BULLETIN OF THE
INSTITUTE OF MATHEMATICS
ACADEMIA SINICA
Volume 29, Number 4, December 2001

ON SHARPENING AND CHARACTERIZING OPERATOR INEQUALITIES

BY

C.-S. LIN (林嘉祥)

Dedicated to Professor Hong-Buey Lim on her retirement

Abstract. In [14] we initiated a study of covariance and variance for two bounded linear operators on a Hilbert space, proving, there, that the c-v inequality holds which is equivalent to the Cauchy-Schwarz inequality. In this paper we show that the notion of the c-v inequality is an important concept in the study of sharpening and characterizing operator inequalities. We shall introduce, finally, the extended covariance and variance.

1. Notations and introduction. In what follows H is a Hilbert space over the field **C** of complex numbers; B(H) is the algebra of all bounded linear operators on H; I denotes the identity operator in B(H); and the complex conjugate of $\alpha \in \mathbf{C}$ is $\overline{\alpha}$. In [14] we initiated a study of the covariance-variance inequality (the c-v inequality in short). Let us repeat some definitions from [14]. The covariance of S and $T \in$ B(H) is a mapping $\mathrm{Cov}_y(S,T)$: H $\to \mathbf{C}$ defined by

$$\mathrm{Cov}_y(S,T) = \|y\|^2(Sx,Tx) - (Sx,y)(y,Tx)$$

Received by the editors August 10, 2000 and in revised form November 23, 2000.
AMS 2000 Subject Classification: 47A63.
Key words and phrases: Covariance-variance inequality, Reid's inequality, Furuta inequality, Löwner-Heinz formula, extended covariance-variance inequality.

for every $x, y \in$ H. It is called the variance of S if $T = S$, i.e.,

$$\mathrm{Var}_y(S) = \mathrm{Cov}_y(S, S) = \|y\|^2\|Sx\|^2 - |(Sx, y)|^2.$$

The covariance and variance enjoy the following properties. For $S, T, R \in$ B(H) and $y \in$ H: (1) $\mathrm{Var}_y(S) \geq 0$; (2) $\mathrm{Var}_y(S) = 0$ if and only if y and Sx are proportional; (3) $\mathrm{Cov}_y(R + S, T) = \mathrm{Cov}_y(R, T) + \mathrm{Cov}_y(S, T)$; (4) $\mathrm{Cov}_y(\lambda S, T) = \lambda \mathrm{Cov}_y(S, T)$ for $\lambda \in \mathbf{C}$; and (5) $\overline{\mathrm{Cov}_y(T, S)} = \mathrm{Cov}_y(S, T)$. It was proved in [14, Theorem 1], among others, that the c-v inequality

$$|\mathrm{Cov}_y(S, T)|^2 \leq \mathrm{Var}_y(S)\mathrm{Var}_y(T)$$

holds for every $x, y \in$ H, which is equivalent to the Cauchy-Schwarz inequality, i.e., $|(x, y)| \leq \|x\|\|y\|$. Also, the c-v equality holds if and only if $y = (S - \lambda T)x$ for $\lambda \in \mathbf{C}$.

The primary object of the present paper is to show that the c-v inequality could be used to sharpen and characterize operator inequalities. We first start with basic sharpening theorems and obtain their consequences; and then using Theorem 1 obtained from the c-v inequality we are able to establish our goal. We prove that inequalities in Theorem 1 are equivalent to inequalities expressed in terms of covariance and variance. Finally, we shall introduce extended covariance and variance.

2. Basic sharpening theorems and their consequences. A special operator $P_e \in$ B(H) will be used throughout this paper, which is defined by $P_e x = (x, e)e$ for every $x \in$ H and e is a fixed nonzero vector. We shall make frequent use of a simple observation that for every nonzero $x \in$ H there exists a vector orthogonal to x; take $z \in$ H for example with $z = y - \frac{(y,x)x}{\|x\|^2}$, $0 \neq y \in$ H. It should be pointed out that (1) of Theorem 1 below is crucial in the sense that it is the main formula used to sharpen and characterize operator inequalities and equalities in this paper.

Theorem 1. *Let* $S, T \in \mathrm{B(H)}$, $e, x, y \in \mathrm{H}$ *with* $\|e\| = 1$. *Then*

$$|(Sx,Ty)-(Sx,e)(e,Ty)|^2 \le [\|Sx\|^2-|(Sx,e)|^2][\|Ty\|^2-|(Ty,e)|^2],$$
$$\text{or, } |(Sx,Ty)-(e,Ty)(Sx,e)| \le \|Sx-(Sx,e)e\|\|Ty-(Ty,e)e\|$$

The equality holds if and only if $Ty = Sx - \lambda e$, $\lambda \in \mathbf{C}$.

Moreover, there are two particular cases:

(1) *If* $(Ty, e) = 0$, *then*

$$|(Sx,Ty)|^2+\|Ty\|^2|(Sx,e)|^2 \le \|Sx\|^2\|Ty\|^2,$$
$$\text{or, } |(Sx,Ty)| \le \|Sx-(Sx,e)e\|\|Ty\|.$$

The equality holds if and only if $Ty = Sx - (Sx, e)e$.

(2) *If* $(Sx, Ty) = 0$, *then*

$$\|Sx\|^2|(Ty,e)|^2+\|Ty\|^2|(Sx,e)|^2 \le \|Sx\|^2\|Ty\|^2,$$
$$\text{or, } |(e,Ty)(Sx,e)| \le \|Sx-(Sx,e)e\|\|Ty-(Ty,e)e\|.$$

The equality holds if and only if $Ty = Sx - \frac{\|Sx\|^2 e}{(e,Sx)}$, $(e, Sx) \neq 0$.

