

## BEREZIN-LIEB INEQUALITY: AN EXTENSION TO NORMAL OPERATORS

N Sookia\*  
Faculty of Science,  
University of Mauritius,  
Réduit  
Email: *sookian@uom.ac.mu*

P Nunkoo Gonpot  
Faculty of Science,  
University of Mauritius,  
Réduit  
Email: *pgonpot@uom.ac.mu*

*Paper Accepted on 06 January 2011*

---

### **Abstract**

The Berezin-Lieb inequality, which is a Jensen's type inequality for convex functions of self-adjoint operators, is considered. We find a similar type of inequality for convex functions  $\phi$  of normal operators. The normal operator is assumed to be bounded and acting on an infinite dimensional separable Hilbert space  $\mathcal{H}$ . Finally the minimal closed convex set upon which it is sufficient to define the convex function  $\phi$  is determined.

Keywords: Berezin-Lieb inequality, normal operator, convex combination, closed convex hull.

*\*For correspondences and reprints*

## 1. INTRODUCTION

If  $f$  is a convex function,  $B$  is a self-adjoint (not necessarily bounded) operator in a Hilbert space  $\mathcal{H}$  and  $B_P = PBP$ , where  $P$  is an orthogonal projection in  $\mathcal{H}$ , then

$$\mathrm{Tr}Pf(B_P)P \leq \mathrm{Tr}Pf(B)P \quad (1.1)$$

provided that  $Pf(B)P$  belongs to the trace class  $\mathbf{S}_1$ . This inequality was proved in the early seventies by F. Berezin (1) and, independently, by E. Lieb (6) (see also (10)). In (4) A. Laptev and Yu. Safarov obtained a more general version of (1.1). In particular, they only assumed that

$$Pf(B)P - Pf(B_P)P \in \mathbf{S}_1 \quad (1.2)$$

and considered the case where  $P$  is a contraction operator.

The inequalities (1.1), (1.2) proved to be very useful and have already been applied to various spectral problems, see (4), (6), (11). Inequalities of this type are being actively studied (See (15) and (16)) and have yielded a number of physically motivated applications in Mathematical Physics, (7), (8), (12), (13). The aim of this paper is to obtain an analogue of (1.1) for normal operators  $B$ .

In the main part of the paper we shall use various notions and results from operator theory which can be found in for example (2), (3). We state, without proof, the Berezin-Lieb inequality as in (5) but include some valuable remarks. We have included in section 2 the necessary notions of convexity, and Jensen's inequality which is in essence a generalization of the defining inequality for a convex function.

We manage to obtain an analogue of (1.1) for a bounded normal operator  $B$  acting on  $\mathcal{H}$ . We again use the theory of convexity and convex sets to give precise information about the set upon which we define our convex function  $\phi$  in the statement of our main theorem.

## 2. THE SETTING

Let  $\mathcal{H}$  denote a separable Hilbert space and  $\mathcal{B}(\mathcal{H})$  the class of bounded operators acting in  $\mathcal{H}$ . We assume the spectrum of  $B \in \mathcal{B}(\mathcal{H}), \sigma(B)$ , to be discrete. Let  $\{\lambda_i\}$  be the set of all distinct eigenvalues of  $B$  with corresponding eigenspaces  $\{M_i\}$  and orthogonal complements  $M_i^\perp$ , and let  $\pi_i$  denote the orthogonal projection on the corresponding eigenspace  $M_i$  of  $\lambda_i$ .

Denote by  $\mathcal{C}(\sigma(t))$ , the Banach space of all continuous complex-valued functions  $f$  on  $\sigma(T)$ , with the norm of  $f$  given by

$$\|f\| = \sup\{|f(\lambda)| : \lambda \in \sigma(T)\}. \quad (2.3)$$

Let  $\mathbf{S}_1$  be the Banach space of nuclear operators  $T$  in  $\mathcal{B}(\mathcal{H})$  with trace norm given by

$$\|T\|_1 = \sum_k |\langle T\xi_k, \xi_k \rangle|, \quad (2.4)$$

where  $\{\xi_k\}$  is any orthonormal basis in  $\mathcal{H}$ . It is well-known that the trace of  $T$  coincides with the trace of the matrix representation of  $T$ , that is,

$$\text{Tr}T = \sum_i t_{ii}. \quad (2.5)$$

Let  $\mathbb{R}, \mathbb{C}$  denote the fields of real and complex numbers respectively and  $A$  a non-empty, convex set in  $\mathbb{C}$ . Then

**Theorem 2.1.** *Every convex combination of points of  $A$  in  $\mathbb{C}$  belongs to that set.*

*Proof:* The proof is by induction. See (14) for details. □

We recall the following definition from (14).

