

JENSEN'S TRACE INEQUALITY IN SEVERAL VARIABLES

FRANK HANSEN & GERT K. PEDERSEN

4th March, 2003

ABSTRACT. For a convex, real function f we present a simple proof of the formula

$$\mathrm{Tr}(f(\sum_{k=1}^m a_k^* x_k a_k)) \leq \mathrm{Tr}(\sum_{k=1}^m a_k^* f(x_k) a_k),$$

valid for each tuple (x_1, \dots, x_m) of symmetric matrices in \mathbb{M}_n and every unital column (a_1, \dots, a_m) of matrices, i.e. $\sum_{k=1}^m a_k^* a_k = \mathbf{1}$. This is the standard Jensen trace inequality. If $f \geq 0$ it holds also for the unbounded trace on $\mathbb{B}(\mathfrak{H})$, where \mathfrak{H} is an infinite-dimensional Hilbert space. We then investigate the more general case where τ is a densely defined, lower semi-continuous trace on a C^* -algebra \mathcal{A} and f is a convex, continuous function of n variables, and show that we have the inequality

$$\tau(f(\sum_{k=1}^m a_k^* \underline{x}_k a_k)) \leq \tau(\sum_{k=1}^m a_k^* f(\underline{x}_k) a_k)$$

for every family of abelian n -tuples $\underline{x}_k = (x_{1k}, \dots, x_{nk})$, i.e. tuples of self-adjoint elements in \mathcal{A} such that $[x_{ik}, x_{jk}] = 0$ for all i, j and k , where $1 \leq k \leq m$, and every unital m -column (a_1, \dots, a_m) in $M(\mathcal{A})$, provided that the elements $y_i = \sum_{k=1}^m a_k^* x_{ik} a_k$ also form an abelian n -tuple. We even establish this result for weak* measurable, self-adjoint, abelian fields $(x_{it})_{t \in T}$, $1 \leq i \leq n$, i.e. $[x_{it}, x_{jt}] = 0$ for all i, j and t , and a weak* measurable, unital column field $(a_t)_{t \in T}$ in $M(\mathcal{A})$ paired with any trace or trace-like functional φ , i.e. one that contains the n -tuple (presumed abelian) with elements $y_i = \int_T a_t^* x_{it} a_t d\mu(t)$ in its centralizer. This takes the form of the inequality

$$\varphi(f(\int_T a_t^* \underline{x}_t a_t d\mu(t))) \leq \varphi(\int_T a_t^* f(\underline{x}_t) a_t d\mu(t)).$$

We also study functions of n variables that are monotone increasing in each variable, and show in two important cases that $\varphi(f(\underline{x})) \leq \varphi(f(\underline{y}))$ whenever $\underline{x} = (x_1, \dots, x_n)$ and $\underline{y} = (y_1, \dots, y_n)$ are abelian n -tuples with $x_i \leq y_i$ for each i and φ is a trace or a trace-like functional.

1. Introduction. Several important concepts in operator theory, in quantum statistical mechanics (the entropy, the relative entropy and Gibbs' free energy), in electrical engineering and in mathematical economics involve the trace of a function of a self-adjoint operator. This has motivated a considerable amount of abstract research about such functions in the last fifty years. An important subset of questions concern the convexity of trace functions with respect to their argument, and the generalizations of this known as Jensen trace inequalities.

1991 *Mathematics Subject Classification.* Primary 46L05; Secondary 46L10, 47A60, 46C15.

Key words and phrases. Operator algebras, trace functions, trace inequalities, Jensen inequality.

©2001 by the authors. This paper may be reproduced, in its entirety, for non-commercial purposes.

The convexity of the function $x \rightarrow \operatorname{Tr}(f(x))$, when f is a convex function of one variable and x is a self-adjoint operator, was known to von Neumann, cf. [21, V.3. p. 390]. An early proof for $f(x) = \exp(x)$ can be found in [29, 2.5.2]. A proof found by E.H. Lieb in the early seventies describes the number $\operatorname{Tr}(f(x))$, where f is convex, as a supremum (taken over all possible choices of orthonormal bases of the Hilbert space) of the sum of the values of f at the diagonal elements of the matrix for x . Obviously, then, this is a convex function of x . The proof was communicated to B. Simon, who used the method to give an alternative proof of the second Berezin-Lieb inequality in [30, Theorem 2.4], see also [31, Lemma II.10.4]. Simon only considers the exponential function, but the argument is valid for any convex function, cf. [17, Proposition 3.1]. The general case for an arbitrary normal trace on a von Neumann algebra was established by D. Petz in [28, Theorem 4], using the theory of spectral dominance (spectral scale).

When a convex combination $\sum_{k=1}^m \lambda_k x_k$ of matrices (or operators) with coefficients $(\lambda_1, \dots, \lambda_m)$ is replaced by the non-commutative version $\sum_{k=1}^m a_k^* x_k a_k$, where (a_1, \dots, a_m) is a unital m -column, i.e. an m -tuple of matrices (or operators) such that $\sum_{k=1}^m a_k^* a_k = \mathbf{1}$, we obtain a generalization known as Jensen's operator inequality. For an *operator convex* function, i.e. a function f such that $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ for any pair of self-adjoint matrices x and y (of arbitrary high order), this result was found by the first author in [6], and used by the two of us in [9] to give a concise review of Löwner's and Bendat-Sherman's theory of operator monotone and operator convex functions. With hindsight we must admit that we unfortunately chose the contractive form $f(a^* x a) \leq a^* f(x) a$ for $a^* a \leq \mathbf{1}$, this being the seemingly most attractive version at the time. However, this necessitated the further conditions that $0 \in I$ and $f(0) \leq 0$, conditions that have haunted the theory since then, and which become a real obstacle when we pass to several variables. The Jensen inequality for a trace on a von Neumann algebra and an arbitrary convex function f was found by Brown and Kosaki in [5], still in the contractive version. Elementary proofs of these results can now be found in [11].

We begin the paper with the simple proof of the full Jensen trace inequality for matrices taken from [11], which uses ideas from Lieb's proof mentioned above. Although this result follows from the more general theorem later on in the paper we feel that an elementary proof of the most applicable version would be a convenience for the (not too specialized) reader. Also, the simple proof contains all the basic ideas in the more elaborate versions and thus makes it easier to grasp these.

2. Theorem. *If $f: I \rightarrow \mathbb{R}$ is a convex function defined on an interval I the inequality*

$$\operatorname{Tr} \left(f \left(\sum_{k=1}^m a_k^* x_k a_k \right) \right) \leq \operatorname{Tr} \left(\sum_{k=1}^m a_k^* f(x_k) a_k \right) \quad (1)$$

holds for each m -tuple of self-adjoint $n \times n$ matrices (x_1, \dots, x_m) with spectra in I , every unital m -tuple (a_1, \dots, a_m) of $n \times n$ matrices and all natural numbers n .

