GENERALIZATIONS OF KY FAN'S DOMINANCE THEOREM AND SOME RESULTS OF SO AND ZIETAK

CHARLES L. DOLBERRY* AND TIN-YAU TAM†

Abstract. A generalization of the Dominance Theorem of Ky Fan on the unitarily invariant norms is obtained. We also extend some results of So and Zietak on unitarily invariant norms including the characterization of the set of the dual matrices of a given matrix, the extreme points and the faces of the unit ball.

Key words. Ky Fan's dominance theorem, Eaton triple, reduced triple, dual norm, dual element, face, unit ball

AMS subject classifications. 49J52, 15A18

1. Introduction. A norm $\|\cdot\| : \mathbb{C}_{p\times q} \to \mathbb{R}$ is said to be unitarily invariant if for any $A \in \mathbb{C}_{p\times q}$, $\|UAV\| = \|A\|$ for all $U \in U(p)$, $V \in U(q)$, where $U(p)$ denotes the group of $p \times p$ unitary matrices. The characterization of unitarily invariant norms is well-known and is due to von Neumann [20, 2]. Among the unitarily invariant norms, Ky Fan's k-norms, $\|\cdot\|_k$, are the most important ones due to the following result of Ky Fan [2]. The Ky Fan k-norms $\|\cdot\|_k : \mathbb{C}_{p\times p} \to \mathbb{R}$, defined by $\|A\|_k = \sum_{i=1}^k s_i(A)$, where $s_1(A) \geq \cdots \geq s_n(A)$ are the singular values of A. Our main purpose in the next section is to extend the following result.

Theorem 1.1. (Ky Fan) *Let $A, B \in \mathbb{C}_{n\times n}$. Then $\|A\| \leq \|B\|$ for all unitarily invariant norms if and only if $\|A\|_k \leq \|B\|_k$ for all $k = 1, \ldots, n$.*

In sections 3 and 4 generalizations of the results of Zietak [23] are obtained, namely, the characterization of the dual matrices of a given matrix and the study of the faces of the unit ball, both with respect to a unitarily invariant norm. In section 5, a result of So [15] is generalized.

Here is a framework for our study which only requires basic knowledge of linear algebra. Let G be a closed subgroup of the orthogonal group on a finite dimensional real inner product space V. The triple (V, G, F) is an *Eaton triple* if $F \subset V$ is a nonempty closed convex cone such that

(A1) $Gx \cap F$ is nonempty for each $x \in V$.

(A2) $\max_{g\in G}(x, gy) = (x, y)$ for all $x, y \in F$.

Example 1.2. Consider the *symmetric group* S_n. It can be thought of as a subgroup of the group $O_n(\mathbb{R})$ of $n \times n$ orthogonal matrices in the following way. Make a permutation act on $\mathbb{R}^n$ by permuting the standard basis vectors $e_1, \ldots, e_n$ (permute the subscripts). Observe that the transposition (ij) acts as a reflection, sending $e_i - e_j$ to its negative and fixing pointwise the orthogonal complement, which consists of all vectors in $\mathbb{R}^n$ having equal *ith* and *jth* components. Since S_n is generated by

*Department of Mathematics, 218 Parker Hall, Auburn University, Auburn, AL 36849-5310, USA (email: dolbecl@auburn.edu).

†Department of Mathematics, 218 Parker Hall, Auburn University, Auburn, AL 36849-5310, USA (email: tamtiny@auburn.edu) homepage: http://www.auburn.edu/∼tamtiny.

transpositions, it is a reflection group. The triple $(\mathbb{R}^n, S_n, F)$ is an Eaton triple, where $F = \{(a_1, \ldots, a_n) \in \mathbb{R}^n : a_1 \geq \cdots \geq a_n\}$.

The Eaton triple (W, H, F) is called a *reduced* triple of the Eaton triple (V, G, F) if it is an Eaton triple and $W := \operatorname{span} F$ and $H := \{g|_W : g \in G,\ gW = W\} \subset O(W)$, the orthogonal group of W. For $x \in V$, let $F(x)$ denote the unique element of the singleton set $Gx \cap F$. It is known that H is a finite reflection group [13]. Also see [11, 16] for the normal decomposition systems and normal decomposition subsystems and their relation to Eaton triples and reduced triples.

Let us recall some rudiments of finite reflection groups [8]. Let V be a finite dimensional real inner product space. A reflection s_α on V is an element of $O(V)$, which sends some nonzero vector α to its negative and fixes pointwise the hyperplane H_α orthogonal to α, that is $s_\alpha \lambda := \lambda - 2(\lambda, \alpha)/(\alpha, \alpha)\, \alpha,\ \lambda \in V$. A finite group G generated by reflections is called a *finite reflection group*. A *root system* of G is a finite set of nonzero vectors in V, denoted by Φ, such that $\{s_\alpha : \alpha \in \Phi\}$ generates G, and satisfies

(R1) $\Phi \cap \mathbb{R}\alpha = \{\pm\alpha\}$ for all $\alpha \in \Phi$.

(R2) $s_\alpha \Phi = \Phi$ for all $\alpha \in \Phi$.

The elements of Φ are called *roots*. We do not require that the roots are of equal length. A root system Φ is *crystallographic* if it satisfies the additional requirement:

(R3) $2\frac{(\alpha,\beta)}{(\beta,\beta)} \in \mathbb{Z}$ for all $\alpha, \beta \in \Phi$,

and the group G is known as the *Weyl group* of Φ.

A (open) *chamber* C is a connected component of $V \setminus \cup_{\alpha \in \Phi} H_\alpha$. Given a *total order* $<$ in V [8, p.7], $\lambda \in V$ is said to be *positive* if $0 < \lambda$. Certainly, there is a total order in V: Choose an arbitrary ordered basis $\{\lambda_1, \ldots, \lambda_m\}$ of V and say $\mu > \nu$ if the first nonzero number of the sequence $(\lambda, \lambda_1), \ldots, (\lambda, \lambda_m)$ is positive, where $\lambda = \mu - \nu$. Now $\Phi^+ \subset \Phi$ is called a *positive system* if it consists of all those roots which are positive relative to a given total order. Of course, $\Phi = \Phi^+ \cup \Phi^-$, where $\Phi^- = -\Phi^+$. Now Φ^+ contains [8, p.8] a unique *simple system* Δ, that is, Δ is a basis for $V_1 := \operatorname{span} \Phi \subset V$, and each $\alpha \in \Phi$ is a linear combination of Δ with coefficients all of the same sign (all nonnegative or all nonpositive). The vectors in Δ are called *simple roots* and the corresponding reflections are called *simple reflections*. The finite reflection group G is generated by the simple reflections. Denote by $\Phi^+(C)$ the positive system obtained by the total order induced by an ordered basis $\{\lambda_1, \ldots, \lambda_m\} \subset C$ of V as described above. Indeed $\Phi^+(C) = \{\alpha \in \Phi : (\lambda, \alpha) > 0 \text{ for all } \lambda \in C\}$. The correspondence $C \mapsto \Phi^+(C)$ is a bijection of the set of all chambers onto the set of all positive systems. The group G acts simply transitively on the sets of positive systems, simple systems and chambers. The closed convex cone $F := \{\lambda \in V : (\lambda, \alpha) \geq 0, \text{ for all } \alpha \in \Delta\}$, that is, $F := C^-$ is the closure of the chamber C which defines Φ^+ and Δ, is called a (closed) *fundamental domain* for the action of G on V associated with Δ. Since G acts transitively on the chambers, given $x \in V$, the set $Gx \cap F$ is a singleton set and its element is denoted by x_0. It is known that (V, G, F) is an Eaton triple (see [13]). Let $V_0 := \{x \in V : gx = x \text{ for all } g \in G\}$ be the set of fixed points in V under the action of G. Let $\Delta = \{\alpha_1, \ldots, \alpha_n\}$, that is, $\dim V_1 = n$, where $V_1 = V_0^\perp$. If $\{\lambda_1, \ldots, \lambda_n\}$ denotes the basis of V_1 dual to the basis $\{\beta_i = 2\alpha_i/(\alpha_i, \alpha_i) : i = 1, \ldots, n\}$, that is, $(\lambda_i, \beta_j) = \delta_{ij}$, then $F = \{\sum_{i=1}^n c_i \lambda_i : c_i \geq 0\} \dotplus V_0$. Thus the interior $\operatorname{Int} F = C$ of F

is the nonempty set $\{\sum_{i=1}^n c_i\lambda_i : c_i > 0\} \dot{+} V_0$. The *dual cone* of F in V_1 is the cone

$$\mathrm{dual}_{V_1} F := \{x \in V_1 : (x, u) \geq 0, \text{ for all } u \in F\}$$

induced by F, and is equal to $\{\sum_{i=1}^n c_i\alpha_i : c_i \geq 0\}$. The finite reflection group G is said to be *essential* relative to V if $V_0 = \{0\}$. In this case, $F = \{\sum_{i=1}^n c_i\lambda_i : c_i \geq 0\}$. The space V is said to be *irreducible* if V contains no proper G-invariant subspace.

