Noname manuscript No.
(will be inserted by the editor)

E. Lieb convexity inequalities and noncommutative Bernstein inequality in Jordan-algebraic setting

Leonid Faybusovich

communicated by

Received: date / Accepted: date

Abstract We describe a Jordan-algebraic version of E. Lieb convexity inequalities. A joint convexity of Jordan-algebraic version of quantum entropy is proven. A version of noncommutative Bernstein inequality is proven as an application of one of convexity inequalities. A spectral theory on semi-simple Jordan complex algebras is used as a tool to prove the convexity results. Possible applications to optimization and statistics are indicated.

Keywords Generalized Convexity Euclidean Jordan algebras Quantum Entropy

Mathematics Subject Classification (2000) MSC 90C25 · MSC 17C20

1 Introduction

In [11] E. Lieb proved a number of interrelated convexity inequalities, which found important applications in quantum physics, quantum information theory, statistics and probability. An interesting fact related to these inequalities is that pretty much all of them admit a Jordan-algebraic interpretation. That makes it tempting to generalize them to the setting of Euclidean Jordan algebras. If a given simple Euclidean Jordan algebra admits a representation in Jordan algebra of real symmetric matrices, it is quite straightforward in most of the cases. Unfortunately, it is not always the case. Since an arbitrary Euclidean Jordan algebra is a direct sum of simple ones, to prove the results in general, a different approach is required. While by now a number of different proofs of original results is known, surprisingly (and in contrast with mere reformulation of the results), none of them admits an immediate generalization in Jordan-algebraic setting. In present paper we provide a Jordan-algebraic version of E. Lieb's results. One can consider this paper as an attempt to further develop a version of matrix analysis (in the sense

L. Faybusovich
University of Notre Dame
Department of Mathematics
255 Hurley Hall
Notre Dame,IN 46556 ,USA
E-mail: leonid.faybusovich.1@nd.edu

of, say, [1]) in the context of Euclidean Jordan algebras. The plan of the paper is as follows. In section 2 we introduce the vocabulary of Euclidean Jordan algebras. In section 3 we formulate a Jordan-algebraic version of the main theorem of [11]. We then derive a number of convexity inequalities and, in particular, prove a joint convexity of Jordan-algebraic version on quantum entropy. In section 4 we prove (as an application of one of the E. Lieb inequalities) the noncommutative Bernstein inequality developing some ideas of J. Tropp. In section 5 we prove the main theorem following the scheme of [4]. The section may be of an independent interest, since it shows a deep analogy of spectral theory in semi-simple complex Jordan algebras and **C***-algebras.

2 Jordan-algebraic Concepts

We adhere to the notation of an excellent book [7]. We do not attempt to provide a comprehensive introduction to Jordan algebras but rather describe a vocabulary with references to [7]. Let **F** be the field **R** or **C**. A vector space V over **F** is called an algebra over **F** if a bilinear mapping $(x, y) \to xy$ from $V \times V$ into V is defined. For an element x in V let $L(x) : V \to V$ be the linear map such that

$$L(x)y = xy.$$

An algebra V over **F** is a Jordan algebra if

$$xy = yx, x(x^2 y) = x^2(xy), \forall x, y \in V.$$

In other words, Jordan algebra is always commutative but typically not associative. In an algebra V one defines x^n recursively by $x^n = x \cdot x^{n-1}$. An algebra V is said to be power assotiative if $x^p \cdot x^q = x^{p+q}$ for any $x \in V$ and integers p, q.

Proposition 2.1. *A Jordan algebra is power associative. Besides,*

$$[L(x^p), L(x^q)] = 0, \forall x \in V,$$

and any positive integers p and q. (In other words, corresponding linear operators commute).

This is Proposition II.1.2 in [7]. We will always assume that the Jordan algebra has an identity element e (i.e. , $xe = x, \forall x \in V$). The power associativity of Jordan algebras allows one (among other things) to develop the spectral theory very similar to classical case of linear operators on finite dimensional spaces or finite-dimensional $\mathbf{C}^*$- algebras.

Let V be a finite-dimensional power associative algebra over **F** with an identity element e, and let $F[Y]$ denote the algebra over **F** of polynomials in one variable with coefficients in **F**. For $x \in V$ we define

$$\mathbf{F}[x] = \{p(x) : p \in \mathbf{F}[Y]\}.$$

A polynomial $p \in \mathbf{F}[Y]$ of minimal possible degree such that $p(x) = 0$ is called the minimal polynomial of x. Given $x \in V$, let $m(x)$ be the degree of the minimal polynomial of x. We define the rank of V as

$$r = \max\{m(x) : x \in V\}.$$

An element x is called regular if $m(x) = r$.

Proposition 2.2. *The set of regulr elements is open and dense in V.There exist polynomials $a_1, \ldots a_r$ on V such that the minimal polynomial of every regular element x is given by*

$$f(\lambda; x) = \lambda^r - a_1(x)\lambda^{r-1} + a_2(x)\lambda^{r-2} + \ldots + (-1)^r a_r(x).$$

The polynomials $a_1, \ldots a_r$ are unique and a_j is homogeneous of degree j.

This is Proposition II.2.1 in [7]. The coefficient $a_1(x)$ is called the trace of x and is denoted $tr(x)$ (in particular, trace is linear). The coefficient $a_r(x)$ is called the determinant of x and is denoted $\det(x)$. An element x is said to be invertible if there exists an element $y \in \mathbf{F}[x]$ such that $xy = e$. The set $\lambda \in \mathbf{F}$ such that $x - \lambda e$ is not invertible is called the spectrum of x and is denoted $spec(x)$.

Given $x \in V$, we define

$$P(x) = 2L(x)^2 - L(x^2).$$

The map P is called the quadratic representation of V. We denote $DP(x)y$ by $2P(x, y)$. Here $DP(x)y$ is the Frechet derivative of the map P at point $x \in V$ evaluated on $y \in V$. It is easy to see that

$$P(x, y) = L(x)L(y) + L(y)L(x) - L(xy), x, y \in V.$$

Proposition 2.3. *Let V be a finite-dimensional Jordan algebra over $\mathbf{F}$. An element $x \in V$ is invertible if and only if $P(x)$ is invertible. In this case*

$$P(x)x^{-1} = x, P(x)^{-1} = P(x^{-1}).$$

This is Proposition II.3.1 in [7].

Proposition 2.4. *Let $\mathcal{J}$ be the (open) set of invertible elements in V. The map $x \to x^{-1} : \mathcal{J} \to \mathcal{J}$ is Frechet differentiable and*

i)$D(x^{-1})u = -P(x^{-1})u, x \in \mathcal{J}, u \in V.$

ii) If x and y are invertible, then $P(x)y$ is invertible and $(P(x)y)^{-1} = P(x^{-1})y^{-1}$.

iii)

$$P(P(x)y) = P(x)P(y)P(x), \forall x, y \in V.$$

iv)

$$P(P(x)y, P(x)z) = P(x)P(y, z)P(x), \forall x, y, z \in V.$$

This is Proposition II.3.3 in [7]. A bilinear form β on V is called associative if

$$\beta(xy, z) = \beta(x, yz), \forall x, y, z \in V.$$

Proposition 2.5. *The symmetric bilinear forms $TrL(xy)$ and $tr(xy)$ are associative.*

This is Proposition II.4.3 in [7].

In case , where $\mathbf{F} = \mathbf{R}$ we consider an important class of Euclidean Jordan algebras. A Jordan algebra V over $\mathbf{R}$ is called Euclidean if $tr(x^2) > 0, \forall x \in V \setminus \{0\}$. An element $c \in V$ is called idempotent if $c^2 = c$. Two idempotents are orthogonal if $cd = 0$. A system of idempotents $c_1, \ldots c_k$ is a complete system of orthogonal idempotents If $c_i^2 = c_i, c_i c_j = 0, i \neq j$, and $c_1 + \ldots + c_k = e$.

Theorem 2.1. *Let V be an Euclidean Jordan algebra. Given $x \in V$,there exist unique real numbers $\lambda_1, \dots \lambda_k$, all distinct, and a unique complete system of orthogonal idempotents $c_1, \dots, c_k$ such that*

$$x = \lambda_1 c_1 + \dots + \lambda_k c_k.$$

In this case $spec(x) = \{\lambda_1, \dots, \lambda_k\}, c_1, \dots, c_k \in \mathbf{R}[x]$.

This is Theorem III.1.1 in [7].

An idempotent is primitive if it is non-zero and cannot be written as a sum of two non-zero idempotents. We say that $c_1, \dots, c_m$ is a complete system of orthogonal primitive idempotents, or Jordan frame, if each c_j is primitive idempotent and if

$$c_j c_k = 0, j \neq k, c_1 + \dots + c_m = e.$$

Note that in this case $m = r$ (rank of V).

Theorem 2.2. *Suppose V has rank r. Then for $x \in V$ there exists a Jordan frame $c_1, \dots c_r$ and real numbers $\lambda_1, \dots \lambda_r$ such that*

$$x = \sum_{j=1}^{r} \lambda_j c_j.$$

The numbers λ_j (with multiplicities) are uniquely determined by x. Furthermore,

$$det(x) = \prod_{j=1}^{r} \lambda_j, tr(x) = \sum_{j=1}^{r} \lambda_j.$$

This is Theorem III.1.2 in [7].