Proof. Let $x_k = \sum_{\operatorname{sp}(x_k)} \lambda E_k(\lambda)$ denote the spectral resolution of x_k for $1 \leq k \leq m$. Thus, $E_k(\lambda)$ is the spectral projection of x_k on the eigenspace corresponding to λ if λ is an eigenvalue for x_k ; otherwise $E_k(\lambda) = 0$. For each unit vector ξ in \mathbb{C}^n define

the (atomic) probability measure

$$\mu_\xi(S) = \left(\sum_{k=1}^m a_k^* E_k(S) a_k \xi \mid \xi \right) = \sum_{k=1}^m (E_k(S) a_k \xi \mid a_k \xi) \quad (2)$$

for any (Borel) set S in \mathbb{R} . Note now that if $y = \sum_{k=1}^m a_k^* x_k a_k$ then

$$\begin{aligned} (y\xi \mid \xi) &= \left(\sum_{k=1}^m a_k^* x_k a_k \xi \mid \xi \right) \\ &= \left(\sum_{k=1}^m \sum_{\text{sp}(x_k)} \lambda E_k(\lambda) a_k \xi \mid a_k \xi \right) = \int \lambda d\mu_\xi(\lambda). \end{aligned} \quad (3)$$

If a unit vector ξ is an eigenvector for y then the corresponding eigenvalue is $(y\xi \mid \xi)$ and ξ is also an eigenvector for $f(y)$ with corresponding eigenvalue $(f(y)\xi \mid \xi) = f((y\xi \mid \xi))$. In this case we therefore have

$$\begin{aligned} &\left(f \left(\sum_{k=1}^m a_k^* x_k a_k \right) \xi \mid \xi \right) = (f(y)\xi \mid \xi) = f((y\xi \mid \xi)) \\ &= f \left(\int \lambda d\mu_\xi(\lambda) \right) \leq \int f(\lambda) d\mu_\xi(\lambda) \\ &= \sum_{k=1}^m \left(\sum_{\text{sp}(x_k)} f(\lambda) E_k(\lambda) a_k \xi \mid a_k \xi \right) = \sum_{k=1}^m (a_k^* f(x_k) a_k \xi \mid \xi), \end{aligned} \quad (4)$$

where we used (3) and the convexity of f – in form of the usual Jensen inequality – to get the inequality in (4).

The result in (1) now follows by summing over an orthonormal basis of eigenvectors for y . □

3. Spectral Theory in Several Variables. The really new problems start when we consider a function $f(\underline{\lambda})$ of n real variables (with $\underline{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)$). Naturally, we wish to replace the real variables λ_j by self-adjoint operators x_j as in the one-variable case. An immediate problem that now arises is how to define $f(\underline{x})$ in this case. The spectral theorem which was used in the one-variable case fails here unless the x_j 's commute with one another. This means that the largest domain of definition for f is the set of *abelian* n -tuples in $\mathbb{B}(\mathfrak{H})$, i.e. tuples $\underline{x} = (x_1, \dots, x_n)$ of self-adjoint elements such that $[x_i, x_j] = 0$ for all i and j .

For functions of two variables the spectral theory of abelian tuples (pairs) is equal to spectral theory for normal, instead of self-adjoint operators. [As long as we consider only continuous and not differentiable functions, a complex function is just a function of two real variables !] This theory is markedly more difficult than the one variable case, in particular because the set of normal operators has no linear structure.

To be more specific, consider a C^* -algebra \mathcal{A} of operators on some Hilbert space \mathfrak{H} . For each interval I let M_I denote the convex set of self-adjoint elements in \mathcal{A}

with spectra contained in I . If $\underline{I} = I_1 \times \cdots \times I_n \subset \mathbb{R}^n$ and f is a continuous function on \underline{I} we can for each abelian n -tuple $\underline{x} = \{x_1, \dots, x_n\}$ in $\bigoplus \mathcal{A}_{\text{sa}}^{I_i}$ define an element $f(\underline{x})$ in \mathcal{A} . To see this, let $x_i = \int \lambda dE_i(\lambda)$ be the spectral resolution of x_i for $1 \leq i \leq n$. Since the x_i 's commute, so do their spectral measures. We can therefore define the product spectral measure E on \underline{I} by $E(S_1 \times \cdots \times S_n) = E_1(S_1) \cdots E_n(S_n)$, and then write

$$f(\underline{x}) = \int f(\underline{\lambda}) dE(\underline{\lambda}) = \int f(\lambda_1, \dots, \lambda_n) dE(\lambda_1, \dots, \lambda_n). \quad (5)$$

Of course, if f is a polynomial in the variables $\lambda_1, \dots, \lambda_n$ we simply find $f(\underline{x})$ by replacing each λ_i with x_i . The map $f \rightarrow f(\underline{x})$ so obtained is a $*$ -homomorphism of $C(\underline{I})$ into \mathcal{A} and generalizes the ordinary spectral mapping theory for a single (self-adjoint) operator. In particular, the support of the map (the smallest closed set S such that $f(\underline{x}) = 0$ for every function f that vanishes on S) may be regarded as the “joint spectrum” of the elements x_1, \dots, x_n . In Gelfand language the commutative unital C^* -subalgebra generated by the x_i 's is $*$ -isomorphic to $C(S)$.

4. Convexity in Several Variables. The set of abelian n -tuples in $\mathbb{B}(\mathfrak{H})$ is obviously not a convex set, so at first glance it makes little sense to discuss convexity properties of the operator function $\underline{x} \rightarrow f(\underline{x})$. We shall therefore consider abelian tuples \underline{x} and \underline{y} that are *compatible*, which by definition means that the line segment between them also consists of abelian tuples. It is easily seen that this happens precisely when

$$[x_i, y_j] = [x_j, y_i] \quad \text{for all } i \text{ and } j. \quad (6)$$

Now we can meaningfully ask whether $\text{Tr}(f(\lambda \underline{x} + (1-\lambda)\underline{y})) \leq \text{Tr}(\lambda f(\underline{x}) + (1-\lambda)f(\underline{y}))$ when f is a convex function.

Note from (6) that if $\{\underline{x}_1, \dots, \underline{x}_m\}$ is a set of pairwise compatible, abelian n -tuples, then any linear combination $\sum_{k=1}^m \lambda_k \underline{x}_k$ is again an abelian n -tuple compatible with all the \underline{x}_i 's, so that the set $\text{conv}\{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_m\}$ is a convex domain for the operator function f . This also means that any set \mathcal{S}_o of pairwise compatible, abelian n -tuples in a C^* -subalgebra \mathcal{A} of $\mathbb{B}(\mathfrak{H})$ is contained in a maximal set \mathcal{S} , which by necessity must be a closed, linear subspace of $\mathcal{A}_{\text{sa}}^n$. One may wonder how such maximal sets look like, and a few experiments show that the variety is wide. Let \mathcal{C} be a commutative C^* -subalgebra of \mathcal{A} such that $\mathcal{C}'' = \mathcal{C}$ (where $'$ denotes relative commutant). For example, \mathcal{C} could be the center of \mathcal{A} (in which case $\mathcal{C}' = \mathcal{A}$), or it could be any maximal abelian C^* -subalgebra of \mathcal{A} (in which case $\mathcal{C}' = \mathcal{C}$). Note though, that the condition $\mathcal{C}'' = \mathcal{C}$ means that \mathcal{C} always contains the center of \mathcal{A} . Now fix a non-zero vector $(\varepsilon_1, \dots, \varepsilon_n)$ in \mathbb{R}^n and define

$$\mathcal{S} = \{\underline{x} = (\varepsilon_1 x + c_1, \dots, \varepsilon_n x + c_n) \mid x \in \mathcal{C}'_{\text{sa}}, c_i \in \mathcal{C}_{\text{sa}}\}. \quad (7)$$

Then it is easy to check that \mathcal{S} is a maximal set of pairwise compatible, abelian n -tuples in \mathcal{A} .