If we denote by $\langle x, y\rangle = 2(x, y)/(y, y)$, $y \neq 0$ (depends linearly on x), the matrix $(\langle \alpha_i, \alpha_j\rangle)$ is called the *Cartan matrix*. It is the change of basis matrix from $\{\lambda_j\}$ to $\{\alpha_i\}$: $\alpha_i = \sum_j \langle\alpha_i, \alpha_j\rangle\lambda_j$. Let $\mathfrak{L}$ be the collection of all subsets L of $\{1, \ldots, n\}$ for which there does not exist a nonempty subset $J \subset L$ satisfying $(\alpha_j, \alpha_k) = 0$ for all $j \in J$, $k \in L\backslash J$. So $L \in \mathfrak{L}$ if and only if Φ_L is irreducible in the sense of [1, p56]. If $L \in \mathfrak{L}$, then $d_{ij} > 0$ are positive rational numbers [7, p.72], for all $i, j \in L$ where (d_{ij}) is the inverse of the Cartan matrix.

The following example will yield the results of Ky Fan and Zietak via our results.

Example 1.3. Let $\mathbb{C}_{p\times q}$ denote the space of $p \times q$ complex matrices equipped with the inner product $(A, B) = \mathrm{Re}\ \mathrm{tr}\, AB^*$. For definiteness we assume $p \leq q$. Let G be the group of action of $U(p) \times U(q)$ on $\mathbb{C}_{p\times q}$ defined by $A \mapsto UAV^*$, where $U \in U(p)$, $V \in U(q)$. By the singular value decomposition, for any $A \in \mathbb{C}_{p\times q}$, there exist $U \in U(p)$, $V \in U(q)$ such that $A = U\Sigma V$. A well-known result of von Neumann [20] asserts that

$$\max\{\mathrm{Re}\ \mathrm{tr}\, AUBV : U \in U(p),\ V \in U(q)\} = \sum_{i=1}^p s_i(A)s_i(B),$$

where $s_1(A) \geq s_2(A) \geq \cdots \geq s_p(A)$ are the singular values of A. Thus $(\mathbb{C}_{p\times q}, G, F)$ is an Eaton triple with reduced triple (W, H, F), where W is the space $p \times q$ real "diagonal" matrices and F is the cone of $p \times q$ real "diagonal" matrices with diagonal entries in nonincreasing order. Here H is the group that permutes the diagonal entries of $\Sigma \in W$ and changes signs. If we identify W with $\mathbb{R}^p$, then the simple roots [19] are

$$\alpha_i = e_i - e_{i+1}, \qquad i = 1, \ldots, p-1, \qquad \alpha_p = e_p,$$

and the dual basis consists of

$$\lambda_i = \sum_{k=1}^i e_k, \qquad i = 1, \ldots, p-1. \qquad \lambda_p = \frac{1}{2}\sum_{k=1}^p e_k.$$

The function $f_{\lambda_m}(A) := (\lambda_m, F(A))$ yields the sum of the m largest singular values of the complex matrix A, where $F(A)$ is the unique Σ, where $A = U\Sigma V$, that is, Ky Fan's m-norm when $1 \leq m \leq p-1$ and f_{λ_p} is just half of Ky Fan's p-norm. Similarly one may get the real case.

Example 1.4. Let $V = H_n$ be the space of $n \times n$ Hermitian matrices with inner product

$$(X, Y) = \mathrm{tr}\, XY, \quad X, Y \in H_n.$$

Let $G = \mathrm{Ad}\,(U(n))$ be the group of adjoint action of the unitary group $U(n)$ on H_n, that is, $\mathrm{Ad}\,(U)(A) = UAU^*$ for all $U \in U(n)$. By the spectral theorem for Hermitian matrices, for each $H \in H_n$ there is a $U \in U(n)$ such that

$$UAU^* = \mathrm{diag}(a_1, a_2, \dots, \alpha_n),$$

where $a_1 \geq a_2 \geq \cdots \geq a_n$ are the eigenvalues of A. Let

$$F := \{\mathrm{diag}(a_1, a_2, \dots, a_n) : a_1 \geq a_2 \geq \cdots \geq a_n\}$$

be the set of diagonal matrices in H_n with diagonal elements arranged in nonincreasing order. This ensures that (A1) is satisfied. By a result of Ky Fan [16, Corollary 1.6, p.4-5]: for any two $n \times n$ Hermitian matrices A and B with eigenvalues $a_1 \geq \cdots \geq a_n$, $b \geq \cdots \geq b_n$, respectively,

$$\max\,\{\mathrm{tr}\,AUBU^* : U \in U(n)\} = \sum_{i=1}^{n} a_i b_i,$$

(A2) is satisfied. Let $W := \mathrm{span}\,F$, the space of diagonal matrices. Then (H_n, G, F) is an Eaton triple with reduced triple (W, S_n, F), where S_n is the symmetric group. S_n is not essential relative to W since $W_0 = \mathrm{span}\,\{I_n\}$ is the set of fixed points under the action of S_n. If we identify W with $\mathbb{R}^n$, then the simple roots are

$$\alpha_i = e_i - e_{i+1}, \qquad i = 1, \dots, n-1,$$

where $\{e_i\}$ is the standard basis of $\mathbb{R}^n$. The corresponding λ_i are

$$\lambda_i = \sum_{k=1}^{i} e_k, \qquad i = 1, \dots, n-1.$$

The function $f_{\lambda_m}(z) := (\lambda_m, F(z))$ yields the sum of the largest m eigenvalues of the Hermitian z. The statements remain true for the space of $n \times n$ Hermitian matrices with zero trace. The real case is similar.

2. Generalization of Ky Fan's Dominance Theorem. Let V be a real Euclidean space with the inner product $(\cdot, \cdot)$. The *dual norm* $\varphi^D : V \to \mathbb{R}$ of a norm $\varphi : V \to \mathbb{R}$ is defined as

$$\varphi^D(A) = \max_{\varphi(X) \leq 1} (A, X),$$

that is, the dual norm of A is simply the norm of the linear functional induced by A via the inner product. It is clear that

$$\varphi^D(A) = \max_{\varphi(X) = 1} (A, X), \qquad \text{and} \qquad \varphi = \varphi^{DD}.$$

It is easy to see that φ is G-invariant if and only if φ^D is G-invariant.

Let (V, G, F) be an Eaton triple. For any nonzero $\alpha \in F$, we define

$$f_\alpha(A) = (\alpha, F(A)).$$

Though f_α is G-invariant (since $f_\alpha(gA) = (\alpha, F(gA)) = (\alpha, F(A)) = f_\alpha(A)$ for all $g \in G$), it is not necessarily a norm (see Example 1.4). Very recently Tam and Hill [18] obtained the following result.

THEOREM 2.1. [18] *Let φ be a G-invariant norm on V where (V, G, F) is an Eaton triple. Let $C = \{F(A) : \varphi^D(A) \le 1, A \in V\} \subset F$, a compact set. Then*

$$\varphi(X) = \max\{f_\alpha(X) : \alpha \in C\}, \qquad \textit{for all} \quad X \in V.$$

We are ready to prove the following generalization of Ky Fan's result.