Given a function f which is defined at least on $spec(x)$, we can define

$$f(x) = \sum_{i=1}^{r} f(\lambda_i) c_i,$$

if $x = \sum_{i=1}^{r} \lambda_i c_i$. In particular,

$$\exp(x) = \sum_{i=1}^{r} \exp(\lambda_i) c_i, \ln x = \sum_{i=1}^{r} \ln \lambda_i c_i, \lambda_i > 0.$$

Convexity and differentiability ofsuch functions on Euclidean Jordan algebras have been studied in [**?**],[**?**] (see also [5]). We extensively use these properties in the paper.

Let

$$Q = \{x^2 : x \in V\}.$$

Theorem 2.3. *Let V be an Euclidean Jordan algebra.The interior Ω of Q is a symmetric (i.e. , self-dual, homogeneous) convex cone. Furthermore,Ω is the connected component of e in the set $\mathcal{J}$ of invertible elements, and also Ω is the set of elements x in V for which $L(x)$ is positive definite. In particular, the group of linear automorphisms $GL(\Omega)$ of Ω acts transitively on it. Moreover, $P(x) \in GL(\Omega)$ for any invertible x.*

This is Proposition III.2.2 in [7].

Let $c_1, \dots c_k$ be complete system of orthogonal idempotents. For each idempotent c, denote $V(c,0), V(c,1), V(c,1/2)$ the eigenspaces of $L(c)$ corresponding to eigenvalues $0, 1, 1/2$, respectively. Then $L(c_1), \dots, L(c_k)$ pairwise commute and

$$V = \bigoplus_{1 \le i \le j} V_{ij},$$

where $V_{ii} = V(c_i, 1), V_{ij} = V(c_i, 1/2) \cap V(c_j, 1/2)$. Such a decomposition of V corresponding to a complete system of orthogonal idempotents is called the Peirce decomposition. It is studied in detail in Section 1 of Chapter iV in [7]. A typical example of a Jordan algebra over a field $\mathbf{F}$ is the vector space of symmetric matrices over $\mathbf{F}$ with multiplication operation

$$A \cdot B = \frac{AB + BA}{2},$$

where on the right we have a usual matrix multiplication. In case $\mathbf{F} = \mathbf{R}$ we get an exmple of an Euclidean Jordan algebra.

3 Convexity inequalities

In this section we mostly follow the original paper [11] making necessary Jordan-algebraic adjustments.

Let V be an Euclidean Jordan algebra.

Theorem 3.1. *Let $0 \le p \le 1, k \in V$. The function $f_1 : \Omega \times \Omega \to \mathbf{R}$,*

$$f_1(a,b) = tr((P(k)a^p)b^{1-p})$$

is concave.

Here P is quadratic representation on V. This Theorem is proved in Section 5.

Lemma 3.1. *Given $k, u, v \in V$,*

$$tr(P(k)u)v) = \langle k, P(u,v)k \rangle = \langle P(k)u, v \rangle.$$

Proof. By definition: $P(k)u = 2L(k)^2 u - L(k^2)u$ and hence

$$tr((P(k)u)v) = \langle 2L(k)u, L(k)v \rangle - \langle L(k^2)u, v \rangle.$$

On the other hand,

$$\langle k, P(u,v)k \rangle = \langle k, (L(u)L(v) + L(v)L(u))k \rangle - \langle k, L(uv)k \rangle =$$

$$2\langle L(u)k, L(v)k \rangle - \langle k^2, uv \rangle = 2\langle L(k)u, L(k)v \rangle - \langle L(u)k^2, v \rangle =$$

$$2\langle L(k)u, L(k)v \rangle - \langle L(k^2)u, v \rangle.$$

□

Consider the function $\psi : [0,1] \to \mathbf{R}$,

$$\psi(p) = \langle P(k)a^p, b^{1-p} \rangle.$$

We obviously have:

$$\psi'(p) = \langle P(k)(a^p \ln a), b^{1-p} \rangle - \langle P(k)a^p, b^{1-p} \ln b \rangle.$$

In particular,

$$\psi'(1) = \langle P(k)(a \ln a), e \rangle - \langle P(k)a, lnb \rangle.$$

For $k = e$ we obtain:

$$\psi'(1) = tr(a \ln a - a \ln b).$$

Theorem 3.2. *The function* $(a, b) \to tr(a \ln a - a \ln b)$ *is convex on* $\Omega \times \Omega$.

Proof. $\psi(1) = \langle P(k)a, e \rangle$ is a linear function of (a, b), whereas the function $\psi(1+h)$ is concave in (a, b) for $-1 < h < 0$ by Theorem 3.1. Hence,

$$\Delta(h) = \frac{\psi(1+h) - \psi(1)}{h}$$

is convex for $-1 < h < 0$. Consequently,

$$\psi'(1) = lim \Delta(h), h \to 0^-$$

is convex. □

The function

$$\mathcal{D} : \Omega \times \Omega \to \mathbf{R}$$

$$\mathcal{D}(a, b) = tr(a \ln a - alnb - (a - b))$$

is called quantum relative entropy.

Corollary 3.1. *The quantum relative entropy is (jointly) convex on* $\Omega \times \Omega$.

Lemma 3.2. *Let* $\xi_l, \eta_l : [a, b] \to \mathbf{R}, \alpha_l \in \mathbf{R}, l = 1, \ldots, M$. *If*

$$\sum_{l=1}^{M} \alpha_l \xi_l(\lambda) \eta_l(\mu) \geq 0,$$

for all $\lambda, \mu \in [a, b]$. *Then for* $u, v \in V, spec(u) \subset [a, b], spec(v) \subset [a, b]$,

$$tr(\sum_{l=1}^{M} \alpha_l \xi_l(u) \eta_l(v)) \geq 0.$$

Proof. Let

$$u = \sum_{i=1}^{r} \lambda_i c_i, v = \sum_{i=1}^{r} \mu_i d_i$$

be spectral decompositions u, v,respectively (see Theorem [**?**]). Then

$$\xi_l(u) = \sum_{i=1}^{r} \xi_l(\lambda_i) c_i, \eta_l(v) = \sum_{i=1}^{r} \eta_l(\mu_i) d_i$$

and, consequently

$$tr(\sum_{l}^{M} \alpha_l \xi_l(u) \eta_l(v)) = \sum_{i=1}^{r} \sum_{j=1}^{r} \langle c_i, d_j \rangle \sum_{l=1}^{M} \alpha_l \xi_l(\lambda_i) \eta_l(\mu_j) \geq 0,$$

since $\langle c_i, d_j \rangle \geq 0$, for all i, j. □

Proposition 3.1.

$$\mathcal{D}(a, b) \geq 0, \forall (a, b) \in \Omega \times \Omega.$$

Proof. The function $\phi(\lambda) = \lambda \ln \lambda$ is convex $for \lambda > 0$. Hence,

$$\phi(\lambda) - \phi(\mu) \geq \phi'(\mu)(\lambda - \mu),$$

for any $\lambda, \mu > 0$. Consequently,

$$\lambda \ln \lambda - \lambda \ln \mu - (\lambda - \mu) \geq 0, \lambda, \mu > 0.$$

By Lemma 3.2 $\mathcal{D}(a, b) \geq 0$. □

Proposition 3.2. *Let $b \in \Omega$. Then*

$$tr(b) = \max\{tr(a \ln b - a \ln a + a) : a \in \Omega\}.$$

Proof. Since $\mathcal{D}(a, b) \geq 0, \forall (a, b) \in \Omega \times \Omega$, we have:

$$tr(b) \geq tr(alnb - alna + a), \forall a \in \Omega.$$

But for $a = b$ we obtain the equality. □

Theorem 3.3. *Given $h \in V$, the function $f_2 : \Omega \to \mathbf{R}$,*

$$f_2(a) = tr(exp(h + \ln a))$$

is concave on Ω.

Proof. Take $b = \exp(h + lna)$ in Proposition 3.2. Then:

$$tr(exp(h + lna)) = max\{tr(v(h + lna) - vlnv + v) : v \in \Omega\} =$$

$$max\{tr(vh) - \mathcal{D}(v, a) + tr(a) : v \in \Omega\}. \quad (1)$$

Since the function $\mathcal{D}(v, a)$ is jointly convex in (v, a), (1) shows that $f_2(a)$ is concave on Ω. □

Proposition 3.3. *For $a \in \Omega$*

$$\ln a = \int_0^{+\infty} (\frac{e}{1+\tau} - (a+\tau e)^{-1}) d\tau. \tag{2}$$

Proof. Let

$$a = \sum_{i=1}^{r} \lambda_i c_i$$

be a spectral decomposition of a.See Theorem 2.2. Then

$$\ln a = \sum_{i=1}^{r} \ln \lambda_i c_i.$$

On the other hand, for $R > 0$

$$\int_0^R (\frac{1}{1+\tau} - (a+\tau e)^{-1}) d\tau = \sum_{i=1}^{r} [\int_0^R (\frac{1}{1+\tau} - \frac{1}{\lambda_i + \tau}) d\tau] c_i.$$

But

$$\int_0^R (\frac{1}{\tau+1} - \frac{1}{\lambda_i+\tau}) d\tau = \ln(\frac{1+R}{\lambda_i+R}) + \ln \lambda_i.$$

Taking limit when $R \to +\infty$, we obtain the result. □

The expression (2) allows one to compute (using Proposition 2.4 i)) the Frechét derivative of $\ln a$:

$$D \ln(a) h = [\int_0^{+\infty} P(a+\tau e)^{-1} d\tau] h, h \in V.$$

Following the original paper of E.Lieb [11], we will introduce notation T_a for the linear operator

$$T_a(h) = [\int_0^{+\infty} P(a+\tau e)^{-1} d\tau] h.$$

Note that

$$\langle T_a(h), h \rangle = D^2 \phi(a)(h,h),$$

where $\phi(a) = tr(a \ln a), a \in \Omega$,i.e. T_a is the Hessian of the quantum entropy. In this connection, it is important to calculate the inverse of T_a. Obviously, T_a is positive definite for any $a \in \Omega$.

