J. Math. Soc. Japan
Vol. 62, No. 3 (2010) pp. 949–961
doi: 10.2969/jmsj/06230949

A note on the Jensen inequality for self-adjoint operators

By Tomohiro Hayashi

(Received May 20, 2009)
(Revised June 25, 2009)

Abstract. In this paper we consider a certain order-like relation for self-adjoint operators on a Hilbert space. This relation is defined by using the Jensen inequality. We will show that under some assumptions this relation is antisymmetric.

1. Introduction.

Let $f(t)$ be a continuous, increasing concave function on the real line $\boldsymbol{R}$ and let A be a bounded self-adjoint operator on a Hilbert space $\mathfrak{H}$ with an inner product $\langle \cdot, \cdot \rangle$. Then for each unit vector $\xi \in \mathfrak{H}$, we have the so-called Jensen inequality:

$$\langle f(A)\xi, \xi \rangle \leq f(\langle A\xi, \xi \rangle).$$

If two self-adjoint operators X and Y satisfy $f(X) \leq f(Y)$, then by using the Jensen inequality we have

$$\langle f(X)\xi, \xi \rangle \leq \langle f(Y)\xi, \xi \rangle \leq f(\langle Y\xi, \xi \rangle).$$

Therefore if $\langle f(X)\xi, \xi \rangle \leq f(\langle Y\xi, \xi \rangle)$ for any unit vector $\xi \in \mathfrak{H}$, we may consider that X is dominated by Y in some sense. Keeping this in our minds, we shall consider the following problem: If we have $\langle f(X)\xi, \xi \rangle \leq f(\langle Y\xi, \xi \rangle)$ and $\langle f(Y)\xi, \xi \rangle \leq f(\langle X\xi, \xi \rangle)$ for any unit vector $\xi \in \mathfrak{H}$, can we conclude $X = Y$? (This problem was suggested by Professor Bourin [**2**].)

The main results of this paper consist of two theorems. In Section 2 we will solve the above problem affirmatively when the Hilbert space $\mathfrak{H}$ is finite dimensional. Unfortunately we cannot show this in the infinite dimensional case. But in Section 3 we will solve a modified problem in full generality.

Here we remark that in the paper [**1**], T. Ando considered a similar problem and showed the following theorem: Let $f(t)$ be an operator monotone function.

2000 *Mathematics Subject Classification.* Primary 47A63.

Key Words and Phrases. operator inequality, Jensen inequality.

If two positive invertible operators X and Y satisfy $\langle f(X)\xi,\xi\rangle \le f(\langle Y\xi,\xi\rangle)$ and $f(\langle Y^{-1}\xi,\xi\rangle^{-1}) \le \langle f(X)^{-1}\xi,\xi\rangle^{-1}$ for any unit vector $\xi \in \mathfrak{H}$, then we have $f(X) = f(Y)$.

The author wishes to express his hearty gratitude to Professor Jean-Christophe Bourin. The author is also grateful to Professors Yoshihiro Nakamura and Tsuyoshi Ando for valuable comments. The author would like to thank the referee for careful reading and suggestions.

Throughout this paper we assume that the readers are familiar with basic notations and results on operator theory. We refer the readers to Conway's book [**3**].

We denote by $\mathfrak{H}$ a (finite or infinite dimensional) complex Hilbert space and by $B(\mathfrak{H})$ all bounded linear operators on it. The operator norm of $A \in B(\mathfrak{H})$ is denoted by $\|A\|$. The inner product and the norm for two vectors $\xi,\eta \in \mathfrak{H}$ are denoted by $\langle\xi,\eta\rangle$ and $\|\xi\|$ respectively. We denote the defining function for an interval $[a,b)$ by $\chi_{[a,b)}(t)$.

