

THE NON-COMMUTATIVE LEGENDRE-FENCHEL TRANSFORM

SØREN WINKLER

Abstract

We use the theory of matrix convex sets of Effros-Winkler to introduce a non-commutative version of convex functions, and we show how the Legendre-Fenchel transform generalizes to this situation based on the analogue of the classical theorem that any closed convex function is the supremum over all affine functions dominated by the convex function.

The classical Legendre transform implements the idea that one may view a curve either as a locus of points or as an envelope of tangents. The close connection with convexity was realized by Fenchel ([3]), and it turns out that this duality is a consequence of the fact that a closed convex function is the supremum over all affine functions dominated by the convex function. This is merely restating that a closed convex set is the intersection of all half-spaces containing it, which again is an immediate consequence of the Hahn-Banach theorem.

In [2] E. G. Effros and the author laid the foundations for a non-commutative convexity theory, denoted matrix convexity, building on a definition of Wittstock ([11]). This included several matricial analogues of the various forms of the Hahn-Banach theorem, including a separation-type result corresponding to the above-mentioned. It is therefore only natural to ask whether the theory of convex functions and the Legendre-Fenchel transform generalizes to the non-commutative setting, and in this paper we give an affirmative answer to this question.

We begin by reviewing the theory of matrix convex sets. We proceed by defining the non-commutative analogue of convex functions and discuss several examples, in particular the relation with the so-called operator convex functions (cf. [5]). The next step is to prove the non-commutative version of the fact that a closed convex function is the supremum over all affine functions dominated by the convex function, and this is then used to define

The author was supported by the EPSRC and the EU.
Received January 20, 1997.

the matricial Legendre-Fenchel transform. As an application we devote the final section to a generalization of the close relation between the Legendre-Fenchel transform and the Cauchy-Schwartz inequality.

The point of view of this paper is that of convexity even though the classical Legendre transform is by means of differentiability. The question of generalizations from the point of view of differentiability will be considered in a subsequent paper.

We wish to thank E. G. Effros for many stimulating discussions and helpful remarks.

1. Matrix convexity

All vector spaces in this paper will be assumed to be complex. Let $M_{m,n}(V)$ be the vector space of $m \times n$ matrices over a vector space V and set $M_n(V) = M_{n,n}(V)$. We denote $\mathbf{M}_{m,n} = M_{m,n}(\mathbf{C})$ and $\mathbf{M}_n = M_{n,n}(\mathbf{C})$, which means that we may identify $M_{m,n}(V)$ with the tensor product $\mathbf{M}_{m,n} \otimes V$. We use the standard matrix multiplication and $*$ -operation for compatible matrices, and \mathbf{l}_n for the identity matrix in \mathbf{M}_n .

There are natural matrix operations on the matrix spaces: For $v \in M_m(V)$, $w \in M_n(V)$, and $\alpha \in \mathbf{M}_{n,m}$, $\beta \in \mathbf{M}_{m,n}$, we define

$$\begin{aligned}\alpha v \beta &= [\sum_{j,k} \alpha_{ij} v_{jk} \beta_{kl}] \in M_n(V); \\ v \oplus w &= \begin{bmatrix} v & 0 \\ 0 & w \end{bmatrix} \in M_{m+n}(V).\end{aligned}$$

DEFINITION 1.1. A *matrix convex set* in a vector space V is a collection $\mathbf{K} = (K_n)$ with $K_n \subset M_n(V)$ satisfying

$$\begin{aligned}v \oplus w \in K_{m+n}, & \quad v \in K_m, w \in K_n; \\ \gamma^* v \gamma \in K_n, & \quad v \in K_m, \gamma \in \mathbf{M}_{m,n}, \gamma^* \gamma = \mathbf{l}_n.\end{aligned}$$

Equivalently,

$$\sum_{i=1}^k \gamma_i^* v_i \gamma_i \in K_n$$

whenever $v_i \in K_{n_i}$ and $\gamma_i \in \mathbf{M}_{n_i,n}$ for $i = 1, \dots, k$ satisfies $\sum_i \gamma_i^* \gamma_i = \mathbf{l}_n$. This form bears stronger resemblance to the classical definition of convexity, and it was in this form it appeared in [11]. It is only a matter of taste which form to use but we have chosen to retain the flavor of [2].

EXAMPLE 1.2. For some standard examples of matrix convex sets consider an operator space V , i.e., a linear subspace of $\mathcal{B}(\mathcal{H})$, the bounded operators on a Hilbert space \mathcal{H} . The natural inclusion

$$M_n(V) \hookrightarrow M_n(\mathcal{B}(\mathcal{H})) \simeq \mathcal{B}(\mathcal{H}^n)$$

endows $M_n(V)$ with a norm using the operator norm on $\mathcal{B}(\mathcal{H}^n)$, and it is easy to check that the collection of unit balls

$$B_n = \{v \in M_n(V) \mid \|v\| \leq 1\}$$

is a matrix convex set. If V is an operator system, i.e., a self-adjoint subspace of $\mathcal{B}(\mathcal{H})$ containing the identity operator I , then the above inclusion defines an ordering on $M_n(V)$ via the usual ordering on $\mathcal{B}(\mathcal{H}^n)$. In this case the collection of positive cones

$$P_n = \{v \in M_n(V) \mid v \geq 0\}$$

forms a matrix convex set. We may also consider the collection of *matrix states*

$$\mathcal{CS}_n = \{\varphi : V \rightarrow \mathbb{M}_n \mid \varphi \text{ completely positive, } \varphi(I) = I_n\},$$

and again we get a matrix convex set (in V^*).

One may think of the above three examples as non-commutative analogues of balanced convex sets, positive cones, and compact convex sets, respectively.

For a detailed account of matrix convexity we refer to [2] or [9]. In [2] we proved the following analogue of the separation-type Hahn-Banach theorem.

THEOREM 1.3. *Let V be a locally convex vector space. Assume that $K = (K_r)$ is a matrix convex set with $0 \in K_1$, and such that K_r is closed in the product topology in $M_r(V)$ for all $r \in \mathbb{N}$. Given $v_0 \notin K_n$ for some $n \in \mathbb{N}$, there exists a continuous linear mapping $\Phi : V \rightarrow \mathbb{M}_n$ such that*

$$\operatorname{Re} \Phi_r(v) \leq I_r \otimes I_n$$

for all $r \in \mathbb{N}$, $v \in K_r$, and

$$\operatorname{Re} \Phi_n(v_0) \not\leq I_n \otimes I_n.$$

In the above theorem Φ_r denotes the usual amplification of Φ to $M_r(V)$, i.e.,

$$\Phi_r = \operatorname{id} \otimes \Phi : M_r(V) \simeq M_r \otimes V \rightarrow M_r(M_n) \simeq M_r \otimes M_n,$$

and – as throughout this paper – we use the usual ordering on M_n determined by the positive semi-definite matrices, denoted by M_n^+ . We also remark that for $\alpha \in M_n$, $\operatorname{Re} \alpha$ denotes the self-adjoint part of α , i.e.,

$$\operatorname{Re} \alpha = \frac{1}{2}(\alpha + \alpha^*) \in (M_n)_{\text{sa}}.$$

Given dual vector spaces V and V' with the pairing written as $\langle v, \varphi \rangle \in \mathbf{C}$ for $v \in V$ and $\varphi \in V'$, we also have a *matrix-valued* pairing

$$(1) \quad M_m(V) \times M_n(V') \rightarrow M_m \otimes M_n : (v, \varphi) \mapsto \langle\langle v, \varphi \rangle\rangle = [\langle v_{ij}, \varphi_{kl} \rangle].$$

If we equip V and V' with the corresponding weak topologies and we identify $\varphi \in M_n(V')$ with a weakly continuous linear mapping $\Phi : V \rightarrow M_n$ then

$$\langle\langle v, \varphi \rangle\rangle = \Phi_m(v).$$

Theorem 1.3 may be rewritten accordingly. As a general rule this matrix pairing substitutes the scalar pairing in the transition from the commutative to the non-commutative case.

