Pacific Journal of Mathematics

A PROPERTY OF FREE ENTROPY

S.T. Belinschi and H. Bercovici

Volume 211 No. 1 September 2003

PACIFIC JOURNAL OF MATHEMATICS
Vol. 211, No. 1, 2003

A PROPERTY OF FREE ENTROPY

S.T. Belinschi and H. Bercovici

We show that the restriction on the uniform norms of approximating matricial microstates can be removed when defining free entropy.

1. Introduction.

Denote by $\mathfrak{M}_k$ the algebra of complex $k \times k$ matrices, and by τ_k the normalized trace on $\mathfrak{M}_k$, i.e., $\tau_k(A) = \frac{1}{k}\mathrm{Tr}(A)$ for $A \in \mathfrak{M}_k$. Consider for each k a standard Gaussian Hermitian random matrix X_k. Thus, if E denotes expected value, $E\tau_k(X_k) = 0$ and $E\tau_k(X_k^2) = 1$. It was shown by E. Wigner [9] that, as $k \to \infty$, X_k tends in distribution to a semicircular law, i.e., the limits

$$\mu_p = \lim_{k\to\infty} E\tau_k(X_k^p)$$

exist, and they can be calculated as

$$\mu_p = \frac{1}{2\pi}\int_{-2}^{2} t^p \sqrt{4-t^2}\, dt$$

for $p = 1, 2, \ldots$. If we have several independent standard Gaussian Hermitian random matrices $(X_k(i))_{i=1}^n$, D. Voiculescu [4] proved that, as $k \to \infty$, these sets of variables converge in distribution to a free semicircular family. Briefly, this means that given indices $i_j \in \{1, 2, \ldots, n\}$ such that $i_j \neq i_{j+1}$ for $j = 1, 2, \ldots, m-1$, and given positive integers $p_1, p_2, \ldots, p_m$, the limit

$$\lim_{k\to\infty} E\tau_k(X_k(i_1)^{p_1} X_k(i_2)^{p_2} \ldots X_k(i_m)^{p_m})$$

exists, and

$$\lim_{k\to\infty} E\tau_k[(X_k(i_1)^{p_1} - \mu_{p_1})(X_k(i_2)^{p_2} - \mu_{p_2}) \ldots (X_k(i_m)^{p_m} - \mu_{p_m})] = 0.$$

It is natural to look for large deviation principles associated with these limit laws. For this purpose (and also with motivation from information theory and statistical physics) Voiculescu introduced in [6] (cf. also [5]) the notion of free entropy. The original definition of free entropy, which will be reviewed below, involves a bound $R > 0$ on the operator norm of approximating matricial microstates, and this may perhaps obscure its significance for large deviations. It is our purpose here to show that this bound can be removed — roughly speaking, one can set $R = \infty$ in the definition of free entropy. This result applies to other notions of free entropy which appeared subsequently (see for instance [7] for free entropy in the presence of additional variables, [8] for free entropy using an ultrafilter, and [3] for free entropy of a nonselfadjoint variable). We will only provide the proof for the original quantity χ defined in [6], but it should be obvious how the argument applies in the other situations.

It should be noted that a large deviation theorem for Wigner's result has been proved by G. Ben Arous and A. Guionnet [1], where the natural topology of weak convergence of probability measures on the real line is used. The rate function is closely related with free entropy. For several variables, a thorough study of large deviations was undertaken by T. Cabanal Duvillard and A. Guionnet [2]. The rate function they determine is related with another version of free entropy (microstate free).

2. The main result.

For the remainder of this note we fix a positive integer n. We will denote by I the collection of all multiindices $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ with $m \geq 1$ and $\alpha_j \in \{1, 2, \ldots, n\}$ for all $j = 1, 2, \ldots, m$. In other words, $I = \bigcup_{m=1}^{\infty} \{1, 2, \ldots, n\}^m$. A multiindex of the form $(\alpha, \alpha, \ldots, \alpha)$ will also be denoted α^m. We consider the space $\mathbb{S}$ consisting of all families $(\mu(\alpha))_{\alpha \in I}$ of complex numbers indexed by I. The space $\mathbb{S}$ will be endowed with the topology of componentwise convergence.

