æ

On the eigenvalue distribution of some symmetric random matrices

Ferenc Oravecz and **Dénes Petz***

Department for Mathematical Analysis
Technical University of Budapest
H-1521 Budapest XI. Sztoczek u. 2, Hungary

Abstract. Marchenko and Pastur studied the limit of the eigenvalue distribution of the sum of $p(n)$ rank one random projections in the n dimensional space when $p(n)/n \to a$ and $n \to \infty$ in 1967. More recently this Marchenko-Pastur distribution occured in the free analogue of the Poisson limit theorem. In this paper we derive a recursive as well as an explicite formula for the moments of the Marchenko-Pastur distribution which turn out to be polynomials of a. Moreover, an elementary combinatorial proof is given to the known fact that a variant of the Marchenko-Pastur distribution describes the asymptotical eigenvalue density of sample covariance matrices.

1. Preliminaries and the Marchenko-Pastur distribution. For a positive number a the corresponding Marchenko-Pastur distribution is

$$\mu_a = \begin{cases} (1-a)\delta_0 + \dfrac{\sqrt{4a-(x-1-a)^2}}{2\pi x}\chi & \text{if } a < 1 \\ \dfrac{\sqrt{4a-(x-1-a)^2}}{2\pi x}\chi & \text{if } a \geq 1, \end{cases} \tag{1}$$

where χ stands for the characteristic function of the interval $[(1-\sqrt{a})^2, (1+\sqrt{a})^2]$. This probability distribution appeared in [MP] as the limiting mean spectral density of the random matrices

$$B_n(p) = \sum_{i=1}^{p} P_i,$$

where P_i's are independent rank one projections with unitarily invariant distribution and $p/n \to a$. It is clear that for $a < 1$ zero is in the spectrum of $B_n(p)$ with multiplicity $n-p$ and this yields the atomic part of the limiting distribution.

* Supported by OTKA T016924.

The moments of μ_a are the integrals

$$m_n = \int x^n \, d\mu_a(x) = \frac{1}{2\pi} \int_A^B x^{n-1} \sqrt{4a - (x-1-a)^2} \, dx$$

where $A = (1-\sqrt{a})^2$ and $B = (1+\sqrt{a})^2$. We are unable to compute this integral directly but it is not difficult to find a recursive formula by means of repeated integration by parts. We found

$$m_n = \frac{2n-1}{n+1}(1+a)m_{n-1} - \frac{n-2}{n+1}(1-a)^2 m_{n-2} \tag{2}$$

for $n \geq 2$. It is straightforward that $m_0 = a$ if $0 < a \leq 1$, $m_0 = 1$ if $1 < a$ and $m_1 = a$.

More recently the Marchenko-Pastur distribution was observed in the free analogue of the Poisson limit theorem. Voiculescu introduced a new convolution of probability measures on $\mathbb{R}$, the so-called (additive) free convolution, which is denoted by $\boxplus$. It was proved that

$$\mu_n = \left(\left(1 - \frac{a}{n}\right) \delta_0 + \frac{a}{n} \delta_1 \right)^{\boxplus n} \tag{3}$$

tends to the Marchenko-Pastur distribution as $n \to \infty$. In the course of the proof one observes that the so-called Voiculescu transform of μ_a is

$$\frac{a}{1-z} = \sum_{n=0}^{\infty} a z^n$$

and therefore

$$m_n = \int x^n \, d\mu_a(x) = \left\langle \left(l^* + \sum_{k=0}^{\infty} a l^k \right)^n \Phi, \Phi \right\rangle, \tag{4}$$

holds for the moments of the Marchenko-Pastur distribution, where l and l^* are the creation and annihilation operators and Φ is the vacuum vector in the full Fock space over $\mathbb{C}$. So l is an isometry, $l^* l = I$ and one needs also the fact $\langle l^k \Phi, \Phi \rangle = \langle (l^*)^k \Phi, \Phi \rangle = 0$ $(k = 1, 2, \ldots)$ to compute (4) for a concrete k. It follows from equation (4) that a is the mean of μ_a.

Computing m_n, we have to deal with terms of the form $\langle L_n L_{n-1} \ldots L_2 L_1 \Phi, \Phi \rangle$, where L_i $(1 \leq i \leq n)$ equals either l^* or $a l^k$ for some integer $k \geq 0$. If in such a term the total number of l^*'s is not the same as the total power of l, then that term vanishes. The term vanishes again if there is an index i_0 such that in the subproduct $L_{i_0} L_{i_0 - 1} \ldots L_2 L_1$ the total number of l^* is more than the total power of l. In every other case the term does not vanish and it gives a power of a. Collecting the terms according the power of a in them, we can write (4) in the following form:

$$m_n = \sum_{i=1}^{n} Q_n^{(i)} a^i, \tag{5}$$

where $Q_n^{(i)}$ is the number of those non-vanishing terms in which the number of factors is n and the total power of a is l. It is clear that here l cannot be less than one, as in that case there would be only annihilation operators in the term. On the other hand, l cannot be bigger than n when we take terms of n factors. It is also easy to check that $Q_n^{(1)} = Q_n^{(n)} = 1$ for every $n \geq 1$.

