Journal of Mathematical Sciences, Vol. 168, No. 3, 2010

NUMERICAL EXPERIMENTS IN PROBLEMS OF ASYMPTOTIC REPRESENTATION THEORY

A. Vershik* **and D. Pavlov**† UDC 519.6

The article presents the results of numerical computations of statistics related to Young diagrams, including estimates on the maximum and average (with respect to the Plancherel distribution) dimension of irreducible representations of the symmetric group S_n. The computed limit shapes of two-dimensional and three-dimensional diagrams distributed according to the Richardson statistics are also presented. Bibliography: 14 titles.

1. Introduction

The classical definition of a Young diagram (see [1]) is as follows: a Young diagram of size n is a finite descending ideal in the lattice $\mathbb{Z}_{\leq 0} \times \mathbb{Z}_{\leq 0}$, i.e., a set of cells in the positive quadrant that along with a cell (i, j) contains also all the cells that are less than (i, j) in the natural partial order.

A standard Young tableau of size n is a Young diagram whose cells are filled with the numbers from 1 to n that increase along each row and each column. In other words, a Young tableau is a path in the lattice of Young diagrams that starts at the empty (zero-size) diagram and ends at a given diagram.

Young diagrams correspond to irreducible representations of S_n (see, e.g., [1, 2]), and the number of Young tableaux that fit into a given Young diagram Λ is equal to the dimension of the corresponding irreducible representation. For brevity, we call this number the dimension of the diagram Λ and denote it by $\dim(\Lambda)$.

This paper is organized as follows. In Sec. 2, we present the results of numerical experiments with the asymptotics of the typical (with respect to the Plancherel measure) dimension of an irreducible representation of the symmetric group. This measure was introduced in [3]; for further details, see [4].

Section 3 is devoted to similar computations for the maximum dimension of an irreducible representation of the symmetric group.

The results of Sec. 2 should be regarded as supporting the conjecture from [4] on the existence of the limit value of the normalized dimension of a typical diagram Λ with respect to the Plancherel measure. This hypothetical limit value is denoted by c, and in [4] it was called the "specific entropy of an irreducible representation."

At the same time, on the ground of our computations described in Sec. 3, nothing conclusive can be said about the asymptotic behavior of the maximum dimension. More exactly, we cannot claim that c_n stabilizes for accessible values of n.

In Sec. 4, we consider another distribution on Young diagrams: the Richardson distribution, which was studied in [5]. We give an experimental evidence for the proved theorem on the limit shape of typical Young diagrams. Then we present the results of similar experiments in three dimensions (in which the dimension of a diagram has nothing to do with S_n), and state a conjecture on the three-dimensional limit shape.

2. The asymptotic behavior of the typical dimension of an irreducible representation of S_n with respect to the Plancherel measure

Let $\widehat{S_n}$ be the set of equivalence classes of complex irreducible representations of S_n. For $\Lambda_n \in \widehat{S_n}$, we denote by $\dim \Lambda_n$ the dimension of the representation Λ_n, and by

$$\mu_n(\Lambda_n) = \frac{\dim^2 \Lambda_n}{n!}, \tag{1}$$

its Plancherel measure, see [3]. This is actually a probability measure on $\widehat{S_n}$, as follows from Burnside's formula

$$\sum_{\Lambda_n \in \widehat{S_n}} \dim^2 \Lambda_n = n!\,.$$

*St.Petersburg Department of the Steklov Mathematical Institute, St.Petersburg, Russia, e-mail: vershik@pdmi.ras.ru.

[1]St.Petersburg State Polytechnical University, St.Petersburg, Russia, e-mail: dmitry.pavlov@gmail.com.

Translated from *Zapiski Nauchnykh Seminarov POMI*, Vol. 373, 2009, pp. 77–93. Original article submitted September 9, 2009.

1072-3374/10/1683-0351

The set $\widehat{S_n}$ and the dimension $\dim \Lambda_n$ have the following interpretation (see [1]): $\widehat{S_n}$ is the ensemble of all Young diagrams of size n, and $\dim \Lambda_n$ is the number of standard Young tableaux that fit into the diagram Λ_n. In the present article, we make no difference between Λ_n and the corresponding irreducible representation.

