数理解析研究所講究録
1177 巻 2000 年 52-65

On Free Relative Entropy

Fumio Hiai (日合 文雄) and Masaru Mizuo (水尾 勝)

Graduate School of Information Sciences
Tohoku University
Aoba-ku, Sendai 980-8577, Japan

Abstract

First, Voiculescu's single variable free entropy is generalized in two different ways to the free relative entropy for compactly supported probability measures on the real line; the one is introduced by the integral expression and the other is based on matricial (or microstates) approximation. Their equivalence is shown based on a large deviation result for the empirical eigenvalue distribution of a relevant random matrix. Secondly, the perturbation theory for compactly supported probability measures via free relative entropy is developed on the analogy of the perturabation theory via relative entropy. When the perturbed meausre via relative entropy is suitably arranged on the space of selfadjoint matrices and the matrix size goes to infinity, it is proven that the perturabtion via relative entropy on the matrix space approaches asymptotically to that via free relative entropy.

1 Free relative entropy

1.1 Definition of free relative entropy

For a probability Borel measure μ on $\mathbb{R}$ the *free entropy* $\Sigma(\mu)$ was introduced by Voiculescu [13] as

$$\Sigma(\mu) := \iint \log|x-y|\,d\mu(x)\,d\mu(y)\,, \tag{1.1}$$

and it is indeed the minus sign of the so-called *logarithmic energy* of μ familiar in potential theory [12]. Note that the double integral (1.1) always exists with a value in $[-\infty,+\infty)$ whenever μ is compactly supported. The free entropy functional $\Sigma(\mu)$ is upper semi-continuous in weak topology when the support of μ is restricted in a fixed compact set, and it is strictly concave in the sense that $\Sigma(\lambda\mu_1+(1-\lambda)\mu_2) > \lambda\Sigma(\mu_1)+(1-\lambda)\Sigma(\mu_2)$ if $0<\lambda<1$ and μ_1,μ_2 are compactly supported probability measures such that $\mu_1\neq\mu_2$, $\Sigma(\mu_1)>-\infty$ and $\Sigma(\mu_2)>-\infty$ (see [7, 5.3.2]).

The matricial approach (or the microstates approach) for free entropy was developed in [14]. For each $n \in \mathbb{N}$ let M_n denote the space of all $n \times n$ complex matrices and tr_n the normalized trace functional on M_n. The set of all selfadjoint matrices in M_n is denoted by M_n^{sa}. There is a natural linear bijection between M_n^{sa} and $\mathbb{R}^{n^2}$ which is an isometry for the Hilbert-Schmidt and Euclidean norms, so the "Lebesgue" measure Λ_n on M_n^{sa} is induced by the Lebesgue measure on $\mathbb{R}^{n^2}$ via this isometry. Let μ be a probability Borel measure supported in $[-R, R]$, $R > 0$. For $n, r \in \mathbb{N}$ and $\varepsilon > 0$ define

$$\Gamma_R(\mu; n, r, \varepsilon) := \{A \in M_n^{sa} : \|A\| \le R,\ |\mathrm{tr}_n(A^k) - m_k(\mu)| \le \varepsilon,\ k \le r\}, \tag{1.2}$$

where $\|A\|$ is the operator norm and $m_k(\mu) := \int x^k\, d\mu(x)$, the kth moment of μ. Then the limit

$$\chi_R(\mu; r, \varepsilon) := \lim_{n\to\infty} \left[\frac{1}{n^2} \log \Lambda_n\big(\Gamma_R(\mu; n, r, \varepsilon)\big) + \frac{1}{2}\log n\right] \tag{1.3}$$

exists for every $r \in \mathbb{N}$ and $\varepsilon > 0$, and

$$\lim_{r\to\infty,\, \varepsilon\to+0} \chi_R(\mu; r, \varepsilon) = \Sigma(\mu) + \frac{1}{2}\log(2\pi) + \frac{3}{4}. \tag{1.4}$$

(See [7, 5.6.2] for the existence of the limit in (1.3) while lim was originally lim sup in [14].)

In classical probability theory, the *Boltzmann-Gibbs entropy* $S(\mu)$ of a probability measure μ on $\mathbb{R}$ is given as

$$S(\mu) := -\int \frac{d\mu}{dx} \log \frac{d\mu}{dx}\, dx$$

if μ is absolutely continuous with respect to the Lebesgue measure dx and $\frac{d\mu}{dx}$ is the Radon-Nikodym derivative; otherwise $S(\mu) := -\infty$. The *relative entropy* (or the *Kullback-Leibler divergence*) $S(\mu, \nu)$ of μ with respect to another probability measure ν is defined as

$$S(\mu, \nu) := \int \frac{d\mu}{d\nu} \log \frac{d\mu}{d\nu}\, d\nu = \int \log \frac{d\mu}{d\nu}\, d\mu$$

if μ is absolutely continuous with respect to ν; otherwise $S(\mu, \nu) := +\infty$. If μ and ν are supported in $[-R, R]$, then these entropies have the asymptotic expressions as follows:

$$\begin{aligned} S(\mu) &= \lim_{r\to\infty,\, \varepsilon\to+0} \lim_{n\to\infty} \frac{1}{n} \log L^n\bigg(\Big\{(x_1, \dots, x_n) \in [-R, R]^n : \\ &\qquad \left|\frac{x_1^k + \cdots + x_n^k}{n} - m_k(\mu)\right| \le \varepsilon,\ k \le r\Big\}\bigg), \qquad (1.5) \\ -S(\mu, \nu) &= \lim_{r\to\infty,\, \varepsilon\to+0} \lim_{n\to\infty} \frac{1}{n} \log \nu^n\bigg(\Big\{(x_1, \dots, x_n) \in [-R, R]^n : \\ &\qquad \left|\frac{x_1^k + \cdots + x_n^k}{n} - m_k(\mu)\right| \le \varepsilon,\ k \le r\Big\}\bigg), \qquad (1.6) \end{aligned}$$

where L^n is the n-dimensional Lebesgue measure and ν^n is the n-fold product of ν. These expressions can be derived from Sanov's large deviation theorem for the empirical distribution of i.i.d. random variables (see [7, 5.1.1] for details).

