AIMS Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 9(5): 10462–10477.
DOI: 10.3934/math.2024511
Received: 06 November 2023
Revised: 09 February 2024
Accepted: 23 February 2024
Published: 18 March 2024

Research article

Convexity of nonlinear mappings between bounded linear operator spaces

Messaoud Bounkhel[1,*] **and Ali Al-Tane**[2]

[1] Department of Mathematics, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia; Email: bounkhel@ksu.edu.sa

[2] Department of Basic Sciences, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia; Email: aaltane.c@ksu.edu.sa

* **Correspondence:** Email: bounkhel@ksu.edu.sa.

Abstract: Motivated by the work [7], in which the author studied the convexity of nonlinear mappings defined between bounded linear operator spaces, our research extends this inquiry. In this work, we continue the study of the convexity of nonlinear mappings defined between bounded linear operator spaces and we establish a characterization in terms of the second order directional derivative. We apply the main result to prove the convexity and the nonconvexity of well-known nonlinear mappings. The case of nondifferentiable mappings is also treated in the last section.

Keywords: nonlinear mapping; directional derivative; second order directional derivative; Bounded linear operators; nondifferentiable convex mappings
Mathematics Subject Classification: 47B47, 47A30

1. Introduction

Convexity concept for functions, mappings, and set-valued mappings is a very important tool in various applications. Its significance is most notably demonstrated in the realms of linear programming and convex optimization. Convexity guarantees the existence of global minima or maxima, streamlining the optimization process. Furthermore, it finds utility in signal processing, machine learning, and data analysis, where convex models are employed to handle and manipulate data efficiently. Extensive references to explore these applications can be found in [1–3, 6, 16, 19–21, 27], as well as their associated sources.

The study of real-valued and vector-valued functions has received considerable attention, as evidenced by references like [2, 3, 16, 20, 21]. However, the domain of matrix-valued functions, where inputs and outputs are matrices, holds a special place. Particularly, when these functions exhibit convexity concerning the Lowner partial order, they have wide applications in semi-definite programming and semi-definite complementarity problems, as detailed in [8–13, 15, 17, 22, 26, 27]. These matrix-valued functions also serve as the backbone for specialized numerical methods designed to efficiently solve the associated problems.

A fascinating area of study revolves around mappings that take values in ordered Banach spaces, as illuminated in various papers such as [3–5, 14, 15, 18, 23, 24]. However, the applicability of this general theory to convex matrix-valued functions encounters challenges. The primary constraint lies in the fact that the space of symmetric matrices, when equipped with the Lowner partial order, deviates from the typical characteristics of a vector lattice.

These deviations mean that many established techniques in convex analysis cannot be straightforwardly extended or applied to the study of convex matrix-valued functions. Consequently, alternative and specific techniques are necessitated to effectively analyze such functions.

In this work, our primary objective is to establish a characterization of convexity for matrix-valued functions and, more broadly, for nonlinear mappings defined within the Banach space $\mathcal{B}(\mathcal{H})$ of bounded linear operators. Our work unfolds as follows: Section 3 introduces the essential concepts that underpin the entire paper and is dedicated to articulating and proving our key findings, while we conclude with Section 4, which delves into the case of non-differentiable nonlinear mappings.

2. Preliminary

Let $\mathcal{H}$ be a real Hilbert space and let $\mathcal{B}(\mathcal{H})$ be the Banach space of bounded linear operators defined on $\mathcal{H}$. First, we define nonnegative and positive operators on $\mathcal{B}(\mathcal{H})$. An operator $A \in \mathcal{B}(\mathcal{H})$ is called nonnegative provided that $\langle Af, f\rangle \geq 0$ for all $f \in \mathcal{H}$, and we write $A \geq 0$. An operator A is called positive, and we write $A > 0$ if it is nonnegative and sel-fadjoint ($A = A^*$). An operator $A \in \mathcal{B}(\mathcal{H})$ is said to be nonpositive (resp., negative) if the operator $-A$ is nonnegative (resp., positive). Using this concept of nonnegativity and positivity of operators in $\mathcal{B}(\mathcal{H})$, we gather some important definitions needed in all the paper.

Definition 2.1.

(1) A subset $\mathcal{U}$ of $\mathcal{B}(\mathcal{H})$ is said to be convex if $\alpha X + (1-\alpha)Y \in \mathcal{U}$, whenever $X, Y \in \mathcal{U}$ and $\alpha \in [0, 1]$.

(2) A map $\Psi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is said to be convex (resp., strictly convex, concave, strictly concave) on a convex subset $\mathcal{U}$ of $\mathcal{B}(\mathcal{H})$, provided that $[\alpha\Psi(X) + (1-\alpha)\Psi(Y)] - \Psi(\alpha X + (1-\alpha)Y)$ is nonnegative (resp., positive, nonpositive, negative) for all $X, Y \in \mathcal{U}$ and all $\alpha \in [0, 1]$.

(3) The set of all nonnegative (resp., positive) operators in $\mathcal{B}(\mathcal{H})$ is denoted by $\mathcal{B}_+(\mathcal{H})$ (resp., $\mathcal{B}_{++}(\mathcal{H})$). We notice that $\mathcal{B}_+(\mathcal{H})$ (resp., $\mathcal{B}_{++}(\mathcal{H})$) is a closed convex cone (resp., an open convex cone) in $\mathcal{B}_{sa}(\mathcal{H})$. A set S is said to be a cone provided that $\forall\alpha > 0$, $\forall s \in S$, we have $\alpha s \in S$.

(4) The set of all self-adjoint operators in $\mathcal{B}(\mathcal{H})$ is denoted by $\mathcal{B}_{sa}(\mathcal{H})$. This set is a closed linear subspace of $\mathcal{B}(\mathcal{H})$. Obviously, we have $\mathcal{B}_{++}(\mathcal{H}) \subset \mathcal{B}_+(\mathcal{H})$ and $\mathcal{B}_{++}(\mathcal{H}) \subset \mathcal{B}_{sa}(\mathcal{H})$.

(5) The set of all nonnegative functionals in the topological dual space $\mathcal{B}^*(\mathcal{H})$ of $\mathcal{B}(\mathcal{H})$ is denoted by $\mathcal{B}^*_+(\mathcal{H})$ and defined by

$$\mathcal{B}^*_+(\mathcal{H}) := \{Z \in \mathcal{B}^*(\mathcal{H}) : Z(X) \geq 0, \quad \forall X \in \mathcal{B}_+(\mathcal{H})\}.$$

(6) The negative polar of the closed convex cone $\mathcal{B}_+^*(\mathcal{H})$ is denoted by $[\mathcal{B}_+^*(\mathcal{H})]^-$ and defined as follows:

$$[\mathcal{B}_+^*(\mathcal{H})]^- := \{X \in \mathcal{B}(\mathcal{H}) : Z(X) \leq 0, \quad \forall Z \in \mathcal{B}_+^*(\mathcal{H})\}.$$

Easily, we can check that $[\mathcal{B}_+^*(\mathcal{H})]^- = -\mathcal{B}_+(\mathcal{H})$.

Definition 2.2. Let E and F be two Banach spaces and let Ω be an open set in E. Let $f : \Omega \to F$ be a given mapping.

(1) We define the directional derivative of f at $x \in \Omega$ in the direction h, as follows:

$$\nabla f(x; h) = \lim_{t \to 0^+} \frac{f(x + th) - f(x)}{t}.$$

When f is differentiable at x, we have $\nabla f(x)(h) = f'(x; h)$, $\forall h \in E$.

