APPROXIMATING AN INTEGRAL TRANSFORM OF POSITIVE OPERATORS IN HILBERT SPACES WITH APPLICATIONS FOR JENSEN'S GAP

SILVESTRU SEVER DRAGOMIR[1,2]

Abstract. For a continuous and positive function $w(\lambda)$, $\lambda > 0$ and μ a positive measure on $(0,\infty)$ we consider the following *integral transform*

$$\mathcal{D}(w,\mu)(T) := \int_0^\infty w(\lambda)(\lambda+T)^{-1} d\mu(\lambda),$$

where the integral is assumed to exist for T a positive operator on a complex Hilbert space H. We show among others that, if $A, B > 0$ then

$$\begin{aligned}
&\mathcal{D}(w,\mu)(B) - \mathcal{D}(w,\mu)(A) - D(\mathcal{D}(w,\mu))(A)(B-A) \\
&= 2\int_0^\infty w(\lambda)\left(\int_0^1 (1-t)(\lambda+(1-t)A+tB)^{-1}\right. \\
&\times \Big((B-A)(\lambda+(1-t)A+tB)^{-1}(B-A)\Big) \\
&\left.\times (\lambda+(1-t)A+tB)^{-1} dt\right) d\mu(\lambda) \quad (\geq 0),
\end{aligned}$$

where $D(\mathcal{D}(w,\mu))$ is the Fréchet derivative of $\mathcal{D}(w,\mu)$ as an operator function. We also obtain the following representation for the operator Jensen's gap of the logarithm

$$\begin{aligned}
&\ln\left(\sum_{k=1}^n p_k A_k\right) - \sum_{k=1}^n p_k \ln A_k \\
&= 2\sum_{k=1}^n p_k \int_0^1 (1-t)\left(\int_0^\infty \left(\lambda+(1-t)\sum_{j=1}^n p_j A_j + tA_k\right)^{-1}\right. \\
&\times \left(A_k - \sum_{j=1}^n p_j A_j\right)\left(\lambda+(1-t)\sum_{j=1}^n p_j A_j + tA_k\right)^{-1} \\
&\left.\times \left(A_k - \sum_{j=1}^n p_j A_j\right)\left(\lambda+(1-t)\sum_{j=1}^n p_j A_j + tA_k\right)^{-1} d\lambda\right) dt \ (\geq 0)
\end{aligned}$$

for the n-tuple of positive operators $\mathbf{A} = (A_1,...,A_n)$ and probability density n-tuple $\mathbf{p} = (p_1,...,p_n)$.

1. Introduction

Consider a complex Hilbert space $(H, \langle\cdot,\cdot\rangle)$. An operator T is said to be positive (denoted by $T \geq 0$) if $\langle Tx, x\rangle \geq 0$ for all $x \in H$ and also an operator T is said to be *strictly positive* (denoted by $T > 0$) if T is positive and invertible. A real valued

1991 *Mathematics Subject Classification.* 47A63, 47A60.

Key words and phrases. Operator monotone functions, Operator convex functions, Operator inequalities, Löwner-Heinz inequality. Logarithmic operator inequalities.

RGMIA Res. Rep. Coll. 23 (2020), Art. 100, 16 pp. Received 03/08/20

continuous function f on $(0,\infty)$ is said to be operator monotone if $f(A) \geq f(B)$ holds for any $A \geq B > 0$.

We have the following representation of operator monotone functions [7], see for instance [1, p. 144-145]:

Theorem 1. *A function $f : (0,\infty) \to \mathbb{R}$ is operator monotone in $(0,\infty)$ if and only if it has the representation*

$$f(t) = a + bt + \int_0^\infty \frac{t\lambda}{t+\lambda} d\mu(\lambda), \tag{1.1}$$

where $a \in R$, $b \geq 0$ and a positive measure μ on $(0,\infty)$ such that

$$\int_0^\infty \frac{\lambda}{1+\lambda} d\mu(\lambda) < \infty. \tag{1.2}$$

If f is operator monotone in $[0,\infty)$, then $a = f(0)$ in (1.1).

For some example of operator monotone functions see [3]-[5], [8], [9] and the references therein.

A real valued continuous function f on an interval I is said to be *operator convex (operator concave)* on I if

$$f((1-\lambda)A + \lambda B) \leq (\geq)\, (1-\lambda) f(A) + \lambda f(B) \tag{OC}$$

in the operator order, for all $\lambda \in [0,1]$ and for every selfadjoint operator A and B on a Hilbert space H whose spectra are contained in I. Notice that a function f is operator concave if $-f$ is operator convex.

We have the following representation of operator convex functions [1, p. 147]:

Theorem 2. *A function $f : (0,\infty) \to \mathbb{R}$ is operator convex in $(0,\infty)$ if and only if it has the representation*

$$f(t) = a + bt + ct^2 + \int_0^\infty \frac{t^2\lambda}{t+\lambda} d\mu(\lambda), \tag{1.3}$$

where a, $b \in \mathbb{R}$, $c \geq 0$ and a positive measure μ on $(0,\infty)$ such that (1.2) holds. If f is operator convex in $[0,\infty)$, then $a = f(0)$ and $b = f'_+(0)$, the right derivative, in (1.1).

We have the following integral representation for the power function when $t > 0$, $r \in (0,1]$, see for instance [1, p. 145]

$$t^r = \frac{\sin(r\pi)}{\pi} t \int_0^\infty \frac{\lambda^{r-1}}{\lambda + t} d\lambda. \tag{1.4}$$

Observe that for $t > 0$, $t \neq 1$, we have

$$\int_0^u \frac{d\lambda}{(\lambda+t)(\lambda+1)} = \frac{\ln t}{t-1} + \frac{1}{1-t} \ln\left(\frac{u+t}{u+1}\right)$$

for all $u > 0$.

