Math. Res. Lett. **14** (2007), *no.* 3, 10001–10010

TALAGRAND INEQUALITY FOR THE SEMICIRCULAR LAW AND ENERGY OF THE EIGENVALUES OF BETA ENSEMBLES

Ionel Popescu

Abstract. We give a short proof of an extension of the free Talagrand transportation cost inequality to the semicircular which was originally proved in [1]. The proof is based on a convexity argument and is in the spirit of the original Talagrand's approach for the classical counterpart from [8]. We also discuss the convergence, fluctuations and large deviations of the energy of the eigenvalues of β ensembles, which, as an application of Talagrand inequality gives in particular yet another proof of the convergence of the eigenvalue distribution to the semicircle law.

1. Introduction

In [8], Talagrand proves the transportation cost inequality to the Gaussian measure. The one dimensional version for the Gaussian measure $\gamma(dx) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$ reads as

$$(W_2(\mu, \gamma))^2 \leq 2H(\nu|\gamma), \tag{1.1}$$

where $W_2(\mu, \gamma)$ is the Wasserstein distance defined below by (2.2) and the relative entropy is

$$H(\nu|\gamma) = \begin{cases} \int f(x) \log(f(x)) d\gamma(x) & \text{if } \nu(dx) = f(x)\gamma(dx) \\ \infty & \text{if } \nu \text{ is singular to } \gamma. \end{cases}$$

In the context of free probability, Biane and Voiculescu proved in [1] a free version of this:

$$(W_2(\mu, \sigma))^2 \leq 2(E(\mu) - E(\sigma)), \tag{1.2}$$

where $E(\mu) = \frac{1}{2} \int x^2 \mu(dx) - \iint \log(|x - y|)\mu(dx)\mu(dy)$ is the free energy of μ and $\sigma(dx) = \frac{1}{2\pi} \mathbb{1}_{[-2,2]}(x)\sqrt{4 - x^2}dx$ is the semicircular law, the minimizer of $E(\mu)$ over all probability measures on the real line. The role of the relative entropy is played here by the difference of the free energy of μ and the semicircular.

Using random matrix approximations, Hiai, Petz and Ueda proved in [7] the following extension of (1.2),

$$\rho \left(W_2(\mu, \mu_Q)\right)^2 \leq E^Q(\mu) - E^Q(\mu_Q) \tag{1.3}$$

where $\rho > 0$ and $Q : \mathbb{R} \to \mathbb{R}$ is a function so that $Q(x) - \rho x^2$ is convex and

$$E^Q(\mu) = \int Q(x)\mu(dx) - \iint \log|x - y|\mu(dx)\mu(dy).$$

Received by the editors November 14, 2006.

Here μ_Q is the minimizer of E^Q on the set of all probability measures on the real line. They also prove a version of this for measures supported on the circle $\mathbf{T}$:

$$(\rho+1/4)\left(W_2(\nu,\nu_Q)\right)^2 \le E^Q(\nu)-E^Q(\nu_Q) \tag{1.4}$$

where $Q:\mathbf{T}\to\mathbb{R}$ so that $Q(e^{ix})-\rho x^2$ is convex on $\mathbb{R}$, $\rho>-1/4$ and μ_Q is the minimizer of the functional E^Q on probability measures on the unit circle $\mathbf{T}$.

Another proof of (1.2) is given in [5] via a Brunn-Minkovsky inequality for free probability.

The primary purpose of this note is to give an elementary proof of (1.3) and (1.4) in the spirit of Talagrand's proof to (1.1). The idea is to exploit convexity of the logarithm appearing in the E^Q. We also discuss (see Theorem 2.16 and Proposition 2.20) the discrete version of the transportation cost inequalities and some consequences involving Fekete points.

The second purpose of this note is to discuss the energy of the eigenvalues of β ensembles and in particular the fluctuations and the deviations from the minimum energy (see Theorem 3.1). This is a simple application of Selberg's formula together with elementary estimates on Γ functions. As a consequence, using the the results in the first part we reprove that the distribution of the eigenvalues converges almost surely to the semicircular law.

2. Talagrand Inequalities

The following result is an obvious one but is the key to our problem.

