Trace inequalities in nonextensive statistical mechanics

Shigeru Furuichi

1Department of Electronics and Computer Science, Tokyo University of Science, Yamaguchi, 756-0884, Japan

Abstract. In this short paper, we establish a variational expression of the Tsallis relative entropy. In addition, we derive a generalized thermodynamic inequality and a generalized Peierls-Bogoliubov inequality. Finally we give a generalized Golden-Thompson inequality.

Keywords : Trace inequalities, thermodynamic inequality, Peierls-Bogoliubov inequality, Golden-Thompson inequality and Tsallis relative entropy

2000 Mathematics Subject Classification : 47A63, 94A17, 15A39

1 Introduction

Recently, the matrix trace inequalities in statistical mechanics are studied by Bebiano et.al. in [2]. Their results are generalized by them in [1] via \( \alpha \)-power mean. In addition, the further generalized logarithmic trace inequalities are obtained and their convergences are shown via generalized Lie-Trotter formulae in [7]. Inspired by their works, we generalize the trace inequalities in [2] by means of a parametric extended logarithmic function \( \ln_\lambda \) which will be defined below. Our generalizations are different from those in [1, 7]. In the sense of our generalization, we give a generalized Golden-Thompson inequality. In addition, we give a related trace inequality as concluding remarks.

We denote \( e_\lambda^x \equiv (1 + \lambda x)^{\frac{1}{\lambda}} \) and its inverse function \( \ln_\lambda x \equiv \frac{x^{\lambda-1}}{\lambda} \), for \( \lambda \in (0, 1] \) and \( x \geq 0 \). The functions \( e_\lambda^x \) and \( \ln_\lambda x \) converge to \( e^x \) and \( \log x \) as \( \lambda \to 0 \), respectively. Note that we have the following relations:

\[
e_\lambda^{x+y+\lambda xy} = e_\lambda^x e_\lambda^y, \quad \ln_\lambda xy = \ln_\lambda x + \ln_\lambda y + \lambda \ln_\lambda x \ln_\lambda y.
\] (1)

The Tsallis entropy was originally defined in [20] by \(-\sum_{i=1}^n p_i^q \ln_{1-q} p_i = \frac{\sum_{i=1}^n (p_i^q - p_i)}{1-q}\) for any nonnegative real number \( q \) and a probability distribution \( p_i \equiv p(X = x_i) \) of a given random variable \( X \). Taking the limit as \( q \to 1 \), the Tsallis entropy converges to the Shannon entropy \(-\sum_{i=1}^n p_i \log p_i \). We may regard that the expectation value \( E_q(X) = \sum_{i=1}^n p_i^q x_i \) depending on the parameter \( q \) is adopted in order to define the Tsallis entropy as \( x_i = -\ln_{1-q} p_i \), while the usual expectation value \( E(X) = \sum_{i=1}^n p_i x_i \) is adopted in order to define the Shannon entropy as \( x_i = -\log p_i \). In the sequel we use the parameter \( \lambda \in (0, 1] \) instead of \( q \). There is a relation between these two parameters such that \( q = 1 - \lambda \).

The Tsallis entropy and the Tsallis relative entropy in quantum system (noncommutative system) are defined in the following manner. See [5, 6] for example.

Definition 1.1 The Tsallis entropy is defined by

\[
S_\lambda(\rho) \equiv \frac{\text{Tr}[\rho^{1-\lambda} - \rho]}{\lambda} = -\text{Tr}[\rho^{1-\lambda} \ln_\lambda \rho]
\]

E-mail:furuichi@ed.yama.tus.ac.jp
for a density operator $\rho$ and $\lambda \in (0, 1]$. The Tsallis relative entropy is defined by

$$D_\lambda(\rho|\sigma) \equiv \frac{\text{Tr}[\rho - \rho^{1-\lambda}\sigma^\lambda]}{\lambda} = \text{Tr}[\rho^{1-\lambda}(\ln\rho - \ln\sigma)]$$

