Ann. Funct. Anal. 9 (2018), no. 1, 41–55
https://doi.org/10.1215/20088752-2017-0023
ISSN: 2008-8752 (electronic)
http://projecteuclid.org/afa

A TREATMENT OF STRONGLY OPERATOR-CONVEX FUNCTIONS THAT DOES NOT REQUIRE ANY KNOWLEDGE OF OPERATOR ALGEBRAS

LAWRENCE G. BROWN

Communicated by J.-C. Bourin

Abstract. In a previous article, we proved the equivalence of six conditions on a continuous function f on an interval. These conditions determine a subset of the set of operator-convex functions whose elements are called *strongly operator-convex*. Two of the six conditions involve operator-algebraic semicontinuity theory, as given by Akemann and Pedersen, and the other four conditions do not involve operator algebras at all. Two of these conditions are operator inequalities, one is a global condition on f, and the fourth is an integral representation of f, stronger than the usual integral representation for operator-convex functions. The purpose of this article is to make the equivalence of these four conditions accessible to people who do not know operator algebra theory as well as to operator algebraists who do not know the semicontinuity theory. A treatment of other operator inequalities characterizing strong operator convexity is included.

1. Introduction

A continuous real-valued function f on an interval I is called *operator-monotone* if $h_1 \leq h_2$ and $\sigma(h_i) \subset I$ imply that $f(h_1) \leq f(h_2)$ and *operator-convex* if $\sigma(h_1), \sigma(h_2) \subset I$ and $t \in [0,1]$ imply that $f(th_1+(1-t)h_2) \leq tf(h_1)+(1-t)f(h_2)$. Here h_1 and h_2 are in $B(H)_{\mathrm{sa}}$ (the set of self-adjoint bounded operators on a Hilbert space H), $\sigma(h_i)$ denotes the spectrum, and $f(h_i)$ is defined by the continuous functional calculus. Because of the assumed continuity of f, it is sufficient

Received Dec. 7, 2016; Accepted Feb. 4, 2017.
First published online Jun. 29, 2017.
2010 *Mathematics Subject Classification.* Primary 47A63; Secondary 26A51.
Keywords. operator-convex, strongly operator-convex, operator inequality.

to verify either condition for finite-dimensional H, and operator monotonicity or convexity on the interior of I implies the same on all of I. (The reader is referred to [7] and [9] for further information on these topics and their history.) Davis showed in [6] that f is operator-convex if and only if

$$pf(php)p \le pf(h)p \quad \text{for } h \in B(H)_{\mathrm{sa}} \text{ with } \sigma(h) \subset I \text{ and } p \text{ a projection.} \tag{1}$$

And Hansen and Pedersen showed in [10] that if $0 \in I$ and $f(0) \le 0$, then f is operator-convex if and only if

$$f(a^*ha) \le a^*f(h)a \quad \text{for } h \in B(H)_{\mathrm{sa}} \text{ with } \sigma(h) \subset I \text{ and } \|a\| \le 1. \tag{2}$$

For comparative purposes, we state a theorem from [3], slightly rephrased, as well as the version that will be proved in this article.

Theorem A ([3, Theorem 2.36]). *If f is a continuous real-valued function on an interval I containing zero, then the following are equivalent.*

(i) *If h is a self-adjoint quasimultiplier of E and $\sigma(H) \subset I$, then $f(h)$ is strongly lower-semicontinuous.*
(ii) *If $p, h \in B(H)_{\mathrm{sa}}$ such that p is a projection and $\sigma(h) \subset I$, then $pf(php)p \le f(h)$.*
(iii) *If $p, h \in B(H)_{\mathrm{sa}}$ such that $0 \le p \le \mathbf{1}$ and $\sigma(h) \subset I$, then $f(php) \le f(h) + f(0)(\mathbf{1} - p)$.*
(iv) *The condition in (i) holds if E is replaced by an arbitrary C^*-algebra A.*
(v) *Either $f = 0$ or $f(x) > 0$, $\forall x \in I$, and $-1/f$ is operator-convex.*
(vi) *The function f has a representation*

$$f(x) = c + \int_{r<I} \frac{1}{x-r}\, d\mu_-(r) + \int_{r>I} \frac{1}{r-x}\, d\mu_+(r), \tag{3}$$

where $\mu_\pm$ are positive measures such that $\int \frac{1}{1+|r|}\, d\mu_\pm(r) < \infty$ and $c \ge 0$.

Theorem 3.3. *If f is a continuous real-valued function on an interval I containing zero, then the following are equivalent.*

(i) *If $h \in \mathcal{Q}$ and $\sigma(h) \subset I$, then $f(h) \in \mathcal{S}$.*
(ii) *Same as (ii) of Theorem A.*
(iii) *Same as (iii) of Theorem A.*
(iv) *Same as (v) of Theorem A.*
(v) *Same as (vi) of Theorem A.*

In the above, E is a specific C^*-algebra whose definition will be given below (although we do not really need it), $\mathbf{1}$ is the identity of $B(H)$, and $\mathcal{Q}$ and $\mathcal{S}$ will be defined below in concrete ways. Since Theorem 3.3(i) and Theorem A(i) are in fact equivalent, we will provide references along with some of our preliminary results to the corresponding semicontinuity results, for the benefit of those readers who have some interest in the semicontinuity theory. However, the concepts and proofs below will use only some basic operator theory.

The requirement that zero be in I was discussed in [3, Remark 2.37(a)] and will similarly be discussed in Remark 3.4(i) below. Both operator convexity and strong operator convexity are invariant under translation of the independent variable.

Another characterization of strong operator convexity is given in [5, Theorem 4.8]. Since this seems to depend essentially on operator-algebraic semicontinuity theory, we will not mention it further.

2. Preliminaries

We will work with operators on the Hilbert space ℓ^2. The set of compact operators on ℓ^2 is denoted by $\mathcal{K}$, and $v \times w$, for v, w vectors in ℓ^2, is the operator $u \to (u, w)v$. If $h \in B(\ell^2)_{\mathrm{sa}}$, we can write uniquely $h = h_+ - h_-$, where $h_+, h_- \geq 0$ and $h_+h_- = 0$. Let E be the set of norm-convergent sequences in $\mathcal{K}$, and denote by E^{**} the set of bounded indexed collections $\{t_n\}_{1\leq n\leq\infty}$ with each t_n in $B(\ell^2)$. Algebraic operations on E and E^{**}, including the operation $t \mapsto t^*$, are defined componentwise; and for $k = (k_n)$ in E or $t = \{t_n\}$ in E^{**}, $\|k\| = \sup\{\|k_n\| : 1 \leq n < \infty\}$ and $\|t\| = \sup\{\|t_n\| : 1 \leq n \leq \infty\}$. It is not necessary to know that E^{**} can be identified with the Banach space bidual of E, but it may be helpful to keep in mind that E^{**} is a Banach algebra. We will denote by $\mathbf{1} = \mathbf{1}_{E^{**}}$ the element $\{t_n\}$ of E^{**} such that $t_n = \mathbf{1}_{\ell^2}$ for $1 \leq n \leq \infty$, and for $h = \{h_n\}$ in E^{**}_{sa}, $\sigma(h)$ denotes the set $(\bigcup_{1\leq n\leq\infty} \sigma(h_n))^-$, where E^{**}_{sa} is the set of self-adjoint elements of E^{**}. If f is a continuous function whose domain includes $\sigma(h)$, for $h = \{h_n\}$ in E^{**}_{sa}, then $f(h)$ denotes $\{f(h_n)\}_{1\leq n\leq\infty}$. Finally, if $h' = \{h'_n\}$ and $h'' = \{h''_n\}$ are two elements of E^{**}_{sa}, then $h' \leq h''$ means $h'_n \leq h''_n$ for $1 \leq n \leq \infty$.