2. Matrix convex functions

In classical convexity theory it is common to study a convex function $F : V \rightarrow \mathbf{R}$ on a vector space V in terms of its *super-graph*

$$\mathcal{G}(F) = \{(v, \alpha) \in V \times \mathbf{R} \mid F(v) \leq \alpha\}.$$

For instance, it is well-known that F is convex if and only if $\mathcal{G}(F)$ is convex, and that F is lower semi-continuous if and only if $\mathcal{G}(F)$ is closed. Moreover, one may recover F from $\mathcal{G}(F)$ since

$$(2) \quad F(v) = \inf\{\alpha \in \mathbf{R} \mid (v, \alpha) \in \mathcal{G}(F)\}$$

for $v \in V$. Having built a theory of matrix convex sets as described in the previous section, it is therefore only natural to seek to define a matrix convex function by matrix convexity of an appropriate super-graph.

In trying to generalize (2) to the matricial situation we are faced with the lack of suitable infimum and supremum in M_n . Following Wittstock ([10]), we resolve this by working with *set-valued* maps, which in the scalar case above corresponds to identifying F with the mapping

$$v \mapsto [F(v), \infty[$$

for $v \in V$.

As range of our mappings we shall use certain subsets of M_n . Apart from the usual set-theoretic operations, such as $S \cap T$ for subsets S and T of M_n , we also have subsets

$$\begin{aligned} S + T &= \{\alpha + \beta \mid \alpha \in S, \beta \in T\}; \\ \lambda S &= \{\lambda \alpha \mid \alpha \in S\} \end{aligned}$$

with $\lambda \in \mathbf{C}$. Moreover, with $S \subset M_m$ and $T \subset M_n$, and $\gamma, \delta \in M_{m,n}$, we can also construct new subsets

$$\begin{aligned} S \oplus T &= \{\alpha \oplus \beta \mid \alpha \in S, \beta \in T\} \subset \mathbf{M}_{m+n}; \\ \gamma^* S \delta &= \{\gamma^* \alpha \delta \mid \alpha \in S\} \subset \mathbf{M}_n. \end{aligned}$$

We use the conventions $S + \emptyset = \emptyset$, $\lambda \emptyset = \emptyset$, $S \oplus \emptyset = \emptyset$, and $\gamma^* \emptyset \delta = \emptyset$.

We say that a subset $S \subset (\mathbf{M}_n)_{\text{sa}}$ is a *sector* if for each $\alpha \in S$ we have that $\alpha \leq \beta$ implies that $\beta \in S$, or equivalently, if

$$S = S + \mathbf{M}_n^+.$$

We let \mathbf{S}_n be the collection of all *closed* sectors in \mathbf{M}_n . The sectors will play the rôle of the scalar intervals $[\alpha, \infty[$ in the non-commutative case. The restriction to closed sectors is not a serious one as the closure of a sector is again a sector. The proof of this fact, due to E. G. Effros, is outlined below.

LEMMA 2.1. *If $S \subset (\mathbf{M}_n)_{\text{sa}}$ is a sector, then its closure \bar{S} is again a sector.*

PROOF. Let $\alpha, \beta \in (\mathbf{M}_n)_{\text{sa}}$ be given such that $\beta \geq \alpha$ and $\alpha \in \bar{S}$. First assume that α and thus β is positive and invertible. Then there exists an invertible contraction δ with $\alpha^{1/2} = \delta\beta^{1/2}$. Letting α_r be a sequence in S which converges to α , we may assume that α_r is positive and invertible. It follows that $\gamma_r = \delta^{-1}\alpha_r^{1/2}$ converges to $\beta^{1/2}$, and thus $\gamma_r^*\gamma_r$ converges to β . But $\alpha_r^{1/2} = \delta\gamma_r$, and thus

$$\alpha_r = \gamma_r^* \delta^* \delta \gamma_r \leq \gamma_r^* \gamma_r$$

from which it follows that $\gamma_r^* \gamma_r \in S$, and therefore $\beta \in \bar{S}$.

In the general case we substitute sector $S + (\|\alpha\| + 1)\mathbf{I}_n$ for S , $\alpha + (\|\alpha\| + 1)\mathbf{I}_n$ for α , and $\beta + (\|\alpha\| + 1)\mathbf{I}_n$ for β . By the above $\beta + (\|\alpha\| + 1)\mathbf{I}_n \in \bar{S} + (\|\alpha\| + 1)\mathbf{I}_n$, and therefore $\beta \in \bar{S}$.

For an arbitrary subset $S \subset (\mathbf{M}_n)_{\text{sa}}$ we have that $\overline{S + \mathbf{M}_n^+}$ is the smallest closed sector containing S . In particular, each $\alpha \in (\mathbf{M}_n)_{\text{sa}}$ determines the sector

$$[\alpha, \infty[= \overline{\{\alpha\} + \mathbf{M}_n^+} = \{\beta \in \mathbf{M}_n \mid \beta \geq \alpha\}.$$

We also regard the empty set \emptyset and $(\mathbf{M}_n)_{\text{sa}}$ as sectors which we denote by $+\infty$ and $-\infty$, respectively.

We define a partial ordering “ \preceq ” on the subsets of \mathbf{M}_n by

$$S \preceq T \Leftrightarrow T \subset S$$

for subsets S and T . This is motivated by the fact that for $\alpha, \beta \in (\mathbf{M}_n)_{\text{sa}}$, we have $\alpha \leq \beta$ if and only if $[\alpha, \infty[\preceq [\beta, \infty[$. We also write $S \prec +\infty$ if $S \neq \emptyset$, and $-\infty \prec S$ whenever $S \neq (\mathbf{M}_n)_{\text{sa}}$.

We see now that any function $F : V \rightarrow \mathbf{R}$ on a set V can be written as the sector-valued mapping

$$v \mapsto [F(v), \infty[$$

for $v \in V$.

DEFINITION 2.2. We therefore define a *matrix function* \mathbf{F} on a set V to be a collection $\mathbf{F} = (F_n)$ of sector-valued mappings

$$F_n : M_n(V) \rightarrow \mathbf{S}_n,$$

such that for all $n \in \mathbf{N}$ we have

$$(3) \quad -\infty \prec F_n(v)$$

for all $v \in M_n(V)$, and

$$(4) \quad F_n(v) \prec +\infty$$

for some $v \in M_n(V)$.

We do not exclude the value $+\infty (= \emptyset)$ in the range so we define the *domain* $\mathcal{D}(\mathbf{F})$ of a \mathbf{F} to be the collection of domains $\mathcal{D}(F_n) \subset M_n(V)$ given by

$$\mathcal{D}(F_n) = \{v \in M_n(V) \mid F_n(v) \prec +\infty\},$$

and by (4), $\mathcal{D}(F_n) \neq \emptyset$ for all $n \in \mathbf{N}$. Furthermore, we let the *super-graph* $\mathcal{G}(\mathbf{F})$ of \mathbf{F} be the collection of non-empty super-graphs $\mathcal{G}(F_n) \subset M_n(V) \times \mathbf{M}_n = M_n(V \times \mathbf{C})$ given by

$$\mathcal{G}(F_n) = \{(v, \alpha) \in M_n(V) \times (\mathbf{M}_n)_{\text{sa}} \mid \alpha \in F_n(v)\}.$$

We also say that \mathbf{F} is *single-valued* if F_n is single-valued for all $n \in \mathbf{N}$, i.e., if for all $v \in \mathcal{D}(F_n)$, $F_n(v) = [\alpha, \infty[$ for some $\alpha \in (\mathbf{M}_n)_{\text{sa}}$.

Working with $+\infty$ of course calls for special care – just as in the case of \mathbf{R} – but in practice this is rarely a problem.

With the above notions at hand we may finally define our matricial analogue of a convex function.

