SÉMINAIRE DE PROBABILITÉS (STRASBOURG)

YAO-ZHONG HU

Some operator inequalities

Séminaire de probabilités (Strasbourg), tome 28 (1994), p. 316-333

<http://www.numdam.org/item?id=SPS_1994__28__316_0>

SOME OPERATOR INEQUALITIES

by Yaozhong HU

Introduction. According to a suggestion of Prof. P.A. Meyer, I have collected in this paper a number of interesting inequalities concerning operators. I have tried to include useful results, choosing in the literature the simplest proofs.

The author thanks Prof. P.A. Meyer for his careful reading of preliminary versions of the paper, pointing out several mistakes and simplifying some proofs.

§1. Operator-monotone and operator-convex functions

We denote by $\mathcal{H}$ a complex Hilbert space with scalar product $< .,. >$. In this section we assume $\mathcal{H}$ is finite dimensional, leaving to the reader the extension to (bounded) operators on an infinite dimensional space. We assume the reader is familiar with elementary definitions as positivity, spectrum, trace, etc.

The definition of a continuous function which is monotone non-decreasing (abbreviated below to monotone) or convex on self-adjoint operators is clear, and recalled below. Such a function is of course monotone (convex) in the ordinary sense, but this is far from sufficient. The most important result is Löwner's theorem ([30], 1934) which gives an explicit form for the operator monotone (convex) functions.

We denote by T some interval of $\mathbb{R}$ and by $\mathrm{Sp}^{-1}(T)$ the set of all operators A whose spectrum $\mathrm{Sp}(A)$ is contained in T. These operators are self-adjoint, and the description of the set $\mathrm{Sp}^{-1}(T)$

$$(\,\mathrm{Sp}(A) \subset [a,b]\,) \Longleftrightarrow (\,\forall x : \|\,x\,\| = 1 \quad a \leq < Ax, x > \leq b\,)$$

shows that it is convex.

DEFINITION. *A real (Borel) function f defined on T is called* operator-monotone *if for (any finite-dimensional Hilbert space $\mathcal{H}$ and) any two operators $A \leq B \in \mathrm{Sp}^{-1}(T)$ on $\mathcal{H}$, we have $f(A) \leq f(B)$. It is called* operator-convex, *if for any two operators $A, B \in \mathrm{Sp}^{-1}(T)$, we have*

$$f(\lambda A + (1-\lambda)B) \leq \lambda f(A) + (1-\lambda) f(B) \quad , \quad (0 \leq \lambda \leq 1)\,. \tag{1.1}$$

If f is monotone or convex in T it is so in a smaller interval. On the other hand it is monotone or convex in the ordinary sense, hence locally bounded. Therefore it can be regularized by convolution in the usual way, remaining monotone (convex) on a slightly smaller interval. It will be convenient at some places to deal with C^1 or C^2 functions, but the results extend to full generality.

Here is the main theorem in this section. We break it into three statements for convenience.

THEOREM 1.1 (Löwner [30]). *For every operator-monotone function f on $(-1,1)$, there exists a (unique) probability measure μ on $[-1,1]$ such that*

$$f(t) = f(0) + f'(0) \int_{-1}^{1} \frac{t}{1-xt}\, d\mu(x) \cdot \tag{1.2}$$

THEOREM 1.2. *If f is operator-convex on $T =]-1,1[$ and $f(0)=0$, then $g(t) = f(t)/t$ is operator-monotone on T (and conversely).*

It follows that :

THEOREM 1.3. *For each operator-convex function f on $T = (-1,1)$, there exists a (unique) probability measure μ on $[-1,1]$ such that*

$$f(t) = f(0) + f'(0)\,t + \int_{-1}^{1} \frac{tx}{1-xt}\, d\mu(x) \cdot \tag{1.3}$$

It follows in particular that operator-monotone or convex functions are real analytic, and can be extended analytically outside T. But we will not discuss this important topic (see Donoghue [15]).

There are several proofs of this celebrated theorem, see [1], [7], [13], [15], [22], [26], [30], [37] etc. Three remarkable proofs due to Löwner [30], Bendat and Sherman [7] and Korányi [26]) are included in the book [15]. The proof we give here is adapted from the last remark in [22], where it is given as a simplification of Korányi's proof.

Example. We begin by an example of operator-monotone function which will show the sufficiency of (1.2). First take two operators $0 \le a \le b$, and $\lambda > 0$. Then we have $0 \le \lambda + a \le \lambda + b$, implying $I \le (\lambda+a)^{-1/2}(\lambda+b)(\lambda+a)^{-1/2}$. Taking inverses we get $I \ge (\lambda+a)^{1/2}(\lambda+b)^{-1}(\lambda+a)^{1/2}$ and finally the function $f(t) = 1/(\lambda+t)$ is operator-decreasing. Then $1 - \lambda f(t) = t/(\lambda+t)$ is operator-monotone on $T = [0,\infty[$ and the same follows for any homographic function which is increasing on T, and maps T into itself.

It follows that the mapping $(t-1)/(t+1)$ is a monotone increasing 1-1 mapping from $\mathrm{Sp}^{-1}(0,\infty)$ onto $\mathrm{Sp}^{-1}]-1,1[$. Carrying the result to the new interval we find that homographic increasing maps of $]-1,1[$ into itself are operator-monotone. This is the case for $t/(1-xt)$ with $x \in]-1,1[$, and it follows that (1.2) is indeed operator-monotone.

First characterization of monotone functions. Recall that the ***Hadamard product*** of two matrices (not operators!) $A = (a_{ij})$, $B = (b_{ij})$ is the matrix $A \circ B = (a_{ij}b_{ij})$. Shur's well known theorem asserts that the Hadamard product of a given matrix A with an arbitrary positive matrix B is positive if and only if A is positive.

We use in the whole paper the notation

$$D_t^k f(A,H) = \frac{d^k}{dt^k} f(A+tH)\ , \tag{1.4}$$

whenever the right hand side exists. When t is omitted it is meant that $t=0$.

The operator $f(A+tH)$ is well defined for A, H self-adjoint and f continuous. Which regularity of f implies that $f(A+tH)$ is, say, differentiable? Here is a simple result in that direction. We will need a similar result for second order derivatives, but the proof is given in a paper in the same volume.

LEMMA 1.4. *Let f be a function of class C^1 on some open interval T. Then for $A \in \mathrm{Sp}^{-1}(T)$ and arbitrary self-adjoint H the derivative $Df(A,H)$ exists. In a basis where A is a diagonal matrix with eigenvalues $(\lambda_1, \dots \lambda_n)$, this derivative is the Hadamard product $f^{[1]}(A) \circ H$ where $f^{[1]}(A)$ is the matrix with coefficients*

$$f^{[1]}(A)_{ij} = \begin{cases} (f(\lambda_i) - f(\lambda_j))/(\lambda_i - \lambda_j) & \text{if } \lambda_i \neq \lambda_j \\ f'(\lambda_i) & \text{if } \lambda_i = \lambda_j \end{cases} . \tag{1.5}$$

The function f is operator-monotone on T if and only if we have $Df(A,H) \geq 0$ for $A \in \mathrm{Sp}^{-1}(T)$ and $H \geq 0$.

