Adv. Oper. Theory 2 (2017), no. 1, 50–58
http://doi.org/10.22034/aot.1610.1041
ISSN: 2538-225X (electronic)
http://aot-math.org

ON THE GENERALIZED FREE ENERGY INEQUALITY

N. BEBIANO,[1] and J. DA PROVIDÊNCIA[2]

Communicated by Y. T. Poon

ABSTRACT. The generalized free energy inequality known from statistical mechanics is stated in the finite dimension setting and the maximizing matrix is restored. Our approach uses the maximum-entropy inference principle and numerical range methods.

1. INTRODUCTION

The maximum entropy principle goes back to Boltzmann in the nineteenth century [6] and is one of the standard techniques in quantum inference problems [9].

Pure states of a quantum system are described by vectors of a Hilbert space $\mathcal{H}$, throughout assumed of finite dimension. Quantum observables are represented by Hermitian matrices in the algebra M_n of complex $n \times n$ matrices, $n \in \mathbb{N}$. We focus on the case of two observables encoded into a single matrix $A = H + iK \in M_n$ with *real part*

$$H = \Re(A) = \frac{1}{2}(A + A^*)$$

and *imaginary part*

$$K = \Im(A) = \frac{1}{2i}(A - A^*).$$

The set of Hermitian matrices

$$\mathcal{D}_n := \{\rho \in M_n : \rho \geq 0,\ \mathrm{Tr}\rho = 1\},$$

Date: Received: Oct. 28, 2016; Accepted: Jan. 15, 2017.
*Corresponding author.
2010 *Mathematics Subject Classification.* Primary 47A12; Secondary 62F30, 54C10.
Key words and phrases. Maximum-entropy inference, generalized free energy inequality, von Neumann entropy, numerical range.

where $\rho \geq 0$ means that the matrix ρ is positive semi-definite, is the so called *state space*. In physics, elements of $\mathcal{D}_n$ are called *density matrices* and we use synonymously the terms state and density matrices. The eigenvalues of a density matrix are nonnegative and give the probabilities of the physical states described by the corresponding eigenvectors. Density matrices of rank one represent *pure states* and density matrices of rank greater than one represent *mixed states*. Let $\rho \in M_n,\ \rho \geq 0,\ \mathrm{Tr}\rho = 1$ represent a density matrix. The expected values of H and K in the state ρ are given, respectively, by

$$\mathrm{Tr}H\rho = x,\ \mathrm{Tr}K\rho = y. \tag{1.1}$$

The *von Neumann entropy* of the state $\rho \in \mathcal{D}_n$ is

$$S(\rho) = -\mathrm{Tr}\rho\log\rho, \tag{1.2}$$

where by convention $0\log 0 = 0$. The von Neumann entropy quantifies the degree of disorder of the state ρ [12]. The *equilibrium mixed state* is identified with the most disordered state compatible with the expected values of the observables H and K.

The following problem arises: given $\mathrm{Tr}H\rho$ and $\mathrm{Tr}K\rho$, to determine ρ which describes the equilibrium mixed state.

This problem, known in information theory as the maximum entropy inference problem, consists on the maximization of the entropy $S = -\mathrm{Tr}\rho\log\rho$ with respect to ρ subject to the constraint (1.1). This is equivalent to the maximization, with respect to ρ, of the functional

$$\Xi(\beta, \gamma, \rho) = \beta\mathrm{Tr}\rho H + \gamma\mathrm{Tr}\rho K - \mathrm{Tr}\rho\log\rho, \tag{1.3}$$

where $\beta, \gamma \in \mathbb{R}$ are Lagrange multipliers. This functional is precisely the *generalized free energy*, known from statistical mechanics, multiplied by β.

