Acta Mathematica Sinica, English Series
July, 2003, Vol.19, No.3, pp. 1–11

Acta Mathematica Sinica, English Series

Logarithmic Sobolev Inequalities, Matrix Models and Free Entropy

Philippe BIANE
CNRS, Département de Mathématiques et Applications, École Normale Supérieure, 45,
rue d'Ulm 75005 *Paris, France*
E-mail: Philippe.Biane@ens.fr

Abstract We give two applications of logarithmic Sobolev inequalities to matrix models and free probability. We also provide a new characterization of semi-circular systems through a Poincaré-type inequality.

Keywords
MR(2000)Subject Classification 46L54; 60G15, 94A17

1 Introduction

In this paper we explore several connections between Voiculescu's free entropy [1,2] and random matrices. After recalling some results due to Bakry, Émery and Ledoux in Section 2, we use the logarithmic Sobolev inequality for Gibbs measures on matrix spaces, with a uniformly convex potential in order to generalize an inequality of Voiculescu between the free entropy and the free Fisher information of probability measures on the real line. We use again logarithmic Sobolev inequalities through the Bakry–Émery criterion, this time on unitary groups, and its applications to concentration inequalities, in order to construct non-commutative random variables with prescribed conjugate variables. Finally in the last section we compute the second derivative of the free entropy functional on the trace state space, at some point with finite Fisher information, which we identify with a Dirichlet form on the L^2 space of the trace state. Then we give a characterization of semi-circular systems by a Poincaré inequality for this Dirichlet form.

2 Logarithmic Sobolev Inequalities and Concentration of Measure

If μ is a probability measure and f a positive measurable function on $\mathbb{R}$, let

$$\mathrm{Ent}_\mu(f) = \int_{\mathbb{R}} f \log f \, d\mu - \int_{\mathbb{R}} f \, d\mu \log \left(\int_{\mathbb{R}} f \, d\mu \right).$$

Received February 20, 2002, Accepted October 25, 2002

Theorem 2.1 *Let* V *be a convex function on* $\mathbb{R}^n$ *with* $V'' \geq c\,\mathrm{Id}$, *and let* $\mu(dx) = \frac{1}{Z}e^{-V(x)}dx$ *be a probability measure. Then one has, for every function* f *on* $\mathbb{R}^n$,

$$\mathrm{Ent}_\mu(f^2) \leq \frac{2}{c}\int |\nabla f(x)|^2 \mu(dx).$$

For a proof see Section 2.2 of [3].

In the following (M, g) is a compact Riemannian manifold, with Riemannian measure dx, while Ric is its Ricci curvature, and $\mu(dx) = \frac{1}{Z}e^{-V(x)}dx$ is a probability measure with V a smooth function whose Hessian we denote by $\mathrm{Hess}(V)$.

Theorem 2.2 *Assume that* $\mathrm{Ric} + \mathrm{Hess}(V) \geq c\,g$ *for some positive constant* c. *Then the following logarithmic Sobolev inequality holds for any smooth function* f *on* M,

$$\mathrm{Ent}_\mu(f^2) \leq \frac{2}{c}\int |\nabla f(x)|^2 \mu(dx).$$

This follows from the Bakry–Émery Γ_2 criterion, see [4]. The next result is a consequence of Theorem 2.2 and Corollary 2.6, Section 2.3 of [3].

Theorem 2.3 *In the situation of Theorem* 2.2, *for every function* F *with* $\|\nabla F\|_\infty \leq 1$ *one has* $\mu(F \geq E_\mu(F) + r) \leq e^{-cr^2/4}$.

3 Some Free Logarithmic Sobolev Inequalities

The purpose of this section is to establish a generalization of Voiculescu's inequality between free Fisher information and free entropy for measures on the real line, proved in [2]. The main result is Theorem 3.1 below.

Let μ be a probability measure on $\mathbb{R}$ with a density $p \in L^3(\mathbb{R})$ and let

$$Hp(x) = \lim_{\varepsilon \to 0} 2\int_{\mathbb{R}} \frac{(x-y)}{(x-y)^2 + \varepsilon^2} p(y)dy,$$

where the limit exists in $L^3(\mathbb{R})$. Then the free Fisher information of the measure μ is

$$\Phi(\mu) = \int_{\mathbb{R}} (Hp)^2(x)\mu(dx)$$

and its free entropy is

$$\chi(\mu) = \int\int_{\mathbb{R}} \log|x-y|\mu(dx)\mu(dy) + \frac{3}{4} + \frac{1}{2}\log(2\pi),$$

see [2], where the following inequality is shown to hold between these two quantities:

$$\chi(\mu) \geq \frac{1}{2}\log\left(\frac{2\pi e}{\Phi(\mu)}\right). \tag{3.1}$$

