PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
S 0002-9939(2011)11141-9
Article electronically published on August 4, 2011

FROM JOINT CONVEXITY OF QUANTUM RELATIVE ENTROPY TO A CONCAVITY THEOREM OF LIEB

JOEL A. TROPP

(Communicated by Marius Junge)

ABSTRACT. This paper provides a succinct proof of a 1973 theorem of Lieb that establishes the concavity of a certain trace function. The development relies on a deep result from quantum information theory, the joint convexity of quantum relative entropy, as well as a recent argument due to Carlen and Lieb.

1. INTRODUCTION

In his 1973 paper on trace functions, Lieb establishes an important concavity theorem [Lie73, Thm. 6] concerning the trace exponential.

Theorem 1 (Lieb)**.** *Let H be a fixed self-adjoint matrix. The map*

$$A \longmapsto \operatorname{tr} \exp (H + \log A) \tag{1}$$

is concave on the positive-definite cone.

The most direct proof of Theorem 1 is due to Epstein [Eps73]; see Ruskai's papers [Rus02, Rus05] for a condensed version of this argument. Lieb's original proof develops the concavity of the function (1) as a corollary of another deep concavity theorem [Lie73, Thm. 1]. In fact, many convexity and concavity theorems for trace functions are equivalent with each other, in the sense that the mutual implications follow from relatively easy arguments. See [Lie73, §5] and [CL08, §5] for discussions of this point.

The goal of this paper is to demonstrate that a modicum of convex analysis allows us to derive Theorem 1 directly from another major theorem, the joint convexity of the quantum relative entropy. The literature contains several elegant, conceptual proofs of the latter result; for example, see [Eff09]. These arguments now deliver Lieb's theorem as an easy corollary.

The author's interest in Theorem 1 stems from its striking applications in random matrix theory; refer to the paper [Tro10] for a detailed discussion. Researchers concerned with these developments may find the current approach to Lieb's theorem more transparent than earlier treatments.

Received by the editors January 2, 2011 and, in revised form, January 4, 2011.
2010 *Mathematics Subject Classification.* Primary 52A41.

The main ideas in our presentation are drawn from the work of Carlen and Lieb [CL08], so this dispatch does not contain a truly novel technique. Nevertheless, this paper should be valuable because it provides a geometric intuition for Theorem 1 and connects it to another major result.

2. Background

Our argument rests on the properties of a function, called the quantum relative entropy, which can be interpreted as a measure of dissimilarity between two positive-definite matrices.

Definition 2. Let X, Y be positive-definite matrices. The *quantum relative entropy* of X with respect to Y is defined as

$$\mathrm{D}(X;Y) := \mathrm{tr}(X \log X - X \log Y - (X - Y)).$$

Other appellations for this function include *quantum information divergence* and *von Neumann divergence*.

The quantum relative entropy has a nice geometric interpretation [DT07, §2.2 and §2.6]. Define the quantum entropy function $\varphi(X) := \mathrm{tr}(X \log X)$ for a positive-definite argument. The divergence $\mathrm{D}(X;Y)$ can be viewed as the difference between $\varphi(X)$ and the best affine approximation of the entropy φ at the matrix Y. That is,

$$\mathrm{D}(X;Y) = \varphi(X) - [\varphi(Y) + \langle \nabla\varphi(Y),\ X - Y \rangle].$$

The entropy φ is a strictly convex function, which implies that the affine approximation strictly underestimates φ. This observation yields the following result.

Fact 3. The quantum relative entropy is nonnegative:

$$\mathrm{D}(X;Y) \geq 0.$$

Equality holds if and only if $X = Y$.

In quantum statistical mechanics, Fact 3 is usually called *Klein's inequality*. Another proof proceeds by showing that certain functional relations for scalars extend to matrix trace functions [Pet94, §2].

The convexity properties of quantum relative entropy have paramount importance. We require a major theorem, due to Lindblad [Lin74, Lem. 2], which encapsulates the difficulties of the proof.

