Semicircularity, Gaussianity and Monotonicity of Entropy

Hanne Schultz *

Abstract

S. Artstein, K. Ball, F. Barthe, and A. Naor have shown (cf. [ABBN]) that if \((X_j)_{j=1}^\infty\) are i.i.d. random variables, then the entropy of \(\frac{X_1+\ldots+X_n}{\sqrt{n}}\), \(H\left(\frac{X_1+\ldots+X_n}{\sqrt{n}}\right)\), increases as \(n\) increases.

The free analogue was recently proven by D. Shlyakhtenko in [Sh]. That is, if \((x_j)_{j=1}^\infty\) are freely independent, identically distributed, self-adjoint elements in a noncommutative probability space, then the free entropy of \(\frac{X_1+\ldots+X_n}{\sqrt{n}}\), \(\chi\left(\frac{X_1+\ldots+X_n}{\sqrt{n}}\right)\), increases as \(n\) increases.

In this paper we prove that if \(H(X_1) > -\infty\) (\(\chi(x_1) > -\infty\), resp.), and if the entropy (the free entropy, resp.) is not a strictly increasing function of \(n\), then \(X_1\) (\(x_1\), resp.) must be Gaussian (semicircular, resp.).

1 Introduction.

Shannon’s entropy of a (classical) random variable \(X\) with Lebesgue absolutely continuous distribution \(d\mu_X(x) = \rho(x)dx\), is given by

\[
H(X) = -\int_{\mathbb{R}} \rho(x) \log \rho(x) \, dx,
\]

whenever the integral exists. If the integral does not exist, or if the distribution of \(X\) is not Lebesgue absolutely continuous, then \(H(X) = -\infty\).

The entropy can also be written in terms of score functions and of Fisher information. Take a standard Gaussian random variable \(G\) such that \(X\) and \(G\) are independent. Let \(X^{(t)} = X + \sqrt{t}G\), \(t \geq 0\), and let \(j(X^{(t)}) = \left(\frac{\partial}{\partial x}\right)^* \mathbf{1} \in L^2(\mu_{X^{(t)}})\) denote the score function of \(X^{(t)}\) (cf. [Sh, Section 3]). Then

\[
H(X) = \frac{1}{2} \int_0^\infty \left[ \frac{1}{1 + t} - \|j(X^{(t)})\|_2^2 \right] dt + \frac{1}{2} \log(2\pi e).
\]

The quantity \(\|j(X^{(t)})\|_2^2\) is called the Fisher information of \(X^{(t)}\) and is denoted by \(F(X^{(t)})\). Among all random variables with a given variance, the Gaussians are the (unique) ones with the smallest Fisher information and the largest entropy.

* As a student of the PhD-school OP-ALG-TOP-GEO the author is partially supported by the Danish Research Training Council.
† Partially supported by The Danish National Research Foundation.
A. J. Stam (cf. [St]) was the first to rigorously show that if $X_1$ and $X_2$ are independent random variables of the same variance, with $H(X_1), H(X_2) > -\infty$, then for all $t \in [0, 1]$,

$$H(\sqrt{t}X_1 + \sqrt{1-t}X_2) \geq tH(X_1) + (1-t)H(X_2),$$

with equality iff $X_1$ and $X_2$ are Gaussian. It follows that if $(X_j)_{j=1}^{\infty}$ is a sequence of i.i.d. random variables with finite entropy, then

$$n \mapsto H\left(\frac{X_1 + \cdots + X_{2n}}{2}\right)$$

is an increasing function of $n$, and if it is not strictly increasing, then $X_1$ is necessarily Gaussian.

Knowing about Stam’s result, it seems natural to ask whether the map

$$n \mapsto H\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right)$$

is monotonically increasing as well, or even simpler: Is $H\left(\frac{X_1 + X_2 + X_3}{\sqrt{3}}\right) \geq H\left(\frac{X_1 + X_2}{\sqrt{2}}\right)$? Surprisingly enough, it took more than 40 years for someone to answer these questions. Both questions were answered in the affirmative in [ABBN] in 2004.

In this paper we extend Stam’s result by showing that if $H(X_1) > -\infty$ and if for some $n \in \mathbb{N}$,

$$H\left(\frac{X_1 + \cdots + X_{n+1}}{\sqrt{n+1}}\right) = H\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right),$$

then $X_1$ is necessarily Gaussian (Theorem 3.1).