The maximum entropy inference problem [4, 9, 10] is mathematically related to the concept of *numerical range*, a classical object in operator theory [8]. We recall that the *numerical range* of A is the set of complex numbers

$$W(A) = \{x^*Ax : x \in \mathbb{C}^n, x^*x = 1\},$$

which is convex as asserted by the Toeplitz-Hausdorff Theorem [5, 11]. In other words $W(A)$ is the set of all expected values of H and K in pure states. We notice that $x + iy \in W(H + iK)$, with x, y given by (1.1).

A generalization of the numerical range, introduced by Westwick, is the *c-numerical range* of $A \in M_n$ with $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$, defined by

$$W_c(A) = \{\sum_{j=1}^{n} c_j x_j^* A x_j : \{x_1, \ldots x_n\} \text{ is an orthonormal basis of } \mathbb{C}^n\}.$$

Obviously, $W_c(A)$ reduces to $W(A)$ for $c = (1, 0, \ldots, 0)$.

The $n!$ points

$$z_\sigma = \sum_j c_{\sigma(j)}\alpha_j,\ \sigma \in S_n,$$

where S_n is the symmetric group of degree n and $\alpha_1, \ldots \alpha_n$ denote the eigenvalues of A, are called *σ-points* and they obviously belong to $W_c(A)$. We recall that

$W_c(A)$ is convex and when A is normal, $W_c(A)$ is the convex hull of the σ-points [1, 13]. In our study, we assume that $x + iy$ is in the interior of $W(H + iK)$. Points in a line segment on the boundary of $W(H + iK)$ are generated by linear combinations of the vectors that generate the end points of that line segment, which are readily determined.

The rest of this note is organized as follows. In Section 2 the famous generalized free energy inequality in statistical mechanics is derived. With the help of maximum entropy method, the maximum entropy inference problem is solved and the maximizing matrix is restored. In Section 3, a solution is illustrated with an Example.

2. Generalized free energy inequality

Next, we derive the free energy inequality for two observables, also known as the generalized free energy inequality, and characterize the occurrence of equality. In [2, 3] inequalities in the same framework have been obtained.

Theorem 2.1. *For H, K Hermitian matrices, and any ρ positive semidefinite, and fixed parameters $\beta, \gamma \in \mathbb{R}$ we have*

$$\log \mathrm{Tr} \mathrm{e}^{\beta H+\gamma K} \geq \mathrm{Tr}\rho(\beta H + \gamma K - \log \rho) \tag{2.1}$$

with equality occurring if and only if

$$\rho = \rho_0 := \frac{\mathrm{e}^{\beta H+\gamma K}}{\mathrm{Tr}\mathrm{e}^{\beta H+\gamma K}}. \tag{2.2}$$

Proof. As the trace is invariant under unitary similarity transformations, let us replace ρ by $U\rho U^*$ in (1.3), where U is a unitary matrix. Obviously, $\mathrm{Tr}\rho \log \rho$ remains unchanged. The maximum of

$$\mathrm{Tr}U\rho U^*(\beta H + \gamma K - \log(U\rho U^*)) \tag{2.3}$$

with respect to U, occurs when

$$[U\rho U^*, (\beta H + \gamma K)] = 0,$$

where, as usual, $[X, Y] = XY - YX$ denotes the commutator of X and Y. This easily follows, assuming that the maximum is reached for U in the compact unitary group $\mathcal{U}_n$, replacing U by $\exp(i\epsilon S)U$, where S is an arbitrary Hermitian matrix and ϵ a sufficiently small real number, and expanding (2.3) up to first order in ϵ. Since this term must vanish for any S, we conclude that $[U\rho U^*, (\beta H+\gamma K)] = 0$. Therefore, the Hermitian matrices $U\rho U^*$ and $(\beta H + \gamma K)$ are simultaneously unitarily diagonalizable. Let us denote the real eigenvalues of ρ and $(\beta H + \gamma K)$, respectively, by $\eta_1, \ldots, \eta_n$ and by $\lambda_1, \ldots, \lambda_n$, so that we may write