Proof. Let $I - P_e = R \in \mathrm{B(H)}$, and consider the c-v inequality

$$|\mathrm{Cov}_{Ty}(R,P_e)|^2 \le \mathrm{Var}_{Ty}(R)\mathrm{Var}_{Ty}(P_e)$$

acting on Sx, i.e.,

$$\begin{aligned}&|\|Ty\|^2(R(Sx),P_e(Sx))-(R(Sx),Ty)(Ty,P_e(Sx))|^2\\ \le\ &[\|Ty\|^2\|R(Sx)\|^2-|(R(Sx),Ty)|^2]\end{aligned}$$

$$\cdot[\|Ty\|^2\|P_e(Sx)\|^2 - |(P_e(Sx), Ty)|^2].$$

Notice that $P_e(Sx) = (Sx, e)e$; $R(Sx) = Sx - (Sx, e)e$; and $(R(Sx), P_e(Sx)) = 0$. Also, it can be shown that $\|Sx - (Sx, e)e\|^2 = \|Sx\|^2 - |(Sx, e)|^2$. Thus,

$$\begin{aligned} & |(Sx, Ty) - (Sx, e)(e, Ty)|^2|(Sx, e)|^2|(Ty, e)|^2 \\ \leq\ & \{\|Ty\|^2\ [\|Sx\|^2 - |(Sx, e)|^2] - |(Sx, Ty) - (Sx, e)(e, Ty)|^2\} \\ & \cdot\{\|Ty\|^2|(Sx, e)|^2 - |(Sx, e)(e, Ty)|^2\}, \end{aligned}$$

and the required inequality follows.

The equality holds if and only if $Ty = (R - \alpha P_e)Sx$, $\alpha \in \mathbf{C}$, or, $Ty = Sx - (Sx, e)e - \alpha(Sx, e)e = Sx - \lambda e$, $\lambda \in \mathbf{C}$.

In particular, the inequality (1) is easily obtained. Since $0 = (Ty, e) = (Sx, e) - \lambda$, so $\lambda = (Sx, e)$ which yields the equality condition.

(2) The required inequality is obtained by letting $(Sx, Ty) = 0$ in the original inequality. As for equality condition, since $0 = (Ty, Sx) = (Sx - \lambda e, Sx) = \|Sx\|^2 - \lambda(e, Sx)$ which yields $\lambda = \frac{\|Sx\|^2}{(e,Sx)}$, $(e, Sx) \neq 0$. If $(Sx, Ty) = (e, Sx) = 0$, then Ty and e are proportional.

We remark firstly that a much shorter proof due to the extended covariance and variance is indicated at the end of this paper. Secondly, direct proofs of (1) and (2) in Theorem 1 are possible. In both cases just consider the c-v inequality $|\mathrm{Cov}_{Ty}(S, P_e)|^2 \leq \mathrm{Var}_{Ty}(S)\mathrm{Var}_{Ty}(P_e)$ and simplify it.

Observe that the Ostrowski inequality in terms of vectors in H is

$$\frac{\|a\|^2}{\|a\|^2\|b\|^2 - |(a, b)|^2} \leq \|z\|^2$$

if $(a, z) = 0$ and $(b, z) = 1$ for nonzero vectors a, b and z. And the equality holds if and only if $a = b - \frac{z}{\|z\|^2}$, equivalently, $z = \frac{\|a\|^2 b - (b,a)a}{\|a\|^2\|b\|^2 - |(a,b)|^2}$ (this condition is shown in [4, Theorem 4.1]). It is indeed a special case of (1) in Theorem 1, where we set $Ty = a$, $Sx = b$, and $e = z/\|z\|$. A different expression of the inequality above is: $\|z\|^2|(a,b)| \leq \|\|z\|^2 b - z\|\|a\|$ by (1) in Theorem 1. It can be checked that the two conditions for the equality are actually equivalent to each other due to assumptions $(a,z) = 0$ and $(b,z) = 1$; the first condition is due to (1) in Theorem 1, and the second one was proved in [4, Theorem 4.1]. The Ostrowski inequality was used to produce operator inequalities in [4].

For the next result which is an extension of (1) in Theorem 1 we note that $\|2(a,b)b - \|b\|^2 a\| = \|a\|\,\|b\|^2$ for $a, b \in \mathrm{H}$, by a straightforward computation.

Corollary 1. *Let $R, S, T \in \mathrm{B(H)}$, $e, x, y, z \in \mathrm{H}$, with $\|e\| = 1$, and $(Ty, e) = 0$. Then*

$$\begin{aligned} & |2(Rz,Sx)(Sx,Ty) - \|Sx\|^2(Rz,Ty)|^2 \\ & + \|Ty\|^2 |2(Rz,Sx)(Sx,e) - \|Sx\|^2(Rz,e)|^2 \\ \leq \; & \|Ty\|^2\|Rz\|^2\|Sx\|^4. \end{aligned}$$

The equality holds if and only if

$$Ty = 2(Rz,Sx)[Sx - (Sx,e)e] + \|Sx\|^2[(Rz,e)e - Rz].$$

In particular,

$$|(Rz,Sx)(Sx,Ty)| \leq \frac{\|Rz\|\|Ty\| + |(Rz,Ty)|}{2}\|Sx\|^2.$$

Proof. The inequality is due to a replacement of Sx by $2(Rz,Sx)Sx - \|Sx\|^2 Rz$ in (1) of Theorem 1, and notice that $\|2(Rz,Sx)Sx - \|Sx\|^2 Rz\| = \|Rz\|\|Sx\|^2$ by a remark above; $(2(Rz,Sx)Sx - \|Sx\|^2 Rz, Ty) = 2(Rz,Sx)(Sx,Ty) - \|Sx\|^2(Rz,Ty)$; and $(2(Rz,Sx)Sx - \|Sx\|^2 Rz, e) = 2(Rz,Sx)(Sx,e) - \|Sx\|^2(Rz,e)$.

The equality holds if and only if

$$Ty = 2(Rz, Sx)Sx - \|Sx\|^2 Rz - [2(Rz, Sx)(Sx, e) - \|Sx\|^2(Rz, e)]e$$

by (1) in Theorem 1, and we have the desired condition after simplification.

In particular, let the second term on the left side equal to 0. Then, $2|(Rz, Sx)(Sx, Ty)| - \|Sx\|^2|(Rz, Ty)| \leq |2(Rz, Sx)(Sx, Ty) - \|Sx\|^2(Rz, Ty)| \leq \|Sx\|^2\|Rz\|\|Ty\|$ and the desired particular case follows.