for density operators $\rho$, $\sigma$ and $\lambda \in (0, 1]$. The Tsallis entropy and the Tsallis relative entropy converge to the von Neumann entropy $S(\rho) \equiv -\text{Tr}[\rho \ln \rho]$ and the relative entropy $D(\rho|\sigma) \equiv \text{Tr}[\rho(\ln \rho - \ln \sigma)]$ as $\lambda \to 0$, respectively. See [15] for details on the theory of quantum entropy. Two Tsallis entropies have nonadditivities such that

$$S_\lambda(\rho_1 \otimes \rho_2) = S_\lambda(\rho_1) + S_\lambda(\rho_2) + \lambda S_\lambda(\rho_1)S_\lambda(\rho_2), \quad (2)$$

and

$$D_\lambda(\rho_1 \otimes \rho_2|\sigma_1 \otimes \sigma_2) = D_\lambda(\rho_1|\sigma_1) + D_\lambda(\rho_2|\sigma_2) - \lambda D_\lambda(\rho_1|\sigma_1)D_\lambda(\rho_2|\sigma_2), \quad (3)$$
due to the nonadditivity Eq. (1) of the function $\ln\$. Thus the field of the study using these entropies is often called the nonextensive statistical physics and many research papers have been published in mainly statistical physics [21]. In addition, for the relative Rényi entropy of order $\lambda$

$$R_\lambda(\rho|\sigma) \equiv \frac{1}{\lambda} \log \text{Tr}[\rho^{1-\lambda}\sigma^\lambda], \quad \lambda \in (0, 1]$$

we have the following relation between $R_\lambda(\rho|\sigma)$ and $D_\lambda(\rho|\sigma)$:

$$\lambda D_\lambda(\rho|\sigma) + \exp[\lambda R_\lambda(\rho|\sigma)] = 1.$$

See our previous papers [5, 6] on the mathematical properties of the Tsallis entropy and the Tsallis relative entropy.

## 2 A variational expression of Tsallis relative entropy

In this section, we derive a variational expression of the Tsallis relative entropy as a parametric extension of that of the relative entropy in Lemma 1.2 of [10]. A variational expression of the relative entropy has been studied in the general setting of von Neumann algebras [18, 12]. In the sequel, we consider $n \times n$ complex matrices in the finite quantum system. A Hermitian matrix $A$ is called a nonnegative matrix (and denoted by $A \geq 0$) if $\langle x, Ax \rangle \geq 0$ for all $x \in \mathbb{C}^n$. A nonnegative matrix $A$ is called a positive matrix (and denoted by $A > 0$) if it is invertible. Throughout this paper, $\text{Tr}$ means the usual matrix trace. A nonnegative matrix $A$ is called a density matrix if $\text{Tr}[A] = 1$. In the below, we sometimes relax the condition of the unital trace for the matrices in the definition of the Tsallis relative entropy $D_\lambda(\cdot|\cdot)$, since it is not essential in the mathematical studies of the entropic functionals.

**Theorem 2.1** For $\lambda \in (0, 1]$, we have the following relations.

1. If $A$ and $Y$ are nonnegative matrices, then

$$\ln_\lambda \text{Tr}[e^{A+\ln_\lambda Y}] = \max \left\{ \text{Tr}[X^{1-\lambda}A] - D_\lambda(X|Y) : X \geq 0, \text{Tr}[X] = 1 \right\}.$$

2. If $X$ is a positive matrix with $\text{Tr}[X] = 1$ and $B$ is a Hermitian matrix, then

$$D_\lambda(X|e^B) = \max \left\{ \text{Tr}[X^{1-\lambda}A] - \ln_\lambda \text{Tr}[e^{A+B}] : A \geq 0 \right\}.$$
Proof: The proof is almost similar to that of Lemma 1.2 in [10].

