THE NONCOMMUTATIVE HILBERT TRANSFORM APPROACH TO FREE ENTROPY

DAN VOICULESCU

Department of Mathematics, University of California
Berkeley, California 94720-3840, U.S.A.
E-mail: dvv@math.berkeley.edu

Free entropy is the entropy quantity with the right behavior with respect to free independence. We began the study of free entropy with the one-variable case [7], taking as definition, based on random matrix heuristics, the negative of the logarithmic energy of the distribution. We then found in [8] a general definition of $\chi(X_1, \ldots, X_n)$, the free entropy of an n-tuple of noncommutative random variables in a tracial W^*-probability space, using matricial microstates. The parallel to classical entropy suggests a number of natural properties free entropy should satisfy and several of these have been established ([7, 8, 10, 4, 6]). This new machinery has had striking applications to the solution of some old problems on von Neumann algebras ([9, 2, 3, 1]). On the other hand, from the point of view of this parallel to classical entropy, the theory is still incomplete, which keeps certain further applications to von Neumann algebras out of reach. The main reason for the difficulties is that we know very little about matrix-approximants to elements in type II_1 von Neumann algebras. Even the general existence question for such approximants, which would be a first step in this direction, is unsolved and coincides with Alain Connes' well known problem about embedding type II_1 factors into the ultraproduct of the hyperfinite II_1 factor. Under these circumstances we began in [11] to look for another approach to free entropy which avoids matricial microstates. This should not be viewed only as an alternative to microstates, since some of the results actually can be used within the microstates approach.

This note summarizes the lecture we intended to give at the Gdańsk Quantum Probability meeting about our current work towards a "microstates-free" approach to free entropy [11].

---

1. Background: the definition via microstates. Throughout, (M, τ) will denote a tracial W^*-probability space, i.e. a von Neumann algebra M endowed with a faithful normal trace state τ. We recall here the definition we gave in [8] of the free entropy $\chi(X_1, \ldots, X_n)$ where $X_j = X_j^* \in M$, $1 \leq j \leq n$. The set of microstates $\Gamma_R(X_1, \ldots, X_n; m, k, \varepsilon)$ consists of n-tuples $(A_1, \ldots, A_n) \in (\mathcal{M}_k^{sa})^n$ of self-adjoint $k \times k$-matrices such that $\|A_j\| < R$ and $|k^{-1}\mathrm{Tr}A_{i_1} \ldots A_{i_p} - \tau(X_{i_1} \ldots X_{i_p})| < \varepsilon$, $(1 \leq p \leq m)$. The free entropy $\chi(X_1, \ldots, X_n)$ is obtained by taking

$$\limsup_{k \to \infty}(k^{-2} \log \mathrm{vol}\, \Gamma_R(X_1, \ldots, X_n; m, k, \varepsilon) + \frac{n}{2} \log k)$$

(vol denoting the euclidean volume with respect to the Hilbert-Schmidt norm) followed by

$$\sup_{R>0} \inf_{m \in \mathbb{N}} \inf_{\varepsilon>0}.$$

The definition of $\chi(X_1, \ldots, X_n)$ via microstates is inspired by Boltzmann's formula $S = k \log W$ and the fact that free independence occurs asymptotically in random matrices ([5]). Note that classical entropy can also be obtained along similar lines. If (M, τ) is commutative, i.e. the L^∞-space over some probability measure space (Ω, μ) we define the microstates Γ similarly, with the only modification that $(A_1, \ldots, A_n) \in (\Delta_k^{sa})^n$ where Δ_k^{sa} denotes the diagonal self-adjoint $k \times k$-matrices. Going again through an appropriate normalized limit, etc., we obtain the classical entropy

$$H(X_1, \ldots, X_n) = - \int_{\mathbb{R}^n} p(t_1, \ldots, t_n) \log p(t_1, \ldots, t_n) dt_1 \ldots dt_n$$

if τ (spectral measure of $(X_1, \ldots, X_n)$) has density $p(t_1, \ldots, t_n)$ with respect to Lebesgue measure.