Below a variant of the Marchenko-Pastur distribution will play a role. Set

$$\bar{\mu}_a = \begin{cases} (1-a^{-1})\delta_0 + \dfrac{\sqrt{4a-(x-1-a)^2}}{2\pi a x}\chi & \text{if } 1 < a \\ \dfrac{\sqrt{4a-(x-1-a)^2}}{2\pi a x}\chi & \text{if } 0 < a \ \leq 1, \end{cases} \tag{5}$$

where, as above, χ denotes for the characteristic function of the interval $[(1-\sqrt{a})^2, (1+\sqrt{a})^2]$. The moments of this variant are rather similar to those of μ_a. If m_n ($\bar{m}_n$) is the nth moment of μ_a ($\bar{\mu}_n$), then we have

$$m_n = a\bar{m}_n \qquad (n = 2, 3, \ldots) \tag{6}$$

while $\bar{m}_1 = 1$.

2. A limit theorem. The limit theorem we treat concerns the matrix TT^t, where t stands for the transpose and T is a random matrix with identically distributed independent standardized entries. The Wishart matrix is a particular, but important example, see [R]. A stronger form of result we prove here was called well-known in [Ba] and credit was given to [Wa]. Our main concern here is the presentation of an elementary combinatorial proof based on the method of moments. This technique is known, for example, from the proof of the central limit theorem and was adapted to random matrices in the classic paper [Wi], see also [F-K].

Theorem 2.1. Assume that the entries of the $p(n) \times n$ random matrix $T^{(n)} = (\xi_{ij}^{(n)})$ are independent and identically distributed with mean zero and variance 1 and all moments are bounded. Moreover, assume that $\lim_{n\to\infty} p(n)/n = a$. Then the mean spectral density of the symmetric $p(n) \times p(n)$ random matrix $\Lambda^{(n)} = T^{(n)}(T^{(n)})^t/n$ tends to the distribution $\bar{\mu}_a$ given in (5).

We prove convergence in the sense of moments. Set for an $m \times m$ random matrix X

$$\tau_m(X) = \frac{1}{m}\sum_i E(X_{ii}).$$

Then the statement of the theorem is the limit relation

$$\tau_{p(n)}((\Lambda^{(n)})^k) \to \int x^k \, d\bar{\mu}_a(x).$$

The proof of this result will be put in a series of combinatorial lemmas. First we prove that for a fixed k the limit exists and it is a polynomial of a. Then we give different combinatorial characterizations of the coefficients of that polynomial and finally the limit will be identified with moment of the Marchenko-Pastur distribution.

Let us say that the chain of $2k$ ordered pairs

$$(m_1, n_1)\,(m_2, n_1)\,(m_2, n_2)\,(m_3, n_2)\ldots(m_k, n_k)\,(m_1, n_k) \tag{2.1}$$

of integers corresponds to a certain non-crossing pair partition $\mathcal{P}$ of the set $\{1, 2, \ldots, 2k\}$ if and only if

(a) $(m_i, n_i) = (m_j, n_{j-1})$ if $\{2i-1, 2j-2\} \in \mathcal{P}$,

(b) $(m_i, n_i) = (m_1, n_k)$ if $\{2i-1, 2k\} \in \mathcal{P}$,

(c) all the other m_i's of different indices and, on the other hand all the other n_i-s of different indices, for which rules (a) and (b) do not say anything, are different.

To explain the above correspondence more clearly, we arrange each of the $2k$ pairs vertically:

$$\binom{m_1}{n_1}\binom{m_2}{n_1}\binom{m_2}{n_2}\binom{m_3}{n_2}\cdots\binom{m_k}{n_k}\binom{m_1}{n_k}$$

The rule above says that $(i, j) \in \mathcal{P}$ if and only if the ith vertical pair is the same as the jth vertical pair. Not every chain of ordered pairs corresponds to a pair partition and several chains may correspond to the same pair partition. For example, the chains

$$\binom{1}{2}\binom{4}{2}\binom{4}{2}\binom{3}{2}\binom{3}{2}\binom{1}{2}$$

and

$$\binom{2}{2}\binom{4}{2}\binom{4}{2}\binom{3}{2}\binom{3}{2}\binom{2}{2}$$

correspond to the pair partition $\{\{1, 6\}, \{2, 3\}, \{4, 5\}\}$.