2. Finite dimensional case.

Theorem 2.1. *Let $f(t)$ be a continuous strictly increasing (or decreasing) convex function on an interval I and let $X,Y \in M_n(\boldsymbol{C})$ be two hermitian matrices whose numerical ranges are contained in I. If X and Y satisfy*

$$\langle f(X)\xi,\xi\rangle \ge f(\langle Y\xi,\xi\rangle)$$

and

$$\langle f(Y)\xi,\xi\rangle \ge f(\langle X\xi,\xi\rangle)$$

for any unit vector $\xi \in \boldsymbol{C}^n$, then we have $X = Y$.

Proof. Replacing $f(t)$ by $f(t)+c$ for some positive constant c if necessary, we may assume that $f \ge 0$ on I. Then $f(X)$ and $f(Y)$ are positive semidefinite matrices. Take unit eigenvectors $\xi,\eta \in \boldsymbol{C}^n$ satisfying $f(X)\xi = \|f(X)\|\xi$ and $f(Y)\eta = \|f(Y)\|\eta$. Then for rank-one projections $P = \xi\otimes\xi^c$ and $Q = \eta\otimes\eta^c$ we have $XP = PX$, $YQ = QY$, $f(X)P = \|f(X)\|P$ and $f(Y)Q = \|f(Y)\|Q$. Then we see that $\langle f(X)\eta,\eta\rangle Q = Qf(X)Q$ and $f(\langle Y\eta,\eta\rangle)Q = \|f(Y)\|Q$. Therefore by the assumption we have $Qf(X)Q \ge \|f(Y)\|Q$ and hence $\|f(X)\|Q \ge Qf(X)Q \ge \|f(Y)\|Q$. By the similar way we see that $\|f(Y)\|P \ge Pf(Y)P \ge \|f(X)\|P$. Hence we get $\|f(X)\| = \|f(Y)\|$ and $Qf(X)Q = \|f(X)\|Q$. Since

$$0 = Q(\|f(X)\| - f(X))Q = Q(\|f(X)\| - f(X))^{\frac{1}{2}}(\|f(X)\| - f(X))^{\frac{1}{2}}Q,$$

we have

$$Qf(X) = f(X)Q = \|f(X)\|Q = \|f(Y)\|Q = f(Y)Q$$

and hence $QX = XQ = YQ$. (Here we use the existence of $f^{-1}(t)$.) Since two matrices $X(1-Q)$ and $Y(1-Q)$ satisfy the same assumptions on $(1-Q)\boldsymbol{C}^n$, we can repeat this argument. Therefore we get $X = Y$. □

Corollary 2.2. *Let $f(t)$ be a continuous strictly increasing (or decreasing) concave function on an interval I and let $X, Y \in M_n(\boldsymbol{C})$ be two hermitian matrices whose numerical ranges are contained in I. If X and Y satisfy*

$$\langle f(X)\xi, \xi\rangle \leq f(\langle Y\xi, \xi\rangle)$$

and

$$\langle f(Y)\xi, \xi\rangle \leq f(\langle X\xi, \xi\rangle)$$

for any unit vector $\xi \in \boldsymbol{C}^n$, then we have $X = Y$.

Proof. Apply the previous theorem to the function $-f(t)$. □

Remark 2.1. If $f(X)$ and $f(Y)$ are of the forms

$$f(X) = \sum_{i=1}^{\infty} \lambda_i P_i \quad f(Y) = \sum_{j=1}^{\infty} \mu_j Q_j$$

where $\{P_i\}_i$ and $\{Q_j\}_j$ are orthogonal families of projections and $\lambda_1 \geq \lambda_2 \geq \cdots$ and $\mu_1 \geq \mu_2 \geq \cdots$, then Theorem 2.1 holds by the same proof. For example, if both X and Y are compact positive and $f(t)$ is strictly increasing, then $f(X)$ and $f(Y)$ are of the above forms.

