A GENERALIZATION OF THE BEREZIN–LIEB INEQUALITY

BY
A.Laptev and Yu. Safarov

In the early seventies both F. Berezin [B] and E. Lieb [L] (see also [S]) independently obtained a Jensen's type inequality for convex functions of selfadjoint operators. This inequality turns out to be very useful and has been applied to various spectral problems, see for example [BSh].

If φ is a convex function, B_P is a selfadjoint operator (not necessarily bounded) in a Hilbert space H, and moreover the operator B_P can be represented as $B_P = PBP$, where P is an orthogonal projection in H then the Berezin inequality states that

$$\operatorname{Tr} P\varphi(B_P)P \le \operatorname{Tr} P\varphi(B)P,$$

provided that the right hand side is finite.

Applying this inequality to the spectral analysis of pseoudodifferential operators we were interested in two sides estimates of the trace $\operatorname{Tr} P\psi(B_P)P$ when the function ψ is not necessarily a convex function. In Theorem 12 of this paper we prove a trace estimate for such functions. This estimate implies a more general version of the Berezin inequality (see Corollary 13). In particular we prove the inequality

$$\operatorname{Tr}\left(P\varphi(B)P - P\varphi(B_P)P\right) \ge 0,$$

assuming only that the difference $P\varphi(B)P - P\varphi(B_P)P$ is from the trace class. We also obtain inequalities where P is a contraction operator.

1. The operator P^*BP. Let H and H_0 be Hilbert spaces, B be a selfadjoint operator in H, and $P : H_0 \to H$ be a bounded operator such that $\|P\|_{H_0\to H} \le 1$. The operator B is allowed to be unbounded, and then we denote by $\mathcal{D}(B)$ its domain. We are going to consider the operator P^*BP acting in the space H_0. When B is bounded, this operator is well-defined and selfadjoint. However, when B is unbounded, the natural definition of P^*BP might make no sense (for example, if $\mathcal{D}(B) \cap PH_0 = \{0\}$). In this case we need some additional assumptions.

Let $(\cdot,\cdot)$, $\|\cdot\|$ and $(\cdot,\cdot)_0$, $\|\cdot\|_0$ be scalar products and norms in H and H_0 respectively. We denote by $E_B(\lambda)$ the spectral measure of the operator B, and consider the skew-linear form

$$Q[\xi,\eta] \overset{\text{def}}{=} Q_{B,P}[\xi,\eta] \;=\; \int \lambda\,(dE_B(\lambda)P\xi, P\eta), \qquad \xi,\eta \in H_0,$$

and the corresponding quadratic form

$$Q[\xi] = Q_{B,P}[\xi] \overset{\text{def}}{=} \int \lambda\,(dE_B(\lambda)P\xi, P\xi), \qquad \xi \in H_0. \tag{1}$$

Typeset by $\mathcal{AMS}$-TEX

If B is bounded then $Q[\xi,\eta]=(BP\xi,P\eta)$ and the form $Q[\xi]$ is defined on the whole space H_0. In general situation the domain of Q is the linear set

$$\mathcal{D}(Q)\stackrel{\mathrm{def}}{=}\{\ \xi\in H_0\ :\ \int|\lambda|\,(dE_B(\lambda)P\xi,P\xi)<\infty\ \}. \tag{2}$$

Obviously, we have

$$\mathcal{D}(Q)=\{\ \xi\in H_0 : P\xi\in\mathcal{D}(|B|^{1/2})\ \} \tag{3}$$

and

$$\int|\lambda|\,(dE_B(\lambda)P\xi,P\xi)=\|\,|B|^{1/2}P\xi\|^2. \tag{4}$$

Generally speaking, the set (2) may also be very poor. Besides, even if that is not true, Q might not generate a selfadjoint operator. Therefore we introduce the following two conditions which are assumed to be fulfilled throughout all the paper.

(C_1) The set $\mathcal{D}(Q)$ is dense in H_0.
(C_2) The form $Q[\cdot]$ is semi-bounded and can be closable in H_0.

Let $\overline{Q}[\cdot]$ be the closure of the form $Q[\cdot]$. This closure is defined on some dense set $\mathcal{D}(\overline{Q})\subset H_0$ containing $\mathcal{D}(Q)$, and it defines a Hilbert structure on $\mathcal{D}(\overline{Q})$. We denote this Hilbert space by H_1, $H_1\subset H_0$.

Let H' be a closed subspace of H_1 which is also dense in H_0, and $Q'[\cdot]$ be the restriction of the form $\overline{Q}[\cdot]$ to H'. Then $Q'[\cdot]$ is a closed quadratic form in H_0, and so it generates some selfadjoint operator B_P.

Obviously, if B is bounded then $H'=H_1=H_0$ and $B_P=P^*BP$. If B is an unbounded operator, then B_P is not defined uniquely. Each H' takes care of a selfadjoint operator B_P, which can be considered as a selfadjoint realization of P^*BP. All further results are valid for any such realization. Through all over the paper we assume H' to be fixed and deal with the corresponding selfadjoint operator B_P.

The condition (C_2) is not effective. The following lemma gives the equivalent condition which is more convenient to deal with.

