ON EQUALITY CONDITION FOR TRACE JENSEN INEQUALITY IN SEMI-FINITE VON NEUMANN ALGEBRAS

Harada, Tetsuo
Faculty of Mathematics, Kyushu University

Kosaki, Hideki
Faculty of Mathematics, Kyushu University

KYUSHU UNIVERSITY

ON EQUALITY CONDITION FOR TRACE JENSEN INEQUALITY IN SEMI-FINITE VON NEUMANN ALGEBRAS

TETSUO HARADA AND HIDEKI KOSAKI

ABSTRACT. Let τ be a faithful semi-finite normal trace on a semi-finite von Neumann algebra, and $f(t)$ be a convex function with $f(0) = 0$. The trace Jensen inequality states $\tau\big(f(a^*xa)\big) \leq \tau\big(a^*f(x)a\big)$ for a contraction a and a self-adjoint operator x. Under certain strict convexity assumption on $f(t)$ we will study when this inequality reduces to the equality.

1. INTRODUCTION

Jensen-type operator inequalities for operator convex functions have been investigated by many authors (see §2.1 for information). For example it was shown in [9] that an operator convex function $f(t)$ with $f(0) = 0$ satisfies $f(a^*xa) \leq a^*f(x)a$ for a self-adjoint operator x and a contraction a, and hence the numerical inequality $\tau(f(a^*xa)) \leq \tau(a^*f(x)a)$ holds true with a trace τ on a von Neumann algebra. Operator convexity here is quite a strong condition, and it is known that the ordinary convexity is actually sufficient for such a trace inequality. Indeed, the trace inequality

$$\tau\big(a^*g(x)a\big) \leq \tau\big(g(a^*xa)\big)$$

was shown (as long as the both sides are well-defined) in [3] for a concave function with $g(0) = 0$ and a positive τ-measurable operator x. The closely related inequality

$$\tau\big(\alpha(f(x))\big) \geq \tau\big(f(\alpha(x))\big)$$

for a positive contractive map $\alpha : \mathcal{M} \to \mathcal{M}$ and a convex function $f(t)$ with $f(0) = 0$ (see [18]) is also known. In our recent article [11] among other results the following trace Jensen inequality was shown:

$$(1) \qquad \tau\big(f(a^*xa)\big) \leq \tau\big(a^*f(x)a\big)$$

(as long as the both sides are well-defined) for a convex function $f(t)$ with $f(0) = 0$ and a semi-bounded τ-measurable operator x (i.e., $x \geq -m$ or $x \leq m$ for some $m \in \mathbf{R}_+$).

The purpose of the present article is to investigate when the trace Jensen inequality (1) reduces to the equality. Of course certain strict convexity of $f(t)$ has to be assumed here. The "ordinary strict convexity" (i.e., $f(\lambda t_1 + (1 - \lambda)t_2) < \lambda f(t_1) + (1 - \lambda)f(t_2)$ as long as $t_1 \neq t_2$, $0 < \lambda < 1$) is desirable, However, we have to require stronger strict convexity in our operator algebra setting and "absence of eigenvectors" is responsible for this difficulty.

We explain our approach to the above problem along with the organization of the article. Let $g(t)$ be a non-affine operator convex function on a suitable interval I ($\ni 0$) satisfying $g(0) = 0$. We at first study when the operator Jensen inequality $a^*g(x)a \geq g(a^*xa)$ reduces to the equality, and $a^*x^2a = (a^*xa)^2$ turns out to be a necessary and sufficient condition (whose meaning is clarified in Remark 4). This fact (almost implicit in [4, 18]) is proved in Corollary 3 (in §3 for an interval I of finite length) and Theorem 9 (in §4 for $I = [0, \infty)$), and might be of independent interest. Let us assume that a function $f(t)$ is of the form $f = \phi \circ g$ with a convex and strictly increasing function $\phi(t)$ for instance. The equality in (1) for this composition $f(t)$ is shown to imply $a^*g(x)a = g(a^*xa)$ and hence $a^*x^2a = (a^*xa)^2$. Here, monotonicity properties on trace values obtained in [11, §5] (which are summarized in **2.2**) play crucial roles. An interval of finite length (on which $f(t)$ is defined) is considered in §3 whereas we deal with the case of the half line $[0, \infty)$ in §4. For example, when I is a finite interval, a C^2-function $f(t)$ with strictly positive $f''(t)$ can be expressed in the form $\phi \circ g$ as above (after adding a suitable linear term). Our investigation is somewhat indirect, but our equality condition will be shown to cover quite many situations and to have an abundance of applications.

Validity of our equality condition under the ordinary strict convexity seems to deserve investigation, and the authors hope to come back to this problem in future. As a supporting evidence, in Appendix we will show that it is indeed the case for matrices or more generally for compact operators.

2. PRELIMINARIES

2.1. Operator monotone and operator convex functions. A continuous function $f(t)$ on an interval I is said to be operator monotone (on I) when

$$f(x) \leq f(y)$$

holds true for any pair (x, y) of self-adjoint operators with spectra included in I satisfying $x \leq y$. Löwner's theory (see [6] for details) shows that an operator monotone function $f(t)$ is characterized as a Pick function, i.e., $f(t)$ can be extended to an analytic function on the open upper half plane with positive imaginary part.

A continuous function $f(t)$ on an interval I is said to be operator convex (on I) when we have

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

for each $\lambda \in [0, 1]$ and self-adjoint operators x, y with spectra included in I. When $-f(t)$ is operator convex, we say that $f(t)$ is operator concave. It is known (see [1, Theorem 3.3]) that $f(t)$ is operator convex if and only if

$$\frac{f(t) - f(t_0)}{t - t_0}$$

is operator monotone for each $t_0 \in I$.

It was shown in [4, 5] that an operator convex function $f(t)$ satisfies

$$\pi(f(x)) \geq f(\pi(x))$$

for a positive unital map π (which is closely related to the classical Kadison inequality [14]). Another related result is the operator Jensen inequality ([9]) stating

$$a^* f(x) a \geq f(a^* x a),$$

where x is a self-adjoint operator (with the spectrum $\sigma(x)$ sitting in I) and a is a contraction (i.e., $||a|| \leq 1$). The possibility $a = 0$ is not excluded here so that $0 \in I$ has to be assumed in order that the right side makes sense. Note that $f(0) \leq 0$ has to be also assumed (indeed, the inequality means $f(0)a^*a \geq f(0)1$ for $x = 0$). A readable account on these and related subjects can be found in [1, 9, 10] for instance.

An equality condition for the first inequality (for $f(t)$ non-affine operator convex) was studied in [4, 17]. Namely, it reduces to the equality for every self-adjoint x if and only if π is a C^*-homomorphism, i.e., $\pi(x^2) = \pi(x)^2$ ([4, Theorem 2.5]). On the other hand, it was shown in [17] that the equality holds for a single self-adjoint x if and only if π is multiplicative on the algebra generated by x. Certain integral representations for operator monotone and operator convex functions are known, which are essential tools for these characterizations as well as our Theorem 2 and Theorem 9 in later sections.

2.2. Monotonicity on trace values. Let $\mathcal{M}$ be a semi-finite von Neumann algebra equipped with a faithful semi-finite normal trace τ (throughout the article). A densely-defined closed operator x affiliated with $\mathcal{M}$ is said to be τ-measurable if for any $\varepsilon > 0$ one can find a projection $e \in \mathcal{M}$ such that the subspace cut by e is included in the domain of x and $\tau(1 - e) \leq \varepsilon$. Basic facts on τ-measurable operators can be found in [16, 20] for instance. The only thing needed in this article is the following standard fact: the set of all τ-measurable operators forms a $*$-algebra and one can treat them without worrying "domain problems".

