



**HAL**  
open science

# On the Vershik–Kerov Conjecture Concerning the Shannon–McMillan–Breiman Theorem for the Plancherel Family of Measures on the Space of Young Diagrams

Alexander I. Bufetov

► **To cite this version:**

Alexander I. Bufetov. On the Vershik–Kerov Conjecture Concerning the Shannon–McMillan–Breiman Theorem for the Plancherel Family of Measures on the Space of Young Diagrams. *Geometric And Functional Analysis*, 2012, 10.1007/s00039-012-0169-4 . hal-01256210

**HAL Id: hal-01256210**

**<https://hal.science/hal-01256210v1>**

Submitted on 14 Jan 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**ON THE VERSHIK-KEROV CONJECTURE CONCERNING  
THE SHANNON-MCMILLAN-BREIMAN THEOREM  
FOR THE PLANCHEREL FAMILY OF MEASURES  
ON THE SPACE OF YOUNG DIAGRAMS.**

ALEXANDER I. BUFETOV

*To Alesha*

CONTENTS

1. Introduction.	2
1.1. The Vershik-Kerov Conjecture.	2
1.2. The Shannon-McMillan-Breiman Theorem	3
1.3. Outline of the Proof of Theorem 1.1.	4
1.4. Acknowledgements.	4
2. The Vershik-Kerov Variational Formula	5
2.1. The Limit Shape of Plancherel Young Diagrams.	5
2.2. The Quadratic Variation of the Hook Integral	6
2.3. Proof of Lemma 2.4.	7
3. Poissonization.	8
3.1. Diagrams and Sequences.	8
3.2. The Bessel Point Process.	8
3.3. Depoissonization.	9
3.4. The Debye Asymptotics.	11
4. Local Patterns in Plancherel Young Diagrams.	12
4.1. The Discrete Sine-Process.	12
4.2. The Variance of the Discrete Sine-Process.	12
4.3. Decay of Correlations for the Plancherel Measure.	14
4.4. Frequency of Local Patterns.	14
5. Proof of Lemmas 4.3, 4.4, 4.5, 2.1, 2.2.	15
5.1. Decay of Correlations for the Bessel Point Process.	15
5.2. Proof of Lemma 4.3.	16
5.3. Proof of Lemma 4.4.	17
5.4. Proof of Lemma 2.1.	18
5.5. Proof of Lemmas 2.2, 4.5.	19
6. Proof of Lemma 2.3.	19
6.1. Outline of the Proof.	19
6.2. Proof of Lemma 6.1	22

6.3.	Average Variance of the Discrete Bessel Process.	25
6.4.	Estimates on the Bessel Kernel.	26
6.5.	Proof of Lemma 6.2.	28
7.	Proofs of Estimates for the Discrete Bessel Kernel.	31
7.1.	Proof of Lemma 6.7	31
7.2.	Proof of Lemma 6.3.	35
	References	38

## 1. INTRODUCTION.

**1.1. The Vershik-Kerov Conjecture.** Let  $n \in \mathbb{N}$  and let  $\mathbb{Y}_n$  be the set of Young diagrams with  $n$  cells. For  $\lambda \in \mathbb{Y}_n$  let  $\dim \lambda$  be the dimension of the irreducible representation of the symmetric group on  $n$  elements corresponding to  $\lambda$ . The *Plancherel* probability measure  $\mathbb{P}l^{(n)}$  on  $\mathbb{Y}_n$  is given by the formula

$$\mathbb{P}l^{(n)}(\lambda) = \frac{\dim^2 \lambda}{n!}.$$

In 1985 Vershik and Kerov [17] showed that there exist two positive constants  $\alpha_1, \alpha_2$  such that

$$\lim_{n \rightarrow \infty} \mathbb{P}l^{(n)}\{\lambda \in \mathbb{Y}_n : \alpha_1 \sqrt{n} \leq -\log \mathbb{P}l^{(n)}(\lambda) \leq \alpha_2 \sqrt{n}\} = 1.$$

and conjectured that the sequence of random variables  $\frac{-\log \mathbb{P}l^{(n)}(\lambda)}{\sqrt{n}}$  converges to a constant according to the Plancherel measure.

The main result of this paper is the proof of the Vershik-Kerov conjecture:

**Theorem 1.1.** *There exists a constant  $H > 0$  such that for any  $\varepsilon > 0$  we have*

$$(1) \quad \lim_{n \rightarrow \infty} \mathbb{P}l^{(n)} \left\{ \lambda \in \mathbb{Y}_n : \left| H + \frac{\log \mathbb{P}l^{(n)}(\lambda)}{\sqrt{n}} \right| \leq \varepsilon \right\} = 1.$$

Theorem 1.1 immediately implies

**Corollary 1.2** (Asymptotic Equidistribution for the Plancherel Measure). *For any  $\varepsilon > 0$  there exists  $n_0 > 0$  such that for any  $n \in \mathbb{N}$ ,  $n > n_0$ , there exists a subset  $\mathbb{Y}_n(\varepsilon) \subset \mathbb{Y}_n$  with the following properties:*

(1) *The cardinality  $\#\mathbb{Y}_n(\varepsilon)$  of the set  $\mathbb{Y}_n(\varepsilon)$  satisfies the inequality*

$$e^{(H-\varepsilon)\sqrt{n}} \leq \#\mathbb{Y}_n(\varepsilon) \leq e^{(H+\varepsilon)\sqrt{n}}.$$

(2) *For each  $\lambda \in \mathbb{Y}_n(\varepsilon)$  we have*

$$e^{-(H+\varepsilon)\sqrt{n}} \leq \mathbb{P}l^{(n)}(\lambda) \leq e^{-(H-\varepsilon)\sqrt{n}}.$$

Vershik and Kerov have suggested to call  $H$  the *entropy* of the Plancherel measure. Numerical experiments allowing to estimate the entropy  $H$  of the Plancherel measure are given in [19]. An explicit formula for  $H$  is given below in (15).

**1.2. The Shannon-McMillan-Breiman Theorem.** The term “entropy” is suggested by the following analogy. Let  $W(N)$  be the set of all binary words of length  $N$ . Take  $p \in (0, 1)$ , and let  $\mathbb{P}_{N,p}$  be the Bernoulli measure which to a word  $w \in W(N)$  with  $k$  zeros assigns the weight

$$\mathbb{P}_{N,p}(w) = p^k \cdot (1-p)^{N-k}.$$

Let

$$H(p) = -p \log p - (1-p) \log(1-p)$$

be the entropy of the Bernoulli measure. Shannon’s Theorem then says that for any  $p \in (0, 1)$  and any  $\varepsilon > 0$  we have

$$\lim_{N \rightarrow \infty} \mathbb{P}_{N,p} \left( \left\{ w \in W(N) : \left| \frac{\log \mathbb{P}_{N,p}(w)}{N} + H(p) \right| > \varepsilon \right\} \right) = 0.$$

Using Kolmogorov’s Strong Law of Large Numbers, Shannon’s Theorem can be strengthened to a pointwise statement, the Shannon-McMillan-Breiman Theorem (for a detailed discussion of the Shannon-McMillan-Breiman Theorem see e.g. [6], [7], [14]). Similarly, Vershik and Kerov have given a pointwise analogue of their conjecture. Let  $\mathbb{T}(\mathbb{Y})$  be the space of infinite *Young tableaux*, that is, infinite directed paths in the Young graph starting at the origin. The sequence  $\left( \mathbb{P}^{(n)} \right)_{n \in \mathbb{N}}$  gives rise to a natural Markov measure on  $\mathbb{T}(\mathbb{Y})$ ; that measure is denoted  $\mathbb{P}^l$  and called the Plancherel measure on  $\mathbb{T}(\mathbb{Y})$  (see [17] for details). Vershik and Kerov conjectured that for  $\mathbb{P}^l$ -almost all paths

$$(\lambda^{(n)})_{n \in \mathbb{N}} \in \mathbb{T}(\mathbb{Y})$$

we have

$$\lim_{n \rightarrow \infty} \frac{-\log \mathbb{P}^{(n)}(\lambda^{(n)})}{\sqrt{n}} = H.$$

The pointwise conjecture remains open.

Convergence in  $L_p$  for  $p < \infty$ , on the other hand, can be obtained simultaneously with the convergence in measure. Let  $\mathbb{E}_{\mathbb{P}^{(n)}}$  stand for the expectation with respect to the Plancherel measure.

**Corollary 1.3.** *There exists a constant  $H > 0$  such that for any  $p$ ,  $0 < p < \infty$ , we have*

$$\lim_{n \rightarrow \infty} \mathbb{E}_{\mathbb{P}^{(n)}} \left| H + \frac{\log \mathbb{P}^{(n)}(\lambda)}{\sqrt{n}} \right|^p = 0.$$

Indeed, by the Euler-Hardy-Ramanujan Formula, the number of Young diagrams with  $n$  cells does not exceed  $\exp(2\pi\sqrt{n}/\sqrt{6})$ , whence

$$\mathbb{P}^{l^{(n)}} \left\{ \lambda : \frac{-\log \mathbb{P}^{l^{(n)}}(\lambda)}{\sqrt{n}} > K \right\} \leq \exp((-K + 2\pi/\sqrt{6})\sqrt{n}),$$

and Corollary 1.3 is immediate from Theorem 1.1 .

**1.3. Outline of the Proof of Theorem 1.1.** The first step, due to Vershik and Kerov [17], is a variational formula for the normalized logarithm of the Plancherel measure (see Subsection 2.2). Using the hook formula, Vershik and Kerov represent the normalized logarithm of the Plancherel measure as a special double integral, called *the hook integral*. The hook integral admits a unique minimum – the Vershik-Kerov-Logan-Shepp limit shape. The Vershik-Kerov variational formula (2) is an explicit expression for the quadratic variation of the hook integral.

The next step is the Theorem established, independently and simultaneously, by Borodin, Okounkov and Olshanski [3] and Johansson [9], which claims that the poissonization of the Plancherel measure is the discrete Bessel determinantal point process. Using this Theorem, Borodin, Okounkov and Olshanski showed that local patterns in the bulk of a Plancherel Young diagram are governed by the discrete sine-process.

The Vershik-Kerov Variational Formula has two types of terms: the *local* terms and the *nonlocal* terms. For the local terms, the Borodin-Okounkov-Olshanski Theorem is averaged along the boundary of the Young diagram, and it is shown that the normalized number of appearances of a given local pattern in a Young diagram converges to a constant with respect to the Plancherel measure (Lemma 4.4). In particular, for  $k \in \mathbb{N}$ , it is shown that the normalized number of cells with hook length  $k$  converges to a constant according to the Plancherel measure (Lemma 2.1). The proof relies on a simple upper estimate for the decay of correlations of the Plancherel measure (Lemma 4.3).

The final step is to show that the nonlocal terms of the Vershik-Kerov formula converge to 0 according to the Plancherel measure (Lemma 2.3, proved in Section 6). The proof relies on upper estimates for the variance of the Bessel point process and the Plancherel measure, which are obtained using the classical contour integral representations for Bessel functions and the Okounkov contour integral representation for the discrete Bessel kernel (Section 7).

**1.4. Acknowledgements.** Grigori Olshanski posed the problem to me and suggested the poissonization approach; I am deeply grateful to him. I am deeply grateful to Alexei Borodin who suggested the use of Okounkov's

contour integral for the discrete Bessel kernel. I am deeply grateful to Elena Rudo for her careful reading of the manuscript and for many very helpful suggestions on improving the presentation. I am deeply grateful to Sevak Mkrtchyan, Fedor Petrov, Alexander Soshnikov, Konstantin Tolmachov and Anatoly M. Vershik for helpful discussions. I am deeply grateful to the referee for many useful comments. I am deeply grateful to Nikita Kozin for typesetting parts of the manuscript. This work was supported in part by an Alfred P. Sloan Research Fellowship, by the Grant MK-4893.2010.1 of the President of the Russian Federation, by the Programme on Mathematical Control Theory of the Presidium of the Russian Academy of Sciences, by the Programme 2.1.1/5328 of the Russian Ministry of Education and Research, by the RFBR-CNRS grant 10-01-93115, by the RFBR grant 11-01-00654, by the Edgar Odell Lovett Fund at Rice University and by the National Science Foundation under grant DMS 0604386.

## 2. THE VERSHIK-KEROV VARIATIONAL FORMULA

**2.1. The Limit Shape of Plancherel Young Diagrams.** Take a Young diagram  $\lambda = (\lambda_1, \lambda_2, \dots)$  (setting  $\lambda_i = 0$  for all large  $i$ ). Introduce a piecewise-linear function  $\Phi_\lambda$  in the following way: we set  $\Phi'_\lambda|_{(k,k+1)} = -1$  if  $k = \lambda_i - i$  for some  $i$ , we set  $\Phi'_\lambda|_{(k,k+1)} = 1$  otherwise, and we require that the equality  $\Phi_\lambda(t) = |t|$  hold for all sufficiently large  $t$  (it is easy to see that the continuous function  $\Phi_\lambda$  is uniquely defined by these requirements; it is differentiable except at integer points).

The function  $\Phi_\lambda$  admits the following combinatorial interpretation. Assume that the cells of our diagram are squares with diagonal 2. Following Vershik and Kerov [17], rotate the diagram  $\lambda$  by  $\pi/4$ ; the boundary of the rotated diagram forms the graph of  $\Phi_\lambda$ , while “beyond” the diagram, for all sufficiently large  $|t|$ , we have  $\Phi_\lambda(t) = |t|$  (see Fig. 1 on p. 482 in [3]).

Following Vershik and Kerov, introduce the function

$$\Omega(t) = \begin{cases} \frac{2}{\pi}(t \arcsin(t/2) + \sqrt{4-t^2}), & \text{if } |t| \leq 2; \\ |t|, & \text{if } |t| > 2, \end{cases}$$

and denote

$$F_\lambda(t) = \Phi_\lambda(t) - \sqrt{n}\Omega(t/\sqrt{n}).$$

By definition, the function  $F_\lambda$  has compact support. The functions  $\Phi_\lambda$  and  $\Omega$  are Lipschitz with constant 1, therefore the function  $F_\lambda$  is Lipschitz with constant 2.

Vershik and Kerov [16] and, independently and simultaneously, Logan and Shepp [11] have shown that for any  $\varepsilon > 0$  we have

$$\lim_{n \rightarrow \infty} \mathbb{P}^{(n)} \{ \lambda \in \mathbb{Y}_n : |F_\lambda(t)/\sqrt{n}| \leq \varepsilon \} = 1.$$

**2.2. The Quadratic Variation of the Hook Integral.** Recall that the *hook length* of a cell in a Young diagram is the number of cells to the right of it and under it (including the cell itself) and let  $h_k(\lambda)$  stand for the number of cells in  $\lambda$  with hook length  $k$ .