THEOREM 2.2. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). such that H is essential relative to W. Let $A, B \in V$. If $f_{\lambda_i}(A) \le f_{\lambda_i}(B)$ for all $i = 1, \dots n$, then $\varphi(A) \le \varphi(B)$ for all G-invariant norms φ.*

Proof. Let φ be a G-invariant norm. By Theorem 2.1 there exists a compact set C such that $\varphi(A) = \max_{\alpha \in C} f_\alpha(A)$. Since C is compact, the maximum must be attained at some $\alpha \in C$. Since $C \subset F$ and H is essential relative to W, $\alpha = \sum_{i=1}^n c_i \lambda_i$, where $c_i \ge 0$ for all $i = 1, \dots, n$. Thus

$$\varphi(A) = (\sum_{i=1}^n c_i \lambda_i, F(A)) = \sum_{i=1}^n c_i f_{\lambda_i}(A) \le \sum_{i=1}^n c_i f_{\lambda_i}(B) = (\alpha, F(B)) \le \varphi(B).$$

□

EXAMPLE 2.3. The result yields Ky Fan's Dominance Theorem via Example 1.3, where H is essential relative to W. It is not the case for Example 1.4 in which the symmetric group S_n is not essential relative to W. If $A = \operatorname{diag}(\frac{1}{n-1}, \dots, \frac{1}{n-1}, -1)$ and $B = \operatorname{diag}(1, 0, \dots, 0)$, then $f_{\lambda_i}(A) = \frac{i}{n-1} \le 1 = f_{\lambda_i}(B)$ for all $i = 1, \dots, n-1$ which mean the first largest $n-1$ eigenvalues of A are majorized by the first largest $n-1$ eigenvalues of B and indeed it is the case. However, $\|A\|_n = 2 > 1 = \|B\|_n$. Theorem 2.2 applies for the traceless case of Example 1.4 for if the eigenvalues of a traceless Hermitian A are majorized by the eigenvalues of a traceless Hermitian B, then the absolute values of the eigenvalues of A are weakly majorized by the absolute values of the eigenvalues of B [2, p.42]. Thus $\|A\|_n \le \|B\|_n$ and Theorem 2.2 says so.

REMARK 2.4. In Theorem 2.2, the condition that H is essential relative to W is necessary. If $W_0 \ne \{0\}$, then choose $A, B \in W_0$ with $A \ne B$. Since $A, B \in W_0$, $F(A) = A$, $F(B) = B$ and thus $f_{\lambda_i}(A) = f_{\lambda_i}(B) = 0$ for all $i = 1, \dots, n$. Consider the G-invariant norm $\varphi(A) = (A, A)^{1/2}$ and we have $\varphi(A) \ne \varphi(B)$. The condition $f_{\lambda_i}(A) \le f_{\lambda_i}(B)$ for all $i = 1, \dots n$, amounts to $F(A) \in \operatorname{conv} HF(B)$, or equivalently, $F(B) - F(A) \in \operatorname{dual}_W F$ [17].

Though f_{λ_i} is a convex G-invariant function, it may not be a norm. If f_{λ_i}, $i = 1, \dots, n$, are norms, then the converse of Theorem 2.2 is clearly true. The necessary and sufficient condition for f_λ being a norm is given in [18]. That $-1 \in H$ is a sufficient condition and holds for the case $\mathbb{C}_{p\times q}$ which yields Ky Fan's Dominance Theorem.

Due to the importance of the functions f_{λ_i}, $i=1,\dots,n$, we want to compute the dual of f_{λ_i}, $i=1,\dots,n$, if they are norms.

THEOREM 2.5. *Let (V,G,F) be an Eaton triple with reduced triple (W,H,F). Suppose W is irreducible. If $f_{\lambda_i}: V\to\mathbb{R}$ defined by $f_{\lambda_i}(x)=(F(x),\lambda_i)$, $x\in V$, $i=1,\dots,n$, are norms, then*

$$f^D_{\lambda_k}(x)=\max_{j=1,\dots,n}(F(x),\lambda_j)/(\lambda_k,\lambda_j), \qquad \textit{for all} \quad x\in V.$$

Proof. Since $f^D_{\lambda_k}$ is G-invariant, we may assume $x\in F$. Now $f^D_{\lambda_k}(x)=\max\{(x,y): f_{\lambda_k}(y)=1\}$ and by (A2) we may assume $y\in F$. By [18, Theorem 7], H is essential relative to W and thus $y=\sum_{j=1}^n c_j\lambda_j$ for some $c_j\ge 0$. So

$$f^D_{\lambda_k}(x)=\max\{\sum_{j=1}^n c_j(x,\lambda_j):\sum_{j=1}^n c_j(\lambda_k,\lambda_j)=1,\ c_j\ge 0, j=1,\dots,n\}.$$

Let (d_{kj}) be the inverse of the Cartan matrix $(\langle\alpha_i,\alpha_j\rangle)$. Since $\lambda_k=\sum_{j=1}^n d_{kj}\alpha_j$, where $d_{kj}>0$ for all k,j [7, p.72], $(\lambda_k,\lambda_j)=d_{kj}(\alpha_j,\alpha_j)/2>0$. The set $S=\{\sum_{j=1}^n c_j\lambda_j : \sum_{j=1}^n c_j(\lambda_k,\lambda_j)=1\}$ is an affine hyperplane. The intersection $S\cap F$ is evidently a convex set. Since λ_j, $j=1,\dots,n$, are the generators of F, the maximum is attained among $c_j\lambda_j$ such that $c_j(\lambda_k,\lambda_j)=1$, $j=1,\dots,n$. Explicitly each $\sum_{j=1}^n c_j\lambda_j\in S\cap F$ can be rewritten as $\sum_{j=1}^n c_j(\lambda_k,\lambda_j)[\lambda_j/(\lambda_k,\lambda_j)]$, a convex combination of $\lambda_j/(\lambda_k,\lambda_j)$. □

EXAMPLE 2.6. With respect to Example 1.3 and $p=q=n$, the symmetric matrix $((\lambda_k,\lambda_j))$ is:

$$\Lambda=\begin{pmatrix}
1 & 1 & 1 & 1 & \cdots & 1 & \frac12\\
1 & 2 & 2 & 2 & \cdots & 2 & 1\\
1 & 2 & 3 & 3 & \cdots & 3 & \frac32\\
1 & 2 & 3 & 4 & \cdots & 4 & 2\\
 & & & \cdots & \cdots & \cdots & \\
1 & 2 & 3 & 4 & \cdots & n-1 & \frac{n-1}{2}\\
\frac12 & 1 & \frac32 & 2 & \cdots & \frac{n-1}{2} & \frac{n}{4}
\end{pmatrix}$$

Direct observation leads to

$$\begin{aligned}&f^D_{\lambda_k}(A)\\ &=\begin{cases}\max\{f_{\lambda_1}(A),2f_{\lambda_n}(A)/k\}=\max\{s_1(A),(\sum_{i=1}^n s_i(A))/k\} & \text{if } 2\le k\le n-1\\ 2f_{\lambda_1}(A)=2s_1(A) & \text{if } k=n.\end{cases}\end{aligned}$$

So the dual of the Ky Fan k-norm is $\max\{s_1(A),(\sum_{i=1}^n s_i(A))/k\}$, $k=1,\dots,n$, [2, p.90] (the definition of dual norm there involves taking absolute value but it makes no difference in our case).

The irreducible root systems associated with the finite reflection groups are well known [8]. One can readily compute the dual norms of f_{λ_j}, $j=1,\dots,n$, of other types (the previous example is of type B_n). The following is a slight extension of the previous theorem and the proof is omitted.

THEOREM 2.7. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Suppose that $H \cong H_1 \times \cdots \times H_r$ is a decomposition of H into its irreducible components, where $H_i = H|_{W_i}$, with $H_0 = \{id\}$, $i = 1, \dots, r$, and $W = W_1 + \cdots + W_r$ is an orthogonal direct sum so that $F = F_1 + \cdots + F_r$. Let $\{\alpha_{j1}, \dots, \alpha_{jn_j}\}$ be the simple roots for H_j, $j = 1, \dots, r$. Suppose $f_{\lambda_{jk}} : V \to \mathbb{R}$, $1 \le j \le r$, $1 \le k \le n_j$ defined by $f_{\lambda_{jk}}(x) = (\lambda_{jk}, F(x))$ is a norm. Then*

$$f^D_{\lambda_{jk}}(x) = \max_{t=1,\dots,n_j} (F_j(x), \lambda_{jk})/(\lambda_{jk}, \lambda_{jt}), \qquad \text{for all} \quad x \in V.$$

We remark that the assumption that H is essential relative to W in Theorem 2.7 is necessary for each $f_{\lambda_{jk}}$ being a norm [18, Theorem 7].