Proposition 3.4.

$$T_a^{-1} = \int_0^1 P(a^{1-\tau}, a^{\tau}) d\tau.$$

Proof. Let

$$a = \sum_{i=1}^{k} \lambda_i c_i$$

be the spectral decomposition of a such that $\lambda_1 > \lambda_2 > \ldots > \lambda_k > 0$. (see Theorem 2.1). Let, further,

$$V = \bigoplus_{1 \le i \le j \le k} V_{ij}$$

be the corresponding Peirce decomposition. Then $P(a + \tau e)^{-1}$ restricted to V_{ij} acts by multiplication by

$$\frac{1}{(\lambda_i + \tau)(\lambda_j + \tau)}.$$

Hence, for $h \in V_{ij}$:

$$T_a(h) = \int_0^{+\infty} \frac{d\tau}{(\lambda_i + \tau)(\lambda_j + \tau)} h =$$

$$\frac{\ln \lambda_j - \ln \lambda_i}{\lambda_j - \lambda_i} h, i \neq j, \tag{3}$$

$$T_a(h) = \frac{h}{\lambda_i}, i = j.$$

Consider

$$I_a = \int_0^1 P(a^{1-\tau}, a^{\tau}) d\tau.$$

For $h \in V_{ij}$;

$$P(a^{1-\tau}, a^{\tau})h = [2L(a^{1-\tau})L(a^{\tau}) - L(a)]h = \frac{\lambda_i^{1-\tau}\lambda_j^{\tau} + \lambda_i^{\tau}\lambda_j^{1-\tau}}{2} h.$$

Hence,

$$I_a(h) = \frac{\lambda_j - \lambda_i}{\ln \lambda_j - \ln \lambda_i} h, i \neq j, \tag{4}$$

$$I_a(h) = \lambda_i h, i = j.$$

Comparing this with (3), we conclude that $I_a = T_a^{-1}$. □

Proposition 3.5. *The function* $q : V \times \Omega \to \mathbf{R}$,

$$q(a, h) = \langle h, T_a(h) \rangle$$

is jointly convex in (h, a).

Proof. Fix $a, b \in \Omega, 0 < \lambda < 1$. Denote $\lambda a + (1-\lambda)b$ by c. Consider quadratic forms:

$$\Gamma_1(u,v) = \lambda q(a,u) + (1-\lambda)q(b,v),$$

$$\Gamma_2(u,v) = q(c, \lambda u + (1-\lambda)v)$$

on $V \times V$. Note that Γ_1 is positive definite. Consider an optimization problem

$$\phi(u,v) = \frac{\Gamma_2(u,v)}{\Gamma_1(u,v)} \to \max,$$

$(u,v) \in V \times V \setminus \{0,0\}$. Let γ be the maximal value of ϕ. If $\gamma \leq 1$ (for all choices of $a, b \in \Omega, 0 < \lambda < 1$), then the result follows. The stationary points (u^*, v^*) of the optimization problem should satisfy the equation:

$$D\Gamma_2(u^*, v^*)(g,h) - \phi(u^*, v^*) D\Gamma_1(u^*, v^*)(g,h) = 0$$

for all $(g,h) \in V \times V$. This leads to equations:

$$\langle \lambda T_c(\lambda u^* + (1-\lambda)v^*), g\rangle + \langle (1-\lambda)T_c(\lambda u^* + (1-\lambda)v^*), h\rangle =$$

$$\phi(u^*, v^*)(\langle \lambda T_a(u^*), g\rangle + \langle (1-\lambda)T_b(v^*), h\rangle),$$

and hence

$$\Delta = T_c(w) = \gamma T_a(u^*), T_c(w) = \gamma T_b(v^*), \tag{5}$$

where $\gamma = \phi(u^*, v^*), w = \lambda u^* + (1-\lambda)v^*$. If $\gamma = 0$ (and consequently less or equal than one), we are done. If not,

$$u^* = \frac{1}{\gamma} T_a^{-1}(\Delta), v^* = \frac{1}{\gamma} T_b^{-1}(\Delta)$$

by (5). Note that $\Delta \neq 0$ (otherwise, $(u^*, v^*) = (0,0)$.). Since $T_c^{-1}(\Delta) = \lambda u^* + (1-\lambda)v^*$, we obtain:

$$\lambda T_a^{-1}(\Delta) + (1-\lambda)T_b^{-1}(\Delta) = \gamma T_c^{-1}(\Delta). \tag{6}$$

By Proposition 3.4 the relationship (5) means:

$$\int_0^1 [\lambda P(a^\tau, a^{1-\tau}) + (1-\lambda)P(b^\tau, b^{1-\tau}) - \gamma P(c^\tau, c^{1-\tau})] \Delta d\tau = 0. \tag{7}$$

However,

$$\langle \Delta, [\lambda P(a^\tau, a^{1-\tau}) + (1-\lambda)P(b^\tau, b^{1-\tau})]\Delta\rangle \leq \langle \Delta, P(c^\tau, c^{1-\tau})\Delta\rangle,$$

$0 \leq \tau \leq 1$ by Theorem 3.1. Hence, $\gamma \leq 1$. □

Recall that

$$T_a(h) = D\ln(a)h = [\int_0^{+\infty} P(a+\tau e)^{-1} d\tau]h.$$

Since

$$DP(a)g = \frac{1}{2}P(a,g), a, g \in V,$$

we can calculate the second Frechet derivative of ln using the chain rule. Let

$$\psi(a) = (a+\tau e)^{-1}, \phi(a) = P(\psi(a))h.$$

Then

$$D\phi(a)g = (DP(\psi(a))D\psi(a)g)h = 2P(\psi(a), -P(a+\tau e)^{-1}g)h =$$

$$-2P((a+\tau e)^{-1}, P(a+\tau e)^{-1}g)h =$$

$$-2P(P((a+\tau e)^{-1/2})e, P(a+\tau e)^{-1/2}P(a+\tau e)^{-1/2}g)h =$$

$$-2P(a+\tau e)^{-1/2}P(e, P(a+\tau e)^{-1/2}g)P(a+\tau e)^{-1/2}h =$$

$$-2P(a+\tau e)^{-1/2}[L(P(a+\tau e)^{-1/2})g)(P(a+\tau e)^{-1/2}h)].$$

Hence,

$$D^3(tr(a\ln a))(h_1,h_2,h_3) = \langle h_1, D^2\ln(a)(h_2,h_3\rangle = \tag{8}$$

$$-2\int_0^{+\infty} tr[(M(a,\tau)h_1)(M(a,\tau)h_2)(M(a,\tau)h_3)]d\tau,$$

$a \in \Omega, h_1, h_2, h_3 \in V, M(a,\tau) = P(a+\tau e)^{-1/2}$.

Lemma 3.3. *Let C be a convex cone in a vector space and let $F : C \to \mathbf{R}$ be a convex function such that*

$$\lim \frac{F(a+tb) - F(a)}{t}, t \to 0^+,$$

exists and is denoted by $G(a,b)$ for all $a, b \in C$. Assume that F is homogeneous of order 1, i.e. , $F(\lambda a) = \lambda F(a), a \in C, \lambda > 0$. Then

$$G(a;b) \leq F(b), \forall a, b \in C.$$

Proof. For $t > 0, a, b \in C$:

$$F(a+tb) = F((1+t)(\frac{a}{1+t}+\frac{t}{1+t}b)) = (1+t)F(\frac{a}{1+t}+\frac{t}{1+t}b) \leq (1+t)(\frac{1}{1+t}F(a)+\frac{t}{1+t}F(b)) =$$

$$F(a) + tF(b).$$

Hence,

$$\frac{F(a+tb) - F(a)}{t} \leq F(b).$$

□

Note that the function $q(a,h) = \langle h, T_a(h)\rangle$ is homogeneous of order 1 on the-cone $\Omega \times V$. Indeed,

$$q(\lambda a, \lambda h) = \int_0^{+\infty} \langle \lambda h, P(\lambda a + \tau e)^{-1} \lambda h\rangle d\tau =$$

$$\int_0^{+\infty} \langle h, P(a + \frac{\tau}{\lambda} e)^{-1} h\rangle d\tau = \lambda q(a,h).$$

The last equality is obtained by making the change of variables $\tilde{\tau} = \frac{\tau}{\lambda}$. Applying Lemma 3.3 to q, we obtain

$$Dq(a,h)(b,g) = 2\langle T_a(h), g\rangle - \langle R_a(h), b\rangle \leq \langle g, T_b(g)\rangle \tag{9}$$

for all $a, b \in \Omega, g, h \in V$. Here

$$R_a(h) = 2\int_0^{+\infty} P(a+\tau e)^{-1/2}[P(a+\tau e)^{-1/2}h]^2 d\tau,$$

(see (8)). The relationship (9) is crucial in [11] for proving various convexity inequalities.

