SIAM J. MATRIX ANAL. APPL.
Vol. 31, No. 4, pp. 1606–1620

FROM SUBADDITIVE INEQUALITIES OF SINGULAR VALUES TO TRIANGLE INEQUALITIES OF CANONICAL ANGLES*

YANXIA ZHANG† AND LI QIU‡

Abstract. The singular values of matrices $A, B, C \in \mathbb{C}^{m\times n}$ with $C = A+B$ satisfy an extensive list of subadditive inequalities discovered by K. Fan, V.B. Lidskii, H. Wielandt, R.C. Thompson, A. Horn, and so on. These inequalities still hold when we apply a nonnegative concave function to each of the singular values involved, as shown recently by M. Uchiyama and J.C. Bourin. The main purpose of this paper is to show that all of these singular value inequalities can be translated into canonical angle inequalities. The bridge between the singular values and the canonical angles is given by a "multiplicative Pythagorean identity" relating the direct rotations between three subspaces.

Key words. triangle inequalities, canonical angles, subadditive inequalities, singular values, direct rotation

AMS subject classifications. 15A18, 15A03

DOI. 10.1137/09074841X

1. Introduction.

It is well known that every matrix $A \in \mathbb{C}^{m\times n}$ has a singular value decomposition of the form $A = U\Sigma V^H$, where U and V are unitary matrices and Σ is an $m \times n$ diagonal matrix with nonnegative real diagonal entries ordered decreasingly. The diagonal entries of Σ are called the singular values of A and are denoted by $\sigma_1(A), \ldots, \sigma_{\min\{n,m\}}(A)$. In the following, we also use a vector $\sigma(A)$ to denote the decreasing $\min\{n, m\}$-tuple $(\sigma_1(A), \ldots, \sigma_{\min\{n,m\}}(A))$.

Denote the set of m-dimensional subspaces of $\mathbb{C}^n$ by $\mathcal{G}_{m,n}$. This set is called a Grassmann manifold or simply Grassmannian. Let $\mathcal{X}, \mathcal{Y} \in \mathcal{G}_{m,n}$. Suppose the columns of $X_1, Y_1 \in \mathbb{C}^{n\times m}$ form orthonormal bases for $\mathcal{X}$ and $\mathcal{Y}$, respectively. Then the singular values of $X_1^H Y_1$ lie in $[0, 1]$ and are independent of the particular choices of the orthonormal bases. The *canonical angles* between $\mathcal{X}$ and $\mathcal{Y}$ are defined to be [23]

$$\theta_i(\mathcal{X}, \mathcal{Y}) = \arccos \sigma_{m-i+1}(X_1^H Y_1) \in [0, \pi/2], \quad i = 1, 2, \ldots, m.$$

In the following, we use a vector $\theta(\mathcal{X}, \mathcal{Y})$ to denote the decreasingly ordered m-tuple $(\theta_1(\mathcal{X}, \mathcal{Y}), \ldots, \theta_m(\mathcal{X}, \mathcal{Y}))$.

The discovery of canonical angles [13] followed shortly after that of singular values [12]; both were made by French mathematician C. Jordan. Their applications in statistics came one after the other, pioneered by American mathematician and statistician H. Hotelling, leading to principal component analysis (PCA) [10] and canonical correlation analysis (CCA) [11], respectively. However, the theory of canonical angles has lagged far behind that of singular values since then.

*Received by the editors February 2, 2009; accepted for publication (in revised form) by N. J. Higham November 20, 2009; published electronically February 17, 2010.
http://www.siam.org/journals/simax/31-4/74841.html

†Department of Mathematics, Beijing Institute of Technology, Beijing 100081, China (yxzhang@bit.edu.cn). The research of this author was supported by the Excellent Young Scholars Research Fund of the Beijing Institute of Technology through grant 2006Y0714 and by the Natural Science Foundation of China through grant 60704002.

‡Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China (eeqiu@ust.hk). The research of this author was supported by the Hong Kong Research Grants Council under grant 602107 and by the Shanghai Key Laboratory of Contemporary Applied Mathematics of Fudan University during his visit there.

It is our view that canonical angles share a large number of properties with singular values. The support of this view can be seen in the recent paper [22], where some canonical angle inequalities analogous to certain singular value inequalities were proved; a family of unitarily invariant intrinsic metrics in $\mathcal{G}_{m,n}$ resembling the family of unitarily invariant matrix norms was introduced; bounds on canonical angle variations when the subspaces were perturbed were given, which are similar to those on singular value variations when the matrix was perturbed; and the stability of nullity and deficiency under subspace perturbations was analyzed in a parallel way to the analysis of the stability of matrix rank under matrix perturbations. Very recently, we became aware that two of the main results on unitarily invariant metrics and canonical angle inequalities in [22], Theorems 3.1 and 3.2, were proved earlier in [20] using a geometric approach, in contrast to the algebraic approach in [22].

In this paper, we start with a review of singular value subadditive inequalities with some small extensions in section 2. We then present without proof the analogous canonical angle inequalities in section 3. The preparation for the proof of canonical angle inequalities is done in section 4 by further studying the direct rotations between two subspaces. A "multiplicative Pythagorean identity" is stated and proved in section 4 and is very powerful in, among other things, bridging the singular values and the canonical angles. Section 5 is devoted to the proof of canonical angle inequalities based on the preparation in section 4.

The notation used in this paper is mostly standard. For real vectors x and y, the operations and relations such as $|x|$, $x \le y$, $f(x)$, where f is a scalar function, are interpreted in an elementwise manner. For $n \times n$ Hermitian matrices A and B, inequality $A \le B$ means that $B - A$ is positive semidefinite. For $n \times n$ Hermitian matrix A and scalar function f, $f(A)$ is defined using the standard spectral calculus.

2. Singular value inequalities—a review and a little more. The study of singular values often relies on the study of eigenvalues of Hermitian matrices due to the fact that the eigenvalues of an $(m+n) \times (m+n)$ Hermitian matrix

$$\begin{bmatrix} 0 & A \\ A^H & 0 \end{bmatrix} \tag{2.1}$$

are $\pm\sigma_1(A), \pm\sigma_2(A), \ldots, \pm\sigma_{\min\{m,n\}}$ together with an $|m-n|$ number of 0's. The matrix in (2.1) is sometimes called the Jordan–Wielandt matrix corresponding to A.