Denote

$$\|F_\lambda\|_{1/2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left( \frac{F_\lambda(t) - F_\lambda(s)}{t - s} \right)^2 dt ds = 2 \int_0^{\infty} \int_{-\infty}^{\infty} \left( \frac{F_\lambda(t+h) - F_\lambda(t)}{h} \right)^2 dt dh.$$

The Vershik-Kerov Variational Formula (see [17], Lemma 1 and formulas (5), (8), (9)) is the equality

$$(2) \quad \frac{-\log \mathbb{P}l^{(n)}(\lambda)}{\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{k=1}^{\infty} h_k(\lambda) \left( \sum_{l=1}^{\infty} \frac{1}{l(l+1)(2l+1)k^{2l}} \right) + \frac{1}{8\sqrt{n}} \|F_\lambda\|_{1/2} + \frac{1}{\sqrt{n}} \int_{|t| \geq 2\sqrt{n}} F_\lambda(t) \operatorname{arccosh} \left( \frac{t}{2\sqrt{n}} \right) dt - \varepsilon_n,$$

where  $\varepsilon_n$  only depends on  $n$  (not on  $\lambda$ ) and tends to 0 as  $n \rightarrow \infty$ .

It will be convenient for us to adopt the following terminology. Assume that for each  $n \in \mathbb{N}$  we are given a random variable  $\xi_n$  on  $\mathbb{Y}_n$ . If there exists  $\beta$  such that for any  $\varepsilon > 0$  we have

$$\lim_{n \rightarrow \infty} \mathbb{P}l^{(n)} \{ \lambda \in \mathbb{Y}_n : |\xi_n(\lambda) - \beta| \leq \varepsilon \} = 1,$$

then we say that  $\xi_n$  converges to the constant  $\beta$  according to the Plancherel measure.

If

$$\lim_{n \rightarrow \infty} \mathbb{P}l^{(n)} \{ \lambda \in \mathbb{Y}_n : \xi_n(\lambda) < \beta \} = 1,$$

we say that  $\xi_n$  is asymptotically majorated by  $\beta$  according to the Plancherel measure.

We shall analyze the terms of the Vershik-Kerov Variational Formula one by one.

**Lemma 2.1.** *For any  $k \in \mathbb{N}$  the random variables  $\frac{h_k(\lambda)}{\sqrt{n}}$  converge to the constant  $\frac{32k^2}{(4k^2-1)\pi^2}$  according to the Plancherel measure.*

**Lemma 2.2.** *For any  $h_0 > 0$  the random variables*

$$\frac{1}{\sqrt{n}} \int_0^{h_0} \int_{-\infty}^{\infty} \left( \frac{F_\lambda(t+h) - F_\lambda(t)}{h} \right)^2 dt dh$$

*converge to a constant according to the Plancherel measure.*

**Lemma 2.3.** *For any  $\varepsilon > 0$  there exists  $h_0 > 0$  such that the random variables*

$$\frac{1}{\sqrt{n}} \int_{h_0}^{\infty} \int_{-\infty}^{\infty} \left( \frac{F_{\lambda}(t+h) - F_{\lambda}(t)}{h} \right)^2 dt dh$$

*are asymptotically majorated by  $\varepsilon$  according to the Plancherel measure.*

**Lemma 2.4.** *The random variables*

$$(3) \quad \frac{1}{\sqrt{n}} \int_{|t| \geq 2\sqrt{n}} F_{\lambda}(t) \operatorname{arccosh} \left( \frac{t}{2\sqrt{n}} \right) dt$$

*converge to 0 according to the Plancherel measure.*

As before, let  $\lambda_1$  be the length of the first row of  $\lambda$ , and let  $\lambda'_1$  be the length of the first column of  $\lambda$ . Vershik and Kerov [17] established that for any  $\varepsilon > 0$  we have

$$(4) \quad \lim_{n \rightarrow \infty} \mathbb{P}l^{(n)} \left( \{ \lambda : \lambda_1 < (2 + \varepsilon)\sqrt{n}, \lambda'_1 < (2 + \varepsilon)\sqrt{n} \} \right) = 1.$$

Theorem 1.1 is now immediate from the bound (4), the Vershik-Kerov Variational Formula and the Lemmas 2.1, 2.2, 2.3, 2.4.

We proceed to the proof of the Lemmas.

**2.3. Proof of Lemma 2.4.** We shall need a more precise estimate than (4).

**Proposition 2.5.** (1) *For any  $\delta_0 > 1/6$  there exists constants  $C > 0$ ,  $\tilde{\gamma} > 0$  such that for all  $\delta$  satisfying  $\delta_0 \leq \delta \leq 1/2$  we have*

$$(5) \quad \lim_{n \rightarrow \infty} \mathbb{P}l^{(n)} \left( \{ \lambda : \lambda_1 > 2\sqrt{n} + n^{\delta} \text{ or } \lambda'_1 > 2\sqrt{n} + n^{\delta} \} \right) \leq C \exp \left( -\tilde{\gamma} n^{3\delta/2 - 1/4} \right).$$

(2) *For any  $\varepsilon_1 > 0$  there exists  $\tilde{\gamma}_1 > 0$  depending only on  $\varepsilon_1$  such that for any  $\varepsilon > \varepsilon_1$  we have*

$$(6) \quad \lim_{n \rightarrow \infty} \mathbb{P}l^{(n)} \left( \{ \lambda : \lambda_1 > (2 + \varepsilon)\sqrt{n} \text{ or } \lambda'_1 > (2 + \varepsilon)\sqrt{n} \} \right) \leq C \exp \left( -\tilde{\gamma}_1 \varepsilon n \right).$$

Proposition 2.5 is well-known. For completeness of the exposition a proof is given below (see Proposition 6.11).

Now, using Proposition 2.5, choose  $\delta > \frac{1}{6}$  and assume that

$$F_{\lambda}(t) = 0 \quad \text{for } |t| > 2\sqrt{n} + n^{\delta}$$

In this case for  $|t| \in [2\sqrt{n}, 2\sqrt{n} + n^{\delta}]$  we have:

$$|F_{\lambda}(t)| \leq 2n^{\delta}, \quad \left| \operatorname{arccosh} \left( \frac{t}{2\sqrt{n}} \right) \right| \leq 2n^{\frac{\delta}{2} - \frac{1}{4}},$$



whence

$$\frac{1}{\sqrt{n}} \left| \int_{|t| \geq 2\sqrt{n}} F_\lambda(t) \operatorname{arccosh} \left( \frac{t}{2\sqrt{n}} \right) dt \right| \leq 32n^{\frac{5\delta}{2} - \frac{3}{4}},$$

and, as soon as  $\delta < \frac{3}{10}$ , we are done.

We proceed to the analysis of the remaining terms.

### 3. POISSONIZATION.

**3.1. Diagrams and Sequences.** Let  $\Omega_2 = \{0, 1\}^{\mathbb{Z}}$  be the space of bi-infinite sequences of the symbols 0, 1:

$$\Omega_2 = \{\omega = \dots \omega(-n) \dots \omega(n) \dots, \omega(n) \in \{0, 1\}\}.$$

To a sequence  $\omega$  we assign a continuous piecewise-linear function  $\Phi_\omega$  in the following way: we set  $\Phi_\omega(0) = 0$ ,  $\Phi'_\omega|_{(k, k+1)} = 1$  if  $\omega_k = 0$ ,  $\Phi'_\omega|_{(k, k+1)} = -1$  if  $\omega_k = 1$  (it is easy to see that the continuous function  $\Phi_\omega$  is uniquely defined by these requirements; it is differentiable except at integer points).

Take a Young diagram  $\lambda = (\lambda_1, \lambda_2, \dots)$  (setting  $\lambda_i = 0$  for all large  $i$ ) and introduce a sequence  $c(\lambda) \in \Omega_2$  by the rule  $c_k(\lambda) = 1$  if  $k = \lambda_i - i$  for some  $i$  and  $c_k(\lambda) = 0$  otherwise. It is clear from the definitions that the difference  $\Phi_\lambda - \Phi_{c(\lambda)}$  is a constant.

Take an integer vector  $\vec{m} = (m_1, \dots, m_r)$  all whose coordinates are distinct and for a Young diagram  $\lambda$  denote

$$c_{\vec{m}}(\lambda) = c_{m_1}(\lambda) \dots c_{m_r}(\lambda).$$

Similarly, for  $\omega \in \Omega_2$  write

$$c_{\vec{m}}(\omega) = \omega_{m_1} \dots \omega_{m_r}.$$

In what follows, when we speak of integer vectors, we shall always assume that all their coordinates are distinct.

**3.2. The Bessel Point Process.** Set

$$\mathbb{Y} = \bigcup_{n=1}^{\infty} \mathbb{Y}_n,$$

and for  $\eta > 0$  let

$$Pois_\eta = \exp(-\eta) \sum_{n=0}^{\infty} \frac{\eta^n}{n!} \mathbb{P}l^{(n)}$$

be the  $\eta$ -poissonized Plancherel measure on  $\mathbb{Y}$ .

At the centre of our argument lies a theorem obtained by Borodin, Okounkov and Olshanski in [3] and Johansson in [9] which states that the

measure  $Pois_\eta$  naturally induces the Bessel determinantal point process on the space of sequences of two symbols. We proceed to the exact formulation and start by recalling the definition of a determinantal point process on  $\Omega_2$  (for a more detailed exposition, see, e.g., [15]).

Let  $\mathcal{K} : \ell_2(\mathbb{Z}) \rightarrow \ell_2(\mathbb{Z})$  be a self-adjoint positive contraction, or, in other words, a self-adjoint linear operator satisfying

$$0 \leq \langle f, \mathcal{K}f \rangle \leq \langle f, f \rangle.$$

Set

$$\mathcal{K}(x, y) = \langle \mathcal{K}\delta_y, \delta_x \rangle.$$

There exists a unique probability measure  $\mathbb{P}_{\mathcal{K}}$  on  $\Omega_2$  such that

$$\mathbb{E}_{\mathbb{P}_{\mathcal{K}}}(c_{\vec{m}}) = \det(\mathcal{K}(m_i, m_j)) \Big|_{i,j=1,\dots,r}.$$

Now set  $\eta = \theta^2$  (assuming  $\theta > 0$ ) and, for  $x \neq y$ , write

$$\mathcal{J}(\theta^2; x, y) = \theta \frac{J_x(2\theta)J_{y+1}(2\theta) - J_{x+1}(2\theta)J_y(2\theta)}{x - y}.$$

The expression  $\mathcal{J}(\theta^2; x, x)$  is defined using the l'Hospital Rule. The kernel  $\mathcal{J}(\theta^2)$  is called *the discrete Bessel kernel*, and the resulting measure  $\mathbb{P}_{\mathcal{J}(\theta^2)}$  on  $\Omega_2$  is called *the Bessel point process*.

Recall that to a Young diagram  $\lambda$  we have assigned a sequence  $c(\lambda) \in \Omega_2$ . Slightly abusing notation, we denote the push-forward of the measure  $Pois_\eta$  on  $\mathbb{Y}$  under the map  $\lambda \rightarrow c(\lambda)$  by the same symbol  $Pois_\eta$ .

The theorem of Borodin, Okounkov and Olshanski [3] and Johansson [9] states that the measure  $Pois_\eta$  defined above is precisely the Bessel point process with parameter  $\eta = \theta^2$ .

**3.3. Depoissonization.** Information about the Plancherel measure will be derived from the corresponding properties of the Bessel point process with the use of the following lemma of Borodin, Okounkov and Olshanski (a slight modification of Lemma 3.1 in [3]).

**Lemma 3.1** (Borodin, Okounkov, Olshanski). *Let  $0 < \alpha < 1/4$ . Let  $\{f_n\}$  be a sequence of entire functions*

$$(7) \quad f_n(z) = \exp(-z) \sum_{k \geq 0} \frac{f_{nk}}{k!} z^k, \quad n = 1, \dots$$

*and assume that there exist constants  $f_\infty, \gamma, C_1, C_2$  such that*

- (1)  $\max_{|z|=n} |f_n(z)| \leq C_1 \exp(\gamma\sqrt{n})$ ;
- (2)  $\max_{|z-n| < n^{1-\alpha}} |f_n(z) - f_\infty| \exp(-\gamma|z-n|/\sqrt{n}) \leq C_2$ .

Then there exists a constant  $C = C(\gamma, C_1, C_2)$  such that for all  $n > 0$  we have

$$|f_{nn} - f_\infty| \leq C.$$

The proof is identical to the proof of Lemma 3.1 in [3] except that  $o(1)$  in the last two formulas on page 495 must be replaced by  $O(1)$ .

To use Lemma 3.1 we must allow complex values of the poissonization parameter  $\theta^2$ : in this case, expressions such as  $\mathbb{E}_{g(\theta^2)}$  are understood formally (by analytic continuation).

Lemma 3.1 can be equivalently reformulated as follows.

**Lemma 3.2.** *Let  $\delta > 0$  be arbitrary, let  $\alpha$  satisfy  $0 < \alpha < \frac{1}{4}$ . Assume that there exist constants  $f_\infty, \gamma_1, \gamma_2, \gamma_3, C_1, C_2, C_3 > 0$  such that*

- 1)  $\max_{|z-n| < n^\delta} |f_n(z) - f_\infty| e^{-\frac{\gamma_1|z-n|}{\sqrt{n}}} \leq C_1,$
- 2)  $\max_{|z-n| < n^{1-\alpha}} |f_n(z)| e^{-\frac{\gamma_2|z-n|}{\sqrt{n}}} \leq C_2,$
- 3)  $\max_{|z|=n} |f_n(z)| \leq C_3 e^{\gamma_3 \sqrt{n}}.$

Then there exists a constant  $C = C(\gamma_1, \gamma_2, \gamma_3, C_1, C_2, C_3)$  such that for all  $n > 0$  we have

$$|f_{nn} - f_\infty| < C.$$

**Corollary 3.3.** *Assume that the sequence  $f_n$  of entire functions defined by (7) satisfies conditions 2) and 3) of Lemma 3.2. Let  $\tilde{C}_1 > 0$ , and let  $a_n$  be a sequence of positive numbers satisfying  $|a_n| \leq \tilde{C}_1$ . If*

$$\max_{|z-n| < n^\delta} |f_n(z) - f_\infty| e^{-\frac{\gamma_1|z-n|}{\sqrt{n}}} \leq C_1 a_n,$$

then for all  $n > 0$  we have:

$$|f_{nn} - f_\infty| \leq C a_n.$$

Here  $C$ , again, is a constant depending only on  $\gamma_1, \gamma_2, \gamma_3, C_1, \tilde{C}_1, C_2, C_3$ .

*Proof.* Follows by applying Lemma 3.2 to the sequence

$$\frac{f_n(z) - f_\infty}{a_n}.$$

□

We shall be mainly concerned with depoissonization of various polynomials of Bessel functions, and it is useful to note that in this case Conditions 2 and 3 in Lemma 3.2 and Corollary 3.3 hold automatically. More precisely, we have the following proposition.