(1) For \( \lambda = 1 \), we have \( \ln \lambda \text{Tr}[e^{A+\ln \lambda}Y] = \text{Tr}[X^{-\lambda}A] - D_\lambda(X|Y) \) by easy calculations with the usual convention \( X^0 = I \). Thus we assume \( \lambda \in (0, 1) \). Let us denote

\[
F_\lambda(X) = \text{Tr}[X^{-\lambda}A] - D_\lambda(X|Y)
\]

for a nonnegative matrix \( X \) with \( \text{Tr}[X] = 1 \). If we take the Schatten decomposition \( X = \sum_{j=1}^{n} r_j E_j \), where all \( E_j \), \( (j = 1, 2, \cdots, n) \) are projections of rank one with \( \sum_{j=1}^{n} E_j = I \) and \( r_j \geq 0 \), \( (j = 1, 2, \cdots, n) \) with \( \sum_{j=1}^{n} r_j = 1 \), then we rewrite

\[
F_\lambda \left( \sum_{j=1}^{n} r_j E_j \right) = \sum_{j=1}^{n} \left\{ r_j^{1-\lambda} \text{Tr}[E_j A] + \frac{1}{\lambda} r_j^{1-\lambda} \text{Tr}[E_j Y^\lambda] - \frac{1}{\lambda} r_j \text{Tr}[E_j] \right\}.
\]

Since we have

\[
\frac{\partial}{\partial r_j} F_\lambda \left( \sum_{j=1}^{n} r_j E_j \right) \bigg|_{r_j=0} = +\infty
\]

\( F_\lambda(X) \) attains its maximum at a nonnegative matrix \( X_0 \) with \( \text{Tr}[X_0] = 1 \). Then for any Hermitian matrix \( S \) with \( \text{Tr}[S] = 0 \) (since \( \text{Tr}[X_0 + tS] \) must be 1), we have

\[
0 = \frac{d}{dt} F_\lambda(X_0 + tS)|_{t=0} = (1 - \lambda) \text{Tr}[S(X_0^{-\lambda}A + \frac{1}{\lambda} X_0^{-\lambda}Y^\lambda)],
\]

so that \( X_0^{-\lambda}A + \frac{1}{\lambda} X_0^{-\lambda}Y^\lambda = cI \) for \( c \in \mathbb{R} \). Thus we have

\[
X_0 = \frac{e^{A+\ln \lambda}Y}{\text{Tr}[e^{A+\ln \lambda}Y]}
\]

by the normalization condition. By the formulae \( \ln \lambda \frac{y}{x} = \ln \lambda y + y^\lambda \ln \lambda \frac{1}{x} \) and \( \ln \lambda \frac{1}{x} = -x^{-\lambda} \ln \lambda x \), we have

\[
F_\lambda(X_0) = \ln \lambda \text{Tr}[e^{A+\ln \lambda}Y].
\]

(2) For \( \lambda = 1 \), we have \( D_\lambda(X|e^B) = \text{Tr}[X^{-\lambda}A] - \ln \lambda \text{Tr}[e^{A+B}] \) by easy calculations. Thus we assume \( \lambda \in (0, 1) \). It follows from (1) that the functional \( g(A) = \ln \lambda \text{Tr}[e^{A+B}] \) defined on the set of all nonnegative matrices is convex, due to triangle inequality on max. Now let \( A_0 = \ln \lambda X - B \), and denote

\[
G_\lambda(A) = \text{Tr}[X^{-\lambda}A] - \ln \lambda \text{Tr}[e^{A+B}],
\]

which is concave on the set of all nonnegative matrices. Then for any nonnegative matrix \( S \), there exists a nonnegative matrix \( A_0 \) such that

\[
\frac{d}{dt} G_\lambda(A_0 + tS)|_{t=0} = 0.
\]

Therefore \( G_\lambda(A) \) attains the maximum \( G_\lambda(A_0) = D_\lambda(X|e^B) \).

Taking the limit as \( \lambda \to 0 \), Theorem [24] recovers the similar form of Lemma 1.2 in [10] under the assumption of nonnegativity of \( A \). If \( Y = I \) and \( B = 0 \) in (1) and (2) of Theorem [24] respectively, then we obtain the following corollary.
Corollary 2.2  (1) If $A$ is a nonnegative matrix, then
\[ \ln \lambda \text{Tr}[e_A^\lambda] = \max \left\{ \text{Tr}[X^{1-\lambda}A] : X \geq 0, \text{Tr}[X] = 1 \right\}. \]

(2) For a density matrix $X$, we have
\[ -S_\lambda(X) = \max \left\{ \text{Tr}[X^{1-\lambda}A] - \ln \lambda \text{Tr}[e_A^\lambda] : A \geq 0 \right\}. \]

Taking the limit as $\lambda \to 0$, Corollary 2.2 recovers the similar form of Theorem 1 in [2] under the assumption of nonnegativity of $A$.