Lemma 2.1. *Let $f:[0,1]\to\mathbb{R}$ be a convex function with the property that $f(0)=0$ and there exists $a\ge 0$ so that*

$$f(t)\ge -at^2 \quad \textit{for} \quad t\in[0,1].$$

Then

$$f(t)\ge 0 \quad \textit{for all} \quad t\in[0,1].$$

Proof. It follows from the assumptions that for any $\epsilon>0$, if $\delta_\epsilon=\min(1,\epsilon/a)$, then $f(t)\ge -t\epsilon$ for $t\in[0,\delta_\epsilon]$. Now, since f is convex, one gets $f(mt)\ge mf(t)\ge -mt\epsilon$ for any integer m with $mt\le 1$, and therefore, $f(t)\ge -\epsilon t$ for any $t\in[0,1]$. Since this is true for any $\epsilon>0$, we get $f(t)\ge 0$ for any $t\in[0,1]$. □

In the following, $\mathcal{P}(\Omega)$ denotes the set of all probability measures on Ω, and for two probability measures with finite second moment on $\mathcal{P}(\mathbb{R})$ or $\mathcal{P}(\mathbf{T})$, where $\mathbf{T}=\{z\in\mathbb{C}:|z|=1\}$, we define $W_2(\mu,\nu)$, the Wasserstein distance by

$$W_2(\mu,\nu):=\sqrt{\inf_{\pi\in\Pi(\mu,\nu)}\iint |x-y|^2 d\pi(x,y)}. \tag{2.2}$$

Here $\Pi(\mu,\nu)$ is the set of probability measures on $\mathbb{R}^2$ with marginal distributions μ and ν, and it can be shown that there is at least one solution $\pi\in\Pi(\mu,\nu)$ to this minimization problem.

If μ and ν are two measures on $\mathbb{R}$ with F and G their cumulative distribution functions (i.e. $F(x)=\mu((-\infty,x])$), then Theorem 2.18 in [9] states that

$$(W_2(\mu,\nu))^2=\int_0^1 |F^{-1}(t)-G^{-1}(t)|^2 dt \tag{2.3}$$

where F^{-1} denotes the generalized inverse of F.

Theorem 2.4. *Let $Q : \mathbb{R} \to \mathbb{R}$ be a function so that $Q(x) - \rho x^2$ is convex for a certain $\rho > 0$. If μ_Q is a solution to the minimization problem*

$$I^Q := \inf_{\mu \in \mathcal{P}(\mathbb{R})} E^Q(\mu), \tag{2.5}$$

where

$$E^Q(\mu) = \int Q(x)\mu(dx) - \iint \log|x-y|\mu(dx)\mu(dy), \tag{2.6}$$

then for any $\mu \in \mathcal{P}(\mathbb{R})$, we have

$$\rho \left(W_2(\mu, \mu_Q)\right)^2 \leq E^Q(\mu) - I^Q. \tag{2.7}$$

In particular, the minimization problem (2.5) *has a unique solution.*

Proof. There exist constants c_1 and c_2 so that

$$Q(x) - \rho x^2 \geq c_1 \quad \text{and} - \log(|x-y|) \geq -\frac{\rho}{4}(x^2+y^2) + c_2.$$

Then for a certain C, we get that

$$\frac{1}{2}(Q(x)+Q(y)) - \log(|x-y|) \geq \frac{\rho}{4}(x^2+y^2) + C \geq C, \tag{2.8}$$

and this in turn implies that the infimum in (2.5) is finite (since $E^Q(\mu)$ is finite for μ the uniform distribution on $[0,1]$) and in particular $\int Q(x)d\mu_Q(x)$, and $\iint \log|x-y|d\mu_Q(x)d\mu_Q(y)$ are finite, which means that μ_Q has finite second moment and no atoms.

Since $E^Q(\mu) > -\infty$, we may assume that $E^Q(\mu)$ is finite, otherwise there is nothing to prove. Then, $\iint \log|x-y|\mu(dx)\mu(dy)$ and $\int Q(x)\mu(dx)$ are finite. In particular, μ has finite second moment and no atoms.

Taking F_μ and F_{μ_Q}, the cumulative distributions of μ, μ_Q and F^{-1}, F_Q^{-1} their generalized inverses, set $\theta(x) = F^{-1}(F_Q(x))$. According to [9, Theorem 2.18] and the discussion following thereafter, the minimizing measure π from (2.2) is the distribution of $x \to (x, \theta(x))$ under μ_Q. In this case, the inequality we want to prove becomes

$$\rho \iint |x-\theta(x)|^2\mu_Q(dx) \leq \int Q(x)\mu(dx) - \iint \log|x-y|\mu(dx)\mu(dy) - I^Q.$$