**Proposition 3.4.** *Let  $K > 0$ ,  $k \in \mathbb{N}$ . Let  $P$  be a polynomial in  $k$  variables. Let*

$$x_n^{(1)}, \dots, x_n^{(k)}, n \in \mathbb{N},$$

*be  $k$  sequences of integers satisfying*

$$|x_n^{(1)}|, \dots, |x_n^{(k)}| \leq K\sqrt{n}.$$

*Then there exist constants*

$$C_1, C_2, \gamma_1, \gamma_2 > 0$$

*depending only on  $k$ ,  $K$  and  $P$  such that*

- 1)  $\max_{|z|=n} \left| P \left( J_{x_n^{(1)}}(2\sqrt{z}), \dots, J_{x_n^{(k)}}(2\sqrt{z}) \right) \right| \leq C_1 e^{\gamma_1 \sqrt{n}}.$
- 2)  $\max_{|z-n| < n^{1-\alpha}} \left| P \left( J_{x_n^{(1)}}(2\sqrt{z}), \dots, J_{x_n^{(k)}}(2\sqrt{z}) \right) \right| e^{-\frac{\gamma_2 |z-n|}{\sqrt{n}}} \leq C_2.$

The Proposition is immediate from the contour integral representation of the Bessel functions.

We also note that in the depoisonization arguments that follow, weaker assumptions on  $\theta$  than those of [3] are sufficient: namely, we shall always assume that  $\theta = \sqrt{z}$  satisfies

$$\left| \frac{\theta}{\sqrt{n}} - 1 \right| \leq \varepsilon_0,$$

where  $\varepsilon_0$  is sufficiently small.

**3.4. The Debye Asymptotics.** For depoisonization we need the asymptotics of Bessel functions when both order and argument are large. First results of this type are due to Carlini; we shall use the asymptotics due to Debye, following the exposition by Watson [21].

Take  $\varepsilon > 0$ . Set

$$(8) \quad N_n^{(\varepsilon)} = \{k \in \mathbb{Z} : \frac{|k|}{\sqrt{n}} < 2 - \varepsilon\}.$$

Take  $x \in N_n^{(\varepsilon)}$ . Then there exists  $\varepsilon_0 > 0$  depending only on  $\varepsilon$  such that for any  $\theta \in \mathbb{C}$  satisfying  $|\theta/\sqrt{n} - 1| < \varepsilon_0$  the following is true. Introduce  $u$  by the formula  $\cos u = x/2\theta$ ,  $0 < \Re(u) < \pi$ . Then we have the following representation for Bessel functions, asymptotic in the sense of Poincaré:

$$(9) \quad J_x(2\theta) = \frac{\cos \left( 2\theta(\tan u - u) - \frac{\pi}{4} \right)}{\sqrt{\pi\theta \tan u}} \left( 1 + \sum_{m=1}^{\infty} \frac{\alpha_m(u)}{\theta^m} \right),$$

where for any  $\varepsilon' > 0$  there exists  $\delta' > 0$  such that all  $\alpha_m(u)$  are holomorphic in  $u$  in the strip  $[\varepsilon', \pi - \varepsilon'] \times [-\delta', \delta']$ .

## 4. LOCAL PATTERNS IN PLANCHEREL YOUNG DIAGRAMS.

**4.1. The Discrete Sine-Process.** Take  $a \in (-2, 2)$  and introduce *the discrete sine-kernel* by the formula

$$\mathcal{S}(k, a) = \begin{cases} \frac{\sin(\arccos(a/2)k)}{\arccos(a/2)}, & \text{if } k \neq 0; \\ \frac{\pi k}{\pi}, & \text{if } k = 0. \end{cases}$$

Introduce a measure  $\mathcal{S}(a)$  on  $\Omega_2$  by setting

$$(10) \quad \mathbb{E}_{\mathcal{S}(a)}(c_{\vec{m}}) = \det(\mathcal{S}(m_i - m_j, a)) \Big|_{i,j=1,\dots,r}.$$

The measure  $\mathcal{S}(a)$  is called *the discrete sine-process*.

For  $x \in \mathbb{Z}$ ,  $\vec{m} \in \mathbb{Z}^r$ , denote  $x + \vec{m} = (x + m_1, \dots, x + m_r)$ .

The theorem of Borodin, Okounkov and Olshanski [3] says that for any  $a \in (-2, 2)$ , any integer vector  $\vec{m}$  and any sequence  $x_n \in \mathbb{Z}$  satisfying

$$\lim_{n \rightarrow \infty} \frac{x_n}{\sqrt{n}} = a,$$

we have

$$(11) \quad \lim_{n \rightarrow \infty} \mathbb{E}_{\mathbb{P}^{(n)}}(c_{x_n + \vec{m}}(\lambda)) = \mathbb{E}_{\mathcal{S}(a)}(c_{\vec{m}}).$$

**4.2. The Variance of the Discrete Sine-Process.** We shall need the following simple estimate.

**Proposition 4.1.** *There exists a positive constant  $C$  such that for any  $s \in [0, 1]$ , any  $h > 1$  and any  $a \in (-2, 2)$  we have*

$$\mathbb{E}_{\mathcal{S}(a)} \left( \Phi_\omega(s+h) - \Phi_\omega(s) - \frac{2}{\pi} \arcsin\left(\frac{a}{2}\right) h \right)^2 \leq C(1 + \log h).$$

*Proof.* First, recall that for any  $a \in (-2, 2)$  the operator  $\mathcal{S}_a : l_2(\mathbb{Z}) \rightarrow l_2(\mathbb{Z})$  given by the formula

$$\mathcal{S}_a f(x) = \sum_{k \in \mathbb{Z}} \mathcal{S}(k, a) f(x+k),$$

is an orthogonal projection (this is easy to check by taking the Fourier transform).

Second, recall the well-known

**Proposition 4.2.** *Let  $\mathcal{K} : l_2(\mathbb{Z}) \rightarrow l_2(\mathbb{Z})$*

$$\mathcal{K}f(x) = \sum_{y \in \mathbb{Z}} \mathcal{K}(x, y) f(y),$$

be an orthogonal projection, and let  $\mathbb{P}_{\mathcal{K}}$  be the corresponding determinantal measure on  $\Omega_2$ . Then for any  $k_1, k_2 \in \mathbb{Z}$ ,  $k_1 < k_2$ , we have

$$\mathrm{Var}_{\mathbb{P}_{\mathcal{K}}} \left( \sum_{n=k_1}^{k_2} c_n \right) = \sum_{x \in [k_1, k_2]} \sum_{y \notin [k_1, k_2]} |\mathcal{K}(x, y)|^2$$

The proof is a straightforward computation using the formula

$$\mathcal{K}(x, x) = \sum_{y \in \mathbb{Z}} |\mathcal{K}(x, y)|^2,$$

which holds for any  $x \in \mathbb{Z}$ .

In the remainder of the proof,  $C$  stands for a positive constant that does not depend on  $a \in (-2, 2)$ .

By definition, for any  $k \neq 0$  and all  $a \in (-2, 2)$  we have  $|\mathcal{S}(k, a)| \leq \frac{1}{|k|}$ . Therefore, for any  $N > 0$ , we have

$$\mathrm{Var}_{\mathcal{S}(a)} \left( \sum_{n=0}^N c_n \right) \leq C(1 + \log N).$$

Since  $\Phi_\omega(0) = 0$  and

$$\Phi_\omega(n+1) - \Phi_\omega(n) = 1 - 2c_n(\omega),$$

for any  $N > 0$  we have

$$\Phi_\omega(N) = N - 2 \sum_{n=0}^{N-1} c_n(\omega).$$

It follows that

$$\mathbb{E}_{\mathcal{S}(a)} \Phi_\omega(N) = \frac{2}{\pi} N \arcsin \left( \frac{a}{2} \right),$$

$$\mathrm{Var}_{\mathcal{S}(a)} \Phi_\omega(N) \leq C(1 + \log N).$$

In other words,

$$\mathbb{E}_{\mathcal{S}(a)} \left( \Phi_\omega(N) - \frac{2}{\pi} N \arcsin \left( \frac{a}{2} \right) \right)^2 \leq C(1 + \log N).$$

Recall that the function  $\Phi_\omega$  is Lipschitz with constant 1. Using the inequality  $(c+d)^2 \leq 2(c^2+d^2)$ , for any  $s \in (0, 1)$  and any  $h > 1$  we finally obtain

$$\mathbb{E}_{\mathcal{S}(a)} \left( \Phi_\omega(s+h) - \Phi_\omega(s) - \frac{2}{\pi} \arcsin \left( \frac{a}{2} \right) h \right)^2 \leq C(1 + \log h),$$

which is what we had to prove.  $\square$

**4.3. Decay of Correlations for the Plancherel Measure.** Borodin, Okounkov and Olshanski [3] have also shown that if  $a \neq b$  and

$$\lim_{n \rightarrow \infty} \frac{x_n}{\sqrt{n}} = a, \quad \lim_{n \rightarrow \infty} \frac{y_n}{\sqrt{n}} = b,$$

then for any integer vectors  $\vec{l}, \vec{m}$  we have

$$(12) \quad \lim_{n \rightarrow \infty} \mathbb{E}_{\mathbb{P}_L(n)} \left( c_{x_n + \vec{m}}(\lambda) \cdot c_{y_n + \vec{l}}(\lambda) \right) = \mathbb{E}_{S(a)C_{\vec{m}}} \cdot \mathbb{E}_{S(b)C_{\vec{l}}}.$$

Distant local patterns in a Young diagram are thus asymptotically independent. We shall need an estimate for the decay of correlations of the Plancherel measure.

For an integer vector  $\vec{m}$ , let  $|\vec{m}|$  stand for the maximum of absolute values of its coordinates.

**Lemma 4.3.** *For any  $\varepsilon > 0$ ,  $L > 0$ , there exists a constant  $C = C(\varepsilon, L)$  such that for any  $n > 0$ , any  $x, y \in \mathbb{Z}$  such that*

$$\frac{|x|}{\sqrt{n}}, \frac{|y|}{\sqrt{n}} < 2 - \varepsilon$$

and any integer vectors  $\vec{l}, \vec{m}$  satisfying  $|\vec{l}| \leq L$ ,  $|\vec{m}| \leq L$ , we have

$$(13) \quad \left| \mathbb{E}_{\mathbb{P}_L(n)}(c_{x+\vec{l}} \cdot c_{y+\vec{m}}) - \mathbb{E}_{\mathbb{P}_L(n)}(c_{x+\vec{l}}) \cdot \mathbb{E}_{\mathbb{P}_L(n)}(c_{y+\vec{m}}) \right| \leq C(\varepsilon, L) \left( \left( \frac{1}{|x-y|+1} \right)^2 + \frac{1}{\sqrt{n}} \right).$$

**4.4. Frequency of Local Patterns.** Lemma 4.3 will be used to in the next section to prove the following

**Lemma 4.4.** *For any continuous bounded function  $f : \mathbb{R} \rightarrow \mathbb{C}$  and any integer vector  $\vec{m}$ , the sequence of random variables*

$$(14) \quad \frac{1}{\sqrt{n}} \sum_{k=-\infty}^{\infty} f\left(\frac{k}{\sqrt{n}}\right) c_{k+\vec{m}}(\lambda)$$

converges, as  $n \rightarrow \infty$ , to the constant

$$\int_{-2}^2 f(a) \mathbb{E}_{S(a)C_{\vec{m}}} da$$

according to the Plancherel measure.

We shall see in the next section that Lemma 2.1 is a simple corollary of Lemma 4.4. Lemma 2.2 admits the following more precise version, which will also be derived from Lemma 4.4.

**Lemma 4.5.** *For any  $h_0 > 0$  the random variables*

$$\frac{1}{\sqrt{n}} \int_0^{h_0} \int_{-\infty}^{\infty} \left( \frac{F_\lambda(t+h) - F_\lambda(t)}{h} \right)^2 dt dh$$

*converge to the constant*

$$\int_{-2}^2 \int_0^1 \int_0^{h_0} \mathbb{E}_{S(a)} \left( \frac{\Phi_\omega(s+h) - \Phi_\omega(s)}{h} - \frac{2}{\pi} \arcsin(a/2) \right)^2 dh ds da$$

*according to the Plancherel measure.*

For the constant  $H$ , the entropy of the Plancherel measure, we now obtain

(15)

$$\begin{aligned} H = & \frac{1}{4} \int_{-2}^2 \int_0^1 \int_0^{\infty} \mathbb{E}_{S(a)} \left( \frac{\Phi_\omega(s+h) - \Phi_\omega(s)}{h} - \frac{2}{\pi} \arcsin(a/2) \right)^2 dh ds da + \\ & + \frac{32}{\pi^2} \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \frac{1}{l(l+1)(2l+1)k^{2l-2}(4k^2-1)}. \end{aligned}$$

Convergence of the integral in  $h$  is clear from Proposition 4.1.

## 5. PROOF OF LEMMAS 4.3, 4.4, 4.5, 2.1, 2.2.

**5.1. Decay of Correlations for the Bessel Point Process.** Given a measure  $\mathbb{P}$  on  $\Omega_2$ , any natural  $x, y$ , and any integer vectors  $\vec{l} = (l_1, \dots, l_r)$ ,  $\vec{m} = (m_1, \dots, m_s)$ , denote

$$\text{Cov}_{\mathbb{P}}(x, \vec{l}; y, \vec{m}) = \mathbb{E}_{\mathbb{P}}(c_{x+\vec{l}} \cdot c_{y+\vec{m}}) - \mathbb{E}_{\mathbb{P}}(c_{x+\vec{l}}) \cdot \mathbb{E}_{\mathbb{P}}(c_{y+\vec{m}}).$$

For complex  $\theta$  the expression  $\text{Cov}_{\mathcal{J}(\theta^2)}(x, \vec{l}; y, \vec{m})$  is defined formally, by analytic continuation. Our next aim is to estimate this quantity from above.