PROOF. The crucial point in the proof is that, if $f(t) = t^k$, we may write the matrix (1.5) as

$$f^{[1]}(A)_{ij} = \sum\nolimits_{p=0}^{k-1} \lambda_i^p \lambda_j^{k-1-p} .$$

It is clear that $f(A+tH)$ is well defined for $A \in \mathrm{Sp}^{-1}(T)$ and t small enough. When $f(t) = t^k$ all derivatives exist

$$\frac{d}{dt}(A+tH)^k = \sum_{p=0}^{k-1}(A+tH)^p H (A+tH)^{k-1-p} .$$

We take $t = 0$ and represent operators by matrices in a basis where A is diagonal as stated. The elements of the last matrix are equal to $h_{ij} \sum_p \lambda_i^p \lambda_j^{k-1-p}$, and we get the corresponding Hadamard product. Thus the formula is proved for a polynomial. Then it is extended to $f \in C^1$ by approximation, because the Hadamard product is continuous on a finite dimensional Hilbert space. Note the Hadamard product is basis dependent, and no explicit formula is given in an arbitrary basis.

From the Hadamard product formula it follows that

$$\| Df(A,H) \| \leq C \| f' \|_A \| H \|$$

where C depends on the dimension of $\mathcal{H}$, $\| f' \|_A$ is the uniform norm of f' on the spectral interval of A. Then it also follows that (assuming f is defined on $\mathbb{R}$ for simplicity)

$$\| D_t f(A+tH) \| \leq C \| f' \| \| H \|$$

where $\| f' \|$ is computed on some large compact set. But then approximating f in C^1 norm by polynomials we see that $f(A+tH)$ is continuously differentiable. The last statement is nearly obvious.

COROLLARY. *A function f of class C^1 on T is operator monotone if and only if the kernel*

$$\widehat{f}(x,y) = \frac{f(y) - f(x)}{y - x} \qquad (f'(x) \text{ if } y = x)$$

is of positive type on T.

This follows at once from the Hadamard product formula, and Shur's theorem.

Proof of Theorem 1.1. The representation (1.2) will be proved for operator-monotone functions of class C^1. The extension of the representation to arbitrary functions will be left to the reader.

We may assume that $f(0) = 0$, so that $g(t) = f(t)/t$ is continuous. We will need the following property, to be proved later :

(1.6) the functions $f \pm g$ are operator-monotone.

Then let us sketch the proof. We consider the reproducing kernel Hilbert space E associated with the kernel $\widehat{f}(x,y)$ on $T =]-1,1[$. That is, we consider the linear space of measures with finite support in $]-1,1[$ with the scalar product $<\varepsilon_x, \varepsilon_y> = \widehat{f}(x,y)$, and we complete it. We now define a symmetric bilinear form on the space of finite measures by the formula $<\varepsilon_x, \varepsilon_y>' = \widehat{g}(x,y)$, and the fact that $f \pm g$ is monotone means that on this subspace we have $-<u,u> \leq <u,u>' \leq <u,u>$. Therefore there is an operator G of norm ≤ 1 such that $<u,v>' = <u, Gv>$. For $t \in]-1,1[$ define $u = (I - tG)\,\varepsilon_t$. Then we have

$$<u, \varepsilon_y> = <(I-tG)\varepsilon_t, \varepsilon_y> = \widehat{f}(x,y) - t\widehat{g}(t,y)$$
$$= \frac{f(t)-f(y)}{t-y} - t\frac{\frac{f(t)}{t} - \frac{f(y)}{y}}{t-y} = \frac{f(y)}{y} = g(y)$$

Therefore $u = g$ doesn't depend on t. Consider now the spectral decomposition $G = \int_{-1}^{1} s dE_s$ (there may be masses at ± 1) and introduce the positive measure $\mu(ds) = <u, dE_s u>$. For $|t| < 1$ $I - tG$ is invertible and we have $\varepsilon_t = (I-tG)^{-1}u$, therefore

$$\int_{-1}^{1} (1-ts)^{-1}\mu(ds) = <u, (I-tG)^{-1}u> = <u, \varepsilon_t> = g(t) = f(t)/t\ .$$

Multiplying by t we get Löwner's representation (1.2).

Operator convex functions.

The following result is theorem 2.1 from [22].

LEMMA 1.5. *Let* f *be continuous on* T, *any interval containing* 0. *The following properties are equivalent :*

1) f *is operator-convex and* $f(0) \leq 0$.

2) $f(a^* x a) \leq a^* f(x)\, a$ *for* $\|a\| \leq 1$, $x \in S(T)$.

3) $f(a^* x a + b^* y b) \leq a^* f(x)\, a + b^* y b$ *for* $a^* a + b^* b \leq I$, $x, y \in S(T)$.

4) *Like 2), but* a *is a projection.*

PROOF. We consider the operateurs on $\mathcal{H} \oplus \mathcal{H}$ given by

$$X = \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix} \quad , \quad U = \begin{pmatrix} a & b \\ c & -a^* \end{pmatrix} \quad , \quad V = \begin{pmatrix} a & -b \\ c & a^* \end{pmatrix}\ .$$

with $b = (1 - aa^*)^{1/2}$, $c = (1 - a^*a)^{1/2}$. Then U and V are unitary (this amounts to saying $ac - ba = 0 = ca^* - a^*b$, and the first equality suffices. We have $ah(a^*a) = h(aa^*)a$ for $h(x) = x^n$, then for a polynomial, and finally for $h(x) = (1-x)^{1/2}$.) Then we have

$$U^*XU = \begin{pmatrix} a^*xa & a^*xb \\ bxa & bxb \end{pmatrix} \quad , \quad V^*XV = \begin{pmatrix} a^*xa & -a^*xb \\ -bxa & bxb \end{pmatrix} .$$

If $\mathrm{Sp}(x) \subset T$, the same is true for X, U^*XU, V^*XV (we need here to know that $0 \in T$). If f is operator-convex, we have

$$\begin{aligned} \begin{pmatrix} f(a^*xa) & 0 \\ 0 & f(bxb) \end{pmatrix} &= f\begin{pmatrix} a^*xa & 0 \\ 0 & bxb \end{pmatrix} = f(\tfrac{1}{2}(U^*XU + V^*XV)) \\ &\le \tfrac{1}{2}(f(U^*XU) + f(V^*XV)) = \tfrac{1}{2}(U^*f(X)U + V^*f(X)V) \\ &= \tfrac{1}{2}U^*\begin{pmatrix} f(x) & 0 \\ 0 & f(0) \end{pmatrix}U + \tfrac{1}{2}V^*\begin{pmatrix} f(x) & 0 \\ 0 & f(0) \end{pmatrix}V \\ &\le \tfrac{1}{2}U^*\begin{pmatrix} f(x) & 0 \\ 0 & 0 \end{pmatrix}U + \tfrac{1}{2}V^*\begin{pmatrix} f(x) & 0 \\ 0 & 0 \end{pmatrix}V \\ &= \begin{pmatrix} a^*f(x)a & 0 \\ 0 & bf(x)b \end{pmatrix} . \end{aligned}$$

In particular, we get $f(a^*xa) \le a^*f(x)a$.