Let $f : [0,1] \to \mathbb{R}$ be given by

$$\begin{aligned} f(t) = -\rho t^2 \int |\theta(x)-x|^2\mu_Q(dx) + \int Q(t\theta(x)+(1-t)x)\mu_Q(dx) \\ - \iint \log(|t(\theta(x)-\theta(y)) + (1-t)(x-y)|)\mu_Q(dx)\mu_Q(dy) - I^Q. \end{aligned}$$

Notice here that f is well defined. Indeed, Q is convex, hence bounded below and because $\int Q(\theta(x))\mu_Q(dx) = \int Q(x)\mu(dx)$ and $\int Q(x)\mu_Q(dx)$ are both finite, one concludes that $\int Q(t\theta(x)+(1-t)x)\mu_Q(dx)$ is finite too. One the other hand, there is a $C > 0$ so that for any $t \in [0,1]$,

$$-\log(|t(\theta(x)-\theta(y)) + (1-t)(x-y)|) \geq -C(\theta(x)^2+\theta(y)^2+x^2+y^2) - C,$$

which, combined with the finiteness of the second moment of μ and μ_Q, results with (for a constant C)

$$-\iint \log(|t(\theta(x)-\theta(y))+(1-t)(x-y)|)\mu_Q(dx)\mu_Q(dy) > C \quad \text{for all} \quad t\in[0,1].$$

Now, since θ is a nondecreasing function we can write

$$-\iint \log(|t(\theta(x)-\theta(y))+(1-t)(x-y)|)\mu_Q(dx)\mu_Q(dy) = \\ -2\iint_{x>y} \log(t(\theta(x)-\theta(y))+(1-t)(x-y))\mu_Q(dx)\mu_Q(dy),$$

which combined with the convexity of $-\log$ on $(0,\infty)$ and the finiteness of $\iint \log|x-y|\mu_Q(dx)\mu_Q(dy)$ and $\iint \log|x-y|\mu(dx)\mu(dy)$, yields the fact that

$$(**)\qquad t\to -\iint \log(|t(\theta(x)-\theta(y))+(1-t)(x-y)|)\mu_Q(dx)\mu_Q(dy)$$

is well defined and convex.

The inequality (2.7) is now equivalent to $f(1)\geq 0$. To show this, we apply Lemma 2.1. The convexity follows easily from the convexity of $Q(x)-\rho x^2$ and (**). Now if ν_t is the distribution of $x\to t\theta(x)+(1-t)x$ under μ_Q, then the minimization property of μ_Q implies that

$$f(t)\geq -\rho t^2\iint |\theta(x)-x|^2\mu_Q(dx) \quad \text{for} \quad t\in[0,1],$$

and then, Lemma 2.1 shows that $f(t)\geq 0$ for any $t\in[0,1]$.

The existence statement follows from the lower continuity of E^Q. For a proof of the existence and compactness of the support of μ_Q, see for instance Chapter 6 in [2]. □

Remark 2.9. *What was essential during the proof was the convexity of* $-\log$ *on* $(0,\infty)$ *and the fact that for any* $a>0$*, there is a* $C(a)$ *so that* $-\log|x-y|\geq -a(x^2+y^2)+C(a)$*. Therefore if we replace the* log *in the statement of this theorem by any kernel* $K(|x-y|)$ *with the property that* K *on* $(0,\infty)$ *is concave and that for any* $a>0$*, there is a* $C(a)$ *so that* $-K(|x-y|)\geq -a(x^2+y^2)+C(a)$*, then the result still holds. Other examples of such kernels are* $-1/x^\alpha$*,* $\alpha>0$ *and* $1/\log(x^2+1)$*.*

If we take $Q(x)=\frac{x^2}{2}$, and keep in mind that the minimizing measure μ_Q for E^Q is the semicircular law, one gets the following result proved in [1].

Corollary 2.10. *Let* $\sigma(dx)=\frac{1}{2\pi}\mathbb{1}_{[-2,2]}(x)\sqrt{4-x^2}dx$ *be the semicircular law on* $[-2,2]$*. Then for any* $\mu\in\mathcal{P}(\mathbb{R})$*,*

$$\frac{1}{2}(W_2(\mu,\sigma))^2\leq \frac{1}{2}\int x^2\mu(dx)-\iint \log(|x-y|)\mu(dx)\mu(dy)-\frac{3}{4}.$$

The next theorem is just inequality (1.4).