Let us denote the decreasingly ordered n-tuple of eigenvalues of an $n \times n$ Hermitian matrix A by a vector $\lambda(A) := (\lambda_1(A), \ldots, \lambda_n(A))$. In 1949, Fan [6] proved that for any $n \times n$ Hermitian matrices A, B, C with $C = A + B$, their eigenvalues satisfy

$$\sum_{l=1}^{k} \lambda_l(C) \le \sum_{l=1}^{k} \lambda_l(A) + \sum_{l=1}^{k} \lambda_l(B) \tag{2.2}$$

for $1 \le k \le n-1$ in addition to the obvious equality

$$\sum_{l=1}^{n} \lambda_l(C) = \sum_{l=1}^{n} \lambda_l(A) + \sum_{l=1}^{n} \lambda_l(B). \tag{2.3}$$

A more trendy way of writing inequalities (2.2) and equality (2.3) is by using majorization. For $x, y \in \mathbb{R}^n$, x is said to be majorized by y, denoted by $x \preceq y$, if $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ and

$$\max_{1 \le i_1 < \cdots < i_k \le n} \{x_{i_1} + \cdots + x_{i_k}\} \le \max_{1 \le i_1 < \cdots < i_k \le n} \{y_{i_1} + \cdots + y_{i_k}\} \tag{2.4}$$

for $1 \le k \le n-1$. Furthermore, x is said to be weakly majorized by y, denoted by $x \preceq_w y$, if $\sum_{i=1}^n x_i \le \sum_{i=1}^n y_i$ and (2.4) holds for $1 \le k \le n-1$. Majorization and weak majorization are partial orders not on $\mathbb{R}^n$ but on the quotient space of $\mathbb{R}^n$ over the set of all permutations on $\mathbb{R}^n$. With the majorization language, inequalities (2.2) and equality (2.3) can be succinctly written as

$$\lambda(C) \preceq \lambda(A) + \lambda(B).$$

Applying (2.2) to the Jordan–Wielandt matrices corresponding to $A, B, C \in \mathbb{C}^{m\times n}$ with $C = A + B$, we immediately obtain

$$\sum_{l=1}^{k} \sigma_l(C) \le \sum_{l=1}^{k} \sigma_l(A) + \sum_{l=1}^{k} \sigma_l(B) \tag{2.5}$$

for $1 \le k \le \min\{m, n\}$, as Fan did in 1951 [7]. Inequalities (2.5) can also be simply written as

$$\sigma(C) \preceq_w \sigma(A) + \sigma(B).$$

One implication of inequalities (2.5) is that a family of matrix norms is defined by

$$\|A\|_{(k)} = \sum_{l=1}^{k} \sigma_l(A)$$

since inequalities (2.5) are nothing but the triangle inequalities of the norms. These norms are unitarily invariant in the sense that $\|UAV\|_{(k)} = \|A\|_{(k)}$ for all $A \in \mathbb{C}^{m\times n}$ and for all unitary matrices U and V. The implication of (2.5) actually goes far beyond. A function $\Phi : \mathbb{R}^p \to \mathbb{R}$ is called a symmetric gauge function if it is a norm on $\mathbb{R}^p$ satisfying the additional properties that it is symmetric, i.e., $\Phi(P\xi) = \Phi(\xi)$ for any $\xi \in \mathbb{R}^p$ and permutation matrix P, and that it is absolute, i.e., $\Phi(|\xi|) = \Phi(\xi)$. The Fan inequalities can be used to establish that an extended family of unitarily invariant norms is defined by

$$\|A\|_\Phi = \Phi[\sigma(A)]$$

for a symmetric gauge function Φ, as von Neumann discovered earlier [27]. It was also pointed out by von Neumann that all unitarily invariant matrix norms are given in such a way.

Inequalities (2.2) were extended to a bigger family by Lidskii [18] in 1950 and Wielandt [28] in 1955. They established that for $n \times n$ Hermitian matrices $A, B, C = A + B$,

$$\sum_{l=1}^{k} \lambda_{i_l}(C) \le \sum_{l=1}^{k} \lambda_{i_l}(A) + \sum_{l=1}^{k} \lambda_l(B) \tag{2.6}$$

for all $1 \le i_1 < \cdots < i_k \le n$ and $1 \le k \le n-1$. Inequalities (2.6), together with the obvious equality (2.3), can be expressed as

$$\lambda(C) - \lambda(A) \preceq \lambda(B).$$

Again, via the Jordan–Wielandt matrices, we easily obtain that for $A, B, C \in \mathbb{C}^{m\times n}$ with $C = A + B$,

$$\sum_{l=1}^{k} \sigma_{i_l}(C) \le \sum_{l=1}^{k} \sigma_{i_l}(A) + \sum_{l=1}^{k} \sigma_l(B) \tag{2.7}$$

for all $1 \le i_1 < \cdots < i_k \le \min\{m, n\}$ and $1 \le k \le \min\{m, n\}$, or in a short way,

$$\sigma(C) - \sigma(A) \preceq_w \sigma(B).$$

Clearly, the Lidskii–Wielandt eigenvalue and singular value inequalities (2.6) and (2.7) include the Fan inequalities (2.2) and (2.5) as special cases.

Mirsky [19] observed in 1960 that applying the Lidskii–Wielandt eigenvalue inequalities to the Jordan–Wielandt matrices in a more sophisticated way actually yields that for $A, B \in \mathbb{C}^{m\times n}$,

$$\sum_{l=1}^{k} |\sigma_{i_l}(A) - \sigma_{i_l}(B)| \le \sum_{l=1}^{k} \sigma_l(A - B)$$

for all $1 \le i_1 < \cdots < i_k \le \min\{m, n\}$, namely,

$$|\sigma(A) - \sigma(B)| \preceq_w \sigma(A - B),$$

which consequently leads to

$$\Phi(\sigma(A) - \sigma(B)) \le \|A - B\|_\Phi$$

for any symmetric gauge function Φ.

In 1971, Thompson and Freede extended the Lidskii–Wielandt eigenvalue inequalities further by showing that for $n \times n$ Hermitian matrices $A, B, C = A + B$,

$$\sum_{l=1}^{k} \lambda_{i_l+j_l-l}(C) \le \sum_{l=1}^{k} \lambda_{i_l}(A) + \sum_{l=1}^{k} \lambda_{j_l}(B) \tag{2.8}$$

for each $1 \le i_1 < \cdots < i_k \le n$, $1 \le j_1 < \cdots < j_k \le n$ with $i_k + j_k - k \le n$.

Thompson [25] then presented in 1973 the singular value version: For $A, B, C \in \mathbb{C}^{m\times n}$ with $C = A + B$,

$$\sum_{l=1}^{k} \sigma_{i_l+j_l-l}(C) \le \sum_{l=1}^{k} \sigma_{i_l}(A) + \sum_{l=1}^{k} \sigma_{j_l}(B) \tag{2.9}$$

for each $1 \le i_1 < \cdots < i_k \le \min\{m, n\}$, $1 \le j_1 < \cdots < j_k \le \min\{m, n\}$ with $i_k + j_k - k \le \min\{m, n\}$.

After finding various inequalities governing the eigenvalues of Hermitian matrices $A, B, C = A + B$, researchers had been trying to give a complete characterization of these eigenvalues. In 1962, Horn [9] defined sets T_r^n, $r = 1, \ldots, n$, of triples (I, J, K), where I, J, K are certain subsets of $\{1, 2, \ldots, n\}$ of the same cardinality r, and gave a remarkable conjecture that the inequalities

$$\sum_{k\in K} \gamma_k \le \sum_{i\in I} \alpha_i + \sum_{j\in J} \beta_j \tag{2.10}$$

for all these triples would give the necessary and sufficient conditions for three n-tuples α, β, and γ with $\alpha: \ \alpha_1 \geq \cdots \geq \alpha_n, \ \ \beta: \ \beta_1 \geq \cdots \geq \beta_n, \ \ \gamma: \ \gamma_1 \geq \cdots \geq \gamma_n$ to arise as eigenvalues of $n \times n$ Hermitian matrices A, B, and C with $C = A + B$.