It is clear that 2) $\Longrightarrow$ 4). To show 2) $\Longrightarrow$ 3), apply 2) in $\mathcal{H} \oplus \mathcal{H}$ with

$$A = \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix} \quad , \quad X = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} .$$

It remains to show that 4) $\Longrightarrow$ 1). Given $x, y \in S(T)$ and $t \in [0,1]$ we put

$$X = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} \quad , \quad U = \begin{pmatrix} \sqrt{t} & -\sqrt{1-t} \\ \sqrt{1-t} & \sqrt{t} \end{pmatrix} \quad , \quad P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Then $X \in S(T)$, U is unitary, thus $U^*XU \in S(T)$, P is a projection. We write $f(PU^*XUP) \le Pf(U^*XU)P = PU^*f(X)UP$, whence

$$\begin{pmatrix} f(tx + (1-t)y) & 0 \\ 0 & f(0) \end{pmatrix} \le \begin{pmatrix} tf(x) + (1-t)f(y) & 0 \\ 0 & 0 \end{pmatrix}$$

and 1) follows.

LEMMA 1.6. *1) A function f is operator-monotone in $T =]0, \alpha[$ if and only if $h(t) = tf(t)$ is operator-convex in T.*

2) Let $T =]-1, 1[$ and f be operator-monotone. Then $(t+\lambda)f(t)$ is operator-convex in T for $\lambda \in T$.

3) Let $T =]-1, 1[$, assume f is operator-monotone and $f(0) = 0$, and put $g(t) = f(t)/t$. Then $f + \lambda g$ is operator-monotone for $t \in T$.

Proof. 1) Assume f is operator-monotone. Let P be a projection. For $X \in \mathrm{Sp}^{-1}(T)$ we have $X^{1/2}PX^{1/2} \leq X$, therefore

$$f(X^{1/2}PX^{1/2}) \leq f(X)$$
$$PX^{1/2}f(X^{1/2}PX^{1/2})X^{1/2}P \leq PX^{1/2}f(X)X^{1/2}P = PXf(X)P = Ph(X)P$$
$$PX^{1/2}X^{1/2}P\,f(PXP) \leq Ph(X)P$$

Here we have used the identity $f(X^{1/2}PX^{1/2})X^{1/2}P = X^{1/2}P\,f(PXP)$, which is proved first for $f(t) = t^k$ and extended to continuous f. The left hand side is $h(PXP)$ and it follows that h is operator-convex.

Conversely, assume h is operator-convex. Take $X, Y \in \mathrm{Sp}^{-1}(T)$ with $X \leq Y$. They are invertible, and the operator $A = Y^{-1/2}X^{1/2}$ has a norm ≤ 1. Writing that $h(A^*YA) \leq h(Y)$ proves that $f(X) \leq f(Y)$.

Let us prove 2). According to 1) $t \to tf(t-1)$ is operator-convex on $]0,2[$, hence $t \to (1+t)f(t)$ is operator convex on T. Applying this result to the operator-monotone function $t \to -f(-t)$ we have that $t \to -(t+1)f(-t)$ is operator-convex. But the mapping $t \to -t$ preserves convexity, thus $t \to (t-1)f(t)$ is operator-convex. Taking a convex combination we get that $(t+\lambda)f$ is operator convex for $\lambda \in [-1,1]$.

To prove 3) — which plays an essential role in the proof of Löwner's theorem — we will assume the operator-convex function f on $T =]-1,1[$ such that $f(0) = 0$ belongs to $\mathcal{C}^2$. Then $g \in \mathcal{C}^1$ and it is sufficient to prove that $Dg(A,H) \geq 0$ for $A \in \mathrm{Sp}^{-1}(T)$ and $H \geq 0$ small enough. Consider the following operators on $\mathcal{H} \oplus \mathcal{H}$

$$X = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 0 & \sqrt{H} \\ \sqrt{H} & 0 \end{pmatrix} \quad , \quad P = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$$

Then we will prove that

$$PD^2f(X,B)P = \begin{pmatrix} Dg(A,H) & 0 \\ 0 & 0 \end{pmatrix}$$

Since f is operator-convex, the left hand side is positive, and therefore the result will follow. To prove this formula, it is sufficient to deal with polynomials, and then with $f(t) = t^k$, $k > 0$. Then we have

$$D^2f(X,B) = \sum_{p+q+r=k-2} X^pBX^qBX^r \, .$$

On the other hand, if $q \neq 0$ we have $X^pBX^q = 0$. Therefore this sum reduces to $\sum_{p+r=k-2} X^pB^2X^r = Dg(X,B^2)$. Applying P on both sides we get the desired formula.

From 2) and 3), the mapping $f(t) \leftrightarrow g(t) = f(t)/t$ sets a 1–1 correspondence between operator-convex functions on $]-1,1[$ such that $f(0) = 0$ and operator-monotone functions. Thus the two remaining parts of the main theorem are proved.

Addition. One of the consequences of Löwner's theorem is that if $A \geq B \geq 0$ we have $A^r \geq B^r$ for $0 \leq r \leq 1$. Under special hypotheses on the exponents it is possible to prove that $(B^r A^s B^r)^t \geq B^{(2r+s)t}$ (Furota's inequality, see additional references at the end).

§2. Concavity related to the trace

The topics in this section are loosely related to those of the preceding one, by the use of Pick functions. A *Pick function* (Donoghue [15]) is a holomorphic function $f = u+iv$ in the upper half-plane which has a positive (i.e. ≥ 0) imaginary part v. A *Herglotz function* is defined in the same way, but in the unit disc (Epstein [16] uses this name also in the upper half-plane). Since v is a positive harmonic function it has a Poisson representation,

$$v(x,y) = \alpha + by + \int \frac{y\, d\theta(t)}{(t-x)^2 + y^2}$$

with $b \geq 0$, $\alpha \geq 0$, and θ is a positive measure on $\mathbb{R}$ such that $1/(1+t^2)$ is integrable. Let us write θ as $(1+t^2)\,\rho$, a bounded measure. We have

$$(1+t^2)\,\frac{y}{(t-x)^2+y^2} = \Im m\,\frac{1+tz}{t-z}$$

and therefore

$$f(z) = a + bz + \int_{-\infty}^{\infty} \frac{1+tz}{t-z}\, d\rho(t)$$

where ρ is a positive bounded measure, $\Im m(a) \geq 0$ and $b \geq 0$.