Free entropy, which is the proper free analogue of Shannon’s entropy, was defined by Voiculescu in [V1]. If $x$ is a self-adjoint element in a finite von Neumann algebra $\mathcal{M}$ with faithful normal tracial state $\tau$ and if $\mu_x \in \text{Prob}(\mathbb{R})$ denotes the distribution of $x$ with respect to $\tau$, then the free entropy of $x$, $\chi(x) \in [-\infty, \infty]$, is given by

$$\chi(x) = \int \int \log|s - t|d\mu_x(s)d\mu_x(t) + \frac{3}{4} + \frac{1}{2}\log(2\pi).$$

Exactly as in the classical case, $\chi(x)$ may be written in terms of the free analogue of the score function (the conjugate variable) and the free Fisher information. That is, if $s$ is a $(0,1)$-semicircular element which is freely independent of $x$ and if we let

$$x^{(t)} = x + \sqrt{ts}, \quad t \geq 0,$$

then

$$\chi(x) = \frac{1}{2} \int_0^\infty \left[ \frac{1}{1 + t} - \Phi(x^{(t)}) \right] dt + \frac{1}{2}\log(2\pi e), \quad \text{(1.3)}$$

where $\Phi(x^{(t)})$ is the free Fisher information of $x^{(t)}$. In [V2] Voiculescu defines for a (non-scalar) self-adjoint variable $y$ in $(\mathcal{M}, \tau)$ a derivation $\partial_y : \mathbb{C}[y] \rightarrow \mathbb{C}[y] \otimes \mathbb{C}[y]$ by

$$\partial_y(1) = 0 \quad \text{and} \quad \partial_y(y) = 1 \otimes 1.$$

Then the conjugate variable of $y$, if it exists, is the unique vector $\mathcal{J}(y) \in L^2(W^*(y))$ satisfying that for all $k \in \mathbb{N}$,

$$\langle \mathcal{J}(y), y^k \rangle = \langle 1 \otimes 1, \partial_y(y^k) \rangle. \quad \text{(1.4)}$$
That is, \( J(y) = (\partial_y)^*(1 \otimes 1) \). The conjugate variable is the free analogue of the score function, and the free Fisher information of \( y \) is exactly \( \|J(y)\|_2^2 \), so that

\[
\chi(x) = \frac{1}{2} \int_0^\infty \left[ \frac{1}{1 + t} - \|J(x(t))\|_2^2 \right] dt + \frac{1}{2} \log(2\pi e). \tag{1.5}
\]

Note that if \( J(y) = y \), then the moments of \( y \) are determined by (1.4), and it is not hard to see that \( y \) is necessarily (0,1)-semicircular.

In [Sh] D. Shlyakhtenko showed that if \( (x_j)_{j=1}^\infty \) are freely independent, identically distributed self-adjoint elements in \((M, \tau)\), then the map

\[
n \mapsto \chi\left( x_1 + \cdots + x_n \sqrt{n} \right)
\]

is monotonically increasing in \( n \). In fact, the method used in [Sh] applies to the classical case as well. In this paper we will dig into the proof of the inequality

\[
\chi\left( \frac{x_1 + \cdots + x_{n+1}}{\sqrt{n+1}} \right) \geq \chi\left( \frac{x_1 + \cdots + x_n}{\sqrt{n}} \right) \tag{1.6}
\]

and find out what it means for all of the estimates obtained in the course of the proof to be equalities. We conclude that if \( \chi(x_1) > -\infty \) and if (1.6) is an equality for some \( n \), then \( x_1 \) is necessarily semicircular. With a few modifications, our method applies to the classical case as well.

Acknowledgements. I would like to thank Dimitri Shlyakhtenko for introducing me to the problem on monotonicity of entropy and for fruitful discussions about the subject. In addition, I would like to thank the UCLA Math Department, especially the functional analysis group, for their kind hospitality during my stay at UCLA.

2 The Free Case.

Recall that the \((0,1)\)-semicircle law is the Lebesgue absolutely continuous probability measure on \( \mathbb{R} \) with density

\[
d\sigma_{0,1}(t) = \frac{1}{2\pi} \sqrt{4 - t^2} \, 1_{[-2,2]}(t) \, dt.
\]

More generally, for \( \mu, \gamma \in \mathbb{R} \) with \( \gamma > 0 \), the \((\mu, \gamma)\)-semicircle law is the Lebesgue absolutely continuous probability measure on \( \mathbb{R} \) with density

\[
d\sigma_{\mu,\gamma}(t) = \frac{1}{2\pi\gamma} \sqrt{4\gamma - (t - \mu)^2} \, 1_{[\mu - 2\sqrt{\gamma}, \mu + 2\sqrt{\gamma}]}(t) \, dt.
\]

The parameters \( \mu \) and \( \gamma \) refer to the first moment and the variance of \( \sigma_{\mu,\gamma} \), respectively.

Throughout this section, \( M \) denotes a finite von Neumann algebra with faithful, normal, tracial state \( \tau \). We are going to prove:

2.1 Theorem. Let \( n \in \mathbb{N} \) and let \( x_1, \ldots, x_{n+1} \) be freely independent, identically distributed self-adjoint elements in \((M, \tau)\). Then

\[
\chi\left( \frac{x_1 + \cdots + x_{n+1}}{\sqrt{n+1}} \right) \geq \chi\left( \frac{x_1 + \cdots + x_n}{\sqrt{n}} \right). \tag{2.1}
\]

Moreover, if \( \chi(x_1) > -\infty \), then equality holds in (2.1) iff \( x_1 \) is semicircular.
Monotonicity of free entropy was already proven in [Sh]. Likewise, most of the results stated in this section consist of two parts: An inequality which was proven in [Sh] or in [ABBN] and a second part which was proven by us.

