathcal{R})$ be the set of all fromal $\mathbf{Z}$-linear combinations of the vertices of graph $\mathcal{R}$. Now we can define a multiplication in this abelian group. Because $\mathcal{R}$ is the space of functions on C with values in set of Young diagrams we can define a *pointvise multiplication* on $\mathbf{Z}(\mathcal{R})$ using multiplication in the space $\mathbf{Z}(\mathbf{Y})$ mentioned above. We will call this multiplication as *parabolic multiplication* in $\mathbf{Z}(\mathcal{R})$. By definition it is commutative associative multiplication. From the definition we have the following facts:

Theorem 3 (Multiplicative structure of the $\mathbf{Z}(\mathcal{R})$).

1. Additive subgroup of the ring $\mathbf{Z}(\mathcal{R})$ generated by vertices of virtual Young subgraph (or generated by primary vertices associated with a caspidal vertex c) is a subring of $\mathbf{Z}(\mathcal{R})$. We call them caspidal subrings $R_c = \mathbf{Z}\mathbf{Y}_c)$. The subring $R_{\mathbf{1}} = \mathbf{Z}\mathbf{Y}_{bf1})$ which is generated by the set of principle vertices we will call principle subring.

2. As the ring $\mathbf{Z}(\mathcal{R})$ is the tensor product over $\mathbf{Z}$ of the caspidal subrings

$$\mathbf{Z}(\mathcal{R}) = \bigotimes_{c \in C} R_c$$

So each element of $\mathbf{Z}(\mathcal{R})$ has unique expression as a linear combination over $\mathbf{Z}$ of the parabolic product of the primary vertices.

3. Each additive subgroup G_ϕ of the ring $\mathbf{Z}(\mathcal{R})$ generated by the true Young subgraphs with initial vertex ϕ is the module over principle subring:

$$G_\phi = R_{\mathbf{1}} \cdot \phi$$

(parabolic product of the princile subring and initial element).

4.3 The following link with the representation theory of linear groups over finite field is not trivial and moreover plays the central role in the whole theory:

Theorem 4 (Parabolic multiplication). *Under identification of the vertices of graph $\mathcal{R}$ with irreducible representations of the groups $GL(n, F_q), n = 1, 2, \ldots$ the parabolic multiplication in the sense of the representation theory coinsides with parabolic multiplication defined above.*

Recall that parabolic induction (multiplication) of the irredicible representation π_1 of the group $GL(n, F_q)$ and the irreducible representation π_2 of the group $GL(m, F_q)$ defined as follows. This is the representation (in general reducible) of the group of $GL(n+m, F_q)$ which is induced representation from tensor product of $\pi_1 \bigotimes \pi_2$ of the natural parabolic subgroup of $GL(n+m, F_q)$ with two diagonal blocks n and m onto whole group $GL(n+m, F_q)$. Even the commutavity and associativity of this multiplication is not evident. The description and the properties of that parabolic induction could be found in the papers [5, 2, 34].

Corollary. *Grothendic group K_0 of the algebra $\mathcal{A}$ coinsides with additive group K_0 of the ring $\mathbf{Z}(\mathcal{R})$ and consequently has the ring structure. From the item 3 of the previous theorem it follows that K_0 is the multiplicative K-functor in the sense of [10]. We have*

$$K_0(\mathcal{A}) = \mathbf{Z}(\mathcal{R}).$$

4.4 Recall that the character (trace) on the *-algebra L is nonnegative (in the sense of star-structure) linear functional: $\chi : L \to \mathbf{C}$ which is central: $\chi(ab) = \chi(ba)$. Because our algerba is nonunital we have no normalization of the characters. The characters generated the convex cone in the dual space to algebra and the points of the extremal rays of the cone called *indecomposable character*. Each character $\mathcal{A}$ defines the additive functional on the K_0-functor of $\mathcal{A}$ or on the $\mathbf{Z}(\mathcal{R})$. Because each linear functional is the some of its restrictions on the ideals the studying of the characters on $\mathcal{A}$ reduces to the study of characters on each ideal. In terms of K_0 it means that we can consider the character on each module over principle subring of the ring $\mathbf{Z}(\mathcal{R})$. But because of the item 3 of Theorem3 above each character on such module is the shift of the character on the principle subring, in another words if ϕ is initial vertex and *chi* is a character on $R_\mathbf{1}$ then $\chi(\phi \cdot \dot{)}$ is the character on the module generated by ϕ.

Now we can claim the main thesis of the asymptotic theory of the characters of the group GLB and more specifically of algerba $\mathcal{A}$: all the problems reduced to the studying of the characters and representatins of the *principle series* and to description of the caspidal representations of $GL(n, F_q)$. This means that first of all we must find the characters of the Hecke subalgebra of algebra $\mathcal{A}$. On the first glance it means that we completely reduce the problem to the case of infinite symmetric group because the infinite Hecke algebra is isomorphic to the group algebra of symmetric group. But this is not precisely true because that subalgebra is not unital and we must deform the theory of characters of infinite symmetric group to the case where we add to Hecke algebra a unity in a special way. This is the subject of the future lectures.

References

1. Borodin, A.: Multiplicative central measures on the Schur graph. In: Vershik, A. M. (ed) Representation theory, dynamical systems, combinatorial and algorithmic methods II. Zapiski Nauchnykh Seminarov POMI, **240**, 44–52 (1997) (Russian); English translation: J. Math. Sci., **96**, no. 5, 3472–3477 (1999)
2. Faddeev, D. K.: Complex representations of general linear group over finite field. Zap. Nauchn. Sem. LOMI, **46**, 64–88 (1974)
3. Fulman, J.: Random Matrix Theory over Finite Fields. Bull. AMS, **39**, 51–85 (2002)
4. Goncharov, V. L.: One combinatorial problem. Izv. AN SSSR, **8**, no. 1, 3–48 (1944)
5. Green, J. A.: The characters of the finite general linear groups. Trans. Amer. Math. Soc., **80**, 402–447 (1955)
6. Hewitt, E., Ross, K.: Abstract Harmonic Analysis, Vol. 2. Springer–Verlag, Berlin Heidelberg New York (1970)
7. Kerov, S., Vershik, A.: Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young Tableaux. Sov. Dokl., **233**, no. 6, 1024–1027 (1977)
8. Kerov, S., Vershik, A.: Characters and factor representations of infinite symmetric group. Dokl. AN SSSR, **257**, 1037–1040 (1981)
9. Kerov, S., Vershik, A.: Asymptotics of maximal and typical dimensions of irreducible representations of a symmetric group (with S.V.Kerov). Funkts. Anal. Prilozh., **19**, no. 1, 25–36 (1985). English translation: Funct. Anal. Appl. 19, 21–31 (1985).
10. Kerov, S., Vershik, A.: Locally semisimple algebras, combinatorial theory and K-functor Current problems in Mathematics. Newest results. Itogi Nauki i Tehninki. VINITI **26**, 3–56 (1985) English translation: Journ. of Sov. Math., **38**, 1701–1733 (1987)
11. Kerov, S., Vershik, A.: On the group of infinite matricaes over finite field. Funct. Anal., **32**, no. 3 (1998)
12. Logan, B. F., Shepp, L. A.: A variational problem for random Young tableaux. Adv. Math., **26**, 206–222 (1997)
13. Macdonald, I. G.: Symmetric functions and Hall polynomials, 2nd edition. Oxford University Press (1995)
14. von Neumann, J.: Approximative Properties of Matrices of High Finite Order. Portugaliae Math., **3**, 1–62 (1942)
15. Okounkov, A.: Thoma's theorem and representations of infinite bisymmetric group. Funktsion. Anal. Prilozhen., **28**, no. 2, 31–40 (1994) (Russian); English translation: Funct. Anal. Appl., **28**, no. 2, 101–107 (1994)
16. Okounkov, A., Vershik, A.: A new approach to representation theory of symmetric groups. Selecta Math., **2**, no. 4, 581–605, (1996)
17. Olshanski, G.: Unitary representations of (G, K)-pairs connected with the infinite symmetric group $S(\infty)$. Algebra i Analiz, **1**, no. 4, 178–209 (1989) (Russian); English translation: Leningrad Math. J. **1**, 983–1014 (1990)
18. Pushkarev, I.: On the representation theory of wreath products of finite group and symmetric group. In: Vershik, A. M. (ed) Representation theory, dynamical systems, combinatorial and algorithmic methods II. Zapiski Nauchnykh Seminarov POMI, **240**, 44–52 (1997) (Russian); English translation: J. Math. Sci., **96**, no. 5 (1999)

19. Schmidt, A., Vershik, A.: Symmetric groups of higher degree. Sov. Dokl., **206** no. 2, 269–272 (1972)
20. Schmidt, A., Vershik, A.: Limit measures arising in the asymptotic theory of symmetric groups. I, II. Teor. Verojatn. i Prim., **22**, no. 1, 72–88 (1977); **23**, no. 1, 42–54 (1978) (Russian); English translation: Theory of Probab. Appl., **22**, 70–85 (1977); **23**, 36–49 (1978)
21. Skudlarek, H. L. Die unzerlegbaren Charaktere einiger diskreter Gruppen. Math. Ann., **223**, 213–231 (1976)
22. Thoma, E.: Die unzerlegbaren, positiv-difiniten Klassenfunktionen der abzählbar unendlichen, symmetrischen Gruppe. Math. Z., **85**, no. 1, 40–61 (1964)
23. Thoma, E.: Die Einschränkung der Charactere von $GL(n,q)$ auf $GL(n-1,q)$. Math. Z., **119**, 321–338 (1971)
24. Tsilevich N. V. Distributions of mean values for some random measures. In: Vershik, A. M. (ed) Representation theory, dynamical systems, combinatorial and algorithmic methods II. Zapiski Nauchnykh Seminarov POMI, **240**, 268–279 (1997) (Russian); English translation: J. Math. Sci., **96**, no. 5, (1999)
25. Tsilevich, N., Vershik, A., Yor, M.: An infinite-dimensional analogue of the Lebesgue measure and distinguished properties of the gamma process. J. Funct. Anal., **185**, no. 1, 274–296 (2001).
26. Vershik, A.: Asymptotical distribution of decompositions of natural numbers on prime divisors. Dokl. Acad.Nauk SSSR, **289**, no. 2, 269–272 (1986)
27. Vershik, A.: Local algebras and a new version of Young's orthogonal form. In: Topics in Algebra, part 2: Commutative Rings and Algebraic Groups (Warsaw 1988), Banach Cent. Publ., **26**, 467-473 (1990).
28. Vershik, A. M.: Statistical mechanics of the combinatorial partitions and their limit configurations. Funct. Anal. i Pril., **30**, no. 2, 19–30 (1996)
29. Vershik, A. M.: Limit distribution of the energy of a quantum ideal gas from the point of view of the theory of partitions of natural numbers. Uspekhi Mat. Nauk **52**, no. 2, 139–146 (1997) (Russian); English translation: Russian Math. Surveys, **52**, no. 2, 379–386 (1997)
30. Vershik, A.: Asymptotic aspects of the representation theory of symmetric groups. Selecta Math. Sov., **11**, no. 2, 159–180 (1992)
31. Vershik, A. (ed): Representation Theory, Dynamical Systems, Combinatorial and Algorithmic Methods. Part 6. Zapiski Nauchn. Sem. POMI, **283** (2001)
32. Vershik, A., Yakubovich, Yu.: Limit shape and fluctuations of random partitions of naturals with fixed number of summands. Moscow Math. J., **3** (2001)
33. Weyl, H.: Phylosophy of mathematical and natural sceince (1949)
34. Zelevinsky, A. V.: Representations of Finite Classical Groups. Lecture Notes in Math., **869**, 1–184 (1981)

Characters of symmetric groups and free cumulants

Philippe Biane

CNRS
Département de Mathématiques et Applications
École Normale Supérieure
45, rue d'Ulm 75005 Paris
FRANCE
Philippe.Biane@ens.fr

Summary. We investigate Kerov's formula expressing the normalized irreducible characters of symmetric groups evaluated on a cycle, in terms of the free cumulants of the associated Young diagrams.

1 Introduction

Let μ be a probability measure on $\mathbb{R}$, with compact support. Its Cauchy transform has the expansion

$$G_\mu(z) = \int_{\mathbb{R}} \frac{1}{z-x}\mu(dx) = z^{-1} + \sum_{k=1}^{\infty} M_k\, z^{-k-1} \tag{1}$$

where the M_k are the moments of the measure μ. This Laurent series has an inverse for composition $K(z)$, with an expansion

$$K_\mu(z) = z^{-1} + \sum_{k=1}^{\infty} R_k z^{k-1}. \tag{2}$$

The R_k are called the free cumulants of μ and can be expressed as polynomials in terms of the moments. Free cumulants show up in the asymptotic behaviour of characters of large symmetric groups. More precisely, let λ be a Young diagram, to which we associate a piecewise affine function $\omega : \mathbb{R} \to \mathbb{R}$, with slopes ± 1, such that $\omega(x) = |x|$ for $|x|$ large enough, as in Fig. 1 below, which corresponds to the partition $8 = 4+3+1$. Alternatively we can encode the Young diagram using the local minima and local maxima of the function ω, denoted by $x_1, \ldots, x_m$ and $y_1, \ldots, y_{m-1}$ respectively, which form two interlacing sequences of integers. These are (-3,-1,2,4) and (-2,1,3) respectively in

A. Borodin et al.: LNM 1817, A.M. Vershik (Ed.), pp. 185–200, 2003.

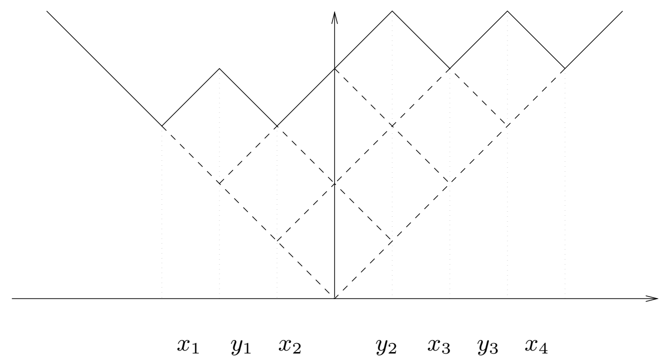


Fig. 1.

the picture. Associated with the Young diagram there is a unique probability measure μ_ω on the real line, such that

$$\int_{\mathbb{R}} \frac{1}{z-x}\mu_\omega(dx) = \frac{\prod_{i=1}^{m-1}(z-y_i)}{\prod_{i=1}^{m}(z-x_i)} \quad \text{for all } z \in \mathbb{C} \setminus \mathbb{R}. \tag{3}$$

This probability measure is supported by the set $\{x_1, \dots, x_k\}$ and is called the transition measure of the diagram, see [K1]. We shall denote by $R_j(\omega)$ its free cumulants. Let $\sigma \in S_n$ be a permutation with k_2 cycles of length 2, k_3 of length 3, etc. We shall keep $k_2, k_3, ...$ fixed, and denote $r = \sum_{j=2}^{\infty} jk_j$, while we let $n \to \infty$. The normalized character χ_ω associated to a Young diagram with n cells has the following asymptotic evaluation from [B]

$$\chi_\omega(\sigma) = \prod_{j=2}^{\infty} R_{j+1}^{k_j}(\omega) n^{-r} + O(n^{-\frac{r+1}{2}}). \tag{4}$$

Here the O term is uniform over all Young diagrams whose numbers of rows and columns are $\leq A\sqrt{n}$ for some constant A, and all permutations with $r \leq r_0$ for some r_0.

As remarked by S. Kerov [K2], free cumulants can be used to get universal, exact formulas for character values. More precisely consider the following quantities

$$\Sigma_k(\omega) = n(n-1)\dots(n-k+1)\chi_\omega(c_k)$$

for $k \geq 1$ where c_k is a cycle of order k (with $c_1 = e$).

Theorem 1 (Kerov's formula for characters). *There exist universal polynomials $K_1, K_2, \dots, K_m, \dots$, with integer coefficients, such that the following identities hold for any n and any Young diagram ω with n cells*

$$\Sigma_k(\omega) = K_k(R_2(\omega), R_3(\omega), \dots, R_{k+1}(\omega)).$$

We list the few first such polynomials

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

The coefficients of Kerov's polynomials seem to have some interesting combinatorial significance, although the situation is far from being understood. In this paper we shall give a proof of the above theorem, compute some of the coefficients in the formula, as well as give some insight into this problem.

This paper is organized as follows. In Section 2 we gather some information on free cumulants, Boolean cumulants and their combinatorial significance. In Section 3 we introduce some elements in the center of the symmetric group algebra. These are used in Section 4 to give a combinatorial proof of Theorem 1. In Section 5 we give another proof, based on a formula of Frobenius, which yields a computationally efficient formula for computing Kerov's polynomials. In Section 6 we compute the coefficients of the linear terms of Kerov's polynomials, as well as some coefficients of degree 2. We make some remarks in Section 7 on the possible combinatorial significance of the coefficients of Kerov's polynomials. This involves in a natural way the Cayley graph of the symmetric group. Finally in Section 8 we list the values of Kerov polynomials up to Σ_{11}.

I would like to thank A. Okounkov and R. Stanley for useful communication, as well as G. Olshanski for providing me a copy of [IO].

2 Noncrossing partitions, moments and free cumulants

From the relation between moments and cumulants given by

$$G_\omega = K_\omega^{\langle -1\rangle} \tag{5}$$

we obtain by Lagrange inversion formula that

$$R_k = -\frac{1}{k-1}[z^{-1}]\, G_\omega(z)^{-k+1} \tag{6}$$

(where $[z^{-1}]\, L(z)$ denotes the coefficient of z^{-1} in the expansion of a Laurent series $L(z)$). From this we get that the coefficient of $M_1^{l_1}\dots M_r^{l_r}$ in R_k is equal to

$$(-1)^{1+l_1+\dots+l_r}\frac{(k-2+\sum_i l_i)!}{l_1!\dots l_r!(k-1)!}. \tag{7}$$

if $k=\sum_j jl_j$, and to 0 if not.

Conversely one has

$$M_k = \frac{1}{k+1}[z^{-1}]\, K(z)^{k+1}$$

and the coefficient of $R_1^{l_1} \dots R_r^{l_r}$ in M_k, with $k = \sum_i il_i$ is equal to

$$\frac{k!}{l_1! \dots l_r!(k+1-\sum_i l_i)!} \tag{8}$$

It will be also interesting to introduce the series

$$H_\omega(z) = 1/G_\omega(z) = z - \sum_{k=1}^{\infty} B_k z^{1-k}$$

The coefficients B_k in this formula are called Boolean cumulants [SW] and the coefficient for $B_1^{l_1} \dots B_r^{l_r}$ in M_k, with $\sum_j jl_j = k$ is the multinomial coefficient

$$\frac{(l_1 + l_2 + \ldots + l_r)!}{l_1! l_2! \ldots l_r!} \tag{9}$$

A combinatorial interpretation of these formulas is afforded by R. Speicher's work [Sp] which we recall now. A noncrossing partition of $\{1, \dots, k\}$ is a partition such that there are no a, b, c, d with $a < b < c < d$, a and c belong to some block of the partition and c, d belong to some other block. The noncrossing partitions form a ranked lattice which will be denoted by $NC(k)$. We shall use the order opposite to the refinement order so that the rank of a non-crossing partition is $k+1-$(number of parts). The relation between moments and free cumulants now reads

$$M_k = \sum_{\pi \in NC(k)} R[\pi] \tag{10}$$

where, for a noncrossing partition $\pi = (\pi_1, \pi_2, \dots \pi_r)$, one has $R[\pi] = \prod_i R_{|\pi_i|}$, where $|\pi_i|$ is the number of elements of the part π_i. It follows from (9) that the coefficient of $R_1^{l_1} \dots R_r^{l_r}$ in the expression of M_k (with $k = \sum il_i$) is equal to the number of non-crossing partitions in $NC(k)$ with l_i parts of i elements, and is given by (8).

A parallel development can be made for the connection between Boolean cumulants and moments. A partition of $\{1, \dots, k\}$ is called an interval partition if its parts are intervals. The interval partitions form a lattice $B(k)$, which is isomorphic to the lattice of all subsets of $\{1, \dots, k-1\}$ (assign to an interval partition the complement in $\{1, \dots, k\}$ of the set of largest elements in the parts of the partition). The formula for expressing the M_k in terms of the B_k is

$$M_k = \sum_{\pi \in B(k)} B[\pi] \tag{11}$$

3 On central elements in the group algebra of the symmetric group

Let λ be a Young diagram, and for $n \geq |\lambda|$ let ϕ be a one to one map from the cells of λ to the set $\{1, \ldots, n\}$. Consider the associated permutation σ_ϕ whose cycles are given by the rows of the map ϕ. For example the following map with $n \geq 24$, gives the permutation with cycle decomposition

$$(4\,7\,11\,2\,19\,12)(20\,21\,1\,14\,24\,16)(5\,3)(22\,13)(2\,9)$$

4	7	11	2	19	12
20	21	1	14	24	16
5	3				
22	13				
2	9				
10					
8					

Fig. 2.

If Φ_λ is the set of such maps defined on λ, we shall call $a_{\lambda;n}$ the element in the group algebra of S_n given by

$$a_{\lambda;n} = \sum_{\phi \in \Phi_\lambda} \sigma_\phi$$

see [KO]. If λ has one row, of length l, we call $a_{l;n}$ the corresponding element. Note that $a_{1;n} = n.e$.

Lemma 1. *There exists universal polynomials P_λ with integer coefficients such that, for all n, one has*

$$a_{\lambda;n} = P_\lambda(a_{1;n}, \ldots, a_{|\lambda|;n})$$

and $\sum_j j \deg_{P_\lambda}(a_{j;n}) \leq |\lambda|$.

