LOGARITHMIC ENERGY AS AN ENTROPY FUNCTIONAL

Dénes Petz and Fumio Hiai

ABSTRACT. This paper is twofold. On the one hand, a short introduction is given to noncommutative random variables and a concise review of some areas of Voiculescu's analysis is presented, especially concentrating on the relation to random matrices. The main goal is to show that the negative logarithmic energy shares some properties with classical entropy functionals. On the other hand, we present a large deviation theorem for the empirical eigenvalue distribution of some not necessarily selfadjoint Gaussian random matrices with a full proof. This extends the first large deviation result due to Ben Arous and Guionnet for selfadjoint Gaussian matrices.

The concept of entropy originated from thermodynamics and became a mathematical notion in the work of Gibbs and Boltzmann. Later it got importance in information theory and in the statistical problem of testing hypothesis. The Boltzmann entropy $-\int g(x)\log g(x)\,dx$ of a probability density g appears mostly in limit theorems. The reason for the observation that entropy shows up in so many limit problems is the fact that this quantity governs often the asymptotics of some probabilities. This is very clear in the large deviation theory which concerns limit theorems with exponential convergence. The rate of the convergence is described by an entropy functional.

In the last 15 years Voiculescu systematically has developed an analysis to create new invariants to study free group factors. He recognized that the algebraic-probabilistic free relation may be modeled asymptotically by random matrices and heuristically studied the asymptotics of the Boltzmann entropy of large random matrices. In this way he arrived at the appropriate entropy concept from the point of view of free relation. In our opinion, his main achievement is the discovery of a concept of entropy which becomes additive in case of freeness while the subadditivity holds in general. (This property provides motivation to call the new concept free entropy, but this free entropy does not have much to do with free entropy in the context of thermodynamics.) Although the highlight of his theory is in the multivariable case, we confine ourselves in this paper mostly to a single variable, which can be a probability measure. In the context of probability measures, the

---

1991 *Mathematics Subject Classification.* 15A52, 46L50, 60F10.

Written form of the invited lecture of the first author in *The* 1997 *Georgia Tech - UAB International Conference on Differential Equations and Mathematical Physics*, March, 23-29 1997. Published in *Advances in Differential Equations and Mathematical Physics*, eds. E. Carlen, E.M. Harell, M. Loss, Contemporary Math. **217**(1998), 205–221.

free entropy is the logarithmic energy, at least up to a sign, which is familiar from two-dimensional potential theory.

The free entropy appears as an important component in the rate function in large deviation theorems for random matrices. Wigner's original result was the convergence of the mean spectral density of a certain Gaussian random matrix to the semicircle law. His result was improved in [1] by showing that the empirical eigenvalue density converges almost surely. The large deviation result was hinted in the work of Voiculescu but first proven by Ben Arous and Guionnet [2]. According to their result the convergence to the semicircle law is exponentially fast.

The present paper is twofold. On the one hand, we give a rather short review of certain areas of Voiculescu's analysis; especially we concentrate on the relation to random matrices. Our main goal is to show that the negative logarithmic energy shares some properties with classical entropy functionals. On the other hand, we present large deviation theorems for the empirical eigenvalue distribution of some random matrices. We mostly consider non-selfadjoint Gaussian matrices in which the real part and the imaginary part play non-symmetric role. We call such matrix elliptic. The large deviation theorem is presented with full proof for elliptic Gaussian matrices. In the elliptic Gaussian case we can determine the limit distribution (i.e. the integrated density of states) which is, of course, the minimizer of the rate function. It is the uniform distribution on an ellipse and called the elliptic law. Note that our large deviation result implies almost sure convergence to the elliptic law. Almost sure convergence to the circular law has been known in the case when real part and imaginary part have the same variance; [11] attributes the proof to J.W. Silverstein. Also, the general elliptic case where random matrices are not necessarily Gaussian was treated by Girko [5, 6]. We explicitly have the joint eigenvalue distribution in the elliptic Gaussian case. Our approach, as well as Silverstein's proof and the method of Ben Arous and Guionnet, is based on the explicit form of the joint distribution of the eigenvalues and does not extend to more general non-Gaussian examples. Finally, we discuss a large deviation result for the eigenvalue distribution of some unitary random matrices, whose proof will be published elsewhere. It is remarkable that the limit distribution of this convergence result was identified earlier by Gross and Witten [7].

1. Noncommutative random variables, free relation and random matrix models

Let A be a bounded or unbounded selfadjoint operator on a Hilbert space $\mathcal{H}$, and let $A = \int \lambda\, dE(\lambda)$ be its spectral decomposition by means of a projection-valued measure E on the Borel sets of the real line. In a state vector $f \in \mathcal{H}$ the distribution of A is the probability measure $H \mapsto \langle E(H)f, f\rangle$. This concept of distribution originated from quantum mechanics. When $\mathcal{H} = L^2(\mathbb{R})$, the position operator Q is defined on the domain $\mathcal{D}(Q) = \{f \in L^2(\mathbb{R}) : xf(x) \in L^2(\mathbb{R})\}$ and $(Qf)(x) = xf(x)$. A unit vector $f \in L^2(\mathbb{R})$ is usually called a wave function in quantum mechanics. The distribution measure of Q at the vector f has the density $|f|^2$ and $\int_H |f(x)|^2\, dx$ gives the probability that the quantum particle of one degree of freedom is confined to a subset $H \subset \mathbb{R}$. This example appeared in von Neumann's famous book "Mathematical Foundation of Quantum Mechanics" published 1932 and it contains the essence of the statistical interpretation of the theory. The momentum P can be regarded as a random variable as well and its probability density in the state f is $|\hat{f}(x)|^2$, where $\hat{f}$ is for the Fourier transform of f. It is well understood that P and Q cannot be regarded as Kolmogorovian random variables on the same probability space because their joint distribution cannot be defined. The reason for this is very simply $PQ \neq QP$. Nevertheless, we need to have a probability theory in which the position and the momentum can be treated at the same time. A solution is the algebraic probability theory. An algebraic or noncommutative probability space is a pair $(\mathcal{A}, \varphi)$ such that $\mathcal{A}$ is an algebra with unit I and $\varphi : \mathcal{A} \to \mathbb{C}$ is a linear functional with $\varphi(I) = 1$, and this gives the expectation value of the noncommutative random variables, elements of $\mathcal{A}$. In case of position and momentum, all polynomials of P and Q can be taken as $\mathcal{A}$. The only good choice for domain is the Schwartz space $\mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R})$ and the functional φ is $a \mapsto \langle af, f\rangle$, where the state vector f is from $\mathcal{S}(\mathbb{R})$.

Random matrices form a very important class of noncommutative random variables and they have been often used in theoretical physics [14]. Let us restrict ourselves to $n \times n$ matrices with entries in an algebra of classical random variables, and set the expectation functional τ_n as $X \mapsto n^{-1} \sum_{i=1}^{n} E(X_{ii})$. A concrete noncommutative random variable is the *standard selfadjoint Gaussian random matrix* $W(n)$:

(i) $\{\mathrm{Re}\, W_{ij}(n) : 1 \leq i \leq j \leq n\} \cup \{\mathrm{Im}\, W_{ij}(n) : 1 \leq i < j \leq n\}$ is an independent family of Gaussian random variables,
(ii) $E(W_{ij}(n)) = 0$ for $1 \leq i \leq j \leq n$, $E(W_{ii}(n)^2) = 1/n$ for $1 \leq i \leq n$, and $E((\mathrm{Re}\, W_{ij})^2) = E((\mathrm{Im}\, W_{ij})^2) = 1/2n$ for $1 \leq i < j \leq n$.