3. Infinite dimensional case.

Let $f(t)$ and $g(t)$ be positive, strictly increasing, concave C^2-functions on $(0, \infty)$ and continuous on $[0, \infty)$. For a positive operator A, by the Jensen inequality we have

$$\langle (g \circ f)(A)\xi, \xi\rangle \leq g(\langle f(A)\xi, \xi\rangle) \leq (g \circ f)(\langle A\xi, \xi\rangle)$$

for any unit vector $\xi \in \mathfrak{H}$. We would like to consider the "converse" of this fact.

THEOREM 3.1. *Let $f(t)$ and $g(t)$ be positive, strictly increasing, concave C^2-functions on $(0,\infty)$ and continuous on $[0,\infty)$. If two positive operators X and Y on $\mathfrak{H}$ satisfy*

$$\langle (g \circ f)(X)\xi, \xi\rangle \le g(\langle f(Y)\xi, \xi\rangle) \le (g \circ f)(\langle X\xi, \xi\rangle)$$

for any unit vector $\xi \in \mathfrak{H}$, then we have $X = Y$.

For example consider the case $f(t) = g(t) = \sqrt{t}$. Then we have;

EXAMPLE 3.1. If two positive operators X and Y on $\mathfrak{H}$ satisfy

$$\langle X^{\frac{1}{4}}\xi, \xi\rangle \le \langle Y^{\frac{1}{2}}\xi, \xi\rangle^{\frac{1}{2}} \le \langle X\xi, \xi\rangle^{\frac{1}{4}}$$

for any unit vector $\xi \in \mathfrak{H}$, then we have $X = Y$.

The strategy for our proof is essentially same as that of [**1**], [**4**].

LEMMA 3.2 (Ando [**1**]). *Let $h(t)$ be a positive, strictly increasing, concave C^2-function on $(0,\infty)$ and continuous on $[0,\infty)$. For positive operators A and B, the inequality*

$$\langle h(A)\xi, \xi\rangle \le h(\langle B\xi, \xi\rangle)$$

holds for any unit vector $\xi \in \mathfrak{H}$ if and only if we have

$$h(A) \le h'(\lambda)B - \lambda h'(\lambda) + h(\lambda)$$

for any positive number λ.

PROOF. First we will show the "only if" part. Since $h(t)$ is concave, we have

$$h(t) \le h'(\lambda)t - \lambda h'(\lambda) + h(\lambda).$$

(The right-hand side is the tangent line to $h(t)$ at $t = \lambda$.) Letting $t = \langle B\xi, \xi\rangle$, we get

$$h(\langle B\xi, \xi\rangle) \le h'(\lambda)\langle B\xi, \xi\rangle - \lambda h'(\lambda) + h(\lambda) = \langle \{h'(\lambda)B - \lambda h'(\lambda) + h(\lambda)\}\xi, \xi\rangle.$$

Combining this with the inequality $\langle h(A)\xi, \xi\rangle \le h(\langle B\xi, \xi\rangle)$, we see that

$$h(A) \le h'(\lambda)B - \lambda h'(\lambda) + h(\lambda).$$

Conversely if

$$h(A) \le h'(\lambda)B - \lambda h'(\lambda) + h(\lambda)$$

holds for any $\lambda > 0$, we see that for any unit vector $\xi \in \mathfrak{H}$

$$\langle h(A)\xi, \xi\rangle \le \langle (h'(\lambda)B - \lambda h'(\lambda) + h(\lambda))\xi, \xi\rangle = h'(\lambda)\langle B\xi, \xi\rangle - \lambda h'(\lambda) + h(\lambda).$$

Then it is easy to see that the minimal value of the right-hand side over $\lambda > 0$ is equal to $h(\langle B\xi, \xi\rangle)$. □

LEMMA 3.3. *Under the assumptions in Theorem* 3.1*, we have*

$$\begin{aligned}\frac{(g\circ f)(X) + f(\lambda)g'(f(\lambda)) - g(f(\lambda))}{g'(f(\lambda))} &\le f(Y)\\ &\le f'(\lambda)X - \lambda f'(\lambda) + f(\lambda)\end{aligned}$$

for any positive number λ.

