journal of inequalities in pure and applied mathematics

http://jipam.vu.edu.au

issn: 1443-5756

Volume 8 (2007), Issue 1, Article 1, 6 pp.

FREE AND CLASSICAL ENTROPY OVER THE CIRCLE

GORDON BLOWER

DEPARTMENT OF MATHEMATICS AND STATISTICS
LANCASTER UNIVERSITY
LANCASTER LA1 4FY
ENGLAND UK.
g.blower@lancaster.ac.uk

Received 12 September, 2006; accepted 17 February, 2007
Communicated by F. Hansen

ABSTRACT. Relative entropy with respect to normalized arclength on the circle is greater than or equal to the negative logarithmic energy (Voiculescu's negative free entropy) and is greater than or equal to the modified relative free entropy. This note contains proofs of these inequalities and related consequences of the first Lebedev–Milin inequality.

Key words and phrases: Transportation inequality; Free probability; Random matrices.

2000 *Mathematics Subject Classification.* 60E15; 46L54.

1. INTRODUCTION AND DEFINITIONS

In this note we consider inequalities between various notions of relative entropy and related metrics for probability measures on the circle. The introduction contains definitions and brief statements of results which are made precise in subsequent sections.

Definition 1.1. For μ and ν probability measures on $\mathbb{T}$ with ν absolutely continuous with respect to μ, let $d\nu/d\mu$ be the Radon–Nikodym derivative. The (classical) *relative entropy* of ν with respect to μ is

$$\mathrm{Ent}(\nu \mid \mu) = \int_{\mathbb{T}} \log \frac{d\nu}{d\mu}\, d\nu; \tag{1.1}$$

note that $0 \leq \mathrm{Ent}(\nu \mid \mu) \leq \infty$ by Jensen's inequality; we take $\mathrm{Ent}(\nu \mid \mu) = \infty$ when ν is not absolutely continuous with respect to μ.

Definition 1.2. Let ρ be a probability measure on $\mathbb{R}$ that has no atoms. If the integral

$$\Sigma(\rho) = \iint_{\mathbb{R}^2} \log |x - y|\, \rho(dx)\rho(dy) \tag{1.2}$$

converges absolutely, then ρ has *free entropy* $\Sigma(\rho)$, that is, the logarithmic energy.

I am grateful to Profs. Hiai and Ueda for helpful communications and to Prof Ledoux for pointing out some references. The research was partially supported by the project 'Phenomena in High Dimensions MRTN-CT-2004-511953'.

249-06

Voiculescu [14] introduced this along with other concepts of free probability; see also [3], [5], [6], [8], where various notations and constants are employed.

In Theorem 2.1 we compare free with relative entropy with respect to arclength measure $d\theta/2\pi$ on $\mathbb{T}$ and show that $\rho(d\theta) = p(e^{i\theta})d\theta/2\pi$ satisfies

$$-\Sigma(\rho) \leq \text{Ent}(\rho \mid d\theta/2\pi). \tag{1.3}$$

The proof involves the sharp Hardy–Littlewood–Sobolev inequality.

Definition 1.3. Suppose that f and g are probability density functions with respect to $d\theta/2\pi$, and let

$$\Sigma(f,g) = \iint_{\mathbb{T}^2} \log \frac{1}{|e^{i\theta} - e^{i\phi}|} \big(f(e^{i\theta}) - g(e^{i\theta})\big)\big(f(e^{i\phi}) - g(e^{i\phi})\big) \frac{d\theta}{2\pi}\frac{d\phi}{2\pi} \tag{1.4}$$

be the *modified relative free entropy* as in [5], [6], [7], [8].