The sets T_r^n of triples (I, J, K) were defined using the following inductive procedure. Set

$$U_r^n = \left\{ (I, J, K) \mid \sum_{i \in I} i + \sum_{j \in J} j = \sum_{k \in K} k + r(r+1)/2 \right\}. \tag{2.11}$$

When $r = 1$, set $T_1^n = U_1^n$. In general,

$$T_r^n = \Bigg\{ (I, J, K) \in U_r^n \mid \text{for all } p \leq r \text{ and all } (F, G, H) \text{ in } T_p^r,$$

$$\sum_{f \in F} i_f + \sum_{g \in G} j_g \leq \sum_{h \in H} k_h + p(p+1)/2 \Bigg\}. \tag{2.12}$$

Recently, with the works of a few researchers, including Klyachko [14] and Knutson and Tao [15, 16], Horn's conjecture was proved to be true. See also the survey paper [8]. Again, via the Jordan–Wielandt matrices, necessary and sufficient inequality conditions can be obtained for a triple of $\min\{m, n\}$ nonnegative numbers to arise as singular values of matrices $A, B, C \in \mathbb{C}^{m \times n}$ with $C = A + B$. Details of the conditions can be found in [8]. Here we list only part of the necessary inequalities:

$$\sum_{k \in K} \sigma_k(C) \leq \sum_{i \in I} \sigma_i(A) + \sum_{j \in J} \sigma_j(B)$$

for every (I, J, K) in $T_r^{\min\{m,n\}}$, $r = 1, 2, \ldots, \min\{m, n\}$.

In addition to comparing singular values, it is sometimes desired to compare the functions of singular values. In this regard, Uchiyama [26] recently proved that for each concave function $f : [0, \infty) \to [0, \infty)$ and for each $k \leq \min\{m, n\}$,

$$\sum_{l=1}^{k} f[\sigma_l(C)] \leq \sum_{l=1}^{k} f[\sigma_l(A)] + \sum_{l=1}^{k} f[\sigma_l(B)]. \tag{2.13}$$

In short, we can rewrite it as

$$f[\sigma(C)] \preceq_w f[\sigma(A)] + f[\sigma(B)].$$

It is easy to see that in order for (2.13) to hold even for the trivial case when the matrices are actually scalars, the function f has to be subadditive, i.e., $f(\alpha + \beta) \leq f(\alpha) + f(\beta)$. It turns out that a concave function $f : [0, \infty) \to [0, \infty)$ is always subadditive, because the concavity of f indicates

$$f(\alpha) = f\left(\frac{\alpha}{\alpha + \beta}(\alpha + \beta) + \frac{\beta}{\alpha + \beta} \cdot 0 \right) \geq \frac{\alpha}{\alpha + \beta} f(\alpha + \beta) + \frac{\beta}{\alpha + \beta} f(0),$$
$$f(\beta) = f\left(\frac{\beta}{\alpha + \beta}(\alpha + \beta) + \frac{\alpha}{\alpha + \beta} \cdot 0 \right) \geq \frac{\beta}{\alpha + \beta} f(\alpha + \beta) + \frac{\alpha}{\alpha + \beta} f(0),$$

which yields

$$f(\alpha + \beta) \leq f(\alpha) + f(\beta) - f(0) \leq f(\alpha) + f(\beta).$$

It also turns out that a concave function $f : [0, \infty) \to [0, \infty)$ is always nondecreasing.

On the other hand, it is possible to construct a nondecreasing subadditive, even continuous, function $f : [0, \infty) \to [0, \infty)$ which is not concave. Here is an example:

$$f(\alpha) = \begin{cases} \sin\alpha, & \alpha \in [0, \pi/2], \\ 2\alpha/\pi, & \alpha \in (\pi/2, \infty). \end{cases}$$

Hence, we might be curious to know whether we can relax the concavity requirement of the permissible functions in (2.13) and replace it by requiring the function to be subadditive, or to be nondecreasing and subadditive. It turns out that we cannot do so as soon as we go beyond the trivial scalar case. Consider

$$A = \begin{bmatrix} \alpha & \sqrt{\alpha\beta} \\ \sqrt{\alpha\beta} & \beta \end{bmatrix}, \quad B = \begin{bmatrix} \alpha & -\sqrt{\alpha\beta} \\ -\sqrt{\alpha\beta} & \beta \end{bmatrix}, \quad C = A + B = \begin{bmatrix} 2\alpha & 0 \\ 0 & 2\beta \end{bmatrix}.$$

Then $\sigma(A) = \{\alpha+\beta, 0\}, \sigma(B) = \{\alpha+\beta, 0\}, \sigma(C) = \{2\alpha, 2\beta\}$. If $f(0) = 0$, inequalities (2.13) require

$$f(2\alpha) + f(2\beta) \le 2f(\alpha + \beta),$$

which implies that f is concave.

To carry out matrix analysis, we may in principle use a metric (distance) which is not necessarily given by a norm. A metric ρ on $\mathbb{C}^{m\times n}$ is said to be unitarily invariant if $\rho(UAV, UBV) = \rho(A, B)$ for all $A, B \in \mathbb{C}^{m\times n}$ and for all unitary matrices U and V, and it is said to be translationally invariant if $\rho(A + C, B + C) = \rho(A, B)$ for all $A, B, C \in \mathbb{C}^{m\times n}$. Clearly, a unitarily and translationally invariant metric $\rho(A, B)$ depends only on the singular values of $A - B$. Unlike unitarily invariant norms, it is still unknown how all unitarily invariant metrics can be characterized. Nevertheless, inequalities (2.13) imply that a class of unitarily and translationally invariant metrics is given by

$$\rho_{\Phi,f}(A, B) = \Phi(f(\sigma(A - B))), \tag{2.14}$$

where $f : [0, \infty) \to [0, \infty)$ is a concave function satisfying $f(0) = 0$. Then the following question naturally arises: Why do people prefer to use a norm instead of one of the metrics given above by a nonlinear function f? This is partially because the norms, corresponding to the cases when f is a linear function, lead to intrinsic metrics, whereas other metrics do not [2]. A metric ρ on $\mathbb{C}^{m\times n}$ is said to be intrinsic if for each $A, B \in \mathbb{C}^{m\times n}$, there exists a continuous function $C : [0, 1] \to \mathbb{C}^{m\times n}$ such that $C(0) = A$, $C(1) = B$, and $\rho(A, B) = \rho(A, C(\lambda)) + \rho(C(\lambda), B)$ for all $\lambda \in [0, 1]$. Roughly speaking, a metric being intrinsic means that the distance between two points is given by the length of the shortest curve connecting the two points. The property of being intrinsic brings significant advantages in the analysis.

Polya once proved in [21] that for $x, y \in \mathbb{R}$, $x \preceq_w y$ implies $g(x) \preceq_w g(y)$ for all *convex* nondecreasing functions $g : \mathbb{R} \to \mathbb{R}$. However, $z \preceq_w x + y$ does not imply $f(z) \preceq_w f(x) + f(y)$ for all concave functions f. Inequalities (2.13) hold not only because inequalities (2.5) do, but also because $\sigma(A)$, $\sigma(B)$, and $\sigma(C)$ occur as the singular values of A, B, and $C = A + B$, respectively.

Inequalities (2.13) can be considered as an f-version of the Fan inequalities (2.5). It is natural to ask whether there is an f-version of the Lidskii–Wielandt inequalities (2.7) or an f-version of the Thompson inequalities (2.9). The answers are affirmative. Here, we state a theorem which summarizes the known inequalities (2.5), (2.7), (2.9), (2.13) and extends them to an f-version of Horn inequalities.