All this shows that the microstates approach is legitimate and rooted in statistical physics. The main drawback is that it is a rather "philosophical" definition. To prove, for instance, the additivity of free entropy for free random variables requires the full power of the asymptotic freeness results for random matrices ([8, 5, 6]). Also, for the classical entropy, the basic properties are established using the formula $- \int p \log p \, dt$, rather than going back to microstates.

2. Fisher information versus entropy. The classical Fisher information $J(f)$ is defined by:

$$J(f) = \lim_{\varepsilon \downarrow 0} \varepsilon^{-1}(H(f + \varepsilon^{\frac{1}{2}}g) - H(f))$$

where f and g are independent random variables and g is (0,1)-Gaussian. If the distribution of f has density p, then

$$J(f) = \int \frac{(p'(t))^2}{p(t)} \, dt \, .$$

Correspondingly, the free Fisher information [7] for one variable is

$$\Phi(X) = \lim_{\varepsilon \downarrow 0} \varepsilon^{-1}(\chi(X + \varepsilon^{\frac{1}{2}}S) - \chi(X))$$

where $X = X^* \in M$, $S \in M$ and X, S are free and S is (0,1)-semicircular. It follows (see [7]) that

$$\Phi(X) = \tfrac{2}{3} \int (p(t))^3 dt$$

when the distribution μ_X of X has density $p(t)$ and otherwise $\phi(X) = +\infty$.

Thus, the Fisher information (classical or free) is a kind of derivative of the entropy. Hence entropy and Fisher information are equivalent quantities, in the sense that a theory of one quantity can be obtained from a theory of the other.

The microstates-free approach to free entropy will have the free Fisher information as the primary quantity.

3. The idea of the microstates-free approach for one scalar variable. The formula $\Phi(X) = \tfrac{2}{3} \int (p(t))^3 dt$ involves the Lebesgue integral for which there is not an obvious generalization. The way out is provided by an identity for the Hilbert transform:

$$\int p^3(t) = 3 \int (Hp)^2 p \, dt.$$

Here Hp is the Hilbert transform of p, i.e., if $d\mu = p \, dt$, then

$$Hp(s) = c \int \frac{d\mu(t)}{s - t}$$

where c is a constant and the integral is in the principal value sense). Hence up to constants $\Phi(X)$ coincides with $\|Hp\|^2_{L^2(\mu)}$.

Moreover the Hilbert transform of p viewed as an element of $L^2(\mu)$ can be expressed using the difference quotient derivation:

$$(\partial f)(s, t) = \frac{f(s) - f(t)}{s - t}$$

viewed as an unbounded operator from $L^2(\mu)$ to $L^2(\mu) \otimes L^2(\mu)$. Indeed, then we have

$$H\mu = \partial^*(1 \otimes 1).$$

This simple idea, suitably generalized, is the basis for the microstates-free approach.

4. Noncommutative Hilbert transforms ([11]). If $1 \in B \subset M$ is a $*$-subalgebra and $X = X^* \in M$ let $B[X]$ denote the subalgebra generated by B and X. If X and B are algebraically free (i.e. there is no nontrivial algebraic relation between X and B) let

$$\partial_X : B[X] \to B[X] \otimes B[X]$$

be the derivation of $B[X]$ into the $B[X]$-bimodule $B[X] \otimes B[X]$ (here $m_1(a_1 \otimes a_2)m_2 = m_1 a_1 \otimes a_2 m_2$) so that $\partial_X(X) = 1 \otimes 1$ and $\partial_X(B) = 0$. This means

$$\partial_X(b_0 X b_1 X \dots b_n) = \sum_{j=1}^{n} b_0 X \dots b_{j-1} \otimes b_j X \dots b_n.$$

DEFINITION 1. An affiliated operator with $W^*(B[X])$, $\xi \in L^1(W^*(B[X]))$, is called the conjugate of X with respect to B, denoted $\mathcal{J}(X : B)$, if

$$\tau(\xi m) = (\tau \otimes \tau)(\partial_X m)$$

for all $m \in B[X]$.

R e m a r k s. 1) The definition is equivalent to $\xi = \partial_X^*(1 \otimes 1)$.