3 Generalized logarithmic trace inequalities

In this section, we derive some trace inequalities in terms of the results obtained in the previous section. From (1) of Corollary 2.2, we have the generalized thermodynamic inequality:
\[ \ln \lambda \text{Tr}[e_H^\lambda] \geq \text{Tr}[D^{1-\lambda}H] + S_\lambda(D), \]  
for a density matrix $D$ and a nonnegative matrix $H$. Putting $D = A$ and $H = \ln \lambda A$ in Eq. (4) for $A \geq 0$ and $B \geq I$, we have the generalized Peierls-Bogoliubov inequality (cf. Theorem 3.3 of [2]):
\[ (\text{Tr}[A])^{1-\lambda} (\ln \lambda \text{Tr}[A] - \ln \lambda \text{Tr}[B]) \leq \text{Tr}[A^{1-\lambda}(\ln \lambda A - \ln \lambda B)], \]
for nonnegative matrices $A$ and $B \geq I$.

Lemma 3.1 The following statements are equivalent.

(1) $F_\lambda(A) = \ln \lambda \text{Tr}[e_A^\lambda]$ is convex in a Hermitian matrix $A$.

(2) $f_\lambda(t) = \ln \lambda \text{Tr}[e_A^{1+tB}]$ is convex in $t \in R$.

Proof: Putting $A_1 = A + xB$ and $A_2 = A + yB$ in
\[ \ln \lambda \text{Tr}[e_A^{\mu A_1 + (1-\mu)A_2}] \leq \mu \ln \lambda \text{Tr}[e_A^{A_1}] + (1-\mu) \ln \lambda \text{Tr}[e_A^{A_2}], \]
we find the convexity of $f_\lambda(t)$. Thus (1) implies (2).
Conversely, putting $A = A_2$, $B = A_1 - A_2$ and $x = 1, y = 0$ in
\[ \ln \lambda \text{Tr}[e_A^{A + \mu A_2}] \leq \mu \ln \lambda \text{Tr}[e_A^{A_2}] + (1-\mu) \ln \lambda \text{Tr}[e_A^{A_2}], \]
we find the convexity of $F_\lambda(A)$. Thus (2) implies (1).

Corollary 3.2 For nonnegative matrices $A$ and $B$, we have
\[ \ln \lambda \text{Tr}[e_A^{A+B}] - \ln \lambda \text{Tr}[e_A^A] \geq \frac{\text{Tr}[B(e_A^{1-\lambda})]}{(\text{Tr}[e_A^A])^{1-\lambda}}. \]

Proof: From Theorem 2.1 $F_\lambda(A) = \ln \lambda \text{Tr}[e_A^\lambda]$ is convex. Due to Lemma 3.1, $f_\lambda(t) = \ln \lambda \text{Tr}[e_A^{1+tB}]$ is convex, which implies that $f_\lambda(1) \geq f_\lambda(0) + f_\lambda'(0)$. Thus we have the corollary by simple calculations.
4 Generalized exponential trace inequality

For nonnegative real numbers $x, y$ and $0 < \lambda \leq 1$, the relations $e^{x+y} \leq e^{x+y+\lambda xy} = e^x e^y$ hold. These relations naturally motivate us to consider the following inequalities in the noncommutative case.

**Proposition 4.1** For nonnegative matrices $X$ and $Y$, and $0 < \lambda \leq 1$, we have

$$\text{Tr}[e^{X+Y}] \leq \text{Tr}[e^{X+Y+\lambda XY^{1/2}X^{1/2}}].$$

**(Proof)** Since $Y^{1/2}XY^{1/2} \geq 0$ for nonnegative matrices $X$ and $Y$, we have

$$I + \lambda (X + Y) \leq I + \lambda \left\{ X + Y + \lambda (Y^{1/2}XY^{1/2}) \right\}.$$

It is known \[17\] that $A \leq B$ implies $\text{Tr}[f(A)] \leq \text{Tr}[f(B)]$, if $A$ and $B$ are Hermitian matrices and $f : \mathbb{R} \to \mathbb{R}$ is an increasing function. Thus we have

$$\text{Tr}[(I + \lambda(X + Y))^{1/\lambda}] \leq \text{Tr}((I + \lambda \left\{ X + Y + \lambda (Y^{1/2}XY^{1/2}) \right\})^{1/\lambda}).$$

for $0 < \lambda \leq 1$.