DEFINITION 2.3. We say that a matrix function $\mathbf{F} = (F_n)$ on a vector space V is *matrix convex* if the super-graph $\mathcal{G}(\mathbf{F})$ is a matrix convex set in $V \times \mathbf{C}$. Equivalently, \mathbf{F} is matrix convex if for all $n, m \in \mathbf{N}$,

$$\begin{aligned} F_{m+n}(v \oplus w) &\preceq F_m(v) \oplus F_n(w), & v \in M_m(V), w \in M_n(V); \\ F_n(\gamma^* v \gamma) &\preceq \gamma^* F_m(v) \gamma, & \gamma \in \mathbf{M}_m, n, \gamma^* \gamma = \mathbf{1}_n, v \in M_m(V). \end{aligned}$$

If V is a topological space then we say that \mathbf{F} is *closed* if $\mathcal{G}(F_n)$ is closed for all $n \in \mathbf{N}$.

We remark that it follows from the above that the domain of a matrix convex function is a matrix convex set. Furthermore, it is easily seen from

the matrix convexity of the super-graph, that conditions (3) and (4) hold for all $n \in \mathbb{N}$ if and only if they hold for $n = 1$.

We also remark that the above definitions of matrix functions and matrix convex functions can be applied to sector valued mappings with values in $M_n(W)$ over a matrix ordered space W rather than just over \mathbb{C} . This is actually the case for the so-called “matrix sublinear functionals” studied by Wittstock in [10]. Since the results of this paper only deals with the case of \mathbb{C} we have made no attempt to include the more general case.

We begin with some examples.

EXAMPLE 2.4. Let V and V' be dual vector spaces with matrix-valued pairing defined by (1). Given $\varphi \in M_r(V')$ and $\beta \in (M_r)_{\text{sa}}$, we define a matrix function $F^{\varphi, \beta}$ on V by

$$F_n^{\varphi, \beta}(v) = \{\alpha \in (M_n)_{\text{sa}} \mid \text{Re}\langle\langle v, \varphi \rangle\rangle \leq \alpha \otimes \mathbf{1}_r + \mathbf{1}_n \otimes \beta\}$$

for $v \in M_n(V)$. This is clearly a closed sector, and $-\infty \prec F_n^{\varphi, \beta}(v) \prec +\infty$ for all $v \in M_n(V)$.

We claim that that $F^{\varphi, \beta}$ is a closed matrix convex function. Indeed, to show that

$$F_n^{\varphi, \beta}(\gamma^* v \gamma) \preceq \gamma^* F_m^{\varphi, \beta}(v) \gamma$$

for given $v \in M_m(V)$, $\gamma \in M_{m, n}$ with $\gamma^* \gamma = \mathbf{1}_n$, let $\alpha \in F_m^{\varphi, \beta}(v)$ and observe that,

$$\begin{aligned} \text{Re}\langle\langle \gamma^* v \gamma, \varphi \rangle\rangle &= (\gamma \otimes \mathbf{1}_r)^* \text{Re}\langle\langle v, \varphi \rangle\rangle (\gamma \otimes \mathbf{1}_r) \\ &\leq (\gamma \otimes \mathbf{1}_r)^* (\alpha \otimes \mathbf{1}_r + \mathbf{1}_m \otimes \beta) (\gamma \otimes \mathbf{1}_r) \\ &= \gamma^* \alpha \gamma \otimes \mathbf{1}_r + \mathbf{1}_n \otimes \beta, \end{aligned}$$

i.e., $\gamma^* \alpha \gamma \in F_n^{\varphi, \beta}(\gamma^* v \gamma)$.

A similar argument shows that $F_{m+n}^{\varphi, \beta}(v \oplus w) \preceq F_m^{\varphi, \beta}(v) \oplus F_n^{\varphi, \beta}(w)$ for $v \in M_m(V)$ and $w \in M_n(V)$, and it is easy to see that $F^{\varphi, \beta}$ is closed.

We call the matrix functions $F^{\varphi, \beta}$ defined above for the *matrix affine* functions, since they are the natural matricial analogue of the classical affine functions $v \mapsto \langle v, \varphi \rangle - \beta$ with $\varphi \in V'$ and $\beta \in \mathbb{R}$.

EXAMPLE 2.5. Given a matrix convex subset $\mathbf{K} = (K_n)$ of a vector space V , we define the *matrix indicator* function $\chi^{\mathbf{K}} = (\chi_n^{\mathbf{K}})$ on V by letting

$$\chi_n^{\mathbf{K}}(v) = \begin{cases} \mathbf{M}_n^+ & v \in K_n, \\ +\infty & v \notin K_n, \end{cases}$$

for $v \in M_n(X)$. $\chi^{\mathbf{K}}$ is clearly matrix convex with $\mathcal{D}(\chi^{\mathbf{K}}) = \mathbf{K}$, and $\chi^{\mathbf{K}}$ is closed if and only if \mathbf{K} is closed.

EXAMPLE 2.6. Given dual vector spaces V and V' and a matrix convex set $\mathbf{K} = (K_n)$, we define the *matrix support* function $\Sigma^{\mathbf{K}} = (\Sigma_n^{\mathbf{K}})$ on V' by letting

$$\Sigma_n^{\mathbf{K}}(\varphi) = \{\beta \in (\mathbf{M}_n)_{\text{sa}} \mid \operatorname{Re}\langle v, \varphi \rangle \leq \mathbf{l}_r \otimes \beta \text{ for all } v \in K_r, r \in \mathbf{N}\},$$

for $\varphi \in M_n(V')$. It is understood that $\Sigma_n^{\mathbf{K}}(\varphi) = +\infty$ when no such β exists, but at least $\Sigma_n^{\mathbf{K}}(0) \prec +\infty$. It is straightforward to see that $\Sigma^{\mathbf{K}}$ is closed and matrix convex.

EXAMPLE 2.7. We claim that the single-valued matrix convex functions on \mathbf{R} are exactly the so-called operator convex functions (cf. [5]). Recall, that given a real-valued function F defined on an interval $I \subset \mathbf{R}$ we may define $F_n(v)$ for any $v \in (\mathbf{M}_n)_{\text{sa}}$ with spectrum contained in I . Indeed, writing

$$v = u^* \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} u$$

for some unitary u and $\lambda_1, \dots, \lambda_n \in I$, we define

$$F_n(v) = u^* \begin{bmatrix} F(\lambda_1) & & \\ & \ddots & \\ & & F(\lambda_n) \end{bmatrix} u.$$

F is operator convex if

$$F_n(tv + (1-t)w) \leq tF_n(v) + (1-t)F_n(w)$$

for all $n \in \mathbf{N}$, $v, w \in (\mathbf{M}_n)_{\text{sa}}$ with spectra in I , and $t \in [0, 1]$. The claim is that if F is operator convex then $\mathbf{F} = (F_n)$ is matrix convex, defining $F_n(v) = +\infty$ when the spectrum of v is not contained in I . Since $F_{m+n}(v \oplus w) \leq F_m(v) \oplus F_n(w)$ is obvious it remains to show that $F_n(\gamma^* v \gamma) \leq \gamma^* F_m(v) \gamma$ for $\gamma \in \mathbf{M}_{m,n}$ such that $\gamma^* \gamma = \mathbf{l}_n$, and $v \in (\mathbf{M}_m)_{\text{sa}}$ with spectrum in I .

In the case where $F(0) \leq 0$, this is proved in [5, Theorem 2.1]. The general case is handled by translation. Choose $v_0 \in I$ and define a new function G on $I - v_0$ by

$$G(v) = F(v + v_0) - F(v_0).$$

Since $G_n(v) = F_n(v + v_0 \mathbf{l}_n) - F(v_0) \mathbf{l}_n$ for any $v \in (\mathbf{M}_n)_{\text{sa}}$ with spectrum in $I - v_0$ we see that G is operator convex. Since $G(0) \leq 0$, we get that $G_n(\gamma^* v \gamma) \leq \gamma^* G_m(v) \gamma$. Applying this to $v - v_0 \mathbf{l}_m$ for $v \in (\mathbf{M}_m)_{\text{sa}}$ with spectrum in I yields the desired result.

Conversely, given a single-valued matrix convex function $\mathbf{F} = (F_n)$ on \mathbf{R} , it immediately follows that F_n satisfies the inequality for operator convexity.

