J. OPERATOR THEORY
63:1(2010), 129–150

TRACE JENSEN INEQUALITY AND RELATED WEAK MAJORIZATION IN SEMI-FINITE VON NEUMANN ALGEBRAS

TETSUO HARADA and HIDEKI KOSAKI

Communicated by William Arveson

ABSTRACT. Let $\mathcal{M}$ be a semi-finite von Neumann algebra equipped with a faithful semi-finite normal trace τ, and we assume that $f(t)$ is a convex function with $f(0) = 0$. The trace Jensen inequality $\tau(f(a^*xa)) \leqslant \tau(a^*f(x)a)$ is proved for a contraction $a \in \mathcal{M}$ and a self adjoint operator $x \in \mathcal{M}$ (or more generally for a semi-bounded τ-measurable operator) together with an abundance of related weak majorization-type inequalities. Notions of generalized singular numbers and spectral scales are used to express our results. Monotonicity properties for the map: $x \in \mathcal{M}_{\mathrm{sa}} \to \tau(f(x))$ are also investigated for an increasing function $f(t)$ with $f(0) = 0$.

KEYWORDS: *Generalized singular number, Jensen inequality, semi-finite von Neumann algebra, τ-measurable operator, trace, weak majorization.*

MSC (2000): Primary 46L51, 46L10; Secondary 47A63, 15A42.

1. INTRODUCTION

A continuous function $f(t)$ on an interval I is said to be operator convex when $f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda)f(y)$ is valid for each $\lambda \in [0,1]$ and self-adjoint operators x, y with spectra included in I. It was shown in [8], [9] that such an operator convex function $f(t)$ satisfies

$$\pi(f(x)) \geqslant f(\pi(x))$$

for a positive unital map π. A closely related inequality is the so-called operator Jensen inequality ([13], [14], [15]) stating

$$a^*f(x)a \geqslant f(a^*xa).$$

Here, a is a contraction (i.e., $\|a\| \leqslant 1$) and both of $0 \in I$, $f(0) \leqslant 0$ have to be assumed. A readable account on these and related subjects can be found in [5], [14], [15].

Operator convexity is much stronger than the usual convexity, and for trace inequalities we expect (or at least hope) that usual convexity (or concavity) is sufficient to get estimates in similar nature. Indeed, the Jensen-type trace inequality

$$\tau(a^* g(x) a) \leqslant \tau(g(a^* x a))$$

was obtained in [7] for a semi-finite von Neumann algebra $\mathcal{M}$ with a trace τ, where $g(t)$ is a continuous concave function with $g(0) = 0$. Here, $a \in \mathcal{M}$ is a contraction again and in [7] (as well as in [24]) $x \in \mathcal{M}$ was assumed to be positive. Whenever such trace inequalities are considered, we will assume $g(0) = 0$ (otherwise $\tau(|g(a^* x a)|) = \infty$ for $x = 0$ unless $\tau(1) < \infty$). The notion of spectral dominance (see Section 2.1) played an important role in this work. The closely related inequality

$$\tau(\alpha(f(x))) \geqslant \tau(f(\alpha(x)))$$

for a positive contractive map $\alpha : \mathcal{M} \to \mathcal{M}$ and a convex function $f(t)$ with $f(0) = 0$ was also proved in [24]. It is easy to see that these inequalities actually remain valid for self-adjoint operators x (see Theorem 3.4 in Section 3).

In recent years some convexity inequalities of weak majorization-type were obtained for eigenvalues of Hermitian matrices (see [4], [29] for instance). Weak majorization for matrices deals with partial sums of eigenvalues, and usefulness of this technique in the matrix and/or operator setting is concisely explained in the survey article [3]. The purpose of the present article is to prove the Jensen-type trace inequality for self-adjoint operators at first and then to obtain many related weak majorization-type inequalities in the semi-finite von Neumann algebra setting.

In Section 2 we will collect some basic notions (such as generalized singular numbers and spectral scales) needed in the article. In Section 3 at first we will prove

$$\tau(a^* f(x) a) \geqslant \tau(f(a^* x a))$$

for a convex function $f(t)$ with $f(0) = 0$ and $x \in \mathcal{M}_{\mathrm{sa}}$ (more precisely for a semi-bounded τ-measurable operator), which will be referred to as the trace Jensen inequality. Then, by closely examining its proof, we will study Jensen-type weak majorization in the (semi-finite) von Neumann algebra setting. Here, "eigenvalues" make no sense and we will use the notion of generalized singular numbers ([10], [11]) to formulate our results. We will also obtain a certain comparison result between $|x + y|$ and $|x| + |y|$ (i.e., an operator valued triangle inequality). In Section 4 we will study Jensen-type weak majorization with the notion of spectral scales [23]. Thus, τ has to be a finite trace, but results are more satisfactory in the sense that many results known in the matrix setting can be proved. In Section 5 we will study monotonicity properties for the map: $x \in \mathcal{M}_{\mathrm{sa}} \to \tau(f(x))$ for a continuous increasing function $f(t)$ satisfying $f(0) = 0$. As expected, the map is indeed monotone relative to the ordinary order $\leqslant$ on $\mathcal{M}_{\mathrm{sa}}$ although required arguments are somewhat tricky. Results here are used in our related analysis [16] and could be useful for other purposes as well.

2. PRELIMINARIES

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 $\tau(1-e) \leqslant \varepsilon$. It is known that the set $\overline{\mathcal{M}}$ $(\supset \mathcal{M})$ of all τ-measurable operators forms a $*$-algebra (and one can treat them without worrying "domain problems"). For $\varepsilon, \delta > 0$ we set

$$V(\varepsilon,\delta)=\{x\in\overline{\mathcal{M}} : \text{there exists a projection } e\in\mathcal{M} \text{such that } \|xe\|\leqslant\varepsilon \text{ and } \tau(1-e)\leqslant\delta\}.$$

The linear topology on $\overline{\mathcal{M}}$ whose fundamental system of neighborhoods around 0 is given by $V(\varepsilon,\delta)$'s is known as the measure topology. It is known that $\overline{\mathcal{M}}$ is a complete $*$-algebra relative to this topology. Basic facts on τ-measurable operators and the measure topology can be found in [21] (see also [25]).

2.1. SPECTRAL DOMINANCE. For positive operators $x, y \in \mathcal{M}$ (or rather positive τ-measurable operators) the spectral dominance $x \gtrsim y$ (or $y \lesssim x$) means

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

in the Murray–von Neumann sense. Here (and in the rest), the spectral projection of x (corresponding to a subset $I \subseteq \mathbb{R}$) will be denoted by e^x_I (or simply $e_I(x)$).

The following facts will be repeatedly used:

(a) For positive τ-measurable operators x, y with $x \leqslant y$ we have the spectral dominance $x \lesssim y$ (see Lemma 3(i) of [7]).

(b) For positive τ-measurable operators x, y with $x \lesssim y$ we have $g(x) \lesssim g(y)$ for any continuous increasing function $g(t)$ on $[0,\infty)$ satisfying $g(0) = 0$.

(c) If self-adjoint τ-measurable operators x, y satisfy $x \leqslant y$, then we have $x_+ \lesssim y_+$. Indeed, with the support projection e of x_+ we compute

$$x_+ = exe \leqslant eye = ey_+e - ey_-e \leqslant ey_+e.$$

On the other hand, since $e_{(s,\infty)}(ey_+e)$, $e_{(s,\infty)}((y_+)^{1/2}e(y_+)^{1/2})$ are equivalent projections and $(y_+)^{1/2}e(y_+)^{1/2} \leqslant y_+$, we have $x_+ \lesssim y_+$.