$$\mathrm{Tr} U\rho U^*(\beta H + \gamma K - \log(U\rho U^*)) = \sum_j (\eta_j \lambda_j - \eta_j \log \eta_j)$$
$$= \sum_j \eta_j (\log \mathrm{e}^{\lambda_j} - \log \eta_j) = -\sum_j \eta_j \left(\log \left(\eta_j \mathrm{e}^{-\lambda_j} \sum_k \mathrm{e}^{\lambda_k} \right) - \log \sum_k \mathrm{e}^{\lambda_k} \right)$$
$$= -\sum_j \frac{\mathrm{e}^{\lambda_j}}{\sum_k \mathrm{e}^{\lambda_k}} \left(\eta_j \mathrm{e}^{-\lambda_j} \sum_k \mathrm{e}^{\lambda_k} \right) \log \left(\eta_j \mathrm{e}^{-\lambda_j} \sum_k \mathrm{e}^{\lambda_k} \right) + \log \sum_j \mathrm{e}^{\lambda_j}$$
$$\leq -\sum_j \frac{\mathrm{e}^{\lambda_j}}{\sum_k \mathrm{e}^{\lambda_k}} \left(\eta_j \mathrm{e}^{-\lambda_j} \sum_k \mathrm{e}^{\lambda_k} - 1 \right) + \log \sum_j \mathrm{e}^{\lambda_j}$$
$$= \log \sum_j \mathrm{e}^{\lambda_j} = \log \mathrm{Tr} \mathrm{e}^{\beta H + \gamma K},$$

where the inequality follows because $x \log\, x \geq x - 1$. Thus, we get the inequality in (2.1). It is obvious that the equality occurs if and only if $\eta_j = \mathrm{e}^{\lambda_j} / \sum_k \mathrm{e}^{\lambda_k}$. □

It is convenient to introduce the function $\Psi : \mathbb{R}^2 \to \mathbb{R}$ defined as

$$\Psi(\beta, \gamma) = \log \mathrm{Tr}\ \mathrm{e}^{\beta H + \gamma K}. \tag{2.4}$$

Having in mind (1.3) and (2.1), we get $\Psi(\beta, \gamma) \geq \Xi(\beta, \gamma, \rho)$, for any $(\beta, \gamma) \in \mathbb{R}^2$.

Theorem 2.2. *For ρ_0 given by (2.2), H, K Hermitian commuting matrices and $\Psi : \mathbb{R}^2 \to \mathbb{R}$ defined in (2.4), the following holds*

$$\mathrm{Tr}\rho_0 H = \frac{\partial \Psi}{\partial \beta},\ \mathrm{Tr}\rho_0 K = \frac{\partial \Psi}{\partial \gamma}.$$

Proof. From $[H, K] = 0$ it follows that

$$\mathrm{e}^{\beta H + \gamma K} = \mathrm{e}^{\beta H} \mathrm{e}^{\gamma K}.$$

Thus

$$\frac{\partial(\mathrm{e}^{\beta H + \gamma K})}{\partial \beta} = H \mathrm{e}^{\beta H + \gamma K}, \quad \frac{\partial(\mathrm{e}^{\beta H + \gamma K})}{\partial \gamma} = K \mathrm{e}^{\beta H + \gamma K},$$

so that

$$\frac{\partial \Psi}{\partial \beta} = \frac{\mathrm{Tr} H \mathrm{e}^{\beta H + \gamma K}}{\mathrm{Tr} \mathrm{e}^{\beta H + \gamma K}},\ \frac{\partial \Psi}{\partial \gamma} = \frac{\mathrm{Tr} K \mathrm{e}^{\beta H + \gamma K}}{\mathrm{Tr} \mathrm{e}^{\beta H + \gamma K}}.$$

The result follows trivially. □

Theorem 2.3. *For ρ_0 defined in (2.2) and H, K Hermitian matrices, then $\mathrm{Tr}\rho_0(H + iK)$ is a boundary point of $W_{\rho_0}(H + iK)$.*