Suppose now ρ does not charge $T = (-1,1)$. Then $f(z)$ is meaningful for $z \in T$ real, and in fact can be continued analytically across T. Putting $t = 1/s$ we get from ρ a bounded measure τ on $[-1,1]$ which doesn't charge 0, and then we have

$$f(z) = a + bz + \int_{-1}^{1} \frac{(s+z)\, d\tau(s)}{1-sz} = a' + bz + \int_{-1}^{1} \frac{z\,(1+s^2)\, d\tau(s)}{1-sz}$$

where a' is a new constant with positive imaginary part (the $'$ will be omitted from now on). The unit masses at ± 1 yield the two functions $(1+z)/(1-z) = 1-2z/(1-z)$, $(z-1)/(z+1) = -1 + 2z/(1+z)$, which are operator-monotone functions. Note also that allowing τ to have a mass at 0 we may take the function bz into the integral, and then we have the Löwner representation — except that the constant a must be real in the operator-monotone case.

Now let us quote Epstein's theorem. Let $\mathcal{A}$ be a (complex) C^*-algebra; the mappings $\Re e(a)$ and $\Im m(a)$ are defined as for scalars. Let D be the open set of all elements $a \in \mathcal{A}$ such that

$$\text{for some } \theta \in [-\tfrac{\pi}{2}, \tfrac{\pi}{2}] \text{ and some } \varepsilon > 0,\ \Re e(e^{-i\theta} a) \geq \varepsilon\,.$$

Let D_s be the set of self-adjoint elements of D — they are positive and invertible.

Let f be a complex valued holomorphic function in D, positively homogeneous of degree s, $0 < s \leq 1$, which has the same property as Löwner's functions :

$$\text{If } a \in D \text{ and } \Im m(a) > 0 \ (< 0), \text{ then } \Im m(f(a)) \geq 0 \ (\leq 0).$$

By continuity, f is real on D_s.

THEOREM 2.1. *Under these hypotheses, the restriction of f to D_s is concave.*

EXAMPLES 2.2. Take for $\mathcal{A}$ the algebra of matrices. For B fixed, the following functions satisfy these hypotheses. In this way, Epstein unifies a number of results of Lieb.

1) $f(A) = \operatorname{Tr}\exp(B + \log A)$ (B self-adjoint).
2) $f(A) = \operatorname{Tr}(BA^{1/n}B)^n$ (n integer, $B \geq 0$).
3) $f(A) = \operatorname{Tr}(A^pBA^qB^*)^{1/(p+q)}$ (B arbitrary, $0 \leq p, q$, $p + q \leq 1$).
4) $f(A) = \operatorname{Tr}(A^pBA^qB^*)$

§3. Some trace inequalities

We will now prove some inequalities related to the trace of the operators. First we have

THEOREM 3.1. *If If A and B are two positive operators on some Hilbert space $\mathcal{H}$ (of finite dimension n). Let $0 \leq a_1 \leq \ldots \leq a_n$, $0 \leq b_1 \leq \ldots \leq b_n$ be their eigenvalues. Then for m positive integer*

$$\sum_{i=1}^{n} a_i^m b_{n-i}^m \leq \operatorname{Tr}(AB)^m \leq \operatorname{Tr}(A^mB^m) \leq \sum_{i=1}^{n} a_i^m b_i^m \tag{3.1}$$

REMARK 3.1. The inequality $\operatorname{Tr}(AB)^m \leq \operatorname{Tr}(A^mB^m)$ was proved by Lieb and Thirring [29]. The other parts were proved by Couteur [12] and by Bushell and Trustrum [10].

PROOF. We only prove that $\operatorname{Tr}(AB)^m \leq \operatorname{Tr}(A^mB^m)$ (from [29]). By a unitary transformation we may suppose that A is diagonal. Put $C = B^m \geq 0$ and $f(C) = \operatorname{Tr}(AC^{1/m})^m - \operatorname{Tr}(A^mC)$. Let $C = D + C'$ where D is the diagonal of C, and $C_\lambda = D + \lambda C' = \lambda C + (1-\lambda)D$ for $\lambda \in [0,1]$. Put $f(C_\lambda) = R(\lambda)$. We want to show that $R(1) \leq 0$. Now it is elementary to see that $R(0) \leq 0$ and by the preceding section, second example in 2.2 we know that $R(\lambda)$ is a concave function. Thus it lies below its tangent at 0 and it is sufficient to prove that

$$\frac{d}{d\lambda}\Big|_{\lambda=0} R(\lambda) = 0\,. \tag{3.2}$$

Recall that A and D are diagonal and C' has vanishing diagonal elements. So

$$\frac{d}{d\lambda}\Big|_{\lambda=0} \operatorname{Tr}(A^m(D + \lambda C')) = \operatorname{Tr}(A^mC') = 0$$

For the other term we have

$$\frac{d}{d\lambda}\Big|_{\lambda=0} \operatorname{Tr}[A(D+\lambda C')^{1/m}]^m = m \operatorname{Tr}[A D^{1/m}]^{m-1} \operatorname{Tr}[A \frac{d}{d\lambda}\Big|_{\lambda=0} (D+\lambda C')^{1/m}] .$$

We may compute this derivative (since D is diagonal) by the result of section 1 using Hadamard products, and see its diagonal elements vanish. Thus the trace of the product with A (diagonal) is 0.

By the same techniques as Lieb and Thirring's, one can prove

THEOREM 3.2. *If A and B are two selfadjoint operators, then*

$$\operatorname{Tr} e^{A+B} \leq \operatorname{Tr}(e^A e^B) . \tag{3.3}$$

This inequality was discovered by Golden [18] and Thompson [38] and further studied by Deift [14], Lenard [27], Thompson [39] etc.

ADDITION. The trace is used in the L^p norm of operators, $\| A \|_p = (\operatorname{Tr}[(A^*A)p/2])^{1/p}$. An excellent exposition of the results on these norms, including deep new inequalities (non-trivial even in the commutative case) can be found in a preprint by Ball, Carlen and Lieb (see the additional references).

§4. Inequalities concerning absolute value

The absolute value of a non-necessarily selfadjoint operator A is defined by $|A| = (A^*A)^{1/2}$. Any operator A has a polar decomposition $A = U|A|$ with U a partial isometry. Generally it is not true that $|A+B| \leq |A|+|B|$. An example (from B. Simon) is

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}$$
$$|A+B| = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}, \qquad |A|+|B| = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$$

However, we have ($\| . \|_2$ is the Hilbert-Schmidt norm, i.e. $\|A\|_2^2 = \operatorname{Tr}(A^*A)$).