2.2. GENERALIZED SINGULAR NUMBER. For $x \in \mathcal{M}$ positive and $t > 0$ we set

$$\mu_t(x) = \inf\{s \geqslant 0;\ \tau(e^x_{(s,\infty)}) \leqslant t\}.$$

A positive operator x affiliated with $\mathcal{M}$ is τ-measurable exactly when $\tau(e^x_{(s,\infty)}) < \infty$ for some $s > 0$ (and consequently $\lim_{s\to\infty} \tau(e^x_{(s,\infty)}) = 0$ by the dominated convergence theorem). Thus, the quantity $\mu_t(x) < \infty$ (for each $t > 0$) makes sense for each positive τ-measurable operator x. It is known as the "t-th" generalized singular number, and it also admits the "min-max" representation

$$\mu_t(x) = \inf \sup\{(x\xi,\xi);\ \xi \in p\mathcal{H} \text{ and } \|\xi\| = 1\},$$

where the infimum is taken over all projections $p \in \mathcal{M}$ satisfying $\tau(1-p) \leqslant t$. Actually, the quantity $\mu_t(x)$ $(= \mu_t(|x|))$ is defined for an arbitrary $x \in \overline{\mathcal{M}}$, and $\{\mu_t(\cdot)\}_{t>0}$ serves as a continuous analog for singular numbers $\{\mu_n(\cdot)\}_{n=0,1,2,\dots}$ (see [12], [26]), i.e., the decreasing rearrangement of (positive) eigenvalues of the absolute value part of a matrix in question.

The following properties are useful:

(a) The spectral dominance $x \gtrsim y$ (for $x, y \geqslant 0$) implies $\mu_t(x) \geqslant \mu_t(y)$ (for $t > 0$).

(b) The trace value can be computed as

$$\tau(|x|) = \int_0^\infty \mu_s(x)\, \mathrm{d}s.$$

(c) We have $x \in V(\varepsilon, \delta) \iff \mu_\delta(x) \leqslant \varepsilon$ (see Lemma 3.1 of [11]). Thus, a sequence $\{x_n\}$ in $\overline{\mathcal{M}}$ tends to x in the measure topology if and only if

$$\lim_{n\to\infty} \mu_t(x - x_n) = 0 \quad (\text{for each } t > 0).$$

(d) If a sequence $\{x_n\}$ in $\overline{\mathcal{M}}$ tends to x in the measure topology, then we have

$$\mu_t(x) \leqslant \liminf_{n\to\infty} \mu_t(x_n)$$

for each $t > 0$ (see Lemma C in Appendix of [20] or Lemma 3.4 of [11]).

2.3. SPECTRAL SCALE. We assume that $\mathcal{M}$ is a finite von Neumann algebra equipped with a faithful normal trace τ satisfying $\tau(1) < \infty$. For $y \in \mathcal{M}_{\mathrm{sa}}$ the quantity

$$\lambda_t(y) = \inf\{s \in \mathbb{R};\ \tau(e^y_{(s,\infty)}) \leqslant t\} \quad (t \in (0, \tau(1)))$$

is known as the ("t-th") spectral scale of y. This notion is a continuous analog of the decreasing rearrangement of (real) eigenvalues of a Hermitian matrix.
The above three notions will play important roles throughout, and their details (as well as more information) can be found in [7], [10], [11] and [23] respectively.

2.4. WELL-DEFINEDNESS OF TRACE VALUES. Let y be a self-adjoint τ-measurable operator. We say that $\tau(y)$ is well-defined if either $\tau(y_+) < \infty$ or $\tau(y_-) < \infty$. In this case we can set

$$\tau(y) = \tau(y_+) - \tau(y_-) \quad (\in [-\infty, \infty]).$$

Let us summarize basic properties (see Lemma 8 and Lemma 9 of [7]):

(a) Let E be a τ-preserving conditional expectation (E to be used in Section 3 is very explicit). We compute

$$E(y)_+ - E(y)_- = E(y) = E(y_+ - y_-) = E(y_+) - E(y_-)$$

with $E(y_\pm) \geqslant 0$. Thus, minimality of the Jordan decomposition (relative to the spectral dominance) guarantees $E(y)_\pm \lesssim E(y_\pm)$ (see Lemma of [7] or the reasoning in 2.1,(c)) and consequently

$$\tau(E(y)_\pm) \leqslant \tau(E(y_\pm)) = \tau(y_\pm) \quad (\leqslant \infty).$$

(b) In particular, if $\tau(y)$ is well-defined, then so is $\tau(E(y))$ and we have of course

$$\tau(y) = \tau(E(y)).$$

(c) We assume that $\tau(y_1)$, $\tau(y_2)$ are well-defined for self-adjoint $y_i \in \overline{\mathcal{M}}$. If $\tau(y_1) + \tau(y_2)$ is well-defined (in the sense that "$\infty - \infty$" does not occur), then $\tau(y_1 + y_2)$ is also well-defined and

$$\tau(y_1 + y_2) = \tau(y_1) + \tau(y_2).$$

This can be easily shown based on $(y_1 + y_2)_\pm \lesssim (y_1)_\pm + (y_2)_\pm$, which is another consequence of minimality of the Jordan decomposition.

2.5. MISCELLANEOUS FACTS. Some facts needed in later sections will be collected here. We begin with the next fact (that is pointed out in Chapter II Section 5 of [12] for compact operators based on quite different arguments).

PROPOSITION 2.1. *Let E be a τ-preserving conditional expectation. For a τ-measurable operator x (in $\mathcal{M} + L^1(\mathcal{M};\tau)$) we have*

$$\int_0^t \mu_s(E(x))\,\mathrm{d}s \leqslant \int_0^t \mu_s(x)\,\mathrm{d}s \quad \textit{for each } t > 0.$$

Proof. We observe $\|E(x)\| \leqslant \|x\|$ and $\|E(x)\|_1 \leqslant \|x\|_1$, where $\|\cdot\|_1$ is the trace norm. The first inequality is obvious while for the second we note

$$\|E(y)\|_1 = \sup|\tau(E(y)x)| = \sup|\tau(yE(x))|.$$

Here the supremum is over all x's in the unit ball $\mathcal{M}_1$. But, for $x \in \mathcal{M}_1$ we estimate

$$|\tau(yE(x))| \leqslant \|y\|_1\|E(x)\| \leqslant \|y\|_1\|x\| \leqslant \|y\|_1.$$

Therefore, the desired result follows from the variational expression

$$\int_0^t \mu_s(x)\,\mathrm{d}s = \inf\{t\|x_0\| + \|x_1\|_1\}$$

(as a "K-functional"), where the infimum is taken over all decompositions $x = x_0 + x_1$ (see p. 289 of [11] and [22]) ∎

In the rest of the subsection we will assume $\tau(1) < \infty$ and deal with spectral scales $\lambda_t(\cdot)$ ($t \in (0, \tau(1))$). The next characterization is well-known for Hermitian matrices. The standard proof (presented in Theorem 1.1 of [3] for instance) can be easily modified to cover type II_1 von Neumann algebras (see Proposition 1.2 of [17]), whose details are presented here for the reader's convenience.

PROPOSITION 2.2. *For $x, y \in \mathcal{M}_{\mathrm{sa}}$ we have the weak majorization*

$$\int_0^t \lambda_s(x)\, \mathrm{d}s \leqslant \int_0^t \lambda_s(y)\, \mathrm{d}s \quad (t \in (0, \tau(1)))$$

if and only if $\tau((x - r1)_+) \leqslant \tau((y - r1)_+)$ for each $r \in \mathbb{R}$.

Proof. We assume the weak majorization. If $r > \lambda_0(x)$, then we have $(x - r1)_+ = 0$ and hence $\tau((x - r1)_+) = 0 \leqslant \tau((y - r1)_+)$. On the other hand, if $r \leqslant \lambda_{\tau(1)}(x)$, then we have $(x - r1)_+ = x - r1$ and

$$\tau((x-r1)_+) = \tau(x-r1) = \tau(x) - r\tau(1) \leqslant \tau(y) - r\tau(1) = \tau(y-r1) \leqslant \tau((y-r1)_+).$$

Finally, if either $r = \lambda_{s_0}(x)$ or $\lim\limits_{s \to s_0^-} \lambda_s(x) \geqslant r > \lambda_{s_0}(x)$, then we estimate

$$\begin{aligned}
\tau((x - r1)_+) &= \int_0^{\tau(1)} (\lambda_s(x) - r)_+ \mathrm{d}s = \int_0^{s_0} (\lambda_s(x) - r) \mathrm{d}s \\
&\leqslant \int_0^{s_0} (\lambda_s(y) - r) \mathrm{d}s \quad \text{(by the assumption)} \\
&\leqslant \int_0^{s_0} (\lambda_s(y) - r)_+ \mathrm{d}s \leqslant \int_0^{\tau(1)} (\lambda_s(y) - r)_+ \mathrm{d}s = \tau((y - r1)_+).
\end{aligned}$$