Proof. Let us consider the Hermitian matrix

$$\Re(\mathrm{e}^{-i\theta}(H + iK)) = \cos\theta H + \sin\theta K,\ \ \theta \in \mathbb{R}.$$

We recall that that $W_{\rho_0}(\cos\theta H + \sin\theta K)$ gives the projection of the convex set $W_{\rho_0}(H + iK)$ in the direction θ. Consider now the direction θ_0 such that

$$\cos\theta_0 = \frac{\beta}{\sqrt{\beta^2 + \gamma^2}},\ \sin\theta_0 = \frac{\gamma}{\sqrt{\beta^2 + \gamma^2}}.$$

Let $\lambda_1(\theta_0) \geq \ldots \geq \lambda_n(\theta_0)$, denote the eigenvalues of $\cos\theta_0 H + \sin\theta_0 K$ and let $\beta_0 = \sqrt{\beta^2+\gamma^2}$. The eigenvalues of $\beta H + \gamma K = \beta_0(\cos\theta_0 H + \sin\theta_0 K)$ are

$$\beta_0\lambda_1(\theta_0) \geq \ldots \geq \beta_0\lambda_n(\theta_0).$$

From (2.2), it follows that the eigenvalues of ρ_0 are

$$\chi\exp(\beta_0\lambda_1(\theta_0)) \geq \ldots \geq \chi\exp(\beta_0\lambda_n(\theta_0)),$$

where $\chi = (\mathrm{e}^{\beta_0\lambda_1(\theta_0)} + \ldots + \mathrm{e}^{\beta_0\lambda_n(\theta_0)})^{-1}$. The projection of $W_{\rho_0}(H+iK)$ in the direction θ_0 is the line segment with endpoints

$$\chi(\lambda_1(\theta_0)\exp(\beta_0\lambda_1(\theta_0)) + \ldots + \lambda_n(\theta_0)\exp(\beta_0\lambda_n(\theta_0)))$$

and

$$\chi(\lambda_n(\theta_0)\exp(\beta_0\lambda_1(\theta_0)) + \ldots + \lambda_1(\theta_0)\exp(\beta_0\lambda_n(\theta_0))).$$

Notice that the eigenvalues of ρ_0 and $\cos\theta_0 H + \sin\theta_0 K$ are ordered in exactly the same manner if $-\pi/2 \leq \theta_0 \leq \pi/2$ and in the opposite way if $\pi/2 \leq \theta_0 \leq 3\pi/2$, because the exponential is an increasing function of its argument. Since

$$\lambda_1(\theta_0)\exp(\beta_0\lambda_1(\theta_0)) + \ldots + \lambda_n(\theta_0)\exp(\beta_0\lambda_n(\theta_0))$$

is an endpoint of the projection of $W_{\rho_0}(H+iK)$ in the direction θ_0, it is obvious that $\mathrm{Tr}\rho_0(H+iK)$ is a boundary point of $W_{\rho_0}(H+iK)$, as it lies on the supporting line of $W_{\rho_0}(H+iK)$ perpendicular to the direction θ_0. □

In the case of commuting H and K, the following holds.

Theorem 2.4. *For ρ_0 given by (2.2), and if H, K commute, then $\mathrm{Tr}\rho_0(H+iK)$ is a σ-point on the boundary of $W_{\rho_0}(H+iK)$.*

Proof. We clearly have

$$[\rho_0, H] = [\rho_0, K] = [H, K] = 0.$$

Therefore, we may assume that the Hermitian matrices ρ_0, H and K are in diagonal form. Thus, the result follows. □