The proof is by induction on the number of cells of λ. This is clear by definition if λ has one row. If λ has more than one row then let λ' be λ with the last row deleted, and let k be the length of this row. One has

$$a_{\lambda';n} a_{k;n} = \sum_{\phi_1, \phi_2} \sigma_{\phi_1} \sigma_{\phi_2}$$

where ϕ_2 is a map on the diagram λ' and ϕ_2 a map on the diagram (k) with one row of length k. For any pair (ϕ_1, ϕ_2) there is a unique pair A, B where A is a subset of cells of λ', B is a subset of cells of (k), and a bijection τ from A to B which tells on which cells the two maps ϕ_1 and ϕ_2 coincide. The cycle structure of $\sigma_{\phi_1}\sigma_{\phi_2}$ depends only on λ', k, A, B, τ, and not on the values taken by the maps ϕ_1, ϕ_2. Let $\Lambda_{\lambda',k,A,B,\tau}$ be the diagram with $|\lambda| - |A|$ boxes (putting some one-box rows if necessary) of this conjugacy class. For each (A, B, τ) take some corresponding (ϕ_1, ϕ_2), and take a map on the diagram $\Lambda_{\lambda',k,A,B,\tau}$, which realizes the permutation $\sigma_{\phi_1}\sigma_{\phi_2}$. If necessary put the fixed points in the one-box rows. Now extend this to all pairs of maps (ϕ_1, ϕ_2) covariantly with respect to the action of S_n. Each map on $\Lambda_{\lambda',k,A,B,\tau}$ is obtained exactly once, and if $A = B = \emptyset$ then clearly $\Lambda_{\lambda',k,A,B,\tau} = \lambda$. It follows that

$$a_{\lambda';n}a_{k;n} = a_{\lambda;n} + \sum_{A,B,\tau,|A|\geq 1} a_{\Lambda_{\lambda',A,B,\tau};n} \tag{12}$$

For all terms in the sum one has $|\Lambda_{\lambda',A,B,\tau}| < |\lambda|$. The proof follows by induction. The condition on degrees is checked also by induction. $\square$

4 Jucys–Murphy elements and Kerov's formula

Consider the symmetric group S_n acting on $\{1, 2, \ldots, n\}$ and let $*$ be a new symbol. We imbed S_n into S_{n+1} acting on $\{1, 2, \ldots, n\} \cup \{*\}$. In the group algebra $\mathbb{C}(S_{n+1})$, consider the Jucys–Murphy element

$$J_n = (1\,*) + (2\,*) + \ldots + (n\,*)$$

where $(i\,j)$ denotes the transposition exchanging i and j. Let E_n denote the orthogonal projection from $\mathbb{C}(S_{n+1})$ onto $\mathbb{C}(S_n)$, i.e. $E_n(\sigma) = \sigma$ if $\sigma \in S_n$, and $E_n(\sigma) = 0$ if not. If we endow $\mathbb{C}(S_{n+1})$ with its canonical trace $\tau(\sigma) = \delta_{e\sigma}$ (i.e. τ is the linear extension of the normalized character of the regular representation), then E_n is the conditional expectation onto $\mathbb{C}(S_n)$, with respect to τ. We define the moments of the Jucys–Murpy elements by

$$\mathcal{M}_k = E_n(J_n^k) \tag{13}$$

To this sequence of moments we can associate a sequence of free cumulants through the construction of section 2. We call $\mathcal{R}_k$ these free cumulants. By construction, the $\mathcal{M}_k$ and $\mathcal{R}_k$ belong to the center of the group algebra $\mathbb{C}(S_n)$ (even to $\mathbb{Z}(S_n)$). The relevance of these moments and cumulants is the following

Lemma 2. *For any n and any Young diagram ω with n boxes, one has*

$$\chi_\omega(\mathcal{R}_k) = R_k(\omega)$$

Since χ_ω is an irreducible character, it is multiplicative on the center of the symmetric group algebra and therefore it is enough to check that $\chi_\omega(\mathcal{M}_k) = M_k(\omega)$. Let χ^*_ω be the induced character on S_{n+1}, then $\chi_\omega(\mathcal{M}_k) = \chi^*_\omega(J^k_n)$ and the result follows from the computation of eigenvalues of Jucys–Murphy elements. See e.g. [B], Section 3. □

One has

$$J^k_n = \sum_{i_1,\ldots,i_k \in \{1,\ldots,n\}} (*\, i_1)\ldots(*\, i_k) \tag{14}$$

A term in this sum gives a non trivial contribution to $\mathcal{M}_k$ if and only if the permutation $\sigma = (*\, i_1)\ldots(*\, i_k)$ fixes $*$. In order to see when this happens we have to follow the images of $*$ by the successive partial products of transpositions. Let $j_1 = \sup\{l < k \,|\, i_l = i_k\}$. If this set is empty then $\sigma(*) = i_k \neq *$. If not then one has $\sigma = (*i_1)\ldots(*i_{j-1})\sigma'$ where $\sigma'(*) = *$, hence we can continue and look for $j_2 = \sup\{l < j_1 - 1 \,|\, i_l = i_{j_1 - 1}\}$. In this way we construct a sequence $j_1, j_2, \ldots$. If, and only if, the last term of this sequence is 1 then we get a non trivial contribution. Let π be the partition of $1, \ldots, k$ such that l and m belong to the same part if and only if $i_l = i_m$. The fact that $(*\, i_1)\ldots(*\, i_k)$ fixes $*$ depends only on this partition, and we call admissible partitions the ones for which $(*\, i_1)\ldots(*\, i_k)$ fixes $*$. Furthermore, the conjugacy class, in S_n, of $(*\, i_1)\ldots(*\, i_k)$ depends only on the partition. Let $\lambda(\pi)$ be the Young diagram formed with the nontrivial cycles of this conjugacy class. Let

$$\mathcal{Z}_\pi = \sum_{i_1,\ldots,i_k \sim \pi} (*\, i_1)\ldots(*\, i_k)$$

where $i_1, \ldots, i_k \sim \pi$ means that the partition associated to the sequence $i_1, \ldots, i_k$ is π, then we have

$$\mathcal{M}_k = \sum_{\pi \text{ admissible}} \mathcal{Z}_\pi$$

Let $c(\pi)$ be the number of parts of π, then the number of k-tuples $i_1, \ldots, i_k \sim \pi$ is equal to $(n)_{c(\pi)}$ (where, as usual $(n)_k = n(n-1)\ldots(n-k+1)$), and one has $|\lambda(\pi)| \leq c(\pi)$, therefore one has

$$\mathcal{Z}_\pi = \frac{(n)_{c(\pi)}}{(n)_{|\lambda|}}\, a_{\lambda(\pi);n} = (n - |\lambda|)\ldots(n - c(\pi) + 1) a_{\lambda(\pi);n}$$

In order that $\lambda(\pi)$ be a cycle of length $k-1$, it is necessary an sufficient that π be the partition $\{1,k\}, \{2\}, \{3\}, \ldots, \{k-1\}$. All other admissible partitions have $c(\pi) < k-1$. We deduce that

$$\mathcal{M}_k = a_{k-1;n} + \sum_{\pi \ \text{admissible}, \, c(\pi) < k-1} Z_\pi$$

It follows from Section 3 that $\mathcal{M}_k$ is a polynomial with integer coefficients, independent of n, in the $a_{j;n}$, of the form

$$a_{k-1;n} + (\text{polynomial in } a_{j;n}; j < k-1).$$

We can thus invert this polynomial relation and get

$$a_{k-1;n} = \mathcal{M}_k + (\text{polynomial in } \mathcal{M}_j; j < k).$$

with polynomial with integer coefficients. Since $\mathcal{M}_k$ can be expanded as polynomials with integer coefficients in the $\mathcal{R}_j$ we thus have

$$a_{k-1;n} = \mathcal{R}_k + (\text{polynomial in } \mathcal{R}_j; j < k).$$

Applying χ_ω to both sides of this equation and using Lemma 2, we obtain Theorem 1. □

5 Frobenius formula and free cumulants

Let $\lambda = (\lambda_1, \lambda_2, \ldots)$ be a partition of n, with $\lambda_1 \geq \lambda_2 \geq \ldots$, and $\mu_i = \lambda_i + n - i$. Let

$$\varphi(z) = \prod (z - \mu_i),$$

then the value of the normalized character χ_λ on a cycle of length k is given by Frobenius' formula

$$(n)_k \chi_\lambda(c_k) = -\frac{1}{k}[z^{-1}]\, z(z-1)\ldots(z-k+1)\varphi(z-k)/\varphi(z). \tag{15}$$

See [M], I.7, Example 7, pages 117–118 (beware that characters are not normalized in Macdonald's book). Now we remark that

$$z\varphi(z-1)/\varphi(z) = 1/G_\lambda(z+n-1) = H_\lambda(z+n-1)$$

therefore

$$(n)_k \chi_\lambda(c_k) = -\frac{1}{k}[z^{-1}]\, H_\lambda(z+n-1)\ldots H_\lambda(z+n-k)$$

Using the invariance of the residue under translation of the variable one gets

$$(n)_k \chi_\lambda(c_k) = -\frac{1}{k}[z^{-1}]\, H_\lambda(z)\ldots H_\lambda(z-k+1).$$

Comparing with (6) we deduce the following formula for Kerov's polynomials.

Theorem 2. *Consider the formal power series*

$$H(z) = z - \sum_{j=2}^{\infty} B_j z^{1-j}.$$

Define

$$\Sigma_k = -\frac{1}{k}[z^{-1}]H_\lambda(z)\dots H_\lambda(z-k+1)$$

and

$$R_{k+1} = -\frac{1}{k}[z^{-1}]\,H_\lambda(z)^k$$

then the expression of Σ_k in terms of the R_k's is given by Kerov's polynomials.

This formula for computing Kerov's polynomials was shown to me by A. Okounkov [O]. It seems plausible that S. Kerov was aware of this (see especially the account of Kerov's central limit theorem in [IO]). It is much easier to implement, than the algorithm given by the proof in Section 4. We give the result of some Maple computations in Section 8.

6 Computation of some coefficients of Kerov's polynomials

For a term $R_2^{k_2}\dots R_r^{k_r}$ define its *degree* by $\sum_j k_j$ and its *weight* is $\sum_j j\,k_j$. It is clear from sign considerations that in the expansion of Σ_k only terms of weight having the opposite parity of k occur. The term of highest weight is R_{k+1} and it is the only term with this weight, as follows from Theorem 2.

We shall first be interested in the terms of degree one.

Theorem 3. *The coefficient of R_{k+1-2l} in Σ_k is equal to the number of cycles $c \in S_k$, of length k, such that $(1\,2\dots k)\,c$ has $k-2l$ cycles.*

In order to prove the theorem we shall compute the generating function for the linear coefficients, using Theorem 2 formula. Since we are interested only in linear terms, we see that the formula expressing Σ_k in terms of R_j is the same as the one in terms of B_j. Put $B_i = tx^{i-1}$. We shall find the coefficient of z^{-1} in $H(z)H(z-1)\dots H(z-k+1)$, keeping only the terms with degree one in t. One has $H(z) = z - tx/(z-x)$, therefore

$$[t]\,H(z)H(z-1)\dots H(z-k+1) = x\sum_{j=0}^{k-1}\frac{z(z-1)\dots(z-k+1)}{(z-k)(z-x-j)}$$

Using again invariance of residue by translation one obtains

$$\sum_l x^{2l}[R_{k+1-2l}]\,\Sigma_k = \frac{1}{k}\sum_{j=0}^{k-1}\prod_{l=0}^{k-1}(x+j-l) = \frac{1}{k}\sum_{l=0}^{k-1}Q_k(x-l) \qquad (16)$$

where $Q_k(x) = x(x+1)\dots(x+k-1)$.

Denote by d_λ the dimension of the irreducible representation with Young diagram λ.

Lemma 3. *Let λ be a Young diagram with k cells, let*

$$P_\lambda(x) = \sum_{\square \in \lambda} (x + c(\square))$$

be its content polynomial, and $c(\sigma)$ = number of cycles of σ, then

$$\sum_{\sigma \in S_k} d_\lambda \chi_\lambda(\sigma) x^{l(\sigma)} = P_\lambda(x)$$

See [M], I.1, Example 11, I.3, Example 4, and (7.7).

Let c_k be the cycle $(1\,2\dots k)$, by the orthogonality relations for characters and Lemma 3, one has

$$\frac{1}{k!}\sum_\lambda d_\lambda \chi_\lambda(c_k) P_\lambda(x) P_\lambda(y) = \sum_{\sigma \in S_k} x^{l(\sigma^{-1}c_k)} y^{l(\sigma)} \tag{17}$$

Let us compute the coefficient of y in the left hand side of (17). Only hook diagrams $\lambda = (k-l, 1^l)$ for $l = 0, \dots, k-1$, contribute and for such a diagram $d_\lambda \chi_\lambda(c_k) = \binom{k-1}{l}(-1)^l$. One has $P_\lambda(x) = Q(x-l)$, the coefficient of y in $\frac{1}{k!}P_\lambda(y)$ is $\frac{1}{k}\binom{k-1}{l}^{-1}$, and $P_\lambda(x) = Q(x-l)$ therefore we find formula (16) for the left hand side. Comparing with the right hand side we get Theorem 3. □

Theorem 3 has also been proved by R. Stanley [St1] by a closely related method.

Theorem 4. *The coefficient of $R_{k-3}R_2$ in Σ_k (for $k > 5$) is $(k+1)k(k-1)(k-4)/12$.*

Again this follows from Theorem 2 through some lengthy, but straightforward computations which are omitted.

More generally, based on numerical investigations, we conjecture the following formula for terms of weight $k-1$.

Conjecture 1. The coefficient of $R_2^{l_2} \dots R_s^{l_s}$ in Σ_k, with $k = 2l_2 + 3l_3 + \dots + sl_s + 1$ is equal to

$$\frac{(k+1)k(k-1)}{24} \frac{(l_2 + \dots + l_s)!}{l_2! \dots l_s!} \prod_{j=2}^{s} (j-1)^{l_j}$$

The validity of this conjecture has been checked up to $k = 15$. A proof for the other cases (at least for degree two terms) can presumably be given using Theorem 2 but the computations become quickly very involved. No such simple product formula seems to be available for the general term.

Another natural conjecture is that all coefficients in Kerov's formula are non negative integers, which also has been checked up to $k = 15$. See the next Section for more on this.

7 Connection with the Cayley graph of symmetric group

Let us explore more thoroughly the connections between the M_k, B_k, R_k and Σ_k. Formulas (9) and (10) provide a natural combinatorial model for expressing moments in terms of free or Boolean cumulants. We will be looking for similar models for expressing the other connections. Observe first that for all measures associated with Young diagrams one has $M_1 = B_1 = R_1 = 0$. We will restrict ourselves to this case in the following.

Let us apply Kerov's formula to the trivial character. As we shall see, this will give a lot of information. The probability measure associated with the trivial character is $\frac{n}{n+1}\delta_{-1} + \frac{1}{n+1}\delta_n$ with Cauchy transform

$$G(z) = \frac{z-n+1}{(z+1)(z-n)}.$$

The corresponding moments are

$$M_k = \frac{n}{n+1}(-1)^k + \frac{1}{n+1}n^k = n\frac{n^{k-1}-(-1)^{k-1}}{n-(-1)} = (-1)^{k-1}\sum_{j=1}^{k-1}(-n)^j$$

We will find it useful to interpret this as the generating function for the rank in a totally ordered set with $k-1$ elements. We take this totally ordered set as the set I_{k-1} of partitions of $\{1,\dots,k-1\}$ given by $\{1,2,\dots,j\};\{j+1\};\{j+2\};\dots\{k-1\}$ and the rank is the number of parts.

$$M_k = (-1)^{k-1}\sum_{\pi\in I_{k-1}}(-n)^{|\pi|} \tag{18}$$

The k^{th} free cumulant of this measure is the generating function

$$R_k = (-1)^{k-1}\sum_{l=1}^{k-1}\frac{1}{k-1}\binom{k-1}{l-1}\binom{k-1}{l}(-n)^l = (-1)^{k-1}\sum_{\pi\in NC(k-1)}(-n)^{|\pi|}. \tag{19}$$

For the Boolean cumulants one finds

$$B_k = n(n-1)^{k-2} = (-1)^{k-1}\sum_{j=1}^{k-1}\binom{k-2}{j-1}(-n)^j = (-1)^{k-1}\sum_{\pi\in B(k-1)}(-n)^{|\pi|}. \tag{20}$$

The evaluation of the trivial character on Σ_k gives $n(n-1)\dots(n-k+1)$. This is the generating function (recall that $l(\sigma)$ is the number of cyles of σ)

$$\Sigma_k = (-1)^{k-1}\sum_{\sigma\in S_k}(-n)^{l(\sigma)}$$

Let us now concentrate on the formula expressing the free cumulants in terms of Boolean cumulants, and the Boolean cumulants in terms of moments. The first such expressions are

$$\begin{aligned}
R_2 &= B_2\\
R_3 &= B_3\\
R_4 &= B_4 - B_2^2\\
R_5 &= B_5 - 3B_2B_3\\
R_6 &= B_6 - 4B_2B_4 - 2B_3^2 + 2B_2^3\\
R_7 &= B_7 - 5B_2B_5 - 5B_3B_4 + 10B_2^2B_3
\end{aligned}$$

$$\begin{aligned}
B_2 &= M_2\\
B_3 &= M_3\\
B_4 &= M_4 - M_2^2\\
B_5 &= M_5 - 2M_2M_3\\
B_6 &= M_6 - 2M_2M_4 - M_3^2 + M_2^3\\
B_7 &= M_7 - 2M_2M_5 - 2M_3M_4 + 3M_2^2M_3
\end{aligned}$$

These formulas contain signs but we can make all coefficients positive by an overall sign change of all variables. We are going to give a combinatorial interpretation of these coefficients. Let us replace moments, free cumulants and Boolean cumulants by the values (18), (19) and (20). It seems natural to try to interpret the formulas expressing free cumulants as providing a decomposition of the lattice of noncrossing partitions

$$NC(k-1) = \cup_\pi B[\pi]$$

into a disjoint union of subsets which are products of Boolean lattices, whereas the formula expressing Boolean cumulants should come from a decomposition of the Boolean lattice

$$B(k-1) = \cup_\pi I[\pi]$$

into a union of products of totally ordered sets.

It turns out that such decompositions exist, and we shall now describe them. First we look at the formula for expressing Boolean cumulants in terms of moments. On the Boolean lattice of interval partitions, let us put a new, stronger order relation. For this new order a covers b if and only if b can be obtained from a by deleting the last element of some interval, if this interval had at least three elements, or if it is the interval $[1, 2]$. The resulting decomposition of the Boolean lattice of interval partitions of $\{1, 2, 3, 4, 5\}$ is shown in the following picture. It corresponds to the formula for B_6 above. The first line in the picture corresponds to the interval M_6, the second and third line account for the two intervals M_2M_4, the fourth line for the product interval M_3^2, and the last line for the point corresponding to M_2^3. The ranking is horizontal. The proof that this decompostion yields the right interpretation of the Boolean cumulant-moment formula is easy and left to the reader.

Now let us decompose the lattice of non-crossing partitions into a union of Boolean lattices. For this we put the following new order on noncrossing partitions. A noncrossing partition a covers a noncrossing partition b if and only if b can be obtained from a by cutting a part of b between two successive

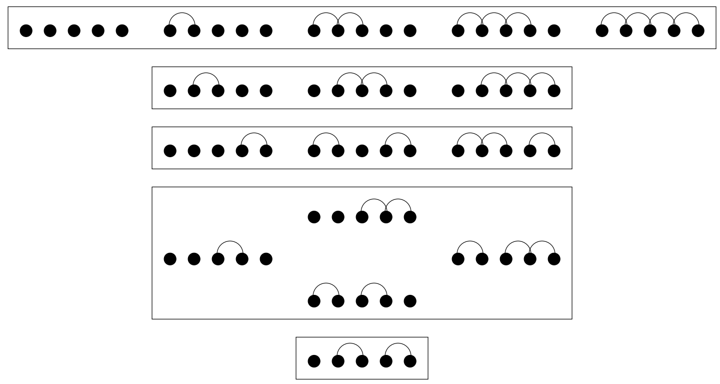


Fig. 3.

elements $i, i+1$. For example we give here the list of the Boolean intervals obtained in this way in $NC(5)$, corresponding to the formula for R_6. Next to each interval we give the term to which it corresponds in the formula. Consider the Cayley graph of S_k with respect to the generating set of all transpositions. The lattice $NC(k)$ can be embedded into S_k, as the subset of elements lying on a geodesic from e to the full cycle $(12...k)$. In this embedding, the cycles of a permutation correspond to the parts of a partition, hence the functions $|\pi|$ and $l(\sigma)$. We give in the notation the cycle structure of a noncrossing partition as an element of S_k. An interval $[\sigma, \sigma']$ in which $\sigma^{-1}\sigma'$ has a cycle structure $2^{l_2} \dots k^{l_k}$ corresponds to a term $B_2^s B_3^{l_2} B_4^{l_3} \dots B_{k+1}^{l_k}$, where s is such that the rank of σ is $s + l_2 + \dots + l_k$.

$$\begin{array}{cc}
[e, (1\,2\,3\,4\,5)] & B_6 \\
[(1\,3), (1\,3\,4\,5)] & B_2B_4 \\
[(1\,4), (1\,4\,5)(2\,3)] & B_3^2 \\
[(1\,5), (1\,5)(2\,3\,4)] & B_2B_4 \\
[(2\,4), (1\,2\,4\,5)] & B_2B_4 \\
[(2\,5), (1\,2\,5)(3\,4)] & B_3^2 \\
[(3\,5), (1\,2\,3\,5)] & B_2B_4 \\
[(1\,3\,5)] & B_2^3 \\
[(1\,5)(2\,4)] & B_2^3
\end{array}$$

Observe that each interval above is in fact an interval for the Bruhat order.