Let $Rz = Sx$ in Corollary 1. Then we get precisely (1) of Theorem 1. In other words, Corollary 1 and (1) in Theorem 1 are equivalent. On the other hand, let $R = S = T = I$ in the special case of Corollary 1. Then

$$|(z, x)(x, y)| \leq \frac{\|z\|\|y\| + |(z, y)|}{2}\|x\|^2$$

which is an extension of the Cauchy-Schwarz inequality, and was first proved in [1]. Later on the proof was simplified in [3] by way of an orthogonal projection; and there some applications of the inequality can be found.

The well-known Selberg inequality is as follows: $\sum_i \frac{|(x,z_i)|^2}{\sum_j |(z_i,z_j)|} \leq \|x\|^2$, $x \in \mathrm{H}$ and nonzero vectors $\{z_i\}_i \subseteq \mathrm{H}$, $i = 1, 2, \ldots, n$. The inequality is refined in [5, Lemma 1] and [2, Lemma 3.9]. More precisely, if $x, y \in \mathrm{H}$ and $(y, z_i) = 0$ for given nonzero vectors $\{z_i\}_i \subseteq \mathrm{H}$, $i = 1, 2, \ldots, n$, then

$$|(x, y)|^2 + \|y\|^2 \sum_i \frac{|(x, z_i)|^2}{\sum_j |(z_i, z_j)|} \leq \|x\|^2\|y\|^2.$$

Our next result is a further improvement of the above, and it will be used in Corollary 3 to improve another inequality.

Corollary 2. *Let $x, y \in \mathrm{H}$ and $(y, z_i) = 0$ for given nonzero vectors*

$\{z_i\}_i \subseteq$ H, $i = 1, 2, \ldots, n$. *Then, for any* $k \in \{1, 2, \ldots, n\}$,

$$\begin{aligned} & |(x,y)|^2 + \|y\|^2[\|z_k\|^{-2}|(x, z_k) \\ & -(\sum_i \frac{(x,z_i)}{\sum_j |(z_i,z_j)|} z_i, z_k)|^2 + \sum_i \frac{|(x,z_i)|^2}{\sum_j |(z_i,z_j)|}] \\ \leq \ & \|x\|^2\|y\|^2. \end{aligned}$$

Proof. Let $u = x - \sum_i \frac{(x,z_i)}{\sum_j |(z_i,z_j)|} z_i$. Then

$$\|u\|^2 \leq \|x\|^2 - \sum_i \frac{|(x,z_i)|^2}{\sum_j |(z_i,z_j)|}$$

by the proof in [5, Lemma 1]. Now, let $S = I = T$, $x = u$, and $e = z_k/\|z_k\|$ in (1) of Theorem 1. Then $(u, y) = (x, y)$ as $(z_i, y) = 0$, $i = 1, 2, \ldots, n$, and so

$$\begin{aligned} & |(x,y)|^2 + \|z_k\|^{-2}\|y\|^2|(u, z_k)|^2 \leq \|u\|^2\|y\|^2. \\ \leq \ & \|y\|^2[\|x\|^2 - \sum_i \frac{|(x,z_i)|^2}{\sum_j |(z_i,z_j)|}]. \end{aligned}$$

The desired inequality thus follows.

The type of inequality in Theorem 2 below was first appeared in [10, Theorem 1 and 4], and was subsequently improved in [2, Theorem 3.1, 3.3 and 3.4]. Now, we have a more general case.

Theorem 2. *Let* $S, T \in$ B(H), $x, y \in$ H, *and let* Sz_i *be unit vectors with* $(Ty, Sz_i) = 0$, $z_i \in$ H, $i = 1, 2, \ldots, n$. *Then*

$$\begin{aligned} & |(Sx,Ty)|^2 + \|Ty\|^2[|(Su_n, Sz_k)|^2 + \sum_i |(Su_{i-1}, Sz_i)|^2] \\ \leq \ & \|Sx\|^2\|Ty\|^2, \\ & \text{or, } |(Sx,Ty)| \leq \|Su_n - (Su_n, Sz_k)Sz_k\|\|Ty\|, \end{aligned}$$

$k \in \{1, 2, \ldots, n\}$, *where* $u_0 = x$ *and* $u_i = u_{i-1} - (Su_{i-1}, Sz_i)z_i$, $i = 1, 2, \ldots, n$.

The equality holds if and only if $Ty = Su_n - (Su_n, Sz_k)Sz_k$.

In particular, if $\{Sz_i\}_i$ *is a set of orthonormal vectors, then*

$$|(Sx, Ty)|^2 + \|Ty\|^2 \sum_i |(Sx, Sz_i)|^2 \leq \|Ty\|^2 \|Sx\|^2,$$

$$\text{or,} \qquad |\ (Sx, Ty)\ | \leq \|\ Ty\ \| \|\ Sx - \sum_i (Sx, Sz_i)Sz_i\ \|\ .$$

The equality holds if and only if $Ty = Sx - \sum_i (Sx, Sz_i)Sz_i$.

Proof. Let $i = 1, 2, \ldots, n$ in the definition of u_i. Then

$$u_n = x - \sum_i (Su_{i-1}, Sz_i)z_i.$$

Due to the definition of u_i we have

$$\|Su_i\|^2 = \|Su_{i-1}\|^2 - |(Su_{i-1}, Sz_i)|^2.$$

Setting $i = 1, 2, \ldots, n$ in above yields

$$\|Su_n\|^2 = \|Sx\|^2 - \sum_i |(Su_{i-1}, Sz_i)|^2.$$

Now, put $x = u_n$ and $e = Sz_k$ in (1) of Theorem 1, $k \in \{1, 2, \ldots, n\}$, then

$$\begin{aligned} & |(Su_n, Ty)|^2 + \|Ty\|^2 |(Su_n, Sz_k)|^2 \\ \leq\ & \|Ty\|^2 \|Su_n\|^2 \\ =\ & \|Ty\|^2 [\|Sx\|^2 - \sum_i |(Su_{i-1}, Sz_i)|^2], \end{aligned}$$

and the equality holds if and only if $Ty = Su_n - (Su_n, Sz_k)Sz_k$. Since $(Su_n, Ty) = (Sx, Ty) - (\sum_i (Su_{i-1}, Sz_i)Sz_i, Ty) = (Sx, Ty)$, the required inequality follows.