Since the exponential is the inverse of logarithm, we have:

$$\ln(exp(a)) = a, a \in V.$$

Using the chain rule, we obtain:

$$D(ln(\exp(a))(D(\exp(a))h = h,$$

$h, a \in V$. Consequently,

$$T_{\exp(a)}(D(\exp(a))h) = h,$$

or

$$D(\exp(a))h = T^{-1}_{\exp(a)}(h) =$$

$$[\int_0^1 P(\exp(a\tau),\ exp(a(1-\tau)))d\tau](h), \tag{10}$$

where in the last equality we used Proposition 3.4.

We say that $a, b \in V$ commute, if

$$[L(a), L(b)] = 0.$$

Proposition 3.6. *The elements $a, b \in V$ commute if and only if, there exists a Jordan frame $c_1, \ldots, c_r$ in V and $\lambda_i, \mu_i \in \mathbf{R}, i = 1, \ldots r$, such that*

$$a = \sum_{i=1}^r \lambda_i c_i, b = \sum_{i=1}^r \mu_i c_i.$$

This is Lemma X.2.2 in [7]. For a detailed discussion of commutavity in the above sense see [12]. The following Theorem is a Jordan-algebraic version of the Golden-Thomson inequality.

Theorem 3.4. *Let $u, v, w \in V$. Then*

$$tr(\exp(w)T_{\exp(-u)}(\exp(v))) \geq tr(\exp(u+v+w)).$$

If u commutes with v, then

$$tr(\exp(u)\exp(v)\exp(w)) \geq tr(\exp(u+v+w)).$$

Remark 3.1. *Recall that*

$$tr(u(vw)) = tr((uv)w), \forall u, v, w \in V.$$

See Proposition 2.5.

Proof. Let $a = \exp(-u), b = \exp(v), l = u + w$. By Theorem 3.3 the function $\phi_l : \Omega \to \mathbf{R}$,

$$\phi_l(c) = -tr(\exp(l + \ln c))$$

is convex. It is also clear that $\phi_l(\lambda c) = \lambda \phi_l(c)$ for any $\lambda > 0$. Hence, by Lemma 3.3:

$$D\phi_l(a)b \leq \phi_l(b).$$

Note that:

$$D\phi_l(a)b = -tr(T_a(b)\exp(l + \ln(a)).$$

Substituting expressions for a, b, l we obtain:

$$tr(exp(u+v+w)) \leq tr(\exp(w)T_{\exp(-u)}(\exp(v))).$$

If u commutes with v, then computing the corresponding integral in common for u and v Jordan frame, we obtain:

$$T_{\exp(-u)}(\exp(v)) = \int_0^{+\infty} (\exp(-u) + \tau e)^{-2} \exp(v) d\tau = \exp(u)\exp(v).$$

□

4 Noncommutative Bernstein inequality

Let V be an Euclidean Jordan algebra. Suppose that $v_1, \ldots, v_M$ are independent random variables on a probability space X (with probability measure Pr defined on σ-algebra $\mathcal{A}$ of subsets of X) with values in V. We denote by $\mathcal{E}$ the mathematical expectation with respect to Pr. In other words, if $v : X \to V$ is a random variable, then

$$\mathcal{E}[v] = \int_X v(\omega) dPr(\omega).$$

Given $v \in V$ with $spec(v) = \{\lambda_1, \ldots \lambda_s\}$ and $\lambda_1 > \lambda_2 > \ldots > \lambda_s$, then $\lambda_{max}(v) := \lambda_1$ and

$$\|v\|_\infty := \max\{\lambda_{max}(v), \lambda_{max}(-v)\}.$$

Note that $\|v\|_\infty$ defines a norm on V invariant under the action of the group of automorphisms of V (see [**?**],[5]). In this section we prove the following result.

Theorem 4.1. *Let $v_1, \dots v_M : X \to V$ be independent random variables such that $\mathcal{E}[v_i] = 0, i = 1, \dots M$. Suppose that*

$$\lambda_{max}(v_i) \leq K$$

almost surely for all $i = 1, \dots M$. Here K is a fixed positive number. Denote

$$\sigma^2 = \| \sum_{i=1}^{M} \mathcal{E}[v_i^2] \|_\infty. \tag{11}$$

Then, for $t > 0$

$$Pr(\lambda_{max}(\sum_{i=1}^{M} v_i) \geq t) \leq r \exp(-\frac{\sigma^2}{K^2} h(\frac{Kt}{\sigma^2})) \leq$$

$$r \exp(-\frac{t^2}{\sigma^2 + Kt/3}).$$

Here r is the rank of V and

$$h(\lambda) = (1 + \lambda) \ln(1 + \lambda) - \lambda, \lambda \geq 0.$$

In case where V is the Jordan algebra of complex Hermitian matrices, this result is due to [14]. Note that we do not assume that V is simple. One can even consider infinite-dimensional spin-factors (as in [3]) as irreducible components. It does not effect the proof.

Corollary 4.1. *Let $v_1, \dots v_M : X \to V$ be independent random variables such that $\mathcal{E}[v_i] = 0, i = 1, \dots M$. Suppose that*

$$\|v_i\|_\infty \leq K$$

almost surely forall $i = 1, 2, \dots M$. Then, for $t > 0$

$$Pr(\| \sum_{i=1}^{M} v_i \|_\infty \geq t) \leq 2r \exp(-\frac{\sigma^2}{K^2} h(\frac{Kt}{\sigma^2})) \leq$$

$$2r \exp(-\frac{t^2}{\sigma^2 + Kt/3}).$$

In our proof ofTheorem 4.1 we follow [8], making necessary Jordan-algebraic adjustments.

Proposition 4.1. *Let $v : X \to V$ be a random variable. Then, given $t > 0$,*

$$Pr(\lambda_{max}(v) \geq t) \leq inf\{\exp(-\theta)\mathcal{E}[tr(\exp(\theta v)] : \theta > 0\}.$$

Proof. We have:

$$Pr(\lambda_{max}(v) \geq t) = Pr(\exp(\lambda_{max}(\theta v)) \geq \exp(\theta t)) \leq$$

$$\exp(-\theta t)\mathcal{E}[\exp(\lambda_{max}(\theta v)].$$

The last inequality is just the standard Markov inequality. Furthermore, given $\omega \in X$,

$$\exp(\lambda_{max}(\theta v(\omega))) = \lambda_{max}(\exp(\theta v(\omega))) \leq$$

$$\sum_{j=1}^{r} \lambda_j(\exp(\theta v(\omega))) = tr(\exp(\theta v(\omega))).$$

Here $\lambda_j(\exp(\theta v(\omega))$ are eigenvalues of $\exp(\theta v(\omega))$. Hence,

$$Pr(\lambda_{\max}(v) \geq t) \leq \exp(-\theta t)\mathcal{E}[tr \exp(\theta v)],$$

for every $\theta > 0$. □

Proposition 4.2. *Let $h \in V$ and $v : X \to V$ be a random variable. Then*

$$\mathcal{E}[tr \exp(h + v)] \leq tr \exp(h + \ln(\mathcal{E}[\exp(v)])).$$

Proof. By Theorem 3.3 the function $\phi_h : \Omega \to \mathbf{R}$,

$$\phi_h(a) = tr \exp(h + \ln a)$$

is concave. By Jensen's inequality

$$\mathcal{E}[\phi_h(\exp v)] \leq \phi_h(\mathcal{E}[\exp v]),$$

i.e.,

$$\mathcal{E}[tr \exp(h + v)] \leq tr \exp(h + \ln \mathcal{E}[\exp(v)]).$$

□

Proposition 4.3. *Let $v_1, \ldots, v_M : X \to V$ be independent random variables. Then for any $\theta \in \mathbf{R}$*

$$\mathcal{E}[tr(\exp(\theta \sum_{i=1}^{M} v_i))] \leq tr \exp(\sum_{i=1}^{M} \ln \mathcal{E}[\exp(\theta v_i)]).$$

Proof. Without loss of generality we may assume $\theta = 1$. Let

$$h_i = \ln \mathcal{E}[\exp(v_i)], i = 1, \ldots M.$$

Since v_i are independent, we can write $\mathcal{E}_{v_i}$ for the expectation with respect to v_i (i.e., the expectation conditional on $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots v_M$). Using Fubini theorem, we obtain:

$$\Delta = \mathcal{E}[tr(\exp(\sum_{i=1}^{M} v_i))] = \mathcal{E}_{v_1}\mathcal{E}_{v_2} \ldots \mathcal{E}_{v_M}[tr \exp(\sum_{i=1}^{M-1} v_i + v_M)].$$

By Proposition 4.2

$$\Delta \leq \mathcal{E}_{v_1} \dots \mathcal{E}_{v_{M-1}}[tr\exp(\sum_{i=1}^{M-1} v_i + \ln \mathcal{E}[\exp(v_M)])] =$$

$$\mathcal{E}_{v_1} \dots \mathcal{E}_{V_{M-1}}[tr\exp(\sum_{i=1}^{M-2} v_i + h_M + v_{M-1})] \leq$$

$$\mathcal{E}_{v_1} \dots \mathcal{E}_{v_{M-2}}[tr\exp(\sum_{i=1}^{M-2} v_i + h_M + h_{M-1})] \leq \dots \leq tr\exp(\sum_{i=1}^{M} h_i),$$

where we repeatedly used Proposition 4.2. □

Proposition 4.4. *Let $v_1, \dots v_M \to V$ be independent random variables. Suppose that there exists a function $g : (0, \infty) \to [0, \infty)$ and fixed $u_1, \dots u_M \in V$ such that*

$$\mathcal{E}[\exp(\theta v_i)] \preceq \exp(g(\theta)u_i), i = 1, \dots M,$$

($u \preceq v$ for $u, v \in V$ means that $v - u \in \bar{\Omega}$). Let $\rho = \lambda_{max}(\sum_{i=1}^{M} u_i)$. Then

$$Pr(\lambda_{max}(\sum_{i=1}^{M} v_i) \geq t) \leq r \inf\{\exp(-\theta t + g(\theta)\rho) : \theta > 0\}.$$

Lemma 4.1. *Let $x, y \in V$ and $x \succeq y \succ 0, i.e., x - y \in \bar{\Omega}, y \in \Omega$. Then*

$$\ln x \succeq \ln y$$

.