PROOF. By the assumptions we have two inequalities

$$\langle g(f(X))\xi, \xi\rangle \le g(\langle f(Y)\xi, \xi\rangle)$$

and

$$\langle f(Y)\xi, \xi\rangle \le f(\langle X\xi, \xi\rangle)$$

for any unit vector $\xi \in \mathfrak{H}$. So by the previous lemma we get

$$g(f(X)) \le g'(\mu)f(Y) - \mu g'(\mu) + g(\mu)$$

and

$$f(Y) \le f'(\lambda)X - \lambda f'(\lambda) + f(\lambda)$$

for any positive numbers μ and λ. Letting $\mu = f(\lambda)$ we get the desired inequality. □

Lemma 3.4. *Fix two positive numbers $0 < a < b$. Then there exists a positive constant c (depending on the choice of a, b) satisfying*

$$f'(\lambda)t - \lambda f'(\lambda) + f(\lambda) - \left\{ \frac{(g \circ f)(t) + f(\lambda)g'(f(\lambda)) - g(f(\lambda))}{g'(f(\lambda))} \right\} \leq c(t-\lambda)^2$$

for any $a \leq \lambda \leq b$ and $a \leq t \leq b$.

Proof. Set

$$\begin{aligned} k(t) = k_\lambda(t) &= c(t-\lambda)^2 - f'(\lambda)t + \lambda f'(\lambda) - f(\lambda) \\ &\quad + \left\{ \frac{(g \circ f)(t) + f(\lambda)g'(f(\lambda)) - g(f(\lambda))}{g'(f(\lambda))} \right\}. \end{aligned}$$

with c to be determined later. Fix λ and we consider $k(t)$ as a function of one variable. Then we see that

$$k'(t) = 2c(t-\lambda) - f'(\lambda) + \frac{(g' \circ f)(t)f'(t)}{g'(f(\lambda))}$$

and

$$k''(t) = 2c + \frac{(g'' \circ f)(t)f'(t)^2 + (g' \circ f)(t)f''(t)}{g'(f(\lambda))}.$$

By the assumptions we can take c with $k''(t) > 0$ for any $a \leq \lambda \leq b$ and $a \leq t \leq b$. Then since $k'(\lambda) = 0$, we have $k'(t) \leq 0$ $(t \leq \lambda)$ and $k'(t) \geq 0$ $(t \geq \lambda)$. Hence we have $k(t) \geq k(\lambda) = 0$. □

Choose $b > 0$ satisfying $\|X\|, \|Y\| < b$ at first. Then, choose and fix a with $0 < a < b$. We can find a positive number α (depending on the choice of a, b) satisfying

$$\frac{(g \circ f)(t) + f(\lambda)g'(f(\lambda)) - g(f(\lambda))}{g'(f(\lambda))} + \alpha \geq 1$$

for any $a \leq \lambda \leq b$ and $a \leq t \leq b$.

Lemma 3.5. *There exists a positive constant c satisfying*

$$\left\{\frac{(g\circ f)(t)+f(\lambda)g'(f(\lambda))-g(f(\lambda))}{g'(f(\lambda))}+\alpha\right\}^{-1}-\{f'(\lambda)t-\lambda f'(\lambda)+f(\lambda)+\alpha\}^{-1}$$
$$\leq c(t-\lambda)^2$$

for any $a\leq\lambda\leq b$ and $a\leq t\leq b$. The constant c is same as that of the previous lemma.