If $\{Sz_i\}_i$ is a set of orthonormal vectors in particular, then

$$\begin{aligned}(Su_n, Sz_k) &= (Sx - \sum_i (Su_{i-1}, Sz_i)Sz_i, Sz_k) \\ &= (Sx, Sz_k) - (Su_{k-1}, Sz_i) \\ &= (Sx, Sz_k) - (Sx - \sum_i (Su_{i-1}, Sz_i)Sz_i, Sz_k) \\ &= 0.\end{aligned}$$

Also, $(Su_{i-1}, Sz_i) = (Sx, Sz_i)$, $i = 1, 2, \dots, n$, by a similar computation, and we have the particular case. The equality holds if and only if $Ty = Su_n$, i.e., $Ty = Sx - \sum_i (Sx, Sz_i)Sz_i$.

As we see, Theorem 2 is obtained from (1) of Theorem 1. Conversely, let $n = 1$ and $Sz_1 = e$ in Theorem 2, then we have (1) of Theorem 1. Hence the two results are equivalent to each other. Now, if we set $S = I = T$ in Theorem 2, then we should obtain generalized and sharpened Cauchy-Schwarz inequality and Bessel's inequality as well, which are as follows.

$$|(x, y)|^2 + \|y\|^2[|(u_n, z_k)|^2 + \sum_i |(u_{i-1}, z_i)|^2] \le \|y\|^2\|x\|^2,$$

and

$$|(x, y)|^2 + \|y\|^2 \sum_i |(x, z_i)|^2 \le \|y\|^2\|x\|^2$$

with appropriate assumtions as in Theorem 2.

We cite the following order preserving operator inequalities. Let $A, B \in$ B(H).

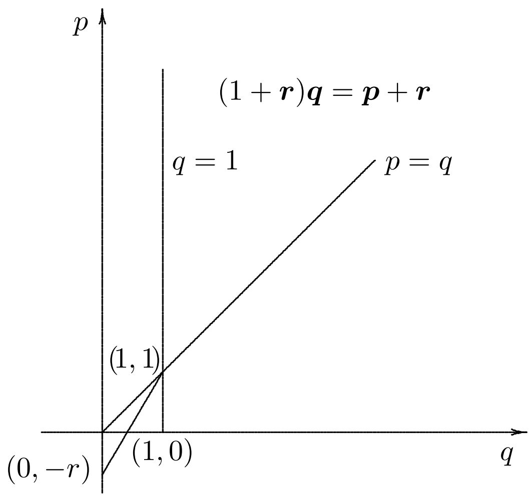
Theorem F. (Furuta inequality). *If $A \ge B \ge O$, then for each $r \ge 0$,*

(i) $(B^{r/2}A^pB^{r/2})^{1/q} \ge (B^{r/2}B^pB^{r/2})^{1/q}$ *and*

(ii) $(A^{r/2}A^pA^{r/2})^{1/q} \ge (A^{r/2}B^pA^{r/2})^{1/q}$ *hold for $p \ge 0$ and $q \ge 1$ with $(1 + r)q \ge p + r$.*

The original proof of Theorem F is in [6], and one page proof in [7]. The domain drawn for p, q and r in the Figure is the best possible one for Theorem F [16].

We also recall the following altered Furuta inequality [8] by the Löwner-Heinz formula.



Figure

Theorem F′. *If* $A \geq B \geq O$, *then for each* $r \geq 0$,

(i) $(B^{r/2}A^pB^{r/2})^{\frac{(1+r)\alpha}{p+r}} \geq B^{(1+r)\alpha}$ *and*

(ii) $A^{(1+r)\alpha} \geq (A^{r/2}B^pA^{r/2})^{\frac{(1+r)\alpha}{p+r}}$ *hold for any* $p \geq 1$ *and* $\alpha \in [0,1]$.

The Furuta inequality is an excellent extension of the Löwner-Heinz formula, i.e., if $A \geq B \geq O$, then $A^\alpha \geq B^\alpha$ for $\alpha \in [0,1]$; but the inequality does not hold in general for $\alpha > 1$. We also recall that $T = U|T|$ is the polar decomposition of T with $U \in$ B(H) the partial isometry, whereas $|T|$ and $|T^*|$ are positive square roots of operators T^*T and TT^*, respectively.

By applying Corollary 2 and the altered Furuta inequality we have the next corollary which is an extension of [5, Theorem 8]; whereas Theorem 8 in [5] in turn is a generalization of operator inequalities in [2, 4, 5, 8]. The inequality in Theorem 8 [5] which is a simultaneous extension of the Selberg and the Hienz-Kato-Furuta inequalities is precisely as follows.

$$\begin{aligned}
& |(T|T|^{(1+2r)\alpha+(1+2s)\beta-1}x, y)|^2 \\
& +\||T^*|^{(1+2s)\beta}y\|^2 \sum_i \frac{|(|T|^{2(1+2r)\alpha}x, z_i)|^2}{\sum_j |(|T|^{2(1+2r)\alpha}z_i, z_j)|} \\
\leq\ & ((|T|^{2r}A^{2p}|T|^{2r})^{\frac{(1+2r)\alpha}{p+2r}}x, x)((|T^*|^{2s}B^{2q}|T^*|^{2s})^{\frac{(1+2s)\beta}{q+2s}}y, y)
\end{aligned}$$

with exactly the same hypotheses in Corollary 3 below.

Corollary 3. *Let $A, B, T \in \mathrm{B(H)}$, and let A and B be positive operators such that $|T|^2 \leq A^2$ and $|T^*|^2 \leq B^2$. Then, for each $r, s \geq 0$,*

$$\begin{aligned}
& |(T|T|^{(1+2r)\alpha+(1+2s)\beta-1}x, y)|^2 \\
& +\||T^*|^{(1+2s)\beta}y\|^2[\||T|^{(1+2r)\alpha}z_k\|^{-2}|W|^2 + \sum_i \frac{|(|T|^{2(1+2r)\alpha}x, z_i)|^2}{\sum_j |(|T|^{2(1+2r)\alpha}z_i, z_j)|}] \\
\leq\ & (|T|^{2(1+2r)\alpha}x, x)(|T^*|^{2(1+2s)\beta}y, y) \\
\leq\ & ((|T|^{2r}A^{2p}|T|^{2r})^{\frac{(1+2r)\alpha}{p+2r}}x, x)((|T^*|^{2s}B^{2q}|T^*|^{2s})^{\frac{(1+2s)\beta}{q+2s}}y, y)
\end{aligned}$$

for $p, q \geq 1$, $\alpha, \beta \in [0, 1]$ with $(1+2r)\alpha + (1+2s)\beta \geq 1$ and $x, y, z_1, \ldots, z_n \in$ H such that $Tz_i \neq 0$ and $(T|T|^{(1+2r)\alpha+(1+2s)\beta-1}z_i, y) = 0$, $i = 1, \ldots, n$, where

$$W = (|T|^{2(1+2r)\alpha}x, z_k) - (\sum_i \frac{(|T|^{2(1+2r)\alpha}x, z_i)}{\sum_j |(|T|^{2(1+2r)\alpha}z_i, z_j)|}|T|^{2(1+2r)\alpha}z_i, z_k),$$

and $k \in \{1, 2, \ldots, n\}$.