2.2 Proposition. Let \( n \in \mathbb{N} \) and let \( x_1, \ldots, x_{n+1} \) be freely independent self-adjoint elements in \((\mathcal{M}, \tau)\) with \( \tau(x_j) = 0 \) and \( \|x_j\|_2 = \|x_1\|_2, 1 \leq j \leq n + 1 \). Let \( a_1, \ldots, a_{n+1} \in \mathbb{R} \) with \( \sum_j a_j^2 = 1 \), and let \( b_1, \ldots, b_{n+1} \in \mathbb{R} \) such that \( \sum_j b_j \sqrt{1 - a_j^2} = 1 \). Then

\[
\Phi\left(\sum_{j=1}^{n+1} a_j x_j\right) \leq n \sum_{j=1}^{n+1} b_j^2 \Phi\left(\frac{1}{\sqrt{1 - a_j^2}} \sum_{i \neq j} a_i x_i\right). \tag{2.2}
\]

Moreover, if \( \Phi(\sum_{i \neq j} a_i x_i) \) is finite for all \( j \), then equality in (2.2) implies that

\[
\mathcal{J}\left(\sum_{j=1}^{n+1} a_j x_j\right) = \frac{1}{\|x_1\|_2} \sum_{j=1}^{n+1} a_j x_j, \tag{2.3}
\]

so that \( \sum_{j=1}^{n+1} a_j x_j \) is \((0, \|x_1\|_2^2)\)-semicircular.

2.3 Lemma. Let \( P_1, \ldots, P_m \) be commuting projections on a Hilbert space \( \mathcal{H} \). If \( \xi_1, \ldots, \xi_m \in \mathcal{H} \) satisfy that for all \( 1 \leq i \leq m \),

\[
P_1 P_2 \cdots P_m \xi_i = 0,
\]

then

\[
\|P_1 \xi_1 + \ldots + P_m \xi_m\|^2 \leq (m - 1) \sum_{i=1}^m \|\xi_i\|^2. \tag{2.4}
\]

Moreover, if equality holds in (2.4), then \( \xi_i \in \bigoplus_{j \neq i} \mathcal{H}_j \), where

\[
\mathcal{H}_j := \{ \xi \in \mathcal{H} \mid P_k \xi = \xi, \ k \neq j, \ P_j \xi = 0 \} = \left( \bigcap_{k \neq j} P_k(\mathcal{H}) \right) \bigcap P_j^\perp(\mathcal{H}).
\]

Proof. The inequality (2.4) is the content of [ABBN, Lemma 5]. The starting point of their proof is to write each \( \xi_i \) as an orthogonal sum,

\[
\xi_i = \sum_{\varepsilon \in \{0,1\}^m \setminus (1,1,\ldots,1)} \varepsilon_i \xi_{\varepsilon_i},
\]

where for \( \varepsilon \in \{0,1\}^m \setminus \{1,1,\ldots,1\} \),

\[
\xi_{\varepsilon_i} \in \mathcal{H}_\varepsilon := \{ \xi \in \mathcal{H} \mid P_j \xi = \varepsilon_j \xi, \ 1 \leq j \leq m \}.
\]

Then

\[
P_1 \xi_1 + \ldots + P_m \xi_m = \sum_{\varepsilon \in \{0,1\}^m \setminus (1,1,\ldots,1)} \sum_{\varepsilon_i = 1} P_i \xi_{\varepsilon_i},
\]

and

\[
\|P_1 \xi_1 + \ldots + P_m \xi_m\|^2 = \sum_{\varepsilon \in \{0,1\}^m \setminus (1,1,\ldots,1)} \left\| \sum_{\varepsilon_i = 1} P_i \xi_{\varepsilon_i} \right\|^2.
\]
For fixed \(\varepsilon \neq (1,1,\ldots,1)\) there can be at most \(m-1\) \(i\)'s for which \(\varepsilon_i = 1\). Thus, by the Cauchy-Schwarz inequality,
\[
\left\| \sum_{\varepsilon_i = 1} P_i \xi^i \right\|^2 \leq \left( \sum_{\varepsilon_i = 1} \| P_i \xi^i \| \right)^2 \leq (m-1) \sum_{\varepsilon_i = 1} \| P_i \xi^i \|^2, \tag{2.5}
\]
with the second inequality being an equality iff the vector \((\| P_i \xi^i \|)_{\varepsilon_i = 1} = (\| \xi^i \|)_{\varepsilon_i = 1}\) has \(m-1\) coordinates and is parallel to the vector \(v = (1,1,\ldots,1) \in \mathbb{R}^{m-1}\). In particular, if the second inequality in (2.5) is an equality for some \(\varepsilon \in \{0,1\}^m\) with more than one coordinate which is zero, then \((\| P_i \xi^i \|)_{i=1}^m\) must consist of zeros only. It follows now that
\[
\| P_1 \xi_1 + \ldots + P_m \xi_m \|^2 \leq (m-1) \sum_{\varepsilon \in \{0,1\}^m \setminus \{1,1,\ldots,1\}} \sum_{\varepsilon_i = 1} \| P_i \xi^i \|^2 \tag{2.6}
\]
\[
= (m-1) \sum_{\varepsilon \in \{0,1\}^m \setminus \{1,1,\ldots,1\}} \sum_{i=1}^m \| \xi^i \|^2 \tag{2.7}
\]
\[
\leq (m-1) \sum_{i=1}^m \sum_{\varepsilon \in \{0,1\}^m \setminus \{1,1,\ldots,1\}} \| \xi^i \|^2 \tag{2.8}
\]
\[
= (m-1) \sum_{i=1}^m \| \xi_i \|^2. \tag{2.9}
\]
Moreover, equality in (2.4) implies that all the inequalities (2.5), (2.6) and (2.8) are equalities. Hence,