Lemma 2.2. For the above defined $\Lambda^{(n)}$ we have

$$\lim_{n\to\infty} \tau((\Lambda^{(n)})^k) = \sum_{i=1}^{k} S_k^{(i)} a^{i-1},$$

where $S_k^{(i)}$ means the number of such non-crossing pair partitions of the set $\{1, 2, \ldots, 2k\}$ for which in the corresponding chain of pairs the cardinality of the set $\{m_1, m_2, \ldots, m_k\}$ is i.

Proof. For the sake of simpler writing, we put ξ_{ij} in place of $\xi_{ij}^{(n)}$. Calculating the diagonals of the kth power of $\Lambda^{(n)}$, we have

$$\begin{aligned}
\left((\Lambda^{(n)})^k\right)_{ii} &= \sum_{r_1,r_2,\ldots r_{k-1}=1}^{p(n)} \Lambda_{ir_1}\Lambda_{r_1r_2}\Lambda_{r_2r_3}\ldots\Lambda_{r_{k-1}i} \\
&= \sum_{r_1,r_2,\ldots r_{k-1}=1}^{p(n)} \left(\frac{1}{n}\sum_{s_1=1}^{n}\xi_{is_1}\xi_{r_1s_1}\right)\left(\frac{1}{n}\sum_{s_2=1}^{n}\xi_{r_1s_2}\xi_{r_2s_2}\right)\cdots\left(\frac{1}{n}\sum_{s_1=1}^{n}\xi_{r_{k-1}s_k}\xi_{is_k}\right) \\
&= \frac{1}{n^k}\sum_{r_1,r_2,\ldots r_{k-1}=1}^{p(n)}\sum_{s_1,s_2,\ldots s_k=1}^{n}\xi_{is_1}\xi_{r_1s_1}\xi_{r_1s_2}\xi_{r_2s_2}\ldots\xi_{r_{k-1}s_k}\xi_{is_k},
\end{aligned}$$

and for the moment of the mean spectral density we obtain

$$\begin{aligned}
\tau(\Lambda^k) &= \frac{1}{p(n)}\sum_{i=1}^{p(n)}E((\Lambda^k)_{ii}) \\
&= \frac{1}{p(n)n^k}\sum_{m_1,m_2,\ldots m_k=1}^{p(n)}\sum_{n_1,n_2,\ldots n_k=1}^{n}E(\xi_{m_1n_1}\xi_{m_2n_1}\xi_{m_2n_2}\xi_{m_3n_2}\ldots\xi_{m_kn_k}\xi_{m_1n_k}).
\end{aligned}$$

Let us introduce two numbers, l_1 and l_2 as follows:

$$l_1 = \#\{m_1,m_2,\ldots,m_k\} \quad \text{and} \quad l_2 = \#\{n_1,n_2\ldots,n_k\} \qquad (2.2).$$

It is easy to see that the number of chains (2.1) satisfying condition (2.2) is at most

$$\binom{p(n)}{l_1} l_1^k \binom{n}{l_2} l_2^k.$$

If we group the terms $E(\xi_{m_1n_1}\xi_{m_2n_1}\xi_{m_2n_2}\xi_{m_3n_2}\ldots\xi_{m_kn_k}\xi_{m_1n_k})$ according to the cardinalities of the sets $\{m_1,m_2,\ldots,m_k\}$ and $\{n_1,n_2\ldots,n_k\}$, we get

$$\tau(\Lambda^k) = \frac{1}{p(n)n^k}\sum_{l_1=1}^{k}\sum_{l_2=1}^{k}\sum_{(l_1,l_2)}E(\xi_{m_1n_1}\xi_{m_2n_1}\xi_{m_2n_2}\xi_{m_3n_2}\ldots\xi_{m_kn_k}\xi_{m_1n_k}), \qquad (2.3)$$

where the last summation is for those terms, where l_1 and l_2 have certain fixed values.

Now we will show that those parts of the summation in (2.3) where $l = l_1 + l_2 < k+1$ vanish.

Due to the Hölder inequality

$$\begin{aligned}
&\left|E(\xi^{(n)}_{m_1n_1}\xi^{(n)}_{m_2n_1}\xi^{(n)}_{m_2n_2}\xi^{(n)}_{m_3n_2}\cdots\xi^{(n)}_{m_kn_k}\xi^{(n)}_{m_1n_k})\right| \\
&\quad\le E(\left|\xi^{(n)}_{m_1n_1}\right|^k)^{1/k}\ldots E(\left|\xi^{(n)}_{m_1n_k}\right|^k)^{1/k} \le C_k(n)
\end{aligned}$$

