Probability Measures on Representations of Unitary Groups

UROP+ Final Paper, Summer 2016

Zachary Izzo
Graduate Mentor: Christopher Ryba
Faculty Supervisor: Professor Vadim Gorin

September 1, 2016

Abstract

We study a family of probability measures arising from extreme characters of the infinite unitary group. We explore the connection between such measures and random lozenge tilings of the plane, then compute the asymptotic behavior of the moments of these distributions.

1 Introduction

This paper extends the work of [3] and explores certain properties of probability measures arising from extreme characters of the infinite unitary group.

The irreducible representations of unitary groups are parameterized by finite weakly decreasing sequences of integers. Their characters can be expressed in terms of Schur symmetric polynomials. Passing to the infinite unitary group, one may define a notion of "normalized character," i.e. a continuous function on $U(\infty)$ which is conjugation invariant, satisfies a positive-definiteness condition, and takes the value 1 at the identity of $U(\infty)$ (like a usual character). There is a description of such functions (corresponding to "irreducible") in terms of eigenvalues of elements of $U(\infty)$. Restricting to a finite unitary group $U(N)$, we obtain a linear combination of normalized characters of irreducible representations of $U(N)$. By resolving the restricted representation in this way, we obtain a probability measure on the set of irreducible representations of $U(N)$. Since these irreducible representations are parameterized by partitions, we arrive at a distribution on partitions.

In this article, we will first explain the motivation for studying such distributions, namely their relation to random lozenge tilings of the plane. We will then delve deeper into how the distributions arise from characters of the infinite unitary group. With the aim of calculating the asymptotic behavior of the moments of these distributions, we introduce a family of differential operators which will drive our calculations. We will then compute the asymptotic moments for a special case, and conclude by suggesting directions for future research.

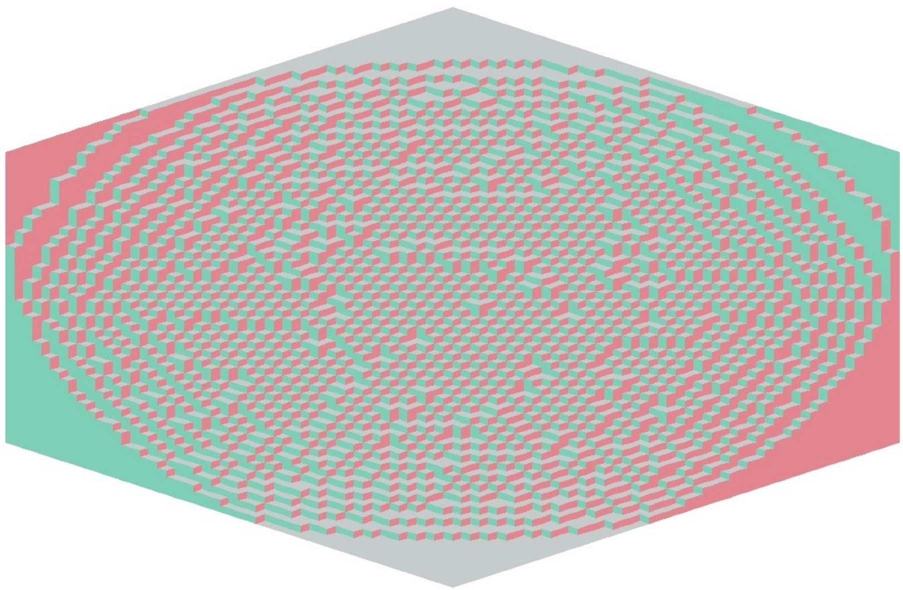

Figure 1: A uniform random tiling of a hexagon. The objects we study can help us understand these tilings quantitatively. For instance, one can prove a "limit shape theorem" explaining the appearance of what appears to be an inscribed circle in this diagram. Figure from [4].

Section 2 gives a survey of the relevant background information, namely the application to lozenge tilings and an explanation of how these distributions arise from characters of $U(\infty)$, as well as the introduction of the aforementioned differential operators. Section 3 contains calculations of the moments in question. Section 4 covers suggestions for future research.

2 Background information

2.1 Lozenge tilings

The motivation for this project comes from its connection to lozenge tilings, which we will briefly explain in this section. For more technical information, refer to [3], section 3.5.