We just need to show that $F_{m+n}(v \oplus w) = F_m(v) \oplus F_n(w)$ in order to see that F_n is actually given by the construction above using $F = F_1$. Writing

$$F_{m+n}(v \oplus w) = \begin{bmatrix} a & c \\ c^* & b \end{bmatrix},$$

and noticing that

$$F_m(v) = F_m\left([I_m \ 0](v \oplus w) \begin{bmatrix} I_m \\ 0 \end{bmatrix}\right) \leq [I_m \ 0]F_{m+n}(v \oplus w) \begin{bmatrix} I_m \\ 0 \end{bmatrix} = a,$$

we have that

$$\begin{bmatrix} a & c \\ c^* & b \end{bmatrix} = F_{m+n}(v \oplus w) \leq F_m(v) \oplus F_n(w) \leq \begin{bmatrix} a & \\ & b \end{bmatrix}.$$

Hence $c = 0$ and therefore $F_m(v) = a$ and $F_n(w) = b$.

3. Affine approximation

In this section we prove our main technical result which generalizes the classical statement that any closed convex function F can be obtained as the pointwise supremum over all affine functions dominated by F ; a statement that is crucial for defining the Legendre-Fenchel transform.

Given matrix convex functions F and G on a vector space V , we write $F \preceq G$ if

$$F_n(v) \preceq G_n(v)$$

for all $n \in \mathbf{N}$ and $v \in M_n(V)$.

Recall from Example 2.4 how we defined the matrix affine function $F^{\varphi, \beta}$ for $\varphi \in M_m(V')$ and $\beta \in (M_m)_{\text{sa}}$ by

$$F_n^{\varphi, \beta}(v) = \{\alpha \in (M_n)_{\text{sa}} \mid \operatorname{Re}\langle v, \varphi \rangle \leq \alpha \otimes I_m + I_n \otimes \beta\}$$

for $v \in M_n(V)$.

Given a matrix convex function F on V we see that $F^{\varphi, \beta} \preceq F$ if and only if

$$\operatorname{Re}\langle v, \varphi \rangle \leq F_n(v) \otimes I_m + I_n \otimes \beta$$

for all $v \in M_n(V)$ and $n \in \mathbf{N}$.

THEOREM 3.1. *Let V and V' be dual vector spaces, equipped with the corresponding weak topologies, and let $F = (F_n)$ be a closed matrix convex function on V . Define a collection $\mathbf{K} = (K_m)$ of subsets $K_m \subset M_m(V' \times \mathbf{C})$ by*

$$K_m = \{(\varphi, \beta) \in M_m(V') \times (M_m)_{\text{sa}} \mid F^{\varphi, \beta} \preceq F\}.$$

Then \mathbf{K} is a non-empty, closed matrix convex set in $V' \times \mathbf{C}$, and

$$(5) \quad F_n(v) = \bigcap_{(\varphi, \beta) \in \mathbf{K}} F_n^{\varphi, \beta}(v).$$

for all $n \in \mathbf{N}$ and $v \in M_n(V)$.

PROOF. It is straightforward to prove that \mathbf{K} satisfies the axioms of a closed matrix convex set, and it will follow from the proof below that $K_m \neq \emptyset$ for all $m \in \mathbf{N}$.

We first consider the case where $0 \in F_1(0)$. To prove (5), let $v_0 \in M_n(V)$ be given. It is immediate that $F_n(v_0) \subset \bigcap_{(\varphi, \beta) \in \mathbf{K}} F_n^{\varphi, \beta}(v_0)$. Assume that $\alpha_0 \notin F_n(v_0)$. We wish to show that $\alpha_0 \notin \bigcap_{(\varphi, \beta) \in \mathbf{K}} F_n^{\varphi, \beta}(v_0)$.

The proof consists of two applications of Theorem 1.3. We first find a special element $(\varphi, \beta) \in K_1$ using essentially only the classical theorem, and then we use this element combined with matricial theorem to find elements $(\varphi_\varepsilon, \beta_\varepsilon) \in K_n$ for $\varepsilon > 0$ such that $\alpha_0 \notin \bigcap_{\varepsilon > 0} F_n^{\varphi_\varepsilon, \beta_\varepsilon}(v_0)$.

By the definition of a matrix function, we may choose $v_1 \in V$ such that $F_1(v_1) \prec +\infty$. Furthermore, since $-\infty \prec F_1(v_1)$ we also have $\alpha_1 \in \mathbf{R}$ such that $\alpha_1 \notin F_1(v_1)$.

Since $\mathcal{G}(\mathbf{F})$ is a closed matrix convex set containing 0 and $(v_1, \alpha_1) \notin \mathcal{G}(F_1)$, we may apply Theorem 1.3 to give a continuous linear mapping $\Phi: V \times \mathbf{C} \rightarrow \mathbf{C}$ such that

$$\operatorname{Re} \Phi_r(v, \alpha) \leq 1_r$$

for all $r \in \mathbf{N}$ and $(v, \alpha) \in \mathcal{G}(F_r)$ and

$$\operatorname{Re} \Phi(v_1, \alpha_1) > 1.$$

Identifying V' and the weakly continuous linear functionals on V , define $\varphi_1 \in V'$ by $\langle v, \varphi_1 \rangle = \Phi(v, 0)$, and set $\beta_1 = -\operatorname{Re} \Phi(0, 1) \in \mathbf{R}$. Then

$$\operatorname{Re} \Phi_r(v, \alpha) = \operatorname{Re} \langle v, \varphi_1 \rangle - \alpha \beta_1$$

for all $v \in M_r(V)$ and $\alpha \in (M_r)_{\text{sa}}$. By assumption $(0, \mathbf{R}^+) \subset \mathcal{G}(F_1)$, so $-\mathbf{R}^+ \beta_1 \leq 1$ which implies that $\beta_1 \geq 0$.

We also have that $\beta_1 \neq 0$. If $\beta_1 = 0$ then $S = \{\alpha \in \mathbf{R} \mid \operatorname{Re} \Phi(v_1, \alpha) \leq 1\} = \{\alpha \in \mathbf{R} \mid \operatorname{Re} \langle v_1, \varphi_1 \rangle \leq 1\}$ is either empty or contains all of \mathbf{R} , and this is a contradiction since $\alpha_1 \notin S$ and $\emptyset \neq F_1(v_1) \subset S$.

We then define $\varphi \in V'$ and $\beta \in \mathbf{R}$ by

$$\begin{aligned} \beta &= \beta_1^{-1}; \\ \varphi &= \beta_1^{-1} \varphi_1. \end{aligned}$$

For any $v \in M_r(V)$ we have that

$$\begin{aligned}
& \{\alpha \in (\mathbf{M}_r)_{\text{sa}} \mid \text{Re } \Phi_r(v, \alpha) \leq \mathbf{l}_r\} \\
&= \{\alpha \in (\mathbf{M}_r)_{\text{sa}} \mid \text{Re} \langle v, \varphi_1 \rangle - \alpha \beta_1 \leq \mathbf{l}_r\} \\
&= \{\alpha \in (\mathbf{M}_r)_{\text{sa}} \mid \text{Re} \langle v, \varphi \rangle \leq \alpha + \beta \mathbf{l}_r\} \\
&= F_r^{\varphi, \beta}(v).
\end{aligned}$$

Since $F_r(v) \subset \{\alpha \mid \text{Re } \Phi_r(v, \alpha) \leq \mathbf{l}_r\}$, this shows that $(\varphi, \beta) \in K_1$.