The representation (9) implies, in particular, the existence of constants  $C, \gamma, \varepsilon_0$  depending only on  $\varepsilon$  such that

$$(16) \quad |\theta(J_x(2\theta)J_{y+1}(2\theta) - J_y(2\theta)J_{x+1}(2\theta))| \leq C \exp(\gamma|\theta - \sqrt{n}|),$$

provided  $x, y \in N_n^{(\varepsilon)}$ ,  $|\theta/\sqrt{n} - 1| < \varepsilon_0$ , whence, if  $x \neq y$ , we have

$$|\mathcal{J}(x, y; \theta^2)| \leq \frac{C \exp(\gamma|\theta - \sqrt{n}|)}{|x - y|}.$$



The function  $\mathcal{J}(x, y, \theta^2)$  is entire in  $x, y$  and, in the same way as in (3.7) on p.498 in [3], write

$$\mathcal{J}(x, x, \theta^2) = \frac{1}{2\pi} \int_0^{2\pi} \mathcal{J}(x, x + r \exp(it), \theta^2) dt,$$

where  $r$  is arbitrary. This representation shows that in the case  $x = y$  we also have

$$(17) \quad |\mathcal{J}(x, x; \theta^2)| \leq C \exp(\gamma|\theta - \sqrt{n}|)$$

We have established the following

**Lemma 5.1.** *For any  $\varepsilon > 0$ ,  $L > 0$ , there exist positive constants  $C = C(\varepsilon, L)$ ,  $\gamma = \gamma(\varepsilon, L)$  and  $\varepsilon_0$  depending only on  $\varepsilon$ , such that for any  $n > 0$ , any  $\theta$  satisfying  $|\theta| = \sqrt{n}$ ,  $|\theta/\sqrt{n} - 1| < \varepsilon_0$ , any  $x, y \in N_n^{(\varepsilon)}$ , and any integer vectors  $\vec{l}, \vec{m}$  with absolute values not exceeding  $L$ , we have*

$$(18) \quad \text{Cov}_{\mathcal{J}(\theta^2)}(x, \vec{l}; y, \vec{m}) \leq \frac{C(\varepsilon, L) \exp(\gamma|\theta - \sqrt{n}|)}{(|x - y| + 1)^2}.$$

**5.2. Proof of Lemma 4.3.** The Debye asymptotics (9) immediately yields (see [3] for details) that for any  $\varepsilon > 0$  there exists  $\varepsilon_0 > 0$  and for any  $l \in \mathbb{N}$  constants  $C = C(l, \varepsilon)$ ,  $\gamma = \gamma(l, \varepsilon)$  such that if  $\theta$  satisfies  $|\theta| = \sqrt{n}$ ,  $|\theta/\sqrt{n} - 1| < \varepsilon_0$ , then for any  $x \in N_n^{(\varepsilon)}$  we have

$$(19) \quad |\mathcal{J}(x, x + l, \theta^2) - \mathfrak{S}\left(\frac{x}{2\theta}, l\right)| \leq \frac{C \exp(\gamma|\theta - \sqrt{n}|)}{\sqrt{n}}.$$

By definition of a determinantal process, the estimate (19) implies the following

**Proposition 5.2.** *For any  $\varepsilon > 0$ ,  $L > 0$ , there exist positive constants  $C = C(\varepsilon, L)$ ,  $\gamma = \gamma(\varepsilon, L)$  and  $\varepsilon_0$  depending only on  $\varepsilon$ , such that for any  $n > 0$ , any  $\theta$  satisfying  $|\theta| = \sqrt{n}$ ,  $|\theta/\sqrt{n} - 1| < \varepsilon_0$ , any  $x \in N_n^{(\varepsilon)}$ , and any integer vector  $\vec{l}$  satisfying  $|\vec{l}| \leq L$ , we have*

$$(20) \quad |\mathbb{E}_{\mathcal{J}(\theta^2)}(c_{x+\vec{l}}) - \mathbb{E}_{\mathfrak{S}(x/2\sqrt{n})}(c_{\vec{l}})| \leq \frac{C \exp(\gamma|\theta - \sqrt{n}|)}{\sqrt{n}}.$$

Depoissonizing by Lemma 3.1, we obtain

$$(21) \quad |\mathbb{E}_{\mathbb{P}_l^{(n)}}(c_{x+\vec{l}}) - \mathbb{E}_{\mathfrak{S}(x/\sqrt{n})}(c_{\vec{l}})| \leq \frac{C(\varepsilon, L)}{\sqrt{n}}.$$

As a simple example, taking  $l = 0$  in (19) and depoissonizing by Lemma 3.1 yields

$$(22) \quad |\mathbb{E}_{\mathbb{P}_l^{(n)}}(c_x) - \frac{1}{\pi} \arccos(x/2\sqrt{n})| \leq \frac{C(\varepsilon)}{\sqrt{n}}.$$

Substituting (20) into (18), we obtain

$$\begin{aligned} & \left| \mathbb{E}_{\mathcal{J}(\theta^2)}(c_{x+l} \cdot c_{y+m}) - \mathbb{E}_{\mathcal{S}(\frac{x}{\sqrt{n}})}(c_{x+l}) \cdot \mathbb{E}_{\mathcal{S}(\frac{y}{\sqrt{n}})}(c_{y+m}) \right| \leq \\ & \leq C(\varepsilon, L) \exp(\gamma|\theta - \sqrt{n}|) \left( \left( \frac{1}{|x-y|+1} \right)^2 + \frac{1}{\sqrt{n}} \right), \end{aligned}$$

whence, by the dePoissonization Lemma 3.1, we have

$$(23) \quad \left| \mathbb{E}_{\mathbb{P}l(n)}(c_{x+l} \cdot c_{y+m}) - \mathbb{E}_{\mathcal{S}(\frac{x}{\sqrt{n}})}(c_l) \cdot \mathbb{E}_{\mathcal{S}(\frac{y}{\sqrt{n}})}(c_m) \right| \leq \\ \leq C(\varepsilon, L) \left( \left( \frac{1}{|x-y|+1} \right)^2 + \frac{1}{\sqrt{n}} \right).$$

Finally, using (21), we write

$$(24) \quad \left| \mathbb{E}_{\mathbb{P}l(n)}(c_{x+l} \cdot c_{y+m}) - \mathbb{E}_{\mathbb{P}l(n)}(c_{x+l}) \cdot \mathbb{E}_{\mathbb{P}l(n)}(c_{y+m}) \right| \leq \\ \leq C(\varepsilon, L) \left( \left( \frac{1}{|x-y|+1} \right)^2 + \frac{1}{\sqrt{n}} \right),$$

and Lemma 4.3 is proved.

**5.3. Proof of Lemma 4.4.** As before, let  $f : \mathbb{R} \rightarrow \mathbb{C}$  be continuous and bounded, let  $\vec{m} = (m_1, \dots, m_r)$  be an integer vector and denote

$$(25) \quad S_n(\vec{m}, f, \lambda) = \frac{1}{\sqrt{n}} \sum_{k=-\infty}^{\infty} f\left(\frac{k}{\sqrt{n}}\right) c_{k+\vec{m}}(\lambda).$$

Take  $\varepsilon > 0$  and consider a modification of the function  $S_n(\vec{m}, f, \lambda)$  defined by the formula

$$(26) \quad S_n(\vec{m}, f, \lambda, \varepsilon) = \frac{1}{\sqrt{n}} \sum_{k:|k| \leq (2-\varepsilon)\sqrt{n}} f\left(\frac{k}{\sqrt{n}}\right) c_{k+\vec{m}}(\lambda).$$

**Proposition 5.3.** *For any  $\varepsilon > 0$ , any continuous  $f$  on  $[-2, 2]$  and any integer vector  $\vec{m}$  we have*

$$(27) \quad \lim_{n \rightarrow \infty} \mathbb{E}_{\mathbb{P}l(n)} \left| S_n(\vec{m}, f, \lambda, \varepsilon) - \int_{\varepsilon-2}^{2-\varepsilon} f(a) (\mathbb{E}_{\mathcal{S}(a)} c_{\vec{m}}) da \right|^2 = 0$$

*Proof.* We begin by estimating

$$(28) \quad \mathbb{E}_{\mathbb{P}l(n)} \left| S_n(\vec{m}, f, \lambda, \varepsilon) - \frac{1}{\sqrt{n}} \sum_{k:|k| \leq (2-\varepsilon)\sqrt{n}} f\left(\frac{k}{\sqrt{n}}\right) (\mathbb{E}_{\mathcal{S}(\frac{k}{\sqrt{n}})} c_{\vec{m}}) \right|^2.$$

By Lemma 4.3 and Proposition 5.2, boundedness of  $f$  implies the estimate

$$\left| f\left(\frac{k}{\sqrt{n}}\right) f\left(\frac{l}{\sqrt{n}}\right) \mathbb{E}_{\mathbb{P}^{l(n)}} \left( (c_{k+\vec{m}} - \mathbb{E}_{\mathcal{S}(\frac{k}{\sqrt{n}})c_{\vec{m}}}) \cdot (c_{l+\vec{m}} - \mathbb{E}_{\mathcal{S}(\frac{l}{\sqrt{n}})c_{\vec{m}}}) \right) \right| \leq \frac{C(f, \varepsilon)}{|k-l|+1}.$$

Summing in  $k$  and  $l$ , we obtain that the integral (28) is bounded above by  $C(f, \varepsilon) \log^2 n / \sqrt{n}$ . Now observe that the quantity

$$\frac{1}{\sqrt{n}} \sum_{k: |k| \leq (2-\varepsilon)\sqrt{n}} f\left(\frac{k}{\sqrt{n}}\right) \left( \mathbb{E}_{\mathcal{S}(\frac{k}{\sqrt{n}})c_{\vec{m}}} \right)$$

is a Riemann sum for the integral

$$\int_{\varepsilon-2}^{2-\varepsilon} f(a) \left( \mathbb{E}_{\mathcal{S}(a)c_{\vec{m}}} \right) da.$$

Since the function  $f(a) \left( \mathbb{E}_{\mathcal{S}(a)c_{\vec{m}}} \right)$  is continuous on  $[\varepsilon-2, 2-\varepsilon]$ , the Riemann sums converge to the integral, and Proposition 5.3 is proved.

To derive Lemma 4.4 from Proposition 5.3, note that if  $\lambda \in \mathbb{Y}_n$  satisfies  $\lambda_1 < (2 + \varepsilon_1)\sqrt{n}$ ,  $\lambda'_1 < (2 + \varepsilon_1)\sqrt{n}$ , then we have

$$(29) \quad |S_n(\vec{m}, f, \lambda) - S_n(\vec{m}, f, \lambda, \varepsilon)| \leq C(\vec{m}, f)(\varepsilon + \varepsilon_1),$$

where  $C(\vec{m}, f)$  is a constant depending only on  $\vec{m}$  and  $f$ . Since  $\varepsilon$  and  $\varepsilon_1$  can be chosen arbitrarily small, Lemma 4.4 follows now from Proposition 2.5.

**5.4. Proof of Lemma 2.1.** Observe the clear identity

$$(30) \quad h_k(\lambda) = \sum_{i=-\infty}^{\infty} (c_i(\lambda) - c_i(\lambda)c_{i-k}(\lambda))$$

(note that only finitely many terms in the right-hand side are nonzero). Lemma 2.1 is now immediate from Lemma 4.4. For the constant, compute

$$(31) \quad \int_{-2}^2 \left( \frac{1}{\pi} \arccos\left(\frac{a}{2}\right) - \frac{1}{\pi^2} \arccos^2\left(\frac{a}{2}\right) + \frac{1}{k^2\pi^2} \sin^2\left(k \arccos\left(\frac{a}{2}\right)\right) \right) da = \\ = \frac{32k^2}{(4k^2 - 1)\pi^2}.$$

**5.5. Proof of Lemmas 2.2, 4.5.** For Lemma 4.5, take  $\varepsilon > 0$  and observe that for  $t$  and  $h$  satisfying

$$\frac{|t|}{\sqrt{n}} < 2 - \varepsilon, \quad 0 < h \leq h_0$$

we have

$$(32) \quad \left| \frac{\sqrt{n}}{h} \left( \Omega \left( \frac{t+h}{\sqrt{n}} \right) - \Omega \left( \frac{t}{\sqrt{n}} \right) \right) - \frac{2}{\pi} \arcsin \left( \frac{t}{2\sqrt{n}} \right) \right| \leq \frac{C(\varepsilon, h_0)}{\sqrt{n}},$$

where the constant  $C(\varepsilon, h_0)$  only depends on  $\varepsilon$  and  $h_0$ .

Now take  $s \in (0, 1)$  and consider the expression

$$(33) \quad \frac{1}{\sqrt{n}} \int_0^{h_0} \left( \sum_{k: |k| \leq (2-\varepsilon)\sqrt{n}} \left( \frac{\Phi_\lambda(s+k+h) - \Phi_\lambda(s+k)}{h} - \frac{2}{\pi} \arcsin \left( \frac{s+k}{2\sqrt{n}} \right) \right)^2 \right) dh$$

For any  $h_0 \geq 0$  there exists  $N_0 = N_0(h_0)$ , and for all  $h : 0 \leq h \leq h_0$ , numbers  $\alpha_i^{(h)}$ ,  $i \in \mathbb{Z}$ , satisfying  $|\alpha_i^{(h)}| \leq 1$  and  $\alpha_i^{(h)} = 0$  for  $|i| > N_0(h_0)$  such that for any  $\omega \in \Omega_2$  we have

$$\frac{\Phi_\omega(s+k+h) - \Phi_\omega(s+k)}{h} = \sum_{i=-\infty}^{\infty} \alpha_i^{(h)} \omega_{i+k}.$$

By Proposition 5.3, the sum (33) converges, with respect to the Plancherel measure, to the constant

$$\int_0^{h_0} \int_{\varepsilon-2}^{2-\varepsilon} \mathbb{E}_{S(a)} \left( \frac{\Phi_\omega(s+h) - \Phi_\omega(s)}{h} - \frac{2}{\pi} \arcsin(a/2) \right)^2 da dh.$$

Taking  $\varepsilon$  to 0 (the transition to the limit is justified in the same way as in (29)) and integrating in  $s$  from 0 to 1, we obtain Lemma 4.5.

## 6. PROOF OF LEMMA 2.3.

**6.1. Outline of the Proof.** The first step in proving Lemma 2.3 is to reduce integrals to sums and to observe that summation need only take place “away from the edge”.

More precisely, let  $\delta \in \mathbb{R}$ ,  $0 < \delta < \frac{1}{4}$  and let  $K > 0$ . Denote

$$\mathbb{Y}_n(K, \delta) = \{\lambda \in \mathbb{Y}_n : \lambda_1 \leq 2\sqrt{n} + Kn^\delta, \lambda'_1 \leq 2\sqrt{n} + Kn^\delta\}.$$

Denote

$$F_\lambda^{(L, \delta)}(t) = \begin{cases} F_\lambda(t), & \text{if } |t| \leq 2\sqrt{n} - Ln^\delta \\ 0, & \text{otherwise.} \end{cases}$$

**Lemma 6.1.** *For any  $\delta$  satisfying  $0 < \delta < \frac{1}{4}$ , any  $K > 0$ ,  $L > 0$  and any  $\varepsilon > 0$ , there exists a number  $h_0 > 1$  depending only on  $\delta, K, L, \varepsilon$  and such that for any  $n \in \mathbb{N}$  and any  $\lambda \in \mathbb{Y}_n(K, \delta)$  we have the inequality*

$$(34) \quad \frac{1}{\sqrt{n}} \int_{h_0}^{+\infty} \int_{-\infty}^{+\infty} \left( \frac{F_\lambda(t+h) - F_\lambda(t)}{h} \right)^2 dt dh \leq \\ \leq \frac{4}{\sqrt{n}} \sum_{l > h_0 - 1} \sum_{k = -\infty}^{+\infty} \left( \frac{F_\lambda^{(L, \delta)}(k+l) - F_\lambda^{(L, \delta)}(k)}{l} \right)^2 + \varepsilon.$$

We postpone its proof to the following subsection.