Fact 4 (Lindblad). The quantum relative entropy is a jointly convex function. That is,

$$\mathrm{D}(tX_1 + (1-t)X_2;\, tY_1 + (1-t)Y_2) \leq t \cdot \mathrm{D}(X_1;Y_1) + (1-t)\cdot \mathrm{D}(X_2;Y_2) \quad \text{for } t \in [0,1],$$

where X_i and Y_i are positive definite for $i = 1, 2$.

Fact 4 follows easily from Lieb's main concavity theorem [Lie73, Thm. 1]; Bhatia's book [Bha97, §IX.6 and Prob. IX.8.17] offers a clear account of this approach. The literature contains several other elegant proofs; see the papers [Uhl77, PW78, And79, Han06]. We single out Effros' work [Eff09] because it is accessible to researchers with experience in matrix theory and convex analysis.

Our final tool is a basic result from convex analysis which ensures that partial maximization of a concave function produces a concave function [CL08, Lem. 2.3]. We include the simple proof.

Proposition 5. *Let $f(\,\cdot\,;\,\cdot\,)$ be a jointly concave function. Then the function $y \mapsto \max_x f(x;y)$ obtained by partial maximization is concave, assuming the maximum is always attained.*

Proof. For each pair of points y_1 and y_2, there are points x_1 and x_2 that satisfy

$$f(x_1;y_1) = \max_x f(x;y_1) \quad \text{and} \quad f(x_2;y_2) = \max_x f(x;y_2).$$

For each $t \in [0,1]$, the joint concavity of f implies that

$$\begin{aligned}\max_x f(x;ty_1+(1-t)y_2) &\geq f(tx_1+(1-t)x_2;ty_1+(1-t)y_2)\\ &\geq t\cdot f(x_1;y_1)+(1-t)\cdot f(x_2;y_2)\\ &= t\cdot \max_x f(x;y_1)+(1-t)\cdot \max_x f(x;y_2).\end{aligned}$$

In words, the partial maximum is a concave function. □

3. Proof of Lieb's Theorem

We begin with a variational representation of the trace, which is a restatement of the fact that quantum relative entropy is nonnegative. The symbol $\succ$ denotes the positive-definite order.

Lemma 6 (Variational Formula for Trace)**.** *Let Y be a positive-definite matrix. Then*

$$\operatorname{tr} \mathsf{Y} = \max_{\mathsf{X}\succ 0} \operatorname{tr}(\mathsf{X}\log\mathsf{Y} - \mathsf{X}\log\mathsf{X} + \mathsf{X}).$$

Proof. Introduce the definition of the quantum relative entropy into Fact 3 and rearrange to reach

$$\operatorname{tr} \mathsf{Y} \geq \operatorname{tr}(\mathsf{X}\log\mathsf{Y} - \mathsf{X}\log\mathsf{X} + \mathsf{X}).$$

When $\mathsf{X} = \mathsf{Y}$, both sides are equal, which yields the advertised result. □

The main result follows quickly using the variational formula and the other tools we have assembled. As noted, the structure of this argument is parallel with the approach of Carlen–Lieb to another concavity theorem [CL08, Thm. 1.1].

Proof of Theorem 1. In the variational formula, Lemma 6, select $\mathsf{Y} = \exp(\mathsf{H} + \log\mathsf{A})$ to obtain

$$\operatorname{tr}\exp(\mathsf{H}+\log\mathsf{A}) = \max_{\mathsf{X}\succ 0} \operatorname{tr}(\mathsf{X}(\mathsf{H}+\log\mathsf{A}) - \mathsf{X}\log\mathsf{X} + \mathsf{X}).$$

The latter expression can be written compactly using the quantum relative entropy:

$$\operatorname{tr}\exp(\mathsf{H}+\log\mathsf{A}) = \max_{\mathsf{X}\succ 0} \,[\operatorname{tr}(\mathsf{XH}) - (\mathrm{D}(\mathsf{X};\mathsf{A}) - \operatorname{tr}\mathsf{A})]. \tag{2}$$

For each self-adjoint matrix H, Fact 4 implies that the bracket is a jointly concave function of the variables A and X. It follows from Proposition 5 that the right-hand side of (2) defines a concave function of A. This observation establishes the theorem. □

Remark 7. The expression (2) states that the function $f : \mathsf{H} \mapsto \operatorname{tr}\exp(\mathsf{H}+\log\mathsf{A})$ is the Fenchel conjugate of $\mathrm{D}(\,\cdot\,;\mathsf{A}) - \operatorname{tr}\mathsf{A}$. This observation implies that f is convex.