The more useful examples occur, however, at the other extreme of the situation above. We assume that the C^* -algebra \mathcal{A} comes equipped with a set of pairwise commuting C^* -subalgebras $\mathcal{A}_1, \dots, \mathcal{A}_n$. Then the subspace

$$\bigoplus_{i=1}^n (\mathcal{A}_i)_{\text{sa}} = \{\underline{x} = (x_1, \dots, x_n) \mid x_i \in (\mathcal{A}_i)_{\text{sa}}\} \quad (8)$$

consists of pairwise compatible, abelian n -tuples; and under the mild extra condition that each \mathcal{A}_i equals the relative commutant in \mathcal{A} of the C^* -algebra generated by the \mathcal{A}_j 's for $j \neq i$, (i.e. $(\bigcup_{j \neq i} \mathcal{A}_j)' = \mathcal{A}_i$) the space is also maximal. This condition may be achieved by replacing in turn each of the algebras \mathcal{A}_i by $(\bigcup_{j \neq i} \mathcal{A}_j)'$.

This frame applies readily to the seminal situation where $\mathcal{A} = \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$ in $\mathbb{B}(\mathfrak{H})$. Indeed, most authors that have considered operator functions of several variables have followed Korányi's lead and used the functions only on tensor products, cf. [14].

In the setting of compatible, abelian tuples we are going to replace the trace Tr on the Hilbert space by a densely defined, lower semi-continuous trace τ on an abstract C^* -algebra \mathcal{A} ; i.e. a functional defined on the set \mathcal{A}_+ of positive elements with values in $[0, \infty]$, such that $\tau(x^*x) = \tau(xx^*)$ for all x in \mathcal{A} . Thus we shall consider the function $\underline{x} \rightarrow \tau(f(\underline{x}))$ on a set of compatible, abelian n -tuples in \mathcal{A}_{sa} .

Some of our results have appeared in more primitive versions before. The tracial convexity of the function $\underline{x} \rightarrow f(\underline{x})$ on the space of n -tuples in $\bigoplus_{i=1}^n (\mathcal{A}_i)_{\text{sa}}$ with values in $\bigotimes_{i=1}^n \mathcal{A}_i$ was proved by the first author for matrix algebras in [8]. His result was extended to general operator algebras and traces by the second author in [27]. Both proofs rely on Fréchet differentiability and somewhat intricate manipulations with first and second order differentials. It was then realized by Lieb that his proof, mentioned above, could be extended to the case of several variables with only marginal changes, and the improved version appeared in [18]. The present version generalizes and subsumes the previous papers. In particular we show that the function $\underline{x} \rightarrow \tau(f(\underline{x}))$ is convex on any set of the form $\text{conv}\{\underline{x}_1, \dots, \underline{x}_m\}$, where the \underline{x}_k 's are pairwise compatible, abelian n -tuples in \mathcal{A}_{sa} .

5. Measurable Fields of Operators. Let \mathcal{A} be a (separable) C^* -algebra of operators on some (separable) Hilbert space \mathfrak{H} and T a locally compact metric space equipped with a Radon measure μ . We say that a field $(a_t)_{t \in T}$ of operators in the multiplier algebra $M(\mathcal{A})$ of \mathcal{A} , i.e. the C^* -algebra of elements a in $\mathbb{B}(\mathfrak{H})$ such that $x\mathcal{A} + \mathcal{A}x \subset \mathcal{A}$, is *weak* measurable* if each function $t \rightarrow \varphi(a_t)$, where $\varphi \in \mathcal{A}^*$, is μ -measurable. It is worth noticing that $(a_t)_{t \in T}$ is weak* measurable if (and only if) for each vector ξ in \mathfrak{H} the function $t \rightarrow a_t \xi$ is weakly (equivalently strongly) measurable (because the set of linear combinations of vector functionals is weak* dense in \mathcal{A}^*). It follows that if both $(a_t)_{t \in T}$ and $(b_t)_{t \in T}$ are weak* measurable fields then also $(a_t b_t)_{t \in T}$ is a measurable field.

If the function $t \rightarrow \varphi(a_t)$ is integrable for all states φ and $\int_T |\varphi(a_t)| d\mu(t) \leq \gamma$ for some constant γ , in particular if the function $t \rightarrow \|a_t\|$ is integrable, there is a unique element in $M(\mathcal{A})$, designated by $\int_T a_t d\mu(t)$, such that

$$\varphi \left(\int_T a_t d\mu(t) \right) = \int_T \varphi(a_t) d\mu(t) \quad \varphi \in \mathcal{A}^*, \tag{10}$$

cf. [26, 2.5.15]. We say in this case that the field $(a_t)_{t \in T}$ is *integrable*. If all the a_t 's belong to \mathcal{A} then also $\int_T a_t d\mu(t)$ belongs to \mathcal{A} . If the weak* measurable field $(a_t^* a_t)_{t \in T}$ is integrable with integral $\mathbf{1}$ we say that $(a_t)_{t \in T}$ is a *unital column field*.

6. Final Notations. Consider now an n -tuple of weak* measurable, bounded fields $(x_{it})_{t \in T}$, each consisting of self-adjoint elements in \mathcal{A} with spectra in some fixed interval I , and assume that $[x_{ij} - x_{ji}] = 0$ for all i, j and t . Thus each vector

$\underline{x}_t = (x_{1t}, \dots, x_{nt})$ is an abelian n -tuple. Furthermore, consider a unital column field $(a_t)_{t \in T}$ in $M(\mathcal{A})$, i.e. $\int_T a_t^* a_t d\mu(t) = \mathbf{1}$. Assume finally that the elements $y_i = \int_T a_i^* x_{it} a_t d\mu(t)$ in \mathcal{A} form an abelian n -tuple.

The commutation condition above for the y_i 's depends on intricate relations between the two measurable fields $(x_{it})_{t \in T}$ and $(a_t)_{t \in T}$. It is, however, satisfied if the fields satisfy the following extension of the commutativity condition in (5):

$$[a_i^* x_{it} a_t, a_s^* x_{js} a_s] = [a_i^* x_{jt} a_t, a_s^* x_{is} a_s] \quad \text{for all } s \text{ and } t. \quad (11)$$

Thus in particular if $x_{it} a_t a_s^* x_{js} = x_{jt} a_t a_s^* x_{is}$ for all s and t .