The second step is to estimate the expectation of the quantity

$$\sum_{l > h_0 - 1} \sum_{k = -\infty}^{+\infty} \left( \frac{F_\lambda^{(L, \delta)}(k+l) - F_\lambda^{(L, \delta)}(k)}{l} \right)^2$$

with respect to the Plancherel measure.

We start with estimates for the poissonized Plancherel measure, the Bessel point process. Take  $\delta > 1/6$  and let

$$\mathcal{N}(n, \delta) = \{x \in \mathbb{Z} : |x| \leq 2\sqrt{n} - n^\delta\}.$$

**Lemma 6.2.** *For any  $\delta > \frac{1}{6}$  there exist constants  $C > 0$ ,  $\gamma > 0$ ,  $\varepsilon > 0$  such that the following holds.*

*For any  $l_0 > 1$  there exists  $n_0 > 0$  such that for all  $n > n_0$ , and all  $\theta$  satisfying*

$$\left| \frac{\theta}{\sqrt{n}} - 1 \right| < \varepsilon$$

*we have*

$$(35) \quad \frac{1}{\sqrt{n}} \sum_{l > l_0} \sum_{\substack{k \in \mathcal{N}(n, \delta) \\ k+l \in \mathcal{N}(n, \delta)}} \left| \frac{\text{Var}_{J(\theta^2)}(c_k + \dots + c_{k+l-1})}{l^2} \right| \leq \frac{C \log l_0}{l_0} e^{\gamma|\theta - \sqrt{n}|}.$$

Lemma 6.2 is again essentially a straightforward computation using simple estimates on the discrete Bessel kernel. We prove Lemma 6.2 in the following subsection. Now we conclude the proof of Lemma 2.3.

Rewrite formula (35) as follows

$$(36) \quad \frac{1}{\sqrt{n}} \sum_{l>l_0} \sum_{\substack{k \in \mathcal{N}(n,\delta) \\ k+l \in \mathcal{N}(n,\delta)}} \left| \frac{1}{l} (\mathbb{E}_{J(\theta^2)} \left( \sum_{r=k}^{k+l-1} (c_r - J(r, r; \theta^2)) \right)) \right|^2 \leq \\ \leq \frac{C \log l_0}{l_0} e^{\gamma|\theta - \sqrt{n}|}.$$

Now write

$$(37) \quad \begin{aligned} F_\lambda(k+1) - F_\lambda(k) &= 1 - 2c_k(\lambda) - \sqrt{n} \left( \Omega \left( \frac{k+1}{\sqrt{n}} \right) - \Omega \left( \frac{k}{\sqrt{n}} \right) \right) = \\ &= 2 \left( \frac{\arccos \frac{k}{2\sqrt{n}}}{\pi} - c_k(\lambda) \right) + \frac{2}{\pi} \arcsin \left( \frac{k}{2\sqrt{n}} \right) - \sqrt{n} \left( \Omega \left( \frac{k+1}{\sqrt{n}} \right) - \Omega \left( \frac{k}{\sqrt{n}} \right) \right) = \\ &= 2 \left( J(k, k; \theta^2) - c_k(\lambda) \right) + 2 \left( \frac{\arccos \frac{k}{2\sqrt{n}}}{\pi} - J(k, k; \theta^2) \right) + \\ &\quad + \left( \frac{2}{\pi} \arcsin \left( \frac{k}{2\sqrt{n}} \right) - \sqrt{n} \left( \Omega \left( \frac{k+1}{\sqrt{n}} \right) - \Omega \left( \frac{k}{\sqrt{n}} \right) \right) \right). \end{aligned}$$

From the Taylor formula applied to the function  $\Omega$  we have, for  $|k| < 2\sqrt{n}$ , the estimate

$$(38) \quad \left| \sqrt{n} \left( \Omega \left( \frac{k+1}{\sqrt{n}} \right) - \Omega \left( \frac{k}{\sqrt{n}} \right) \right) - \frac{2}{\pi} \arcsin \left( \frac{k}{2\sqrt{n}} \right) \right| \leq \frac{10}{\sqrt{4n - k^2}}.$$

To estimate the quantity

$$\left| J(k, k; \theta^2) - \frac{\arccos \frac{k}{2\sqrt{n}}}{\pi} \right|$$

we use the following Lemma.

**Lemma 6.3.** *There exists  $\varepsilon_0 > 0$  such that the following holds. For any  $\delta_0 > \frac{1}{6}$  there exist constants  $C > 0$ ,  $\gamma > 0$  such that for all  $n \in \mathbb{N}$ , all  $x \in \mathbb{Z}$  satisfying  $|x| \leq 2\sqrt{n} - n^{\delta_0}$  and all  $\theta \in \mathbb{C}$  satisfying*

$$\left| \frac{\theta}{\sqrt{n}} - 1 \right| < \varepsilon_0$$

we have

$$\left| J(x, x; \theta^2) - \frac{1}{\pi} \arccos \frac{x}{2\sqrt{n}} \right| \leq \frac{C}{2\sqrt{n} - |x|} e^{\gamma|\theta - \sqrt{n}|}.$$

Observe that it suffices to prove Lemma 6.3 for  $x > 0$ , as the other case follows by symmetry. Lemma 6.3 is again a relatively straightforward estimate using Okounkov's contour integral representation for the discrete Bessel kernel. For the reader's convenience, we give the proof in the last Section.

Using Lemma 6.3, we obtain from (37), (38) the estimate

$$|(F_\lambda(k+1) - F_\lambda(k)) - 2(J(k, k; \theta^2) - c_k(\lambda))| \leq \frac{C}{2\sqrt{n} - |k|} e^{\gamma|\theta - \sqrt{n}|}.$$

Observe now the following simple inequality

$$(39) \quad \frac{1}{\sqrt{n}} \sum_{l > l_0} \sum_{\substack{k \in \mathcal{N}(n, \delta) \\ k+l \in \mathcal{N}(n, \delta)}} \left( \frac{\sum_{x=k}^{k+l} \frac{1}{2\sqrt{n}-|x|}}{l} \right)^2 \leq \frac{C \log^2 l_0}{l_0}.$$

From (36) and (39) we now obtain

$$\frac{1}{\sqrt{n}} \sum_{l > l_0} \sum_{\substack{k \in \mathcal{N}(n, \delta) \\ k+l \in \mathcal{N}(n, \delta)}} \left| \mathbb{E}_{J(\theta^2)} \left( \frac{F_\lambda(k+l) - F_\lambda(k)}{l} \right)^2 \right| \leq \frac{C \log^2 l_0}{l_0} e^{\gamma|\theta - \sqrt{n}|}.$$

Depoissonizing, we have

$$(40) \quad \frac{1}{\sqrt{n}} E_{\mathbb{P}l(n)} \left( \sum_{l > l_0} \sum_{\substack{k \in \mathcal{N}(n, \delta) \\ k+l \in \mathcal{N}(n, \delta)}} \left( \frac{F_\lambda(k+l) - F_\lambda(k)}{l} \right)^2 \right) \leq \frac{C \log^2 l_0}{l_0}.$$

The estimate (40), together with Lemma 6.1, completes the proof of Lemma 2.3.

**6.2. Proof of Lemma 6.1.** The first step is to pass from integrals in  $t$  and in  $h$  to sums in  $k$  and  $l$ . Since the function  $F_\lambda$  is Lipschitz with the Lipschitz constant 2, for any  $t \in \mathbb{R}$ ,  $h \in \mathbb{R}_+$  we have

$$(F_\lambda(t+h) - F_\lambda(t))^2 \leq 2(F_\lambda(k+l) - F_\lambda(k))^2 + 16,$$

where  $k = [t]$ ,  $l = [h]$ . Integrating in  $t$ , we have

$$\int_{-\infty}^{+\infty} (F_\lambda(t+h) - F_\lambda(t))^2 dt \leq 2 \sum_{k=-\infty}^{+\infty} (F_\lambda(k+l) - F_\lambda(k))^2 + 40\sqrt{n},$$

where, as before  $l = [h]$ .

Now, integrating in  $h$  from  $l$  to  $l + 1$ , we arrive at the inequality

$$\int_l^{l+1} \int_{-\infty}^{+\infty} (F_\lambda(t+h) - F_\lambda(t))^2 dt dh \leq 2 \sum_{k=-\infty}^{+\infty} (F_\lambda(k+l) - F_\lambda(k))^2 + 40\sqrt{n},$$

whence, for any  $h_0 > 1$  we have

$$(41) \quad \frac{1}{\sqrt{n}} \int_{h_0}^{+\infty} \int_{-\infty}^{+\infty} \left( \frac{F_\lambda(t+h) - F_\lambda(t)}{h} \right)^2 dt dh \leq \\ \leq \frac{2}{\sqrt{n}} \sum_{l>h_0-1} \sum_{k=-\infty}^{+\infty} \left( \frac{F_\lambda(k+l) - F_\lambda(k)}{l} \right)^2 + 40 \cdot \sum_{l>h_0-1} \frac{1}{l^2}.$$

For  $h_0 > 1$ , we thus arrive at the inequality

$$(42) \quad \frac{1}{\sqrt{n}} \int_{h_0}^{+\infty} \int_{-\infty}^{+\infty} \left( \frac{F_\lambda(t+h) - F_\lambda(t)}{h} \right)^2 dt dh \leq \\ \leq \frac{2}{\sqrt{n}} \sum_{l>h_0-1} \sum_{k=-\infty}^{+\infty} \left( \frac{F_\lambda(k+l) - F_\lambda(k)}{l} \right)^2 + 40 \left( \frac{1}{h_0-1} + \frac{1}{(h_0-1)^2} \right),$$

which concludes the first step of the argument.

To prove Lemma 6.1, it suffices now to establish the following Lemma.

**Lemma 6.4.** *For any  $\delta$  satisfying  $0 < \delta < \frac{1}{4}$ , any  $K, L > 0$  there exists a positive constant  $C(K, L, \delta)$  such that for any  $n \in \mathbb{N}$  any  $\lambda \in \mathbb{Y}_n(K, \delta)$  and any  $h \geq 1$  we have:*

$$\frac{2}{\sqrt{n}} \sum_{l \geq h} \sum_{k=-\infty}^{+\infty} \left( \frac{F_\lambda(k+l) - F_\lambda(k)}{l} \right)^2 \leq \\ \leq \frac{2}{\sqrt{n}} \sum_{l \geq h} \sum_{k=-\infty}^{+\infty} \left( \frac{F_\lambda^{(L,\delta)}(k+l) - F_\lambda^{(L,\delta)}(k)}{l} \right)^2 + C(K, L, \delta) \cdot n^{2\delta - \frac{1}{2}}.$$

The next step is to pass from  $F_\lambda$  to  $F_\lambda^{(L,\delta)}$ . Assume  $\lambda \in \mathbb{Y}_n(K, \delta)$ . Denote

$$\check{F}_\lambda^{(L,\delta)}(t) = F_\lambda(t) - F_\lambda^{(L,\delta)}(t).$$

The support of  $\check{F}_\lambda^{(L,\delta)}$  consists of two intervals, each of length at most  $(K + L) \cdot n^\delta$ . Write

$$\check{I}_\lambda^{(L,\delta)} = \{t : \check{F}_\lambda^{(L,\delta)}(t) \neq 0\},$$



and, for  $l \in \mathbb{N}$ , denote

$$\check{I}_{\lambda,l}^{(L,\delta)} = \{k \in \mathbb{Z} : [k, k+l] \cap \check{I}_{\lambda}^{(L,\delta)} \neq \emptyset\}.$$

By definition, for the cardinality of  $\check{I}_{\lambda,l}^{(L,\delta)}$  we have:

$$\#\check{I}_{\lambda,l}^{(L,\delta)} \leq 2(2l + (K + L)n^{\delta}),$$

whence, using the clear inequality

$$(F_{\lambda}(k+l) - F_{\lambda}(k))^2 \leq 4l^2,$$

we arrive, for any  $l \in \mathbb{N}$ , at the inequality

$$\sum_{k \in \check{I}_{\lambda,l}^{(L,\delta)}} (F_{\lambda}(k+l) - F_{\lambda}(k))^2 \leq 8l^2 (2l + (K + L)n^{\delta}),$$

and, consequently, at the inequality

(43)

$$\begin{aligned} \sum_{k=-\infty}^{+\infty} (F_{\lambda}(k+l) - F_{\lambda}(k))^2 &\leq \sum_{k=-\infty}^{+\infty} \left( F_{\lambda}^{(L,\delta)}(k+l) - F_{\lambda}^{(L,\delta)}(k) \right)^2 + \\ &\quad + 8l^2 (2l + (K + L)n^{\delta}). \end{aligned}$$

Furthermore, again using the fact that  $\check{I}_{\lambda}^{(L,\delta)}$  has measure at most  $2(K + L)n^{\delta}$ , and the Lipschitz property of  $\check{F}_{\lambda}^{(L,\delta)}$  for any  $l \in \mathbb{N}$  we have the inequality

$$\sum_{k=-\infty}^{+\infty} \left( \check{F}_{\lambda}^{(L,\delta)}(k+l) - \check{F}_{\lambda}^{(L,\delta)}(k) \right)^2 \leq 64 \cdot (K + L)^3 \cdot n^{3\delta}.$$

Using the clear inequality

$$\begin{aligned} (44) \quad &\sum_{k=-\infty}^{+\infty} (F_{\lambda}(k+l) - F_{\lambda}(k))^2 \leq \\ &\leq 2 \sum_{k=-\infty}^{+\infty} \left( F_{\lambda}^{(L,\delta)}(k+l) - F_{\lambda}^{(L,\delta)}(k) \right)^2 + 2 \sum_{k=-\infty}^{+\infty} \left( \check{F}_{\lambda}^{(L,\delta)}(k+l) - \check{F}_{\lambda}^{(L,\delta)}(k) \right)^2 \end{aligned}$$

we obtain, for any  $l \in \mathbb{N}$ , the inequality

$$\begin{aligned} (45) \quad &\sum_{k=-\infty}^{+\infty} (F_{\lambda}(k+l) - F_{\lambda}(k))^2 \leq \\ &\leq 2 \sum_{k=-\infty}^{+\infty} \left( F_{\lambda}^{(L,\delta)}(k+l) - F_{\lambda}^{(L,\delta)}(k) \right)^2 + 128 \cdot (K + L)^3 \cdot n^{3\delta}. \end{aligned}$$

We proceed to summing in  $l$ . First we sum in  $l \in [h_0, n^\delta]$  using inequality (43), and then we sum in  $l \in [n^\delta, +\infty)$  using inequality (45). From (43) we immediately obtain

$$(46) \quad \sum_{h_0 \leq l \leq n^\delta} \sum_{k=-\infty}^{\infty} \left( \frac{F_\lambda(k+l) - F_\lambda(k)}{l} \right)^2 \leq \\ \leq \sum_{h_0 \leq l \leq n^\delta} \sum_{k=-\infty}^{\infty} \left( \frac{F_\lambda^{(L,\delta)}(k+l) - F_\lambda^{(L,\delta)}(k)}{l} \right)^2 + 8 \cdot (K+L+2) \cdot n^{2\delta}.$$

From (45) we have

$$(47) \quad \sum_{l > n^\delta} \sum_{k=-\infty}^{\infty} \left( \frac{F_\lambda(k+l) - F_\lambda(k)}{l} \right)^2 \leq \\ \leq 2 \sum_{l > n^\delta} \sum_{k=-\infty}^{\infty} \left( \frac{F_\lambda^{(L,\delta)}(k+l) - F_\lambda^{(L,\delta)}(k)}{l} \right)^2 + 128 \cdot (K+L)^3 \cdot (n^{2\delta} + n^\delta).$$

Indeed, to prove (47), it suffices to observe that

$$\sum_{l > n^\delta} \frac{1}{l^2} < \frac{1}{n^{2\delta}} + \frac{1}{n^\delta}.$$

Combining (46) and (47), we conclude the proof of the Lemma.