Consider now a tracial W^*-probability space $(\mathfrak{A}, \tau)$. That is, $\mathfrak{A}$ is a von Neumann algebra, and τ is a normal trace state on $\mathfrak{A}$. We will write $\mathfrak{A}^{\mathrm{sa}}$ for the space of selfadjoint elements of $\mathfrak{A}$. Given an n-tuple $X = (X_1, X_2, \ldots, X_n) \in (\mathfrak{A}^{\mathrm{sa}})^n$, its distribution $\mu_X \in \mathbb{S}$ is defined by

$$\mu_X(\alpha) = \tau(X_\alpha),$$

where $X_\alpha = X_{\alpha_1} X_{\alpha_2} \ldots X_{\alpha_m}$ for $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m) \in I$. This notation applies in particular to n-tuples of selfadjoint matrices in $\mathfrak{M}_k$. Voiculescu's entropy measures the extent to which the distribution of X can be approximated by distributions of the form μ_A with $A \in (\mathfrak{M}_k^{\mathrm{sa}})^n$. Note first that $\mathfrak{M}_k^{\mathrm{sa}}$ is a real Hilbert space with the Hilbert-Schmidt norm $\|A\|_2 = \mathrm{Tr}(A^2)$, and λ_k will denote the corresponding Lebesgue measure (i.e., a cube whose sides form an orthonormal basis has measure equal to one). On the space $(\mathfrak{M}_k^{\mathrm{sa}})^n$ we have the product measure $\lambda_k^{\otimes n}$.

Given $X \in (\mathfrak{A}^{\mathrm{sa}})^n$, and a neighborhood U of μ_X in $\mathbb{S}$, we set

$$\Gamma(X; k, U) = \{A \in (\mathfrak{M}_k^{\mathrm{sa}})^n : \mu_A \in U\}.$$

Given in addition a positive number R,

$$\Gamma_R(X; k, U) = \{A \in \Gamma(X; k, U) : \|A_j\| < R \text{ for all } j\}.$$

We can then define the quantities

$$\chi_R(X;U) = \liminf_{k\to\infty}\left[\frac{1}{k^2}\log\lambda_k^{\otimes n}(\Gamma_R(X;k,U)) + \frac{n}{2}\log k\right],$$

and

$$\chi_R(X) = \inf_U \chi_R(X;U),$$

where U runs over a neighborhood base of μ_X in $\mathbb{S}$. Finally, the free entropy is defined as

$$\chi(X) = \sup_{R>0}\chi_R(X).$$

We also set

$$\chi_\infty(X;U) = \liminf_{k\to\infty}\left[\frac{1}{k^2}\log\lambda_k^{\otimes n}(\Gamma(X;k,U)) + \frac{n}{2}\log k\right],$$

and $\chi_\infty(X) = \inf_U \chi_\infty(X;U)$. This quantity was introduced in the concluding remarks of [6], where other possible definitions of free entropy are discussed briefly. The inequalities

$$\chi_R(X) \le \chi(X) \le \chi_\infty(X)$$

are obvious for $R > 0$, and Proposition 2.4 of [6] states that $\chi_R(X) = \chi(X)$ if R is sufficiently large; $R > \max_j \|X_j\|$ will suffice. Our main result is as follows:

Proposition 2.1. *For every $X \in (\mathfrak{A}^{\mathrm{sa}})^n$ we have $\chi(X) = \chi_\infty(X)$.*

The proof of this result is a refinement of the proof of Proposition 2.4 in [6]. We begin by considering the diffeomorphism f of the real line onto $(-2, 2)$ defined by $f(t) = t$ for $t \in [-1, 1]$, $f(t) = 2 - \frac{1}{t}$ for $t > 1$, and $f(t) = -2-\frac{1}{t}$ for $t < -1$. Observe that f' does not have any local minimum, and therefore

$$\frac{f(s) - f(t)}{s - t} \ge \min\{f'(s), f'(t)\} \ge f'(s)f'(t)$$

for all s and t. The function $F_n : (\mathfrak{M}_k^{\mathrm{sa}})^n \to (\mathfrak{M}_k^{\mathrm{sa}})^n$ defined by

$$F_n(A_1, A_2, \dots, A_n) = (f(A_1), f(A_2), \dots, f(A_n))$$

is also differentiable, and we need to estimate the Jacobian determinant $(JF_n)(A)$. Since

$$(JF_n)(A) = (JF_1)(A_1)(JF_1)(A_2)\dots(JF_1)(A_n),$$

it suffices to do this in one variable. As pointed out in [6], if A is a $k \times k$ matrix with eigenvalues $\mu_1, \mu_2, \dots, \mu_k$, we have

$$(JF_1)(A) = \left(\prod_{i\ne j}\frac{f(\mu_i) - f(\mu_j)}{\mu_i - \mu_j}\right)\cdot\prod_{i=1}^{k} f'(\mu_i).$$