We now repeat the application of the Theorem 1.3 using the element $(v_0, \alpha_0) \in M_n(V) \times (\mathbf{M}_n)_{\text{sa}}$. Since $\mathcal{G}(\mathbf{F})$ is a closed matrix convex set containing 0 and $(v_0, \alpha_0) \notin \mathcal{G}(F_n)$, there exists a continuous linear mapping $\Psi : V \times \mathbf{C} \rightarrow \mathbf{M}_n$ such that

$$\text{Re } \Psi_r(v, \alpha) \leq \mathbf{l}_r \otimes \mathbf{l}_n$$

for all $r \in \mathbf{N}$ and $(v, \alpha) \in \mathcal{G}(F_r)$ and

$$\text{Re } \Psi_n(v_0, \alpha_0) \not\leq \mathbf{l}_n \otimes \mathbf{l}_n.$$

Identifying $M_n(V')$ and the weakly continuous linear maps $V \rightarrow \mathbf{M}_n$, define $\varphi_0 \in M_n(V')$ by $\langle v, \varphi_0 \rangle = \Psi(v, 0)$, and set $\beta_0 = -\text{Re } \Psi(0, 1) \in (\mathbf{M}_n)_{\text{sa}}$. Then

$$\text{Re } \Psi_r(v, \alpha) = \text{Re} \langle v, \varphi_0 \rangle - \alpha \otimes \beta_0$$

for all $v \in M_r(V)$ and $\alpha \in (\mathbf{M}_r)_{\text{sa}}$. By assumption $(0, \mathbf{R}^+) \subset \mathcal{G}(F_1)$, so $-\mathbf{R}^+ \beta_0 \leq \mathbf{l}_n$ which implies that $\beta_0 \in \mathbf{M}_n^+$, but β_0 is not necessarily invertible as in the case above.

For $\varepsilon > 0$ we define $\varphi_\varepsilon \in M_n(V')$ and $\beta_\varepsilon \in (\mathbf{M}_n)_{\text{sa}}$ by

$$\begin{aligned}
\beta_\varepsilon &= ((1 - \varepsilon)\beta_0 + \varepsilon\beta_1 \mathbf{l}_n)^{-1}; \\
\varphi_\varepsilon &= \beta_\varepsilon^{1/2}((1 - \varepsilon)\varphi_0 + \varepsilon\varphi_1 \mathbf{l}_n)\beta_\varepsilon^{1/2}.
\end{aligned}$$

using that $\beta_0 \in \mathbf{M}_n^+$ and $\beta_1 > 0$. We claim that $(\varphi_\varepsilon, \beta_\varepsilon) \in K_n$. Indeed, given $(v, \alpha) \in \mathcal{G}(F_r)$ we have

$$\begin{aligned}
& \text{Re} \langle v, (1 - \varepsilon)\varphi_0 + \varepsilon\varphi_1 \mathbf{l}_n \rangle - \alpha \otimes ((1 - \varepsilon)\beta_0 + \varepsilon\beta_1 \mathbf{l}_n) \\
&= (1 - \varepsilon)(\text{Re} \langle v, \varphi_0 \rangle - \alpha \otimes \beta_0) + \varepsilon(\text{Re} \langle v, \varphi_1 \mathbf{l}_n \rangle - \alpha \otimes \beta_1 \mathbf{l}_n) \\
&= (1 - \varepsilon)\text{Re } \Psi_r(v, \alpha) + \varepsilon \text{Re } \Phi_r(v, \alpha) \otimes \mathbf{l}_n \\
&\leq (1 - \varepsilon)\mathbf{l}_r \otimes \mathbf{l}_n + \varepsilon \mathbf{l}_r \otimes \mathbf{l}_n = \mathbf{l}_r \otimes \mathbf{l}_n,
\end{aligned}$$

which is equivalent to

$$\text{Re} \langle v, \varphi_\varepsilon \rangle \leq \alpha \otimes \mathbf{l}_n + \mathbf{l}_r \otimes \beta_\varepsilon.$$

Hence $(\varphi_\varepsilon, \beta_\varepsilon) \in K_n$.

To prove that $\alpha_0 \notin \bigcap_{\varepsilon > 0} F_n^{\varphi_\varepsilon, \beta_\varepsilon}(v_0)$, assume the converse, i.e., that

$$\text{Re} \langle v_0, \varphi_\varepsilon \rangle \leq \alpha_0 \otimes \mathbf{l}_n + \mathbf{l}_n \otimes \beta_\varepsilon$$

for all $\varepsilon > 0$. Reversing the above calculation this implies that

$$(1 - \varepsilon)\mathbf{Re} \Psi_n(v_0, \alpha_0) + \varepsilon\mathbf{Re} \Phi_n(v_0, \alpha_0) \otimes \mathbf{l}_n \leq \mathbf{l}_n \otimes \mathbf{l}_n,$$

for all $\varepsilon > 0$, which leads to the contradiction that $\mathbf{Re} \Psi_n(v_0, \alpha_0) \leq \mathbf{l}_n \otimes \mathbf{l}_n$. This concludes the proof in the case where $0 \in F_1(0)$.

The general case is handled by translation. It is straightforward but for the convenience of the reader we include the proof. Choose $(v', \alpha') \in \mathcal{G}(F_1)$, and define a new closed matrix convex function $\mathbf{G} = (G_n)$ on V by

$$G_n(v) = F_n(v + v'\mathbf{l}_n) - \alpha'\mathbf{l}_n$$

for $v \in M_n(V)$. Since $0 \in G_1(0)$ the first part of the proof implies that

$$G_n(v) = \bigcap_{(\varphi, \beta) \in \mathbf{L}} F_n^{\varphi, \beta}(v),$$

where $(\varphi, \beta) \in L_m \subset M_m(V') \times (M_m)_{\text{sa}}$ if and only if $F^{\varphi, \beta} \preceq \mathbf{G}$. Since $F^{\varphi, \beta} \preceq \mathbf{G}$ if and only if

$$\begin{aligned} \mathbf{Re}\langle v, \varphi \rangle &\leq G_r(v) \otimes \mathbf{l}_m + \mathbf{l}_r \otimes \beta \\ &= F_r(v + v'\mathbf{l}_r) \otimes \mathbf{l}_m + \mathbf{l}_r \otimes (\beta - \alpha'\mathbf{l}_m) \end{aligned}$$

for all $v \in M_r(V)$ and $r \in \mathbf{N}$, which again is equivalent to

$$\mathbf{Re}\langle v + v'\mathbf{l}_r, \varphi \rangle \leq F_r(v + v'\mathbf{l}_r) \otimes \mathbf{l}_m + \mathbf{l}_r \otimes (\beta - \alpha'\mathbf{l}_m + \mathbf{Re}\langle v', \varphi \rangle),$$

we see that

$$(\varphi, \beta) \in \mathbf{L} \Leftrightarrow (\varphi, \beta') \in \mathbf{K}.$$

where $\beta \mapsto \beta' : (M_m)_{\text{sa}} \rightarrow (M_m)_{\text{sa}}$ is the bijection given by

$$\beta' = \beta - \alpha'\mathbf{l}_m + \mathbf{Re}\langle v', \varphi \rangle.$$

Furthermore, we have that

$$F_n^{\varphi, \beta}(v - v'\mathbf{l}_n) + \alpha'\mathbf{l}_n = F_n^{\varphi, \beta'}(v).$$

Hence

$$\begin{aligned} F_n(v) &= G_n(v - v'\mathbf{l}_n) + \alpha'\mathbf{l}_n \\ &= \bigcap_{(\varphi, \beta) \in \mathbf{L}} F_n^{\varphi, \beta}(v - v'\mathbf{l}_n) + \alpha'\mathbf{l}_n \\ &= \bigcap_{(\varphi, \beta') \in \mathbf{K}} F_n^{\varphi, \beta'}(v), \end{aligned}$$

which concludes the proof.

4. The non-commutative Legendre-Fenchel transform

Let us recall that given dual vector spaces V and V' the classical *Legendre-Fenchel transform* (or conjugate) of a convex function $F : V \rightarrow \mathbb{R} \cup \{+\infty\}$ is the convex function $F' : V' \rightarrow \mathbb{R} \cup \{+\infty\}$ given by

$$F'(\varphi) = \sup\{\operatorname{Re}\langle v, \varphi \rangle - F(v) \mid v \in \mathcal{D}(F)\}$$

for $\varphi \in V'$. Equivalently, F' is described by its super-graph $\mathcal{G}(F')$ given by

$$\mathcal{G}(F') = \{(\varphi, \beta) \in V' \times \mathbb{R} \mid \operatorname{Re}\langle v, \varphi \rangle \leq F(v) + \beta \text{ for all } v \in V\}.$$

The Legendre-Fenchel transform of F is a closed convex function, and if F is closed then $F = (F')'$. It is immediate that F and F' satisfy the *Fenchel's inequality* which states that

$$\operatorname{Re}\langle v, \varphi \rangle \leq F(v) + F'(\varphi)$$

for all $v \in V$ and $\varphi \in V'$. For more on the classical theory of the Legendre-Fenchel transform, see [8] and [6].