One can recast this inequality by using the semi-circular distribution $\sigma(dx) = \frac{1}{2\pi}\sqrt{4-x^2}\,dx$ on $[-2, +2]$. Recall that σ is the unique distribution, among probability measures on $\mathbb{R}$, which maximizes the quantity

$$\mathcal{E}(\mu) = \chi(\mu) - \frac{1}{2}\int_{\mathbb{R}} x^2\,\mu(dx).$$

Let us introduce the following modified quantities:

$$\tilde{\Sigma}(\mu) = \mathcal{E}(\sigma) - \mathcal{E}(\mu) = \chi(\sigma) - \frac{1}{2}\int_{\mathbb{R}} x^2\sigma(dx) - \chi(\mu) + \frac{1}{2}\int_{\mathbb{R}} x^2\mu(dx) \tag{3.2}$$

and

$$I(\mu)=\int_{\mathbb{R}}(Hp(x)-x)^2\mu(dx)=\Phi(\mu)+\int_{\mathbb{R}}x^2\mu(dx)-\Phi(\sigma)-\int_{\mathbb{R}}x^2\sigma(dx). \tag{3.3}$$

It is shown in [5] that the inequality (3.1) can be recast in the equivalent form

$$\tilde{\Sigma}(\mu)\le\frac{1}{2}I(\mu)\qquad\text{for all }\ \mu$$

and in this form it has been used in [6]. We shall generalize the quantities (3.2) and (3.3) replacing the semi-circular distribution by measures coming from a convex potential. Let V be a C^1 function on $\mathbb{R}$, and define

$$\mathcal{E}_V(\nu)=\int\int_{\mathbb{R}}\log|x-y|\nu(dx)\nu(dy)-\int_{\mathbb{R}}V(x)\nu(dx).$$

Let ν_V be the unique probability measure on $\mathbb{R}$, which maximizes the functional $\mathcal{E}_V(\nu)$. As soon as V has a sufficient growth rate at infinity, this measure exists and it has compact support and a density q which satisfies

$$Hq(x)=V'(x)\quad\text{ for all } x \text{ in the support of }\quad \nu_V. \tag{3.4}$$

Furthermore there exists a constant C such that for all $x\in\mathbb{R}$ one has

$$\int_{\mathbb{R}}\log|x-y|\nu_V(dy)\le V(x)+C$$

with equality for x in the support of ν_V, see [7] for a discussion of these results.

Such measures appear in a random matrix theory through the following computation. Let H_N be the space of $N\times N$ Hermitian matrices with the Euclidean structure given by $\langle A,B\rangle=\mathrm{Tr}(AB)$, and the associated Lebesgue measure dM. Consider the probability measure

$$\frac{1}{Z_N(V)}e^{-N\mathrm{Tr}(V(M))}\,dM\qquad\left(\text{with } Z_N(V)=\int_{H_N}e^{-N\mathrm{Tr}(V(M))}\,dM\right)$$

on H_N. Then for any bounded continuous function G one has

$$\frac{1}{Z_N(V)}\int_{H_N}\frac{1}{N}\mathrm{Tr}(G(M))e^{-N\mathrm{Tr}(V(M))}\,dM\to_{N\to\infty}\int_{\mathbb{R}}G(x)\nu_V(dx). \tag{3.5}$$

Furthermore one has

$$\lim_{N\to\infty}\frac{1}{N^2}\log Z_N(V)=\mathcal{E}_V(\nu_V), \tag{3.6}$$

see [8, Chapter 6]. In particular, the semi-circular distribution corresponds to the function $V(x)=x^2/2$ and the matrix model result is known as Wigner's theorem. Let μ be a measure with finite free Fisher information, and define the following quantities:

$$\tilde{\Sigma}_V(\mu)=\mathcal{E}_V(\nu_V)-\mathcal{E}_V(\mu),\quad I_V(\mu)=\int(Hp(x)-V'(x))^2\mu(dx).$$

Since ν_V is a maximum, one has $\tilde{\Sigma}_V(\mu)\ge 0$ for all μ, with equality only for $\mu=\nu$.

Theorem 3.1 *Suppose that $V''\ge c>0$. Then for every measure μ with $I_V(\mu)<\infty$ one has*

$$\tilde{\Sigma}_V(\mu)\le\frac{2}{c}I_V(\mu). \tag{3.7}$$

Proof Let μ be a probability measure with compact support included in $[-A,+A]$ with a smooth density p, and let

$$L(x)=\int_{\mathbb{R}}\log|x-y|\,\mu(dy).$$

Let F be a C^1 function such that $F \geq L$, $F = L$ on $[-A, +A]$ and $F(x) = V(x)$ for $|x|$ large enough. Note that $L(x) = O(\log|x|)$ as $x \to \infty$, while V has at least quadratic growth, so that there exists such an F. Then the quantity $\mathcal{E}_F(\mu)$ attains its maximum at the unique point μ and one has, for all bounded continous functions G,