Looking at the above results and at Theorem 3, it is tempting to try interpreting Kerov polynomials as coming from a decomposition of the symmetric group into "intervals" for some suitable order relation, in which the adjacency relation should be induced by the one of the Cayley graph. The coefficient of $R_2^{l_2} \dots R_s^{l_s}$ would count the number of intervals isomorphic to the ordered set $NC(1)^{l_2} NC(2)^{l_3} \dots NC(s-1)^{l_s}$. Such interval would be of the form $[\sigma, \sigma']$ with $|\sigma| = 1 + (k+1-\sum_j l_j)/2$, and $|\sigma'| = 1 + (k+1+\sum_j l_j)/2$. The obvious choice for the first term R_{k+1} in Σ_k would be to take the set of geodesics from e to $(1\,2\ldots k)$. Observe however that because of sign problems we should get a "signed" covering of S_k, namely each element of S_k would be contained in a certain number of intervals and intervals corresponding to terms of even degree would give a multiplicity one while terms of odd degree would give a multiplicity -1, the sum of multiplicities would then be $+1$ for any $\sigma \in S_k$.

One way to get around this problem of signed covering would be to look at the expression of characters in terms of Boolean cumulants, where this problem disappears. One would then be lead to look for a decomposition of the Cayley graph into a union of products of Boolean lattices. It is here natural to try doing so by using the Bruhat order. Indeed some decompositions of the symmetric group into Boolean lattices have appeared in the literature [LS], [M], but they are not the ones we are looking for, indeed by Theorem 2 one can compute the total number of intervals which should occur in the decomposition of S_k, and even the generating function of the number of terms according to their degrees, it is given by

$$S_{2p-1}(x) = \frac{1}{p}\binom{2p-2}{p-1} x \prod_{j=2}^{p}(x + i(i-1)) \qquad S_{2p}(x) = (2p-1)S_{2p-1}(x)$$

whereas the above decompositions have $(k-1)!$ intervals.

It has been observed by R. Stanley [St2], that if one evaluates the character of a cycle for a rectangular $p \times q$ Young diagram, then one can improve the formulas (18) to (20) by replacing them by their homogeneous two-variable correspondants, while the character is now given by the rhs of (17), with $x = -p, y = q$. This gives more evidence for the connection with the Cayley graph of symmetric group.

Let us now look at the first values of k. The cases of Σ_k for $k = 2, 3, 4$ do not present difficulty so let us concentrate on $\Sigma_5 = R_6 + 15R_4 + 5R_2^2 + 8R_2$. We already have the interpretation of the terms R_6 and $8R_2$, they should correspond repectively to the interval $[e, (1\,2\,3\,4\,5)]$, of elements such that $d(e,\sigma) + d(\sigma, (1\,2\,3\,4\,5)) = 4$ (d is the distance in the Cayley graph) and the eight one point intervals $[c]$ where c is a 5-cycle whose product with $(1\,2\,3\,4\,5)^{-1}$ is a 5-cycle. It remains to cover the elements of S_5 satisfying $d(e,\sigma) + d(\sigma, (1\,2\,3\,4\,5)) = 6$ by 15 intervals isomorphic to a 3-cycle. This should yield 5 elements with $d(e,\sigma) = 3 = d(\sigma, (1\,2\,3\,4\,5))$ which are counted twice. After some guesswork the following (non unique) decomposition can be found. The 15 intervals are

$$[(1\,3)(2\,4), (1\,3\,4\,2\,5)]$$
$$[(1\,3\,2), (1\,3\,2\,4\,5)]$$
$$[(1\,4\,2), (1\,4\,2\,3\,5)]$$

and their conjugates by $(1\,2\,3\,4\,5)$. The 5 elements counted twice are

$$(1\,3)(2\,5\,4)$$

which appears in the intervals $[(1\,3)(2\,4),(1\,3\,4\,2\,5)]]$ and $[(2\,5\,4),(1\,3\,2\,5\,4)]$, and all its conjugates by $(1\,2\,3\,4\,5)$.

8 Values of Σ_k for $k = 7$ to 11

$$\begin{aligned}
\Sigma_7 &= R_8 + 70\,R_6 + 84\,R_2\,R_4 + 56\,{R_3}^2 + 14\,{R_2}^3 + \\
&\quad 469\,R_4 + 224\,{R_2}^2 + 180\,R_2
\end{aligned}$$

$$\begin{aligned}
\Sigma_8 &= R_9 + 126\,R_7 + 168\,R_2\,R_5 + 252\,R_3\,R_4 + 126\,{R_2}^2\,R_3 \\
&\quad +1869\,R_5 + 2688\,R_2\,R_3 + 3044\,R_3
\end{aligned}$$

$$\begin{aligned}
\Sigma_9 &= R_{10} + 210\,R_8 + 300\,R_2\,R_6 + 480\,R_3\,R_5 + 270\,{R_4}^2 \\
&\quad +360\,R_2\,{R_3}^2 + 270\,{R_2}^2\,R_4 + 30\,{R_2}^4 \\
&\quad +5985\,R_6 + 10548\,R_2\,R_4 + 6714\,{R_3}^2 + 2400\,{R_2}^3 \\
&\quad +26060\,R_4 + 14580\,{R_2}^2 + 8064\,R_2
\end{aligned}$$

$$\begin{aligned}
\Sigma_{10} &= R_{11} + 330\,R_9 + 495\,R_7\,R_2 + 825\,R_3\,R_6 + 990\,R_4\,R_5 \\
&\quad +495\,R_5\,{R_2}^2 + 1485\,R_2\,R_3\,R_4 + 330\,{R_3}^3 + 330\,{R_2}^3\,R_3 \\
&\quad +16401\,R_7 + 32901\,R_2\,R_5 + 46101\,R_3\,R_4 + 33000\,{R_2}^2\,R_3 \\
&\quad +152900\,R_5 + 258060\,R_2\,R_3 + 193248\,R_3
\end{aligned}$$

$$\begin{aligned}
\Sigma_{11} &= R_{12} + 495\,R_{10} + 770\,R_8\,R_2 + 1320\,R_3\,R_7 + 1650\,R_6\,R_4 + 880\,{R_5}^2 \\
&\quad +825\,{R_2}^2\,R_6 + 2640\,R_5\,R_2\,R_3 + 1485\,R_2\,{R_4}^2 + 1980\,{R_3}^2\,R_4 \\
&\quad +660\,{R_2}^3\,R_4 + 1320\,{R_2}^2\,{R_3}^2 + 55\,{R_2}^5 \\
&\quad +39963\,R_8 + 87890\,R_2\,R_6 + 130108\,R_3\,R_5 + 71214\,{R_4}^2 \\
&\quad +105545\,{R_2}^2\,R_4 + 136345\,R_2\,{R_3}^2 + 15400\,{R_2}^4 \\
&\quad +696905\,R_6 + 1459700\,R_2\,R_4 + 902440\,{R_3}^2 + 386980\,{R_2}^3 \\
&\quad +2286636\,R_4 + 1401444\,{R_2}^2 + 604800\,R_2
\end{aligned}$$

References

[B] Biane, P.: Representations of symmetric groups and free probability. Adv. Math., **138**, 126–181 (1998)

[IO] Ivanov, V., Olshanski, G.: Kerov's central limit theorem for the Plancherel measure on Young diagrams. Preprint (June 2001)

[K1] Kerov, S. V.: Transition probabilities of continual Young diagrams and the Markov moment problem. Funct. Anal. Appl., **27**, 104–117 (1993)

[K2] Kerov, S. V.: Talk at IHP conference (January 2000)

[KO] Kerov, S. V., Olshanski, G.: Polynomial functions on the set of Young diagrams. C. R. Acad. Sci. Paris Sér. I Math., **319**, no. 2, 121–126 (1994)

[LS] Lascoux, A., Schützenberger, M. P.: Treillis et bases des groupes de Coxeter. Electron. J. Combin. **3**, no. 2, Research paper 27, 35 pp. (1996)

[M] Macdonald, I. G.: Symmetric functions and Hall polynomials, Second Edition. Oxford Univ. Press, Oxford (1995)

[Mo] Molev, A. I.: Stirling partitions of the symmetric group and Laplace operators for the orthogonal Lie algebra. Proceedings of the 7th Conference on Formal Power Series and Algebraic Combinatorics (Noisy-le-Grand, 1995). Discrete Math., **180**, no. 1–3, 281–300 (1998)

[O] Okounkov, A.: Private communication (January 2001)

[Sp] Speicher, R.: Combinatorial Theory of the Free Product with Amalgamation and Operator-Valued Free Probability Theory. Memoirs of the AMS, **627** (1998)

[St1] Stanley, R.: Private communication (May 2001)

[St2] Stanley, R.: Irreducible symmetric group characters of rectangular shape. Preprint (September 2001)

[SW] Speicher, R., Woroudi, R.: Boolean independence. In: Voiculescu, D. (ed) Free probability, Fields Institue Communications, 267–280 (1996)

Algebraic length and Poincaré series on reflection groups with applications to representations theory

Marek Bożejko* and Ryszard Szwarc**

Institute of Mathematics
Wrocław University
pl. Grunwaldzki 2/4
50–384 Wrocław, Poland
bozejko@math.uni.wroc.pl
and
szwarc@math.uni.wroc.pl

Summary. Let W be a reflection group generated by a finite set of simple reflections S. We determine sufficient and necessary condition for invertibility and positive definitness of the Poincaré series $\sum_w q^{\ell(w)} w$, where $\ell(w)$ denotes the algebraic length on W relative to S. Generalized Poincaré series are defined and similar results for them are proved.

In case of finite W, representations are constructed which are canonically associated with the algebraic length. For crystallographic groups (Weyl groups) these representations are decomposed into irreducible components. Positive definitness of certain functions involving generalized lengths on W is proved. The proofs don't make use of the classification of finite reflection groups. Examples are provided.

Key words: reflection group, Coxeter group, root system, Poincaré series

1 Introduction

We start with an account of definitions and results on reflection groups that we need in this paper. For the details we refer to the book by Humphreys [20], whose notation we follow, and to [2, Chapter VI].

A reflection in a finite dimensional vector space V, endowed with a nondegenerate symmetric bilinear form (x, y), is a linear operator which sends some nonzero vector to its negative and fixes the orthogonal complement to this vector. If s_x denotes the reflection about the vector x then s_x acts by the rule

* Partially supported by KBN (Poland) under grant 2 P03A 054 15.
** Partially supported by KBN (Poland) under grant 2 P03A 048 15.

A. Borodin et al.: LNM 1817, A.M. Vershik (Ed.), pp. 201–221, 2003.

$$s_x y = y - 2\frac{(x,y)}{(x,x)}x. \tag{1}$$

Let the vectors $\Delta = \{r_1, r_2, \ldots, r_n\}$ form a basis for V and let $S = \{s_1, s_2, \ldots s_n\}$ denote the set of corresponding reflections. A reflection group is a group W generated by S.

The set

$$\Phi = \{wr_i \mid w \in W,\, i = 1, 2 \ldots, n\}$$

is called a root system whenever it can be decomposed as $\Phi = \Pi \cup -\Pi$, where Π denotes the subset of Φ consisting of the vectors $\sum_1^n a_i r_i$, with $a_i \geq 0$. The elements of Π are called positive roots, while those of $-\Pi$ are called negative roots. The elements of Δ are called the simple roots and the elements of S are called simple reflections.

If the group W is finite (which means that Φ is a finite set) then the order $\mathrm{m}(i,j)$ of any product $s_i s_j$ has to be finite. It turns out that the numbers $\mathrm{m}(i,j)$ determine the group W up to an isomorphic equivalence. The group W can be described in algebraic way as generated by S with

$$s_i^2 = 1, \qquad (s_i s_j)^{\mathrm{m}(i,j)} = 1,$$

as the only relations among elements in S. A convenient way of encoding this information is the Coxeter graph of W. The set of generators is the vertex set of this graph. If $\mathrm{m}(s,s') \geq 3$ we join s and s' by an edge labelled with $\mathrm{m}(s,s')$.

Any element $w \in W$ is a product of simple reflections, say $w = s_{i_1} s_{i_2} \ldots s_{i_k}$. The smallest k for which this occur will be called the length $\ell(w)$ of w. The length has important geometric interpretation. There holds

$$\ell(w) = |\Pi \cap w^{-1}(-\Pi)|. \tag{2}$$

This means that $\ell(w)$ is precisely the number of positive roots sent by w to negative roots.

If we select some of the reflections $J \subseteq S$, then the group generated by them is also a reflection group denoted by W_J. The corresponding simple roots are $\Delta_J = \{r_i \mid s_i \in J\}$. The set

$$\Phi_J = \{wr_i \mid w \in W_J,\ s_i \in J\}$$

is a corresponding root system. The set Φ_J decomposes into subsets of positive Π_J and negative $-\Pi_J$ roots. The length of w as element in W_J coincides with the one defined by the ambient system (W, S). Moreover for elements in $w \in W_J$ we have

$$\Pi_J \cap w^{-1}(-\Pi_J) = \Pi \cap w^{-1}(-\Pi). \tag{3}$$

(see [20, Proposition 1.10]).

The group W is finite if and only if the bilinear form associated with the $n \times n$ matrix

$$a(i,j) = -\cos\frac{\pi}{\mathrm{m}(i,j)}$$

is positive definite. Moreover finite reflection groups are classified into 5 series and some exceptional cases.

Part of our result will hold only for so called crystallographic reflection groups. These are finite reflection groups such that $2a(i,j)$ is an integer for any i, j. It turns out that these are precisely the Weyl groups of simple Lie algebras. The finite reflections groups that are left behind are dihedral groups and two exceptional ones, known as H_3 and H_4.

In Section 2 we study the invertibility of Poincaré series of (W, S). This has many important applications in noncommutative probability theories. The Poincaré series (or Poincaré polynomial in case of finite W) is a formal power series of the form $\sum\limits_{w\in W} q^{\ell(w)}w$, where q is a complex number. For finite Coxeter groups (W, S) we determine the values of Q for which $P_q(W)$ is invertible in the group algebra of W. In particular, for the permutation group S_n, we obtain a result of Zagier [28] that $P_q(S_n)$ is invertible if and only if $q^{(i-1)i} \neq 1$ for any $i = 1, 2, \ldots, n$. Zagier uses the invertibility properties of $P_q(S_n)$ in constructing models of q–oscillators in infinite statistics. On the other hand Bożejko et al. [6, 7, 8, 9] make use of positive definitness of $P_q(S_n)$ this time in constructing generalized Brownian motions and in finding realizations of noncommutative Gausssian processes in connection with quantum Young–Baxter equation (see also [22]). Dykema and Nica [16] use Fock realizations of canonical q–commutation relations found in [8]. They derive positivity properties of $P_q(S_n)$ to show stability properties of the generalized Cuntz algebras generated by q–oscillator (see [6, 8, 17, 22]). The positive definitness of $P_q(W)$ for infinite Coxeter groups was a key point in the proof [4] that such groups do not have Kazhdan property T.

In this paper we study also positive definitness and invertibility of the generalized Poincaré series for (W, S). This is any function $P(\cdot)$ which is multiplicative in the following sense: if the product uv is reduced for $u, v \in W$, then $P(uv) = P(u)P(v)$. We also require that $P(s) = P(s')$ for $s, s' \in S$, whenever s and s' are conjugate in W. We show that the generalized Poincaré series $P = \sum P(w)\, w$ is positive definite if and only if $-1 \leq P(s) \leq 1$, for any $s \in S$ (see [7], where the finite Coxeter groups were considered).

The positive definitness of the Poincaré series is also connected with the problem of determining which locally compact groups are so called weakly amenable (see [14]). Roughly a group G is weakly amenable if the Fourier algebra $A(G)$ admits an approximate unit bounded with respect to completely bounded multiplier norm (this is a norm weaker than the norm of $A(G)$ but stronger than the multiplier norm on $A(G)$). Partial results were obtained for special groups (see [5, 21, 23, 24, 27]). In particular our Theorem 1 played a crucial role in [21]. We conjecture that all Coxeter groups are weakly amenable. We think that the results in this paper constitute a step towards proving this conjecture. Another step, from the geometrical viewpoint, has been made in [25].

The proof of invertibility of $P_q(W)$ relies on a peculiar geometrical property of the Coxeter complex associated with (W, S). Namely, if w_0 denotes the longest element in W, then

$$w_0 W_{J_1} \cap W_{J_2} = \emptyset,$$

where J_1 and J_2 are arbitrary proper subsets of S. In other words, any facet about the identity element of the group e is disjoint from any facet about the element w_0.

In Section 3 we construct representations of W associated with the length $\ell(w)$. We show that the correspondence

$$w \mapsto \tfrac{1}{2}|\Pi| - \ell(w)$$

is a positive definite function on W. As a sideffect we obtain another proof of positive definitness of the Poincaré series $P_q(W)$ in case of finite Coxeter group. It would be of great interest to find a canonical representation of W associted with $P_q(W)$. This has been done only for special cases (see [23, 24]).

The main result in Section 3.1 is a decomposition of the representations corresponding to the length function into irreducible components in case of crystallographic groups. For noncrystallographic case, the decomposition holds but it does not yield irreducible components. In Section 3.2 we compute the decomposition of the function $\varphi(w) = \frac{1}{2}|\Pi| - \ell(w)$ into irreducible positive definite functions. Th eprrof don't make use of classification of finite reflection groups. The key point is a characterization of kernels $k(x, y)$ defined on $\Phi \times \Phi$, invariant for the action of the group W.

Throughout the paper by $A \subset B$ we will mean that A is properly contained in B. Otherwise we will write $A \subseteq B$. The symbol $|A|$ will denote the number of elements in the set A. The symbol $A \triangle B$ will denote the set $(A \setminus B) \cup (B \setminus A)$.

2 Positive definitness and invertibility of the Poincaré series

Let (W, S) be a Coxeter system with finite generator set S. For an element $w \in W$ let $\ell(w)$ denote the length of w.

For a subset J of S the symbol W_J will denote the subgroup of W generated by J. Let

$$W^J = \{w \in W \mid \ell(ws) > \ell(w), \text{ for all } s \in J\}. \tag{4}$$

By [20, page 19] (see also [2, Problem IV.1.3]) any element $w \in W$ admits a unique decomposition

$$w = w_J w^J, \quad \ell(w) = \ell(w_J) + \ell(w^J), \tag{5}$$

where $w_J \in W_J$ and $w^J \in W^J$.

Let $\underline{q} = \{q_s\}_{s\in S}$ be a family of complex numbers such that $q_s = q_{s'}$, if s and s' are conjugate in W. For an element $w \in W$, let

$$\underline{q}^w = q_{s_1} q_{s_2} \dots q_{s_n} \quad \text{if} \quad w = s_1 s_2 \dots s_n,\ \ell(w) = n.$$

By [2, Proposition 1.5.5] the function $w \mapsto \underline{q}^w$ is well defined. Observe that if $q_s = q$ for all $s \in S$, then $\underline{q}^w = q^{\ell(w)}$. The function $\underline{q}^w$ is multiplicative in the following sense.

$$\underline{q}^{w_1 w_2} = \underline{q}^{w_1} \underline{q}^{w_2}\ ,\ \text{if } \ell(w_1 w_2) = \ell(w_1) + \ell(w_2).$$

The conjugation relation yields a decomposition of S into equivalence classes say

$$S = A_1 \cup A_2 \cup A_m. \tag{6}$$

Hence $\underline{q}$ takes m values $q_1, q_2, \dots, q_m$. Let $w = s_1 s_2 \dots s_n$ be a reduced representation of w. Then

$$\underline{q}^w = \prod_{i=1}^{m} q_i^{\ell_i(w)}, \tag{7}$$

where

$$\ell_i(w) = |\{j \mid s_j \in A_i\}|. \tag{8}$$

By aforementioned [2, Proposition 1.5.5] the function $\ell_i(w)$ does not depend on representation of w in reduced form. This will also follow from the next proposition.

Proposition 1.

(i) Two simple reflections s and s' in S are conjugate in W if and only if there exists $w \in W$ such that $wr = r'$, where r and r' are the corresponding simple roots.

(ii) For any $w \in W$ and $i = 1, 2, \dots, m$

$$\ell_i(w) = |\Pi_i \cap w^{-1}(-\Pi_i|,$$

where $\Pi_i = \Pi \cap \{wr_j \mid s_j \in A_i\}$.

Proof. Part (i) follows immediately from the identity

$$s_{wr} = w s_r w^{-1}.$$

Let $w = s_1 s_2 \dots s_n$ be a reduced representation of w. Define roots θ_i, $i = 1, 2, \dots, n$ by the rule

$$\theta_i = s_n s_{n-1} \dots s_{i+1}(r_i).$$

It is easy to check that

$$\Pi \cap w^{-1}\Pi = \{\theta_1, \theta_2, \dots, \theta_n\},$$

(see [20, Exercise 5.6.1, page 115]). Observe that $s_j \in A_i$ if and only if $\theta_j \in \Pi_i$. This yields the conclusion.

Lemma 1. *Assume a group W acts transitively on a set Ω, and there exists a subset $A \subset \Omega$, such that $(wA)\triangle A$ is a finite set for any $w \in G$. Then the function $w \mapsto q^{|(wA)\triangle A|}$ is positive definite on W, for any $-1 \leq q \leq 1$.*

Proof. Observe that

$$|(v^{-1}wA)\triangle A| = |(wA)\triangle(vA)| = \sum_{x\in\Omega} |\chi_{wA} - \chi_{vA}|^2.$$

By [1, page 81] the correspondence $w \mapsto |(wA)\triangle A|$ is so called negative definite function on W. Thus by Schoenberg's theorem (see [1, Theorem 2.2, page 74]) we get the conclusion for $0 \leq q \leq 1$. Observe that the real valued function

$$w \mapsto (-1)^{|(wA)\triangle A|}$$

is multiplicative on W. Hence it is positive definite on W. Using Schur's theorem, that the product of positive definite functions yields another such function, completes the proof of the lemma.