Let y be a self-adjoint τ-measurable operator (with the Jordan decomposition $y = y_+ - y_-$). When $\tau(y_+) < \infty$ or $\tau(y_-) < \infty$, we say that $\tau(y)$ is well-defined and we set

$$\tau(y) = \tau(y_+) - \tau(y_-).$$

Note that $\tau(y)$ is well-defined as long as $y \in L^1(\mathcal{M}; \tau)$, i.e., $\tau(|y|) < \infty$. The following monotonicity properties (obtained in [11, Theorem 16]) will be repeatedly used in sequel:

Theorem 1. *Let $f(t)$ be a continuous increasing function (on $\mathbf{R}$) with $f(0) = 0$, and we assume that self-adjoint τ-measurable operators x, y satisfy $x \leq y$.*

(i) *The trace inequality $\tau(f(x)) \leq \tau(f(y))$ holds true as long as the both sides are well-defined.*

(ii) *We further assume strict increasingness of $f(t)$ and integrability x, y, $f(x)$, $f(y) \in L^1(\mathcal{M}; \tau)$. Then, we must have $x = y$ as long as the equality $\tau(f(x)) = \tau(f(y))$ holds.*

(iii) *We have the same conclusion as (ii) under convexity and strict increasingness of $f(t)$ and integrability $f(x), f(y) \in L^1(\mathcal{M}; \tau)$.*

2.3. Spectral dominance. For positive τ-measurable operators x, y the spectral dominance $x \gtrsim y$ (or $y \lesssim x$) means that their spectral projections satisfy

$$e^x_{(s,\infty)} \gtrsim e^y_{(s,\infty)} \quad (s > 0)$$

in the Murray-von Neumann sense. This means that there exists a partial isometry $u \in \mathcal{M}$ satisfying $e^x_{(s,\infty)} \geq uu^*$ and $u^*u = e^y_{(s,\infty)}$. Note that $e^x_{(s,\infty)} \gtrsim e^y_{(s,\infty)}$ yields $\tau\left(e^x_{(s,\infty)}\right) \geq \tau\left(e^y_{(s,\infty)}\right)$. From this we easily conclude that the spectral dominance $x \gtrsim y$ implies $\tau(g(x)) \geq \tau(g(y))$ for each continuous increasing function $g(t)$ on $[0, \infty)$ satisfying $g(0) = 0$ (see [3, Lemma 4]).

3. Equality condition (finite interval)

In this section (and the next) we will investigate when the trace Jensen inequality (1) reduces to the equality. We begin by pointing out that some modifications of reasoning in [4, 17] enable us to prove the next equality condition for operator Jensen inequality. This result (as well as Corollary 3) could be useful by itself, on which our analysis here will heavily rely.

Theorem 2. *We assume that $g(t)$ is a non-affine operator convex function on an interval (α, β) containing the origin and $g(0) = 0$. For a self-adjoint operator x (with $\sigma(x) \subset (\alpha, \beta)$) and a projection e the operator Jensen inequality reduces to the equality $g(exe) = eg(x)e$ if and only if $ex = xe$.*

Proof. The commutativity $ex = xe$ obviously implies $g(exe) = eg(x)e$ (due to $g(0) = 0$). To show the converse, let us assume $g(exe) = eg(x)e$. We set

$$r = \frac{\beta - \alpha}{2}, \quad t_0 = \frac{\alpha + \beta}{2}$$

for convenience, i.e., $(\alpha, \beta) = (t_0 - r, t_0 + r)$ and make use of the well-known integral expression

$$g(t) = p + qt + \int_{-1/r}^{1/r} \frac{(t - t_0)^2}{1 - s(t - t_0)} \, d\mu(s)$$

(see [1, Corollary 2 in p.69] or [9, Theorem 4.5] for instance). Here, μ is a measure on $[-1/r, 1/r]$ (with total mass $g''(t_0)/2$) and $p, q \in \mathbf{R}$. Note $\mu \neq 0$ (since $g(t)$ is non-affine) and

$$0 = g(0) = p + \int_{-1/r}^{1/r} \frac{t_0^2}{1 + st_0} \, d\mu(s).$$

Subtracting this from the expression for $g(t)$, we arrive at

$$g(t) = qt + \int_{-1/r}^{1/r} g_s(t) \, d\mu(s) \quad \text{with} \quad g_s(t) = \frac{(t - t_0)^2}{1 - s(t - t_0)} - \frac{t_0^2}{1 + st_0}.$$

Each integrand is operator convex on $(t_0 - r, t_0 + r)$ (i.e., $eg_s(x)e \geq g_s(exe)$ in general $\cdots$ see also the paragraph right after the proof), and hence the assumption $g(exe) = eg(x)e$ means

$$eg_s(x)e = g_s(exe) \quad \text{for } s \in [-1/r, 1/r] \ \mu\text{-a.e.}$$

We note

$$eg_s(x)e = e(x - t_0)^2 \left(1 - s(x - t_0)\right)^{-1} e - t_0^2 (1 + st_0)^{-1} e.$$

Notice that $g_s(exe)$ also lives on the subspace cut by the projection e because of $g(0) = 0$. Thus, $g_s(exe) \ (= eg_s(exe)e)$ is given by

$$
\begin{aligned}
g_s(exe) &= e\left[(exe - t_0)^2 \left(1 - s(exe - t_0)\right)^{-1} - t_0^2 (1 + st_0)^{-1}\right] e \\
&= (exe - t_0 e)^2 \left(e - s(exe - t_0 e)\right)^{-1} - t_0^2 (1 + st_0)^{-1} e \\
&= \left(e(x - t_0)e\right)^2 \left(e\left(1 - s(x - t_0)\right)e\right)^{-1} - t_0^2 (1 + st_0)^{-1} e
\end{aligned}
$$

with inverses as operators acting on the subspace. Thus, the above operator equality means

$$(2) \qquad e(x - t_0)^2 \left(1 - s(x - t_0)\right)^{-1} e = \left(e(x - t_0)e\right)^2 \left(e\left(1 - s(x - t_0)\right)e\right)^{-1}$$

for $s \in [-1/r, 1/r] \ \mu\text{-a.e.}$

   If $\mu(\{0\}) > 0$, then we must have

$$e(x - t_0)^2 e = \left(e(x - t_0)e\right)^2$$

(i.e., (2) with $s = 0$). This says $ex^2e = (exe)^2$, implying $ex = xe$ due to

$$0 = ex^2e - (exe)^2 = ex(1 - e)xe.$$

On the other hand, when $\mu(\{0\}) = 0$, (2) is valid for some $s \in [-1/r, 1/r] \setminus \{0\}$. In this case we set $y = 1 - s(x - t_0)$ (i.e., $x - t_0 = s^{-1}(1 - y)$). We note that y (≥ 0) admits a bounded inverse and the both sides of (2) are computed as follows:

$$
\begin{aligned}
e(x - t_0)^2\left(1 - s(x - t_0)\right)^{-1}e &= s^{-2}e(1 - y)^2y^{-1}e \\
&= s^{-2}e(1 - 2y + y^2)y^{-1}e = s^{-2}\{ey^{-1}e - 2e + eye\}, \\
\left(e(x - t_0)e\right)^2\left(e\left(1 - s(x - t_0)\right)e\right)^{-1} &= s^{-2}(e(1 - y)e)^2(eye)^{-1} \\
&= s^{-2}\{e - 2eye + (eye)^2\}(eye)^{-1} = s^{-2}\{(eye)^{-1} - 2e + eye\}.
\end{aligned}
$$

We thus get $ey^{-1}e = (eye)^{-1}$ from which $ey = ye$ is to be shown.