$$\frac{1}{Z_N(F)} \int_{H_N} \frac{1}{N} \mathrm{Tr}(G(M)) e^{-N\mathrm{Tr}(F(M))}\, dM \to_{N\to\infty} \int_{\mathbb{R}} G(x)\mu(dx), \tag{3.8}$$

see [8, Chapter 6]. The function $W(M) = \mathrm{Tr}(V(M))$ is convex on H_N and satisfies $W'' \geq c\,\mathrm{Id}$, we can therefore apply Theorem 1.1 to the probability measure

$$\frac{1}{Z_N(V)} \exp\{-N\mathrm{Tr}(V(M))\}\, dM$$

and to the function

$$f^2(M) = \exp\{-N\mathrm{Tr}(F(M) - V(M))\}.$$

Note that

$$\nabla f(M) = -\frac{N}{2}(F'(M) - V'(M)) \exp\left\{-\frac{N}{2}\mathrm{Tr}(F(M) - V(M))\right\}$$

and

$$\|\nabla f(M)\|^2 = \frac{N^2}{4}\mathrm{Tr}((F'(M) - V'(M))^2) \exp\{-N\mathrm{Tr}(F(M) - V(M))\}.$$

We get

$$\begin{aligned}\frac{1}{Z_N(V)} \int_{H_N} N\mathrm{Tr}(V(M) - F(M)) e^{-N\mathrm{Tr}(F(M))} dM - \frac{Z_N(F)}{Z_N(V)} \log\left(\frac{Z_N(F)}{Z_N(V)}\right) \\ \leq \frac{2}{cNZ_N(V)} \int_{H_N} \frac{N^2}{4} \mathrm{Tr}((F'(M) - V'(M))^2) e^{-N\mathrm{Tr}(F(M))} dM.\end{aligned}$$

We multiply both sides of the inequality by $\frac{Z_N(V)}{N^2 Z_N(F)}$ and apply (3.5), using the fact that $F - V$ has compact support, and (3.4), (3.6) for the function F, to get (3.7). This gives the result for measures with smooth densities. Measures μ and ν corresponding to more general V and F are obtained by approximation.

Let us note the following corollary of Theorem 3.1:

Corollary 3.2 *Let V be a convex potential. Then there is a unique measure μ with a density p satisfying $Hp(x) = V'(x)$ for all x in the support of μ.*

Indeed it follows from the hypothesis that $I_V(\mu) = 0$, so that by Theorem (3.1) one has $\mathcal{E}_V(\mu) = \mathcal{E}_V(\nu_V)$, and hence $\mu = \nu_V$.

Note that there exist non-convex potentials such that the equation $Hp_i(x) = V'(x)$ on the support of μ_i holds for distinct measures $\mu_i = p_i(x)dx$. The simplest example is a two-well potential V satisfying $V(x) = \frac{1}{2}(x - a_i)^2$ in $[a_i - 2, a_i + 2]$ for points a_1, a_2 with $|a_1 - a_2| > 4$, then the semi-circular measures centered on a_1 and a_2, respectively, satisfy $Hp_i(x) = V'(x)$ on their respective supports.

4 Variables with Prescribed Conjugate Variables

4.1 Conjugate variables

Denote by $\mathbb{C}_{\langle n\rangle} = \mathbb{C}\langle X_1, \ldots, X_n\rangle$ the free algebra on n generators, endowed with the antilinear involution for which $X_j^* = X_j$. Let $\partial_j : \mathbb{C}_{\langle n\rangle} \to \mathbb{C}_{\langle n\rangle} \otimes \mathbb{C}_{\langle n\rangle}$ be the derivation such that $\partial_j X_k = \delta_{jk} 1 \otimes 1$ (where the bi-modular structure on $\mathbb{C}_{\langle n\rangle} \otimes \mathbb{C}_{\langle n\rangle}$ is given by $a.(x \otimes y).b = ax \otimes by$). Let $(\mathcal{A}, \tau)$ be a tracial noncommutative probability space, and $X_1, \ldots, X_n \in \mathcal{A}$ be self-adjoint elements. The n-tuple $(X_1, \ldots, X_n)$ admits a conjugate variable if there exists $(\mathcal{J}_{X_1}, \ldots, \mathcal{J}_{X_n}) \in L^2(\mathcal{A}, \tau)$ such that for all $P \in \mathbb{C}_{\langle n\rangle}$ and $j \in \{1, \ldots, n\}$ one has

