

QUASI-ENTROPIES FOR FINITE QUANTUM SYSTEMS

DÉNES PETZ

Mathematical Institut HAS, Budapest, Hungary

(Received September 24, 1984)

Convexity properties of entropy-like functionals on states of a finite dimensional algebra are discussed. The treatment covers both the quantum mechanical and the classical cases. The purpose is to generalize Lieb's convexity theorem and the monotonicity of the relative entropy using the Jensen inequality of operator convex functions. From the quasi-entropies defined here the quantum version of Rényi's α -entropies can be deduced.

Introduction

In this paper we discuss the convexity properties of entropy-like functionals. Our treatment covers both the quantum mechanical and the classical cases, however, the main emphasis is on quantum mechanics. Mathematically, the objects investigated are positive linear functionals on finite dimensional operator algebras. In the commutative case we obtain simple theorems in probability and the matrix case is related to quantum mechanics.

When John von Neumann introduced the entropy

$$S(\varrho) = -\text{Tr } \varrho \log \varrho$$

of a density matrix ϱ [18] it turned out immediately that important properties (for example, concavity and increasing under measurement) are due to the concavity of the function $f(t) = -t \log t$. The functional $S_f(\varrho) = \text{Tr } f(\varrho)$ has been widely studied [1, 6, 16, 19, 27] and the expression "quasi-entropy" was proposed in [27]. In this paper we carry out an analogous generalization in case of the relative entropy. (Of course, entropy as the relative entropy with respect to a trace remains a particular case.) The Kullback–Leibler information for discrimination appeared in the operator algebra setting in the pioneering work of Umegaki [25], and Lindblad [15, 16] discussed properties of this quantity (called already relative entropy)

$$S(\varrho_1, \varrho_2) = \text{Tr } \varrho_2 (\log \varrho_2 - \log \varrho_1).$$

The culminating point was surely Lieb's convexity theorem (sometimes referred as Wigner–Yanase–Dyson–Lieb convexity) which easily implies the concavity of the relative entropy ([13], see also [8, 20, 23, 26]).

The forerunner of the present paper in classical information theory is [4] which serves as a guideline for us. (We note that in [4] and [5] the term f -divergence is used instead of our quasi-entropy.) We shall define quasi-entropies

$$S_f^k(\varphi, \omega)$$

for positive functionals φ and ω . Here f is a function on \mathbf{R}^+ and k is a fixed operator in a given algebra. When f is operator convex, $S_f^k(\varphi, \omega)$ yields the convexity and monotonicity properties of the usual relative entropy corresponding to the choice $k = 1$ and $f(t) = -\log t$. The properties above will be proved by simple application of the Jensen operator inequality. Both the integral representation of operator convex functions and interpolation theories are avoided (cf. [3, 11, 24]). The quantum version of Rényi's α -entropies [21, 22, 26] can be deduced from quasi-entropies. Quasi-entropies are not additive and subadditive, hence their physical importance is restricted. We do not touch here on the infinite dimensional case but a treatment in a von Neumann algebra context will follow in a subsequent paper.

The author is grateful to I. Csiszár and J. Fritz for useful conversations and it is a pleasure to thank E. Lieb and A. Wehrl for comments on the first draft of the manuscript.

1. Definition

Let \mathcal{A} be a finite dimensional C^* -algebra with a faithful trace τ . In other words

- (i) $\tau \in \mathcal{A}_+^*$,
- (ii) $\tau(ab) = \tau(ba)$, ($a, b \in \mathcal{A}$),
- (iii) $a \in \mathcal{A}_+$ and $\tau(a) = 0$ imply $a = 0$.

Such a τ always exists but it is not unique. Let φ and ω be positive functionals on \mathcal{A} . Then there are density operators ϱ_φ and ϱ_ω such that $\varphi(\cdot) = \tau(\varrho_\varphi \cdot)$ and $\omega(\cdot) = \tau(\varrho_\omega \cdot)$. For technical simplicity we assume that ω is faithful, i.e. ϱ_ω is invertible. With the trace inner product

$$\langle a, b \rangle_\tau = \tau(b^* a).$$

\mathcal{A} becomes a Hilbert space \mathcal{H} and $a \mapsto L_a$ is a representation of \mathcal{A} on \mathcal{H} ($L_a b = ab$). We identify $L(\mathcal{A})$ with \mathcal{A} . Thus φ and ω arise as follows:

$$\varphi(a) = \langle a\phi, \phi \rangle, \quad \omega(a) = \langle a\Omega, \Omega \rangle,$$

where $\phi = \varrho_\phi^{1/2}$ and $\Omega = \varrho_\omega^{1/2}$. Here Ω is a cyclic and separating vector for \mathcal{A} . There exists a unique positive operator $\Delta_{\phi, \Omega}$ such that

$$\langle \Delta_{\phi, \Omega} a \Omega, a \Omega \rangle = \varphi(a a^*).$$

It is easy to check that $\Delta_{\phi, \Omega} = L_{\varrho_\phi} R_{\varrho_\omega}^{-1}$, where $R_{\varrho_\omega} a = a \varrho_\omega$.

To define the quasi-entropies we choose a continuous function $f: (0, \infty) \rightarrow \mathbb{R}$ and an operator $k \in \mathcal{A}$. We set

$$S_f^k(\varphi, \omega) = \langle f(\Delta_{\phi, \Omega}) k \Omega, k \Omega \rangle_\tau.$$

This quantity does not depend on the trace τ . If τ' is another functional satisfying conditions (i)–(iii) then there is an invertible element c in the centre of \mathcal{A} such that $\tau'(a) = \tau(ca)$ ($a \in \mathcal{A}$). Then $\varrho'_\phi = c^{-1} \varrho_\phi$ and $\varrho'_\omega = c^{-1} \varrho_\omega$. So $\Delta_{\phi', \Omega'} = \Delta_{\phi, \Omega}$ and we find that S_f^k is independent of τ .

In the really interesting cases f will be an operator concave function or operator monotone function. Let us consider some examples.

$$1. f(t) = -\log t$$

$$S_f^k(\varphi, \omega) = \tau(\varrho_\omega^{1/2} k^* k \varrho_\omega^{1/2} \log \varrho_\omega - k \varrho_\omega k^* \log \varrho_\phi).$$

$$2. f(t) = t^\alpha$$

$$S_f^k(\varphi, \omega) = \tau(\varrho_\omega^{1-\alpha} k^* \varrho_\phi^\alpha k).$$

$$3. f(t) = t \log t$$

$$S_f^k(\varphi, \omega) = \tau(\varrho_\phi \log \varrho_\phi k k^* - \varrho_\phi k \log \varrho_\omega k^*).$$

$$4. f(t) = ut + v$$

$$S_f^k(\varphi, \omega) = u\varphi(k k^*) + v\omega(k^* k).$$

One can recognize immediately that Lieb's convexity theorem is related to Example 2. Examples 1 and 3 give the usual relative entropy (up to order of the functionals in the case $f(t) = t \log t$). The linear case is quite uninteresting. Following [21] and [22] we can define the α -entropies:

$$S_\alpha(\varphi, \omega) = \frac{1}{\alpha - 1} \log \tau(\varrho_\phi^{1-\alpha} \varrho_\omega^\alpha), \quad 0 < \alpha < 1.$$

(Compare with [26] where $\omega = \tau = \text{Tr}$.) Now

$$S_1(\varphi, \omega) = \lim_{\alpha \rightarrow 1} S_\alpha(\varphi, \omega) = \tau(\varrho_\omega [\log \varrho_\omega - \log \varrho_\phi])$$

by the same calculation as in the commutative case.

We note that if $f_1(t) = f(t) + c$, then

$$S_{f_1}^k(\varphi, \omega) = S_f^k(\varphi, \omega) + c\omega(k^* k).$$

Hence it is not an essential restriction if we fix $f(0)$.

When $k = 1$ we write simply $S_f(\varphi, \omega)$.