Lemma 1. *The condition (C_2) is fulfilled if and only if there exists a constant C such that*

$$\|\,|B|^{1/2}P\xi\|^2\le C\,(|\,Q[\xi]\,|+\|\xi\|_0^2),\qquad \forall\xi\in\mathcal{D}(Q). \tag{5}$$

Proof. By lemma 10.1.6 from [BS] the form $Q[\cdot]$ can be closed if and only if for any sequence $\xi_k\in\mathcal{D}(Q)$, $k=1,2,\dots$, such that $\|\xi_k\|_0\to0$, $k\to\infty$, and

$$Q[\xi_k-\xi_j]\to0,\qquad j,k\to\infty, \tag{6}$$

we have

$$Q[\xi_k,\eta]\to0,\qquad \forall\eta\in\mathcal{D}(Q). \tag{7}$$

By (3) we can write

$$Q[\xi_k,\eta] = \Big((I+|B|)^{1/2}P\xi_k, B(I+|B|)^{-1/2}P\eta\Big).$$

Therefore the form $Q[\cdot]$ can be closed if and only if the sequence $(I+|B|)^{1/2}P\xi_k$ weakly tends to zero in H.

The condition (6) implies that $Q[\xi_k]$ are uniformly bounded. Hence, from (5) it follows that $\|(I+|B|)^{1/2}P\xi_k\|$ are also uniformly bounded. For any $u\in\mathcal{D}(|B|^{1/2})$ we have

$$\Big((I+|B|)^{1/2}P\xi_k, u\Big) = \Big(P\xi_k, (I+|B|)^{1/2}u\Big) \to 0.$$

Thus, the sequence $(I+|B|)^{1/2}P\xi_k$ is bounded and weakly tends to zero on the set $\mathcal{D}(|B|^{1/2})$ which is dense in H. It implies that this sequence weakly tends to zero. So (5) yields (C_2).

If the estimate (5) does not hold, then there exists a sequence $\xi_k\in\mathcal{D}(Q)$ such that $\|\xi_k\|_0\to 0$, $Q[\xi_k]\to 0$, $k\to\infty$, but $\|(I+|B|)^{1/2}P\xi_k\|\to\infty$. For these ξ_k the sequence $(I+|B|)^{1/2}P\xi_k$ does not weakly converge, and therefore the form $Q[\cdot]$ cannot be closed. The proof is complete.

2. Functional spaces. In what follows we always assume all functions to be measurable. Moreover, we are going to deal only with functions from the class $BV^1(\mathbf{R})$ which is defined as follows.

Definition 2. Complex function $\psi\in C(\mathbf{R})$ is from the class $BV^1(\mathbf{R})$ if its second derivatives ψ'' coincides with a complex measure ρ_ψ on $\mathbf{R}$ in the sense of distribution theory.

Obviously, the complex measure ρ_ψ is defined uniquely by the function ψ. For example, the class $BV^1(\mathbf{R})$ contains all linear functions for which $\psi''=\rho_\psi=0$. Inversely, for each complex measure ρ there exists a function $\psi\in BV^1(\mathbf{R})$ such that $\rho=\rho_\psi$. This function is defined uniquely modulo a linear function. We denote by ψ^* the class of functions which differ from the function ψ by a linear function. Then we have one-to-one corespondence between complex measures and factor classes ψ^*, $\psi\in BV^1(\mathbf{R})$.

Remark 3. The first derivatives of functions from $BV^1(\mathbf{R})$ are functions with locally bounded variation, which explains the notation BV^1. In particular, for $\psi\in BV^1(\mathbf{R})$ the first derivative ψ' is a locally bounded function which is continuous almost everywhere and has limits $\psi'(s-0)$, $\psi'(s+0)$ for every $s\in\mathbf{R}$. Therefore $BV^1(\mathbf{R})\subset W^1_{\infty,\mathrm{loc}}(\mathbf{R})$, where $W^1_{\infty,\mathrm{loc}}(\mathbf{R})$ is the Sobolev space.

Real function φ defined on $\mathbf{R}$ is said to be convex if

$$\varphi(\alpha s_1+(1-\alpha)s_2)\le\alpha\,\varphi(s_1)+(1-\alpha)\,\varphi(s_2)$$

for any $s_1,s_2\in\mathbf{R}$ and $\alpha\in[0,1]$. This immediatly implies that for convex fuctions

$$\varphi(\alpha s)\le(1-\alpha)\,\varphi(0)+\alpha\,\varphi(s) \tag{8}$$

and

$$\varphi(s+t)+\varphi(s-t)-2\,\varphi(s)\ge 0 \tag{9}$$

for all $s,t\in\mathbf{R}$ and $\alpha\in[0,1]$.

The next lemma characterizes the class of convex functions (see [Hö], v.1, Theorem 4.1.6). We prove it here for the sake of completeness.

Lemma 4. *Function φ is convex if and only if $\varphi \in BV^1(\mathbf{R})$ and ρ_φ is a positive measure.*

Proof. Let φ be convex. Then in view of (9) for a real non-negative test function $f \in \mathcal{D}(\mathbf{R})$ we have

$$0 \le \int [\varphi(s+t) + \varphi(s-t) - 2\,\varphi(s)]\, f(s)\, ds$$
$$= \int \varphi(s)\, [f(s+t) + f(s-t) - 2\, f(s)]\, ds.$$

Dividing by t^2 when $t \to 0$ we obtain $\langle \varphi'', f\rangle \ge 0$. Since a positive distribution is a positive measure this proves the first part of the lemma.