(i) \(\xi^i = P_i \xi^i\) for all \(\varepsilon \neq (1,1,\ldots,1)\) and all \(1 \leq i \leq m\) (cf. (2.7) and (2.8)), and

(ii) by the Cauchy-Schwarz argument, for all \(\varepsilon \in \{0,1\}^m\) with more than one coordinate which is zero, \(\| \xi^i \|^{(0)} = \| P_i \xi^i \| = 0\) for all \(i\).

Thus, if equality holds in (2.4), then \(\xi_i \in P_i(\mathcal{H})\) and \(\xi_i \in \bigoplus_{j \neq i} \mathcal{H}_j\), as claimed. \qed

Proof of Proposition 2.2. (2.2) is the content of [Sh, Lemma 2]. Now, assume that equality holds in (2.2) and that \(\Phi(\sum_{i \neq j} a_i x_i)\) is finite for all \(j\). We are going to "backtrack" the proof of [Sh, Lemma 2] to show that (2.3) holds. We will assume that \(\| x_j \|_2 = 1\) for all \(j\).

With
\[
\xi_j = b_j \left( \frac{1}{\sqrt{1-a_j^2}} \sum_{i \neq j} a_i x_i \right), \quad 1 \leq j \leq n+1,
\]
equality in (2.2) implies (cf. [Sh, proof of Lemma 2]) that
\[
\Phi \left( \sum_{j=1}^{n+1} a_j x_j \right) = \left\| \sum_{j=1}^{n+1} \xi_j \right\|^2 \leq n \sum_{j=1}^{n+1} \| \xi_j \|^2. \tag{2.10}
\]
Let \(M = W^*\langle x_1, \ldots, x_{n+1}\rangle\). We now apply Lemma 2.3 to the projections \(E_1, \ldots, E_{n+1} \in B(L^2(M))\) introduced in [Sh, proof of Lemma 2]. That is, \(E_j\) is the projection onto \(L^2(W^*\langle x_1, \ldots, x_j, \ldots, x_{n+1}\rangle)\). Note that the subspace \(\mathcal{H}_j\) defined in Lemma 2.3,
\[
\mathcal{H}_j = \{ \xi \in L^2(M) \mid E_k \xi = \xi, \ k \neq j, E_j \xi = 0 \},
\]
is in this case exactly $L^2(W^*(x_j))$. Thus, the second identity in (2.10) and the fact that $\xi_j \perp \mathbb{C} 1$, implies that
\[
\xi_j \in \bigoplus_{i \neq j} (L^2(W^*(x_i)) \oplus \mathbb{C} 1).
\tag{2.11}
\]

With $E : L^2(M) \to L^2(M)$ the projection onto $L^2(W^*(\sum_j a_j x_j))$ we have (cf. [Sh, proof of Lemma 2]):
\[
\mathcal{J}\left(\sum_{j=1}^{n+1} a_j x_j\right) = E\left(\sum_{j=1}^{n+1} \xi_j\right).
\tag{2.12}
\]

The first identity in (2.10) then implies that $E\left(\sum_{j=1}^{n+1} \xi_j\right) = \sum_{j=1}^{n+1} \xi_j$, and so
\[
\mathcal{J}\left(\sum_{j=1}^{n+1} a_j x_j\right) = \sum_{j=1}^{n+1} \xi_j \in \bigoplus_{i=1}^{n+1} (L^2(W^*(x_i)) \oplus \mathbb{C} 1).
\]

Now choose elements $\eta_j \in L^2(W^*(x_j)) \oplus \mathbb{C} 1, 1 \leq j \leq n + 1$, such that
\[
\mathcal{J}\left(\sum_{j=1}^{n+1} a_j x_j\right) = \sum_{j=1}^{n+1} \eta_j.
\tag{2.13}
\]

Then
\[
0 = \left[\sum_{i=1}^{n+1} a_i x_i, \sum_{j=1}^{n+1} \eta_j\right] = \sum_{i \neq j} (a_i x_i \eta_j - \eta_i a_j x_j).
\]