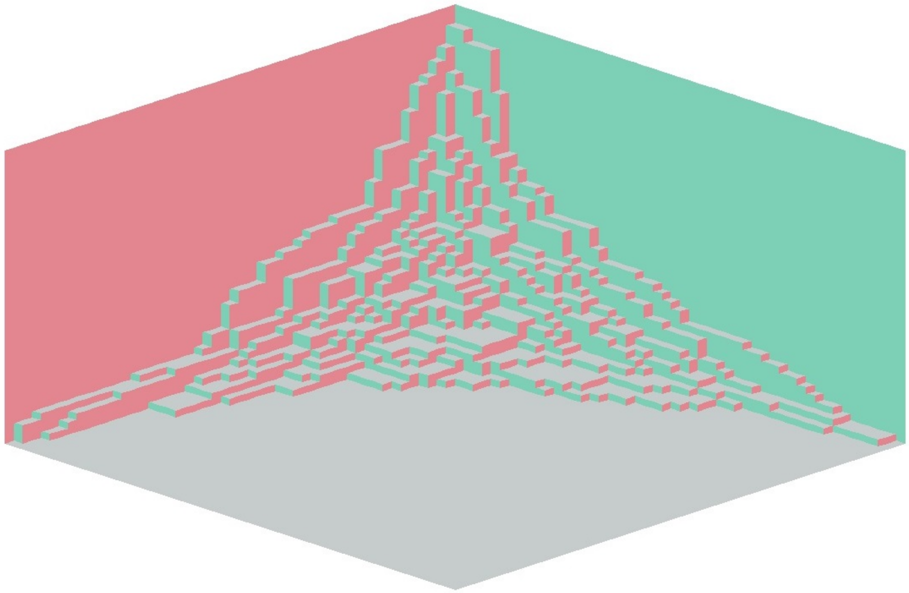

Figure 2: Another example of a randomly tiled hexagon. This sample is drawn from a volume-weighted measure. Figure from [4].

We are interested in studying properties of random tilings of regions in the plane by "lozenges," which are simply rhombuses rhombuses in one of three orientations. On the N-th level, we consider a strip of length N in the plane. (See figure 3.) On the rightmost part of the strip, we choose the location of N horizontal lozenges. Note that above the topmost horizontal lozenge, all of the lozenges must be in the same orientation as the dark green lozenges in figure 3; similarly, below the bottommost horizontal lozenge, all of the lozenges must be in the same orientation as the light green lozenges in figure 3. It suffices, then, to only consider the trapezoidal region prescribed by the top- and bottom-most lozenges, and the left-hand side of the strip. To get the random tiling, we draw uniformly at random from amongst the valid tilings of this trapezoidal region, conditional on the position of the N horizontal lozenges on the right side of the strip.

There is a convenient way of encoding the location of these N horizontal lozenges. Let $y_1 > \ldots > y_N$ be the coordinates of the N horizontal lozenges along the y-axis shown in figure 4. To this set of lozenges, we associate the partition $\lambda = (y_1-N+1, y_2-N+2, \ldots, y_N)$. (Note that, since the y_is are strictly decreasing, the parts of λ are weakly decreasing. Also, since $y_i \geq N-i$, we will have $\lambda_i \geq 0$, so λ is indeed a partition.) From a partition $\lambda = (\lambda_1, \ldots, \lambda_N)$, we obtain the positions of the rightmost lozenges by placing the i-th lozenge at y-coordinate $y_i = \lambda_i + N - i$. In this way, we have defined a bijection.

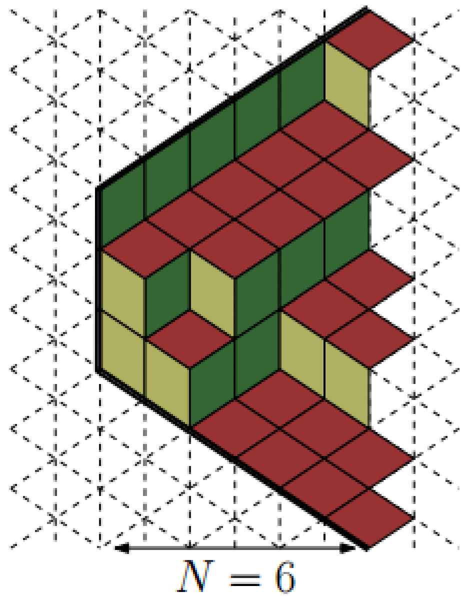

Figure 3: The case $N = 6$. Note that all lozenges above and below the colored trapezoidal region have their orientations determined, regardless of the choice of tiling for the colored region. Figure from [3], p. 24.