The following normalization of the dimension of diagrams with n cells (see [4]) allows one to study the asymptotics of the dimension as $n \to \infty$:

$$c(\Lambda_n) = \frac{-2}{\sqrt{n}} \log \frac{\dim \Lambda_n}{\sqrt{n!}}. \tag{2}$$

We call $c(\Lambda_n)$ the *normalized dimension.* From [4], the following two-sided estimates are known for $c(\Lambda_n)$:

$$\lim_{n\to\infty} \mu_n\{\Lambda_n : c_0 < c(\Lambda_n) < c_1\} = 1$$

$$\left(c_0 = \frac{2}{\pi} - \frac{4}{\pi^2} \approx 0.2313,\ c_1 = \frac{2\pi}{\sqrt{6}} \approx 2.5651\right).$$

In other words, asymptotically almost all diagrams are of dimension lying in the following range:

$$\sqrt{n!}e^{-\frac{c_1}{2}\sqrt{n}} < \dim \Lambda_n < \sqrt{n!}e^{-\frac{c_0}{2}\sqrt{n}}.$$

Vershik and Kerov [4] put forward the conjecture that the limit $\lim\limits_{n\to\infty} c(\Lambda_n)$ exists almost everywhere (with respect to the Plancherel measure) on the set of infinite Young tableaux. An infinite Young tableau is an infinite sequence of nested Young diagrams $\Lambda_1 \subset \Lambda_2 \subset \Lambda_3 \subset \dots$ of sizes $1, 2, 3, \dots$.

We study the behavior of the coefficient $c(\Lambda_n)$ with respect to n. Some experiments in this area were undertaken in [6], where the expectation and variance of $c(\Lambda_n)$ are given for five values of n, with maximum of 1600 (for $n = 1600$, the sample under investigation contained only 14 diagrams). We performed the same computations on modern hardware, which is much more powerful than 25 years ago.

Note. After we had carried out the described experiments, we learned about a new unpublished paper by Alexander Bufetov, which contains a proof of the existence of the limit C of $c(\Lambda_n)$ in $L^2(Y)$ with respect to the Plancherel measure:

$$\lim_{n\to\infty} \int (c(\Lambda_n) - C)^2 d(\mu_n(\Lambda_n)) = 0.$$

The actual value of C is not likely to be obtainable by the methods used in Bufetov's proof. In should also be noted that the conjecture from [4] about the existence of the a.e. limit of $c(\Lambda_n)$ is still open.

In the next two sections of the paper, we describe well-known auxiliary routines: generating random diagrams with the Plancherel distribution by the RSK algorithm and counting the dimension of a diagram.

2.1. The hook-length formula

The hook-length formula (see, for example, [1, 7]) allows one to compute the dimension $\dim \Lambda$ of a Young diagram Λ without enumerating all Young tableaux that fit into Λ:

$$\dim \Lambda = \frac{n!}{\prod\limits_{(i,j)\in\Lambda} h_Y(i,j)}, \tag{3}$$

where (i, j) is a cell of Λ and $h_Y(i, j)$ is the length of the hook associated with this cell. The hook associated with a cell (i, j) consists of the cell itself and all cells that lie in the jth row to the right of (i, j) or in the ith column above (i, j) (see Fig. 1).

1				
5	3	2	1	
7	5	4	3	1

FIG. 1. The Hook lengths of a Young diagram.

2.2. RSK-random diagrams

Even having the hook-length formula, the straightforward generation of random diagrams distributed according to the Plancherel measure would be very computationally expensive. The Robinson–Schensted–Knuth (RSK) correspondence and the row insertion algorithm come to the rescue.

The RSK algorithm [1] takes as input an arbitrary permutation $s \in S_n$, performs a sequence of row insertions, and produces a pair of standard Young tableaux (P, Q) with the same diagrams; moreover, there is a one-to-one (RSK) correspondence between such pairs and permutations. Hence the uniform distribution (Haar measure) on S_n transforms into the Plancherel measure on the set of left (or right) Young tableaux. Applying the RSK correspondence to a random permutation in S_n and taking the Young diagram $Y(P)$ of the left Young tableau in the pair, we obtain a random Young tableau distributed according to the Plancherel measure.