Proof. The required first inequality may be easily obtained by substitutions. In other words, in Corollary 2 replace x by $U|T|^{(1+2r)\alpha}x$, y by $|T^*|^{(1+2s)\beta}y$, and z_i by $U|T|^{(1+2r)\alpha}z_i$, and use well-known relations $U^*U = I$, and $|T^*|^a = U|T|^a U^*$ for $a > 0$. Also, by Corollary 2,

$$\begin{aligned}
& (x, z_k) - (\sum_i \frac{(x, z_i)}{\sum_j |(z_i, z_j)|} z_i, z_k) \\
= \ & (|T|^{2(1+2r)\alpha} x, z_k) - (\sum_i \frac{(|T|^{2(1+2r)\alpha} x, z_i)}{\sum_j |(|T|^{2(1+2r)\alpha} z_i, z_j)|} |T|^{2(1+2r)\alpha} z_i, z_k) \\
= \ & W
\end{aligned}$$

by substitutions in above. The required last inequality is due to the altered Furuta inequality for both cases $O \le |T|^2 \le A^2$ and $O \le |T^*|^2 \le B^2$.

We recall the Reid's inequality: $|(RKx, x)| \le \|K\|(Rx, x)$ for every $x \in$ H, where $R \ge O$ and RK is Hermitian. In the next result we shall use (1) of Theorem 1 to sharpen a generalized Reid's inequality in [13, Theorem 1], which is the inequality in Corollary 4 below except without the second term $\||RK|^{(1+2s)\beta} y\|^2 |(V|RK|^{(1+2r)\alpha} x, e)|^2$ on the left side of the inequality.

Corollary 4. *Let $R \ge O$, RK be Hermitian, and let $RK = V|RK|$ be the polar decomposition, $r, s \ge 0$, and $p, q \ge 1$ with $(1+2r)\alpha + (1+2s)\beta \ge 1$. If $x, y \in$ H and $|RK|^{(1+2s)\beta} y$ $(\ne 0)$ is orthogonal to a unit vector e, then*

$$\begin{aligned}
& |(RK|RK|^{(1+2r)\alpha+(1+2s)\beta-1} x, y)|^2 \\
& + \| \ |RK|^{(1+2s)\beta} y\|^2 |(V|RK|^{(1+2r)\alpha} x, e)|^2 \\
\le \ & (|RK|^{2(1+2r)\alpha} x, x)(|RK|^{2(1+2s)\beta} y, y) \\
\le \ & \|K\|^a (|RK|^{2r} R^{2p} |RK|^{2r})^{\frac{(1+2r)\alpha}{p+2r}} x, x)(|RK|^{2s} R^{2q} |RK|^{2s})^{\frac{(1+2s)\beta}{q+2s}} y, y),
\end{aligned}$$

where $a = \frac{2p(1+2r)\alpha}{p+2r} + \frac{2q(1+2s)\beta}{q+2s}$.

The first inequality becomes equality if and only if

$$|RK|^{(1+2s)\beta} y = V|RK|^{(1+2r)\alpha} x - (V|RK|^{(1+2r)\alpha} x, e)e.$$

Proof. In (1) of Theorem 1 replace S by $V|RK|^{(1+2r)\alpha}$, T by $|RK|^{(1+2s)\beta}$

($= V|RK|^{(1+2s)\beta}V^*$ as RK is Hermitian. cf. a known relation in the proof of Corollary 3), and then use a known inequality in [13, Proof of Theorem 1]:

$$|RK|^{2(1+2r)\alpha} \leq \|K\|^{\frac{2p(1+2r)\alpha}{p+2r}} (|RK|^{2r} R^{2p} |RK|^{2r})^{\frac{(1+2r)\alpha}{p+2r}}$$

for $R \geq O$, a Hermitian operator RK, the polar decomposition $RK = V|RK|$, $r \geq 0$, $p \geq 1$, and $\alpha \in [0,1]$.

3. Equivalent relations. As we mentioned in the introduction Theorem 1 is essential in this paper. In this section we shall prove that inequalities in Theorem 1 are equivalent to inequalities expressed in terms of covariances and variances. In fact, their expressions are quite analogous.

Theorem 3. *Let* $S, T \in \mathrm{B(H)}$, $e, x, y \in \mathrm{H}$ *with* $\|e\| = 1 = Var_{Ty}(P_e)$. *Then the following are equivalent.*

(i) $|(Sx, Ty) - (Sx, e)(e, Ty)|^2 \leq [\|Sx\|^2 - |(Sx, e)|^2][\|Ty\|^2 - |(Ty, e)|^2]$.

The equality holds if and only if $Ty = Sx - \lambda e$, $\lambda \in \mathbf{C}$.

(ii) $|\mathrm{Cov}_{Ty}(S,T) - \mathrm{Cov}_{Ty}(S,P_e)\mathrm{Cov}_{Ty}(P_e,T)|^2$

$\leq [\mathrm{Var}_{Ty}(S) - |\mathrm{Cov}_{Ty}(S,P_e)|^2][\mathrm{Var}_{Ty}(T) - |\mathrm{Cov}_{Ty}(P_e,T)|^2]$.

The equality holds if and only if $Ty = Sx + \alpha Tx + \beta e$, $\alpha, \beta \in \mathbf{C}$.