(2) We define the second order directional derivative of f at x in the directions (h, k), as follows:

$$f''(x; (h, k)) = \lim_{t \to 0^+} \frac{\nabla f(x + tk)(h) - \nabla f(x)(h)}{t}.$$

When f is twice differentiable at x, we have $\nabla^2 f(x)(h)(k) = f''(x; (h, k))$, $\forall h, k \in E$.

3. Main results

The following result is well known and can be found in any calculus book. For completeness of our work, we state it without proof.

Theorem 3.1. Let S be an open convex subset of a given Banach space E and let $f : S \to \mathbb{R}$ be a twice differentiable real function on S. Then for any $x, y \in U$, $\exists z_0 \in (x, y) := \{z = ty + (1 - t)x, \text{ for some } t \in (0, 1)\}$ such that

$$f(y) = f(x) + \nabla f(x)(y - x) + \frac{1}{2} \nabla^2 f(z_0)(y - x)(y - x).$$

We use this result to prove the following characterizations of the convexity of real-valued functions defined on open convex sets in Banach spaces.

Theorem 3.2. Let E be a given Banach space and let $f : S \to \mathbb{R}$ be a twice differentiable function on an open convex subset S in E. Then the following assertions are equivalent:

1- f is convex on S;

2- $(\nabla f(x) - \nabla f(y))(x - y) \geq 0, \forall x, y \in S$;

3- $\nabla^2 f(x)(h)(h) \geq 0, \forall x \in S, \forall h \in E$.

Proof. (1) $\Rightarrow$ (2): Assume that f is convex on S. Let x and y be any two points in S. For any $\lambda \in (0, 1)$, we have

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

so

$$\frac{f(y+\lambda(x-y))-f(y)}{\lambda} \leq f(x)-f(y).$$

Taking $\lambda \to 0$ gives

$$\nabla f(y)(x-y) \leq f(x)-f(y).$$

Interchanging the roles of x and y, we get the inequality:

$$\nabla f(x)(y-x) \leq f(y)-f(x).$$

Adding these two inequalities, we get

$$(\nabla f(x)-\nabla f(y))(x-y) \geq 0, \forall x, y \in S,$$

which completes the proof of (2).

(2) $\Rightarrow$ (3): Let x be any point in S and h be any point in E. Since S is an open set, we have $x+th \in S$ for $t > 0$ small enough. Then by (2), we obtain

$$\nabla f(x+th)(th)-\nabla f(x)(th) = (\nabla f(x+th)-\nabla f(x))(th) \geq 0.$$

Dividing by t^2 and taking $t \to 0$, we get

$$\nabla^2 f(x)(h)(h) = \lim_{t\to 0} \frac{\nabla f(x+th)(h)-\nabla f(x)(h)}{t} \geq 0,$$

which proves (3).

(3) $\Rightarrow$ (1): Let x and y be any two points in S and let $\lambda \in [0,1]$. Using Theorem 3.1, we have, for some $z_0 \in (y,x) \subset S$,

$$f(y) = f(x)+\nabla f(x)(y-x)+\frac{1}{2}\nabla^2 f(z_0)(y-x)(y-x). \tag{3.1}$$

Since $z_0 \in S \subset E$, we deduce from (3),

$$\nabla^2 f(z_0)(y-x)(y-x) \geq 0, \quad \forall x, y \in S. \tag{3.2}$$

This inequality with (3.1) ensures

$$f(y)-f(x) \geq \nabla f(x)(y-x), \quad \forall x, y \in S. \tag{3.3}$$

Interchanging the roles of x and y, we obtain

$$f(x)-f(y) \geq \nabla f(y)(x-y), \quad \forall x, y \in S. \tag{3.4}$$

Adding inequalities (3.3) and (3.4) gives

$$(\nabla f(y)-\nabla f(x))(y-x) \geq 0, \quad \forall x, y \in S, \tag{3.5}$$

and so (2) is satisfied. We continue showing (1), i.e., the convexity of the function f on S. To do that, we define a scalar function $\varphi : [0,1] \to \mathbb{R}$ by $\varphi(t) := f(y + t(x-y))$, $\forall t \in [0,1]$. Obviously, we have for any $t \in (0,1)$, $\varphi'(t) = \nabla f(y + t(x-y))(x-y)$. By convexity of S, we have for any $\alpha,\beta \in (0,1)$, $y_\alpha = y + \alpha(x-y)$, and $y_\beta = y + \beta(y-x)$ will stay in S, so by (3.5), we get

$$\begin{aligned}\varphi'(\beta) - \varphi'(\alpha) &= \nabla f(y+\beta(x-y))(x-y) - \nabla f(y+\alpha(x-y))(x-y)\\ &= (\nabla f(y_\beta) - \nabla f(y_\alpha))(x-y) = \frac{(\nabla f(y_\beta) - \nabla f(y_\alpha))(y_\beta - y_\alpha)}{\beta-\alpha} \geq 0,\end{aligned}$$

which implies that φ' is increasing on $(0,1)$; and φ is convex on $(0,1)$. Hence, for any $\alpha \in (0,1)$, we obtain

$$f(\alpha x + (1-\alpha)y) = \varphi(\alpha) \leq \alpha\varphi(1) + (1-\alpha)\varphi(0) = \alpha f(x) + (1-\alpha) f(y).$$

This ensures that f is convex on S, and so the proof is complete. □

The next proposition is our main tool in the study of the convexity of nonlinear mappings defined on $\mathcal{B}(\mathcal{H})$.

Proposition 3.3. Let $\mathcal{U}$ be an open convex subset of the Banach space of bounded linear operators $\mathcal{B}(\mathcal{H})$. Let $\Psi : \mathcal{U} \to \mathcal{B}(\mathcal{H})$ be a given nonlinear mapping. Then Ψ is convex on $\mathcal{U}$ if, and only if, for any $Z \in \mathcal{B}_+^*(\mathcal{H})$, the real-valued function $\phi_Z := \langle Z, \Psi\rangle$ is convex on $\mathcal{U}$.

Proof. Assume that Ψ is convex on $\mathcal{U}$. Fix any $t \in [0,1]$ and any $X, Y \in \mathcal{U}$, we set $Z_t := tX + (1-t)Y$. By convexity of $\mathcal{U}$ and Ψ, we have $Z_t \in \mathcal{U}$ and $\Psi(Z_t) \leq t\Psi(X) + (1-t)\Psi(Y)$, that is, $t\Psi(X) + (1-t)\Psi(Y) - \Psi(Z_t) \in \mathcal{B}_+(\mathcal{H})$. Using the fact that $\mathcal{B}_+(\mathcal{H}) = -\left[\mathcal{B}_+^*(\mathcal{H})\right]^-$, we can write

$$\langle Z; t\Psi(X) + (1-t)\Psi(Y) - \Psi(Z_t)\rangle \geq 0, \quad \forall Z \in \mathcal{B}_+^*(\mathcal{H}).$$

So,

$$t\phi_Z(X) + (1-t)\phi_Z(Y) - \phi_Z(Z_t) = t\langle Z, \Psi(X)\rangle + (1-t)\langle Z; \Psi(Y)\rangle - \langle Z, \Psi(Z_t)\rangle \geq 0,$$

that is,

$$\phi_Z(Z_t) \leq t\phi_Z(X) + (1-t)\phi_Z(Y),$$

which mean that ϕ_Z is convex on $\mathcal{U}$; for any $Z \in \mathcal{B}_+^*(\mathcal{H})$.

