Journal of Convex Analysis
Volume 26 (2019), No. 2, 573–592

Ky Fan's Dominance Theorem and Eaton Triple

Charles L. Dolberry
Math/Science Division, Andrew College, Cuthbert, GA 39840, U.S.A.
charlesdolberry@andrewcollege.edu

Tin-Yau Tam*
Department of Mathematics and Statistics, University of Nevada, Reno, NV 89557, U.S.A.
ttam@unr.edu

Received: September 30, 2017
Revised manuscript received: June 18, 2018
Accepted: June 20, 2018

A generalization of the Dominance Theorem of Ky Fan on the unitarily invariant norm is obtained in the content of Eaton triple. The results of Zietak on the characterization of the set of dual matrices of a given matrix and the faces of the unit ball, both with respect to a unitarily invariant norm, are also extended in the same context. The notion of dual matrices of a given matrix A with respect to a unitarily invariant norm is related to the subdifferential of the norm at A. A generalization of So's characterization of the extreme points of the unit ball is given.

Keywords: Ky Fan's Dominance Theorem, Eaton triple, unitarily invariant norm, extreme point, unit ball.

2010 Mathematics Subject Classification: 90C31, 15A18

1. Introduction

Let G be a closed subgroup of the orthogonal group $\mathrm{O}(V)$ on a finite dimensional real inner product space V. So G is compact. The triple (V, G, F) is called an *Eaton triple* if there exists a nonempty closed convex cone $F \subset V$ such that

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

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

It follows from (A2) that $Gx \cap F$ is a singleton set. Now, for each $x \in V$, define $\psi_F\colon V \to F$ by $\psi_F(x) = gx \in Gx \cap F$. Clearly the function $\psi_F\colon V \to F$ is idempotent, that is, $\psi_F^2 = \psi_F$. From now on we will denote the sole element of $Gx \cap F$ by $\psi_F(x)$. The notion of Eaton triple was first introduced by Eaton [4, p. 155] in 1987. See [4, p. 157–158] for examples. 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 $H := \{g|_W : g \in G,\ gW = W\} \subset \mathrm{O}(W)$ [23]. It is known that H is a finite reflection group [17]. See [9] for basic principles of finite reflection groups. We denote the convex hull of the underlying set S by $\operatorname{conv} S$. Niezgoda

*Corresponding author.

ISSN 0944-6532 / $ 2.50

[17, Theorem 3.2, p. 21] obtained the following characterization of Eaton triples with reduced triples.

Theorem 1.1. (Niezgoda, 1998) *Let V be a finite dimensional real inner product space and let G be a closed subgroup of $\mathrm{O}(V)$. Suppose that (V, G, F) is an Eaton triple. Set $W := \operatorname{span} F$ and $H := \{g|_W : g \in G,\ gW = W\} \subset \mathrm{O}(W)$ (with respect to the induced inner product). Then the following conditions are equivalent.*

(1) *For any $x, y \in W$, $y \in \operatorname{conv}(Gx)$ if and only if $y \in \operatorname{conv}(Hx)$.*
(2) *(W, H, F) is a reduced Eaton triple of the Eaton triple (V, G, F).*
(3) *For any $x \in V$, $\pi(x) \in \operatorname{conv}(H\psi_F(x))$, where $\pi\colon V \to W$ is the orthogonal projection onto W.*

In addition, if any of the above conditions holds, then H is a finite reflection group.

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.

We now mention some examples from matrix theory.

Example 1.2. Let $V = H_n$ be the space of $n \times n$ Hermitian matrices with inner product

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

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

$$UHU^* = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigenvalues of H. Let

$$F := \{\operatorname{diag}(h_1, h_2, \ldots, h_n) : h_1 \geq h_2 \geq \cdots \geq h_n\}$$

be the set of diagonal matrices in H_n with diagonal elements arranged in non-increasing order. This ensures that (A1) is satisfied. By a result of Ky Fan [22, Corollary 1.6, p. 4–5]: for any two $n \times n$ Hermitian matrices A and B with eigenvalues $\alpha_1 \geq \cdots \geq \alpha_n$, $\beta_1 \geq \cdots \geq \beta_n$,

$$\max\{\operatorname{tr}(AUB^*U^*) : U \in U(n)\} = \sum_{i=1}^n \alpha_i\beta_i,$$

(A2) is satisfied. Let $W = \operatorname{span} F$, the set of diagonal matrices. Then (V, G, F) is an Eaton triple with reduced triple (W, S_n, F). S_n is not essential relative to W since $W_0 = \operatorname{span}\{I_n\}$ is the set of fixed points under the action of S_n. If we identify W with $\mathbb{R}^n$, then the simple roots [25] are $\alpha_i = e_i - e_{i+1}$, $i = 1, \ldots, n-1$, where $\{e_i\}$ is the standard basis of $\mathbb{R}^n$. The corresponding λ_i are