Proof. (i) implies (ii). Our main tool is the c-v inequality, since (i) implies the Cauchy-Schwarz inequality by (1) in Theorem 1 when $S = T = I$, which in turn is equivalent to the c-v inequality by a remark in section 1. Now, let $u = \mathrm{Cov}_{Ty}(S,T) - \mathrm{Cov}_{Ty}(S,P_e)\mathrm{Cov}_{Ty}(P_e,T)$ and $S + cuT = R \in \mathrm{B(H)}$, where c is any real number. Then, as $\mathrm{Var}_{Ty}(P_e) = 1$,

$$\begin{aligned} 0 &\leq \mathrm{Var}_{Ty}(R)\mathrm{Var}_{Ty}(P_e) - |\mathrm{Cov}_{Ty}(R,P_e)|^2 \\ &= \mathrm{Cov}_{Ty}(S+cuT, S+cuT) - \mathrm{Cov}_{Ty}(S+cuT, P_e)\mathrm{Cov}_{Ty}(P_e, S+cuT) \\ &= \mathrm{Var}_{Ty}(S) + c\bar{u}\mathrm{Cov}_{Ty}(S,T) + cu\mathrm{Cov}_{Ty}(T,S) + c^2|u|^2\mathrm{Var}_{Ty}(T) \end{aligned}$$

$$
\begin{aligned}
&-[\mathrm{Cov}_{Ty}(S,P_e)+cu\mathrm{Cov}_{Ty}(T,P_e)][\mathrm{Cov}_{Ty}(P_e,S)+c\bar{u}\mathrm{Cov}_{Ty}(P_e,T)] \\
= \;& [\mathrm{Var}_{Ty}(S)-|\mathrm{Cov}_{Ty}(S,P_e)|^2]+2c|u|^2 \\
&+c^2|u|^2[\mathrm{Var}_{Ty}(T)-|\mathrm{Cov}_{Ty}(P_e,T)|^2].
\end{aligned}
$$

Now, if $\mathrm{Var}_{Ty}(T)-|\mathrm{Cov}_{Ty}(P_e,T)|^2=0$, then $\mathrm{Cov}_{Ty}(S,T)-$ $\mathrm{Cov}_{Ty}(S,P_e)$ $\mathrm{Cov}_{Ty}(P_e,T)=u=0$ as $[\mathrm{Var}_{Ty}(S)-|\mathrm{Cov}_{Ty}(S,P_e)|^2]+2c|u|^2\geq 0$, and the relation becomes trivial. If $\mathrm{Var}_{Ty}(T)-|\mathrm{Cov}_{Ty}(P_e,T)|^2\neq 0$, then we set $c=\frac{-1}{\mathrm{Var}_{Ty}(T)-|\mathrm{Cov}_{Ty}(P_e,T)|^2}$ to get the required inequality.

The equality holds if and only if, by a remark in section 1, $Ty=(R-\lambda P_e)x=Sx+cuTx-\lambda(x,e)e=Sx+\alpha Tx+\beta e$, $\alpha,\beta\in\mathbf{C}$.

(ii) implies (i). We may choose a unit vector e such that Cov $_{Ty}(S,P_e)=0$, i.e., $(Sx,e)=0=(Ty,e)$. Then, clearly (ii) implies the c-v inequality, which in turn implies (i) by Theorem 1.

Consequently, (1) and (2) in Theorem 1 may be similarly characterized respectively as follows without proof.

Corollary 5. *Let* $S,T\in$ B(H), $e,x,y\in$ H *with* $\|e\|=1=Var_{Ty}(P_e)$ *and* $(Ty,e)=0=Cov_{Ty}(T,P_e)$. *Then the following are equivalent.*

(i) $|(Sx,Ty)|^2+\|Ty\|^2|(Sx,e)|^2\leq\|Sx\|^2\|Ty\|^2$.

The equality holds if and only if $Ty=Sx-(Sx,e)e$.

(ii) $|Cov_{Ty}(S,T)|^2+Var_{Ty}(T)|Cov_{Ty}(S,P_e)|^2\leq\ Var_{Ty}(S)\,Var_{Ty}(T)$.

The equality holds if and only if $Ty=Sx+\alpha Tx-[(Sx,e)+\alpha(Tx,e)]e$.

Corollary 6. *Let* $S,T\in$ B(H), $e,x,y\in$ H *with* $\|e\|=1=Var_{Ty}(P_e)$ *and* $(Sx,Ty)=0=Cov_{Ty}(S,T)$. *Then the following are equivalent.*

(i) $\|Sx\|^2|(Ty,e)|^2+\|Ty\|^2|(Sx,e)|^2\leq\|Sx\|^2\|Ty\|^2$.

The equality holds if and only if $Ty=Sx-\frac{\|Sx\|^2e}{(e,Sx)}$, $(e,Sx)\neq 0$.

(ii) $\mathrm{Var}_{Ty}(S)|\mathrm{Cov}_{Ty}(P_e,T)|^2+\mathrm{Var}_{Ty}(T)||\mathrm{Cov}_{Ty}(S,P_e)|^2$

$\leq \mathrm{Var}_{Ty}(S)\mathrm{Var}_{Ty}(T)$.

The equality holds if and only if $Ty = Sx + \alpha Tx - \frac{[\|Sx\|^2 + \alpha(Tx,Sx)]e}{(e,Sx)}$.

We have to remark here that the vector Ty in (ii) of Theorem 3, and in (ii) of Corollary 5 and 6 may be replaced by any vector $z \neq e$ (since $\mathrm{Var}_e(P_e) = 0 \neq 1$) without affecting equivalent relations. However, the equality conditions should be changed accordingly.

4. Characterizations of operator inequalities. In each of the following corollaries we shall apply (1) of Theorem 1 to characterize a known operator inequality. Observe that recently there are many papers written about characterizing the Furuta inequality [2, 4, 5, 8, 13, 14]. However, we shall present different types of characterizations below. First of all let $A, B, S \in \mathrm{B(H)}$ and $\alpha, \beta \in [0,1]$.