A standard application of freeness shows that for $(i, j) \neq (k, l)$, the terms $a_i x_i \eta_j - \eta_i a_j x_j$ and $a_k x_k \eta_l - \eta_k a_l x_l$ are perpendicular elements of $L^2(M)$. Thus, the above identity implies that for all $i \neq j$,
\[
a_i x_i \eta_j = a_j \eta_i x_j.
\tag{2.14}
\]

With $L^2(W^*(x_j))^0 = L^2(W^*(x_j)) \oplus \mathbb{C} 1, 1 \leq j \leq n + 1$, consider the free product of Hilbert spaces
\[
\mathbb{C} 1 \oplus \left(\bigoplus_{p \geq 1} \left(\bigoplus_{1 \leq i_1, \ldots, i_p \leq n+1, i_1 \neq i_2 \neq \cdots \neq i_p} L^2(W^*(x_{i_1}))^0 \otimes L^2(W^*(x_{i_2}))^0 \otimes \cdots \otimes L^2(W^*(x_{i_p}))^0\right)\right),
\]

and notice that $x_i \in L^2(W^*(x_i))^0$ and $\eta_j \in L^2(W^*(x_j))^0$. It follows from unique decomposition within the free product that there is only one way that (2.14) can be fulfilled, namely when $\eta_j$ is proportional to $x_j$. That is, there exist $c_1, \ldots, c_{n+1} \in \mathbb{R}$ such that $\eta_j = c_j x_j$ and hence,
\[
\mathcal{J}\left(\sum_{j=1}^{n+1} a_j x_j\right) = \sum_{j=1}^{n+1} c_j x_j.
\tag{2.15}
\]

We can assume that $a_1, \ldots, a_{n+1} > 0$, and then by (2.14),
\[
c_j = c_1 a_j / a_1, \quad 1 \leq j \leq n + 1.
\]

In particular, all the $c_j$’s have the same sign. Taking inner product with $\sum_{j=1}^{n+1} a_j x_j$ in (2.15), we find that
\[
\sum_{j=1}^{n+1} a_j c_j = 1,
\tag{2.16}
\]
so that the $c_j$’s must be positive. Also, since $\sum_j a_j^2 = 1$, we have that $\sum_j c_j^2 \geq 1$. But
\[
\sum_{j=1}^{n+1} c_j^2 = \frac{c_1^2}{a_1^2},
\]
and so $c_1 \geq a_1$, and in general, $c_j \geq a_j$. Then by (2.16), $c_j = a_j$, and (2.3) holds. As mentioned in the introduction, this implies that $\sum_{j=1}^{n+1} a_j x_j$ is $(0,1)$-semicircular (when $\|x_1\|_2 = 1$).

2.4 Corollary. Let $x_1, \ldots, x_{n+1}$ be as in Proposition 2.2 and let $a_1, \ldots, a_{n+1} \in \mathbb{R}$ with $\sum_j a_j^2 = 1$. Then
\[
\chi\left( \sum_{j=1}^{n+1} a_j x_j \right) \geq \sum_{j=1}^{n+1} \frac{1 - a_j^2}{n} \chi\left( \frac{1}{\sqrt{1-a_j^2}} \sum_{i \neq j} a_i x_i \right). \tag{2.17}
\]
Moreover, if $\chi(\sum_{i \neq j} a_i x_i) > -\infty$ for all $j$, then equality in (2.17) implies that $\sum_j a_j x_j$ is semicircular.

Proof. The inequality (2.17) was proven by D. Shlyakhtenko in [Sh, Theorem 2]. Now, assume that $\chi(\sum_{i \neq j} a_i x_i) > -\infty$ for all $j$ and that
\[
\chi\left( \sum_{j=1}^{n+1} a_j x_j \right) = \sum_{j=1}^{n+1} \frac{1 - a_j^2}{n} \chi\left( \frac{1}{\sqrt{1-a_j^2}} \sum_{i \neq j} a_i x_i \right).
\]
Take $(0,1)$-semicirculars $s_1, \ldots, s_{n+1}$ such that $x_1, \ldots, x_{n+1}, s_1, \ldots, s_{n+1}$ are free, and put
\[
x_j(t) = x_j + \sqrt{t} s_j.
\]
Then by assumption,
\[
\int_0^\infty \left[ \sum_{j=1}^{n+1} \frac{1 - a_j^2}{n} \Phi\left( \frac{1}{\sqrt{1-a_j^2}} \sum_{i \neq j} a_i x_i^{(t)} \right) - \Phi\left( \sum_{j=1}^{n+1} a_j x_j^{(t)} \right) \right] dt = 0. \tag{2.18}
\]
Applying Proposition 2.2 with $b_j = \frac{1}{n} \sqrt{1-a_j^2}$, we see that the integrand in (2.18) is positive. Thus, (2.18) can only be fulfilled if for a.e. $t > 0$,
\[
\sum_{j=1}^{n+1} \frac{1 - a_j^2}{n} \Phi\left( \frac{1}{\sqrt{1-a_j^2}} \sum_{i \neq j} a_i x_i^{(t)} \right) = \Phi\left( \sum_{j=1}^{n+1} a_j x_j^{(t)} \right). \tag{2.19}
\]
In fact, since both sides of (2.19) are right continuous functions of $t$ (cf. [V2]), we have equality for all $t > 0$. Then by Proposition 2.2, $\sum_{j=1}^{n+1} a_j x_j^{(t)}$ is semicircular. By additivity of the $\mathcal{R}$-transform, this can only happen if $\sum_{j=1}^{n+1} a_j x_j$ is semicircular.