By taking the limit over $u \to \infty$ in this equality, we derive

$$\frac{\ln t}{t-1} = \int_0^\infty \frac{d\lambda}{(\lambda+t)(\lambda+1)},$$

which gives the representation for the logarithm

$$\ln t = (t-1) \int_0^\infty \frac{d\lambda}{(\lambda+1)(\lambda+t)} \tag{1.5}$$

for all $t>0$.

Motivated by these representations, we introduce, for a continuous and positive function $w(\lambda)$, $\lambda>0$, the following *integral transform*

$$\mathcal{D}(w,\mu)(t):=\int_0^\infty \frac{w(\lambda)}{\lambda+t}d\mu(\lambda),\ t>0, \tag{1.6}$$

where μ is a positive measure on $(0,\infty)$ and the integral (1.6) exists for all $t>0$.

For μ the Lebesgue usual measure, we put

$$\mathcal{D}(w)(t):=\int_0^\infty \frac{w(\lambda)}{\lambda+t}d\lambda,\ t>0. \tag{1.7}$$

If we take μ to be the usual Lebesgue measure and the kernel $w_r(\lambda)=\lambda^{r-1}$, $r\in(0,1]$, then

$$t^r=\frac{\sin(r\pi)}{\pi}t\mathcal{D}(w_r)(t),\ t>0. \tag{1.8}$$

For the same measure, if we take the kernel $w_{\ln}(\lambda)=(\lambda+1)^{-1}$, $t>0$, we have the representation

$$\ln t=(t-1)\mathcal{D}(w_{\ln})(t),\ t>0. \tag{1.9}$$

Assume that $T>0$, then by the continuous functional calculus for selfadjoint operators, we can define the positive operator

$$\mathcal{D}(w,\mu)(T):=\int_0^\infty w(\lambda)(\lambda+T)^{-1}d\mu(\lambda), \tag{1.10}$$

where w and μ are as above. Also, when μ is the usual Lebesgue measure, then

$$\mathcal{D}(w)(T):=\int_0^\infty w(\lambda)(\lambda+T)^{-1}d\lambda, \tag{1.11}$$

for $T>0$.

In this paper, we show among others that, if A, $B>0$ then

$$\begin{aligned}
&\mathcal{D}(w,\mu)(B)-\mathcal{D}(w,\mu)(A)-D(\mathcal{D}(w,\mu))(A)(B-A)\\
&=2\int_0^\infty w(\lambda)\left(\int_0^1(1-t)(\lambda+(1-t)A+tB)^{-1}\right.\\
&\times\Big((B-A)(\lambda+(1-t)A+tB)^{-1}(B-A)\Big)\\
&\left.\times(\lambda+(1-t)A+tB)^{-1}dt\right)d\mu(\lambda)\\
&\geq 0,
\end{aligned}$$

where $D(\mathcal{D}(w,\mu))$ is the Fréchet derivative of $\mathcal{D}(w,\mu)$ as an operator function. We also obtain the following representation for the operator Jensen's gap of the logarithm

$$
\begin{aligned}
&\ln\left(\sum_{k=1}^{n} p_k A_k\right) - \sum_{k=1}^{n} p_k \ln A_k \\
&= 2\sum_{k=1}^{n} p_k \int_0^1 (1-t)\left(\int_0^\infty \left(\lambda + (1-t)\sum_{j=1}^{n} p_j A_j + tA_k\right)^{-1}\right. \\
&\times \left(A_k - \sum_{j=1}^{n} p_j A_j\right)\left(\lambda + (1-t)\sum_{j=1}^{n} p_j A_j + tA_k\right)^{-1} \\
&\times \left.\left(A_k - \sum_{j=1}^{n} p_j A_j\right)\left(\lambda + (1-t)\sum_{j=1}^{n} p_j A_j + tA_k\right)^{-1} d\lambda\right) dt \\
&\geq 0,
\end{aligned}
$$

for the n-tuple of positive operators $\mathbf{A} = (A_1, ..., A_n)$ and probability density n-tuple $\mathbf{p} = (p_1, ..., p_n)$.

2. Main Results

We consider the inverse function $\ell^{-1}(t) = t^{-1}$, $t > 0$ and the corresponding operator inverse function $\ell^{-1}(A) = A^{-1}$ defined for operators $A > 0$. Denote by $S(H)$ the class of all selfadjoint operators on H.

Lemma 1. *The operator inverse function ℓ^{-1} is Fréchet differentiable in $A > 0$ and*

$$D\left(\ell^{-1}\right)(A)(V) = -A^{-1}VA^{-1} \tag{2.1}$$

for all $V \in S(H)$.

Proof. Let $A > 0$ and $V \in S(H)$. Then there exists a small open interval $(-\delta, \delta)$ such that for $t \in (-\delta, \delta)$, $A + tV$ is invertible. Then

$$
\begin{aligned}
\ell^{-1}(A+tV) - \ell^{-1}(A) &= (A+tV)^{-1} - A^{-1} = (A+tV)^{-1}(A - (A+tV))A^{-1} \\
&= -t(A+tV)^{-1}VA^{-1},
\end{aligned}
$$

which implies that

$$\frac{\ell^{-1}(A+tV) - \ell^{-1}(A)}{t} = -(A+tV)^{-1}VA^{-1},\ t \neq 0.$$

By taking the limit over $t \to 0$, we get

$$
\begin{aligned}
D\left(\ell^{-1}\right)(A)(V) &= \lim_{t\to 0}\frac{\ell^{-1}(A+tV) - \ell^{-1}(A)}{t} = -\lim_{t\to 0}(A+tV)^{-1}VA^{-1} \\
&= -A^{-1}VA^{-1}
\end{aligned}
$$

and the statement is proved. □

Lemma 2. *The operator inverse function ℓ^{-1} is twice Fréchet differentiable in $A > 0$ and*

$$D^2\left(\ell^{-1}\right)(A)(V,V) = 2A^{-1}\left(VA^{-1}V\right)A^{-1} \geq 0 \tag{2.2}$$

for all $V \in S(H)$.