2.3. Results

For $n \leq 120$, the expectations of $c(\Lambda_n)$ for the Plancherel measure were computed directly by the formula

$$c_n = \sum_{\Lambda_n \in \widehat{S_n}} c(\Lambda_n) \frac{\dim^2 \Lambda_n}{n!}.$$

The results are presented in the next section: see Table 3 and Fig. 5.

n	sample size	$\approx c_n$	$\approx \sigma_n$
1000	2000	1.6984314	0.10431497
2000	2000	1.746588	0.091339454
3000	2000	1.7644972	0.08351989
4000	2000	1.7750576	0.07747431
5000	2000	1.7873781	0.07282907
6000	2000	1.7917556	0.07022077
7000	2000	1.7969893	0.06630529
8000	2000	1.8000197	0.06586118
9000	2000	1.8070668	0.06243244
10000	10000	1.8102994	0.061589677
11000	10000	1.8118591	0.059796795
12000	10000	1.8147597	0.057941828
13000	10000	1.8162445	0.05743194
14000	10000	1.8187699	0.056453623
15000	20000	1.820125	0.05504108
16000	20000	1.8181555	0.054255717
17000	20000	1.8197316	0.053651392
18000	20000	1.8249108	0.052745327

TABLE 1. The expected values and standard deviation of $c(\Lambda_n)$.

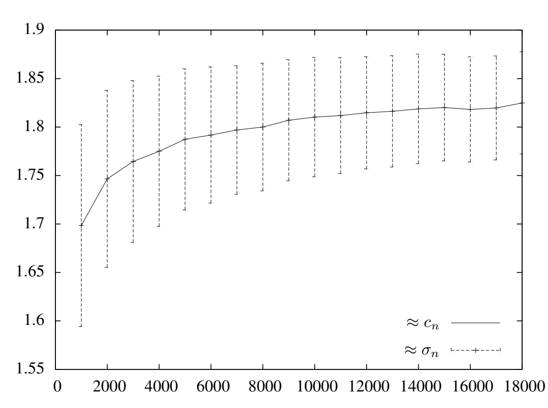

FIG. 2. The expected values of $c(\Lambda_n)$. For each n, the height of the vertical line is equal to σ_n.

There are 1,844,349,560 Young diagrams of size 120. For larger n, the expectations of $c(\Lambda_n)$ were calculated by the Monte-Carlo method, using a sample of RSK-random diagrams. The normalized dimension of each diagram was computed by formula (2), using the hook-length formula (3) for obtaining $\dim\Lambda$. The procedure was run for various n in the range from 1000 to 18000. The expectation $c_n = E(c(\Lambda_n))$ and the standard deviation $\sigma_n = \sigma(c(\Lambda_n))$ of $c(\Lambda_n)$ are listed in Table 1 and Fig. 2. From these results we see that the values of c_n asymptotically increase and presumably have a limit. To be quite honest, c_n do not increase monotonically in the selected range: for example, they decrease from $n = 15000$ to $n = 16000$. This fact was rechecked and confirmed with a sample of size 40000. The third decimal place remained constant after the size of the sample reached 20000.

2.4. The individual evolution of the dimension of a typical diagram

The Plancherel measure on infinite Young tableaux is defined as a Markov measure having the following property: the corresponding measure on Young diagrams of size n is the Plancherel measure (1) on diagrams. It is not difficult to find the transition probabilities for this Markov measure. Given that tableaux with equal diagrams have equal measures, we can say that the measure of a λ-shaped Young tableau is $\frac{\dim\lambda}{n!}$. Therefore, the probability of the transition from λ to Λ is

$$P(\Lambda|\lambda) = \frac{\dim(\Lambda)}{(n+1)\dim(\lambda)} \tag{4}$$

(see, for example, [8]). Thus the conjecture from [4] on the existence of the a.e. limit of $c(\Lambda_n)$ means that for almost all infinite Young tableaux $\{\Lambda_n, n = 1, 2, \dots\}$ generated by the described Markov process, $c(\Lambda_n)$ converges to some common limit value, which is obviously equal to the limit value for the expectation of the normalized dimension. Using formula (4) for the transition probability, we simulate the Markov process and obtain a sequence of Young diagrams of increasing size, each distributed according to the Plancherel measure and containing all the previous ones.

Our experiments show that the behavior of the normalized dimension of such a sequence is very chaotic, which probably indicates that the Vershik–Kerov conjecture on the existence of the a.e. limit of $c(\lambda_n)$ is not easy to prove nor disprove.