3. The maximum entropy inference method

In Theorem 2.1 we have proved that

$$\rho_0(\beta,\gamma) = \frac{\exp(\beta H + \gamma K)}{\mathrm{Tr}\exp(\beta H + \gamma K)}$$

gives the equilibrium mixed state, for the values of the parameters β, γ that produce x, y according to (1.1). Substituting in (1.1) ρ by ρ_0 given by (2.2), we conclude that

$$\frac{\mathrm{Tr}H\mathrm{e}^{\beta H+\gamma K}}{\mathrm{Tr}\mathrm{e}^{\beta H+\gamma K}} = x(\beta,\gamma), \quad \frac{\mathrm{Tr}K\mathrm{e}^{\beta H+\gamma K}}{\mathrm{Tr}\mathrm{e}^{\beta H+\gamma K}} = y(\beta,\gamma). \tag{3.1}$$

To solve the maximum entropy inference problem, we have to determine (β,γ), from the knowledge of $x(\beta,\gamma)$ and $y(\beta,\gamma)$, that is, we determine the pre-image (β,γ) of the function $f : \mathbb{R}^2 \to \mathbb{R}^2$, such that $(\beta,\gamma) \to (x(\beta,\gamma), y(\beta,\gamma))$. This allows the restorement of the maximizing matrix in (2.2), as illustrated in the next example.

Theorem 2.2 states the relation between the parameters β, γ and the constraints on $\mathrm{Tr}H\rho$ and $\mathrm{Tr}K\rho$. We may observe that the set of points $(x(\beta,\gamma), y(\beta,\gamma))$ in (3.1), associated with the Lagrange multipliers β, γ, for $\beta = \beta_0 \cos\theta,\ \gamma = \beta_0 \sin\theta,\ 0 \leq \theta \leq 2\pi$, and fixed β_0, tends to $\partial W(H+iK)$ when $\beta_0 \to +\infty$. It is instructive to investigate the maximum entropy inference problem for $\beta_0(\cos\theta H + \sin\theta K)$, for fixed β_0 and arbitrary θ, and for fixed θ and arbitrary β_0. This is done in the next Example. The procedure is valid for any finite number of dimensions and may be easily implemented.

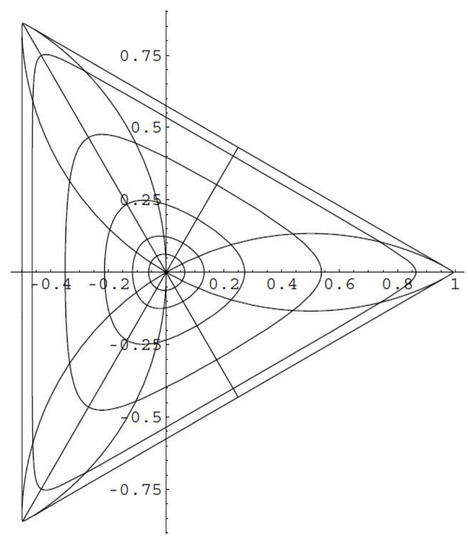

FIGURE 1. The curves described by $x(\beta,\gamma), y(\beta,\gamma)$ with $\beta = \beta_0 \cos\theta,\ \gamma = \beta_0 \sin\theta$, for $\beta_0 = 0.125,\ 0.25,\ 0.5,\ 1,\ 2, +\infty$, with $0 \leq \theta < 2\pi$, and when $\theta = n\pi/6,\ n = 1, 2, \ldots, 12$, with $0 \leq \beta_0 < +\infty$. The horizontal and the vertical axes, represent, respectively, $x(\beta,\gamma)$ and $y(\beta,\gamma)$.