3. Dual elements and facial structure. Let V be a real inner product space and let $O(V)$ denote the group of orthogonal linear operators on V. Motivated by [23] we introduce the notion of the φ-dual elements to $A \in V$, where $\varphi : V \to \mathbb{R}$ is a norm (not necessarily G-invariant): An element $K \in V$ with $\varphi(K) \le 1$ is said to be a φ-dual element to A if $\varphi^D(A) = (A, K)$. It is clear that if $A \ne 0$, then the φ-dual elements K of A are in the unit sphere $S_\varphi = \{A \in V : \varphi(A) = 1\}$. The set of φ-dual elements to A is a compact convex set (possibly empty) and is denoted by $\mathbb{D}_V(A : \varphi)$, that is,

$$\mathbb{D}_V(A : \varphi) := \{K \in V : \varphi^D(A) = (A, K),\ \varphi(K) \le 1\}.$$

If $A \ne 0$, then $\mathbb{D}_V(A : \varphi) := \{K \in V : \varphi^D(A) = (A, K),\ \varphi(K) = 1\}$. Clearly, $\mathbb{D}_V(\alpha A : \varphi) = \mathbb{D}_V(A : \varphi)$ for any $\alpha > 0$ and $\mathbb{D}_V(\alpha A : \varphi) = -\mathbb{D}_V(A : \varphi)$ for any $\alpha < 0$. We will denote by $B_\varphi = \{A \in V : \varphi(A) \le 1\}$ the unit ball in V associated with the norm φ. Clearly $\mathbb{D}_V(0 : \varphi) = B_\varphi$. A norm $\varphi : V \to \mathbb{R}$ is said to be *strictly convex* if X_1, X_2, $X := \frac{1}{2}(X_1 + X_2) \in S_\varphi$ implies $X_1 = X_2 = X$. For example, the Schatten p-norms, $1 < p < \infty$, are strictly convex. We remark that $\mathbb{D}_V(A : \varphi) \subset S_\varphi$ is a singleton set if $A \ne 0$ and φ is strictly convex.

Given a norm φ on $V \ne \{0\}$, a convex set $\mathbb{F} \subset B_\varphi$ is called a *face* of B_φ if $B, C \in \mathbb{F}$ whenever $\alpha B + (1-\alpha)C \in \mathbb{F}$ for some $0 < \alpha < 1$, and $B, C \in B_\varphi$ [23]. In other words, every closed line segment in B_φ with a relative interior point in $\mathbb{F}$ has both endpoints in $\mathbb{F}$ [14, p.162]. The empty set and B_φ itself are faces of B_φ, known as the trivial faces. Extreme points of B_φ are simply zero-dimensional faces of B_φ. Since B_φ is compact, so are its faces [14, Corollary 18.11]. A nontrivial face $\mathbb{F}$ is called an *maximal face* of B_φ if there is no other faces of B_φ containing $\mathbb{F}$ properly.

EXAMPLE 3.1.

1. Consider $V = \mathbb{R}^3$ equipped with the max norm φ. Then B_φ is simply the unit cube. The nontrivial faces of B_φ are the corners, the edges (notice that the faces of a face of B_φ are faces of B_φ) and the walls. The walls are the maximal faces.
2. Consider $V = \mathbb{R}^3$ equipped with the 2-norm φ. Then B_φ is simply the usual unit ball. The points on the unit sphere are the extreme points and there are

no other nontrivial faces. This is true for any strictly convex norm φ on a vector space V and [15, Theorem 2] is merely a particular case.

EXAMPLE 3.2. Let φ be a norm (not necessarily G-invariant) on V. Each $\mathbb{D}_V(A:\varphi)$, $A \neq 0$, is a face of B_φ: if $\alpha K_1 + (1-\alpha)K_2 \in \mathbb{D}_V(A:\varphi) \subset S_\varphi$, $0<\alpha<1$, and $K_1, K_2 \in S_\varphi$,

$$\varphi^D(A) = (A, \alpha K_1 + (1-\alpha)K_2) = \alpha(A, K_1) + (1-\alpha)(A, K_2) \leq \varphi^D(A),$$

since $(A, K_i) \leq \varphi^D(A)\varphi(K_i) \leq \varphi^D(A)$. Thus $(A, K_i) = \varphi^D(A)$, for $i = 1, 2$.

The following is an extension of [23, Theorem 4.1] and the idea of the proof is from [15, 23]. Also see [15, 22]. It gives some relationship between the facial structure of B_φ and the sets of dual elements associated with a norm φ on V.

THEOREM 3.3. *Let $V \neq \{0\}$ be a real inner product space and let φ be a norm on V. Let $\mathbb{F} \subset S_\varphi$ be a nontrivial face of the unit ball B_φ. Then $\cap_{H\in\mathbb{F}}\mathbb{D}_V(H:\varphi^D) \neq \phi$. Hence*

1. *there exists $A \in V$ such that $\mathbb{F} \subset \mathbb{D}_V(A:\varphi)$.*
2. *each maximal face of B_φ is of the form $\mathbb{D}_V(A:\varphi)$ for some $A \neq 0$.*

Proof. Each $\mathbb{D}_V(H:\varphi^D)$ is closed in the compact set B_φ so it is sufficent to show that the family $\{\mathbb{D}_V(H:\varphi^D) : H \in \mathbb{F}\}$ has the finite intersection property [12, p.170]. Let $H_1, \ldots, H_k \in \mathbb{F}$. Since $\mathbb{F} \subset S_\varphi$, $\varphi(H_i) = 1$ for all $i = 1, \ldots, k$. By the convexity of $\mathbb{F}$, $\hat{H} := \frac{1}{k}(H_1 + \cdots + H_k) \in \mathbb{F}$ so that $\varphi(\hat{H}) = 1$. Let $K \in \mathbb{D}_V(\hat{H}:\varphi^D)$. So

$$1 = \varphi(\hat{H}) = (\hat{H}, K) = \frac{1}{k}\sum_{i=1}^{k}(H_i, K) \leq \frac{1}{k}\sum_{i=1}^{k}\varphi(H_j)\varphi^D(K) \leq 1,$$

since $\varphi^D(K) = 1$. Thus $(H_i, K) = 1$ for all $i = 1, \ldots, k$. Hence $K \in \cap_{i=1}^k\mathbb{D}_V(H_i:\varphi^D)$.

1. Any $A \in \cap_{H\in\mathbb{F}}\mathbb{D}_V(H:\varphi^D)$ satisfies $\mathbb{F} \subset \mathbb{D}_V(A:\varphi)$ since for each $H \in \mathbb{F}$, $\varphi^D(A) = 1$ and $1 = \varphi(H) = \varphi^{DD}(H) = (H, A)$.
2. It follows from Example 3.2 and the definition of maximal face.

□

In view of the above theorem, we remark that if $\mathbb{F} = B_\varphi$, then simply set $A = 0$ and the case is trivial.

4. Characterization of the dual elements. The following is a generalization of [23, Theorem 3.1] (see Example 1.2).

PROPOSITION 4.1. *Let V be a real inner product space and let $G \subset O(V)$. Suppose $A \in V$ and $A = gB$, for some $g \in G$. Let φ be a G-invariant norm on V. Then $\mathbb{D}_V(A:\varphi) = g\mathbb{D}_V(B:\varphi)$.*

Proof. Since φ^D is G-invariant as well as φ and g is orthogonal,

$$\begin{aligned} K \in \mathbb{D}_V(A:\varphi) &\Leftrightarrow (A, K) = \varphi^D(A),\ \varphi(K) = 1 \\ &\Leftrightarrow (B, g^{-1}K) = \varphi^D(B),\ \varphi(g^{-1}K) = 1 \\ &\Leftrightarrow g^{-1}K \in \mathbb{D}_V(B:\varphi). \end{aligned}$$

□

The notion of dual elements of $A \in V$ is related to the *subdifferential* of φ at A.

$$\partial\varphi(A) = \{K \in V : \varphi(B) \geq \varphi(A) + (K, B - A) \text{ for all } B \in V\}.$$

It is easy to see that $\partial\varphi(A)$ is a closed convex set and $0 \in \partial\varphi(A)$ implies $A = 0$. More generally the subdifferential can be defined for convex functions $\varphi : V \to \mathbb{R}$:

$$\partial\varphi(x) = \{\xi : \varphi(x') \geq \varphi(x) + (\xi, x' - x) \text{ for all } x' \in V\},$$

and the elements of the subdifferential of φ at $x \in V$ are called *subgradients* of φ at x. Geometrically $\varphi(x') \geq \varphi(x) + (\xi, x' - x)$ for all $x' \in V$ means that the affine function $h(x') := \varphi(x) + (\xi, x' - x)$ is a nontrivial supporting hyperplane to the convex set $\operatorname{epi}\varphi$ (the epigraph of φ) at the point $(x, \varphi(x))$ [14, p.215].