Since our distribution on tilings is uniform after we have chosen this set of N rightmost lozenges, we can restrict our attention to studying the distribution of these lozenges—or, equivalently, distributions on partitions of length N.

2.2 Distributions on partitions and connection to $U(\infty)$

We will specifically consider distributions on partitions arising from representations of the infinite unitary group $U(\infty) = \cup_{N=1}^{\infty} U(N)$ and their characters.

Definition 2.1. (From [3], section 3.4) A **character** of $U(\infty)$ is a continuous function $\chi : U(\infty) \to \mathbb{C}$ with the following properties.

1. $\chi(e) = 1$, where e is the identity of $U(\infty)$
2. $\chi(ghg^{-1}) = \chi(h)$ for any elements g and h of $U(\infty)$
3. $[\chi(g_i g_j^{-1}]_{i,j=1}^{n}$ is a Hermitian, positive-definite matrix for any $n \geq 1$ and any choice of $g_1, \ldots, g_n \in U(\infty)$.

For $u \in U(\infty)$ and χ a character of $U(\infty)$, it can be shown that $\chi(u)$ is a symmetric function in the eigenvalues of u. We can decompose the restriction

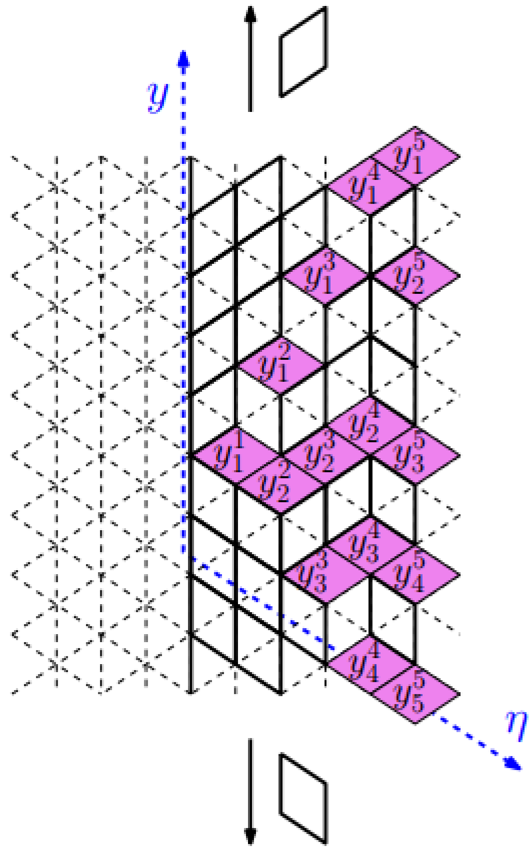

Figure 4: Diagram depicting the two axes used. In this case, $N = 5$, so our the rightmost lozenges are a distance 5 along the η-axis. The y-coordinates of these lozenges are 9, 7, 4, 2, and 0, which corresponds to the partition $(5, 4, 2, 1, 0)$. Figure from [3], p. 23.

of χ to $U(N)$ as a linear combination of Schur functions:

$$\chi|_{U(N)} = \sum_{\ell(\lambda)\leq N} c_\lambda \frac{s_\lambda(u_1, \ldots, u_N)}{s_\lambda(1^N)}.$$

Since $\chi|_{U(N)}(1^N) = 1$, we have that $\sum_\lambda c_\lambda = 1$. Furthermore, one has due to properties of $U(N)$ that $c_\lambda \geq 0$. Thus, setting $\rho(\lambda) = c_\lambda$ defines a probability measure ρ on partitions. Our primary concern is with distributions arising in this way. We can also go in reverse: given a distribution ρ on partitions, we define **character generating function**

$$S_\rho^{U(N)} := \sum_{\ell(\lambda)\le N} \rho(\lambda)\frac{s_\lambda(u_1,\dots,u_N)}{s_\lambda(1^N)}.$$

Once we have a distribution on partitions, we pass to a distribution on $\mathbb{R}$ by choosing first choosing a partition according to our distribution, then choosing a part of the chosen partition uniformly at random. We now turn our attention to the tools required to compute the moments of such a distribution. Our end goal is to compute the moments of the so-called "difference distribution": given a sequence of measures $p(N)$ arising in this way, we want to examine the difference

$$Np(N) - (N-1)p(N-1).$$

This will provide some insight into the relation between the limiting behavior for $U(N)$ and $U(N1)$ coming from measures constructed in such a way.