Example 3.1. Let us consider the observables

$$H = \frac{1}{2}\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad K = \frac{i}{2}\begin{bmatrix} 0 & -i & i \\ i & 0 & -i \\ -i & i & 0 \end{bmatrix}$$

and

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad A = H + iK.$$

It may be seen that

$$\sigma(\beta H + \gamma K) = \{\beta, (-\beta - \sqrt{3}\gamma)/2, (-\beta + \sqrt{3}\gamma)/2\},\ \beta, \gamma \in \mathbb{R},$$

and so $W(A)$ is a regular triangle having one vertical side. From (2.4) it follows that

$$\Psi(\beta,\gamma) = \log\left(e^{\beta} + e^{(-\beta-\sqrt{3}\gamma)/2} + e^{(-\beta+\sqrt{3}\gamma)/2}\right),$$

and so

$$\frac{\partial\Psi}{\partial\beta} = \frac{2e^{\beta} - e^{(-\beta-\sqrt{3}\gamma)/2} - e^{(-\beta+\sqrt{3}\gamma)/2}}{2\left(e^{\beta} + e^{(-\beta-\sqrt{3}\gamma)/2} + e^{(-\beta+\sqrt{3}\gamma)/2}\right)} = x(\beta,\gamma),$$

$$\frac{\partial\Psi}{\partial\gamma} = \frac{(-\sqrt{3}e^{(-\beta-\sqrt{3}\gamma)/2} + \sqrt{3}e^{(-\beta+\sqrt{3}\gamma)/2}}{2\left(e^{\beta} + e^{(-\beta-\sqrt{3}\gamma)/2} + e^{(-\beta+\sqrt{3}\gamma)/2}\right)} = y(\beta,\gamma). \quad (3.2)$$

Next, we consider $x(\beta,\gamma),\ y(\beta,\gamma)$ in (3.2), for

$$\beta = \beta_0\cos\theta,\ \gamma = \beta_0\sin\theta.$$

Fixing β_0 and varying θ we obtain a closed curve surrounding the origin. Fixing θ

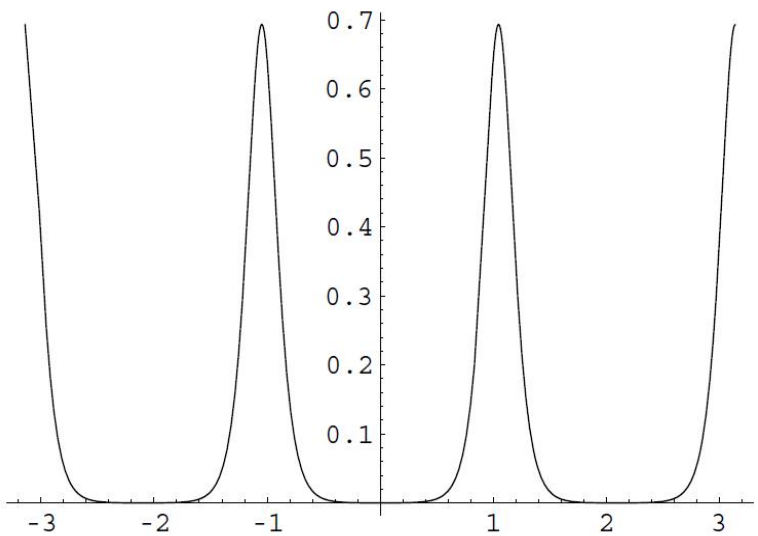

FIGURE 2. The maximum entropy and the angle θ are presented, respectively, on the vertical and horizontal axis, for $\beta = 8\cos\theta,\ \gamma = 8\sin\theta$. Notice the pronounced peaks for $\theta = \pm\pi/3, \pi$. See Example 3.1

and varying β_0, we obtain curves connecting the origin with corners of $W(H+iK)$. These curves are displayed in Fig. 1 for $\beta_0 = 0.125,\ 0.25,\ 0.5,\ 1,\ 2$, in the limit $\beta_0 \to +\infty$, and for $\theta = n\pi/6,\ n = 1, 2, \ldots, 12$. In the limit $\beta_0 \to +\infty$, $\partial W(H + iK))$ is obtained i.e., the limit of the solution ρ_0 corresponds for any θ to a pure state, with entropy $S = 0$, meaning that ρ_0 is unitarily similar to $\mathrm{diag}(1, 0, 0)$, except at the angles $\theta = \pm\pi/3, \pi$, where $S = \log 2$, so that then ρ_0 is unitarily similar to $\mathrm{diag}(1/2, 1/2, 0)$. From the matrix ρ_0, and having in mind (1.2) and (2.4), the values of the entropy and of the generalized free energy follow. In Fig 2, the entropy S is represented for $\beta_0 = 8$. In fig 3, the generalized free energy times β is represented for $\beta_0 = 8$.