Now assume that $\varphi \in BV^1(\mathbf{R})$ and that φ'' coincides with a positive measure. Let $s_1 < s_2$, $\alpha \in [0,1]$, and

$$f(s) \stackrel{\text{def}}{=} \begin{cases} 0, & \text{for } s \le s_1 \text{ and } s \ge s_2, \\ \alpha(s - s_1), & \text{for } s_1 \le s \le \alpha s_1 + (1-\alpha)s_2, \\ (1-\alpha)(s_2 - s), & \text{for } \alpha s_1 + (1-\alpha)s_2 \le s \le s_2. \end{cases}$$

The function f is non-negative and continuous, and

$$f''(s) = \alpha\, \delta(s - s_1) + (1-\alpha)\, \delta(s - s_2) - \delta(s - \alpha s_1 - (1-\alpha)s_2),$$

where $\delta(\cdot)$ is the delta-function. Therefore

$$\alpha\, \varphi(s_1) + (1-\alpha)\, \varphi(s_2) - \varphi(\alpha s_1 + (1-\alpha)s_2)$$
$$= \int \varphi(s)\, f''(s)\, ds = \langle \varphi'', f\rangle = \int f\, d\rho_\varphi \ge 0.$$

This completes the proof.

Obviously Lemma 4 can be reformulated in the following way : the function φ is convex if and only if $\varphi \in BV^1(\mathbf{R})$ and the first derivative of φ is a non-decreasing function. Now we introduce

Definition 5. Let $\psi \in BV^1(\mathbf{R})$ and φ be a convex function. We say that the function ψ is dominated by φ if $d\rho_\psi = g\, d\rho_\varphi$ with some density $g \in L_\infty(\mathbf{R}, \rho_\varphi)$. In this case we denote $|\psi|_\varphi \stackrel{\text{def}}{=} \|g\|_{L_\infty(\mathbf{R},\rho_\varphi)}$.

Obviously if ψ is dominated by φ then any of the representative from the class ψ^* is dominated by every function from φ^*.

Lemma 6. *Let $\psi \in BV^1(\mathbf{R})$ be dominated by a non-negative convex function φ. Then there exists a linear function l such that*

$$|\, \psi(s) - l(s)\, | \;\le\; |\psi|_\varphi\, \varphi(s), \qquad \forall s \in \mathbf{R}. \tag{10}$$

Proof. Assume first that there exists a point s_0 such that $\varphi'(s_0 - 0) \le 0$ and $\varphi'(s_0 + 0) \ge 0$. Without loss of generality we assume $|\psi|_\varphi = 1$, otherwise we replace φ by $|\psi|_\varphi\, \varphi$. Then $|\rho_\psi(I)| \le \rho_\varphi(I)$ for any bounded interval I. Therefore,

$$|\, \psi'(s \pm 0) - \psi'(s_0 + 0)\, | \le \varphi'(s \pm 0) - \varphi'(s_0 + 0), \qquad s_0 < s, \tag{11}$$

$$|\,\psi'(s\pm 0)-\psi'(s_0-0)\,|\le \varphi'(s_0-0)-\varphi'(s\pm 0),\qquad s<s_0, \tag{12}$$

and for arbitrary $s_1\le s_2$

$$|\,\psi'(s_2)-\psi'(s_1)\,|\le \varphi'(s_2)-\varphi'(s_1). \tag{13}$$

Let us show that there is a constant $C\in\mathbf{R}$, such that

$$|\,\psi'(s)-C\,|\ \le\ |\,\varphi'(s)\,|,\qquad \forall s\in\mathbf{R}. \tag{14}$$

We introduce two intervals I_1 and I_2 such that

$$I_1=[-\psi'(s_0+0)-\varphi'(s_0+0),\ -\psi'(s_0+0)+\varphi'(s_0+0)],$$

$$I_2=[-\psi'(s_0-0)+\varphi'(s_0-0),\ -\psi'(s_0-0)-\varphi'(s_0-0)]. \tag{15}$$

If in (13) we substitute $s_2=s_0+0$ and $s_1=s_0-0$ we have

$$-\psi'(s_0-0)+\varphi'(s_0-0)\le -\psi'(s_0+0)+\varphi'(s_0+0),$$
$$-\psi'(s_0+0)-\varphi'(s_0+0)\le -\psi'(s_0-0)-\varphi'(s_0-0).$$

In particular, this implies that the intersection of I_1 and I_2 is not empty. From (11) we obtain that (14) is satisfied for any $s_0<s$ and $C\in I_1$. Respectively, (14) follows from (12) for any $s<s_0$ and $C\in I_2$. If now $C\in I_1\cap I_2$, then the inequalify (14) holds for all $s<s_0$, $s_0<s$ and therefore for $s=s_0-0$ and $s=s_0+0$.

The inequality (14) implies

$$\begin{aligned}|\,\psi(s)-C\,(s-s_0)-\psi(s_0)|&=|\int_{s_0}^{s}(\psi'(t)-C)\,dt\,|\\ &\le\int_{s_0}^{s}\varphi'(t)\,dt=\varphi(s)-\varphi(s_0)\le\varphi(s),\qquad s>s_0,\end{aligned}$$

$$\begin{aligned}|\,\psi(s)-C\,(s-s_0)-\psi(s_0)|&=|\int_{s}^{s_0}(\psi'(t)-C)\,dt\,|\\ &\le-\int_{s}^{s_0}\varphi'(t)\,dt=\varphi(s)-\varphi(s_0)\le\varphi(s),\qquad s<s_0,\end{aligned}$$

and we obtain (10) with $l(s)=C\,(s-s_0)+\psi(s_0)$.