THEOREM 4.1. *For any two (non-necessarily self-adjoint) operators A and B we have*

$$\| |A| - |B| \|_2 \leq \sqrt{2} \|A-B\|_2 \tag{4.1}$$

and when A and B are selfadjoint we have furthermore

$$\| |A| - |B| \|_2 \leq \|A-B\|_2 . \tag{4.2}$$

PROOF (from [5] and [6]). First we have the Schwarz inequality :

$$2|\operatorname{Tr}(SR)| \leq 2 \operatorname{Tr}(S^*S)^{1/2} \operatorname{Tr}(R^*R)^{1/2} \leq \|S\|_2^2 + \|R\|_2^2$$

Thus for any $X \geq 0$, $Y \geq 0$ and Q such that $\|Q\| \leq 1$ we have

$$\begin{aligned}4|\operatorname{Tr}(QXY)| &\leq 4\operatorname{Tr}((Y^{1/2}QX^{1/2})(X^{1/2}Y^{1/2}))\\ &\leq 2(\operatorname{Tr}(X^{1/2}Q^*YQX^{1/2}) + \operatorname{Tr}(Y^{1/2}XY^{1/2}))\\ &\leq 2\operatorname{Tr}(XQ^*YQ) + 2\operatorname{Tr}(XY)\\ &\leq \operatorname{Tr}(Q^*Y^2Q) + \operatorname{Tr}(QX^2Q^*) + 2\operatorname{Tr}(XY)\\ &\leq \operatorname{Tr}(X^2+Y^2+XY+YX) = \operatorname{Tr}((X+Y)^2)\end{aligned}$$

Now let $A = U|A|$ and $B = V|B|$ be the polar decompositions of A and B. Applying the above inequality for $X = |A|$, $Y = |B|$ and $Q = V^*U$ we have

$$\begin{aligned}2\|A-B\|_2^2 &= 2\operatorname{Tr}(|A|^2+|B|^2) - 2\Re e\operatorname{Tr}(|B|V^*QU|A|)\\ &\geq 2\operatorname{Tr}(|A|^2+|B|^2) - \operatorname{Tr}(|A|+|B|)^2\\ &= \operatorname{Tr}(|A|-|B|)^2 = \||A|-|B|\|_2^2.\end{aligned}$$

To prove (4.2) it suffices to prove

$$\operatorname{Tr}(|A||B|) \geq \operatorname{Tr}(AB)$$

First we may suppose that A is diagonal $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Note by b_{ii} the diagonal elements of B and c_{ii} those of $|B|$. Since B is self-adjoint we have $|B| \geq B \geq -|B|$, hence $c_{ii} \geq |b_{ii}|$. Consequently,

$$\begin{aligned}\operatorname{Tr}(AB) &= \sum \lambda_i b_{ii} \leq \sum |\lambda_i||b_{ii}|\\ &\leq \sum |\lambda_i| c_{ii} = \operatorname{Tr}|A||B|.\end{aligned}$$

THEOREM 4.2. *If $f(t)$ is a non-negative operator-monotone function on $[0,\infty)$, then for $A, B \geq 0$ we have*

$$\operatorname{Tr}[f(A) - f(B)] \leq \operatorname{Tr}[f(|A-B|)]\,. \tag{4.3}$$

PROOF (due to Ando [2]). Let $C = A - B$. Since C is self-adjoint we have $C \leq |C|$ hence $0 \leq A = B + C \leq B + |C|$. So $f(A) \leq f(B+|C|)$ and

$$f(A) - f(B) \leq f(B+|C|) - f(B)$$

and we are reduced to the case where C is positive. By Löwner's theorem (for the interval $[0,\infty[$) we have

$$f(t) = \alpha + \beta t + \int_{-1}^{1} \frac{xt}{1+xt} d\mu(x) \tag{4.4}$$

for some $\alpha, \beta \geq 0$ and a positive measure μ on $[0,\infty)$ such that $\int_0^\infty \frac{x}{1+x}\mu(dx) \leq \infty$. So it is sufficient to consider the function $f(t) = t/(x+t)$, and we may take $x = 1$. Since $f(t) = 1 - (1+t)^{-}1$ we are reduced to proving that for $B, C \geq 0$

$$\operatorname{Tr}[(B+1)^{-1} - (B+C+1)^{-1}] \leq \operatorname{Tr}[1 - (C+1)^{-1}]\,. \tag{4.5}$$

Put $1/\sqrt{I+B)} = H \leq 1$. Then $HDH \leq D$ for $D \geq 0$ and the operator on the left is

$$H(1 - \frac{1}{1+HCH})H \leq 1 - \frac{1}{1+HCH} \leq 1 - \frac{1}{1+C} \cdot$$

Then applying Tr we get the result.

5. Grothendieck's Inequality

Grothendieck's inequality [19] has been the starting point of the modern theory of Banach spaces (see the paper by Lindenstrauss and Pelczynsky in the additional references. For the history see [32]). It has been the subject of many publications ([9], [20], [21], [25], [31], [32], [33]...). Here we present Krivine's proof which uses a probability language and gives the best (known) estimate for the real case Grothendieck constant. Then we point out an equivalent form which can be extended to the non-commutative case.

The elementary form of Grothendieck's inequality is the following : let T be a finite set, and let $\mathcal{C} = \mathcal{C}(T)$ be the finite dimensional space of real valued functions on T with the *sup* norm.

THEOREM 5.1. *Let u be a bilinear form on $\mathcal{C} \times \mathcal{C}$ of norm ≤ 1*

$$u(a,b) = \sum_{s,t \in T} u(s,t)\, a(s) b(t) \quad \textit{with} \quad |u(a,b)| \leq \sup_s |a(s)| \sup_t |b(t)| \,. \tag{5.1}$$

There exists an absolute constant K (called the real Grothendieck constant) such that, if A, B now take values in a real Hilbert space $\mathcal{H}$

$$\Big| \sum_{s,t \in T} u(s,t) < A(s)\,, B(t) > \Big| \leq K \sup_s \| A(s) \| \sup_t \| B(t) \| \,. \tag{5.2}$$

Replacing real functions and Hilbert spaces by complex ones defines the complex Grothendieck constant (which is smaller). The exact value of these constants is not known, though the estimates are rather precise.

The general case is as follows :

THEOREM 5.1$'$. *Let S and T be two compact spaces, u a bilinear form of norm ≤ 1 on $\mathcal{C}(S) \times \mathcal{C}(T)$. Let $\mathcal{H}$ be a real Hilbert space and u be extended to a bilinear form on $(\mathcal{C}(S) \otimes \mathcal{H}) \times (\mathcal{C}(T) \otimes \mathcal{H})$ as*

$$u(a \otimes h, b \otimes k) = u(a,b) < h,k > \,.$$

Then u can be extended to $\mathcal{C}(S,\mathcal{H}) \times \mathcal{C}(T,\mathcal{H})$ in such a way that

$$|u(A,B)| \leq K \sup_s \| A(s) \| \sup_t \| B(t) \| \,. \tag{5.2$'$}$$

The fact that S may be different from T is not important. Taking a basis for $\mathcal{H}$, another way of stating (5.2′) is : given finite families $a_i \in \mathcal{C}(S), b_i \in \mathcal{C}(T)$, we have

$$(5.2'') \qquad |\sum_i u(a_i, b_i)| \leq K \sup_s (\textstyle\sum_i a_i(s)^2)^{1/2} \sup_t (\textstyle\sum_i b_i(t)^2)^{1/2} .$$

We will prove the elementary form of the theorem. Working on a finite set instead of compact spaces will preserve the essential idea of the proof, but spare some technical details. To help the reader imagine the general proof, we put between braces a few words which are useless in the finite case. One can also deduce the general case from the elementary case.