Conversely, assume that ϕ_Z is convex on $\mathcal{U}$ for all $Z \in \mathcal{B}_+^*(\mathcal{H})$. Fix any $t \in [0,1]$ and any $X, Y \in \mathcal{U}$ and set $Z_t := tX + (1-t)Y$. Then $\phi_Z(Z_t) \leq t\phi_Z(X) + (1-t)\phi_Z(Y)$, that is,

$$\langle Z; t\Psi(X) + (1-t)\Psi(Y) - \Psi(Z_t)\rangle \geq 0, \forall Z \in \mathcal{B}_+^*(\mathcal{H}).$$

This means that

$$-[t\Psi(X) + (1-t)\Psi(Y) - \Psi(Z_t)] \in [\mathcal{B}_+^*(\mathcal{H})]^- \quad \text{(the negative polar of } \mathcal{B}_+^*(\mathcal{H})).$$

Recall that $[\mathcal{B}_+^*(\mathcal{H})]^- = -\mathcal{B}_+(\mathcal{H})$. Hence, we obtain $t\Psi(X) + (1-t)\Psi(Y) - \Psi(Z_t) \in \mathcal{B}_+(\mathcal{H})$, that is ,

$$\Psi(Z_t) \leq t\Psi(X) + (1-t)\Psi(Y), \quad \forall t \in [0,1], \forall X, Y \in \mathcal{U}.$$

This means that Ψ is convex on $\mathcal{U}$ and, hence, the proof is complete. □

Using this proposition and Theorem 3.2, we establish some characterizations of the convexity of mappings defined on open convex sets in $\mathcal{B}(\mathcal{H})$ and having images in $\mathcal{B}(\mathcal{H})$.

Theorem 3.4. Let $\mathcal{U}$ be an open convex subset in $\mathcal{B}(\mathcal{H})$ and let $\Psi : \mathcal{U} \to \mathcal{B}(\mathcal{H})$ be twice differentiable on $\mathcal{U}$. Then the following assertions are equivalent:

(1) Ψ is convex on $\mathcal{U}$;

(2) $\nabla\Psi$ is monotone on $\mathcal{U}$, i.e., $(\nabla\Psi(X) - \nabla\Psi(Y))(X - Y) \geq 0, \forall X, Y \in \mathcal{U}$;

(3) $\nabla^2\Psi(X)(V)(V) \geq 0, \forall X \in \mathcal{U}, \forall V \in \mathcal{B}(\mathcal{H})$.

Proof. The proof of the implications (1) $\Rightarrow$ (2) and (2) $\Rightarrow$ (3) follow the same lines as in the proof of Theorem 3.2. We have to check (3) $\Rightarrow$ (1). Assume that $\nabla^2\Psi(X)(V)(V)\rangle \geq 0, \forall X \in \mathcal{U}, \forall V \in \mathcal{B}(\mathcal{H})$. Fix any Z in $\mathcal{B}^*_+(\mathcal{H})$ and associate with Z and Ψ the real-valued function $\phi_Z : \mathcal{U} \to \mathbb{R}$ by $\phi_Z(X) := \langle Z; \Psi(X)\rangle$. Observe that

$$\nabla^2\phi_Z(X)(V)(V) = \langle Z; \nabla^2\Psi(X)(V)(V)\rangle, \forall X \in \mathcal{U}, \forall V \in \mathcal{B}(\mathcal{H}).$$

By assertion (3) and the definition of $\mathcal{B}^*_+(\mathcal{H})$, we get $\langle Z; \nabla^2\Psi(X)(V)(V)\rangle \geq 0$ and so $\langle\nabla^2\phi_Z(X, V); V)\rangle \geq 0, \forall X \in \mathcal{U}, \forall V \in \mathcal{B}(\mathcal{H})$. Using now Theorem 3.2 for the convex real-valued function ϕ_Z, we get $\phi_Z$is convex on $\mathcal{U}$, for all $Z \in \mathcal{B}^*_+(\mathcal{H})$. This ensures by Proposition 3.3 the convexity of Ψ on $\mathcal{U}$ and, hence, the proof is achieved. □

We use this characterization to study the convexity of some examples of nonlinear mappings defined on $\mathcal{B}(\mathcal{H})$ that we cannot prove using the definition. We start with the following two examples, in which we prove the convexity of the nonlinear mappings by both the definition and by Theorem 3.4.

Remark 3.5. An inspection of the proof of Theorem 3.4 shows that its conclusion is still valid when we replace $\mathcal{B}(\mathcal{H})$ with $\mathcal{B}_{sa}(\mathcal{H})$ or any closed subspace of $\mathcal{B}(\mathcal{H})$ containing $\mathcal{B}_{++}(\mathcal{H})$. Consequently, we can write the following characterization and omit its proof.

Proposition 3.6. Let $\mathcal{U}$ be an open convex subset in $\mathcal{B}_{sa}(\mathcal{H})$ and let $\Psi : \mathcal{U} \to \mathcal{B}(\mathcal{H})$ be twice differentiable on $\mathcal{U}$. Then Ψ is convex on $\mathcal{U}$ if, and only if, $\nabla^2\Psi(X)(V)(V) \geq 0, \forall X \in \mathcal{U}, \forall V \in \mathcal{B}_{sa}(\mathcal{H})$.

Example 3.7. Define the following two nonlinear mappings:

(1) $\Psi_1 : \mathcal{B}_{inv}(\mathcal{H}) \to \mathcal{B}_{++}(\mathcal{H})$ defined by $\Psi_1(X) = X^{-1}$, where $\mathcal{B}_{inv}(\mathcal{H}) := \{X \in \mathcal{B}_{++}(\mathcal{H}) : X^{-1}$ exists $\}$.

(2) $\Psi_2 : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ defined by $\Psi(X) = X^*AX$, where A is a nonnegative operator.

Then, Ψ_1 is strictly convex on $\mathcal{B}_{inv}(\mathcal{H})$ and Ψ_2 is convex on all the space $\mathcal{B}(\mathcal{H})$. If, in addition, A is positive, then Ψ_2 is strictly convex on $\mathcal{B}(\mathcal{H})$.

Proof. First, we prove the strict convexity of Ψ_1 on $\mathcal{B}_{inv}(\mathcal{H})$ using the definition. Let X and Y be any two operators in $\mathcal{B}_{inv}(\mathcal{H})$ and let $r, s \in [0, 1]$ with $r + s = 1$. Set $R := X^{-1} \in \mathcal{B}_{inv}(\mathcal{H})$, $S := Y^{-1} \in \mathcal{B}_{inv}(\mathcal{H})$, and $T := (rS + sR)^{-1} \in \mathcal{B}_{inv}(\mathcal{H})$. Observe that

$$STR = S(rS + sR)^{-1}R = \left(R^{-1}(rS + sR)S^{-1}\right)^{-1} = \left(rR^{-1}SS^{-1} + sR^{-1}RS^{-1}\right)^{-1} = \left(rR^{-1} + sS^{-1}\right)^{-1}$$

and by the same way, we get $RTS = \left(rR^{-1} + sS^{-1}\right)^{-1}$. Thus,

$$r\Psi_1(X) + s\Psi_1(Y) - \Psi_1(rX + sY) = sX^{-1} + sY^{-1} - (rX + sY)^{-1} = rR + sS - \left(rR^{-1} + sS^{-1}\right)^{-1}$$

$$
\begin{aligned}
&= \frac{1}{2}\Big[rR + rR + sS + sS - \big(rR^{-1} + sS^{-1}\big)^{-1} - \big(rR^{-1} + sS^{-1}\big)^{-1}\Big]\\
&= \frac{1}{2}\Big[rRTT^{-1} + rT^{-1}TR + sSTT^{-1} + sT^{-1}TS - STR - RTS\Big]\\
&= \frac{1}{2}\Big[2rs(RTR + STS) - 2rs(RTS + STR)\Big]\\
&= rs\Big[RTR + STS - RTS - STR\Big]\\
&= rs(R - S)T(R - S).
\end{aligned}
$$