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

Example 1.3. Take $V = \mathfrak{o}(n)$, the space of $n \times n$ real skew symmetric matrices with inner product $(X, Y) = \mathrm{tr}\,(XY^T)$ for all $X, Y \in K_n$. Let $G = \mathrm{Ad}\,(\mathrm{O}(n))$ be the adjoint action (conjugation) of the orthogonal group $\mathrm{O}(n)$ on $\mathfrak{o}(n)$, that is, $\mathrm{Ad}\,(U)(H) = UHU^T$ for all $U \in \mathrm{O}(n)$. By the spectral theorem for $n \times n$ real skew symmetric matrices, for each $H \in \mathfrak{o}(n)$ there is a $U \in \mathrm{O}(n)$ such that

$$UHU^T = \begin{cases} \begin{bmatrix} 0 & D \\ -D & 0 \end{bmatrix} & \text{if } \ n = 2m \\ \\ \begin{bmatrix} 0 & D \\ -D & 0 \end{bmatrix} \oplus 0 & \text{if } \ n = 2m+1, \end{cases}$$

where $D =$ diag $(\lambda_1, \lambda_2, \ldots, \lambda_m)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$. Of course $\pm i\lambda_1, \pm i\lambda_2, \ldots, \pm i\lambda_n$ are the eigenvalues of H (0 is included if $n = 2m+1$). Let

$$\begin{aligned} F &:= \left\{ \begin{bmatrix} 0 & D \\ -D & 0 \end{bmatrix} : h_1 \geq h_2 \geq \cdots \geq h_m \geq 0 \right\} \qquad \text{if } n = 2m, \\ F &:= \left\{ \begin{bmatrix} 0 & D \\ -D & 0 \end{bmatrix} \oplus 0 : h_1 \geq h_2 \geq \cdots \geq h_m \geq 0 \right\} \text{ if } n = 2m+1, \end{aligned}$$

in which $D = \mathrm{diag}\,(h_1, h_2, \ldots, h_m)$. Again this ensures that (A1) is satisfied by the spectral theorem for real skew symmetric matrices, and (A2) is satisfied by the result of Ky Fan [22, Corollary 1.6, p. 4–5]. Let $W = \mathrm{span}\, F$. Then (V, G, F) is an Eaton triple with reduced triple (W, S_n, F) if we view S_n as acting on D.

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

Example 1.4. 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 $\mathrm{U}(p) \times \mathrm{U}(q)$ on $\mathbb{C}_{p\times q}$ defined by $A \mapsto UAV$, where $U \in \mathrm{U}(p)$, $V \in \mathrm{U}(q)$. By the singular value decomposition, for any $A \in \mathbb{C}_{p\times q}$, there exist $U \in \mathrm{U}(p)$, $V \in \mathrm{U}(q)$ such that $A = U\Sigma V$, where $\Sigma = \mathrm{diag}\,(s_1(A), \ldots s_p(A))$ and $s_1(A) \geq s_2(A) \geq \cdots \geq s_p(A)$ are the singular values of A. Thus (A1) is satisfied. If F is the cone of $p \times q$ real "diagonal" matrices with diagonal entries in nonincreasing order, and $A, B \in F$, then clearly $(A, B) = \sum_{i=1}^p s_i(A)s_i(B)$. By a well-known result of von Neumann [26],

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

This shows that (A2) is satisfied if $(A, gB) = \mathrm{Re}\ (\mathrm{tr}\,(A(UBV)^*))$, where $g := (U, V)$. Thus $(\mathbb{C}_{p\times q}, G, F)$ is an Eaton triple with reduced triple (W, H, F), where W is the space of $p \times q$ 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 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

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

In the coming sections, the notions of Eaton triples and finite reflection groups are used to extend some results. In Section 2, we extend Ky Fan's Dominance Theorem [6]. In Sections 3 and 4 some results of Zietak on the characterization of dual matrices and on the facial structure of the unit ball are generalized. In Section 5, a result of So is extended.

Remark 1.5. In [13], Lewis introduced the notion of normal decomposition system (V, G, γ) for the study of optimization and programming problems, where V is a real inner product space and G is a closed subgroup of $O(V)$ and the map $\gamma\colon V \to V$ has the following properties:

(1) γ is idempotent, i.e., $\gamma^2 = \gamma$;
(2) γ is G-invariant, i.e., $\gamma(gx) = \gamma(x)$ for all $x \in V$
(3) for any $x \in V$, there is $g \in G$ such that $x = g\gamma(x)$.
(4) if $x, y \in V$, then $(x, y) \le (\gamma x, \gamma y)$ with equality if and only if $x = g\gamma(x)$ and $y = g\gamma(y)$ for some $g \in G$.