2. Properties

In this section we establish the concavity and the monotonicity of the quantity S_f^k . Lieb's convexity theorem will be a consequence. Our main tool is the Jensen inequality of operator concave functions. We recall some definitions and basic facts. A function $f: (0, \infty) \rightarrow \mathbf{R}$ is called operator monotone if $f(A) \leq f(B)$ provided that A and B are positive definite matrices and $A \leq B$. The function f is operator concave if for positive definite matrices A, B we have $f(\lambda A + (1-\lambda)B) \geq \lambda f(A) + (1-\lambda)f(B)$ for $0 < \lambda < 1$. f is operator convex if $-f$ is operator concave. Every operator monotone function (on the interval $(0, \infty)$) is operator concave (Theorem III.3 in [2]) and every operator concave function is C^* -concave in the following sense. If $\sum_{i=1}^n C_i^* C_i = 1$, then

$$f\left(\sum_{i=1}^n C_i^* A_i C_i\right) \geq \sum_{i=1}^n C_i^* f(A_i) C_i.$$

(Here A_1, A_2, \dots, A_n are positive definite matrices and C_1, C_2, \dots, C_n are arbitrary.) We shall mostly need the consequence of the inequality above. Namely, if $f(0) \geq 0$ and $\|v\| \leq 1$, then

$$f(vAv^*) \geq vf(A)v^*.$$

For notions related to operator inequalities we refer to [2], [7] and [9]. We note that a non-linear operator concave function is strictly concave ($f'' < 0$).

THEOREM 1. *Assume that f is operator convex. Let \mathcal{A}_0 be a subalgebra of the finite dimensional algebra \mathcal{A} and let φ, ω be positive functionals on \mathcal{A} . If ω is faithful and $k \in \mathcal{A}_0$, then*

$$S_f^k(\varphi_0, \omega_0) \leq S_f^k(\varphi, \omega),$$

where φ_0, ω_0 denote the restrictions of φ and ω to \mathcal{A}_0 .

Proof: We fix faithful traces τ and τ_0 on \mathcal{A} and \mathcal{A}_0 , respectively. \mathcal{H} and \mathcal{H}_0 denote the corresponding inner product spaces. With obvious notations

$$\varphi(a) = \langle a\phi, \phi \rangle, \quad \varphi_0(a_0) = \langle a_0\phi_0, \phi_0 \rangle,$$

$$\omega(a) = \langle a\Omega, \Omega \rangle, \quad \omega_0(a_0) = \langle a_0\Omega_0, \Omega_0 \rangle.$$

The linear operator $v: a_0\Omega_0 \rightarrow a_0\Omega$ is an isometry. We check that

$$\Delta_{\phi_0, \Omega_0} = v^* \Delta_{\phi, \Omega} v.$$

Indeed,

$$\langle \Delta_{\phi_0, \Omega_0} a_0 \Omega_0, a_0 \Omega_0 \rangle = \varphi_0(a_0 a_0^*) = \langle \Delta_{\phi, \Omega} a_0 \Omega, a_0 \Omega \rangle = \langle v^* \Delta_{\phi, \Omega} v a_0 \Omega_0, a_0 \Omega_0 \rangle.$$

By the Jensen inequality

$$f(\Delta_{\phi_0, \Omega_0}) \leq v^* f(\Delta_{\phi, \Omega}) v$$

and in particular we have

$$\langle f(\Delta_{\phi_0, \Omega_0}) k \Omega_0, k \Omega_0 \rangle \leq \langle f(\Delta_{\phi, \Omega}) k \Omega, k \Omega \rangle.$$

COROLLARY 2. *If f is operator convex, then*

$$S_f(\varphi, \omega) \geq f(\varphi(1)/\omega(1))\omega(1).$$

Proof: By the monotonicity (Theorem 1) $S_f(\varphi, \omega)$ is not smaller than the quasi-entropy of the restrictions to $C \cdot 1$.

We can formulate a stronger result using an estimate in [4].

THEOREM 3. *Let φ and ω be faithful states and f a non-linear operator convex function. Then*

$$S_f(\varphi, \omega) \geq f(1)$$

and equality holds if and only if $\varphi = \omega$. More precisely, there exists a $C > 0$ such that if δ is small enough, then

$$S_f(\varphi, \omega) - f(1) < \delta$$

implies

$$\|\varphi - \omega\| \leq C\delta^{1/2}.$$

Proof: Let $\psi_1 - \psi_2$ be the Jordan decomposition of $\varphi - \omega$ and e the support of ψ_1 . If \mathcal{A}_0 is the subalgebra generated by e and φ_0, ω_0 are the restrictions of φ, ω to \mathcal{A}_0 , then

$$\|\varphi - \omega\| = \|\varphi_0 - \omega_0\|$$

and by the monotonicity

$$S_f(\varphi_0, \omega_0) \leq S_f(\varphi, \omega).$$

Since $f''(1) > 0$ we can apply 2.1. Theorem of [4] (see also (4.1) in [5]) to φ_0 and ω_0 and it completes the proof.