To do so, we set

$$y = \begin{bmatrix} y_{11} & z \\ z^* & y_{22} \end{bmatrix}$$

(relative to the decomposition $1 = e + (1 - e)$). Note that y_{11}, y_{22} are positive and invertible (since so is y), and let us recall the well-known computation

$$y = \begin{bmatrix} 1 & zy_{22}^{-1} \\ 0 & 1 \end{bmatrix}\begin{bmatrix} y_{11} - zy_{22}^{-1}z^* & 0 \\ 0 & y_{22} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ y_{22}^{-1}z^* & 1 \end{bmatrix}$$

(to deal with the "Schur complement" $y_{11} - zy_{22}^{-1}z^*$ $\cdots$ see [13, §7.7] for instance). We note

$$\begin{bmatrix} 1 & zy_{22}^{-1} \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -zy_{22}^{-1} \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 \\ y_{22}^{-1}z^* & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -y_{22}^{-1}z^* & 1 \end{bmatrix}.$$

Therefore, the $(1, 1)$ block $y_{11} - zy_{22}^{-1}z^*$ in the above middle matrix is also invertible (and positive) and we have

$$y^{-1} = \begin{bmatrix} 1 & 0 \\ -y_{22}^{-1}z^* & 1 \end{bmatrix}\begin{bmatrix} \left(y_{11} - zy_{22}^{-1}z^*\right)^{-1} & 0 \\ 0 & y_{22}^{-1} \end{bmatrix}\begin{bmatrix} 1 & -zy_{22}^{-1} \\ 0 & 1 \end{bmatrix}.$$

We thus compute

$$
\begin{aligned}
ey^{-1}e &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -y_{22}^{-1}z^* & 1 \end{bmatrix} \begin{bmatrix} \left(y_{11} - zy_{22}^{-1}z^*\right)^{-1} & 0 \\ 0 & y_{22}^{-1} \end{bmatrix} \begin{bmatrix} 1 & -zy_{22}^{-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \left(y_{11} - zy_{22}^{-1}z^*\right)^{-1} & 0 \\ 0 & y_{22}^{-1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} \left(y_{11} - zy_{22}^{-1}z^*\right)^{-1} & 0 \\ 0 & 0 \end{bmatrix}.
\end{aligned}
$$

Thus, if $ey^{-1}e = (eye)^{-1}$, then we have $zy_{22}^{-1}z^* = 0$, implying $z = 0$ and $[y, e] = 0$. $\square$

From the above computation we generally have

$$
ey^{-1}e \geq (eye)^{-1}.
$$

Here, the left side is understood as an operator acting on the subspace while the inverse in the right side should be taken as an operator on this subspace. Observe that $ey^{-1}e \geq (eye)^{-1}$ implies $eg_s(x)e \geq g_s(exe)$ (by reversing the preceding arguments).

A similar characterization for operator convex functions on intervals of the form $[0, \alpha)$ (with possibly $\alpha = \infty$) will be shown later (Theorem 9). The next generalization is based on the standard "dilation technique" (used in [9, p.230] for instance), which will be repeatedly used.

Corollary 3. *We assume that $g(t)$ is a non-affine operator convex function on an interval (α, β) containing the origin and $g(0) = 0$. For a contraction a and a self-adjoint operator x with $\sigma(x) \subset (\alpha, \beta)$ the following conditions are equivalent:*

   (i) $g(a^*xa) = a^*g(x)a$;
   (ii) $a^*x^2a = (a^*xa)^2$.

Proof. We set

$$
U = \begin{bmatrix} a & b \\ c & -a^* \end{bmatrix} \quad \text{with } b = (1 - aa^*)^{1/2}, \ c = (1 - a^*a)^{1/2},
$$

and we also set

$$
T = \begin{bmatrix} x & 0 \\ 0 & 0 \end{bmatrix}, \quad E = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.
$$

We note

$$
g(EU^*TUE) = g\left(\begin{bmatrix} a^*xa & 0 \\ 0 & 0 \end{bmatrix} \right) = \begin{bmatrix} g(a^*xa) & 0 \\ 0 & g(0) \end{bmatrix} = \begin{bmatrix} g(a^*xa) & 0 \\ 0 & 0 \end{bmatrix}.
$$

On the other hand, since U is a unitary operator, we compute

$$Eg(U^*TU)E = EU^*g(T)UE = \begin{bmatrix} a^*g(x)a & 0 \\ 0 & 0 \end{bmatrix}.$$

Theorem 2 says that $g(a^*xa) = a^*g(x)a$ if and only if $U^*TUE = EU^*TU$, but we compute

$$U^*TUE = \begin{bmatrix} a^*xa & 0 \\ b^*xa & 0 \end{bmatrix}, \quad EU^*TU = \begin{bmatrix} a^*xa & a^*xb \\ 0 & 0 \end{bmatrix}.$$

Thus, the latter condition is the same as $a^*xb = 0$, or equivalently

(3) $0 = a^*xb^2xa = a^*x(1 - aa^*)xa.$

□

Remark 4. *The condition $a^*x^2a = (a^*xa)^2$ (in the preceding corollary) is also equivalent to*

$$a^*x(1 - a^*a)^{1/2} \quad \text{or equivalently} \quad a^*x(1 - a^*a) = 0$$

(see (3)). This condition implies $aa^*x(1 - aa^*) = 0$, i.e., $aa^*x = aa^*xaa^*$ $(= xaa^*)$. Let $a^* = u|a^*|$ be the polar decomposition and $\{e^{|a^*|}\}$ be the spectral projections for $|a^*|$ $(0 \le |a^*| \le 1)$. We have $[|a^*|, x] = 0$, and the condition $a^*x^2a = (a^*xa)^2$ is obviously equivalent to $|a^*|x^2|a^*| = (|a^*|x|a^*|)^2$, that is,

$$x^2|a^*|^2(1 - |a^*|^2) = 0.$$

The projection onto the closure of the range of $|a^*|^2(1 - |a^*|^2)$ being the spectral projection $e^{|a^*|}_{(0,1)}$ (corresponding to the open interval $(0,1)$), the self-adjoint operator x must be of the form

$$x = \begin{bmatrix} x_0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & x_1 \end{bmatrix}$$

relative to the orthogonal decomposition $1 = e^{|a^*|}_{\{0\}} + e^{|a^*|}_{(0,1)} + e^{|a^*|}_{\{1\}}$. Conversely, when x is of this form, we obviously have $|a^*|x^2|a^*| = (|a^*|x|a^*|)^2$ and $a^*x^2a = (a^*xa)^2$ (or more generally $a^*f(x)a = f(a^*xa)$ for any $f(t)$ with $f(0) = 0$).

In this article, a function $f(t)$ on $[\alpha, \beta]$ is called a **strictly convex** function if its domain can be extended to $(\alpha - \delta_1, \beta + \delta_1)$ (for some $\delta_1 > 0$) in such a way that (the extended) $f(t)$ is of class C^2 on $(\alpha - \delta_1, \beta + \delta_1)$ and $f''(t) > 0$ on $[\alpha, \beta]$.

Lemma 5. *Let $f(t)$ be a strictly convex and strictly increasing function on $[\alpha, \beta]$, and we assume $0 \in (\alpha, \beta)$ and $f(0) = 0$. For $\delta > 0$ small $(\alpha + \delta < 0 < \beta - \delta)$ we can find two functions $g(t)$ and $\phi(t)$ satisfying*

(i) *$g(t)$ is a non-affine operator convex function on $[\alpha + \delta, \beta - \delta]$ with $g(0) = 0$ and $g([\alpha + \delta, \beta - \delta]) \subseteq [\alpha + \delta/2, \beta - \delta/2]$;*

    (ii) $\phi(t)$ *is a convex and strictly increasing on* $[\alpha + \delta/2, \beta - \delta/2]$ *with* $\phi(0) = 0$;

    (iii) $f = \phi \circ g$ *on* $[\alpha + \delta, \beta - \delta]$.

Proof. For $\varepsilon \in (0, 1]$ we set

(4) $$g(t) = (t + 2L)^{1+\varepsilon} - (2L)^{1+\varepsilon} \quad \text{with } L = \max\{\beta - \alpha, 1\}.$$

It is of course operator convex on $[\alpha, \beta]$, and the inverse is given by

$$g^{-1}(t) = \left(t + (2L)^{1+\varepsilon}\right)^{1/(1+\varepsilon)} - 2L.$$

We compute

$$(g^{-1})'(t) = \frac{\left(t + (2L)^{1+\varepsilon}\right)^{-\varepsilon/(1+\varepsilon)}}{1 + \varepsilon}, \quad (g^{-1})''(t) = -\frac{\varepsilon\left(t + (2L)^{1+\varepsilon}\right)^{-\varepsilon/(1+\varepsilon)-1}}{(1 + \varepsilon)^2},$$

and note

(5) $$\begin{cases} g(\beta - \delta) = (\beta - \delta + 2L)^{1+\varepsilon} - (2L)^{1+\varepsilon} \searrow \beta - \delta, \\ g(\alpha + \delta) = (\alpha + \delta + 2L)^{1+\varepsilon} - (2L)^{1+\varepsilon} \nearrow \alpha + \delta \end{cases}$$

as $\varepsilon \searrow 0$. Monotonicity of the convergence here is guaranteed by

$$\beta - \delta + 2L \geq 2L \geq 2 \quad \text{and} \quad 2L \geq \alpha + \delta + 2L \geq L \geq 1.$$

Thus, for ε sufficiently small we have

$$g\left([\alpha + \delta, \beta - \delta]\right) \subseteq [\alpha + \delta/2, \beta - \delta/2].$$

We will show that $\varepsilon > 0$ can be chosen in such a way that the composition

$$\phi(t) = f \circ g^{-1}(t) \left(= f\left(\left(t + (2L)^{1+\varepsilon}\right)^{1/(1+\varepsilon)} - 2L\right)\right)$$

is convex on $[\alpha + \delta/2, \beta - \delta/2]$. Note that $\phi(t)$ is strictly increasing since so are $f(t)$ and $g^{-1}(t)$.