Proof of Theorem 2.1. The inequality (2.1) was proven by D. Shlyakhtenko in [Sh]. Now, assume that $\chi(x_1) > -\infty$ and that
\[
\chi\left( \frac{x_1 + \cdots + x_{n+1}}{\sqrt{n+1}} \right) = \chi\left( \frac{x_1 + \cdots + x_n}{\sqrt{n}} \right).
\]
If we replace \( x_j \) by \( \frac{x_j - \tau(x_j)}{\|x_j - \tau(x_j)\|_2} \), we will still have equality. Hence, we will assume that \( \tau(x_j) = 0 \) and that \( \|x_j\|_2 = 1 \). Now,

\[
\chi\left(\frac{x_1 + \cdots + x_{n+1}}{\sqrt{n+1}}\right) = \frac{1}{n+1} \sum_{j=1}^{n+1} \chi\left(\frac{1}{\sqrt{n}} \sum_{i \neq j} x_i\right),
\]

and by application of Corollary 2.4 with \( a_j = \frac{1}{\sqrt{n+1}} \frac{x_1 + \cdots + x_{n+1}}{\sqrt{n+1}} \) must be semicircular. Additivity of the \( \mathcal{R} \)-transform tells us that this can only happen if \( x_1 \) is semicircular. ■

We would like to thank Serban Belinschi for pointing out to us the following consequence of Theorem 2.1:

2.5 Corollary. Among the freely stable compactly supported probability measures on \( \mathbb{R} \), the semicircle laws are the only ones with finite free entropy.

Proof. By definition, a compactly supported probability measure \( \mu \) on \( \mathbb{R} \) is freely stable if for all \( n \in \mathbb{N} \), there exist \( a_n > 0, b_n \in \mathbb{R} \), such that if \( x_1, \ldots, x_n \) are freely independent self-adjoint elements which are distributed according to \( \mu \), then

\[
\frac{1}{a_n} (x_1 + \cdots + x_n) + b_n
\]

has distribution \( \mu \). Note that the set of freely stable laws is invariant under transformations by the affine maps \( (\phi_{s,r})_{s \in \mathbb{R}, r > 0} \), where

\[
\phi_{s,r}(t) = \frac{t - s}{r}, \quad t \in \mathbb{R}.
\]

Also, by [VDN, p. 27], the semicircle laws are freely stable.

Suppose now that \( \mu \) is a freely stable compactly supported probability measure on \( \mathbb{R} \). By the above remarks, we can assume that \( \mu \) has first moment 0 and variance 1.

Let \( x_1, x_2 \) be freely independent self-adjoint elements in distributed according to \( \mu \). Since \( \mu \) is freely stable, \( \frac{x_1 + x_2}{\sqrt{2}} \) has distribution \( \mu \) as well (by the assumptions on \( \mu \), \( a_2 = \sqrt{2} \) and \( b_2 = 0 \)). But then

\[
\chi\left(\frac{x_1 + x_2}{\sqrt{2}}\right) = \chi(x_1),
\]

and by Theorem 2.1, either \( \chi(x_1) = -\infty \), or \( x_1 \) is semicircular. ■

3 The Classical Case.

In this section we are going to prove the classical analogue of Theorem 2.1:

3.1 Theorem. Let \( n \in \mathbb{N} \), and let \( X_1, \ldots, X_{n+1} \) be i.i.d. random variables. Then

\[
H\left(\frac{X_1 + \cdots + X_{n+1}}{\sqrt{n+1}}\right) \geq H\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right).
\]

Moreover, if \( H(X_1) > -\infty \) and if (3.1) is an equality, then \( X_1 \) is Gaussian.
3.2 Lemma. Let $n \in \mathbb{N}$. Then for every $m \in \mathbb{N}$, the $m$'th Hermite polynomial, $H_m$, satisfies:

$$n^m H_m\left(\frac{x_1 + \cdots + x_n}{\sqrt{n}}\right) = \sum_{k_1, \ldots, k_n \geq 0, \sum_j k_j = m} \frac{m!}{k_1! \cdots k_n!} H_{k_1}(x_1) H_{k_2}(x_2) \cdots H_{k_n}(x_n).$$  \hfill (3.2)