2) If $B = \mathbb{C}$, then $\partial_X : \mathbb{C}[X] \to \mathbb{C}[X] \otimes \mathbb{C}[X]$ coincides with the difference quotient.

Here is a list of the main results we have obtained concerning conjugates (i.e. non-commutative Hilbert transforms).

FACTS. 1. If $B = \mathbb{C}$ and μ is the distribution of X and if $d\mu = p\, dt$ with $p \in L^3$ with respect to Lebesgue measure, then $\mathcal{J}(X : \mathbb{C}) = g(X)$, where $g = 2\pi\, Hp$.

2. If $1 \in B \subset M$, $1 \in C \subset M$ are $*$-subalgebras and C and $B[X]$ are free in (M, τ), then

$$\mathcal{J}(X : B) = \mathcal{J}(X : W^*(B \cup C))$$

3. If $1 \in B \subset M$, $1 \in C \subset M$ are $*$-subalgebras, $X_j = X_j^* \in M$, $Y_j = Y_j^* \in M$, $1 \leq j \leq n$ and $B[X_1, \ldots, X_n]$ and $C[Y_1, \ldots, Y_n]$ are free, then

$$\mathcal{J}(X_1 + Y_1 : (B \vee C)[X_2 + Y_2, \ldots, X_n + Y_n])$$
$$= \quad E_{(B \vee C)[X_1 + Y_1, \ldots, X_n + Y_n]} \mathcal{J}(X_1 : B[X_2, \ldots, X_n]).$$

(Here $B \vee C$ is the algebra generated by B and C and E denotes the conditional expectation.)

4. If S is (0,1)-semicircular and S and $B[X]$ are free in (M, τ), then $\mathcal{J}(X + \varepsilon S : B) = \varepsilon^{-1} E_{B[X + \varepsilon S]} S$. In particular,

$$\|\mathcal{J}(X + \varepsilon S : B)\| \leq 2\varepsilon^{-1} \ .$$

5. Let S be (0,1)-semicircular and free with respect to $B[X]$ and assume $\|\mathcal{J}(X : B)\| < \infty$. Then

$$\tau(b_0(X + \frac{\varepsilon}{2}\mathcal{J}(X : B))b_1(X + \frac{\varepsilon}{2}\mathcal{J}(X : B)) \ldots b_n)$$
$$= \quad \tau(b_0(X + \varepsilon^{\frac{1}{2}} S)b_1(X + \varepsilon^{\frac{1}{2}} S) \ldots b_n) + O(\varepsilon^2).$$

6. Let $X_j = X_j^* \in M$, $1 \leq j \leq n$ and assume $\chi(X_1, \ldots, X_n) > -\infty$ and $\mathcal{J}(X_k : \mathbb{C}[X_1, \ldots, X_{k-1}, X_{k+1}, \ldots, X_n])$ for $1 \leq k \leq n$ exist. Let further $P_j = P_j^* \in \mathbb{C}[X_1, \ldots, X_n]$. Then

$$\frac{d}{d\varepsilon}\chi(X_1 + \varepsilon P_1, \ldots, X_n + \varepsilon P_n)|_{\varepsilon=0}$$
$$= \quad \sum_{1 \leq j \leq n} \tau(P_j \mathcal{J}(X_j : \mathbb{C}[X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_n]).$$

Some comments on these properties are in order. 1 shows that $\mathcal{J}(X : B)$ is a generalization of the usual Hilbert transform. 2 and 3 show that $\mathcal{J}(X : B)$ has certain nice properties with respect to free independence. 4 is quite important technically: small semicircular perturbations regularize the noncommutative Hilbert transform and assure its existence and boundedness. 5 connects $\mathcal{J}(X : B)$ with the free Brownian motion starting at X. 6 establishes an infinitesimal connection between free entropy and noncommutative Hilbert transforms. The partial noncommutative Hilbert transform $(\mathcal{J}(X_j : \mathcal{C}[X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_n])_{1 \leq j \leq n}$ acts as a gradient of the free entropy.