We will need some preliminary explanations.

1) Since T is finite, $\mathcal{C}\otimes\mathcal{C}$, the set of all functions $F(s,t) = \sum_i a_i(s)\, b_i(t)$, $(a_i, b_i \in \mathcal{C})$ is the set of all functions of two variables. If T were compact, it would merely be dense in $\mathcal{C}(T \times T)$.

2) There is a norm on the space $\mathcal{C} \otimes \mathcal{C}$, called the *projective norm,* such that the conjugate space is that of (bounded) bilinear functionals on $\mathcal{C} \times \mathcal{C}$ with its usual norm. It can be computed as

$$(5.3) \qquad \| F \|_\pi = \inf \textstyle\sum_i \| a_i \| \, \| b_i \|$$

over all decompositions $F(s,t) = \sum_i a_i(s)\, b_i(t)$.

3) Let us define another norm on functions of two variables. We denote by $\mathcal{E}$ the set of all functions $F(s,t)$ that can be represented as

$$F(s,t) = \langle X(s), Y(t) \rangle$$

for a pair of (continuous) functions X, Y taking values in some Hilbert space $\mathcal{H}$, and such that for all s, t

$$(5.4) \qquad \| X(s) \| = \| Y(t) \| = \rho \quad \text{(some constant)} \quad .$$

The smallest possible value of ρ^2 is denoted $\| F \|_*$. Of course it is larger than the uniform norm of F. We prove a few elementary facts.

— Given $F \in \mathcal{E}$, then $-F \in \mathcal{E}$ (change $X \to -X$) and $t^2 F \in \mathcal{E}$ (change $X \to tX$ and $Y \to tY$).

— Given $F, F' \in \mathcal{E}$, then $F + F' \in \mathcal{E}$ (use $X \oplus X', Y \oplus Y'$ taking values in $\mathcal{H} \oplus \mathcal{H}'$).

— Given $F, F' \in \mathcal{E}$, then $FF' \in \mathcal{E}$ (use $X \otimes X', Y \otimes Y'$ taking values in $\mathcal{H} \otimes \mathcal{H}'$).

It is not difficult to see that $\| . \|_*$ is a norm on $\mathcal{E}$. To cover the general case it is necessary also to prove that $\mathcal{E}$ is complete (hence a Banach algebra). This is one of the points we may skip.

— Let $F(s,t) = \langle X(s), Y(t) \rangle$ with $\| X(s) \| \leq \sigma$ and $\| Y(t) \| \leq \tau$. Then $F \in \mathcal{E}$ and $\| F \|_* \leq \sigma\tau$.

To see this, first add to $\mathcal{H}$ two vectors ξ, η orthogonal to each other and to $\mathcal{H}$, and replace X, Y by $X(s) + u(s)\,\xi$, $Y(t) + v(t)\,\eta$ so that the norms are increased to a, b without changing $\langle X, Y \rangle$. Then replace X, Y by $X\sqrt{b/a}, Y\sqrt{a/b}$ so that the norms are equal.

— As a consequence, taking $\mathcal{H} = \mathbb{R}$, $F(s,t) = a(s)\,b(t)$ belongs to $\mathcal{E}$ with $\| F \|_* \leq \sup_s |a(s)| \sup_t |b(t)|$.

Then it follows that $\mathcal{C} \otimes \mathcal{C} \subset \mathcal{E}$ and the $\|\,.\,\|_\pi$ norm is larger than $\|\,.\,\|_*$.

More generally, we get the following result, which will be useful later on :

$$\| \textstyle\sum_i a_i \otimes b_i \|_* \leq \| (\sum_i a_i^2)^{1/2} \|_\infty \| (\sum_i b_i^2)^{1/2} \|_\infty \, . \tag{5.5}$$

If we remember now that $\|\,.\,\|_\pi$ is the conjugate norm of the usual norm of bilinear functionals, the Grothendieck inequality can be read as :

Theorem 5.2. *We have on* $\mathcal{C} \otimes \mathcal{C}$

$$\| F \|_\pi \leq K \| F \|_* \, . \tag{5.6}$$

with $K \leq \pi/(2\log(1+\sqrt{2})) = 1,782\ldots$

Proof. Krivine's argument relies on the following probabilistic result : *let X, Y be two normalized real jointly Gaussian random variables. Then we have*

$$\mathbb{E}[\operatorname{sign} X \ \operatorname{sign} Y] = \frac{2}{\pi} \arcsin \mathbb{E}[XY] \, . \tag{5.7}$$

It is proved as follows (from [32]). Put $\theta_0 = \mathbb{E}[XY] \in [\pi/2, \pi/2]$. Then we may write, denoting by Z a normalized Gaussian r.v. independent from X

$$\mathbb{E}[\operatorname{sign} X \ \operatorname{sign} Y] = \int \operatorname{sign} X \operatorname{sign}(X \sin\theta_0 + Z\cos\theta_0)\, e^{-(x^2+y^2)/2} \frac{dx\,dy}{2\pi}$$

Computing this in polar coordinates we get

$$\int\!\!\int \operatorname{sign}(\cos\theta) \ \operatorname{sign}(\sin(\theta+\theta_0))\, e^{-r^2/2} r\,dr \, \frac{d\theta}{2\pi}$$
$$= \frac{1}{2\pi}\int_0^{2\pi} \operatorname{sign}(\cos\theta) \ \operatorname{sign}((\theta+\theta_0))\, d\theta = \frac{2\theta_0}{\pi} \, .$$

Next, we remark that any Hilbert space $\mathcal{H}$ is isomorphic to a Hilbert space of Gaussian random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $F(s,t)$ belong to $\mathcal{C}\otimes\mathcal{C}$ with a norm $\| F \|_* \leq 1$ (in particular, the uniform norm of F is at most 1). Then since $\mathcal{E}$ is a Banach algebra, for $b > 0$ $\sin(aF(s,t))$ belongs to $\mathcal{E}$ with a norm

$$\| \sin(aF) \|_* \leq \sinh a \, .$$

Take $a < \log(1+\sqrt{2}) < 1$, so that $\sinh a < 1$. Then $\sin(aF)$ has a representation using normalized Gaussian r.v.'s

$$\sin(aF(s,t)) = \mathbb{E}[X_s Y_t]$$

Since $a < 1$, $a|F| \leq \pi/2$ and we can invert, computing $F(s,t)$ as

$$\frac{1}{a}\arcsin\sin(aF(s,t)) = \frac{\pi}{2a}\mathbb{E}[\operatorname{sign} X_s \ \operatorname{sign} Y_t] = \frac{\pi}{2a}\int \operatorname{sign} X_s(\omega) \ \operatorname{sign} Y_t(\omega)\, d\mathbb{P}(\omega) \, .$$

On the right hand side we have an average of functions of (s,t) depending on ω, whose norm $\|\,.\,\|_\pi$ is smaller than 1. Thus the norm of F is at most $\pi/2a$ and the theorem is proved.