$$\tau(\mathcal{J}_{X_j} P) = \tau \otimes \tau(\partial_j P),$$

see [2]. We want to prove the existence of a large set of n-tuples with prescribed conjugate variables. We shall find it easier to deal with unitary operators instead of self-adjoint ones, so we shall define a suitable notion of conjugate variables for unitary elements. Let $\mathbb{C}(F_n)$ be the group algebra of the free group with n generators $g_1, \ldots, g_n$, endowed with the antilinear involution $g_i^* = g_i^{-1}$. If $(U_1, \ldots, U_n)$ are some unitary operators and $P \in \mathbb{C}(F_n)$, we call $P(U_1, \ldots, U_n)$ the result of the substitution of g_i by U_i. Introduce the left derivations $\partial_i : \mathbb{C}(F_n) \to \mathbb{C}(F_n) \otimes \mathbb{C}(F_n)$ satisfying

$$\partial_i(1) = 0; \quad \partial_i(g_j) = \delta_{ij}\, g_i \otimes 1$$

and

$$\partial_i g_j^{-1} = -\delta_{ij}\, 1 \otimes g_i^{-1}.$$

We also define the cyclic gradient $D_i = m \circ \sigma \circ \partial_i$, where σ is the flip automorphism of $\mathbb{C}(F_n) \otimes \mathbb{C}(F_n)$ and $m : \mathbb{C}(F_n) \otimes \mathbb{C}(F_n) \to \mathbb{C}(F_n)$ the multiplication. We say that a family of unitaries in some noncommutative probability space $(\mathcal{A}, \tau)$ admits $(\mathcal{J}_1, \ldots, \mathcal{J}_n) \in L^2(\mathcal{A}, \tau)$ as a conjugate variable if for all j and all $P \in \mathbb{C}(F_n)$ one has

$$\tau(\mathcal{J}_j P(U_1, \ldots, U_n)) = \tau \otimes \tau(\partial_j P(U_1, \ldots, U_n)).$$

For example, let $U_1, \ldots, U_n$ be n free Haar unitaries. Then they admit $(0, \ldots, 0)$ as conjugate variables. In fact it is easy to check that this is a characterization of n-tuples of free Haar unitaries. We shall prove that there exists small deformations of these free n-tuples of Haar unitaries, whose conjugate variables are prescribed cyclic gradients. Let us put the following semi-norm on $\mathbb{C}(F_n)$:

$$\Big\| \sum_{w \in F_n} a_w w \Big\|_\Delta = \frac{1}{2} \sum_{w \in F_n} l(w)^2 |a_w|.$$

Theorem 4.1 *Let $P \in \mathbb{C}(F_n)$ with $P = P^*$ and $\|P\|_\Delta < 1$. Then there exist a tracial noncommutative probability space (M, τ) and unitary elements $(U_1, \ldots, U_n) \in M$ with conjugate variables $(\mathcal{J}_1, \ldots, \mathcal{J}_n)$ such that, for all j, one has*

$$\mathcal{J}_j = D_j P(U_1, \ldots, U_n).$$

Proof For N a positive integer, consider the compact semi-simple Lie group $G = SU(N)$, with its Killing metric k and normalized Haar measure dU. Its Lie algebra $su(N)$ consists of $N \times N$ traceless antihermitian matrices, and the Killing form is given by $k(A, B) = -2N\mathrm{Tr}(AB)$. The group $SU(N)$ is a Riemannian manifold with constant curvature, the Ricci curvature being equal to the Killing form k, see e.g. [9]. Let μ_P denote the probability measure

$$d\mu_P(U_1,\dots,U_n) = \frac{1}{Z_N} e^{-N\mathrm{Tr}(P(U_1,\dots,U_n))} dU_1 \dots dU_n$$

on $SU(N)^n$. Denote by tr the normalized trace on $M_N(\mathbb{C})$. Let $R = \sum_i S_i \otimes T_i \in M_N(\mathbb{C}) \otimes M_N(\mathbb{C})$ and $A \in M_N(\mathbb{C})$. Then we define $R\#A = \sum_i S_i A T_i$.

Using the invariance by right multiplication of the Haar measure on $SU(N)^n$ we get, for any $Q \in \mathbb{C}(F_n)$ and matrices $A, B \in su(N)$,

$$\begin{aligned} 0 &= \frac{d}{dt}_{|t=0} \int_{SU(N)^n} \mathrm{tr}\left[Q^{j,t} B\right] e^{-N\mathrm{Tr}(P^{j,t})} dU_1 \dots dU_n \\ &= \int_{SU(N)^n} \big(\mathrm{tr}\left[(\partial_i Q(U_1,\dots,U_n)\#A)B\right] \\ &\quad - N^2 \mathrm{tr}\left[Q(U_1,\dots,U_n)B\right] \mathrm{tr}\left[D_j P(U_1,\dots,U_n)A\right]\big) d\mu_P(U_1,\dots,U_n), \end{aligned}$$