**6.3. Average Variance of the Discrete Bessel Process.** Let  $l_0 > 1$  and for  $r \in \mathbb{N}$  denote

$$\psi(l_0, r) = \begin{cases} 2 \left( \frac{1}{l_0} + \dots + \frac{1}{r-1} + r \left( \sum_{n=r}^{+\infty} \frac{1}{n^2} \right) \right), & \text{if } r > l_0 \\ 2r \left( \sum_{n=l_0}^{+\infty} \frac{1}{n^2} \right), & \text{if } r \leq l_0 \end{cases}$$

**Proposition 6.5.** For any  $\theta \in \mathbb{C}$  we have

$$\sum_{l=l_0}^{+\infty} \sum_{k=-\infty}^{+\infty} \frac{\text{Var}_{J(\theta^2)}(c_k + \dots + c_{k+l-1})}{l^2} = \sum_{r=1}^{+\infty} \psi(l_0, r) \cdot \left( \sum_{\substack{x, y \in \mathbb{Z} \\ y-x=r}} (J(x, y; \theta^2))^2 \right).$$

*Proof.* From the well-known identity

$$J(x, x; \theta^2) = \sum_{y \in \mathbb{Z}} (J(x, y; \theta^2))^2$$

we derive

$$\mathrm{Var}_{J(\theta^2)}(c_k + \dots + c_{k+l-1}) = \sum_{\substack{x \in \mathbb{Z} \\ x \in [k, k+l]}} \sum_{\substack{y \in \mathbb{Z} \\ y \notin [k, k+l]}} (J(x, y; \theta^2))^2.$$

Summing in  $k$ , we obtain

$$\sum_{k=-\infty}^{+\infty} \mathrm{Var}_{J(\theta^2)}(c_k + \dots + c_{k+l-1}) = 2 \sum_{\substack{x, y \in \mathbb{Z} \\ x < y}} \min(y - x, l) \cdot (J(x, y; \theta^2))^2.$$

Dividing by  $l^2$  and summing in  $l$  we obtain the Proposition.  $\square$

Recall that  $\mathcal{N}(n, \delta) = \{x \in \mathbb{Z} : |x| \leq 2\sqrt{n} - n^\delta\}$ .

Summing only over indices belonging to  $\mathcal{N}(n, \delta)$ , we obtain

**Corollary 6.6.** *For any  $\delta > 0$ ,  $l > 1$  we have*

$$\sum_{l > l_0} \sum_{\substack{k \in \mathcal{N}(n, \delta) \\ k+l \in \mathcal{N}(n, \delta)}} \left| \frac{\mathrm{Var}_{J(\theta^2)}(c_k + \dots + c_{k+l-1})}{l^2} \right| \leq \sum_{r=1}^{+\infty} \psi(l_0, r) \cdot \sum_{\substack{x \in \mathcal{N}(n, \delta), y \in \mathbb{Z} \\ |x-y|=r}} |J(x, y; \theta^2)|^2$$

**6.4. Estimates on the Bessel Kernel.** To estimate the average variance of the discrete Bessel process using Corollary 6.6, we need estimates for the discrete Bessel kernel for various values of the parameters. We formulate these estimates in this subsection and postpone their routine proofs until the last Section. Very simple estimates on the Bessel kernel are quite sufficient for our purposes.

We start with the following estimate for the Bessel function.

**Lemma 6.7.** *There exists  $\varepsilon_0 > 0$  such that the following holds. For any  $\delta > \frac{1}{6}$  there exist constants  $C > 0$ ,  $\gamma > 0$  depending only on  $\delta$  such that for all  $x \in \mathbb{N}$  satisfying*

$$0 \leq x \leq 2\sqrt{n} - n^\delta$$

and all  $\theta \in \mathbb{C}$  satisfying  $\left| \frac{\theta}{\sqrt{n}} - 1 \right| < \varepsilon_0$  we have

$$(1) \quad |J_x(2\theta)| \leq \frac{C e^{\gamma|\theta - \sqrt{n}|}}{n^{1/8} \sqrt[4]{2\sqrt{n} - x}};$$

$$(2) \quad |J_{x+1}(2\theta) - J_x(2\theta)| \leq \frac{C e^{\gamma|\theta - \sqrt{n}|} \sqrt[4]{2\sqrt{n} - x}}{n^{3/8}}$$

**Remark.** By symmetry, for  $x < 0$ ,  $|x| \leq 2\sqrt{n} - n^\delta$ , we have

$$|J_x(2\theta)| \leq \frac{C e^{\gamma|\theta - \sqrt{n}|}}{n^{1/8} \sqrt[4]{2\sqrt{n} - |x|}};$$

$$|J_{x+1}(2\theta) + J_x(2\theta)| \leq \frac{C e^{\gamma|\theta - \sqrt{n}|} \sqrt[4]{2\sqrt{n} - |x|}}{n^{3/8}}.$$

This is easily proved using the contour integral for the Bessel function; for the reader's convenience, the proof is given in the last Section.

**Corollary 6.8.** *Let  $\delta > \frac{1}{6}$ . There exist constants  $C > 0$ ,  $\gamma > 0$ ,  $\varepsilon > 0$  such that for all  $n \in \mathbb{N}$ , all  $\theta \in \mathbb{C}$  satisfying*

$$\left| \frac{\theta}{\sqrt{n}} - 1 \right| < \varepsilon$$

and all  $x, y \in \mathbb{N}$ ,  $x \neq y$ ,  $x, y \in \mathcal{N}(n, \delta)$  we have

$$|J(x, y; \theta^2)| \leq \frac{C}{|x - y|} \left( \sqrt[4]{\frac{2\sqrt{n} - x}{2\sqrt{n} - y}} + \sqrt[4]{\frac{2\sqrt{n} - y}{2\sqrt{n} - x}} \right) \cdot e^{\gamma|\theta - \sqrt{n}|}.$$

*Proof.* Immediate from the preceding Lemma and the formula

$$(48) \quad \begin{aligned} J_{x+1}(2\theta)J_y(2\theta) - J_{y+1}(2\theta)J_x(2\theta) &= \\ &= (J_{x+1}(2\theta) - J_x(2\theta))J_y(2\theta) - (J_{y+1}(2\theta) - J_y(2\theta))J_x(2\theta). \end{aligned}$$

□

**Corollary 6.9.** *Let  $\delta > \frac{1}{6}$ . There exist constants  $C > 0$ ,  $\gamma > 0$ ,  $\varepsilon > 0$  such that for all  $n \in \mathbb{N}$ , all  $\theta \in \mathbb{C}$  satisfying*

$$\left| \frac{\theta}{\sqrt{n}} - 1 \right| < \varepsilon$$

and all  $x, y \in \mathbb{Z}$ ,  $x \neq y$ ,  $x, y \in \mathcal{N}(n, \delta)$  we have

$$|J(x, y; \theta^2)| \leq C \cdot \frac{\sqrt[4]{n} \cdot e^{\gamma|\theta - \sqrt{n}|}}{\sqrt[4]{(2\sqrt{n} - |x|)(2\sqrt{n} - |y|)} |x - y|}.$$

These are immediate from Lemma 6.7.

For the values of  $x$  that are “close the edge”, we use the following estimate.

**Proposition 6.10.** *There exist constants  $C > 0$ ,  $\varepsilon > 0$ ,  $\gamma > 0$  such that for all  $x \in \mathbb{Z}$ , all  $n \in \mathbb{N}$  and all  $\theta \in \mathbb{C}$  satisfying*

$$\left| \frac{\theta}{\sqrt{n}} - 1 \right| < \varepsilon$$

we have

$$|J_x(2\theta)| \leq \frac{C e^{\gamma|\theta - \sqrt{n}|}}{n^{\frac{1}{6}}}.$$

Proposition 6.10 is immediate from the uniform asymptotic estimates of Olver for Bessel functions (see, e.g., Abramowitz and Stegun [1], 9.3.5, 9.3.6, 9.3.38, 9.3.39).

For values of  $x$  “beyond the edge”, we use the following stretched exponential estimate.

**Proposition 6.11.** *There exists constants  $\varepsilon_1 > 0, \varepsilon_2 > 0$  such that the following is true. For any  $\delta > \frac{1}{6}$  there exist constants  $C > 0, \varepsilon > 0, \gamma > 0, \tilde{\gamma} > 0$  such that for all  $\delta : \delta_0 \leq \delta \leq \frac{1}{2}$ , all  $n \in \mathbb{N}$ , all  $\theta \in \mathbb{C}$  satisfying  $\left| \frac{\theta}{\sqrt{n}} - 1 \right| < \varepsilon$  and all  $x \in \mathbb{N}$  satisfying  $2\sqrt{n} + n^\delta < x < (2 + \varepsilon_1)\sqrt{n}$ , we have*

$$|J_x(2\theta)| \leq C \exp \left( -\tilde{\gamma} \frac{(x - 2\sqrt{n})^{3/2}}{n^{1/4}} + \gamma |\theta - \sqrt{n}| \right),$$

while for  $x > (2 + \varepsilon_1)\sqrt{n}$  we have

$$|J_x(2\theta)| \leq C \exp \left( -\varepsilon_2(x - 2\sqrt{n}) + \gamma |\theta - \sqrt{n}| \right).$$

Proposition 6.11 is immediate from the contour integral representation of Bessel functions. Note that Proposition 2.5 is immediate from Proposition 6.11 by depossionization.

**6.5. Proof of Lemma 6.2.** We now derive Lemma 6.2 from Corollary 6.6.

*Proof.* First, note the clear inequality

$$\psi(l_0, r) < \begin{cases} 2 \left( \log \left( \frac{r}{l_0-1} \right) + \frac{r}{r-1} \right) & \text{for } l_0 \leq r \\ \frac{2r}{l_0-1} & \text{for } l_0 > r \end{cases}$$

We further assume that  $n$  is large enough so that  $n^\delta > l_0$ . We consider several cases.

*Case 1. The Bulk.*

First, take  $\varepsilon > 0$  and denote

$$\mathcal{N}_\varepsilon = \{ (x, y) \in \mathbb{Z}^2 : |x|, |y| \leq (2 - \varepsilon)\sqrt{n} \}.$$

In this case, the Debye asymptotics implies

$$|J(x, y; \theta^2)|^2 \leq \frac{C e^{\gamma|\theta - \sqrt{n}|}}{(|x - y| + 1)^2},$$

and we consequently have

$$\frac{1}{\sqrt{n}} \sum_{(x,y) \in \mathcal{N}_\varepsilon} \psi(l_0, r) |J(x, y; \theta^2)|^2 \leq \frac{C \log l_0}{l_0} e^{\gamma|\theta - \sqrt{n}|}.$$

We proceed to the analysis of the remaining terms. Recall our notation

$$\mathcal{N}(n, \delta) = \{ x \in \mathbb{Z} : |x| \leq 2\sqrt{n} - n^\delta \}, \quad r = |x - y|$$

and denote

$$\mathcal{N}_2(n, \delta) = \left\{ y \in \mathbb{Z} : |y| \leq 2\sqrt{n} - \frac{1}{2}n^\delta \right\}.$$

*Case 2. Close Points.*

Denote

$$\mathcal{N}_2 = \left\{ (x, y) \in \mathbb{Z}^2 : x \in \mathcal{N}(n, \delta), y \in \mathcal{N}_2(n, \delta), |x - y| \leq n^\delta \right\}.$$

In this case, again, we have

$$|J(x, y; \theta^2)|^2 \leq \frac{C e^{\gamma|\theta - \sqrt{n}|}}{(|x - y| + 1)^2}.$$

Indeed, if  $x$  and  $y$  are both positive, then the estimate follows from Corollary 6.8; if both are negative, then the estimate follows by symmetry; while if  $x$  and  $y$  have different signs, then  $(x, y) \in \mathcal{N}_\varepsilon$ , and the estimate also holds. Consequently, again we have

$$\frac{1}{\sqrt{n}} \sum_{(x, y) \in \mathcal{N}_2} \psi(l_0, r) |J(x, y; \theta^2)|^2 \leq \frac{C \log l_0}{l_0} e^{\gamma|\theta - \sqrt{n}|}.$$

**Remark.** Of course,  $\mathcal{N}_\varepsilon \cap \mathcal{N}_2 \neq \emptyset$ , but that does not matter since we are only concerned with upper estimates.

*Case 3. Distant Positive Points.*

Set

$$\mathcal{N}_3 = \left\{ (x, y) \in \mathbb{Z}^2 : x \geq 0, y \geq 0, x \in \mathcal{N}(n, \delta), y \in \mathcal{N}_2(n, \delta), |x - y| > n^\delta \right\}.$$

For definiteness, let  $x > y$ . Corollary 6.8 implies the bound

$$|J(x, y; \theta^2)|^2 \leq \frac{C e^{\gamma|\theta - \sqrt{n}|}}{r^2} \left( 1 + \sqrt{\frac{r}{2\sqrt{n} - x}} \right)$$

and, consequently, there exists  $\delta_2 > 0$  depending only on  $\delta$  such that

$$\frac{1}{\sqrt{n}} \sum_{(x, y) \in \mathcal{N}_3} \psi(l_0, r) \cdot |J(x, y; \theta^2)|^2 \leq C n^{-\delta_2} \cdot e^{\gamma|\theta - \sqrt{n}|}.$$

*Case 4. Distant Negative Points.* Set

$$\mathcal{N}_4 = \left\{ (x, y) \in \mathbb{Z}^2 : x \leq 0, y \leq 0, x \in \mathcal{N}(n, \delta), y \in \mathcal{N}_2(n, \delta), |x - y| > n^\delta \right\}.$$

This case is similar to the previous one.