Proof. Let $A > 0$ and $V \in S(H)$. Then there exists a small open interval $(-\delta, \delta)$ such that for $t \in (-\delta, \delta)$, $A + tV$ is invertible and put

$$\begin{aligned} U_t &:= D\left(\ell^{-1}\right)(A+tV)(V) - D\left(\ell^{-1}\right)(A)(V) \\ &= A^{-1}VA^{-1} - (A+tV)^{-1}V(A+tV)^{-1}. \end{aligned}$$

If we multiply both sides of U_t with $A + tV$, then we get

$$\begin{aligned} (A+tV)U_t(A+tV) &= (A+tV)A^{-1}VA^{-1}(A+tV) - V \\ &= \left(1+tVA^{-1}\right)V\left(1+tA^{-1}V\right) - V \\ &= \left(V+tVA^{-1}V\right)\left(1+tA^{-1}V\right) - V \\ &= V + tVA^{-1}V + tVA^{-1}V + t^2VA^{-1}VA^{-1}V - V \\ &= 2tVA^{-1}V + t^2\left(VA^{-1}\right)^2V \\ &= t\left[2VA^{-1}V + t\left(VA^{-1}\right)^2V\right]. \end{aligned}$$

By multiplying this equality both sides with $(A+tV)^{-1}$ we get

$$U_t = t(A+tV)^{-1}\left[2VA^{-1}V + t\left(VA^{-1}\right)^2V\right](A+tV)^{-1}.$$

Therefore

$$\begin{aligned} D^2\left(\ell^{-1}\right)(A)(V,V) &= \lim_{t\to 0}\frac{D\left(\ell^{-1}\right)(A+tV)(V) - D\left(\ell^{-1}\right)(A)(V)}{t} \\ &= \lim_{t\to 0}\frac{t(A+tV)^{-1}\left[2VA^{-1}V + t\left(VA^{-1}\right)^2V\right](A+tV)^{-1}}{t} \\ &= \lim_{t\to 0}(A+tV)^{-1}\left[2VA^{-1}V + t\left(VA^{-1}\right)^2V\right](A+tV)^{-1} \\ &= 2A^{-1}\left(VA^{-1}V\right)A^{-1}. \end{aligned}$$

Since $A^{-1} > 0$, then by multiplying both sides by V we get $VA^{-1}V \geq 0$ and by multiplying both sides by A^{-1} we get $A^{-1}\left(VA^{-1}V\right)A^{-1} \geq 0$, and the proof of (2.2) is thus proved. □

Lemma 3. *For all* C, $E > 0$ *we have the representation*

$$\begin{aligned} E^{-1} - C^{-1} &= -C^{-1}(E-C)C^{-1} \\ &\quad + 2\int_0^1 (1-t)\left((1-t)C+tE\right)^{-1} \\ &\quad \times \Big((E-C)\left((1-t)C+tE\right)^{-1}(E-C)\Big)\left((1-t)C+tE\right)^{-1}dt. \end{aligned} \tag{2.3}$$

Proof. We use the Taylor's type formula with integral remainder, see for instance [2, p. 112],

$$\begin{aligned} f(E) &= f(C) + D(f)(C)(E-C) \\ &\quad + \int_0^1 (1-t)D^2(f)\left((1-t)C+tE\right)(E-C, E-C)\,dt \end{aligned} \tag{2.4}$$

that holds for functions f which are of class C^2 on an open and convex subset $\mathcal{O}$ in the Banach algebra $S(H)$ and C, $E \in \mathcal{O}$.

Therefore, if we write this formula for ℓ^{-1} and the class of strictly positive operators, we get

$$\begin{aligned} E^{-1} &= C^{-1} - C^{-1}(E-C)C^{-1} \\ &+ 2\int_0^1 (1-t)\left((1-t)C+tE\right)^{-1} \\ &\times \Big((E-C)\left((1-t)C+tE\right)^{-1}(E-C)\Big)\left((1-t)C+tE\right)^{-1} dt \end{aligned}$$

and the representation (2.3) is proved. □

Lemma 4. *For all $A>0$,*

$$D(\mathcal{D}(w,\mu))(A)(V) = -\int_0^\infty w(\lambda)(\lambda+A)^{-1}V(\lambda+A)^{-1}d\mu(\lambda) \tag{2.5}$$

for all $V\in S(H)$.

Proof. By the definition of $\mathcal{D}(w,\mu)$ we have for t in a small open interval around 0 that

$$\begin{aligned} &\mathcal{D}(w,\mu)(A+tV) - \mathcal{D}(w,\mu)(A) \\ &= \int_0^\infty w(\lambda)\left[(\lambda+A+tV)^{-1} - (\lambda+A)^{-1}\right]d\mu(\lambda) \\ &= \int_0^\infty w(\lambda)\left[(\lambda+A+tV)^{-1}(\lambda+A-\lambda-A-tV)(\lambda+A)^{-1}\right]d\mu(\lambda) \\ &= -t\int_0^\infty w(\lambda)\left[(\lambda+A+tV)^{-1}V(\lambda+A)^{-1}\right]d\mu(\lambda). \end{aligned}$$

Therefore,

$$\begin{aligned} &\lim_{t\to0}\frac{\mathcal{D}(w,\mu)(A+tV) - \mathcal{D}(w,\mu)(A)}{t} \\ &= -\lim_{t\to0}\int_0^\infty w(\lambda)\left[(\lambda+A+tV)^{-1}V(\lambda+A)^{-1}\right]d\mu(\lambda) \\ &= -\int_0^\infty w(\lambda)\left[(\lambda+A)^{-1}V(\lambda+A)^{-1}\right]d\mu(\lambda) \end{aligned}$$

and the identity (2.5) is obtained. □

We have the following approximation result of $\mathcal{D}(w,\mu)(B)$ by

$$\mathcal{D}(w,\mu)(A) + D(\mathcal{D}(w,\mu))(A)(B-A)$$

with an explicit error term expressed as a double integral:

Theorem 3. *For all $A,\ B>0$, we have the representation*

$$\begin{aligned} &\mathcal{D}(w,\mu)(B) - \mathcal{D}(w,\mu)(A) - D(\mathcal{D}(w,\mu))(A)(B-A) \\ &= 2\int_0^\infty w(\lambda)\Big(\int_0^1 (1-t)(\lambda+(1-t)A+tB)^{-1} \\ &\times \Big((B-A)(\lambda+(1-t)A+tB)^{-1}(B-A)\Big) \\ &\times (\lambda+(1-t)A+tB)^{-1}dt\Big)d\mu(\lambda) \quad (\geq 0). \end{aligned} \tag{2.6}$$

Proof. From the identity (2.3) we get for $E=\lambda+B$ and $C=\lambda+A$ that

$$(\lambda+B)^{-1}-(\lambda+A)^{-1}=-(\lambda+A)^{-1}(B-A)(\lambda+A)^{-1} + 2\int_0^1 (1-t)(\lambda+(1-t)A+tB)^{-1} \times\Big((B-A)(\lambda+(1-t)A+tB)^{-1}(B-A)\Big) \times(\lambda+(1-t)A+tB)^{-1}dt \tag{2.7}$$

for $\lambda\geq 0$.

If we multiply this equality by $w(\lambda)\geq 0$ and integrate, then we get

$$\int_0^\infty w(\lambda)(\lambda+B)^{-1}d\mu(\lambda)-\int_0^\infty w(\lambda)(\lambda+A)^{-1}d\mu(\lambda) = -\int_0^\infty w(\lambda)(\lambda+A)^{-1}(B-A)(\lambda+A)^{-1} + 2\int_0^\infty w(\lambda)\Big(\int_0^1 (1-t)(\lambda+(1-t)A+tB)^{-1} \times\Big((B-A)(\lambda+(1-t)A+tB)^{-1}(B-A)\Big) \times(\lambda+(1-t)A+tB)^{-1}dt\Big)d\mu(\lambda) \tag{2.8}$$

and since, by (2.5),

$$-\int_0^\infty w(\lambda)(\lambda+A)^{-1}(B-A)(\lambda+A)^{-1}=D(\mathcal{D}(w,\mu))(A)(B-A),$$

hence by (2.8) we derive the equality in (2.8).

We observe that for all $\lambda\geq 0$ and $t\in[0,1]$, $(\lambda+(1-t)A+tB)^{-1}>0$. If we multiply this inequality both sides by $(B-A)$ we derive

$$(B-A)(\lambda+(1-t)A+tB)^{-1}(B-A)\geq 0.$$

Further, if we multiply this inequality both sides by $(\lambda+(1-t)A+tB)^{-1}$ we get

$$(1-t)(\lambda+(1-t)A+tB)^{-1} \times\Big((B-A)(\lambda+(1-t)A+tB)^{-1}(B-A)\Big)(\lambda+(1-t)A+tB)^{-1} \geq 0$$

for all $\lambda\geq 0$ and $t\in[0,1]$. If we multiply this inequality with $w(\lambda)\geq 0$ and integrate, then we obtain the inequality in (2.6). □

Corollary 1. *Assume that function $f:[0,\infty)\to\mathbb{R}$ is operator monotone in $[0,\infty)$ and has the representation (1.1), then*

$$
\begin{aligned}
&\left(f\left(B\right)-f\left(0\right)\right)B^{-1}-\left(f\left(A\right)-f\left(0\right)\right)A^{-1} \\
&+\left(f\left(A\right)-f\left(0\right)\right)A^{-1}\left(B-A\right)A^{-1}-D\left(f\right)\left(A\right)\left(B-A\right)A^{-1} \\
&=2\int_0^\infty \lambda\left(\int_0^1 \left(1-t\right)\left(\lambda+\left(1-t\right)A+tB\right)^{-1}\right. \\
&\times\left(\left(B-A\right)\left(\lambda+\left(1-t\right)A+tB\right)^{-1}\left(B-A\right)\right) \\
&\left.\times\left(\lambda+\left(1-t\right)A+tB\right)^{-1}dt\right)d\mu\left(\lambda\right) \\
&\geq 0.
\end{aligned}
\tag{2.9}
$$

If $f\left(0\right)=0$, then we have the simpler equality

$$
\begin{aligned}
&f\left(B\right)B^{-1}-f\left(A\right)A^{-1}+f\left(A\right)A^{-1}\left(B-A\right)A^{-1} \\
&-D\left(f\right)\left(A\right)\left(B-A\right)A^{-1} \\
&=2\int_0^\infty \lambda\left(\int_0^1 \left(1-t\right)\left(\lambda+\left(1-t\right)A+tB\right)^{-1}\right. \\
&\times\left(\left(B-A\right)\left(\lambda+\left(1-t\right)A+tB\right)^{-1}\left(B-A\right)\right) \\
&\left.\times\left(\lambda+\left(1-t\right)A+tB\right)^{-1}dt\right)d\mu\left(\lambda\right) \\
&\geq 0.
\end{aligned}
\tag{2.10}
$$

Proof. We have

$$
\mathcal{D}\left(\ell,\mu\right)\left(t\right)=\frac{f\left(t\right)-f\left(0\right)}{t}-b,\ t>0,
$$

where $b\geq 0$ and μ is a positive measure on $(0,\infty)$.