where $Q^{t,j}$ (resp. $P^{t,j}$) denotes the substitution of $U_j e^{tA}$ for U_j in $Q(U_1,\dots,U_n)$ (resp. $P(U_1,\dots,U_n)$). Taking the sum over pairs $A = B = A_j$, where A_j runs through an orthonormal basis of $su(N)$, we obtain the identity

$$\int_{SU(N)^n} \mathrm{tr}\otimes\mathrm{tr}\left[\partial_i Q\right] - \frac{1}{N^2}\mathrm{tr}\left[D_i Q\right]\, d\mu_P \tag{4.1}$$

$$= \int_{SU(N)^n} \mathrm{tr}\left[Q D_i P\right] - \frac{1}{N^2}\mathrm{tr}\left[Q\right]\mathrm{tr}\left[D_i P\right] d\mu_P. \tag{4.2}$$

On $SU(N)^n$ we consider the function

$$V(U_1,\dots,U_n) = N\mathrm{Tr}(Z(U_1,\dots,U_n)),$$

where $Z \in F_n$ is a word of length l. One has, for $H = (H_1,\dots,H_n) \in su(N)^n$,

$$\mathrm{Hess}(V)(U_1,\dots,U_n)(H,H) = \sum_{i,j=1}^{n} \sum_{k=1}^{p_{ij}} N\mathrm{Tr}(V_{ijk} H_i W_{ijk} H_j),$$

where V_{ijk}, W_{ijk} are unitary operators and $p_{ij} \le l^2$; therefore, by Hölder's inequality,

$$|\mathrm{Hess}(V)(U_1,\dots,U_n)(H,H)| \le -l^2\, N \sum_{i=1}^{n} \mathrm{Tr}(H_i^2) = \frac{l^2}{2} k^{(n)}(H,H),$$

$k^{(n)}$ being the Killing metric on $SU(N)^n$. Replacing Z by some $P \in \mathbb{C}(F_n)$ we get $|\mathrm{Hess}(V)| \le \|P\|_\Delta k^{(n)}$.

It follows that Theorems 2.2 and 2.3 apply on the manifold $SU(N)^n$, with $c = 1 - \|P\|_\Delta$, as soon as $c > 0$ (the Ricci tensor is again equal to $k^{(n)}$).

For any word $R \in \mathbb{C}(F_n)$ let $F(U_1,\dots,U_n) = \mathrm{tr}(R(U_1,\dots,U_n))$ then one has

$$F'(U_1,\dots,U_n).(H_1,\dots,H_n) = \sum_{i=1}^{n} \mathrm{tr}(D_i R(U_1,\dots,U_n) H_i),$$

so that

$$\nabla F(U_1,\dots,U_n) = -\frac{1}{2N^2}(D_1 R(U_1,\dots,U_n),\dots,D_n R(U_1,\dots,U_n));$$

therefore $\|\nabla F\|_\infty \le \frac{\sqrt{n}l}{2N^2}$. Applying Theorem 2.3 one has therefore

$$\mu_P(|\mathrm{tr}(R) - E_{\mu_P}(\mathrm{tr}(R))| \ge r) \le 2e^{-r^2N^2/nl^2}.$$

Denote $\tau_N = \mu_P \otimes \mathrm{tr}$ on $L^\infty(SU(N)^n) \otimes M_N(\mathbb{C})$. Using the above concentration property of the measures $d\mu_P$ we see that for any $P_1, P_2 \in \mathbb{C}(F_n)$ one has

$$\int_{SU(N)^n} \mathrm{tr} \otimes \mathrm{tr}(P_1 \otimes P_2) d\mu_P = \tau_N(P_1) \otimes \tau_N(P_2) + o(1) \tag{4.3}$$

as $N \to \infty$; therefore

$$\int_{SU(N)^n} \mathrm{tr} \otimes \mathrm{tr}(\partial_i Q) d\mu_P = \tau_N \otimes \tau_N(\partial_i Q) + o(1), \tag{4.4}$$

for any $Q \in \mathbb{C}(F_n)$. By the compactness of the space of states on $\mathbb{C}(F_n)$, we can extract a convergent subsequence of the states τ_N, which yields a trace state τ on $\mathbb{C}(F_n)$. From (4.1) and (4.4) we have, for all $Q \in \mathbb{C}(F_n)$,

$$\tau_N \otimes \tau_N(\partial_i Q) = \tau_N(Q D_i P) + o(1),$$

hence taking the convergent subsequence

$$\tau \otimes \tau(\partial_i Q) = \tau(Q D_i P),$$

i.e. $(U_1, \ldots, U_n)$ admit $(D_1 P, \ldots, D_n P)$ as conjugate variables in $(\mathbb{C}(F_n)'', \tau)$.