Remarks and extensions

We first indicate a consequence of theorem 5.1′. Let f_i, g_i be elements of $\mathcal{C}(S), \mathcal{C}(T)$ (finitely many) with $\| g_i \|_\infty \le 1$ and apply (5.2″) to the functions $a_i = f_i$, $b_i = \alpha_i g_i$ where $\sum_i \alpha_i^2 = 1$, and take a supremum over (α_i). Then we get

$$(\textstyle\sum_i u(f_i, g_i)^2)^{1/2} \le K \sup_s (\sum_i f_i(s)^2)^{1/2} .$$

Take now a *sup* over (g_i). Calling U the operator from $\mathcal{C}(S)$ to $\mathcal{C}(T)^*$ associated with u, this can be written

$$(\sum_i \| U f_i \|^2)^{1/2} \le K \sup_\mu (\sum_i \mu(f_i)^2)^{1/2}$$

where μ ranges over the unit ball of $\mathcal{C}(S)^*$. In the technical language of Banach spaces, one says that U is a *2-summing operator*.

There are other versions of Grothendieck's theorem. The following one (Theorem 5.5 of [32]) is rather striking. The constant K is the Grothendieck constant.

THEOREM 5.3. *Let S and T be two compact spaces, u a bounded bilinear functional on $\mathcal{C}(S) \times \mathcal{C}(T)$, of norm ≤ 1. There exist two probability measures λ on S and μ on T such that*

$$|u(f,g)| \le K \lambda(f^2)^{1/2} \mu(g^2)^{1/2} . \tag{5.8}$$

Let us show how this implies Theorem (5.1′). Let $(f_i), (g_i)$ be finite sequences of elements of $\mathcal{C}(S)\mathcal{C}(T)$. Then applying the Schwarz inequality to (5.8), we have

$$\frac{1}{K} | \sum_i u(f_i, g_i) | \le \sum_i \lambda(f_i^2)^{1/2} \mu(g_i^2)^{1/2} \le (\sum_i \lambda(f_i^2))^{1/2} (\sum_i \mu(g_i^2))^{1/2}$$

We may replace each integral by a *sup* since λ, μ are probability laws, and we get (5.2″).

The proof that Theorem (5.2) implies Theorem (5.3) is due to Amemiya and Shiga (*Kodai Math. Sem. Reports,* **9**, 1957) and very interesting. We just sketch it. We begin with the case of $S = T$. We do not use Grothendieck's theorem, but only the assumption (which follows from $\| u \|_* \le 1$, see (5.5)) that for $g_i \in \mathcal{C}(S)$

$$| \textstyle\sum_i u(g_i, g_i) | \le \| \sum_i g_i^2 \|_\infty . \tag{5.9}$$

Then Theorem 5.2 links this property to the hypothesis of Theorem 5.3. We will deduce from (5.9) and the Hahn-Banach theorem the existence of a probability law μ such that

$$|u(f,f)| \le \int f^2 \, d\mu . \tag{5.10}$$

For $f \in \mathcal{C}(S)$, we put

$$p(f) = \inf \left(\| f + \sum_i g_i^2 \| - | \sum_i u(g_i, g_i) | \right) .$$

the *inf* ranging over finite families $(g_i) \in \mathcal{C}(S)$. Then the assumption (5.9) implies that $-\| f \| \leq p(f) \leq \| f \|$. Next p is shown to be a sublinear function ($p(tf) = tp(f)$ for $t > 0$, $p(f+g) \leq p(f) + p(g)$) and by the Hahn-Banach theorem there exists a linear functional μ dominated by p. The obvious relation $p(-f^2) \leq -|u(f,f)|$ then shows μ is a positive measure, which then is shown to have a mass ≤ 1 and to satisfy (5.10).

It remains to dominate $u(f,g)$ instead of $u(f,f)$. To this end we put $R = S + T$ and define a bilinear form

$$v(f+g, f'+g') = \frac{1}{2}(u(f,g') + u(f',g)) \qquad (f, f' \in \mathcal{C}(S)\,,\, g, g' \in \mathcal{C}(T))\,.$$

which from (5.5) is easily shown to satisfy (5.9). Therefore there is a probability measure on R (i.e. a pair of probability measures λ on S and μ on T, and a number $t \in [0,1]$) such that

$$| \sum_i v(f_i + g_i, f_i + g_i) \leq t\lambda(\sum_i f_i^2) + (1-t)\mu(\sum_i g_i^2) .$$

Replace f_i by cf_i, g_i by g_i/c, the left hand side does not change. Then minimize over c to get

$$2\sqrt{t(1-t)}\,\lambda(\sum_i f_i^2))^{1/2}\,\mu(\sum_i g_i^2))^{1/2} ,$$

and conclude since $2\sqrt{t(1-t)} \leq 1$.

On the other hand, Theorem 5.3 can be generalized to the non-commutative analogue of spaces $\mathcal{C}(T)$, *i.e.* C^*-algebras. This answered a conjecture of Grothendieck. The first result in this direction was Pisier [31], under a special assumption on u, which was lifted by Haagerup [20]. The result is sharp, and thus the "non-commutative complex Grothendieck constant" is known, while the commutative one is not.

THEOREM 5.4. *Let $\mathcal{A}$ and $\mathcal{B}$ be two C^*-algebras, u a bounded bilinear form on $\mathcal{A} \times \mathcal{B}$, of norm 1. There exist four states λ_1, λ_2 on $\mathcal{A}$, μ_1, μ_2 on $\mathcal{B}$, such that $|u(a,b)|$ is dominated by*

$$(\lambda_1(a^*a) + \lambda_2(aa^*))^{1/2} + (\mu_1(b^*b) + \mu_2(bb^*))^{1/2} .$$

ADDITION. In his proof [31] of the non-commutative extension of Gr.'s theorem, Pisier has a very interesting lemma, concerning a bounded linear operator u from $\mathcal{A}$ to $\mathcal{B}$:

THEOREM. Given elements a_i of $\mathcal{A}$, we have (C being a universal constant)

$$\| (\sum_i u(a_i)^* u(a_i))^{1/2} \| \leq C\,|u|\,\sup(\| (\sum_i a_i^* a_i)^{1/2} \|\, \| (\sum_i a_i a_i^*)^{1/2} \|) .$$

The proof has been simplifed by Haagerup (additional references), and the result has been extended in a preprint by Haagerup and Pisier.

REFERENCES

[1] T. ANDO. *Topics on Operator Inequalities*, mimeographed lecture notes, Hokkaido Univ., Sapporo, 1978.

[2] T. ANDO. Comparison of the norms $\| f(A) - f(B) \|$ and $\| f(|A - B|) \|$, *Math. Z.*, **197**, 1988, 403–409.

[3] T. ANDO. On some operator inequalities, *Math. Ann.*, **279**, 1987, 157–159.