Conversely, when $\tau((x - r1)_+) \leqslant \tau((y - r1)_+)$ ($r \in \mathbb{R}$), for $t \in (0, \tau(1))$ we estimate

$$\begin{aligned}
\int_0^t \lambda_s(y)\, \mathrm{d}s &= \int_0^t \{\lambda_s(y) - \lambda_t(y)\} \mathrm{d}s + t\lambda_t(y) = \tau((y - \lambda_t(y)1)_+) + t\lambda_t(y) \\
&\geqslant \tau((x - \lambda_t(y)1)_+) + t\lambda_t(y) \quad \text{(by the assumption)} \\
&= \int_0^{\tau(1)} (\lambda_s(x) - \lambda_t(y))_+ \mathrm{d}s + t\lambda_t(y) \geqslant \int_0^t (\lambda_s(x) - \lambda_t(y))_+ \mathrm{d}s + t\lambda_t(y) \\
&\geqslant \int_0^t \{(\lambda_s(x) - \lambda_t(y)\} \mathrm{d}s + t\lambda_t(y) = \int_0^t \lambda_s(x)\, \mathrm{d}s,
\end{aligned}$$

and we are done. ∎

For $x \in \mathcal{M}$ we set

$$\widehat{x} = \begin{bmatrix} x & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{M} \otimes M_2(\mathbb{C}).$$

A generalized singular number $\mu_t(\cdot)$ satisfies $\mu_t(\widehat{x}) = \mu_t(x)$ (for $t < \tau(1)$), where $\mu_t(\widehat{x})$ is relative to the product trace $\tau \otimes \mathrm{Tr}_{M_2(\mathbb{C})}$ with the unnormalized trace $\mathrm{Tr}_{M_2(\mathbb{C})}$. Although a spectral scale $\lambda_t(\cdot)$ does not possess this property, we have

COROLLARY 2.3. *For $x, y \in \mathcal{M}_{\mathrm{sa}}$ we have*

$$\int_0^t \lambda_s(x)\,\mathrm{d}s \leqslant \int_0^t \lambda_s(y)\,\mathrm{d}s \quad (t \in (0, \tau(1)))$$

if and only if

$$\int_0^t \lambda_s(\widehat{x})\,\mathrm{d}s \leqslant \int_0^t \lambda_s(\widehat{y})\,\mathrm{d}s \quad (t \in (0, 2\tau(1))).$$

Proof. We note

$$(\widehat{x} - r1)_+ = \begin{bmatrix} (x - r1)_+ & 0 \\ 0 & (-r1)_+ 1 \end{bmatrix},$$
$$(\tau \otimes \mathrm{Tr}_{M_2(\mathbb{C})})((\widehat{x} - r1)_+) = \tau((x - r1)_+) + (-r)_+\tau(1)$$

(and similarly for y's) so that the result follows from Proposition 2.2. ∎

3. WEAK MAJORIZATION (SEMI-FINITE CASE)

In this section the trace Jensen inequality (in Section 1) for $x \in \mathcal{M}_{\mathrm{sa}}$ will be proved at first. More precisely it will be shown for semi-bounded τ-measurable operators (see Remark 3.3). Arguments here will actually enable us to obtain some weak majorization inequalities with the notion of generalized singular numbers (explained in 2.2).

Throughout the section $f(t)$ is a continuous convex function with $f(0) = 0$, and let us recall two lemmas from [7]:

LEMMA 3.1 ([7], Lemma 9). *Let $a \in \mathcal{M}$ be a contraction and we assume $x \in \mathcal{M}_{\mathrm{sa}}$, the self-adjoint operators in $\mathcal{M}$. For a unit vector ξ we have*

$$f((a^*xa\xi, \xi)) \leqslant (a^*f(x)a\xi, \xi).$$

LEMMA 3.2 ([7], Lemma 10). *Let $I \subset \mathbb{R}$ be an interval on which $f(t)$ is monotone (either increasing or decreasing). For a contraction $a \in \mathcal{M}$ and a self-adjoint $x \in \mathcal{M}$ we set $p = e_I^{a^*xa}$.*

(i) *If $f(t) \geqslant 0$ on I, then we have $pa^*f(x)ap \geqslant 0$ and moreover*

$$pa^*f(x)ap \gtrsim pf(a^*xa)p.$$

(ii) *If $f(t) \leqslant 0$ on I, then the negative part $(pa^*f(x)ap)_-$ $(\geqslant 0)$ of the Jordan decomposition satisfies*

$$-pf(a^*xa)p \gtrsim (pa^*f(x)ap)_-\,.$$

The former is just the Jensen inequality applied to the probability measure

$$\mu(I) = \|e_I^x a\xi\|^2 + (1 - \|a\xi\|^2)\, \delta_0(I) \quad \text{(for a subset } I \subset \mathbb{R}).$$

The latter (proved based on the former) was the main technical ingredient in [7] (where $x \geqslant 0$ and the concavity of $f(t)$ were assumed). Proofs there work in the current case as well with obvious modifications so that details are left to the reader.

REMARK 3.3. For a general (unbounded) $x = x^*$ the notion of an "inner product" $(a^*xa\xi, \xi)$ makes no sense. To avoid this difficulty, let us assume that a self-adjoint τ-measurable operator x is *semi-bounded*, i.e.,

$$\text{either } -m \leqslant x \quad \text{or} \quad x \leqslant m \quad \text{(for some } m \in \mathbb{R}_+).$$

We assume $-m \leqslant x$ for instance, and we set $x_n = x\chi_{[-m,n]}(x) \in \mathcal{M}_{\rm sa}$ for each $n \in \mathbb{N}$. Since $f(t)$ is monotone (increasing or decreasing) for t large, we can set

$$(a^* f(x)a\xi, \xi) = \lim_{n\to\infty} (a^* f(x_n)a\xi, \xi)$$

belonging to $(-\infty, \infty]$ (respectively $[-\infty, \infty)$) in the increasing (respectively decreasing) case. Note that $(a^*xa\xi, \xi) \in (-\infty, \infty]$ is well-defined due to lower semi-boundedness $-m \leqslant x$. When $(a^*xa\xi, \xi) = \infty$, we set

$$f((a^*xa\xi, \xi)) = \lim_{t\to\infty} f(t).$$

With the conventions so far, Lemma 3.1 obviously remains valid for $x \geqslant -m$ (by the obvious limiting argument). Of course we can play a similar game for a upper semi-bounded $x \leqslant m$.

Lemma 3.2 is just based on Lemma 3.1 (together with careful analysis on Murray–von Neumann equivalence of relevant spectral projections), and consequently it remains valid for semi-bounded (self-adjoint) τ-measurable operators.

Let us consider the case where $f(t)$ is decreasing at the origin in this section. (The opposite case can be handled by considering $f(-t)$ and $-x$.) We assume that $f(t)$ is

(i) positive and decreasing on $I_1 = (-\infty, 0)$,
(ii) negative and decreasing on $I_2 = [0, t_1)$,
(iii) negative and increasing on $I_3 = [t_1, t_2)$,
(iv) positive and increasing on $I_4 = [t_2, \infty)$.

Some of these intervals could be $\emptyset$. For example, we have $I_2 = I_3 = \emptyset$ for $f(t) \geqslant 0$, and $I_3 = I_4 = \emptyset$ for $f(t)$ decreasing.