For notational convenience, we identify an absolutely continuous probability measure with its probability density function and write $\mathbb{I}$ for the constant function 1. In Theorem 2.2 we show that $\Sigma(f,\mathbb{I}) \leq \text{Ent}(f \mid \mathbb{I})$. The proof uses the first Lebedev–Milin inequality for functions in the Dirichlet space over unit disc $\mathbb{D}$. Let $u : \mathbb{D} \to \mathbb{R}$ be a harmonic function such that $\|\nabla u(z)\|^2$ is integrable with respect to area measure, let v be its harmonic conjugate with $v(0) = 0$ and $g = (u+iv)/2$. Then by [10], u satisfies

$$\log \int_{\mathbb{T}} \exp\big(u(e^{i\theta})\big)\frac{d\theta}{2\pi} \leq \frac{1}{4\pi}\iint_{\mathbb{D}} \|\nabla u(re^{i\theta})\|^2 r dr d\theta + \int_{\mathbb{T}} u(e^{i\theta})\frac{d\theta}{2\pi}; \tag{1.5}$$

thus $\exp g$ belongs to the Hardy space $\mathrm{H}^2(\mathbb{D})$. One can interpret this inequality as showing that $\mathrm{H}^2(\mathbb{D})$ is the symmetric Fock space of Dirichlet space, which is reflected by the reproducing kernels, as in [12].

Definition 1.4. Let μ and ν be probability measures on $\mathbb{T}$. Then the *Wasserstein p metric* for $1 \leq p < \infty$ and the cost function $|e^{i\theta} - e^{i\phi}|^p/p$ is

$$W_p(\mu,\nu) = \inf_{\omega}\left\{\left(\frac{1}{p}\iint_{\mathbb{T}^2} |e^{i\theta} - e^{i\phi}|^p \omega(d\theta d\phi)\right)^{\frac{1}{p}}\right\}, \tag{1.6}$$

where ω is a probability measure on $\mathbb{T}^2$ that has marginals μ and ν. See [13].

Let $u : \mathbb{T} \to \mathbb{R}$ be a 1-Lipschitz function in the sense that $|u(e^{i\theta}) - u(e^{i\phi})| \leq |e^{i\theta} - e^{i\phi}|$ for all $e^{i\theta}, e^{i\phi} \in \mathbb{T}$, and suppose further that $\int_{\mathbb{T}} u(e^{i\theta})d\theta/2\pi = 0$. Then by (1.6), as reformulated in (3.2) below, we have

$$\int_{\mathbb{T}} \exp\big(tu(e^{i\theta})\big)\frac{d\theta}{2\pi} \leq \exp\left(\frac{t^2}{2}\right) \qquad (t \in \mathbb{R}). \tag{1.7}$$

Bobkov and Götze have shown that the dual form of this concentration inequality is the transportation inequality $W_1(\rho, d\theta/2\pi)^2 \leq 2\text{Ent}(\rho \mid d\theta/2\pi)$ for all probability measures ρ of finite relative entropy with respect to $d\theta/2\pi$, as in [13], 9.3. In Section 3 we provide a free transportation inequality $W_1(\rho,\nu)^2 \leq 2\Sigma(\rho,\nu)$ which generalizes and strengthens this dual inequality.

2. FREE VERSUS CLASSICAL ENTROPY WITH RESPECT TO ARCLENGTH

For completeness, we recall the following result of Beckner and Lieb [2].

Theorem 2.1. *Suppose that f is a probability density function on $\mathbb{R}$ such that $f\log f$ is integrable. Then f has finite free entropy and*

$$\iint_{\mathbb{R}^2} \log\frac{1}{|x-y|} f(x)f(y)\,dxdy \le \log 2\pi + \int_{\mathbb{R}} f(x)\log f(x)\,dx. \tag{2.1}$$

Proof. The sharp form of the Hardy–Littlewood–Sobolev inequality, due to Lieb [2], gives

$$\iint_{\mathbb{R}^2} \frac{f(x)f(y)}{|x-y|^{\lambda}}\,dxdy \le \pi^{3/2-2/p}\frac{\Gamma(1/p-1/2)}{\Gamma(1/p)}\left(\int_{\mathbb{R}} |f(x)|^p\,dx\right)^{\frac{2}{p}}, \tag{2.2}$$

for $\lambda = 2(1-1/p)$ with $1 \le p < 2$, and with equality when $p = 1$. Hence the derivative at $p = 1+$ of the left-hand side is less than or equal to the derivative of the right-hand side. By differentiating Legendre's duplication formula $\Gamma(2x)\Gamma(1/2) = 2^{2x-1}\Gamma(x)\Gamma(x+1/2)$ at $x = 1/2$, we obtain

$$\Gamma'(1)/\Gamma(1) = 2\log 2 + \Gamma'\left(\frac{1}{2}\right)\Big/\Gamma\left(\frac{1}{2}\right), \tag{2.3}$$

and hence we obtain the derivative of the numerical factor in (2.2).