Lemma 2.1. *If $h \in B(\ell^2)_{\mathrm{sa}}$, then the following are equivalent.*

(i) *We have $h_- \in \mathcal{K}$.*
(ii) *There is an increasing sequence (k_n) in $\mathcal{K}$ such that $k_n \to h$ weakly (equivalently, strongly).*
(iii) *There is $k \in \mathcal{K}$ such that $h \geq k$.*

Proof. (i) $\Rightarrow$ (ii): Let (p_n) be an increasing sequence of finite-rank projections such that $p_n \to \mathbf{1}$ weakly, and take $k_n = h_+^{\frac{1}{2}} p_n h_+^{\frac{1}{2}} - h_-$.

(ii) $\Rightarrow$ (iii): Let $k = k_1$.

(iii) $\Rightarrow$ (i): Since $h_+ - h_- \geq k$, then $h_- \leq h_+ - k$. Therefore, $h_-^3 = h_-h_-h_- \leq h_-h_+h_- - h_-kh_- = -h_-kh_-$. The facts that $0 \leq h_-^3 \leq -h_-kh_-$ and $-h_-kh_- \in \mathcal{K}$ imply $h_-^3 \in \mathcal{K}$, whence $h_- \in \mathcal{K}$. □

Lemma 2.2 (cf. [3, Exercise 5.3])**.** *Let h in $B(\ell^2)_{\mathrm{sa}}$ be the weak limit of an increasing sequence (k_n) with each k_n in $\mathcal{K}$, let $k \leq h$ for $k \in \mathcal{K}$, and let $\epsilon > 0$. Then $k \leq k_n + \epsilon\mathbf{1}$ for n sufficiently large.*

Proof. If this is false, then there are unit vectors v_n such that $(kv_n, v_n) > (k_nv_n, v_n) + \epsilon$, $\forall n$. Choose a subsequence (v_{n_i}) which converges weakly to a vector v. Since $(k_nv, v) \to (hv, v) \geq (kv, v)$, it holds that $(k_nv, v) > (kv, v) - \frac{\epsilon}{3}$ for n sufficiently large. Choose one such n. Since $kv_{n_i} \to kv$ in norm and $k_nv_{n_i} \to k_nv$ in norm, then $(kv_{n_i}, v_{n_i}) \to (kv, v)$ and $(k_nv_{n_i}, v_{n_i}) \to (k_nv, v)$. Therefore, for i sufficiently large, $n_i \geq n$, $|(kv_{n_i}, v_{n_i}) - (kv, v)| < \frac{\epsilon}{3}$, and $|(k_nv_{n_i}, v_{n_i}) - (k_nv, v)| < \frac{\epsilon}{3}$. Choose one such i. Thus, $(kv, v) > (kv_{n_i}, v_{n_i}) - \frac{\epsilon}{3} > (k_{n_i}v_{n_i}, v_{n_i}) + \frac{2\epsilon}{3} \geq (k_nv_{n_i}, v_{n_i}) + \frac{2\epsilon}{3} > (k_nv, v) + \frac{\epsilon}{3}$, which is a contradiction. □

Definition 2.3. We denote by $\mathcal{S}$ the set of elements $h = \{h_n\}$ in E^{**}_{sa} such that

(i) h_n satisfies the conditions in Lemma 2.1 for $1 \le n \le \infty$, and
(ii) if $k \in \mathcal{K}$, $k \le h_\infty$, and $\epsilon > 0$, then $k \le h_n + \epsilon\mathbf{1}$ for n sufficiently large.

Note that for $\lambda \in \mathbb{R}$, $\lambda\mathbf{1} \in \mathcal{S}$ if and only if $\lambda \ge 0$.

Corollary 2.4 (cf. [3, Remark (i) after Section 5.13])**.** *Let $h = \{h_n\}$ be an element of E^{**}_{sa} which satisfies Definition 2.3(i), and let (k_m) be an increasing sequence in $\mathcal{K}$ which converges weakly to h_∞. If for each m and each $\epsilon > 0$ we have $k_m \le h_n + \epsilon\mathbf{1}$ for n sufficiently large, then $h \in \mathcal{S}$.*

Proof. Given k in $\mathcal{K}$ with $k \le h_\infty$ and $\epsilon > 0$, apply Lemma 2.2 with $\epsilon/2$ in place of ϵ. □

Lemma 2.5. *If $h = \{h_n\}$ is in E^{**}_{sa}, then the following are equivalent.*

(i) *For each vector v, $(h_\infty v, v) \le \liminf(h_n v, v)$.*
(ii) *For each weak cluster point h' of the sequence (h_n), $h_\infty \le h'$.*
(iii) *For each finite-rank projection p and each $\epsilon > 0$, $ph_\infty p \le ph_n p + \epsilon p$ for n sufficiently large.*

Proof. (i) ⇒ (ii): For each vector v, $(h'v, v)$ is a cluster point of $((h_n v, v))$. Therefore, $(h'v, v) \ge \liminf(h_n v, v) \ge (h_\infty v, v)$, whence $h' \ge h_\infty$.

(ii) ⇒ (iii): If false, there is a subsequence (h_{n_i}) such that the relation $ph_\infty p \le ph_{n_i}p + \epsilon p$ is false, $\forall i$. Passing to a further subsequence, we may assume that $h_{n_i} \to h'$ weakly for some h'. Then $ph_{n_i}p \to ph'p$ in norm. Therefore, for i sufficiently large, $ph_{n_i}p \ge ph'p - \epsilon p \ge ph_\infty p - \epsilon p$, which is a contradiction.

(iii) ⇒ (i): For a unit vector v, let p be the rank 1 projection $v \times v$. Since $ph_\infty p = (h_\infty v, v)p$ and $ph_n p = (h_n v, v)p$, the given relation implies that $\forall\epsilon > 0$, and we have $(h_\infty v, v) \le (h_n v, v) + \epsilon$ for n sufficiently large. Therefore, $(h_\infty v, v) \le \liminf(h_n v, v)$. □

Definition 2.6. Denote by $\mathcal{W}$ the set of h in E^{**}_{sa} satisfying the conditions in Lemma 2.5, and denote by $\mathcal{Q}$ the set of $h = \{h_n\}$ in E^{**}_{sa} such that $h_n \to h_\infty$ weakly. Thus, $h \in \mathcal{Q}$ if and only if $h \in \mathcal{W}$ and $-h \in \mathcal{W}$. Note that $\lambda\mathbf{1} \in \mathcal{Q} \subset \mathcal{W}, \forall\lambda \in \mathbb{R}$.

If t is in $B(\ell^2)$ and if $h = \{h_n\}$ is in E^{**}_{sa}, then t^*ht denotes the element $\{t^*h_n t\}$ of E^{**}_{sa}.

Proposition 2.7. *The sets $\mathcal{S}$ and $\mathcal{W}$ are closed in the norm topology and are closed under addition, multiplication by nonnegative scalers, and the operation $h \mapsto t^*ht$, $t \in B(\ell^2)$. Also $\mathcal{S} \subset \mathcal{W}$.*

Proof. It follows easily from Lemma 2.5(i) or 2.5(iii) that $\mathcal{W}$ is norm-closed. Suppose that $h^{(m)} \in \mathcal{S}$ for $m = 1, 2, \ldots$ and that $h^{(m)} \to h$ in the norm of E^{**}. Since the map $t \mapsto t_-$ is norm-continuous on $B(\ell^2)_{\text{sa}}$, it is clear that h_n satisfies Lemma 2.1(i) for $1 \le n \le \infty$. Now let p be a finite-rank projection, and let $k = (h_{\infty+})^{\frac{1}{2}}p(h_{\infty+})^{\frac{1}{2}} - h_{\infty-}$ and $k^{(m)} = (h^{(m)}_{\infty+})^{\frac{1}{2}}p(h^{(m)}_{\infty+})^{\frac{1}{2}} - h^{(m)}_{\infty-}$. Then $k, k^{(m)} \in \mathcal{K}$, $k \le h_\infty$, $k^{(m)} \le h^{(m)}_\infty$, and $k^{(m)} \to k$ in norm. Let $\epsilon > 0$, and choose an m such that $\|h^{(m)} - h\| < \frac{\epsilon}{3}$ and $\|k^{(m)} - k\| < \frac{\epsilon}{3}$. Then $\exists N$ such that $n > N \Rightarrow k^{(m)} \le h^{(m)}_n + \frac{\epsilon}{3}\mathbf{1}$.