(The existence of such $C_k(n)$-s is trivial since the ξ_{ij} matrix elements are identically distributed and all the moments are bounded.) This estimate implies that

$$\begin{aligned}
&\frac{1}{p(n)n^k}\left|\sum_{(l_1,l_2)} E(\xi_{m_1n_1}\xi_{m_2n_1}\xi_{m_2n_2}\xi_{m_3n_2}\dots\xi_{m_kn_k}\xi_{m_1n_k})\right| \\
&\le \frac{1}{p(n)n^k}\sum(l_1,l_2)\left|\sum_{(l_1,l_2)} E(\xi_{m_1n_1}\xi_{m_2n_1}\xi_{m_2n_2}\xi_{m_3n_2}\dots\xi_{m_kn_k}\xi_{m_1n_k})\right| \\
&\le \frac{1}{p(n)n^k}\sum_{(l_1,l_2)} C_k(n) \le \frac{1}{p(n)n^k}\binom{p(n)}{l_1} l_1^k \binom{n}{l_2} l_2^k C_k(n) \\
&= \frac{p(p-1)\dots(p-l_1+1)n(n-1)\dots(n-l_2+1)}{pn^k}\frac{l_1^k l_2^k}{l_1!l_2!}C_k(n)
\end{aligned}$$

and the latest tends to zero if $n\to\infty$ whenever $l = l_1 + l_2 < k+1$.

Next we show that if $l = l_1 + l_2 > k+1$ then

$$E(\xi_{m_1n_1}\xi_{m_2n_1}\xi_{m_2n_2}\xi_{m_3n_2}\dots\xi_{m_kn_k}\xi_{m_1n_k}) = 0. \tag{2.4}$$

We can prove this by induction. The case $k = 1$ is trivial as in this case l cannot be bigger then 2. When $k = 2$ the only possibility for $l > k+1$ is the sequence $\xi_{m_1n_1}\xi_{m_2n_1}\xi_{m_2n_2}\xi_{m_1n_2}$, where and $n_1 \ne n_2$. Then all ξ_{ij} are without repetition, and because of their zero mean and independence (2.4) holds true. This is the first case where the reason why such terms vanish can be seen.

Now let us suppose that for k and for $l = k+2$ there is at least one ξ_{ij} in the corresponding sequence without repetition. This would assure that (2.4) is true. Then the statement is true for this fixed k and for arbitrary $l > k+1$. Namely if for an $l = k+2+j$ every ξ_{ij} appears with repetition, than changing j "new" indeces for "old" ones, this would be the case for $l = k+2$, too, in contrary with our assumption.

Keeping this in mind, it is enough to prove the induction step for $k+1$ and for $l = k+3$. Let us suppose that in a sequence satisfying this there is no ξ_{ij} without repetition. This means that there is at least one m_p or n_p index which appears only in two neighboring ξ_{ij}, otherwise each index would appear in at least four ξ_{ij} according to the structure of the sequence, but $4(k+3)$ indexes cannot fit in $4k+4$ sites!

As this m_p or n_p index appears only in two ξ_{ij} and according to our assumption there is no ξ_{ij} without repetition, this two ξ_{ij} form a pair:

$$\xi_{m_{p-1}n_{p-1}}\xi_{m_pn_{p-1}}\xi_{m_pn_p}\xi_{m_{p+1}n_p}\xi_{m_{p+1}n_{p+1}}$$

If m_p is the index appearing only in two ξ_{ij} than, as in this case the second and the third ξ_{ij} form a pair, $n_{p-1} = n_p$, and if n_p is the index appearing only in two ξ_{ij} than, as in this case the third and the fourth ξ_{ij} form a pair, $m_p = m_{p+1}$. Whichever case we have, eliminating the corresponding pair from the sequence, we get a shorter sequence which corresponds to k, $l = k + 2$, and where there is still no ξ_{ij} without repetition, as the eliminated pair could not be the member of a triplet as m_p (or n_p) appeared only in two ξ_{ij}. This would be in contradiction with our assumption.

This means that we have proven the following: only those terms have an affect on the summation in (2.2), where $l = k + 1$.

Next we shall show that for every k and $l = k + 1$, if each ξ_{ij} appears at least twice than each of them appears exactly twice. We prove it by induction, in a similar way as above. The case $k = 1$ is trivial: the only possible sequence is $\xi_{m_1 n_1} \xi_{m_1 n_1}$. Let us suppose that the statement is true for k. Than for any sequence corresponding to $k + 1$, $l = k + 2$ and satisfying the condition that each ξ_{ij} appears at least twice in it, just as above, there is at least one m_p or n_p index which appears only in two neighboring ξ_{ij}. These form a pair again, from which $n_{p-1} = n_p$ (or $m_p = m_{p+1}$) follows. Eliminating this pair we get a shorter sequence which corresponds to k, $l = k + 2$, and where each ξ_{ij} appears exactly twice according to our assumption, since there still cannot be "lonely" ξ_{ij}. But as the eliminated ξ_{ij}'s cannot be identical to any part of this shorter sequence since m_p (or n_p) appeared only in this two ξ_{ij}, in the sequence corresponding to $k+1$, $l = k+2$ each ξ_{ij} appears also exactly twice.