Proof. By Proposition 3.3

$$\ln a = \int_0^{+\infty} [\frac{e}{1+\tau} - (a + \tau e)^{-1}] d\tau, a \in \Omega.$$

Hence, it suffices to show that

$$(x + \tau e)^{-1} \preceq (y + \tau e)^{-1} \tag{12}$$

for any $\tau \geq 0$. But $x - y \in \bar{\Omega}$ is equivalent to $(x + \tau e) - (y + \tau e) \in \bar{\Omega}$. Consequently,

$$P(y + \tau e)^{-1/2}(x + \tau e) - e \in \bar{\Omega}, \tag{13}$$

since $P(u) \in GL(\Omega), \forall u \in V$.(see Theorem 2.3. Similarly, (12) is equivalent to

$$e - P(y + \tau e)^{1/2}(x + \tau e)^{-1} \in \bar{\Omega}. \tag{14}$$

Let

$$P(y + \tau e)^{-1/2}(x + \tau e) = \sum_{i=1}^{r} \lambda_i c_i, \lambda_i > 0, i = 1, \dots r,$$

be the spectral decomposition. Then (13) is equivalent to $\lambda_i \geq 1, i = 1, \dots, r$ whereas (14) is equaivalent to $1 - 1/\lambda_i \geq 0, i = 1, \dots r$, since $P(y + \tau e)^{-1/2}(x + \tau e) = P(y + \tau e)^{-1/2}(x + \tau e) = P(y + \tau e)^{1/2}(x + \tau e)^{-1}]^{-1}$by Proposition 2.4 ii). However, these are the same conditions. □

Lemma 4.2. *If $x, y \in V, x \succeq y$, then*

$$tr \exp(x) \succeq tr \exp(y). \tag{15}$$

Proof. Each Euclidean Jordan algebra is a direct sum of simple Euclidean Jordan algebras, i.e. ,

$$V = V_1 \oplus \ldots \oplus V_s,$$

where V_i are simple Euclidean Jordan algebras. Note that

$$\Omega = \Omega_1 \oplus \ldots \oplus \Omega_s,$$

where Ω_l is the cone of invertible squares in V_l. Moreover, if

$$x = \sum_{i=1}^{s} x_l,$$

the corresponding decomposition of $x \in V$, then

$$tr(\exp(x) = \sum_{l=1}^{s} tr(\exp(x_l).$$

Thus, to prove (15) it suffices to consider the case where V is simple. Let $spec(x)$ (resp. $spec(y)) = \{\lambda_1(x), \ldots \lambda_r(x)\}$ (resp. $\{\lambda_1(y), \ldots \lambda_r(y)\}$, where $\lambda_1(x) \geq \lambda_2(x) \ldots \lambda_r(x)$ (resp. $\lambda_1(y) \geq \lambda_2(y) \ldots \lambda_r(y)$). Then $x \succeq y$ implies $\lambda_i(x) \geq \lambda_i(y), i = 1, \ldots r$ (see [10]). Consequently,

$$tr \exp(x) = \sum_{i=1}^{r} \exp(\lambda_i(x) \geq \sum_{i=1}^{r} \exp(\lambda_i(y)) = tr(\exp(y).$$

□

We are now in position to prove Proposition 4.4.

Proof. By Propositions 4.1,4.3

$$\Delta = Pr(\lambda_{\max}(\sum_{i=1}^{M} v_i) \geq t) \leq \inf\{\exp(-\theta t) tr \exp(\sum_{i=1}^{M} \ln \mathcal{E}[\exp(\theta v_i)] : \theta > 0\}.$$

Using Lemma 4.1 and Lemma 4.2, we obtain:

$$\Delta \leq \inf\{\exp(-\theta t) tr \exp(\sum_{i=1}^{M} g(\theta) u_i) : \theta > 0\}.$$

Now,

$$tr(\exp(g(\theta) \sum_{i=1}^{M} u_i) \leq r \lambda_{max}(\exp(g(\theta) \sum_{i=1}^{M} u_i)) =$$

$$r \exp(g(\theta) \lambda_{max}(\sum_{i=1}^{r} u_i)) = r \exp(\rho g(\theta)).$$

□

Fix $\theta > 0$ and consider the function

$$f(\lambda) = \lambda^{-2}(\exp(\theta\lambda) - \theta\lambda - 1), \tag{16}$$

for $\lambda \neq 0, f(0) = \theta^2/2$.

Lemma 4.3. *Function f is monotonically nondecreasing on* $\mathbf{R}$.

For a proof see [8], p. 222.

Hence, $f(\lambda) \leq f(1)$ if $\lambda \leq 1$. We will assume that in the formulation of Theorem 4.1 $K = 1$ (otherwise, substitute v_i by v_i/K). Since all eigenvaues of v_i are bounded by one from above, we have:

$$f(v_i(\omega)) \leq f(1)e, \omega \in X, i = 1, \ldots M.$$

The identity $\exp(\theta\lambda) = 1 + \theta\lambda + \lambda^2 f(\lambda)$ implies

$$\exp(\theta v_i(\omega)) = e + \theta v_i(\omega) + v_i^2(\omega) f(v_i(\omega)) =$$

$$e + \theta v_i(\omega) + P(v_i(\omega)) f(v_i(\omega)) \preceq e + \theta v_i(\omega) + P(v_i(\omega)) f(1)e =$$

$$e + \theta v_i(\omega) + f(1) v_i^2(\omega).$$

Hence,

$$\Delta = \mathcal{E}[\exp(\theta v_i)] \preceq 1 + f(1)\mathcal{E}[v_i^2] \preceq \exp(f(1)\mathcal{E}[v_i^2]),$$

where we used an obvious inequality $\exp(\lambda \geq 1 + \lambda, \lambda \in \mathbf{R}$. Recalling the definition of f (see (16), we obtain:

$$\Delta \preceq \exp((\exp(\theta) - \theta - 1)\mathcal{E}[v_i^2]), i = 1, \ldots M.$$

By Proposition 4.4:

$$Pr(\lambda_{\max}(\sum_{i=1}^{M} v_i) \geq t) \leq r \inf\{\exp(-\theta t + g(\theta)\sigma^2) : \theta > 0\}, \tag{17}$$

where $g(\theta) = \exp(\theta) - \theta - 1$. Here

$$\sigma^2 = \lambda_{\max}(\sum_{i=1}^{M} \mathcal{E}[v_i^2]) = \| \sum_{i=1}^{M} \mathcal{E}[v_i^2] \|_\infty.$$

Lemma 4.4. *Let* $h(\lambda) = (1 + \lambda)\ln(1 + \lambda) - \lambda, \lambda > -1$,

$$g(\theta) = \exp(\theta) - \theta - 1.$$

Then, for $\mu > 0, \eta \geq 0$

$$\inf\{\theta\eta + g(\theta)\mu : \theta > 0\} = -\mu h(\eta/\mu)$$

and

$$h(\lambda) \geq \frac{\lambda^2/2}{1 + \lambda/3}, \lambda \geq 0.$$

For a proof see [8], Lemma 8.21 . Combining (17) and Lemma 4.4, we obtain Theorem 4.1.

5 Proof of the main theorem

Several proofs of the original version of Theorem 3.1 are known (see [13],[11],[4],[6]). However, it seems none of them admits an immediate generalization to Jordan-algebraic setting. We have chosen an approach developed in[4] mostly for the case of finite-dimensional $\mathbf{C}^*$-algebras. $\mathbf{C}^*$-algebras are associative but not necessarily commutative, whereas Jordan algebras are commutative but typically nonassociative. However, both classes are power associative which makes spectral theory quite similar for both of them. We provide (almost) all details for the Jordan-algebraic case. Let V be an Euclidean Jordan algebra. We define its complexification $V^{\mathbf{C}}$ as the set V with the following operations:

$$\alpha + i\beta)(x, y) = (\alpha x - \beta y, \beta x + \alpha y), \alpha, \beta \in \mathbf{R}, i = \sqrt{-1},$$

$$(x, y) + (x', y') = (x + x', y + y').$$

Then $V^{\mathbf{C}}$ is a vector space over $\mathbf{C}$ (and hence it makes sense to talk about holomorpic functions on open subsets of $V^{\mathbf{C}}$. One considers V as a subset of $V^{\mathbf{C}}$ under the identification $x \sim (x, 0)$. The elements of $V^{\mathbf{C}}$ can be written as $x + iy$ with $x, y \in V$. The vector space $V^{\mathbf{C}}$ has a distinguished conjugation operation:

$$\overline{x + iy} = x - iy.$$

We define on $V^{\mathbf{C}}$ the structure of Jordan algebra over $\mathbf{C}$:

$$(x + iy)((x' + iy') = (xx' - yy') + i(yx' + xy').$$

Each $\mathbf{R}$-linear map $A : V \to V$ can be extended to $\mathbf{C}$-linear map:

$$A(x + iy) = A(x) + iA(y).$$

Recall that on V there exists the canonical scalar product:

$$\langle x, x' \rangle = tr(xx').$$

We can extend it to $\mathbf{C}$-bilinear form on $V^{\mathbf{C}}$:

$$\langle x + iy, x' + iy' \rangle = (\langle x, x' \rangle - \langle y, y' \rangle) + i(\langle x, y' \rangle + \langle y, x' \rangle).$$

We define a Hermitian scalar product on $V^{\mathbf{C}}$:

$$\langle\langle w, w' \rangle\rangle = \langle w, \bar{w}' \rangle, w, w' \in V^{\mathbf{C}}.$$

Then

$$\|w\| = \langle\langle w, w \rangle\rangle^{1/2}.$$

Consider

$$T_\Omega = V + i\Omega \subset V^{\mathbf{C}}.$$

Each $w \in V^{\mathbf{C}}$ has a unique representation

$$w = \Re w + i\Im w$$

with $\Re w, \Im w \in V$. Hence, $w \in T_\Omega$ if and only if $\Im w \in \Omega$.