The nonregular behavior of the normalized dimension is illustrated on Fig. 3, which depicts the values of $c(\lambda_n)$ for two Markov sequences of Young diagrams. The values are computed for $n \in [100...7000]$ (only multiples of 100 were taken).

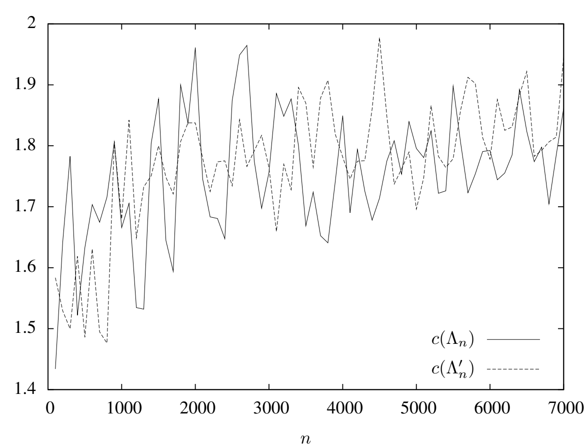

FIG. 3. The values of $c(\lambda_n)$ for two random sequences of Young diagrams distributed according to the Plancherel measure.

3. The asymptotic behavior of the maximum dimension of an irreducible representation of S_n

In this section, we study the behavior of the *maximum dimension* of a diagram of size n,

$$m_n = \max_{\Lambda_n \in \widehat{S_n}} \dim \Lambda_n,$$

and its normalized value

$$\overline{c_n} = c(\overline{\Lambda}_n),$$

where $\overline{\Lambda}_n$ is the diagram of size n that has the maximum dimension over all diagrams of size n.

The problem of computing the maximum dimension was posed in 1968 (see [9]). McKay [10] presented the values of $\max \dim \Lambda_n$ for n up to 75. Relying on these results, he conjectured that

$$\max \frac{\dim \lambda_n}{\sqrt{n!}} \leq \frac{1}{n}. \tag{5}$$

This was opposite to an alternative conjecture, according to which there are irreducible representations of arbitrary large dimension for which inequality (5) is not true. Just before his paper was published, McKay sadly admitted that the alternative conjecture is true for $n = 81$. Nevertheless, as shown in [4], McKay's conjecture is asymptotically true, and an even stronger fact holds: as $n \to \infty$, $\frac{\max \dim \Lambda_n}{\sqrt{n!}}$ decreases as $e^{-c\sqrt{n}}$, i.e., not only faster than $1/n$, but faster than any polynomial fraction. The estimates on the normalized dimension given in [4] for a typical Young diagram coincide with the estimates for the maximum dimension; while both have the same logarithmic order, experiments show that the constants are different.

For n up to 130, we find $\max \dim \Lambda_n$ by enumerating all Young diagrams of size n. In fact, we enumerate not diagrams, but *partitions* of n, using the trivial correspondence between Young diagrams and partitions of integers (see [1]). The dimension of each diagram is computed using the hook-length formula (3).

There are 5,371,315,400 Young diagrams of size $n = 130$. For larger n, since we have limited computational resources and cannot enumerate all Young diagrams, we considered only symmetric diagrams, *or diagrams that can be obtained from symmetric ones by adding one cell.* Often, this restriction does not affect the final result; but, for example, for $n = 14$ the diagram of maximum dimension does not satisfy it. However, this restriction is not substantial. Table 2 contains the values of $\overline{c_n}$. For $n = 310$, we enumerated 151,982,627 diagrams.

n	$\overline{c_n}$	n	$\approx \overline{c_n}$
10	0.57453286	140	1.05010306
20	0.8198125	150	1.0839802
30	0.7912792	160	1.05304872
40	0.86301332	170	1.0784368
50	0.90097636	180	1.0775954
60	0.94780416	190	1.0940416
70	0.98343194	200	1.0953336
80	0.96466594	210	1.1026434
90	0.9749938	220	1.11596048
100	1.035376	230	1.1106038
110	1.02168428	240	1.1273114
120	1.02246392	250	1.11251032
130	1.0514124	260	1.11878812
		270	1.1175388
		280	1.1173389
		290	1.13589692
		300	1.12641788
		310	1.148327

TABLE 2. The values of $c(\Lambda_n)$ for the diagrams of maximum dimension. The first column contains exact values, while the second one contains the values obtained by enumerating the restricted sets of diagrams.