Then for $n > N$, $k \le k^{(m)} + \frac{\epsilon}{3}\mathbf{1} \le h_n^{(m)} + \frac{2\epsilon}{3}\mathbf{1} \le h_n + \epsilon\mathbf{1}$. Thus, by Corollary 2.4 and the proof of (i) $\Rightarrow$ (ii) in Lemma 2.1, we conclude that $h \in \mathcal{S}$.

It is obvious that $\mathcal{S}$ and $\mathcal{W}$ are closed under multiplication by nonnegative scalars. Let $h', h'' \in \mathcal{S}$ and $h = h' + h''$. Clearly h_n satisfies Lemma 2.1(iii) for $1 \le n \le \infty$. If (k'_m) and (k''_m) are increasing sequences in $\mathcal{K}$ such that $k'_m \to h'_\infty$ and $k''_m \to h''_\infty$ weakly, then apply Corollary 2.4 with $k_m = k'_m + k''_m$ to conclude that $h \in \mathcal{S}$. The situation is similar for the operator $h \mapsto t^*ht$. It is obvious for $\mathcal{W}$ (note that $(t^*h_ntv, v) = (h_ntv, tv)$), and for $\mathcal{S}$ we use the sequence (t^*k_mt), which increases to $t^*h_\infty t$ if (k_m) increases to h_∞. If $k_m \le h_n + \epsilon\mathbf{1}$, then $t^*k_mt \le t^*h_nt + \epsilon\|t\|^2\mathbf{1}$.

Finally, let $h \in \mathcal{S}$, and choose a vector v. If $k \in \mathcal{K}$ and $k \le h_\infty$, then the fact that $\forall\epsilon > 0$, $k \le h_n + \epsilon\mathbf{1}$ for n sufficiently large implies that $(kv, v) \le \liminf(h_nv, v) + \epsilon\|v\|^2$. Since ϵ is arbitrary, this implies that $(kv, v) \le \liminf(h_nv, v)$. And since k is arbitrary and h_∞ satisfies Lemma 2.1(ii), this implies that $(h_\infty v, v) \le \liminf(h_nv, v)$. □

Proposition 2.8 (cf. [1, Proposition 3.5], which is slightly rephrased in [3, Proposition 2.1(a)])**.** *Assume that $h \in E^{**}_{\mathrm{sa}}$ and $h \ge \eta\mathbf{1}$ for some $\eta > 0$. Then $h \in \mathcal{S}$ if and only if $-h^{-1} \in \mathcal{W}$.*

Proof. By replacing h with $h_\infty^{-\frac{1}{2}}hh_\infty^{-\frac{1}{2}}$, we reduce to the case $h_\infty = \mathbf{1}$. Now assume that $h \in \mathcal{S}$, that p is a finite-rank projection, and that $\epsilon > 0$. Choose $\delta > 0$ such that $p(h_n + 2\delta\mathbf{1})^{-1}p \ge ph_n^{-1}p - \epsilon p$, $\forall n$. Then $\exists N$ such that $p \le h_n + \delta\mathbf{1}$ for $n > N$. Therefore, $p + \delta\mathbf{1} \le h_n + 2\delta\mathbf{1}$ for $n > N$, whence $(1+\delta)^{-1}p + \delta^{-1}(\mathbf{1} - p) = (p + \delta\mathbf{1})^{-1} \ge (h_n + 2\delta\mathbf{1})^{-1}$ for $n > N$. Therefore, $p \ge (1+\delta)^{-1}p \ge p(h_n + 2\delta\mathbf{1})^{-1}p \ge ph_n^{-1}p - \epsilon p$ for $n > N$. Thus, $-h^{-1}$ satisfies Lemma 2.5(iii).

Next assume that $-h^{-1} \in \mathcal{W}$, that p is a finite-rank projection, and that $\epsilon > 0$. Choose $\delta > 0$ such that $(1+2\delta)h_n \le h_n + \epsilon\mathbf{1}$, $\forall n$. Then $\exists N$ such that $ph_n^{-1}p \le p + \delta p$ for $n > N$. It follows that for some $\lambda > 0$, $h_n^{-1} \le p + 2\delta p + \lambda(\mathbf{1} - p)$ for $n > N$. To see this, it is convenient to represent elements of $B(\ell^2)_{\mathrm{sa}}$ by 2×2 matrices $\left(\begin{smallmatrix} a & b \\ b^* & c \end{smallmatrix}\right)$, where $a \in pB(\ell^2)p$, $b \in pB(\ell^2)(\mathbf{1} - p)$, and so on. If $a \ge \eta_1 p$ and $c \ge \eta_2(\mathbf{1} - p)$ for $\eta_1, \eta_2 > 0$, then this matrix is positive if and only if $\|a^{-\frac{1}{2}}bc^{-\frac{1}{2}}\| \le 1$. Now we have $h_n \ge (p + 2\delta p + \lambda(\mathbf{1} - p))^{-1} = (1 + 2\delta)^{-1}p + \lambda^{-1}(\mathbf{1} - p) \ge (1 + 2\delta)^{-1}p$ for $n > N$. Thus, $p \le h_n + \epsilon\mathbf{1}$ for $n > N$. Since there is a sequence of finite-rank projections which increases to $\mathbf{1}$ and since $h \ge 0$, it follows that $h \in \mathcal{S}$. □

Lemma 2.9. *If $h \in \mathcal{Q}$, then $h^2 \in \mathcal{W}$.*

Proof. If $v \in \ell^2$, then $h_nv \to h_\infty v$ weakly. Also $(h_n^2v, v) = \|h_nv\|^2$ and $(h_\infty^2v, v) = \|h_\infty v\|^2$. It is well known that, for vectors in a Hilbert space, $w_n \to w$ weakly implies that $\|w\| \le \liminf\|w_n\|$. □

3. Main results

One direction of the equivalence in the next theorem is needed for the proof of Theorem 3.3. We prove both directions because of the intrinsic interest.

Theorem 3.1 (cf. [3, Propositions 2.34 and 2.35(b)])**.**

(a) *If f is a continuous real-valued function on a compact interval $[a,b]$, then f is operator-convex if and only if whenever $h_n \to h$ weakly, where $h_n \in B(\ell^2)_{\mathrm{sa}}$ and $\sigma(h_n) \subset [a,b]$, $\forall n$, and $v \in \ell^2$, then $(f(h)v,v) \leq \liminf(f(h_n)v,v)$.*

(b) *Equivalently, if f is a continuous real-valued function on an interval I, then f is operator-convex if and only if $h \in \mathcal{Q}$ and $\sigma(h) \subset I$ imply $f(h) \in \mathcal{W}$.*

Remark 3.2. (i) As is well known, the strong convergence of a sequence (h_n) to h in $B(\ell^2)_{\mathrm{sa}}$ implies that $f(h_n) \to f(h)$ strongly for any continuous function f, but there is no similar implication for weak convergence. Version (a) says that operator convexity is characterized by the fact that $h_n \to h$ weakly implies "half" of what is needed to conclude that $f(h_n) \to f(h)$ weakly. In fact, if the operator-convex function f is nonlinear, it is impossible that $f(h_n) \to f(h)$ weakly unless $h_n \to h$ strongly. This follows from [3, Proposition 2.59(a)]. The original plan for this article was to include a nonoperator-algebraic proof of this, but it turns out that the result has nothing to do with operator convexity. If f is merely a continuous strictly convex function, then $h_n \to h$ strongly. This is proved in [4, Theorem 2.1] in an elementary way. (Of course, every nonlinear operator-convex function is strictly convex.)