Proof. Set

$$p(t)=f'(\lambda)t-\lambda f'(\lambda)+f(\lambda)+\alpha$$

and

$$q(t)=\frac{(g\circ f)(t)+f(\lambda)g'(f(\lambda))-g(f(\lambda))}{g'(f(\lambda))}+\alpha.$$

Since f and g are concave we have

$$f(t)\leq f'(\lambda)t-\lambda f'(\lambda)+f(\lambda)$$

and

$$g(s)\leq g'(\mu)s-\mu g'(\mu)+g(\mu).$$

Letting $\mu=f(\lambda)$ and $s=f(t)$, we get

$$g(f(t))\leq g'(f(\lambda))f(t)-f(\lambda)g'(f(\lambda))+g(f(\lambda))$$

and hence

$$\frac{(g\circ f)(t)+f(\lambda)g'(f(\lambda))-g(f(\lambda))}{g'(f(\lambda))}\leq f(t).$$

Therefore we have $p(t)\geq f(t)+\alpha\geq q(t)\geq 1$. Then by the previous lemma we have $p(t)-q(t)\leq c(t-\lambda)^2$. So we get

$$q(t)^{-1}-p(t)^{-1}=q(t)^{-1}p(t)^{-1}(p(t)-q(t))\leq c(t-\lambda)^2. \qquad \square$$

Proof of Theorem 3.1. Take a spectral projection P of X. By Lemma 3.3 we have

$$\begin{aligned}&\left\{\frac{(g\circ f)(X)+f(\lambda)g'(f(\lambda))-g(f(\lambda))}{g'(f(\lambda))}+\alpha\right\}P\\&\quad\le P(f(Y)+\alpha)P\\&\quad\le \big\{(f'(\lambda)X-\lambda f'(\lambda)+f(\lambda))+\alpha\big\}P\end{aligned}$$

for any positive number λ. On the other hand we have

$$\begin{aligned}&\left\{\frac{(g\circ f)(X)+f(\lambda)g'(f(\lambda))-g(f(\lambda))}{g'(f(\lambda))}+\alpha\right\}P\\&\quad\le (f(X)+\alpha)P\\&\quad\le \big\{(f'(\lambda)X-\lambda f'(\lambda)+f(\lambda))+\alpha\big\}P\end{aligned}$$

for any positive number λ. Combining these with Lemma 3.4 we get

$$\|(f(X)+\alpha)P-P(f(Y)+\alpha)P\|\le c\|XP-\lambda P\|^2 \tag{1}$$

whenever $P\le \chi_{[a,b)}(X)$ and $a\le\lambda\le b$.

Similarly since we have two inequalities

$$\begin{aligned}&\big\{(f'(\lambda)X-\lambda f'(\lambda)+f(\lambda))+\alpha\big\}^{-1}P\\&\quad\le P(f(Y)+\alpha)^{-1}P\\&\quad\le \left\{\frac{(g\circ f)(X)+f(\lambda)g'(f(\lambda))-g(f(\lambda))}{g'(f(\lambda))}+\alpha\right\}^{-1}P\end{aligned}$$

and

$$\begin{aligned}&\big\{(f'(\lambda)X-\lambda f'(\lambda)+f(\lambda))+\alpha\big\}^{-1}P\\&\quad\le (f(X)+\alpha)^{-1}P\\&\quad\le \left\{\frac{(g\circ f)(X)+f(\lambda)g'(f(\lambda))-g(f(\lambda))}{g'(f(\lambda))}+\alpha\right\}^{-1}P,\end{aligned}$$

by lemma 3.5 we get

$$\left\|(f(X)+\alpha)^{-1}P - P(f(Y)+\alpha)^{-1}P\right\| \le c\|XP-\lambda P\|^2$$

whenever $P \le \chi_{[a,b)}(X)$ and $a \le \lambda \le b$. Let us use $(P(f(Y)+\alpha)^{-1}P)^{-1}$ to denote the inverse of $P(f(Y)+\alpha)^{-1}P$ on $P\mathfrak{H}$. Then we have