For ease of notation we shall write $\underline{I} = I_1 \times \dots \times I_n$ and $\underline{\lambda} = (\lambda_1, \dots, \lambda_n)$ if $\underline{\lambda} \in \underline{I}$. Moreover, we regard the vector space of n -tuples in \mathcal{A} as a bimodule over $M(\mathcal{A})$ and write $\underline{x}_t = (x_{1t}, \dots, x_{nt})$ and $a_t^* \underline{x}_t a_t = (a_t^* x_{1t} a_t, \dots, a_t^* x_{nt} a_t)$, so that $\underline{y} = (y_1, \dots, y_n) = \int_T a_t^* \underline{x}_t a_t d\mu(t)$.

We finally recall that the *centralizer* of a positive functional φ on \mathcal{A} is the C^* -subalgebra $\mathcal{A}^\varphi = \{y \in \mathcal{A} \mid \forall x \in \mathcal{A}: \varphi(xy) = \varphi(yx)\}$. If φ is unbounded, but lower semi-continuous on \mathcal{A}_+ and finite on the minimal dense ideal $K(\mathcal{A})$ of \mathcal{A} , we define $\mathcal{A}^\varphi = \{y \in \mathcal{A} \mid \forall x \in K(\mathcal{A}): \varphi(xy) = \varphi(yx)\}$.

7. Theorem. *Let $(\underline{x}_t)_{t \in T}$ be a bounded, weak* measurable field of abelian n -tuples in a C^* -algebra \mathcal{A} , with $\text{sp}(x_{it}) \subset I_i$ for $1 \leq i \leq n$, and let $(a_t)_{t \in T}$ be a unital column field in $M(\mathcal{A})$ such that the elements $y_i = \int_T a_i^* x_{it} a_t d\mu(t)$ form an abelian n -tuple. Then for each continuous, convex function f defined on the cube $\underline{I} = I_1 \times \dots \times I_n$ in \mathbb{R}^n and every positive functional φ that contains the y_i 's in its centralizer \mathcal{A}^φ , i.e. $\varphi(xy_i) = \varphi(y_i x)$ for all x in \mathcal{A} and every y_i , we have the inequality:*

$$\varphi \left(f \left(\int_T a_t^* \underline{x}_t a_t d\mu(t) \right) \right) \leq \varphi \left(\int_T a_t^* f(\underline{x}_t) a_t d\mu(t) \right). \quad (12)$$

If φ is unbounded, but lower semi-continuous on \mathcal{A}_+ and finite on the minimal dense ideal $K(\mathcal{A})$ of \mathcal{A} , the result still holds if $f \geq 0$, even though the function may now attain infinite values.

Proof. Let $\mathcal{C} = C_o(S)$ denote the commutative C^* -subalgebra of \mathcal{A} generated by y_1, \dots, y_n , and let μ_φ be the finite Radon measure on the locally compact, metric space S defined, via the Riesz representation theorem, by

$$\int_S y(s) d\mu_\varphi(s) = \varphi(y) \quad y \in \mathcal{C} = C_o(S). \quad (13)$$

Since for all (x, y) in $M(\mathcal{A})_+ \times \mathcal{C}_+$ we have $\varphi(xy) = \varphi(y^{1/2} x y^{1/2})$ it follows that

$$0 \leq \varphi(xy) \leq \|x\| \varphi(y). \quad (14)$$

Consequently the functional $y \rightarrow \varphi(xy)$ on \mathcal{C} defines a Radon measure on S dominated by a multiple of μ_φ , hence determined by a unique element $\Phi(x)$ in $L_{\mu_\varphi}^\infty(S)$. By linearization this defines a conditional expectation $\Phi: M(\mathcal{A}) \rightarrow L_{\mu_\varphi}^\infty(S)$ (i.e. a positive, unital module map) such that

$$\int y(s) \Phi(x)(s) d\mu_\varphi(s) = \varphi(yx), \quad y \in \mathcal{C} \quad x \in M(\mathcal{A}). \quad (15)$$

Inherent in this formulation is the fact that if $y \in \mathcal{C} = C_o(S)$, then $\Phi(y)$ is the natural image of y in $L_{\mu_\varphi}^\infty(S)$. In particular, $y(s) = \Phi(y)(s)$ for almost all s in S .

Observe now that since the C^* -algebra $C_o(\underline{I})$ is separable we can for almost every s in S define a Radon measure μ_s on \underline{I} by

$$\int_{\underline{I}} g(\underline{\lambda}) d\mu_s(\underline{\lambda}) = \Phi \left(\int_T a_t^* g(\underline{x}_t) a_t d\mu(t) \right) (s) \quad g \in C_o(\underline{I}). \quad (16)$$

As $\int_T a_t^* a_t d\mu(t) = \mathbf{1}$ this is actually a probability measure.

If we put $g_i(\underline{\lambda}) = \lambda_i$ for $1 \leq i \leq n$ then

$$\int_{\underline{I}} g_i(\underline{\lambda}) d\mu_s(\underline{\lambda}) = \Phi \left(\int_T a_t^* x_{it} a_t d\mu(t) \right) (s) = \Phi(y_i)(s) = y_i(s). \quad (17)$$

Since $y_i \in \mathcal{C}$ for all i we get by (17) – using the convexity of f in form of the standard Jensen inequality – that

$$\begin{aligned} f(\underline{y})(s) &= f(y(s)) = f(y_1(s), \dots, y_n(s)) \\ &= f \left(\int_{\underline{I}} g_1(\underline{\lambda}) d\mu_s(\underline{\lambda}), \dots, \int_{\underline{I}} g_n(\underline{\lambda}) d\mu_s(\underline{\lambda}) \right) \\ &\leq \int_{\underline{I}} f(g_1(\underline{\lambda}), \dots, g_n(\underline{\lambda})) d\mu_s(\underline{\lambda}) \\ &= \int_{\underline{I}} f(\underline{\lambda}) d\mu_s(\underline{\lambda}) = \Phi \left(\int_T a_t^* f(\underline{x}_t) a_t d\mu(t) \right) (s). \end{aligned} \quad (18)$$

Integrating over s now gives the desired result:

$$\begin{aligned} \varphi(f(\underline{y})) &= \int_S f(\underline{y})(s) d\mu_\varphi(s) \\ &\leq \int_S \Phi \left(\int_T a_t^* f(\underline{x}_t) a_t d\mu(t) \right) (s) d\mu_\varphi(s) \\ &= \int_T \int_S \Phi(a_t^* f(\underline{x}_t) a_t) (s) d\mu_\varphi(s) d\mu(t) \\ &= \int_T \varphi(a_t^* f(\underline{x}_t) a_t) d\mu(t) = \varphi \left(\int_T a_t^* f(\underline{x}_t) a_t d\mu(t) \right). \end{aligned} \quad (19)$$