Notice the peculiar behavior for $\theta = \pm\pi/3, \pi$, where the entropy attains its maximum value, always greater than $\log 2$, and the generalized free energy multiplied by β attains its minimum value. On the other hand, at the angles $\pm 2\pi/3, 0$, the maximum entropy attains its minimum value and the free energy multiplied by β attains its maximum value. If β_0 is large enough, the entropy is very close to 0 for almost any $0 \le \theta \le 2\pi$, except for $\theta = \pi, \pm\pi/3$, where it is very close to $\log 2$ suggesting a discontinuous behavior of density matrix. The maximum of the entropy becomes increasingly sharper, and the minimum increasingly broader as β_0 increases. We notice that the limit of ρ_0 as $\beta_0 \to +\infty$ is diag$(0,1,0)$ and diag$(0,0,1)$, diag$(1,0,0)$, respectively for $-\pi/3 < \theta < \pi/3$, $-\pi < \theta < -\pi/3$, $\pi/3 < \theta < \pi$, while it becomes diag$(1/2,0,1/2)$, diag$(1/2,1/2,0)$ and diag$(0,1/2,1/2)$, since Tr$\rho_0 = 1$, respectively for $\theta = \pi/3$, $\theta = -\pi/3$, $\theta = \pi$.

More specifically, we find

$$\lim_{\beta_0 \to +\infty} \rho_0(\beta_0, \theta) = \begin{cases} \text{diag}(1,0,0) \text{ for } -\pi/3 < \theta < \pi/3 \\ \text{diag}(0,1,0) \text{ for } \pi/3 < \theta < \pi \\ \text{diag}(0,0,1) \text{ for } -\pi < \theta < -\pi/3 \end{cases}$$

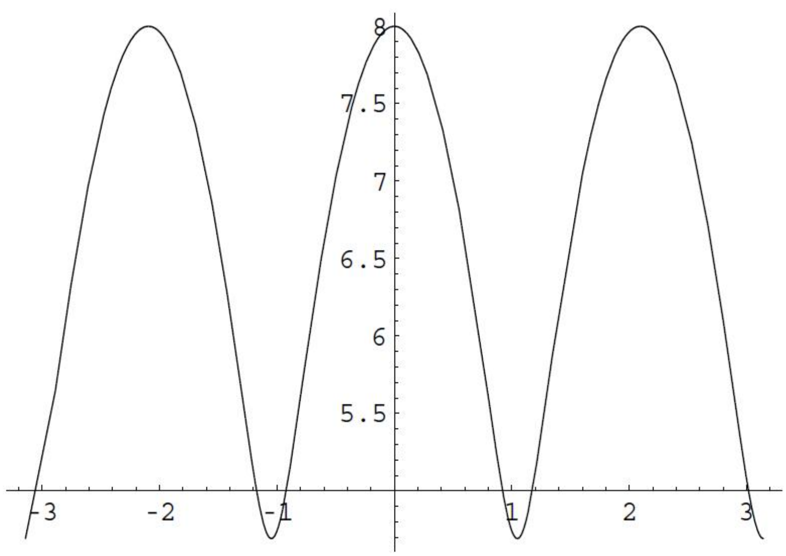

FIGURE 3. The generalized free energy multiplied by β, i.e. Ψ, and the angle θ are represented, respectively, on the vertical and horizontal axis, for $\beta = 8\cos\theta$, $\gamma = 8\sin\theta$, Notice the peculiar behavior for $\theta = \pm 2\pi/3, 0$, and for $\theta = \pm\pi/3, \pi$.