Next we show that for each k, $l = k + 1$, if each ξ_{ij} appears at least twice, then there is a one to one correspondence between these sequences and the non-crossing pair partitions of the set $\{1, 2, \ldots 2k\}$. To prove this, we use the explicitly given correspondence described before the lemma was stated.

What we have to show is that for every "correct" sequence $\mathcal{P}$ is a non-crossing pair partition. (The rest of the statement is clear.) For this we use again mathematical induction in k.

The case $k = 1$ is clear:

$$\xi_{m_1 n_1} \xi_{m_1 n_1} \longrightarrow (m_1, n_1) = (m_1, n_k) \Leftrightarrow \{1, 2\} \in \nu.$$

Let us suppose that the statement is true for k. For $k + 1$, $l = k + 2$ in every "correct" sequence there is again at least one m_p or n_p index which appears only in two neighboring ξ_{ij}. These form a pair again, from which $n_{p-1} = n_p$ (or $m_p = m_{p+1}$) follows. Eliminating this pair we get a shorter sequence which corresponds to k, $l = k+2$. According to our assumption there is a non-crossing pair partition ν of the set $\{1, 2, \ldots, 2k\}$, corresponding to this shorter sequence. Adding $\{2p-1, 2p-2\}$ (or $\{2p-1, 2p\}$) to ν and renumbering the set to fit these new elements in the right place, we clearly get a non-crossing pair partition of the set $\{1, 2, \ldots, 2k+2\}$ corresponding to our longer sequence.

So what we have now is

$$\lim_{k\to\infty} \tau((A^{(n)})^k) = \lim_{k\to\infty} \frac{1}{p(n)n^k} \sum_{(a),(b)} E(\xi_{m_1 n_1} \xi_{m_2 n_1} \xi_{m_2 n_2} \xi_{m_3 n_2} \ldots \xi_{m_k n_k} \xi_{m_1 n_k}),$$

where the summation is over all $1 \leq m_1, m_2, \ldots, m_k \leq p(n)$ and $1 \leq n_1, n_2, \ldots, n_k \leq n$ satisfying the following two conditions:

(a) $l = l_1 + l_2 = k + 1$ $(l_1, l_2 \geq 1)$,

(b) every element of the sequence $\{m_1, n_1\}, \{m_2, n_1\} \ldots \{m_1, n_k\}$ appears at least twice.

As the variance of each (independent) $\xi_{ij}^{(n)}$ is 1 and according to the above proven fact that in every sequence each ξ_{ij} appears exactly twice we get

$$\begin{aligned}
\lim_{k\to\infty} \tau(\Lambda^k) &= \lim_{k\to\infty} \frac{1}{p(n)n^k} \sum_{(a),(b)} 1 \\
&= \lim_{k\to\infty} \frac{1}{p(n)n^k} \sum_{i=1}^{k} S_k^{(i)} p(p-1)\ldots(p-i+1)n(n-1)\ldots(n-k+i),
\end{aligned}$$

where $S_k^{(i)}$ means the number of such non-crossing pair partitions of the set $\{1, 2, \ldots, 2k\}$ for which in the corresponding sequence (that is to say in the corresponding chain of ordered pairs) $l_1 = i$ and $l_2 = k + 1 - i$. If $n \to \infty$ and $p(n)/n \to a$, so for the moments of the limit distribution we get

$$\lim_{n\to\infty} \tau(\Lambda(n)^k) = \sum_{i=1}^{k} S_k^{(i)} a^{i-1}$$

which was stated. □

We note that for $p(n) = n$ (which means $a = 1$), we have

$$m_k = \sum_{i=1}^{k} S_k^{(i)} = \frac{1}{k+1} \binom{2k}{k},$$

and in this case we arrive at the number of all non-crossing pair partitions, which is, on the other hand is the $2k$th moment of the standard Wigner-law. Due to relation (4) this means that in this special case the limit distribution is the Marchenko-Pastur distribution. To see this for $a \neq 1$ is unfortunately more difficult.

Lemma 2.3. $S_k^{(i)}$ equals the number of syntactically correct bracket sequences in which there are k pairs of brackets and the number of opening brackets on even sites is $i - 1$.