Note that we have the matrix inequality:

$$e^{X+Y}\leq e^{X+Y+\lambda XY^{1/2}X^{1/2}}$$

for $\lambda \geq 1$ by the application of the Löwner-Heinz inequality \[14\] [9] [16].

**Proposition 4.2** For nonnegative matrices $X, Y,$ and $\lambda \in (0, 1]$, we have

$$\text{Tr}[e^{X+Y+\lambda XY}] \leq \text{Tr}[e^X e^Y].$$

**(7)**

**Proof:** For $\lambda \in (0, 1]$, we have

$$\text{Tr}[(AB)^{1/\lambda}] \leq \text{Tr}[A^{1/\lambda}B^{1/\lambda}]$$

from the special case of the Lieb-Thirring inequality \[13\] (see also \[11\] [22]), $\text{Tr}[(AB)^{\alpha}] \leq \text{Tr}[A^{\alpha}B^{\alpha}]$ for nonnegative matrices $A, B$ and any real number $\alpha \geq 1$. Putting $A = I + \lambda X$ and $B = I + \lambda Y$ in Eq.(8), we have

$$\text{Tr}[(I + \lambda X)(I + \lambda Y)]^{1/\lambda} \leq \text{Tr}[(I + \lambda X)^{1/\lambda}(I + \lambda Y)^{1/\lambda}],$$

which is the desired one.

Notice that Golden-Thompson inequality \[8\] [19],

$$\text{Tr}[e^{X+Y}] \leq \text{Tr}[e^X e^Y]$$

which holds for Hermitian matrices $X$ and $Y$, is recovered by taking the limit as $\lambda \to 0$ in Proposition \[4.2\] in particular case of nonnegative matrices $X$ and $Y$. 


5 Concluding remarks

Since $\text{Tr}[HZHZ] \leq \text{Tr}[H^2Z^2]$ for Hermitian matrices $H$ and $Z$ \cite{18, 4}, we have for nonnegative matrices $X$ and $Y$,

$$\text{Tr}[(I + X + Y + Y^{1/2}XY^{1/2})^2] \leq \text{Tr}[(I + X + Y)^2]$$

by easy calculations. This implies the inequality

$$\text{Tr}[e^{X+Y+1/2Y^{1/2}XY^{1/2}}] \leq \text{Tr}[e^{X+Y+1/2XY}].$$

Thus we have

$$\text{Tr}[e^{X/Y}] \leq \text{Tr}[e^{X/Y}] \leq \text{Tr}[e^{X/Y}] \leq \text{Tr}[e^{X/Y}]$$

from Proposition 4.1 and Proposition 4.2. Putting $B = \ln_{1/2} Y$ and $A = \ln_{1/2} Y^{-1/2}XY^{-1/2}$ in (2) of Theorem 2.1 under the assumption of $I \leq Y \leq X$ and using Eq. (9), we have

$$D_{1/2}(X|Y) = D_{1/2}(X|\ln_{1/2} Y)$$

$$\geq \text{Tr}[X^{1/2}A] - \ln_{1/2} \text{Tr}[e^{A+B}]$$

$$\geq \text{Tr}[X^{1/2}A] - \ln_{1/2} \text{Tr}[e^{A+B}]$$

$$\leq \text{Tr}[X^{1/2} \ln_{1/2} Y^{-1/2}XY^{-1/2}] - \ln_{1/2} \text{Tr}[Y^{-1/2}XY^{-1/2}]$$

which gives a lower bound of the Tsallis relative entropy in the case of $\lambda = 1/2$ and $I \leq Y \leq X$.