(ii) The forward implication in version (a) can be strengthened by replacing ℓ^2 with an arbitrary (possibly nonseparable) Hilbert space and replacing the sequence (h_n) with a (necessarily bounded) net. Essentially the same proof works, or the stronger version can be deduced from the version stated.

Proof of Theorem 3.1. We prove version (b). We reduce to the case $0 \in I$ and $f(0) = 0$ by replacing f with $f(\cdot + x_0) - f(x_0)$ for some x_0 in I. This does not affect either half of the claimed equivalence.

If f is operator-convex, then f has a representation

$$f(x) = ax^2 + bx + c + \int_{r<I} \frac{(x-x_0)^2}{(x-r)(x_0-r)^2}\, d_{\mu_-}(r) + \int_{r>I} \frac{(x-x_0)^2}{(r-x)(r-x_0)^2}\, d_{\mu_+}(r), \tag{4}$$

where x_0 can be any interior point of I, $a \geq 0$, $b, c \in \mathbb{R}$, and $\mu_\pm$ are positive measures such that $\int \frac{1}{1+|r|^3}\, d_{\mu_\pm}(r) < \infty$. If I contains one or both of its endpoints, then convergence of (4) at such endpoint(s) imposes an additional condition on $\mu_\pm$. If $h \in \mathcal{Q}$ and $\sigma(h) \subset I$, then $f(h)$ is obtained by substituting h for x in (4), thus obtaining a Bochner integral. (Note that the integrands in (4) give a continuous function from $\mathbb{R}\backslash I$ to the Banach space E^{**}.) Because of the properties of $\mathcal{W}$ proved in Proposition 2.7, it is enough to show that each value of the integrand and each term $ah^2, bh, c\mathbf{1}$ is in $\mathcal{W}$. Now Proposition 2.8 implies that $(r\mathbf{1} - h)^{-1}$, for $r > I$, and $(h - r\mathbf{1})^{-1}$, for $r < I$, are in $\mathcal{S} \subset \mathcal{W}$. Also, the integrands in (4) are obtained from $1/(r-x)$ or $1/(x-r)$ by subtracting its first degree Taylor polynomial at $x = x_0$. Since the linear terms are in $\mathcal{Q} \subset \mathcal{W}$, and since Lemma 2.9 covers the ah^2 term, we conclude that $f(h) \in \mathcal{W}$.

Now assume that $h \in \mathcal{Q}$ and $\sigma(h) \subset I$ imply that $f(h) \in \mathcal{W}$. We will prove that f is operator-convex by proving (1), and we begin with a matrix version.

For natural numbers k, l, consider $(k+l)\times(k+l)$ self-adjoint matrices

$$t = \begin{pmatrix} a & b \\ b^* & c \end{pmatrix} \qquad \text{and} \qquad p = \begin{pmatrix} \mathbf{1}_k & 0 \\ 0 & 0 \end{pmatrix},$$

where a is $k\times k$, b is $k\times l$, and so on, and $\sigma(t)\subset I$. Let $f(t) = \left(\begin{smallmatrix} a' & b' \\ b'^* & c' \end{smallmatrix}\right)$. Then the desired relation, $pf(ptp)p \le pf(t)p$, amounts to $f(a)\le a'$. Let $e_1, e_2, \ldots$ be the standard orthonormal basis vectors for ℓ^2, and define $h = \{h_n\}$ by $h_n = \sum a_{ij}e_i\times e_j + \sum b_{ii'}e_i \times e_{n+k+i'} + \sum \bar{b}_{ii'}e_{n+k+i'}\times e_i + \sum c_{i'j'}e_{n+k+i'}\times e_{n+k+j'}$, for $n<\infty$, and $h_\infty = \sum a_{ij}e_i\times e_j$. Here $i,j = 1,\ldots,k$ and $i',j' = 1,\ldots,l$. Then $\sigma(h) = \sigma(t)\cup\sigma(a)\cup\{0\}\subset I$, and $h\in\mathcal{Q}$. So $f(h)\in\mathcal{W}$. If $f(h) = \{s_n\}_{1\le n\le\infty}$, then for finite n, s_n has a similar formula to h_n with a,b,c replaced by a',b',c'. Thus, (s_n) converges weakly to $\sum a'_{ij}e_i\times e_j$, and our desired relation follows from $s_\infty \le \lim s_n$.

The general case, where $t\in B(H)_{\mathrm{sa}}$ and p is a projection in $B(H)$, follows by a standard argument. Let (p_i) and (q_i) be nets of finite-rank projections such that $p_i\le p$, $q_i\le \mathbf{1}-p$, $p_i\to p$, and $q_i\to\mathbf{1}-p$, with convergence in the strong operator topology. Then $(p_i+q_i)t(p_i+q_i)\to t$ strongly and $\sigma((p_i+q_i)t(p_i+q_i))\subset I$. So it is enough to prove (1) for p_i and $(p_i+q_i)t(p_i+q_i)$, and this follows from the matrix version. □

Theorem 3.3 (cf. [3, Theorem 2.36]). *If f is a continuous real-valued function on an interval I containing zero, then the following are equivalent.*

(i) *If $h\in\mathcal{Q}$ and $\sigma(h)\subset I$, then $f(h)\in\mathcal{S}$.*
(ii) *If $p,t\in B(H)_{\mathrm{sa}}$ such that p is a projection and $\sigma(t)\subset I$, then $pf(ptp)p\le f(t)$.*
(iii) *If $p,t\in B(H)_{\mathrm{sa}}$ such that $0\le p\le\mathbf{1}$ and $\sigma(t)\subset I$, then $f(ptp)\le f(t)+f(0)(\mathbf{1}-p)$.*
(iv) *Either $f=0$ or $f(x)>0$, $\forall x\in I$, and $-1/f$ is operator-convex.*
(v) *The funcion f has a representation*

$$f(x) = c + \int_{r<I}\frac{1}{x-r}\,d_{\mu_-}(r) + \int_{r>I}\frac{1}{r-x}\,d_{\mu_+}(r), \tag{3}$$

where $\mu_\pm$ are positive measures such that $\int \frac{1}{1+|r|}\,d_{\mu_\pm}(r)<\infty$ and $c\ge 0$.

Proof. (i) ⇒ (ii): As in the proof of Theorem 3.1, we first prove a matrix version, and we use the same choices of h and the same notation as in the second part of the proof of Theorem 3.1. It is no longer true that $f(0)=0$, but since $f(x)\mathbf{1} = f(x\mathbf{1})\in\mathcal{S}$ for $x\in I$, (i) implies that $f\ge 0$ on I. Let $f(a) = a''$ and $k = \sum a''_{ij}e_i\times e_j$. Then k is a compact operator and $k\le s_\infty$, whence $\forall\epsilon>0$, $k\le s_n+\epsilon\mathbf{1}$ for n sufficiently large. For any n, the last relation amounts to the matrix inequality $\left(\begin{smallmatrix} f(a) & 0 \\ 0 & 0 \end{smallmatrix}\right) \le \left(\begin{smallmatrix} a'+\epsilon\mathbf{1}_k & b' \\ b'^* & c'+\epsilon\mathbf{1}_l \end{smallmatrix}\right)$. Since ϵ is arbitrary, we conclude that $\left(\begin{smallmatrix} f(a) & 0 \\ 0 & 0 \end{smallmatrix}\right)\le f(t)$, the matrix version of (ii). The general version follows from the matrix version, just as in the proof of Theorem 3.1.