If there is no such point s_0 then either $\varphi(s)\to 0$ as $s\to-\infty$ or $\varphi(s)\to 0$ as $s\to+\infty$. Let, for example, we have the first case. Then φ' is positive, $\varphi'(s)\to 0$ as $s\to-\infty$ and $\varphi(s)=\int_{-\infty}^{s}\varphi'(t)\,dt$. From the inequality, obtained by analogy with (11), we have

$$|\,\psi'(s)-\psi'(s_1+0)\,|\le\varphi'(s)-\varphi'(s_1+0),\qquad s_1\le s.$$

This implies that there exists the limit $C=\lim_{s_1\to-\infty}\psi'(s_1+0)$ and

$$|\,\psi'(s)-C\,|\le\varphi'(s).$$

Therefore if $C_1=\lim_{s\to-\infty}(\psi(s)-Cs)$ we have

$$|\,\psi(s)-Cs-C_1)|=|\int_{-\infty}^{s}(\psi'(t)-C)\,dt\,|\le\int_{-\infty}^{s}\varphi'(t)\,dt=\varphi(s),$$

and we have (10) with $l(s)=Cs+C_1$. The lemma is proved.

The next proposition characterizes the dominating property not via measures but via functions themselves.

Proposition 7. *Function $\psi \in BV^1(\mathbf{R})$ is dominated by the convex function φ if and only if*

$$(16)\quad |\,\psi(s+t)+\psi(s-t)-2\,\psi(s)\,| \le C\,(\varphi(s+t)+\varphi(s-t)-2\,\varphi(s)), \quad \forall s,t\in\mathbf{R},$$

for some constant C. The minimal constant C satisfying (16) coincides with $|\psi|_\varphi$.

Proof. Let us assume first that (16) is fulfilled with some constant $C \ge 0$. Let $\psi_1 = \operatorname{Re}\psi$, $\psi_2 = \operatorname{Im}\psi$. Then for any real non-negative test function f we have

$$\begin{aligned} -\,C_k \int [\varphi(s+t)+\varphi(s-t)-2\,\varphi(s)]\,f(s)\,ds \\ \le \int [\psi_k(s+t)+\psi_k(s-t)-2\,\psi_k(s)]\,f(s)\,ds \\ \le C_k \int [\varphi(s+t)+\varphi(s-t)-2\,\varphi(s)]\,f(s)\,ds, \end{aligned}$$

where $k=1,2$ and C_k are some constants such that $C=\sqrt{C_1^2+C_2^2}$. Dividing by t^2 when $t\to 0$ we obtain

$$(17)\qquad -C_k \int f\,d\rho_\varphi \le \int f\,d\rho_{\psi_k} \le C_k \int f\,d\rho_\varphi, \qquad k=1,2.$$

This implies that the measure $\rho_\psi = \rho_{\psi_1} + i\,\rho_{\psi_2}$ is absolutely continuous with respect to ρ_φ. Therefore, by the Radon–Nikodym theorem we have $d\rho_\psi = g\,d\rho_\varphi$ with some complex density $g \in L_{1,\mathrm{loc}}(\mathbf{R},\rho_\varphi)$.

Now from (17) it also follows that $|\int f\,d\rho_\psi| = |\int f\,g\,d\rho_\varphi| \le C\int |f|\,d\rho_\varphi$ for any (not necesserily non-negative) test function f. Hence, the function g defines a linear continuous functional on the space $L_1(\mathbf{R},\rho_\varphi)$ which norm is estimated by C, and then $g\in L_\infty(\mathbf{R},\rho_\varphi)$, $\|g\|_{L_\infty(\mathbf{R},\rho_\varphi)} \le C$.

It remains to prove the necessity. Let $d\rho_\psi = g\,d\rho_\varphi$ with $g\in L_\infty(\mathbf{R},\rho_\varphi)$, and

$$C_1 = \|\operatorname{Re} g\|_{L_\infty(\mathbf{R},\rho_\varphi)}, \qquad C_2 = \|\operatorname{Im} g\|_{L_\infty(\mathbf{R},\rho_\varphi)}.$$

Then the functions

$$(18)\qquad \psi_1^\pm \overset{\text{def}}{=} C_1\,\varphi \pm \operatorname{Re}\psi, \qquad \psi_2^\pm \overset{\text{def}}{=} C_2\,\varphi \pm \operatorname{Im}\psi$$

are convex because their second derivatives are positive measures, and so for each of them (9) holds. These estimates altogether mean exactly that

$$|\operatorname{Re}\psi(s+t)+\operatorname{Re}\psi(s-t)-2\operatorname{Re}\psi(s)\,| \le C_1\,(\varphi(s+t)+\varphi(s-t)-2\,\varphi(s)), \quad \forall s,t\in\mathbf{R},$$

$$|\operatorname{Im}\psi(s+t)+\operatorname{Im}\psi(s-t)-2\operatorname{Im}\psi(s)\,| \le C_2\,(\varphi(s+t)+\varphi(s-t)-2\,\varphi(s)), \quad \forall s,t\in\mathbf{R},$$

which implies

$$|\,\psi(s+t)+\psi(s-t)-2\,\psi(s)\,| \le C_0\,(\varphi(s+t)+\varphi(s-t)-2\,\varphi(s)), \qquad \forall s,t\in\mathbf{R}$$

with $C_0=\sqrt{C_1^2+C_2^2}=\|g\|_{L_\infty(\mathbf{R},\rho_\varphi)}$. The proof is complete.