Since X and Y are positive, so is T, hence, $(R-S)T(R-S)$ is positive. Indeed, for any $f \in \mathcal{H}$, we have

$$
\langle (R-S)T(R-S)f, f\rangle = \langle T(R-S)f, (R-S)^* f\rangle = \langle T(R-S)f, (R-S)f\rangle > 0.
$$

Therefore,

$$
r\Psi_1(X) + s\Psi_1(Y) - \Psi_1(rX + sY) > 0, \quad \forall r, s \in [0,1] \text{ with } r + s = 1,\ \forall X, Y \in \mathcal{B}_{inv}(\mathcal{H}).
$$

This means that Ψ_1 is strictly convex on $\mathcal{B}_{inv}(\mathcal{H})$.
The convexity of Ψ_2 using the definition has been proved in Example 3.2 in [7]. Now, we turn to the application of Theorem 3.4 to prove the convexity of Ψ_1 and Ψ_2. We need to compute the second order derivative of both mappings Ψ_1 and Ψ_2. We begin with the first directional derivative. Simple computations give for any $V \in \mathcal{B}(\mathcal{H})$:

$$
\nabla\Psi_1(X)(V) = -X^{-1}VX^{-1}, \forall X \in \mathcal{B}_{inv}(\mathcal{H}) \text{ and } \nabla\Psi_2(X)(V) = X^*AV + V^*AX, \forall X \in \mathcal{B}(\mathcal{H}).
$$

We notice that for any $V \in \mathcal{B}(\mathcal{H})$ and any $X \in \mathcal{B}_{inv}(\mathcal{H})$ we get $X + tV \in \mathcal{B}_{inv}$, for t small enough, so $\Psi_1(X + tV)$ is well-defined. We can write

$$
\begin{aligned}
\nabla^2\Psi_1(X)(V)(V) &= \lim_{t\to 0}\frac{1}{t}\Big[\nabla\Psi_1(X + tV)(V) - \nabla\Psi_1(X)(V)\Big]\\
&= -\lim_{t\to 0}\frac{1}{t}\Big[(X + tV)^{-1}V(X + tV)^{-1} - X^{-1}VX^{-1}\Big].
\end{aligned}
$$

Note that

$$
\begin{aligned}
(X + tV)^{-1}V(X + tV)^{-1} - X^{-1}VX^{-1} &= X^{-1}VX^{-1}\Big[XV^{-1}X - (X + tV)V^{-1}(X + tV)\Big](X + tV)^{-1}V(X + tV)^{-1}\\
&= X^{-1}VX^{-1}\Big[-2tX - t^2V\Big](X + tV)^{-1}V(X + tV)^{-1}\\
&= -tX^{-1}VX^{-1}\Big[2X + tV\Big](X + tV)^{-1}V(X + tV)^{-1}.
\end{aligned}
$$

Thus, for any $V \in \mathcal{B}(\mathcal{H})$ and any $X \in \mathcal{B}_{inv}(\mathcal{H})$,

$$
\begin{aligned}
\nabla^2\Psi_1(X)(V)(V) &= \lim_{t\to 0} X^{-1}VX^{-1}\Big[(2X + tV\Big](X + tV)^{-1}V(X + tV)^{-1}\\
&= X^{-1}VX^{-1}\Big[2X\Big]X^{-1}VX^{-1}\\
&= 2X^{-1}VX^{-1}VX^{-1}.
\end{aligned}
$$

For the second nonlinear mapping Ψ_2, we have for any $X, V \in \mathcal{B}(\mathcal{H})$,

$$
\nabla^2\Psi_2(X)(V)(V) = \lim_{t\to 0}\frac{1}{t}\Big[\nabla\Psi_2(X + tV)(V) - \nabla\Psi_2(X)(V)\Big]
$$

$$= \lim_{t\to 0} \frac{1}{t}\Big[(X+tV)^*AV + V^*A(X+tV) - X^*AV - V^*AX\Big]$$
$$= \lim_{t\to 0} \frac{1}{t}\Big[2tV^*AV\Big] = 2V^*AV.$$

Therefore, we obtain for any $V \in \mathcal{B}(\mathcal{H})$,

$$\nabla^2\Psi_1(X)(V)(V) = 2X^{-1}VX^{-1}VX^{-1}, \forall X \in \mathcal{B}_{inv}(\mathcal{H}) \text{ and } \nabla^2\Psi_2(X)(V)(V) = 2V^*AV, \forall X \in \mathcal{B}(\mathcal{H}).$$

Obviously, we have $\nabla^2\Psi_1(X)(V)(V) > 0$, $\forall X \in \mathcal{B}_{inv}(\mathcal{H})$ and $\forall V \in \mathcal{B}_{sa}(\mathcal{H})$, and so by Proposition 3.6, the nonlinear mapping Ψ_1 is strictly convex on $\mathcal{B}_{inv}(\mathcal{H})$. Also, we have $\nabla^2\Psi_2(X)(V)(V) > 0$ $\forall X, V \in \mathcal{B}(\mathcal{H})$, whenever A is positive. This ensures by Theorem 3.4 the strict convexity of Ψ_2 on $\mathcal{B}(\mathcal{H})$. When A is nonnegative, we get $\nabla^2\Psi_2(X)(V)(V) \geq 0$ $\forall X, V \in \mathcal{B}(\mathcal{H})$, which gives the convexity of Ψ_2 on $\mathcal{B}(\mathcal{H})$.

Now, by using our characterizations of the convexity in Theorem 3.4 and Proposition 3.6, we are going to study the convexity of some well-known nonlinear mappings on $\mathcal{B}(\mathcal{H})$. Let us start with the power mapping defined by $\Psi_n : X \mapsto \Psi_n(X) = X^n$, $n \geq 1$. To do that, we need to compute the second directional derivative of Ψ_n. First, we recall a very important result on two noncommutative binomial theorems established in Walter Wyss [28].

Theorem 3.8. Let E be an associative algebra, not necessarily commutative, with identity 1. For any X and Y in E, we have

$$(X+Y)^n = \sum_{k=0}^{n} \binom{n}{k} \{(X+\delta_Y)^k \cdot 1\} Y^{n-k}, \text{ where } \delta_Y(X) = YX - XY. \tag{3.6}$$

Set $D_n(Y, X) := (X + \delta_Y)^n \cdot 1 - X^n$. Obviously, we have $D_0(Y, X) = D_1(Y, X) = 0$. It has been shown in [28] that D_n satisfies the iteration formula for any $n \geq 1$.

$$D_{n+1}(Y, X) = \delta_Y(X^n) + (X + \delta_Y)D_n(Y, X). \tag{3.7}$$

Using this important result, we compute the first and second directional derivatives of the nonlinear mapping Ψ_n.