This gives (2.1); to deduce (1.3), we take $f(\theta) = p(e^{i\theta})\mathbb{I}_{[0,2\pi]}(\theta)/2\pi$ where $\rho(d\theta) = p(e^{i\theta})d\theta/2\pi$. □

In [7] the authors assert that the relative and free entropies with respect to arclength are incomparable, contrary to Theorem 2.2 below and (1.3). Whereas the values of the entropies of their attempted counterexample are correct on [7, p. 220] and [5, p. 204], the limit on [7, p. 220, line 7] should be 1 and not 0; so the calculation fails. The calculation on [7, p. 219] does show that (1.3) has no reverse inequality.

Definition 2.1. With real α and Fourier coefficients $\hat{f}(n) = \int_{\mathbb{T}} f(e^{i\theta})e^{-in\theta}d\theta/2\pi$, let $H^{\alpha}(\mathbb{T})$ be the subspace of $L^2(\mathbb{T})$ consisting of those f such that

$$\|f\|_{H^{\alpha}(\mathbb{T})} = \left(\sum_{n\in\mathbb{Z}}(1+|n|^{2\alpha})|\hat{f}(n)|^2\right)^{\frac{1}{2}} \tag{2.4}$$

is finite, and let $\dot{H}^{\alpha}(\mathbb{T})$ be the completion of the subspace $\{f \in H^{\alpha}(\mathbb{T}) : \hat{f}(0) = 0\}$ for the norm

$$\|f\|_{\dot{H}^{\alpha}(\mathbb{T})} = \left(\sum_{n\in\mathbb{Z}\setminus\{0\}} |n|^{2\alpha}|\hat{f}(n)|^2\right)^{\frac{1}{2}}; \tag{2.5}$$

we use the notation $\|f\|_{\dot{H}^{\alpha}(\mathbb{T})}$ to indicate the semi-norm defined by this sum for typical elements of $H^{\alpha}(\mathbb{T})$.

There is a natural pairing of $\dot{H}^{\alpha}(\mathbb{T})$ with $\dot{H}^{-\alpha}(\mathbb{T})$ whereby $g(e^{i\theta}) \sim \sum_{n\in\mathbb{Z}\setminus\{0\}} b_n e^{in\theta}$ in $\dot{H}^{-\alpha}(\mathbb{T})$ defines a bounded linear functional on $\dot{H}^{\alpha}(\mathbb{T})$ by

$$\sum_{n\in\mathbb{Z}\setminus\{0\}} a_n e^{in\theta} \mapsto \sum_{n\in\mathbb{Z}\setminus\{0\}} a_n\bar{b}_n. \tag{2.6}$$

When p and q are probability density functions of finite relative free entropy, their difference $f = p - q$ belongs to $\dot{H}^{-1/2}(\mathbb{T})$ and is real; so when we take the Taylor expansion of the kernel in (1.4) we deduce that

$$\|p-q\|^2_{\dot H^{-1/2}(\mathbb{T})} = \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{\hat f(n)\hat f(-n)}{|n|} = 2\sum_{n=1}^\infty \frac{|\hat f(n)|^2}{n} = 2\Sigma(p,q), \tag{2.7}$$

as in [8, p. 716].