Theorem 1. *Let (W, S) be a Coxeter system and let $\underline{q} = \{q_s\}_{s\in S}$ be such that $-1 \leq q_s \leq 1$ and $q_s = q_{s'}$, whenever s, s' are conjugate in W. Then the function $w \mapsto \underline{q}^w$ is positive definite.*

Proof. By (7) it suffices to show that $w \mapsto q_i^{\ell_i(w)}$ is positive definite. This follows immediately from Lemma 1, Proposition 1 and the fact that

$$\ell_i(w) = |\Pi_i \cap w^{-1}(-\Pi)| = \frac{1}{2}|(w\Pi_i)\triangle\Pi_i|.$$

The functions $w \mapsto \ell_i(w)$ will play essential role in the next section.

The function $w \mapsto \underline{q}^w$ will be called the Poincaré series of W and denoted by $P_{\underline{q}}(W)$. It can be expressed as the power series

$$P_{\underline{q}}(W) = \sum_{w\in W} \underline{q}^w\, w.$$

In the sequel we will identify $P_{\underline{q}}(W)$ with the convolution operator by this function on $\ell^2(W)$. We are interested when this operator is invertible.

For a subset $A \subset W$, let

$$P_{\underline{q}}(A) = \sum_{w\in A} \underline{q}^w\, w\ .$$

By (4) and (5) we immediately get

$$P_{\underline{q}}(W) = P_{\underline{q}}(W_J) P_{\underline{q}}(W^J)\ . \tag{9}$$

The following formula has important consequences (cf. [20, Proposition 1.11]).

Proposition 2.
(i) Let (W,S) be a finite Coxeter group. Then

$$\underline{q}^{w_0} w_0 = \sum_{J \subseteq S} (-1)^{|J|} P_{\underline{q}}(W^J) \ ,$$

where w_0 is the unique longest element in W.
(ii) If $P_{\underline{q}}(W)$ is an invertible operator, then

$$\sum_{J \subset S} (-1)^{|J|} P_{\underline{q}}(W_J)^{-1} = \{\underline{q}^{w_0}\, w_0 - (-1)^{|S|}\, e\} P_{\underline{q}}(W)^{-1} \ .$$

Proof. By (9) the operator $P_{\underline{q}}(W_J)$ is invertible, if $P_{\underline{q}}(W)$ is invertible. Thus it suffices to show (i). We have

$$\begin{aligned} \sum_{J \subseteq S} (-1)^{|J|} P_{\underline{q}}(W^J) &= \sum_{J \subseteq S} (-1)^{|J|} \left(\sum_{w \in W^J} \underline{q}^w \lambda(w) \right) \\ &= \sum_{w \in W} \left(\sum_{\substack{J \\ w \in W^J}} (-1)^{|J|} \right) \underline{q}^w \lambda(w) \ . \end{aligned}$$

Let

$$J_w = \{ s \in S \mid \ \ell(ws) > \ell(w) \}.$$

Observe that $w \in W^J$ if and only if $J \subseteq J_w$. Therefore

$$\sum_{\substack{J \\ w \in W^J}} (-1)^{|J|} = \sum_{J \subset J_w} (-1)^{|J|} = (1-1)^{|J_w|} = \begin{cases} 0 & \text{if } \ J_w \neq \emptyset \\ 1 & \text{if } \ J_w = \emptyset \end{cases}$$

However $J_w = \emptyset$ if and only if $w = w_0$. Thus

$$\sum_{J \subset S} (-1)^{|J|} P_{\underline{q}}(W_J)^{-1} = \underline{q}^{w_0} w_0.$$

The subset $I \subseteq S$ will be called connected if I is connected in the Coxeter graph of (W,S). If W_I is finite group the unique longest element of W_I will be denoted by $w_0(I)$.

For a subset $J \subseteq S$ let

$$T(J) = \{ \underline{q} \mid \ (\underline{q}^{w_0(I)})^2 = 1 \text{ for some connected } I \subseteq J \}.$$

Proposition 3. *Let (W,S) be a finite Coxeter system. If $\underline{q} \notin T(S)$, then the convolution with $P_{\underline{q}}(W)$ is an invertible operator on $\ell^2(W)$.*

Proof. We prove the assertion by induction on $n = |S|$. For $n = 1$ we have $S = \{s\}$ and $P(W) = 1 + qs$. Hence $P_{\underline{q}}(W)^{-1} = (1 - q^2)^{-1}(1 - qs)$ exists as long as $q^2 \neq 1$.

Assume S is not connected. Then $S = S' \cup S''$, $s's'' = s''s'$ for any $s' \in S'$ and $s'' \in S''$. Hence $W = W_{S'}W_{S''}$. This implies

$$P_{\underline{q}}(W) = P_{\underline{q}}(W_{S'})P_{\underline{q}}(W_{S''}).$$

Thus, with no loss of generality, we can restrict ourselves to the case S is connected. By definition we have

$$T(J) \subseteq T(S).$$

Therefore, if $\underline{q} \notin T(S)$, then $\underline{q} \notin T(J)$ for $J \subset S$. By induction hypothesis the inverse $P_{\underline{q}}(W_J)^{-1}$ exists for each $J \subset S$. Thus by Proposition 2

$$\begin{aligned} \underline{q}^{w_0}w_0 - (-1)^{|S|}e &= \sum_{J \subset S} (-1)^{|J|} P_{\underline{q}}(W^J) \\ &= \left[\sum_{J \subset S} (-1)^{|J|} P_{\underline{q}}(W_J)^{-1}\right] P_{\underline{q}}(W) \ . \end{aligned}$$

Using the fact that $w_0^2 = 1$ (see [20, page 16]) gives

$$\begin{aligned} (\underline{q}^{w_0})^2 - 1 &= (\underline{q}^{w_0}w_0 + (-1)^{|S|}e)(\underline{q}^{w_0}w_0 - (-1)^{|S|}e) \\ &= (\underline{q}^{w_0}w_0 + (-1)^{|S|}e)\left[\sum_{J \subset S} (-1)^{|J|} P_{\underline{q}}(W_J)^{-1}\right] P_{\underline{q}}(W). \end{aligned} \tag{10}$$

Since by assumption $(\underline{q}^{w_0})^2 \neq 1$ thus thus $P_{\underline{q}}(W)$ is left invertible, and also right invertible as it is a finite dimensional operator.

The converse implication will be shown in Theorem 3. To this end we need more information about parabolic subgroups of (W, S). A Coxeter group (W, S) is called irreducible if the Coxeter graph is connected; i.e. the set S cannot be decomposed into two disjoint subsets S_1 and S_2 commuting with each other.

The following result is interesting for its own sake.

Theorem 2. *If (W, S) is a finite irreducible Coxeter group, and J, J' are proper subsets of S, then*

$$w_0 W_J \cap W_{J'} = \emptyset \ .$$

We start with a lemma.

Lemma 2. *Let $\{s_1, s_2, \ldots, s_n\} = S$ be such that $\{s_1, s_2, \ldots, s_k\}$ is connected for any $1 \leq k \leq n$. Then the root $r = s_n s_{n-1} \ldots s_2(r_1)$ is positive and*

$$r = \sum_{i=1}^{n} \alpha_i r_i, \quad \alpha_i > 0.$$

Proof. We use induction on n. Assume

$$s_{n-1}\dots s_2(r_1)=\sum_{i=1}^{n-1}\alpha_i r_i \quad \alpha_i>0.$$

Then

$$\begin{aligned} r = s_n s_{n-1}\dots s_2(r_1) &= s_n\left(\sum_{i=1}^{n-1}\alpha_i r_i\right) \\ &= \sum_{i=1}^{n-1}\alpha_i r_i - 2\sum_{i=1}^{n-1}\alpha_i\frac{(r_i,r_n)}{(r_n,r_n)}r_n. \end{aligned}$$

We have $(r_i,r_n)\leq 0$ (see [20]) and $(r_i,r_n)<0$ for at least one value of i, where $i=1,2,\dots,n-1$. Thus the lemma follows.

We return to the proof of Theorem 2. We will use the fact that w_0 sends all positive roots to negative roots. This implies

$$\Pi\cap(w_0w)^{-1}(-\Pi)=\Pi\cap w^{-1}\Pi.$$

Hence

$$\Pi=[\Pi\cap w^{-1}(-\Pi)]\cup[\Pi\cap(w_0w)^{-1}(-\Pi)]. \tag{11}$$

Consider the positive root constructed in Lemma 2. If r belongs to the first summand in (11) then it is a linear combination of those r_i, for which $s_i\in J$ (see Introduction). This is a contradiction. Similarly r cannot belong to the second summand, if $w_0w\in W_{J'}$. □

Remark. Theorem 2 does not hold if the group (W,S) is not irreducible. For example we can take $W=\mathbb{Z}_2\times\mathbb{Z}_2$ and $S=\{s_1,s_2\}$. Let $W_J=\{e,s_1\}$ and $W_{J'}=\{e,s_2\}$. Then $w_0=s_1s_2$ and

$$w_0W_J\cap W_{J'}=\{s_2\}.$$

Theorem 3. *Let (W,S) be a finite Coxeter group. Then the convolution operator $P_{\underline{q}}(W)$ is invertible if and only if $\underline{q}\notin T(S)$.*

Proof. It suffices to consider the case when the system (W,S) is irreducible. The "if" part has been shown in Proposition 3. We only need to show that if $P_{\underline{q}}(W)$ is invertible, then $\underline{q}\notin T(S)$. Assume for a contradiction that $\underline{q}\in T(S)$. Then there exists a connected subset $J\subset S$, such that $(\underline{q}^{w_0(J)})^2=1$. We will show that $P_{\underline{q}}(W_J)$ is not invertible, which implies that $P_{\underline{q}}(W)$ is not so either, in view of (9). Thus with no loss of generality we may assume that $J=S$, i.e. $(\underline{q}^{w_0})^2=1$. By (10) we get

$$(\underline{q}^{w_0}w_0+(-1)^{|S|}e)\left[\sum_{J\subset S}(-1)^{|J|}P_{\underline{q}}(W_J)^{-1}\right]P_{\underline{q}}(W)=0.$$

Since $P_{\underline{q}}(W)$ is invertible we have

$$(\underline{q}^{w_0} w_0 + (-1)^{|S|} e) \left[\sum_{J \subset S} (-1)^{|J|} P_{\underline{q}}(W_J)^{-1} \right] = 0. \tag{12}$$

We will show that (12) is impossible. The support of $P_{\underline{q}}(W_J)^{-1}$ is a subset of W_J. This is based on the general fact that the convolution inverse to a function f on a group is supported by a subgroup generated by the support of f. Therefore

$$\mathrm{supp}(f) = G \subset \bigcup_{J \subset S} W_J \ .$$

The equation (12) implies $w_0 G \cap G \neq \emptyset$, which yields

$$w_0 W_J \cap W_{J'} \neq \emptyset$$

for some $J, J' \subset S$. In view of Theorem 2, this is a contradiction.

Remark. In view of Theorems 1 and 3 we get that for every choice of $q_s \in (-1, 1)$, where $s \in S$, the convolution operator with $P_{\underline{q}}(W)$ is strictly positive definite; i.e.

$$(P_{\underline{q}}(W) * f, f) \geq c(f, f),$$

for some $c > 0$, where $(f, g) = \sum_{w \in W} f(w) g(w)$.

If (W, S) has the property that all generators in S are conjugate, then $q_s = q$ does not depend on s, and $\underline{q}^w = q^{\ell(w)}$.

Corollary 1. *Let $W = S_n$ be the permutation group generated by the set of transpositions $S = \{(i, i+1) \mid i = 1, 2, \ldots, n\}$. The convolution operator with the Poincaré polynomial*

$$P_q(S_n) = \sum_{g \in S_n} q^{\ell(g)} g$$

is invertible if and only if

$$q^{\frac{i(i-1)}{2}} \neq 1 \quad \textit{for every } i = 1, 2, \ldots, n.$$

3 Canonical representations associated with length

3.1 Irreducible root systems and invariant kernels

Let Φ be a root system in an Euclidean space V (see [20, Section 1.2]). Any vector x in Φ defines the reflection s_x according to (1). The group generated by all reflections s_x will be denoted by W. The elements of W permute the vectors in Φ. We will assume that the system Φ is irreducible; i.e. it cannot be decomposed into two nonempty root systems which are orthogonal to each other with respect to the inner product in V.

Vectors in Φ may have different lengths. For a real number λ let Φ_λ consist of all roots in Φ of length λ, and let V_λ be the linear span of Φ_λ. Since the elements of W act on V by isometries they permute the vectors in Φ_λ.

Lemma 3. *Let* $k_\lambda = \dim V_\lambda$ *and* $d_\lambda = |\Phi_\lambda|$. *Then*

$$\sum_{z\in\Phi_\lambda}(x,z)(y,z) = \lambda^2\frac{d_\lambda}{k_\lambda}(x,y)$$

for any $x, y \in V_\lambda$.

Proof. It suffices to prove the formula for $x, y \in \Phi_\lambda$ since both its sides represent bilinear forms on V_λ. We define a new inner product in V_λ as follows.

$$[x,y] = \sum_{z\in\Phi_\lambda}(x,z)(y,z).$$

Any element $g \in W$ permute the roots in Φ_λ, hence

$$[gx, gy] = [x, y] \qquad \text{for any} \qquad x, y \in V_\lambda,\ g \in W.$$

By [2, Prop V.1.11] the group W acts transitively on Φ_λ, beacuse the system Φ is irreducible. This implies

$$[x, x] = [y, y] \qquad \text{for any} \qquad x, y \in \Phi_\lambda.$$

Let $c = \lambda^{-2}[x, x]$ for $x \in \Phi_\lambda$. Then since $\lambda^2 = (x, x)$ we get

$$\begin{aligned}[x,y] &= [s_x x, s_x y] = [-x, y - 2\frac{(x,y)}{(x,x)}x] \\ &= -[x,y] + 2c(x,y),\end{aligned}$$

for $x, y \in \Phi_\lambda$. Thus

$$[x, y] = c(x, y) \qquad \text{for} \qquad x, y \in \Phi_\lambda. \tag{13}$$

The equality extends linearly to all $x, y \in V_\lambda$.

It suffices to determine the constant c. Let $\{e_i\}_{i=1}^{k_\lambda}$ be an orthonormal basis for V_λ relative the inner product $(\,\cdot,\ \cdot)$. Then by (13) we have

$$\begin{aligned}ck_\lambda &= \sum_{i=1}^{k_\lambda} c(e_i, e_i) = \sum_{i=1}^{k_\lambda}[e_i, e_i] \\ &= \sum_{z\in\Phi_\lambda}\sum_{i=1}^{k_\lambda}(e_i, z)^2 = \sum_{z\in\Phi_\lambda}(z, z) = \lambda^2 d_\lambda.\end{aligned}$$

Let $\mathcal{H}_\lambda$ be the linear space of all real valued functions on Φ_λ with the property $f(-x) = -f(x)$, for $x \in \Phi_\lambda$. We endow $\mathcal{H}_\lambda$ with the inner product

$$(f_1, f_2)_{\mathcal{H}} = \sum_{x \in \Phi_\lambda} f_1(x) f_2(x). \tag{14}$$

The action

$$\pi_\lambda(g) f(x) = f(g^{-1}x)$$

yields a unitary representation of W on $\mathcal{H}_\lambda$. We are going to decompose this representation into irreducible components. In the next subsection we will also study its matrix coefficients.

By Schur's lemma the problem of decomposition of a given unitary representation can be reduced to studying linear selfadjoint operators commuting with this representation. Any selfadjoint operator K on $\mathcal{H}_\lambda$ commuting with the action of W on $\mathcal{H}_\lambda$ corresponds to a real valued matrix $k(x,y)$, $x, y \in \Phi_\lambda$, such that

$$k(gx, gy) = k(x,y), \tag{15}$$
$$k(x,y) = k(y,x), \tag{16}$$
$$k(-x,y) = -k(x,y). \tag{17}$$

The correspondence is given by

$$k(x,y) = (Ke_x, e_y),$$

where

$$e_v(w) = 2^{-1/2}(\delta_{v,w} - \delta_{-v,w}), \qquad v, w \in \mathcal{H}_\lambda,$$

where

$$\delta_{v,w} = \begin{cases} 1 & \text{if } v = w, \\ 0 & \text{if } v \neq w. \end{cases}$$

From now on we assume that the root system Φ is crystallographic; i.e. the quantity

$$\mathrm{n}(x,y) = 2\frac{(x,y)}{(y,y)}$$

takes only integer values. This property implies (see [2, Proposition VI.1.12]) that the roots can take two different lengths at most. Moreover the following lemma will be very useful (see [2, Section VI.1.3]).

Lemma 4 ([2]). *Let $x, y \in \Phi_\lambda$. Then $\mathrm{n}(x,y)$ can take only the values 0, ± 1 and ± 2. In particular $\mathrm{n}(x,y) = \pm 2$ if and only if $x = \pm y$.*

Proposition 4. *Let a matrix $k(x,y)$ satisfy (15), (16) and (17). Then there exist real constants α and β such that*

$$k(x,y) = \alpha(x,y) + \beta\{\delta_{x,y} - \delta_{-x,y}\}.$$

Proof. First we will show that the value $k(x,y)$ depends only on (x,y) or equivalently on $\mathrm{n}(x,y)$. By (17) we may restrict ourselves to the case $n(x,y) \geq 0$. We have three possibilities.

(1) $\mathrm{n}(x,y)=0$.

Then $s_x y = y$ and

$$-k(x,y) = k(-x,y) = k(s_x x, s_x y) = k(x,y).$$

Thus $k(x,y)=0$.

(2) $n(x,y)=2$.

Thus by Lemma 2 we get $x=y$. Therefore it suffices to make sure that $k(x,x)$ does not depend on x. This holds true because W acts transitively on Φ_λ (see [2, Prop V.1.11]) and $k(gx,gx)=k(x,x)$ for $g \in W$.

(3) $\mathrm{n}(x,y)=1$.

Let $\mathrm{n}(x',y')=1$. We have to show that $k(x',y')=k(x,y)$. In view of the properties of $k(x,y)$ and $\mathrm{n}(x,y)$ satisfy it suffices to show that there exists $w \in W$ such that

$$wx = \pm x' \quad \text{and} \quad wy = \pm y'.$$

Firstly, there exists g such that $gx=x'$. Let $y''=gy$. If $y''=y'$ we are done. Thus we may assume $y'' \neq y'$. We have

$$\mathrm{n}(x',y'') = \mathrm{n}(gx,gy) = \mathrm{n}(x,y) = \mathrm{n}(x',y') = 1.$$

This implies $y'' \neq -y'$. Thus $y'' \neq \pm y'$. We will further break the resaoning into three subcases.

(3a) $\mathrm{n}(y',y'')=1$.

Letting $h = s_{y''} s_{y'} s_{y''}$ gives

$$h = s_{y'-y''}, \qquad hx' = x', \qquad hy' = y''.$$

Let $w=hg$. Then $wx = hx' = x'$ and $wy = hy'' = y'$.

(3b) $n(y',y'')=0$.

Letting $h = s_{y'} s_{x'} s_{y''} s_{x'} s_{y'} s_{x'}$ gives

$$h = s_{y'+y''-x'} s_{x'}, \qquad hx' = -x', \qquad hy'' = -y'.$$

Let $w=hg$. Then $wx = hx' = -x'$ and $wy = hy'' = -y'$.

(3c) $n(y',y'')=-1$.

We have

$$\mathrm{n}(s_{x'}y'', y') = \mathrm{n}(y''-x', y') = \mathrm{n}(y'',y') - \mathrm{n}(x',y') = -2.$$

In view of Lemma 2 this implies $s_{x'}y'' = -y'$. Let $w = s_{x'}g$. Then $wx = s_{x'}x' = -x'$ and $wy = s_{x'}y'' = -y'$.

We are now in position to determine the constants α and β. Assume $x \not\perp y$, $x \neq \pm y$ and let $\alpha = (x,y)^{-1}k(x,y)$, if such a pair exists, and $\alpha = 0$, otherwise. Set $\beta = k(x,x) - \alpha(x,x)$. The lemma is then satisfied.

Proposition 1 implies that the space of operators commuting with action of W on V_λ is at most two dimensional. Hence the representation π_λ can be decomposed into two irreducible subrepresentations or it is irreducible itself. The latter holds if and only if the positive roots in Φ_λ are mutually orthogonal.

Let P_λ be defined on $\mathcal{H}_\lambda$ by

$$P_\lambda f(x) = \frac{k_\lambda}{\lambda^2 d_\lambda} \sum_{z \in \Phi_\lambda} (x,z) f(z). \tag{18}$$

Lemma 3 implies that the operator P_λ is a projection. Clearly it commutes with the action of W. Let us determine the subspace $P_\lambda \mathcal{H}_\lambda$.

Lemma 5. $P_\lambda \mathcal{H}_\lambda = \{f \in \mathcal{H}_\lambda \mid f(\,\cdot\,) = (\,\cdot\,,\ w)$ *for some* $w \in V_\lambda\}$.