Proposition 3.9. Let E be an associative algebra, not necessarily commutative, with identity 1. Then

(i) $\nabla\Psi_n(X)(V) = \sum\limits_{k=0}^{n-1} X^k V X^{n-k-1}$;

(ii) $\nabla^2\Psi_n(X)(V)(V) = \sum\limits_{k=1}^{n-1}\sum\limits_{j=0}^{k-1} \Big[X^{n-1-k}VX^jVX^{k-1-j} + X^jVX^{k-1-j}VX^{n-1-k}\Big].$

Proof. (i) By definition of the first order directional derivative, we have

$$\nabla\Psi_n(X)(V) = \lim_{t\to 0} \frac{1}{t}\Big[\Psi_n(X+tV) - \Psi_n(X)\Big]$$
$$= \lim_{t\to 0} \frac{1}{t}\Big[(X+tV)^n - X^n\Big].$$

For $(X+tV)^n$, we can write by using Theorem 3.8

$$(X+tV)^n = \{(X+\delta_{tV})^n \cdot 1\} + n\{(X+\delta_{tV})^{n-1} \cdot 1\}tV + \sum_{k=0}^{n-2} \binom{n}{k} \{(X+\delta_{tV})^k \cdot 1\}(tV)^{n-k}.$$

Set $E(t) := t^{-1} \sum_{k=0}^{n-2} \binom{n}{k} \{(X+\delta_{tV})^k \cdot 1\}(tV)^{n-k}$, so we get

$$(X+tV)^n = \{(X+\delta_{tV})^n \cdot 1\} + nt\{(X+\delta_{tV})^{n-1} \cdot 1\}V + tE(t),$$

and so

$$\begin{aligned}
\frac{(X+tV)^n - X^n}{t} &= \frac{\{(X+\delta_{tV})^n \cdot 1\} - X^n}{t} + n\{(X+\delta_{tV})^{n-1} \cdot 1\}V + E(t) \\
&= \frac{D_n(tV,X)}{t} + n\{(X+t\delta_V)^{n-1} \cdot 1\}V + E(t). \qquad (3.8)
\end{aligned}$$

Taking $t \downarrow 0$ gives

$$\nabla\Psi_n(X)(V) = \lim_{t\to 0} \frac{(X+tV)^n - X^n}{t} = \lim_{t\to 0} \frac{D_n(tV,X)}{t} + nX^{n-1}V + \lim_{t\to 0} E(t).$$

First, we show that $\lim_{t\to 0} E(t) = 0$. We have, by definition,

$$\begin{aligned}
E(t) &= t^{-1} \sum_{k=0}^{n-2} \binom{n}{k} \{(X+\delta_{tV})^k \cdot 1\}(tV)^{n-k} \\
&= \sum_{k=0}^{n-2} \binom{n}{k} \{(X+\delta_{tV})^k \cdot 1\} t^{n-k-1}(V)^{n-k} \\
&= \binom{n}{0} t^{n-1}(V)^n + \binom{n}{1} \{(X+\delta_{tV}) \cdot 1\} t^{n-2}(V)^{n-1} + ... + \binom{n}{n-2} \{(X+\delta_{tV})^{n-1} \cdot 1\} t(V)^2 \\
&= t\left[t^{n-2}(V)^n + n\{(X+\delta_{tV}) \cdot 1\} t^{n-3}(V)^{n-1} + ... + \frac{n(n-1)}{2} \{(X+\delta_{tV})^{n-1} \cdot 1\}(V)^2 \right].
\end{aligned}$$

This ensures that $E(t)$ goes to zero as $t \to 0$. On the other hand, we compute the limit $\lim_{t\to 0} \frac{D_n(tV,X)}{t}$. First, we prove by induction that $D_n(tV,X) = t \sum_{k=0}^{n-1} (X+\delta_{tV})^k \delta_V(X^{n-1-k})$. Using the iteration formula (3.7), we obtain

$$\begin{aligned}
D_n(tV,X) &= t\delta_V(X^{n-1}) + (X+t\delta_V)D_{n-1}(tV,X) \\
&= t\delta_V(X^{n-1}) + (X+t\delta_V)[t\delta_V(X^{n-2}) + (X+t\delta_V)D_{n-2}(tV,X)] \\
&= t\delta_V(X^{n-1}) + t(X+t\delta_V)\delta_V(X^{n-2}) + (X+t\delta_V)^2 D_{n-2}(tV,X) \\
&= t\delta_V(X^{n-1}) + t(X+t\delta_V)\delta_V(X^{n-2}) + (X+t\delta_V)^2[t\delta_V(X^{n-3}) + (X+t\delta_V)D_{n-3}(tV,X)] \\
&= t\delta_V(X^{n-1}) + t(X+t\delta_V)\delta_V(X^{n-2}) + t(X+t\delta_V)^2\delta_V(X^{n-3}) + (X+t\delta_V)^3 D_{n-3}(tV,X) \\
&= t\delta_V(X^{n-1}) + t(X+t\delta_V)\delta_V(X^{n-2}) + ... + (X+t\delta_V)^{n-2} D_2(tV,X)
\end{aligned}$$

$$
\begin{aligned}
&= t\delta_V(X^{n-1}) + t(X + t\delta_V)\delta_V(X^{n-2}) + ... + t(X + t\delta_V)^{n-2}\delta_V(X) \\
&= t\Big[\delta_V(X^{n-1}) + (X + t\delta_V)\delta_V(X^{n-2}) + ... + (X + t\delta_V)^{n-2}\delta_V(X)\Big] \\
&= t\sum_{k=0}^{n-1}(X + t\delta_V)^k\delta_V(X^{n-1-k}).
\end{aligned}
$$

This implies

$$
\lim_{t\to 0}\frac{D_n(tV, X)}{t} = \lim_{t\to 0}\sum_{k=0}^{n-1}(X + t\delta_V)^k\delta_V(X^{n-1-k}) = \sum_{k=0}^{n-1} X^k\delta_V(X^{n-1-k}) = \sum_{k=0}^{n-1}[X^kVX^{n-1-k} - X^{n-1}].
$$

Therefore,

$$
\begin{aligned}
\nabla\Psi_n(X)(V) &= \sum_{k=0}^{n-1}[X^kVX^{n-1-k} - X^{n-1}] + nX^{n-1}V \\
&= \sum_{k=0}^{n-1}[X^kVX^{n-1-k} - X^{n-1}V] + nX^{n-1}V \\
&= \sum_{k=0}^{n-1} X^kVX^{n-1-k} - nX^{n-1}V + nX^{n-1}V \\
&= \sum_{k=0}^{n-1} X^kVX^{n-1-k}.
\end{aligned}
$$

(ii) By definition of the second order directional derivative of Ψ_n at X in the directions (V, V), we have

$$
\nabla^2\Psi_n(X)(V)(V) = \lim_{t\to 0}\frac{1}{t}\Big[\nabla\Psi_n(X + tV)(V) - \nabla\Psi_n(X)(V)\Big].
$$

First, note that

$$
\begin{aligned}
\nabla\Psi_n(X)(V) &= \sum_{k=0}^{n-1} X^kVX^{n-k-1} \\
&= VX^{n-1} + X^{n-1}V + \sum_{k=1}^{n-2} X^kVX^{n-k-1}.
\end{aligned}
$$