By Theorem 2.4 of [13], the range of γ is a closed convex cone and we denote it by F. Thus a normal decomposition system (V, G, γ) gives an Eaton triple (V, G, F). Conversely if (V, G, F) is an Eaton triple, then there exists a unique operator $\gamma\colon V \to F$ satisfying the above four conditions [17, p. 14] except the equality case.

The reduced triple is almost identical to the normal decomposition (sub)system [14, Assumption 4.1] while the difference is only the equality case. It is interesting to notice that the Eaton triple arose from probability while normal decomposition arises from optimization and programming.

2. Generalization of Ky Fan's Dominance Theorem

A norm $\|\cdot\| : \mathbb{C}_{p\times q} \to \mathbb{R}$ is said to be *unitarily invariant* if $\|UAV\| = \|A\|$ for all $A \in \mathbb{C}_{p\times q}$, $U \in \mathrm{U}(p)$ and $V \in \mathrm{U}(q)$. The characterization of unitarily invariant norms is well-known and is due to von Neumann [26, 1]. The Ky Fan k-norms $\|\cdot\|_k : \mathbb{C}_{p\times q} \to \mathbb{R}$, defined by $\|A\|_k = \sum_{i=1}^{k} s_i(A)$, for $k = 1, ..., r$, where $s_1(A) \ge \cdots \ge s_r(A)$ are the singular values of A where $r = \min\{p, q\}$. Among the unitarily invariant norms, they are probably the most important ones due to the following result of Ky Fan [6], which is what we will generalize.

Theorem 2.1. (Ky Fan's Dominance Theorem, 1951) *Let $A, B \in \mathbb{C}_{p\times q}$. Then $\|A\| \le \|B\|$ for all unitarily invariant norms if and only if $\|A\|_k \le \|B\|_k$ for all $k = 1, \ldots, r$ where $r = \min\{p, q\}$.*

To prove the extension, we recall a generalization of a result of Horn and Mathias [7] on the unitarily invariant norms, due to Tam and Hill [24]. Horn and Mathias proved that there is a compact set $C \subset \mathbb{R}^r_+\!\downarrow := \{(a_1, \dots, a_r) : a_1 \geq \cdots \geq a_r \geq 0\}$, where $r = \min\{p, q\}$, such that

$$\|X\| = \max\{\|X\|_\alpha : \alpha \in C\} \quad \text{for all } X \in \mathbb{C}_{p\times q}.$$

Here $\|X\|_\alpha = \sum_{i=1}^r \alpha_i s_i(X)$, $\alpha \in \mathbb{R}^r_+\!\downarrow$, in which $s_1(X) \geq \cdots \geq s_r(X) \geq 0$ are the singular values of X (the nonnegative square roots of the eigenvalues of X^*X or XX^*), and is called the *generalized spectral norm* if $\alpha \neq 0$. Indeed, C may be taken as $C = \{(s_1(A), \dots, s_r(A)) : \|A\|^D \leq 1, A \in \mathbb{C}_{p\times q}\}$ where $\|A\|^D = \max_{\|X\|\leq 1} |\mathrm{tr}\,(AX^*)|$ is the dual norm of A. The Horn-Mathias quasi-linear representation of unitarily invariant norms indicates the importance of the generalized spectral norm.

Let V be a real Euclidean space with inner product $(\cdot,\cdot)$. Then the *dual norm* $\|\cdot\|^D \colon V \to \mathbb{R}$ of a norm $\|\cdot\| \colon V \to \mathbb{R}$ is defined as

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

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

$$\|A\|^D = \max_{\|X\|=1} (A, X), \quad \|\cdot\| = \|\cdot\|^{DD}.$$

It is easy to see that $\|\cdot\|$ is G-invariant if and only if $\|\cdot\|^D$ is G-invariant.

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

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

Though f_α is G-invariant, since $f_\alpha(gA) = (\alpha, \psi_F(gA)) = (\alpha, \psi_F(A))$ for all $g \in G$, it is not necessarily a norm. The criteria for f_α being a norm are given in the following theorem obtained by Tam and Hill [24].

Theorem 2.2. (Tam and Hill, 2001) *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Suppose that $H \cong H_0 \times H_1 \times \cdots \times H_k$ is a decomposition of H into its irreducible components where $H_i = H|_{W_i}$, with $H_0 = \{id\}$, $i = 1, \dots, k$, and $W = W_0 + W_1 + \cdots + W_k$ is an orthogonal direct sum. Let $\alpha \in F$ be a nonzero vector. Then $\|\cdot\|_\alpha \colon V \to \mathbb{R}$ defined by $\|x\|_\alpha = (\alpha, \psi_F(x))$ is a norm if and only if H is essential relative to W, $\pi_i(\alpha) \neq 0$, where $\pi_i \colon W \to W_i$ is the orthogonal projection onto W_i, for all $i = 1, \dots, k$, and $\omega_0(\alpha) = -\alpha$ where ω_0 is the longest element of H.*

One can extend the definition of f_α to nonzero $B \in V$: $f_B(A) = (\psi_F(B), \psi_F(A))$. In other words, $\alpha = \psi_F(B)$. By (A2), $f_B(A) = \max_{g\in G}(A, gB)$.