We use the word standard because $\tau_n(W(n)) = 0$ and $\tau_n(W(n)^2) = 1$. The variance of the entries is chosen such a way that the distribution of $W(n)$ is invariant under unitary conjugation. A famous theorem of Wigner tells us that $(W(n), \tau_n)$ converges in distribution to the semicircle law w_2, in other words

$$\lim_{n\to\infty} \tau_n(W(n)^k) = \frac{1}{2\pi} \int_{-2}^{2} x^k \sqrt{4 - x^2}\, dx \quad \text{for} \quad k \in \mathbb{N}.$$

Another class of examples of noncommutative random variables can be shown on the full Fock space $\mathcal{F}(\mathcal{H})$ over a Hilbert space $\mathcal{H}$. For every $h \in \mathcal{H}$ an operator $\ell(h)$ is determined on $\mathcal{F}(\mathcal{H})$:

$$\ell(h)\xi = h \otimes \xi \quad \text{for} \quad \xi \in \mathcal{F}(\mathcal{H}).$$

The vacuum expectation φ is defined as $A \mapsto \langle A\Phi, \Phi\rangle$ on the operators acting on $\mathcal{F}(\mathcal{H})$, Φ being the vacuum vector. (If $\mathcal{H} = \mathbb{C}$, then $\mathcal{F}(\mathcal{H}) = l^2(\mathbb{Z}^+)$ and $\ell(1)$ is simply the right shift.) The distribution of $\ell(h) + \ell(h)^*$ is the standard semicircle law w_2 when h is a unit vector. Hence, we can say that the sequence $(W(n), \tau_n)$ is a *random matrix model* for $\ell(h) + \ell(h)^*$.

Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space and let $\mathcal{A}_i$ $(i \in I)$ be subalgebras of $\mathcal{A}$. We say, after Voiculescu, that the family $\{\mathcal{A}_i : i \in I\}$ is *in free relation* (or shortly free) with respect to φ if

$$\begin{aligned} &\varphi(a_1 a_2 \cdots a_n) = 0 \quad \text{whenever} \\ &\qquad a_k \in \mathcal{A}_{i(k)},\ \ i(k) \in I,\ \ \varphi(a_k) = 0,\ \ i(k) \neq i(k+1). \end{aligned}$$

The elements $a, b \in \mathcal{A}$ are free if the algebras generated by them are free [26]. The free relation is more understandable in the context of the free product of groups, but the full Fock space provides an instructive example as well [26].

EXAMPLE 1. Let $h_i \in \mathcal{H}$ $(i = 1, 2)$. Then $\ell(h_1)$ and $\ell(h_2)$ are free with respect to the vacuum state if and only if $\langle h_1, h_2 \rangle = 0$.

Another example tells us that the spectrum of uv can be continuous even if u and v have discrete spectrum.

EXAMPLE 2. Let u and v be free unitaries with 0 expectation. Then the distribution of uv is the Haar measure on the unit circle.

The free relation is analogous to the independence of classical random variables, for example, because a kind of central limit theorem holds as stated in the following theorem of Voiculescu [21]. Since the limit distribution in the free central limit theorem is the semicircle law, this was the first sign that there must be some connection between freeness and random matrices.

THEOREM 3. *Let $a_1, a_2, \dots \in \mathcal{A}$ and assume that they are free in $\mathcal{A}$ with respect to a state φ. If $\varphi(a_n) = 0$ and $\varphi(a_n^2) = 1$ for every $n \in \mathbb{N}$ and $\sup_n |\varphi(a_n^k)| < +\infty$ for all $k \in \mathbb{N}$, then*

$$\frac{a_1 + a_2 + \dots + a_n}{\sqrt{n}}$$

converges in distribution to the semicircle law w_2.

The next argument is almost a proof of the theorem. Let

$$a_n = \ell(h_n)^* + \ell(h_n) + \sum_{i=2}^{\infty} \alpha_i \ell(h_n)^i ,$$

which gives a Toeplitz operator on $\mathcal{F}(\mathcal{H})$ if $\sum_i |\alpha_i| < \infty$. If we choose the unit vectors $h_1, h_2, \dots$ to be pairwise orthogonal, then $a_1, a_2, \dots$ are free according to Example 1 and all conditions of Theorem 3 are fulfilled. Since the left creation operators $\ell(h)$ are linearly depending on h, in our example $(a_1 + a_2 + \dots + a_n)/\sqrt{n}$ is written as

$$\ell((h_1 + h_2 + \dots + h_n)/\sqrt{n})^* + \ell((h_1 + h_2 + \dots + h_n)/\sqrt{n}) + \sum_{i=2}^{\infty} \alpha_i \sum_{j=1}^{n} \frac{\ell(h_j)^i}{\sqrt{n}} .$$

It is a matter of elementary algebraic computation that the distribution of the last operator is the same as that of

$$\begin{aligned} \ell((h_1 + h_2 + \dots + h_n)/\sqrt{n})^* \quad &+ \quad \ell((h_1 + h_2 + \dots + h_n)/\sqrt{n}) \\ &+ \quad \sum_{i=2}^{\infty} \frac{\alpha_i}{n^{(i-1)/2}} \ell((h_1 + h_2 + \dots + h_n)/\sqrt{n})^i . \end{aligned}$$

We may replace the unit vector $(h_1 + h_2 + \dots + h_n)/\sqrt{n}$ by h_1, which obviously does not affect the distribution. In this way we arrive at

$$\ell(h_1)^* + \ell(h_1) + \sum_{i=2}^{\infty} \frac{\alpha_i}{(\sqrt{n})^{i-1}} \ell(h_1)^i$$

that has the strong limit as $n \to \infty$, namely $\ell(h_1)^* + \ell(h_1)$, which is known to be semicircle-distributed already.

The free relation can be modeled by independent random matrices. Let $Y(n)$ be a sequence of diagonal random matrices such that $(Y(n), \tau_n)$ has a limit distribution f and $Y(n)$ is independent of $W(n)$. Then $Y(n)$ and $W(n)$ become free in the limit

$n \to \infty$. This means that if $p_1, p_2, \dots, p_k$ and $q_1, q_2, \dots, q_k$ are polynomials such that

$$\lim_{n\to\infty} \tau_n\big(p_i(Y(n))\big) = 0 \quad \text{and} \quad \lim_{n\to\infty} \tau_n\big(q_i(W(n))\big) = 0,$$

then

$$\lim_{n\to\infty} \tau_n\big(p_1(Y(n))q_1(W(n))p_2(Y(n)) \cdots q_k(W(n))\big) = 0$$

holds. This discovery allows to use method of freeness to compute the limit distribution g of the sequence $(Y(n) + W(n), \tau_n)$ which is a statement for the limit of the eigenvalue distribution of the random matrix $Y(n) + W(n)$ [26, 18]. In fact, this has been done much earlier than the breakthrough of Voiculescu's analysis. Marchenko and Pastur studied this subject in late 60's and they called a *deformed Wigner law* for the limiting density g [13]. In the language of the free probability theory one says that g is the *additive free convolution* of the limit distributions of $(Y(n), \tau_n)$ and $(W(n), \tau_n)$ [26].

The Wigner theorem concerning the random matrices $W(n)$ has stronger form than the one stated above. Assume that $Y_n(\omega)$ is a random $n \times n$ matrix with eigenvalues $\lambda_1(\omega), \lambda_2(\omega), \dots, \lambda_n(\omega)$. Then the *empirical eigenvalue distribution* of $Y_n(\omega)$ is the random atomic measure

$$P_n(\omega) := \frac{\delta(\lambda_1(\omega)) + \delta(\lambda_2(\omega)) + \cdots + \delta(\lambda_n(\omega))}{n},$$

where $\delta(\zeta)$ denotes the Dirac measure at ζ. The above stated Wigner theorem nearly means that $E(P_n)$ converges to w_2 in weak topology when P_n is the empirical eigenvalue density of $W(n)$ (and $E(P_n)$ is the mean eigenvalue density). However, $P_n(\omega)$ converges to the semicircle law for almost every ω [1]. Some people expresses this fact by saying that w_2 is the *integrated density of states.*

The *standard non-selfadjoint Gaussian random matrix* $X(n)$ is defined in the following way:

(i) $\{\mathrm{Re}\, X_{ij}(n) : 1 \le i, j \le n\} \cup \{\mathrm{Im}\, X_{ij}(n) : 1 \le i, j \le n\}$ is an independent family of Gaussian random variables,
(ii) $E(X_{ij}(n)) = 0$ for $1 \le i, j \le n$ and $E((\mathrm{Re}\, X_{ij})^2) = E((\mathrm{Im}\, X_{ij})^2) = 1/2n$ for $1 \le i, j \le n$.