**Definition 2.2.** *The convex hull of  $A \subset \mathbb{C}$ , denoted by  $\mathcal{C}(A)$ , is defined to be the set of all convex combinations of points of  $A$ . It follows that if  $a_1, \dots, a_m \in \mathbb{C}$ , then*

$$\mathcal{C}\{a_1, \dots, a_m\} = \left\{ \sum_{i=1}^m \alpha_i a_i : \alpha_1, \dots, \alpha_m \geq 0, \sum_{i=1}^m \alpha_i = 1 \right\}. \quad (2.6)$$

$\mathcal{C}(A)$ , is convex, and is the intersection of all convex sets that contain  $A$ . Hence given  $a_1, \dots, a_m \in A$ ,  $\mathcal{C}\{a_1, \dots, a_m\}$  is the polygon which these  $m$  points determine. It is the minimal convex set which contains all these  $m$  points.

The closed convex hull of a set, is simply defined as the closure of its convex hull, and is thus the smallest closed convex set containing  $A$ .

The convexity of a function  $\phi$  can be related with its second-order partial derivatives. Since every complex number is a function of two real variables (the real part and the imaginary part), one can identify, in one way or another, the set  $\mathbb{C}$  with  $\mathbb{R}^2$  and thus write  $A \subset \mathbb{R}^2$  instead of  $A \subset \mathbb{C}$ . Indeed  $\phi$  will then be a convex function defined on  $A \subset \mathbb{R}^2$ .

**Lemma 2.3.** *Let  $\phi$  be a convex function defined on  $A \subset \mathbb{R}^2$  and  $\mathbf{b} = (b_1, b_2) \in \mathbb{R}^2$ . Let  $\mathbf{a} \in A$  and  $I = \{\alpha \in \mathbb{R} : \mathbf{a} + \alpha\mathbf{b} \in A\}$ . Then  $I$  is an interval of  $\mathbb{R}$  containing the origin. If we define  $g : I \rightarrow \mathbb{R}$  as*

$$g(\alpha) = \phi(\mathbf{a} + \alpha\mathbf{b}), \quad \text{all } \alpha \in I, \quad (2.7)$$

then  $g$  is convex.

*Proof:* For all  $p, q \in I$  and  $\alpha \in [0, 1]$ ,

$$\begin{aligned} \phi(\mathbf{a} + (\alpha p + (1 - \alpha)q)\mathbf{b}) &= \phi(\alpha(\mathbf{a} + p\mathbf{b}) + (1 - \alpha)(\mathbf{a} + q\mathbf{b})) \\ &\leq \alpha\phi(\mathbf{a} + p\mathbf{b}) + (1 - \alpha)\phi(\mathbf{a} + q\mathbf{b}). \end{aligned}$$

□

Define the real-valued function  $Q$  with domain  $A \times \mathbb{R}^2$  by

$$Q(\mathbf{a}, \mathbf{z}) = \sum_{i,j=1}^2 \phi_{ij}(\mathbf{a})z_i z_j, \quad (2.8)$$

where  $\phi_{ij}(\mathbf{a}) = \frac{\partial^2 \phi}{\partial a^j \partial a^i}(\mathbf{a})$  denotes the partial derivative of order 2 of  $\phi$  at  $\mathbf{a}$ .