Proof. For $f \in \mathcal{H}_\lambda$ we have

$$P_\lambda f(x) = (x,w), \quad \text{where} \quad w = \frac{k_\lambda}{\lambda^2 d_\lambda} \sum_{z \in \Phi_\lambda} f(z)\, z.$$

Conversely, if $f(x) = (x,w)$ for some $w \in V_\lambda$, then by Lemma 3

$$P_\lambda f(x) = \frac{k_\lambda}{\lambda^2 d_\lambda} \sum_{z \in \Phi_\lambda} (x,z)(z) = (x,w) = f(x).$$

$\square$

Lemma 3 implies that $P_\lambda \mathcal{H}_\lambda$ is isomorphic to V_λ and the representation π_λ restricted to $P_\lambda \mathcal{H}_\lambda$ is equivalent to the action of W on V_λ. If Φ_λ contains two roots which are not parallel nor perpendicular, then $P_\lambda \neq I$ and $\mathcal{H}_\lambda$ has nontrivial decomposition

$$\mathcal{H}_\lambda = (I - P_\lambda)\mathcal{H}_\lambda \oplus P_\lambda \mathcal{H}_\lambda.$$

Thus we arrived at the following.

Theorem 4. *The representation π_λ of W on $\mathcal{H}_\lambda$ has two irreducible subspaces $P_\lambda\mathcal{H}_\lambda$ and $(I - P_\lambda)\mathcal{H}_\lambda$, provided that Φ_λ contains two roots x and y such that $x \neq \pm y$ and $x \not\perp y$. Otherwise the representation π_λ is irreducible itself. The representation π_λ restricted to $P_\lambda\mathcal{H}_\lambda$ is equivalent to the action of W on V_λ.*

3.2 Matrix coefficients of π_λ

For an irreducible root system Φ in V let Δ denote the set of simple roots. Φ decomposes into two disjoint subset Π and $-\Pi$ consisting of positive and negative roots, respectively. The group W is generated by the reflections s_x, where $x \in \Delta$. As was mentioned in the Introduction the algebraic length of elements $g \in W$ with respect to the generators s_x, $x \in \Delta$ can be expressed in geometric manner as

$$\ell(g) = |g\Pi \cap -\Pi|. \tag{19}$$

This formula has been used in [4] to show that the correspondence $g \mapsto \ell(g)$ is a negative definite function on W. We will extend this result in the case of finite Coxeter groups by showing that the function $g \mapsto \frac{1}{2}|\Pi| - \ell(g)$ is positive definite. We will also give a decomposition of it into pure positive definite functions; i.e. positive definite functions which are coefficients of irreducible representations.

Let $\mathcal{H}$ consists of all real valued functions on Φ such that $f(-x) = -f(x)$, for $x \in \Phi$, endowed with the inner product (14). The group W acts on $\mathcal{H}$ by isometries

$$\pi(g)f(x) = f(g^{-1}x).$$

Obviously the representations π_λ defined in Section 2 are subreprepresentations of π.

Proposition 5. *Let $\xi = \frac{1}{2}(\chi_\Pi - \chi_{-\Pi})$, where χ_A denotes the indicator function of a set A. Then*

$$(\pi(g)\xi, \xi)_{\mathcal{H}} = \tfrac{1}{2}|\Pi| - \ell(g).$$

In particular the mapping $g \mapsto \frac{1}{2}|\Pi| - \ell(g)$, is positive definite function on W.

Proof. We have

$$\begin{aligned} 4(\pi(g)\xi, \xi)_{\mathcal{H}} = &|g\Pi \cap \Pi| + |g(-\Pi) \cap (-\Pi)| \\ &-|g\Pi \cap (-\Pi)| - |g(-\Pi) \cap \Pi|. \end{aligned}$$

We have

$$\begin{aligned} |g\Pi \cap \Pi| &= |g(-\Pi) \cap (-\Pi)| = |\Pi| - |g\Pi \cap (-\Pi| = |\Pi| - \ell(g), \\ |g(-\Pi) \cap \Pi| &= |g\Pi \cap (-\Pi)| = \ell(g). \end{aligned}$$

This gives the conclusion.

The set Φ_λ decomposes in natural way into the subsets Π_λ and $-\Pi_\lambda$ of positive and negative, respectively, roots of length λ. Let us define the length function on W relative to λ as

$$\ell_\lambda(g) = |g\Pi_\lambda \cap -\Pi_\lambda|.$$

Similarly to Proposition 2 we get the following.

Proposition 6. *Let* $\xi_\lambda = \frac{1}{2}(\chi_{\Pi_\lambda} - \chi_{-\Pi_\lambda})$. *Then*

$$(\pi_\lambda(g)\xi_\lambda, \xi_\lambda)_{\mathcal{H}} = \frac{1}{2}|\Pi_\lambda| - \ell_\lambda(g).$$

In particular the mapping $g \mapsto \frac{1}{2}|\Pi_\lambda| - \ell_\lambda(g)$, *is positive definite function on* W.

Remark. By [2, Proposition V.1.11] if two simple roots r_i and r_j have equal length then there exists $w \in W$ such that $wr_i = r_j$. This implies $ws_iw^{-1} = s_j$; i.e. the reflections s_i and s_j are conjugate to each other. The converse is also true. In this way the set Δ_λ corresponds to a conjugate class A_i in (6). Thus the length ℓ_λ coincides with one of the functions ℓ_i defined in Proposition 1.

Using Schur's theorem and power series expansion of e^x gives that the exponent of a positive function is again such function. Thus Proposition 6 implies that for every choice of $0 < q < 1$ the correspondence

$$w \mapsto q^{\ell_\lambda(w)}$$

is positive definite. This gives an alternate proof of Theorem 1 in the case of finite reflection groups.

From now on we assume that the root system Φ is crystallographic. By [2] Φ can have roots of at most two different lengths. If $\Phi = \Phi_\lambda \cup \Phi_{\lambda'}$ then $\ell(g) = \ell_\lambda(g) + \ell_{\lambda'}(g)$ and $\pi = \pi_\lambda \oplus \pi_{\lambda'}$. If $\Phi = \Phi_\lambda$, then of course $\ell(g) = \ell_\lambda(g)$ and $\pi = \pi_\lambda$. Thus $\pi = \pi_\lambda$ or it decomposes into representations π_λ and $\pi_{\lambda'}$. In Section 3.1 we solved the problem of decomposing the representation π_λ. Now we will compute the corresponding decomposition of the positive definite functions

$$g \mapsto \tfrac{1}{2}|\Pi_\lambda| - \ell_\lambda(g).$$

Lemma 6. *Let* v_λ *be the sum of all positive roots of length* λ; *i.e.*

$$v_\lambda = \sum_{x \in \Pi_\lambda} x.$$

Then

$$(r, v_\lambda) = \begin{cases} 0 & r \in \Delta \setminus \Delta_\lambda, \\ \lambda^2 & r \in \Delta_\lambda, \end{cases} \tag{20}$$

where Δ_λ *is the set of simple roots of length* λ.

Proof. Let $r \in \Delta \setminus \Delta_\lambda$. Then $s_r v_\lambda = v_\lambda$, because s_r permutes the roots in Π_λ. Hence

$$(r, v_\lambda) = (r, s_r v_\lambda) = (s_r r, v_\lambda) = -(r, v_\lambda).$$

Thus $(r, v_\lambda) = 0$.

Let $r \in \Delta_\lambda$. Then $s_r v_\lambda = v_\lambda - 2r$, beacuse $s_r r = -r$. Hence

$$(r, v_\lambda) = (s_r r, s_r v_\lambda) = -(r, v_\lambda) + 2(r, r).$$

Therefore

$$(r, v_\lambda) = (r, r) = \lambda^2.$$

Any root x in Φ can be uniquely represented as a linear combination of simple roots

$$x = \sum_{r \in \Delta} \alpha_r(x) r. \tag{21}$$

The coefficients $\alpha_r(x)$, $r \in \Delta$, are all nonnegative or all nonpositive according to whether the root x is positive or negative. Define the function $\mathrm{n}_\lambda(x)$ for $x \in \Phi_\lambda$ by

$$\mathrm{n}_\lambda(x) = \sum_{r \in \Delta_\lambda} \alpha_r(x),$$

where $\Delta_\lambda = \{r \in \Delta \mid \|r\| = \lambda\}$.

Lemma 7. *For $x \in \Phi_\lambda$ we have*

$$\sum_{z \in \Pi_\lambda} (x, z) = \lambda^2 \mathrm{n}_\lambda(x).$$

Proof. By Lemma 4 and by (21) we have

$$\begin{aligned} \sum_{z \in \Pi_\lambda} (x, z) = (x, v_\lambda) &= \sum_{r \in \Delta} \alpha_r(x)(r, v_\lambda) \\ &= \sum_{r \in \Delta_\lambda} \lambda^2 \alpha_r(x) = \lambda^2 \mathrm{n}_\lambda(x). \end{aligned}$$

Theorem 5. *The mappings*

$$\varphi_\lambda(g) = \mathrm{n}_\lambda(g v_\lambda) \tag{22}$$

$$\psi_\lambda(g) = \tfrac{1}{4} d_\lambda - \ell_\lambda(g) - \frac{k_\lambda}{d_\lambda} \mathrm{n}_\lambda(g v_\lambda) \tag{23}$$

are pure positive definite functions on W, where $d_\lambda = |\Phi_\lambda|$ and $k_\lambda = \dim V_\lambda$.

Proof. We will show that

$$(\pi_\lambda(g) P_\lambda \xi_\lambda, P_\lambda \xi_\lambda) = \frac{k_\lambda}{\lambda^2 d_\lambda} \mathrm{n}_\lambda(g v_\lambda)$$

$$(\pi_\lambda(g)(I - P_\lambda)\xi_\lambda, (I - P_\lambda)\xi_\lambda) = \tfrac{1}{4} d_\lambda - \ell_\lambda(g) - \frac{k_\lambda}{\lambda^2 d_\lambda} \mathrm{n}_\lambda(g v_\lambda).$$

This will yield that both functions are positive definite. Since the restrictions of π_λ to either $P_\lambda \mathcal{H}_\lambda$ or $(I - P_\lambda)\mathcal{H}_\lambda$ are irreducible, the above functions are pure positive definite functions on W.

By definition of P_λ we have

$$P_\lambda \delta_x(y) = \frac{k_\lambda}{\lambda^2 d_\lambda}(x, y).$$

Combining this, Lemma 5 and the fact that P_λ is an orthogonal projection commuting with $\pi_\lambda(g)$ gives

$$\begin{aligned}(\pi_\lambda(g)P_\lambda v_\lambda, P_\lambda v_\lambda) = (\pi_\lambda(g)P_\lambda v_\lambda, v_\lambda) &= \frac{1}{4}\sum_{x,y\in\Pi_\lambda}(P_\lambda\{\delta_{gx} - \delta_{-gx}\}, \delta_y - \delta_{-y}) \\ &= \frac{k_\lambda}{\lambda^2 d_\lambda}\sum_{x,y\in\Pi_\lambda}(gx, y) = \frac{k_\lambda}{\lambda^2 d_\lambda}\sum_{x\in\Pi_\lambda}\mathrm{n}_\lambda(gx) \\ &= \frac{k_\lambda}{\lambda^2 d_\lambda}\mathrm{n}_\lambda(gv_\lambda).\end{aligned}$$

Furthermore, since $I - P_\lambda$ is a projection commuting with $\pi_\lambda(g)$, we obtain

$$\begin{aligned}(\pi_\lambda(g)(I - P_\lambda)\xi_\lambda, (I - P_\lambda)\xi_\lambda) &= (\pi_\lambda(g)(I - P_\lambda)\xi_\lambda, \xi_\lambda) \\ &= (\pi_\lambda(g)\xi_\lambda, \xi_\lambda) - (\pi_\lambda(g)P_\lambda\xi_\lambda, \xi_\lambda) \\ &= \tfrac{1}{4}d_\lambda - \ell_\lambda(g) - \frac{k_\lambda}{\lambda^2 d_\lambda}\mathrm{n}_\lambda(gv_\lambda).\end{aligned}$$

3.3 Examples

(a) Groups of type A_n.

We have

$$V = \{x \in \mathbb{R}^{n+1} \mid \sum_{i=1}^{n+1} x_i = 0\} \qquad \text{and} \qquad \Phi = \{e_i - e_j \mid i \neq j\}.$$

The roots have equal length $\lambda = \sqrt{2}$. The positive roots and simple roots are as follows.

$$\begin{aligned}\Pi &= \{e_i - e_j \mid i < j\}, \\ \Delta &= \{e_1 - e_2,\ e_2 - e_3,\ \ldots,\ e_n - e_{n+1}\}.\end{aligned}$$

We also have

$$\begin{aligned}d_\lambda &= n(n+1), \\ k_\lambda &= n, \\ v_\lambda &= \sum_{i<j}(e_i - e_j), \\ \mathrm{n}_\lambda(e_i - e_j) &= j - i.\end{aligned}$$

The group W can be identified with the permutation group S_{n+1}. In this case we have

$$\varphi_\lambda(\sigma) = \sum_{i<j}[\sigma(j) - \sigma(i)]$$
$$\psi_\lambda(\sigma) = \frac{n(n+1)}{4} - \ell(\sigma) - \frac{1}{2(n+1)}\sum_{i<j}[\sigma(j) - \sigma(i)]$$

By Theorem 5 both functions are pure positive definite functions on S_{n+1}.

It can be easily determined that the irreducible representation of S_{n+1} corresponding to the function φ_λ has the diagram $(2, 1, 1, \ldots, 1)$ int he Young tableaux, while the representation corresponding to ψ_λ has the diagram $(3, 1, 1, \ldots, 1)$.

(b) Groups of type B_n.

In that case $V = \mathbb{R}^n$ and

$$\Phi = \{\pm e_i \pm e_j \mid i \neq j = 1, 2, \ldots, n\} \cup \{\pm e_i \mid i = 1, 2, \ldots, n\}.$$

The positive roots and simple roots are

$$\Pi = \{e_i \pm e_j \mid i < j\} \ \cup \ \{e_i \mid i = 1, 2, \ldots, n\},$$
$$\Delta = \{e_1 - e_2,\ e_2 - e_3,\ \ldots,\ e_{n-1} - e_n, e_n\}.$$

There are two lengths 1 and $\sqrt{2}$. Let $\lambda = \sqrt{2}$. Then

$$\Pi_\lambda = \{e_i \pm e_j \mid i < j\},$$
$$\Delta_\lambda = \{e_1 - e_2,\ e_2 - e_3,\ \ldots,\ e_{n-1} - e_n\}.$$

We thus have

$$\begin{aligned} d_\lambda &= 2n^2, \\ k_\lambda &= n, \\ v_\lambda &= 2\sum_{i=1}^{n-1}(n-i)e_i, \\ \mathrm{n}_\lambda(e_i) &= n - i. \end{aligned}$$

The group W can be identified with $\mathbb{Z}_2^n \rtimes S_n$. An element (τ, σ), where $\tau \in \mathbb{Z}_2^n$ and $\sigma \in S_n$, acts on V by the rule

$$(\tau, \sigma)e_i = (-1)^{\tau(i)} e_{\sigma(i)}.$$

We have

$$\varphi_\lambda(\tau,\sigma) = 2\sum_{i=1}^{n-1}(-1)^{\tau(i)}(n-i)[n-\sigma(i)]$$
$$\psi_\lambda(\tau,\sigma) = \frac{n^2}{2} - \ell(\tau,\sigma) - \frac{1}{2n}\sum_{i=1}^{n-1}(-1)^{\tau(i)}(n-i)[n-\sigma(i)].$$

(c) Groups of type D_n.

In that case $V = \mathbb{R}^n$ and

$$\Phi = \{\pm e_i \pm e_j \mid i \neq j = 1, 2, \ldots, n\}.$$

The roots have equal length $\lambda = \sqrt{2}$. The positive roots and simple roots are

$$\Pi = \{e_i \pm e_j \mid i < j\},$$
$$\Delta = \{e_1 - e_2,\ e_2 - e_3,\ \ldots,\ e_{n-1} - e_n, e_{n-1} + e_n\}.$$

We have

$$d_\lambda = 2n(n-1),$$
$$k_\lambda = n,$$
$$v_\lambda = 2\sum_{i=1}^{n-1}(n-i)e_i,$$
$$\mathrm{n}_\lambda(e_i) = n - i.$$

The group W can be identified with the semidirect product of S_n and the subgroup A of $\mathbb{Z}_2^n$ consisting of $\tau = (\tau(1), \tau(2), \ldots, \tau(n))$ such that $\sum \tau(i)$ is an even number. The subgroup A is normal in W. Then

$$\varphi_\lambda(\tau,\sigma) = 2\sum_{i=1}^{n-1}(-1)^{\tau(i)}(n-i)[n-\sigma(i)]$$
$$\psi_\lambda(\tau,\sigma) = \frac{n(n-1)}{2} - \ell(\tau,\sigma) - \frac{1}{2(n-1)}\sum_{i=1}^{n-1}(-1)^{\tau(i)}(n-i)[n-\sigma(i)].$$

References

1. Berg, C., Christensen, J. P.: Harmonic Analysis on Semigroups: Theory of Positive Definite Functions. Graduate Texts in Mathematics, **100**, Springer Verlag, New York, (1984)
2. Bourbaki, N.: Groupes et algèbres de Lie, Ch. 4–6. Hermann, Paris, (1968)
3. Bożejko, M.: Positive and negative kernels on discrete groups. Lectures at Heidelberg University, (1987)

4. Bożejko, M., Januszkiewicz, T., Spatzier, R. J.: Infinite Coxeter groups do not have Kazhdan's property. J. Operator Theory, **19**, 63–68 (1988)
5. Bożejko, M., Picardello, M. A.: Weakly amenable groups and amalgamated products. Proc. Amer. Math. Soc., **117**, 1039–1046 (1993)
6. Bożejko, M., Speicher, R.: An example of a generalized Brownian motion. Comm. Math. Phys., **137**, 519–531 (1991)
7. Bożejko, M., Speicher, R.: Completely positive maps on Coxeter groups, deformed commutation relations and operator spaces. Math. Ann., **300**, 97–120 (1994)
8. Bożejko, M., Speicher, R.: Interpolation between bosonic and fermionic relations given by generalized Brownian motions. Math. Z., **222**, 135–160 (1996)
9. Bożejko, M., Kümerer, B., Speicher, R.: q-Gaussian processes: nonclassical and classical aspects. Comm. Math. Phys. (to appear)
10. Carter, R.: Simple Groups of Lie Type. Wiley–Interscience, London (1972)
11. Coxeter, H. S. M.: Regular Complex Polytopes. Cambridge University Press, Cambridge, (1974)
12. Coxeter, H. S. M.: Discrete groups generated by reflections. Annals of Math., **35**, 588–621 (1934)
13. Coxeter, H. S. M., Moser, W. O.: Generators and Relations for Discrete Groups, 2nd ed. Springer–Verlag, New York, (1965)
14. de Canniere, J., Haagerup, U.: Multipliers of the Fourier algebra of some simple Lie groups and their discrete subgroups. Amer. Math. Soc., **107**, 455–500 (1983)
15. de la Harpe, P.: Groupes de Coxeter infinites non affines. Expo. Math., **5**, 91–96 (1987)
16. Dykema, K., Nica, A.: On the Fock representation of the q-commutation relations. J. Reine Angew. Math., **440**, 201–212 (1993)
17. Greenberg, O. W.: Particles with small violations of Fermi or Bose statistics. Phys. Rev. D, **43**, 4111–4120 (1991)
18. Grove, L. C., Benson, C. T.: Finite Reflection Groups, 2nd ed. Springer-Verlag, New York (1985)
19. Hiller, H.: Geometry of Coxeter Groups. Pitman, London, (1982)
20. Humphreys, J. E.: Reflection groups and Coxeter groups. Cambridge University Press, Cambridge, (1990)
21. Januszkiewicz, T.: For right angled Coxeter groups $z^{|g|}$ is a coefficient of uniformly bounded representation. Proc. Amer. Math. Soc., **119**, 1115–1119 (1993)
22. Jorgensen, P. E. T., Schmitt, L M., Werner, R. F.: q-canonical commutation relations and stability of the Cuntz algebra. Pacific Math. J., **165**, 131–151 (1994)
23. Pytlik, T., Szwarc, R.: An analytic series of uniformly bounded representations of free groups. Acta Math., **157**, 287–309 (1986)
24. Szwarc, R.: Groups acting on trees and approximation properties of the Fourier algebra, J. Funct. Anal., **95**, 320–343 (1991)
25. Szwarc, R.: Structure of geodesics in the Cayley graph of infinite Coxeter groups. Preprint (1996)
26. Tits, J.: Buildings of Spherical Type and Finite BN–pairs. Lecture Notes in Math. **386**, Springer-Verlag, (1974)
27. Valette, A.: Weak amenability of right-angled Coxeter groups. Proc. Amer. Math. Soc., **119**, 1331–1334 (1993)
28. Zagier, D.: Realizabilty of a model in infinite statistics. Comm. Math. Phys., **147**, 199–210 (1992)

Mixed hook-length formula for degenerate affine Hecke algebras

Maxim Nazarov

Department of Mathematics
University of York
York YO10 5DD
England
mln1@york.ac.uk

Summary. Take the degenerate affine Hecke algebra H_{l+m} corresponding to the group GL_{l+m} over a p-adic field. Consider the H_{l+m}-module W induced from the tensor product of the evaluation modules over the algebras H_l and H_m. The module W depends on two partitions λ of l and μ of m, and on two complex numbers. There is a canonical operator J acting in W, it corresponds to the Yang R-matrix. The algebra H_{l+m} contains the symmetric group algebra $\mathbb{C}S_{l+m}$ as a subalgebra, and J commutes with the action of this subalgebra in W. Under this action, W decomposes into irreducible subspaces according to the Littlewood–Richardson rule. We compute the eigenvalues of J, corresponding to certain multiplicity-free irreducible components of W. In particular, we give a formula for the ratio of two eigenvalues of J, corresponding to the maximal and minimal irreducible components. As an application of our results, we derive the well-known hook-length formula for the dimension of the irreducible $\mathbb{C}S_l$-module corresponding to λ.