Theorem 2.11. *Assume* $Q:\mathbf{T}\to\mathbb{R}$ *is a function so that* $Q(e^{ix})-\rho x^2$ *is convex on* $\mathbb{R}$ *for a given* $\rho>-1/4$*. If* μ_Q *is a solution to the minimization problem*

$$(2.12)\qquad I^Q:=\inf_{\mu\in\mathcal{P}(\mathbf{T})} E^Q(\mu),$$

where

$$E^Q(\nu) = \int Q(z)\nu(dz) - \iint_{\mathbf{T}\times\mathbf{T}} \log|z-z'|\nu(dz)\nu(dz'), \tag{2.13}$$

then, for any $\nu \in \mathbf{T}$, *we have*

$$(\rho + 1/4)\,(W_2(\nu,\nu_Q))^2 \leq E^Q(\nu) - I^Q. \tag{2.14}$$

In particular, there is a unique solution for the minimization problem (2.12).

Proof. We identify $[-\pi,\pi)$ with $\mathbf{T}$ via the exponential map $x \to e^{ix}$ and move the measure ν to μ and ν_Q to μ_Q. We then follow the proof of 2.4 with the necessary adjustments needed. The function $f(t)$ there becomes here

$$f(t) = -(\rho+1/4)t^2 \int |\theta(x)-x|^2\nu_Q(dx) + \int Q(e^{i(t\theta(x)+(1-t)x)})\nu_Q(dx)$$
$$- \iint \log(|e^{i(t(\theta(x)+(1-t)x))} - e^{i(t\theta(y)+(1-t)(y))}|)\nu_Q(dx)\nu_Q(dy) - I^Q.$$

Now, $|e^{ia}-e^{ib}|^2 = 4\sin^2((a-b)/2)$ for a, b real numbers and

$$\int |\theta(x)-x|^2\nu_Q(dx) = \frac{1}{2}\iint ((\theta(x)-x)-(\theta(y)-y))^2\nu_Q(dx)\nu_Q(dy).$$

Next, set $\theta_t(x) = t\theta(x) + (1-t)x$ and notice that

$$\begin{aligned} g(t) :=& -\frac{t^2}{4}\int |\theta(x)-x|^2\nu_Q(dx) - \iint \log(|e^{it\theta_t(x))} - e^{i\theta_t(y))}|)\nu_Q(dx)\nu_Q(dy) \\ =& -\iint \frac{t^2}{8}((\theta(x)-x)-(\theta(y)-y))^2\nu_Q(dx)\nu_Q(dy) \\ & - \iint \log|2\sin((\theta_t(x)-\theta_t(y))/2)|\,\nu_Q(dx)\nu_Q(dy) \\ =& -2\iint_{x>y} \frac{t^2}{8}((\theta(x)-x)-(\theta(y)-y))^2\nu_Q(dx)\nu_Q(dy) \\ & -2\iint_{x>y} \log(2\sin((\theta_t(x)-\theta_t(y))/2))\,\nu_Q(dx)\nu_Q(dy), \end{aligned}$$

where in the last line we used the fact that θ is a nondecreasing function. Since $x, y, \theta(x), \theta(y) \in [-\pi,\pi)$ and for $0 < a < b < \pi$, we have

$$\begin{aligned} &\frac{d^2}{dt^2}\left(-\frac{t^2}{8}(a-b)^2 - \log\left(\sin\left(\frac{ta+(1-t)b}{2}\right)\right)\right) \\ &\quad = \frac{(a-b)^2}{4}\left(\frac{1}{\sin^2\left(\frac{ta+(1-t)b}{2}\right)} - 1\right) \geq 0, \end{aligned}$$

which implies that the function g is convex on $[0,1]$. This coupled with the convexity of $Q(e^{ix}) - \rho x^2$ concludes that f is a convex function. Finally

$$f(t) \geq -(\rho+1/4)t^2 \int |\theta(x)-x|^2\nu_Q(dx),$$

and thus, Lemma 2.1 shows that $f(1) \geq 0$, which is (2.14).

The existence of a minimizer follows from the fact that E^Q is lower semicontinuous. □

For $Q = 0$ and $\rho = 0$, the minimizer of (2.12) is the Haar measure on $\mathbf{T}$. One can check this by showing directly that the uniform measure satisfy the variational form of (2.12).

Corollary 2.15. *For any $\mu \in \mathcal{P}(\mathbf{T})$*

$$\frac{1}{4}\left(W_2\left(\mu, \frac{dx}{2\pi}\right)\right)^2 \leq -\iint_{\mathbf{T}\times\mathbf{T}} \log|z - z'|\mu(dz)\mu(dz').$$

Using the same argument as in the proof of Theorem 2.4, we can also prove a discrete version of it.