By the estimate for difference quotients shown above,

$$\begin{aligned}(JF_1)(A) &\geq \left(\prod_{i\neq j} f'(\mu_i)f'(\mu_j)\right)\cdot\prod_{i=1}^{k} f'(\mu_i)\\ &= \prod_{i=1}^{k} f'(\mu_i)^{2k-1} = \prod_{|\mu_i|>1} \mu_i^{-2(2k-1)}.\end{aligned}$$

Denoting $\log^+(t) = \max\{\log t, 0\}$, we obtain

$$\begin{aligned}\log(JF_1)(A) &\geq -2(2k-1)\sum_{i=1}^{k}\log^+\mu_i\\ &= -2k(2k-1)\frac{1}{k}\sum_{i=1}^{k}\log^+\mu_i\\ &= -2k(2k-1)\tau_k(\log^+|A|).\end{aligned}$$

We have therefore proved the following estimate:

Lemma 2.2. *Given $A = (A_1, A_2, \ldots, A_n) \in (\mathfrak{M}_k^{\mathrm{sa}})^n$, we have*

$$(JF_n)(A) \geq \exp\left[-2k(2k-1)\sum_{j=1}^{n}\tau_k(\log^+|A_j|)\right].$$

Note for further use that, for a selfadjoint $k\times k$ matrix A, $\tau_k(\log^+|A|)$ can be estimated in terms of the moments $\tau_k(A^{2p})$, $p\geq 1$. In fact, $\log^+ t = \frac{1}{2p}\log^+ t^{2p} \leq \frac{1}{2p}t^{2p}$, and therefore

$$\tau_k(\log^+|A|) \leq \frac{1}{2p}\tau_k(A^{2p}).$$

We need one more ingredient.

Lemma 2.3. *Let $X \in (\mathfrak{A}^{\mathrm{sa}})^n$ satisfy $\max_j\|X_j\| < 1$, and let U be a neighborhood of μ_X in $\mathbb{S}$. There exists a neighborhood V of μ_X in $\mathbb{S}$ such that*

$$F_n(\Gamma(X;k,V)) \subset \Gamma_2(X;k,U) \text{ for all } k.$$

Proof. Clearly it suffices to prove the lemma for neighborhoods of the form

$$U = \{\mu\in\mathbb{S} : |\mu(\alpha) - \tau(X_\alpha)| < \varepsilon\},$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m) \in I$ and $\varepsilon > 0$ are fixed. Using the Hölder inequality

$$\|A_\alpha\|_1 \leq \|A_{\alpha_1}\|_m\|A_{\alpha_2}\|_m\ldots\|A_{\alpha_m}\|_m,$$

we see that it is sufficient to choose V so that, for all $A \in \Gamma(X;k,V)$, we have $|\tau_k(A_\alpha) - \tau(X_\alpha)| < \varepsilon/2$, $\|A_j\|_m \leq 1$, and $\|A_j - f(A_j)\|_m < \varepsilon/2m$ for

$j = 1, 2, \ldots, n$. Choose a number $r < 1$ so that $r > \|X_j\|$ for all j, and choose an even integer $q > m$ such that $r^{q/m} < \varepsilon/2m$. Define next

$$V = \left\{ \mu \in \mathbb{S} : |\mu(\alpha) - \tau(X_\alpha)| < \frac{\varepsilon}{2} \text{ and } |\mu(j^q)| < r^q \text{ for } j = 1, 2, \ldots, n \right\};$$

recall that j^q denotes the q-multiindex with all entries equal to j. Consider now $A \in \Gamma(X; k, V)$, and note that the inequalities $|\tau_k(A_\alpha) - \tau(X_\alpha)| < \varepsilon/2$ are obviously satisfied. Also,

$$\|A_j\|_m \leq \|A_j\|_q = \tau(A_j^q)^{1/q} \leq r < 1.$$

Finally, if $\mu_1, \mu_2, \ldots, \mu_k$ are the eigenvalues of A_j,

$$\begin{aligned} \|A_j - f(A_j)\|_m &\leq \left(\frac{1}{k} \sum_{|\mu_j| \geq 1} |\mu_j|^m \right)^{1/m} \leq \left(\frac{1}{k} \sum_{|\mu_j| \geq 1} |\mu_j|^q \right)^{1/m} \\ &\leq (\tau_k(A_j^q))^{1/m} \leq r^{q/m}, \end{aligned}$$