   To choose ε making $\phi(t)$ convex, we need to compute $\phi''(t)$. We have

$$\begin{aligned} \phi''(t) &= f''(g^{-1}(t))\left((g^{-1})'(t)\right)^2 + f'(g^{-1}(t))(g^{-1})''(t) \\ &= \left((g^{-1})'(t)\right)^2 \left[f''(g^{-1}(t)) + \frac{f'(g^{-1}(t))(g^{-1})''(t)}{\left((g^{-1})'(t)\right)^2}\right] \end{aligned}$$

while the above computations for $(g^{-1})'(t)$, $(g^{-1})''(t)$ mean

$$\frac{(g^{-1})''(t)}{\left((g^{-1})'(t)\right)^2} = -\varepsilon\left(t + (2L)^{1+\varepsilon}\right)^{-1/(1+\varepsilon)} = -\frac{\varepsilon}{g^{-1}(t) + 2L}.$$

Hence, we get

$$\phi''(t) = \left((g^{-1})'(t)\right)^2 \left[f''(g^{-1}(t)) - \frac{\varepsilon f'(g^{-1}(t))}{g^{-1}(t) + 2L}\right].$$

We note

$$g^{-1}(\beta - \delta/2) = \left(\beta - \delta/2 + (2L)^{1+\varepsilon}\right)^{1/(1+\varepsilon)} - 2L \;\rightarrow\; \beta - \delta/2,$$

$$g^{-1}(\alpha + \delta/2) = \left(\alpha + \delta/2 + (2L)^{1+\varepsilon}\right)^{1/(1+\varepsilon)} - 2L \;\rightarrow\; \alpha + \delta/2$$

as $\varepsilon \to 0$. Thus, for $\varepsilon > 0$ sufficiently small (that might be even smaller than the previously chosen one $\cdots$ note monotonicity in the convergence (5)) we have

$$g^{-1}\left([\alpha + \delta/2, \beta - \delta/2]\right) \subseteq [\alpha, \beta].$$

We set

$$\begin{aligned}
I &= \min\{f''(s); \, s \in [\alpha, \beta]\}, \\
II &= \max\left\{\frac{f'(s)}{s + 2L}; \, s \in [\alpha, \beta]\right\}.
\end{aligned}$$

Strict increasingness of $f(t)$ says $I, II > 0$, and for $\varepsilon < I/II$ we indeed have $\phi''(t) > 0$ on $[\alpha + \delta/2, \beta - \delta/2]$. $\quad\square$

Remark 6. *Let $f(t)$ be a strictly convex and strictly increasing function on $[\alpha, \beta]$ containing 0 and we assume $f(0) = 0$. From the definition we observe that the function has the same property on a slightly extended interval $[\alpha - \delta_1, \beta + \delta_1]$ (for $\delta_1 > 0$ small enough). For $f(t)$ considered on this extended interval we make use of Lemma 5. Namely, by choosing δ (in the lemma) smaller than δ_1, we see that $f(t)$ can be expressed as a composition $\phi \circ g$ (on $[\alpha, \beta]$) in such a way that*

   (a) *$g(t)$ is a non-affine operator convex function on $[\alpha, \beta]$ with $g(0) = 0$,*
   (b) *$\phi(t)$ is defined on a suitable interval (larger than $g([\alpha, \beta])$), and it is convex and strictly increasing there with $\phi(0) = 0$,*

The function $g(t)$ here is actually operator convex on an interval somewhat larger than $[\alpha, \beta]$. Indeed, the function $g(t)$ used in the proof of the lemma is explicitly given by (4), and it is operator convex on $[-2L, \infty)$.

We are now ready to prove the next main result in the section.

Theorem 7. *Let f be a strictly convex function with $f(0) = 0$. For a self-adjoint operator x in $\mathcal{M} \cap L^1(\mathcal{M}; \tau)$ and a contraction $a \in \mathcal{M}$ the following conditions are equivalent:*

   (i)   *$\tau(f(a^*xa)) = \tau(a^*f(x)a)$;*
   (ii) *$a^*x^2a = (a^*xa)^2$;*
   (iii) *$a^*f(x)a = f(a^*xa)$.*

The assumption here (and the next corollary) means the following: $f(t)$ is a strictly convex function on an interval I (of finite length) containing 0, and $x \in \mathcal{M}_{sa}$ satisfies $\sigma(x) \subset I$ and integrability $\tau(|x|) < \infty$. Note that the function $f(t)$ (and other functions appearing in the proof below) is in modulus majorized by $k|t|$ (on I) for some k. Theorem 1 is to be used in the proof below, and the integrability requirement $f(a^*xa), a^*f(x)a \in L^1(\mathcal{M}; \tau)$ there is automatic.

Proof. The implication (ii) $\Rightarrow$ (iii) was pointed out in Remark 4 so that (i) $\Rightarrow$ (ii) has to be proved. If the function $f(t)$ in question satisfies the trace equality, then so does $f(t) + \lambda t$ for $\lambda \in \mathbf{R}$ (due to $x \in L^1(\mathcal{M}; \tau)$). By changing $f(t)$ to $f(t) + \lambda t$ with λ sufficiently large, we may and do assume strict increasingness of $f(t)$ from the beginning. Thanks to Lemma 5 and Remark 6 we can find functions $g(t)$, $\phi(t)$ satisfying (a),(b) (in Remark 6) and $f = \phi \circ g$.

One can prove $a^*g(x)a = g(a^*xa)$ based on Theorem 1. Indeed, since $\phi(t)$ is convex, the trace Jensen inequality (1) shows

$$\tau(a^*f(x)a) = \tau(a^*\phi(g((x))a) \geq \tau(\phi(a^*g(x)a)).$$

On the other hand, we have $a^*g(a)a \geq g(a^*xa)$ by operator convexity of $g(t)$ (see **2.1**), and consequently

$$\tau(\phi(a^*g(a)a)) \geq \tau(\phi(g(a^*xa))) = \tau(f(a^*xa))$$

by increasingness of $\phi(t)$ (see Theorem 1,(i)). Hence, the assumption $\tau(f(a^*xa)) = \tau(a^*f(x)a)$ says

$$\tau(\phi(a^*g(a)a)) = \tau(\phi(g(a^*xa))),$$

and $a^*g(x)a = g(a^*xa)$ follows from Theorem 1,(ii) or (iii). Recall that $g(t)$ can be considered as an operator convex function on a slightly larger open interval (see the last part in Remark 6). Thus, we can use Corollary 3 and conclude $a^*x^2a = (a^*xa)^2$. $\square$

In the above proof the assumption $x \in L^1(\mathcal{M}; \tau)$ was used only when we change $f(t)$ to $f(t) + \lambda t$ (which is strictly increasing for λ large). Therefore, if $f(t)$ is assumed to be increasing in addition, then this procedure is not required and this integrability condition can be dropped. But of course we have to assume integrability of $f(a^*xa), a^*f(x)a$ instead (besides the boundedness $x \in \mathcal{M}$).

Corollary 8. *Let f be a strictly convex function with $f(0) = 0$. For self-adjoint operators x, y in $\mathcal{M} \cap L^1(\mathcal{M}; \tau)$ and operators $a, b \in \mathcal{M}$ satisfying $a^*a + b^*b \leq 1$ the following conditions are equivalent:*

   (i) $\tau(f(a^*xa + b^*yb)) = \tau(a^*f(x)a + b^*f(y)b)$;
   (ii) $(a^*xa + b^*yb)^2 = a^*x^2a + b^*y^2b$.