Theorem 5.1. *The map $w \to -w^{-1}$ is an involutive holomorphic automorphism of T_Ω, having ie as its unique fixed point. In particular, $w \in T_\Omega$ implies w is inverible and $-\Im(w^{-1}) \in \Omega$.*

This is Theorem X.1.1 in [7]. Let $w \in T_\Omega$ and $\lambda \in \mathbf{C}, \Im\lambda \leq 0$. Then $\Im(w-\lambda e) = \Im w - (\Im\lambda)e \in \Omega$,i.e., $w - \lambda e \in T_\Omega$ and consequently is invertible. In particular,

$$spec(w) \subset \{\lambda \in \mathbf{C} : \Im\lambda > 0\}. \tag{18}$$

Let $\mathbf{R}_- = \{\lambda \in \mathbf{C} : \Im\lambda = 0, \Re\lambda \leq 0\}$, and

$$U = \{v \in V^{\mathbf{C}} : spec(v) \subset \mathbf{C} \setminus \mathbf{R}_-\}.$$

Theorem 5.2. *Let $f : U \to \mathbf{C}$ be a holomorphic function with the following properties:*
(i)$\Im \underline{f(v)} \geq 0$, if $\Im v \in \Omega$;
(ii) $\overline{f(v)} = f(\bar{v}), v \in U$;
(iii) $f(\rho v) = \rho f(v), \rho > 0, v \in U$.
Then the restriction of f on Ω is concave. More precisely, let $a \in \Omega, h \in V$. Then for sufficiently small real t and integers $n \geq 1$

$$\frac{d^{2n}\phi}{dt^{2n}}(t) \leq 0,$$

where $\phi(t) = f(a + th)$.

Remark 5.1. *A more general version of this Theorem is considered in [4] in $\mathbf{C}^*$-algebras settings. The corresponding Jordan-algebraic counterpart is also true.*

Proof. Given $a \in \Omega, h \in V$, consider two holomorphic functions

$$F(\lambda) = f(a + \lambda h), G(\lambda) = f(h + \lambda a).$$

Note that F is defined for $\lambda \in \mathbf{C}$ such that $| \lambda |< 1/\tau$ and G is defined for $\lambda \in \mathbf{C}$ such that $\Re\lambda > \tau$ or $\Im\lambda \neq 0$. Here

$$\tau = \|h\|\|a^{-1}\|.$$

Indeed, consider

$$\Delta(\mu) = a + \lambda h - i\mu e, \mu \in \mathbf{R}.$$

Then $\Re\Delta(\mu) = a + \Re\lambda h = P(a^{1/2})(e + \Re\lambda P(a^{-1/2})h)$. Let $\lambda_1, \dots \lambda_r$ be (real!) eigenvalues of $P(a^{-1/2})h$. Then $1+\Re\lambda_1, \dots 1+\Re\lambda\lambda_r$ are eigenvalues of $P(a^{-1/2})\Re\Delta(\mu)$. If $| \Re\lambda || \lambda_j |< 1$ for all j, then $\Re\Delta(\mu) \in \Omega$ and hence $\Delta(\mu)$ is invertible by Theorem 5.1. Hence, $a + \lambda h \in U$. The conditions $| \Re\lambda || \lambda_j |< 1$ for all j are satisfied if

$$| \lambda |< \frac{1}{\max\{| \lambda_j |: j \in [1, r]\}}.$$

But

$$\max\{| \lambda_j |: j \in [1, r]\} \leq \|P(a^{-1/2})h\| \leq \|P(a^{-1/2})\|\|h\| \leq \|h\|\|a^{-1}\|$$

Hence, $a + \lambda h \in U$ if $| \lambda |< \frac{1}{\|h\|\|a^{-1}\|}$. Similarly, for $\lambda a + h$, consider

$$\Delta_1(\mu) = \lambda a + h - i\mu e, \mu \in \mathbf{R}.$$

Since
$$\Re\Delta_1(\mu) = \Re\lambda a + h = P(a^{1/2})(\Re\lambda e + P(a^{-1/2})h),$$
we have: $\Re\Delta_1(\mu) \in \Omega$ if $\Re\lambda + \lambda_j > 0$ for all j. This condition is satisfied if $\Re\lambda > \max\{\mid \lambda_j \mid : j \in [1,r]\}$. But $\max\{\mid \lambda_j \mid : j \in [1,r]\} < \tau$. Consequently, $\Delta_1(\mu)$ is invertible for $\Re\lambda > \tau$. This means that $\lambda a + h \in U$ if $\Re\lambda > \tau$. Furthermore,

$$\Im\Delta_1(\mu) = \Im\lambda a - \mu e.$$

Thus,for $\mu \leq 0, \Im\lambda > 0$, we have: $\Im\Delta_1(\mu) \in \Omega$ and hence,$\Delta_1(\mu)$ is invertible by Theorem 5.1. This means that $\lambda a+h \in U$, if $\Im\lambda > 0$. But then $\overline{\lambda a + h} = \bar{\lambda}a+h \in U$, i.e. , $\lambda a + h \in U$, if $\Im\lambda \neq 0$.

Note that due to condition (iii)

$$G(\rho) = \rho F(\rho^{-1}), \rho > 0.$$

Hence, by the principle of analytic continuation

$$G(\lambda) = \lambda F(\lambda^{-1}), \tag{19}$$

if $\Re\lambda > \tau$ (both functions are analytic for $\Re\lambda > \tau$ and coincide for real λ greater than τ). Note,further, that the function $\lambda \to \lambda F(\lambda^{-1})$ is analytic for $\mid \lambda \mid > \tau$ and hence G can be analytically continued across the real axis from $-\infty$ to $-\tau$. Consequently, G is analytic in the complement of the cut

$$\{\lambda \in \mathbf{C} : \Im\lambda = 0, \mid \lambda \mid \leq \tau\}.$$

Due to condition i), G is also the Herglotz function (i.e., $\Im\lambda > 0$ implies $\Im G(\lambda) \geq 0$). Due to (19) G is bounded by a constant times $\mid \lambda \mid$ at the infinity. Hence, (see e.g. [1], section V.4)
$$G(\lambda) = \int_{-\tau}^{\tau} \frac{d\nu(t)}{t - \lambda} + \xi\lambda + \eta,$$
for all λ in the complement of the cut $\{\lambda \in \mathbf{C} : \Im\lambda = 0, \mid \lambda \mid \leq \tau\}$. Here ν is a positive finite measure with support in $[-\tau, \tau]$, and ξ, η are some constants. However,
$$F(\lambda) = \lambda G(\lambda^{-1}) = \int_{-\tau}^{\tau} \frac{\lambda^2 d\nu(t)}{\lambda t - 1} + \eta\lambda + \xi$$
for all λ in the complement of the cut$\{\lambda \in \mathbf{C} : \Im\lambda = 0, \mid \lambda \mid \geq \tau^{-1}\}$. But then for $n \geq 2$
$$\frac{d^n F}{d\lambda^n}(\lambda) = -n! \int_{-\tau}^{\tau} \frac{t^{n-2} d\nu(t)}{(1 - t\lambda)^{n+1}},$$
which is nonpositive when n is even and λ is real and $\mid \lambda \mid < \tau^{-1}$. □

Let Z be a Jordan algebra over C. Given $x \in Z$, let $p(Y) \in \mathbf{C}[Y]$ be the minimal polynomial of x,

$$p(Y) = \prod_{j=1}^{k} (Y - \lambda_j)^{\nu_j}.$$

Proposition 5.1. *There exists a complete system of orthogonal idempotents* $c_1 \dots c_k$ *in* $\mathbf{C}[Y]$, *i.e.,* $c_j^2 = c_j, c_j c_l = 0, j \neq l, c_1 + \dots c_k = e$, *such that for any polynomial* $q \in \mathbf{C}[Y]$,

$$q(x) = \sum_{j=1}^{k} \sum_{l=0}^{\nu_j - 1} \frac{(x - \lambda_j e)^l}{l!} q^{(l)}(\lambda_j) c_l.$$

Futhermore,

$$(x - \lambda_j e)^{\nu_j} c_j = 0, j = 1, \dots k.$$

This is Proposition 8.3.2 from [7].Note that $spec(x) = \{\lambda_1, \dots, \lambda_k\}$.