2.3 Differential operators

Here we review section 4.3 of [2], as it is crucial for our calculations. In order to compute the asymptotic behavior of the moments, we make use of the differential operators

$$\mathcal{D}_k^{U(N)} = \frac{1}{V^{U(N)}} \circ \sum_{i=1}^{N} \left(u_i \frac{\partial}{\partial u_i}\right)^k \circ V^{U(N)}.$$

They will allow us to compute the moments without directly computing the values of $\rho(\lambda)$. Here, $V^{U(N)}$ is to be interpreted as multiplication by the $N \times N$ Vandermonde determinant.

For a partition λ, let $\ell(\lambda)$ denote the length of λ. Consider a character generating function

$$S_\rho^{U(N)}(u_1,\dots,u_N) = \sum_{\ell(\lambda)\le N} \rho(\lambda)\frac{s_\lambda(u_1,\dots,u_N)}{s_\lambda(1^N)},$$

where we have expanded the function in terms of Schur polynomials s_λ. The differential operators are linear, so we examine what happens when we apply it to a Schur polynomial:

$$\mathcal{D}_k^{U(N)} s_\lambda = \frac{1}{V^{U(N)}} \circ \sum_{j=1}^{N} \left(u_j \frac{\partial}{\partial u_j}\right)^k \circ V^{U(N)} \left(\frac{\det(u_j^{\lambda_i+n-i})_{i,j=1}^N}{V^{U(N)}}\right)$$

First, the factors of $V^{U(N)}$ cancel. Then, each time we apply $u_i \frac{\partial}{\partial u_j}$, the derivative gives us a factor of $\lambda_i + N - i$, and then multiplication by u_j "resets" the power of u_j, preserving $\det(u_j^{\lambda_i+n-i})_{i,j=1}^N$. Thus, when we sum over all j and divide again by the Vandermonde, we are left with

$$\begin{aligned}\mathcal{D}_k^{U(N)} s_\lambda &= \left(\sum_{i=1}^N (\lambda_i + N - i)^k\right) \frac{\det(u_j^{\lambda_i+n-i})_{i,j=1}^N}{V^{U(N)}} \\ &= \sum_{i=1}^N (\lambda_i + N - i)^k s_\lambda.\end{aligned}$$

Thus, by the linearity of the operator, we have

$$\mathcal{D}_k^{U(N)} \left(\sum_{\ell(\lambda)\le N} \rho(\lambda) \frac{s_\lambda(u_1,\dots,u_N)}{s_\lambda(1^N)} \right) = \sum_{\ell(\lambda)\le N} \rho(\lambda) \cdot \left[\left(\sum_{i=1}^N (\lambda_i + N - i)^k \right) \frac{s_\lambda(u_1,\dots,u_N)}{s_\lambda(1^N)} \right].$$

Evaluating this expression when all of the u_i are 1 then yields

$$\begin{aligned}\mathcal{D}_k^{U(N)} s_\lambda \Big|_{u_i=1} &= \sum_{\ell(\lambda)\le N} \rho(\lambda) \cdot \left(\sum_{i=1}^N (\lambda_i + N - i)^k \right) \\ &= \mathbb{E}_\rho[X^k],\end{aligned}$$

which is precisely the value of the k-th moment which we seek to compute.

3 Second order asymptotics and difference moments

In this section, we present and prove our main results, namely the second order asymptotic behavior of the moments of the original distribution, and the leading order asymptotic behavior of the difference moments.

3.1 Notation and supporting lemmas

To aid in our analysis later, we will assume that, for all k,

$$\lim_{N\to\infty} \frac{1}{N} \ln(S_{\rho(N)}^{U(N)}(u_1,\dots,u_k,1^{N-k})) = Q(u_1) + \ldots + Q(u_n)$$

for some function Q analytic in a neighborhood of 1.

We make use of lemma 5.5 from [3]. The proof of this lemma can be found there as well.