Example 8. For the convex function $\varphi(s) = s^2/2$ the measure ρ_φ coincides with the Lebesgue measure on $\mathbf{R}$. In this case $\psi \in BV^1(\mathbf{R})$ is dominated by φ if only if $\psi \in W^2_{\infty,\mathrm{loc}}(\mathbf{R})$ and $\psi'' \in L_\infty(\mathbf{R})$, and $|\psi|_\varphi = \|\psi''\|_{L_\infty(\mathbf{R})}$.

Futher on we use the following well known result.

Theorem 9 (Jensen inequality). *Let ν be a positive measure on $\mathbf{R}$ such that $\nu(\mathbf{R}) = 1$ and $\int s\,d\nu < \infty$, and φ be a convex function from $L_1(\mathbf{R}, \nu)$. Then*

$$\int \varphi(s)\,d\nu \;-\; \varphi\big(\int s\,d\nu\big) \;\geq\; 0.$$

Corollary 10. *Let us assume that in Theorem 8 $\nu(\mathbf{R}) \overset{\mathrm{def}}{=} c_\nu \leq 1$. Then*

$$(1 - c_\nu)\,\varphi(0) + \int \varphi(s)\,d\nu - \varphi\big(\int s\,d\nu\big) \;\geq\; 0. \tag{19}$$

Proof. If we apply (8) and the Jensen inequality we have

$$\varphi\big(\int s\,d\nu\big) \leq (1 - c_\nu)\,\varphi(0) + \varphi\big(\int s\,c_\nu^{-1}\,d\nu\big) \leq (1 - c_\nu)\,\varphi(0) + \int \varphi(s)\,d\nu,$$

which proves the corollary.

Corollary 11. *Let ν be a positive measure on $\mathbf{R}$ such that $\nu(\mathbf{R}) \overset{\mathrm{def}}{=} c_\nu \leq 1$, $\int s\,d\nu < \infty$, and $\psi \in BV^1(\mathbf{R}) \cap L_1(\mathbf{R}, \nu)$ be dominated by a convex function $\varphi \in L_1(\mathbf{R}, \nu)$. Then*

$$\begin{aligned} &\big|\,(1 - c_\nu)\,\psi(0) + \int \psi(s)\,d\nu - \psi\big(\int s\,d\nu\big)\,\big| \\ &\qquad\leq\; |\psi|_\varphi\,\Big((1 - c_\nu)\,\varphi(0) + \int \varphi(s)\,d\nu - \varphi\big(\int s\,d\nu\big)\Big). \end{aligned} \tag{20}$$

Proof. As in the proof of Proposition 7 we introduce the convex function (18), and apply to each of them the inequality (19). Then we obtain the inequalities

$$\begin{aligned} &\big|\,(1 - c_\nu)\,\mathrm{Re}\,\psi(0) + \int \mathrm{Re}\,\psi(s)\,d\nu - \mathrm{Re}\,\psi\big(\int s\,d\nu\big)\,\big| \\ &\qquad\leq C_1\,\Big((1 - c_\nu)\,\varphi(0) + \int \varphi(s)\,d\nu - \varphi\big(\int s\,d\nu\big)\Big), \end{aligned}$$

$$\begin{aligned} &\big|\,(1 - c_\nu)\,\mathrm{Im}\,\psi(0) + \int \mathrm{Im}\,\psi(s)\,d\nu - \mathrm{Im}\,\psi\big(\int s\,d\nu\big)\,\big| \\ &\qquad\leq C_2\,\Big((1 - c_\nu)\,\varphi(0) + \int \varphi(s)\,d\nu - \varphi\big(\int s\,d\nu\big)\Big), \end{aligned}$$

which are equivalent to (20).

3. Berezin–Lieb inequality. We study operators of the form

$$G(B,P;\psi) \stackrel{\text{def}}{=} \psi(0)\,(I-P^*P) + P^*\psi(B)P - \psi(B_P),$$

where $\psi \in BV^1(\mathbf{R})$. Note that under the conditions (C_1) and (C_2) the operator $G(B,P;\psi)$ is well defined and equal to zero for linear functions ψ. When B is unbounded, for some functions ψ the expression $P^*\psi(B)P$ or $G(B,P;\psi)$ might make no sense. Therefore we introduce an additional restriction.

(C_3) The set $\mathcal{D}_\psi = \{\, \xi \in H_0 : P\xi \in \mathcal{D}(\psi(B)) \,\} \cap \mathcal{D}(\psi(B_P)) \cap \mathcal{D}(B_P)$ is dense in H_0 and the operator $G(B,P;\psi)$ defined on $\mathcal{D}_\psi$ is bounded.