The free entropy $\Sigma(\mu)$ is considered as the free probabilistic analogue of the Boltzmann-Gibbs entropy $S(\mu)$, and the asymptotic expression given in (1.2)–(1.4) (with scale n^{-2}) is the "free" counterpart of the expression (1.5) (with scale n^{-1}). Now, naturally arises the following question: What is the free analogue of the relative entropy $S(\mu,\nu)$? It turned out that the *free relative entropy* $\Sigma(\mu,\nu)$ of μ with respect to ν can be defined as

$$\Sigma(\mu,\nu) = -\iint \log|x-y|\,d(\mu-\nu)(x)\,d(\mu-\nu)(y)\,, \tag{1.7}$$

which is the logarithmic energy of a signed measure $\mu-\nu$ (see [8]). Here the following two definitions are available for precise meaning of (1.7):

(A) $\Sigma(\mu,\nu)$ is well-defined by (1.7) if $\log|x-y|$ is integrable with respect to the total variation measure $d|\mu-\nu|(x)\,d|\mu-\nu|(y)$; otherwise $\Sigma(\mu,\nu) := +\infty$.

(B) Based on the fact that $\varepsilon > 0 \mapsto -\iint \log(|x-y|+\varepsilon)\,d(\mu-\nu)(x)\,d(\mu-\nu)(y)$ is increasing as $\varepsilon \downarrow 0$ ([8, Lemma 3.6]), define

$$\Sigma(\mu,\nu) := \lim_{\varepsilon\to+0}\left[-\iint \log(|x-y|+\varepsilon)\,d(\mu-\nu)(x)\,d(\mu-\nu)(y)\right].$$

Note that if $\log|x-y|$ is integrable with respect to $d|\mu-\nu|(x)\,d|\mu-\nu|(y)$, then the definitions (A) and (B) are the same; this is the case in particular when $\Sigma(\mu) > -\infty$ and $\Sigma(\nu) > -\infty$.

In [8] the asymptotic expression of the free relative entropy $\Sigma(\mu,\nu)$ was obtained in the microstates approach. We here give a brief summary on some large deviation result related to random matrices, which is a basis of deriving the asymptotic expression of $\Sigma(\mu,\nu)$ and indeed play a crucial role in Sect. 2 as well.

Let $R > 0$ and Q be a real continuous function on $[-R,R]$. For each $n \in \mathbb{N}$ define the probability distribution $\tilde\lambda_n(Q;R)$ on $\mathbb{R}^n$ by

$$\begin{aligned}\tilde\lambda_n(Q;R) \;&:=\; \frac{1}{Z_n(Q;R)}\exp\Bigl(-n\sum_{i=1}^n Q(x_i)\Bigr)\prod_{i<j}|x_i-x_j|^2 \\ &\qquad\times\prod_{i=1}^n \chi_{[-R,R]}(x_i)\,dx_1dx_2\cdots dx_n\,,\end{aligned} \tag{1.8}$$

where $Z_n(Q:R)$ is the normalizing constant:

$$Z_n(Q;R) := \int_{-R}^{R}\cdots\int_{-R}^{R}\exp\Bigl(-n\sum_{i=1}^n Q(x_i)\Bigr)\prod_{i<j}|x_i-x_j|^2\,dx_1\cdots dx_n\,. \tag{1.9}$$

Moreover, let $\lambda_n(Q;R)$ be the probability distribution on M_n^{sa} which is invariant under unitary conjugation and whose joint eigenvalue distribution on $\mathbb{R}^n$ is $\tilde{\lambda}_n(Q;R)$; more explicitly,

$$\lambda_n(Q;R) := \big(dU \otimes \tilde{\lambda}_n(Q;R)\big) \circ \Phi_n^{-1}, \tag{1.10}$$

where dU is the Haar probability measure on the n-dimensional unitary group $\mathcal{U}_n$ and $\Phi_n : \mathcal{U}_n \times \mathbb{R}^n \to M_n^{sa}$ is defined as

$$\Phi_n(U,(x_1,\dots,x_n)) := U\mathrm{diag}(x_1,\dots,x_n)U^*.$$

One can consider $\lambda_n(Q;R)$ as the distribution of an $n\times n$ random selfadjoint matrix, or more explicitly $\lambda_n(Q;R)$ itself as a random matrix. The support of $\lambda_n(Q;R)$ is

$$(M_n^{sa})_R := \{A \in M_n^{sa} : \|A\| \le R\}. \tag{1.11}$$

The *empirical eigenvalue distribution* of this random matrix is

$$\frac{\delta(x_1)+\delta(x_2)+\cdots+\delta(x_n)}{n}$$

where $\delta(x)$ is the point measure at x and the $\mathbb{R}^n$-vector $(x_1,x_2,\dots,x_n)$ is distributed subject to the distribution (1.8). Let $\mathcal{M}([-R,R])$ denote the set of all probability measures supported in $[-R,R]$ equipped with the weak topology. Then we have the following large deviation theorem which is a matricial counterpart of the famous Sanov large deviation theorem ([2, 3]).