Finally we note that it is possible to use very similar arguments in order to construct perturbations of semi-circular systems having prescribed conjugate variables of the form $\mathcal{J}_{X_j} = X_j + D_j \Psi$, where Ψ is in the algebra generated by $f(X_j)$; $j = 1, \ldots, n$, $f \in C_c^\infty(\mathbb{R})$, and is sufficiently small so that $\mathrm{Tr}(\sum X_j^2 + \Psi(X_1, \ldots, X_n))$ is a uniformly convex function on $(H_N)^n$ and one can apply Theorem 2.1 and its concentration of measure consequences. We shall not get into the details.

5 Poincaré Inequality and Semi-circular Systems

In this section we shall compute the Hessian of the free entropy, at a point where the free Fisher information is finite, and show that it is the complete Dirichlet form associated with the free gradient. Then we prove that semi-circular systems are characterized by a Poincaré inequality with respect to this Dirichlet form.

We start with a finite-dimensional analogy. Let $\Phi(x) = \frac{1}{N^2} \sum_{i \neq j} \log |x_i - x_j|$ for $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$. Let us identify the tangent space at a point $x \in \mathbb{R}^N$ with the L^2 space of the measure $\mu_x = \frac{1}{N} \sum_{i=1}^N \delta_{x_i}$. Then the gradient of the above function is

$$\nabla \Phi(x) = (\mathcal{J}(x_1), \ldots, \mathcal{J}(x_N)),$$

where

$$\mathcal{J}(x_j) = \frac{1}{N} \sum_{i \neq j} \frac{1}{x_i - x_j}.$$

Furthermore the second derivative of Φ is

$$\Phi''(x)(f, g) = \frac{1}{N^2} \sum_{i \neq j} \frac{(f(x_i) - f(x_j))(g(x_i) - g(x_j))}{(x_i - x_j)^2}.$$

This is a positive quadratic form, which is positive definite on the space orthogonal to constants, hence the function Φ is convex outside the hyperplanes $x_i = x_j$, it is constant under translation by multiples of $(1, \ldots, 1)$, and its restrictions to the hyperplanes $\sum_i x_i = cste$ are strictly convex. We observe that for each x, $\Phi(x)(.,.)$ is a Dirichlet form on the L^2 space of μ_x.

We want to consider the "large N limit" of $\mathbb{R}^N$ where each element of $\mathbb{R}^N$ is identified with the measure μ_x. It is natural to think of this space as the space of probability measures on $\mathbb{R}$, with a manifold structure which is not the usual linear structure coming from the convex combinations of probability measures. Instead the manifold structure is the one induced from the action of the diffeomorphism group of $\mathbb{R}$ on $P(\mathbb{R})$. This makes this space into a highly singular "infinite-dimensional manifold". We shall not try to develop a theory of such manifold structures, rather we shall only remark that we can tell what is a tangent vector at a point μ: this is a vector field on $\mathbb{R}$, which is defined up to an addition of a vector field of μ-divergence 0 (indeed the diffeomorphism group generated by a μ-divergence 0 vector field leaves the measure μ invariant). So essentially this is a function on the support of μ and it is natural to put on such tangent vectors the Riemannian structure given by the $L^2(\mu)$ inner product.

Consider now the space of trace states on the algebra $\mathbb{C}\langle X_1, \dots, X_n\rangle$ of noncommutative polynomials in n self-adjoint indeterminates, and think of this as a multidimensional, noncommutative analogue of the manifold of probability measures on $\mathbb{R}$. Again we can define the action of an infinitesimal transformation of a trace state τ by a polynomial change of variables $X_j \mapsto X_j + \epsilon V_j(X_1, \dots, X_n)$ where V_j are self-adjoint polynomials. Thus the tangent space at a point τ is the space of n-tuples of self-adjoint polynomials $(V_1, \dots, V_n)$, modulo the τ-divergence free polynomials. The condition of leaving τ infinitesimally invariant (i.e. the τ-divergence 0 condition) is easily stated: If

$$\frac{\partial}{\partial\epsilon}_{|\epsilon=0} \tau(P(X_j + \epsilon V_j)) = 0,$$

for all polynomials P, then $(V_1, \dots, V_n)$ is orthogonal to the space of cyclic gradients i.e.

$$\sum_{j=1}^{n} \tau(V_j D_{X_j} P) = 0,$$

for all polynomials P. Here $D_{X_j} = m \circ \sigma \circ \partial_{X_j}$ is the i^{th} component of the cyclic gradient. Therefore we can identify the tangent space at τ with the closure of the space of all cyclic gradients in $L^2(\tau)^n$. See [10] for more on these matters.