Let x be a self-adjoint operator in $\mathcal{M}$, and we set

$$p_i = e_{I_i}^{a^*xa} \quad (i = 1, 2, 3, 4) \quad \text{and} \quad E(y) = \sum_{i=1}^{4} p_i y p_i.$$

Since p_i's are spectral projections, we have $p_i f(a^*xa) p_i = f(p_i a^*xa p_i)$ due to $f(0) = 0$. It is convenient and intuitive to express $f(a^*xa)$ $(= E(f(a^*xa)))$ in the matrix form:

$$f(a^*xa) = \begin{bmatrix} p_1 f(a^*xa)p_1 & 0 & 0 & 0 \\ 0 & p_2 f(a^*xa)p_2 & 0 & 0 \\ 0 & 0 & p_3 f(a^*xa)p_3 & 0 \\ 0 & 0 & 0 & p_4 f(a^*xa)p_4 \end{bmatrix}$$

with

$$p_i f(a^*xa)p_i \geqslant 0 \quad (i = 1,4) \quad \text{and} \quad p_i f(a^*xa)p_i \leqslant 0 \quad (i = 2,3).$$

On the other hand, $a^*f(x)a$ is not necessarily diagonal and we have

$$E(a^*f(x)a) = \begin{bmatrix} p_1 a^*f(x)ap_1 & 0 & 0 & 0 \\ 0 & p_2 a^*f(x)ap_2 & 0 & 0 \\ 0 & 0 & p_3 a^*f(x)ap_3 & 0 \\ 0 & 0 & 0 & p_4 a^*f(x)ap_4 \end{bmatrix}$$

with

$$p_i a^*f(x)ap_i \geqslant 0 \quad (i = 1,4)$$

(thanks to the first part of Lemma 3.2(i)). The Jordan decomposition

$$f(a^*xa) = f(a^*xa)_+ - f(a^*xa)_-$$

(with $f(a^*xa)_\pm \geqslant 0$ and orthogonal supports) is obviously given by

$$f(a^*xa)_+ = \begin{bmatrix} p_1 f(a^*xa)p_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_4 f(a^*xa)p_4 \end{bmatrix},$$

$$f(a^*xa)_- = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -p_2 f(a^*fxa)p_2 & 0 & 0 \\ 0 & 0 & -p_3 f(a^*xa)p_3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

On the other hand, the Jordan decomposition of $a^*f(x)a$ is difficult to describe. However, the one for $E(a^*f(x)a)$ is simply given by

$$(E(a^*f(x)a))_+ = \begin{bmatrix} p_1 a^*f(x)ap_1 & 0 & 0 & 0 \\ 0 & (p_2 a^*f(x)ap_2)_+ & 0 & 0 \\ 0 & 0 & (p_3 a^*f(x)ap_3)_+ & 0 \\ 0 & 0 & 0 & p_4 a^*f(x)ap_4 \end{bmatrix},$$

$$(E(a^*f(x)a))_- = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & (p_2 a^*f(x)ap_2)_- & 0 & 0 \\ 0 & 0 & (p_3 a^*f(x)ap_3)_- & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Lemma 3.2 enables us to compare the diagonal blocks of $f(a^*xa)_\pm$ with those of $(E(a^*f(x)a))_\pm$:

$$\begin{aligned} p_ia^*f(x)ap_i &\gtrsim p_if(a^*xa)p_i \quad \text{for } i=1,4, \\ -p_if(a^*xa)p_i &\gtrsim (p_ia^*f(x)ap_i)_- \quad \text{for } i=2,3. \end{aligned}$$

The spectral dominance is preserved under taking a direct sum and we have

$$(E(a^*f(x)a))_+ \geqslant \begin{bmatrix} p_1a^*f(x)ap_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_4a^*f(x)ap_4 \end{bmatrix} \gtrsim f(a^*xa)_+, \tag{3.1}$$

$$f(a^*xa)_- \gtrsim (E(a^*f(x)a))_- \,. \tag{3.2}$$

Note that (3.1) implies

$$(E(a^*f(x)a))_+ \gtrsim f(a^*xa)_+ \quad \text{and} \quad \tau((E(a^*f(x)a))_+) \geqslant \tau(f(a^*xa)_+)$$

while (3.2) says

$$\tau(f(a^*xa)_-) \geqslant \tau((E(a^*f(x)a))_-).$$

Let us assume that both of $\tau(a^*f(x)a)$, $\tau(f(a^*xa))$ are well-defined (see 2.4). Then, so is $\tau(E(a^*f(x)a))$, and from the above inequalities on trace values we have

$$\begin{aligned} \tau(E(a^*f(x)a)) &= \tau((E(a^*f(x)a))_+) - \tau((E(a^*f(x)a))_-) \\ &\geqslant \tau(f(a^*xa)_+) - \tau(f(a^*xa)_-) = \tau(f(a^*xa)). \end{aligned}$$

But, since E preserves τ, the above means the following trace Jensen inequality:

$$\tau(a^*f(x)a) \geqslant \tau(f(a^*xa)).$$

Note that the middle part in (3.1) is

$$(E(a^*f(x)a))_+ - (p_2a^*f(x)ap_2)_+ - (p_3a^*f(x)ap_3)_+ \;(\gtrsim f(a^*xa)_+).$$

Thus, it is possible to strengthen the above trace inequality as follows:

$$\tau(a^*f(x)a) \geqslant \tau(f(a^*xa)) + \tau((p_2a^*f(x)ap_2)_+) + \tau((p_3a^*f(x)ap_3)_+).$$

Discussions so far obviously remain valid for semi-bounded τ-measurable operators (see Remark 3.3), and hence we have shown the next result.

THEOREM 3.4. *We assume that $a \in \mathcal{M}$ is a contraction and x is a semi-bounded τ-measurable operator. For a continuous convex function $f(t)$ with $f(0)=0$ the trace Jensen inequality*

$$\tau(f(a^*xa)) \leqslant \tau(a^*f(x)a)$$

holds true as long as the both sides are well-defined.

The semi-boundedness requirement here can be dropped in certain cases. For example, when $f(t) \geqslant 0$ on $\mathbb{R}$, for any self-adjoint τ-measurable operator x we have

$$\tau(f(a^*xa)) \leqslant \tau(a^*f(x)a) \quad (\leqslant \infty).$$

To prove this based on "Fatou's lemma for traces" (see Theorem 3.5 of [11] for instance) is an easy exercise, and the next Theorem 3.5 actually covers this fact.

THEOREM 3.5. *Let $f(t)$ be a continuous convex function satisfying $f(0) = 0$. For a contraction $a \in \mathcal{M}$ and a self-adjoint τ-measurable operator x the following weak majorization holds true*:

$$\int_0^t \mu_s(f(a^*xa)_+)\,\mathrm{d}s \leqslant \int_0^t \mu_s((a^*f(x)a)_+)\,\mathrm{d}s \quad (\textit{for each } t > 0).$$

In particular, by letting $t \to \infty$, we always have

$$\tau(f(a^*xa)_+) \leqslant \tau((a^*f(x)a)_+) \ (\leqslant \infty).$$

Proof. Let us begin with the case when x is a semi-bounded τ-measurable operator. We have the spectral dominance

$$(E(a^*f(x)a))_+ \lesssim E((a^*f(x)a)_+)$$

(see 2.4(a)), showing (together with (3.1))

$$f(a^*xa)_+ \lesssim E((a^*f(x)a)_+).$$

Thus, the weak majorization in this case follows from Proposition 2.1.

Let us move to the general case. It suffices to consider the following three situations:

(i) $f(t) \geqslant 0$ on $\mathbb{R}$,
(ii) $f(t)$ is monotone decreasing on $\mathbb{R}$,
(iii) $f(t)$ is monotone decreasing on $(-\infty, t_0]$ with $t_0 > 0$, increasing on $[t_0, \infty)$ and $\lim\limits_{t\to\infty} f(t) = \infty$.

Note that the case when $f(t)$ is increasing at $t = 0$ can be reduced to the above (ii) or (iii) by considering $f(-t)$ and $-x$. Let us approximate x by the following semi-bounded operators:

$$x_n = \begin{cases} x\chi_{[-n,n]}(x) & \text{in case (i),} \\ x\chi_{[-n,\infty)}(x) & \text{in cases (ii) and (iii).} \end{cases}$$

Note $f(x_n) \leqslant f(x)$ in all the cases (since $f(t) \geqslant 0$ on $(-\infty, 0]$ in cases (ii) and (iii)). We thus have the spectral dominance $(a^*f(x_n)a)_+ \lesssim (a^*f(x)a)_+$ (see 2.1(c)). Since $f(a^*x_na)_+ \to f(a^*xa)_+$ in measure (see [28]), lower semi-continuity of $\mu_s(\cdot)$

relative to the measure topology (see 2.2(d)) and Fatou's lemma guarantee

$$\int_0^t \mu_s(f(a^*xa)_+)\,\mathrm{d}s \leqslant \int_0^t \liminf_{n\to\infty} \mu_s(f(a^*x_na)_+)\,\mathrm{d}s \leqslant \liminf_{n\to\infty} \int_0^t \mu_s(f(a^*x_na)_+)\,\mathrm{d}s$$
$$\leqslant \liminf_{n\to\infty} \int_0^t \mu_s((a^*f(x_n)a)_+)\,\mathrm{d}s \leqslant \int_0^t \mu_s((a^*f(x)a)_+)\,\mathrm{d}s.$$