THEOREM 2.1. *Let $f : [0, \infty) \to [0, \infty)$ be a concave function. Then for $A, B, C \in \mathbb{C}^{m\times n}$ with $C = A + B$,*

$$\sum_{k\in K} f\big(\sigma_k(C)\big) \leq \sum_{i\in I} f\big(\sigma_i(A)\big) + \sum_{j\in J} f\big(\sigma_j(B)\big) \tag{2.15}$$

for every (I, J, K) in $T_r^{\min\{m,n\}}$, $r = 1, 2, \ldots, \min\{m, n\}$.

The proof of Theorem 2.1 requires the nonsquare extension of a matrix-valued triangle inequality [3].

LEMMA 2.2. *Let $f : [0, \infty) \to [0, \infty)$ be a concave function. Then for $A, B, C \in \mathbb{C}^{m\times n}$ with $C = A + B$, there exist unitary matrices U and V such that*

$$f(|C|) \leq U f(|A|)U^H + V f(|B|)V^H, \tag{2.16}$$

where $|A|$ means $(A^H A)^{1/2}$.

Proof. For the case when $m = n$, a recent proof is given in [3]. Now, we will show that the result can be extended to the nonsquare case. For the case when $m < n$, we may consider

$$A_\sharp = \begin{bmatrix} A \\ 0_{(n-m)\times n} \end{bmatrix}, \quad B_\sharp = \begin{bmatrix} B \\ 0_{(n-m)\times n} \end{bmatrix}, \quad C_\sharp = \begin{bmatrix} C \\ 0_{(n-m)\times n} \end{bmatrix}.$$

Since $A_\sharp, B_\sharp, C_\sharp$ are all square matrices with $C_\sharp = A_\sharp + B_\sharp$, and $|A_\sharp| = |A|$, $|B_\sharp| = |B|$, $|C_\sharp| = |C|$, inequality (2.16) immediately follows from the square case.

For the case when $m > n$, we may find a unitary matrix Q such that

$$QC = \begin{bmatrix} C_\flat \\ 0_{(m-n)\times n} \end{bmatrix}, \quad QA = \begin{bmatrix} A_\flat \\ A_\natural \end{bmatrix}, \quad QB = \begin{bmatrix} B_\flat \\ B_\natural \end{bmatrix},$$

and $A_\flat$, $B_\flat$, $C_\flat$ are all $n \times n$ matrices. Since $C_\flat = A_\flat + B_\flat$ and $|C| = |C_\flat|$, there exist unitary matrices U_1, V_1 such that

$$f(|C|) = f(|C_\flat|) \leq U_1 f(|A_\flat|)U_1^H + V_1 f(|B_\flat|)V_1^H. \tag{2.17}$$

It follows from $A_\flat^H A_\flat \leq A^H A$ and $B_\flat^H B_\flat \leq B^H B$ that $\lambda(|A_\flat|) \leq \lambda(|A|)$ and $\lambda(|B_\flat|) \leq \lambda(|B|)$. Since f is nondecreasing on $[0, \infty)$, we have

$$\lambda\big(f(|A_\flat|)\big) = f\big(\lambda(|A_\flat|)\big) \leq f\big(\lambda(|A|)\big) = \lambda\big(f(|A|)\big),$$

and similarly, $\lambda\big(f(|B_\flat|)\big) \leq \lambda\big(f(|B|)\big)$. Then it follows from an exercise statement in [1, p. 74], and can be shown by simple diagonalization argument, that there exist unitary matrices U_2, V_2 such that

$$f(|A_\flat|) \leq U_2 f(|A|)U_2^H, \quad f(|B_\flat|) \leq V_2 f(|B|)V_2^H. \tag{2.18}$$

Combining (2.17) and (2.18), we obtain (2.16). ∎

Proof of Theorem 2.1. From Lemma 2.2, we know that (2.16) holds. Let $\hat{C}$ be the matrix on the right-hand side of (2.16). Then for any $1 \leq k \leq n$, we have

$$f(\sigma_k(C)) = \lambda_k\left(f(|C|)\right) \leq \lambda_k(\hat{C}).$$

Here, we have used the fact that f is nondecreasing on $[0,\infty)$. Since any (I,J,K) belonging to $T_r^{\min\{m,n\}}$, $r=1,2,\ldots,\min\{m,n\}$, also belongs to T_r^n, it follows from (2.10) that

$$\begin{aligned}\sum_{k\in K} f\big(\sigma_k(C)\big) &\le \sum_{k\in K}\lambda_k(\hat{C})\\ &\le \sum_{i\in I}\lambda_i\left(f\big(|A|\big)\right)+\sum_{j\in J}\lambda_j\left(f\big(|B|\big)\right)\\ &= \sum_{i\in I} f\big(\sigma_i(A)\big)+\sum_{j\in J} f\big(\sigma_j(B)\big). \qquad \square\end{aligned}$$

From the f-version of the Lidskii–Wielandt singular value inequalities, the f-version of the Mirsky inequalities follows.

COROLLARY 2.3. *Let Φ be a symmetric gauge function, and let $f:[0,\infty)\to[0,\infty)$ be a concave function with $f(0)=0$. Then for $A,B\in\mathbb{C}^{m\times n}$,*

$$\Phi[f(\sigma(A))-f(\sigma(B))]\le \rho_{\Phi,f}(A,B),$$

where $\rho_{\Phi,f}(A,B)$ is given by (2.14).

3. Canonical angle inequalities—new results. Recently, it was proved in [20] using a geometric approach, and proved again in [22] (without knowledge of [20]) using an algebraic approach, that

$$\Phi[\theta(\mathcal{X},\mathcal{Z})]\le \Phi[\theta(\mathcal{X},\mathcal{Y})]+\Phi[\theta(\mathcal{Y},\mathcal{Z})] \tag{3.1}$$

and

$$\Phi\left[\theta(\mathcal{X},\mathcal{Z})-\theta(\mathcal{X},\mathcal{Y})\right]\le \Phi[\theta(\mathcal{Y},\mathcal{Z})] \tag{3.2}$$

for $\mathcal{X},\mathcal{Y},\mathcal{Z}\in\mathcal{G}_{m,n}$ and all symmetric gauge functions Φ.

Owing to (3.1), a family of unitarily invariant metrics on $\mathcal{G}_{m,n}$, called angular metrics, is defined by

$$\rho_\Phi(\mathcal{X},\mathcal{Y})=\Phi[\theta(\mathcal{X},\mathcal{Y})]$$

for all symmetric gauge functions Φ. It has also been shown in [22] that these metrics are intrinsic; that is, for each pair $\mathcal{X}$ and $\mathcal{Y}$, there exists a continuous function $\mathcal{Z}:[0,1]\to\mathcal{G}_{m,n}$ such that $\mathcal{Z}(0)=\mathcal{X}$, $\mathcal{Z}(1)=\mathcal{Y}$, and $\rho_\Phi(\mathcal{X},\mathcal{Y})=\rho_\Phi(\mathcal{X},\mathcal{Z}(\lambda))+\rho_\Phi(\mathcal{Z}(\lambda),\mathcal{Y})$ for all $\lambda\in[0,1]$. We also conjecture that every unitarily invariant intrinsic metric on $\mathcal{G}_{m,n}$ is given by $\Phi[\theta(\mathcal{X},\mathcal{Y})]$ for some symmetric gauge function Φ.