So,

$$
\begin{aligned}
\nabla\Psi_n(X + tV)(V) - \nabla\Psi_n(X)(V) &= V(X + tV)^{n-1} - VX^{n-1} + (X + tV)^{n-1}V - X^{n-1}V \\
&+ \sum_{k=1}^{n-2}\Big[(X + tV)^kV(X + tV)^{n-k-1} - X^kVX^{n-1-k}\Big] \\
&= V\Big[(X + tV)^{n-1} - X^{n-1}\Big] + V\Big[(X + tV)^{n-1}V - X^{n-1}V] \\
&+ \sum_{k=1}^{n-2}\Big[(X + tV)^kV(X + tV)^{n-k-1} - (X + tV)^kVX^{n-k-1}
\end{aligned}
$$

$$
\begin{aligned}
&+ \quad (X+tV)^k V X^{n-k-1} - X^k V X^{n-1-k}\Big] \\
&= \quad V\Big[(X+tV)^{n-1} - X^{n-1}\Big] + V\Big[(X+tV)^{n-1}V - X^{n-1}V\Big] \\
&+ \quad \sum_{k=1}^{n-2}(X+tV)^k V\Big[(X+tV)^{n-k-1} - X^{n-k-1}\Big] \\
&+ \quad \sum_{k=1}^{n-2}\Big[(X+tV)^k - X^k\Big] V X^{n-k-1}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\lim_{t\to 0} t^{-1}\Big[\nabla\Psi_n(X+tV)(V) - \nabla\Psi_n(X)(V)\Big] &= V\nabla\Psi_{n-1}(X)(V) + \nabla\Psi_{n-1}(X)(V)V \\
&+ \sum_{k=1}^{n-2} X^k V\nabla\Psi_{n-1-k}(X)(V) + \sum_{k=1}^{n-2}\nabla\Psi_k(X)(V)VX^{n-1-k} \\
&= V\nabla\Psi_{n-1}(X)(V) + \nabla\Psi_{n-1}(X)(V)V \\
&+ \sum_{k=1}^{n-2} X^{k-1-k} V\nabla\Psi_k(X)(V) + \sum_{k=1}^{n-2}\nabla\Psi_k(X)(V)VX^{n-1-k} \\
&= \sum_{k=1}^{n-1}\Big[X^{k-1-k}V\nabla\Psi_k(X)(V) + \nabla\Psi_k(X)(V)VX^{n-1-k}\Big] \\
&= \sum_{k=1}^{n-1}\sum_{j=0}^{k-1}\Big[X^{n-1-k}VX^jVX^{k-1-j} + X^jVX^{k-1-j}VX^{n-1-k}\Big].
\end{aligned}
$$

□

Using this result and our characterization of the convexity in Proposition 3.6, we are going to study the convexity of the following particular cases:

Case n=2. For this case, we have

$$\nabla^2\Psi_2(X)(V)(V) = 2V^2, \quad \forall V \in \mathcal{B}(\mathcal{H}).$$

Obviously, we have $\nabla^2\Psi_2(X)(V)(V) = 2V^2 \geq 0, \forall V \in \mathcal{B}_{sa}(\mathcal{H})$, which ensures the convexity of $\Psi_2(X) = X^2$ on $\mathcal{B}_{sa}(\mathcal{H})$. We notice that $2V^2$ is not always nonnegative for any $V \in \mathcal{B}(\mathcal{H})$. This fact ensures that the $\Psi_2(X) = X^2$ is not convex on all the space $\mathcal{B}(\mathcal{H})$.

Case n=3. For this case, we have

$$\nabla^2\Psi_3(X)(V)(V) = 2[V^2X + X^2V + VXV], \quad \forall V \in \mathcal{B}(\mathcal{H}).$$

Obviously, $\nabla^2\Psi_3(X)(V)(V)$ is not necessarily nonnegative, even for $V \in \mathcal{B}_{sa}(\mathcal{H})$ and $X \in \mathcal{B}_{++}(\mathcal{H})$. Take, for instance, the finite dimensional case $\mathcal{B}(\mathcal{H})_{sa} = \mathbf{S}^2$ (the space of 2×2 real symmetric matrices), $X = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} \in \mathbf{S}^2_{++}$ (the set of matrices positive definite in $\mathbf{S}^2$), and $V = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \in \mathbf{S}^2$.

For these two matrices X and V, we have $\nabla^2\Psi_3(X)(V)(V) = \begin{pmatrix} 425.0006 & 248.4804 \\ 244.4000 & 142.1600 \end{pmatrix}$, which is not nonnegative. This conclusion of nonconvexity of Ψ_3 on $\mathbf{S}^{\mathbf{2}}_{++}$ can be confirmed using the definition of convexity by taking $X = \begin{pmatrix} 4.01 & 2 \\ 2 & 1 \end{pmatrix}$ and $Y = \begin{pmatrix} 4 & 0 \\ 0 & 8 \end{pmatrix}$. Both matrices are in $\mathbf{S}^{\mathbf{2}}_{++}$, but $(\frac{1}{2}X + \frac{1}{2}Y)^3 \nleq \frac{1}{2}X^3 + \frac{1}{2}Y^3$, that is, Ψ_3 is not convex on $\mathbf{S}^{\mathbf{2}}_{++}$.

Case n=4. For this case, we have

$$\nabla^2\Psi_4(X)(V)(V) = 2[V^2X^2 + X^2V^2 + (XV)^2 + (VX)^2 + XV^2X + VX^2V], \quad \forall V \in \mathcal{B}(\mathcal{H})$$

As in the case $n = 3$, we can check for $X = \begin{pmatrix} 4.05 & 2 \\ 2 & 1 \end{pmatrix} \in \mathbf{S}^2_{++}$ and $V = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \in \mathbf{S}^2$; that $\nabla^2\Psi_4(X)(V)(V) = \begin{pmatrix} 4125.9 & 2301.3 \\ 2301.3 & 1278.4 \end{pmatrix} \ngeq 0$. This ensures by Proposition 3.6 that Ψ_4 is not convex on S^2_{++}.

According to what we showed previously about the nonlinear mapping Ψ_n ($n \geq 3$) being not convex, even on $\mathcal{B}_{++}(\mathcal{H})$, the question now is to find the convex subsets in $\mathcal{B}(\mathcal{H})$ or in $\mathcal{B}_{sa}(\mathcal{H})\}$, on which we have the convexity of Ψ_n. To do that, we define the set Ω_n in $\mathcal{B}_{sa}(\mathcal{H})$ as follows:

$$\Omega_n := \{X \in \mathcal{B}_{sa}(\mathcal{H}) : \nabla^2\Psi_n(X)(V)(V) \geq 0, \quad \forall V \in \mathcal{B}_{sa}(\mathcal{H})\}. \tag{3.9}$$

Using Proposition 3.6, we obtain the convexity of Ψ_n over any open convex subset $\mathcal{U}$ in Ω_n. So, the first question is whether Ω_n contains at least one open convex subset $\mathcal{U}$. The second question is the characterization of such open convex sets. For example, we take the case $n = 4$. We have

$$\Omega_4 := \{X \in \mathcal{B}_{sa}(\mathcal{H}) : V^2X^2 + X^2V^2 + (XV)^2 + (VX)^2 + XV^2X + VX^2V \geq 0, \forall V \in \mathcal{B}_{sa}(\mathcal{H})\}. \tag{3.10}$$

Assume that there exists at least one element in the topological interior of Ω_4, that is, there exists some $A \in \mathcal{B}_{sa}(\mathcal{H})$ such that $V^2A^2 + A^2V^2 + (AV)^2 + (VA)^2 + AV^2A + VA^2V > 0, \quad \forall V \in \mathcal{B}_{sa}(\mathcal{H})$. Thus, there exists some $\epsilon > 0$ for which $A + \epsilon\mathbb{B} \subset \Omega_4$, and then we can conclude that the mapping Ψ_4 is convex on the open convex set $\mathcal{U} := A + \epsilon\mathbb{B}$. This means that instead of looking at the open convex sets $\mathcal{U}$ on which we have the convexity of Ψ_n, we look for the solution of the following operator strict inequality: Find $X_0 \in \mathcal{B}_{sa}(\mathcal{H})$ such that

$$V^2X_0^2 + X_0^2V^2 + (X_0V)^2 + (VX_0)^2 > 0, \quad \forall V \in \mathcal{B}_{sa}(\mathcal{H}).$$

Since we always have $XV^2X \geq 0$ and $VX^2V \geq 0, \quad \forall X, V \in \mathcal{B}_{sa}(\mathcal{H})$, then any solution of the above operator strict inequality belongs to the open interior of Ω_4, and then we are done.