Theorem 2.2. *Let f be a probability density function on $\mathbb{T}$ that has finite relative entropy with respect to $d\theta/2\pi$. Then*

$$\Sigma(f,\mathbb{I}) \le \mathrm{Ent}(f \mid \mathbb{I}). \tag{2.8}$$

Proof. We consider harmonic extensions of $L^2(\mathbb{T})$ to the unit disc. Let $u_\phi(e^{i\theta}) = u(e^{i\theta - i\phi})$ and let $u(re^{i\theta}) = \int_{\mathbb{T}} P_r(e^{i\phi})u_\phi(e^{i\theta})d\phi/2\pi$ be the Poisson extension of u, where $P_r(e^{i\theta}) = \sum_{n\in\mathbb{Z}} r^{|n|}e^{in\theta}$. The dual space of $\dot H^{-1/2}(\mathbb{T})$ under the pairing of (2.6) is $\dot H^{1/2}(\mathbb{T})$, which we identify with the Dirichlet space G of harmonic functions $u:\mathbb{D}\to\mathbb{R}$ such that $\int_{\mathbb{T}} u(e^{i\theta})d\theta/2\pi = 0$ and

$$\iint_{\mathbb{D}} \|\nabla u\|^2 dxdy/\pi < \infty. \tag{2.9}$$

By the joint convexity of relative entropy [4], any pair of probability density functions of finite relative entropy satisfies

$$\mathrm{Ent}(f \mid u) = \int_{\mathbb{T}} P_r(e^{i\phi})\mathrm{Ent}(f_\phi \mid u_\phi)\frac{d\phi}{2\pi} \ge \mathrm{Ent}(P_rf \mid P_ru); \tag{2.10}$$

so, in particular,

$$\mathrm{Ent}(f \mid \mathbb{I}) \ge \mathrm{Ent}(P_rf \mid \mathbb{I}) \quad (0 \le r < 1). \tag{2.11}$$

Hence it suffices to prove the theorem for P_rf instead of f, and then take limits as $r \to 1-$. For notational simplicity, we shall assume that f has a rapidly convergent Fourier series so that various integrals converge absolutely.

Suppose that u is a real function in $H^{1/2}(\mathbb{T})$ that has $\int_{\mathbb{T}} u(e^{i\theta})d\theta/2\pi = -t$ and $\|u\|_{\dot H^{1/2}(\mathbb{T})} = s$; by adding a constant to u if necessary, we can assume that $s^2/2 = t$. Then by (1.5) we have

$$\int_{\mathbb{T}} \exp u(e^{i\theta})\frac{d\theta}{2\pi} \le \exp\left(\frac{s^2}{2} - t\right) = 1, \tag{2.12}$$

and consequently by the dual formula for relative entropy

$$\begin{aligned} \int_{\mathbb{T}} f(e^{i\theta})\log f(e^{i\theta})\,\frac{d\theta}{2\pi} &= \sup\left\{\int_{\mathbb{T}} h(e^{i\theta})f(e^{i\theta})\frac{d\theta}{2\pi} : \int_{\mathbb{T}} \exp h(e^{i\theta})\frac{d\theta}{2\pi} \le 1\right\} \\ &\ge \int_{\mathbb{T}} f(e^{i\theta})u(e^{i\theta})\frac{d\theta}{2\pi}. \end{aligned} \tag{2.13}$$

Recalling the dual pairing of $\dot H^{-1/2}(\mathbb{T})$ with $\dot H^{1/2}(\mathbb{T})$, we write

$$\langle f, u\rangle = \int_{\mathbb{T}} f(e^{i\theta})u(e^{i\theta})\frac{d\theta}{2\pi} - \int_{\mathbb{T}} f(e^{i\theta})\frac{d\theta}{2\pi}\int_{\mathbb{T}} u(e^{i\theta})\frac{d\theta}{2\pi}, \tag{2.14}$$

so that by (2.13)

$$\langle f, u\rangle \le t + \int_{\mathbb{T}} f(e^{i\theta})\log f(e^{i\theta})\frac{d\theta}{2\pi}. \tag{2.15}$$