An element $x \in Z$ is said to be semi-simple if its minimal polynomial has only simple roots. For such an element

$$x = \sum_{j=1}^{k} \lambda_j c_j, q(x) = \sum_{j=1}^{k} q(\lambda_j) c_j, q \in \mathbf{C}[Y].$$

An element $x \in Z$ is said to be nilpotent if $x^m = 0$ for some integer m.

Proposition 5.2. *Every element* $x \in Z$ *can be uniquely written in the form*

$$x = x' + x''$$

with $x', x'' \in \mathbf{C}[x], x'$ *is semisimple and* x'' *is nilpotent.*

If f is holomorphic in an open set U of $\mathbf{C}$ containing $spec(x)$, we can define (following [7], p. 152)

$$f(x) = \sum_{j=1}^{k} \sum_{l=0}^{\nu_j - 1} \frac{(x - \lambda_j e)^l}{l!} f^{(l)}(\lambda_j) c_j. \tag{20}$$

Note that$f(x) \in \mathbf{C}[x]$ and if f, g are two such functions, then

$$(fg)(x) = f(x)g(x).$$

Proposition 5.3.

$$f(x) = \frac{1}{2\pi i} \int_C f(z)(ze - x)^{-1} dz,$$

where C *is a closed contour in* U *surrounding* $spec(x)$.

Proof. Consider the function

$$\phi_z(\lambda) = \frac{1}{z - \lambda}, \lambda \in \mathbf{C} \setminus \{z\}.$$

Then

$$\phi_z^{(l)}(\lambda) = \frac{l!}{(z - \lambda)^{l+1}}, l = 1, 2 \dots.$$

Hence, according to (20)

$$(ze - x)^{-1} = \sum_{j=1}^{k} \sum_{l=0}^{\nu_j - 1} \frac{(x - \lambda_j e)^l}{l!} \frac{l!}{(z - \lambda_j)^{l+1}} c_j =$$

$$\sum_{j=1}^{k}\sum_{l=0}^{\nu_j-1}\frac{(x-\lambda_j e)^l}{(z-\lambda_j)^{l+1}}c_j.$$

Hence,

$$\frac{1}{2\pi i}\int_C f(z)(ze-x)^{-1}dz=\sum_{j=1}^{k}\sum_{l=0}^{\nu_j-1}(x-\lambda_j e)^l c_j\frac{1}{2\pi i}\int_C\frac{f(z)dz}{(z-\lambda_j)^{l+1}}.$$

By residue theorem:

$$\frac{1}{2\pi i}\int_C\frac{f(z)dz}{z-\lambda_j)^{l+1}}=\frac{f^{(l)}(\lambda_j)}{l!}.$$

□

Remark 5.2. *Due to Proposition 5.3 one can develop a standard functional calculus on Z similar to,e.g. , [9], chapter 9.*

Consider a holomorhic branch

$$\ln\lambda=\ln\mid\lambda\mid+iarg\lambda, \tag{21}$$

where $-\pi < arg\lambda < \pi$. We ,then, can define

$$\lambda^p=\exp(p\ln\lambda), p\in\mathbf{C},\lambda\in\mathbf{C}\setminus\mathbf{R}_-.$$

Recall that $\mathbf{R}_-=\{\lambda\in\mathbf{C}:\Im\lambda=0,\Re\lambda\le 0\}$. If $U=\{z\in Z: spec(z)\subset\mathbf{C}\setminus\mathbf{R}_-$, we can define $\ln z, z^p$, using Proposition 5.1 or Proposition 5.3. Since, according to our definitions,

$$\lambda^p=\sum_{j=0}^{\infty}\frac{p^j(\ln\lambda)^j}{j!},\lambda\in\mathbf{C}\setminus\mathbf{R}_-,$$

we will have correspondingly (according to standard functional calculus; see e.g.[9], chapter 9).

$$z^p=\sum_{j=0}^{\infty}\frac{p^j(\ln z^j)}{j!}=\exp(p\ln z), z\in U.$$

We will need yet another characterization of functions $\ln z, z^p$ on U.

Proposition 5.4. *We have:*

$$lnz=\int_0^{+\infty}[\frac{e}{\tau+1}-(\tau e+z)^{-1}]d\tau, \tag{22}$$

$$z^\alpha=\frac{\sin(\pi\alpha)}{\alpha}\int_0^{+\infty}\tau^\alpha(\frac{e}{\tau}-(\tau e+z)^{-1})d\tau, \tag{23}$$

$z\in U, 0<\alpha<1$.

Proof. We will prove (23). Note that

$$\lambda^{\alpha} = \frac{\sin(\pi\alpha)}{\alpha}\int_0^{+\infty} \tau^{\alpha}(\frac{1}{\tau} - \frac{1}{\tau+\lambda})d\tau,$$

$\lambda \in \mathbf{C}\setminus\mathbf{R}_-, 0<\alpha<1$, since both sides are holomorphic functions on $\mathbf{C}\setminus\mathbf{R}_-$ which coincide for real positive λ (see e.g. [2], p. 106). Let $z\in U$,i.e., $spec(z) = \{\lambda_1,\ldots,\lambda_r\}\subset \mathbf{C}\setminus\mathbf{R}_-$. Then by Proposition 5.1:

$$(\tau e + z)^{-1} = \sum_{j=1}^{r}\sum_{l=0}^{\nu_{j-1}} \frac{(z-\lambda_j e)^l}{l!}\psi_{\tau}^{(l)}(\lambda_j)c_j,$$

where

$$\psi_{\tau}(\lambda) = \frac{1}{\tau+\lambda}.$$

Consequently,

$$\Delta = \frac{\sin(\pi\alpha)}{\alpha}\int_0^{+\infty}\tau^{\alpha}(\frac{e}{\tau} - (\tau e+z)^{-1})d\tau =$$

$$\frac{\sin(\pi\alpha)}{\alpha}[\sum_{j=1}^{r}\int_0^{+\infty}\tau^{\alpha}(\frac{1}{\tau}-\psi_{\tau}(\lambda_j))d\tau - \sum_{j=1}^{r}(\sum_{l=1}^{\nu_j-1}\frac{(z-\lambda_j e)^l}{l!}\int_0^{+\infty}\tau^{\alpha}\psi_{\tau}^{(l)}(\lambda_j)d\tau)c_j. \tag{24}$$

By (23):

$$\frac{d^n(\lambda^{\alpha})}{d\lambda^n} = -\frac{\sin(\pi\alpha)}{\alpha}\int_0^{+\infty}[\frac{d^n}{d\lambda^n}\psi_{\tau}(\lambda)]\tau^{\alpha}d\tau, \tag{25}$$

$n\geq 1$. Combining (23),(24),(25), we obtain:

$$\Delta = \sum_{j=1}^{r}\lambda_j^{\alpha}c_j + \sum_{j=1}^{r}\sum_{l=1}^{\nu_j-1}\frac{(z-\lambda_j e)^l}{l!}(\lambda^{\alpha})^{(l)}c_j = z^{\alpha},$$

where the last equality is due to Proposition 5.1. □

Lemma 5.1. *For $z\in U$,*

$$\ln(z^{-1}) = -\ln(z).$$

Proof. By (22)

$$\ln(z^{-1}) = \int_0^{+\infty}[\frac{e}{\tau+1} - (\tau e + z^{-1})^{-1}]d\tau = \int[\frac{e}{\tau+1} - \frac{z}{\tau}(z+\frac{e}{\tau}(z+\frac{e}{\tau})^{-1}]d\tau.$$

Further,

$$z(z+\frac{e}{\tau})^{-1} = e - \frac{1}{\tau}(z+\frac{e}{\tau})^{-1}.$$

Consequently,

$$\ln(z^{-1}) = \int_0^{+\infty}\frac{1}{\tau^2}[-\frac{e}{1+1/\tau} + (z+e/\tau)^{-1}]d\tau.$$

Making change of variables $\tilde{\tau} = 1/\tau$, we obtain the result. □

We now return to the case $Z = V^{\mathbf{C}}$ (i.e. , the complexification of an Euclidean Jordan algebra).

Lemma 5.2. *If* $z \in T_\Omega$, *then* $\ln(-z) = -i\pi e + \ln z$. *If* $-z \in T_\Omega$, *then* $\ln(-z) = i\pi e + \ln z$.

Proof. If $p(Y) \in \mathbf{C}[Y]$ isa minimal polynomial for z, then $p(-Y)$ is a minimal polynomial for $-z$. Consequently, by Proposition 5.1, if

$$\ln z = \sum_{j=1}^{r} \sum_{l=0}^{\nu_j - 1} \frac{(z - \lambda_j)^l}{l!} \ln^{(l)}(\lambda_j) c_j, spec(z) = \{\lambda_1, \ldots, \lambda_r\},$$

then

$$\ln(-z) = \sum_{j=1}^{r} \sum_{l=0}^{\nu_j - 1} \frac{(-z + \lambda_j)^l}{l!} \ln^{(l)}(-\lambda_j) c_j.$$

But

$$\ln^{(l)}(\lambda) = \frac{(-1)^{l-1}(l-1)!}{\lambda^l}, l \geq 1, \lambda \in \mathbf{C} \setminus \mathbf{R}_-.$$

Consequently,

$$\ln(z) - \ln(-z) = \sum_{j=1}^{r} [\ln \lambda_j - \ln(-\lambda_j)] c_j.$$

Since by (18), $\Im \lambda_j > 0, j = 1, \ldots, r$ for $z \in T_\Omega$, the result follows (see (21). □

Proposition 5.5. *Given* $z \in T_\Omega$, *we have:*

$$\ln z \in T_\Omega, i\pi e - \ln z \in T_\Omega.$$

Proof. If $z \in T_\Omega$ then $\tau e + z \in T_\Omega$ for all real τ. Hence, $-(\tau e + z)^{-1} \in T_\Omega$ by Theorem 5.1. But then $\ln z \in T_\Omega$ by (22). We also have that $-z^{-1} \in T_\Omega$. Hence,

$$\ln(-z^{-1}) = -\ln(-z) = -(-i\pi e + \ln z) = i\pi e - \ln z \in T_\Omega,$$

where we used Lemmas 5.1,5.2. □

Proposition 5.6. *If* $z \in T_\Omega, 0 < \alpha < 1$, *then* $z^\alpha \in T_\Omega, -\exp(i\alpha\pi) z^\alpha \in T_\Omega$.