Proposition 4.2. *Let V be a real inner product space and let $A \in V$. Suppose φ is a norm. Then $K \in \partial\varphi(A)$ if and only if $\varphi(A) = (A, K)$ and $\varphi^D(K) \leq 1$. Thus $\partial\varphi^D(A) = \mathbb{D}_V(A, \varphi)$.*

Proof. Suppose $\varphi(A) = (A, K)$ and $\varphi^D(K) \leq 1$. Then $\varphi(A) + (B - A, K) = (B, K) \leq \varphi^D(K)\varphi(B) \leq \varphi(B)$ for all B. Suppose $K \in \partial\varphi(A)$, that is, $\varphi(B) \geq \varphi(A) + (B - A, K)$ for all $B \in V$. Notice that setting $B = 0$ and $B = 2A$ yield $\varphi(A) = (A, K)$. Thus $\varphi(B) \geq (B, K)$ for all B, which means $\varphi^D(K) \leq 1$ by letting B run over the unit ball B_φ. Then notice $\partial\varphi^D(A) = \{K : (A, K) = \varphi^D(A), \varphi(K) \leq 1\} = \mathbb{D}_V(A, \varphi)$. □

Proposition 4.3.

1. *Let V be a real inner product space and let $x \in V$ and $g \in O(V)$. Let $\varphi : V \to \mathbb{R}$ be a convex function such that $\varphi(gx) = \varphi(x)$ for all $x \in V$. Then $\partial\varphi(gx) = g\partial\varphi(x)$.*
2. *Let (V, G, F) be an Eaton triple. Let $\varphi : V \to \mathbb{R}$ be a G-invariant convex function. Then $\partial\varphi(x) = g\partial\varphi(F(x))$, where $gF(x) = x$, $x \in V$.*

Proof. The second part follows immediately from the first part. Now

$$\begin{aligned}
\xi \in \partial\varphi(x) &\Leftrightarrow \varphi(x') \geq \varphi(x) + (\xi, x' - x) \text{ for all } x' \in V \\
&\Leftrightarrow \varphi(gx') \geq \varphi(gx) + (g\xi, gx' - gx) \text{ for all } x' \in V \\
&\Leftrightarrow \varphi(y) \geq \varphi(gx) + (g\xi, y - gx) \text{ for all } y \in V \\
&\Leftrightarrow g\xi \in \partial\varphi(gx).
\end{aligned}$$

□

Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). It is clear that the restriction $\hat\varphi : W \to \mathbb{R}$ of a G-invariant norm $\varphi : V \to \mathbb{R}$ on W is also a norm and H-invariant. On the other hand, if $\hat\varphi : W \to \mathbb{R}$ is a H-invariant norm, then one can define $\varphi : V \to \mathbb{R}$ by $\varphi(x) = \hat\varphi(F(x))$ which is a G-invariant norm. This is a generalization [18] of von Neumann's well-known result on the one-to-one correspondence between unitarily invariant norms and symmetric gauge functions [20] as well as the result of Davis [4]. Indeed it is true for G-invariant convex functions [18]. Given $\gamma \in F$, how do we obtain $\mathbb{D}_W(\gamma : \hat\varphi)$ from $\mathbb{D}_V(\gamma : \varphi)$? We intend to give the answer in the following proposition.

Lemma 4.4. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi : V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat{\varphi} : W \to \mathbb{R}$ the restriction of φ on W. Then $(\widehat{\varphi^D}) = (\hat{\varphi})^D$, that is, the dual of the restriction is the restriction of the dual.*

Proof. Notice that φ^D is also a G-invariant norm. For any $x \in W$,

$$(\widehat{\varphi^D})\,(x) = \varphi^D(x) = \max_{\varphi(u)\le 1,\ u\in V}(u,x) \ge \max_{\hat{\varphi}(z)\le 1,\ z\in W}(z,x) = (\hat{\varphi})^D(x).$$

In other words, $(\widehat{\varphi^D}) \ge (\hat{\varphi})^D$. Moreover we may assume that $x \in F$. Let $u \in V$ such that $\varphi(u) \le 1$ and $(\widehat{\varphi^D})\,(x) = (u, x)$. By (A2), $(u, x) \le (gu, x)$, where $gu \in F$, $g \in G$. Since $\varphi(u) = \varphi(gu)$, we can replace u by $w := gu \in F \subset W$, that is, $(\widehat{\varphi^D})\,(x) = (w, x)$, where $w \in F$ with $\varphi(w) \le 1$. So $(\widehat{\varphi^D})\,(x) \le (\hat{\varphi})^D(x)$ and thus $(\widehat{\varphi^D}) = (\hat{\varphi})^D$. □

Proposition 4.5. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi : V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat{\varphi} : W \to \mathbb{R}$ the restriction of φ on W.*

1. *Then for each $\gamma \in W$, $\mathbb{D}_V(\gamma : \varphi) \cap W = \mathbb{D}_W(\gamma : \hat{\varphi})$.*
2. *If $X \in \mathbb{D}_V(A : \varphi)$, then $F(X) \in \mathbb{D}_W(F(A) : \hat{\varphi})$.*

Proof.

1.

$$\begin{aligned}
&X \in \mathbb{D}_W(\gamma : \hat{\varphi})\\
\Leftrightarrow\ & X \in W,\ \hat{\varphi}(X) \le 1,\ (\gamma, X) = \hat{\varphi}^D(\gamma)\\
\Leftrightarrow\ & X \in W,\ \varphi(X) \le 1,\ (\gamma, X) = (\widehat{\varphi^D})\,(X) \qquad \text{since} \quad (\widehat{\varphi^D}) = (\hat{\varphi})^D\\
\Leftrightarrow\ & X \in W,\ \varphi(X) \le 1,\ (\gamma, X) = \varphi^D(X)\\
\Leftrightarrow\ & X \in W,\ X \in \mathbb{D}_V(\gamma : \varphi).
\end{aligned}$$

2. If $X \in \mathbb{D}_V(A : \varphi)$, then $\varphi(X) = 1$ and $\varphi^D(A) = (X, A) \le (F(X), F(A)) \le \hat{\varphi}^D(F(A)) = \widehat{\varphi^D}\,(F(A)) = \varphi^D(F(A)) = \varphi^D(A)$. Thus $F(X) \in \mathbb{D}_V(F(A) : \hat{\varphi})$ since $\hat{\varphi}(F(X)) = \varphi(X) = 1$.

□

Proposition 4.5 is an extension of [22, Theorem 4.1]. Proposition 4.5 (1) enables us to compute $\mathbb{D}_W(\gamma : \hat{\varphi})$ if we know $\mathbb{D}_V(\gamma : \varphi)$. On the other hand, we want to know how to recover $\mathbb{D}_V(\gamma : \varphi)$ if we know $\mathbb{D}_W(\gamma : \hat{\varphi})$. We now proceed to tackle the problem. To this end we need to recall some basics about the Clarke generalized gradient. Let Y be a subset of V. A function $f : Y \to \mathbb{R}$ is said to be *Lipschitz* [3, p.25] on Y with Lipschitz constant K if for some $K \ge 0$,

$$|f(y) - f(y')| \le K\sqrt{(y - y', y - y')}, \qquad y, y' \in Y.$$

We say that f is Lipschitz near x if for some $\epsilon > 0$, f satisfies the Lipschitz condition on the set $x + \epsilon B$, where B is the open unit ball with respect to the inner product. Let f be Lipschitz near a given $x \in V$ and let $0 \neq v \in V$. The *Clarke directional derivative* [3, p.25] of f at x in the direction v is defined as

$$f^o(x; v) = \limsup_{y \to x \ t \downarrow 0} \frac{f(y + tv) - f(y)}{t},$$

The *Clarke generalized gradient* of f at x, denoted by $\partial_C f(x)$, is defined as

$$\partial_C f(x) := \{\xi \in V : f^o(x; v) \geq (\xi, v) \text{ for all } v \in V\}.$$

THEOREM 4.6. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi : V \to \mathbb{R}$ be a G-invariant convex Lipschitz function and denote by $\hat{\varphi} : W \to \mathbb{R}$ the restriction of φ on W. Then*

$$\partial\varphi(A) = conv\, G_A \partial\hat{\varphi}(F(A)),$$

where $G_A := \{g \in G : A = gF(A)\}$.