Under this conditions we extend the operator $G(B,P;\psi)$ to the whole Hilbert space H_0, and then $P^*\psi(B)P$ is a well defined selfadjoint operator with domain $\mathcal{D}(\psi(B_P))$. Obviously, if the condition (C_3) is satisfied for a function ψ then it is also satisfied for any $\psi_1 \in \psi^*$ and $\mathcal{D}_{\psi_1} = \mathcal{D}_\psi$, $G(B,P;\psi_1) = G(B,P;\psi)$. Besides, if for some convex function φ the set $\mathcal{D}_\varphi$ is dense then in view of Lemma 6 for any ψ dominated by φ the set $\mathcal{D}_\psi$ is also dense.

We denote by $\sigma(B_P)$ the spectrum of the selfadjoint operator B_P and by $\sigma_c(B_P)$ its continuous part. Let $\operatorname{ch}\sigma_c(B_P)$ be the closed convex hull of $\sigma_c(B_P)$, and $\operatorname{Int}\operatorname{ch}\sigma_c(B_P)$ be its interior. (The last set coincides with the interior of the minimal interval containing $\sigma_c(B_P)$.)

Theorem 12. *Let the conditions (C_1)–(C_2) be fulfiled. Let $\psi \in BV^1(\mathbf{R})$ be dominated by a convex function φ such that $\rho_\varphi\big(\operatorname{Int}\operatorname{ch}\sigma_c(B_P)\big) = 0$. Assume that the condition (C_3) is fulfilled for both φ and ψ and that the operators $G(B,P;\varphi)$, $G(B,P;\psi)$ are from the trace class $\mathfrak{S}_1$. Then*

$$|\operatorname{Tr} G(B,P;\psi)| \;\le\; |\psi|_\varphi \operatorname{Tr} G(B,P;\varphi). \tag{21}$$

Proof. Let $\varphi_0 \in \varphi^*$ be a non-negative representative, and $\psi_0 \in \psi^*$ be such representative that $|\psi_0| \le |\psi|_\varphi\,\varphi_0$ (see Lemma 6). If $\operatorname{Int}\operatorname{ch}\sigma_c(B_P)$ is not empty we assume in addition that $\varphi = 0$ on $\operatorname{ch}\sigma_c(B_P)$. Then ψ is also equal to zero $\operatorname{ch}\sigma_c(B_P)$.

For every $\xi \in \mathcal{D}_{\varphi_0}$ we have

$$\begin{aligned} \int \varphi_0(\lambda)\,(dE_B(\lambda)P\xi,P\xi) &= (\varphi_0(B)P\xi,P\xi) \\ &= (G(B,P;\varphi_0)\xi,\xi)_0 + (\varphi_0(B_P)\xi,\xi)_0. \end{aligned} \tag{22}$$

Since the function φ_0 is non-negative and the operator $G(B,P;\varphi_0)$ is bounded, then (22) can be extended on $\xi \in \mathcal{D}(\varphi_0(B_P))$. For chosen representative ψ_0 we have $\mathcal{D}(\varphi_0(B_P)) \subset \mathcal{D}(\psi_0(B_P))$ and

$$\begin{aligned} \int \psi_0(\lambda)\,(dE_B(\lambda)P\xi,P\xi) &= (\psi_0(B)P\xi,P\xi)_0 \\ &= (G(B,P;\psi_0)\xi,\xi)_0 + (\psi_0(B_P)\xi,\xi)_0 \end{aligned} \tag{23}$$

is also valid for $\xi \in \mathcal{D}(\varphi_0(B_P))$.

Let Π_c be the spectral projection of the operator B_P corresponding to the closed interval $\operatorname{ch}\sigma_c(B_P)$. We choose an orthonormed basis $\{\xi_k\}$ in the subspace $(I-\Pi_c)H_0$ formed by eigenfunctions ξ_k of the operator B_P with eigenvalues λ_k lying outside $\operatorname{ch}\sigma_c(B_P)$. It is clear that ξ_k are contained in $\mathcal{D}(\varphi_0(B_P)) \subset \mathcal{D}(\psi_0(B_P))$. We choose also an orthonormed basis $\{\eta_j\}$ in the subspace $\Pi_c H_0$ with $\eta_j \in \mathcal{D}(\varphi_0(B_P))$. Then $\{\xi_k, \eta_j\}$ form an orthonormed basis in the whole space H_0.

Let ν_k be the positive measures with $d\nu_k = (dE_B(\lambda)P\xi_k, P\xi_k)$. Then

$$(\varphi_0(B_P)\xi_k, \xi_k)_0 = \varphi_0\big((B_P\xi_k, \xi_k)_0\big) = \varphi_0(\lambda_k),$$

$$(\psi_0(B_P)\xi_k, \xi_k)_0 = \psi_0\big((B_P\xi_k, \xi_k)_0\big) = \psi_0(\lambda_k),$$

and by (22), (23)

$$(\varphi_0(B)P\xi_k, P\xi_k) = \int \varphi_0(\lambda)\, d\nu_k,$$

$$(\psi_0(B)P\xi_k, P\xi_k) = \int \psi_0(\lambda)\, d\nu_k.$$

Therefore, applying (20) we obtain

$$|\,(G(B,P;\psi_0)\xi_k, \xi_k)_0\,| \;\le\; |\psi|_\varphi\,\big(\,(G(B,P;\varphi_0)\xi_k, \xi_k)_0\,\big). \tag{24}$$