It was proved in [10] by the same method that

$$\|\varphi - \omega\|^2 \leq 2S(\varphi, \omega).$$

Now we are going to obtain a more general monotonicity related to

Uhlmann's theorem [24]. Instead of interpolation we use the Jensen inequality in adequate form. We recall that a linear map $\alpha: \mathcal{A}_0 \rightarrow \mathcal{A}$ is a Schwarz map if it satisfies the inequality $\alpha(a_0)^* \alpha(a_0) \leq \alpha(a_0^* a_0)$.

THEOREM 4. *Let \mathcal{A}_0 and \mathcal{A} be finite dimensional algebras and $\alpha: \mathcal{A}_0 \rightarrow \mathcal{A}$ a Schwarz map. Assume that φ_0, ω_0 and φ, ω are faithful positive functionals on \mathcal{A}_0 and \mathcal{A} , respectively, such that*

$$\omega \circ \alpha \leq \omega_0, \quad \varphi \circ \alpha \leq \varphi_0.$$

Then for every operator monotone function f with $f(0) = 0$ and for any $k \in \mathcal{A}_0$ we have

$$S_f^k(\varphi_0, \omega_0) \geq S_f^{\alpha(k)}(\varphi, \omega).$$

Proof: \mathcal{A}_0 and \mathcal{A} become Hilbert spaces with trace inner products. So $\varphi_0, \omega_0, \varphi, \omega$ can be represented by means of vectors as follows:

$$\begin{aligned} \varphi(a) &= \langle a\phi, \phi \rangle, & \varphi_0(a_0) &= \langle a_0\phi_0, \phi_0 \rangle, \\ \omega(a) &= \langle a\Omega, \Omega \rangle, & \omega_0(a_0) &= \langle a_0\Omega_0, \Omega_0 \rangle. \end{aligned}$$

We set $v: \mathcal{H}_0 \rightarrow \mathcal{H}$ by the formula

$$va_0\Omega_0 = \alpha(a_0)\Omega$$

and obtain a contraction. We show that

$$v^* \Delta_{\phi, \Omega} v \leq \Delta_{\phi_0, \Omega_0}$$

by a simple calculation

$$\begin{aligned} \langle v^* \Delta_{\phi, \Omega} va_0\Omega_0, a_0\Omega_0 \rangle &= \langle \alpha(a_0)^* \phi, \alpha(a_0)^* \phi \rangle = \varphi(\alpha(a_0)\alpha(a_0)^*) \\ &\leq \varphi_0(a_0 a_0^*) = \langle \Delta_{\phi_0, \Omega_0} a_0\Omega_0, a_0\Omega_0 \rangle. \end{aligned}$$

Now an invocation to the Jensen inequality completes the proof

$$\begin{aligned} S_f^k(\varphi_0, \omega_0) &= \langle f(\Delta_{\phi_0, \Omega_0}) k\Omega_0, k\Omega_0 \rangle \geq \langle f(v^* \Delta_{\phi, \Omega} v) k\Omega_0, k\Omega_0 \rangle \\ &\geq \langle v^* f(\Delta_{\phi, \Omega}) vk\Omega_0, k\Omega_0 \rangle = S_f^{\alpha(k)}(\varphi, \omega). \end{aligned}$$

The joint convexity is based on the monotonicity of the quasi-entropy and the following construction.

LEMMA 5. *Let $\varphi_1, \varphi_2, \omega_1, \omega_2$ be faithful positive functionals on \mathcal{A} and $\mathcal{B} = \mathcal{A} \oplus \mathcal{A}$. We consider the functionals*

$$\begin{aligned} \varphi_{12}(a_1 \oplus a_2) &= \lambda \varphi_1(a_1) + (1-\lambda) \varphi_2(a_2), \\ \omega_{12}(a_1 \oplus a_2) &= \lambda \omega_1(a_1) + (1-\lambda) \omega_2(a_2) \end{aligned}$$

on \mathcal{B} when $0 < \lambda < 1$. Then

$$S_f^{k \oplus k}(\varphi_{12}, \omega_{12}) = \lambda S_f^k(\varphi_1, \omega_1) + (1 - \lambda) S_f^k(\varphi_2, \omega_2).$$