(ii) ⇒ (i): By applying (ii) with $t = x\mathbf{1}, x\in I$, we deduce $f\ge 0$. Now let (p_m) be an increasing sequence of finite-rank projections in $B(\ell^2)$ which converges weakly (and strongly) to $\mathbf{1}$, and let $h\in\mathcal{Q}$ with $\sigma(h)\subset I$. Apply (ii) with

p_m for p and $p_{m+1}h_\infty p_{m+1}$ for t to deduce $p_m f(p_m h_\infty p_m)p_m \leq f(p_{m+1}h_\infty p_{m+1})$. Multiplying on both sides by p_{m+1}, we find that $k_m = p_m f(p_m h_\infty p_m)p_m \leq p_{m+1}f(p_{m+1}h_\infty p_{m+1})p_{m+1} = k_{m+1}$. Since $k_m \to f(h_\infty)$, it is sufficient to show that $\forall m, \forall \epsilon > 0, k_m \leq f(h_n)+\epsilon\mathbf{1}$ for n sufficiently large. (Condition (i) of Definition 2.3 follows from the fact that $f \geq 0$.) Since $h_n \to h_\infty$ weakly, $p_m h_n p_m \to p_m h_\infty p_m$ in norm, and hence $p_m f(p_m h_n p_m)p_m \to k_m$ in norm. Thus, for sufficiently large n, $k_m \leq p_m f(p_m h_n p_m)p_m + \epsilon\mathbf{1} \leq f(h_n) + \epsilon\mathbf{1}$.

(i) and (ii) $\Rightarrow$ (iii): Let $0 \leq p \leq \mathbf{1}$, and choose k in $\mathcal{K}_{\text{sa}}$ with $\sigma(k) \subset I$. Choose a sequence (p_n) of projections such that $p_n \to p$ weakly. (The possibility of this was proved by Halmos in [8].) Then define $h = \{h_n\}$ in $\mathcal{Q}$ by $h_n = p_n k p_n, n < \infty$, and $h_\infty = pkp$. Since $\sigma(h) \subset I$, (i) implies that $f(h) \in \mathcal{S} \subset \mathcal{W}$. So if t is a weak cluster point of $(f(h_n))$, then $t \geq f(pkp)$. But by (ii) and the fact that $f(p_n k p_n) = p_n f(p_n k p_n)p_n + f(0)(\mathbf{1}-p_n)$, $f(h_n) \leq f(k)+f(0)(\mathbf{1}-p_n)$, and $f(k)+f(0)(\mathbf{1}-p_n) \to f(k)+f(0)(\mathbf{1}-p)$ weakly. Thus, $f(pkp) \leq t \leq f(k)+f(0)(\mathbf{1}-p)$. As above, this is sufficient to establish the general case of (iii).

(iii) $\Rightarrow$ (ii): Apply (iii) with p a projection, and note that $f(ptp)-f(0)(\mathbf{1}-p) = pf(ptp)p$.

(i) $\Rightarrow$ (iv): We have already seen that (i) implies that $f \geq 0$. If $f \neq 0$, let J be an open subinterval of I such that $f(x) > 0, \forall x \in J$. If $h \in \mathcal{Q}$ and $\sigma(h) \subset J$, then $f(h) \in \mathcal{S}$ and Proposition 2.8 implies that $-f(h)^{-1} \in \mathcal{W}$. Thus Theorem 3.1 implies that $-1/f$ is operator-convex on J. In particular, $-1/f$ is convex, and a convex function cannot approach $-\infty$ at a finite endpoint of its interval of definition. Therefore, if either endpoint of J is in I, then f does not vanish at that endpoint.

Now let $J_0 = \{x \in I^0 : f(x) > 0\}$, where I^0 is the interior of I. Then J_0 is the disjoint union of open intervals, and the above implies that none of these intervals can have an endpoint in I^0. It follows that $J_0 = I^0$, and another application of the above shows that $f(x) > 0, \forall x \in I$. Now it is clear that $-1/f$ is operator-convex on all of I.

(iv) $\Rightarrow$ (v): We may assume that $0 \in I^0$, since neither (iv) nor (v) is affected by a translation of the independent variable. Assume that $f \neq 0$. Let $\varphi(x) = -1/x$ for $x < 0$. Since φ is both operator-monotone and operator-convex, and since $-1/f$ is operator-convex, then $f = \varphi(-1/f)$ is operator-convex. Let f be represented as in (4) with $x_0 = 0$. The analytic extension of f into the nonreal part of the complex plane is obtained by replacing x by z in (4).

Our first task is to show that $\int 1/|r|\, d\mu_\pm(r) < \infty$. Let $g(x) = (f(x) - f(0))/x$. One of the main results of Bendat and Sherman [2, Theorem 3.2] implies that g is operator-monotone, and also that $\frac{g(x)}{f(0)f(x)} = \frac{-\frac{1}{f(x)}-(-\frac{1}{f(0)})}{x}$ is operator-monotone. By Löwner's theorem, each of these operator monotone functions is either a constant or it carries the upper half-plane into itself. If g were a nonzero constant, then $-1/f$ would not be operator-convex; and if $g = 0$, then f is a positive constant, a trivial case of (3). And if $g(x)/f(0)f(x)$ is a constant, then f is the reciprocal of a linear function, another trivial case of (3).

Thus, we may assume that both g and g/f carry the upper half-plane into itself (since $f(0) > 0$). If $\operatorname{Im} z = y > 0$, then

$$\operatorname{Im} \frac{g(z)}{f(z)} = \operatorname{Im} \frac{g(z)}{f(0) + zg(z)} = \frac{f(0) \operatorname{Im} g(z) - y|g(z)|^2}{\text{positive}}.$$

So $f(0) \operatorname{Im} g(z) > y|g(z)|^2 \geq y|\operatorname{Im} g(z)|^2$, whence $\operatorname{Im} g(z) < f(0)/y$. From (4) we obtain

$$g(x) = \int_{r>I} \Big(\frac{1}{r-x} - \frac{1}{r}\Big) \frac{1}{r}\, d\mu_+(r) - \int_{r<I} \Big(\frac{1}{x-r} - \frac{1}{|r|}\Big) \frac{1}{|r|}\, d\mu_-(r) + ax + b, \quad (5)$$

and

$$\operatorname{Im} g(z) = \int_{r>I} \frac{y}{|r-z|^2} \frac{1}{r}\, d\mu_+(r) + \int_{r<I} \frac{y}{|z-r|^2} \frac{1}{|r|}\, d\mu_-(r) + ay. \quad (6)$$

This implies that $ay^2 + \int_{r>I} \frac{y^2}{|r-z|^2} \frac{1}{r}\, d\mu_+(r) + \int_{r<I} \frac{y^2}{|z-r|^2} \frac{1}{|r|}\, d\mu_-(r) < f(0)$. If $\operatorname{Re} z = 0$, then $|r-z|^2 = r^2 + y^2$, and we can let $y \to \infty$ and apply the monotone convergence theorem to the two integrals. We conclude that $a = 0$ and $\int \frac{1}{|r|}\, d\mu_\pm(r) < \infty$.