*Case 5. Distant Points of Opposite Signs.*

Set

$$\mathcal{N}_5 = \left\{ (x, y) \in \mathbb{Z}^2 : x \in \mathcal{N}(n, \delta), y \in \mathcal{N}_2(n, \delta), (x, y) \notin \mathcal{N}_\varepsilon, x \text{ and } y \text{ have opposite signs} \right\}.$$

In this case  $|x - y| \geq \sqrt{n}$ , and Corollary 6.9 implies

$$|J(x, y; \theta^2)| \leq \frac{Ce^{\gamma|\theta - \sqrt{n}|}}{\sqrt{n} \cdot \sqrt[4]{(2\sqrt{n} - |x|)(2\sqrt{n} - |y|)}},$$

whence

$$\frac{1}{\sqrt{n}} \sum_{(x,y) \in \mathcal{N}_4} \psi(l_0, r) \cdot |J(x, y; \theta^2)|^2 \leq \frac{Ce^{\gamma|\theta - \sqrt{n}|}}{\sqrt{n}}.$$

*Case 6. Large Values of  $y$ .*

Set

$$\mathcal{N}_6 = \{(x, y) \in \mathbb{Z}^2 : x \in \mathcal{N}(n, \delta), |y| \geq 2\sqrt{n} + n^\delta\}.$$

In this case, by Proposition 6.11, there exists  $\delta_6 > 0$  depending only on  $\delta$  such that

$$|J(x, y; \theta^2)| \leq Ce^{\gamma|\theta - \sqrt{n}| - (y - 2\sqrt{n})^{\delta_6}}$$

and, consequently, there exists  $\delta_7 > 0$  depending only on  $\delta_6$  such that we have

$$\sum_{(x,y) \in \mathcal{N}_6} \psi(l_0, r) \cdot |J(x, y; \theta^2)|^2 \leq Ce^{\gamma|\theta - \sqrt{n}| - n^{\delta_7}}.$$

*Case 7. The Point  $y$  on the Edge.*

Set

$$\mathcal{N}_7 = \left\{ (x, y) \in \mathbb{Z}^2 : x \in \mathcal{N}(n, \delta), 2\sqrt{n} - \frac{1}{2}n^\delta \leq |y| \leq 2\sqrt{n} + n^\delta \right\}.$$

In this case, Lemma 6.7 and Proposition 6.10 give

$$|J(x, y; \theta^2)| \leq \frac{Ce^{\gamma|\theta - \sqrt{n}|} \cdot n^{\frac{5}{24}}}{(2\sqrt{n} - |x|)^{\frac{5}{4}}},$$

whence

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{(x,y) \in \mathcal{N}_7} \psi(l_0, r) \cdot |J(x, y; \theta^2)|^2 &\leq \\ &\leq Ce^{\gamma|\theta - \sqrt{n}|} \cdot \frac{n^{\frac{5}{12}}}{\sqrt{n}} \cdot n^\delta \cdot \sum_{x \in \mathcal{N}(n, \delta)} \frac{1}{(2\sqrt{n} - |x|)^{\frac{5}{2}}} \leq \\ &\leq Ce^{\gamma|\theta - \sqrt{n}|} \cdot n^{-\frac{1}{12} - \frac{\delta}{2}}. \end{aligned}$$

The Lemma is proved completely.  $\square$

## 7. PROOFS OF ESTIMATES FOR THE DISCRETE BESSEL KERNEL.

7.1. **Proof of Lemma 6.7.** It is convenient to prove the following equivalent reformulation of Lemma 6.7.

**Lemma 7.1.** *There exists  $\varepsilon_0 > 0$  such that the following holds.*

*For any  $\delta_0 > \frac{1}{6}$  and any  $K > 0$  there exist constants  $C > 0$ ,  $\gamma > 0$  depending only on  $\delta_0$  and  $K$ , such that for all  $\delta : \delta_0 \leq \delta \leq \frac{1}{2}$ ,  $n \in \mathbb{N}$ , all  $x \in \mathbb{N}$  satisfying*

$$x \leq 2\sqrt{n} - Kn^\delta$$

*and all  $\theta \in \mathbb{C}$  satisfying  $\left| \frac{\theta}{\sqrt{n}} - 1 \right| < \varepsilon_0$  we have*

- (1)  $|J_x(2\theta)| \leq \frac{Ce^{\gamma|\theta - \sqrt{n}|}}{n^{\frac{\delta}{4} + \frac{1}{8}}}$ ;
- (2)  $|J_{x+1}(2\theta) - J_x(2\theta)| \leq \frac{Ce^{\gamma|\theta - \sqrt{n}|}}{n^{\frac{3}{8} - \frac{\delta}{4}}}$

Throughout the proof, the symbols  $C$  and  $\gamma$  will denote constants depending only on  $\delta_0$  and  $K$ . Let  $\mathcal{K}$  be a contour going around 0 counterclockwise once. We then have the following integral representation for the Bessel function:

$$J_x(2\theta) = \frac{1}{2\pi i} \oint_{\mathcal{K}} e^{\theta(z-z^{-1})} \cdot z^{-x-1} dz$$

We choose  $\mathcal{K}$  as follows. Take  $\varepsilon > 0$  sufficiently small. Denote  $u = \frac{x+1}{\sqrt{n}}$  and introduce the angle  $\varphi_u$ ,  $0 \leq \varphi_u \leq \frac{\pi}{2}$ , by the formula

$$2 \cos \varphi_u = u.$$

Introduce the arcs  $I^+$ ,  $I^-$  by the formulas:

$$\begin{aligned} I^+ &= e^{i\varphi_u} + \frac{1+i}{\sqrt{2}} e^{i\varphi_u} \cdot t, & |t| &\leq n^{-\frac{1}{8} - \frac{\delta}{4} + \varepsilon} \\ I^- &= e^{-i\varphi_u} + \frac{1-i}{\sqrt{2}} e^{-i\varphi_u} \cdot t, & |t| &\leq n^{-\frac{1}{8} - \frac{\delta}{4} + \varepsilon} \end{aligned}$$

We now complete the contour  $\mathcal{K}$  by drawing a circle arc  $O^+$  counterclockwise from the outer endpoint of  $I^+$  to the outer endpoint of  $I^-$ , and, similarly, drawing a circle arc  $O^-$  counterclockwise from the inner endpoint of  $I^-$  to the inner endpoint of  $I^+$  (see Fig. 1). We estimate the contribution of each arc separately and show that the arcs  $I^+$ ,  $I^-$  give the main contribution, while the contribution of the arcs  $O^+$ ,  $O^-$  is negligible. We start with the arc  $I^+$ . Write

$$e^{\theta(z-z^{-1})} \cdot z^{-x-1} = e^{\sqrt{n}(z-z^{-1}-u \log z)} \cdot e^{(\theta - \sqrt{n})(z-z^{-1})},$$

and denote

$$S(z) = z - z^{-1} - u \log z.$$



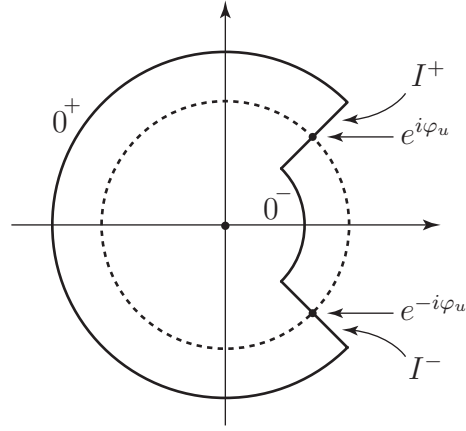


FIGURE 1.

Write

$$z_t = e^{i\varphi_u} + \frac{1+i}{\sqrt{2}} e^{i\varphi_u} \cdot t$$

and denote  $\tilde{S}(t) = S(z_t)$ .

Since

$$\frac{dS}{dz} \Big|_{z=e^{i\varphi_u}} = 0, \quad \frac{d^2S}{dz^2} \Big|_{z=e^{i\varphi_u}} = 2i \sin \varphi_u \cdot e^{-2i\varphi_u}$$

and since there exists a constant  $C$  such that

$$\left| \frac{d^3S}{dz^3} \right| \leq C \quad \text{for all } z \in I^+,$$

one can write

$$\tilde{S}(t) = -\sin \varphi_u \cdot t^2 + A(t)t^3,$$

where  $|A(t)| \leq C$  provided  $|t| \leq n^{-\frac{1}{8}-\frac{\delta}{4}+\varepsilon}$ .

Consequently,

$$\begin{aligned} \int_{I^+} e^{\theta(z-z^{-1})} \cdot z^{-x-1} dz &= \int_{I^+} e^{(\theta-\sqrt{n})(z_t-z_t^{-1})} \cdot e^{\sqrt{n}\tilde{S}(t)} dt = \\ &= \int_{I^+} e^{(\theta-\sqrt{n})(z_t-z_t^{-1})} \cdot e^{-\sqrt{n}\sin \varphi_u \cdot t^2} \cdot e^{\sqrt{n}A(t) \cdot t^3} dt. \end{aligned}$$

Noting that  $z_t - z_t^{-1}$  is bounded on  $I^+$ , that we have

$$\left| \sqrt{n} \cdot A(t) \cdot t^3 \right| \leq C \cdot n^{\frac{1}{8}-\frac{3\delta}{4}+3\varepsilon},$$

where the exponent is negative as long as  $\varepsilon < \frac{\delta_0}{4} - \frac{1}{24}$  (here we use the condition  $\delta_0 > \frac{1}{6}$ ), and that

$$\sin \varphi_u \geq C \cdot n^{\frac{\delta}{2} - \frac{1}{4}},$$

we arrive at the estimate

$$\left| \int_{I^+} e^{\theta(z-z^{-1})} \cdot z^{-x-1} dz \right| \leq \frac{C e^{\gamma|\theta - \sqrt{n}|}}{n^{\frac{1}{8} + \frac{\delta}{4}}}.$$

The contribution of the arc  $I^-$  is estimated in the same way. It remains to estimate the contributions of the circular arcs  $O^+$  and  $O^-$ . Our aim is to show that there exists  $\varepsilon_0 > 0$  such that

$$\Re(S(z)) < -Cn^{\varepsilon_0}.$$

for all  $z \in O^+ \cup O^-$ . We only show it for  $O^+$ , as the case of  $O^-$  is completely similar.

For  $z \in O^+$  write

$$z = r_z e^{i\varphi_z}.$$

There exist positive constants  $C_1, C_2, C_3$  such that

$$(49) \quad \begin{aligned} r_z &\geq 1 + C_1 n^{-\frac{\delta}{4} - \frac{1}{8} + \varepsilon}, \\ r_z &\leq 1 + C_2 n^{-\frac{\delta}{4} - \frac{1}{8} + \varepsilon}, \\ \varphi_z - \varphi_u &\geq C_3 n^{-\frac{\delta}{4} - \frac{1}{8} + \varepsilon}. \end{aligned}$$

Consider the function

$$S^\#(t) = \Re S(te^{i\varphi_z}), \quad t \in [1, r_z].$$

Note that  $S^\#(1) = 0$ . From (49) we have:

$$\frac{dS^\#}{dt} \Big|_{t=1} \leq -C_{11} n^{\frac{\delta}{4} - \frac{3}{8} + \varepsilon}, \quad \left| \frac{d^2 S^\#}{dt^2} \Big|_{t=1} \right| \leq C_{12} n^{\frac{\delta}{2} - \frac{1}{4}}, \quad \max_{t \in [1, r_z]} \left| \frac{d^3 S^\#}{dt^3} \right| \leq C_{14},$$

whence there exists  $\varepsilon_0 > 0$  such that

$$S^\#(r_z) \leq -C_{17} n^{\varepsilon_0}$$

for all  $z \in O^+$ . We conclude that

$$\left| \int_{O^+} e^{\theta(z-z^{-1})} \cdot z^{-x-1} dz \right| \leq e^{\gamma|\theta - \sqrt{n}| - Cn^{\varepsilon_0}}.$$

The case of  $O^-$  is similar, and the first claim of the Lemma is proved completely.

We proceed to the proof of the second claim. Write

$$J_x(2\theta) - J_{x-1}(2\theta) = \frac{1}{2\pi i} \oint_{\mathcal{K}} e^{\theta(z-z^{-1})} \cdot z^{-x-1} \cdot (1-z) dz.$$

The contour  $\mathcal{K}$  stays the same. We first estimate the contribution of the interval  $I^+$ . Write

$$1-z = 1 - e^{i\varphi u} + e^{i\varphi u} - z,$$

and note that, since

$$|1 - e^{i\varphi u}| \leq Cn^{\frac{2\delta-1}{4}},$$

we have

$$\left| \int_{I^+} e^{\theta(z-z^{-1})} \cdot z^{-x-1} \cdot (1 - e^{i\varphi u}) dz \right| \leq \frac{Ce^{\gamma|\theta-\sqrt{n}|}}{n^{\frac{3}{8}-\frac{\delta}{4}}}.$$

This is the main contribution. We proceed to the analysis of the remaining terms and estimate

$$\int_{I^+} e^{\theta(z-z^{-1})} \cdot z^{-x-1} \cdot (e^{i\varphi u} - z) dz.$$

Write

$$e^{\theta(z-z^{-1})} \cdot z^{-x-1} = e^{\theta(z-z^{-1}-u \log z)} \cdot e^{(\theta-\sqrt{n})u \log z}.$$

The notation  $t$ ,  $z_t$  and  $A(t)$  has the same meaning as before, and we write

$$\begin{aligned} (50) \quad & e^{\theta(z_t-z_t^{-1}-u \log z_t)} \cdot e^{(\theta-\sqrt{n})u \log z_t} \cdot (e^{i\varphi u} - z_t) = \\ & = e^{-\theta \sin \varphi_u \cdot t^2} \left( e^{i(\theta-\sqrt{n})u \varphi_u} \right) \cdot \frac{1+i}{\sqrt{2}} e^{i\varphi_u} \cdot t + \\ & + e^{-\theta \sin \varphi_u \cdot t^2} \left( e^{(\theta-\sqrt{n})u \log z_t + A(t) \cdot t^3} - e^{i(\theta-\sqrt{n})u \varphi_u} \right) \times \frac{1+i}{\sqrt{2}} e^{i\varphi_u} \cdot t. \end{aligned}$$

Recall that

$$dz_t = \frac{1+i}{\sqrt{2}} e^{i\varphi_u} dt$$

and note that the first summand in the right hand side of (50) is an *odd* function of  $t$ , so its integral over  $I^+$  is zero. We now estimate the second summand in absolute value. Note that

$$e^{(\theta-\sqrt{n})u \log z_t + A(t) \cdot t^3} \Big|_{t=0} = e^{i(\theta-\sqrt{n})u \varphi_u}$$

and that on  $I^+$  we have an estimate

$$\left| \frac{d}{dt} e^{(\theta-\sqrt{n})u \log z_t + A(t) \cdot t^3} \right| \leq C_{20} e^{\gamma|\theta-\sqrt{n}|}.$$

Furthermore,

$$\left| e^{-\theta \sin \varphi_u \cdot t^2} \right| \leq e^{-\frac{\sqrt{n}}{2} \sin \varphi_u \cdot t^2}.$$

In view of all the above, we have

$$\left| e^{-\theta \sin \varphi_u \cdot t^2} \left( e^{(\theta - \sqrt{n})u \log z_t + A(t) \cdot t^3} - e^{i(\theta - \sqrt{n})u \varphi_u} \right) \right| \leq C e^{-\frac{\sqrt{n}}{2} \sin \varphi_u \cdot t^2} \cdot |t| \cdot e^{\gamma |\theta - \sqrt{n}|}.$$

Noting that

$$\int_{|t| \leq n^{\frac{\delta}{2} - \frac{1}{4} + \varepsilon}} e^{-\frac{\sqrt{n}}{2} \sin \varphi_u \cdot t^2} \cdot t^2 dt \leq \frac{C}{n^{\frac{3}{8} + \frac{3\delta}{4}}},$$

we conclude that the contribution of the second summand in (50) is bounded above by

$$\frac{C e^{\gamma |\theta - \sqrt{n}|}}{n^{\frac{3}{8} + \frac{3\delta}{4}}},$$

and so is negligible compared to the main contribution. Finally, we have

$$\left| \int_{I^+} e^{\theta(z-z^{-1})} \cdot z^{-x-1} \cdot (1-z) dz \right| \leq \frac{C e^{\gamma |\theta - \sqrt{n}|}}{n^{\frac{3}{8} - \frac{\delta}{4}}}.$$

The contribution of  $I^-$  is estimated in the same way, and the contribution of the circular arcs is shown to be negligible in exactly the same way as in the proof of the first Claim. The Lemma is proven completely.