Corollary 7. *Let* $A \geq B \geq O$, $p, q \geq 1$ *and* $r, s \geq 0$. *If* $y \in \mathrm{H}$ *and* $B^{\frac{(1+r)\alpha}{2}}y$ $(\neq 0)$ *is orthogonal to a unit vector* e, *then the following are equivalent.*

(i) $B^{(1+r)\alpha} \leq (B^{r/2}A^pB^{r/2})^{\frac{(1+r)\alpha}{p+r}}$, *the altered Furuta inequality* [8];

(ii) $|(B^{\frac{(1+r)\alpha}{2}+\frac{(1+s)\beta}{2}}x, y)|^2 + \|B^{\frac{(1+r)\alpha}{2}}y\|^2|(B^{\frac{(1+s)\beta}{2}}x, e)|^2$

$\leq (B^{(1+r)\alpha}y, y)(B^{(1+s)\beta}x, x)$

$\leq ((B^{r/2}A^pB^{r/2})^{\frac{(1+r)\alpha}{p+r}}y, y)(B^{s/2}A^qB^{s/2})^{\frac{(1+s)\beta}{q+s}}x, x)$

for every $x \in \mathrm{H}$.

Proof. (i) implies (ii). In (1) of Theorem 1 replace S by $B^{\frac{(1+s)\beta}{2}}$ and T by $B^{\frac{(1+r)\alpha}{2}}$, then the required inequality follows due to (i).

(ii) implies (i). Let $y = x$, $\beta = \alpha$, $s = r$, and $q = p$ in (ii). Then $|(B^{(1+r)\alpha}x, x)|^2 \leq (B^{(1+r)\alpha}x, x)^2 \leq ((B^{r/2}A^pB^{r/2})^{\frac{(1+r)\alpha}{p+r}}x, x)^2$ since $(B^{\frac{(1+s)\beta}{2}}$ $x, e) = (B^{\frac{(1+r)\alpha}{2}}y, e) = 0$, and hence (1).

Corollary 8. *Let* $A \geq B \geq O$, $p, q, r, s \geq 0$, *and* $u, v \geq 1$ *with* $(1+r)u \geq p+r$ *and* $(1+s)v \geq q+s$. *If* $y \in \mathrm{H}$ *and* $B^{\frac{p+r}{2u}}y$ $(\neq 0)$ *is orthogonal to a unit vector* e, *then the following are equivalent.*

(i) $B^{\frac{p+r}{u}} \leq (B^{r/2}A^pB^{r/2})^{\frac{1}{u}}$, *the original Furuta inequality* [6];

(ii) $|(B^{\frac{p+r}{2u}+\frac{q+s}{2v}}x, y)|^2 + \|B^{\frac{p+r}{2u}}y\|^2|(B^{\frac{q+s}{2v}}x, e)|^2$

$\leq (B^{\frac{p+r}{u}}y, y)(B^{\frac{q+s}{v}}x, x)$

$\leq ((B^{r/2}A^pB^{r/2})^{\frac{1}{u}}y, y)(B^{s/2}A^qB^{s/2})^{\frac{1}{v}}x, x)$.

for every $x \in \mathrm{H}$.

Proof. (i) implies (ii). In (1) of Theorem 1 replace S by $B^{\frac{q+s}{2v}}$ and T by $B^{\frac{p+r}{2u}}$, and use (i) to get (ii).

(ii) implies (i). Let $y = x$, $q = p$, $s = r$, and $v = u$ in (ii) so that $(B^{\frac{q+s}{2v}}x, e) = 0$. Hence, $|(B^{\frac{p+r}{u}}x, x)|^2 \leq (B^{\frac{p+r}{u}}x, x)^2 \leq ((B^{r/2}A^pB^{r/2})^{\frac{1}{u}}x, x)^2$ which implies (i).

Consequently we have a characterization of the Löwner-Heinz formula as follows: If $A \geq B \geq O$, $B^{\frac{\alpha}{2}}y \neq 0$, $(B^{\frac{\alpha}{2}}y, e) = 0$ for a unit vector e, and $\alpha, \beta \in [0, 1]$, then the following are equivalent: (i) $A^\alpha \geq B^\alpha$; (ii) $|(B^{\frac{\alpha+\beta}{2}}x, y)|^2 + \|B^{\frac{\alpha}{2}}y\|^2|(B^{\frac{\beta}{2}}x, e)|^2 \leq (B^\alpha y, y)(B^\beta x, x) \leq (A^\alpha y, y)(A^\beta x, x)$ for every $x \in \mathrm{H}$. This is obtained by letting $p = \alpha$, $q = \beta$, $r = s = 0$, and $u = v = 1$ in Corollary 6.

Corollary 9. *Let* $A \geq B \geq O$, $p, q, r, s \geq 0$, *and* $u, v \geq 1$ *with* $(1+r)u \geq p+r$ *and* $(1+s)v \geq q+s$. *If* $y \in \mathrm{H}$ *and* $(A^{r/2}B^pA^{r/2})^{\frac{1}{2u}}y$ $(\neq 0)$ *is orthogonal to a unit vector* e, *then the following are equivalent.*

(i) $(A^{r/2}B^pA^{r/2})^{\frac{1}{u}} \leq A^{\frac{p+r}{u}}$, *the original Furuta inequality* [6];

(ii) $|((A^{s/2}B^qA^{s/2})^{\frac{1}{2v}}x, (A^{r/2}B^pA^{r/2})^{\frac{1}{2u}}y)|^2$

$+\|(A^{r/2}B^pA^{r/2})^{\frac{1}{2u}}y\|^2|((A^{s/2}B^qA^{s/2})^{\frac{1}{2v}}x, e)|^2$

$$\leq ((A^{r/2}B^pA^{r/2})^{\frac{1}{u}}y, y)((A^{s/2}B^qA^{s/2})^{\frac{1}{v}}x, x)$$
$$\leq (A^{\frac{p+r}{u}}y, y)(A^{\frac{q+s}{v}}x, x).$$

for every $x \in$ H.

Proof. (i) implies (ii). In (1) of Theorem 1 replace S by $(A^{s/2}B^qA^{s/2})^{\frac{1}{2v}}$ and T by $(A^{r/2}B^pA^{r/2})^{\frac{1}{2u}}$, and apply (i) to get (ii).

(ii) implies (i). Let $y = x$, $q = p$, $s = r$, and $v = u$ in (ii) and thus $(A^{s/2}B^qA^{s/2})^{\frac{1}{2v}}x, e) = 0$. So $|((A^{r/2}B^pA^{r/2})^{\frac{1}{u}}x, x)|^2 \leq ((A^{r/2}B^pA^{r/2})^{\frac{1}{u}}x, x)^2$ $\leq (A^{\frac{p+r}{u}}x, x)^2$ which implies (i).