and this quantity is less than $\varepsilon/2m$. □

Proof of Proposition 2.1. It suffices to prove the proposition in case $\|X_j\| < 1$ for all j. From the results of [**6**] we know that $\chi_2(X) = \chi(X)$, and clearly $\chi_2(X) \leq \chi_\infty(X)$. To prove the opposite inequality $\chi_2(X) \geq \chi_\infty(X)$, let U be a neighborhood of μ_X in $\mathbb{S}$, and let V be the neighborhood of μ_X furnished by Lemma 2.3, i.e., $F_n(\Gamma(X; k, V)) \subset \Gamma_2(X; k, U)$ for all $k \geq 1$. Given a positive integer p, we may also assume that $\tau_k(A_j^{2p}) \leq 1$ whenever $A = (A_1, A_2, \ldots, A_n) \in \Gamma(X; k, V)$. It follows then from Lemma 2.2 (and the remark following its statement) that

$$(JF_n)(A) \geq \exp\left[-2k(2k-1)\frac{n}{2p} \right]$$

for all $A \in \Gamma(X; k, V)$. Since the function F_n is one-to-one, we deduce that

$$\begin{aligned} \lambda_k^{\otimes n}(\Gamma_2(X; k, U)) &\geq \lambda_k^{\otimes n}(F_n(\Gamma(X; k, V))) \\ &\geq \exp\left[-2k(2k-1)\frac{n}{2p} \right] \lambda_k^{\otimes n}(\Gamma(X; k, V)). \end{aligned}$$

Therefore

$$\begin{aligned} &\frac{1}{k^2} \log \lambda_k^{\otimes n}(\Gamma_2(X; k, U)) + \frac{n}{2} \log k \\ &\geq \frac{1}{k^2} \log \lambda_k^{\otimes n}(\Gamma(X; k, V)) + \frac{n}{2} \log k - \left(2 - \frac{1}{k}\right) \frac{n}{p}, \end{aligned}$$

and as $k \to \infty$ this yields

$$\chi_2(X; U) \geq \chi_\infty(X; V) - \frac{2n}{p}.$$

Since p is arbitrary, we deduce that $\chi_2(X;U) \geq \chi_\infty(X;V) \geq \chi_\infty(X)$, and the proof is concluded by taking the infimum over U. □

We remark that a suitable modification of the above proof yields directly that $\chi_\infty(X) = \chi_R(X)$ if $\|X_j\| < R$. One needs an appropriate version of the function f, and that is easily constructed.

References

[1] G. Ben Arous and A. Guionnet, *Large deviations for Wigner's law and Voiculescu's noncommutative entropy*, Probab. Theory Related Fields, **108** (1997), 517-542, MR 98i:15026, Zbl 0954.60029.

[2] T. Cabanal Duvillard and A. Guionnet, *Large deviations upper bounds for the laws of matrix-valued processes and non-commutative entropies*, Ann. Probab., **29** (2001), 1205-1261, MR 2003a:60040.

[3] P. Śniady, *Inequality for Voiculescu's free entropy in terms of Brown measure*, Int. Math. Res. Notices, (2003), 51-64, CMP 1 935 566.

[4] D. Voiculescu, *Limit laws for random matrices and free products*, Invent. Math., **104** (1991), 201-220, MR 92d:46163, Zbl 0736.60007.

[5] ———, *The analogues of entropy and Fisher's information measure in free probability*, I, Comm. Math. Phys., **155** (1993), 71-92, MR 94k:46137, Zbl 0781.60006.

[6] ———, *The analogues of entropy and Fisher's information measure in free probability*, II, Invent. Math., **1118** (1994), 411-440, MR 96a:46117, Zbl 0820.60001.

[7] ———, *The analogues of entropy and Fisher's information measure in free probability. III. The absence of Cartan algebras*, Geom. Funct. Anal., **6** (1996), 172-199, MR 96m:46119, Zbl 0856.60012.

[8] ———, *A strengthened asymptotic freeness result for random matrices with applications to free entropy*, Internat. Math. Res. Notices, (1998), 41-63, MR 2000d:46080, Zbl 0895.60004.

[9] E. Wigner, *On the distribution of the roots of certain symmetric matrices*, Ann. Math., **67** (1958), 325-327, MR 20 #2029, Zbl 0085.13203.

Received May 6, 2002 and revised June 20, 2002. The second author was supported in part by a grant from the National Science Foundation.

Institute of Mathematics
Romanian Academy
P. O. Box 1-764
Bucharest RO-70700
Romania
E-mail address: sbelinsc@indiana.edu

Mathematics Department
Indiana University
Bloomington, IN 47405
E-mail address: bercovic@indiana.edu