Remark 3.10. Using the same reasoning mentioned above, we can prove that the following nonlinear mappings are not convex over $\mathcal{B}_{++}(\mathcal{H})$: $X \mapsto X^{-n}, \quad n \geq 2$, $X \mapsto \exp(X)$. For instance, for the mapping $\Psi : X \mapsto X^{-2}$, we have that the second order directional derivative of Ψ is found to be equal to $\nabla^2\Psi(X)(V)(V) = 2[X^{-2}VX^{-1}VX^{-1} + X^{-1}VX^{-2}VX^{-1} + X^{-1}VX^{-1}VX^{-2}], \quad \forall X \in \mathcal{B}_{++}(\mathcal{H}), \forall V \in \mathcal{B}_{sa}(\mathcal{H})$. Take $X = \begin{pmatrix} 40.5 & 20 \\ 20 & 10 \end{pmatrix} \in \mathbf{S}^2_{++}$ and $V = \begin{pmatrix} 0.3 & 0.2 \\ 0.2 & 0.1 \end{pmatrix} \in \mathbf{S}^2$. Simple computations give $\nabla^2\Psi(X)(V)(V) = \begin{pmatrix} 4.800 & -9.630 \\ -9.630 & 19.322 \end{pmatrix}$, which is not nonnegative.

Using the previous results, we study of convexity of the following nonlinear mapping. For a given positive operator B and any three operators A, C, and D in $\mathcal{B}(\mathcal{H})$, we define the nonlinear mapping Ψ as follows:

$$\Psi(X) := X^2 + A^*(X+B)^{-1}A - CXD.$$

This nonlinear mapping is well-defined on the open convex cone $\mathcal{U} := \{X \in \mathcal{B}_{sa}(\mathcal{H}) : X + B > 0\}$. Obviously, this cone contains all nonnegative operators, that is, $\mathcal{B}_+(\mathcal{H}) \subset \mathcal{U}$, and it also contains some negative and nonpositive operators. Take, for instance, $X := -\alpha B \in \mathcal{U}$, (for any $\alpha \in (0,1)$). It may also contain some self-adjoint operators, which are not neither nonpositive nor nonnegative. For instance, take $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $X = \begin{pmatrix} 1 & 0 \\ 0 & -0.5 \end{pmatrix}$. Obviously, B is positive, and X is self-adjoint and is not neither nonnegative nor nonpositive. However, $X + B = \begin{pmatrix} 2 & 0 \\ 0 & 0.5 \end{pmatrix}$ is obviously positive, and so $X \in \mathcal{U}$. We are going to show the convexity of Ψ over $\mathcal{U}$. Using our previous results on the second order directional derivative, we can write:

$$\nabla^2\Psi(X)(V)(V) = 2[V^2 + A^*((X+B)^{-1}V(X+B)^{-1}V(X+B)^{-1})A],$$

for any $X \in \mathcal{U}$ and any $V \in \mathcal{B}_{sa}(\mathcal{H})$. This ensures that $\nabla^2\Psi(X)(V)(V) > 0, \forall X \in \mathcal{U}, \forall V \in \mathcal{B}_{sa}(\mathcal{H})$, and so by Proposition 3.6, the mapping Ψ is strictly convex on $\mathcal{U}$. Similarly, we get the convexity of Ψ on the closure of $\mathcal{U}$, that is, on the set $\{X \in \mathcal{B}_{sa}(\mathcal{H}) : X + B \geq 0\}$.

We close this section with the study of the convexity of the following nonlinear mapping:

$$\Psi(X) = \begin{cases} X^{-1} & \text{if } X > 0; \\ (-X)^{-1} & \text{if } X < 0. \end{cases}$$

Clearly, the mapping Ψ is well-defined on the open cone $\mathcal{K} := \{X \in \mathcal{B}_{sa}(\mathcal{H}) : X > 0 \text{ or } X < 0\}$, which is the union of the two disjoint open convex cones $\mathcal{B}_{++}(\mathcal{H})$ and $\mathcal{B}_{--}(\mathcal{H})$. Following our reasoning in Example 3.7, we obtain that Ψ is twice differentiable on $\mathcal{K}$ in all directions $V \in \mathcal{B}_{sa}(\mathcal{H})$ and

$$\nabla^2\Psi(X)(V)(V) = \begin{cases} 2X^{-1}VX^{-1}VX^{-1} & \text{if } X > 0; \\ 2(-X)^{-1}V(-X)^{-1}V(-X)^{-1} & \text{if } X < 0. \end{cases}$$

Obviously, in both cases $X \in \mathcal{B}_{++}(\mathcal{H})$ and $X \in \mathcal{B}_{--}(\mathcal{H})$, the second directional derivatives are positive for any $V \in \mathcal{B}_{sa}(\mathcal{H})$. However, the nonlinear mapping Ψ is not convex on $\mathcal{K}$. This follows from the nonconvexity of the cone $\mathcal{K}$, so the convexity of the domain of definition of the nonlinear mapping is necessary to apply our characterization in Proposition 3.6.

4. Convex nondifferentiable mappings defined between bounded linear operator spaces

In this last section, we assume that $\mathcal{U}$ is an open convex subset of $\mathcal{B}(\mathcal{H})$ and that $\Psi : \mathcal{U} \to \mathcal{B}(\mathcal{H})$ is not necessarily differentiable, and we assume that it is only directionally differentiable on $\mathcal{U}$ for any direction $V \in \mathcal{U} - \mathcal{U} := \{V = U_1 - U_2 : U_1, U_2 \in \mathcal{U}\}$, that is,

$$\Psi'(X;V) = \lim_{t \downarrow 0} t^{-1}[\Psi(X+tV) - \Psi(X)] \text{ exists }, \quad \forall X \in \mathcal{U}, \quad \forall V \in \mathcal{U} - \mathcal{U}.$$

Using the same arguments and reasoning in the proof of Theorem 3.4, we can state the following result and we omit its proof.

Proposition 4.1. Under the above assumptions on $\mathcal{U}$ and Ψ, we have Ψ is convex on $\mathcal{U}$ if, and only if, $\Psi'(X, X-Y) \geq \Psi'(Y, X-Y)$, $\forall X, Y \in \mathcal{U}$.