Now, as mentioned in the proof of Theorem 3.1, the integrands in (4) are $\pm(\frac{1}{r-x} - \frac{1}{r} - \frac{x}{r^2})$. We now know that each of these three terms is integrable, so we can drop the linear terms from the integrals and absorb the integrals of the linear terms into $bx + c$, obtaining

$$f(x) = \int_{r>I} \frac{1}{r-x}\, d\mu_+(r) + \int_{r<I} \frac{1}{x-r}\, d\mu_-(r) + bx + c, \quad (4')$$

where b and c no longer have the same values as in (4), and

$$g(x) = \int_{r>I} \frac{1}{r(r-x)}\, d\mu_+(r) - \int_{r<I} \frac{1}{|r|(x-r)}\, d\mu_-(r) + b, \quad (5')$$

$$\operatorname{Im} g(z) = \int_{r>I} \frac{y}{r|r-z|^2}\, d\mu_+(r) + \int_{r<I} \frac{y}{|r||z-r|^2}\, d\mu_-(r). \quad (6')$$

Then the inequality $f(0) \operatorname{Im} g(z) > y|g(z)|^2$ yields

$$f(0)\Big(\int_{r>I} \frac{1}{r|r-z|^2}\, d\mu_+(r) + \int_{r<I} \frac{1}{|r||z-r|^2}\, d\mu_-(r)\Big) > \Big|b + \int_{r>I} \frac{1}{r(r-z)}\, d\mu_+(r) - \int_{r<I} \frac{1}{|r|(z-r)}\, d\mu_-(r)\Big|^2.$$

If we let $z \to \infty$ so that $\operatorname{Im} z$ is bounded away from zero, the dominated convergence theorem applies and gives $f(0) \cdot 0 \geq |b|^2$, whence $b = 0$.

Now we calculate $\lim_{\substack{y\to\infty \\ \operatorname{Re} z = 0}} y \operatorname{Im} g(z) = \lim_{y\to\infty} \int \frac{y^2}{r^2+y^2} \frac{1}{|r|}\, d\mu(r)$, where $\mu = \mu_+ + \mu_-$. The monotone convergence theorem applies and yields $\lim y \operatorname{Im} g(z) = \int \frac{1}{|r|}\, d\mu(r)$. Since $\operatorname{Im} g(z) < f(0)/y$, we conclude that $\int \frac{1}{|r|}\, d\mu(r) \leq f(0) = c + \int \frac{1}{|r|}\, d\mu(r)$, whence $c \geq 0$.

(v) $\Rightarrow$ (i): As in the proof of Theorem 3.1, we can calculate $f(h)$, for $h \in \mathcal{Q}$ and $\sigma(h) \subset I$, by substituting h for x in (3), thus obtaining a Bochner integral. Because of the properties of $\mathcal{S}$ established in Proposition 2.7, it is enough to show that each value of the integrand is in $\mathcal{S}$ and observe that $c\mathbf{1} \in \mathcal{S}$. The first fact follows from Proposition 2.8. □

Remark 3.4. (i) The assumption that $0 \in I$ was made because condition (iii) does not make sense otherwise. But conditions (i), (iv), and (v) all make sense for arbitrary intervals I and are unaffected by translations of the independent variable. Also, condition (ii) can easily be interpreted to make sense for arbitrary I, and then it too is unaffected by translation. For example, we can extend f arbitrarily to $I \cup \{0\}$, define $f(ptp)$ by the Borel functional calculus, and observe that $pf(ptp)p$ depends only on $f_{|I}$.

Thus, for f defined on an arbitrary interval I, we can define strong operator convexity by applying any of conditions (i), (ii), (iv), or (v) directly to f, or by applying condition (iii) to $f(\cdot + x_0)$ for some x_0 in I. The last yields

(iii)$_{x_0}$ $f(ptp + x_0(\mathbf{1} - p^2)) \leq f(t) + f(x_0)(\mathbf{1} - p)$, for $p, t \in B(H)_{\mathrm{sa}}$, $0 \leq p \leq \mathbf{1}$, and $\sigma(t) \subset I$.

So what we have proved shows that (iii)$_{x_0}$ is independent of the choice of x_0 in I.

(ii) If f is real-analytic on I^0 and strongly operator-convex on some nonempty open subinterval J of I, then f (still assumed continuous and real-valued on I) is strongly operator-convex on I. This can be proved easily from condition (iv) (also easily from (v)). The assumptions and the principle of uniqueness of analytic continuation imply that there is a holomorphic function $\widetilde{f}$ on $I^0 \cup \{z \in \mathbb{C} : \operatorname{Im} z \neq 0\}$ which agrees with f on I^0. So (iv) for f, J implies that $\widetilde{f}(z) \neq 0$ for $\operatorname{Im} z > 0$ and $\frac{-\frac{1}{\widetilde{f}(z)} - (-\frac{1}{\widetilde{f}(x_0)})}{z - x_0}$ is either constant or maps the upper half-plane into itself. And this fact about $\widetilde{f}$ implies that f is strongly operator-convex on any open subinterval J of I which contains x_0 and does not contain any zeros of f. Then part of the proof above that (i) $\Rightarrow$ (iv) applies to show that $f(x) > 0$, $\forall x \in I$, whence the conclusion.

(iii) If an operator monotone or operator-convex function f on an open interval I is extended to one or both endpoints of I so that it is still monotone or convex but no longer continuous, then the operator inequalities used to define operator monotonicity or operator convexity will still hold for the extended function (using the Borel functional calculus to define $f(t_i)$). Let f be the function on $[0, \infty)$ with $f(0) = 1$ and $f(x) = 0$ for $x > 0$. Then for $t \geq 0$ in $B(H)$, $f(t)$ is the kernel projection of t. Then f satisfies conditions (ii) and (iii) of the theorem but fails conditions (i), (iv), and (v). With regard to (i), if h is a positive element of $\mathcal{Q}$, $f(h)$ need not even be in $\mathcal{W}$. The easiest way to prove (ii) and (iii) for f is to note that $f(x) = \lim_{\epsilon \to 0^+} \epsilon/(\epsilon + x)$ and the function $f_\epsilon(x) = \epsilon/(\epsilon + x)$ is strongly operator-convex on $[0, \infty)$. The dominated convergence theorem and the spectral theorem imply that $f_\epsilon(t) \to f(t)$ strongly. For the function $1 + f$, (ii), (iii), and (iv) hold, and (i) and (v) still fail.

(iv) It is easy to construct strongly operator-convex functions from operator-convex functions with the help of condition (iv). If g is operator-convex on an interval I, choose $\lambda \in \mathbb{R}$ such that $g(x) < \lambda$ for some x in I. Then let J be a subinterval of I such that $g(x) < \lambda$, $\forall x \in J$, and let $f = 1/(\lambda - g_{|J})$.

4. Additional operator inequalities

Part of the original motivation for condition (iii) in Theorem 3.3 (and [3, Theorem 2.36]) was to obtain a condition which is related to (ii) in the same way as (2) relates to (1). However, it is not at all clear that (iii) accomplishes this. Although it is obvious that (iii) $\Rightarrow$ (ii) $\Rightarrow$ (1), it is not obvious that (iii) $\Rightarrow$ (2). The only way we know to prove this is to repeat part of the proof of [10, Theorem 2.1] and deduce (2) from (1). We do not understand how (iii) fits into the general scheme of things and note that it is a somewhat peculiar looking condition. (We did make direct use of (iii) in [3] in the proof that (iii) $\Rightarrow$ (iv) in Theorem 2.36, but this was not a true application of (iii), since we could have proved (iv) just as efficiently in a different way.) And condition $(\text{iii})_{x_0}$ in Remark 3.4(i) looks even more peculiar.

In this section, we take a different approach to obtain some operator inequalities for strongly operator-convex functions which are similar to but stronger than some operator inequalities for (general) operator-convex functions. The idea is to find a way to deduce an inequality from (1), and then see what stronger inequality we get if we use Theorem 3.3(ii) instead of (1). We always exclude the case $f = 0$ in what follows.