Lemma 3.1. ([3], lemma 5.5) Take $n > 0$ and a function $g(u)$ analytic in a neighborhood of 1. Then

$$\begin{aligned}\lim_{z_i\to 1} \Bigg(&\frac{g(u_1)}{(u_1-u_2)(u_1-u_3)\cdots(u_1-u_n)} + \frac{g(u_2)}{(u_2-u_1)(u_2-u_3)\cdots(u_2-u_n)} \\ &+ \ldots + \frac{g(u_n)}{(u_n-u_1)(u_n-u_2)\cdots(u_n-u_{n-1})} \Bigg) = \frac{\partial^{n-1}}{\partial u^{n-1}} \left(\frac{g(u)}{(n-1)!} \right) \Bigg|_{u=1}.\end{aligned}$$

3.2 Asymptotics for $Q(u) = \gamma(u-1)$

For simplicity, we'll work in the scenario where $S^{U(N)}_{\rho(N)} = \prod_{i=1}^N \exp(\gamma N(u_i - 1))$, so $Q(u) = \gamma(u-1)$. We wish to compute

$$\mathcal{D}_k^{U(N)} S_\rho^{U(N)} \bigg|_{u_i=1}.$$

Using the fact that $\partial u_i \ u_i = 1 + u_i \partial u_i$, we can expand $(u_i \partial u_i)^k$ as a sum $\sum_{j=0}^k c_j u_i^j \partial u_i^j$.

By the product rule, applying $u_a^k \partial u_a^k$ breaks down into terms of the form

$$u_a^k \frac{\partial u_a^\ell V}{V} \partial u_a^{k-\ell} \exp\left(\gamma N \sum_{i=1}^N (u_i - 1)\right).$$

We can directly compute

$$\partial u_a^{k-\ell} \exp\left(\gamma N \sum_{i=1}^N (u_i - 1)\right) = \gamma^{k-\ell} N^{k-\ell} \exp\left(\gamma N \sum_{i=1}^N (u_i - 1)\right).$$

Since the value of the exponential is 1 when all of the u_is are sent to 1, we can ignore it. As we've seen before, $\frac{\partial u_a^\ell V}{V}$ gives us terms of the form

$$\frac{1}{\prod_{j \in P} (u_a - u_i)}$$

for sets P of indices not containing a with $|P| = \ell$. By the symmetry of the expansion, for every choice of $\ell + 1$ elements $P \cup \{a\}$, we get a sum of terms of the form indicated in lemma 5.5 of [2] with $g(z) = z^k$ and $n = \ell + 1$. There are $\binom{N}{\ell+1}$ ways to choose these $\ell + 1$ variables, which gives us a factor of $\binom{N}{\ell+1}$. Furthermore, there are $k(k-1)\cdots(k-\ell+1)$ ways to apply ℓ of the k factors to the Vandermonde, so we pick up this factor as well. Applying the result of lemma 5.5 leaves us with

$$\frac{N(N-1)\cdots(N-\ell)}{\ell+1} \binom{k}{\ell}^2 \gamma^{k-\ell} N^{k-\ell}.$$

The coefficient on the N^ℓ term of $N(N-1)\cdots(N-\ell)$ is $-k(k+1)/2$, so when we sum over all choices of ℓ, we get a second order contribution of

$$-\sum_{\ell=0}^k \frac{\ell(\ell+1)}{2(\ell+1)} \binom{k}{\ell}^2 \gamma^{k-\ell} N^k.$$

When there are $N-1$ variables in play, all of the calculations done here are the same, except that we must choose $\ell + 1$ variables from amongst $N-1$. We therefore want the coefficient of N^ℓ in $(N-1)(N-2)\cdots(N-\ell-1)$, which is $-(\ell+2)(\ell+1)/2$. We get therefore get a second order contribution of

$$-\sum_{\ell=0}^{k}\frac{(\ell+2)}{2}\binom{k}{\ell}^2\gamma^{k-\ell}N^k.$$

The only other contributions to second order terms will come from the highest order term gained from $u_i^{k-1}\partial u_i^{k-1}$. (As we have seen from the previous two cases, the maximum order of N obtained by applying r derivatives is N^r.) However, it's easy to check that this is the same for levels N and $N-1$. As before, the only difference between the two is that we have a choice of $\binom{N}{\ell+1}$ variables on level N and $\binom{N-1}{\ell+1}$ on level $N-1$. But the coefficient on the highest order term for these two cases is the same, so these terms will cancel. The scaled k-th difference moment is therefore

$$\frac{1}{N^k}\left[-\sum_{\ell=0}^{k}\frac{\ell}{2}\binom{k}{\ell}^2\gamma^{k-\ell}N^k-\left(-\sum_{\ell=0}^{k}\frac{(\ell+2)}{2}\binom{k}{\ell}^2\gamma^{k-\ell}N^k\right)\right].$$