Equivalently, one can say that $X(n) = (W^{(1)}(n) + \mathrm{i}\, W^{(2)}(n))/\sqrt{2}$, where $W^{(1)}(n)$ and $W^{(2)}(n)$ are independent standard selfadjoint Gaussian matrices.

Let $u, v \in \mathbb{R}$ such that $u, v \ge 0$ and $u^2 + v^2 = 1$, and for $n \in \mathbb{N}$ set

$$Y(n) = uX(n) + vX(n)^*, \tag{1}$$

what we call the *elliptic Gaussian matrix.* This is a kind of interpolation between the standard selfadjoint Gaussian and the standard non-selfadjoint Gaussian cases. Namely, the choices $u = v$ and $v = 0$ recover those examples.

We need the joint distribution $\bar{\nu}_n$ of the eigenvalues of $Y(n)$. The probability measure on $M_n(\mathbb{C})$ induced by $X(n)$ is

$$\frac{1}{Z_n} \exp(-n \mathrm{Tr}\, A^* A)\, d\Lambda_n(A),$$

where Λ_n is the Lebesgue measure on $M_n(\mathbb{C}) \cong \mathbb{R}^{2n^2}$. The measure induced by $Y(n)$ is the image measure of the above via the transformation $B = uA + vA^*$. This is a non-singular linear transformation when $u \neq v$. Its inverse is $A = aB + bB^*$ and therefore the measure induced by $Y(n)$ is

$$\begin{aligned}\nu_n : &= \frac{1}{Z_n}\exp(-n\mathrm{Tr}\,|aA + bA^*|^2)\, d\Lambda_n(A)\\ &= \frac{1}{Z_n}\exp\Big[-\frac{n}{1-\tau^2}\mathrm{Tr}\,(A^*A - \tau\mathrm{Re}\,A^2)\Big]\, d\Lambda_n(A)\,,\end{aligned}$$

where

$$a := \frac{u}{u^2-v^2}\,, \quad b := -\frac{v}{u^2-v^2}\,, \quad \tau := 2uv\,,$$

and Z_n is a new normalizing constant (including the Jacobian of the above linear transformation).

LEMMA 4. *If $u \neq v$ and hence $-1 < \tau < 1$, then $\bar\nu_n$ has the following joint probability density with respect to $d\zeta_1 \cdots d\zeta_n$ ($d\zeta_i$ is the Lebesgue measure on $\mathbb{C}$):*

$$\frac{1}{Z_n}\exp\Big[-n\sum_{i=1}^n\Big(\frac{\mathrm{Re}\,\zeta_i)^2}{1+\tau} + \frac{(\mathrm{Im}\,\zeta_i)^2}{1-\tau}\Big)\Big]\prod_{i<j}|\zeta_i - \zeta_j|^2\,.$$

PROOF. The proof is on the lines similar to that for $X(n)$ in [14, A.35]. Take a unitary U such that $U^*AU = T$ is upper triangular. Then $\zeta_i := T_{ii}$ are the eigenvalues of A and $dH := -\mathrm{i}\,U^*dU$ is Hermitian. We may impose $dH_{ii} = 0$ $(1 \le i \le n)$. Since

$$\mathrm{Tr}\,|aA + bA^*|^2 = \sum_{i=1}^n\frac{(\mathrm{Re}\,\zeta_i)^2}{1+\tau} + \sum_{i=1}^n\frac{(\mathrm{Im}\,\zeta_i)^2}{1-\tau} + \frac{1}{1-\tau^2}\sum_{i<j}|T_{ij}|^2\,,$$

we have

$$\begin{aligned}\nu_n &= \frac{1}{Z_n}\exp\Big[-n\sum_{i=1}^n\Big(\frac{(\mathrm{Re}\,\zeta_i)^2}{1+\tau} + \frac{(\mathrm{Im}\,\zeta_i)^2}{1-\tau}\Big)\Big]\prod_{i<j}|\zeta_i-\zeta_j|^2\prod_{i=1}^n d\zeta_i\\ &\quad\times\exp\Big(-\frac{n}{1-\tau^2}\sum_{i<j}|T_{ij}|^2\Big)\prod_{i<j}d(\mathrm{Re}\,T_{ij})\,d(\mathrm{Im}\,T_{ij})\\ &\quad\times\prod_{i<j}d(\mathrm{Re}\,H_{ij})\,d(\mathrm{Im}\,H_{ij})\,.\end{aligned}$$

Integrating with respect to $\prod d(\mathrm{Re}\,T_{ij})\,d(\mathrm{Im}\,T_{ij})$ and $\prod d(\mathrm{Re}\,H_{ij})\,d(\mathrm{Im}\,H_{ij})$ we obtain the desired result. □

2. Logarithmic energy versus Voiculescu's entropy

Let μ be a measure on $\mathbb{C}$. In a physical picture the measure μ may be thought as the distribution of electric charges in a two-dimensional universe. The double integral

$$I(\mu) = -\iint \log|x-y|\,d\mu(x)\,d\mu(y)$$

yields the Coulomb energy of the two dimensional electrostatic field due to electrostatic repulsion if the repulsion force between charges is proportional to the inverse of the distance. $I(\mu)$ has been called the *logarithmic energy* of μ [12, 19].

On the other hand, Voiculescu's work opened a completely new perspective for the logarithmic energy [23, 24]. Before discussing this, we recall the classical relative entropy, or Kullback-Leibler information. If μ_i $(i = 1, 2)$ are measures on

$\mathbb{R}$ with density $f_i(x)$, then the *relative entropy* of μ_1 with respect to μ_2 is defined as

$$S(\mu_1, \mu_2) = \int f_1(x)\big(\log f_1(x) - \log f_2(x)\big)\, dx\,.$$

This quantity is non-negative when μ_1 and μ_2 are probability measures. It is worthwhile to note that $S(\mu_1, \mu_2)$ is not symmetric in its two variables; μ_1 and μ_2 play really different roles.

Let μ and ν be probability measures on $\mathbb{R}$ and ν^n denote the n-fold product measure $\nu \times \nu \times \cdots \times \nu$. Furthermore, let $m_k(\mu)$ be the kth moment of μ, i.e. $m_k(\mu) = \int x^k\, d\mu(x)$ for $k \in \mathbb{N}$. For $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ we denote by κ_x the discrete measure

$$\frac{1}{n}\big(\delta(x_1) + \delta(x_2) + \cdots + \delta(x_n)\big).$$

THEOREM 5. *Assume that μ and ν are probability measures on $\mathbb{R}$ and μ is compactly supported. Then*

$$\lim_{\substack{r\to\infty \\ \varepsilon\to+0}} \limsup_{n\to\infty} \frac{1}{n} \log \nu^n\big(\{x \in \mathbb{R}^n : |m_k(\kappa_x) - m_k(\mu)| \le \varepsilon,\ k \le r\}\big) = -S(\mu, \nu)\,.$$

Note that this theorem follows from the so-called Sanov theorem [4, 10]. In particular, if ν is the standard Gaussian measure, then the limit is

$$S(\mu) - (\log 2\pi + m_2(\mu))/2$$

where $m_2(\mu)$ is the second moment. So the main term is the *Boltzmann entropy* $S(\mu)$.