Theorem 1.1 *Let Q and Q_n ($n \in \mathbb{N}$) be real continuous functions on $[-R,R]$ such that $Q_n(x) \to Q(x)$ uniformly on $[-R,R]$. For each $n \in \mathbb{N}$ define the probability distribution $\tilde{\lambda}_n(Q_n;R)$ supported on $[-R,R]^n$ by (1.8) and the normalizing constant $Z_n(Q_n;R)$ by (1.9) with Q_n in place of Q. Then the finite limit*

$$B(Q;R) := \lim_{n\to\infty} \frac{1}{n^2} \log Z_n(Q_n;R)$$

exists, and if $(x_1,\dots,x_n) \in [-R,R]^n$ is distributed with the joint distribution $\tilde{\lambda}_n(Q_n;R)$, then the empirical distribution $\frac{1}{n}(\delta(x_1)+\cdots+\delta(x_n))$ satisfies the large deviation principle in the scale n^{-2} with the good rate function:

$$I(\mu) := -\Sigma(\mu) + \mu(Q) + B(Q;R) \quad \text{for} \quad \mu \in \mathcal{M}([-R,R]).$$

There exists a unique minimizer μ_Q of I with $I(\mu_Q)=0$ and $B(Q;R)$ is determined only by Q independently of $\{Q_n\}$. Furthermore, the above empirical distribution converges almost surely to μ_Q as $n \to \infty$ in weak topology.

Now let us return to the free relative entropy. Let ν be a compactly supported probability measure on $\mathbb{R}$, and assume that the function

$$Q_\nu(x) := 2\int \log|x-y|\,d\nu(y) \tag{1.12}$$

is finite and continuous (as a function on $\mathbb{R}$) at every $x \in \operatorname{supp}\nu$, where $\operatorname{supp}\nu$ means the support of ν. Then Q_ν is a continuous function on the whole $\mathbb{R}$, because Q_ν is always continuous on $\mathbb{R}\setminus\operatorname{supp}\nu$. For instance, this is the case when ν is absolutely continuous with respect to dx and $\frac{d\nu}{dx}$ is bounded. For $R > 0$ define the probability distribution $\lambda_n(\nu;R)$ on M_n^{sa} by putting $Q = Q_\nu$ in (1.8) and (1.10): $\lambda_n(\nu;R) := \lambda_n(Q_\nu;R)$. Then the next theorem was proved in [8, Theorem 3.8] by appealing to the above large deviation theorem in the case $Q_n = Q = Q_\nu$.

Theorem 1.2 *Let μ,ν be compactly supported probability measures, and assume that $Q_\nu(x)$ in (1.12) is continuous on $\mathbb{R}$. Then for any $R > 0$ with $\operatorname{supp}\mu$, $\operatorname{supp}\nu \subset [-R,R]$,*

$$-\Sigma(\mu,\nu) = \lim_{r\to\infty,\,\varepsilon\to+0}\ \lim_{n\to\infty}\frac{1}{n^2}\log\lambda_n(\nu;R)\big(\Gamma_R(\mu;n,r,\varepsilon)\big) \tag{1.13}$$

in either definition (A) or (B) for $\Sigma(\mu,\nu)$, where $\Gamma_R(\mu;n,r,\varepsilon)$ is as in (1.2).

The above expression (1.13) is the free analogue of (1.6). The reference measure $\lambda_n(\nu;R)$ on M_n^{sa} is a bit more complicated than the product ν^n on $\mathbb{R}^n$ in (1.6), but it is the right one in free (or matricial) probability. In fact, Theorem 1.1 (together with Lemma 2.1) says that the empirical eigenvalue distribution of the $n\times n$ selfadjoint random matrix having the distribution $\lambda_n(\nu;R)$ converges almost surely to ν, the minimizer of the rate function, as $n\to\infty$ in weak topology (hence in the distribution sense). In this way, Theorem 1.2 gives a justification for our free relative entropy $\Sigma(\mu,\nu)$. Another (more decisive) justification will be presented in Sect. 2.

1.2 Properties of free relative entropy

We will examine properties of $\Sigma(\mu,\nu)$ in either case of the definitions (A) or (B). They are summarized in the following. The free relative entropy differs from the classical one in the first property. But the other important properties are common. For a compact subset K of $\mathbb{R}$, let $\mathcal{M}(K)$ denote the set of all probability Borel measures supported in K. Also let $\mathcal{M}_\Sigma(K) := \{\mu\in\mathcal{M}(K) : \Sigma(\mu) > -\infty\}$.

Proposition 1.3 *Let μ,ν be compactly supported probability measures on $\mathbb{R}$.*

(1) *Symmetry:* $\Sigma(\mu,\nu) = \Sigma(\nu,\mu)$.

(2) *Strict positivity:* $\Sigma(\mu,\nu)\ge 0$, *and* $\Sigma(\mu,\nu) = 0$ *if and only if* $\mu=\nu$.

(3) *Joint convexity: If* $\Sigma(\mu_i) > -\infty$ *and* $\Sigma(\nu_i) > -\infty$ $(i = 1, 2)$*, then*

$$\Sigma(\alpha\mu_1 + (1-\alpha)\mu_2, \alpha\nu_1 + (1-\alpha)\nu_2) \le \alpha\Sigma(\mu_1, \nu_1) + (1-\alpha)\Sigma(\mu_2, \nu_2) \quad (1.14)$$

for $0 < \alpha < 1$*. Furthermore, in the case (B), (1.14) holds without the conditions* $\Sigma(\mu_i),\ \Sigma(\nu_i) > -\infty$ $(i = 1, 2)$.