In this article we work with the degenerate affine Hecke algebra H_l corresponding to the general linear group GL_l over a local non-Archimedean field. This algebra was introduced by V. Drinfeld in [3], see also [8]. The complex associative algebra H_l is generated by the symmetric group algebra $\mathbb{C}S_l$ and by the pairwise commuting elements $x_1, \ldots, x_l$ with the cross relations for $p = 1, \ldots, l-1$ and $q = 1, \ldots, l$

$$\begin{aligned} \sigma_{p,p+1}\, x_q &= x_q\, \sigma_{p,p+1}\,, \quad q \neq p, p+1\,; \\ \sigma_{p,p+1}\, x_p &= x_{p+1}\, \sigma_{p,p+1} - 1\,. \end{aligned}$$

Here and in what follows $\sigma_{pq} \in S_l$ denotes the transposition of p and q.

For any partition $\lambda = (\lambda_1, \lambda_2, \ldots)$ of l let V_λ be the corresponding irreducible $\mathbb{C}S_l$-module. There is a homomorphism $H_l \to \mathbb{C}S_l$ identical on the subalgebra $\mathbb{C}S_l \subset H_l$ such that $x_p \mapsto \sigma_{1p} + \ldots + \sigma_{p-1,p}$ for each $p = 1, \ldots, l-1$. So V_λ can be regarded as a H_l-module. For any number $z \in \mathbb{C}$ there is also an automorphism of H_l identical on the subalgebra

A. Borodin et al.: LNM 1817, A.M. Vershik (Ed.), pp. 223–236, 2003.

$\mathbb{C}S_l \subset H_l$ such that $x_p \mapsto x_p + z$ for each $p = 1, \ldots, l$. We will denote by $V_\lambda(z)$ the H_l-module obtained by pulling V_λ back through this automorphism. The module $V_\lambda(z)$ is irreducible by definition.

Consider the algebra $H_l \otimes H_m$ where both l and m are positive integers. It is isomorphic to the subalgebra in H_{l+m} generated by the transpositions σ_{pq} where $1 \leqslant p < q \leqslant l$ or $l+1 \leqslant p < q \leqslant l+m$, along with all the elements $x_1, \ldots, x_{l+m}$. For any partition μ of m and any number $w \in \mathbb{C}$ take the corresponding H_m-module $V_\mu(w)$. Now consider the H_{l+m}-module W induced from the $H_l \otimes H_m$-module $V_\lambda(z) \otimes V_\mu(w)$. Also consider the H_{l+m}-module W' induced from the $H_m \otimes H_l$-module $V_\mu(w) \otimes V_\lambda(z)$. Suppose that $z - w \notin \mathbb{Z}$, then the modules W and W' are irreducible and equivalent; see [2]. So there is a unique, up to scalar multiplier, H_{l+m}-intertwining operator $I : W \to W'$. For a certain particular realization of the modules W and W' we will give an explicit expression for the operator I. In particular, this will fix the normalization of I.

Following [2], we will realize W and W' as certain left ideals in the group algebra $\mathbb{C}S_{l+m}$. We will have $W' = W\tau$ where $\tau \in S_{l+m}$ is the permutation

$$(1, \ldots, m, m+1, \ldots, m+l) \mapsto (l+1, \ldots, l+m, 1, \ldots, l).$$

Let $J : W \to W$ be the composition of the operator I and the operator $W' \to W$ of the right multiplication by τ^{-1}. In §2 we give an explicit expression for the operator I. Using this expression makes the spectral analysis of the operator J an arduous task; cf. [1]. However, our results are based on this expression.

The subalgebra $\mathbb{C}S_{l+m} \subset H_{l+m}$ acts in $W \subset \mathbb{C}S_{l+m}$ via left multiplication. Under this action the space W splits into irreducible components according to the Littlewood–Richardson rule [9]. The operator J commutes with this action. Hence J acts via multiplication by a certain complex number in every irreducible component of W appearing with multiplicity one. In this article we compute these numbers for certain multiplicity-free components of W, see Theorems 1 and 2.

For example, there are two distinguished irreducible components of the $\mathbb{C}S_{l+m}$-module W which always have multiplicity one. They correspond to the partitions

$$\lambda + \mu = (\lambda_1 + \mu_1, \lambda_2 + \mu_2, \ldots) \quad \text{and} \quad (\lambda' + \mu')' = (\lambda'_1 + \mu'_1, \lambda'_2 + \mu'_2, \ldots)'$$

where as usual $\lambda' = (\lambda'_1, \lambda'_2, \ldots)$ denotes the conjugate partition. Denote by $h_{\lambda\mu}(z,w)$ the ratio of the corresponding two eigenvalues of J, this ratio does not depend on the normalization of the operator I. Theorems 1 and 2 have

Corollary. *We have the equality*

$$h_{\lambda\mu}(z,w) \;=\; \prod_{i,j} \frac{z - w - \lambda'_j - \mu_i + i + j - 1}{z - w + \lambda_i + \mu'_j - i - j + 1}$$

where the product is taken over all $i, j = 1, 2, \ldots$ *such that* $j \leqslant \lambda_i, \mu_i$.

We derive this corollary in §6. Now identify partitions with their Young diagrams. The condition $j \leqslant \lambda_i\,,\mu_i$ means that the box $(i\,,j)$ belongs to the intersection of the diagrams λ and μ. If $\lambda=\mu$ the numbers $\lambda_i+\lambda'_j-i-j+1$ are called [9] the *hook-lengths* of the diagram λ. If $\lambda\neq\mu$ the numbers in the above fraction

$$\lambda_i+\mu'_j-i-j+1 \quad \text{and} \quad \lambda'_j+\mu_i-i-j+1$$

may be called the *mixed hook-lengths* of the first and of the second kind respectively. Both these numbers are positive for any box $(i\,,j)$ in the intersection of λ and μ.

Let h_λ denote the product of all l hook-lengths of the Young diagram λ. This product appears in the well-known *hook-length formula* [9] for the dimension of the irreducible $\mathbb{C}S_l$-module V_λ: $\dim V_\lambda = l\,!\,/\,h_\lambda$. In §7 we derive this formula from our Theorem 2. At the end of §7 we discuss applications of our Theorems 1 and 2 to the representaton theory of affine Hecke algebras, cf. [6].

The present work arose from my conversations with B. Leclerc and J.-Y. Thibon. I am very grateful to them, and to the EPSRC for supporting their visits to the University of York. I am also grateful to S. Kumar for a useful discussion. I have been supported by the EC under the TMR grant FMRX-CT97-0100.

1 Young bases

Here we will collect several known facts about the irrdeducible $\mathbb{C}S_l$-modules. Fix the chain $S_1\subset S_2\subset\cdots\subset S_l$ of subgroups with the standard embeddings. There is a distinguished basis in the space V_λ associated with this chain, called the *Young basis*. Its vectors are labeled by the standard tableaux [9] of shape λ. For every such a tableau Λ the basis vector $v_\Lambda\in V_\lambda$ is defined, up to a scalar multiplier, as follows. For any $p=1\,,\,\ldots\,,\,l-1$ take the tableau obtained from Λ by removing each of the numbers $p+1\,,\,\ldots\,,\,l$. Let the Young diagram π be its shape. Then the vector v_Λ is contained in an irreducible $\mathbb{C}S_p$-submodule of V_λ corresponding to π. Fix an S_l-invariant inner product $\langle\;,\;\rangle$ in V_λ. The vectors v_Λ are then pairwise orthogonal. We will agree that $\langle v_\Lambda, v_\Lambda\rangle=1$ for every Λ.

There is an alternative definition [5] of the vector $v_\Lambda\in V_\lambda$. For each $p=1\,,\,\ldots\,,\,l$ put $c_p=j-i$ if the number p appears in the i-row and the j-th column of the tableau Λ. The number c_p is called the *content* of the box of the diagram λ occupied by p. Here on the left we show the column tableau of shape $\lambda=(3\,,2)$:

1	3	5
2	4	

0	1	2
-1	0	

On the right we indicated the contents of the boxes of the Young diagram $\lambda = (3,2)$. So here we have $(c_1, \ldots, c_5) = (0,-1,1,0,2)$. For each $p = 1, \ldots, l$ consider the image $\sigma_{1p} + \ldots + \sigma_{p-1,p}$ of the generator x_p under the homomorphism $H_l \to \mathbb{C}S_l$.

Proposition 1. [5] *The vector v_Λ of the $\mathbb{C}S_l$-module V_λ is determined, up to a scalar multiplier, by equations $(\sigma_{1p}+\ldots+\sigma_{p-1,p})\cdot v_\Lambda = c_p v_\Lambda$ for $p = 1, \ldots, l$.*

Take the diagonal matrix element of V_λ corresponding to the vector v_Λ

$$F_\Lambda = \sum_{\sigma \in S_l} \langle v_\Lambda, \sigma \cdot v_\Lambda \rangle \sigma .$$

As a general property of matrix elements, we have the equality $F_\Lambda^2 = l! \big/ \dim V_\lambda \cdot F_\Lambda$. There is an alternative expression for the element $F_\Lambda \in \mathbb{C}S_l$, it goes back to [2]. For any distinct $p,q = 1, \ldots, l$ introduce the rational function of two complex variables u,v

$$f_{pq}(u,v) = 1 - \sigma_{pq}/(u-v) .$$

These functions take values in the algebra $\mathbb{C}S_l$ and satisfy the relations

$$f_{pq}(u,v)\, f_{pr}(u,w)\, f_{qr}(v,w) = f_{qr}(v,w)\, f_{pr}(u,w)\, f_{pq}(u,v), \tag{1}$$

$$f_{pq}(u,v)\, f_{qp}(v,u) = 1 - (u-v)^{-2} \tag{2}$$

for all pairwise distinct indices p,q,r. Introduce l complex variables $z_1, \ldots, z_l$. Order lexicographically the pairs (p,q) with $1 \leqslant p < q \leqslant l$. Define the rational function $F_\Lambda(z_1, \ldots, z_l)$ as the ordered product of the functions $f_{pq}(z_p + c_p, z_q + c_q)$ over all the pairs (p,q). Let $\mathcal{Z}_\Lambda$ be the vector subspace in $\mathbb{C}^l$ consisting of all l-tuples $(z_1, \ldots, z_l)$ such that $z_p = z_q$, whenever the numbers p and q appear in the same row of the tableau Λ.

Proposition 2. [10] *The restriction of the rational function $F_\Lambda(z_1, \ldots, z_l)$ to $\mathcal{Z}_\Lambda$ is regular at the origin $(0, \ldots, 0) \in \mathbb{C}^l$, and takes there the value F_Λ.*

2 Intertwining operators

Let us choose any standard tableau Λ of shape λ. The H_l-module $V_\lambda(z)$ can be realized as the left ideal in $\mathbb{C}S_l$ generated by the element F_Λ. The subalgebra $\mathbb{C}S_l \subset H_l$ acts here via left multiplication. Due to Proposition 1 and to the defining relations of H_l, the action of the generators $x_1, \ldots, x_l$ in this left ideal can be then determined by setting $x_p \cdot F_\Lambda = (c_p + z) F_\Lambda$ for each $p = 1, \ldots, l$.

Also fix any standard tableau M of shape μ. Let d_q be the content of the box of the diagram μ occupied by $q = 1, \ldots, m$ in M. We will denote by $\overline{G}$ the image of any element $G \in \mathbb{C}S_m$ under the embedding $\mathbb{C}S_m \to \mathbb{C}S_{l+m}$: $\sigma_{pq} \mapsto \sigma_{l+p,l+q}$. The H_{l+m}-module W induced from the $H_l \otimes H_m$-module

$V_\lambda(z)\otimes V_\mu(w)$ can be realized as the left ideal in $\mathbb{C}S_{l+m}$ generated by the product $F_\Lambda\overline{F}_{\mathrm{M}}$. The action of the generators $x_1,\ldots,x_{l+m}$ in the latter left ideal can be determined by setting

$$\begin{aligned} x_p\cdot F_\Lambda\overline{F}_{\mathrm{M}} &= (c_p+z)\,F_\Lambda\overline{F}_{\mathrm{M}} \quad \text{for each} \quad p=1,\ldots,l\,; \\ x_{l+q}\cdot F_\Lambda\overline{F}_{\mathrm{M}} &= (d_q+w)\,F_\Lambda\overline{F}_{\mathrm{M}} \quad \text{for each} \quad q=1,\ldots,m\,. \end{aligned} \tag{3}$$

Now introduce the ordered products in the symmetric group algebra $\mathbb{C}S_{l+m}$

$$\begin{aligned} R_{\Lambda\mathrm{M}}(z,w) &= \prod_{p=1,\ldots,l}^{\longrightarrow}\Bigl(\prod_{q=1,\ldots,m}^{\longleftarrow} f_{p,l+q}(c_p+z\,,d_q+w)\Bigr)\,, \\ R'_{\Lambda\mathrm{M}}(z,w) &= \prod_{p=1,\ldots,l}^{\longleftarrow}\Bigl(\prod_{q=1,\ldots,m}^{\longrightarrow} f_{p,l+q}(c_p+z\,,d_q+w)\Bigr)\,; \end{aligned}$$

the arrows indicate the ordering of (non-commuting) factors. We keep to the assumption $z-w\notin\mathbb{Z}$. Applying Proposition 2 to the tableaux Λ,M and using (1) repeatedly, we get

$$F_\Lambda\overline{F}_{\mathrm{M}}\,R_{\Lambda\mathrm{M}}(z,w)=R'_{\Lambda\mathrm{M}}(z,w)\,F_\Lambda\overline{F}_{\mathrm{M}}\,. \tag{4}$$

Hence the right multiplication in $\mathbb{C}S_{l+m}$ by $R_{\Lambda\mathrm{M}}(z,w)$ preserves the left ideal W.

The H_{l+m}-module W' induced from the $H_m\otimes H_l$-module $V_\mu(w)\otimes V_\lambda(z)$ can be then realized as the left ideal in $\mathbb{C}S_{l+m}$ generated by the element $\tau^{-1}F_\Lambda\overline{F}_{\mathrm{M}}\tau$, so that $W'=W\tau$. The action of the generators $x_1,\ldots,x_{l+m}$ in W' is determined by

$$x_q\cdot\tau^{-1}F_\Lambda\overline{F}_{\mathrm{M}}\,\tau=(d_q+w)\,\tau^{-1}F_\Lambda\overline{F}_{\mathrm{M}}\,\tau \quad \text{for each} \quad q=1,\ldots,m\,; \tag{5}$$

$$x_{m+p}\cdot\tau^{-1}F_\Lambda\overline{F}_{\mathrm{M}}\,\tau=(c_p+z)\,\tau^{-1}F_\Lambda\overline{F}_{\mathrm{M}}\,\tau \quad \text{for each} \quad p=1,\ldots,l\,. \tag{6}$$

Consider the operator of right multiplication in $\mathbb{C}S_{l+m}$ by $R_{\Lambda\mathrm{M}}(z,w)\tau$. Denote by I the restriction of this operator to the subspace $W\subset\mathbb{C}S_{l+m}$. Due to (4) the image of the operator I is contained in the subspace W'.

Proposition 3. *The operator $I\,:\,W\to W'$ commutes with the action of H_{l+m}.*

Proof. The subalgebra $\mathbb{C}S_{l+m}\subset H_{l+m}$ acts in W,W' via left multiplication; the operator I commutes with this action by definition. The left ideal W is generated by the element $F_\Lambda\overline{F}_{\mathrm{M}}$, so it suffices to check that $x_p\cdot I\bigl(F_\Lambda\overline{F}_{\mathrm{M}}\bigr)=I\bigl(x_p\cdot F_\Lambda\overline{F}_{\mathrm{M}}\bigr)$ for each $p=1,\ldots,l+m$. Firstly consider the case $1\leqslant p\leqslant l$, then by (3, 4, 5, 6)

$$\begin{aligned} x_p\cdot I\bigl(F_\Lambda\overline{F}_{\mathrm{M}}\bigr) &= x_p\cdot\bigl(R'_{\Lambda\mathrm{M}}(z,w)\,\tau\bigr)\bigl(\tau^{-1}F_\Lambda\overline{F}_{\mathrm{M}}\,\tau\bigr) \\ &= \bigl(R'_{\Lambda\mathrm{M}}(z,w)\,\tau\bigr)\bigl(x_{\tau^{-1}(p)}\cdot\tau^{-1}F_\Lambda\overline{F}_{\mathrm{M}}\,\tau\bigr) \\ &= \bigl(R'_{\Lambda\mathrm{M}}(z,w)\,\tau\bigr)(c_p+z)\bigl(\tau^{-1}F_\Lambda\overline{F}_{\mathrm{M}}\,\tau\bigr)=I\bigl(x_p\cdot F_\Lambda\overline{F}_{\mathrm{M}}\bigr)\,; \end{aligned}$$

here we also used the defining relations of the algebra H_{l+m}. For more details of this argument see [8]. The case $l+1 \leqslant p \leqslant l+m$ is similar. □

Consider the operator of the right multiplication in $\mathbb{C}S_{l+m}$ by $R_{\Lambda\mathrm{M}}(z,w)$. This operator preserves the subspace W due to (4). The restriction of this operator to W will be denoted by J. The subalgebra $\mathbb{C}S_{l+m} \subset H_{l+m}$ acts in W via left multiplication, and J commutes with this action. Regard W as a $\mathbb{C}S_{l+m}$-module only. Let ν be any partition of $l+m$ such that the irreducible $\mathbb{C}S_{l+m}$-module V_ν appears in W with multiplicity one. The operator $J : W \to W$ preserves the subspace $V_\nu \subset W$ and acts there as multiplication by a certain number from $\mathbb{C}$. Denote this number by $r_\nu(z,w)$; it depends on z and w as a rational function of $z-w$, and does not depend on the choice of the tableaux Λ and M. Our aim is to compute the eigenvalues $r_\nu(z,w)$ of J for certain ν.

Before performing the computation, let us observe one general property of the eigenvalues $r_\nu(z,w)$. Similarly to the defintion of the H_{l+m}-intertwining operator $I : W \to W'$, one can define an operator $I' : W' \to W$ as the restriction to W' of the operator of the right multiplication in $\mathbb{C}S_{l+m}$ by $R_{\mathrm{M}\Lambda}(w,z)\,\tau^{-1}$. The operator I' commutes with the action of H_{l+m} as well. One can also consider the operator $J' : W' \to W'$, defined as the restriction to W' of the operator of right multiplication in $\mathbb{C}S_{l+m}$ by $R_{\mathrm{M}\Lambda}(w,z)$. There is a unique irreducible $\mathbb{C}S_{l+m}$-submodule in $V'_\nu \subset W'$ equivalent to $V_\nu \subset W$; actually here we have $V'_\nu = V_\nu\,\tau$. Consider the corresponding eigenvalue $r'_\nu(w,z)$ of the operator J'.

Proposition 4. *We have the relation*

$$r_\nu(z,w)\, r'_\nu(w,z) = \prod_{i=1}^{\lambda'_1} \prod_{k=1}^{\mu'_1} \frac{(z-w+\lambda_i-i+k)\,(z-w-\mu_k-i+k)}{(z-w+\lambda_i-\mu_k-i+k)\,(z-w-i+k)}\,.$$

Proof. The product $r_\nu(z,w)\, r'_\nu(w,z)$ is the eigenvalue of the composition $I' \circ I : W \to W$, corresponding to the subspace $V_\nu \subset W$. By definition, this composition is the operator of right multiplication in $W \subset \mathbb{C}S_{l+m}$ by the element

$$\begin{aligned}
R_{\Lambda\mathrm{M}}(z,w)\,\tau\, R_{\mathrm{M}\Lambda}(w,z)\,\tau^{-1} &= \prod_{p=1,\dots,l}^{\longrightarrow} \Bigl(\prod_{q=1,\dots,m}^{\longleftarrow} f_{p,l+q}(c_p+z\,, d_q+w) \Bigr) \cdot \tau \\
&\quad \times \prod_{q=1,\dots,m}^{\longrightarrow} \Bigl(\prod_{p=1,\dots,l}^{\longleftarrow} f_{q,m+p}(d_q+w\,, c_p+z) \Bigr) \cdot \tau^{-1} \\
= \prod_{p=1,\dots,l}^{\longrightarrow} \Bigl(\prod_{q=1,\dots,m}^{\longleftarrow} & f_{p,l+q}(c_p+z\,, d_q+w) \Bigr) \\
&\quad \times \prod_{p=1,\dots,l}^{\longleftarrow} \Bigl(\prod_{q=1,\dots,m}^{\longrightarrow} f_{l+q,p}(d_q+w\,, c_p+z) \Bigr)
\end{aligned}$$

$$= \prod_{p=1}^{l} \prod_{q=1}^{m} \Big(1 - (z - w + c_p - d_q)^{-2} \Big) .$$

The last equality has been obtained by using repeatedly the relations (2), it shows that the composition $I' \circ I$ is a scalar operator. Now recall that the contents of the boxes in the same row of a Young diagram increase by 1, when moving from left to right. For the i-th row of λ, the contents of the leftmost and rightmost boxes are $1-i$ and $\lambda_i - i$ respectively. For the k-th row of μ, the contents of the leftmost and rightmost boxes are respectively $1-k$ and $\mu_k - k$. Hence the right hand side of the last equality can be rewritten as

$$\prod_{p=1}^{l} \prod_{q=1}^{m} \left(\frac{z - w + c_p - d_q + 1}{z - w + c_p - d_q} \cdot \frac{z - w + c_p - d_q - 1}{z - w + c_p - d_q} \right)$$

$$= \prod_{i=1}^{\lambda'_1} \prod_{q=1}^{m} \left(\frac{z - w + \lambda_i - i - d_q + 1}{z - w + 1 - i - d_q} \cdot \frac{z - w - i - d_q}{z - w + \lambda_i - i - d_q} \right)$$

$$= \prod_{i=1}^{\lambda'_1} \prod_{k=1}^{\mu'_1} \left(\frac{z - w + \lambda_i - i + k}{z - w + \lambda_i - i - \mu_k + k} \cdot \frac{z - w - i - \mu_k + k}{z - w - i + k} \right). \qquad \square$$

3 Multiplicity-free components

Choose any sequence $a_1, \ldots, a_{\lambda'_1} \in \{1, 2, \ldots\}$ of pairwise distinct indices, we emphasize that this sequence needs not to be increasing. Here λ'_1 is the number of non-zero parts in the partition λ. Consider the partition μ as an infinite sequence with finitely many non-zero terms. Define an infinite sequence $\gamma = (\gamma_1, \gamma_2, \ldots)$ by

$$\gamma_{a_i} = \mu_{a_i} + \lambda_i \,, \quad i = 1, \ldots, \lambda'_1 \,;$$
$$\gamma_a = \mu_a \,, \quad a \neq a_1, \ldots, a_{\lambda'_1} \,.$$

Suppose we have $\gamma_1 \geqslant \gamma_2 \geqslant \ldots$, so that γ is a partition of $l+m$. Then the irreducible $\mathbb{C}S_{l+m}$-module V_γ corresponding to the partition γ appears in W with multiplicity one. Indeed, the multiplicity of V_γ in W equals the multiplicity of V_λ in the $\mathbb{C}S_l$-module corresponding to the skew Young diagram γ/μ. The latter multiplicity is one by the definition of γ; see for instance [9].