Inequalities (3.2) can be rewritten as

$$\Phi\left[\theta(\mathcal{X},\mathcal{Z})-\theta(\mathcal{X},\mathcal{Y})\right]\le \rho_\Phi(\mathcal{Y},\mathcal{Z});$$

that is, the perturbation of canonical angles is bounded by the perturbation of the subspaces involved, respectively, measured by a symmetric gauge function and the corresponding angular metric. This inequality is instrumental in the perturbation analysis of linear subspaces, as demonstrated in [22].

Specializing inequalities (3.1) and (3.2) to the Fan k-norms gives the following canonical angle counterparts of the Fan inequalities:

$$\sum_{l=1}^{k}\theta_l(\mathcal{X},\mathcal{Z})\le \sum_{l=1}^{k}\theta_l(\mathcal{X},\mathcal{Y})+\sum_{l=1}^{k}\theta_l(\mathcal{Y},\mathcal{Z}) \tag{3.3}$$

and the following canonical angle counterparts of the Lidskii–Wielandt inequalities:

$$\sum_{l=1}^{k}\theta_{i_l}(\mathcal{X},\mathcal{Z})\le\sum_{l=1}^{k}\theta_{i_l}(\mathcal{X},\mathcal{Y})+\sum_{l=1}^{k}\theta_l(\mathcal{Y},\mathcal{Z}) \tag{3.4}$$

for $\mathcal{X},\mathcal{Y},\mathcal{Z}\in\mathcal{G}_{m,n}$, $1\le k\le m$, and $1\le i_1<\cdots<i_k\le m$. Again, in the majorization language, inequalities (3.3) can be rewritten as

$$\theta(\mathcal{X},\mathcal{Z})\preceq_w\theta(\mathcal{X},\mathcal{Y})+\theta(\mathcal{Y},\mathcal{Z}),$$

and inequalities (3.4) imply

$$\theta(\mathcal{X},\mathcal{Z})-\theta(\mathcal{X},\mathcal{Y})\preceq_w\theta(\mathcal{Y},\mathcal{Z}).$$

In fact, the Horn inequalities also have canonical angle counterparts, as had been conjectured in [20] and [22]. Generally, we now have a result for canonical angles which is completely analogous to Theorem 2.1 for singular values.

THEOREM 3.1. *Let $f:[0,\pi/2]\to[0,\infty)$ be a nondecreasing concave function. Then for $\mathcal{X},\mathcal{Y},\mathcal{Z}\in\mathcal{G}_{m,n}$,*

$$\sum_{k\in K}f\big(\theta_k(\mathcal{X},\mathcal{Z})\big)\le\sum_{i\in I}f\big(\theta_i(\mathcal{X},\mathcal{Y})\big)+\sum_{j\in J}f\big(\theta_j(\mathcal{Y},\mathcal{Z})\big) \tag{3.5}$$

for every $(I,J,K)\in T_r^m$, $r=1,2,\dots,m$.

The proof of Theorem 3.1 will be given in section 5. Since the inequalities cover inequalities (3.3) as well as inequalities (3.4) as special cases, the proof in section 5 also proves again the Fan and Mirsky counterparts in a better way.

Theorem 3.1 can be specialized as follows to yield the canonical angle counterparts of the f-version of the Fan, Lidskii–Wielandt, and Thompson inequalities:

1. For $k=1,2,\dots,m$,

$$\sum_{l=1}^{k}f\big(\theta_l(\mathcal{X},\mathcal{Z})\big)\le\sum_{l=1}^{k}f\big(\theta_l(\mathcal{X},\mathcal{Y})\big)+\sum_{l=1}^{k}f\big(\theta_l(\mathcal{Y},\mathcal{Z})\big). \tag{3.6}$$

2. For $1\le k\le m$ and $1\le i_1<\cdots<i_k\le m$,

$$\sum_{l=1}^{k}f(\theta_{i_l}(\mathcal{X},\mathcal{Z}))\le\sum_{l=1}^{k}f(\theta_{i_l}(\mathcal{X},\mathcal{Y}))+\sum_{l=1}^{k}f(\theta_l(\mathcal{Y},\mathcal{Z})). \tag{3.7}$$

3. For each $1\le i_1<\cdots<i_k\le m$, $1\le j_1<\cdots<j_k\le m$ with $i_k+j_k-k\le m$,

$$\sum_{l=1}^{k}f\left(\theta_{i_l+j_l-l}(\mathcal{X},\mathcal{Z})\right)\le\sum_{l=1}^{k}f\left(\theta_{i_l}(\mathcal{X},\mathcal{Y})\right)+\sum_{l=1}^{k}\left(\theta_{j_l}(\mathcal{Y},\mathcal{Z})\right). \tag{3.8}$$

Inequalities (3.6) imply that a big family of unitarily invariant metrics is given by

$$\rho_{\Phi,f}(\mathcal{X},\mathcal{Y})=\Phi[f(\theta(\mathcal{X},\mathcal{Y}))]$$

for symmetric gauge functions Φ and nondecreasing concave functions $f:[0,\pi/2]\to[0,\infty)$ with $f(0)=0$. Two such functions are $f(\alpha)=\sin\alpha$ and $f(\alpha)=2\sin(\alpha/2)$. It has been well known that $\Phi[f(\theta(\mathcal{X},\mathcal{Y}))]$ defines unitarily invariant metrics for these two particular functions, which were established in completely different ways, and are called gap metrics and Hausdorff metrics, respectively. Actually, before the introduction of the angular metrics $\Phi[\theta(\mathcal{X},\mathcal{Y})]$, the gap metrics and the Hausdorff metrics were commonly used. Now, we have at our disposal a vast family of unitarily invariant metrics which includes gap, Hausdorff, and new angular metrics as special cases. In spite of many possible choices of f, we hold the conviction that the angular metrics given by $f(\alpha)=\alpha$ are the most favorable since they are intrinsic and are conjectured to be the only intrinsic ones. For a proof of why metrics $\Phi[\theta(\mathcal{X},\mathcal{Y})]$ are intrinsic and a two-dimensional example showing why the intrinsic property is desirable, see [22].

The family of unitarily invariant metrics $\rho_{\Phi,f}$ mentioned above can be further enlarged. Let $g:[0,\infty)\to[0,\infty)$ be a nondecreasing subadditive function with $g(0)=0$. Then

$$\rho_{g,\Phi,f}=g\{\Phi[f(\theta(\mathcal{X},\mathcal{Y}))]\}$$

also defines a metric on $\mathcal{G}_{m,n}$. Does this family exhaust all unitarily invariant metrics or those that are interesting? The answer is not clear.

Inequalities (3.7) imply that

$$\Phi[f(\theta(\mathcal{X},\mathcal{Z}))-f(\theta(\mathcal{X},\mathcal{Y}))]\le\rho_{\Phi,f}(\mathcal{Y},\mathcal{Z})$$

for nondecreasing concave functions $f:[0,\pi/2]\to[0,\infty)$ with $f(0)=0$. When f is given by $f(\alpha)=\sin\alpha$, these inequalities were proved in [17] using a completely different method.

4. Direct rotations and their squares. In this section, we first review the theory of direct rotations laid out by Davis and Kahan [5] and then present a nice relation of direct rotations between three subspaces.