Proof. We first notice that $\hat{\varphi} : W \to \mathbb{R}$ is also convex and Lipschitz and $\varphi = \hat{\varphi} \circ F$. Since φ is convex and Lipschitz, the subdifferentials of φ and $\hat{\varphi}$ coincide with the Clarke's generalized gradients of φ and $\hat{\varphi}$ [3, Proposition 2.2.7] respectively, and thus we have

$$\begin{aligned}
\partial\varphi(A) &= g\, \partial\varphi(F(A)) \qquad \text{by Proposition 4.3, where } g \in G_A \\
&= g\, \partial_C\varphi(F(A)) \\
&= g\, \partial_C(\hat{\varphi} \circ F)(F(A)) \qquad \text{since } \hat{\varphi} \circ F = \varphi \\
&= g \operatorname{conv} G_{F(A)} \partial_C \hat{\varphi}(F(A)) \qquad \text{by [19, Lemma 3.11]} \\
&= \operatorname{conv} G_A \partial\hat{\varphi}(F(A)) \qquad \text{since } g\, G_{F(A)} = G_A.
\end{aligned}$$

□

THEOREM 4.7. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi : V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat{\varphi} : W \to \mathbb{R}$ the restriction of φ on W. Then $\partial\varphi(A) = conv\, G_A \partial\hat{\varphi}(F(A))$, where $G_A := \{g \in G : A = gF(A)\}$. Hence $\mathbb{D}_V(A : \varphi) = conv\, G_A \mathbb{D}_W(F(A) : \hat{\varphi})$. In particular when $\gamma \in F$, we have $\mathbb{D}_V(\gamma : \varphi) = conv\, G_\gamma \mathbb{D}_W(\gamma : \hat{\varphi})$ and $G_\gamma = \{g \in G : g\gamma = \gamma\}$.*

Proof. As a norm φ is clearly convex. So it suffices to show that φ is Lipschitz. Now for any $y, y' \in V$, we have

$$\begin{aligned}
|\varphi(y) - \varphi(y')| &\leq \varphi(y - y') \\
&= \max_{\alpha \in C} (\alpha, F(y - y')) \qquad \text{by Theorem 2.1} \\
&\leq \max_{\alpha \in C} (\alpha, \alpha)^{1/2} (F(y - y'), F(y - y'))^{1/2} \\
&\leq K(y - y', y - y')^{1/2},
\end{aligned}$$

where $K := \max_{\alpha\in C}(\alpha,\alpha)^{1/2}$ is the Lipschitz constant for φ. Hence

$$
\begin{aligned}
\mathbb{D}_V(A:\varphi) &= \partial\varphi^D(A) && \text{by Proposition 4.2}\\
&= \operatorname{conv} G_A\partial\,(\overset{\wedge}{\varphi^D})\,(F(A)) && \text{by Theorem 4.6}\\
&= \operatorname{conv} G_A\mathbb{D}_W(F(A):((\overset{\wedge}{\varphi^D}))^D) && \text{by Proposition 4.2}\\
&= \operatorname{conv} G_A\mathbb{D}_W(F(A):(\hat\varphi)^{DD}) && \text{by } (\overset{\wedge}{\varphi^D})=(\hat\varphi)^D\\
&= \operatorname{conv} G_A\mathbb{D}_W(F(A):\hat\varphi) && \text{since } (\hat\varphi)^{DD}=\hat\varphi.
\end{aligned}
$$

□

We remark that the last statement of Theorem 4.7 generalizes [23, Theorem 3.2] if one considers the fact that $G_\gamma\mathbb{D}_W(\gamma,\hat\varphi)$ is a convex set for Example 1.3 [23]. See [21, Theorem 2]. It is also the case for Example 1.4 [10, Theorem 3.12]. See Remark 3.12 of [18].

Example 4.8. When $\gamma=0$, $G_\gamma=G$, and $\mathbb{D}_V(\gamma:\varphi)=\operatorname{conv} G_\gamma\mathbb{D}_W(\gamma:\hat\varphi)$ becomes $B_\varphi=\operatorname{conv} GB_{\hat\varphi}$, where B_φ is the unit ball in V with respect to the norm φ and $B_{\hat\varphi}$ is the unit ball in W with respect to $\hat\varphi$. Indeed $B_\varphi=GB_{\hat\varphi}$

Example 4.9.

1. With respect to Example 1.4, let $\varphi(A)=\sum_{i=1}^n|\lambda_i(A)|$, that is the sum of singular values of the Hermitian matrix A (Ky Fan's n-norm). Evidently it is a unitary similarity norm. It follows that [2] $\varphi^D(A)=\max_{i=1,\ldots,n}|\lambda_i(A)|$, the operator norm. Let $\gamma=\operatorname{diag}(\gamma_1,\ldots,\gamma_n)$, where $\gamma_1>\cdots>\gamma_n$. Then G_γ is the group of similarity via the group $U(1)\oplus\cdots\oplus U(1)$ which denotes the group of diagonal unitary matrices. Now $K\in\mathbb{D}_W(\gamma:\hat\varphi)$ means that K is a real diagonal matrix, $\sum_{i=1}^n|k_i|=1$ (since $\gamma\neq0$ and thus $\mathbb{D}_W(\gamma:\hat\varphi)\subset S_\varphi$) with $\max_{i=1,\ldots,n}|\gamma_i|=\sum_{i=1}^n\gamma_ik_i$. So

Case 1 $|\gamma_1|\geq|\gamma_n|$ (so $\gamma_1\geq0$): $k_1=1$, and $k_2=\cdots=k_n=0$. So $G_\gamma\mathbb{D}_W(\gamma:\hat\varphi)=\mathbb{D}_W(\gamma:\hat\varphi)=\{\operatorname{diag}(1,\ldots,0)\}$. On the other hand, $X\in\mathbb{D}_V(\gamma:\varphi)$ means $\gamma_1=\sum_{i=1}^n\gamma_ix_{ii}$ with $\sum_{i=1}^n s_i(X)=1$. Notice that the diagonal element of X, $(|x_{11}|,\cdots,|x_{nn}|)$ is weakly majorized by the vector of singular values of X, $(s_1(X),\ldots,s_n(X))$. So $x_{11}=1$ and $x_{22}=\cdots=x_{nn}=0$ and thus $X=\operatorname{diag}(1,0,\ldots,0)$. So $\mathbb{D}_V(\gamma:\varphi)=\mathbb{D}_W(\gamma:\hat\varphi)$.

Case 2 $|\gamma_1|<|\gamma_n|$ (so $\gamma_n\leq0$): $k_n=-1$, and $k_1=\cdots=k_{n-1}=0$. So $G_\gamma\mathbb{D}_W(\gamma:\hat\varphi)=\mathbb{D}_W(\gamma:\hat\varphi)=\{\operatorname{diag}(0,\ldots,0,-1)\}$. Similarly $\mathbb{D}_V(\gamma:\varphi)=\mathbb{D}_W(\gamma:\hat\varphi)$.

2. Continuing with the above example, let $\gamma=I_n$ instead. Clearly G_γ is the group of similarity via the whole unitary group. Now $K\in\mathbb{D}_W(\gamma:\hat\varphi)$ means that K is a real diagonal matrix, $\sum_{i=1}^n|k_i|=1$ and $1=\sum_{i=1}^n k_i$. Thus $k_i\geq0$ for all $i=1,\ldots,n$. Hence

$$\mathbb{D}_W(\gamma:\hat\varphi)=\{\operatorname{diag}(k_1,\ldots,k_n):k_i\geq0,\ i=1,\ldots,n,\ \sum_{i=1}^n k_i=1\},$$

so that

$$G_\gamma \mathbb{D}_W(\gamma : \hat{\varphi}) = \{U \operatorname{diag}(k_1, \ldots, k_n) U^{-1} : U \in U(n),\ k_i \geq 0,$$
$$i = 1, \ldots, n,\ \sum_{i=1}^n k_i = 1\}.$$

On the other hand, $X \in \mathbb{D}_V(\gamma : \varphi)$ means that X is a Hermitian matrix, $\sum_{i=1}^n x_{ii} = 1$ with $\sum_{i=1}^n s_i(X) = 1$. Notice that $(|x_{11}|, \cdots, |x_{nn}|)$ is weakly majorized by $(s_1(X), \ldots, s_n(X))$. So $0 \leq x_{ii}$ for all $i = 1, \ldots, n$ and $\sum_{i=1}^n x_{ii} = 1$. Thus the eigenvalues of X must be nonnegative. So X must be of the form $U \operatorname{diag}(k_1, \ldots, k_n) U^{-1}$ for some $U \in U(n)$, where $k_i \geq 0$, for all $i = 1, \ldots, n$, with $\sum_{i=1}^n k_i = 1$. Hence $\mathbb{D}_V(\gamma : \varphi) = G_\gamma \mathbb{D}_W(\gamma : \hat{\varphi})$. Indeed by [17, Theorem 11],

$$\{U \operatorname{diag}(k_1, \ldots, k_n) U^{-1} : U \in U(n),\ k_i \geq 0,\ i = 1, \ldots, n,\ \sum_{i=1}^n k_i = 1\}$$
$$= \operatorname{conv}\{U \operatorname{diag}(1, 0, \ldots, 0) U^{-1} : U \in U(n)\}.$$

We remark that we have similar result when φ is the operator norm.