**7.2. Proof of Lemma 6.3.** We start with Okounkov's integral formula for the discrete Bessel kernel [12]. We take any positive numbers  $\alpha_1 > \alpha_2 > 0$  and write

$$(51) \quad J(x, y; \theta^2) = \frac{1}{(2\pi i)^2} \int_{|z|=\alpha_1} \int_{|w|=\alpha_2} \frac{e^{\theta(z-z^{-1}-w+w^{-1})}}{(z-w)z^{x+1}w^{-y}} dz du.$$

As before, we define  $\varphi_u$  by formula  $2 \cos \varphi_u = u$ , and we set

$$S(z, u) = z - z^{-1} - u \log z$$

(the principal branch of the logarithm is taken here). Setting  $x = y = u\sqrt{n}$ , we rewrite (51) as follows

$$J(x, x; \theta^2) = \frac{1}{(2\pi i)^2} \int_{|z|=\alpha_1} \int_{|w|=\alpha_2} \frac{e^{\theta(S(z,u)-S(w,u))+(\theta-\sqrt{n})u(\log z - \log w)}}{(z-w)z} dz dw.$$

Now, following Okounkov [12], we deform the contour of integration and obtain an integral representation for the quantity

$$J(x, x; \theta^2) - \frac{\varphi_u}{\pi} = J(x, x; \theta^2) - \frac{\arccos \frac{x}{2\sqrt{n}}}{\pi}.$$

We take  $\varepsilon > 0$  sufficiently small and introduce the intervals  $I_z^+$ ,  $I_z^-$  as before:

$$I_z^\pm = e^{\pm i\varphi_u} + \frac{1 \pm i}{\sqrt{2}} e^{\pm i\varphi_u} \cdot t, \quad |t| \leq n^{\frac{1}{8} - \frac{\delta}{4} + \varepsilon}$$

As before, we complete the contour by drawing a circular arc  $O_z^+$  joining the outer endpoints of  $I_z^\pm$  and a circular arc  $O_z^-$  joining the inner endpoints of  $I_z^\pm$ . The resulting contour is denoted  $\mathcal{K}_z$ .

We now introduce the intervals  $I_w^+$ ,  $I_w^-$  by the formulas

$$\begin{aligned} I_w^+ &= e^{i\varphi_u} = \frac{1-i}{\sqrt{2}} e^{i\varphi_u} t & |t| \leq n^{-\frac{1}{8} - \frac{\delta}{4} + \varepsilon}; \\ I_w^- &= e^{-i\varphi_u} = \frac{1+i}{\sqrt{2}} e^{i\varphi_u} t & |t| \leq n^{-\frac{1}{8} - \frac{\delta}{4} + \varepsilon}. \end{aligned}$$

In a similar way, we join the outer and the inner endpoints of  $I_w^\pm$ , respectively, by circle arcs  $O_w^+$  and  $O_w^-$ .

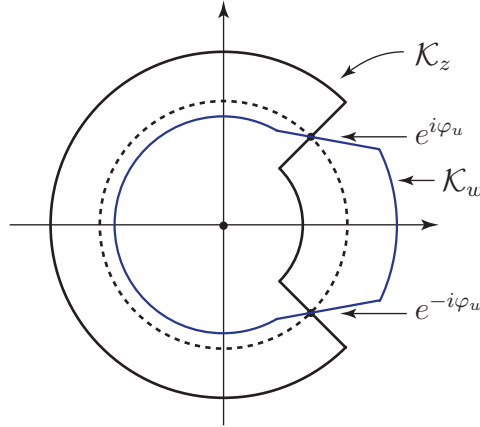


FIGURE 2.

Okounkov [12] showed that

$$(52) \quad J(x, x; \theta^2) - \frac{\arccos \frac{x}{2\sqrt{n}}}{\pi} = \frac{1}{(2\pi i)^2} \int_{\mathcal{K}_z} \int_{\mathcal{K}_w} \frac{e^{\theta(S(z,u)-S(w,u))+(\theta-\sqrt{n})u(\log z-\log w)}}{z(z-w)} dz dw.$$

We estimate the right-hand side of (52), and we begin by estimating

$$(53) \quad \int_{I_z^+} \int_{I_w^+} \frac{e^{\theta(S(z,u)-S(w,u))+(\theta-\sqrt{n})u(\log z-\log w)}}{z(z-w)} dw dz.$$

As before, we write

$$\begin{aligned} z_t &= e^{i\varphi_u} + \frac{1+i}{\sqrt{2}} e^{i\varphi_u} t, \\ w_s &= e^{i\varphi_u} + \frac{1-i}{\sqrt{2}} e^{i\varphi_u} s, \end{aligned}$$

and

$$\begin{aligned} S(z_t, u) &= -\sin \varphi_u \cdot t^2 + A(t) \cdot t^3, \\ S(w_s, u) &= \sin \varphi_u \cdot s^2 + \tilde{A}(s) \cdot s^3. \end{aligned}$$

We set

$$I(n) = \left[ -n^{-\frac{1}{8}-\frac{\delta}{4}+\varepsilon}, n^{-\frac{1}{8}-\frac{\delta}{4}+\varepsilon} \right]$$

and estimate the integral

$$\text{Int}(n) = \int_{I(n)} \int_{I(n)} \frac{e^{-\theta \sin \varphi_u (t^2+s^2) - \theta(A(t) \cdot t^3 + \tilde{A}(s) \cdot s^3) + (\theta - \sqrt{n})u(\log z_t - \log w_s)}}{(t+is) \left(1 + \frac{1+i}{\sqrt{2}} t\right)} ds dt.$$

Since  $I(n)$  is symmetric around the origin, we have

$$\int_{I(n)} \int_{I(n)} \frac{e^{-\theta \sin \varphi_u (t^2+s^2)}}{t+is} ds dt = 0,$$

whence

$$\text{Int}(n) = \int_{I(n)} \int_{I(n)} \frac{e^{-\theta \sin \varphi_u (t^2+s^2)}}{(t+is)} \left[ \frac{e^{(A(t) \cdot t^3 + \tilde{A}(s) \cdot s^3) + (\theta - \sqrt{n})(\log z_t - \log w_s)}}{1 + \frac{1+i}{\sqrt{2}} t} - 1 \right] ds dt.$$

We estimate the integrand in absolute value. It is clear that

$$\begin{aligned} \left| \frac{\partial}{\partial t} \left( \frac{\exp \left( -\theta(A(t) \cdot t^3 + \tilde{A}(s) \cdot s^3) + (\theta - \sqrt{n})(\log z_t - \log w_s) \right)}{1 + \frac{1+i}{\sqrt{2}} t} \right) \right| &\leq C e^{-\gamma|\theta - \sqrt{n}|}, \\ \left| \frac{\partial}{\partial s} \left( \frac{\exp \left( -\theta(A(t) \cdot t^3 + \tilde{A}(s) \cdot s^3) + (\theta - \sqrt{n})(\log z_t - \log w_s) \right)}{1 + \frac{1+i}{\sqrt{2}} t} \right) \right| &\leq C e^{-\gamma|\theta - \sqrt{n}|}, \\ \left| \frac{\exp \left( -\theta \sin \varphi_u \cdot (s^2 + t^2) \right)}{t+is} \right| &\leq \frac{\exp \left( \frac{-\sqrt{n}}{2} \sin \varphi_u \cdot (s^2 + t^2) \right)}{\sqrt{t^2 + s^2}}, \end{aligned}$$

whence

$$|\text{Int}(n)| \leq C e^{\gamma|\theta - \sqrt{n}|} \cdot \int_{I(n)} \int_{I(n)} \frac{e^{\frac{-\sqrt{n}}{2} \sin \varphi_u \cdot (t^2+s^2)}}{\sqrt{t^2 + s^2}} (|t| + |s|) dt ds \leq \frac{\tilde{C} e^{\gamma|\theta - \sqrt{n}|}}{2\sqrt{n-x}}.$$

The contributions of the integrals along the remaining rectangular arcs are estimated in the same way, whereas the contribution of those parts of contours where either  $z \in O_z^\pm$  or  $w \in O_w^\pm$  is immediately seen to be majorated by

$$C \exp(\gamma_1 |\theta - \sqrt{n}| - \gamma_2 n^{\delta_2})$$

with  $\delta_2 > 0$  depending only on  $\delta$ . Lemma 6.3 is proved completely.

Theorem 1.1 is proved completely.

#### REFERENCES

- [1] Handbook of mathematical functions with formulas, graphs, and mathematical tables, edited by M. Abramowitz and I. A. Stegun. Tenth printing, online at <http://people.math.sfu.ca/~cbm/aands/>.
- [2] J. Baik, P. Deift, K. Johansson, On the distribution of the length of the longest increasing subsequence of random permutations, *J. Amer. Math. Soc.* 12(1999), 1119–1178.
- [3] Alexei Borodin; Andrei Okounkov; Grigori Olshanski, Asymptotics of Plancherel measures for symmetric groups, *J. Amer. Math. Soc.* 13 (2000), 481–515.
- [4] Borodin, Alexei; Olshanski, Grigori. Distributions on partitions, point processes, and the hypergeometric kernel. *Comm. Math. Phys.* 211 (2000), no. 2, 335–358.
- [5] Bogachev, L. V.; Su, Ch. G. A central limit theorem for random partitions with respect to the Plancherel measure. (Russian) *Dokl. Akad. Nauk* 414 (2007), no. 3, 295–298.
- [6] Cornfeld, I. P.; Fomin, S. V.; Sinai, Ya. G. Ergodic theory. Translated from the Russian by A. B. Sosinskii. *Grundlehren der Mathematischen Wissenschaften*, 245. Springer-Verlag, New York, 1982.
- [7] Cover, Thomas M.; Thomas, Joy A. Elements of information theory. Second edition. Wiley-Interscience, Hoboken, NJ, 2006.
- [8] M. V. Fedoryuk, Metod Perevala (The Saddle-Point Method), “Nauka”, Moscow, 1977.
- [9] Johansson, Kurt Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann. of Math. (2)* 153 (2001), no. 1, 259–296.
- [10] Johansson, Kurt On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.* 91 (1998), no. 1, 151–204.
- [11] Logan, B. F.; Shepp, L. A. A variational problem for random Young tableaux. *Advances in Math.* 26 (1977), no. 2, 206–222.
- [12] Okounkov, Andrei. Symmetric functions and random partitions. *Symmetric functions 2001: surveys of developments and perspectives*, 223–252, NATO Sci. Ser. II Math. Phys. Chem., 74, Kluwer Acad. Publ., Dordrecht, 2002.
- [13] Olshanski, Grigori, An introduction to harmonic analysis on the infinite symmetric group, *Asymptotic combinatorics with applications to mathematical physics* (St. Petersburg, 2001), 127–160, *Lecture Notes in Mathematics*, 1815, Springer, Berlin, 2003.
- [14] Petersen, Karl, Ergodic theory. *Cambridge Studies in Advanced Mathematics*, 2. Cambridge University Press, Cambridge, 1983.
- [15] Soshnikov, Alexander, Determinantal random point fields. *Uspekhi Mat. Nauk* 55 (2000), no. 5(335), 107–160; translation in *Russian Math. Surveys* 55 (2000), no. 5, 923–975.
- [16] Vershik, A. M.; Kerov, S. V. Asymptotic behaviour of the Plancherel measure of the symmetric group and the limit form of Young tableaux. *Dokl. Akad. Nauk SSSR* 233 (1977), no. 6, 1024–1027.

- [17] Vershik, A. M.; Kerov, S. V. Asymptotic behaviour of the maximum and generic dimensions of irreducible representations of the symmetric group. *Funktsional. Anal. i Prilozhen.* 19 (1985), no. 1, 25–36, 96.
- [18] Vershik, A. M.; Kerov, S. V. Asymptotic theory of the characters of a symmetric group. (Russian) *Funktsional. Anal. i Prilozhen.* 15 (1981), no. 4, 15–27.
- [19] A.M. Vershik, D. Pavlov, Numerical experiments in problems of asymptotic representation theory, *Zap.Nauchn.Sem. POMI*, 373, 2009, 77-93.
- [20] Vershik A.M., Kerov, S.V., Characters and factor representations of the infinite symmetric group, *Dokl. Akad. Nauk SSSR* 257 (1981), no.5, 1037-1040.
- [21] G.N. Watson, *A treatise on the theory of Bessel functions*, Cambridge University Press, 1944.

RICE UNIVERSITY, HOUSTON

THE STEKLOV INSTITUTE OF MATHEMATICS, MOSCOW

THE INSTITUTE FOR INFORMATION TRANSMISSION PROBLEMS, MOSCOW

THE HIGHER SCHOOL OF ECONOMICS, MOSCOW

THE INDEPENDENT UNIVERSITY OF MOSCOW