Let us now investigate in this formalism what is the gradient of the free entropy. Using the change of variable formula (see Proposition 3.5 of [1], Section 1 of [11], Proposition 3.10 of [2]) for an n-tuple $(X_1, \dots, X_n)$ with free Fisher information, one has

$$\begin{aligned}\frac{\partial}{\partial\varepsilon}\chi(X_1 + \varepsilon V_1, \dots, X_n + \varepsilon V_n)_{|\varepsilon=0} &= \frac{\partial}{\partial\epsilon}\log|\mathcal{J}|(X_1 + \varepsilon V_1, \dots, X_n + \varepsilon V_n)_{|\varepsilon=0} \\ &= \tau\otimes\tau\Big(\sum_{i=1}^{n}\partial_{X_i} V_i\Big) = \tau\Big(\sum_{i=1}^{n}\mathcal{J}_{X_i} V_i\Big).\end{aligned}$$

Therefore the gradient of the free entropy is the projection onto the space of cyclic gradients of the vector $(\mathcal{J}_{X_1}, \dots, \mathcal{J}_{X_n})$, where $\mathcal{J}_{X_i} = \partial^*_{X_i}(1\otimes 1)$ are the conjugate variables. The second derivative can also be computed

$$\chi''(X_1, \dots, X_n)(V, W) = \sum_{i,j} \tau\otimes\tau(\partial_{X_j} V_j \partial_{X_j} W_j),$$

i.e. $\chi'' = \mathrm{Id} \otimes Q$ on $\mathbb{C}^n \otimes L^2(\tau)$, where Q is the quadratic form

$$Q(g) = \sum_{j=1}^{n} \|\partial_{X_j} g\|^2_{L^2(\tau\otimes\tau)}.$$

Note that Q, as the L^2-norm of a derivation, is a Dirichlet form (see [12], and [13] for the notion of a complete Dirichlet form), i.e. the associated Laplacian $\partial^*\partial$ generates a semi-group of completely positive maps $P_t = e^{-t\partial^*\partial}$ on $L^2(\tau)$.

Semi-circular systems are characterized by the maximization of the free entropy or minimization of the free Fisher information with given variance, see [11], [14]. This is an analogue of the well-known properties of Gaussian families in classical probability theory. I shall give here a similar characterization, using the free Poincaré inequality, which also has a classical analogue, see [15].

Let $(\mathcal{A}, \tau)$ be a noncommutative probability space, and let $X_1, \dots, X_n \in \mathcal{A}$ be centered elements, with covariance I_n. Then one can define on $\mathbb{C}_{\langle n \rangle}$ the Hermitian quadratic form

$$Q_X(P) = \sum_{i=1}^{n} \|\partial_{X_i} P(X_1, \dots, X_n)\|^2_{L^2(\tau\otimes\tau)}. \tag{5.1}$$

If $(X_1, \dots, X_n)$ have finite free Fisher information, then this yields a densely defined closable complete Dirichlet form on $L^2(\mathcal{A}, \tau)$.

In the case of semi-circular systems, this quadratic form is given by the number operator on the free Fock space, and it generates the free Ornstein–Uhlenbeck semi-group, see [16]. In particular, its eigenvalues are $0, 1, 2, \dots$, the eigenvalue 0 being simple with associated eigenvector 1. We deduce from this the following Poincaré inequality for a semi-circular system:

$$\mathrm{Var}(P(X_1, \dots, X_n)) \le Q_X(P) \qquad \text{for all polynomials} \quad P \in \mathbb{C}_{\langle n \rangle}. \tag{5.2}$$

Conversely we prove:

Theorem 5.1 *Let $X_1, \dots, X_n$ be centered elements, with covariance I_n, in some noncommutative probability space $(\mathcal{A}, \tau)$. Let Q_X be the associated quadratic form defined by* (5.1)*, and assume that* (5.2) *is satisfied. Then X is a semi-circular system.*

Proof Let $(P_n; n \ge 0)$ be the Tchebychev polynomials of the second kind, orthogonal with respect to the semi-circle distribution. It is enough to check that

$$\tau(P_{i_1}(X_{l_1}) \dots P_{i_k}(X_{l_k})) = 0,$$

for all values of $i_1, \dots, i_k \ge 1$ and $l_1 \ne l_2 \ne \dots \ne l_k$. The assumptions on the covariance of $X_1, \dots, X_n$ imply that this equality holds for $k = 1$ and $i_1 = 1$ or 2. Assume by induction that it is true for terms of total degree less than m, and let $P_{i_1}(X_{l_1}) \dots P_{i_k}(X_{l_k})$ be a term of degree $m + 1$. Since

$$X_{l_1} P_{i_1 - 1}(X_{l_1}) = P_{i_1}(X_{l_1}) + P_{i_1 - 2}(X_{1_1}),$$

it is enough to check that

$$\tau(X_{l_1} P_{i_1 - 1}(X_{l_1}) P_{i_2}(X_{l_2}) \dots P_{i_k}(X_{l_k})) = 0,$$

where we can assume that either $k \ge 2$ or $k = 1$ and $i_1 \ge 3$. Let

$$W = P_{i_1 - 1}(X_{l_1}) \dots P_{i_k}(X_{l_k}).$$

By the Poincaré inequality (5.2) one has, for all $a \in \mathbb{C}$,

$$\begin{aligned} \operatorname{Var}(aX_{l_1} + W) &= a^2 \operatorname{Var}(X_{l_1}) + 2\Re(a \operatorname{Cov}(X_{l_1}, W)) + \operatorname{Var}(W) \\ &\leq Q_X(aX_{l_1} + W) \\ &\leq a^2 Q_X(X_{l_1}) + 2\Re(a Q_X(X_{l_1}, W)) + Q(W). \end{aligned}$$