Theorem 2.16. *Let $Q : \mathbb{R} \to \mathbb{R}$ be a function so that $Q(x) - \rho x^2$ is convex for a certain $\rho > 0$. For $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, set the energy of $\mathbf{x}$ to be given by*

$$E_n^Q(\mathbf{x}) = \frac{1}{n}\sum_{k=1}^n Q(x_i) - \frac{2}{n(n-1)}\sum_{1\leq i<j\leq n} \log|x_i - x_j|.$$

If $\Delta_n^Q = E_n^Q(\mathbf{y}) = \inf\{E_n^Q(\mathbf{x}) : x \in \mathbb{R}^n\}$, then for any $\mathbf{x} \in \mathbb{R}^n$,

$$\rho(W_2(\mu(\mathbf{x}), \mu(\mathbf{y})))^2 \leq E_n^Q(\mathbf{x}) - E_n^Q(\mathbf{y}) = E_n^Q(\mathbf{x}) - \Delta_n^Q \tag{2.17}$$

where $\mu(\mathbf{x}) = \frac{1}{n}\sum_{k=1}^n \delta_{x_k}$. Moreover,

$$\Delta_n^Q \leq \Delta_{n+1}^Q. \tag{2.18}$$

The only statement that needs to be clarified here is (2.18). If $\mathbf{y}_{n+1}$ is a minimum point for E_{n+1}^Q and $\mathbf{y}_{n+1}^i$ denotes the n dimensional vector obtained from $\mathbf{y}_{n+1}$ by removing the ith component, then $\Delta_{n+1}^Q = \frac{1}{n+1}\sum_{i=1}^{n+1} E_n^Q(\mathbf{y}_{n+1}^i)$, which is obviously $\geq \Delta_n^Q$.

The minimum points of E_n^Q are called Fekete points in the literature. It is known (see for instance chapter 6 in [2]) that $\lim_{n\to\infty} \Delta_n^Q = I^Q$, with I^Q defined in (2.5). We will reprove this fact below in Proposition 2.20.

For $Q(x) = x^2$, the formula [6, A.6.11] with the appropriate scaling gives the formula for computing $\Delta_n = \Delta_n^Q$ as

$$\Delta_n = \frac{1}{2}(1+\log(n-1)) - \frac{1}{n(n-1)}\sum_{j=1}^n j\log j = \frac{1}{2} - \frac{\log n}{n-1} - \frac{1}{n}\sum_{j=1}^{n-1}\frac{j}{n-1}\log\left(\frac{j}{n-1}\right). \tag{2.19}$$

The next statement is a similar result to Theorems 2.4 and 2.16.

Proposition 2.20. *Assume $Q : \mathbb{R} \to \mathbb{R}$ is a function so that $Q(x) - \rho x^2$ is convex for a certain $\rho > 0$. Then for any $\nu \in \mathcal{P}(\mathbb{R})$ and $\mathbf{y} \in \mathbb{R}^n$ a Fekete point for E_n^Q, we have*

$$\rho(W_2(\nu, \mu(\mathbf{y})))^2 \leq E^Q(\nu) - \Delta_n^Q. \tag{2.21}$$

Furthermore, if μ_Q is the minimizing measure of E^Q, and $\mathbf{y}_n \in \mathbb{R}^n$ is a Fekete point for E_n^Q, then

$$\lim_{n\to\infty} \Delta_n^Q = I^Q \quad \text{and} \quad \lim_{n\to\infty} W_2(\mu_Q, \mu(\mathbf{y}_n)) = 0, \tag{2.22}$$

hence, $\mu(\mathbf{y}_n) \underset{n\to\infty}{\longrightarrow} \mu_Q$ weakly.

Proof. In the first place there is nothing to prove if $E^Q(\nu) = \infty$. Therefore we assume that $E^Q(\nu) < \infty$. Integrating (2.17) with respect to $\nu(dx_1)\nu(dx_2)\dots\nu(dx_n)$, one gets that

$$\rho \int (W_2(\mu(\mathbf{x}), \mu(\mathbf{y})))^2 \nu(dx_1)\nu(dx_2)\dots\nu(dx_n) \le E^Q(\nu) - \Delta_n^Q.$$

We finish the proof of (2.21) by showing that

$$\int (W_2(\mu(\mathbf{x}), \mu(\mathbf{y})))^2 \nu(dx_1)\nu(dx_2)\dots\nu(dx_n) = (W_2(\nu, \mu(\mathbf{y})))^2. \tag{*}$$

To do this, we proceed by induction. For $n = 1$, this statement becomes

$$\int (W_2(\delta_x, \delta_y))^2 \nu(dx) = (W_2(\nu, \delta_y))^2$$

which, cf. (2.3), is equivalent to the following (here F_ν is the cumulative distribution function of ν)

$$\int |x-y|^2 \nu(dx) = \int_0^1 |y - F_\nu^{-1}(t)|^2 dt.$$

This can be checked by changing the variable in the second integral.