[4] T. ANDO. Concavity of certain maps on positive definite matrices and applications to Hadamard products, *Linear Alg. Appl.*, **26**, 1979, 203–241.

[5] H.ARAKI. On quasi-free states of CAR and Bogoliubov automorphisms, *Publ. RIMS Kyoto Univ.*, **6**, 1971, 385–442.

[6] H. ARAKI and S. YAMAGAMI. An inequality for Hilbert–Schmidt norm, *Comm. Math. Phys.*, **81**, 1981, 89–96.

[7] J. BENDAT and S. SHERMAN. Monotone and convex operator functions, *Trans. Amer. Math. Soc.*, **79**, 1955, 58–71.

[8] F.A. BEREZIN. Convex operator functions, *Math. Sbornik*, **17**, 1972, 269–277.

[9] R.C. BLEI. An elementary proof of the Grothendieck inequality, *Proc. Amer. Math. Soc.*, **100**, 1987, 58–60.

[10] P.J. BUSHELL and G.B. TRUSTRUM. Trace inequalities for positive definitive power products, *Linear Alg. Appl.*, **132**, 1990, 173–178.

[11] M. BREITENECKER and H.R. GRÜMM. Note on trace inequalities, *Comm. Math. Phys.*, **26**, 1972, 276–279.

[12] K.J. COUTEUR. Representation of the function $\mathrm{Tr}(\exp(A - \lambda B))$ as a Laplace transformation with positive weight and some matrix inequalities, *J. Phys. A.*, **13**, 1980, 3147–3149.

[13] G. DAVIS. Notions generalizing convexity for functions defined on space of matrices, *Proc. Symposia in Pure Math. : "Convexity"*, **1963**, 159–170.

[14] P. DEIFT. Applications of a commutation formula, *Duke Math. J.*, **45**, 1978, 267–310.

[15] W. DONOGHUE. *Monotone Matrix Functions and Analytic Continuation*, Springer, 1974.

[16] H. EPSTEIN. Remarks on two theorems of E. Lieb, *Comm. Math. Phys.*, **31**, 1973, 317–325.

[17] S. FRIEDLAND and M. KATZ. On a matrix inequality, *Linear Alg. Appl.*, **85**, 1987, 185–190.

[18] S. GOLDEN. Lower bounds for the Herglotz functions, *Phys. Rev.*, **137 B**, 1965, 1127–1128.

[19] A. GROTHENDIECK. Résumé de la théorie métrique des produits tensoriels topologiques, *Bol. Mat. Sao Paulo*, **8**, 1956, 1–79.

[20] U. HAAGERUP. The Grothendieck inequality for the bilinear forms on C^*-algebras, *Adv. in Math.*, **56**, 1985, 93–116.

[21] U. HAAGERUP. A new upper bound for the complex Grothendieck constant, *Israel J. Math.*, **60**, 1987, 199–224.

[22] F. HANSEN and G.K.PEDERSEN. Jensen's inequality for operators and Löwner's theorem. *Math. Ann.*, **258**, 1982, 229–241.

[23] R.A. HORN. The Hadamard product, *Proc. Symp. in Applied Math.*, AMS, **40**, 1990, 87–169.

[24] Y.Z. HU. Calculs formels sur les E.D.S. de Stratonovitch, *Sém. Prob. ' XXIV*,s LNM **1426**, Springer, 1990, 453–460.

[25] J.L. KRIVINE. Constante de Grothendieck et fonctions de type positif sur les sphères, *Adv. Math.*, **31**, 1979, 16–30.

[26] A. KORÁNYI. On a theorem of Löwner and its connection with resolvents of selfadjoint transformations, *Acta Sci. Math. (Szeged)*, **17**, 1956, 63–70.

[27] A. LENARD. Generalization of the Golden-Thompson inequality $\mathrm{Tr}(e^A e^B) \geq \mathrm{Tr}(e^{A+B})$, *Indiana Univ. Math. J.*, **21**, 1971, 457–467.

[28] E.H. LIEB. Convex trace functions and the Wigner-Yanase-Dyson conjecture, *Adv. in Math.*, **11**, 1973, 267–288.

[29] E.H. LIEB and W.E. THIRRING. Inequalities for the moments of the eigenvalues of the Schrödinger Hamiltonian and their relation to Sobolev inequalities, *Essays in honour of V. Bargman,* Princeton, N.J., 1976.

[30] K. LÖWNER. Über monotone Matrixfunctionen, *Math. Z.*, **38**, 1934, 117–216.

[31] G. PISIER. Grothendieck's theorem for noncommutative C^*-algebras, with an appendix on Grothendieck's constants, *J. Funct. Anal.*, **29**, 1978, 397–415.

[32] G. PISIER. *Factorization of Linear Operators and Geometry of Banach Spaces,* CBMS Regional Conf. Series in Math., **60**, AMS, 1986.

[33] R. REITZ. A proof of the Grothendieck inequality, *Israel J. Math.*, **19**, 1974, 271–276.

[34] M. ROSENBLUM and J. ROVNYAK. *Hardy Classes and Operator Theory*, Oxford Univ. Press, 1985.

[35] B. SIMON. *Trace Ideals and Their Applications*, London Math. Soc. Lect. Notes Series 35, Cambridge Univ. Press, 1979.

[36] G. SPARR. A new proof of Löwner's theorem on monotone matrix functions, *Math. Scand.*, **47**, 1980, 266–274.

[37] B. Sz. NAGY. Remarks on the preceding paper of A. Korányi, *Acta Sci. Math. (Szeged)* **17**, 1956, 71–75.

[38] C. THOMPSON. Inequalities with applications in statistical mechanics, *J. Math. Phys.*, **6**, 1965, 1812–1813.

[39] C. THOMPSON. Inequalities and partial orders on matrix spaces, *Indiana Univ. Math. J.*, **21**, 1971, 469–480.

ADDITIONAL REFERENCES

The following papers have been known to us after the paper was written.

J. LINDENSTRAUSS and A. PELCZYNSKI. Absolutely summing operators in $\mathcal{L}_p$ spaces and their applications, *Studia Math.*, **19**, 1968, 275–326.

K. BALL, E.A. CARLEN and E.H. LIEB. Sharp uniform convexity and smoothness inequalities for trace norms, *preprint,* 1993.

T. FURUTA. A proof by operator means of an order preserving inequality. *Linear Alg. Appl.*, 113 :129–130 (1989).

M. FUJII, T. FURUTA, K. KAMEI. Operator functions associated with Furuta's inequality, *Linear Alg. Appl.*, 149 :91–96 (1991).

U. HAAGERUP. Solution of the similarity próblem for cyclic representations of C^*-algebras, *Ann. of M.*, **118**, 1983, 215–240.

U. HAAGERUP and G. PISIER. Bounded linear operators between C^*-algebras, *preprint* (1993).

Institute of Math. Science, Academia Sinica,
Wuhan 430071, China
and
University of Oslo, POB 1053 Blindern N-0316 Oslo