Proof: If τ is a faithful trace on \mathcal{A} , then $\tau_{12} = \tau \oplus \tau$ is a faithful trace on \mathcal{B} . With our usual notations we have

$$\mathcal{H}_{12} = \mathcal{H} \oplus \mathcal{H}, \quad \Omega_{12} = \sqrt{\lambda} \Omega_1 \oplus \sqrt{1 - \lambda} \Omega_2, \quad \phi_{12} = \sqrt{\lambda} \phi_1 \oplus \sqrt{1 - \lambda} \phi_2.$$

Hence $\Delta_{\phi_{12}, \Omega_{12}} = \Delta_{\phi_1, \Omega_1} \oplus \Delta_{\phi_2, \Omega_2}$ is independent of λ . Now simple calculation gives the statement

$$\begin{aligned} S_f^{k \oplus k}(\varphi_{12}, \omega_{12}) &= \langle f(\Delta_{\phi_{12}, \Omega_{12}})(k \oplus k)(\sqrt{\lambda} \Omega_1 \oplus \sqrt{1 - \lambda} \Omega_2), (k \oplus k)(\sqrt{\lambda} \Omega_1 \oplus \sqrt{1 - \lambda} \Omega_2) \rangle \\ &= \lambda \langle f(\Delta_{\phi_1, \Omega_1}) k \Omega_1, k \Omega_1 \rangle + (1 - \lambda) \langle f(\Delta_{\phi_2, \Omega_2}) k \Omega_2, k \Omega_2 \rangle. \end{aligned}$$

THEOREM 6. *Let f be operator-convex. Then for any $k \in \mathcal{A}$*

$$S_f^k(\varphi, \omega)$$

is a jointly convex function of φ and ω .

Proof: We are going to use the construction and the notation of the previous lemma. We have

$$S_f^{k \oplus k}(\varphi_{12}, \omega_{12}) = \lambda S_f^k(\varphi_1, \omega_1) + (1 - \lambda) S_f^k(\varphi_2, \omega_2).$$

$\mathcal{A}_0 = \{a \oplus a : a \in \mathcal{A}\}$ is a subalgebra of \mathcal{B} . Set $\varphi_0 = \varphi_{12}|_{\mathcal{A}_0}$ and $\omega_0 = \omega_{12}|_{\mathcal{A}_0}$. Evidently, \mathcal{A}_0 is isomorphic to \mathcal{A} and the isomorphism carries φ_0 into $\lambda \varphi_1 + (1 - \lambda) \varphi_2$ and ω_0 into $\lambda \omega_1 + (1 - \lambda) \omega_2$. Therefore

$$S_f^{k \oplus k}(\varphi_0, \omega_0) = S_f^k(\lambda \varphi_1 + (1 - \lambda) \varphi_2, \lambda \omega_1 + (1 - \lambda) \omega_2)$$

and an appeal to Theorem 1 completes the proof.

An alternative version of Theorem 6 is the following, where f is not only operator concave but even operator monotone (compare with [11]).

THEOREM 7. *Assume that f is operator-monotone and $f(0) = 0$. Let $\varphi_1, \varphi_2, \omega_1, \omega_2, \varphi, \omega$ be faithful positive functionals on the finite dimensional algebra \mathcal{A} and let $k \in \mathcal{A}$, $0 < \lambda < 1$.*

If

$$\lambda \varphi_1 + (1 - \lambda) \varphi_2 \leq \varphi, \quad \lambda \omega_1 + (1 - \lambda) \omega_2 \leq \omega,$$

then

$$\lambda S_f^k(\varphi_1, \omega_1) + (1 - \lambda) S_f^k(\varphi_2, \omega_2) \leq S_f^k(\varphi, \omega).$$

Proof: The proof is essentially the same as that of Theorem 6. The only difference that at the last point one has to invoke to Theorem 4 instead of Theorem 1.