Having proved the finite case, let us now assume that φ is unbounded, but lower semi-continuous on \mathcal{A}_+ and finite on the minimal dense ideal $K(\mathcal{A})$ of \mathcal{A} . Such functionals were termed C^* -integrals in [23] and [24]. This – by definition – means that $\varphi(x) < \infty$ if $x \in \mathcal{A}_+$ and $x = xe$ for some e in \mathcal{A}_+ , because $K(\mathcal{A})$ is the hereditary $*$ -subalgebra of \mathcal{A} generated by such elements, cf. [25, 5.6.1]. Restricting φ to \mathcal{C} we therefore obtain a unique Radon measure μ_φ on S such that

$$\int_S y(t) d\mu_\varphi(t) = \varphi(y) \quad y \in \mathcal{C}. \quad (20)$$

Inspection of the proof above now shows that the Jensen trace inequality still holds if only $f \geq 0$, even though ∞ may now occur in the inequality. \square

8. Remarks. The second condition in Theorem 7, that the elements y_i are mutually commuting, is not easy to verify. There are, however, a few cases that can be handled with ease. In the first we simply set $n = 1$, so that we obtain the one-variable extension of Theorem 2. This is done in Corollary 9. In the second case we let each a_t be a positive scalar and set $a_t^* a_t = \lambda(t)$. Then if $[x_{it}, x_{js}] = [x_{jt}, x_{is}]$ for all i, j, s and t (so, in particular $[x_{it}, x_{jt}] = 0$), the elements $y_i = \int_T x_{it} \lambda(t) d\mu(t)$ will form an abelian n -tuple. Thus in Corollary 11 we obtain an extremely strong version of the convexity of the trace function, proved in weaker forms in [8, 27, 18].

9. Corollary. *For each convex, continuous function f on an interval I , every bounded, weak* measurable field $(x_t)_{t \in T}$ in $\mathcal{A}_{\text{sa}}^I$ and every unital column field $(a_t)_{t \in T}$ in $M(\mathcal{A})$ we have the inequality*

$$\varphi \left(f \left(\int_T a_t^* x_t a_t d\mu(t) \right) \right) \leq \varphi \left(\int_T a_t^* f(x_t) a_t d\mu(t) \right) \quad (21)$$

for every positive functional φ that contains the element $y = \int_T a_t^* x_t a_t d\mu(t)$ in its centralizer.

If φ is unbounded, but lower semi-continuous and finite on the minimal dense ideal $K(\mathcal{A})$ of \mathcal{A} , the result still holds if $f \geq 0$. \square

For continuous fields this result was proved in [11, Theorem 4.1].

10. Corollary. *For each convex, continuous function f on a cube $\underline{I} = I_1 \times \cdots \times I_n$ in \mathbb{R}^n , each probability measure μ on a locally compact, metric space T and every n -tuple of bounded, weak* measurable fields $(x_{it})_{t \in T}$, where $x_{it} \in \mathcal{A}_{\text{sa}}^{I_i}$ for $1 \leq i \leq n$, such that $[x_{it}, x_{js}] = [x_{jt}, x_{is}]$ for all i, j, s and t we have*

$$\varphi \left(f \left(\int_T \underline{x}_t d\mu(t) \right) \right) \leq \varphi \left(\int_T (f(\underline{x}_t)) d\mu(t) \right), \quad (22)$$

for every positive functional φ on \mathcal{A} that contains the elements $y_i = \int_T x_{it} d\mu(t)$ in its centralizer. \square

Specializing to convex combinations (discrete probability measures) and traces we obtain the following version of Corollary 10:

11. Corollary. *For each convex, continuous function f on a cube $\underline{I} = I_1 \times \cdots \times I_n$ in \mathbb{R}^n , and every trace τ on a C^* -algebra \mathcal{A} the function*

$$\underline{x} \rightarrow \tau(f(\underline{x})) \quad (23)$$

is convex on the set of compatible pairs of abelian n -tuples $\underline{x} = (x_1, \dots, x_n)$ in \mathcal{A} with $\text{sp}(x_i) \in I_i$ for all i . \square

The condition in Theorem 7 that the elements x_{it} and x_{jt} commute mutually is also rather awkward. An easy and important solution to this problem is to assume from the outset that the C^* -algebra \mathcal{A} comes equipped with mutually commuting C^* -subalgebras $\mathcal{A}_1, \dots, \mathcal{A}_n$ and then require that $x_{it} \in \mathcal{A}_i$ for all i and t . Now the domain of definition of f is the convex set $\bigoplus_{i=1}^n (\mathcal{A}_i)_{\text{sa}}^{I_i}$ and we can state the Jensen trace inequality for f in ordinary terms

12. Corollary. *For each convex, continuous function f on a cube $\underline{I} = I_1 \times \cdots \times I_n$ in \mathbb{R}^n and every C^* -algebra \mathcal{A} with mutually commuting C^* -subalgebras $\mathcal{A}_1, \dots, \mathcal{A}_n$ we have the inequality*

$$\varphi \left(f \left(\int_T a_t^* \underline{x}_t a_t d\mu(t) \right) \right) \leq \varphi \left(\int_T a_t^* f(\underline{x}_t) a_t d\mu(t) \right) \quad (24)$$

for each bounded, weak* measurable field $(\underline{x}_t)_{t \in T} = ((x_{1t})_{t \in T}, \dots, (x_{nt})_{t \in T})$ in $\bigoplus_{i=1}^n (\mathcal{A}_i)_{\text{sa}}^{I_i}$, and every unital column field $(a_t)_{t \in T}$ in $M(\mathcal{A})$, provided that the elements $y_i = \int_T a_t^* x_{it} a_t d\mu(t)$ form an abelian n -tuple and the functional φ contains these elements in its centralizer.

If φ is unbounded, but lower semi-continuous and finite on the minimal dense ideal $K(\mathcal{A})$ of \mathcal{A} , the result still holds if $f \geq 0$. \square

This result generalizes both [11, Theorem 4.1] and [27, Theorem 2].