Proof. By (23), $z^\alpha \in T_\Omega$. Besides, $u = -z^{-1} \in T_\Omega$. Hence, $u^\alpha \in T_\Omega$. Consequently, $(-u^\alpha)^{-1} \in T_\Omega$. However,

$$u^\alpha = \exp(\alpha \ln u) = \exp(\alpha \ln(-z^{-1})) = \exp(\alpha(i\pi e - \ln z)) =$$

$$\exp(\alpha\pi i) \exp(-\alpha \ln z).$$

Hence,

$$(-u^\alpha)^{-1} = -\exp(-\alpha\pi i) \exp(\alpha \ln z) = -\exp(-i\alpha\pi) z^\alpha.$$

Thus, $-\exp(-i\alpha\pi) z^\alpha \in T_\Omega$. □

Proposition 5.7. *Let* $u, v \in T_\Omega, -\exp(-i\alpha) u \in T_\Omega, -\exp(-i\beta) v \in T_\Omega, \alpha > 0, \beta > 0, \alpha + \beta < \pi$. *Then*

$$tr(uv) \subset \{\lambda = \rho \exp(i\theta), \rho > 0, 0 < \theta < \alpha + \beta\}.$$

Proof. Let $u = u_1 + iu_2, v = v_1 + iv_2; u_j, v_j \in V, j = 1, 2$. Then

$$\Im tr(uv) = tr(u_1 v_2) + tr(u_2 v_1).$$

Furthermore,

$$\exp(-i\alpha)u = (\cos\alpha + \sin\alpha) + i(\cos\alpha u_2 - \sin\alpha u_1).$$

Hence, the assumptions imply:

$$u_1 - \cot\alpha u_2 \in \Omega, u_2 \in \Omega,$$

and similarly

$$v_1 - \cot\beta v_2 \in \Omega, v_2 \in \Omega.$$

Consequently,

$$tr(u_1 v_2) = tr(P(v_2)^{1/2} u_1) > \cot\alpha tr(P(v_2)^{1/2} u_2) = \cot\alpha tr(u_2 v_2).$$

Similarly,

$$tr(u_2 v_1) > \cot\beta tr(u_2 v_2).$$

Hence,

$$\Im tr(uv) > (\cot\alpha + cot\beta) tr(u_2 v_2) = \frac{\sin(\alpha+\beta)}{\sin\alpha\sin\beta} tr(u_2 v_2) > 0.$$

Consider

$$u' = \exp(-i\alpha)u, v' = \exp(-i\beta)v.$$

Then by assumptions: $-\Im u' \in \Omega, -\Im v' \ in\Omega, \Im(\exp(i\alpha)u') \in \Omega, \Im(\exp(i\beta)v') \in \Omega$. Consequently,

$$\overline{u'} = \exp(i\alpha)\bar{u}, \overline{v'} = \exp(i\beta)\bar{v}$$

satisfy original assumptions. Hence, by what we have already proved:

$$\Im tr(\overline{u'v'}) > 0,$$

or

$$\Im tr(u'v') < 0,$$

i.e., $\Im tr(\exp(-i(\alpha+\beta))uv) < 0$. Let $tr(uv) = \rho\exp(i\theta), -\pi < \theta \leq \pi$. Since $\Im tr(uv) > 0$, we have $0 < \theta < \pi$. Then

$$tr(\exp(-i(\alpha+\beta))uv) = \rho\exp(i(\theta - (\alpha+\beta)),$$

$$\Im tr(\exp(-i(\alpha+\beta))uv) = \rho\sin(\theta - (\alpha+\beta)) < 0$$

implies $\theta < \alpha + \beta$. □

Remark 5.3. *Note that, if assumptions of proposition are satisfied, they also satisfied for $\alpha - \epsilon, \beta - \epsilon$ for some small positive ϵ. Consequently, proposition holds true , if $\alpha + \beta = \pi$.*

Example 5.1. *Consider*

$$f_1(v) = tr(\exp(h + \ln v), h \in V, v \in U \subset V^{\mathbf{C}}.$$

If $v \in T_\Omega$, *then* $\ln v, i\pi e - \ln v \in T_\Omega$ *by Proposition 5.5. But then, since* $\Im h = 0$, *we also have* $h + \ln v, i\pi e - (lnv + h) \in T_\Omega$. *By (18)*

$$spec(h + \ln v) \subset \{\lambda \in \mathbf{C} : \pi > \Im\lambda > 0\}. \tag{26}$$

If (26) is satisfied, then

$$spec(\exp(h + \ln v)) \subset \{\lambda \in \mathbf{C} : \Im\lambda > 0\}.$$

This obviously implies that $\Im f_1(v) > 0$, *if* $\Im v \in \Omega$. *It is also clear that* $f_1(\rho v) = \rho f_1(v), \forall \rho > 0, v \in U$. *By Theorem 5.2 the restriction of* f_1 *on* Ω *is concave. This is out Theorem 3.3.*

Example 5.2. *Let* $f_2(u, v) = tr((P(k)u^p)v^{1-p}), (u, v) \in U \times U \subset V^{\mathbf{C}} \times V^{\mathbf{C}} \cong (V \times V)^{\mathbf{C}}$. *Here* $0 \leq p \leq 1, k \in V$ *are fixed. If* $(u, v) \in T_\Omega \times T_\Omega$, *then* $u^p \in T_\Omega, v^{1-p} \in T_\Omega, -\exp(-ip\pi)u^p \in T_\Omega, -\exp(-i(1-p)\pi)v^{1-p} \in T_\Omega$ *by Proposition 5.6. It is clear that* $P(k)u^p$ *possesses the same properties as* u^p. *Hence, by Proposition 16* $\Im f_2(u, v) > 0$ *for* $(u, v) \in T_\Omega \times T_\Omega = T_{\Omega \times \Omega}$. *It is also clear that* $f_2(\bar{u}, \bar{v}) = \overline{f_2(u, v)}$ *and* $f_2(\rho u, \rho v) = \rho f_2(u, v)$ *for* $\rho > 0$. *Consequently, the restriction of* f_2 *on* $\Omega \times \Omega$ *is concave. This is our main Theorem 3.1.*

6 Conclusions

In this paper we developed a Jordan-algebraic version of E. Lieb inequalities. As an application, we proved a version of noncommutative Bernstein inequality. Possible further applications include optimization, statistics and quantum information theory through the Jordan-algebraic version of quantum entropy. It also would be interesting to see what asymptotic properties of random matrix ensembles admit Jordan-algebraic generalizations.

7 Acknowledgments

This research was supported in part by Simmons Foundation Grant 275013.

References

1. Bhatia, R.: Matrix Analysis. Springer, New York(1997)
2. Cartan,H.:Elementary theory of analytic functions of one or several comlex variables. Dover Publications, New York (1995)
3. Faybusovich, L., Tsuchiya, T.: Primal-dual algorithms and infinite-dimensional Jordan algebras of finite rank. Mathematical Programming, 97 , no. 3, 471-493(2003)
4. Epstein, H.: Remarks on two theorems of E. Lieb. Comm. Math. Phys. 31, 317-325 (1973)
5. Faybusovich,L:Several Jordan-algebraic aspects of optimization. Optimization, 57,no 3,379-393(2008)
6. Effros, E:A matrix convexity approach to some celebrated quantum inequalities. Proc. Natl. Acad. Sci. 106 (4),1006-1008(2009)
7. Faraut, J. , Koranyi, A.: Analysis on Symmetric Cones. Clarendon Press,Oxford(1994)

8. Foucart,S. ,Holger, R.:A Mathematical introduction to compressive sensing. Birkhauser,New York (2013)
9. Hille, E.: Methods in classical and functional analysis. Addison-Wesley, Reading(1972)
10. Hirzebruch,U.: Der Min-Max Satz von E. Fisher fur formal-reel Jordan-Algebren. Math. Z,94,387-390 (1970)
11. Lieb, E.:Convex trace functions and the Wigner-Yanase-Dyson conjecture. Adv. Math. 11,267-288(1973)
12. Lim,Y.,Kim,J.,Faybusovich, L.:Simaltaneous diagonalization on simple Euclidean Jordan algebras and its applications. Forum Math., 15, no. 4, 639-644(2003)
13. Simon, B.: Trace ideals and their applications. Cambridge University Press, Cambridge (1979)
14. Tropp, J. A.:User-friendly tail bounds for sums of random matrices. Found. Comput. Math. 12(4),389-434(2012)