Acknowledgement

We would like to thank the reviewers for providing valuable comments to improve the manuscript. This work was supported by the Japanese Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Encouragement of Young scientists (B), 17740068.

References

[1] N.Bebiano, R.Lemos and J.da Providência Jr., Inequalities for quantum relative entropy, Linear Alg. Appl., Vol.401, pp.159-172 (2005).

[2] N.Bebiano, J.da Providência Jr. and R.Lemos, Matrix inequalities in statistical mechanics, Linear Alg. Appl., Vol.376, pp.265-273 (2004).

[3] J.-C.Bourin, Some inequalities for norms on matrices and operators, Linear Alg.Appl., Vol.292, pp.139-154 (1999).

[4] J.I.Fujii, A trace inequality arising from quantum information theory, Linear Alg. Appl., Vol.400, pp.141-146 (2005).

[5] S.Furuichi, K.Yanagi and K.Kuriyama, Fundamental properties of Tsallis relative entropy, J.Math.Phys.,Vol.45,pp.4868-4877 (2004).

[6] S.Furuichi, K.Yanagi and K.Kuriyama, A note on operator inequalities of Tsallis relative operator entropy, Linear Alg. Appl.,Vol.407,pp.19-31 (2005).
[7] T. Furuta, Convergence of logarithmic trace inequalities via generalized Lie-Trotter formulae, Linear Alg. Appl., Vol. 396, pp. 353-372 (2005).

[8] S. Golden, Lower bounds for the Helmholtz function, Phys. Rev., Vol. 137, pp. B1127-B1128 (1965).

[9] E. Heinz, Beiträge zur Störungstheorie der Spektralzerlegung, Math. Ann., Vol. 123, pp. 415-438 (1951).

[10] F. Hiai and D. Petz, The Golden-Thompson trace inequality is complemented, Linear Alg. Appl., Vol. 181, pp. 153-185 (1993).

[11] Y. Hong and R. A. Horn, The Jordan canonical form of a product of a Hermitian and a positive semidefinite matrix, Linear Alg. Appl., Vol. 147, pp. 373-386 (1991).

[12] H. Kosaki, Relative entropy for states: a variational expression, J. Operator Theory, Vol. 16, pp. 335-348 (1986).

[13] E. H. Lieb and W. Thirring, Inequalities for the moments of the eigenvalues of the Schrödinger Hamiltonian and their relation to Sobolev inequalities, in E. Lieb, B. Simon and A. S. Wightman eds., “Studies in Mathematical Physics. Essays in Honor of Valentine Bargmann”, Princeton University Press, Princeton, NJ, pp. 269-303 (1976).

[14] K. Löwner, Über monotone Matrixfunktionen, Math. Z., Vol. 38, pp. 177-216 (1934).

[15] M. Ohya and D. Petz, Quantum Entropy and its Use, Springer-Verlag, 1993.

[16] G. K. Pedersen, Some operator monotone functions, Proc. Amer. Math. Soc., Vol. 36, pp. 309-310 (1972).

[17] D. Petz, A survey of certain trace inequalities, in Functional Analysis and Operator Theory, pp. 287-298, Banach Center Publications, 30 (Warszawa 1994).

[18] D. Petz, A variational expression for the relative entropy of states of a von Neumann algebra, Commun. Math. Phys., Vol. 114, pp. 345-349 (1988).

[19] C. J. Thompson, Inequality with applications in statistical mechanics, J. Math. Phys., Vol. 6, pp. 1812-1813 (1965).

[20] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, J. Stat. Phys., Vol. 52, pp. 479-487 (1988).

[21] C. Tsallis et al., *Nonextensive Statistical Mechanics and Its Applications*, edited by S. Abe and Y. Okamoto (Springer-Verlag, Heidelberg, 2001); see also the comprehensive list of references at [http://tsallis.cat.cbpf.br/biblio.htm](http://tsallis.cat.cbpf.br/biblio.htm).

[22] B. Y. Wang and F. Zhang, Trace and eigenvalue inequalities for ordinary and Hadamard products of positive semidefinite Hermitian matrices, SIAM J. Matrix. Anal. Appl., Vol. 16, No. 4, pp. 1173-1183 (1995).