The third inequality here of course follows from the first half of the proof. ∎

Another weak majorization akin to Theorem 3.5 is also possible: We have

$$(p_1+p_4)a^*f(x)a(p_1+p_4) = \begin{bmatrix} p_1a^*f(x)ap_1 & 0 & 0 & p_1a^*f(x)ap_4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ p_4a^*f(x)ap_1 & 0 & 0 & p_4a^*f(x)ap_4 \end{bmatrix}.$$

By cutting off-diagonal blocks, i.e., $E((p_1+p_4)a^*f(x)a(p_1+p_4))$, we get the middle part (which majorizes $f(a^*xa)_+$ in the sense of spectral dominance) in (3.1). Since the sum $e = p_1 + p_4$ is nothing but the support projection of $f(a^*xa)_+$, we conclude

$$\int_0^t \mu_s(f(a^*xa)_+)\,\mathrm{d}s \leqslant \int_0^t \mu_s(e(a^*f(x)a)e)\,\mathrm{d}s \quad \text{for each } t > 0.$$

Related estimates for the special convex function $f(t) = |t|^r$ with $r \geqslant 1$ (and for compact operators) were studied in [18], [19].

COROLLARY 3.6. *Let $f(t)$ be a continuous convex function with $f(0) = 0$. For self-adjoint τ-measurable operators x, y and $a, b \in \mathcal{M}$ satisfying $a^*a + b^*b \leqslant 1$ we have*

$$\int_0^t \mu_s(f(a^*xa+b^*yb)_+)\,\mathrm{d}s \leqslant \int_0^t \mu_s((a^*f(x)a+b^*f(y)b)_+)\,\mathrm{d}s \quad (\textit{for each } t > 0).$$

Proof. We set

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

and observe that $A \in \mathcal{M} \otimes M_2(\mathbb{C})$ is a contraction. It is also elementary to see

$$f(A^*XA)_+ = \begin{bmatrix} f(a^*xa+b^*yb)_+ & 0 \\ 0 & 0 \end{bmatrix}, \quad (A^*f(X)A)_+ = \begin{bmatrix} (a^*f(x)a+b^*f(y)b)_+ & 0 \\ 0 & 0 \end{bmatrix},$$

so that (from the very definition of $\mu_t(\cdot)$) we have

$$\mu_t(f(A^*XA)_+) = \mu_t(f(a^*xa+b^*yb)_+),$$
$$\mu_t((A^*f(X)A)_+) = \mu_t((a^*f(x)a+b^*f(y)b)_+),$$

(where the left side $\mu_t(\cdot)$ is relative to the product trace $\tau \otimes \mathrm{Tr}_{M_2(\mathbb{C})}$ with the unnormalized trace $\mathrm{Tr}_{M_2(\mathbb{C})}$). Thus, the result follows from Theorem 3.5. ∎

Notice that majorization of the form

$$\int_0^t \mu_s(f(a^*xa))\,\mathrm{d}s \leqslant \int_0^t \mu_s(a^*f(x)a)\,\mathrm{d}s \quad (t > 0)$$

cannot be expected (since $\|f(a^*xa)\| \leqslant \|a^*f(x)a\|$ is simply false). The fact that $\mu_s(\cdot)$ destroys information on "negative eigenvalues" is responsible for this failure, and majorization of the above form holds true for spectral scales $\lambda_s(\cdot)$ (attached to a finite trace τ) as will be clarified in the next section (see Theorem 4.2).

Operator triangle inequalities (for the absolute value part $|x| = \sqrt{x^*x}$) of the form

$$|x+y| \leqslant u|x|u^* + v|y|v^*$$

(with u, v unitaries or something alike) have been studied by several authors (see [1], [27]). Note that the operator inequality $|x+y| \leqslant |x| + |y|$ is false even for Hermitian matrices. For example, 2×2 matices $x = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $y = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}$ give us

$$|x+y| = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \quad \text{and} \quad |x| + |y| = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix},$$

and hence $|x+y| \not\leqslant |x| + |y|$ (see p. 1 of [26]). The eigenvalus of the latter are

$$\mu_1(|x| + |y|) = 2 + \sqrt{2}, \quad \mu_2(|x| + |y|) = 2 - \sqrt{2},$$

and hence the spectral dominance $|x+y| \lesssim |x| + |y|$ is not valid either.

On the other hand, Corollary 3.6 applied to the positive convex function $f(t) = |t|$ and $a = b = 1/\sqrt{2}$ yields

$$\int_0^t \mu_s(|x+y|)\,\mathrm{d}s \leqslant \int_0^t \mu_s(|x| + |y|)\,\mathrm{d}s \tag{3.3}$$

for self-adjoint τ-measurable operators x, y. For $x = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $y = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ we compute

$$\|x+y\| \;(= \sqrt{2}) \not\leqslant \|\,|x| + |y|\,\| \;(= 1).$$

showing that the weak majorization (3.3) cannot be expected for general x, y.

THEOREM 3.7. *For $a, b \in \mathcal{M}$ we have*

$$\int_0^t f(\mu_s(a+b))\,\mathrm{d}s \leqslant \int_0^t f(\mu_s(|a| + |b|)^{1/2}\mu_s(|a^*| + |b^*|)^{1/2})\mathrm{d}s \quad (t > 0)$$

for a continuous increasing function $f(t)$ on $[0,\infty)$ such that $t \to f(e^t)$ is convex and $f(0)=0$. In particular, (with $f(t)=t$) we have

$$\int_0^t \mu_s(a+b)\,\mathrm{d}s \leqslant \int_0^t \mu_s(|a|+|b|)^{1/2}\mu_s(|a^*|+|b^*|)^{1/2}\mathrm{d}s$$
$$\leqslant \frac{1}{2}\int_0^t [\mu_s(|a|+|b|)+\mu_s(|a^*|+|b^*|)]\mathrm{d}s.$$

Proof. For $a \in \mathcal{M}$ (with the polar decomposition $a = u|a|$) we have

$$\begin{bmatrix} |a| & a^* \\ a & |a^*| \end{bmatrix} = \begin{bmatrix} |a| & |a|u^* \\ u|a| & u|a|u^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & u \end{bmatrix}\begin{bmatrix} |a| & |a| \\ |a| & |a| \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & u \end{bmatrix}^* \geqslant 0$$

(and positivity of the corresponding matrix for b), and consequently

$$\begin{bmatrix} |a|+|b| & a^*+b^* \\ a+b & |a^*|+|b^*| \end{bmatrix} \geqslant 0.$$

Hence, there exists a contraction $c \in \mathcal{M}$ satisfying

$$a+b = (|a^*|+|b^*|)^{1/2} c\,(|a|+|b|)^{1/2}$$

(see [2] for instance). Since $\Lambda_t(\cdot) = \exp\left(\int_0^t \log \mu_s(\cdot)\mathrm{d}s\right)$ is submultiplicative (see [10] or Theorem 4.2 of [11]) and $\Lambda_t(c) \leqslant 1$, we get

$$\Lambda_t(a+b) \leqslant \Lambda_t((|a^*|+|b^*|)^{1/2})\Lambda_t(c)\Lambda_t((|a|+|b|)^{1/2})$$
$$\leqslant \Lambda_t((|a|+|b|))^{1/2}\Lambda_t((|a^*|+|b^*|))^{1/2}$$

(for each $t>0$), which means

$$\int_0^t \log \mu_s(a+b)\,\mathrm{d}s \leqslant \int_0^t \frac{1}{2}[\log \mu_s(|a|+|b|) + \log \mu_s(|a^*|+|b^*|)]\mathrm{d}s.$$

The result thus follows from standard majorization theory. ∎

Note that we have $\Lambda_t(a+b) \leqslant \Lambda_t(|a|+|b|)$ for a, b normal, a considerable strengthening of (3.3). Also, standard approximation arguments (similar to those in the proof of Theorem 3.5) enable us to get the above estimates for general τ-measurable operators.

4. WEAK MAJORIZATION (FINITE TRACE CASE)

In this section $\mathcal{M}$ is a finite von Neumann algebra without minimal projections equipped with a faithful normal trace τ satisfying $\tau(1) < \infty$. With the notion of spectral scales (in 2.3) we will investigate weak majorization inequalities.