Proof. First we note that there is a one to one correspondence between the non-crossing pair partitions of the set $\{1, 2, \ldots 2k\}$ and the syntactically correct sequences of brackets., where, if $\{i, j\} \in \nu$, $(i < j)$ than on the ith site there is an opening bracket with its closing pair being on the jth site. Together with the proof of Lemma 2.2 this means that there is a one to one correspondence between the "correct" sequences of ξ_{ij}'s and the syntactically correct sequences of brackets.

For $S_1^{(1)}$ the statement of the lemma is true. It is also true for $S_k^{(1)}$ for every $k \geq 1$. This can be seen as follows: To each pair partition corresponding to $S_k^{(1)}$ corresponds a unique sequence of ξ_{ij}-s, for which $l_1 = 1$. This means that every m_i indices are the same, let us say m. But then each n_i have to be different, otherwise the condition $l_1 + l_2 = k + 1$ could not be satisfied. So the sequence is:

$$\xi_{mn_1}\xi_{mn_1}\xi_{mn_2}\ldots\xi_{mn_k}.$$

From here it can be seen that the only possibility for ν is $\{\{1,2\}\{3,4\}\ldots\{2k-1,12k\}\}$, which corresponds to the sequence of brackets: $()\,()\ldots()$. It is easy to see that this is the only possible bracket-sequence, where there is no opening bracket on even site. This justifies the statement and also shows that for every $k \leq 1$ $S_k^{(1)} = 1$.

Let us suppose that for a fixed k the statement of the lemma is true for every $\tilde{k}, i$ satisfying that $1 \leq \tilde{k} \leq k$, $1 \leq i \leq \tilde{k}$. (For $k = 1$ this is also true.) For $k+1$ let us select the shortest pair of brackets for which the opening bracket is on even site (if there are several among them of the same size, let us chose the earliest one for example). If there is no such a pair, it means that we have the case $S_k^{(1)}$ which is already clear. If there is one, we have two possibilities.

In the first case the two brackets of the pair are beside each other. Then eliminating them, we get a sequence corresponding to $S_k^{(i-1)}$,that is to say corresponding to a certain sequence of ξ_{ij}-s of the length of $2k$ for which $l_1 = i-1$ according to our assumption. Now let us replace the eliminated pair of brackets on their original site. We can see now that the original sequence of brackets corresponds to such a sequence of ξ_{ij}-s where $l_1 = i$, because the replacing of these brackets according to the proof of lemma 2.1 is equivalent to placing a $\xi_{m_i n_{i-1}}\xi_{m_i n_i}$ pair in the above mentioned sequence of ξ_{ij}-s, where, as the two new ξ_{ij}-s form a pair, obviously $n_{i-1} = n_i$, which, as this pair joints the rest of the sequence by its n_i indices according to the property of sequences of ξ_{ij}-s, means that m_i have to be a new index which did not appear in the chain before (that is to say l_1 is increasing by one in this procedure), otherwise $\xi_{m_i n_i}$ would appear at least three times in the sequence, which is impossible for any "correct" sequence.

In the second case the two brackets of the pair are not beside each other. Now eliminating the part of the sequence being between this two brackets, we get a sequence corresponding to $S_{k+1-j}^{(i)}$,that is to say corresponding to a certain sequence of ξ_{ij}-s of the length of $2\,(k+1-j)$ for which $l_1 = i$. This is true as in the eliminated part of bracket sequence there could not be any opening bracket in even site, otherwise our chosen shortest pair would not be the shortest (recall that the sequence of brackets have to be syntactically correct!) and according to our assumption, where j is a positive integer. (Here we have to show that i cannot be bigger than $k+1-j$. If it was we would get a contradiction in the following way: In the original sequence of brackets the number of even sites is $k+1$. The number of opening brackets in these sites is i. This means that there are $k+1-i$ closing brackets in them. But if j is grater then $k+1-i$, than there is not enough closing brackets for the j sites between our two chosen brackets, which means that there must be at least one opening bracket there, but in this case our selected pair cannot be the shortest!)

On the other hand, the eliminated part corresponds to such a sequence of ξ_{ij}-s of the length of $2j$ for which $l_1 = 1$, or as we have seen it above, where every m_i index is the same. Now let us replace the eliminated part of the bracket-sequence on its original site. We can see now that the original sequence of brackets corresponds to such a sequence of ξ_{ij}-s where $l_1 = i$, because the replacing of these brackets according to the proof of lemma 2.1 is equivalent to placing the above sequence of ξ_{ij}-s corresponding to the replaced part of bracket-sequence into a certain point (which is behind an even site and before the following odd site) of the sequence of ξ_{ij}-s corresponding to the rest of the brackets, and as this new part joints the rest of the sequence by its (identical) m_i indices according to the property of sequences of ξ_{ij}-s, which means that in this new sequence $l_1 = i$ still holds.