One way of interpretation of Voiculescu's work is that he replaced the product measure ν^n in Theorem 5 by another $\bar{\nu}_n$ whose joint probability distribution is of the form

$$\frac{1}{Z^{(n)}_{\beta,\sigma}} \exp\left(-\frac{n}{4\sigma^2}\sum_{i=1}^{n} \lambda_i^2\right) \prod_{i<j} |\lambda_i - \lambda_j|^{2\beta}\,, \tag{2}$$

where $\beta, \sigma^2 > 0$ are fixed and $Z^{(n)}_{\beta,\sigma}$ is a constant for normalization. This measure is an extension of the joint distribution of the eigenvalues of the standard selfadjoint Gaussian matrix $W(n)$, $\beta = 1$ and $\sigma^2 = 1/2$.

THEOREM 6. *Let μ be a measure on $\mathbb{R}$ with density $f(x)$ of compact support and denote $\bar{\nu}^n$ the measure of density* (2) *on $\mathbb{R}^n$. Then*

$$\begin{aligned}&\lim_{\substack{r\to\infty \\ \varepsilon\to+0}} \limsup_{n\to\infty} \frac{1}{n^2} \log \bar{\nu}^n\{\lambda \in \mathbb{R}^n : |m_k(\kappa_\lambda) - m_k(\mu)| \le \varepsilon,\ k \le r\} \\ &\quad = \beta \iint \log|x-y| f(x) f(y)\, dx\, dy - \frac{1}{4\sigma^2}\int x^2 f(x)\, dx - \frac{\beta}{2}\log(2\beta\sigma^2) + \frac{3\beta}{4}\,.\end{aligned}$$

If we compare Theorems 5 and 6, we observe that the logarithmic energy of the measure μ is replacing the Boltzmann entropy term. This is a completely new aspect of the logarithmic energy. We call

$$\Sigma(\mu) = \iint \log|x-y|\, d\mu(x) d\mu(y)$$

the *free entropy* of the measure μ.

Basic properties of $\Sigma(\mu)$ are determined by the fact that the logarithmic kernel $(x,y)\mapsto \log|x-y|$ is strictly negative definite. This yields the following concavity result [8].

THEOREM 7. *The free entropy functional $\Sigma(\mu)$ is strictly concave and weakly upper semicontinuous restricted to the set of probability measures supported in a compact subset of $\mathbb{C}$.*

Voiculescu's definition for the entropy of the multivariables $(X_1, X_2, \dots, X_n)$ goes as follows. $X_1, X_2, \dots, X_n$ are selfadjoint elements of a von Neumann algebra endowed with a faithful normal tracial state τ.

For $m, k \in \mathbb{N}$, $R>0$, and $\varepsilon>0$ set:

$$\begin{aligned}\Gamma_R(X_1,\dots,X_n;m,k,\varepsilon) \;=\; & \{(A_1,\dots,A_n)\in (M_k(\mathbb{C})^{sa})^n : \|A_i\|\le R,\\ & |\tau(X_{i_1}\cdots X_{i_p})-\tau_k(A_{i_1}\cdots A_{i_p})|\le\varepsilon\\ & \text{for all } (i_1,\dots,i_p)\in\{1,\dots,n\}^p,\ 1\le p\le m\}\,,\end{aligned}$$

which is a set of n-tuples of $k\times k$ selfadjoint matrices approximating the given n-tuple $(X_1,\dots,X_n)$ in a set of joint moments.

$$\chi_R(X_1,\dots,X_n;m,\varepsilon)=\limsup_{k\to\infty}\left[\frac{1}{k^2}\log\Lambda(\Gamma_R(X_1,\dots,X_n;m,k,\varepsilon))+\frac{n}{2}\log k\right],$$

where Λ is the Lebesgue measure on $(M_k(\mathbb{C})^{sa})^n$ and the correction $(n/2)\log k$ gives a chance for a finite lim sup. After the matrix size went to infinity, we take the limit in the other parameters:

$$\chi_R(X_1,\dots,X_n)=\lim_{\substack{m\to\infty\\ \varepsilon\to+0}}\chi_R(X_1,\dots,X_n;m,\varepsilon)\,,$$

$$\chi(X_1,\dots,X_n)=\sup_{R>0}\chi_R(X_1,\dots,X_n)\,.$$

Finally $\chi(X_1,\dots,X_n)$ is the (multiple) free entropy of the n-tuple $(X_1,\dots,X_n)$.

For the case of a single X, we can take $R\ge\|X\|$ and

$$\chi_R(X;m,\varepsilon)=\lim_{k\to\infty}\left[\frac{1}{k^2}\log\Lambda(\Gamma_R(X;m,k,\varepsilon))+\frac{1}{2}\log k\right]$$

since the lim sup becomes a limit, and

$$\chi(X)=\chi_R(X)=\Sigma(X)+\frac{3}{4}+\frac{1}{2}\log 2\pi\,.$$

The additive constant is not important, it could be removed by an appropriate renormalization of the measures used on matrix algebras.

The subadditivity

$$\chi(X_1,\dots,X_n)\le\chi(X_1,\dots,X_k)+\chi(X_{k+1},\dots,X_n).$$

holds for every $1\le k\le n$. The key property of the multiple free entropy is the additivity under freeness [24, 25]:

THEOREM 8. *If $X_1,\dots,X_n$ are in free relation, then*

$$\chi(X_1,\dots,X_n)=\chi(X_1)+\dots+\chi(X_n)\,.$$

Conversely, if $\chi(X_i)>-\infty$ for $1\le i\le n$ and the above additivity holds, then $X_1,\dots,X_n$ are free.

If (X,τ) is a not necessarily selfadjoint noncommutative random variable, then by the free entropy of X we can mean simply $\chi(X_1,X_2)$ where $X = X_1 + \mathrm{i}\,X_2$. Under the constraint $\tau(X^*X) = 1$ the entropy $\chi(X_1,X_2)$ is maximal if X_1 and X_2 are free and semicircular-distributed. Such X is called a *circular* element by Voiculescu. While the semicircular element is the analogue of real Gaussian random variable, the circular resembles complex Gaussian variable. The standard non-selfadjoint Gaussian matrix is a random matrix model for the circular variable; in fact it converges in distribution to the circular law. In the next section we shall prove that the convergence is exponential, that is, there is a large deviation in the background.

3. Large deviation for empirical eigenvalue distribution

Since we are interested in large deviation for the empirical eigenvalue distribution of certain random matrices, we recall the concept of large deviation in the appropriate setting. See [3,4] for more general definitions. Let $\mathcal{M}(\mathbb{C})$ be the space of all probability Borel measures on $\mathbb{C}$. $\mathcal{M}(\mathbb{C})$ is a Polish space when endowed with weak topology. Let P_n be a measure on $\mathcal{M}(\mathbb{C})$, in other words P_n is a random measure on $\mathbb{C}$ or a measure-valued random variable, and we write sometimes $P_n(\omega)$ to emphasize this side. The sequence P_n satisfies the *large deviation principle* if

$$\liminf_{n\to\infty} n^{-2}\log P_n(G) \ge -\inf\{I(\mu) : \mu \in G\}$$

holds for every open set $G \subset \mathcal{M}(\mathbb{C})$ and

$$\limsup_{n\to\infty} n^{-2}\log P_n(F) \le -\inf\{I(\mu) : \mu \in F\}$$

holds for every closed set $F \subset \mathcal{M}(\mathbb{C})$ with a lower semicontinuous function $I : \mathcal{M}(\mathbb{C}) \to \mathbb{R}^+ \cup \{\infty\}$. When the level sets $\{\mu : I(\mu) \le c\}$ are compact for all $c \ge 0$, it is said that I is a *good rate function.*

The first large deviation theorem for the empirical eigenvalue distribution of selfadjoint Gaussian random matrices was inspired by the work of Voiculescu [23, 24] and proven by Ben Arous and Guionnet [2]. We do not state their result now because it will be recovered from our generalization.

