Mathematical Research Letters **10**, 185–189 (2003)

ON AN EQUIVALENT DEFINITION OF FREE ENTROPY

Serban Teodor Belinschi

Abstract. In his second paper on free entropy, Voiculescu suggested five other possible definitions of the entropy of an n-tuple of selfadjoint random variables. We show that two of them are equivalent to the original one.

1. Introduction

Let $X_1, \dots, X_n$ be n selfadjoint random variables in the tracial W^*- probability space $(\mathcal{M}, \tau)$. Denote by $\mathcal{M}_k^{sa}$ the space of $k \times k$ selfadjoint matrices over $\mathbb{C}$, let Tr_k be the (unnormalized) trace on $\mathcal{M}_k^{sa}$, and denote by τ_k the normalized trace, $\tau_k = \frac{1}{k}\mathrm{Tr}_k$. Then $(\mathcal{M}_k^{sa})^n$ becomes a real Hilbert space with the norm

$$\|(A_1, \dots, A_n)\|_2 = (\mathrm{Tr}_k(A_1^2 + \cdots + A_n^2))^{1/2}.$$

Denote by λ the Lebesgue measure corresponding to this norm (i.e. the Lebesgue measure normalized so that a cube with edges of length one has measure equal to one).

Let $\varepsilon > 0, m > 1, R > \max\{\|X_1\|, \dots \|X_n\|\}$, and let $\Psi_{m,k} : (\mathcal{M}_k^{sa})^n \to \mathbb{R}_+$ be defined by

$$\Psi_{m,k}(A_1, \dots, A_n) = \sum_{q=1}^{m} \sum_{1 \le i_1, \dots, i_q \le n} |\tau_k(A_{i_1} \dots A_{i_q}) - \tau(X_{i_1} \dots X_{i_q})|.$$

Following Voiculescu's definition of microstate-based free entropy $\chi(X_1, \dots, X_n)$, let

$$\begin{aligned}
\Gamma_R(X_1, \dots, X_n; m, k, \varepsilon) &= \{(A_1, \dots, A_n) \in (\mathcal{M}_k^{sa})^n : \Psi_{m,k}(A_1, \dots, A_n) < \varepsilon, \\
&\qquad \|A_j\| \le R\} \\
\chi_R(X_1, \dots, X_n; m, k, \varepsilon) &= \log \lambda(\Gamma_R(X_1, \dots, X_n; m, k, \varepsilon)) \\
\chi_R(X_1, \dots, X_n; m, \varepsilon) &= \limsup_{k \to \infty}(k^{-2}\chi_R(X_1, \dots, X_n; m, k, \varepsilon) + 2^{-1} n \log k) \\
\chi_R(X_1, \dots, X_n) &= \lim_{\varepsilon \to 0, m \to \infty} \chi_R(X_1, \dots, X_n; m, \varepsilon) \\
\chi(X_1, \dots, X_n) &= \sup_{R > 0} \chi_R(X_1, \dots, X_n).
\end{aligned}$$

Received November 19, 2002.
2000 *Mathematics Subject Classification.* Primary: 46L54; Secondary: 94A17.

In Section 7.1 of [3], two alternative quantities were defined by Voiculescu as follows. Let $\mathcal{D}_{1,R}(X_1,\ldots,X_n;m,k,\varepsilon)$ be the set of probability densities $f : (\mathcal{M}_k^{sa})^n \to \mathbb{R}_+$ (i.e. $\int f d\lambda = 1$) so that $\int f(A_1,\ldots,A_n)\Psi_{m,k}(A_1,\ldots,A_n)d\lambda \leq \varepsilon$, and $f(A_1,\ldots,A_n) = 0$ unless $\|A_j\| \leq R$ $(1 \leq j \leq n)$. Then

$$\chi_R^{(1)}(X_1,\ldots,X_n;m,k,\varepsilon) = \sup\left\{-\int f\log f d\lambda : f \in \mathcal{D}_{1,R}(X_1,\ldots,X_n;m,k,\varepsilon)\right\}.$$

$\chi^{(1)}$ is now defined following the same steps as in the definition of χ. One can also define χ_∞ and $\chi_\infty^{(1)}$ by setting $R = \infty$ from the beginning, that is, by leaving out the conditions on the norms of the A_j's. Since the characteristic function of $\Gamma_R(X_1,\ldots,X_n;m,k,\varepsilon)$ (properly normalized) belongs to $\mathcal{D}_{1,R}(X_1,\ldots,X_n;m,k,\varepsilon)$, we conclude that

$$\chi_R^{(1)}(X_1,\ldots,X_n;m,k,\varepsilon) \geq \chi_R(X_1,\ldots,X_n;m,k,\varepsilon).$$

It follows that $\chi \leq \chi^{(1)}$, and, likewise, $\chi_\infty \leq \chi_\infty^{(1)}$.