As we can do this there-and-back procedure to every bracket sequences corresponding to $k+1$, we have proven that the number of syntactically correct bracket-sequences in which there are k pairs of brackets and the number of opening brackets on even sites is $i-1$ is less or equal than S_i^k. What we still have to show is that there in no such sequence of ξ_{ij}-s of the length of $2k+2$ and where $l_1 = i$, for which in the corresponding sequence of brackets the number of opening brackets on even sites is $\tilde{i} \neq i-1$. Let us suppose that the opposite is true. For this sequence of brackets applying our procedure, first we get a bracket sequence that, according our assumption, corresponds to a sequence of ξ_{ij}-s of the length of $2k$ and where $l_1 = \tilde{i}-1$, in the first case, or $l_1 = \tilde{i}$, in the second case. Now applying the second step of the above procedure we get a sequence of ξ_{ij}-s of the length of $2k+2$ where $l_1 = \tilde{i}$ (in both cases). As applying the whole procedure we always get back the starting sequences, and as $\tilde{i} \neq i$ this is a contradiction. This means that the statement of the lemma is true for $k+1$, too, hence it is true in general. □

Lemma 2.4. $S_{k+1-i}^k = Q_i^k$ for ever $k \geq 1$ and $1 \geq i \geq k$.

Proof. To prove the statement of the lemma we will give a one to one correspondence between the "good" terms (recall what it means from \$1) corresponding to Q_i^k and the bracket-sequences corresponding to S_{k+1-i}^k. First we give this correspondence, then we will show that it is well defined.

We note that in such a "good" term there are k factors, and in such a bracket sequence there are k pairs of brackets. From now we will use the notation of \$1 again.

Let us define the correspondence as follows:

1. For some i L_i is aI in a term if and only if in the corresponding bracket-sequence in the sites $2i-1$ and $2i$ there is a pair of brackets: () (in this order).

2. For some i L_i is l^* in a term if and only if in the corresponding bracket-sequence in the site $2i-1$ there a closing bracket. In this case we say that in the site $2i$ there is a "space" (which can be either an opening or a closing bracket, being interesting only later on). So if there is a closing bracket in any odd site, the next site is always considered a "space".

3. For some i L_i is al^n $(1 \leq n \leq k-1)$ in a term corresponding to Q_j^k $(j = 1 \ldots k)$ if and only if the following conditions are satisfied: in the corresponding bracket-sequence in the sites $2i-1$ and $2i$ there are two opening brackets, the next $n-1$ "free spaces" that are not determined yet by earlier al^n-type segments are also opening brackets and the following "free space" is a closing bracket. (So any coding or decoding should be started in the beginning of the bracket-sequence or at the L_1-end of the given term.)

4. The "spaces" not defined by the above rule are closing brackets.

It is easy to check that starting from different terms we get different bracket-sequences from these rules, and inversely, starting from different bracket-sequences we get different terms. So what we only have to show is that starting from any "good" term corresponding to Q_i^k, the corresponding bracket-sequence is syntactically correct and corresponds to S_{k+1-i}^k; and inversely, starting from any syntactically correct bracket-sequence corresponding to S_{k+1-i}^k, the corresponding term is "good" and corresponds to Q_i^k.

Let us introduce four functions as follows. Let $A(i)$ mean the total number of l^*-s in the expression $L_i L_{i-1} \ldots L_2 L_1$, let $C(i)$ mean the total power of l in the expression $L_i L_{i-1} \ldots L_2 L_1$, let $Op(i)$ mean the number of opening brackets in the first i sites of any bracket-sequences and

let $Cl(i)$ mean the number of closing brackets in the first i sites of any bracket-sequences.

We have seen in \$1 that a term corresponding to Q_i^k is "good" if and only if $A(i) \leq C(i)$ for every $0 \leq i < k$ and $A(k) = C(k)$. (In the point $i = 0$ all of these functions are zero, so it is not too interesting here, but later it will be convenient to have them defined, too.)

On the other hand, it is easy to see, that a bracket-sequence corresponding to S_i^k is syntactically correct if and only if $Cl(i) \leq Op(i)$ for every $1 \leq i < 2k$ and $Cl(2k) = Op(2k)$.

Now we show that starting from a "good" term $Cl(i) > Op(i)$ cannot occur for any i. It is clear from the above rules (and especially from rule 3.) that whenever

$A(i) = C(i)$ holds $Cl(2i) = Op(2i)$ also holds. This can be seen from that fact that each al^n segment is "neutralized" by exactly n l^* segments in both senses, that is to say from the sight of the term (this is trivial) and from that of the bracket-sequence (here we need rule 3.), moreover in such a way that $Op > Cl$ and $C > A$ hold as long as the "naturalization" is not complete. (From the sight of the bracket-sequence the advantage of the opening brackets comes from that rule that every al^n-type fragment take immediately two opening brackets in the sequence.) But if $A(i) = C(i)$ holds for some i than there cannot be more l^* in the term as long as there is no more al^n-type segment in it (otherwise the term would not be "good"), so $Cl(i) > Op(i)$ can really never occur. $Cl(2k) = Op(2k)$ is clear from the fact that $A(k) = C(k)$.