In the next case let the parameter space be $\mathbb{N}_n \times T$, where \mathbb{N}_n denotes the finite subset $\{1, 2, \dots, n\}$. So instead of the index t we now have the double index (j, t) . We then assume that $x_{ijt} = x_{it}$ if $j = i$ and that $x_{ijt} = 0$ if $j \neq i$. Furthermore we assume that $a_{jt} \in M(\mathcal{A}_j)$ for all j , so that we have the elements $b_j = \int_T a_{jt}^* a_{jt} d\mu(t)$ in $M(\mathcal{A}_j)$ with $\sum_{j=1}^n b_j = \mathbf{1}$. Note now that

$$y_i = \sum_{j=1}^n \int_T a_{jt}^* x_{ijt} a_{jt} d\mu(t) = \int_T a_{it}^* x_{it} a_{it} d\mu(t) \in \mathcal{A}_i, \quad (25)$$

so the commutativity condition is trivially satisfied. Consequently we have the following result:

13. Corollary. *For each convex, continuous function f on a cube $\underline{I} = I_1 \times \cdots \times I_n$, where $0 \in I_i$ for each i , every n -tuple of bounded, weak* measurable fields $(x_{it})_{t \in T} \subset (\mathcal{A}_i)_{\text{sa}}^{I_i}$ and every n -tuple of integrable column fields $(a_{it})_{t \in T}$ in $M(\mathcal{A}_i)$ with $\sum_{i=1}^n \int_T a_{it}^* a_{it} d\mu(t) = \mathbf{1}$ we have the inequality:*

$$\begin{aligned} & \varphi \left(f \left(\int_T a_{1t}^* x_{1t} a_{1t} d\mu(t), \dots, \int_T a_{nt}^* x_{nt} a_{nt} d\mu(t) \right) \right) \\ & \leq \varphi \left(\sum_{i=1}^n \int_T a_{it}^* f(0, \dots, x_{it}, \dots, 0) a_{it} d\mu(t) \right) \end{aligned} \quad (26)$$

for every positive functional φ that contains the elements $y_i = \int_T a_{it}^* x_{it} a_{it} d\mu(t)$ in its centralizer. \square

In the last case we again use the parameter space $\mathbb{N}_n \times T$ and take a_{jt} in $M(\mathcal{A}_j)$ for all j , but we now put $x_{ijt} = x_i$ constantly for some fixed x_i in $(\mathcal{A}_i)_{\text{sa}}$. Then

$$y_i = \sum_{j=1}^n \int_T a_{jt}^* x_i a_{jt} d\mu(t) = \int_T a_{it}^* x_i a_{it} d\mu(t) + (\mathbf{1} - b_i) x_i \in \mathcal{A}_i, \quad (27)$$

so again we have the desired relations. This gives the following result:

14. Corollary. *For each convex, continuous function f on a cube \underline{I} , every n -tuple \underline{x} with elements x_i in $(\mathcal{A}_i)_{\text{sa}}^{I_i}$ and every n -tuple of integrable column fields $(a_{it})_{t \in T}$ in $M(\mathcal{A}_i)$, with $\sum_{i=1}^n b_i = \mathbf{1}$, where $b_i = \int_T a_{it}^* a_{it} d\mu(t)$, we have the inequality:*

$$\begin{aligned} & \varphi \left(f \left(\int_T a_{1t}^* x_1 a_{1t} d\mu(t) + (\mathbf{1} - b_1)x_1, \dots, \int_T a_{nt}^* x_n a_{nt} d\mu(t) + (\mathbf{1} - b_n)x_n \right) \right) \\ & \leq \varphi \left(\sum_{i=1}^n \int_T a_{it}^* f(x_1, \dots, x_n) a_{it} d\mu(t) \right) \end{aligned} \quad (28)$$

for every positive functional φ that contains the elements $y_i = \int_T a_{it}^* x_i a_{it} d\mu(t) + (\mathbf{1} - b_i)x_i$ in its centralizer. \square

15. Monotonicity. We conclude the paper with some results about monotonicity of operator functions under a trace or a trace-like functional. The tendency is that if f is monotone increasing in each variable and $\underline{x} = (x_1, \dots, x_n)$ and $\underline{y} = (y_1, \dots, y_n)$ are abelian tuples, so that we have a chance to define $f(\underline{x})$ and $f(\underline{y})$, then $\varphi(f(\underline{x})) \leq \varphi(f(\underline{y}))$ if only $x_i \leq y_i$ for all i . This result may or may not be true in general. We can prove it when f is either convex or concave, or when \underline{x} and \underline{y} are compatible.

16. Theorem. *Let $f: \underline{I} \rightarrow \mathbb{R}$ be a continuous function on a cube $\underline{I} = [\alpha_1, \beta_1] \times \dots \times [\alpha_n, \beta_n]$ in \mathbb{R}^n , and assume that f is monotone increasing in each variable. Then for any two abelian n -tuples (x_1, \dots, x_n) and (y_1, \dots, y_n) of self-adjoint elements in a C^* -algebra \mathcal{A} with $\alpha_i \mathbf{1} \leq x_i \leq y_i \leq \beta_i \mathbf{1}$ for all i we have the inequality*

$$\varphi(f(\underline{x})) \leq \varphi(f(\underline{y})) \quad (29)$$

for any positive functional φ on \mathcal{A} that contains the elements x_1, \dots, x_n in its centralizer, provided that f is also convex. If instead f is concave the result holds if the elements y_1, \dots, y_n belong to the centralizer of φ .

Proof. Let $\mathcal{C} = C_o(S)$ denote the commutative C^* -subalgebra of \mathcal{A} generated by the x_i 's. As in the proof of Theorem 7 we then obtain a Radon measure μ_φ on S and a conditional expectation $\Phi: M(\mathcal{A}) \rightarrow L_{\mu_\varphi}^\infty(S)$ such that

$$\int_S z(s) \Phi(y)(s) d\mu_\varphi(s) = \varphi(zy), \quad z \in \mathcal{C} \quad y \in M(\mathcal{A}), \quad (30)$$

where $\Phi(z)(s) = z(s)$ almost everywhere on S for each z in \mathcal{C} .

Since \underline{I} is separable we can for almost every s in S define a probability measure μ_s on \underline{I} by the formula

$$\int_{\underline{I}} g(\underline{\lambda}) d\mu_s(\underline{\lambda}) = \Phi(g(\underline{y}))(s) \quad g \in C(\underline{I}). \quad (31)$$

If we set $g_i(\underline{\lambda}) = \lambda_i$ for each i , this means that

$$\int g_i(\underline{\lambda}) d\mu_s(\underline{\lambda}) = \Phi(g_i(\underline{y}))(s) = \Phi(y_i)(s). \quad (32)$$

Assume now that f - in addition to being monotone increasing - is also convex on the cube \underline{I} . Then, using that $f(\underline{x}) \in \mathcal{C}$ it follows that

$$\begin{aligned}
 \Phi(f(\underline{x}))(s) &= f(\underline{x})(s) = f(x_1(s), \dots, x_n(s)) \\
 &= f(\Phi(x_1)(s), \dots, \Phi(x_n)(s)) \leq f(\Phi(y_1)(s), \dots, \Phi(y_n)(s)) \\
 &= f\left(\int_{\underline{I}} g_1(\underline{\lambda}) d\mu_s(\underline{\lambda}), \dots, \int_{\underline{I}} g_n(\underline{\lambda}) d\mu_s(\underline{\lambda})\right) \\
 &\leq \int_{\underline{I}} f(g_1(\underline{\lambda}), \dots, g_n(\underline{\lambda})) d\mu_s(\underline{\lambda}) = \int_{\underline{I}} f(\lambda_1, \dots, \lambda_n) d\mu_s(\underline{\lambda}) \\
 &= \int_{\underline{I}} f(\underline{\lambda}) d\mu_s(\underline{\lambda}) = \Phi(f(\underline{y}))(s),
 \end{aligned} \tag{33}$$

where we used the monotonicity of f to obtain the first inequality sign in (33) and the convexity of f - in form of the usual Jensen inequality - to obtain the second inequality sign. Integrating over s now yields the desired result:

$$\begin{aligned}
 \varphi(f(\underline{x})) &= \int_S \Phi(f(\underline{x}))(s) d\mu_\varphi(s) \\
 &\leq \int_S \Phi(f(\underline{y}))(s) d\mu_\varphi(s) = \varphi(f(\underline{y})).
 \end{aligned} \tag{34}$$