It has already been proved in [1] that $\chi_\infty = \chi$. In this paper we shall prove that $\chi_\infty^{(1)} = \chi_\infty$, which will obviously imply that $\chi_\infty^{(1)} = \chi^{(1)} = \chi = \chi_\infty$. As an immediate consequence, we obtain a new way to define free entropy replacing $\chi_R(X_1,\ldots,X_n;m,k,\varepsilon)$ by:

$$\chi_\infty^{(1)}(X_1,\ldots,X_n;m,k,\varepsilon) = \log\int e^{-s(\Psi_{m,k}(A_1,\ldots,A_n)-\varepsilon)}d\lambda,$$

and continuing as in the original definition. Here s is an appropriate positive number depending on $X_1, X_2, \ldots, X_n, m, \varepsilon$ and k.

2. The main result

Given a probability density $f : \mathbb{R}^p \to [0,+\infty)$, i.e. $\int_{\mathbb{R}^p} f(x)d\lambda(x) = 1$, we denote by $I(f) = -\int_{\mathbb{R}^p} f(x)\log f(x)d\lambda(x)$ the classical entropy of f. A density of special interest is the Gaussian density $G_{a^2}^{(p)}$ defined on $\mathbb{R}^p$ by $G_{a^2}^{(p)}(x) = \left(\frac{1}{2\pi a^2}\right)^{\frac{p}{2}} e^{-\frac{\|x\|_2^2}{2a^2}}$.

We will need the following extension of Shannon's classical inequality (see [2], pp. 55–57).

Lemma 2.1. *Let $\Omega \subset \mathbb{R}^p$ be a set of positive Lebesgue measure, and $\varepsilon > 0$ be fixed. If φ is a positive continuous function on Ω such that $\int_\Omega \varphi \geq \varepsilon\lambda(\Omega)$, $\varphi^{-1}([0,\varepsilon))$ is nonempty, and both $e^{-\varphi}$ and $\varphi^2 e^{-\varphi}$ are integrable on Ω, then there is a number $s \in (0,\infty)$ depending on φ, p, and ε such that $M :=$*

$$\max\left\{I(f) : f \geq 0, f = 0 \text{ on } \mathbb{R}^n \setminus \Omega, \int_\Omega f(x)d\lambda(x) = 1, \int_\Omega f(x)\varphi(x)d\lambda(x) = \varepsilon\right\}$$

is attained for $f = (\int_\Omega e^{-s\varphi(x)}d\lambda(x))^{-1}e^{-s\varphi}$ and equals $\log\int_\Omega e^{-s(\varphi(x)-\varepsilon)}d\lambda$.

Proof. For each $s > 0$, let $f_s(x) = (\int_\Omega e^{-s\varphi(y)} d\lambda(y))^{-1} e^{-s\varphi(x)}$. Consider the function $h : (0, \infty) \to (0, \infty)$ given by $h(s) = \int_\Omega \varphi(x) f_s(x) d\lambda(x)$. Since both $e^{-\varphi}$ and $\varphi^2 e^{-\varphi}$ are integrable on Ω, we can differentiate under the integral sign, and a straightforward calculation shows that

$$h'(s) = -\int_\Omega \left(\varphi(x) - \int_\Omega \varphi(y) f_s(y) d\lambda(y)\right)^2 f_s(x) d\lambda(x),$$

so $h'(s) < 0$, We also have that

$$\lim_{s\to\infty} h(s) = \min_{x\in\Omega} \varphi(x),$$

which is strictly less than ε, and $\lim_{s\to 0} h(s)$ equals $\frac{1}{\lambda(\Omega)} \int_\Omega \varphi$ if $\lambda(\Omega) < \infty$ and infinity if $\lambda(\Omega) = \infty$, which are both greater than or equal to ε. Hence the existence of s is assured by the monotonicity of h. Moreover, since φ is nonconstant, s is unique with this property.