(4) *Single strict convexity: If* $Q_\nu(x)$ *is continuous, then*

$$\Sigma(\alpha\mu_1 + (1-\alpha)\mu_2, \nu) \le \alpha\Sigma(\mu_1, \nu) + (1-\alpha)\Sigma(\mu_2, \nu) \quad (1.15)$$

for $0 < \alpha < 1$*. If* $\Sigma(\mu_i, \nu) < +\infty$ $(i = 1, 2)$ *and* $\mu_1 \neq \mu_2$*, (1.15) can be replaced by strict inequality. Furthermore, in the case (B), (1.15) holds without the continuity of* $Q_\nu(x)$.

(5) *Joint lower semicontinuity: Let* K *be any compact subset of* $\mathbb{R}$*. Then* $\Sigma(\mu, \nu)$ *is weakly jointly lower semicontinuous on* $\mathcal{M}_\Sigma(K)$*. Furthermore, in the case (B), it is weakly jointly lower semicontinuous on* $\mathcal{M}(K)$.

(6) *Single lower semicontinuity: Let* K *be any compact subset of* $\mathbb{R}$*. If* $Q_\nu(x)$ *is continuous, then* $\Sigma(\mu, \nu)$ *is weakly lower semicontinuous in* μ *on* $\mathcal{M}(K)$.

2 Perturbation via free relative entropy

2.1 Free Perturbation Theory

Let K be a fixed compact subset of $\mathbb{R}$ having positive capacity. Let $C_\mathbb{R}(K)$ denote the space of all real continuous functions on K. For $\mu \in \mathcal{M}(K)$ and $h \in C_\mathbb{R}(K)$ we write $\mu(h)$ for $\int_K h\,d\mu$. Throughout this section, let $\nu \in \mathcal{M}(K)$ be such that the function $Q = Q_\nu$ given in (1.12) is continuous on K. We adopt (B) as the definition of $\Sigma(\mu, \nu)$ in this section, but no crucial difference between (A) and (B) will occur; in fact, $\Sigma(\mu, \nu)$ is uniquely determined by (1.13) whenever the assumption on ν in Theorem 1.2 is supposed. For given $h \in C_\mathbb{R}(K)$ define the *weighted energy integral*

$$E_h(\mu) := \iint \log\frac{1}{|x-y|}\,d\mu(x)\,d\mu(y) + \int h\,d\mu = -\Sigma(\mu) + \mu(h)$$

for $\mu \in \mathcal{M}(K)$. We state the fundamental result in the theory of weighted potentials ([12, I.1.3 and I.3.1]) in a reduced form of the next lemma, which plays a key role in the sequel.

Lemma 2.1 *For every* $h \in C_\mathbb{R}(K)$ *the following assertions hold:*

(i) *There exists a unique* $\mu_h \in \mathcal{M}(K)$ *such that*

$$E_h(\mu_h) = \inf\{E_h(\mu) : \mu \in \mathcal{M}(K)\}.$$

(ii) *$E_h(\mu_h)$ and $\Sigma(\mu_h)$ are finite.*

(iii) *The minimizer μ_h is characterized as $\mu_h \in \mathcal{M}(K)$ such that for some $B \in \mathbb{R}$*

$$2\int \log|x-y|\,d\mu_h(y) \begin{cases} \geq h(x)+B & \text{for all } x \in \operatorname{supp}\mu_h, \\ \leq h(x)+B & \text{for quasi-every } x \in K. \end{cases}$$

In this case, $B = -2E_h(\mu_h) + \mu_h(h)$.

For $\nu \in \mathcal{M}(K)$ fixed as above, the *Legendre transform* of $\mu \in \mathcal{M}(K) \mapsto \Sigma(\mu,\nu)$ is defined as

$$c(h,\nu) := \sup\{-\mu(h) - \Sigma(\mu,\nu) : \mu \in \mathcal{M}(K)\}$$

for each $h \in C_{\mathbb{R}}(K)$.

Theorem 2.2 *With the above definitions, the following assertions hold:*

(i) *$c(\cdot,\nu)$ is a convex function on $C_{\mathbb{R}}(K)$ satisfying*

$$-\nu(h) \leq c(h,\nu) \leq \|h\|$$

(in particular, $c(0,\nu) = 0$) where $\|h\|$ is the sup-norm, and it is decreasing, i.e. $c(h_1,\nu) \geq c(h_2,\nu)$ if $h_1 \leq h_2$. Moreover,

$$|c(h_1,\nu) - c(h_2,\nu)| \leq \|h_1 - h_2\|$$

for all $h_1, h_2 \in C_{\mathbb{R}}(K)$.

(ii) *For every $\mu \in \mathcal{M}(K)$,*

$$\Sigma(\mu,\nu) = \sup\{-\mu(h) - c(h,\nu) : h \in C_{\mathbb{R}}(K)\}. \tag{2.16}$$

(iii) *For every $h \in C_{\mathbb{R}}(K)$ there exists a unique $\nu^h \in \mathcal{M}(K)$ such that*

$$-\nu^h(h) - \Sigma(\nu^h,\nu) = c(h,\nu).$$

Moreover, $\Sigma(\nu^h)$ is finite and

$$c(h,\nu) = \Sigma(\nu^h) + \Sigma(\nu) - \nu^h(Q+h).$$

(iv) *For every $h \in C_{\mathbb{R}}(K)$ and $\mu \in \mathcal{M}(K)$, $\mu = \nu^h$ if and only if*

$$c(h+k,\nu) \geq c(h,\nu) - \mu(k) \quad \text{for all} \quad k \in C_{\mathbb{R}}(K).$$

We call ν^h in Theorem 2.2 the *perturbed probability measure* of ν by h (via free relative entropy). Note that the variational expression (2.16) of $\Sigma(\mu,\nu)$ is valid for any choice of a compact $K \subset \mathbb{R}$ such that $K \supset \operatorname{supp}\mu, \operatorname{supp}\nu$. Clearly, $\nu^{h+\alpha} = \nu^h$ and $c(h+\alpha,\nu) = c(h,\nu) - \alpha$ for $\alpha \in \mathbb{R}$.