We have for $A>0$, $V\in S\left(H\right)$ that

$$
\begin{aligned}
&\mathcal{D}\left(\ell,\mu\right)\left(A+tV\right)-\mathcal{D}\left(\ell,\mu\right)\left(A\right) \\
&=\left(f\left(A+tV\right)-f\left(0\right)\right)\left(A+tV\right)^{-1}-\left(f\left(A\right)-f\left(0\right)\right)A^{-1} \\
&=\left(f\left(A+tV\right)-f\left(0\right)\right)\left(A+tV\right)^{-1}-\left(f\left(A+tV\right)-f\left(0\right)\right)A^{-1} \\
&+\left(f\left(A+tV\right)-f\left(0\right)\right)A^{-1}-\left(f\left(A\right)-f\left(0\right)\right)A^{-1} \\
&=\left(f\left(A+tV\right)-f\left(0\right)\right)\left[\left(A+tV\right)^{-1}-A^{-1}\right] \\
&+\left[\left(f\left(A+tV\right)-f\left(0\right)\right)-\left(f\left(A\right)-f\left(0\right)\right)\right]A^{-1} \\
&=\left(f\left(A+tV\right)-f\left(0\right)\right)\left[\left(A+tV\right)^{-1}-A^{-1}\right] \\
&+\left[f\left(A+tV\right)-f\left(A\right)\right]A^{-1}
\end{aligned}
$$

for t in a small interval around 0.

Then

$$
\begin{aligned}
& D(\mathcal{D}(w,\mu))(A)(V) \\
& = \lim_{t\to 0} \frac{\mathcal{D}(\ell,\mu)(A+tV) - \mathcal{D}(\ell,\mu)(A)}{t} \\
& = \lim_{t\to 0} \left[(f(A+tV) - f(0)) \frac{(A+tV)^{-1} - A^{-1}}{t} \right] \\
& + \lim_{t\to 0} \left[\frac{f(A+tV) - f(A)}{t} \right] A^{-1} \\
& = (f(A) - f(0)) \left(-A^{-1} V A^{-1} \right) + D(f)(A)(V) A^{-1} \\
& = D(f)(A)(V) A^{-1} - f(A) A^{-1} V A^{-1} + f(0) A^{-1} V A^{-1}.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
& \mathcal{D}(w,\mu)(B) - \mathcal{D}(w,\mu)(A) - D(\mathcal{D}(w,\mu))(A)(B-A) \\
& = (f(B) - f(0)) B^{-1} - (f(A) - f(0)) A^{-1} - D(f)(A)(B-A) A^{-1} \\
& + f(A) A^{-1} (B-A) A^{-1} - f(0) A^{-1} (B-A) A^{-1} \\
& = (f(B) - f(0)) B^{-1} - (f(A) - f(0)) A^{-1} \\
& + (f(A) - f(0)) A^{-1} (B-A) A^{-1} - D(f)(A)(B-A) A^{-1}
\end{aligned}
$$

and by (2.6) we get (2.9). □

Remark 1. *We consider the representation obtained from (1.8) for the operator $A > 0$ and the power $r \in (0,1]$,*

$$
T^{r-1} = \mathcal{D}(\tilde{w}_r)(T)
$$

with the kernel $\tilde{w}_r(\lambda) := \frac{\sin(r\pi)}{\pi} \lambda^{r-1}$, $r \in (0,1]$. Then by (2.6) we get for A, $B > 0$ that

$$
\begin{aligned}
& \mathcal{D}(\tilde{w}_r)(B) - \mathcal{D}(\tilde{w}_r)(A) - D(\mathcal{D}(\tilde{w}_r))(A)(B-A) \\
& = 2 \frac{\sin(r\pi)}{\pi} \int_0^\infty \lambda^{r-1} \left(\int_0^1 (1-t) (\lambda + (1-t) A + tB)^{-1} \right. \\
& \times \Big((B-A) (\lambda + (1-t) A + tB)^{-1} (B-A) \Big) \\
& \left. \times (\lambda + (1-t) A + tB)^{-1} dt \right) d\lambda \\
& \geq 0,
\end{aligned}
$$

namely

$$
\begin{aligned}
& B^{r-1} - A^{r-1} - D\left(\ell^{r-1}\right)(A)(B-A) \\
& = 2 \frac{\sin(r\pi)}{\pi} \int_0^\infty \lambda^{r-1} \left(\int_0^1 (1-t) (\lambda + (1-t) A + tB)^{-1} \right. \\
& \times \Big((B-A) (\lambda + (1-t) A + tB)^{-1} (B-A) \Big) \\
& \left. \times (\lambda + (1-t) A + tB)^{-1} dt \right) d\lambda \\
& \geq 0,
\end{aligned}
$$

where $\ell(\lambda) = \lambda$, $\lambda \geq 0$.

We also observe that

$$D\left(\ell^{r-1}\right)(A)(V)=-\frac{\sin(r\pi)}{\pi}\int_0^\infty \lambda^{r-1}(\lambda+A)^{-1}V(\lambda+A)^{-1}d\lambda. \tag{2.11}$$

We have the following representation of operator Jensen's gap for the n-tuple of positive operators $\mathbf{A}=(A_1,...,A_n)$ and probability density n-tuple $\mathbf{p}=(p_1,...,p_n)$,

$$J(\mathbf{A},\mathbf{p},\mathcal{D}(w,\mu)):=\sum_{k=1}^n p_k\mathcal{D}(w,\mu)(A_k)-\mathcal{D}(w,\mu)\left(\sum_{k=1}^n p_kA_k\right).$$

Theorem 4. *We have the representation*

$$\begin{aligned}
&J(\mathbf{A},\mathbf{p},\mathcal{D}(w,\mu)) \\
&=2\sum_{k=1}^n p_k\int_0^\infty w(\lambda)\left(\int_0^1(1-t)\left(\lambda+(1-t)\sum_{j=1}^n p_jA_j+tA_k\right)^{-1}\right. \\
&\times\left(A_k-\sum_{j=1}^n p_jA_j\right)\left(\lambda+(1-t)\sum_{j=1}^n p_jA_j+tA_k\right)^{-1} \\
&\times\left.\left(A_k-\sum_{j=1}^n p_jA_j\right)\left(\lambda+(1-t)\sum_{j=1}^n p_jA_j+tA_k\right)^{-1}dt\right)d\mu(\lambda) \\
&\geq 0.
\end{aligned} \tag{2.12}$$

This also shows that $\mathcal{D}(w,\mu)$ *is operator convex on* $(0,\infty)$.