We have the following extension of the first part of Example 4.9. The result takes care of the regular points $\gamma \in W$, that is, the points in $\operatorname{Int}_W F$.

THEOREM 4.10. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi : V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat{\varphi} : W \to \mathbb{R}$ the restriction of φ on W. If $F(A) \in \operatorname{Int}_W F$, then $\mathbb{D}_V(A : \varphi) = g\mathbb{D}_W(F(A) : \hat{\varphi})$, where $A = gF(A)$. In particular, if $\gamma \in \operatorname{Int}_W F$, then $\mathbb{D}_V(\gamma : \varphi) = \mathbb{D}_W(\gamma : \hat{\varphi}) \subset F$.*

Proof. In view of proposition 4.1, it suffices to show the last statement. Suppose $\gamma \in \operatorname{Int}_W F$, the interior of F in W. Let $X \in \mathbb{D}_V(\gamma : \varphi)$. We may assume $\varphi^D(\gamma) = 1$ since $\mathbb{D}_V(\gamma : \varphi)$ is invariant under positive scaling of γ and the case $\gamma = 0$ is trivial (Example 4.8). So $1 = \varphi^D(\gamma) = (\gamma, X)$ and $1 = \varphi(X) = \max_{\alpha \in C}(\alpha, F(X))$, where $C = \{\beta \in F : \varphi^D(\beta) \leq 1\}$. Let $p(X)$ be the projection of X under the orthogonal projection $p : V \to W$ with respect to the inner product. Since $\gamma \in C \subset F \subset W$, we have $(\gamma, X) = (\gamma, p(X)) = (\gamma, F(X))$. By [13, Theorem 3.2], $p(X)$ is in $\operatorname{conv} HF(X)$, where H is a finite reflection group, that is, $p(X) = \sum_{i=1}^k c_i h_i F(X)$ where $h_i \in H$ and $c_i > 0$ with $\sum_{i=1}^k c_i = 1$, and $h_i F(X)$, $i = 1, \ldots, k$, are distinct. Thus $\sum_{i=1}^k c_i(\gamma, h_i F(X)) = (\gamma, F(X))$. By (A2) $(\gamma, h_i F(X)) \leq (\gamma, F(X))$ and thus $(\gamma, h_i F(X)) = (\gamma, F(X))$ for all $i = 1, \ldots, k$. Suppose $h_i F(X) \neq F(X)$. By [8, p.22], each nonzero $F(X) - h_i F(X)$ is a nonnegative combination of the simple roots, that is, for each $i = 1, \ldots, k$, there exist nonnegative numbers $d_j \geq 0$, $j = 1, \ldots, n$, not all zero, such that $F(X) - h_i F(X) = \sum_{j=1}^n d_j \alpha_j$. Now $\gamma \in \operatorname{Int}_W F$, we have $(\gamma, \alpha_j) > 0$ for all $j = 1, \ldots, n$ so that $(\gamma, F(X) - h_i F(X)) > 0$, a contradiction. So $p(X) = F(X)$. By considering $(X, X) = (F(X), F(X))$, we have $X = F(X)$. Thus $\mathbb{D}_V(\gamma : \varphi) \subset F$. By Proposition 4.5, $\mathbb{D}_V(\gamma : \varphi) = \mathbb{D}_W(\gamma : \hat{\varphi})$. ☐

In view of Example 4.9 and Theorem 4.10, one may guess that if $\gamma \in \operatorname{Int}_W F$, then $\mathbb{D}_V(\gamma : \varphi) = \mathbb{D}_W(\gamma : \hat{\varphi}) \subset F$ is a singleton set. The following example shows that it is not true.

Example 4.11. With respect to Example 1.3, let $\varphi(A) = \sum_{i=1}^{k} s_i(A)$, the Ky Fan k-norm on $\mathbb{C}_{n\times n}$. Notice that $\varphi^D(A) = \max\{s_1(A), (\sum_{i=1}^{n} s_i(A))/k\}$ [2, p.90]. Choose appropriate n, k, and $\gamma \in \mathrm{Int}_W F$ ($\gamma_1 > \gamma_2 > \cdots > \gamma_n > 0$) such that $\gamma_1 = (\sum_{i=1}^{n} \gamma_i)/k$. Now $A \in \mathbb{D}_W(\gamma : \hat\varphi)$ means that $A = \mathrm{diag}(a_1, \ldots, a_n)$ with $a_1 \geq \cdots \geq a_n \geq 0$ by Theorem 4.10, $\sum_{i=1}^{k} a_i = 1$, and

$$\gamma_1 = \hat\varphi^D(\gamma) = \sum_{i=1}^{n} \gamma_i a_i.$$

One may have more than one $A \in F$ satisfying the above condition. For example $n = 3$, $\gamma_1 = 1$, $\gamma_2 = 2/3$ and $\gamma_3 = 1/3$ and $k = 2$. Then both $A = \mathrm{diag}(1, 0, 0)$ and $A' = \mathrm{diag}(1/2, 1/2, 1/2)$ satisfy the conditions. Thus $\mathbb{D}_W(\gamma : \hat\varphi)$ is not a singleton set in F. One may view it as a way to show that the corresponding Ky Fan's k-norm is not strictly convex.

5. On the unit balls. Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi : V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat\varphi : W \to \mathbb{R}$ the restriction of φ on W. Then we evidently have $B_{\hat\varphi} = B_\varphi \cap W$, that is, the unit ball of $\hat\varphi$ in W is the intersection of W and the unit ball of φ in V. Thus one can determine $B_{\hat\varphi}$ from B_φ easily. On the other hand one can easily show that $B_\varphi = GB_{\hat\varphi}$ (see Example 4.8). We summarize them as

Proposition 5.1. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi : V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat\varphi : W \to \mathbb{R}$ the restriction of φ on W. Then*

1. $B_{\hat\varphi} = B_\varphi \cap W$.
2. $B_\varphi = GB_{\hat\varphi}$.

Recall that a norm $\varphi : V \to \mathbb{R}$ is said to be *strictly convex* if X_1, X_2, $X := \frac{1}{2}(X_1 + X_2) \in S_\varphi$ implies $X_1 = X_2 = X$. The following is an extension of [22, Theorem 3.1].

Theorem 5.2. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi : V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat\varphi : W \to \mathbb{R}$ the restriction of φ on W. Then φ is strictly convex if and only if $\hat\varphi$ is strictly convex. In this event, $\mathbb{D}_V(\gamma : \varphi) = \mathbb{D}_W(\gamma : \hat\varphi) \subset F$ is a singleton set for any $\gamma \in F$.*

Proof. Suppose that φ is strictly convex. If $X, X_1, X_2 \in S_{\hat\varphi}$ with $X = \frac{1}{2}(X_1 + X_2)$, then they all belong to S_φ and thus $X = X_1 = X_2$.