EXAMPLE 4.1. A standard example of a conjugate pair of closed convex functions on \mathbb{R} (considered in duality with itself) is

$$F(\alpha) = \frac{1}{p} |\alpha|^p;$$

$$F'(\beta) = \frac{1}{q} |\beta|^q,$$

with $1 < p, q < +\infty$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$. In this case the Fenchel inequality reduces to

$$\alpha\beta \leq \frac{1}{p} |\alpha|^p + \frac{1}{q} |\beta|^q,$$

which is often called *Young's inequality*.

Let V and V' be dual vector spaces and let $\mathbf{F} = (F_n)$ be a matrix convex function on V . We define the *matricial Legendre-Fenchel transform* of \mathbf{F} to be the matrix function \mathbf{F}' on V' given by the closed sector

$$F'_n(\varphi) = \{\beta \in (\mathbf{M}_n)_{\text{sa}} \mid \operatorname{Re}\langle\langle v, \varphi \rangle\rangle \leq F_r(v) \otimes \mathbf{1}_n + \mathbf{1}_r \otimes \beta \text{ for all } v \in M_r(V), r \in \mathbf{N}\}$$

for $\varphi \in M_n(V')$. In particular $F'_n(\varphi) = \infty (= \emptyset)$ if no such β exists. Note that for $n = 1$ this reduces to the classical Legendre-Fenchel transform.

We observe that

$$\begin{aligned}
 (6) \quad F'_n(\varphi) &= \bigcap_{(v,\alpha) \in \mathcal{G}(F)} \{ \beta \in (\mathbf{M}_n)_{\text{sa}} \mid \text{Re} \langle v, \varphi \rangle \leq \alpha \otimes \mathbf{1} + \mathbf{1} \otimes \beta \} \\
 &= \bigcap_{(v,\alpha) \in \mathcal{G}(F)} F_n^{v,\alpha}(\varphi),
 \end{aligned}$$

for $\varphi \in M_n(V')$, and also that $(\varphi, \beta) \in \mathcal{G}(F'_n)$ if and only if $F^{\varphi,\beta} \preceq F$. This shows that $\mathcal{G}(F')$ is exactly the closed matrix convex set \mathbf{K} of Theorem 3.1. In particular, this implies that F' is a closed matrix convex function.

Notice that the definition of the Legendre-Fenchel transform implies the following matricial version of the Fenchel inequality,

$$\text{Re} \langle v, \varphi \rangle \leq F_n(v) \otimes \mathbf{1}_m + \mathbf{1}_n \otimes F'_m(\varphi)$$

for all $v \in M_n(V)$, $\varphi \in M_m(V')$, and $m, n \in \mathbf{N}$.

COROLLARY 4.2. *Let V and V' be dual vector spaces, equipped with the corresponding weak topologies. If $F = (F_n)$ is a closed matrix convex function on V then*

$$F = (F')'.$$

PROOF. Let $v \in M_n(V)$. By (6) we have that

$$F''_n(v) = \bigcap_{(\varphi,\beta) \in \mathcal{G}(F')} F_n^{\varphi,\beta}(v).$$

and by Theorem 3.1

$$F_n(v) = \bigcap_{(\varphi,\beta) \in \mathbf{K}} F_n^{\varphi,\beta}(v).$$

But as remarked above, $\mathbf{K} = \mathcal{G}(F')$ so it follows that $F''_n(v) = F_n(v)$.

EXAMPLE 4.3. If $\mathbf{K} = (K_n)$ is a closed matrix convex set in V then $(\chi^{\mathbf{K}})' = \Sigma^{\mathbf{K}}$ and $(\Sigma^{\mathbf{K}})' = \chi^{\mathbf{K}}$ (cf. Examples 2.5 and 2.6). Indeed, for $\varphi \in M_n(V')$ we have that

$$\begin{aligned}
 &(\chi^{\mathbf{K}})'_n(\varphi) \\
 &= \{ \beta \in (\mathbf{M}_n)_{\text{sa}} \mid \text{Re} \langle v, \varphi \rangle \leq \chi_r^{\mathbf{K}}(v) \otimes \mathbf{1}_n + \mathbf{1}_r \otimes \beta \text{ for all } v \in M_r(V), r \in \mathbf{N} \} \\
 &= \{ \beta \in (\mathbf{M}_n)_{\text{sa}} \mid \text{Re} \langle v, \varphi \rangle \leq \mathbf{M}_r^+ \otimes \mathbf{1}_n + \mathbf{1}_r \otimes \beta \text{ for all } v \in K_r, r \in \mathbf{N} \} \\
 &= \{ \beta \in (\mathbf{M}_n)_{\text{sa}} \mid \text{Re} \langle v, \varphi \rangle \leq \mathbf{1}_r \otimes \beta \text{ for all } v \in K_r, r \in \mathbf{N} \} \\
 &= \Sigma_n^{\mathbf{K}}(\varphi),
 \end{aligned}$$

i.e., $(\chi^{\mathbf{K}})' = \Sigma^{\mathbf{K}}$. The second statement follows from the first and the fact that $\chi^{\mathbf{K}}$ is closed:

$$(\Sigma^K)' = (\chi^K)'' = \chi^K.$$

EXAMPLE 4.4. It is not true that the conjugate of a single-valued matrix convex function is single-valued; this fails even in the case of single-valued matrix convex function on \mathbb{R} , i.e., in the case of operator convex functions (cf. Example 2.7). Using the fact that $v \mapsto v^{1/2}$ is operator monotone ([1, Example 3]) on \mathbb{R}^+ and that $v \mapsto v^2$ is not ([1, Example 2]), combined with [5, Theorem 2.4] we obtain that $F : v \mapsto \frac{2}{3}v^{3/2}$ is operator convex, and that $G : v \mapsto \frac{1}{3}v^3$ is not. The classical Legendre-Fenchel transform of F is G since $1/\frac{3}{2} + 1/3 = 1$, so in particular $F'_1 = G$. But F' cannot be single-valued since this would imply that F' is determined by G as in Example 2.7 and hence that is G operator convex.

5. The Legendre-Fenchel transform and the Cauchy-Schwartz inequality

Let $V = \mathcal{H}$ be a Hilbert space with inner product $(\cdot | \cdot)$ and norm $\|\cdot\|$, regarded in duality with the conjugate Hilbert space $\bar{\mathcal{H}}$ via the duality

$$\langle \xi, \bar{\eta} \rangle = (\xi | \eta)$$

for $\xi \in \mathcal{H}$, $\bar{\eta} \in \bar{\mathcal{H}}$. Here we denote elements in $\bar{\mathcal{H}}$ by $\bar{\eta}$ with $\eta \in \mathcal{H}$ to distinguish $\bar{\mathcal{H}}$ from \mathcal{H} . Let the convex function $F : \mathcal{H} \rightarrow \mathbb{R}$ be given by

$$F(\xi) = 1/2\|\xi\|^2$$

for $\xi \in \mathcal{H}$. Observing that the Cauchy-Schwarz inequality is equivalent to the inequality

$$\operatorname{Re}\langle \xi, \bar{\eta} \rangle \leq 1/2\|\xi\|^2 + 1/2\|\eta\|^2 = F(\xi) + F(\eta)$$

for $\xi, \eta \in \mathcal{H}$, we see that the Cauchy-Schwartz inequality is equivalent to the statement that the Legendre-Fenchel transform $F' : \bar{\mathcal{H}} \rightarrow \mathbb{R}$ of F satisfies $F'(\bar{\eta}) \leq F(\eta)$. We actually have that $F'(\bar{\eta}) = F(\eta)$ since $F'(\bar{\eta}) \geq \operatorname{Re}\langle \eta, \bar{\eta} \rangle - F(\eta) = F(\eta)$. Moreover, it is easy to see that the condition $F'(\bar{\eta}) = F(\eta)$ determines F uniquely.