The following inequality, due to Hansen and Pedersen though not quite explicitly stated in [10], holds if f is operator-convex on I, $a_1, \dots, a_n \in B(H)$, $t_1, \dots, t_n \in B(H)_{\text{sa}}$, $\sigma(t_i) \subset I$, and $\sum a_i^* a_i = \mathbf{1}$:

$$f\left(\sum a_i^* t_i a_i\right) \le \sum a_i^* f(t_i) a_i. \tag{7}$$

To deduce (7) from (1), consider the isometry from H into $H \oplus \cdots \oplus H$ given by the column

$$v = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

If the range of v is M, let H' be a Hilbert space of the same dimension as $M^\perp$, and find an isometry

$$w = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

from H' onto $M^\perp$. Thus, $u = (v\ w)$ is a unitary from $H \oplus H'$ to $H \oplus \cdots \oplus H$. Then (7) results from applying (1) to $t = u^*(t_1 \oplus \cdots \oplus t_n)u \in B(H \oplus H')_{\text{sa}}$ and p is the projection with range H. If f is strongly operator-convex and we instead apply Theorem 3.3(ii) with the same data, we obtain the following, where the 2×2 matrices represent elements of $B(H \oplus H')$:

$$\begin{pmatrix} f(\sum a_i^* t_i a_i) & 0 \\ 0 & 0 \end{pmatrix} \le \begin{pmatrix} \sum a_i^* f(t_i) a_i & \sum a_i^* f(t_i) b_i \\ \sum b_i^* f(t_i) a_i & \sum b_i^* f(t_i) b_i \end{pmatrix}. \tag{8}$$

Obviously (8) implies (7), and we can deduce an inequality in $B(H)$ by applying the following principle:

$$\begin{pmatrix} a & b \\ b^* & c \end{pmatrix} \geq 0 \qquad \Leftrightarrow \qquad a \geq bc^{-1}b^*, \tag{9}$$

provided that c is positive and invertible. Since $f(x) > 0, \forall x \in I$, (9) does apply in our situation, and we obtain

$$\begin{aligned} f\Big(\sum a_i^* t_i a_i\Big) \leq \sum a_i^* f(t_i) a_i \\ - \Big(\sum a_i^* f(t_i) b_i\Big)\Big(\sum b_i^* f(t_i) b_i\Big)^{-1}\Big(\sum b_i^* f(t_i) a_i\Big). \end{aligned} \tag{10}$$

Of course (8) and (10) are not fully explicit because we have not given formulas for $b_1, \dots, b_n$, but this could be remedied, at the cost of more complicated notation, as follows. The projection q with range $M^\perp$ is given by the $n \times n$ operator matrix $\mathbf{1}_n - vv^*$. Let $H' = M^\perp$, and let w be the inclusion of $M^\perp$ into $H \oplus \cdots \oplus H$. It would actually be easiest, and permissible, to replace w by q, so that u becomes a coisometry given by an $n \times (n+1)$ matrix with entries in $B(H)$.

We prefer instead to stick with (8) and (10) and will consider a couple of special cases where $b_i, \dots, b_n$ can be calculated more easily. Since strongly operator convex functions do not satisfy the condition $f(0) \leq 0$, replace (2) by:

$$f(a^* t a) \leq a^* f(t) a + f(0)(\mathbf{1} - a^* a). \tag{2$'$}$$

This is the special case of (7) where $n = 2$, $t_2 = 0$, and $a_2 = (\mathbf{1} - a_1^* a_1)^{\frac{1}{2}} = (\mathbf{1} - a^* a)^{\frac{1}{2}}$. Then, by a well-known formula, we can take $H' = H$, $b_1 = (\mathbf{1} - aa^*)^{\frac{1}{2}}$, and $b_2 = -a^*$. Thus, when f is strongly operator-convex, we get the following strengthening of (2$'$):

$$\begin{aligned} f(a^* t a) \leq a^* f(t) a + f(0)(\mathbf{1} - a^* a) \\ - a^*\big(f(t) - f(0)\mathbf{1}\big)(\mathbf{1} - aa^*)^{\frac{1}{2}} \\ \times \big((\mathbf{1} - aa^*)^{\frac{1}{2}} f(t)(\mathbf{1} - aa^*)^{\frac{1}{2}} + f(0) aa^*\big)^{-1} \\ \times (\mathbf{1} - aa^*)^{\frac{1}{2}}\big(f(t) - f(0)\mathbf{1}\big) a. \end{aligned} \tag{2$''$}$$

A better comparison with (2) can be obtained by rewriting (2$''$) in terms of $g = f - f(0)$. (Here f is still strongly operator-convex but g is not.) We have

$$\begin{aligned} g(a^* t a) \leq a^* g(t) a - a^* g(t)(\mathbf{1} - aa^*)^{\frac{1}{2}} \\ \times \big(f(0)\mathbf{1} + (\mathbf{1} - aa^*)^{\frac{1}{2}} g(t)(\mathbf{1} - aa^*)^{\frac{1}{2}}\big)^{-1}(\mathbf{1} - aa^*)^{\frac{1}{2}} g(t) a. \end{aligned} \tag{2$'''$}$$

If a is a projection, then (2$''$) implies Theorem 3.3(ii), whence (2$''$) is equivalent to strong operator convexity.

Next take $n = 2$, $a_1 = \lambda^{\frac{1}{2}}\mathbf{1}$, and $a_2 = (\mathbf{1} - \lambda)^{\frac{1}{2}}\mathbf{1}$ for $\lambda \in (0,1)$, so that (7) becomes the defining relation for operator convexity. Since $a_2 \geq 0$, as above there are simple formulas, $b_1 = (1-\lambda)^{\frac{1}{2}}\mathbf{1}$ and $b_2 = -\lambda^{\frac{1}{2}}\mathbf{1}$. Then for f strongly operator-convex, (10) becomes

$$f\big(\lambda t_1 + (1-\lambda)t_2\big) \le \lambda f(t_1) + (1-\lambda) f(t_2) \\ - \lambda(1-\lambda)\big(f(t_1) - f(t_2)\big)\big((1-\lambda)f(t_1) + \lambda f(t_2)\big)^{-1} \\ \times \big(f(t_1) - f(t_2)\big). \tag{11}$$

Note that (11) is not satisfied by arbitrary operator-convex functions, for example, by $f(x) = x^2$, even if t_1 and t_2 are scalar operators.

By using the same construction that Davis used in [6] to show that operator convexity implies (1), we can show that (11) implies Theorem 3.3(ii). So (11) is equivalent to strong operator convexity.

Next we consider the case where the subspace $M^\perp$ of $H \oplus \cdots \oplus H$ is 1-dimensional. This implies, if $H = \ell^2$, that the C^*-algebra generated by $a_1, \dots, a_n$ is an extension of $\mathcal{K}$ by the Cuntz algebra $\mathcal{O}_n$. This extension does not represent the usual generator of $\mathrm{Ext}_s(\mathcal{O}_n)$, the group of extensions of $\mathcal{K}$ by $\mathcal{O}_n$ with strong equivalence, but rather its negative. Let $v = u_1 \oplus \cdots \oplus u_n$ be a unit vector in $M^\perp$. Then for f strongly operator-convex, (10) becomes

$$f\Big(\sum a_i^* t_i a_i\Big) \le \sum a_i^* f(t_i) a_i \\ - \Big(\sum \big(f(t_i)u_i, u_i\big)\Big)^{-1}\Big(\sum a_i^* f(t_i) u_i\Big) \times \Big(\sum a_i^* f(t_i) u_i\Big). \tag{12}$$

Finally, consider the special case of Theorem 3.3(ii) where the projection p has rank 1. (This is not an additional inequality in the sense meant by the title of this section.) This amounts to the following situation. We are given a probability measure μ supported on a compact subset of I, $H = L^2(\mu)$, t is multiplication by the identity function on I, and the range of p is the set of constant functions in L^2. In this case, (1) becomes

$$f\Big(\int x\, d\mu(x)\Big) \le \int f(x)\, d\mu(x). \tag{13}$$

Of course (13) is just the classical Jensen's inequality, which is valid for arbitrary convex functions. But Theorem 3.3(ii) yields, when f is strongly operator-convex,

$$f\Big(\int x\, d\mu(x)\Big) \le 1\Big/\int f(x)^{-1}\, d\mu(x). \tag{14}$$

Also, Remark 3.4(iii) yields

$$f\Big(s^2 \int d\mu(x)\Big) \le 1\Big/\int f(x)^{-1}\, d\mu(x) + f(0)(1-s), \quad 0 \le s \le 1. \tag{15}$$

That (14) implies (13) is just the fact that the harmonic mean is less than or equal to the arithmetic mean. And (14) is not true for arbitrary operator-convex functions. We have not found any applications of (14), only new proofs of already known facts, but conceivably (14) could have an interesting consequence if applied to a particularly interesting strongly operator-convex function f. If the measure

μ is supported by a two-point set, then (14) is the same as the specialization of (11) to scalar operators. And we have found no interesting consequences of (15).