Our main goal is to discuss the large deviation principle for the empirical eigenvalue distribution of the elliptic Gaussian matrix $Y(n)$ given in (1). Let λ_τ denote the *elliptic law* uniformly distributed on the ellipse

$$E_\tau := \Big\{x + \mathrm{i}\,y : \frac{x^2}{(1+\tau)^2} + \frac{y^2}{(1-\tau)^2} \le 1\Big\}.$$

THEOREM 9. *Let P_n be the empirical eigenvalue distribution of the elliptic Gaussian random matrix $Y(n)$. Then*

$$\lim_{n\to\infty} \frac{1}{n^2}\log Z_n = -\frac{3}{4},$$

and P_n satisfies the large deviation principle with good rate function

$$I(\mu) := -\Sigma(\mu) + \int \Big(\frac{x^2}{1+\tau} + \frac{y^2}{1-\tau}\Big)\,d\mu(\zeta) - \frac{3}{4} \qquad (\zeta = x + \mathrm{i}\,y)$$

for $\mu \in \mathcal{M}(\mathbb{C})$. Furthermore, the elliptic law λ_τ is a unique minimizer of I so that $I(\lambda_\tau) = 0$.

The theorem implies that the empirical eigenvalue distribution of $Y(n)$ converges to the elliptic law λ_τ almost surely, or the integrated density of states of $Y(n)$ is λ_τ, while this is rather known [5]. In particular, in the selfadjoint case where $u = v = 1/\sqrt{2}$ (so $\tau = 1$), the large deviation result was obtained by Ben Arous and Guionnet, and the limit distribution is the semicircle law w_2 $(= \lambda_1)$. In the circular case where $v = 0$ (so $\tau = 0$), the limit distribution is the circular law.

From now on we will treat a more general probability on $M_n(\mathbb{C})^{sa}$. Let $Q(\zeta)$ be a real continuous function on $\mathbb{C}$ such that

$$\lim_{|\zeta|\to\infty} |\zeta| \exp(-\varepsilon Q(\zeta)) = 0 \quad \text{for any} \quad \varepsilon > 0. \tag{3}$$

For each $n \in \mathbb{N}$ let ν_n be a probability measure on $M_n(\mathbb{C})$ and assume that the induced measure $\bar{\nu}_n$ on the space $\mathbb{C}^n$ of eigenvalues has the joint probability density

$$\frac{1}{Z_n} \exp\left(-n \sum_{i=1}^n Q(\zeta_i)\right) \prod_{i<j} |\zeta_i - \zeta_j|^{2\beta}, \tag{4}$$

where $\beta > 0$ is fixed (independent of n) and Z_n is a normalization constant. Thanks to Lemma 4 we may take $Q(\zeta) = x^2/(1+\tau) + y^2/(1-\tau)$ $(\zeta = z + \mathrm{i}\,y)$ and $\beta = 1$ in the elliptic Gaussian case. Assumption (3) is satisfied when $- < \tau < 1$. (But the discussions below are valid even when $\tau = 1$, the selfadjoint case). In this case $Q(\zeta)$ means

$$Q(x + \mathrm{i}\,y) = \begin{cases} x^2/2 & \text{if } y = 0, \\ \infty & \text{if } y \neq 0, \end{cases}$$

and we may restrict probability measures to those supported on $\mathbb{R}$. The situation when $\tau = -1$ is of course similar.)

To prove the theorem, set

$$\begin{aligned} F(\zeta, \eta) &= -\beta \log|\zeta - \eta| + \frac{1}{2}(Q(\zeta) + Q(\eta)), \\ F_\alpha(\zeta, \eta) &= \min\{F(\zeta, \eta), \alpha\} \quad \text{for} \quad \alpha > 0. \end{aligned}$$

Since F is bounded from below thanks to (3), $F_\alpha(\zeta, \eta)$ is bounded and continuous, so that

$$\mu \in \mathcal{M}(\mathbb{C}) \mapsto \iint F_\alpha(\zeta, \eta)\, d\mu(\zeta)\, d\mu(\eta)$$

is continuous and

$$\begin{aligned} &-\beta\Sigma(\mu) + \int Q(\zeta)\, d\mu(\zeta) \\ &\quad = \iint F(\zeta, \eta)\, d\mu(\zeta)\, d\mu(\eta) = \sup_{\alpha>0} \iint F_\alpha(\zeta, \eta)\, d\mu(\zeta)\, d\mu(\eta) \end{aligned}$$

is lower semicontinuous in the weak topology on $\mathcal{M}(\mathbb{C})$. For simplicity write

$$\kappa_z = \frac{1}{n} \sum_{i=1}^n \delta(z_i) \quad \text{for} \quad z = (z_1, \dots, z_n) \in \mathbb{C}^n.$$

LEMMA 10.

$$\limsup_{n\to\infty} \frac{1}{n^2} \log Z_n \le - \inf_{\mu\in\mathcal{M}(\mathbb{C})} \iint F(\zeta, \eta)\, d\mu(\zeta)\, d\mu(\eta).$$

PROOF. We get

$$\begin{aligned} Z_n &= \int \cdots \int \exp\left(-\sum_{i=1}^n Q(z_i)\right) \exp\left(-2\sum_{i<j} F(z_i, z_j)\right) dz_1 \cdots dz_n \\ &\le \int \cdots \int \exp\left(-\sum_{i=1}^n Q(z_i)\right) \\ &\qquad \times \exp\left(-n^2 \iint_{\{\zeta \ne \eta\}} F(\zeta,\eta)\, d\kappa_z(\zeta)\, d\kappa_z(\eta)\right) dz_1 \cdots dz_n \\ &\le \exp\left(-n^2 \inf_\mu \iint_{\{\zeta\ne\eta\}} F(\zeta,\eta)\, d\mu(\zeta)\, d\mu(\eta)\right) \\ &\qquad \times \int \cdots \int \exp\left(-\sum_{i=1}^n Q(z_i)\right) dz_1 \cdots dz_n \\ &= \left(\int e^{-Q(z)}\, dz\right)^n \exp\left(-n^2 \inf_\mu \iint F(\zeta,\eta)\, d\mu(\zeta)\, d\mu(\eta)\right), \end{aligned}$$

implying the conclusion. □

LEMMA 11. *For every $\mu \in \mathcal{M}(\mathbb{C})$,*

$$\begin{aligned} &\inf_G \left[\limsup_{n\to\infty} \frac{1}{n^2} \log P_n(G)\right] \\ &\qquad \le -\iint F(\zeta,\eta)\, d\mu(\zeta)\, d\mu(\eta) - \liminf_{n\to\infty} \frac{1}{n^2} \log Z_n, \end{aligned}$$

where G runs over a neighborhood base of μ.