Assume (*) is true for $n-1$, $n \ge 2$. A simple application of (2.3) gives that $(W_2(\mu(\mathbf{x}), \mu(\mathbf{y})))^2 = \frac{1}{n}\sum_{i=1}^n |x_{\sigma(i)} - y_{\tau(i)}|^2$, where σ and τ are permutations of $\{1, 2, \dots, n\}$ so that $x_{\sigma(1)} \le x_{\sigma(2)} \dots \le x_{\sigma(n)}$ and $y_{\tau(1)} \le y_{\tau(2)} \dots \le y_{\tau(n)}$. If we denote by $\mathbf{x}_i$ the vector $\mathbf{x}$ with the ith component removed and similarly for $\mathbf{y}_i$, one deduces

$$(W_2(\mu(\mathbf{x}), \mu(\mathbf{y})))^2 = \frac{1}{n}\sum_{i=1}^n (W_2(\mu(\mathbf{x}_i), \mu(\mathbf{y}_i)))^2. \tag{\#}$$

On the other hand,

$$(W_2(\nu, \mu(\mathbf{y})))^2 = \sum_{k=0}^{n-1} \int_{k/n}^{(k+1)/n} |y_{\tau(k)} - F_\nu^{-1}(t)|^2 dt,$$

which can be used to argue that

$$(W_2(\nu, \mu(\mathbf{y})))^2 = \frac{1}{n}\sum_{i=1}^n (W_2(\nu, \mu(\mathbf{y}_i)))^2. \tag{\#\#}$$

Putting together (#) and (##) and the induction hypothesis one finishes the proof of (*).

To prove (2.22), we first point out that (2.21) applied to μ_Q yields that $I^Q \ge \Delta_n^Q$ for any $n \ge 1$. In particular this means that Δ_n^Q is bounded. Since $-\log|x-y| \ge -\frac{\rho}{4}(x^2+y^2) + c$ for a certain constant c, we get that $\Delta_n^Q \ge \frac{\rho}{4n}\sum_{i=1}^n x_i^2 - C$, where C is a constant. This implies that the sequence $\{\int x^2 \mu(\mathbf{y}_n)(dx)\}_{n\ge1}$ is bounded, whose consequence is that the sequence of measures $\mu(\mathbf{y}_n)$ is tight, therefore there is a weak convergent subsequence $\mu(\mathbf{y}_{n_k})$ to a measure ν. Now, for any $L>0$, we have

$$\int \min\{((Q(x)+Q(y))/2-\log|x-y|),L\}\mu(\mathbf{y}_{n_k})(dx)\mu(\mathbf{y}_{n_k})(dy)\le \Delta^Q_{n_k}+L/n_k$$

and this demonstrates that for any $L>0$,

$$\int \min\{((Q(x)+Q(y))/2-\log|x-y|),L\}\nu(dx)\nu(dy)\le I^Q$$

and, after passing $L\to\infty$, this yields

$$E^Q(\nu)\le I^Q.$$

This together with (2.18) and the uniqueness of μ_Q from Theorem 2.4 ends the proof of $\lim_{n\to\infty}\Delta^Q_n=I^Q$. The rest follows. □

3. Discrete Energy for β-Ensembles

In this section we deal with β-ensembles, which are studied in [4]. These are tridiagonal matrices with independent entries of the form

$$A_n=\frac{1}{\sqrt{\beta n}}\begin{bmatrix} N(0,2) & \chi_{(n-1)\beta} & & & \\ \chi_{(n-1)\beta} & N(0,2) & \chi_{(n-2)\beta} & & \\ & \ddots & \ddots & \ddots & \\ & & \chi_{2\beta} & N(0,2) & \chi_\beta \\ & & & \chi_\beta & N(0,2)\end{bmatrix}.$$

Here $N(0,2)$ stands for a normal with mean 0 and variance 2, while χ_γ is the χ-distribution with parameter γ. The joint distribution of the eigenvalues is

$$\frac{1}{Z_{\beta,n}}\prod_{1\le i<j\le n}|x_i-x_j|^\beta \exp\left(-\beta n\sum_{i=1}^n x_i^2\right)$$

where here $Z_{\beta,n}$ is a normalization constant.