As an application of the theory of matrix convex functions we present a proper “quantization” of F . It is clear from the above that this must involve a matricial version of the Cauchy-Schwartz inequality, which in effect is the content of the lemma below. For $\alpha \in \mathbb{M}_m$, $\bar{\alpha} \in \mathbb{M}_m$ is obtained by conjugating of all the entries of α .

LEMMA 5.1. *Let $\alpha_1, \dots, \alpha_k \in \mathbb{M}_m$ and $\beta_1, \dots, \beta_k \in \mathbb{M}_n$, and assume that $\alpha \in (\mathbb{M}_m)_{\text{sa}}$ and $\beta \in (\mathbb{M}_n)_{\text{sa}}$ satisfies*

$$\begin{aligned} \operatorname{Re} \sum_{i=1}^k \alpha_i \otimes \bar{\alpha}_i &\leq \alpha \otimes \mathbf{l}_m + \mathbf{l}_m \otimes \bar{\alpha}, \\ \operatorname{Re} \sum_{i=1}^k \beta_i \otimes \bar{\beta}_i &\leq \beta \otimes \mathbf{l}_n + \mathbf{l}_n \otimes \bar{\beta}. \end{aligned}$$

Then

$$\operatorname{Re} \sum_{i=1}^k \alpha_i \otimes \bar{\beta}_i \leq \alpha \otimes \mathbf{l}_n + \mathbf{l}_m \otimes \bar{\beta}.$$

PROOF. Let $\mathcal{H}\mathcal{S}_{m,n}$ denote the Hilbert space of Hilbert-Schmidt operators in $M_{m,n}$ with inner product $(x|y)_{\mathcal{H}\mathcal{S}} = \operatorname{Tr}(xy^*)$ for $x, y \in M_{m,n}$, where Tr denotes the trace on M_m . The *-isomorphism of $M_m \otimes M_n$ onto $\mathcal{B}(\mathcal{H}\mathcal{S}_{m,n})$ mapping $\alpha \otimes \bar{\beta}$ to $x \mapsto \alpha x \beta^*$ with $x \in \mathcal{H}\mathcal{S}_{m,n}$ is an order isomorphism. We therefore have to show that

$$\sum_{i=1}^k \operatorname{Re} \operatorname{Tr}(\alpha_i x \beta_i^* x^*) \leq \operatorname{Tr}(\alpha x x^*) + \operatorname{Tr}(x \beta x^*)$$

for all $x \in M_{m,n}$.

Let $x = \nu|x|$ be the polar decomposition of x with $\nu \in M_{m,n}$ and $|x| \in M_n$. Set

$$x_1 = \nu|x|^{1/2} \in M_{m,n}, \quad x_2 = |x|^{1/2} \in M_n.$$

Then we have that

$$x = x_1 x_2, \quad |x| = x_2^* x_2, \quad |x^*| = x_1 x_1^*.$$

Applying the Cauchy-Schwartz inequality twice,

$$\begin{aligned} \sum_{i=1}^k \operatorname{Re} \operatorname{Tr}(\alpha_i x \beta_i^* x^*) &= \sum_{i=1}^k \operatorname{Re} \operatorname{Tr}(x_1^* \alpha_i x_1 (x_2 \beta_i^* x_2^*)) \\ &\leq \sum_{i=1}^k \operatorname{Tr}(x_1^* \alpha_i x_1 (x_1^* \alpha_i x_1)^*)^{1/2} \operatorname{Tr}(x_2 \beta_i x_2^* (x_2 \beta_i x_2^*)^*)^{1/2} \\ &= \sum_{i=1}^k \operatorname{Tr}(\alpha_i |x^*| \alpha_i^* |x^*|)^{1/2} \operatorname{Tr}(\beta_i |x| \beta_i^* |x|)^{1/2} \\ &\leq \left(\sum_{i=1}^k \operatorname{Tr}(\alpha_i |x^*| \alpha_i^* |x^*|) \right)^{1/2} \left(\sum_{i=1}^k \operatorname{Tr}(\beta_i |x| \beta_i^* |x|) \right)^{1/2} \\ &\leq 1/2 \sum_{i=1}^k \operatorname{Tr}(\alpha_i |x^*| \alpha_i^* |x^*|) + 1/2 \sum_{i=1}^k \operatorname{Tr}(\beta_i |x| \beta_i^* |x|). \end{aligned}$$

By the assumption we get that

$$\begin{aligned} \sum_{i=1}^k \operatorname{Re} \operatorname{Tr}(\alpha_i x \beta_i^* x^*) &\leq 1/2(\operatorname{Tr}(\alpha |x^*|^2) + \operatorname{Tr}(|x^*| \alpha |x^*|)) \\ &\quad + 1/2(\operatorname{Tr}(\beta |x|^2) + \operatorname{Tr}(|x| \beta |x|)) \\ &= \operatorname{Tr}(\alpha x x^*) + \operatorname{Tr}(x \beta x^*), \end{aligned}$$

which proves the lemma.

The above proof is modelled after Haagerups proof of the normed version of the above statement ([4, Lemma 2.4]), which says that

$$\left\| \sum_{i=1}^k \alpha_i \otimes \bar{\beta}_i \right\| \leq \left\| \sum_{i=1}^k \alpha_i \otimes \bar{\alpha}_i \right\|^{1/2} \left\| \sum_{i=1}^k \beta_i \otimes \bar{\beta}_i \right\|^{1/2}.$$

This inequality plays an important rôle in Pisier's operator Hilbert space (see [7]). It is not hard to prove that Lemma 5.1 implies the above inequality using a standard 2×2 matrix trick.

Given $\xi = \xi_{i,j} \in M_m(\mathcal{H})$ and $\bar{\eta} = \bar{\eta}_{i,j} \in M_n(\bar{\mathcal{H}})$, choose a basis e_1, \dots, e_k in \mathcal{H} for the subspace spanned by $\xi_{i,j}$ and $\eta_{i,j}$. We may then find matrices $\alpha_1, \dots, \alpha_k \in M_m$ and $\beta_1, \dots, \beta_k \in M_n$ such that

$$\begin{aligned} \xi &= \sum_{i=1}^k \alpha_i e_i, \\ \eta &= \sum_{i=1}^k \beta_i e_i, \end{aligned}$$

from which we observe that

$$\langle\langle \xi, \bar{\eta} \rangle\rangle = \sum_{i=1}^k \alpha_i \otimes \bar{\beta}_i.$$

We may therefore reformulate Lemma 5.1 to say the following.

LEMMA 5.2. *Let \mathcal{H} be a Hilbert space, regarded in duality with the conjugate Hilbert space $\bar{\mathcal{H}}$. If $\xi \in M_m(\mathcal{H})$, $\bar{\eta} \in M_n(\bar{\mathcal{H}})$ satisfies*

$$\begin{aligned} \operatorname{Re} \langle\langle \xi, \bar{\xi} \rangle\rangle &\leq \alpha \otimes \mathbf{l}_m + \mathbf{l}_m \otimes \bar{\alpha}, \\ \operatorname{Re} \langle\langle \eta, \bar{\eta} \rangle\rangle &\leq \beta \otimes \mathbf{l}_n + \mathbf{l}_n \otimes \bar{\beta}, \end{aligned}$$

for $\alpha \in (M_m)_{\text{sa}}$ and $\beta \in (M_n)_{\text{sa}}$ then

$$\operatorname{Re} \langle\langle \xi, \bar{\eta} \rangle\rangle \leq \alpha \otimes \mathbf{l}_n + \mathbf{l}_m \otimes \bar{\beta}.$$

We consider the above lemma the matricial version of the Cauchy-Schwartz inequality since it enables us to prove the existence and uniqueness of the proper “quantization” of $\xi \mapsto 1/2\|\xi\|^2$ as in the next proposition.