PROOF. For any neighborhood G of $\mu \in \mathcal{M}(\mathbb{R})$ put

$$\tilde{G} = \{z \in \mathbb{C}^n : \kappa_z \in G\}.$$

As in the proof of Lemma 10 we get

$$\begin{aligned} P_n(G) &= \bar{\nu}_n(\tilde{G}) \\ &= \frac{1}{Z_n} \int \cdots \int_{\tilde{G}} \exp\left(-\sum_{i=1}^n Q(z_i)\right) \exp\left(-2\sum_{i<j} F(z_i,z_j)\right) dz_1 \cdots dz_n \\ &\le \frac{1}{Z_n} \int \cdots \int_{\tilde{G}} \exp\left(-\sum_{i=1}^n Q(z_i)\right) \\ &\qquad \times \exp\left(-n^2 \iint F_\alpha(\zeta,\eta)\, d\kappa_z(\zeta)\, d\kappa_z(\eta) + n\alpha\right) dz_1 \cdots dz_n \\ &= \frac{1}{Z_n} \left(\int e^{-Q(z)}\, dz\right)^n \exp\left(-n^2 \inf_{\mu' \in G} \iint F_\alpha(\zeta,\eta)\, d\mu'(\zeta)\, d\mu'(\eta) + n\alpha\right). \end{aligned}$$

Therefore

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n^2} \log P_n(G) \\ &\qquad \le -\inf_{\mu'\in G} \iint F_\alpha(\zeta,\eta)\, d\mu'(\zeta)\, d\mu'(\eta) - \liminf_{n\to\infty} \frac{1}{n^2} \log Z_n. \end{aligned}$$

Thanks to weak continuity of $\mu' \mapsto \iint F_\alpha(\zeta,\eta)\, d\mu'(\zeta)\, d\mu'(\eta)$ we get

$$\inf_G \left[\limsup_{n\to\infty} \frac{1}{n^2} \log P_n(G)\right] \le -\iint F_\alpha(\zeta,\eta)\, d\mu(\zeta)\, d\mu(\eta) - \liminf_{n\to\infty} \frac{1}{n^2} \log Z_n.$$

Letting $\alpha \to +\infty$ yields the desired inequality. □

LEMMA 12. *For every $\mu \in \mathcal{M}(\mathbb{C})$,*

$$\liminf_{n\to\infty} \frac{1}{n^2} \log Z_n \ge - \iint F(\zeta,\eta)\, d\mu(\zeta)\, d\mu(\eta)$$

and

$$\inf_G \left[\liminf_{n\to\infty} \frac{1}{n^2} \log P_n(G) \right] \ge - \iint F(\zeta,\eta)\, d\mu(\zeta)\, d\mu(\eta) - \limsup_{n\to\infty} \frac{1}{n^2} \log Z_n \,,$$

where G runs over a neighborhood base of μ.

PROOF. First, a suitable regularization process can be performed. It is clear that

$$\mu \in \mathcal{M}(\mathbb{C}) \mapsto \inf\left\{ \liminf_{n\to\infty} \frac{1}{n^2} \log P_n(G) : G \text{ is a neighborhood of } \mu \right\}$$

is upper semicontinuous. Since $F(\zeta,\eta)$ is bounded below, if $\iint F(\zeta,\eta)\, d\mu(\zeta)\, d\mu(\eta) < +\infty$ and μ_k is the conditional measure to the disk of radius k, then

$$\iint F(\zeta,\eta)\, d\mu(\zeta)\, d\mu(\eta) = \lim_{k\to\infty} \iint F(\zeta,\eta)\, d\mu_k(\zeta)\, d\mu_k(\eta)\,.$$

So we may assume that μ has a compact support. For $\varepsilon > 0$ let φ_ε be a nonnegative C^∞-function supported in the disk of radius ε such that $\int \varphi_\varepsilon(\zeta)\, d\zeta = 1$, and $\varphi_\varepsilon * \mu$ be the convolution of μ with φ_ε. Thanks to concavity and upper semicontinuity of $\Sigma(\mu)$ restricted on probability measures with uniformly bounded supports, it is easy to see that

$$\Sigma(\varphi_\varepsilon * \mu) \ge \Sigma(\mu)\,.$$

Also

$$\lim_{\varepsilon\to+0} \int Q(\zeta)\, d(\varphi_\varepsilon * \mu)(\zeta) = \int Q(\zeta)\, d\mu(\zeta)\,.$$

Hence we may assume that μ has a continuous density with compact support. Moreover, let λ be the uniform distribution on a rectangle including $\operatorname{supp}\mu$. Then it suffices to show the required inequalities for each $(1-\delta)\mu + \delta\lambda$ $(0 < \delta < 1)$. After all, we may assume that μ is supported on a rectangle R, it has a continuous density f on that rectangle and $\delta \le f(\zeta) \le \delta^{-1}$ for some $\delta > 0$ on R.

Next we divide R into n smaller rectangles $R_i^{(n)}$, $1 \le i \le n$. The following assumption can be made: $\mu(R_i^{(n)}) = 1/n$ and the ratio of the horizontal and vertical length of $R_i^{(n)}$ is uniformly bounded from above and below. Then we get

$$\lim_{n\to\infty} \left(\max_{1\le i\le n} \operatorname{diam}(R_i^{(n)}) \right) \to 0\,. \tag{5}$$

In each rectangle $R_i^{(n)}$ we take a smaller one $S_i^{(n)}$ by dividing $R_i^{(n)}$ into 9 congruent rectangles and selecting the one in the middle, so that

$$\int_{S_i^{(n)}} d\zeta \ge \frac{\delta}{9n}\,.$$

Now we set

$$\Delta_n = \{(\zeta_1, \dots, \zeta_n) \in \mathbb{C}^n : \zeta_i \in S_i^{(n)},\ 1 \le i \le n\}\,.$$

For any neighborhood G of μ, it is easy to check that

$$\Delta_n \subset \{\zeta \in \mathbb{C}^n : \kappa_\zeta \in G\}$$

for all n large enough. For such n we have

$$\begin{aligned} P_n(G) &\geq \bar{\nu}_n(\Delta_n) \\ &\geq \frac{1}{Z_n}\int\cdots\int_{\Delta_n}\exp\left(-n\sum_{i=1}^n Q(\zeta_i)\right)\prod_{i<j}|\zeta_i-\zeta_j|^{2\beta}\,d\zeta_1\cdots d\zeta_n \\ &\geq \frac{1}{Z_n}\exp\left(-n\sum_{i=1}^n \max_{\zeta\in S_i^{(n)}} Q(\zeta_i)\right) \\ &\qquad\times\prod_{i<j}\left(\min_{\zeta\in S_i^{(n)},\eta\in S_j^{(n)}}|\zeta-\eta|^{2\beta}\right)\int\cdots\int_{\Delta_n}d\zeta_1\cdots d\zeta_n \\ &\geq \frac{1}{Z_n}\left(\frac{\delta}{9n}\right)^n\exp\left(-n\sum_{i=1}^n\max_{\zeta\in S_i^{(n)}}Q(\zeta_i)\right)\prod_{i<j}\left(\min_{\zeta\in S_i^{(n)},\eta\in S_j^{(n)}}|\zeta-\eta|^{2\beta}\right). \end{aligned}$$

So, to obtain the required inequalities, it suffices to show that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n\left(\max_{\zeta\in S_i^{(n)}}Q(\zeta_i)\right)=\int Q(\zeta)f(\zeta)\,d\zeta \tag{6}$$

and

$$\liminf_{n\to\infty}\frac{2}{n^2}\sum_{i<j}\log\left(\min_{\zeta\in S_i^{(n)},\eta\in S_j^{(n)}}|\zeta-\eta|\right)\geq\iint f(\zeta)f(\eta)\log|\zeta-\eta|\,d\zeta\,d\eta\,. \tag{7}$$

But (6) is clear from (5). We have

$$\begin{aligned} &\iint f(\zeta)f(\eta)\log|\zeta-\eta|\,d\zeta\,d\eta \\ &\leq 2\sum_{i<j}\int_{R_i^{(n)}}\int_{R_j^{(n)}}f(\zeta)f(\eta)\log|\zeta-\eta|\,d\zeta\,d\eta \\ &\leq 2\sum_{i<j}\log\left(\max_{\zeta\in R_i^{(n)},\eta\in R_j^{(n)}}|\zeta-\eta|\right)\int_{R_i^{(n)}}f(\zeta)\,d\zeta\int_{R_j^{(n)}}f(\eta)\,d\eta \\ &= \frac{2}{n^2}\sum_{i<j}\log\left(\max_{\zeta\in R_i^{(n)},\eta\in R_j^{(n)}}|\zeta-\eta|\right). \end{aligned}$$