It is instructive to consider the perturbed measure ν^h in comparison with the similar perturbation via relative entropy. For any $\nu \in \mathcal{M}(K)$ and $h \in C_{\mathbb{R}}(K)$, it is well-known that

$$\log \nu(e^{-h}) = \sup\{-\mu(h) - S(\mu,\nu) : \mu \in \mathcal{M}(K)\}$$

and the probability measure $\mu_0 := \frac{e^{-h}}{\nu(e^{-h})}\nu$ (i.e. $\frac{d\mu_0}{d\nu} = \frac{e^{-h}}{\nu(e^{-h})}$) is a unique maximizer of $-\mu(h) - S(\mu,\nu)$ for $\mu \in \mathcal{M}(K)$. In fact, this can be easily verified by using the strict positivity of $S(\mu,\mu_0)$. Moreover, for every $\mu \in \mathcal{M}(K)$,

$$S(\mu,\nu) = \sup\{-\mu(h) - \log \nu(e^{-h}) : h \in C_{\mathbb{R}}(K)\}.$$

The probability measure μ_0 perturbed from ν via the relative entropy $S(\mu,\nu)$ is the so-called *Gibbs ensemble.* The above $c(h,\nu)$ is considered as the "free" counterpart of $\log \nu(e^{-h})$, and the characterization of ν^h in the above (iv) is the "free" analogue of the so-called *variational principle* for Gibbs ensembles ([11]). It is worth noting that this type of perturbation theory via relative entropy was developed even in the quantum probabilistic setting on operator algebras ([10], [4], [9, Sect. 12]).

We shall write $\nu^{h,\Sigma}$ for ν^h in Theorem 2.2 and $\nu^{h,S}$ for the above μ_0, when both perturbed measures via $\Sigma(\mu,\nu)$ and $S(\mu,\nu)$ are simultaneously treated. A simple expression of $c(h,\nu)$ such as $\log \nu(e^{-h})$ is not available; nevertheless we shall give an asymptotic expression of $c(h,\nu)$ in Sect. 2.3.

Proposition 2.3 *For every* $\mu \in \mathcal{M}(K)$,

$$\Sigma(\mu,\nu^h) \le \Sigma(\mu,\nu) + \mu(h) + c(h,\nu).$$

Moreover, if $\operatorname{supp}\mu \subset \operatorname{supp}\nu^h$, *then*

$$\Sigma(\mu,\nu^h) = \Sigma(\mu,\nu) + \mu(h) + c(h,\nu).$$

Corollary 2.4 *For every* $h \in C_{\mathbb{R}}(K)$,

$$\Sigma(\nu^h,\nu) \le \frac{\nu(h) - \nu^h(h)}{2} \le \|h\|,$$

$$c(h,\nu) \ge -\nu(h) + \Sigma(\nu^h,\nu) \ge -\frac{\nu(h) + \nu^h(h)}{2}.$$

Furthermore, if $\operatorname{supp}\nu \subset \operatorname{supp}\nu^h$, *then*

$$\Sigma(\nu^h,\nu) = \frac{\nu(h) - \nu^h(h)}{2},$$

$$c(h,\nu) = -\nu(h) + \Sigma(\nu^h,\nu) = -\frac{\nu(h) + \nu^h(h)}{2}.$$

The next proposition is the chain rule for the perturbation $\nu \mapsto \nu^h$.

Proposition 2.5 *Let $h, k \in C_{\mathbb{R}}(K)$. If $Q_{\nu^h}(x) := 2\int \log|x-y|\,d\nu^h(y)$ as well as $Q = Q_\nu$ is continuous on K and* $\mathrm{supp}\,(\nu^h)^k \subset \mathrm{supp}\,\nu^h$, *then*

$$(\nu^h)^k = \nu^{h+k},$$

$$c(h+k,\nu) = c(h,\nu) + c(k,\nu^h).$$

In particular, these hold if $\mathrm{supp}\,\nu^h = K$ *and* $Q_{\nu^h} = Q + h$.

Corollary 2.6 *Assume either* (a) *or* (b) *in the following:*

(a) *$\mu \in \mathcal{M}(K)$ is such that Q_μ as well as Q_ν is continuous on K, and $h := Q_\mu - Q_\nu$,*

(b) *$h \in C_{\mathbb{R}}(K)$ and $\mu := \nu^h$ satisfies* $\mathrm{supp}\,\nu \subset \mathrm{supp}\,\mu$.

Then for each $0 \le \lambda \le 1$,

$$\nu^{\lambda h} = (1-\lambda)\nu + \lambda\mu,$$

$$\Sigma(\nu^{\lambda h},\nu) = \lambda^2\Sigma(\mu,\nu),$$

$$c(\lambda h,\nu) = \lambda\nu(h) + \lambda^2\Sigma(\mu,\nu).$$

As for the perturbation $\nu \mapsto \nu^{h,S}$ via relative entropy, $\mathrm{supp}\,\nu^{h,S} = \mathrm{supp}\,\nu$ is obvious and the formulas