For $\mathcal{X},\mathcal{Y}\in\mathcal{G}_{m,n}$, let us denote the orthogonal projections onto $\mathcal{X}$, $\mathcal{X}_\perp$, $\mathcal{Y}$, and $\mathcal{Y}_\perp$ by $P_\mathcal{X}$, $P_{\mathcal{X}_\perp}$, $P_\mathcal{Y}$, and $P_{\mathcal{Y}_\perp}$, respectively. A direct rotation from $\mathcal{X}$ to $\mathcal{Y}$ is a unitary matrix R satisfying $R\mathcal{X}=\mathcal{Y}$, $P_\mathcal{X}RP_\mathcal{X}+P_{\mathcal{X}_\perp}RP_{\mathcal{X}_\perp}\ge 0$, and $P_{\mathcal{X}_\perp}RP_\mathcal{X}+P_\mathcal{X}RP_{\mathcal{X}_\perp}$ is skew-Hermitian. A less explicit, but more concise, definition is that the direct rotations from $\mathcal{X}$ to $\mathcal{Y}$ are exactly the minimizers of $\|I-V\|_F$, the Frobenius norm of $I-V$, among all unitary transformations V mapping $\mathcal{X}$ onto $\mathcal{Y}$. This latter definition, which is not directly stated in [5] but can be easily proved using the techniques developed therein, reveals that a direct rotation transforms $\mathcal{X}$ into $\mathcal{Y}$ in the most economic way in terms of the Frobenius norm. Here the Frobenius norm is special since some unitarily invariant norms, such as the trace norm, of $I-V$ are in general not minimized by the direct rotations, and some other unitarily invariant norms, such as the spectral norm, of $I-V$ have additional minimizers to the direct rotations.

Suppose the columns of $X_1,Y_1\in\mathbb{C}^{n\times m}$ form orthonormal bases for $\mathcal{X}$ and $\mathcal{Y}$. Also suppose $X_2,Y_2\in\mathbb{C}^{n\times(n-m)}$ have columns forming orthonormal bases for $\mathcal{X}_\perp$ and $\mathcal{Y}_\perp$, the orthogonal complements of $\mathcal{X}$ and $\mathcal{Y}$, respectively. Then $X=[X_1\ X_2]$ and $Y=[Y_1\ Y_2]$ are unitary matrices. By the CS decomposition of unitary matrices, as presented in [23, Theorem 5.2], there exist $n\times n$ unitary matrices $E=\operatorname{diag}(E_1,E_2)$, $F=\operatorname{diag}(F_1,F_2)$ partitioned consistently with X and Y such that

$$E^HX^HYF=\operatorname{diag}\left[I_{m-l}\,,\ \begin{bmatrix}\Gamma & -\Sigma\\ \Sigma & \Gamma\end{bmatrix},\ I_{n-m-l}\right],$$

where $l=\min\{m,n-m\}$, and Γ and Σ are diagonal matrices with the diagonal entries in $[0,1]$ satisfying $\Gamma^2+\Sigma^2=I_l$. Hence, the diagonal entries of Γ are the cosines of the first l canonical angles between $\mathcal{X}$ and $\mathcal{Y}$; the diagonal entries of Σ are the sines of the first l canonical angles. The rest of the $m-l$ canonical angles between $\mathcal{X}$ and $\mathcal{Y}$ are always zero.

We can always assume $2m \le n$. Otherwise, we can augment $\mathcal{X}$ and $\mathcal{Y}$ as the subspaces of a higher dimensional space $\mathbb{C}^{2m}$ through

$$\tilde{\mathcal{X}} = \operatorname{span}\begin{bmatrix} X_1 \\ 0_{(2m-n)\times m} \end{bmatrix}, \quad \tilde{\mathcal{Y}} = \operatorname{span}\begin{bmatrix} Y_1 \\ 0_{(2m-n)\times m} \end{bmatrix}.$$

Obviously, $\theta(\mathcal{X}, \mathcal{Y}) = \theta(\tilde{\mathcal{X}}, \tilde{\mathcal{Y}})$. Also we can make the same augmentations when we consider three subspaces of $\mathbb{C}^n$ if $2m > n$. Therefore, for convenience, we always assume $2m \le n$ in the following.

Set $\hat{X} = XE$ and $\hat{Y} = YF$. We have

$$\hat{X}^H \hat{Y} = \begin{bmatrix} \Gamma & -\Sigma & 0 \\ \Sigma & \Gamma & 0 \\ 0 & 0 & I_{n-2m} \end{bmatrix}.$$

A direct rotation from $\mathcal{X}$ to $\mathcal{Y}$ is given by

$$R = \hat{X} \begin{bmatrix} \Gamma & -\Sigma & 0 \\ \Sigma & \Gamma & 0 \\ 0 & 0 & I_{n-2m} \end{bmatrix} \hat{X}^H = \hat{X} \exp \begin{bmatrix} 0 & -A \\ A^H & 0 \end{bmatrix} \hat{X}^H, \tag{4.1}$$

where $A = \begin{bmatrix} \operatorname{diag}\theta(\mathcal{X},\mathcal{Y}) & 0_{m,n-2m} \end{bmatrix}$.

While a direct rotation takes a complicated form, its square can be expressed in terms of orthogonal projections in a simple way. Denote $\hat{X}_1 = X_1 E_1$, $\hat{X}_2 = X_2 E_2$. Then $\hat{X} = [\hat{X}_1\ \hat{X}_2]$, $P_{\mathcal{X}} = \hat{X}_1 \hat{X}_1^H$, and $P_{\mathcal{X}_\perp} = \hat{X}_2 \hat{X}_2^H$. Recall that the reflection in $\mathcal{X}$ is a linear transformation mapping everything in $\mathcal{X}$ to itself and everything in $\mathcal{X}_\perp$ to its negative. Hence the reflection in $\mathcal{X}$ is

$$P_{\mathcal{X}} - P_{\mathcal{X}_\perp} = \begin{bmatrix} \hat{X}_1 & \hat{X}_2 \end{bmatrix} \begin{bmatrix} I_m & 0 \\ 0 & -I_{n-m} \end{bmatrix} \begin{bmatrix} \hat{X}_1^H \\ \hat{X}_2^H \end{bmatrix} = \hat{X} \begin{bmatrix} I_m & 0 \\ 0 & -I_{n-m} \end{bmatrix} \hat{X}^H.$$

From (4.1) and the above, we see at once $R^{-1} = (P_{\mathcal{X}} - P_{\mathcal{X}_\perp}) R (P_{\mathcal{X}} - P_{\mathcal{X}_\perp})$.

Since R is the direct rotation from $\mathcal{X}$ to $\mathcal{Y}$, we can see that the columns of $R\hat{X}_1$ form orthonormal bases for $\mathcal{Y}$, and hence $P_{\mathcal{Y}} = R\hat{X}_1 (R\hat{X}_1)^H = RP_{\mathcal{X}}R^{-1}$. From the fact that $P_{\mathcal{X}_\perp} = I - P_{\mathcal{X}}$ and $P_{\mathcal{Y}_\perp} = I - P_{\mathcal{Y}}$, we have $P_{\mathcal{Y}_\perp} = RP_{\mathcal{X}_\perp}R^{-1}$. Simple algebraic manipulation gives

$$R^2 = (P_{\mathcal{Y}} - P_{\mathcal{Y}^\perp})(P_{\mathcal{X}} - P_{\mathcal{X}^\perp}).$$

Up to this point in this section, everything is from Davis and Kahan [5]. Now, let us bring in the third space $\mathcal{Z}$. Then naturally, we have a direct rotation S from $\mathcal{Y}$ to $\mathcal{Z}$ and a direct rotation T from $\mathcal{X}$ to $\mathcal{Z}$ satisfying

$$S^2 = (P_{\mathcal{Z}} - P_{\mathcal{Z}^\perp})(P_{\mathcal{Y}} - P_{\mathcal{Y}^\perp}),$$
$$T^2 = (P_{\mathcal{Z}} - P_{\mathcal{Z}^\perp})(P_{\mathcal{X}} - P_{\mathcal{X}^\perp}).$$

When we multiply R^2 and S^2 and notice $(P_{\mathcal{Y}} - P_{\mathcal{Y}^\perp})^2 = I$, a miracle occurs and T^2 emerges, as stated in the following theorem.