PROPOSITION 5.3. *Let \mathcal{H} be a Hilbert space, regarded in duality with the conjugate Hilbert space $\bar{\mathcal{H}}$. Define*

$$F_m(\xi) = \{\alpha \in (\mathbf{M}_m)_{\text{sa}} \mid \text{Re}\langle\langle \xi, \bar{\xi} \rangle\rangle \leq \alpha \otimes \mathbf{1}_m + \mathbf{1}_m \otimes \bar{\alpha}\}$$

for $\xi \in M_m(\mathcal{H})$, and define \bar{F}_n on $M_n(\bar{\mathcal{H}})$ by

$$\bar{F}_n(\bar{\eta}) = \overline{F_n(\eta)}$$

for $\bar{\eta} \in M_n(\bar{\mathcal{H}})$. Then $\mathbf{F} = (F_n)$ is the unique everywhere defined matrix convex function on \mathcal{H} satisfying

$$\mathbf{F}' = \bar{\mathbf{F}}.$$

PROOF. Define $\mathbf{F} = (F_n)$ as above. \mathbf{F} is clearly an everywhere defined matrix function, and we claim that \mathbf{F} is matrix convex.

Let $\alpha \in F_m(\xi)$ and $\beta \in F_n(\eta)$ be given with $\xi \in M_m(\mathcal{H})$ and $\eta \in M_n(\mathcal{H})$. By the matricial Cauchy-Schwartz inequality (Lemma 5.2) we have that

$$\begin{aligned} \text{Re}\langle\langle \xi \oplus \eta, \bar{\xi} \oplus \bar{\eta} \rangle\rangle &= \text{Re}\langle\langle \xi, \bar{\xi} \rangle\rangle \oplus \text{Re}\langle\langle \eta, \bar{\eta} \rangle\rangle \oplus \text{Re}\langle\langle \eta, \bar{\xi} \rangle\rangle \oplus \text{Re}\langle\langle \xi, \bar{\eta} \rangle\rangle \\ &\leq (\alpha \otimes \mathbf{1}_m + \mathbf{1}_m \otimes \bar{\alpha}) \oplus (\alpha \otimes \mathbf{1}_n + \mathbf{1}_n \otimes \bar{\beta}) \\ &\quad \oplus (\beta \otimes \mathbf{1}_m + \mathbf{1}_n \otimes \bar{\alpha}) \oplus (\beta \otimes \mathbf{1}_n + \mathbf{1}_n \otimes \bar{\beta}) \\ &= (\alpha \oplus \beta) \otimes \mathbf{1}_{m+n} + \mathbf{1}_{m+n} \otimes (\bar{\alpha} \oplus \bar{\beta}), \end{aligned}$$

i.e., $\alpha \oplus \beta \in F_{m+n}(\xi \oplus \eta)$.

It is straightforward to see that $F_n(\gamma^* \xi \gamma) \preceq \gamma^* F_m(\xi) \gamma$ when $\xi \in M_m(\mathcal{H})$ and $\gamma \in \mathbf{M}_{m,n}$, $\gamma^* \gamma = \mathbf{1}_n$. Hence \mathbf{F} is matrix convex.

To prove $\mathbf{F}' = \bar{\mathbf{F}}$, we apply Lemma 5.2 again to see that

$$\begin{aligned} \overline{F_n(\eta)} \subset \{\beta \in (\mathbf{M}_n)_{\text{sa}} \mid \text{Re}\langle\langle \xi, \bar{\eta} \rangle\rangle \leq F_m(\xi) \otimes \mathbf{1}_n + \mathbf{1}_m \otimes \beta \text{ for all} \\ \xi \in M_m(\mathcal{H}), m \in \mathbf{N}\} = F'_n(\bar{\eta}). \end{aligned}$$

For the converse inclusion, let $\beta \in F'_n(\bar{\eta})$ and observe that since $0 \in F_1(0)$ we have that $\beta \geq 0$. It suffices to prove that $\beta + \varepsilon \mathbf{1}_n \in \overline{F_n(\eta)}$ for $\varepsilon > 0$. Letting $k = 2\varepsilon\|\langle\langle \eta, \bar{\eta} \rangle\rangle\|^{-1}$, we see that

$$\begin{aligned} \text{Re}\langle\langle k^{1/2}\eta, k^{1/2}\bar{\eta} \rangle\rangle &\leq k\|\langle\langle \eta, \bar{\eta} \rangle\rangle\| \mathbf{1}_n \otimes \mathbf{1}_n = 2\varepsilon \mathbf{1}_n \otimes \mathbf{1}_n \\ &\leq \overline{(\beta + \varepsilon \mathbf{1}_n)} \otimes \mathbf{1}_n + \mathbf{1}_n \otimes (\beta + \varepsilon \mathbf{1}_n), \end{aligned}$$

i.e., $\beta + \varepsilon \mathbf{1}_n \in \overline{F_n(k^{1/2}\eta)}$. By the above $\beta + \varepsilon \mathbf{1}_n \in F'_n(\bar{\eta})$ and therefore

$$\text{Re}\langle\langle k^{1/4}\eta, k^{1/4}\bar{\eta} \rangle\rangle = \text{Re}\langle\langle k^{1/2}\eta, \bar{\eta} \rangle\rangle \leq \overline{(\beta + \varepsilon \mathbf{1}_n)} \otimes \mathbf{1}_n + \mathbf{1}_n \otimes (\beta + \varepsilon \mathbf{1}_n),$$

i.e., $\beta + \varepsilon \mathbf{l}_n \in \overline{F_n(k^{1/4}\eta)}$. Inductively we get that $\beta + \varepsilon \mathbf{l}_n \in \overline{F_n(k^{1/2^p}\eta)}$, and since $k^{1/2^p} \rightarrow 1$ as $p \rightarrow \infty$ we get that $\beta + \varepsilon \mathbf{l}_n \in \overline{F_n(\eta)}$. Since $\varepsilon > 0$ was arbitrary, $\beta \in \overline{F_n(\eta)}$.

Thus \mathbf{F} satisfies $\mathbf{F}' = \overline{\mathbf{F}}$. To see uniqueness, let $\mathbf{G} = (G_n)$ be a matrix convex function satisfying $\mathbf{G}' = \overline{\mathbf{G}}$. If $\beta \in G_n(\eta)$ then $\bar{\beta} \in G'_n(\bar{\eta})$, and therefore $\operatorname{Re}\langle\langle \eta, \bar{\eta} \rangle\rangle \leq \beta \otimes \mathbf{l}_n + \mathbf{l}_n \otimes \bar{\beta}$, so $G_n(\eta) \subset F_n(\eta)$. Hence $\mathbf{F} \preceq \mathbf{G}$, and therefore also $\mathbf{G} = \overline{\mathbf{G}'} \preceq \overline{\mathbf{F}'} = \mathbf{F}$.

REFERENCES

1. C. Davis, *Notions generalizing convexity for functions defined on spaces of matrices*, Proc. Amer. Math. Soc. Symposia (Convexity) 7 (1963), 187–201.
2. E. Effros and S. Winkler, *Matrix convexity: Operator analogues of the Bipolar and Hahn-Banach theorems*, to appear in J. Funct. Anal., 1995.
3. W. Fenchel, *On conjugate convex functions*, Canad. J. Math. 1 (1949), 73–77.
4. U. Haagerup, *Injectivity and decomposition of completely bounded maps*, in Operator algebras and their connections with topology and ergodic theory, Lecture Notes in Math. 1132 (1985), 170–222.
5. F. Hansen and G. K. Pedersen, *Jensen's inequality for operators and Löwner's theorem*, Math. Ann. 258 (1982), 229–241.
6. L. Hörmander, *Notions of Convexity*, Birkhäuser, 1994.
7. G. Pisier, *The operator Hilbert space OH, complex interpolation, and tensor norms*, Mem. Amer. Math. Soc. 122 (1996).
8. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.
9. S. Winkler, *Matrix convexity*, Ph. D. thesis, UCLA, 1996.
10. G. Wittstock, *Ein Operatorvertiger Hahn-Banach Satz*, J. Funct. Anal. 40 (1981), 127–150.
11. G. Wittstock, *On Matrix Order and Convexity*, in Functional Analysis, Survey and Recent Results III, Math. Studies 90 (1984), 175–188.

UNIVERSITY OF WALES, SWANSEA
 SINGLETON PARK
 SWANSEA SA2 8PP
 UK

E-mail address: swi@simcorp.dk