$$S(\mu,\nu^{h,S}) = S(\mu,\nu) + \mu(h) + \log\nu(e^{-h}),$$

$$(\nu^{h,S})^{k,S} = \nu^{h+k,S},$$

$$\log\nu(e^{-(h+k)}) = \log\nu(e^{-h}) + \log\nu^{h,S}(e^{-k})$$

generally hold. The relation between ν and $\nu^h = \nu^{h,\Sigma}$ is more complicated than that between ν and $\nu^{h,S}$. However, the formulas in Corollary 2.6 (though they do not generally hold) are quite simple compared with those for $\nu^{\lambda h,S}$; in fact, $\nu^{\lambda h,S}$ $(0 \le \lambda \le 1)$ is not a line segment, and $\frac{d^2}{d\lambda^2}S(\nu^{\lambda h,S},\nu)$ and $\frac{d^2}{d\lambda^2}S(\nu,\nu^{\lambda h,S})$ are non-constant functions of λ. The simple formulas for $\nu^{\lambda h,\Sigma}$ in Corollary 2.6 correspond to the flatness of the Riemannian metric induced by the free entropy ([8, Sect. 4]).

The next proposition gives a simple sufficient condition for $\mu \in \mathcal{M}(K)$ to be a perturbed probability measure of ν.

Proposition 2.7 *If $\mu \in \mathcal{M}(K)$ satisfies $\mu \le \alpha\nu$ for some constant $\alpha \ge 1$, then $Q_\mu(x) := 2\int \log|x-y|\,d\mu(y)$ is continuous on K and there exists an $h \in C_{\mathbb{R}}(K)$ such that $\mu = \nu^h$ and*

$$Q_\mu(x) \ge \alpha Q_\nu(x) + 2(1-\alpha)\log R \qquad (x \in K),$$

where $R := \max\{|x-y| : x, y \in K\}$, *the diameter of K.*

Corollary 2.8 *If $\mu \in \mathcal{M}(K)$ satisfies $\beta\nu \le \mu \le \alpha\nu$ for some constants $0 < \beta \le 1 \le \alpha$, then there exists an $h \in C_{\mathbb{R}}(K)$ such that $\mu = \nu^h$ and*

$$(1-\alpha)(2\log R - Q_\nu) \le h \le (1-\beta)(2\log R - Q_\nu),$$

$$\Sigma(\mu,\nu) \le (\alpha(\alpha-1) + (1-\beta))(\log R - \Sigma(\nu)).$$

where R is the diameter of K. (Note $Q_\nu \le 2\log R$.)

2.2 Convergence of perturbed measures

The aim of this subsection is to show the continuity properties in h of the perturbation ν^h introduced in the previous section. Define

$$d(\mu_1,\mu_2) := \Sigma(\mu_1,\mu_2)^{1/2}\,(\in [0,+\infty))$$

for $\mu_1,\mu_2 \in \mathcal{M}_\Sigma(K)$. The next lemma is an application of the series expansion of the free entropy due to Haagerup [5], and it will play an important role in the proof of the following theorem.

Lemma 2.9 *The above defined $d(\mu_1,\mu_2)$ is a metric on $\mathcal{M}_\Sigma(K)$ and the d-topology is strictly stronger than the weak topology (restricted on $\mathcal{M}_\Sigma(K)$) and $(\mathcal{M}_\Sigma(K),d)$ is a non-compact Polish space.*

Theorem 2.10 *If $h, h_n \in C_{\mathbb{R}}(K)$, $n \in \mathbb{N}$, satisfy $\|h_n - h\| \to 0$, then the following convergences hold:*

(i) $c(h_n,\nu) \to c(h,\nu)$.

(ii) $\Sigma(\nu^{h_n},\mu) \to \Sigma(\nu^h,\mu)$ *for every* $\mu \in \mathcal{M}_\Sigma(K)$; *in particular,* $\Sigma(\nu^{h_n},\nu^h) \to 0$.

(iii) $\nu^{h_n} \to \nu^h$ *weakly.*

(iv) $\nu^{h_n}(h_n) \to \nu^h(h)$.

(v) $\Sigma(\nu^{h_n}) \to \Sigma(\nu^h)$.

Concerning the perturbation $\nu^{h,S}$ via relative entropy, the continuity of $h \mapsto \nu^{h,S}$ can be straightforwardly seen from the explicit formula $\nu^{h,S} = \frac{e^{-h}}{\nu(e^{-h})}\nu$. In fact, when $h_n, h \in C_{\mathbb{R}}(K)$ and $h_n \to h$ boundedly pointwise, i.e. $\sup_n \|h_n\| < +\infty$ and $h_n(x) \to h(x)$ for every $x \in K$, one gets the w*-convergence $\nu^{h_n,S} \to \nu^{h,S}$ by the Lebesgue bounded convergence theorem. However, it is not known whether the w*-convergence $\nu^{h_n,\Sigma} \to \nu^{h,\Sigma}$ follows or not under this convergence $h_n \to h$ weaker than $\|h_n - h\| \to 0$.

The next proposition says that the weak convergence and the d-convergence are equivalent for a sequence $\{\mu_n\}$ in $\mathcal{M}(K)$ such that μ_n's are uniformly dominated by ν.