For simplicity only bounded self-adjoint operators will be dealt with although spectral scales can be defined for unbounded ones as well.

LEMMA 4.1. *Let $f(t)$ be a continuous convex function with $f(0) = 0$. For $x \in \mathcal{M}_{\mathrm{sa}}$ and a contraction $a \in \mathcal{M}$ we have*

$$\tau(f(a^*xa)) \leqslant \tau(a^*f(x)a).$$

It was already proved in Theorem 3.4. In [24] the result is obtained with $x \geqslant 0$ and a positive contraction map $x \to \alpha(x)$. Arguments there can actually deal with self-adjoint operators, and slight modifications give us a completely elementary proof for the lemma. This proof is quite direct in the sense that no technical apparatus such as $\lesssim$ is needed, and it will be presented in the appendix for the reader's convenience.

THEOREM 4.2. *Let $f(t)$ be a continuous convex function with $f(0) \leqslant 0$. For $x \in \mathcal{M}_{\mathrm{sa}}$ and a contraction $a \in \mathcal{M}$ we have*

$$\int_0^t \lambda_s(f(a^*xa))\,\mathrm{d}s \leqslant \int_0^t \lambda_s(a^*f(x)a)\,\mathrm{d}s \quad (\text{for each } t \in (0,\tau(1))).$$

Proof. At first we assume $f(0) = 0$. For each $t \in (0,\tau(1))$ we can find a projection $e \in \mathcal{M}$ commuting with a^*xa such that

$$\int_0^t \lambda_s(f(a^*xa))\,\mathrm{d}s = \tau(ef(a^*xa)e)$$

and $\tau(e) \leqslant t$ (see Lemma 1 of [23] or Lemma 4.1 of [11]). Since $[e, f(a^*xa)] = 0$ and $f(0) = 0$, we have $ef(a^*xa)e = f(ea^*xae)$. Thus, Lemma 4.1 yields

$$\int_0^t \lambda_s(f(a^*xa))\,\mathrm{d}s = \tau(f(ea^*xae)) \leqslant \tau(ea^*f(x)ae) = \tau(a^*f(x)ae),$$

which is majorized by $\int_0^t \lambda_s(a^*f(x)a)\,\mathrm{d}s$ due to $\tau(e) \leqslant t$ (see Theorem 3 of [23]).

When $f(0) \leqslant 0$, with $g(t) = f(t) - f(0)$ (vanishing at 0) we have

$$\lambda_s(f(a^*xa)) = \lambda_s(g(a^*xa) + f(0)1) = \lambda_s(g(a^*xa)) + f(0),$$
$$\lambda_s(a^*f(x)a) \geqslant \lambda_s(a^*g(x)a + f(0)1) = \lambda_s(a^*g(x)a) + f(0),$$

(due to $f(0)a^*a \geqslant f(0)1$) and get the result by applying the first half to $g(t)$. ∎

When the convex function $f(t)$ is monotone (either increasing or decreasing), the conclusion of the theorem can be strengthened to

$$\lambda_t(f(a^*xa)) \leqslant \lambda_t(a^*f(x)a) \quad \text{for each } t \in (0,\tau(1)).$$

In fact, the "min-max" representation in 2.2 is also available for $\lambda_s(\cdot)$ (as is explained in [23]). Thus, by assuming increasingness of $f(t)$ (use $f(-t)$ and $-x$

otherwise), one can repeat arguments in the proof of Lemma 4.5 of [11] together with Lemma 3.1 (which is easily shown to remain valid as long as $f(0) \leqslant 0$).

COROLLARY 4.3. *Let $f(t)$ be a continuous convex function with $f(0) \leqslant 0$. For $x, y \in \mathcal{M}_{\mathrm{sa}}$ and $a, b \in \mathcal{M}$ satisfying $a^*a + b^*b \leqslant 1$ we have*

$$\int_0^t \lambda_s(f(a^*xa + b^*yb))\,\mathrm{d}s \leqslant \int_0^t \lambda_s(a^*f(x)a + b^*f(y)b)\,\mathrm{d}s \quad (\textit{for each } t \in (0, \tau(1))).$$

Moreover, when $a^*a + b^*b = 1$, the above holds true regardless of the parity of $f(0)$.

Proof. When $f(0) = 0$, we can use the 2×2 matrix trick used in the proof of Corollary 3.6 and the desired weak majorization is a consequence of Theorem 4.2. One subtlety here is that $\lambda_t(f(A^*XA)) = \lambda_t(f(a^*xa + b^*yb))$ and $\lambda_t(A^*f(X)A) = \lambda_t(a^*f(x)a + b^*f(y)b)$ are false (which should be compared with the situation for $\mu_t(\cdot)$). However, Corollary 2.3 is in rescue. On the other hand, to deal with the general case $f(0) \leqslant 0$, we can repeat the same trick as in the proof of Theorem 4.2 thanks to $f(0)(a^*a + b^*b) \geqslant f(0)1$ (which comes free when $a^*a + b^*b = 1$). ∎

5. MONOTONICITY

Let $f(t)$ be a continuous increasing function on $\mathbb{R}$ satisfying $f(0) = 0$. Here, we will collect various monotonicity properties on trace values $\tau(f(x))$ for x self-adjoint.

THEOREM 5.1. *Let $f(t)$ be a continuous increasing function (on $\mathbb{R}$) with $f(0) = 0$, and we assume that self-adjoint τ-measurable operators x, y satisfy $x \leqslant y$.*

(i) *We have the spectral dominance $f(x)_+ \precsim f(y)_+$, $f(y)_- \precsim f(x)_-$, and the trace inequality $\tau(f(x)) \leqslant \tau(f(y))$ holds true as long as the both sides are well-defined (in particular when $f(x), f(y) \in L^1(\mathcal{M}; \tau)$).*

(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 have $x = y$ if and only if $\tau(f(x)) = \tau(f(y))$.*

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

Proof. (i) Note $f(x)_+ = f(x_+)$, $f(y)_+ = f(y_+)$ since $f(t)$ is increasing and $f(0) = 0$. On the other hand, with the (increasing) function $g(t) = -f(-t)$ we observe

$$f(x)_- = (-f(x))_+ = g(-x)_+ = g((-x)_+) = g(x_-)$$

and similarly $f(y)_- = g(y_-)$. Hence, the Jordan decompositions of $f(x), f(y)$ are

$$\begin{cases} f(x) = f(x)_+ - f(x)_- = f(x_+) - g(x_-), \\ f(y) = f(y)_+ - f(y)_- = f(y_+) - g(y_-). \end{cases}$$

Here, we have $y_+ \gtrsim x_+$, i.e., $\mu_t(y_+) \geqslant \mu_t(x_+),\ t > 0$ (as was seen in 2.1(c)) while $-x \geqslant -y$ yields $x_- = (-x)_+ \gtrsim (-y)_+ = y_-$. Since $f(t), g(t)$ are increasing, we actually have $f(y_+) \gtrsim f(x_+)$ and $g(x_-) \gtrsim g(y_-)$ (see 2.1(b)). Therefore, we have

$$\tau(f(x_+)) = \int_0^\infty f(\mu_t(x_+))\,\mathrm{d}t \leqslant \int_0^\infty f(\mu_t(y_+))\,\mathrm{d}t = \tau(f(y_+)) \tag{5.1}$$

(and similarly $\tau(g(y_-)) \leqslant \tau(g(x_-))$), and hence

$$\begin{aligned}\tau(f(y)) - \tau(f(x)) &= (\tau(f(y_+)) - \tau(g(y_-))) - (\tau(f(x_+)) - \tau(g(x_-)))\\ &= (\tau(f(y_+)) - \tau(f(x_+))) + (\tau(g(x_-)) - \tau(g(y_-))) \geqslant 0.\end{aligned}$$

(ii) The assumption $\tau(f(x)) = \tau(f(y))$ and the above arguments in (i) force

$$\tau(f(x_+)) = \tau(f(y_+)) \quad \text{and} \quad \tau(g(x_-)) = \tau(g(y_-)). \tag{5.2}$$

From the first equality we have $f(\mu_t(x_+)) = f(\mu_t(y_+))$ (see (5.1)) and $\mu_t(x_+) = \mu_t(y_+)$ (for $f(t)$ strictly increasing). We similarly have $\mu_t(x_-) = \mu_t(y_-)$. We claim

$$x_+ \leqslant y_+ \quad \text{and} \quad y_- \leqslant x_- \tag{5.3}$$

(in the usual positive definite sense), and note that the conclusion $x = y$ is obtained once this claim is shown. Indeed, the obvious computation

$$\tau(y_+ - x_+) = \tau(y_+) - \tau(x_+) = \int_0^\infty \mu_t(y_+)\,\mathrm{d}t - \int_0^\infty \mu_t(x_+)\,\mathrm{d}t = 0$$

and the similar one $\tau(x_- - y_-) = 0$ (with $x, y \in L^1(\mathcal{M}; \tau)$) yield $x_\pm = y_\pm$.