If on the other hand we assume that f is a concave function we simply permute the rôles of the n -tuples \underline{x} and \underline{y} and let \mathcal{C} denote the C^* -subalgebra of \mathcal{A} generated by the y_i 's. The conditional expectation $\Phi: M(\mathcal{A}) \rightarrow L_{\mu_\varphi}^\infty(S)$ now satisfies that $\Phi(y_i)(s) = y_i(s)$ almost everywhere. Similarly we redefine the probability measures μ_s by the new formula $\int_{\underline{I}} g(\underline{\lambda}) d\mu_s(\underline{\lambda}) = \Phi(g(\underline{x}))(s)$, so that now

$$\int_{\underline{I}} g_i(\underline{\lambda}) d\mu_s(\underline{\lambda}) = \Phi(g_i(\underline{x}))(s) = \Phi(x_i)(s). \tag{35}$$

It follows as in (33) that we have the inequalities

$$\begin{aligned}
 \Phi(f(\underline{x}))(s) &= \int_{\underline{I}} f(\underline{\lambda}) d\mu_s(\underline{\lambda}) \\
 &= \int_{\underline{I}} f(\lambda_1, \dots, \lambda_n) d\mu_s(\underline{\lambda}) = \int_{\underline{I}} f(g_1(\underline{\lambda}), \dots, g_n(\underline{\lambda})) d\mu_s(\underline{\lambda}) \\
 &\leq f\left(\int_{\underline{I}} g_1(\underline{\lambda}) d\mu_s(\underline{\lambda}), \dots, \int_{\underline{I}} g_n(\underline{\lambda}) d\mu_s(\underline{\lambda})\right) \\
 &= f(\Phi(x_1)(s), \dots, \Phi(x_n)(s)) \leq f(\Phi(y_1)(s), \dots, \Phi(y_n)(s)) \\
 &= f(y_1(s), \dots, y_n(s)) = f(\underline{y})(s) = \Phi(f(\underline{y}))(s),
 \end{aligned} \tag{36}$$

where we now used the concavity of f to obtain the first inequality in (36). Integrating over s we again get the desired inequality (29). \square

17. Remarks. Evidently we may combine the two conditions in Theorem 16 to show that if f is an increasing function which admits a decomposition $f = f_+ + f_-$, where f_+ and f_- are both increasing and f_+ is convex whereas f_- is concave, then $\varphi(f(\underline{x})) \leq \varphi(f(\underline{y}))$ if $\underline{x} \leq \underline{y}$, provided that all the elements x_1, \dots, x_n and y_1, \dots, y_n belong to the centralizer of φ . However, such a decomposition, even approximately, is not possible in general, not even in the one-variable case. The reader may check that $\sin(t)$, for $-\pi/2 \leq t \leq \pi/2$ can not be approximated by any function $f = f_+ + f_-$, where f_+ is convex and f_- is concave, and both are increasing. In fact, $\|\sin - f\|_\infty > (2\pi)^{-2}$.

The simple function $f(s, t) = st$ is neither convex nor concave, but increases in each variable on the first quadrant. One proves by direct calculations that if x_1, y_1 and x_2, y_2 are positive elements in a C^* -algebra \mathcal{A} with $x_1 \leq x_2$, $y_1 \leq y_2$ then $\tau(x_1 y_1) \leq \tau(x_2 y_2)$ for every trace τ on \mathcal{A} . The simple argument relies on the cyclicity of the trace, which for two factors is equivalent to commutativity, but does not need the commutator equations $[x_1, y_1] = [x_2, y_2] = 0$, which we are prepared to insert to get abelian tuples. This particular argument fails for three factors, so that we are not able to decide whether the function $f(r, s, t) = rst$ is an increasing trace function on positive abelian triples.

Despite this setback one may still hope that the function $\underline{x} \rightarrow \tau(f(\underline{x}))$ is increasing on the set of abelian n -tuples, provided only that f is monotone increasing in each variable; at least when τ is a trace or a trace-like functional. Our last result, an extension of [27, Corollary 5], shows that this is true when the two abelian n -tuples are compatible. The proof uses the Fréchet differential as in [10].

18. Proposition. *Let f be a continuous function on a cube $\underline{I} = [\alpha_1, \beta_1] \times \dots \times [\alpha_n, \beta_n]$ in \mathbb{R}^n , and assume that f is increasing in each variable. Then for any two compatible, abelian n -tuples \underline{x} and \underline{y} in a C^* -algebra \mathcal{A} that satisfy $\alpha_i \leq x_i \leq y_i \leq \beta_i$ for all i , we have the inequality $\varphi(f(\underline{x})) \leq \varphi(f(\underline{y}))$ for any positive functional φ on \mathcal{A} that contains all the elements x_1, \dots, x_n and y_1, \dots, y_n in its centralizer.*

Proof. Put $\underline{h} = \underline{y} - \underline{x}$, and let $g(t) = \varphi(f(\underline{x} + t\underline{h})) = \varphi(f((1-t)\underline{x} + t\underline{y}))$ for $0 \leq t \leq 1$. (Note that this is well defined since \underline{x} and \underline{y} are compatible.) Then

$$\varphi(f(\underline{y})) - \varphi(f(\underline{x})) = \int_0^1 g'(t) dt, \quad (37)$$

provided, of course, that g is differentiable. However, working by approximation – extending f to a bounded increasing function on \mathbb{R}^n and convolving it with a suitable approximate unit for $L^1(\mathbb{R}^n)$ like $(\varepsilon/\pi)^{n/2} \exp(-\varepsilon \underline{s} \cdot \underline{s})$ – we may assume that f is extendable to a Schwartz function on \mathbb{R}^n , whence $f(\underline{u}) = \int_{\mathbb{R}^n} \exp(i(\underline{u} \cdot \underline{s})) \widehat{f}(\underline{s}) d\underline{s}$. Consequently, with $\underline{z} = \underline{x} + t\underline{h}$,

$$\begin{aligned} g'(t) &= \lim \varepsilon^{-1} \varphi(f(\underline{z} + \varepsilon \underline{h}) - f(\underline{z})) \\ &= \lim \varepsilon^{-1} \int_{\mathbb{R}^n} \varphi(\exp(i((\underline{z} + \varepsilon \underline{h}) \cdot \underline{s})) - \exp(i(\underline{z} \cdot \underline{s}))) \widehat{f}(\underline{s}) d\underline{s}. \end{aligned} \quad (38)$$