Proposition 2.11 *Let $\mu_n, \mu \in \mathcal{M}(K)$ for $n \in \mathbb{N}$, and assume that there is an $\alpha \geq 1$ such that $\mu_n \leq \alpha\nu$ for all $n \in \mathbb{N}$. Then $\mu_n \to \mu$ weakly if and only if $\Sigma(\mu_n, \mu) \to 0$. In this case, $\Sigma(\mu_n) \to \Sigma(\mu)$ and $\Sigma(\mu_n, \mu') \to \Sigma(\mu, \mu')$ for every $\mu' \in \mathcal{M}_\Sigma(K)$.*

As for relative entropy, it is known that if μ_n, ν_n are probability measures on $\mathbb{R}$ such that $\|\mu_n - \mu\| \to 0$, $\|\nu_n - \nu\| \to 0$ and there is an $\alpha > 0$ such that $\mu_n \leq \alpha\nu_n$ for all $n \in \mathbb{N}$, then $S(\mu_n, \nu_n) \to S(\mu, \nu)$. (This is true in the operator algebra setting, see [1, Theorem 3.7].) However, this fails to hold for free relative entropy; one can easily provide an example of $\mu_n, \nu_n \in \mathcal{M}_\Sigma(K)$ such that $\|\mu_n - \nu\| \to 0$, $\|\nu_n - \nu\| \to 0$ and $\mu_n \leq \alpha\nu_n$ for all $n \in \mathbb{N}$, but $\Sigma(\mu_n, \nu_n) \not\to 0$.

2.3 From relative entropy to free relative entropy

We consider a sequence of $n \times n$ selfadjoint random matrices naturally perturbed via relative entropy, and show that the perturbed measure ν^h via free relative entropy is the limit distribution of the empirical eigenvalue distributions of perturbed random matrices as the size n goes to ∞. In so doing, we can also express the free relative entropy $\Sigma(\nu^h, \nu)$ as the limit (with normalization) of the relative entropy defined on the matrix space M_n^{sa}.

Throughout this subsection, we assume for simplicity that K is a finite interval $[-R, R]$. Let $\nu \in \mathcal{M}([-R, R])$ be fixed so that $Q = Q_\nu$ in (1.12) is a continuous function on $[-R, R]$. For each $n \in \mathbb{N}$ we simply write $\lambda_n(\nu)$ for the probability measure $\lambda_n(\nu; R) = \lambda_n(Q; R)$ on $(M_n^{sa})_R$ given in (1.8)–(1.11). Here note that $(M_n^{sa})_R$ is a compact subset of M_n^{sa} identified with a Euclidean space $\mathbb{R}^{n^2}$. For a given $h \in C_\mathbb{R}([-R, R])$ and $n \in \mathbb{N}$, let $\phi_n(h)$ denote the real continuous function on $(M_n^{sa})_R$ defined by

$$\phi_n(h)(A) := n^2 \mathrm{tr}_n(h(A)) \quad \text{for} \quad A \in (M_n^{sa})_R,$$

where $h(A)$ is defined via functional calculus and tr_n is the normalized trace on M_n. Then one can get the probability measure $\lambda_n(\nu)^{\phi_n(h),S}$ on $(M_n^{sa})_R$ which is the perturbed measure of $\lambda_n(\nu)$ by $\phi_n(h)$ via relative entropy; namely, $\lambda_n(\nu)^{\phi_n(h),S}$ is a unique maximizer of the functional

$$-\eta(\phi_n(h)) - S(\eta, \lambda_n(\nu)) \quad \text{for} \quad \eta \in \mathcal{M}((M_n^{sa})_R),$$

where $\mathcal{M}((M_n^{sa})_R)$ is the set of all probability Borel measures on $(M_n^{sa})_R$. In fact, as mentioned after Theorem 2.2, it is given by

$$\lambda_n(\nu)^{\phi_n(h),S} = \frac{e^{-\phi_n(h)}}{\lambda_n(\nu)(e^{-\phi_n(h)})}\lambda_n(\nu)$$

and

$$-\lambda_n(\nu)^{\phi_n(h),S}(\phi_n(h)) - S(\lambda_n(\nu)^{\phi_n(h),S}, \lambda_n(\nu)) = \log \lambda_n(\nu)(e^{-\phi_n(h)}).$$

In the sequel we use the following notations for short:

$$\Delta(x) := \prod_{i<j}(x_i - x_j)^2, \quad dx := dx_1 dx_2 \cdots dx_n .$$

Lemma 2.12 *With the above notations,*

$$\lambda_n(\nu)^{\phi_n(h),S} = \lambda_n(Q + h; R),$$

that is, $\lambda_n(\nu)^{\phi_n(h),S}$ is invariant under unitary conjugation and its joint eigenvalue distribution is

$$\tilde{\lambda}_n(Q + h; R) = \frac{1}{Z_n(Q + h; R)} \exp\left(-n \sum_{i=1}^{n}(Q(x_i) + h(x_i))\right)\Delta(x) \prod_{i=1}^{n} \chi_{[-R,R]}(x_i)\, dx,$$

where $Z_n(Q + h; R)$ is defined by (1.9) *with $Q + h$ in place of Q. Furthermore,*

$$\lambda_n(\nu)(e^{-\phi_n(h)}) = \frac{Z_n(Q + h; R)}{Z_n(Q; R)} .$$

The measure $\lambda_n(\nu)^{\phi_n(h),S}$ on $(M_n^{sa})_R$ may be considered as an $n \times n$ selfadjoint random matrix which is a perturbation of $\lambda_n(\nu)$ via relative entropy. The next theorem says that this perturbation of $\lambda_n(\nu)$ via relative entropy on the matrix space approaches asymptotically as $n \to \infty$ to ν^h $(= \nu^{h,\Sigma})$, the perturbation of ν via free relative entropy. In particular, it justifies our formulation of free relative entropy. In the theorem we actually treat a sequence of perturbed measures $\lambda_n(\nu)^{\phi_n(h_n),S}$ determined by separate $h_n \in C_{\mathbb{R}}([-R, R])$ for each n satisfying $\|h_n - h\| \to 0$. The proof is based on the large deviation result presented in Theorem 1.1.