Proof. Theorem 7 says that (i) is satisfied if and only if the 2×2-matrices

(6) $$A = \begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix}$$

(see the proof of [11, Corollary 9]) satisfy

(7) $$A^* X^2 A = (A^* X A)^2.$$

It is straight-forward to compute

$$A^* X^2 A = \begin{bmatrix} a^* x^2 a + b^* y^2 b & 0 \\ 0 & 0 \end{bmatrix}, \quad (A^* X A)^2 = \begin{bmatrix} (a^* x a + b^* y b)^2 & 0 \\ 0 & 0 \end{bmatrix}$$

so that the condition (7) is equivalent to (ii). $\square$

For simplicity let us assume that $a = \lambda 1$ and $b = \mu 1$ with λ, μ scalars in $[0, 1]$. When either $(x, y) = (0, 0)$ or $(\lambda, \mu) = (0, 0), (1, 0), (0, 1)$, we obviously have the trace equality and (7) is indeed satisfied. Thus, let us assume that none of them is satisfied. Then, the condition (7) (see Remark 4 for its meaning) is equivalent to

$$x = y \quad \text{and} \quad \lambda^2 + \mu^2 = 1.$$

This is an easy linear algebra exercise and details are left to the reader.

4. Equality condition (half line $[0, \infty)$)

We have been dealing with intervals of finite length containing 0 so far. This finiteness assumption cannot be avoided (at least in our proof) because our proof relies on Theorem 2. Consequently, our analysis has been restricted to bounded operators. In this section we will prove a counterpart to Theorem 2 for the half line $[0, \infty)$ at first. This will enable us to study equality conditions for various Jensen-type trace inequalities involving positive τ-measurable operators.

Theorem 9. *Let $\alpha \in (0, \infty]$. We assume that $g(t)$ is a non-affine operator convex function on $[0, \alpha)$ satisfying $g(0) = 0$. For a self-adjoint operator x (with $\sigma(x) \subset [0, \alpha)$) and a contraction a the operator Jensen inequality reduces to the equality $g(a^* x a) = a^* g(x) a$ if and only if $(a^* x a)^2 = a^* x^2 a$.*

Proof. Thanks to the dilation technique in the proof of Corollary 3, we may and do assume that $a = e$ is a projection (and we have to show $g(exe) = eg(x)e \Rightarrow ex = xe$). The assumption is equivalent to the property that $g(t)/t$ is operator monotone on $(0, \alpha)$ (see [9, Theorem 2.4]). On the other hand, operator monotone functions are characterizes as Pick functions, and such a function (i.e., $g(t)/t$ in our case) admits an integral expression of the form

$$p + qt + \int_{(-\infty, \infty) \setminus (0, \alpha)} \left[\frac{1}{s - t} - \frac{s}{s^2 + 1} \right] d\mu(s)$$

(see [6, Theorem 1 in p.20 and Lemma 2 in p.26]). Here, $p \in \mathbf{R}$, $q \geq 0$ and μ is a measure on $(-\infty, \infty) \setminus (0, \alpha)$ satisfying $\int (s^2 + 1)^{-1} d\mu(s) < \infty$. Note that the integral is over the half line $(-\infty, 0]$ when $\alpha = \infty$. From these facts we conclude

$$g(t) = pt + qt^2 + \int_{(-\infty,\infty) \setminus [0,\alpha)} g_s(t) \, d\mu(s)$$

with

$$g_s(t) = t \left[\frac{1}{s-t} - \frac{s}{s^2+1} \right].$$

Observe that the assumption $g(0) = 0$ forces $\mu(\{0\}) = 0$. Indeed, otherwise the constant term $-\mu(\{0\})$ would show up from $g_0(t) = -1$. We also note $q > 0$ or $\mu \neq 0$ since $g(t)$ is non-affine.

Let us assume $eg(x)e = g(exe)$. Since

$$\frac{t}{s-t} = \begin{cases} -1 + \dfrac{s}{s-t} & \text{(when } s \geq \alpha), \\[2mm] -1 + \dfrac{s'}{s'+t} & \text{(when } s' = -s > 0), \end{cases}$$

$g_s(t)$ is operator convex on $[0, \alpha)$ for either $s < 0$ or $s \geq \alpha$. Thus, we generally have $eg_s(x)e \geq g_s(exe)$ for these s's (as well as $ex^2e \geq (exe)^2$) and we must have

$$\begin{cases} q(ex^2e) = q(exe)^2, \\ eg_s(x)e = g_s(exe) \text{ for } s \in (-\infty, \infty) \setminus [0, \alpha) \ \mu\text{-a.e.} \end{cases}$$

Let us assume $q = 0$ (since we are done otherwise). Then, $\mu \neq 0$ and we have $eg_s(x)e = g_s(exe)$ for some $s \in (-\infty, \infty) \setminus [0, \alpha)$. Firstly we have

$$eg_s(x)e = ex(s-x)^{-1}e - s(s^2+1)^{-1}e.$$

On the other hand, since $g_s(0) = 0$, the operator $g_s(exe)$ $(= eg_s(exe)e)$ is given by

$$\begin{aligned} g_s(exe) &= e \left(exe(s - exe)^{-1} - s(s^2+1)^{-1} \right) e \\ &= exe(se - exe)^{-1} - s(s^2+1)^{-1}e \\ &= exe \left(e(s-x)e \right)^{-1} - s(s^2+1)^{-1}e. \end{aligned}$$

Therefore, the equality $eg_s(x)e = g_s(exe)$ yields

$$ex(s-x)^{-1}e = exe \left(e(s-x)e \right)^{-1}.$$

When $s \geq \alpha$, we set $y = s - x$ (which is positive and invertible) and compute

$$\begin{aligned} ex(s-x)^{-1}e &= e(s-y)y^{-1}e = sey^{-1}e - e, \\ exe \left(e(s-x)e \right)^{-1} &= (se - eye)(eye)^{-1} = s(eye)^{-1} - e. \end{aligned}$$

On the other hand, when $s < 0$, with $y = -s + x$ (positive and invertible again) instead we similarly get

$$ex(s-x)^{-1}e = -sey^{-1}e - e, \quad exe\big(e(s-x)e\big)^{-1} = -s(eye)^{-1} - e.$$

Thus, we have $ey^{-1}e = (eye)^{-1}$ with y positive and invertible in both cases (due to $s \neq 0$), and consequently $[y, e] = 0$ (i.e., $[x, e] = 0$) by the arguments in the last part of the proof of Theorem 2. $\square$

For simplicity let us concentrate on the special case $\alpha = \infty$. The preceding theorem obviously remains valid for a positive τ-measurable operator x (with the identical proof).

Theorem 10. *Let $f(t)$ be a continuous convex function on $[0, \infty)$ with $f(0) = 0$. We assume that a positive τ-measurable operator x and a contraction $a \in \mathcal{M}$ satisfy $a^*f(x)a, f(a^*xa) \in L^1(\mathcal{M}; \tau)$. Then, from the trace equality $\tau(a^*f(x)a) = \tau(f(a^*xa))$ we can conclude $(a^*xa)^2 = a^*x^2a$ if one of the following conditions is fulfilled:*

(a) $f = \phi \circ g$ *with a non-affine operator convex function $g(t)$ on $[0, \infty)$ with $g(0) = 0$ and a convex strictly increasing function $\phi(t)$ (on the range of $g(t)$) with $\phi(0) = 0$;*

(b) $f(t)$ *is strictly increasing, twice differentiable on $(0, \infty)$ and*

$$(8) \qquad \qquad \inf\left\{\frac{sf''(s)}{f'(s)}; s > 0 \text{ and } f'(s) > 0\right\} > 0.$$

Proof. We assume (a) and begin with the claim: $\phi(a^*g(x)a) \in L^1(\mathcal{M}; \tau)$. Indeed, we observe

$$\tau\left(\phi(a^*g(x)a)_+\right) \leq \tau\left((a^*\phi \circ g(x)a)_+\right) = \tau\left((a^*f(x)a)_+\right) < \infty$$

by the trace inequality in [11, Theorem 9] together with the integrability assumption $a^*f(x)a \in L^1(\mathcal{M}; \tau)$. On the other hand, from the operator inequality $-g(a^*xa) \geq -a^*g(x)a$ we have the spectral dominance

$$g(a^*xa)_- = \left(-g(a^*xa)\right)_+ \gtrsim \left(-a^*g(x)a\right)_+ = (a^*g(x)a)_-.$$

Since $\psi(t) = -\phi(-t)$ is also increasing and $\psi(0) = 0$, we observe

$$\tau\left(\phi(a^*g(x)a)_-\right) = \tau\left(\psi\big((a^*g(x)a)_-\big)\right)$$
$$\leq \tau\left(\psi\big(g(a^*xa)_-\big)\right) = \tau\left(\phi(g(a^*xa))_-\right) = \tau\left(f(a^*xa)_-\right)$$

(see **2.3**). From this spectral dominance and the integrability $f(a^*xa) \in L^1(\mathcal{M}; \tau)$ we get $\phi(a^*g(x)a)_- \in L^1(\mathcal{M}; \tau)$ and the claim is established.