On the other hand, suppose that $\hat\varphi$ is strictly convex. If $X_1, X_2, X \in S_\varphi$, where $X := \frac{1}{2}(X_1 + X_2)$, then

$$\begin{aligned}
1 = \hat\varphi(F(X)) &= \hat\varphi(F(\frac{1}{2}[X_1 + X_2])) \\
&= \hat\varphi(\frac{1}{2}[F(X_1 + X_2)]) \\
&= \max_{\alpha \in \hat C}(\alpha, \frac{1}{2}F(X_1 + X_2)) \qquad \text{by Theorem 2.1} \\
&\leq \frac{1}{2}\max_{\alpha \in \hat C}(\alpha, F(X_1) + F(X_2)) \qquad \text{by [17, Theorem 10]}
\end{aligned}$$

$$\leq \frac{1}{2}\max_{\alpha\in\hat{C}}(\alpha, F(X_1)) + \frac{1}{2}\max_{\alpha\in\hat{C}}(\alpha, F(X_2))$$
$$= \frac{1}{2}\hat{\varphi}(F(X_1)) + \frac{1}{2}\hat{\varphi}(F(X_2)) = 1,$$

where $\hat{C} = \{F(A) : (\hat{\varphi})^D(A) = 1, A \in W\}$. So $\frac{1}{2}[F(X_1) + F(X_2)] \in S_{\hat{\varphi}}$. By the strict convexity of $\hat{\varphi}$, $\frac{1}{2}[F(X_1) + F(X_2)] = F(X_1) = F(X_2)$. We now proceed to show that $F(X) = F(X_1)$ by contradiction. Suppose $F(X) \neq F(X_1)$. We may assume that $F(X), F(X_1) = F(X_2) \in S_{\hat{\varphi}}$. By the strict convexity of $\hat{\varphi}$, the element $Y := \frac{1}{2}(F(X) + F(X_1)) \notin S_{\hat{\varphi}}$. For all $i = 1, \ldots, n$, by [17, Theorem 10],

$$(\lambda_i, F(X)) = \frac{1}{2}(\lambda_i, F(X_1 + X_2)) \leq \frac{1}{2}(\lambda_i, F(X_1) + F(X_2)) = (\lambda_i, Y),$$

and

$$\begin{aligned}(\lambda_i, Y) &= \frac{1}{2}(\lambda_i, F(X) + F(X_1))\\ &= \frac{1}{2}(\lambda_i, F(X)) + \frac{1}{2}(\lambda_i, F(X_1))\\ &= \frac{1}{4}(\lambda_i, F(X_1 + X_2)) + \frac{1}{2}(\lambda_i, F(X_1))\\ &\leq \frac{1}{4}(\lambda_i, F(X_1) + F(X_2)) + \frac{1}{2}(\lambda_i, F(X_1))\\ &= (\lambda_i, F(X_1)),\end{aligned}$$

since $F(X_1) = F(X_2)$. So by Theorem 2.2, $\hat{\varphi}(F(X)) \leq \hat{\varphi}(Y) \leq \hat{\varphi}(F(X_1))$. Thus $Y \in S_{\hat{\varphi}}$, a contradiction.

Hence $F(X) = F(X_1) = F(X_2) \in F$. Now let $\|X\| = (X, X)^{1/2}$. Evidently it is a G-invariant strictly convex norm since it is induced by the inner product. Since X, X_1, X_2 are of the same length with respect to $\|\cdot\|$, $X = X_1 = X_2$ and thus we have the desired result.

If either one of the events happens, $\mathbb{D}_V(\gamma : \varphi)$ and $\mathbb{D}_W(\gamma : \hat{\varphi})$ are singleton sets since φ is strictly convex. By Proposition 4.5 (1), they are identical. Now if $\gamma \neq 0$, $X \in \mathbb{D}_V(\gamma : \varphi)$ means $\varphi^D(\gamma) = (X, \gamma)$ and $\varphi(X) = 1$. Notice that $\varphi^D(\gamma) = (X, \gamma) \leq (F(X), \gamma) \leq \varphi^D(\gamma)$ since $\varphi(F(X)) = 1$. Thus $F(X) \in \mathbb{D}_V(\gamma : \varphi)$ which is a singleton set. So $X = F(X) \in F$. ◻

The following is an extension of [15, Theorem 1].

THEOREM 5.3. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi : V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat{\varphi} : W \to \mathbb{R}$ the restriction of φ on W. Then $Ext(B_\varphi) = GExt(B_{\hat{\varphi}})$, where $Ext(B_{\hat{\varphi}})$ denotes the set of extreme points of $B_{\hat{\varphi}}$.*

Proof. Suppose that $A \in Ext(B_\varphi)$. Let $g \in G$ such that $A = gF(A)$. Let $F(A) = tx + (1-t)y$, for some $x, y \in B_{\hat{\varphi}}$, $0 < t < 1$. Evidently $\varphi(gx) = \varphi(gy) = 1$ and $A = tgx + (1-t)gy$. Thus $gx = gy$ and hence $x = y$. So $F(A) \in Ext(B_{\hat{\varphi}})$ and thus we have the inclusion $Ext(B_\varphi) \subset GExt(B_{\hat{\varphi}})$.

On the other hand, suppose $a \in Ext(B_{\hat{\varphi}})$. We now proceed to show that $ga \in Ext(B_\varphi)$ for all $g \in G$. Without loss of generality we may assume that $a \in F$. Let $ga = tX + (1-t)Y$, for some $X, Y \in B_\varphi$, $0 < t < 1$. Set $x = g^{-1}X$ and $y = g^{-1}Y$. So $a = tx + (1-t)y$ and then $a = tp(x) + (1-t)p(y)$ where $p : V \to W$ is the orthogonal projection. By [13, Theorem 3.2], $p(x)$ is in $\operatorname{conv} HF(X)$ so that $\varphi(p(x)) \le \varphi(F(x)) = \varphi(x) = 1$ by the triangle inequality. Similarly $\varphi(p(y)) \le \varphi(y) = 1$. Now

$$1 = \varphi(a) \le t\varphi(p(x)) + (1-t)\varphi(p(y)) \le t\varphi(x) + (1-t)\varphi(y) = 1.$$

So $\varphi(p(x)) = \varphi(x) = 1$ and $\varphi(p(y)) = \varphi(y) = 1$. Since $a \in Ext(B_{\hat{\varphi}})$, $a = p(x) = p(y)$. We then have $a \in \operatorname{conv} HF(x)$. However, $a \in Ext(B_{\hat{\varphi}})$ implies that $a = hF(x)$ for some $h \in H$ and since $a \in F$, $a = F(x)$ [8, p.22]. Similarly $a = F(y)$. It follows that $x = p(x) = F(x)$ by considering $(F(x), F(x)) = (x, x)$ and similarly $y = p(y) = F(y)$. Hence $g^{-1}X = g^{-1}Y$ so that $X = Y$ which is the desired result. □

REFERENCES

[1] C. T. Benson and L. C. Grove, *Finite reflection groups*, Springer-Verlag, New York, 1985.

[2] R. Bhatia, *Matrix Analysis*, Springer, New York, 1997.

[3] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[4] C. Davis, All convex invariant functions on hermitian matrices, *Arch. Math.*, 8:276-278, 1957.

[5] M.L. Eaton and M.D. Perlman, Reflection groups, generalized Schur functions and the geometry of majorization, *Ann. Probab.* 5:829-860, 1977.

[6] R.R. Holmes and T.Y. Tam, Distance to the convex hull of an orbit under the action of a compact Lie group, *J. Austral. Math. Soc. Ser. A*, 66:331-357, 1999.

[7] J.E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Springer-Verlag, New York, 1972.

[8] J.E. Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge University Press, 1990.

[9] A.W. Knapp, *Lie Groups Beyond an Introduction*, Birkhäuser, Boston, 1996.

[10] A.S. Lewis, Derivatives of spectral functions, *Math. Oper. Res.*, 21:576–588, 1996.

[11] A.S. Lewis, Group invariance and convex matrix analysis, *SIAM J. Matrix Anal. Appl.*, 17:927-949, 1996.

[12] J.R. Munkres, *Topology: a First Course*, Englewood Cliffs, N.J., 1975.

[13] M. Niezgoda, Group majorization and Schur type inequalities, *Linear Algebra Appl.*, 268:9-30, 1998.

[14] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[15] W.S. So, Facial structures of Schatten p-norms, *Linear and Multilinear Algebra*, 27:207-212, 1990.

[16] T.Y. Tam, An extension of a result of Lewis, *Electronic J. Linear Algebra* 5:1-10, 1999.

[17] T.Y. Tam, Group majorization, Eaton triples and numerical range, *Linear and Multilinear Algebra*, 47:1-28, 2000.

[18] T.Y. Tam and W.C. Hill, On G-invariant norms, *Linear Algebra Appl.*, 331:101-112, 2001.

[19] T.Y. Tam and W.C. Hill, Derivatives of orbital functions, an extension of Berezin-Gel'fand's theorem and applications, submitted for publication (2000), pdf file available at http://www.auburn.edu/∼tamtiny/pub.html.

[20] J. von Neumann, Some matrix-inequalities and metrization of matrix-space, Tomsk. Univ. Rev. 1 (1937), 286-300; in *Collected Works*, Pergamon, New York, 4:205-219, 1962.

[21] G.A. Watson, Characterization of the subdifferential of some matrix norms, *Linear Algebra Appl.*, 170:33-45, 1992.

[22] K. Zietak, On the characterization of the maximal points of the unit sphere of matrice, *Linear Algebra Appl.*, 106:57-75, 1988.

[23] K. Zietak, Subdifferential, faces and dual matrices, *Linear Algebra Appl.*, 185:125-141, 1993.