and

$$\lim_{\beta_0 \to +\infty} \rho_0(\beta_0, \theta) = \begin{cases} \frac{1}{2}\text{diag}(1,1,0) \text{ for } \theta = \pi/3 \\ \frac{1}{2}\text{diag}(0,1,1) \text{ for } \theta = \pi \\ \frac{1}{2}\text{diag}(1,0,1) \text{ for } \theta = -\pi/3, \end{cases}$$

since Tr$\rho_0 = 1$, so that the limit of ρ_0 is a pure state except for $\theta = -\pi/3, \pi/3, \pi$, that is, when θ is the angle of the perpendicular direction to one of the line segments in the boundary of $W(H + iK)$. At these points, an effective discontinuity of the limit of the density matrix arises.

Acknowledgments

The authors are most grateful to the Referee for the careful reading of the manuscript.

This work was partially supported by the Centro de Matemática da Universidade de Coimbra (CMUC), funded by the European Regional Development Fund through the program COMPETE and by the Portuguese Government through the FCT - Fundação para a Ciência e a Tecnologia under the project PEst-C/MAT/UI0324/2011.

References

1. N. Bebiano, *Nondifferentiable points of* $\partial W_c(A)$, Linear Algebra Appl. **19** (1986), 249–257.
2. N. Bebiano, R. Lemos, and J. Providência, *Matrix inequalities in statistical mechanics*, Linear Algebra Appl. **376** (2004), 265–273.
3. N. Bebiano, R. Lemos, and J. Providência, *Inequalities for quantum relative entropy*, Linear Algebra Appl. **401** (2005), 159–172.
4. J. Chen, Z. Ji, C.-K. Li, Y.-T. Poon, Y. Shen, N. Yu, B. Zeng, and D. Zhou, *Discontinuity of maximum entropy inference and quantum phase transitions*, New J. Phys. **17** (2015), 083019, 18pp.
5. F. Hausdorff, *Der Wertworrat einer Bilinearform*, Math. Z. **3** (1919), 314–316.
6. E. T. Jaynes, *Information theory and statistical mechanics*, Phys. Rev. **106** (1957), 620–630; **208** (1957) 171–190.
7. M. Joswig and M. Straub, *On the numerical range map*, J. Austral. Math. Soc. Ser. A **65** (1998), no. 2, 267–283.
8. M. Marcus, *Derivations, Plücker Relations and the Numerical Range*, Indiana Univ. Math. J. **22** (1972/73), 1137–1149.
9. L. Rodman, I. M. Spitkovsky, A. Szkola, and S. Weis, *Continuity of the maximum-entropy inference: Convex geometry and numerical ranges approach*, J. Math. Phys. **57** (2016), no.1, 015204, 17pp.
10. I. M. Spitkovsky and S. Weis, *Pre-images of extreme points of the numerical range, and applications*, arXiv:1509.05676.
11. O. Toeplitz, *Das Algebraich Analogon zu einem Satz von Fejér*, Math. Z. **2** (1918), 187–197.
12. J. von Neumann, *Thermodynamik quantenmechanischer Gesamtheiten*, Nachr. Ges. Wiss. Göttingen **1927** (1927), 273–291.
13. T. Westwick, *A theorem on numerical range*, Linear Multiliner Algebra, **2** (1957), 311–315.

[1] University of Coimbra, CMUC, Department of Mathematics, P 3001-454 Coimbra, Portugal.

E-mail address: bebiano@mat.uc.pt

[2] Univerity of Coimbra, CFisUC, Department of Physics, P 3004-516 Coimbra, Portugal.

E-mail address: providencia@teor.fis.uc.pt