Proof. From the identity (2.6) we get

$$\begin{aligned}
&\mathcal{D}(w,\mu)(A_k)-\mathcal{D}(w,\mu)\left(\sum_{j=1}^n p_jA_j\right) \\
&-D(\mathcal{D}(w,\mu))\left(\sum_{j=1}^n p_jA_j\right)\left(A_k-\sum_{j=1}^n p_jA_j\right) \\
&=2\int_0^\infty w(\lambda)\left(\int_0^1(1-t)\left(\lambda+(1-t)\sum_{j=1}^n p_jA_j+tA_k\right)^{-1}\right. \\
&\times\left(A_k-\sum_{j=1}^n p_jA_j\right)\left(\lambda+(1-t)\sum_{j=1}^n p_jA_j+tA_k\right)^{-1} \\
&\times\left.\left(A_k-\sum_{j=1}^n p_jA_j\right)\left(\lambda+(1-t)\sum_{j=1}^n p_jA_j+tA_k\right)^{-1}dt\right)d\mu(\lambda) \\
&\geq 0
\end{aligned}$$

for all $k\in\{1,...,n\}$.

If we multiply this inequality with $p_k \geq 0$, take into account that $\sum_{k=1}^{n} p_k = 1$ and

$$\sum_{k=1}^{n} p_k D\left(\mathcal{D}\left(w,\mu\right)\right)\left(\sum_{j=1}^{n} p_j A_j\right)\left(A_k - \sum_{j=1}^{n} p_j A_j\right)$$
$$= D\left(\mathcal{D}\left(w,\mu\right)\right)\left(\sum_{j=1}^{n} p_j A_j\right)\left(\sum_{k=1}^{n} p_k A_k - \sum_{j=1}^{n} p_j A_j\right)$$
$$= D\left(\mathcal{D}\left(w,\mu\right)\right)\left(\sum_{j=1}^{n} p_j A_j\right)(0) = 0,$$

then we obtain the desired result (2.12). □

The case of operator monotonic functions is as follows:

Corollary 2. *Assume that function $f : [0,\infty) \to \mathbb{R}$ is operator monotone in $[0,\infty)$ and has the representation (1.1), then*

$$\begin{aligned}(2.13)\qquad & \sum_{k=1}^{n} p_k f\left(A_k\right) A_k^{-1} - f\left(\sum_{j=1}^{n} p_j A_j\right)\left(\sum_{j=1}^{n} p_j A_j\right)^{-1} \\ & - f\left(0\right)\left[\sum_{k=1}^{n} p_k A_k^{-1} - \left(\sum_{j=1}^{n} p_j A_j\right)^{-1}\right] \\ & = 2\sum_{k=1}^{n} p_k \int_0^{\infty} \lambda\left(\int_0^1 (1-t)\left(\lambda + (1-t)\sum_{j=1}^{n} p_j A_j + tA_k\right)^{-1}\right. \\ & \times\left(A_k - \sum_{j=1}^{n} p_j A_j\right)\left(\lambda + (1-t)\sum_{j=1}^{n} p_j A_j + tA_k\right)^{-1} \\ & \left.\times\left(A_k - \sum_{j=1}^{n} p_j A_j\right)\left(\lambda + (1-t)\sum_{j=1}^{n} p_j A_j + tA_k\right)^{-1} dt\right) d\mu\left(\lambda\right) \\ & \geq 0,\end{aligned}$$

for the n-tuple of positive operators $\mathbf{A} = (A_1, ..., A_n)$ and probability density n-tuple $\mathbf{p} = (p_1, ..., p_n)$.

The case of operator convex functions is as follows:

Corollary 3. *Assume that function $f : [0,\infty) \to \mathbb{R}$ is operator convex in $[0,\infty)$ and has the representation (1.3), then for the n-tuple of positive operators* $\mathbf{A} =$

$(A_1, ..., A_n)$ and probability density n-tuple $\mathbf{p} = (p_1, ..., p_n)$,

$$
\begin{aligned}
(2.14)\quad & \sum_{k=1}^{n} p_k f(A_k) A_k^{-2} - f\left(\sum_{k=1}^{n} p_k A_k\right)\left(\sum_{k=1}^{n} p_k A_k\right)^{-2} \\
& - f(0)\left[\sum_{k=1}^{n} p_k A_k^{-2} - \left(\sum_{k=1}^{n} p_k A_k\right)^{-2}\right] \\
& - f'_+(0)\left[\sum_{k=1}^{n} p_k A_k^{-1} - \left(\sum_{j=1}^{n} p_k A_k\right)^{-1}\right] \\
& = 2\sum_{k=1}^{n} p_k \int_0^\infty \lambda \left(\int_0^1 (1-t)\left(\lambda + (1-t)\sum_{j=1}^{n} p_j A_j + tA_k\right)^{-1}\right. \\
& \times \left(A_k - \sum_{j=1}^{n} p_j A_j\right)\left(\lambda + (1-t)\sum_{j=1}^{n} p_j A_j + tA_k\right)^{-1} \\
& \times \left.\left(A_k - \sum_{j=1}^{n} p_j A_j\right)\left(\lambda + (1-t)\sum_{j=1}^{n} p_j A_j + tA_k\right)^{-1} dt\right) d\mu(\lambda) \quad (\geq 0).
\end{aligned}
$$