To show the claim, we note

$$\tau(x_+) \leqslant \tau(eye) = \tau(ey_+e) - \tau(ey_-e) \leqslant \tau(ey_+e) \leqslant \tau(y_+)$$

(see 2.1(c)). But, since $\tau(x_+) = \tau(y_+) < \infty$, we actually have

$$\tau(ey_-e) = 0, \quad \tau(y_+) = \tau(ey_+e) \ (= \tau(y_+^{1/2} e y_+^{1/2})),$$

showing $y_-e = 0$ and $y_+(1-e) = 0$. Hence, the support of y_- is majorized by $1-e$ and that of y_+ is majorized by e. Since $y_\pm$ have orthogonal supports, this means

$$y_+ = eye \quad \text{and} \quad y_- = -(1-e)y(1-e). \tag{5.4}$$

The same expressions for $x_\pm$ are obviously valid (always) and (5.3) holds true:

$$y_+ - x_+ = e\,(y-x)\,e \geqslant 0, \quad x_- - y_- = -(1-e)(x-y)(1-e) \geqslant 0.$$

(iii) The equality (5.2) is still valid, and the main issue here is to prove (5.3) in the current setting (i.e., without $x, y \in L^1(\mathcal{M}; \tau)$). We note

$$\begin{aligned}\tau(f(x_+)) &\leqslant \tau(f(eye)) \quad (\text{because of } 0 \leqslant x_+ \leqslant eye)\\ &\leqslant \tau(ef(y)e) \ \leqslant\ \tau(ef(y_+)e) \ \leqslant\ \tau(f(y_+)).\end{aligned} \tag{5.5}$$

The second inequality is the trace Jensen inequality (Theorem 3.4). However, y is not necessarily semi-bounded so that some justification is needed here. To do so, we set $y_n = y\chi_{[-n,\infty)}(y)$ for each $n \in \mathbb{N}$. The semi-bounded operators y_n obviously satisfy $y \leqslant y_n$. Also, the assumption $f(y) \in L^1(\mathcal{M};\tau)$ guarantees $f(y_n) \in L^1(\mathcal{M};\tau)$ and

$$\lim_{n\to\infty} \|f(y_n) - f(y)\|_1 = 0. \tag{5.6}$$

Note $0 \leqslant x_+ = exe \leqslant eye \leqslant ey_ne$ and $f(ey_ne) \geqslant 0$. In particular, $\tau(f(ey_ne))$ is well-defined. Also so is $\tau(ef(y_n)e)$ because of $f(y_n), ef(y_n)e \in L^1(\mathcal{M};\tau)$. Since $f(ey_ne) \to f(eye)$ $(\geqslant 0)$ in measure (see [28]), "Fatou's lemma for traces" (see Theorem 3.5 of [11]), Theorem 3.4 and (5.6) altogether yield the desired estimate:

$$\tau(f(eye)) \leqslant \liminf_{n\to\infty} \tau(f(ey_ne)) \leqslant \liminf_{n\to\infty} \tau(ef(y_n)e) = \tau(ef(y)e).$$

The quantities appearing in (5.5) are all finite, and the equality (5.2) means

$$\tau(eg(y_-)e) = 0 \quad \text{and} \quad \tau(f(y_+)) = \tau(ef(y_+)e).$$

Thus, the support of $g(y_-)$ (respectively $f(y_+)$) is majorized by e (respectively $1-e$). However, since $f(y_+)$, y_+ and $g(y_-)$, y_- have same supports, (5.4) and hence (5.3) remain valid.

We next make use of the estimates

$$\tau(f(x_+)) \leqslant \tau(f(x_+)) + \tau(f(y_+ - x_+)) \leqslant \tau(f(y_+))$$

(see Proposition 4.6(ii) of [11]). The equality (5.2) shows $\tau(f(y_+ - x_+)) = 0$ and consequently $f(y_+ - x_+) = 0$. Finally, $f(t)$ being strictly increasing with $f(0) = 0$, we conclude $y_+ - x_+ = 0$. Similar arguments also yield $x_- - y_- = 0$ and we are done. ∎

APPENDIX A. DIRECT PROOF OF LEMMA 4.1

In this appendix a direct proof of Lemma 4.1 is presented.

We choose and fix an arbitrary $\varepsilon > 0$. We can then choose $\delta > 0$ satisfying

$$|s-t| \leqslant \delta \quad \Longrightarrow \quad |f(s) - f(t)| \leqslant \varepsilon.$$

for $s,t \in [-\|x\|, \|x\|] \cup [-\|a^*xa\|, \|a^*xa\|]$. Let

$$x = \int_{-\|x\|}^{\|x\|} s\, de_s^x \quad \text{and} \quad a^*xa = \int_{-\|a^*xa\|}^{\|a^*xa\|} t\, de_t^{a^*xa}$$

be the spectral decomposition of x and a^*xa respectively. We divide the intervals $[-\|x\|, \|x\|]$ and $[-\|a^*xa\|, \|a^*xa\|]$ into subintervals of length at most δ:

$$s_0 = -\|x\| < s_1 < s_2 < \cdots < s_n = \|x\|,$$
$$t_0 = -\|a^*xa\| < t_1 < t_2 < \cdots < t_m = \|a^*xa\|.$$

Let

$$p_1 = e^x_{[s_0,s_1]}, \quad p_i = e^x_{(s_{i-1},s_i]} \quad (i = 2,3,\cdots,n),$$
$$q_1 = e^{a^*xa}_{[t_0,t_1]}, \quad q_j = e^{a^*xa}_{(t_{j-1},t_j]} \quad (j = 2,3,\cdots,m),$$

be the corresponding spectral projections, and we set $y = \sum_{i=1}^{n} s_i p_i$. We have

$$\|x-y\| \leqslant \delta \tag{A.1}$$

from the construction while uniform continuity guarantees

$$\|f(x)-f(y)\| \leqslant \varepsilon. \tag{A.2}$$

To approximate $\tau(f(a^*xa))$ by a Riemann sum (of the form $\sum_{j=1}^{m} f(\xi_j)\tau(q_i)$), we set

$$\xi_j = \frac{\tau(a^*xaq_j)}{\tau(q_j)} \quad (j = 1,2,\ldots,m)$$

with the following convention: When $\tau(q_j) = 0$, we do not define ξ_i and simply omit j from the sum $\sum_{j=1}^{m}$. The obvious fact $t_{j-1}q_j \leqslant a^*xaq_j \leqslant t_jq_j$ yields

$$t_{j-1}\tau(q_j) \leqslant \tau(a^*xaq_j) \leqslant t_j\tau(q_j) \quad \text{and hence} \quad t_{j-1} \leqslant \xi_j \leqslant t_j.$$

So uniform continuity shows $|f(t)-f(\xi_j)| \leqslant \varepsilon$ on the j-th subinterval, and we have

$$\Big|\tau(f(a^*xa))-\sum_{j=1}^{m} f(\xi_j)\tau(q_j)\Big| = \Big|\int_{-\|a^*xa\|}^{\|a^*xa\|} f(t)\,\mathrm{d}\tau(e^{a^*xa}_t)-\sum_{j=1}^{m} f(\xi_j)\tau(q_j)\Big| \leqslant \varepsilon\tau(1). \tag{A.3}$$

Let us fix j satisfying $\tau(q_j) > 0$. We remark

$$f(\xi_j) = f\Big(\frac{\tau(a^*xaq_j)}{\tau(q_j)}\Big) = f\Big(\frac{\tau(a^*yaq_j)}{\tau(q_j)} + \frac{\tau(a^*(x-y)aq_j)}{\tau(q_j)}\Big),$$
$$|\tau(a^*(x-y)aq_j)| \leqslant \|a^*(x-y)a\|\tau(q_j) \leqslant \|a\|^2\|x-y\|\tau(q_j) \leqslant \delta\tau(q_j),$$