Since $Q_X(X_{l_1}) = 1 = \operatorname{Var}(X_{l_1})$ this inequality can hold for all a only if

$$\operatorname{Cov}(X_{l_1}, W) = \tau(X_{l_1} W) = Q(X_{l_1}, W).$$

Since $\partial_{X_i} X_j = \delta_{ij} 1 \otimes 1$ one has $Q(X_{l_1}, W) = \tau \otimes \tau(\partial_{X_{l_1}} W)$. A direct computation shows that

$$\partial_X P_n(X) = \sum_{k=0}^{n-1} P_k \otimes P_{n-k-1}. \tag{5.3}$$

Using the induction hypothesis we get $\tau \otimes \tau(\partial_{X_{l_1}} W) = 0$; therefore $\tau(X_{l_1} W) = 0$ and the proof follows by induction.

Note that the above proof can be adapted to the classical case using the identity $H_n' = nH_{n-1}$ for Hermite polynomials, instead of (5.3); this is somewhat simpler than the proof in [15].

References

[1] Voiculescu, D. V.: The analogues of entropy and of Fisher's information measure in free probability theory. II. *Invent. Math.*, **118**(3), 411–440 (1994)

[2] Voiculescu, D. V.: The analogues of entropy and of Fisher's information measure in free probability theory. V. Noncommutative Hilbert transforms. *Invent. Math.*, **132**(1), 189–227 (1998)

[3] Ledoux, M.: Concentration of measure and logarithmic Sobolev inequalities Séminaire de Probabilités XXXIII, Lecture Notes in Mathematics, 1709, Springer, 120–216 (1999)

[4] Bakry, D., Émery, M.: Diffusions hypercontractives. *Sminaire de probabilits, XIX*, 1983/84, 177–206, Lecture Notes in Math., **1123**, Springer, Berlin, 1985

[5] Biane, P., Speicher, R.: Free diffusions, free entropy and free fisher information. *Ann. Inst. H. Poincaré, Probabilités et Statistiques, PR*, **37**, 581–606 (2001)

[6] Biane, P., Voiculescu, D.: A free probability analogue of the Wasserstein distance on the trace-state space. to appear in *GAFA*

[7] Saff, E. B., Totik, V.: Logarithmic potentials with external fields. Appendix B by T. Bloom. Grundlehren der Mathematischen Wissenschaften, 316. Springer-Verlag, Berlin (1997)

[8] Deift, P. A.: Orthogonal polynomials and random matrices: a Riemann-Hilbert approach. Courant Lecture Notes in Mathematics, 3. New York University, Courant Institute of Mathematical Sciences, New York (1999)

[9] Abraham, R., Marsden, J. E.: Foundations of mechanics. Second edition. Benjamin/Cummings Publishing Co., Inc., Advanced Book Program, Reading, Mass. (1978)

[10] Voiculescu, D. V.: Cyclomorphy. MSRI preprint # 2001–019.

[11] Voiculescu, D. V.: The analogues of entropy and of Fisher's information measure in free probability theory IV. Maximum entropy and freeness. Free probability theory (Waterloo, ON, 1995), 293–302, Fields Inst. Commun., 12, Amer. Math. Soc., Providence, RI (1997)

[12] Sauvageot, J. L.: Quantum Dirichlet forms, differential calculus and semigroups. Quantum probability and applications, V (Heidelberg, 1988), 334–346, Lecture Notes in Math., 1442, Springer, Berlin (1990)

[13] Davies, E. B., Lindsay, J. M.: Noncommutative symmetric Markov semigroups. *Math. Z.*, **210**(3), 379–411 (1992)

[14] Nica, A., Shlyakhtenko, D., Speicher, R.: Some minimization problems for the free analogue of the Fisher information. *Adv. Math.*, **141**(2), 282–321 (1999)

[15] Borovkov, A. A., Utev, S. A.: An inequality and a characterization of the normal distribution connected with it. *Teor. Veroyatnost. i Primenen.*, **28**(2), 209–218 (1983)

[16] Biane, P.: Free hypercontractivity. *Communications in Mathematical Physics*, **184,** 457–474 (1997)