While the two-sided estimates given by Vershik and Kerov [4] for the maximum dimension are the same as for the typical dimension, these two statistics have different behavior, and the limit of the sequence $\overline{c_n}$ is not likely to exist (see Fig. 4). We also emphasize that the maximum dimension is much greater than the typical one (and vice versa for the normalized values, because of the minus sign in the exponent).

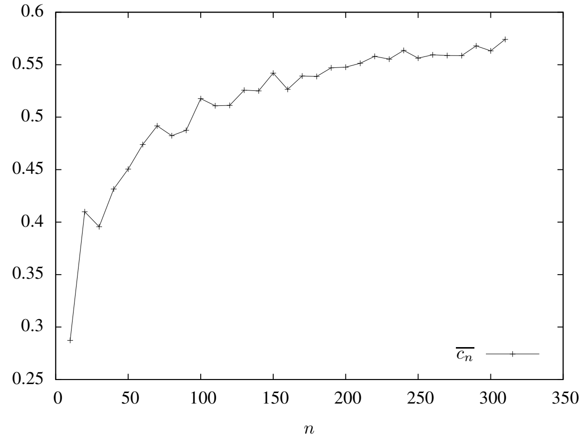

FIG. 4. The values of $c(\Lambda_n)$ for the diagrams of maximum dimension. The values starting from $n = 140$ are approximate, because of the above-mentioned restriction.

n	$\overline{c_n}$	c_n
10	0.57453287	0.9348365
20	0.81981254	1.1238908
30	0.7912792	1.2205664
40	0.8630133	1.283057
50	0.90097636	1.3281072
60	0.94780415	1.3622344
70	0.98343194	1.3878295
80	0.96466595	1.4042087
90	0.9749938	1.4061089
100	1.035376	1.3848866
110	1.0216843	1.3299882
120	1.0224639	1.2363929

TABLE 3. The normalized dimensions, maximum and typical.

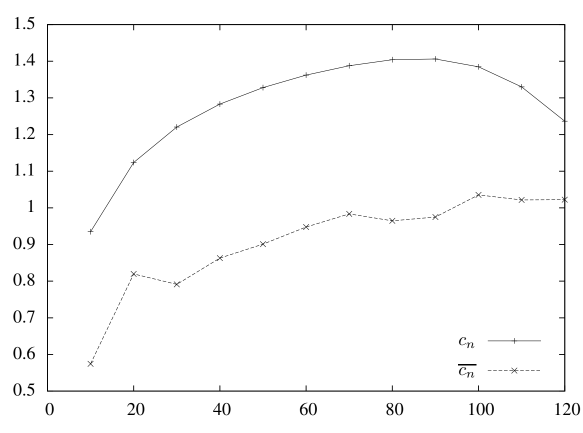

FIG. 5. The normalized dimensions, maximum and typical.

Table 3 and Fig. 5 present the results of comparing the exact values of $\overline{c_n}$ and c_n. Neither function is monotone.

4. Random diagrams with the Richardson distribution

Rost [5] considers a Markov process describing the behavior of a particle in $\{0,1\}^{\mathbb{Z}}$ that can be interpreted as the process of increasing a Young diagram cell by cell starting from the empty diagram. The transition to the next state (increasing the diagram by one cell) is performed in the following way.

Among all diagrams of size $n+1$ that contain the given diagram, one diagram is chosen randomly and uniformly. In other words, from all the k positions for a diagram of size n where a cell can be added, one position is picked with probability $1/k$. This growth process was introduced by Richardson [11]. Rost [5] found and proved the limit shape for Young diagrams distributed according to the Richardson measure (see below).

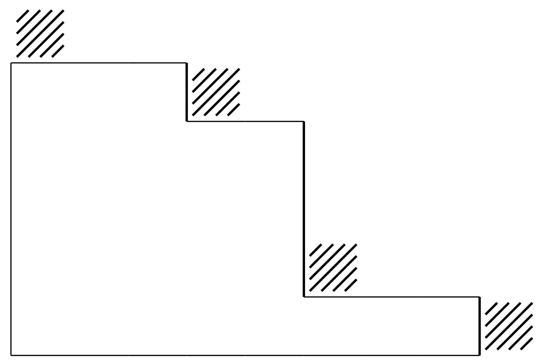

FIG. 6. Building a random Young diagram with the Richardson distribution. Each of the k dashed positions has the probability $1/k$.