Corollary 10. *If* $y \in H$ *and* S^*y $(\neq 0)$ *is orthogonal to a unit vector* e, *then the following are equivalent.*

(i) S *is hyponormal, i.e.,* $S^*S \geq SS^*$;

(ii) $|(SS^*x, y)|^2 + \|S^*y\|^2|(S^*x, e)|^2 \leq (SS^*y, y)(SS^*x, x) \leq \|Sy\|^2\|Sx\|^2$

for every $x \in$ H.

Proof. (i) implies (ii). In (1) of Theorem 1 replace both S and T by S^*, and use (i) to get (ii).

(ii) implies (i). Let $y = x$ in (ii). Then $\|S^*x\| \leq \|Sx\|$ as $(S^*x, e) = 0$, i.e., S is hyponormal.

Immediately we have a characterization of a contraction $S \in$ B(H), i.e., if $y \in$ H and S^*y $(\neq 0)$ is orthogonal to a unit vector e, then the following are equivalent: (i) S is a contraction, i.e., $I \geq SS^*$; (ii) $|(SS^*x, y)|^2 + \|S^*y\|^2|(S^*x, e)|^2 \leq (SS^*y, y)(SS^*x, x) \leq (y, y)(x, x) = \|y\|^2\|x\|^2$ for every $x \in$ H. Note that if S is a contraction, so is S^*, and vice versa.

We remark that by applying (1) of Theorem 1 and adapting the similar techniques as in this section, it is possible to characterize other operator inequalities accordingly.

5. Extended covariance and variance. In the definition of covariance and variance for $S, T \in B(H)$ in section 1, both S and T are acting on the same vector. We may naturally let T act on a different vector, so that we have a slight extension as follows.

Definition. Let $S, T \in B(H)$. The extended covariance of S and T on H is a mapping $\mathrm{Ecov}_z(S,T) : H \times H \to \mathbf{C}$ defined by

$$\mathrm{Ecov}_z(S,T) = \|z\|^2(Sx, Ty) - (Sx, z)(z, Ty)$$

for every $x, y, z \in H$. It is called the extended variance of S on H (written $\mathrm{Evar}_z(S)$) if $Ty = Sx$, i.e.,

$$\mathrm{Evar}_z(S) = \mathrm{Ecov}_z(S,S) = \|z\|^2\|Sx\|^2 - |(Sx, z)|^2,$$

which is exactly the same as $\mathrm{Var}_z(S)$.

By Definition the extended covariance and variance enjoy the five properties as in section 1. Similar to [14, Theorem 1] it can be prove that the extended c-v inequality

$$|\mathrm{Ecov}_z(S,T)|^2 \leq \mathrm{Evar}_z(S)\mathrm{Evar}_z(T)$$

holds, and is equivalent to the Cauchy-Schwarz inequality. Also, the extended c-v equality holds if and only if $z = Sx - \lambda Ty$, $\lambda \in \mathbf{C}$.

We shall leave applications and development of the extended covariance and variance to the readers. We would like to mention, nevertheless, one application in particular as follows. In the extended c-v inequality above, let S act on x, T on y, and replace z by e with $\|e\| = 1$. Then

$$|(Sx, Ty) - (Sx, e)(e, Ty)|^2 \leq [\|Sx\|^2 - |(Sx, e)|^2][\|Ty\|^2 - |(Ty, e)|^2].$$

which is precisely the first inequality in Theorem 1.

References

1. M. L. Buzano, *Generalizzatione della diseguaglianza di Cauchy-Schwarz*, Rend. Sem. Mat. Uni. e Politech. Trimo, **31**(1971/73), 405-409.

2. M. Fujii, *Around the Furuta inequality*, preprint.

3. M. Fujii and F. Kubo, *Buzano's inequality and bounds for roots of algebraic equations*, Proc. Amer. Math. Soc., **117**(1993), 359-361.

4. M. Fujii, C. -S. Lin and R. Nakamoto, *Alternative extensions of Heinz-Kato-Furuta inequality*, Scien. Math., **2**(1999), 215-221.

5. M. Fujii and R. Nakamoto, *Simultaneous extensions of Selberg inequality and Heinz-Kato Furuta inequality*, Nihonkai Math. J., **9**(1998), 219-225.

6. T. Furuta, $A \geq B \geq O$ assures $(B^r A^p B^r)^{1/q} \geq B^{(p+2r)/q}$ for $r \geq 0$, $p \geq 0$, $q \geq 1$ with $(1+2r)q \geq p+2r$, Proc. Amer. Math. Soc., **101**(1987), 85-88.

7. T. Furuta, *Elementary proof of an order preserving operator inequality*, Proc. Japan Acad., **65**(1989), 126.

8. T. Furuta, *Generalization of Heinz-Kato theorem via Furuta inequality*, Operator Theory: Advances and Applications. **62**(1993), 77-83.

9. P. R. Halmos, *Hilbert Space Problem Book*, Van Nostrand, Princeton, N. J., 1967.

10. C. -S. Lin, *Heinz's inequality and Bernstein's inequality*, Proc. Amer. Math. Soc., **125**(1997), 2319-2325.

11. C. -S. Lin, *On Halmos' sharpening of Reid's inequality*, C. R. Math. Rep. Acad. Sci. Canada, **20**(1998), 62-64.

12. C. -S. Lin, *Operator versions of inequalities and equalities on a Hilbert space*, Lin. Alge. and its Appli., **268**(1998), 365-374.

13. C. -S. Lin, *Inequalities of Reid type and Furuta*, Proc. Amer. Math. Soc., to appear.

14. C. -S. Lin, *On variance and covariance for bounded linear operators*, Acta Mathematica Sinica, to appear.

15. C. -S. Lin, *On 2-inner product defined by inner product*, submitted.

16. K. Tanahashi, *Best possibility of the Furuta inequality*, Proc. Amer. Math. Soc., **124**(1996), 141-146.

Department of Mathematics, Bishop's University, Lennoxville, P.Q., J1M 1Z7 Canada.

E-mail: plin@ubishops.ca