Proof. By the representation (1.3) we derive

$$
\mathcal{D}(\ell, \mu)(t) = \frac{f(t) - f(0) - f'_+(0)\, t}{t^2} - ct, \ t > 0.
$$

By utilising Theorem 4 we deduce (2.14). □

Remark 2. *We have the operator power inequality for $r \in (0, 1]$,*

$$
\begin{aligned}
(2.15)\quad & \sum_{k=1}^{n} p_k A_k^{r-1} - \left(\sum_{k=1}^{n} p_k A_k\right)^{r-1} \\
& = 2\frac{\sin(r\pi)}{\pi}\sum_{k=1}^{n} p_k \int_0^\infty \lambda^{r-1}\left(\int_0^1 (1-t)\left(\lambda + (1-t)\sum_{j=1}^{n} p_j A_j + tA_k\right)^{-1}\right. \\
& \times \left(A_k - \sum_{j=1}^{n} p_j A_j\right)\left(\lambda + (1-t)\sum_{j=1}^{n} p_j A_j + tA_k\right)^{-1} \\
& \times \left.\left(A_k - \sum_{j=1}^{n} p_j A_j\right)\left(\lambda + (1-t)\sum_{j=1}^{n} p_j A_j + tA_k\right)^{-1} dt\right) d\lambda \\
& \geq 0,
\end{aligned}
$$

for the n-tuple of positive operators $\mathbf{A} = (A_1, ..., A_n)$ and probability density n-tuple $\mathbf{p} = (p_1, ..., p_n)$.

3. The Case of Logarithm

The case of logarithm is important and can not be obtained from the consideration above.

Theorem 5. *For all* A, $B>0$ *we have*

$$\begin{aligned}(3.1)\quad & \ln B-\ln A-D(\ln)(A)(B-A)\\ &=-2\int_0^1(1-t)\left(\int_0^\infty(\lambda+(1-t)A+tB)^{-1}\right.\\ &\times\left.\Big((B-A)(\lambda+(1-t)A+tB)^{-1}(B-A)\Big)(\lambda+(1-t)A+tB)^{-1}d\lambda\right)dt\\ &\le 0,\end{aligned}$$

where

$$D(\ln)(A)(B-A)=\int_0^\infty(\lambda+A)^{-1}(B-A)(\lambda+A)^{-1}d\lambda. \tag{3.2}$$

Proof. We have from (1.5) for A, $B>0$ that

$$\ln B-\ln A=\int_0^\infty\frac{1}{\lambda+1}\left[(B-1)(\lambda+B)^{-1}-(A-1)(\lambda+A)^{-1}\right]d\lambda. \tag{3.3}$$

Since

$$\begin{aligned}&(B-1)(\lambda+B)^{-1}-(A-1)(\lambda+A)^{-1}\\ &=B(\lambda+B)^{-1}-A(\lambda+A)^{-1}-\Big((\lambda+B)^{-1}-(\lambda+A)^{-1}\Big)\end{aligned}$$

and

$$\begin{aligned}&B(\lambda+B)^{-1}-A(\lambda+A)^{-1}\\ &=(B+\lambda-\lambda)(\lambda+B)^{-1}-(A+\lambda-\lambda)(\lambda+A)^{-1}\\ &=1-\lambda(\lambda+B)^{-1}-1+\lambda(\lambda+A)^{-1}=\lambda(\lambda+A)^{-1}-\lambda(\lambda+B)^{-1},\end{aligned}$$

hence

$$\begin{aligned}&(B-1)(\lambda+B)^{-1}-(A-1)(\lambda+A)^{-1}\\ &=\lambda(\lambda+A)^{-1}-\lambda(\lambda+B)^{-1}-\Big((\lambda+B)^{-1}-(\lambda+A)^{-1}\Big)\\ &=(\lambda+1)\left[(\lambda+A)^{-1}-(\lambda+B)^{-1}\right]\end{aligned}$$

and by (3.3) we get

$$\ln B-\ln A=\int_0^\infty\left[(\lambda+A)^{-1}-(\lambda+B)^{-1}\right]d\lambda. \tag{3.4}$$

Therefore, for $A>0$, $V\in S(H)$ and t in a small interval around 0,

$$\begin{aligned}
&D(\ln)(A)(V)\\
&=\lim_{t\to 0}\frac{\ln(A+tV)-\ln A}{t}=\lim_{t\to 0}\int_0^\infty\left[\frac{(\lambda+A)^{-1}-(\lambda+A+tV)^{-1}}{t}\right]d\lambda\\
&=\int_0^\infty\lim_{t\to 0}\left[\frac{(\lambda+A)^{-1}-(\lambda+A+tV)^{-1}}{t}\right]d\lambda\\
&=\int_0^\infty\lim_{t\to 0}(\lambda+A+tV)^{-1}V(\lambda+A)^{-1}d\lambda=\int_0^\infty(\lambda+A)^{-1}V(\lambda+A)^{-1}d\lambda,
\end{aligned}$$

which gives

$$D(\ln)(A)(B-A)=\int_0^\infty(\lambda+A)^{-1}(B-A)(\lambda+A)^{-1}d\lambda.$$

For $A>0$, $V\in S(H)$ and t in a small interval around 0,

$$\begin{aligned}
&D(\ln)(A+tV)(V)-D(\ln)(A)(V) \quad (3.5)\\
&=\int_0^\infty\left[(\lambda+A+tV)^{-1}V(\lambda+A+tV)^{-1}-(\lambda+A)^{-1}V(\lambda+A)^{-1}\right]d\lambda.
\end{aligned}$$

Put

$$W_t:=(\lambda+A+tV)^{-1}V(\lambda+A+tV)^{-1}-(\lambda+A)^{-1}V(\lambda+A)^{-1}.$$