We computed the values of the normalized dimension $c(\Lambda_n)$ for sequences of nested Young diagrams generated by the Richardson process (see Fig. 7). The difference between these values and the values obtained for the Plancherel measure (Fig. 3) makes it clear that these measures are totally different.

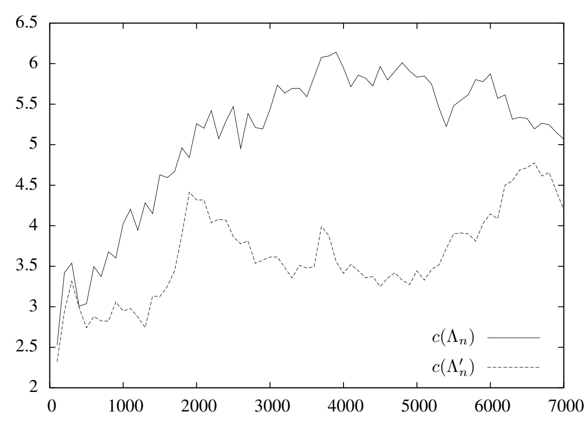

FIG. 7. The values of $c(\lambda_n)$ for two random sequences of Young diagrams distributed according to the Richardson measure.

Let us consider the process of infinite growth of a Young diagram, normalizing it by $\frac{1}{\sqrt{n}}$ along each axis at each step; thus the area of the diagram remains constant. As shown in [5], as $n \to \infty$ the process converges to a limit shape, which is given by the equation $\sqrt{x} + \sqrt{y} = h$. The exact value of h depends on the normalization. In [5], h is equal to one: $\sqrt{x} + \sqrt{y} = 1$. Another normalization that makes sense is the normalization by the area of the resulting figure, which is equal to $1/6$ of the area of the circumscribed square. The side of the square is equal to h^2, so

$$S = \int_0^{h^2} (h - \sqrt{x})^2 dx = h^4/6.$$

If we take the area S as 1, then the value of h is equal to $\sqrt[4]{6}$, and the equation for the limit shape takes the form

$$\sqrt{x} + \sqrt{y} = \sqrt[4]{6}. \tag{6}$$

4.1. d-Dimensional Young diagrams

A d-dimensional Young diagram is a finite descending ideal in the lattice $(\mathbb{Z}_{\leq 0})^d$. Unless specified otherwise, a "Young diagram" means a two-dimensional Young diagram.

Vershik and Kerov [3] introduced a convenient coordinate system for presenting Young diagrams: one should rotate by 45° the so-called French notation of a diagram, which corresponds to the Cartesian coordinates.

Similarly, d-dimensional Young diagrams can be represented as *functions* defined on the $(d-1)$-dimensional hyperplane passing through the origin and orthogonal to the main diagonal. The value of the function at a point of the hyperplane is the length of the interval parallel to the main diagonal starting at this point and ending at the border of the Young diagram.

Having this representation of Young diagrams, we can easily define the *average shape* of a collection of diagrams by averaging the corresponding functions. This definition trivially extends to multidimensional Young diagrams.

Although the function corresponding to a diagram is defined in the "rotated" coordinate system, we still depict Young diagrams and their average shapes in the Cartesian coordinates, by applying the inverse transformation.

The average shape of 2200 random diagrams of size $n = 100000$ is shown in Fig. 8. A visual verification of Rost's theorem can be obtained by plotting the average shape in the coordinates $(\sqrt{x}, \sqrt{y})$ (see Fig. 9).

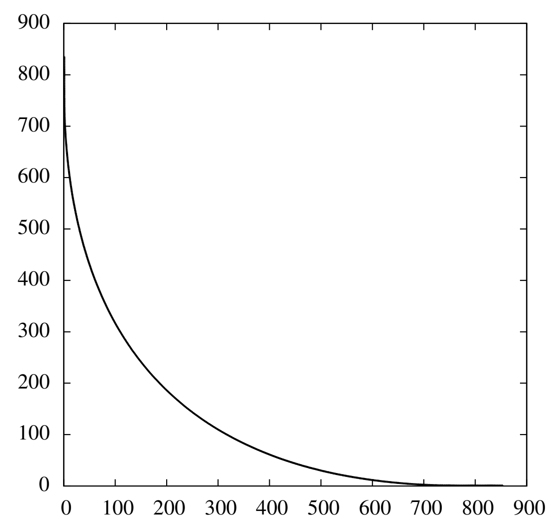

FIG. 8. The average shape of 2200 diagrams of size 100000.