Since the construction of $R_i^{(n)}$ and $S_i^{(n)}$ yields

$$\lim_{n\to\infty}\frac{2}{n^2}\sum_{i<j}\log\left(\frac{\max_{\zeta\in R_i^{(n)},\eta\in R_j^{(n)}}|\zeta-\eta|}{\min_{\zeta\in S_i^{(n)},\eta\in S_j^{(n)}}|\zeta-\eta|}\right)=0\,,$$

we obtain (7). □

LEMMA 13. *The finite limit $L := \lim_{n\to\infty} n^{-2}\log Z_n$ exists.*

PROOF. By Lemmas 10 and 12 we have

$$\limsup_{n\to\infty}\frac{1}{n^2}\log Z_n\leq-\inf_\mu\iint F(\zeta,\eta)\,d\mu(\zeta)\,d\mu(\eta)\leq\liminf_{n\to\infty}\frac{1}{n^2}\log Z_n\,.$$

□

LEMMA 14. *(P_n) is exponentially tight.*

PROOF. For any $\alpha > 0$ set

$$K_\alpha = \left\{\mu \in \mathcal{M}(\mathbb{R}) : \int Q(\zeta)\, d\mu(\zeta) \le \alpha\right\}.$$

Since $Q(\zeta) \to +\infty$ as $|\zeta| \to +\infty$ by assumption (3), it is easy to see that

$$\sup_{\mu \in K_\alpha} (\{\zeta : |\zeta| \ge r\}) \to 0 \quad \text{as} \quad r \to +\infty$$

and hence K_α is compact in weak topology. We have

$$\begin{aligned} P_n(K_\alpha^c) &= \bar{\nu}_n\left(\left\{\zeta \in \mathbb{C}^n : \frac{1}{n}\sum_{i=1}^n Q(\zeta_i) > \alpha\right\}\right) \\ &= \frac{1}{Z_n}\int \cdots \int_{\left\{\frac{1}{n}\sum_{i=1}^n Q(\zeta_i) > \alpha\right\}} \exp\left(-n\sum_{i=1}^n Q(\zeta_i)\right) \\ &\qquad\qquad \times \prod_{i<j} |\zeta_i - \zeta_j|^{2\beta}\, d\zeta_1 \cdots d\zeta_n \\ &\le \frac{1}{Z_n}\exp\left(-\frac{n^2\alpha}{2}\right)\int \cdots \int \exp\left(-\frac{n}{2}\sum_{i=1}^n Q(\zeta_i)\right) \\ &\qquad\qquad \times \prod_{i<j} |\zeta_i - \zeta_j|^{2\beta}\, d\zeta_1 \cdots d\zeta_n\,. \end{aligned}$$

When $Q(\zeta)$ is replaced by $Q(\zeta)/2$, Lemma 13 may be referred to and the finite limit

$$B_2 := \lim_{n\to\infty} \frac{1}{n^2} \log \int \cdots \int \exp\left(-\frac{n}{2}\sum_{i=1}^n Q(\zeta_i)\right) \prod_{i<j} |\zeta_i - \zeta_j|^{2\beta}\, d\zeta_1 \cdots d\zeta_n$$

exists. From the above estimate we conclude

$$\limsup_{n\to\infty} \frac{1}{n^2} \log P_n(K_\alpha^c) \le -B + B_2 - \frac{\alpha}{2}\,.$$

Since $\alpha > 0$ is arbitrary, we have the conclusion. □

THEOREM 15. *Assume that the joint distribution of the random variables* ξ_{n1}, $\xi_{n2}, \dots, \xi_{nn}$ *is given by* (4)*. If assumption* (3) *holds, then the random measure*

$$P_n(\omega) := \frac{\delta(\xi_{n1}(\omega)) + \delta(\xi_{n2}(\omega)) + \cdots + \delta(\xi_{nn}(\omega))}{n}$$

satisfies the large deviation principle with good rate function

$$I(\mu) := -\beta\Sigma(\mu) + \int Q(\zeta)\, d\mu(\zeta) + L$$

for $\mu \in \mathcal{M}(\mathbb{C})$*. The constant* L *is determined by the condition* $\inf I(\mu) = 0$ *and also given by* $L = \lim_{n\to\infty} n^{-2} \log Z_n$*. There exists a unique* $\mu_0 \in \mathcal{M}(\mathbb{C})$ *such that* $I(\mu_0) = 0$.

PROOF. First we note that the measures P_n are exponentially tight due to Lemma 14. By Lemmas 12 and 13, $I(\mu) = \iint F(\zeta, \eta)\, d\mu(\zeta)\, d\mu(\eta) + L \ge 0$ for all $\mu \in \mathcal{M}(\mathbb{C})$ and I becomes a rate function. According to the general theory of large deviations [3], it suffices to prove that

$$\inf_G \left[\limsup_{n\to\infty} n^{-2} \log P_\varepsilon(G)\right] \le -I(\mu) \le \inf_G \left[\liminf_{n\to\infty} n^{-2} \log P_\varepsilon(G)\right]$$

for every $\mu \in \mathcal{M}(\mathbb{C})$, where G runs over a neighborhood base of μ. This holds according to Lemmas 11 and 12.

The minimizer of the rate function is unique, because the minimizer is nothing else but the equilibrium measure for a weighted potential [15]. □

Now, to prove Theorem 9, it remains to specify the minimizer $\mu_0 = \lambda_\tau$ and the constant term $L = -3/4$ for the rate function I in the theorem. According to the generalized Frostman theorem about weighted potentials [15] (also [19, Appendix]), the minimizer μ_0 in Theorem 15 is characterized by the following condition for some real constant C:

$$\int \log|\zeta - \eta|\, d\mu_0(\eta) \begin{cases} = \dfrac{1}{2}Q(\zeta) + C & \text{if } \zeta \in \operatorname{supp}\mu_0, \\[2ex] \le \dfrac{1}{2}Q(\zeta) + C & \text{if } \zeta \in \mathbb{C} \setminus \operatorname{supp}\mu_0, \end{cases}$$

and $L = C - \frac{1}{2}\int Q(\zeta)\, d\mu_0(\zeta)$ in this case. So the next lemma is enough for our purpose.

LEMMA 16. *For* $\zeta = x + \mathrm{i}\, y$,

$$\int \log|\zeta - \eta|\, d\lambda_\tau(\eta) \begin{cases} = \dfrac{1}{2}\Big(\dfrac{x^2}{1+\tau} + \dfrac{y^2}{1-\tau}\Big) - \dfrac{1}{2} & \text{if } \zeta \in E_\tau, \\[2ex] < \dfrac{1}{2}\Big(\dfrac{x^2}{1+\tau} + \dfrac{y^2}{1-\tau}\Big) - \dfrac{1}{2} & \text{if } \zeta \in \mathbb{C} \setminus E_\tau. \end{cases}$$

PROOF. This is proved in detail in [8]; so we just sketch the proof here. In case of $\tau = 1$ the conclusion means

$$\int_{-2}^{2} w_2(y) \log|x - y|\, dy \begin{cases} = \dfrac{x^2}{4} - \dfrac{1}{2} & \text{if } |x| \le 2, \\[2ex] < \dfrac{x^2}{4} - \dfrac{1}{2} & \text{if } |x| > 2, \end{cases}$$

and this is a special case [20, pp. 12–13]. When $-1 < \tau < 1$, one can use the Gauss integral formula (see [5, p. 687], [16, Appendix IV]) to show that

$$\int \frac{d\lambda_\tau(\eta)}{\zeta - \eta} = \begin{cases} \dfrac{x}{1+\tau} - \mathrm{i}\,\dfrac{y}{1-\tau} & \text{if } \zeta \in E_\tau, \\[2ex] \dfrac{\zeta}{2\tau}\left(1 - \sqrt{1 - \dfrac{4\tau}{\zeta^2}}\right) & \text{if } \zeta \in \mathbb{C} \setminus E_\tau. \end{cases}$$