Acknowledgments

The author thanks Eric Carlen for a very illuminating discussion of matrix convexity theorems, including the paper [CL08], as well as comments on an early draft of this paper. Edward Effros contributed insights on quantum information theory, and Elliott Lieb emphasized the equivalences among concavity theorems. This work has been supported in part by ONR awards N00014-08-1-0883 and N00014-11-1-0025, AFOSR award FA9550-09-1-0643, and a Sloan Fellowship. The research was performed while the author attended the IPAM Fall 2010 program on optimization.

References

[And79] T. Ando, *Concavity of certain maps on positive definite matrices and applications to Hadamard products*, Linear Algebra Appl. **26** (1979), 203–241. MR535686 (80f:15023)

[Bha97] R. Bhatia, *Matrix analysis*, Graduate Texts in Mathematics, no. 169, Springer, Berlin, 1997. MR1477662 (98i:15003)

[CL08] E. A. Carlen and E. H. Lieb, *A Minkowski-type trace inequality and strong subadditivity of quantum entropy II: Convexity and concavity*, Lett. Math. Phys. **83** (2008), 107–126. MR2379699 (2009m:47041)

[DT07] I. S. Dhillon and J. A. Tropp, *Matrix nearness problems with Bregman divergences*, SIAM J. Matrix Anal. Appl. **29** (2007), no. 4, 1120–1146. MR2369287 (2008i:90050)

[Eff09] E. G. Effros, *A matrix convexity approach to some celebrated quantum inequalities*, Proc. Natl. Acad. Sci. USA **106** (2009), no. 4, 1006–1008. MR2475796 (2010a:81028)

[Eps73] H. Epstein, *Remarks on two theorems of E. Lieb*, Comm. Math. Phys. **31** (1973), 317–325. MR0343073 (49:7817)

[Han06] F. Hansen, *Extensions of Lieb's concavity theorem*, J. Statist. Phys. **124** (2006), no. 1, 87–101. MR2256619 (2008k:47040)

[Lie73] E. H. Lieb, *Convex trace functions and the Wigner–Yanase–Dyson conjecture*, Adv. Math. **11** (1973), 267–288. MR0332080 (48:10407)

[Lin74] G. Lindblad, *Expectations and entropy inequalities for finite quantum systems*, Comm. Math. Phys. **39** (1974), 111–119. MR0363351 (50:15789)

[Pet94] D. Petz, *A survey of certain trace inequalities*, Functional analysis and operator theory (Warsaw), Banach Center Publications, vol. 30, Polish Acad. Sci., 1994, pp. 287–298. MR1285615 (95c:15038)

[PW78] W. Pusz and S. L. Woronowicz, *Form convex functions and the WYDL and other inequalities*, Lett. Math. Phys. **2** 1977/78, 505–512. MR513118 (80j:47019)

[Rus02] M. B. Ruskai, *Inequalities for quantum entropy: A review with conditions for equality*, J. Math. Phys. **43** (2002), no. 9, 4358–4375. MR1924445 (2004e:82007)

[Rus05] ———, *Erratum: Inequalities for quantum entropy: A review with conditions for equality* [J. Math. Phys. 43, 4358 (2002)], J. Math. Phys. **46** (2005), no. 1, 0199101. MR2113776 (2005h:82010)

[Tro10] J. A. Tropp, *User-friendly tail bounds for sums of random matrices*, ACM Report 2010-01, California Inst. Tech., Pasadena, CA, April 2010.

[Uhl77] A. Uhlmann, *Relative entropy and the Wigner–Yanase–Dyson–Lieb concavity in an interpolation theory*, Comm. Math. Phys. **54** (1977), 21–32. MR0479224 (57:18671)

Department of Computing and Mathematical Sciences, California Institute of Technology, 1200 E. California Boulevard, MC 305-16, Pasadena, California 91125