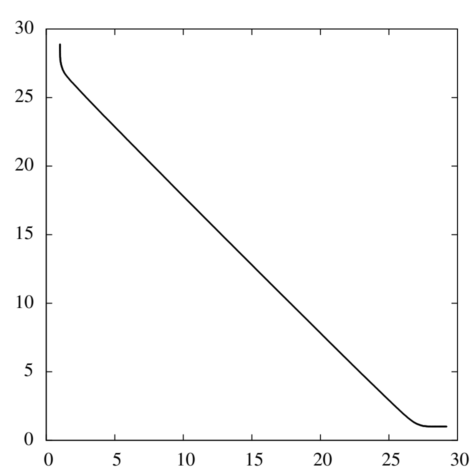

FIG. 9. The average shape of 2200 diagrams of size 100000 in the coordinates $(\sqrt{x}, \sqrt{y})$.

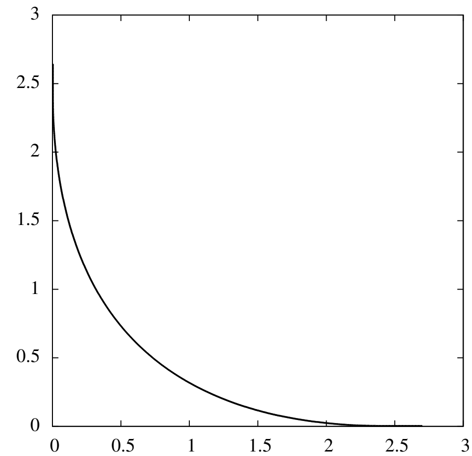

FIG. 10. The average shape of 2200 diagrams of size 100000 scaled by $1/\sqrt{n}$.

Scaling this shape by $1/\sqrt{n}$, we obtain a close approximation to the plot shown in Eq. (6). The area of the shape in Fig. 10 is equal to 1.

4.2. The standard deviation of the main diagonal segment

In the previous section, we showed that the average shape complies with Eq. (6). To verify that the average shape is actually the limit shape, we computed the standard deviation of the so-called *main diagonal segment*, i.e., the length of the interval of the main diagonal starting at the origin and ending at the average shape. In Table 4, the values of the standard deviation $d(n)$ are listed for n from 10000 to 40000, along with the normalized values $d(n)/\sqrt{n}$. Figure 11 shows the decrease of the normalized standard deviation as $n \to \infty$.

n	sample size	$\approx d(n)$	$\approx d(n)/\sqrt{n}$
10000	2000	1.8262177	0.018262176
15000	2000	1.9742892	0.016120004
20000	3000	2.0621564	0.014581648
25000	4000	2.1949573	0.013882129
30000	5000	2.203268	0.012720575
35000	6000	2.3289392	0.012448704
40000	7000	2.3589768	0.011794884

TABLE 4. The standard deviation of the main diagonal segment.

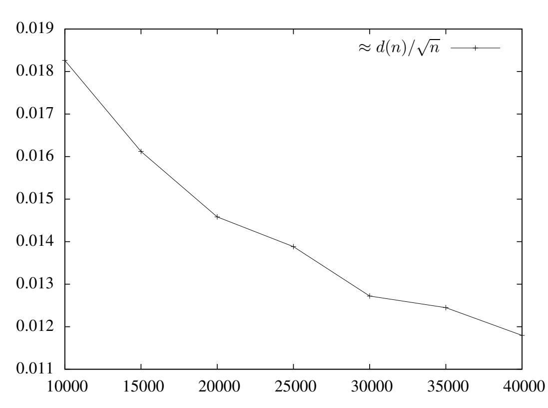

FIG. 11. The normalized standard deviation of the main diagonal segment.