Integrability of all the involved operators are checked, and as in the proof of Theorem 7 we have

$$
\begin{aligned}
\tau\left(a^* f(x)a\right) &= \tau\left(a^* \phi(g(x))a\right) \\
&\geq \tau\left(\phi(a^* g(x)a)\right) \quad \text{(by convexity of } \phi(t)) \\
&\geq \tau\left(\phi(g(a^* xa))\right) \quad \text{(by increasingness of } \phi(t) \text{ and } a^* g(x)a \geq g(a^* xa)) \\
&= \tau\left(f(a^* xa)\right).
\end{aligned}
$$

Convexity of $g(t)$ (on $[0, \infty)$) implies

$$
\text{either} \quad g(t) \leq 0 \quad \text{or} \quad g(t) \geq \min\{g(s); s \geq 0\} > -\infty,
$$

and hence $g(x)$ is always semi-bounded. Thus, the above use of the trace Jensen inequality (1) is certainly legitimate. Hence, when $\tau(a^* f(x)a) = \tau(f(a^* xa))$, we must have $\tau(\phi(a^* g(x)a)) = \tau(\phi(g(a^* xa)))$. Since $\phi(t)$ is strictly increasing, Theorem 1,(iii) says $a^* g(x)a = g(a^* xa)$, and consequently we have $a^* x^2 a = (a^* xa)^2$ thanks to Theorem 9.

We next show that the case (b) can be reduced from (a), by regarding f as the composition $\phi \circ g$ with

$$
g(t) = t^{1+\varepsilon} \quad (\varepsilon \in (0, 1]) \quad \text{and} \quad \phi(t) = f\left(g^{-1}(t)\right) = f\left(t^{1/(1+\varepsilon)}\right).
$$

It is plain to compute

$$
\phi''(t) = \frac{t^{-2\varepsilon/(1+\varepsilon)}}{(1 + \varepsilon)^2} \left[f''\left(t^{1/(1+\varepsilon)}\right) - \frac{\varepsilon f'\left(t^{1/(1+\varepsilon)}\right)}{t^{1/(1+\varepsilon)}} \right].
$$

When $f'\left(t^{1/(1+\varepsilon)}\right) = 0$, we have $\phi''(t) \geq 0$ (due to $f''\left(t^{1/(1+\varepsilon)}\right) \geq 0$. Otherwise (to get $\phi''(t) \geq 0$) we must have

$$
\frac{s f''(s)}{f'(s)} \geq \varepsilon \left(\text{with } s = t^{1/(1+\varepsilon)} \right).
$$

Thus, the condition in (b) (i.e., (8)) guarantees $\phi''(t) \geq 0$. □

Remark 11. *Let $f(t)$ be a continuous convex function on $[0, \infty)$ with $f(0) = 0$ as in the theorem. Then, we have*

(9) $$\tau\left(f(x) + f(y)\right) \leq \tau(f(x + y))$$

for positive τ-measurable operators x, y (as long as the both sides are well-defined, i.e., "$\infty - \infty$" does not occur). This fact can be easily derived from the trace Jensen

inequality (1) *in the following way* ([3, Proposition 15] *or* [7, Proposition 4.6]): *Let* $u, v \in \mathcal{M}$ *be contractions satisfying*

$$x^{1/2} = u(x+y)^{1/2}, \ y^{1/2} = v(x+y)^{1/2},$$
$$u^*u + v^*v = \textit{the support projection of } x+y.$$

We then estimate

$$
\begin{aligned}
\tau(f(x)) + \tau(f(y)) &= \tau\left(f\left(u(x+y)u^*\right)\right) + \tau\left(f\left(v(x+y)v^*\right)\right) \\
&\leq \tau\left(uf(x+y)u^*\right) + \tau\left(vf(x+y)v^*\right) \\
&= \tau\left((u^*u + v^*v)f(x+y)\right).
\end{aligned}
$$

Since $u^*u + v^*v$ *is the support projection of* $x+y$, *the last quantity is nothing but* $\tau\left(f(x+y)\right)$, *showing* (9).

Let us assume that $f(t)$ satisfies one of the conditions in Theorem 10. Then for $x, y \geq 0$ with $f(x), f(y), f(x+y) \in L^1(\mathcal{M}; \tau)$ the inequality (9) reduces to the equality if and only if x, y have orthogonal supports, i.e., $xy = 0$. Indeed, the equality implies $\tau\left(f\left(u(x+y)u^*\right)\right) = \tau\left(uf(x+y)u^*\right)$ so that we must have

$$\left(u(x+y)u^*\right)^2 = u(x+y)^2 u^*$$

by Theorem 10. *The left side here is* x^2 *while the right side can be computed as*

$$u(x+y)^2 u^* = x^{1/2}(x+y)x^{1/2} = x^2 + x^{1/2}yx^{1/2}.$$

Therefore, we conclude $x^{1/2}yx^{1/2} = 0$, *i.e.,* $xy = 0$.

Note that the condition (8) means:

$f'(t)/t^{\varepsilon}$ is increasing for $\varepsilon > 0$ sufficiently small.

Let us consider the convex functions $f(t) = t^r$ $(r > 1)$ on the interval $[0, \infty)$. It is operator convex for $r \in (1, 2]$, and both of the conditions (a),(b) in the theorem can be easily checked for general $r > 1$. Thus, we have

$$\tau((a^*xa)^r) = \tau(a^*x^r a) \iff a^*x^2 a = (a^*xa)^2.$$

However, we have the spectral dominance

(10) $$(a^*xa)^r \lesssim a^*x^r a \ \ (r > 1)$$

since t^r is convex and increasing on $[0, \infty)$ (see [7, Lemma 4.5]). Slight modifications of the preceding arguments together with (10) actually yield the next proposition (which will be useful for dealing with $\| \cdot \|_p$-norm inequalities). Note that for a function $f(t)$ in the proposition we generally have

$$\tau\left(f\left((a^*xa)^r\right)\right) \leq \tau\left(f\left(a^*x^r a\right)\right) \ (\leq \infty)$$

(see [3, Lemma 4] for instance).

Proposition 12. *Let $f(t)$ be a continuous, convex and strictly increasing function on $[0, \infty)$ with $f(0) = 0$. We assume that x is a positive τ-measurable operator and $a \in \mathcal{M}$ is a contraction. Then, the trace equality*

$$\tau\Big(f\big((a^*xa)^r\big)\Big) = \tau\Big(f\big(a^*x^ra\big)\Big) \; (< \infty)$$

holds true for some $r \in (1, \infty)$ if and only if $(a^*xa)^2 = a^*x^2a$.

Proof. When $(a^*xa)^2 = a^*x^2a$, we have $(a^*xa)^r = a^*x^ra$ as was pointed out in Remark 4. To show the converse, we make use of the operator inequality

(11) $$a^*x^ra \geq (a^*xa)^r \quad \text{for } r \in (1, 2],$$

which is a consequence of operator convexity.

At first we assume $r > 2$ and set $\phi(t) = t^{r/2}$. Notice the following estimates:

$$
\begin{aligned}
\tau\left(f(a^*x^ra)\right) &= \tau\left(f\left(a^*\phi\big(x^2\big)a\right)\right) \\
&\geq \tau\left(f \circ \phi\left(a^*x^2a\right)\right) \\
&\geq \tau\left(f \circ \phi\left((a^*xa)^2\right)\right) = \tau((a^*xa)^r).
\end{aligned}
$$

Here, the first inequality follows from increasingness of $f(t)$ and the spectral dominance $a^*\phi(x^2)a \gtrsim \phi(a^*x^2a)$ (i.e., (10) with $r/2$ and x^2 instead) while the second follows from increasingness of $f \circ \phi$ and $a^*x^2a \geq (a^*xa)^2$ (i.e., (11) with $r = 2$). The trace equality thus implies

$$\tau\left(f \circ \phi\left(a^*x^2a\right)\right) = \tau\left(f \circ \phi\left((a^*xa)^2\right)\right).$$

Since $f \circ \phi$ is convex and strictly increasing, we have the operator equality $a^*x^2a = (a^*xa)^2$ thanks to Theorem 1,(iii).