By the Dyson expansion of the operator function $b \rightarrow \exp(a + b)$ we have the expression $\lim \varepsilon^{-1} (\exp(\varepsilon(\underline{z} + \varepsilon \underline{h})) - \exp(\varepsilon \underline{z})) = \int_0^1 \exp(\varepsilon s \underline{z}) \underline{h} \exp((1-s)\varepsilon \underline{z}) ds$ and inserting

this in (38) we get

$$\begin{aligned}
 g'(t) &= \int_{\mathbb{R}^n} \varphi \left(\int_0^1 \exp(i(\underline{z} \cdot \underline{s})r) i(\underline{h} \cdot \underline{s}) \exp(i(\underline{z} \cdot \underline{s})(1-r)) dr \right) \widehat{f}(\underline{s}) d\underline{s} \\
 &= \int_{\mathbb{R}^n} \varphi (\exp(i(\underline{z} \cdot \underline{s})) i(\underline{h} \cdot \underline{s})) \widehat{f}(\underline{s}) d\underline{s} \\
 &= \sum_{k=1}^n \int_{\mathbb{R}^n} \varphi (\exp(i(\underline{z} \cdot \underline{s})) h_k) i s_k \widehat{f}(\underline{s}) d\underline{s} = \sum_{k=1}^n \varphi(f'_k(\underline{z}) h_k),
 \end{aligned} \tag{39}$$

where we used that the element $\underline{z} \cdot \underline{s}$, hence also $\exp(i(\underline{z} \cdot \underline{s})(1-r))$, is in the centralizer of φ . Since $f'_k \geq 0$ and $h_k \geq 0$ for all k it follows that $g' \geq 0$, whence $\varphi(f(\underline{x})) \leq \varphi(f(\underline{y}))$ by (37), as desired. \square

REFERENCES

- [1] Huzihiro Araki, *On an inequality of Lieb and Thirring*, Letters in Mathematical Physics **19** (1990), 167–170.
- [2] Huzihiro Araki & Frank Hansen, *Jensen's operator inequality for functions of several variables*, Proceedings of the American Mathematical Society **128** (2000), 2075–2084.
- [3] Jaspal Singh Aujla, *Matrix convexity of functions of two variables*, Linear Algebra and its Applications **194** (1993), 149–160.
- [4] Julius Bendat & Seymour Sherman, *Monotone and convex operator functions*, Transactions of the American Mathematical Society **79** (1955), 58–71.
- [5] Lawrence G. Brown & Hideki Kosaki, *Jensens inequality in semi-finite von Neumann algebras*, Journal of Operator Theory **23** (1990), 3–19.
- [6] Frank Hansen, *An operator inequality*, Mathematische Annalen **246** (1980), 249–250.
- [7] Frank Hansen, *Operator convex functions of several variables*, Publications of the Research Institute for Mathematical Sciences, Kyoto University **33** (1997), 443–463.
- [8] Frank Hansen, *Convex trace functions of several variables*, Linear Algebra and its Applications **341** (2002), 309–315.
- [9] Frank Hansen & Gert K. Pedersen, *Jensen's inequality for operators and Löwner's theorem*, Mathematische Annalen **258** (1982), 229–241.
- [10] Frank Hansen & Gert K. Pedersen, *Perturbation formulas for traces on C^* -algebras*, Publications of the Research Institute for Mathematical Sciences, Kyoto University **31** (1995), 169–178.
- [11] Frank Hansen & Gert K. Pedersen, *Jensen's operator inequality*, Bulletin of the London Mathematical Society, to appear.
- [12] Fumio Hiai, *Log-majorizations and norm inequalities for exponential operators*, Banach Center Publications **38** (1997), The Polish Academy of Sciences, Warszawa, 119–181.
- [13] Richard V. Kadison & John R. Ringrose, *Fundamentals of the Theory of Operator Algebras*, vol I-II, Academic Press, San Diego, 1986 (Reprinted by AMS in 1997).
- [14] Adam Korányi, *On some classes of analytic functions of several variables*, Transactions of the American Mathematical Society **101** (1961), 520–554.
- [15] Fritz Kraus, *Über convexe Matrixfunktionen*, Mathematische Zeitschrift **41** (1936), 18–42.
- [16] Elliott H. Lieb, *Convex trace functions and the Wigner-Yanase-Dyson conjecture*, Advances in Mathematics **11** (1973), 267–288.
- [17] Elliott H. Lieb, *The classical limit of quantum systems*, Communications in Mathematical Physics **31** (1973), 327–340.
- [18] Elliott H. Lieb & Gert K. Pedersen, *Multivariable convex trace functions*, Reviews in Mathematical Physics **14** (2002), 631–648.

- [19] Elliott H. Lieb & Mary Beth Ruskai, *Some operator inequalities of the Schwarz type*, Advances in Mathematics **26** (1974), 269–273.
- [20] Karl Löwner, *Über monotone Matrixfunktionen*, Mathematische Zeitschrift **38** (1934), 177–216.
- [21] John von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton Press, Princeton NJ, 1955.
- [22] Masanori Ohya & Dénes Petz, *Quantum Entropy and its Use*, Texts and Monographs in Physics, Springer Verlag, Heidelberg, 1993.
- [23] Gert K. Pedersen, *Measure theory for C^* -algebras, I-IV*, Mathematica Scandinavica **19**, **22**, **25**, **25** (1966,1968,1969,1969), 131–145, 63–74, 71–93, 121–127.
- [24] Gert K. Pedersen, *C^* -Integrals, an Approach to Non-Commutative Measure Theory*, Thesis, University of Copenhagen, 1971.
- [25] Gert K. Pedersen, *C^* -Algebras and their Automorphism Groups*, LMS Monographs **14**, Academic Press, San Diego, 1979.
- [26] Gert K. Pedersen, *Analysis Now*, Graduate Texts in Mathematics **118**, Springer Verlag, Heidelberg, 1989, reprinted 1995.
- [27] Gert K. Pedersen, *Convex trace functions of several variables on C^* -algebras*, Journal of Operator Theory, to appear.
- [28] Dénes Petz, *Spectral scale of self-adjoint operators and trace inequalities*, Journal of Mathematical Analysis and Applications **109** (1985), 74–82.
- [29] David Ruelle, *Statistical Mechanics*, The Mathematical Physics Monograph Series, Benjamin, New York, 1969.
- [30] Barry Simon, *The classical limit of quantum partition functions*, Communications in Mathematical Physics **71** (1980), 247–276.
- [31] Barry Simon, *The Statistical Mechanics of Lattice Gases*, vol I, Princeton University Press, Princeton, 1993.

INSTITUTE OF ECONOMICS, UNIVERSITY OF COPENHAGEN, STUDIESTRÆDE 6, DK-1455 COPENHAGEN K, DENMARK & DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COPENHAGEN, UNIVERSITETSPARKEN 5, DK-2100 COPENHAGEN Ø, DENMARK

E-mail address: frank.hansen @ econ.ku.dk & gkped @ math.ku.dk