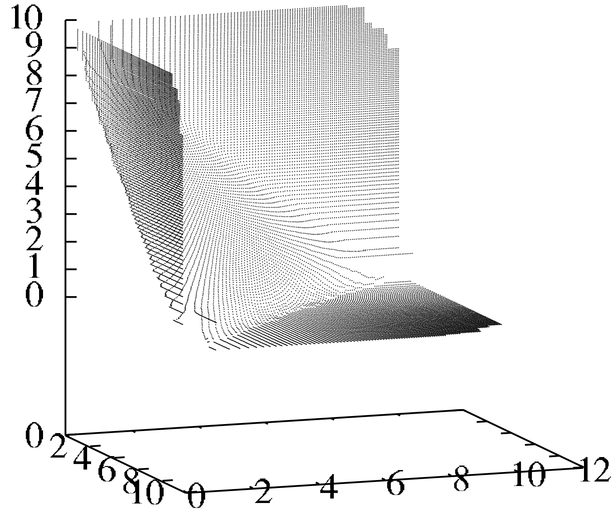

FIG. 12. The average shape of 400 diagrams of size 10000 drawn in the $(\sqrt{x}, \sqrt{y}, \sqrt{z})$ coordinates.

4.3. The average shape in three dimensions

The definition of a random Young diagram with the Richardson distribution can easily be generalized to the three-dimensional case. There are no known results about the limit shape in this case, but Fig. 12, obtained by our computations, leads to the assumption that the limit shape satisfies the analogous equation $\sqrt{x}+\sqrt{y}+\sqrt{z}=h_3$.

This result shows that the limit shape of three-dimensional Young diagrams generated by the Richardson process are probably different from the limit shape for uniformly distributed diagrams, which was studied in [12] and finally found in [13, 14].

The first author is supported by the grants NSh.-2460.2008.1 and RFBR 08-01.00379.

Translated by D. Pavlov.

REFERENCES

1. W. Fulton, *Young Tableaux, with Applications to Representation Theory and Geometry*, Cambridge Univ. Press, Cambridge (1997).
2. A. M. Vershik and A. Yu. Okounkov, "An inductive method of presenting the theory of representations of symmetric groups," *Russian Math. Surveys*, **51**, No. 2, 355–356 (1996).
3. A. M. Vershik and S. V. Kerov, "Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux," *Sov. Math. Dokl.*, **18**, 527–531 (1977).
4. A. M. Vershik and S. V. Kerov, "Asymptotics of maximal and typical dimensions of irreducible representations of a symmetric group," *Funct. Anal. Appl.*, **19**, 21–31 (1985).
5. H. Rost, "Nonequilibrium behaviour of a many particle process: density profile and local equilibria," *Z. Wahrsch. Verw. Gebiete*, **58**, No. 1, 41–53 (1981).
6. A. M. Vershik, A. B. Gribov, and S. V. Kerov, "Experiments in calculating the dimension of a typical representation of the symmetric group," *J. Sov. Math.*, **28**, 568–570 (1983).
7. D. E. Knuth, *The Art of Computer Programming*, Vol. 3, 2nd edition, Addison-Wesley, Reading, Massachusetts (1998).
8. A. M. Vershik and N. V. Tsilevich, "Induced representations of the infinite symmetric group," *Pure Appl. Math. Quart.*, **3**, No. 4, 1005–1026 (2007).
9. R. M. Baer and P. Brock, "Natural sorting over permutation spaces," *Math. Comp.*, **22**, 385–410 (1968).
10. J. McKay, "The largest degrees of irreducible characters of the symmetric group," *Math. Comp.*, **30**, No. 135, 624–631 (1976).
11. D. Richardson, "Random growth in a tessellation," *Proc. Cambridge Philos. Soc.*, **74**, 515–528 (1973).
12. A. Vershik and Yu. Yakubovich, "Fluctuation of maximal particle energy of quantum ideal gas and random partitions," *Comm. Math. Phys.*, **261**, No. 3, 759–769 (2006).

13. R. Cerf and R. Kenyon, "The low-temperature expansion of the Wulff crystal in the 3D Ising model," *Comm. Math. Phys.*, **222**, No. 1, 147–179 (2001).
14. A. Okounkov and N. Reshetikhin, "Correlation function of Schur process with application to local geometry of a random three-dimensional Young diagram," *J. Amer. Math. Soc.*, **16**, 581–603 (2003).