3. Discussion

We have seen that convexity properties of the quasi-entropies are simple consequences of the Jensen operator inequality. Under fairly general conditions on f one can not expect nice additivity and subadditivity properties. However, the α -entropies are additive. Let φ_1, ω_1 and φ_2, ω_2 be faithful states on \mathcal{A}_1 and \mathcal{A}_2 . Then

$$\varrho_{\varphi_1 \otimes \varphi_2}^{1-\alpha} \varrho_{\omega_1 \otimes \omega_2}^{\alpha} = \varrho_{\varphi_1}^{1-\alpha} \varrho_{\omega_1}^{\alpha} \otimes \varrho_{\varphi_2}^{1-\alpha} \varrho_{\omega_2}^{\alpha}$$

and

$$S_{\alpha}(\varphi_1 \otimes \varphi_2, \omega_1 \otimes \omega_2) = S_{\alpha}(\varphi_1, \omega_1) + S_{\alpha}(\varphi_2, \omega_2)$$

follows.

Despite the fact that we always treated faithful functionals, the results can be extended to arbitrary ones by means of continuity arguments. When non-invertible densities may occur it is convenient to use continuous functions f on $[0, \infty)$. Also the usual relative entropy is available in this way through α -entropies or by means of the function $f(t) = t \log t$.

Since our list on related papers is surely incomplete we refer to [26] where a very rich bibliography is given.

REFERENCES

- [1] Alberti, P. M. and Uhlmann, A.: *Stochasticity and Partial Order*, VEB Deutscher Verlag der Wiss., Berlin 1981.
- [2] Ando, T.: *Topics on operator inequalities*, Lecture notes, Hokkaido Univ., Sapporo 1978.
- [3] —: *Lin. Alg. and Appl.* **26** (1979), 203–241.
- [4] Csiszár, I.: *Measures of information type for difference of probability distributions* (in Hungarian), C. Sc. Thesis, Budapest 1966.
- [5] —: *Studia Sci. Math. Hungar.* **2** (1967), 299–318.
- [6] Davies, E. B.: *Non-linear functionals in quantum mechanics*, Lecture notes, 1980.
- [7] Davis, C.: *Proc. Amer. Math. Soc. Symposia (Convexity)* **7** (1963), 187–201.
- [8] Epstein, H.: *Commun. Math. Phys.* **31** (1973), 317–325.
- [9] Hansen, F. and Pedersen, G. K.: *Math. Ann.* **258** (1982), 229–241.
- [10] Hiai, F., Ohya, M. and Tsukada, M.: *Pacific J. Math.* **96** (1981), 99–109.
- [11] Kosaki, H.: *Commun. Math. Phys.* **87** (1982), 315–329.
- [12] Kullback, S.: *Information theory and statistics*, John Wiley, New York 1959.
- [13] Lieb, E. H.: *Adv. Math.* **11** (1973), 267–288.
- [14] —: *Bull. Amer. Math. Soc.* **81** (1975), 1–13.
- [15] Lindblad, G.: *Commun. Math. Phys.* **33** (1973), 305–322.
- [16] —: *ibid.* **39** (1974), 111–119.

- [17] —: *ibid.* **40** (1975), 147–151.
- [18] Neumann, J. von: *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin 1932.
- [19] Petz, D.: Spectral scale of self-adjoint operators and trace inequalities, *J. Math. Anal. Appl.*, to appear.
- [20] Pusz, W. and Woronowicz, S. L.: *Lett. Math. Phys.* **2** (1978), 505–512.
- [21] Rényi, A.: *On measure of entropy and information*, Proc. of the 4th Berkeley Symp. on Math. Stat. and Probability, I, 547–561, Berkeley 1960.
- [22] —: *Wahrscheinlichkeitsrechnung*, VEB Deutscher Verlag der Wiss., Berlin 1962.
- [23] Simon, B.: *Trace ideals and their applications*, London Math. Soc. Lecture Note Ser. 35, Cambridge Univ. Press, Cambridge 1979.
- [24] Uhlmann, A.: *Commun. Math. Phys.* **54** (1977), 21–32.
- [25] Umegaki, H.: *Kôdai Math. Sem. Rep.* **14** (1962), 59–85.
- [26] Wehrl, A.: *Rev. Modern Phys.* **50** (1978), 221–260.
- [27] —: *Found. Phys.* **9** (1979), 939–946.
- [28] Wigner, E. P. and Yanase, M. M.: *Proc. Nat. Acad. Sci. (U.S.)* **49** (1963), 910–918.