5. The relative free Fisher information ([11]). With the noncommutative Hilbert transform at hand we can now pass to the free Fisher information, i.e. to the primary quantity of the microstates-free approach. At present this approach takes the form of a parallel development with some bridges between the two. Further down the road, I expect more results relating the two approaches. In view of this we had to devise some parallel rules for notations. Roughly, the corresponding quantities in the new approach will have an asterisk. So χ denotes entropy defined via microstates, while χ^* will denote the entropy in the new approach and similarly Φ is the free Fisher information that goes with χ, while Φ^* will be the Fisher information of the new approach. Note however that in $\chi(\dots : \cdots)$ and $\chi^*(\dots : \cdots)$ the colon sign will have a different meaning.

DEFINITION 2. If $1 \in B \subset M$ is a $*$-subalgebra and $X_j = X_j^* \in M$, $1 \le j \le n$, the relative free Fisher information of $X_1, \dots, X_n$ with respect to B is given by

$$\Phi^*(X_1, \dots, X_n : B) = \sum_{1 \le j \le n} |\mathcal{J}(X_j : B[X_1, \dots, X_{j-1}, X_{j+1}, \dots X_n])|_2^2$$

if $\mathcal{J}(X_j : B[X_1, \dots, X_{j-1}, X_{j+1}, \dots X_n]) \in L^2(M)$, $(1 \le j \le n)$ and is otherwise equal to $+\infty$.

FACTS. 1. If C and $B[X_1, \dots, X_n]$ are free in (M, τ) then

$$\Phi^*(X_1, \dots, X_n : B \vee C) = \Phi^*(X_1, \dots, X_n : B).$$

2. If $B[X_1, \dots, X_n]$ and $C[Y_1, \dots, Y_m]$ are free, then

$$\Phi^*(X_1, \dots, X_n, Y_1, \dots, Y_m : B \vee C) = \Phi^*(X_1, \dots, X_n : B) + \Phi^*(Y_1, \dots, Y_m : C).$$

3. Without a freeness assumption in the conclusion of (2), the equal sign is replaced by $\ge$.

4. (Free analogue of the Stam inequality.) If $B[X_1, \dots, X_n]$ and $C[Y_1, \dots, Y_n]$ are free, then

$$(\Phi^*(X_1 + Y_1, \dots, X_n + Y_n : B \vee C))^{-1}$$
$$\ge (\Phi^*(X_1, \dots, X_n : B))^{-1} + (\Phi^*(Y_1, \dots, Y_n : C))^{-1}.$$

5. (Free analogue of the Cramer-Rao inequality.)

$$\Phi^*(X_1, \dots, X_n : B)\tau(X_1^2 + \dots + X_1^2) \ge n^2.$$

Equality holds if and only if $B, \{X_1\}, \dots, \{X_n\}$ are free and the X_j's have equal centered semicircular distributions.

6. If each S_j is $(0,1)$ semicircular $(1 \le j \le n)$ and $B[X_1, \dots, X_n]$, $\{S_1\}, \dots, \{S_n\}$ are free, then

$$[0, \infty) \ni t \to \Phi^*(X_1 + t^{\frac{1}{2}}S_1, \dots, X_n + t^{\frac{1}{2}}S_n : B) \in (0, \infty]$$

is decreasing and right continuous. Moreover

$$n^2(C^2 + nt)^{-1} \le \Phi^*(X_1 + t^{\frac{1}{2}}S_1, \dots X_n + t^{\frac{1}{2}}S_n : B) \le nt^{-1}$$

where $C^2 = \tau(X_1^2 + \dots + X_n^2)$.

6. The relative free entropy ([11])

DEFINITION 3. Let $1 \in B \subset M$ be a $*$-subalgebra and $X_j = X_j^* \in M$, $1 \leq j \leq n$. The relative free entropy of $X_1, \ldots, X_n$ with respect to B is given by

$$\chi^*(X_1, \ldots, X_n : B) = \tfrac{1}{2} \int_0^\infty \left(\frac{n}{1+t} - \Phi^*(X_1 + t^{\frac{1}{2}} S_1, \ldots, X_n + t^{\frac{1}{2}} S_n : B) \right) dt + \frac{n}{2} \log 2\pi e$$

where the S_j are (0,1)-semicircular and $B[X_1, \ldots, X_n], \{S_1\}, \ldots, \{S_n\}$ are free.