Sketch of proof. (3.2) holds for $m = 0$ ($H_0 = 1$) and for $m = 1$ ($H_1(x) = 2x$). Now, for general $m \in \mathbb{N}$,

$$H_{m+1}(x) = 2x H_m(x) - 2m H_{m-1}(x).$$

(3.2) then follows by induction over $m$. ■

3.3 Lemma. Let $\mu \in \text{Prob}(\mathbb{R})$ be absolutely continuous w.r.t. Lebesgue measure, and let $\sigma_t \in \text{Prob}(\mathbb{R})$ denote the Gaussian distribution with mean $0$ and variance $t$. Then if $\mu((-\infty,0]) \neq 0$ and $\mu([0,\infty)) \neq 0$, the following inclusion holds:

$$L^2(\mathbb{R}, \mu \ast \sigma_t) \subseteq L^2(\mathbb{R}, \sigma_t).$$  \hfill (3.3)

Proof. Let $f \in L^1(\mathbb{R})$ denote the density of $\mu$ w.r.t. Lebesgue measure. Then the density of $\mu \ast \sigma_t$ is given by

$$\frac{d(\mu \ast \sigma_t)}{ds} (s) = \frac{1}{\sqrt{2\pi t}} \left( \int_{-\infty}^{\infty} f(u) \cdot e^{-\frac{u^2}{2t}} \cdot e^{su} du \right) \cdot e^{-\frac{s^2}{2t}}$$

$$= \phi(s) \cdot \frac{d\sigma_t}{ds} (s),$$

where

$$\phi(s) = \int_{-\infty}^{\infty} f(u) \cdot e^{-\frac{u^2}{2t}} \cdot e^{su} du.$$  \hfill (3.4)

It follows that if $\phi$ is bounded away from 0, then (3.3) holds. For $s \geq 0$ we have that

$$\phi(s) \geq \int_{0}^{\infty} f(u) \cdot e^{-\frac{u^2}{2t}} \cdot e^{su} du$$

$$\geq \int_{0}^{\infty} f(u) \cdot e^{-\frac{u^2}{2t}} du,$$

and similarly for $s \leq 0$:

$$\phi(s) \geq \int_{-\infty}^{0} f(u) \cdot e^{-\frac{u^2}{2t}} du.$$  

Since $\int_{-\infty}^{0} f(u) du > 0$ and $\int_{0}^{\infty} f(u) du > 0$, both of the integrals $\int_{0}^{\infty} f(u) \cdot e^{-\frac{u^2}{2t}} du$ and $\int_{-\infty}^{0} f(u) \cdot e^{-\frac{u^2}{2t}} du$ are strictly positive. This completes the proof. ■

Proof of Theorem 3.1. The inequality (3.1) was proven in [ABBN]. Now, suppose $H(X_1) > -\infty$ and that (3.1) is an equality. We can assume that $X_1$ has first moment $0$ and second moment $1$. Take Gaussian random variables $G_1, \ldots, G_{n+1}$ of mean $0$ and variance $1$ such that $X_1, \ldots, X_{n+1}, G_1, \ldots, G_n, G_{n+1}$ are independent. Then with

$$X_j^{(t)} = X_j + \sqrt{t} G_j,$$
By Lemma 3.2, this implies that
\[ H\left(\frac{X_1+\ldots+X_{n+1}}{\sqrt{n+1}}\right) = \frac{1}{2} \int_0^\infty \left[ \frac{1}{1+t} - \left\| j\left(\frac{X_1^{(t)}+\ldots+X_{n+1}^{(t)}}{\sqrt{n+1}}\right)\right\|_2^2 \right] dt + \frac{1}{2} \log(2\pi e), \tag{3.5} \]
where
\[ j\left(\frac{X_1^{(t)}+\ldots+X_{n+1}^{(t)}}{\sqrt{n+1}}\right) = \left(\frac{d}{dx}\right)^* (1) \in L^2\left(\mathbb{R}, \mu_{\frac{X_1^{(t)}+\ldots+X_{n+1}^{(t)}}{\sqrt{n+1}}}, \phi\right) \tag{3.6} \]
is the score function. Since \( X_1 \) has mean 0 and finite entropy, \( \mu_{X_1} \) and \( \mu_{\frac{X_1+\ldots+X_{n+1}}{\sqrt{n+1}}} \) satisfy the conditions of Lemma 3.3.

For \( t > 0 \), define \( f^{(t)} \in L^2(\mathbb{R}^{n+1}, \otimes_{j=1}^{n+1} \mu_{X_j^{(t)}}) \) by
\[ f^{(t)}(x_1, \ldots, x_{n+1}) = j\left(\frac{X_1^{(t)}+\ldots+X_{n+1}^{(t)}}{\sqrt{n+1}}\right)\left(\frac{x_1^{(t)}+\ldots+x_{n+1}^{(t)}}{\sqrt{n+1}}\right). \]
As in the free case (cf. (2.13)) equality in (3.1) implies that for each \( t > 0 \) there exists a function \( g^{(t)} \in L^2(\mu_{X_1^{(t)}}) \) such that \( \int g^{(t)} d\mu_{X_1^{(t)}} = 0 \) and
\[ f^{(t)}(x_1, \ldots, x_{n+1}) = \sum_{j=1}^{n+1} g^{(t)}(x_j). \tag{3.7} \]
Because of Lemma 3.3 we can now write things in terms of the Hermite polynomials \((H_m)_{m=0}^\infty\). That is, there exist scalars \((\alpha_m)_{m=1}^\infty\) and \((\beta_m)_{m=1}^\infty\) such that
\[ f^{(1)}(x_1, \ldots, x_{n+1}) = \sum_{m=1}^{\infty} \alpha_m H_m\left(\frac{x_1+\ldots+x_{n+1}}{\sqrt{n+1}}\right), \]
and
\[ g^{(1)}(x) = \sum_{m=1}^{\infty} \beta_m H_m(x). \]