We will evaluate the number $r_\gamma(z, w)$ by applying the operator J to a particular vector in the subspace $V_\gamma \subset W$. Assume that $\Lambda = \Lambda^c$ is the column tableau of shape λ; the tableau M will be still arbitrary. The image of the action of the element $F_{\Lambda^c} \overline{F}_{\mathrm{M}}$ in the irreducible $\mathbb{C}S_{l+m}$-module V_γ is a one-dimensional subspace. Let us describe this subspace explicitly. The standard chain of subgroups $S_1 \subset S_2 \subset \cdots \subset S_l$ corresponds to the natural ordering of the numbers $1, 2, \ldots, l$. Now consider the new chain of subgroups

$$S_1 \subset \ldots \subset S_m \subset S_{1+m} \subset \ldots \subset S_{l+m}$$

corresponding to the ordering $l+1, \ldots, l+m, 1, \ldots, l$. Notice that the element $\overline{F}_{\mathrm{M}}$ belongs to the subgroup S_m in this new chain. Take the Young basis in the space V_γ associated with the new chain. In particular, take the basis vector $v_\Gamma \in V_\gamma$ corresponding to the tableau Γ of shape γ defined as follows. The numbers $l+1, \ldots, l+m$ appear in Γ respectively in the same positions as the numbers $1, \ldots, m$ do in tableau M. Now for every positive integer j consider all those parts of λ which are equal to j. These are λ_i with $i = \lambda'_{j+1}+1, \lambda'_{j+1}+2, \ldots, \lambda'_j$. Let $f_1, \ldots, f_{\lambda'_j-\lambda'_{j+1}}$ be the indices a_i with $\lambda_i = j$, arranged in the increasing order. By definition, the numbers appearing in the rows $\lambda'_{j+1}+1, \lambda'_{j+1}+2, \ldots, \lambda'_j$ of the tableau Λ^c, will stand in the rows $f_1, \ldots, f_{\lambda'_j-\lambda'_{j+1}}$ of Γ respectively. For example, here for $\lambda = (3,2)$ and $\mu = (2,1)$ with $a_1 = 2$ and $a_2 = 1$ we show a standard tableau M and the corresponding tableau Γ:

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array} \qquad\qquad \begin{array}{|c|c|c|c|}\hline 6 & 7 & 2 & 4 \\ \hline 8 & 1 & 3 & 5 \\ \hline \end{array}$$

Proposition 5. *With respect to the ordering $l+1, \ldots, l+m, 1, \ldots, l$ the tableau Γ is standard.*

Proof. Reading the rows of the tableau Γ from left to right, or reading its columns downwards, the numbers $l+1, \ldots, l+m$ appear in the increasing order because M is standard. These numbers will also appear before $1, \ldots, l$. Moreover, the numbers $1, \ldots, l$ increase along each row of the tableau Γ by the definition of Λ^c. Now suppose that a column of Γ contains two different numbers $p, q \in \{1, \ldots, l\}$. Let $a, \bar{a}$ be the corresponding rows; assume that $a < \bar{a}$. Then $a = a_i$ and $\bar{a} = a_{\bar{\imath}}$ for certain indices $i, \bar{\imath} \in \{1, \ldots, \lambda'_1\}$. If $\lambda_i \geqslant \lambda_{\bar{\imath}}$ then $p < q$ by definition of Λ^c.

Let $j, \bar{\jmath}$ be the columns corresponding to the numbers p, q in the tableau Λ^c. Suppose that $\lambda_i < \lambda_{\bar{\imath}}$, then $\mu_a > \mu_{\bar{a}}$ because $\mu_a + \lambda_i \geqslant \mu_{\bar{a}} + \lambda_{\bar{\imath}}$. Since p and q stand in the same column of the tableau Γ, we then have $j < \bar{\jmath}$ and $p < q$. ☐

Proposition 6. *The one-dimensional subspace $F_{\Lambda^c}\overline{F}_{\mathrm{M}} \cdot V_\gamma \subset V_\gamma$ is obtained from the space $\mathbb{C}v_\Gamma$ by antisymmetrization relative to the columns of the tableau Λ^c.*

Proof. Let S_λ be the subgroup in S_l consisting of all permutations which preserve the columns of the tableau Λ^c as sets. Let $Q \in \mathbb{C}S_l$ be the alternated sum of all elements from S_λ. Put $V = \overline{F}_{\mathrm{M}} \cdot V_\gamma$. The subspace $V \subset V_\gamma$ is spanned by the Young vectors, corresponding to the tableaux which agree with Γ in the entries $l+1, \ldots, l+m$. The action of the element F_{Λ^c} in V_γ

preserves the subspace V, and the image $F_{\Lambda^c} \cdot V \subset V$ is one-dimensional. Moreover, we have $F_{\Lambda^c} \cdot V = Q \cdot V$; see [4]. It now remains to check that $Q \cdot v_\Gamma \neq 0$.

By our choice of the tableau Γ, it suffices to consider the case when λ consists of one column only. But then the element $Q \in \mathbb{C}S_l$ is central. On the other hand, the vector $v_\Gamma \in V$ is $\mathbb{C}S_l$-cyclic; see [2]. So $Q \cdot V \neq \{0\}$ implies $Q \cdot v_\Gamma \neq 0$. □

4 Main theorem

Let γ be any of the partitions of $l + m$ described in the beginning of §3. Our main result is the following expression for the corresponding eigenvalue $r_\gamma(z,w)$ of the operator $J : W \to W$. This expression will be obtained by applying J to the vector $F_{\Lambda^c}\overline{F}_{\mathrm{M}} \cdot v_\Gamma$ in $V_\gamma \subset W$ and using Proposition 6.

Theorem 1. *We have the equality*

$$r_\gamma(z,w) \;=\; \prod_{(i,j)} \frac{z - w - \lambda'_j - \mu_{a_i} + a_i + j - 1}{z - w - i + j}$$

where the product is taken over all boxes (i,j) of the Young diagram λ.

Proof. Using (4) and applying Proposition 2.12 of [10] to the tableau M, we obtain the equalities in the algebra $\mathbb{C}S_{l+m}$

$$F_{\Lambda^c}\overline{F}_{\mathrm{M}}\, R_{\Lambda^c\mathrm{M}}(z,w) = \dim V_\lambda / l! \cdot F_{\Lambda^c}\overline{F}_{\mathrm{M}}\, R_{\Lambda^c\mathrm{M}}(z,w)\, F_{\Lambda^c} = \dim V_\lambda / l! \,\times$$

$$F_{\Lambda^c}\overline{F}_{\mathrm{M}} \left(\prod_{p=1,\ldots,l}^{\longrightarrow} \frac{\sigma_{l+1,p} + \ldots + \sigma_{l+m,p} - c_p - z + w}{-c_p - z + w} \right) F_{\Lambda^c} = \dim V_\lambda / l! \,\times$$

$$F_{\Lambda^c}\overline{F}_{\mathrm{M}} \left(\prod_{p=1,\ldots,l} \frac{\sigma_{l+1,p} + \ldots + \sigma_{l+m,p} + \sigma_{1p} + \ldots + \sigma_{p-1,p} - 2c_p - z + w}{-c_p - z + w} \right) F_{\Lambda^c};$$

we have also used Proposition 1 of the present article, cf. [12]. Here in the last line the factors corresponding to $p = 1, \ldots, l$ pairwise commute. By the same proposition applied to the partition γ instead of λ, any Young vector in the $\mathbb{C}S_{l+m}$-module V_γ is an eigenvector for the action of the elements

$$\sigma_{l+1,p} + \ldots + \sigma_{l+m,p} + \sigma_{1p} + \ldots + \sigma_{p-1,p}\,; \qquad p = 1, \ldots, l\,.$$

The vector $Q \cdot v_\Gamma \in V_\gamma$ is a linear combination of the Young vectors, corresponding to standard tableaux obtained from Γ by permutations from the subgroup $S_\lambda \subset S_l$. The last expression for $F_{\Lambda^c}\overline{F}_{\mathrm{M}}\, R_{\Lambda^c\mathrm{M}}(z,w)$ now shows, in particular, that the number $r_\gamma(z,w)$ factorizes with respect to the columns of the Young diagram λ.

Firstly suppose that λ consists of one column only. Then the number $r_\gamma(z,w)$ is easy to evaluate; cf. [11]. Here we have $c_i = 1 - i$ for each $i = 1, \ldots, l$. Using the chain of subgroups $S_1 \subset S_2 \subset \cdots \subset S_l$ corresponding to the ordering $l, \ldots, 1$ we then get

$$
\begin{aligned}
&F_{\Lambda^c} \cdot \prod_{i=1,\ldots,l}^{\longrightarrow} \big(\, \sigma_{l+1,i} + \ldots + \sigma_{l+m,i} - 1 + i + u \,\big) \\
= F_{\Lambda^c} \cdot &\prod_{i=1,\ldots,l} \big(\, \sigma_{l+1,i} + \ldots + \sigma_{l+m,i} + \sigma_{li} + \ldots + \sigma_{i+1,i} - 1 + l + u \,\big) \,.
\end{aligned}
$$

Here in the last line the factors corresponding to $i = 1, \ldots, l$ pairwise commute. Their product commutes with any element from the subalgebra $\mathbb{C}S_l \subset \mathbb{C}S_{l+m}$, and acts on the vector $v_\Gamma \in V_\gamma$ as multiplication by the number

$$
\prod_{i=1}^{l} \big(\, \mu_{a_i} - a_i + l + u \,\big) \,. \tag{7}
$$

Let us now apply this result to the j-th column of a general Young diagram λ, consecutively for $j = 1, \ldots, \lambda_1$. For the general λ the content of the box (i,j) is $j - i$. According to our last expression for $F_{\Lambda^c}\overline{F}_{\mathrm{M}}\, R_{\Lambda^c\mathrm{M}}(z,w)$, in the product (7) we then have to replace l, μ_{a_i}, u by $\lambda'_j, \mu_{a_i} + j - 1, 2 - 2j - z + w$ respectively. Hence

$$
r_\gamma(z,w) \;=\; \prod_{j=1}^{\lambda_1} \prod_{i=1}^{\lambda'_j} \frac{\mu_{a_i} + j - 1 - a_i + \lambda'_j + 2 - 2j - z + w}{i - j - z + w} \,.
$$

□

5 Conjugate partitions

Choose any sequence $b_1, \ldots, b_{\lambda_1} \in \{1, 2, \ldots\}$ of pairwise distinct indices. Again, this sequence needs not to be increasing. Let us now regard the partition μ' conjugate to μ as an infinite sequence with finitely many non-zero terms. Determine an infinite sequence $\delta' = (\delta'_1, \delta'_2, \ldots)$ by

$$
\begin{aligned}
\delta'_{b_j} &= \mu'_{b_j} + \lambda'_j \,; \quad j = 1, \ldots, \lambda_1 \,; \\
\delta'_b &= \mu'_b \,; \quad b \neq b_1, \ldots, b_{\lambda_1} \,.
\end{aligned}
$$

Suppose $\delta'_1 \geqslant \delta'_2 \geqslant \ldots$, so that δ' is a partition of $l + m$. Define δ as the partition conjugate to δ'. The irreducible $\mathbb{C}S_{l+m}$-module V_δ appears in W with multiplicity one. Take the corresponding eigenvalue $r_\delta(z,w)$ of the operator $J : W \to W$.

Theorem 2. *We have the equality*

$$r_\delta(z,w) \;=\; \prod_{(i,j)} \frac{z-w+\lambda_i+\mu'_{b_j}-i-b_j+1}{z-w-i+j}$$

where the product is taken over all boxes (i,j) of the Young diagram λ.

Proof. Denote by Z_δ the minimal central idempotent in $\mathbb{C}S_{l+m}$ corresponding to the partition δ. Take the automorphism $*$ of the algebra $\mathbb{C}S_{l+m}$ such that $\sigma^* = \operatorname{sgn}(\sigma)\,\sigma$; we have $Z_\delta^* = Z_{\delta'}$ then. Reflecting the tableaux Λ and M in their main diagonals we get certain standard tableaux of shapes λ' and μ' respectively; denote these tableaux by Λ' and M'. Then we have $F_\Lambda^* = F_{\Lambda'}$ and $F_{\mathrm{M}}^* = F_{\mathrm{M}'}$.

On the other hand, by the definition of the number $r_\delta(z,w)$ we have the equality

$$Z_\delta\, F_\Lambda\, \overline{F}_{\mathrm{M}}\, R_{\Lambda\mathrm{M}}(z,w) \;=\; r_\delta(z,w)\; Z_\delta\, F_\Lambda\, \overline{F}_{\mathrm{M}}\,.$$

By applying the automorphism $*$ to this equality we get

$$Z_{\delta'}\, F_{\Lambda'}\overline{F}_{\mathrm{M}'} R^*_{\Lambda\mathrm{M}}(z,w) = r_\delta(z,w)\; Z_{\delta'}\, F_{\Lambda'}\overline{F}_{\mathrm{M}'}.$$

But $R^*_{\Lambda\mathrm{M}}(z,w) = R_{\Lambda'\mathrm{M}'}(-z,-w)$ by definition. Therefore by applying Theorem 1 to the partitions λ', μ' instead of λ, μ and choosing $\gamma = \delta'$ we get

$$r_\delta(z,w) \;=\; r_{\delta'}(-z,-w) \;=\; \prod_{(i,j)} \frac{z-w+\lambda_j+\mu'_{b_i}-b_i-j+1}{z-w+i-j}$$

where the product is taken over all boxes (i,j) of the diagram λ'. Equivalently, this product may be taken over all boxes (j,i) of the diagram λ. ☐

6 Mixed hook-length formula

Let us now derive the Corollary stated in the beginning of this article. We will use Theorems 1 and 2 in the simplest situation when $a_i = i$ for every $i = 1, \dots, \lambda'_1$ and $b_j = j$ for every $j = 1, \dots, \lambda_1$. Then we have $\gamma = \lambda + \mu$ and $\delta = (\lambda' + \mu')'$. By Theorems 1 and 2, $h_{\lambda\mu}(z,w) = r_{\lambda+\mu}(z,w)\,/\,r_{(\lambda'+\mu')'}(z,w)$ equals the product of the fractions

$$\frac{z-w-\lambda'_j-\mu_i+i+j-1}{z-w+\lambda_i+\mu'_j-i-j+1} \tag{8}$$

taken over all boxes (i,j) of the diagram λ. Consider those boxes of λ which do not belong to μ. These boxes form a skew Young diagram, let us denote it by ω. To obtain the Corollary, it suffices to prove the following

Proposition 7. *The product of the fractions* (8) *over the boxes* $(i\,,j)$ *of* ω *equals* 1.

Proof. We will proceed by induction on the number of boxes in the diagram ω. Let us write u instead of $z-w$ for short. When the diagram ω is empty, the statement to prove is tautological. Now let (a,b) be any box of ω such that by removing it from λ we obtain again a Young diagram; then we have $\lambda_a=b$ and $\lambda'_b=a$. By applying the induction hypothesis to the last diagram instead of λ, we have to show that the product

$$\prod_{j=\mu_a+1}^{b-1}\frac{u+b-1+\mu'_j-a-j+1}{u+b+\mu'_j-a-j+1}\cdot\prod_{i=\mu'_b+1}^{a-1}\frac{u-\mu_i-a+i+b-1}{u-\mu_i-a+1+i+b-1}$$
$$\times\;\frac{u-\mu_a-a+a+b-1}{u+b+\mu'_b-a-b+1}\tag{9}$$

equals 1. Note that here we have $\mu_a<\lambda_a$ and $\mu'_b<\lambda'_b$.

Suppose there is a box $(\bar\imath,\bar\jmath)$ of ω with $\mu_a<\bar\jmath<\lambda_a$ and $\mu'_b<\bar\imath<\lambda'_b$, such that by adding this box to μ we obtain again a Young diagram. Then we have $\mu_{\bar\imath}=\bar\jmath-1$ and $\mu'_{\bar\jmath}=\bar\imath-1$. For the last diagram instead of μ, the product (9) equals 1 by the induction hypothesis. Then it suffices to check the equality to 1 of

$$\frac{u+b-1+\bar\imath-1-a-\bar\jmath+1}{u+b+\bar\imath-1-a-\bar\jmath+1}\cdot\frac{u+b+\bar\imath-a-\bar\jmath+1}{u+b+\bar\imath-a-\bar\jmath}$$
$$\times\;\frac{u-\bar\jmath+1-a+\bar\imath+b-1}{u-\bar\jmath+1-a+1+\bar\imath+b-1}\cdot\frac{u-\bar\jmath-a+\bar\imath+b}{u-\bar\jmath-a+\bar\imath+b-1}\;.$$

But this product has form $\dfrac{v-1}{v}\cdot\dfrac{v+1}{v}\cdot\dfrac{v}{v+1}\cdot\dfrac{v}{v-1}$ with $v=u-a+b+\bar\imath-\bar\jmath\,$.

It remains to consider the case when there is no box $(\bar\imath,\bar\jmath)$ in ω with the above listed properties. Then $\mu'_j=a-1$ for all $j=\mu_a+1,\ldots,b-1$ and $\mu_i=b-1$ for all $i=\mu'_b+1,\ldots,a-1$. Hence in this remaining case the product (9) equals

$$\frac{u+b-1+a-1-a-b+1+1}{u+b+a-1-a-\mu_a-1+1}\cdot\frac{u-b+1-a+\mu'_b+1+b-1}{u-b+1-a+1+a-1+b-1}$$
$$\times\;\frac{u-\mu_a-a+a+b-1}{u+b+\mu'_b-a-b+1}\;=\;1\,.$$

□

7 Concluding remarks

In this final section we derive from Theorem 2 the formula $\dim V_\lambda=l\,!/h_\lambda$ for the dimension of the irreducible $\mathbb{C}S_l$-module V_λ. We will actually show that the coefficient $l\,!/\dim V_\lambda$ in the relation $F_\Lambda^2=l\,!/\dim V_\lambda\cdot F_\Lambda$ equals

h_λ, the product of the l hook-lengths of the Young diagram λ. We will use induction on the number of rows in λ. If there is only one row in λ, then $h_\lambda = l!$ and $\dim V_\lambda = 1$, so the desired equality is clear. Let us now make the inductive assumption for λ, and consider the Young diagram obtained by adding m boxes to λ in the row $\lambda'_1 + 1$. Denote the new diagram by θ, we assume that $m \leqslant \lambda_i$ for any $i = 1, \ldots, \lambda'_1$. Put $\mu = (m, 0, 0, \ldots)$ and consider the eigenvalue $r_\theta(z, w)$ of the operator $J : W \to W$ corresponding to the multiplicity-free component $V_\theta \subset W$.

In our case, there is only one standard tableau M of shape μ. Let Θ be the unique standard tableau of shape θ agreeing with Λ in the entries $1, \ldots, l$; the numbers $l+1, \ldots, l+m$ then appear in the last row of Θ. By definition,

$$F_\Theta \cdot F_\Lambda \overline{F}_{\mathrm{M}} R_{\Lambda\mathrm{M}}(z, w) = r_\theta(z, w) \cdot F_\Theta F_\Lambda \overline{F}_{\mathrm{M}} = h_\lambda \, m! \, r_\theta(z, w) \cdot F_\Theta \, ;$$

the second equality here has been obtained using the inductive assumption. On the other hand, due to Proposition 2 the matrix element F_Θ coincides with the value of the product $F_\Lambda \overline{F}_{\mathrm{M}} R_{\Lambda\mathrm{M}}(z, w)$ at $z - w = \lambda'_1$. To make the inductive step, it now remains to check that h_θ coincides with the value of $h_\lambda \, m! \, r_\theta(z, w)$ at $z - w = \lambda'_1$.