FACTS. 1. $\chi^*(X : \mathbb{C}) = \chi(X)$.

2. $\chi^*(X_1, \ldots, X_n : B) \leq \frac{n}{2} \log(2\pi e n^{-1} C)$ where $C^2 = \tau(X_1^2 + \ldots + X_n^2)$.

3. $\chi^*(X_1, \ldots, X_n, Y_1, \ldots, Y_m : B \vee C) \leq \chi^*(X_1, \ldots, X_n : B) + \chi^*(Y_1, \ldots, Y_m : C)$.

4. If $B[X_1, \ldots, X_n], C[Y_1, \ldots, Y_m]$ are free, then

$$\chi^*(X_1, \ldots, X_n : B) + \chi^*(Y_1, \ldots, Y_m : C) = \chi^*(X_1, \ldots, X_n, Y_1, \ldots, Y_m : B \vee C).$$

5. If $X_j^{(k)} = X_j^{(k)^*} \in M$, $1 \leq j \leq n$, $k \in \mathbb{N}$ and

$$s - \lim_{k \to \infty} X_j^{(k)} = X_j, \qquad (1 \leq j \leq n)$$

then

$$\limsup_{k \to \infty} \chi^*(X_1^{(k)}, \ldots X_n^{(k)} : B) \leq \chi^*(X_1, \ldots, X_n : B)$$

6. If the S_j are (0,1)-semicircular and $B[X_1, \ldots, X_n], \{S_1\}, \ldots, \{S_n\}$ are free let

$$h(t) = \chi^*(X_1 + t^{\frac{1}{2}} S_1, \ldots, X_n + t^{\frac{1}{2}} S_n : B).$$

Then $h : [0, \infty) \to \mathbb{R} \cup \{-\infty\}$ is concave, continuous, increasing and $h(t) \geq \frac{n}{2} \log(2\pi e t)$. Moreover

$$\lim_{\varepsilon \downarrow 0} \varepsilon^{-1}(h(t + \varepsilon) - h(t)) = \tfrac{1}{2} \Phi^*(X_1 + t^{\frac{1}{2}} S_1, \ldots, X_n + t^{\frac{1}{2}} S_n : B).$$

References

[1]   K. D y k e m a, *Two applications of free entropy*, preprint.

[2]   L. G e, *Applications of free entropy to finite von Neumann algebras*, *I*, American J. Math. 119 (1997), 467–485.

[3]   L. G e, *Applications of free entropy to finite von Neumann algebras*, *II*, Ann. of Math. 147 (1998), 143–157.

[4]   S. J. Szarek and D. Voiculescu, *Volumes of restricted Minkowski sums and the free analogue of the entropy power inequality*, Commun. Math. Phys. 178 (1996) 563–570.

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

[6]   D. Voiculescu, *A strengthened asymptotic freeness result for random matrices with applications to free entropy*, Internat. Math. Res. Notices 1998, no. 1, 41–63.

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

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

[9] D. Voiculescu, *The analogues of entropy and of Fisher's information measure in free probability theory*, *III*: *the absence of Cartan subalgebras*, Geometric and Functional Analysis, vol. 6, no. 1 (1996), 172–199.

[10] D. Voiculescu, *The analogues of entropy and of Fisher's information measure in free probability theory*, *IV*: *maximum entropy and freeness*, in: Free Probability, D. Voiculescu (ed.), Fields Institute Communications, Vol. 12, American Mathematical Society, Providence (1996).

[11] D. Voiculescu, *The analogues of entropy and of Fisher's information measure in free probability theory*, *V*: *noncommutative Hilbert transforms*, Invent. Math. 132 (1998), 189–227.

[12] D. Voiculescu, K. Dykema and A. Nica, *Free Random Variables*, CRM Monograph Series, Vol. I, American Mathematical Society (1992).