(see (A.1)) so that (by uniform continuity again) we have

$$f(\xi_j) \leqslant f\Big(\frac{\tau(a^*yaq_j)}{\tau(q_j)}\Big) + \varepsilon. \tag{A.4}$$

Since $\sum_{i=1}^{n} p_i = 1$, we can estimate the above right side as follows:

$$\begin{aligned}
f\Big(\frac{\tau(a^*yaq_j)}{\tau(q_j)}\Big) &= f\Big(\sum_{i=1}^{n} \frac{\tau(a^*yp_iaq_j)}{\tau(q_j)}\Big) \\
&= f\Big(\sum_{i=1}^{n} \Big(s_i \times \frac{\tau(a^*p_iaq_j)}{\tau(q_j)}\Big)\Big) \quad \text{(because of } yp_i = s_ip_i\text{)} \\
&= f\Big(0 \times \frac{\tau(q_j) - \tau(a^*aq_j)}{\tau(q_j)} + \sum_{i=1}^{n} \Big(s_i \times \frac{\tau(a^*p_iaq_j)}{\tau(q_j)}\Big)\Big) \\
&\leqslant f(0) \times \frac{\tau(q_j) - \tau(a^*aq_j)}{\tau(q_j)} + \sum_{i=1}^{n} \Big(f(s_i) \times \frac{\tau(a^*p_iaq_j)}{\tau(q_j)}\Big) \\
&= \sum_{i=1}^{n} \Big(f(s_i) \times \frac{\tau(a^*p_iaq_j)}{\tau(q_j)}\Big) \ = \ \frac{1}{\tau(q_j)} \sum_{i=1}^{n} f(s_i)\tau(a^*p_iaq_j).
\end{aligned}$$

Here, convexity of $f(t)$ and $f(0) = 0$ were used. This estimate (with (A.4)) means

$$f(\xi_j) \leqslant \frac{1}{\tau(q_j)} \sum_{i=1}^{n} f(s_i)\tau(a^*p_iaq_j) + \varepsilon,$$

and consequently

$$\begin{aligned}
\sum_{j=1}^{m} f(\xi_j)\tau(q_j) &\leqslant \sum_{j=1}^{m}\sum_{i=1}^{n} f(s_i)\tau(a^*p_iaq_j) + \varepsilon \sum_{j=1}^{m} \tau(q_j) = \sum_{i=1}^{n} f(s_i)\tau(a^*p_ia) + \varepsilon\tau(1) \\
&= \tau\Big(a^*\Big(\sum_{i=1}^{n} f(s_j)p_i\Big)a\Big) + \varepsilon\tau(1) \ = \ \tau(a^*f(y)a) + \varepsilon\tau(1).
\end{aligned}$$

From this estimate together with (A.3) and the obvious inequality

$$|\tau(a^*f(x)a) - \tau(a^*f(y)a)| \leqslant \|a^*f(x)a - a^*f(y)a\|\tau(1) \leqslant \varepsilon\tau(1)$$

(see (A.2)) we conclude

$$\tau(f(a^*xa)) \leqslant \tau(a^*f(x)a) + 3\varepsilon\tau(1),$$

and we are done.

Acknowledgements. The authors would like to thank the referee for suggesting certain improvement of the original manuscript.

REFERENCES

[1] C.A. AKEMANN, J. ANDERSON , G.K. PEDERSEN, Triangle inequalities in operator algebras, *Linear and Multilinear Algebra* **11**(1982), 167–178.

[2] T. ANDO, *Topics on Operator Inequalities*, Lecture Notes at Ryukyu Univ., Sapporo 1978.

[3] T. ANDO, Majorization, doubly stochastic matrices, and comparison of eigenvalues, *Linear Algebra Appl.* **118**(1989), 163–248.

[4] J.S. AUJLA, F.C. SILVA, Weak majorization inequalities and convex functions, *Linear Algebra Appl.* **369**(2003), 217–233.

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

[6] R. BHATIA, *Matrix Analysis*, Graduate Text in Math., vol. 169, Springer-Verlag, New York 1997.

[7] L.G. BROWN, H. KOSAKI, Jensen's inequality in semi-finite von Neumann algebras, *J. Operator Theory* **23**(1990), 3–19.

[8] M.D. CHOI, A Schwarz inequality for positive linear maps on C^*-algebras, *Illinois J. Math.* **18**(1974), 565–574.

[9] C. DAVIS, A Schwarz inequality for convex operator functions, *Proc. Amer. Math. Soc.* **8**(1957), 42–44.

[10] T. FACK, Sur la notion de valeur caractéristique, *J. Operator Theory* **7**(1982), 307–333.

[11] T. FACK, H. KOSAKI, Generalized s-numbers of τ-measurable operators, *Pacific J. Math.* **123**(1986), 269–300.

[12] I.C. GOHBERG, M.G. KREIN, *Introduction to the Theory of Linear Nonselfadjoint Operators*, Transl. Math. Monographs, vol. 18, Amer. Math. Soc., Providence, RI 1969.

[13] F. HANSEN, An operator inequality, *Math. Ann.* **246**(1980), 249–250.

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

[15] F. HANSEN, G.K. PEDERSEN, Jensen's operator inequality, *Bull. London Math. Soc.* **35**(2003), 553–564.

[16] T. HARADA, H. KOSAKI, On equality condition for trace Jensen inequality in semi-finite von Neumann algebras, *Internat. J. Math.* **19**(2008), 481–501.

[17] F. HIAI, Y. NAKAMURA, Majorization for generalized s-numbers in semi-finite von Neumann algebras, *Math. Z.* **195**(1987), 17–27.

[18] O. HIRZALLAH, F. KITTANEH, Norm inequalities for weighted power means of operators, *Linear Algebra Appl.* **341**(2002), 181–193.

[19] D.R. JOCIĆ, Norm inequalities for self-adjoint derivations, *J. Funct. Anal.* **145**(1997), 24–34.

[20] H. KOSAKI, On the continuity of the map $\varphi \to |\varphi|$ from the predual of a W^*-algebra, *J. Funct. Anal.* **59**(1984), 123–131.

[21] E. NELSON, Notes on non-commutative integration, *J. Funct. Anal.* **15**(1974), 103–116.

[22] J. PEETRE, G. SPARR, Interpolation and non-commutative integration, *Ann. Mat. Pura Appl. (4)* **104**(1975), 187–207.

[23] D. PETZ, Spectral scales of self-adjoint operators and trace inequalities, *J. Math. Anal. Appl.* **109**(1985), 74–82.

[24] D. PETZ, Jensen's inequality for positive contractions on operator algebras, *Proc. Amer. Math. Soc.* **99**(1987), 273–277.

[25] I. SEGAL, A non-commutative extension of abstract integration, *Ann. of Math.* **57**(1953), 491–457.

[26] B. SIMON, *Trace Ideals and their Applications*, London Math. Soc. Lecture Note Ser., vol. 35, Cambridge Univ. Press, Cambridge 1979.

[27] R.C. THOMPSON, Convex and concave functions of singular values of matrix sums, *Pacific J. Math.* **66**(1976), 285–290.

[28] O.E. TIKHONOV, Continuity of operator functions in topologies connected with a trace on a von Neumann algebra [Russian], *Izv, Vyssh. Uchebn. Zaved. Nat.* **1**(1987), 77–79; English *Soviet Math. (Iz. VUZ)* **31**(1987), 110–114.

[29] M. UCHIYAMA, Subadditivity of eigenvalue sums, *Proc. Amer. Math. Soc.* **134**(2006), 1405–1412.

TETSUO HARADA, 9-16-201 HAKOZAKI 1-CHOME, HIGASHI-KU, FUKUOKA 812-0053, JAPAN
E-mail address: tharada@math.kyushu-u.ac.jp

HIDEKI KOSAKI, FACULTY OF MATHEMATICS, KYUSHU UNIVERSITY, HIGASHI-KU, FUKUOKA 812-8581, JAPAN
E-mail address: kosaki@math.kyushu-u.ac.jp

Received April 12, 2007; revised June 4, 2008.