Since $\varphi_0(B_P)\eta_j = 0$ and $\psi_0(B_P)\eta_j = 0$, we have

$$(G(B,P;\varphi_0)\eta_j, \eta_j)_0 = \varphi(0)\,((I - P^*P)\eta_j, \eta_j)_0 + (\varphi_0(B)P\eta_j, P\eta_j),$$

$$(G(B,P;\psi_0)\eta_j, \eta_j)_0 = \psi(0)\,((I - P^*P)\eta_j, \eta_j)_0 + (\psi_0(B)P\eta_j, P\eta_j).$$

Then in view of (22), (23) and the inequality $|\psi_0| \le |\psi|_\varphi\,\varphi_0$ we obtain

$$|\,(G(B,P;\psi_0)\eta_j, \eta_j)_0\,| \;\le\; |\psi|_\varphi\,(G(B,P;\varphi_0)\eta_j, \eta_j)_0.$$

Summing up these inequalities and inequalities (24) we obtain (21). The proof is complete.

If $\psi = \varphi$ then Theorem 12 is a generalization of the inequality obtained in [B] and [L].

Corollary 13 (generalized Berezin–Lieb inequality). *Let the conditions* (C_1)–(C_2) *be fulfiled. Let* φ *be a convex function such that* $\rho_\varphi\big(\operatorname{Int}\operatorname{ch}\sigma_c(B_P)\big) = 0$. *Assume that* ($C_3$) *is fulfilled for the function* φ *and that* $G(B,P;\varphi) \in \mathfrak{S}_1$. *Then*

$$\operatorname{Tr} G(B,P;\varphi) \;\ge\; 0. \tag{25}$$

The conditions of Theorem 12 are rather complicated. But most of them are needed only in order to define the unbounded operators. In particular, if B is bounded then (C_1)–(C_3) are fulfilled automatically, and Theorem 12 can be reformulated in the following way.

Corollary 14. *Let the operator* B *be bounded. Assume that* $\psi \in BV^1(\mathbf{R})$ *is dominated by a convex function* φ *such that* $\rho_\varphi\big(\operatorname{Int}\operatorname{ch}\sigma_c(B_P)\big) = 0$, *and* $G(B,P;\varphi)$, $G(B,P;\psi)$ *are from* $\mathfrak{S}_1$. *Then the estimate (21) holds.*

Let us denote by $\sigma_{\mathrm{ess}}(B_P)$ the essential spectrum of B_P. We have $\sigma_c(B_P) \subset \sigma_{\mathrm{ess}}(B_P)$, and therefore $\operatorname{ch}\sigma_c(B_P) \subset \operatorname{ch}\sigma_{\mathrm{ess}}(B_P)$. The following proposition gives another set of sufficient conditions to Theorem 12.

Proposition 15. *Let conditions (C_1)–(C_2) be fulfilled, and condition (C_3) be fulfilled for a non-negative convex function φ such that the operator $\varphi(0)\,(I-P^*P)+P^*\varphi(B)P$ is from the trace class $\mathfrak{S}_1$. Then*

(1) *φ is equal to zero on the set $\operatorname{ch}\sigma_{\mathrm{ess}}(B_P)$;*
(2) *$\varphi(B_P)\in\mathfrak{S}_1$, and, consequently, $G(B,P;\varphi)\in\mathfrak{S}_1$;*
(3) *for any function $\psi\in BV^1(\mathbf{R})$ dominated by φ the condition (C_3) is fulfilled and $G(B,P;\psi)\in\mathfrak{S}_1$.*

Proof. Let θ_k be eigenfunctions of the operator $\varphi(0)\,(I-P^*P)+P^*\varphi(B)P$ corresponding to eigenvalues μ_k, $|\mu_1|\le|\mu_2|\le\dots$ By (19) for any $\xi\in H_0$ we have

$$
\begin{aligned}
(26)\quad &\varphi(0)\,((I-P^*P)\xi,\xi)_0+(P^*\varphi(B)P\xi,\xi)_0\\
&\qquad=\varphi(0)\,\big(1-\int(dE_B(\lambda)P\xi,P\xi)\big)+\int\varphi(\lambda)\,(dE_B(\lambda)P\xi,P\xi)\\
&\qquad\qquad\ge\varphi\big(\int\lambda\,(dE_B(\lambda)P\xi,P\xi)\big)=\varphi\big((B_P\xi,\xi)_0\big).
\end{aligned}
$$

since the operator $\varphi(0)\,(I-P^*P)+P^*\varphi(B)P$ is compact, (26) implies that there exists a positive sequences $\varepsilon_j\to0$ such that

$$|\varphi\big((B_P\xi,\xi)_0\big)|\le\ \varepsilon_j$$

for any normed vector ξ which is orthogonal to all θ_k with $k\le j$. By the minimax principle (see for example [RS], Theorem XIII.1) it follows now that $\varphi(s)\to0$ as $s\to\pm\infty$ if B_P is unbounded from above or from below respectively, and that $\varphi=0$ on $\sigma_{\mathrm{ess}}(B_P)$. Obviously the set of zeros of a convex function is necesserily convex, and therefore we have proved (1).