THEOREM 4.1. *For $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \in \mathcal{G}_{m,n}$, let R, S, T be the direct rotations from $\mathcal{X}$ to $\mathcal{Y}$, from $\mathcal{Y}$ to $\mathcal{Z}$, and from $\mathcal{X}$ to $\mathcal{Z}$, respectively. Then*

$$S^2 R^2 = T^2. \tag{4.2}$$

The proof of Theorem 4.1 in the arguments above is due to Davis [4].

The product on the left-hand side of (4.2) can be transformed into the exponential of a sum by using a matrix exponential formula of Thompson [24], which asserts that for skew-Hermitian matrices M and N, there exist unitary matrices U and V such that

$$\exp(M)\exp(N) = \exp(UMU^H + VNV^H). \tag{4.3}$$

It is interesting to notice that the proof of the above formula (4.3) relies on the validity of Horn's conjecture in [9], which was only proved to be true over 10 years after Thompson proposed the formula in [24].

By (4.2) and the matrix exponential formula (4.3), we can obtain the following relation of the canonical angles.

COROLLARY 4.2. *For $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \in \mathcal{G}_{m,n}$, let*

$$\begin{aligned} A &= \begin{bmatrix} \operatorname{diag}\theta(\mathcal{X},\mathcal{Y}) & 0_{m,n-2m} \end{bmatrix}, \\ B &= \begin{bmatrix} \operatorname{diag}\theta(\mathcal{Y},\mathcal{Z}) & 0_{m,n-2m} \end{bmatrix}, \\ C &= \begin{bmatrix} \operatorname{diag}\theta(\mathcal{X},\mathcal{Z}) & 0_{m,n-2m} \end{bmatrix}. \end{aligned}$$

Then there exist unitary matrices U and V such that

$$\exp\left(\begin{bmatrix} 0 & -2C \\ 2C^H & 0 \end{bmatrix}\right) = \exp\left(U \begin{bmatrix} 0 & -2A \\ 2A^H & 0 \end{bmatrix} U^H + V \begin{bmatrix} 0 & -2B \\ 2B^H & 0 \end{bmatrix} V^H\right). \tag{4.4}$$

5. Proof of Theorem 3.1.

Define

$$h(\alpha) = \begin{cases} f(\alpha), & \alpha \in [0, \pi/2], \\ f(\pi/2), & \alpha \in (\pi/2, \infty); \end{cases} \tag{5.1}$$

then h is a nondecreasing and concave function on $[0, \infty)$.

As shown in Corollary 4.2, the canonical angles between subspaces $\mathcal{X}, \mathcal{Y}$, and $\mathcal{Z}$ satisfy the relationship (4.4). Hence, the eigenvalues of

$$\begin{bmatrix} 0 & -2C \\ 2C^H & 0 \end{bmatrix} \tag{5.2}$$

and those of

$$U \begin{bmatrix} 0 & -2A \\ 2A^H & 0 \end{bmatrix} U^H + V \begin{bmatrix} 0 & -2B \\ 2B^H & 0 \end{bmatrix} V^H \tag{5.3}$$

are equal modulo $\pm i2\pi$, where $i = \sqrt{-1}$. Since the eigenvalues of matrix (5.2) belong to $i[-\pi, \pi]$, and those of matrix (5.3) are known to belong to $i[-2\pi, 2\pi]$, each eigenvalue of (5.3) either is equal to its corresponding eigenvalue of (5.2) or belongs to $i[-2\pi, -\pi] \cup i[\pi, 2\pi]$, hence having no less absolute value than its corresponding eigenvalue of (5.2). Let

$$\tilde{A} = U \begin{bmatrix} 0 & -A \\ A^H & 0 \end{bmatrix} U^H, \quad \tilde{B} = V \begin{bmatrix} 0 & -B \\ B^H & 0 \end{bmatrix} V^H, \quad \tilde{C} = \begin{bmatrix} 0 & -C \\ C^H & 0 \end{bmatrix}. \tag{5.4}$$

Similarly, each eigenvalue of $\tilde{A} + \tilde{B}$ has no less absolute value than its corresponding eigenvalue of $\tilde{C}$. Since both matrix $\tilde{A} + \tilde{B}$ and matrix $\tilde{C}$ are skew-Hermitian, their singular values are the absolute values of their eigenvalues. Arranging their singular values in decreasing order, respectively, we can see that

$$\sigma_j(\tilde{C}) \leq \sigma_j(\tilde{A} + \tilde{B}), \quad j = 1, 2, \ldots, n. \tag{5.5}$$

Note that the singular values of $\tilde{C}$ are ordered decreasingly as

$$\theta_1(\mathcal{X}, \mathcal{Z}),\, \theta_1(\mathcal{X}, \mathcal{Z}), \ldots, \theta_m(\mathcal{X}, \mathcal{Z}),\, \theta_m(\mathcal{X}, \mathcal{Z}),\, 0, \ldots, 0,$$

and the singular values of $\tilde{A}$ and $\tilde{B}$ are similarly ordered.

Now we are ready to prove the main result, Theorem 3.1. For each $(I, J, K) \in T_r^m$, since the first $2m$ singular values of $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ appear in pairs, respectively, we consider the following triple (I', J', K') of subsets of $\{1, 2, \ldots, 2m\}$ of cardinality $2r$:

$$I' = \{2i-1, 2i | i \in I\}, \quad J' = \{2j-1, 2j | j \in J\}, \quad K' = \{2k-1, 2k | k \in K\}.$$

We will show that (I', J', K') belongs to T_{2r}^{2m}. First, define a one-to-one correspondence between a set $I = \{i_1 < i_2 < \cdots < i_r\}$ of r positive integers and a decreasing sequence $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n)$ of r nonnegative integers as

$$\lambda(I) = (i_r - r,\, i_{r-1} - (r-1), \ldots, i_2 - 2,\, i_1 - 1).$$

Then it follows from [8, Theorem 2] that a triple (I, J, K) is in T_r^m if and only if the corresponding triple $(\lambda(I), \lambda(J), \lambda(K))$ occurs as the eigenvalues of a triple of $r \times r$ Hermitian matrices L, M, N with $N = L + M$. For $(I, J, K) \in T_r^m$, let L, M, $N = L + M$ be a triple of $r \times r$ Hermitian matrices with eigenvalues $(\lambda(I), \lambda(J), \lambda(K))$, respectively. Then we can easily see that $(\lambda(I'), \lambda(J'), \lambda(K'))$ occurs as the eigenvalues of $2r \times 2r$ Hermitian matrices

$$\begin{bmatrix} 2L & 0 \\ 0 & 2L \end{bmatrix}, \begin{bmatrix} 2M & 0 \\ 0 & 2M \end{bmatrix}, \begin{bmatrix} 2N & 0 \\ 0 & 2N \end{bmatrix} = \begin{bmatrix} 2L & 0 \\ 0 & 2L \end{bmatrix} + \begin{bmatrix} 2M & 0 \\ 0 & 2M \end{bmatrix},$$

respectively. Therefore, (I', J', K') is in T_{2r}^{2m} and also in T_{2r}^n.