(Note that this formula was given in [17].) Let $\zeta_0 = x_0 + \mathrm{i}\, y_0$ be on the boundary of E_τ, and set $\phi(t) := \int \log|t\zeta_0 - \eta|\, d\lambda_\tau(\eta)$ for $t \ge 0$. Then it follows that

$$\phi'(t) = x_0 \operatorname{Re}\int \frac{d\lambda_\tau(\eta)}{t\zeta_0 - \eta} - y_0 \operatorname{Im}\int \frac{d\lambda_\tau(\eta)}{t\zeta_0 - \eta} \begin{cases} = \dfrac{tx_0^2}{1+\tau} + \dfrac{ty_0^2}{1-\tau} & \text{if } 0 < t < 1, \\[2ex] < \dfrac{tx_0^2}{1+\tau} + \dfrac{ty_0^2}{1-\tau} & \text{if } t > 1, \end{cases}$$

so that

$$\phi(t)\begin{cases} = \dfrac{1}{2}\Big(\dfrac{t^2x_0^2}{1+\tau}+\dfrac{t^2y_0^2}{1-\tau}\Big)+\phi(0) & \text{if } 0\le t\le 1,\\[2ex] < \dfrac{1}{2}\Big(\dfrac{t^2x_0^2}{1+\tau}+\dfrac{t^2y_0^2}{1-\tau}\Big)+\phi(0) & \text{if } t>1.\end{cases}$$

Finally, $\phi(0)=-1/2$ is an elementary integral computation. □

The above lemma shows that the minimizer of the rate function in Theorem 9 is λ_τ. Furthermore, the constant term L $(=\lim_{n\to\infty} n^{-2}\log Z_n)$ is given as

$$L=-\frac{1}{2}-\frac{1}{2}\int Q(\zeta)\,d\lambda_\tau(\zeta)=-\frac{3}{4},$$

and the proof of Theorem 9 is completed.

It is worthwile to note that the free entropy of the elliptic law λ_τ is $\Sigma(\lambda_\tau)=-1/4$, independently of parameter τ.

Finally we are going to turn to unitary random matrices and we state a large deviation theorem on the unit circle $\mathbb{T}$.

THEOREM 17. *Let $\lambda>0$, and P_n $(n\in\mathbb{N})$ be the empirical eigenvalue distribution of the $n\times n$ random unitary matrix having density*

$$\frac{1}{Z_n}\exp(-n\mathrm{Tr}\,Q(U)),\quad Q(\zeta):=-\frac{2}{\lambda}\mathrm{Re}\,\zeta\quad(\zeta\in\mathbb{T})$$

with respect to the Haar probability measure. Then the finite limit

$$L:=\lim_{n\to\infty}\frac{1}{n^2}\log Z_n$$

exists and P_n satisfies the large deviation principle with good rate function

$$I(\mu):=-\Sigma(\mu)-\frac{2}{\lambda}\int \mathrm{Re}\,\zeta\,d\mu(\zeta)+L\quad\text{for}\quad \mu\in\mathcal{M}(\mathbb{T}).$$

The unique minimizer ρ_λ of the rate function and the value of L were computed by Gross and Witten [7] earlier:

$$\rho_\lambda=\begin{cases}\dfrac{1}{2\pi}\left(1+\dfrac{2}{\lambda}\cos\theta\right)d\theta & \text{if } \lambda\ge 2,\\[2ex] \dfrac{2}{\pi\lambda}\cos\dfrac{\theta}{2}\sqrt{\dfrac{\lambda}{2}-\sin^2\dfrac{\theta}{2}}\,\chi_{[-a,a]}(\theta)\,d\theta & \text{if } 0<\lambda<2,\end{cases}$$

where $\zeta=e^{\mathrm{i}\,\theta}$ $(-\pi\le\theta\le\pi)$ and $a:=2\arcsin\sqrt{\lambda/2}$, and

$$L=\begin{cases}\dfrac{1}{\lambda^2} & \text{if } \lambda\ge 2,\\[2ex] \dfrac{1}{2}\log\dfrac{\lambda}{2}+\dfrac{2}{\lambda}-\dfrac{3}{4} & \text{if } 0<\lambda<2.\end{cases}$$

More details about the minimization of the rate function are found also in [8, 10].

Theorem 17 generalizes to any real continuous function Q on $\mathbb{T}$. The details are found in [9].

Acknowledgement. The authors acknowledge financial supports from the grants OTKA T016924 and AKP 96/2-6782.

References

[1] L. Arnold, On the asymptotic distribution of the eigenvalues of random matrices, *J. Math. Anal. Appl.* **20** (1967), 262–268.

[2] G. Ben Arous and A. Guionnet, Large deviation for Wigner's law and Voiculescu's noncommutative entropy, *Prob. Theory Rel. Fields* **108** (1997), 517–542.

[3] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Jones and Bartlett, Boston-London, 1993.

[4] J.D. Deuschel and D.W. Stroock, *Large Deviations*, Academic Press, 1989.

[5] V.L. Girko, Elliptic law, *Theory Probab. Appl.* **30** (1986), 677–690.

[6] V.L. Girko, Elliptic law: ten years later I, II, *Random Oper. and Stoch. Equ.* **3** (1995), 257–302, 377–398.

[7] D.J. Gross and E. Witten, Possible third-order phase transition in the large-N lattice gauge theory, *Phys. Rev.* **D21** (1980), 446–453.

[8] F. Hiai and D. Petz, Maximizing free entropy, Preprint No. 17/1996, Mathematical Institute HAS, Budapest

[9] F. Hiai and D. Petz, A large deviation theorem for the empirical eigenvalue distribution of random unitary matrices, Preprint No. 17/1997, Mathematical Institute HAS, Budapest

[10] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy*, in preparation

[11] C.R. Hwang, A brief survey on the spectral radius and the spectral distribution of large dimensional random matrices with iid entries, in *Random Matrices and Their Applications*, M.L. Mehta (ed.), Contemp. Math. **50**, Amer. Math. Soc., Providence, 1986, 145–152.

[12] N.S. Landkof, *Foundations of Modern Potential Theory* (translated from the Russian), Springer, Berlin-Heidelberg-New York, 1972.

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

[14] M.L. Mehta, *Random Matrices*, Academic Press, Boston, 1991.

[15] H.N. Mhaskar and E.B. Saff, Weighted analogues of capacity, transfinite diameter, and Chebyshev constant, *Constr. Approx.* **8** (1992), 105–124.

[16] N.I. Muskhelishvili, *Some Basic Problems of the Mathematical Theory of Elasticity*, Noordhoff, Groningen, 1963.

[17] H.J. Sommers, A. Crisanti, H. Sompolinsky and Y. Stein, Spectrum of large random asymmetric matrices, *Phys. Rev. Lett.* **60** (1988), 1895–1898.

[18] R. Speicher, Free convolution and the random sum of matrices. *Publ. Res. Inst. Math. Sci.* **29** (1993), 731-744.

[19] H. Stahl and V. Totik, *General Orthogonal Polynomials*, Cambridge Univ. Press, 1992.

[20] V. Totik, *Weighted Approximation with Varying Weight*, Lecture Notes in Math., **1569**, Springer, Berlin, 1994.

[21] D. Voiculescu, Symmetries of some reduced free product C*-algebras, in *Operator Algebras and Their Connection with Topology and Ergodic Theory*, Lect. Notes in Math. **1132**, Springer, 1985, 556–558.

[22] D. Voiculescu, Limit laws for random matrices and free products, *Invent. Math.* **104** (1991), 201–220.

[23] D. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory, I, *Comm. Math. Phys.* **155** (1993), 71–92.

[24] D. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory, II, *Invent. Math.* **118** (1994), 411–440.

[25] D.V. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory, IV: maximum entropy and freeness, in *Free Probability Theory*, D.V. Voiculescu (ed.), Fields Inst. Comm., Vol. 12, Amer. Math. Soc., 1997, 293–302.

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

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

Department of Mathematical Sciences, Ibaraki University, Mito, Ibaraki 310, Japan