$$\begin{aligned}
&\left\|(f(X)+\alpha)P - (P(f(Y)+\alpha)^{-1}P)^{-1}\right\| \\
&\quad = \left\|(f(X)+\alpha)\{P(f(Y)+\alpha)^{-1}P - (f(X)+\alpha)^{-1}P\}(P(f(Y)+\alpha)^{-1}P)^{-1}\right\| \\
&\quad \le \|f(X)+\alpha\| \cdot \left\|(P(f(Y)+\alpha)^{-1}P)^{-1}\right\| \\
&\qquad \cdot \left\|(P(f(Y)+\alpha)^{-1}P - (f(X)+\alpha)^{-1}P\right\| \\
&\quad \le (f(b)+\alpha)^2 c\|XP-\lambda P\|^2
\end{aligned}$$

and hence

$$\left\|(f(X)+\alpha)P - (P(f(Y)+\alpha)^{-1}P)^{-1}\right\| \le (f(b)+\alpha)^2 c\|XP-\lambda P\|^2. \tag{2}$$

Here we remark that since $f(Y)+\alpha \le f(b)+\alpha$, we have $(f(Y)+\alpha)^{-1} \ge (f(b)+\alpha)^{-1}$ and $P(f(Y)+\alpha)^{-1}P \ge (f(b)+\alpha)^{-1}P$ and hence $(P(f(Y)+\alpha)^{-1}P)^{-1} \le (f(b)+\alpha)P$. Thus we conclude $\|(P(f(Y)+\alpha)^{-1}P)^{-1}\| \le f(b)+\alpha$. We used this inequality in the proof of (2).

Therefore for $P \le \chi_{[a,b)}(X)$ and $a \le \lambda \le b$ by using (1) and (2) we have

$$\left\|P(f(Y)+\alpha)P - (P(f(Y)+\alpha)^{-1}P)^{-1}\right\| \le (1+(f(b)+\alpha)^2)c\|XP-\lambda P\|^2. \tag{3}$$

The rest of the proof is almost same as that of [**1**], [**4**]. We include this for the reader's convenience.

For each integer n, let P_i $(i = 1, 2, \ldots, n)$ be the spectral projections of X corresponding to the interval $[a+(i-1)(b-a)/n, a+i(b-a)/n)$. Then we have $\sum_i P_i = \chi_{[a,b)}(X)$ and

$$\|XP_i - \lambda_i P_i\| \le \frac{b-a}{n}$$

where $\lambda_i = a + \frac{(i-1)(b-a)}{n}$. Then it follows from (1) that

$$\left\|\sum_{i=1}^{n}\left\{(f(X)+\alpha)P_i - P_i(f(Y)+\alpha)P_i\right\}\right\| \le \frac{c(b-a)^2}{n^2}. \tag{4}$$

Similarly it follows from (3) that

$$\left\|P_i(f(Y)+\alpha)P_i-(P_i(f(Y)+\alpha)^{-1}P_i)^{-1}\right\|\le\frac{(1+(f(b)+\alpha)^2)c(b-a)^2}{n^2}.$$

The well-known formula

$$\begin{aligned}&\big(P_i(f(Y)+\alpha\big)^{-1}P_i)^{-1}\\&\quad=P_i(f(Y)+\alpha)P_i-P_i(f(Y)+\alpha)P_i^{\perp}\big(P_i^{\perp}(f(Y)+\alpha)P_i^{\perp}\big)^{-1}P_i^{\perp}(f(Y)+\alpha)P_i\end{aligned}$$