Let ξ_k be the orthonormed eigenfunctions of B_P with eigenvalues λ_k lying outside $\operatorname{ch}\sigma_{\mathrm{ess}}(B_P)$. By (26) we have

$$\varphi(0)\,((I-P^*P)\xi_k,\xi_k)_0+(P^*\varphi(B)P\xi_k,\xi_k)_0\ge\varphi\big((B_P\xi_k,\xi_k)_0\big)=\varphi(\lambda_k).$$

Since $\varphi(0)\,(I-P^*P)+P^*\varphi(B)P\in\mathfrak{S}_1$ the positive series $\sum\varphi(\lambda_k)$ converges, which means that $\varphi(B_P)\in\mathfrak{S}_1$.

To prove the third assertion of the lemma we choose using Lemma 6 a function $\psi_0\in\psi^*$ such that $|\psi|\le|\psi|_\varphi\,\varphi$. Then for any ortonormed basis $\{\zeta_k\}$ in H_0 we have

$$
\begin{aligned}
&|\,\psi_0(0)\,((I-P^*P)\zeta_k,\zeta_k)_0+(P^*\psi_0(B)P\zeta_k,\zeta_k)_0\,|\\
&\qquad\le|\,\psi_0(0)\,((I-P^*P)\zeta_k,\zeta_k)_0\,|+|\,(P^*\psi_0(B)P\zeta_k,\zeta_k)_0\,|\\
&\qquad\qquad\le|\psi|_\varphi\,\big(\varphi(0)\,((I-P^*P)\zeta_k,\zeta_k)_0+(P^*\varphi(B)P\zeta_k,\zeta_k)_0\big),
\end{aligned}
$$

$$|\,(\psi_0(B_P)\zeta_k,\zeta_k)_0\,|\le|\psi|_\varphi\,(\varphi(B_P)\zeta_k,\zeta_k)_0.$$

These estimates imply (see [RS], ch.VI, problem 26) that $\psi_0(0)\,(I-P^*P)+P^*\psi_0(B)P$ and $\psi_0(B_P)$ are from the trace class. Since the operator $G(B,P;\cdot)$ is independent of the choice of representative from the factor-class ψ^*, this completes the proof.

Remark 16. In fact, proving (3) we have obtained a more precise result. Namely, if $\psi\in BV^1(\mathbf{R})$ is dominated by φ then for a representative $\psi_0\in\psi^*$ such that $|\psi_0|\le|\psi|_\varphi\,\varphi$ both operators $\psi_0(0)\,(I-P^*P)+P^*\psi_0(B)P$ and $\psi_0(B_P)$ are from the trace class.

Proposition 15 with $\varphi(s)=s^2/2$ immediately implies

Corollary 16. *Let BP be from the Hilbert–Schmidt class $\mathfrak{S}_2$. Then*

(1) *either $\sigma_{\mathrm{ess}}(B_P) = \{0\}$ or $\sigma_{\mathrm{ess}}(B_P) = \emptyset$;*
(2) *for any function $\psi \in W^2_{\infty,\mathrm{loc}}(\mathbf{R})$ such that $\psi'' \in L_\infty(\mathbf{R})$, the condition ($C_3$) is fulfilled and $G(B,P;\psi) \in \mathfrak{S}_1$.*

From Theorem 12 and Corollary 16 we obtain

Corollary 17. *Let $H_0 = H$ and $P : H \to H$ be an orthogonal projection in H. If the operator BP is from the Hilbert-Schmidt class, then for any function ψ from the Sobolev class $W^2_{\infty,\mathrm{loc}}(\mathbf{R})$ such that $\psi'' \in L_\infty(\mathbf{R})$ we have*

$$|\operatorname{Tr}\Big(P\psi(B)P - P\psi(PBP)P\Big)| \;\le\; \frac{1}{2}\,\|\psi''\|_{L_\infty(K)}\,\|PB(I-P)\|^2_{\mathfrak{S}_2}\,.$$

Remark 18. When we deal with a fixed operator B it is sufficient to define the functions φ and ψ only on the set

$$\bigcup_{0\le t\le 1} t\,\sigma(B) \;\subset\; \mathbf{R}\,.$$

Then all the conditions involving φ and ψ are obviously needed to be fulfilled only on this set.

References

[B] F. Berezin, *Convex functions of operators.*, Mat.sb. **88** (1972), 268–276. (Russian)

[BS] M. Sh. Birman and M. Z. Solomyak, *Spectral theory of selfadjoint operators in Hilbert space.*, D. Reidel Publ. Comp., 1987.

[BSh] F. Berezin and M. Shubin, *The Schrödinger equation*, Kluwer Acad. Publishers, Dordrecht, Boston, 1991.

[Hö] L. Hörmander, *The analysis of Linear Partial Differential operators. I*, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983.

[L] E.H. Lieb, *The classical limit of quantum spin systems*, Comm.Math.Phys. **31** (1973), 327–340.

[RS] Reed and Simon, *Methods of mordern mathematical physics v.4.*, Academic Press, New York, San Francisco, London, 1978.

[S] B. Simon, *The classical limit of quantum partition functions*, Comm.Math.Phys. **71** (1980), 247–276.

Department of Mathematics, Royal Institute of Technology, S-100 44 Stockholm, Sweden

E-mail address: `laptev@math.kth.se`

Department of Mathematics, King's College London, Strand, London WC2R 2LS, UK

E-mail address: `udah@bay.cc.kcl.ac.uk`