Combining these sums termwise, we are left with

$$\sum_{\ell=0}^{k}\binom{k}{\ell}^2\gamma^{k-\ell}.$$

4 Suggestions for future research

Though we have successfully calculated the first order asymptotics of the difference moments, there remain several related questions on the topic to be addressed.

4.1 Asymptotics for general Q

Following the proof of theorem 5.1 in [2], we write

$$S_{\rho(N)}^{U(N)}=\exp\left(\sum_{i=1}^{N}NQ(u_i)\right)T_N(u_1,\ldots,u_N).$$

By the logic from this same proof, any term in which a derivative is applied to T_N will vanish asymptotically, and therefore can be safely ignored.

We can follow the same steps as in section 3.2 until we arrive at

$$\partial u_a^{k-\ell}\exp\left(\sum_{i=1}^{N}NQ(u_i)\right).$$

At this point, one need only keep track of terms generated by the chain rule and product rule. The results of this straightforward (albeit tedious) calculation will be included in a later version of the paper.

4.2 Non-negativity of the difference distribution

Throughout this article, we have referred to a difference "distribution," but at present it remains to be shown that this difference is indeed a probability distribution, rather than a signed measure. If $p(N)$ is the distribution arising from $U(N)$, it is clear that the diffence measure

$$D(N) = Np(N) - (N-1)p(N-1)$$

has total mass 1: since $p(N)$ and $p(N-1)$ are both probability distributions, they both have mass 1, and we have

$$\int_{\mathbb{R}} dD(N) = \int_{\mathbb{R}} N\ dp(N) - \int_{\mathbb{R}} (N-1)\ dp(N-1) = N - (N-1) = 1.$$

This is, of course, not sufficient for $\delta(N)$ to be a probability measure. We must also show that it is nonnegative, and it is not clear a priori that this is the case.

Straightforward algebraic methods do not appear to be sufficient to solve this problem. Even in the simplest cases, the calculations quickly become unwieldy and intractible. Indeed, it may not even be the case that $D(N)$ is a probability measure for all N, but rather only has the non-negativity property for large enough values of N.

4.3 Central limit theorems

Knowing the mean of the k-th power of the difference distribution, it is natural to ask: what can we say about the fluctuation of the random variable about these moments? That is, what can be proven quantitatively about

$$\lim_{N\to\infty} N^{-\alpha}(X^k - \mathbb{E}[X^k]),$$

where X is a random variable distributed according to the N-th difference distribution, and α is an appropriate rescaling constant? In many similar scenarios—see, e.g., [2]—, it turns out that the fluctuations will asymptotically behave like a Gaussian distribution. It seems likely that a similar central limit theorem would hold in this case. (Of course, for any such results to make sense, we assume that that difference distribution is, at least asymptotically, a probability distribution.)

5 Acknowledgements

First, I would like to thank my mentor, Christopher Ryba, for the immense help he provided on this project. I would also like to thank Professor Gorin both for providing the problem and for always being accessible and providing guidance along the way. Lastly, I would like to thank the MIT math department and the UROP+ program for the opportunity to work on this project.

References

[1] Alexei Borodin, Alexey Bufetov, and Grigori Olshanski. *Limit shapes for growing extreme characters of* $U(\infty)$. The Annals of Applied Probability, Vol. 25, No. 4, 2339-2381, 2015. `http://arxiv.org/pdf/1311.5697v2.pdf`.

[2] Alexey Bufetov and Vadim Gorin. *Representations of classical lie groups and quantized free convolution.* `http://arxiv.org/pdf/1311.5780v2.pdf`.

[3] Alexey Bufetov and Vadim Gorin. *Fluctuations of particle systems determined by Schur generating functions.* `http://arxiv.org/pdf/1604.01110v1.pdf`.

[4] Professor Vadim Gorin's homepage. `http://www.mccme.ru/ vadicgor/research.html`.