5. A differential criterion

If f is a smooth function on an open interval I, then the function $t \mapsto f(t)$ for self-adjoint $n \times n$ matrices t with $\sigma(t) \subset I$ is also smooth. In this section, we will denote this function on $\mathbb{M}_{n\mathrm{sa}}$ by F_n. (Up to now we have been casual about the notation.) The first derivative of F_n at t is a linear function from $\mathbb{M}_{n\mathrm{sa}}$ to itself, $h \mapsto F_n'(t) \cdot h$, and the second derivative is a symmetric bilinear function from $\mathbb{M}_{n\mathrm{sa}} \times \mathbb{M}_{n\mathrm{sa}}$ to $\mathbb{M}_{n\mathrm{sa}}$, $(h,k) \mapsto F_n''(t)(h,k)$. A well-known criterion for operator monotonicity is that $F_n'(t) \cdot h \geq 0$ whenever $h \geq 0$, for arbitrary n. And a well-known criterion for operator convexity is that $F_n''(t)(h,h) \geq 0$, for arbitrary n. Of course, these criteria are not complete without information on how to compute F_n' and F_n'' in terms of f, and the reader is referred to the existing literature for this. Condition (iv) of Theorem 3.3 makes it easy to derive a differential criterion for strong operator convexity. It is the following $2n \times 2n$ matrix inequality, which has to hold for arbitrary n:

$$\begin{pmatrix} F_n''(t)(h,h)/2 & F_n'(t) \cdot h \\ F_n'(t) \cdot h & F_n(t) \end{pmatrix} \geq 0. \tag{16}$$

Theorem 5.1. *If f is a continuous real-valued function on an interval I which is C^2 on I^0, then f is strongly operator-convex if and only if (16) holds for all n, for all t in $\mathbb{M}_{n\mathrm{sa}}$ with $\sigma(t) \subset I^0$, and for all h in $\mathbb{M}_{n\mathrm{sa}}$.*

Proof. If J is any open subinterval of I such that $f(x) > 0$, $\forall x \in J$, then $f_{|J}$ is strongly operator-convex if and only if $G_n''(t)(h,h) \geq 0$ for arbitrary n. Here $g = -1/f_{|J}$ and G_n relates to g as F_n to f. Computation shows that $G_n'(t) \cdot h = F_n(t)^{-1}(F_n'(t) \cdot h)F_n(t)^{-1}$, and

$$\begin{aligned} G_n''(t)(h,k) = &-F_n(t)^{-1}\big(F_n'(t) \cdot k\big)F_n(t)^{-1}\big(F_n'(t) \cdot h\big)F_n(t)^{-1} \\ &+ F_n(t)^{-1}\big(F_n''(t)(h,k)\big)F_n(t)^{-1} \\ &- F_n(t)^{-1}\big(F_n'(t) \cdot h\big)F_n(t)^{-1}\big(F_n'(t) \cdot k\big)F_n(t)^{-1}. \end{aligned}$$

Thus,

$$\begin{aligned} G_n''(t)(h,h) = &\, F_n(t)^{-1}\big(F_n''(t)(h,h)\big)F_n(t)^{-1} \\ &- 2F_n(t)^{-1}\big(F_n'(t) \cdot h\big)F_n(t)^{-1}\big(F_n'(t) \cdot h\big)F_n'(t)^{-1}. \end{aligned}$$

So $G_n''(t)(h,h) \geq 0$ if and only if $F_n''(t)(h,h) \geq 2(F_n'(t) \cdot h)F_n(t)^{-1}(F_n'(t) \cdot h)$. This is equivalent to (16), for $\sigma(t) \subset J$, by (9). $\square$

Now it is clear that if f is strongly operator-convex on I, which implies that it is real-analytic on I^0, then (16) holds. Conversely, if (16) holds, then part of the proof that (i) $\Rightarrow$ (iv) in Theorem 3.3 shows that $f(x) > 0$, $\forall x \in I$, if f is not identically zero. (Note that the lower right-hand corner of (16) implies that $f \geq 0$.) Thus, f is strongly operator-convex on I^0 and also on I.

Remark 5.2. The existing literature on operator convexity shows that it is not necessary to prove $F_n''(t)(h,h) \geq 0$ for all pairs (t,h) but only for certain well-chosen pairs. The same applies to Theorem 5.1.

References

1. C. A. Akemann and G. K. Pedersen, *Complications of semicontinuity in C^*-algebra theory*, Duke Math. J. **40** (1973), 785–795. Zbl 0287.46073. MR0358361. DOI 10.1215/S0012-7094-73-04070-2. 45
2. J. Bendat and S. Sherman, *Monotone and convex operator functions*, Trans. Amer. Math Soc. **79** (1955), 58–71. Zbl 0064.36901. MR0082655. DOI 10.2307/1992836. 48
3. L. G. Brown, *Semicontinuity and multipliers of C^*-algebras*, Canad. J. Math. **40** (1988), no. 4, 865–988. Zbl 0647.46044. MR0969204. DOI 10.4153/CJM-1988-038-5. 42, 43, 44, 45, 46, 47, 51
4. L. G. Brown, *Convergence of functions of self-adjoint operators*, Publ. Mat. **60** (2016), no. 2, 551–564. Zbl 06614146. MR3521500. DOI 10.5565/PUBLMAT_60216_09. 46
5. L. G. Brown, *Semicontinuity and closed faces of C^*-algebras*, to appear in Advances in Operator Theory, preprint, arXiv:1312.3624v2 [math.OA]. 43
6. C. Davis, *A Schwarz inequality for convex operator functions*, Proc. Amer. Math. Soc. **8** (1957), 42–44. Zbl 0080.10505. MR0084120. DOI 10.2307/2032808. 42, 53
7. C. Davis, *Notions generalizing convexity for functions defined on spaces of matrices*, Proc. Sympos. Pure Math. **7**, Amer. Math. Soc., Providence, 1963, 187–201. Zbl 0196.30303. MR0155837. 42
8. P. R. Halmos, *Normal dilations and extensions of operators*, Summa Brasil. Math. **2** (1950), 125–134. Zbl 0041.23201. MR0044036. 48
9. F. Hansen, *The fast track to Löwner's theorem*, Linear Algebra Appl. **438** (2013), no. 11, 4557–4571. Zbl 1284.26011. MR3034551. DOI 10.1016/j.laa.2013.01.022. 42
10. F. Hansen and G. K. Pedersen, *Jensen's inequality for operators and Löwner's theorem*, Math. Ann. **258** (1982), no. 3, 229–241. Zbl 0473.47011. MR1513286. DOI 10.1007/BF01450679. 42, 51

Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907, USA.

E-mail address: lgbrown@purdue.edu