Secondly we assume $r \in (1, 2]$. If the trace equality holds, then (11) and strict increasingness of $f(t)$ guarantee $(a^*xa)^r = a^*x^ra$ thanks to Theorem 1,(iii) again and the desired result follows from Theorem 9. $\square$

A version akin to Corollary 8 is also possible, and details are left to the reader (see [15, §6] for related topics). The special choice $f(t) = t^p$ in the proposition gives rise to the following:

Corollary 13. *We assume $p \in [1, \infty)$ and $r \in (1, \infty)$. Let x be a positive τ-measurable operator and $a \in \mathcal{M}$ be a contraction. Then, the $||\cdot||_p$-norm inequality $||(a^*xa)^r||_p \leq ||a^*x^ra||_p$ reduces to the equality*

$$||a^*x^ra||_p = ||(a^*xa)^r||_p \; (< \infty)$$

if and only if $(a^*xa)^2 = a^*x^2a$.

Note that the result is not valid for $p = \infty$. Indeed, let us consider operators e, x of the forms

$$e = E \oplus 1 \quad \text{and} \quad x = X_1 \oplus X_2$$

with a projection E and $X_i \geq 0$. When $||X_1||$ is small and $||X_2||$ is large, the operator norms $||a^*x^r a||$, $||(a^*xa)^r||$ do not see the respective first direct summands.

Corollary 13 (and the next Proposition 14) should be compared with related analysis for compact operators made in [12, §3]. Note that operators of the forms

$$\sum_{i=1}^{n} a_i^* x_i^r a_i, \quad \left(\sum_{i=1}^{n} a_i^* x_i a_i\right)^r \quad (\text{with} \ \ x_i \geq 0 \ \ \text{and} \ \ \sum_{i=1}^{n} a_i^* a_i \leq 1)$$

are treated there, but one can use $n \times n$ matrices analogous to the 2×2 matrices given by (6) (to reduce everything to the case $n = 1$).

Corollary 13 obviously means mutual equivalence of the three conditions under the assumption $a^* x^r a \in L^p(\mathcal{M}; \tau)$ with $r > 1$:

(a) $(a^*xa)^2 = a^*x^2a$, (b) $a^*x^r a = (a^*xa)^r$, (c) $||a^*x^r a||_p = ||(a^*xa)^r||_p$.

Note that Theorem 9 says that for $r \in (1, 2]$ the implication (b) $\Rightarrow$ (a) is valid without the assumption $a^* x^r a \in L^p(\mathcal{M}; \tau)$ (due to operator convexity of t^r with exponents r in this range). The following easy fact seems to be worth pointing out:

Proposition 14. *We assume that x is a positive τ-measurable operator and $a \in \mathcal{M}$ is a contraction. Then, we have $a^* x^r a = (a^*xa)^r$ for some $r > 0$, $r \neq 1$, if and only if $(a^*xa)^2 = a^*x^2a$.*

Proof. The implication $a^*x^2a = (a^*xa)^2 \Rightarrow a^*x^r a = (a^*xa)^r$ is obvious (see Remark 4). To show the converse, we note

(12) $a^* x^2 a = (a^*xa)^2 \iff a^*x^r a = (a^*xa)^r \ (r \in (0, 1))$.

In fact, the operator monotone function t^r is operator concave, and the assertion follows from Theorem 9 applied to the operator convex function $-t^r$. Thus, when $r \in (0, 1)$ (and $ex^r e = (exe)^r$), we are done. On the other hand, when $r \in (1, \infty)$, the assumption $ex^r e = (exe)^r$ implies

$$exe = (ex^r e)^{1/r} \geq e \, (x^r)^{1/r} \, e = exe$$

due to operator concavity of $t^{1/r}$. This shows $(ex^r e)^{1/r} = exe$ and consequently $[x^r, e] = 0$ thanks to (12) with $1/r \in (0, 1)$ instead. We thus have $[x, e] = 0$. $\square$

The equivalence (12) can be also directly seen by arguments similar to those in the proof of Theorem 2 (or Theorem 9) together with the well-known integral expression

$$t^r = \frac{\sin(\pi r)}{\pi} \int_0^\infty \frac{t}{s + t} s^{r-1} ds, \ t \geq 0$$

for $r \in (0, 1)$ (see [2, p.145] for instance), and details are left to the reader

APPENDIX A. EQUALITY CONDITION (MATRIX CASE)

Let $f : \mathbf{R} \to \mathbf{R}$ be a strictly convex function "in the usual sense": For different t_1, t_2 we have $f(\lambda t_1 + (1 - \lambda)t_2) < \lambda f(t_1) + (1 - \lambda)f(t_2)$ as long as $\lambda \in (0, 1)$. In this appendix we will show that for matrices our characterization (such as Theorem 7 and Theorem 10) is available under this usual notion of strict convexity.

We begin with the claim: For $\{t_i\}_{i=1,2,\cdots,n}$ mutually different and $\{\lambda_i\}_{i=1,2,\cdots,n}$ satisfying $\lambda_i \geq 0$, $\sum_{j=1}^{n} \lambda_i = 1$ we have

$$(13) \qquad \sum_{j=1}^{n} \lambda_j f(t_j) = f\left(\sum_{j=1}^{n} \lambda_j t_j\right) \implies \lambda_j = 1 \text{ for some } j.$$

Indeed, it is just the definition for $n = 2$, and let us assume (13) for $n - 1$ (to prove the assertion by induction on n). By permuting λ_i's and x_i's (if necessary), we may and do assume $t_i < t_n$ $(i = 1, 2, \cdots, n - 1)$. We also assume $\lambda_n < 1$ (since we are done otherwise). By setting $\tilde{t} = \sum_{i=1}^{n-1} \lambda_i (1 - \lambda_n)^{-1} t_i$ $(< t_n)$, we estimate

$$
\begin{aligned}
f\left(\sum_{j=1}^{n} \lambda_j t_j\right) &= f\left((1 - \lambda_n)\tilde{t} + \lambda_n t_n\right) \\
&\leq (1 - \lambda_n)f(\tilde{t}) + \lambda_n f(t_n) \\
&\leq (1 - \lambda_n)\left(\sum_{i=1}^{n-1} \lambda_i (1 - \lambda_n)^{-1} f(t_i)\right) + \lambda_n f(t_n) = \sum_{j=1}^{n} \lambda_j f(t_j)
\end{aligned}
$$

by the repeated use of convexity. Thus, if the equality in (13) is satisfied, then we must have

$$f\left((1 - \lambda_n)\tilde{t} + \lambda_n t_n\right) = (1 - \lambda_n)f(\tilde{t}) + \lambda_n f(t_n).$$

Strict convexity thus shows $\lambda_n = 0$, and the induction hypothesis takes care of the rest.

Theorem 15. *Let $f(t)$ be a strictly convex function in the usual sense (as explained above) satisfying $f(0) = 0$, and we assume that $n \times n$ matrices a, x satisfy $x = x^*$ and $||a|| \leq 1$. Then, we have $Tr(a^* f(x)a) = Tr(f(a^* x a))$ if and only if $a^* x^2 a = (a^* x a)^2$.*

Proof. By the dilation technique in the proof of Corollary 3, it suffices to show

$$\mathrm{Tr}\left(ef(x)e\right) = \mathrm{Tr}\left(f(exe)\right) \implies ex = ex$$

for a projection e. We set $k = \mathrm{rank}\, e$ and consider x, e as the following partitioned matrices:

$$x = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}, \qquad e = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

The eigenvalues (with multiplicities counted) of $exe = \begin{bmatrix} X_{11} & 0 \\ 0 & 0 \end{bmatrix}$ consists of the eigenvalues $\{\lambda_i\}_{i=1,2,\cdots,k}$ of X_{11} together with 0 (with multiplicity $n - k$). We choose and fix an orthonormal system $\{\xi_i\}_{i=1,2,\cdots,k}$ ($\subseteq \mathbf{C}^n$) of eivenvectors corresponding to λ_i's so that this system forms a basis for the subspace $e\,\mathbf{C}^n$.