Set $\mu_n=\sum_{k=1}^n\delta_{\lambda_{k,n}}$, the empirical distribution of the eigenvalues $\{\lambda_{k,n}\}_{k=1}^n$ of A_n.

Theorem 3.1. *Set* $E_n=\frac{1}{2n}\sum_{k=1}^n\lambda_k^2-\frac{2}{(n-1)n}\sum_{1\le j<k\le n}\log|\lambda_i-\lambda_j|$ *the energy of the eigenvalues* $\{\lambda_k\}_{k=1}^n$ *of* A_n*. If* Δ_n *is the quantity defined in* (2.19)*, then almost surely,*

$$\lim_{n\to\infty} n(E_n-\Delta_n)=\psi(1+\beta/2)-\log(\beta/2), \tag{3.2}$$

where $\psi(x)=\frac{d}{dx}\log\Gamma(x)$ *and* Γ *is the Gamma function. In addition, we have that*

$$n^{1/2}\Big(n(E_n-\Delta_n)-(\psi(1+\beta/2)-\log(\beta/2))\Big)\xrightarrow[n\to\infty]{} N(0,\psi'(1+\beta/2)), \tag{3.3}$$

where the convergence is in distribution sense.

The large deviations of $n(E_n-\Delta_n)$ *is governed by the rate function*

$$R^*(t)=\sup\{tz-R(z):z\in\mathbb{R}\},$$

$$R(z) = \begin{cases} z + (\beta/2 - z)\log(\beta/2 - z) - \log\left(\frac{\Gamma(1+\beta/2-z)}{\Gamma(1+\beta/2)}\right) - (\beta/2)\log(\beta/2), & z < \beta/2 \\ \infty, & z \geq \beta/2. \end{cases}$$

Proof. The proof is based on a version of Selberg's formula and elementary approximations involving Gamma function.

First, we have

$$\mathbb{E}\left[\exp(z\mathbb{E}_n)\right] = \frac{\int_{\mathbb{R}^n} \prod_{1\leq i<j\leq n} |x_i - x_j|^{\beta - \frac{2z}{n(n-1)}} \exp\left(-(\beta n - \frac{z}{2n})\sum_{j=1}^n x_j^2\right) d\mathbf{x}}{\int_{\mathbb{R}^n} \prod_{1\leq i<j\leq n} |x_i - x_j|^{\beta} \exp\left(-\beta n\sum_{j=1}^n x_j^2\right) d\mathbf{x}}$$

and then, as a consequence of Selberg's formula [6, equation 17.6.7], we get for complex z, that

$$\mathbb{E}\left[e^{z\mathbb{E}_n}\right] = \begin{cases} \frac{(n\beta/2 - z/n)^{-\frac{n}{2}\left[(n-1)(\beta/2 - \frac{z}{n(n-1)}+1)\right]} \prod_{j=1}^n \frac{\Gamma(1+j(\beta/2 - \frac{z}{n(n-1)}))}{\Gamma(1+\beta/2)}}{(n\beta/2)^{-\frac{n}{2}[(n-1)\beta/2+1]} \prod_{j=1}^n \frac{\Gamma(1+j\beta/2)}{\Gamma(1+\beta/2)}}, & \Re(z) < \beta/2 \\ \infty, & \Re(z) \geq \beta/2. \end{cases}$$

We need Stirling formula for approximation of Gamma function in the following form

$$\log\Gamma(t+1) = (t+1/2)\log t - t + \log(2\pi)/2 + \mathcal{O}\left(\frac{1}{1+t}\right) \text{ for } t \geq 0$$

Using this and the above formula for $\mathbb{E}[\exp(zE_n)]$ and (2.19), after some arrangements one gets

$$\begin{aligned} &\log(\mathbb{E}\left[e^{z(E_n - \Delta_n)}\right]) = \frac{z}{n-1} + \frac{z}{2}\log\left(1 + \frac{1}{n-1}\right) \\ &- \frac{z}{n-1}\log\left(\frac{\beta}{2} - \frac{z}{n^2}\right) + \frac{z(n+1)}{2(n-1)}\log\left(1 + \frac{z}{n[(n-1)n\beta/2 - z]}\right) \\ &+ \frac{n[(n-1)\beta+1]}{2}\log\left(1 - \frac{z}{(n-1)[n^2\beta/2 - z]}\right) + \frac{n\beta}{2}\log\left(1 - \frac{2z}{n(n-1)\beta}\right) \\ &- n\left[\log\left(1 + \frac{\beta}{2} - \frac{z}{n(n-1)}\right) - \log\left(1 + \frac{\beta}{2}\right)\right] + \mathcal{O}\left(\frac{z}{n^2}\right). \end{aligned} \tag{3.4}$$