From (5.5) and the nondecreasing property of h, we can see that

$$2 \sum_{k \in K} h\big(\theta_k(\mathcal{X}, \mathcal{Z})\big) = \sum_{k \in K'} h\big(\sigma_k(\tilde{C})\big) \leq \sum_{k \in K'} h\big(\sigma_k(\tilde{A} + \tilde{B})\big). \tag{5.6}$$

By [3], we know that there exist unitary matrices $\tilde{U}$ and $\tilde{V}$ such that

$$h(|\tilde{A} + \tilde{B}|) \leq \tilde{U} h(|\tilde{A}|) \tilde{U}^H + \tilde{V} h(|\tilde{B}|) \tilde{V}^H.$$

From the above and the Horn inequalities on eigenvalues of the sum of Hermitian matrices, we have

$$
\begin{aligned}
\sum_{k\in K'} h\big(\sigma_k(\tilde{A}+\tilde{B})\big) &= \sum_{k\in K'} \lambda_k\big(h(|\tilde{A}+\tilde{B}|)\big)\\
&\le \sum_{k\in K'} \lambda_k\big(\tilde{U}h(|\tilde{A}|)\tilde{U}^H + \tilde{V}h(|\tilde{B}|)\tilde{V}^H\big)\\
&\le \sum_{i\in I'} \lambda_i\big(h(|\tilde{A}|)\big) + \sum_{j\in J'} \lambda_j\big(h(|\tilde{B}|)\big)\\
&= \sum_{i\in I'} h\big(\sigma_i(|\tilde{A}|)\big) + \sum_{j\in J'} h\big(\sigma_j(|\tilde{B}|)\big)\\
&= 2\sum_{i\in I} h\big(\theta_i(\mathcal{X},\mathcal{Y})\big) + 2\sum_{j\in J} h\big(\theta_j(\mathcal{Y},\mathcal{Z})\big). \qquad (5.7)
\end{aligned}
$$

Then from (5.6), (5.7), and the fact that $f = h$ on $[0, \pi/2]$, it immediately follows that

$$\sum_{k\in K} f\big(\theta_k(\mathcal{X},\mathcal{Z})\big) \le \sum_{i\in I} f\big(\theta_i(\mathcal{X},\mathcal{Y})\big) + \sum_{j\in J} f\big(\theta_j(\mathcal{Y},\mathcal{Z})\big).$$

This completes the proof. ∎

Acknowledgments. We thank Professors Jean-Christopher Bourin, Chandler Davis, Chi-Kwong Li, and Xingzhi Zhan for interesting discussions. We also thank the anonymous referees for constructive suggestions that helped us to improve the paper considerably.

REFERENCES

[1] R. Bhatia, *Matrix Analysis*, Springer, New York, 1991.

[2] D. Burago, Y. Burago, and S. Ivanov, *A Course in Metric Geometry*, AMS, Providence, RI, 2001.

[3] J.-C. Bourin and M. Uchiyama, *A matrix subadditivity inequality for $f(A+B)$ and $f(A)+f(B)$*, Linear Algebra Appl., 423 (2007), pp. 512–518.

[4] C. Davis, *Private Email Commmunication to R. Bhatia, W. M. Kahan, C.-K. Li, and L. Qiu*, 2006.

[5] C. Davis and W. M. Kahan, *The rotation of eigenvectors by a perturbation*. III, SIAM J. Numer. Anal., 7 (1970), pp. 1–46.

[6] K. Fan, *On a theorem of Weyl concerning the eigenvalues of linear transformations*. I, Proc. Nat. Acad. Sci. U.S.A., 35 (1949), pp. 652–655.

[7] K. Fan, *Maximum properties and inequalities for the eigenvalues of completely continuous operators*, Proc. Nat. Acad. Sci. U.S.A., 37 (1951), pp. 760–766.

[8] W. Fulton, *Eigenvalues, invariant factors, highest weights, and Schubert calculus*, Bull. Amer. Math. Soc., 37 (2000), pp. 209–249.

[9] A. Horn, *Eigenvalues of sums of Hermitian matrices*, Pacific J. Math., 12 (1962), pp. 225–241.

[10] H. Hotelling, *Analysis of a complex of statistical variables into principal components*, J. Educational Psychology, 24 (1933), pp. 417–441; 498–520.

[11] H. Hotelling, *Relations between two sets of variates*, Biometrika, 28 (1936) pp. 321–377.

[12] C. Jordan, *Mémoire sur les formes bilinéraires*, J. Math. Pures Appl., 19 (1874), pp. 35–54.

[13] C. Jordan, *Essai sur la géométrie à n dimensions*, Bull. Soc. Math., 3 (1875), pp. 103–174.

[14] A. A. Klyachko, *Stable bundles, representation theory and Hermitian operators*, Selecta Math., 4 (1998), pp. 419–445.

[15] A. Knutson and T. Tao, *The Honeycomb model of $GL_n(\mathbb{C})$ tensor products* I: *Proof of the saturation conjecture*, J. American Math. Soc., 12 (1999), pp. 1055–1090.

[16] A. Knutson and T. Tao, *Honeycombs and sums of Hermitian matrices*, Notices Amer. Math. Soc., 48 (2001), pp. 175–186.

[17] A. V. Knyazev and M. E. Argentati, *Majorization for changes in angles between subspaces, Ritz values, and graph Laplacian spectra*, SIAM J. Matrix Anal. Appl., 29 (2006), pp. 15–32.

[18] V. B. Lidskii, *On the proper values of a sum and product of symmetric matrices*, Dokl. Akad. Nauk SSSR (N.S.), 75 (1950), pp. 769–772 (in Russian).

[19] L. Mirsky, *Symmetric gauge functions and unitarily invariant norms*, Quart. J. Math. Oxford Ser. (2), 11 (1960), pp. 50–59.

[20] Y. A. Neretin, *On Jordan angles and the triangle inequality in Grassmann manifolds*, Geom. Dedicata, 86 (2001), pp. 81–92.

[21] G. Polya, *Remark on Weyl's note "Inequalities between the two kinds of eigenvalues of a linear transformation,"* Proc. Nat. Acad. Sci. U.S.A., 36 (1950), pp. 49–51.

[22] L. Qiu, Y. Zhang, and C.-K. Li, *Unitarily invariant metrics on the Grassmann space*, SIAM J. Matrix Anal. Appl., 27 (2005), pp. 507–531.

[23] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.

[24] R. C. Thompson, *Proof of a conjectured exponential formula*, Linear and Multilinear Algebra, 19 (1986), pp. 187–197.

[25] R. C. Thompson, *Singular value inequalities for matrix sums and minors*, Linear Algebra Appl., 11 (1975), pp. 251–269.

[26] M. Uchiyama, *Subadditivity of eigenvalue sums*, Proc. Amer. Math. Soc., 134 (2005), pp. 1405–1412.

[27] J. von Neumann, *Some matrix-inequalities and metrization of matrix-space*, Tomsk. Univ. Rev., 1 (1937), pp. 286–300.

[28] H. Wielandt, *An extremum property of sums of eigenvalues*, Proc. Amer. Math. Soc., 6 (1955), pp. 106–110.