We use this result to prove the convexity of the following nonlinear nondifferentiable mapping: $\Psi : \mathcal{K} \to \mathcal{B}_+(\mathcal{H})$ defined by

$$\Psi(X) = \begin{cases} X & \text{if } X > 0; \\ -X & \text{if } X < 0; \\ 0 & \text{if } X = 0, \end{cases}$$

where $\mathcal{K}$ is given by

$$\mathcal{K} := \mathcal{B}_+(\mathcal{H}) \cup \mathcal{B}_-(\mathcal{H}) \cup \{0\}.$$

Let us compute the directional derivative of Ψ at any $X \in \mathcal{K}$. Let $V \in \mathcal{B}(\mathcal{H})$. At $X = 0$ and $t > 0$, we have $\Psi(X + tV) = \Psi(tV)$ is well-defined only for $V \in \mathcal{K}$, and we have

$$\Psi(X + tV) = \begin{cases} tV & \text{if } V > 0; \\ -tV & \text{if } V < 0; \\ 0 & \text{if } V = 0. \end{cases}$$

This gives $\Psi'(X; V) = \Psi(V)$, for $X = 0$ and $\forall V \in \mathcal{K}$. Assume now that $X > 0$. Then for any $V \in \mathcal{B}(\mathcal{H})$ and for $t > 0$ small enough, we get $X + tV > 0$. Thus,

$$t^{-1}[\Psi(X + tV) - \Psi(X)] = t^{-1}[(X + tV) - (X)] = V,$$

and so $\Psi'(X; V) = V$, for $X > 0$ and $\forall V \in \mathcal{B}(\mathcal{H})$. Similarly, we get, for the negative case, $\Psi'(X; V) = -V$, for $X < 0$ and $\forall V \in \mathcal{B}(\mathcal{H})$. Obviously, $\mathcal{K}$ is not convex, so we consider any open symmetric convex subset $\mathcal{U}$ in $\mathcal{K}$. Since $\mathcal{K}$ is symmetric by definition, we have $\mathcal{U} - \mathcal{U} \subset \mathcal{K}$. Hence, we obtain $\Psi'(0; V)$ exists for any $V \in \mathcal{U} - \mathcal{U}$. Thus, we conclude $\Psi'(X; V) = \Psi(V)$, for any $X \in \mathcal{U}$ and any $V \in \mathcal{U} - \mathcal{U}$. Therefore, we can use Proposition 4.1 as follows: for any $X, Y \in \mathcal{U}$, we have $\Psi'(X, X - Y) = \Psi(X - Y) = \Psi'(Y, X - Y)$. So, Ψ is convex on $\mathcal{U}$.

5. Conclusions

In conclusion, inspired by the findings presented in [7], where the author explored the convexity of nonlinear mappings within bounded linear operator spaces, our present work extends and deepens this investigation. We continue to explore the convexity of nonlinear mappings within these spaces of bounded linear operator, by establishing a characterization based on the second order directional derivative. Through the application of our main result, we substantiate the convex and nonconvex nature of well-known nonlinear mappings. Furthermore, we address the treatment of nondifferentiable mappings in the final section of our study. This work significantly contributes to advancing our understanding of the convex properties of nonlinear mappings within the framework of bounded linear operator spaces.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The first author extends his appreciations to Researchers Supporting Project number (RSPD2024R1001), King Saud University, Riyadh, Saudi Arabia.

Conflict of interest

The authors declare that they have no conflict of interests.

References

1. J. P. Aubin and H. Frankowska, *Set valued analysis*, Birkhauser, Boston, 1990.
2. H. Bauschke and P. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, Springer, 2011.
3. J. M. Borwein and Q. J. Zhu, Variational methods in convex analysis, *J. Glob. Optim.*, **35** (2006), 197–213. https://doi.org/10.1007/s10898-005-3835-3.
4. J. M. Borwein, Subgradients of convex operators, *Math. Oper. Stat. Ser.*, **15** (1984), 179–191. https://doi.org/10.1080/02331938408842921.
5. J. M. Borwein, Continuity and differentiability properties of convex operators, *London Maths. Soc.*, **s3-44** (1982), 420–444. https://doi.org/10.1112/plms/s3-44.3.420.
6. J. M. Borwein, *Convex relations in analysis and optimization*, In: S. Schaible and W. T. Ziemba,(eds.), Generalized Concavity in Optimization and Economics, Academic Press, New York, 1981.
7. M. Bounkhel, Global minimum of nonlinear mappings and orthogonality in C^1-classes, *New Zeal. J. Math.*, **36** (2007), 147–158.
8. J. Brinkhuis, Z. Q. Luo, and S. Zhang, *Matrix convex functions with applications to weighted centers for semidefinite programming*, Technical Report, El 2005-38, Econometric Institute, Erasmus University Rotterdam, 2005.
9. X. Chen, H. Qi, and P. Tseng, Analysis of nonsmooth symmetric-matrix-valued function with applications to semidefinite complementarity problems, *SIAM J. Optim.*, **13** (2003), 960–985. https://doi.org/10.1137/S1052623400380584.
10. M. V. Dolgopolik, DC semidefinite programming and cone constrained DC optimization I: theory, *Comput. Optim. Appl.*, **82** (2022), 649–671. https://doi.org/10.1007/s10589-022-00374-y.
11. M. V. Dolgopolik, DC semidefinite programming and cone constrained DC optimization II: local search methods, *Comput. Optim. Appl.*, **85** (2023), 993–1031. https://doi.org/10.1007/s10589-023-00479-y.
12. M. V. Dolgopolik, Subdifferentials of convex matrix-valued functions, *arxiv:2307.15856*, 2023.
13. F. Hansen and J. Tomiyama, Differential analysis of matrix convex functions, *Linear Algebra Appl.*, **420** (2007), 102–116. https://doi.org/10.1016/j.laa.2006.06.018.
14. N. Kirov, Generic fréchet differentiability of convex operators, *P. Am. Math. Soc.*, **94** (1985), 1.

15. A. G. Kusraev and S. S. Kutateladze, *Subdifferentials: Theory and applications*, Kluwer Academic Publishers, Dordrecht, 1995.

16. B. S. Mordukhovich, *Variational analysis and generalized differentiation I: Basic theory*, Springer-Verlag, Berling, Heidelberg, 2006.

17. K. Nordstrom, Convexity of the inverse and Moore-Penrose inverse, *Linear Algebra Appl.*, **434** (2011), 1489–1512. https://doi.org/10.1016/j.laa.2010.11.023.

18. N. S. Papageorgiou, Nonsmooth analysis on partially ordered vectors spaces: Part 1: Convex case, *Pac. J. Math.*, **107** (1983), 403–458. https://doi.org/10.2140/pjm.1983.107.403.

19. S. M. Robinson, Regularity and stability for convex multivalued functions, *Math. Oper. Res.*, **1** (1976), 130–143. https://doi.org/10.1287/moor.1.2.130.

20. R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, NJ, 1970.

21. R. T. Rockafellar, R. Wets, *Variational analysis*, Springer Verlag, Berlin, 1998.

22. D. Sun and J. Sun, Semismooth matrix-valued functions, *Math. Oper. Res.*, **27** (2002), 150–169. https://doi.org/10.1287/moor.27.1.150.342.

23. M. Théra. Subdifferential calculus for convex operators, *J. Math. Anal. Appl.*, **80** (1981), 78–91.

24. M. Valadier, Sous-differentiabilité des fonctions convexes à valeurs dans un espace vectoriel ordonné, *Math. Scand.*, **30** (1972), 65–74.

25. L. Vesely and L. Zajicek, On differentiability of convex operators, *J. Math. Anal. Appl.*, **402** (2013), 12–22. https://doi.org/10.1016/j.jmaa.2012.12.073.

26. L. Vesely and L. Zajicek, *Delta-convex mappings between Banach spaces and applications*, Warszawa: Instytut Matematyczny Polskiej Akademi Nauk, 1989.

27. W. Watkins, Convex matrix functions, *Proc. Amer. Math. Soc.*, **44** (1974), 31–34. https://doi.org/10.1090 S0002-9939-1974-0340291-3.

28. W. Wyss, Two non-commutative binomial theorems, *arXiv:1707.03861*, 2017.

AIMS Press