Theorem 2.13 *Let $\nu \in \mathcal{M}([-R, R])$ be as above. If $h, h_n \in C_{\mathbb{R}}([-R, R])$, $n \in \mathbb{N}$, satisfy $\|h_n - h\| \to 0$, then the following hold:*

(i) *The empirical eigenvalue distribution of $\lambda_n(\nu)^{\phi_n(h_n),S}$ converges almost surely to ν^h as $n \to \infty$ in weak topology.*

(ii)

$$\nu^h(h) = \lim_{n\to\infty} \frac{1}{n^2} \lambda_n(\nu)^{\phi_n(h_n),S}(\phi_n(h_n)) .$$

(iii)

$$\Sigma(\nu^h, \nu) = \lim_{n\to\infty} \frac{1}{n^2} S(\lambda_n(\nu)^{\phi_n(h_n),S}, \lambda_n(\nu)) .$$

(iv) *With $B(Q;R)$ defined by* (1.1) *and $B(Q+h;R)$ similarly with $Q+h$ in place of Q,*

$$c(h,\nu) = \lim_{n\to\infty} \frac{1}{n^2} \log \lambda_n(\nu)(e^{-\phi_n(h_n)}) = B(Q+h;R) - B(Q;R)\,.$$

(v)

$$\nu(h) - \nu^h(h) - \Sigma(\nu^h,\nu) = \lim_{n\to\infty} \frac{1}{n^2} S(\lambda_n(\nu), \lambda_n(\nu)^{\phi_n(h_n),S})\,.$$

Hence, if $\mathrm{supp}\,\nu \subset \mathrm{supp}\,\nu^h$, *then*

$$\Sigma(\nu^h,\nu) = \lim_{n\to\infty} \frac{1}{n^2} S(\lambda_n(\nu), \lambda_n(\nu)^{\phi_n(h_n),S})\,.$$

Besides its conceptual importance, Theorem 2.13 supplies the asymptotic formulas of $\nu^h(h)$ and $c(h,\nu)$ (when $h_n = h$ for all n); thus we obtain the asymptotic formula of $\Sigma(\nu^h,\nu) = -\nu^h(h) - c(h,\nu)$. In particular, we state the following:

Corollary 2.14 *Let μ,ν be compactly supported probability measures on $\mathbb{R}$ such that Q_μ and Q_ν are continuous. Then for any $R>0$ with* $\mathrm{supp}\,\mu, \mathrm{supp}\,\nu \subset [-R,R]$,

$$\begin{aligned}
&\Sigma(\mu,\nu) \\
&\quad = \lim_{n\to\infty} \frac{\int_{-R}^{R}\cdots\int_{-R}^{R}\left(\frac{1}{n}\sum_{i=1}^{n}(Q_\nu(x_i) - Q_\mu(x_i))\right)\exp\left(-n\sum_{i=1}^{n} Q_\mu(x_i)\right)\Delta(x)\,dx}{\int_{-R}^{R}\cdots\int_{-R}^{R}\exp\left(-n\sum_{i=1}^{n} Q_\mu(x_i)\right)\Delta(x)\,dx} \\
&\qquad + \lim_{n\to\infty} \frac{1}{n^2} \log \frac{\int_{-R}^{R}\cdots\int_{-R}^{R}\exp\left(-n\sum_{i=1}^{n} Q_\nu(x_i)\right)\Delta(x)\,dx}{\int_{-R}^{R}\cdots\int_{-R}^{R}\exp\left(-n\sum_{i=1}^{n} Q_\mu(x_i)\right)\Delta(x)\,dx}\,.
\end{aligned}$$

The free relative entropy $\Sigma(\mu,\nu)$ is symmetric in its two variables unlike the relative entropy, while the formula in Corollary 2.14 is not symmetric in μ and ν. On the other hand, the perturbation via relative entropy is symmetric in the sense that if μ is the perturbation of ν by h, then ν is the perturbation of μ by $-h$. This type of symmetry does not hold in the perturbation via free relative entropy, even though the limiting procedure from the perturbation via relative entropy to that via free relative entropy was established in Theorem 2.13

References

[1] H. Araki, Relative entropy for states of von Neumann algebras II, *Publ. Res. Inst. Math. Sci.* **13** (1977), 173–192.

[2] A. Dembo and O. Zeitouni, *Large Deviation Techniques and Applications*, Second edition, Springer, New York, 1998.

[3] J. D. Deuschel and D. W. Stroock, *Large Deviations*, Academic Press, Boston, 1989.

[4] M. J. Donald, Relative hamiltonian which are not bounded from above, *J. Funct. Anal.* **91** (1990), 143–173.

[5] U. Haagerup, manuscript.

[6] F. Hiai, M. Mizuo and D. Petz, Free relative entropy and perturbation of probability measures, preprint.

[7] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy*, Mathematical Surveys and Monographs, Vol. 77, Amer. Math. Soc., Providence, 2000.

[8] M. Mizuo, Large deviations and microstate free relative entropy, *Interdiscip. Inform. Sci.* **6** (2000), to appear.

[9] M. Ohya and D. Petz, *Quantum Entropy and Its Use*, Springer, Berlin-Heidelberg, 1993.

[10] D. Petz, A variational expression for the relative entropy, *Comm. Math. Phys.* **114** (1988), 345–349.

[11] D. Ruelle, *Thermodynamic Formalism*, Encyclopedia of Math. and Its Appl., Vol. 5, Addison-Wesley, London, 1978.

[12] E. B. Saff and V. Totik, *Logarithmic Potentials with External Fields*, Springer, Berlin-Heidelberg-New York, 1997.

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

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

[15] D.V. Voiculescu, K.J. Dykema and A. Nica, *Free Random Variables*, CRM Monograph Ser., Vol. 1, Amer. Math. Soc., 1992.