**Theorem 2.4.** *Let  $\phi$  be a real-valued function which is defined and has continuous second-order partial derivatives on  $A \subset \mathbb{R}^2$ . Then  $\phi$  is convex if and only if, for every  $\mathbf{a} \in A$ ,*

$$Q(\mathbf{a}, \mathbf{z}) \geq 0, \quad \text{all } \mathbf{z} = (z_1, z_2) \in \mathbb{R}^2. \quad (2.9)$$

*Proof:* Let  $\mathbf{b} \in A$  and  $\mathbf{z} = (z_1, z_2) \in \mathbb{R}^2$ . Define

$$I = \{\alpha \in \mathbb{R} : \mathbf{b} + \alpha\mathbf{z} \in A\}, \quad (2.10)$$

and

$$g(\alpha) = \phi(\mathbf{b} + \alpha\mathbf{z}), \quad \text{all } \alpha \in I. \quad (2.11)$$

Then  $I$  is an interval of  $\mathbb{R}$  and  $g$  is convex if  $\phi$  is convex (see Lemma 2.3). Conversely, suppose that for any choice of  $\mathbf{a}$  and  $\mathbf{b}$ ,  $g(\alpha)$  is convex for  $\alpha \in I$ . Let  $\mathbf{a}, \mathbf{b} \in A$  and  $\alpha \in [0, 1]$ . Write  $\mathbf{z} = \mathbf{a} - \mathbf{b}$ . Since  $g$  is convex, we have

$$\begin{aligned} \phi(\alpha\mathbf{a} + (1 - \alpha)\mathbf{b}) &= g(\alpha) \\ &\leq (1 - \alpha)g(0) + \alpha g(1) \\ &= \alpha\phi(\mathbf{a}) + (1 - \alpha)\phi(\mathbf{b}). \end{aligned}$$

Since  $\phi$  has continuous second-order partial derivatives on  $A$ , each function  $g$  is twice differentiable on  $I$ . Now

$$g''(\alpha) = \sum_{i,j=1}^2 \phi_{ij}(\mathbf{a})z_i z_j = Q(\mathbf{a}, \mathbf{z}), \quad (2.12)$$

where  $\alpha \in I$  and the partial derivatives are evaluated at the point  $\mathbf{a} = \mathbf{b} + \alpha\mathbf{z}$ . Each function  $g$ , being a twice differentiable function, is convex on  $I$  if and only if  $Q(\mathbf{a}, \mathbf{z}) \geq 0$ . □

**Remark 2.5.** *The  $2 \times 2$  matrix whose  $(i, j)^{th}$  element is  $\frac{\partial^2 \phi}{\partial a_i \partial a_j}$  evaluated at  $\mathbf{a}$  of  $A$  is the Hessian matrix of  $\phi$  at  $\mathbf{a}$ . Obviously this matrix is symmetric because  $\phi$  has continuous second-order partial derivatives on  $A$ .*

Hence Theorem 2.4 can be interpreted in this way:

**Theorem 2.6.**  *$\phi$  is convex on  $A$  if and only if its Hessian matrix is non-negative semidefinite at each point of  $A$ .*

We end this section by recalling Jensen's inequality.

**Theorem 2.7.** *(Jensen's Inequality).*

*Let  $A$  be a bounded set having an infinite number of points,  $x_i$ , and  $\phi : A \rightarrow \mathbb{R}$  be a convex function which is bounded. Then*

$$\phi \left( \sum_i \alpha_i x_i \right) \leq \sum_i \alpha_i \phi(x_i), \quad (2.13)$$

where  $\alpha_i \geq 0$ ,  $i = 1, 2, \dots$  and  $\sum_i \alpha_i = 1$ .

*Proof:* The proof is by induction. See (9) for details.  $\square$

**Remark 2.8.** *In a natural way, we identified  $\mathbb{R}^2$  with  $\mathbb{C}$  by identifying each point  $(x,y)$  of  $\mathbb{R}^2$  with the complex number  $x + iy$  and vice versa. This identification has allowed us to refer to the convex hull of a set of complex numbers.*

### 3. BEREZIN INEQUALITY: THE SELF-ADJOINT CASE

Let  $B$  be a self-adjoint operator and  $P$  be any orthogonal projection in  $\mathcal{H}$  with finite rank. So  $P$  is a compact operator. Define

$$B_P = PBP. \quad (3.14)$$

Unlike in (4), we will always deal with  $B_P \in \mathcal{B}(\mathcal{H})$  since  $B$  here is assumed to be bounded. By continuous functional calculus, it follows that

$$\phi(B) = \sum_i \phi(\lambda_i) \pi_i. \quad (3.15)$$

For the sake of completeness, we state the Berezin-Lieb inequality as in (5), without proving it.