From this, replacing z by nz, one immediately obtains that for any $z \in \mathbb{R}$,

$$\log(\mathbb{E}[\exp(zn(E_n - \Delta_n))]) \underset{n\to\infty}{\longrightarrow} z\frac{\Gamma'(1+\beta/2)}{\Gamma(1+\beta/2)} - z\log(\beta/2) = z(\psi(1+\beta/2) - \log(\beta/2)).$$

Applying(3.4) with z replaced by $n^{3/2}z$, one can prove that for any complex z,

$$\log\left(\mathbb{E}\left[\exp\left(zn^{1/2}(n(E_n - \Delta_n) - (\psi'(1+\beta/2) - \log(\beta/2))\right)\right]\right) \underset{n\to\infty}{\longrightarrow} z^2\psi'(1+\beta/2)/2$$

whose consequence is (3.3). This, applied for $z = \pm 1$ together with Chebyshev inequality yields

$$P(|n(E_n - \Delta_n) - (\psi(1+\beta/2) - \log(\beta/2))| \geq \epsilon) \leq Ce^{-\epsilon n^{1/2}}$$

for a certain constant $C > 0$. This and an application of Borel-Cantelli's Lemma prove (3.2). Again applying (3.4) with $n^2 z$ in place of z, one can show that

$$\frac{1}{n}\log\left(\mathbb{E}\left[\exp\left(zn^2(E_n - \Delta_n)\right)\right]\right) \underset{n\to\infty}{\longrightarrow} R(z).$$

for any $z \in \mathbb{R}$. As a consequence of standard large deviations results (see for example Section 2.2 in [3]) we conclude the proof of the last part of the theorem. □

Corollary 3.5. *E_n converges almost surely to $3/4$, the energy of the semicircular law on $[-2,2]$. This implies that the spectral distribution, μ_n of A_n converges almost surely to the semicircular law on $[-2,2]$.*

Proof. The convergence of E_n to $3/4$ follows from (3.2) and the fact that the second expression in (2.19) converges to $1/2 - \int_0^1 x\log(x)dx = 3/4$. Alternatively, we can use Proposition 2.20 for the convergence of Δ_n to the free entropy of the semicircular law. For the converges of the spectral distribution, we use 2.16 and 2.20 with $Q(x) = x^2/2$ plus the triangle inequality to justify that almost surely

$$W_2(\mu_n,\sigma) \leq W_2(\mu_n,\mu(\mathbf{y}_n)) + W_2(\mu(\mathbf{y}_n),\sigma) \leq \sqrt{2(E_n - \Delta_n)} + \sqrt{2(3/4 - \Delta_n)} \underset{n\to\infty}{\longrightarrow} 0.$$

□

References

[1] P. Biane and D. Voiculescu, *A Free Probability Analogue of the Wasserstein Metric on a Trace-State Space*, GAFA - Geometric And Functional Analysis **11** (2001) 1125–1138.

[2] P. Deift, Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach, New York University-Courant Institute of Mathematical Sciences - AMS (2000).

[3] J.-D. Deuschel and D. W. Stroock, Large Deviations, AMS, revised edition (1989).

[4] I. Dumitriu and A. Edelman, *Matrix Models for Beta Ensembles*, Journal of Mathematical Physics **43** (2002), no. 11, 5830–5847.

[5] M. Ledoux, *A (one-dimensional) Free Brunn-Minkowski Inequality*, C. R. Acad. Sciences, Paris **340** (2005) 301–304.

[6] M. L. Mehta, Random Matrices, Academic Press, second edition edition (1991).

[7] F. H. D. Petz and Y. Ueda, *Free Transportation Cost Inequalities via Random Matrix Approximation*, Prob. Theory Rel. Fields **130** (2004) 199–221.

[8] M. Talagrand, *Transportation Cost for Gaussian and other Product Measures*, Geometric and Functional Analysis **6** (1996) 587–600.

[9] C. Villani, Topics in Optimal Transportation, Vol. 58 of *Graduate Studies in Mathematics*, AMS (2003).

Department of Mathematics, Northwestern University, 2033 Sheridan Road, Evanston, IL, 60208, USA

IMAR, 21 Calea Grivitei Street, 010702-Bucharest, Sector 1, Romania
E-mail address: ipopescu@math.northestern.edu