The eigenvalues of $f(exe)$ are $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_k)$ and 0 with multiplicity $n - k$ (due to $f(0) = 0$), and we compute

$$
\begin{aligned}
\mathrm{Tr}\,(f(exe)) &= \sum_{i=1}^{k} f(\lambda_i) = \sum_{i=1}^{k} f\big((exe\xi_i, \xi_i)\big) = \sum_{i=1}^{k} f\big((x\xi_i, \xi_i)\big) \\
&\le \sum_{i=1}^{k} (f(x)\xi_i, \xi_i) = \mathrm{Tr}\,(ef(x)e).
\end{aligned}
$$

Let $\{e_\lambda\}$ be the spectral projections for x, and ν_i be the corresponding (probability) Stieljes measure $d||e_\lambda\xi_i||^2$. Then the above (Jensen-type) inequality can be seen as follows:

$$
f\big((x\xi_i, \xi_i)\big) = f\left(\int \lambda\, d\nu_i(\lambda)\right) \le \int f(\lambda)\, d\nu_i(\lambda) = (f(x)\xi_i, \xi_i).
$$

We now take a closer look at the meaning of this inequality. To do so, let $\{\mu_j\}_{j=1,2,\cdots,m}$ be the distinct eigenvalues of x. We assume that each eigenvalue μ_j is of multiplicity n_j ($\sum_{j=1}^{m} n_j = n$) and $\mu_1 < \mu_2 < \cdots < \mu_m$. Let p_j be the projection (of rank-n_j) onto the subspace $W_{\mu_j} = \{\xi \in \mathbf{C}^n; x\xi = \mu_j\xi\}$. Of course (in our matrix setting) the measure ν_i has just point masses at $\lambda = \mu_j$:

$$
\nu_i(\{\mu_j\}) = ||p_j\xi_i||^2 \quad (j = 1, 2, \cdots, m).
$$

Thus, the above Jensen inequality (for integrals) can be written as

$$
f\left(\sum_{j=1}^{m} \mu_j ||p_j\xi_i||^2\right) \le \sum_{j=1}^{m} f(\mu_j)||p_j\xi_i||^2
$$

in the present case.

The equality $\mathrm{Tr}\,(ef(x)e) = \mathrm{Tr}\,(f(exe))$ means

$$
f\big((x\xi_i, \xi_i)\big) = (f(x)\,\xi_i, \xi_i) \quad \text{for } i = 1, 2, \cdots, k,
$$

and hence strict convexity yields $||p_{j_i}\xi_i||^2 = 1$ for some $j_i \in \{1, 2, \cdots, m\}$ (see (13)), showing $\xi_i \in W_{\mu_{j_i}}$. Since the eigenvectors ξ_i (of x) form a basis for the subspace $e\,\mathbf{C}^n$, x leaves $e\,\mathbf{C}^n$ invariant (i.e,, the block matrix X_{21} vanishes). This means $exe = xe$ and hence we conclude $ex = xe$ (due to $x = x^*$). $\square$

Corollary 16. *Let $f(t)$ be a strictly convex function with $f(0) = 0$. If a non-zero contraction matrix a satisfies $Tr(a^* f(x)a) = Tr(f(a^*xa))$ for all Hermitian matrices x, then a must be a unitary matrix.*

Proof. The theorem shows $a^* x^2 a = (a^* x a)^2$ for every x $(= x^*)$. Then, as was explained in Remark 4, $|a^*|$ $(= \sqrt{aa^*})$ must satisfy $|a^*| x^2 |a^*| = (|a^*| x |a^*|)^2$ and $[|a^*|, x] = 0$ for every x. The commutativity forces $|a^*| = \lambda 1$ with $\lambda \in (0, 1]$, and it suffices to show $\lambda = 1$. However, if we had $\lambda \neq 1$, then $|a^*| x^2 |a^*| = (|a^*| x |a^*|)^2$ would be possible only for $x = 0$, a contradiction. $\square$

The arguments in the proof of Theorem 15 remain valid (with obvious modifications) for a self-adjoint compact operator x under the requirement $a^* f(x) a$, $f(a^* x a) \in \mathcal{C}_1$, the trace class operators. Details are left to the reader as an exercise. Note that the condition $\sum_{j=1}^{\infty} \lambda_j f(t_j) = f\left(\sum_{j=1}^{\infty} \lambda_j t_j\right)$ have to be considered in (13) since x may have countably many eigenvalues. However, the following general fact (for a probability measure ν) resolves this complication:

$$f\left(\int t \, d\nu(t)\right) = \int f(t) \, d\nu(t) \implies \text{the support of } \nu \text{ is singleton}$$

To see this, let us recall the standard proof for the Jensen inequality based on "supporting lines" (see [19, Chap.5 §5] for instance): We set $t_0 = \int t d\nu(t)$ and choose $a, b \in \mathbf{R}$ satisfying

$$f(t) \geq at + b, \quad f(t_0) = at_0 + b.$$

Integration of the both sides of the first inequality gives rise to

$$\int f(t) \, d\nu(t) \geq at_0 + b \int d\nu(t) = at_0 + b = f(t_0).$$

When f is strictly convex, we have $f(t) > at + b$ for $t \neq t_0$. Thus, if the equality occurs in the above, then the support of ν must be $\{t_0\}$.

References

1. J. Bendat and S. Sherman, *Monotone and convex operator functions*, Trans. Amer. Math. Soc., **79** (1955), 58-71.
2. R. Bhatia, *Matrix Analysis*, Springer, 1997.
3. L. G. Brown and H. Kosaki, *Jensen's inequality in semi-finite von Neumann algebras*, J. Operator Theory, **23**(1990), 3-19.
4. M. D. Choi, *A Schwarz inequality for positive linear maps on C^*-algebras*, Ill. J. Math., **18** (1974), 565-574.
5. C. Davis, *A Schwarz inequality for convex operator functions*, Proc. Amer. Math. Soc., **8** (1957), 42-44.
6. W. Donoghue Jr., *Monotone Matrix Functions and Analytic Continuation*, Springer, 1974.
7. T. Fack and H. Kosaki, *Generalized s-numbers of τ-measurable operators*, Pacific J. Math., **123**(1986), 269-300.
8. F. Hansen, *An operator inequality*, Math. Anal., **246** (1980), 249-250.
9. F. Hansen and G. K. Pedersen, *Jensen's inequality for operators and Löwner's theorem*, Math. Ann., **258** (1982), 229-241.
10. F. Hansen and G. K. Pedersen, *Jensen's operator inequality*, Bull. London Math. Soc., **35** (2003), 553-564.

11. T. Harada and H. Kosaki *Trace Jensen inequality and related weak majorization in semi-finite von Neumann algebras*, preprint.
12. O. Hirzallah and F. Kittaneh, *Norm inequalities for weighted power means of operators*, Linear Algebra Appl., **341** (2002), 181-193.
13. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, 1985.
14. R. Kadison, *A generalized Schwarz inequality and algebraic invariants for operator algebras*, Ann. of Math., **56** (1952), 494-503.
15. H. Kosaki, *Applications of uniform convexity of noncommutative L^p-spaces*, Trans. Amer. Math. Soc., **283** (1984), 265-282.
16. E. Nelson, *Notes on non-commutative integration*, J. Funct. Anal., **15** (1974), 103-116.
17. D. Petz, *On the equality in Jensen's inequality for operator convex functions*, Integral Equations Operator Theory, **9** (1986), 744-747.
18. D. Petz, *Jensen's inequality for positive contractions on operator algebras*, Proc. Amer. Math. Soc., **99** (1987), 273-277.
19. H. L. Royden, *Real Analysis (second edition)*, Macmillan, 1968.
20. I. Segal, *A non-commutative extension of abstract integration*, Ann. of Math., **57** (1953), 491-457.

9-16-201 Hakozaki 1-chome, Higashi-ku, Fukuoka 812-0053, Japan

Faculty of Mathematics, Kyushu University, Higashi-ku, Fukuoka 812-8581, Japan