**Theorem 3.1.** *Let  $\phi$  be a convex function on  $K = \bigcup_{0 \leq t \leq 1} t\sigma(B) \subset \mathbb{R}$ . Assume that  $PB$  is a compact operator and that*

$$P\phi(B)P - P\phi(B_P)P \in \mathbf{S}_1. \quad (3.16)$$

*Then*

$$\text{Tr}(P\phi(B)P - P\phi(B_P)P) \geq 0. \quad (3.17)$$

*Proof:* See (4; 5) for details.  $\square$

We make the following remarks:

- (1) Berezin inequality holds under the assumption that  $\phi(0) = 0$ .
- (2) Obviously  $B_P$  is a self-adjoint operator when  $B$ , itself, is self-adjoint.

In fact,

$$B_P^* = (PBP)^* = P^*B^*P^* = PBP = B_P. \quad (3.18)$$

- (3) Since  $K$  contains the set of values  $\langle Bu, u \rangle$ , all  $u \in \mathcal{H}$ , it follows that the spectra of operators  $B$  and  $B_P$  lie in  $K$ . Hence it is enough to define the functions  $\phi$  on  $K$ . Then all the conditions involving  $\phi$  need only hold on this set.
- (4) From Theorem 2.4, we can assume the convex function  $\phi$  to be a non-negative function. Clearly, since there is no restriction on the first partial derivative of  $\phi$ , the convex function  $\phi$  can always be made non-negative by adding a suitable linear function to it. Moreover, in doing so the operator  $P\phi(B)P - P\phi(B_P)P$  remains unchanged.
- (5)  $\langle B_P\xi_k, \xi_k \rangle = \langle B\xi_k, \xi_k \rangle$ , that is, inner product in the whole space is the same as inner product in the subspace as  $P$  is an orthogonal projection in  $\mathcal{H}$ .

#### 4. BEREZIN INEQUALITY: THE NORMAL CASE

Let  $B$  be a bounded normal operator acting on  $\mathcal{H}$ . Then  $B$  has a spectral resolution

$$B = \sum_i \lambda_i \pi_i, \tag{4.19}$$

and by continuous functional calculus,

$$\phi(B) = \sum_i \phi(\lambda_i) \pi_i, \tag{4.20}$$

where  $\phi$  denotes a real-valued, convex function defined on  $\mathbb{C}$ .

**Remark 4.1.** *Functional calculus cannot be applied to the operator  $B_P$  as  $\phi(B_P)$  is not well-defined. This difficulty arises because  $B_P$  is not normal. To overcome such a difficulty, we will first make use of Lidskii's theorem.*

**Theorem 4.2.** *For an operator  $T \in \mathbf{S}_1$ , the matrix trace of the operator  $T$  coincides with its spectral trace, that is,*

$$\text{Tr}(T) = \sum_k \lambda_k(T), \tag{4.21}$$

where the eigenvalues  $\lambda_k(T)$  are counted with their multiplicities.

Note that by multiplicity of the eigenvalue  $\lambda_k(T)$ , we mean the dimension of the eigenspace corresponding to the eigenvalue  $\lambda_k(T)$ .

*Proof:* See (3). □

Assume that  $B_P \in \mathbf{S}_1$  and that  $\sigma(B_P) = \{\mu_j\}_{j=1}^\infty$ . Then by Lidskii's theorem, we have

$$\text{Tr}(B_P) = \sum_j \mu_j. \tag{4.22}$$

Define the trace of  $\phi(B_P)$  as:

$$\text{Tr}(\phi(B_P)) = \sum_j \phi(\mu_j). \tag{4.23}$$

(4.23) holds when  $\phi$  is a polynomial.

**Remark 4.3.** *For any reasonable definition of  $\phi(B_P)$ , the spectrum of this operator must be equal to the set  $\{\phi(\mu_j)\}_{j=1}^\infty$ .*

Since  $B_P$  is not normal, the matrix representation of  $B_P$  with respect to any orthonormal basis in  $\mathcal{H}$  is not a normal matrix and hence it cannot be unitarily similar to a diagonal matrix. However, due to Schur, any arbitrary matrix can be reduced by means of a unitary similarity. This idea is expressed in Schur's lemma, (3):

**Lemma 4.4.** *There exists an orthonormal basis  $\{\xi_k\}$  such that  $B_P$  is a triangular matrix and the diagonal entries of the matrix of the operator  $B_P$ , with respect to the orthonormal basis  $\{\xi_k\}$  consists of eigenvalues  $\mu_k$  of  $B_P$ .*