with $P_i^{\perp}=1-P_i$ (known as the Schur complement) yields

$$\begin{aligned}&\left\|P_i^{\perp}(f(Y)+\alpha)P_i\right\|^2\\&\quad=\left\|(P_i^{\perp}(f(Y)+\alpha)P_i^{\perp})^{1/2}(P_i^{\perp}(f(Y)+\alpha)P_i^{\perp})^{-1/2}P_i^{\perp}(f(Y)+\alpha)P_i\right\|^2\\&\quad\le\|f(Y)+\alpha\|\cdot\left\|(P_i^{\perp}(f(Y)+\alpha)P_i^{\perp})^{-1/2}P_i^{\perp}(f(Y)+\alpha)P_i\right\|^2\\&\quad=\|f(Y)+\alpha\|\cdot\left\|P_i(f(Y)+\alpha)P_i^{\perp}(P_i^{\perp}(f(Y)+\alpha)P_i^{\perp})^{-1}P_i^{\perp}(f(Y)+\alpha)P_i\right\|\\&\quad=\|f(Y)+\alpha\|\cdot\left\|P_i(f(Y)+\alpha)P_i-(P_i(f(Y)+\alpha)^{-1}P_i)^{-1}\right\|\\&\quad\le\frac{(f(b)+\alpha)(1+(f(b)+\alpha)^2)c(b-a)^2}{n^2}.\end{aligned}$$

Therefore we see that

$$\begin{aligned}\left\|\sum_{i=1}^n P_i^{\perp}(f(Y)+\alpha)P_i\right\|^2&=\left\|\left\{\sum_{i=1}^n P_i^{\perp}(f(Y)+\alpha)P_i\right\}\left\{\sum_{j=1}^n P_j(f(Y)+\alpha)P_j^{\perp}\right\}\right\|\\&=\left\|\sum_{i=1}^n P_i^{\perp}(f(Y)+\alpha)P_i(f(Y)+\alpha)P_i^{\perp}\right\|\\&\le\sum_{i=1}^n\left\|P_i^{\perp}(f(Y)+\alpha)P_i(f(Y)+\alpha)P_i^{\perp}\right\|\\&=\sum_{i=1}^n\left\|P_i^{\perp}(f(Y)+\alpha)P_i\right\|^2\\&\le\sum_{i=1}^n\frac{(f(b)+\alpha)(1+(f(b)+\alpha)^2)c(b-a)^2}{n^2}\end{aligned}$$

$$= \frac{(f(b)+\alpha)(1+(f(b)+\alpha)^2)c(b-a)^2}{n}.$$

Thus we get

$$\left\| \sum_{i=1}^{n} P_i^{\perp}(f(Y)+\alpha)P_i \right\| \leq \sqrt{\frac{(f(b)+\alpha)(1+(f(b)+\alpha)^2)c(b-a)^2}{n}}. \tag{5}$$

Since

$$(f(Y)+\alpha)\chi_{[a,b)}(X) = \sum_{i=1}^{n} P_i(f(Y)+\alpha)P_i + \sum_{i=1}^{n} P_i^{\perp}(f(Y)+\alpha)P_i,$$

by using (4) and (5) we see that

$$\begin{aligned}
&\|f(X)\chi_{[a,b)}(X) - f(Y)\chi_{[a,b)}(X)\| \\
&\quad = \|(f(X)+\alpha)\chi_{[a,b)}(X) - (f(Y)+\alpha)\chi_{[a,b)}(X)\| \\
&\quad \leq \left\| \sum_{i=1}^{n} \{(f(X)+\alpha)P_i - P_i(f(Y)+\alpha)P_i\} \right\| + \left\| \sum_{i=1}^{n} P_i^{\perp}(f(Y)+\alpha)P_i \right\| \\
&\quad \leq \frac{c(b-a)^2}{n^2} + \sqrt{\frac{(f(b)+\alpha)(1+(f(b)+\alpha)^2)c(b-a)^2}{n}}.
\end{aligned}$$

By tending $n \to \infty$ we get $f(X)\chi_{[a,b)}(X) = f(Y)\chi_{[a,b)}(X)$. Since a is arbitrary we have $f(X)\chi_{(0,b)}(X) = f(Y)\chi_{(0,b)}(X)$. Therefore, it remains to show $\chi_{\{0\}}(X) = \chi_{\{0\}}(Y)$.