Let

$$g \in \left\{ f : f \geq 0, \int_\Omega f d\lambda = 1, \int_\Omega f\varphi d\lambda = \varepsilon \right\}.$$

We may assume that $g(x) < \infty$ for all $x \in \Omega$. If $I(g) = -\infty$, then it is obvious that $I(f_s) > I(g)$. So suppose that $I(g) > -\infty$. Consider the function $\eta : [0,1] \to \mathbb{R}$ defined by $\eta(t) = I((1-t)f_0 + tg)$. This function is continuous on $[0,1]$, differentiable on $[0,1)$, and twice differentiable on $(0,1)$. Moreover, $\eta'(t) = -\int_\Omega (g - f_s)(\log((1-t)f_s + tg) + 1)$, and $\eta''(t) = -\int_\Omega \frac{(g-f_s)^2}{(1-t)f_s+tg}$. (This follows immediatly from the dominated convergence theorem. Indeed,

$$|(g - f_s)(\log((1-t)f_s + tg) + 1)| \leq g + f_s + |(g - f_s)\log((1-t)f_s + tg)|$$

for all $t \in [0,1)$, and for any compact interval $K \subset [0,1)$, $\sup_{t\in K} |(g-f_s)\log((1-t)f_s + tg)|$ is integrable on Ω. Also,

$$\left|\frac{(g - f_s)^2}{(1-t)f_s + tg}\right| \leq \frac{1}{t} g + \frac{1}{1-t} f_s$$

for all $t \in (0,1)$, and, of course, $\sup_{t\in K} |\frac{1}{t} g + \frac{1}{1-t} f_s|$ is integrable on Ω for any compact interval $K \subset (0,1)$.)

By direct computation, we obtain that $\eta'(0) = 0$. Also, η' is continuous on $[0,1)$, and $\eta''(t) \leq 0$ for all $t \in (0,1)$, so the point $t = 0$ is a global maximum for η. This proves that $I(f_s) = M$ and concludes the proof. □

Remark 2.2. To see how this lemma extends Shannon's inequality, set $\Omega = \mathbb{R}^n$, and $\varphi(x) = \|x\|_2^2$. In this case the extremal density f is $G_\varepsilon^{(p)}$ and $I(G_\varepsilon^{(p)}) = 2^{-1} p \log(2\pi e p^{-1} \varepsilon)$. Also note that for $\lambda(\Omega) < +\infty$ and $\varphi = \varepsilon$, the extremal density is $\frac{1}{\lambda(\Omega)} \chi_\Omega$ and $I(\frac{1}{\lambda(\Omega)} \chi_\Omega) = \log \lambda(\Omega)$.

Now we can prove the main result of the paper.

Theorem 2.1. *Let $X_1, \ldots, X_n$ be selfadjoint random variables in a tracial W^*-probability space $(\mathcal{M}, \tau)$. Then*

$$\chi_\infty(X_1, \ldots, X_n) = \chi_\infty^{(1)}(X_1, \ldots, X_n).$$

Proof. As noted in the introduction, $\chi_\infty(X_1, \ldots, X_n) \leq \chi_\infty^{(1)}(X_1, \ldots, X_n)$, so it is enough to prove the opposite inequality. Let $m \geq 2$ and $\varepsilon \in (0, 1/2)$ be fixed. Let $V_0 = \{(A_1, \ldots, A_n) \in (\mathcal{M}_k^{sa})^n : \Psi_{m,k}(A_1, \ldots, A_n) \leq \varepsilon\}$, and $V_1 = (\mathcal{M}_k^{sa})^n \setminus V_0$. Consider $f_0 \in \mathcal{D}_{1,\infty}(X_1, \ldots, X_n; m, k, \varepsilon^2)$ such that $I(f_0) = \max\{I(f) : f \in \Gamma_\infty^{(1)}(X_1, \ldots, X_n; m, k, \varepsilon^2)\}$, and define $f_0^{(i)} = {c_i}^{-1} f_0 \chi_{V_i}$, where $c_i = \int_{V_i} f,\ i = 0, 1$ depend on $X_1, \ldots, X_n, m, k$ and ε. Then we have:

$$\varepsilon^2 \geq \int f_0 \Psi_{m,k} \geq \int_{V_1} f_0 \Psi_{m,k} \geq \varepsilon \int_{V_1} f_0 = \varepsilon c_1,$$

so $c_1 \leq \varepsilon$, and $c_0 \geq 1 - \varepsilon$.

Since $f_0 \in \mathcal{D}_{1,\infty}(X_1, \ldots, X_n; m, k, \varepsilon^2)$,

$$\int f_0^{(1)} \Psi_{m,k} = \left(\int_{V_1} f_0\right)^{-1} \int_{V_1} f_0 \Psi_{m,k} < \frac{\varepsilon^2}{c_1},$$

which implies that $f_0^{(1)} \in \mathcal{D}_{1,\infty}(X_1, \ldots, X_n; m, k, \frac{\varepsilon^2}{c_1})$.