*Proof:* See (3). □

By Schur's lemma, we have

$$\mu_k = \langle B_P \xi_k, \xi_k \rangle, \tag{4.24}$$

and hence

$$\begin{aligned} \phi(\mu_k) &= \phi(\langle B \xi_k, \xi_k \rangle) \\ &= \phi\left(\sum_i \lambda_i \langle \pi_i \xi_k, \xi_k \rangle\right) \\ &= \phi\left(\sum_i \lambda_i \delta_{ik}\right), \end{aligned}$$

where  $\delta_{ik} = \langle \pi_i \xi_k, \xi_k \rangle$  and  $\sum_i \delta_{ik} = 1$ . Also

$$\begin{aligned} \langle P\phi(B)P\xi_k, \xi_k \rangle &= \langle \phi(B)\xi_k, \xi_k \rangle \\ &= \left\langle \sum_i \phi(\lambda_i)\pi_i \xi_k, \xi_k \right\rangle \\ &= \sum_i \phi(\lambda_i)\delta_{ik}. \end{aligned}$$

In view of Theorem 2.7, we argue that

$$\phi\left(\sum_i \lambda_i \delta_{ik}\right) \leq \sum_i \phi(\lambda_i)\delta_{ik}, \quad (4.25)$$

for any  $\delta_{ik}$  such that  $\sum_{i=1} \delta_{ik} = 1$ . By summing over  $k$ , we see that

$$\text{Tr}(P\phi(B)P) - \text{Tr}\phi(B_P) \geq 0, \quad (4.26)$$

which is the required Berezin inequality in this case.  $\square$

Now we wish to determine the minimal set  $K$ , that includes the complex eigenvalues  $\lambda_i$  and  $\mu_j$  of  $B$  and  $B_P$  respectively, for which (4.25) will hold. First we have the following lemma:

**Lemma 4.5.** *For any operator  $B \in \mathcal{B}(\mathcal{H})$ , denote the set of values  $\langle Bu, u \rangle$ ,  $u \in \mathcal{H}$  by  $\mathcal{E}_B$ . Then  $\mathcal{E}_B$  contains the spectrum of both  $B$  and  $B_P$ .*

*Proof:* Suppose  $u$  is an eigenfunction of  $B$  such that  $\|u\| = 1$  and  $Bu = \lambda u$ , where  $\lambda$  is the eigenvalue corresponding to the eigenfunction  $u$ . Since  $\langle Bu, u \rangle = \lambda$ ,  $\sigma(B) \subset \mathcal{E}_B$ .

To prove the second part, we first note that since we assume  $B_P \in \mathbf{S}_1$ ,  $B_P$  is compact and hence  $0 \neq \mu_j \in \sigma(B_P)$  implies that  $\mu_j$  is an eigenvalue of  $B_P$ . So there exists a normalised eigenfunction  $\psi_j$  such that

$$B_P\psi_j = \mu_j\psi_j, \quad (4.27)$$

and

$$\mu_j = \langle PBP\psi_j, \psi_j \rangle = \langle BP\psi_j, P\psi_j \rangle \in \mathcal{E}_B, \quad (4.28)$$

so that

$$\sigma(B_P) \subset \mathcal{E}_{B_P} \subset \mathcal{E}_B. \quad (4.29)$$

$\square$

**Theorem 4.6.** *Let  $B \in \mathcal{B}(\mathcal{H})$  be a normal operator. Define*

$$A = \sigma(B) \cup \{0\}. \quad (4.30)$$

*Denote by  $\mathfrak{S}(A)$  the closed convex hull of  $A$ . Then  $\mathfrak{S}(A)$  contains both the spectrum of  $B$  and  $B_P$ .*