The number of opening brackets in even sites equals the number of closing brackets in the odd sites which is the number of annihilation operators in the term, that is $k - i$ (recall that i is the total power of a in the term and l^* is the only segment that does not have an a factor with it). So we have shown that starting from any "good" term corresponding to Q_i^k, the corresponding bracket-sequence is syntactically correct and corresponds to S_{k+1-i}^k according to lemma 2.2.

Now let us see a term corresponding to a syntactically correct bracket-sequence corresponding to S_{k+1-i}^k. It is clear that there are k segments in such a term. Due to the rules $A(k) = C(k)$ holds for sure. (The number of closing brackets in the odd sites, that equals the number of the l^*-s in the corresponding term is clearly the same as the number of opening brackets in the even sites, that equals the total power of l.)

Now we show that $A(i) > C(i)$ cannot occur for any i. Let us suppose that there is a bracket-sequence corresponding to S_{k+1-i}^k so, that for the corresponding term $A(i) > C(i)$ holds for some i. Let us take the smallest i that satisfies this condition. As by the l^*-s we can "neutralize" the creation operators of the al^n segments only one by one, and as we have chosen the smallest possible i, for $j = i - 1$ $A(j) = C(j)$ must hold. Now $L_j L_{j-1} \dots L_2 L_1$ is clearly a "good" term. If we represent it now by brackets, we clearly get back the first $2j$ long part of the original bracket-sequence. $A(i) > C(i)$ means that the i-th segment of original term is l^*. This brings a closing bracket to the sequence, but as we have shown above, $A(j) = C(j)$ leads to $Cl(2j) = Op(2j)$, and from these $Cl(2j + 1) > Op(2j + 1)$ follows in contradiction with our assumption that we have a syntactically correct bracket-sequence.

The total power of a in such a term and the number of annihilation operators (which equals the number of closing brackets in the odd sites and so the number of opening brackets in even sites, too, which is $k - i$) together gives k (as there are k segments altogether and l^* is the only segment that does not have an a factor with it). So the total power of a is $k - (k - i) = i$. So we have shown that starting from any syntactically correct bracket-sequence corresponding to S_{k+1-i}^k, the corresponding term is "good" and corresponds to Q_i^k. □

Lemma 2.5. $S_{k+1-i}^k = S_i^k$ for every $k \geq 1$, $1 \leq i \leq k$.

Proof. It is clear that the number of sequences of $l_1 = k+1-i$ and $l_2 = i$ is the same as the number of sequences of $l_1 = i$ and $l_2 = k+1-i$, because up to this point the two types of indices are symmetric. $\square$

Now we are in a position to complete the proof of Theorem 2.1. We have

$$\lim_{n\to\infty} \tau((\Lambda^{(n)})^k) = \sum_{i=1}^{k} S_k^{(i)} a^{i-1} = \sum_{i=1}^{k} S_k^{(k+1-i)} a^{i-1} = \sum_{i=1}^{k} Q_k^{(i)} a^{i-1},$$

where the equalities are from Lemma 2.2, Lemma 2.5 and Lemma 2.4, respectively.

References

[An] T.W. Anderson, *An introduction to multivariate statistical analysis*, J. Wiley, New York, 1958

[Ba] Z.D. Bai, Convergence rate of expected spectral distribution of large random matrices. Part II. Sample covariance matrices, Ann. Prob. **21**(1993), 649–672

[F-K] Z. Füredi, J. Komlós, The eigenvalues of random symmetric matrices, Combinatorica **1**(1981), 233-241

[M-P] V.A. Marchenko, L.A. Pastur, The distribution of eigenvalues in certain sets of random matrices, Math. Sb. **72**(1967), 507–536

[M-Sz] T.F. Móri, J.G. Székely (eds), *Multivariate statistical analysis* (in Hungarian), Műszaki Könyvkiadó, Budapest, 1986

[R] C.R. Rao, *Linear statistical inference and its applications*, J. Wiley, New York, 1973

[VDN] D.V. Voiculescu, K.J. Dykema and A. Nica, *Free Random Variables*, CRM Monograph Ser., Vol. 1, Amer. Math. Soc., 1992.

[Wa] K.W. Wachter, The strong limit of random matrix spectra for sample matrices of independent elements, Ann. Prob. **6**(1978), 1–8

[Wi] E.P. Wigner, On the distribution of the roots of certain symmetric matrices, Ann. Math. **67**(1958), 325–327