Then

$$\begin{aligned}
&(\lambda+A+tV)W_t(\lambda+A+tV)\\
&=V-(\lambda+A+tV)(\lambda+A)^{-1}V(\lambda+A)^{-1}(\lambda+A+tV)\\
&=V-\left(1+tV(\lambda+A)^{-1}\right)V\left(1+t(\lambda+A)^{-1}V\right)\\
&=V-\left(V+tV(\lambda+A)^{-1}V\right)\left(1+t(\lambda+A)^{-1}V\right)\\
&=V-V-tV(\lambda+A)^{-1}V-tV(\lambda+A)^{-1}V-t^2V(\lambda+A)^{-1}V(\lambda+A)^{-1}V\\
&=-2tV(\lambda+A)^{-1}V-t^2\left(V(\lambda+A)^{-1}\right)^2V\\
&=-t\left(2V(\lambda+A)^{-1}V+t\left(V(\lambda+A)^{-1}\right)^2V\right).
\end{aligned}$$

This implies that

$$\begin{aligned}
\frac{1}{t}W_t=&-(\lambda+A+tV)^{-1}\left(2V(\lambda+A)^{-1}V+t\left(V(\lambda+A)^{-1}\right)^2V\right)\\
&\times(\lambda+A+tV)^{-1}
\end{aligned}$$

for $t\neq 0$.

From this, by taking the limit, we get

$$\lim_{t\to 0}\left(\frac{1}{t}W_t\right)=-2(\lambda+A)^{-1}\left(V(\lambda+A)^{-1}V\right)(\lambda+A)^{-1}.$$

Therefore

$$\begin{aligned} D^2(\ln)(A)(V,V) &= \lim_{t\to 0}\frac{D(\ln)(A+tV)(V)-D(\ln)(A)(V)}{t} \\ &= \lim_{t\to 0}\int_0^\infty\left(\frac{1}{t}W_t\right)d\lambda = \int_0^\infty \lim_{t\to 0}\left(\frac{1}{t}W_t\right)d\lambda \\ &= -2\int_0^\infty (\lambda+A)^{-1}\left(V(\lambda+A)^{-1}V\right)(\lambda+A)^{-1}\,d\lambda. \end{aligned}$$

Now, if we write the Taylor's expansion formula (2.4) for the logarithm, we get

$$\begin{aligned} \ln B &= \ln A + D(\ln)(A)(B-A) \\ &+ \int_0^1 (1-t)\,D^2(\ln)((1-t)A+tB)(B-A,B-A)\,dt \\ &= \ln A + D(\ln)(A)(B-A) \\ &- 2\int_0^1 (1-t)\left(\int_0^\infty (\lambda+(1-t)A+tB)^{-1}\right. \\ &\times \left.\left((B-A)(\lambda+(1-t)A+tB)^{-1}(B-A)\right)(\lambda+(1-t)A+tB)^{-1}\,d\lambda\right)dt, \end{aligned}$$

which proves the equality (3.1).

The inequality follows in a similar way as in Theorem 3. □

Corollary 4. *For the n-tuple of positive operators $\mathbf{A}=(A_1,...,A_n)$ and probability density n-tuple $\mathbf{p}=(p_1,...,p_n)$ we have the logarithmic identity for the Jensen's gap,*

$$\begin{aligned} &\ln\left(\sum_{k=1}^n p_kA_k\right)-\sum_{k=1}^n p_k\ln A_k \\ &= 2\sum_{k=1}^n p_k\int_0^1 (1-t)\left(\int_0^\infty\left(\lambda+(1-t)\sum_{j=1}^n p_jA_j+tA_k\right)^{-1}\right. \\ &\times\left(A_k-\sum_{j=1}^n p_jA_j\right)\left(\lambda+(1-t)\sum_{j=1}^n p_jA_j+tA_k\right)^{-1} \\ &\times\left.\left(A_k-\sum_{j=1}^n p_jA_j\right)\left(\lambda+(1-t)\sum_{j=1}^n p_jA_j+tA_k\right)^{-1}d\lambda\right)dt \\ &\geq 0. \end{aligned} \tag{3.6}$$

References

[1] R. Bhatia, *Matrix Analysis*. Graduate Texts in Mathematics, 169. Springer-Verlag, New York, 1997. xii+347 pp. ISBN: 0-387-94846-5.

[2] R. Coleman, *Calculus on Normed Vector Spaces*, Springer, New York, 2012.

[3] J. I. Fujii, Y. Seo, On parametrized operator means dominated by power ones, *Sci. Math.* **1** (1998) 301–306.

[4] T. Furuta, Concrete examples of operator monotone functions obtained by an elementary method without appealing to Löwner integral representation, *Linear Algebra Appl.* **429** (2008) 972–980.

[5] T. Furuta, Precise lower bound of $f(A) - f(B)$ for $A > B > 0$ and non-constant operator monotone function f on $[0, \infty)$. *J. Math. Inequal.* **9** (2015), no. 1, 47–52.
[6] E. Heinz, Beiträge zur Störungsteorie der Spektralzerlegung, *Math. Ann.* **123** (1951) 415–438.
[7] K. Löwner, Über monotone MatrixFunktionen, *Math. Z.* **38** (1934) 177–216.
[8] M. S. Moslehian, H. Najafi, An extension of the Löwner-Heinz inequality, *Linear Algebra Appl.*, **437** (2012), 2359–2365.
[9] H. Zuo, G. Duan, Some inequalities of operator monotone functions. *J. Math. Inequal.* **8** (2014), no. 4, 777–781.

[1]Mathematics, College of Engineering & Science, Victoria University, PO Box 14428, Melbourne City, MC 8001, Australia.
E-mail address: sever.dragomir@vu.edu.au
URL: http://rgmia.org/dragomir

[2]DST-NRF Centre of Excellence in the Mathematical, and Statistical Sciences, School of Computer Science, & Applied Mathematics, University of the Witwatersrand,, Johannesburg, South Africa.