We choose the $\hat u(n)$ for $n \ne 0$ to optimize the left-hand side, and deduce that

$$\|f\|_{\dot H^{-1/2}(\mathbb{T})}\|u\|_{\dot H^{1/2}(\mathbb{T})} = s\|f\|_{\dot H^{-1/2}(\mathbb{T})} \le s^2/2 + \int_{\mathbb{T}} f(e^{i\theta})\log f(e^{i\theta})\,\frac{d\theta}{2\pi}, \tag{2.16}$$

so by choosing s we can obtain the desired result

$$2\Sigma(f,\mathbb{I}) = \|f\|^2_{\dot H^{-1/2}(\mathbb{T})} \le 2\int_{\mathbb{T}} f(e^{i\theta})\log f(e^{i\theta})\,\frac{d\theta}{2\pi}. \tag{2.17}$$

□

The quantity $\mathrm{Ent}(\mathbb{I}\mid w)$ also appears in free probability, and the appearance of the formula (1.5) likewise becomes unsurprising when we recall the strong Szegö limit theorem. Let $w:\mathbb{T}\to\mathbb{R}_+$ be a probability density with respect to $d\theta/2\pi$ such that $u(e^{i\theta}) = \log w(e^{i\theta})$ belongs to $H^{1/2}(\mathbb{T})$, let $D_n = \det[\hat w(j-k)]_{0\le j,k\le n-1}$ be the determinants of the $n\times n$ Toeplitz matrices associated with the symbol w, and let

$$\alpha_n = \exp\left((n+1)\int_{\mathbb{T}} u(e^{i\theta})\frac{d\theta}{2\pi} + \frac{1}{4\pi}\iint_{\mathbb{D}} \|\nabla u(z)\|^2\,dxdy\right) \qquad (n=0,1,\dots). \tag{2.18}$$

Then by (1.5), we have $\alpha_0\ge 1$ since $\int w(e^{i\theta})d\theta/2\pi = 1$; further

$$D_n^{1/n} \to \exp\left(\int_{\mathbb{T}} u(e^{i\theta})\frac{d\theta}{2\pi}\right) = \exp\Big(-\mathrm{Ent}(\mathbb{I}\mid w)\Big) \qquad (n\to\infty) \tag{2.19}$$

by [11, p. 169] and by Ibragimov's Theorem [11, p. 342],

$$D_n/\alpha_n \to 1 \qquad (n\to\infty). \tag{2.20}$$

One can refine the proof given in [1] and prove the following result on the asymptotic distribution of linear statistics. Let f be a real function in $H^{1/2}(\mathbb{T})$ and let $X_n : (U(n),\mu_{U(n)})\to\mathbb{R}$ be the random variable

$$X_n(\gamma) = \mathrm{trace}(f(\gamma)) - n\int_{\mathbb{T}} f(e^{i\theta})\frac{d\theta}{2\pi} \qquad (\gamma\in U(n)), \tag{2.21}$$

where $\mu_{U(n)}$ is the Haar measure on the group $U(n)$ of $n\times n$ unitary matrices. Then (X_n) converges in distribution as $n\to\infty$ to a Gaussian random variable with mean zero and variance $\|f\|^2_{\dot H^{1/2}(\mathbb{T})}$.

3. A SIMPLE FREE TRANSPORTATION INEQUALITY

Theorem 3.1. *Suppose that p and q are probability density functions with respect to $d\theta/2\pi$ such that their relative free entropy is finite. Then*

$$W_1(p,q)^2 \le 2\Sigma(p,q). \tag{3.1}$$

Proof. By the Kantorovich–Rubinstein theorem, as in [13, p. 34],

$$W_1(p,q) = \sup_u\left\{\int_{\mathbb{T}} u(e^{i\theta})\left(p(e^{i\theta}) - q(e^{i\theta})\right)\frac{d\theta}{2\pi} : |u(e^{i\theta}) - u(e^{i\phi})| \le |e^{i\theta} - e^{i\phi}|\right\}. \tag{3.2}$$