Let us use Theorem 2 when $b_j = j$ for every $j = 1, \ldots, \lambda_1$. With our particular choice of μ, we then obtain that $r_\theta(z, w)$ equals the product over $i = 1, \ldots, \lambda'_1$ of

$$\prod_{j=1}^{m} (z - w + \lambda_i - i - j + 2) \cdot \prod_{j=m+1}^{\lambda_i} (z - w + \lambda_i - i - j + 1) \cdot \prod_{j=1}^{\lambda_i} \frac{1}{z - w - i + j} \, .$$

Changing the running index j to $\lambda_i - j + 1$ in the products over $j = 1, \ldots, m$ and over $j = m + 1, \ldots, \lambda_i$ above, we obtain after cancellations the equality

$$r_\theta(z, w) = \prod_{i=1}^{\lambda'_1} \frac{z - w + \lambda_i - i + 1}{z - w + \lambda_i - m - i + 1} \, .$$

This equality shows that the value $r_\theta(z, w)$ at $z - w = \lambda'_1$ coincides with the ratio

$$\frac{h_\theta}{h_\lambda \, m!} = \prod_{i=1}^{\lambda'_1} \prod_{j=1}^{m} \frac{\lambda_i + \lambda'_1 - i - j + 2}{\lambda_i + \lambda'_1 - i - j + 1} = \prod_{i=1}^{\lambda'_1} \frac{\lambda'_1 + \lambda_i - i + 1}{\lambda'_1 + \lambda_i - m - i + 1} \, . \qquad \square$$

Let us make a few concluding remarks. Throughout §§1–6 we assumed that $z - w \notin \mathbb{Z}$, then the H_{l+m}-module W is irreducible. For $z - w \in \mathbb{Z}$, our Corollary implies that the module W is reducible if $z - w$ is a mixed hook-length of the second kind relative to λ and μ, and if $w - z$ is a mixed hook-length of the first kind. When $\lambda = \mu$, our Corollary implies that the module

W is reducible if $|z-w|$ is a hook-length of λ. Moreover, then the module W is irreducible [6] for all remaining values $|z-w|$.

When $\lambda \neq \mu$ the H_{l+m}-module W maybe reducible while neither $z-w$ is a mixed hook-length of the second kind, nor $w-z$ is a mixed hook-length of the first kind. The irreducibility criterion for the module W with arbitrary λ and μ has been also given in [6]. This work shows that the module W is reducible if and only if the difference $z-w$ belongs to a certain finite subset $\mathcal{S}_{\lambda\mu} \subset \mathbb{Z}$ determined in [7]. This subset satisfies the property $\mathcal{S}_{\lambda\mu} = -\mathcal{S}_{\mu\lambda}$.

Denote by $\mathcal{D}_{\lambda\mu}$ the union of the sets of all zeroes and poles of the rational functions $r_{\lambda+\mu}(z,w)/r_\nu(z,w)$ in $z-w$, where ν ranges over all partitions γ and δ described in §3 and §5 respectively. Then $\mathcal{D}_{\lambda\mu} \subset \mathcal{S}_{\lambda\mu}$. Then $-\mathcal{D}_{\mu\lambda} \subset \mathcal{S}_{\lambda\mu}$ also. Using [7] one can demonstrate that if $\lambda'_1, \mu'_1 \leqslant 3$ then $\mathcal{D}_{\lambda\mu} \cup (-\mathcal{D}_{\mu\lambda}) = \mathcal{S}_{\lambda\mu}$. However, $\mathcal{D}_{\lambda\mu} \cup (-\mathcal{D}_{\mu\lambda}) \neq \mathcal{S}_{\lambda\mu}$ for general partitions λ and μ. For example, if $\lambda = (8,3,2,1,0,0,\ldots)$ and $\mu = (6,4,4,0,0,\ldots)$ then $0 \in \mathcal{S}_{\lambda\mu}$ but $0 \notin \mathcal{D}_{\lambda\mu}, \mathcal{D}_{\mu\lambda}$.

For general λ and μ, it would be interesting to point out for every $t \in \mathcal{S}_{\lambda\mu}$ a partition ν of $l+m$, such that the $\mathbb{C}S_{l+m}$-module V_ν appears in W with multiplicity one, and such that the ratio $r_{\lambda+\mu}(z,w)/r_\nu(z,w)$ has a zero or pole at t, as a rational function of $z-w$.

References

1. Alishauskas, S., Kulish, P.: Spectral resolution of $SU(3)$-invariant solutions of the Yang–Baxter equation. J. Soviet Math., **35**, 2563–2574 (1986)
2. Cherednik I.: Special bases of irreducible representations of a degenerate affine Hecke algebra. Funct. Anal. Appl., **20**, 76–78 (1986)
3. Drinfeld, V.: Degenerate affine Hecke algebras and Yangians. Funct. Anal. Appl., **20**, 56–58 (1986)
4. James, G., Kerber, A.: The Representation Theory of the Symmetric Group. Addison-Wesley, Reading MA (1981)
5. Jucys, A.: Symmetric polynomials and the centre of the symmetric group ring. Rep. Math. Phys., **5**, 107–112 (1974)
6. Leclerc, B., Nazarov, M., Thibon, J.-Y.: Induced representations of affine Hecke algebras and canonical bases of quantum groups. In: Joseph, A. et al (eds) Studies in Memory of Issai Schur. Birkhauser, Boston MA, 115–153 (2002)
7. Leclerc, B., Zelevinsky, A.: Quasicommuting families of quantum Plücker coordinates. Amer. Math. Soc. Translat., **181**, 85–108 (1998)
8. Lusztig, G.: Affine Hecke algebras and their graded version. J. Amer. Math. Soc., **2**, 599–635 (1989)
9. Macdonald, I.: *Symmetric Functions and Hall Polynomials*. Clarendon Press, Oxford (1979)
10. Nazarov, M.: Yangians and Capelli identities. Amer. Math. Soc. Translat., **181**, 139–163 (1998)
11. Nazarov, M., Tarasov, V.: On irreducibility of tensor products of Yangian modules. Intern. Math. Research Notices, 125–150 (1998)
12. Okounkov, A.: Young basis, Wick formula, and higher Capelli identities. Intern. Math. Research Notices, 817–839 (1996)

Information about the school

Organizing Committee and Secretary

A. Vershik (chairman; PDMI, St. Petersburg)
K. Kokhas (StPSU, Russia)
Yu. Neretin (ITEP, Russia
E. Novikova
N. Tsilevich (PDMI, Russia)
Yu. Yakubovich

Scientific Committee

A. Vershik (chairman;)
O. Bohigos (Orsay, France)
E. Brésin (ENS, France)
P. Deift (NYU, US)
L. Faddeev (PDMI, Russia)
K. Johansson (KTH, Sweden)
V. Kazakov (ENS, France)
V. Malyshev (INRIA, France)
L. Pastur (Paris VII, France)

Program of the School

July 8, Sunday (Euler International Mathematical Institute, Pesochnaya nab., 10)
15.00–20.00 REGISTRATION

1st day, July 9, Monday (Steklov Mathematical Institute, Fontanka, 27)
08.30–09.45 REGISTRATION
10.00–10.30 Opening Session
10.30–11.30 ***Brezin E.*** An introduction to matrix models – 1
12.00–13.00 ***Vershik A.*** Introduction to asymptotic theory of representations – 1
15.00–15.50 ***Korepin V.*** Quantum spin chains and Riemann Zeta function with odd arguments
16.10–17.00 ***Bozejko M.*** Positive definite functions on Coxeter groups and second quantization of Yang–Baxter type

2nd day, July 10, Tuesday (Euler International Mathematical Institute, Pesochnaya nab., 10)
09.30–10.30 ***Ol'shanski G.*** Harmonic analysis on big groups, and determinantal point processes – 1
10.40–11.40 ***Brezin E.*** An introduction to matrix models – 2
12.10–13.10 ***Malyshev V.*** Combinatorics and probability for maps on two dimensional surfaces
15.00–15.50 ***Hora A.*** An algebraic and combinatorial approach to central limit theorems related to discrete Laplacians
16.00–16.50 ***Nazarov M.*** On the Frobenius rank of a skew Young diagram
17.20–18.10 ***Kenyon R.*** Hyperbolic geometry and the low-temperature expansion of the Wulff shape in the 3D Ising model
19.40 ***Brezin E.*** Informal discussion on matrix models

3rd day, July 11, Wednesday
09.30–10.30 ***Okounkov A.*** Combinatorics and moduli spaces of curves – 1
10.40–11.40 ***Brezin E.*** An introduction to matrix models – 3
12.10–13.10 ***Biane Ph.*** Asymptotics of representations of symmetric groups, random matrices and free cumulants
14.00–17.00 Excursion over St. Petersburg
19.00 Mariinski Theatre

4th day, July 12, Thursday
09.30–10.30 ***Deift P.*** Random matrix theory and combinatorics: a Riemann–Hilbert approach – 1
10.40–11.40 ***Okounkov A.*** Combinatorics and moduli spaces of curves – 2
12.10–13.10 ***Kazakov V.*** Matrix quantum mechanics and statistical physics on planar graphs – 1

15.00–15.50 ***Faddeev L.*** 3-dimensional solitons and knots
16.30–18.10 ***Krichever I.*** τ-functions of conformal maps
19.30 Boat trip

5th day, July 13, Friday
09.30–10.30 ***Deift P.*** Random matrix theory and combinatorics: a Riemann–Hilbert approach – 2
10.40–11.40 ***Okounkov A.*** Combinatorics and moduli spaces of curves – 3
12.10–13.10 ***Ol'shanski G.*** Harmonic analysis on big groups, and determinantal point processes – 2
15.00–15.50 ***Kazakov V.*** Matrix quantum mechanics and statistical physics on planar graphs – 2
16.00–16.50 Round table on combinatorics of the configurations and limit shapes
17.20–18.10 ***Liskovets V.*** Some asymptotic distribution patterns for planar maps

July 14 Excursions, Museums, etc.

July 15 Excursion to Petergof

6th day, July 16, Monday
09.30–10.30 ***Deift P.*** Random matrix theory and combinatorics: a Riemann–Hilbert approach – 3
10.40–11.30 ***Kazakov V.*** Combinatorics of planar graphs in matrix quantum mechanics – 3
12.10–13.10 ***Borodin A.*** Asymptotic representation theory and Riemann–Hilbert problem – 1
15.00–15.50 ***Novikov S.*** On the weakly nonlocal Poisson and symplectic structures
16.00–16.50 ***Spiridonov V.*** Special functions of hypergeometric type associated with elliptic beta integrals
17.20–18.10 ***Missarov M.*** Exactly solvable renormalization group model

7th day, July 17, Tuesday
09.30–10.30 ***Deift P.*** Random matrix theory and combinatorics: a Riemann–Hilbert approach – 4
10.40–11.40 ***Borodin A.*** Asymptotic representation theory and Riemann–Hilbert problem – 2
12.10–13.10 ***Okounkov A.*** Combinatorics and moduli spaces of curves – 4
15.00–15.50 ***Speicher R.*** Free probability and Random matrices
16.00–16.50 ***Korotkin D.*** Riemann–Hilbert problems related to branched coverings of GP^1, τ-function and Liouville action
17.20–18.10 ***Litvinov G.*** Representation theory in Idempotent (asymptotic) Mathematics

8th day, July 18, Wednesday

09.30–10.30 ***Smirnov S.*** Critical percolation is conformally invariant – 1

10.40–11.40 ***Pastur L.*** Eigenvalue distribution of unitary invariant ensembles of random matrices of large order – 1

12.10–13.10 ***Pevzner M.*** On tensor products and Beresin kernels

15.00–17.00 ***Memorial session devoted to Sergei Kerov and Anatoly Izergin***

9th day, July 19, Thursday

09.30–10.30 ***Speicher R.*** Free probability and Random matrices – 2

10.40–11.40 ***Bozejko M.*** White noise associated to the characters of the infinite symmetric group — Hopf–Kerov deformation

12.10–13.00 ***Jacobsen J. L.*** Enumerating coloured tangles

13.05–13.35 ***Sniady P.*** Random matrices and free probability

15.00–15.30 ***Mlotkowski W.*** Λ-free probability

15.30–16.00 ***Petrogradskij V.*** Asymptotical theory of infinite dimensional Lie algebras

16.30–17.00 ***Kuznetsov V.*** On explicit formulae for special Macdonald polynomials

17.05–17.35 ***Dubrovskiy S.*** Moduli space of symmetric connections

17.40–18.10 ***Stukopin V.*** Representations of Yangians of Lie Superalgebras $A(m,n)$ type

10th day, July 20, Friday

09.30–10.30 ***Pastur L.*** Eigenvalue distribution of unitary invariant ensembles of random matrices of large order – 2

10.40–11.10 ***Vershik A.*** Introduction to asymptotic theory of representations – 2

11.15–11.45 ***Yambartsev A.*** Two-dimensional Lorentzian Models

12.15–13.00 ***Round table***: Problems of the theory of integrable operators and determinant processes

13.00–13.30 ***Closing the school***

List of participants

Name	Affiliation	email
Ahmedov Ruslan	Moscow State University, Russia	ahmedov@mccme.ru
Alexandrov Sergej	SPhT, France	alexand@spht.saclay.cea.fr
Aptekarev Alexander	Keldysh Institute, Russia	aptekaa@spp.keldysh.ru
Balog Jozsef	University of Memphis, USA	balogj@msci.memphis.edu
Batchourine Pavel	Moscow State University, Russia	bachurin@ipac.ac.ru
Belogrudov Alexander	Inst. of Mathematics, Ufa Science Centre of RAS; Russia	belan@anrb.ru
Beliaev Dmitri	KTH, Sweden	beliaev@math.kth.se
Biane Philippe	CNRS, ENS, DMA, France	Philippe.Biane@ens.fr
Bloznelis Mindaugas	Vilnius University, Lithuania	mblozn@ieva.maf.vu.lt
Borodin Alexey	University of Pennsylvania, USA	borodine@math.upenn.edu
Bozejko Marek	Instytut Matematyczny, Wrocław University, Poland	bozejko@math.uni.wroc.pl
Brezin Edouard	École Normale Supérieure, France	brezin@corto.lpt.ens.fr
Calkin Neil	Clemson University; USA	calkin@math.clemson.edu
Carlet Guido	SISSA, Italy	carlet@sissa.it
Chebalov Sergej	UIUC, USA	chebalov@math.uiuc.edu
Chou Ching-Sung	Dept. of Math., National Central University, Taiwan, R.O.C.	chou@math.ncu.edu.tw
Cislo Jerzy	University of Wrocław, Institute of Theoretical Physics, Poland	cislo@ift.uni.wroc.pl
Cutler Johnatan	University of Memphis, USA	cutlerj@msci.memphis.edu
Deift Percy	Courant Institute and University of Pennsylvania, USA	deift@math1.cims.nyu.edu
Dobrynin Sergei	St. Petersburg State University, Russia	
Dubrovskiy Stanislav	Northeastern University, USA	dubr@research.neu.edu
Duzhin Sergei	Steklov Institute of Mathematics at St. Petersburg (POMI), Russia	duzhin@pdmi.ras.ru
Erschler Anna	Steklov Institute of Mathematics at St. Petersburg (POMI), Russia	erschler@pdmi.ras.ru
Faddeev Ludvig	Steklov Institute of Mathematics at St. Petersburg (POMI), Russia	faddeev@pdmi.ras.ru
Fedotov Sergej	UMIST, UK	sergei.fedotov@umist.ac.uk
Galkin Sergej	Moscow Independent University, Russia	
Gioev Dimitri	KTH, Sweden	
Gorbulski Alexander	St. Petersburg State University, Russia	alexander@MG4812.spb.edu
Gorbulski Mikhail	St. Petersburg State University, Russia	mike@MG4812.spb.edu
Gordin Mikhail	Steklov Institute of Mathematics at St. Petersburg (POMI), Russia	gordin@pdmi.ras.ru
Grebenkov Denis	École Normale Supérieure, France	dg@pmc.polytechnique.fr
Gurin Alexey	ILT NAF Ukraina, Ukraine	gurin@ilt.kharkov.ua
Hora Akihito	Okayama University, Japan	hora@ems.okayama-u.ac.jp
Igonine Serguei	University of Twente, Netherlands	igonin@mccme.ru
Jacobsen Jesper Lykke	Lab. de Physique Theorique et Modeles Statistiques, France	jacobsen@ipno.in2p3.fr
Kalita Eugen	Inst. of Applied Mathematics and Mechanics NAS, Ukraine	kalita@iamm.ac.donetsk.ua
Kazakov Vladimir	École Normale Supérieure, France	Vladimir.Kazakov@lpt.ens.fr
Kenyon Richard	Universite Paris Sud, France	kenyon@topo.math.u-psud.fr
Khimshiashvili Giorgi	A.Razmadze Mathematical Institute, Georgia	khimsh@rmi.acnet.ge

Name	Affiliation	email
Kirillov Alexander	SUNY at Stony Brook, USA	kirillov@math.sunysb.edu
Kokhas Konstantin	St. Petersburg State University, Russia	kostik@KK1437.spb.edu
Korepin Vladimir	YITP, State University of New York at Stony Brook, USA	korepin@insti.physics.sunysb.edu
Krichever Igor		
Krikun Maxim	Lab. of Large Random Systems, Moscow State University, Russia	
Kuznetcov Vadim	University of Leeds; UK	vadim@amsta.leeds.ac.uk
Larson Craig	University of Houston; USA	clarson@math.uh.edu
Liskovets Valerij	Institute of Mathematics of NAS of Belarus	liskov@im.bas-net.by
Litvinov Grigori	International Sophus Lie Centre, Russia	islc@dol.ru
Lodkin Andrei	St. Petersburg State University, Russia	lodkin@pdmi.ras.ru
Lucena Brian	Brown University: Division of Applied Math., USA	lucena@cfm.brown.edu
Maida Mylene	École Normale Supérieure, France	maida@clipper.ens.fr
Malyshev Vadim	I.N.R.I.A., France	Vadim.Malyshev@inria.fr
Marciniak Macin	Institute of Mathematics, Gdansk University, Poland	matmm@univ.gda.pl
Mattera Massimiliano	Universite Paris Sud, Laboratoire de mathematiques, France	Massimiliano.Mattera@math.u-psud.fr
Matveev Vladimir	Steklov Institute of Mathematics at St. Petersburg (POMI), Russia	matveev@u-bourgogne.fr
Missarov Moukadas	Kazan State University, Department of Applied Mathematics, Laboratory of Probability Theory, Russia	Moukadas.Missarov@ksu.ru
Mlotkowski Wojcieh	Mathematical Institute, The University of Wrocław, Poland	mlotkow@math.uni.wroc.pl
Mnev Nikolai	Steklov Institute of Mathematics at St. Petersburg (POMI), Russia	mnev@pdmi.ras.ru
Momar Dieng	University of California, Davis; USA	momar@math.ucdavis.edu
Nazarov Maxim	Department of Mathematics, University of York, UK	mln1@york.ac.uk
Nikitin Pavel	St. Petersburg State University, Russia	pasha@AN4760.spb.edu
Novikov Sergej	Landau Institute RAN, Moscow, Russia	
Okounkov Andrej	University of California at Berkeley, Dept. of Mathematics, USA	okounkov@math.berkeley.edu
Olshanlki Grigori	Institute for Information Transmission Problems, Russia	olsh@online.ru
Panov Dmitri	École Polytechnique, CMAT, France	panov@ihes.fr
Pastur Leonid	CPT, France/FTINT NANU, Ukraine; France	pastur@cpt.univ=mrs.fr
Petrogradskij Viktor	Ulyanovsk State University, Russia	vmp@mmf.ulsu.ru
Pevzner Micha	ULB; Belgium	mpevzner@ulb.ac.be
Romanovski Yuri	St. Petersburg State University, Russia	
Rider Brian	Dept of E.E., The Technion, Israel	rider@ee.technion.ac.il

Name	Affiliation	email
Sniady Piotr	Instytut Matematyczny, Uniwersytet Wrocławski, Poland	Piotr.Sniady@math.uni.wroc.pl
Speicher Roland	Queen's University Jeffery Hall, Canada	speicher@mast.queensu.ca
Spiridonov Vyacheslav	Joint Institute for Nuclear Research, Dubna, Russia	svp@thsun1.jinr.ru
Spitz Grigory	International Sophus Lie Centre, Moscow, Russia	
Starr Shannon	University of California, Davis, USA	sstarr@math.ucdavis.edu
Stenlund Mikko	Department of Mathematics, University of Helsinki, Finland	mikko.stenlund@helsinki.fi
Stevens Laura	UNC Chapel Hill, USA	laura@jalapeno.math.unc.edu
Strahov Evgeny	Brunel University, UK	strahov@physics.technion.ac.il
Stukopin Vladimir	Don State Technical University, Russia	stukopin@math.rsu.ru
Suidan Toufic	Princeton University; USA	tmsuidan@math.princeton.edu
Szenes Andras	MIT, USA	szenes@math.mit.edu
Tarasov Vitaly	Steklov Institute of Mathematics at St. Petersburg (POMI), Russia	tarasov@pdmi.ras.ru
Terilla John	UNC, USA	terillaj@wlu.edu
Tsilevich Natalia	St. Petersburg State University, Russia	natalia@pdmi.ras.ru
Vasiliev Vladimir	VNIPIET, Russia	
Vershik Anatoly	Steklov Institute of Mathematics at St. Petersburg (POMI), Russia	Vershik@pdmi.ras.ru
Wojakowski Lukazc	University of Wrocław, Mathematical Institute, Poland	Lukasz.Wojakowski@math.uni.wroc.pl
Yakubovich Yuri	Steklov Institute of Mathematics at St. Petersburg (POMI), Russia	yuyakub@mail.ru
Yambartsev Anatoly	Moscow State University, Russia	yambar@lbss.math.msu.su
Yakobson Dmitri		
Zamyatin Andrej	Lab. of Large Random Systems, Moscow State University, Russia.	zamyatin@lbss.math.msu.su
Zinn-Justin Paul	LPTMS, University Paris-Sud, ORSAY, FRANCE	Paul.Zinn-Justin@lpt.ens.fr
Zvonarev Mikhail	Steklov Institute of Mathematics at St. Petersburg (POMI), Russia	zvonarev@pdmi.ras.ru
Zvonkine Dimitri	École Normale Supérieure, France	zvonkine@clipper.ens.fr

Contents of the Proceedings of NATO ASI Asymptotic Combinatorics with Application to Mathematical Physics[1]

[1] Malyshev, V., Vershik, A. (eds), Kluwer Academic Publishers, 2002, 328pp.