For any unit vector $\xi \in \mathfrak{H}$ with $X\xi = 0$, we see that

$$f(0) + \langle (f(Y) - f(0))\xi, \xi \rangle = \langle f(Y)\xi, \xi \rangle \leq f(\langle X\xi, \xi \rangle) = f(0).$$

Therefore $f(Y)\xi = f(0)\xi$ and hence $Y\xi = 0$. Conversely for any unit vector $\xi \in \mathfrak{H}$ with $Y\xi = 0$, we see that

$$\begin{aligned}
&(g \circ f)(0) + \langle ((g \circ f)(X) - (g \circ f)(0))\xi, \xi \rangle \\
&\quad = \langle (g \circ f)(X)\xi, \xi \rangle \leq g(\langle f(Y)\xi, \xi \rangle) = (g \circ f)(0).
\end{aligned}$$

Therefore $(g \circ f)(X)\xi = (g \circ f)(0)\xi$ and hence $X\xi = 0$. □

Remark 3.1.

(i) In lemma 3.4, the assumption $a > 0$ is crucial. For example if we consider the case $a = 0$ and $f(t) = g(t) = \sqrt{t}$, then lemma 3.4 is wrong. Indeed in this case

$$f'(\lambda)t - \lambda f'(\lambda) + f(\lambda) - \left\{ \frac{(g \circ f)(t) + f(\lambda)g'(f(\lambda)) - g(f(\lambda))}{g'(f(\lambda))} \right\}$$
$$= \frac{t}{2\sqrt{\lambda}} + \frac{3\sqrt{\lambda}}{2} - 2\lambda^{\frac{1}{4}}t^{\frac{1}{4}}.$$

It is easy to see that

$$\frac{1}{(t-\lambda)^2}\left\{ \frac{t}{2\sqrt{\lambda}} + \frac{3\sqrt{\lambda}}{2} - 2\lambda^{\frac{1}{4}}t^{\frac{1}{4}} \right\}$$

is unbounded for $0 < \lambda \le b$ and $0 < t \le b$. (Fix $t > 0$ and consider the case $\lambda \to +0$. Then this function tends to ∞.)

(ii) The argument in this section cannot be applied directly to the problem in the previous section. For simplicity, we would like to consider the case $f(t) = \sqrt{t}$. Let X and Y be positive operators on $\mathfrak{H}$. Suppose that they satisfy

$$\left\langle \sqrt{X}\xi, \xi \right\rangle \le \sqrt{\langle Y\xi, \xi \rangle)}$$

and

$$\left\langle \sqrt{Y}\xi, \xi \right\rangle \le \sqrt{\langle X\xi, \xi \rangle}$$

for any unit vector $\xi \in \mathfrak{H}$. Then by lemma 3.2 we have

$$\sqrt{X} \le \frac{1}{2\sqrt{\lambda}}Y + \frac{\sqrt{\lambda}}{2}$$

and

$$\sqrt{Y} \le \frac{1}{2\sqrt{\lambda}}X + \frac{\sqrt{\lambda}}{2}$$

for any $\lambda > 0$. By the first inequality we have

$$2\sqrt{\lambda X} - \lambda \leq Y.$$

Since the left-hand side in this inequality is not positive, we cannot take a square root. This is the main trouble. By this reason we cannot show the statement like Lemma 3.3.

References

[1] T. Ando, Functional calculus with operator-monotone functions, Math. Inequal. Appl., to appear.

[2] J-C. Bourin, private communication.

[3] J. B. Conway, A Course in Operator Theory, Graduate Studies in Mathematics, **21**, American Mathematical Society, Providence, RI, 2000.

[4] T. Hayashi, Non-commutative arithmetic-geometric mean inequality, Proc. Amer. Math. Soc., **137** (2009), 3399–3406.

Tomohiro Hayashi
Nagoya Institute of Technology
Gokiso-cho, Showa-ku, Nagoya
Aichi 466-8555, Japan
E-mail: hayashi.tomohiro@nitech.ac.jp