Any such 1–Lipschitz function u belongs to $H^{1/2}(\mathbb{T})$, since we have

$$\sum_{n\in\mathbb{Z}} |n||\hat u(n)|^2 = \iint_{\mathbb{T}^2}\left|\frac{u(e^{i\theta}) - u(e^{i\phi})}{e^{i\theta} - e^{i\phi}}\right|^2 \frac{d\theta}{2\pi}\frac{d\phi}{2\pi} \le 1, \tag{3.3}$$

by [11, 6.1.58]. Hence by the duality between $\dot H^{1/2}(\mathbb{T})$ and $\dot H^{-1/2}(\mathbb{T})$, we have

$$\begin{aligned} W_1(p,q) &\le \sup_u\left\{\int_{\mathbb{T}} u(e^{i\theta})\left(p(e^{i\theta}) - q(e^{i\theta})\right)\frac{d\theta}{2\pi} : \|u\|_{\dot H^{1/2}(\mathbb{T})} \le 1\right\} \\ &= \|p-q\|_{\dot H^{-1/2}(\mathbb{T})}. \end{aligned} \tag{3.4}$$

□

In [6] and [7], Hiai, Petz and Ueda prove a transportation inequality for W_2 by means of a difficult matrix approximation argument. Whereas transportation inequalities involving W_2 generally imply transportation inequalities for W_1 by the Cauchy–Schwarz inequality, Theorem 3.1 has the merit that it applies to a wide class of p and q and involves the uniform constant 2. Villani [13, p. 234] compares the W_2 metric with the H^{-1} norm, and Ledoux [9] obtains a free logarithmic Sobolev inequality using a proof based upon the Prékopa–Leindler inequality.

REFERENCES

[1] E.L. BASOR, Toeplitz determinants, Fisher–Hartwig symbols and random matrices, pp. 309–336 in *Recent Perspectives in Random Matrix Theory and Number Theory*, Eds. F. Mezzadri and N.C. Snaith, Cambridge University Press, 2005.

[2] W. BECKNER, Sharp Sobolev inequalities on the sphere and the Moser–Trudinger inequality, *Ann. of Math. (2)*, **138** (1993), 213–242.

[3] P. BIANE AND D. VOICULESCU, A free probability analogue of the Wasserstein metric on the trace-state space, *Geom. Funct. Anal.*, **11** (2001), 1125–1138.

[4] E.A. CARLEN AND M.C. CARVALHO, Strict entropy production bounds and stability of the rate of convergence to equilibrium for the Boltzmann equation, *J. Statist. Phys.*, **67** (1992), 575–608.

[5] F. HIAI AND D. PETZ, *The Semicircle Law, Free Random Variables and Entropy*, American Mathematical Society, Rhode Island, 2000.

[6] F. HIAI AND D. PETZ, A free analogue of the transportation cost inequality on the circle, in *Quantum Probability*, Edrs. M. Bozejko, W. Mlotkowsky and J. Wysoczansky, Banach Center Publications, Vol. 73, Warsaw, 2006, 199–206.

[7] F. HIAI, D. PETZ AND Y. UEDA, Free transportation cost inequalities via random matrix approximation, *Probab. Theory Relat. Fields*, **130** (2004), 199–221.

[8] F. HIAI, M. MIZUNO AND D. PETZ, Free relative entropy for measures and a corresponding perturbation theory, *J. Math. Soc. Japan*, **54** (2002), 670–718.

[9] M. LEDOUX, A (one-dimensional) free Brunn–Minkowski inequality, *C. R. Math. Acad. Sci. Paris*, **340** (2005), 301–304.

[10] B. OSGOOD, R. PHILLIPS AND P. SARNAK, Extremals of determinants of Laplacians, *J. Funct. Anal.*, **80**, (1988) 148–211.

[11] B. SIMON, *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*, American Mathematical Society, 2005.

[12] V.I. VASYUNIN AND N.K. NIKOLSKII, Operator-valued measures and coefficients of univalent functions, *St Petersburg Math. J.*, **3** (1992), 1199–1270.

[13] C. VILLANI, *Topics in Optimal Transportation*, American Mathematical Society, 2003.

[14] D. VOICULESCU, The analogues of entropy and of Fisher's information measure in free probability I, *Comm. Math. Phys.*, **115** (1993), 71–92.