By Lemma 3.2, this implies that
\[ \sum_{j=1}^{n+1} \sum_{m=1}^{\infty} \beta_m H_m(x_j) = \sum_{m=1}^{\infty} \frac{\alpha_m}{(n+1)^{\frac{m}{2}}} \sum_{k_1, \ldots, k_{n+1} \geq 0, \sum_j k_j = m} \frac{m!}{k_1!k_2!\cdots k_{n+1}!} H_{k_1}(x_1) H_{k_2}(x_2) \cdots H_{k_{n+1}}(x_{n+1}). \tag{3.8} \]
The functions \((H_{k_1}(x_1) H_{k_2}(x_2) \cdots H_{k_{n+1}}(x_{n+1}))_{k_1,\ldots,k_{n+1}\geq0}\) are mutually perpendicular in \( L^2(\mathbb{R}^{n+1}, \otimes_{j=1}^{n+1} \sigma_1) \).

Fix \( m \geq 2 \), and take \( k_1, \ldots, k_{n+1} \) with \( \sum_j k_j = m \) and \( k_j \geq 1 \) for at least two \( j \)'s. Then take inner product with \( H_{k_1}(x_1) H_{k_2}(x_2) \cdots H_{k_{n+1}}(x_{n+1}) \) on both sides of (3.8) to see that \( \alpha_m \) must be zero. That is,
\[ j\left(\frac{X_1^{(1)}+\ldots+X_{n+1}^{(1)}}{\sqrt{n+1}}\right)\left(\frac{x_1^{(1)}+\ldots+x_{n+1}^{(1)}}{\sqrt{n+1}}\right) = \alpha_1 H_1\left(\frac{x_1^{(1)}+\ldots+x_{n+1}^{(1)}}{\sqrt{n+1}}\right) = 2\alpha_1 \frac{x_1^{(1)}+\ldots+x_{n+1}^{(1)}}{\sqrt{n+1}}. \]
Since the score function of a random variable \( X \), \( j(X) \), satisfies \( \langle j(X), X \rangle_{L^2(\mu_X)} = 1 \), we have that \( \alpha_1 = \frac{1}{2} \), and so
\[ j\left(\frac{X_1^{(1)}+\ldots+X_{n+1}^{(1)}}{\sqrt{n+1}}\right)\left(\frac{x_1^{(1)}+\ldots+x_{n+1}^{(1)}}{\sqrt{n+1}}\right) = \frac{x_1^{(1)}+\ldots+x_{n+1}^{(1)}}{\sqrt{n+1}}. \]
Then \( \frac{X_1^{(1)} + \cdots + X_{n+1}^{(1)}}{\sqrt{n+1}} \) has Fisher information 1, implying that it is standard Gaussian. As in the free case, using additivity of the logarithm of the Fourier transform, this can only happen if \( X_1 \) is Gaussian. ■

References

[ABBN] S. Artstein, K. Ball, F. Barthe, A. Naor, 'Solution of Shannon’s problem on monotonicity of entropy', \textit{Journal Amer. Math. Soc.} \textbf{17} (2004), 975–982.

[Sh] D. Shlyakhtenko, 'A free analogue of Shannon’s problem on monotonicity of entropy'. Preprint at http://xxx.lanl.gov/abs/math.OA/0510103.

[St] A. J. Stam, 'Some inequalities satisfied by the quantities of information of Fisher and Shannon', \textit{Information \\& Control} \textbf{2} (1959), 101–112.

[VDN] D. V. Voiculescu, K. Dykema, A. Nica, 'Free Random Variables', CMR Monograph Series 1, \textit{American Mathematical Society} (1992).

[V1] D. V. Voiculescu, 'The analogues of entropy and of Fisher’s information measure in free probability theory I', \textit{Comm. Math. Phys.} \textbf{155} (1993), 71–92.

[V2] D. V. Voiculescu, 'The analogues of entropy and of Fisher’s information measure in free probability theory V', \textit{Invent. Math.} \textbf{132} (1998), 189–227.

[V3] D. V. Voiculescu, 'Free entropy', \textit{Bull. London Math. Soc.} \textbf{34} (2002), no. 3, 257–278.

Hanne Schultz,
Department of Mathematics and Computer Science,
University of Southern Denmark,
Campusvej 55,
5230 Odense M, Denmark
schultz@imada.sdu.dk