Now,

$$\begin{aligned} I(f_0) &= c_0 I(f_0^{(0)}) + c_1 I(f_0^{(1)}) - c_0 \log c_0 - c_1 \log c_1 \\ &\leq c_0 I\left(\frac{1}{|V_0|}\chi_{V_0}\right) + c_1 I(G_{k(C^2+n\varepsilon^2/c_1)}^{(nk^2)}) - c_0 \log c_0 - c_1 \log c_1 \\ &= c_0 \chi_\infty(X_1, \ldots, X_n; m, k, \varepsilon) + c_1 \frac{nk^2}{2}\left(\log \frac{2\pi e(C^2 + n\varepsilon^2/c_1)}{n} - \log k\right) \\ &\quad - c_0 \log c_0 - c_1 \log c_1, \qquad (1) \end{aligned}$$

where $C^2 = \tau({X_1}^2 + \cdots + {X_n}^2)$. Here we used Lemma 2.1 in the last inequality.

Dividing by k^2 and adding $\frac{n}{2}\log k$ in (1), we obtain:

$$\begin{aligned} \frac{1}{k^2} I(f_0) + \frac{n}{2}\log k &\leq c_0\left(\frac{1}{k^2}\chi_\infty(X_1, \ldots, X_n; m, k, \varepsilon) + \frac{n}{2}\log k\right) \\ &\quad + c_1 \frac{n}{2}\log \frac{2\pi e(C^2 + n\varepsilon^2/c_1)}{n} - \frac{1}{k^2}(c_0 \log c_0 + c_1 \log c_1). \end{aligned}$$

As $k \to \infty$, this yields

$$\begin{aligned} \chi_\infty^{(1)}(X_1, \ldots, X_n; m, \varepsilon^2) &\leq \limsup_{k\to\infty} c_0\left(\frac{1}{k^2}\chi_\infty(X_1, \ldots, X_n; m, k, \varepsilon) + \frac{n}{2}\log k\right) \\ &\quad + \limsup_{k\to\infty} c_1 \frac{n}{2}\log \frac{2\pi e(C^2 + n\varepsilon^2/c_1)}{n} \end{aligned}$$

Since $\lim_{t\to 0, t>0} t \log \frac{2\pi e(C^2+n\varepsilon^2/t)}{n} = 0$, the function

$$s(\varepsilon) := \sup\left\{t\frac{n}{2}\log\frac{2\pi e(C^2+n\varepsilon^2/t)}{n} : t\in(0,\varepsilon]\right\}$$

does not depend on k and m and tends to zero as $\varepsilon \to 0$.

Hence

$$\chi_\infty^{(1)}(X_1,\ldots,X_n;m,\varepsilon^2) \le \limsup_{k\to\infty} c_0\left(\frac{1}{k^2}\chi_\infty(X_1,\ldots,X_n;m,k,\varepsilon)+\frac{n}{2}\log k\right)+s(\varepsilon)$$

On the other hand, since $1 \ge c_0 \ge 1-\varepsilon$, we have that

$$\limsup_{k\to\infty} c_0\left(\frac{1}{k^2}\chi_\infty(X_1,\ldots,X_n;m,k,\varepsilon)+\frac{n}{2}\log k\right) \le \max\left\{(1-\varepsilon)\chi_\infty(X_1,\ldots,X_n;m,\varepsilon),\chi_\infty(X_1,\ldots,X_n;m,\varepsilon)\right\}.$$

This allows us to conclude that

$$\chi_\infty^{(1)}(X_1,\ldots,X_n;m,\varepsilon^2) \le \max\left\{(1-\varepsilon)\chi_\infty(X_1,\ldots,X_n;m,\varepsilon),\chi_\infty(X_1,\ldots,X_n;m,\varepsilon)\right\}+s(\varepsilon).$$

Taking limits when $m\to\infty$ and $\varepsilon\to 0$ in the previous inequality, we obtain

$$\chi_\infty^{(1)}(X_1,\ldots,X_n) = \chi_\infty(X_1,\ldots,X_n).$$

□

Acknowledgement

I wish to thank my advisor Professor Hari Bercovici for the constant help that made this paper possible.

References

[1] S. T. Belinschi, H. Bercovici, *A property of free entropy,* to appear in Pacific J. Math.
[2] C. E. Shannon, *The Mathematical Theory of Communication,* The University of Illinois Press, Urbana, IL, 1949.
[3] D. Voiculescu, *The analogues of entropy and of Fisher's information measure in free probability theory. II,* Invent. Math. **118** (1994), 411–440.

Mathematics Department, Indiana University, Bloomington, Indiana 47405, U.S.A.
E-mail address: sbelinsc@indiana.edu

Institute of Mathematics, Romanian Academy, P.O. Box 1-764, Bucharest, RO-70700, Romania.