*Proof:* By Lemma 4.5 it is enough to prove that  $\mathcal{E}_B \subset \mathfrak{S}(A)$ . We will rather prove that  $\mathcal{E}_B = \mathcal{C}(A)$  since by definition of the closure of a set,  $\mathcal{C}(A) \subset \mathfrak{S}(A)$ . Let  $u \in \mathcal{H}$  be such that  $\|u\| \leq 1$ . We can choose the eigenfunctions  $u_j$  of  $B$  in such a way that they form a complete orthonormal set. Obviously

$$u = \sum_{j=1}^{\infty} c_j u_j, \quad (4.31)$$

for some constants  $c_j$  so that

$$\|u\|^2 = \sum_{j=1}^{\infty} |c_j|^2 \leq 1, \quad (4.32)$$

and

$$\langle Bu, u \rangle = \sum_{j=1}^{\infty} \lambda_j |c_j|^2. \quad (4.33)$$

Since 0 may or may not be an eigenvalue of  $B$ , by including  $\lambda_0 = 0$  and using (4.32), we can always choose  $c_0$  in such a way that

$$\langle Bu, u \rangle = \sum_{j=0}^{\infty} \lambda_j |c_j|^2, \quad (4.34)$$

where  $\sum_{j=0}^{\infty} |c_j|^2 = 1$ .

From Definition 2.2, it follows that  $\mathcal{E}_B = \mathcal{C}(A)$ . □

## 5. CONCLUSION

Let  $B$  be a bounded normal operator acting on an infinite dimensional separable Hilbert space  $\mathcal{H}$ ,  $P$  be an orthogonal projection in  $\mathcal{H}$  and  $K = \mathfrak{S}(A)$ , where  $A$  is the set defined by (4.30). Then the next theorem is a generalization of the Berezin-Lieb inequality for such an operator  $B$ .

**Theorem 5.1.** *Let  $\phi$  be a convex function on  $K$ . Assume that  $\sigma(B)$  is discrete and that  $B_P \in \mathbf{S}_1$ .*

Then

$$\text{Tr}(P\phi(B)P) - \text{Tr}\phi(B_p) \geq 0.$$

*Proof:* Already discussed. □

## 6. ACKNOWLEDGEMENT

We would like to thank Prof. Yu. Safarov for his valuable suggestions.

## REFERENCES

1. F. Berezin, *Convex Functions of Operators*, Math. Sb. **Vol 88** (1972), 268–276.
2. M.Sh. Birman and M.Z. Solomyak, *Spectral theory of self-adjoint operators in hilbert space*, Leningrad, Reidel Publ. Co., 1987, English translation.
3. I.C. Gohberg and M.G. Krein, *Introduction to the theory of nonselfadjoint operators*, American Mathematical society, 1969, Vol.18, translations of Mathematical Monograph.
4. A. Laptev and Yu. Safarov, *A generalization of the Berezin-Lieb inequality*, Amer. Math. Soc. Transl. **Vol 175** (1996), no. 2.
5. ———, *Szego type limit theorems*, J. Funct. Anal. **Vol 138** (1996), no. 2.
6. E. Lieb, *The classical limit of quantum spin systems*, Comm. Math. Phys. **Vol 31** (1973), 327–340.
7. L. Chayes M. Biskup and S. Starr, *Quantum spin systems at positive temperature*, Comm. Math. Phys. **269** (2007).
8. M. Loss R. L. Frank and T. Weidl, *Polya’s conjecture in the presence of a constant magnetic field*, J. Eur. Math. Soc. **Vol 11** (2009), 1365–1383.
9. W. Rudin, *Functional analysis*, New York, McGraw-Hill Book Company, 1973.
10. B. Simon, *Trace ideals and their applications*, Cambridge University Press, London, New York, Melbourne, 1979.
11. ———, *The classical limit of quantum partition functions*, Comm. Math. Phys. **Vol 71** (1980), 247–276.

12. M. V. Travaglia, *On an inequality involving power and contraction of matrices with and without trace*, Journal of Inequalities in Pure and Applied Mathematics **Vol 65** (2006), no. 7.
13. J. T. Lewis W. Cegla and G.A. Raggio, *The free energy of quantum spin systems and large deviations*, Comm. Math. Phys. **Vol 118** (1988), 337–354.
14. R. Webster, *Convexity*, Oxford University Press, Oxford, New York, Tokyo, 1994.
15. T. Weidl, *Improved Berezin-Li-Yau inequalities with a remainder term*, Spectral Theory of Differential Operators, Amer. Math. Soc. Transl **Vol 225** (2008), no. 2, 253–263.
16. S.Y. Yolcu, *An improvement to a Berezin-Li-Yau type inequality*, Proc. Amer. Math. Soc. **Vol 138** (2010), 4059–4066.