Theorem 2.3. (Tam and Hill, 2001) Let $\|\cdot\|$ be a G-invariant norm on V where (V, G, F) is an Eaton triple. Let $C = \{\psi_F(A) : \|A\|^D \leq 1, A \in V\} \subset F$, a compact set. Then

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

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

Proposition 2.4. *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 2.3 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

$$\begin{aligned}\varphi(A) &= (\sum_{i=1}^n c_i \lambda_i, \psi_F(A)) = \sum_{i=1}^n c_i f_{\lambda_i}(A) \\ &\le \sum_{i=1}^n c_i f_{\lambda_i}(B) = (\alpha, \psi_F(B)) \le \varphi(B).\end{aligned}$$ □

Example 2.5. Theorem 2.4 yields Ky Fan's Dominance Theorem via Example 1.4, where H is essential relative to W. It is not the case for Example 1.2 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 means 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 2.4 applies for the traceless case of Example 1.2 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 [1, p. 42]. Thus $\|A\|_n \le \|B\|_n$ and Theorem 2.4 says so.

Remark 2.6. In Theorem 2.4, 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 $A, B \in W_0$, $\psi_F(A) = A$, $\psi_F(B) = B$ 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 $\psi_F(A) \in \operatorname{conv}(H\psi_F(B))$, or equivalently, $\psi_F(B) - \psi_F(A) \in \operatorname{dual}_W F$ [24].

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 2.4 is clearly true. The necessary and sufficient condition for f_λ being a norm is given in Theorem 2.2. 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}, $i = 1, \dots, n$, when they are norms.

Theorem 2.7. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Suppose W is irreducible. If $f_{\lambda_i} \colon V \to \mathbb{R}$ defined by $f_{\lambda_i}(x) = (\psi_F(x), \lambda_i)$, $x \in V$, $i = 1, \dots, n$, are norms, then $f^D_{\lambda_k}(x) = \max_{j=1,\dots,n} (\psi_F(x), \lambda_j)/(\lambda_k, \lambda_j)$, for all $x \in V$.*

Proof. Since $f^D_{\lambda_k}$ is G-invariant, we may assume $x \in F$. Now $f^D_{\lambda_k}(x) = \max\{(x, y) \colon f_{\lambda_k}(y) = 1\}$ and by (A2) we may assume $y \in F$. By Theorem 2.2, H is essential relative to W and thus $y = \sum_{j=1}^n c_j\lambda_j$ for some $c_j \geq 0$. So

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

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 [8, p. 72], $(\lambda_k, \lambda_j) = d_{kj}(\alpha_j, \alpha_j)/2 > 0$, and 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 $f_{\lambda_k}(y) = 1$, $y \in S$. Hence $y \in S \cap F$. Each $y = \sum_{j=1}^n c_j\lambda_j \in S \cap F$ can be rewritten explicitly as $\sum_{j=1}^n c_j(\lambda_k, \lambda_j)[\lambda_j/(\lambda_k, \lambda_j)]$, a convex combination of $\lambda_j/(\lambda_k, \lambda_j)$. So we have

$$\begin{aligned}
(x, y) = (x, \sum_{j=1}^n c_j\lambda_j) &= \ (x, \sum_{j=1}^n c_j(\lambda_k, \lambda_j)[\lambda_j/(\lambda_k, \lambda_j)]) \\
&= \sum_{j=1}^n c_j(\lambda_k, \lambda_j)(x, \lambda_j/(\lambda_k, \lambda_j)) \\
&\leq \sum_{j=1}^n c_j(\lambda_k, \lambda_j)(x, \lambda_i/(\lambda_k, \lambda_i)) = (x, \lambda_i/(\lambda_k, \lambda_i))
\end{aligned}$$

for some $i = 1, \dots, n$. Thus the maximum is attained among $c_j\lambda_j$ such that $c_j(\lambda_k, \lambda_j) = 1$, $j = 1, \dots, n$. Therefore

$$\begin{aligned}
f^D_{\lambda_k}(x) &= \max\{(x, y) : f_{\lambda_k}(y) = 1\} \\
&= \max\{(x, \sum_{j=1}^n c_j(\lambda_k, \lambda_j)[\lambda_j/(\lambda_k, \lambda_j)]) : \sum_{j=1}^n c_j(\lambda_k, \lambda_j) = 1\} \\
&= \max_{j=1,\dots,n} (x, \lambda_j)/(\lambda_k, \lambda_j).
\end{aligned}$$

□

Example 2.8. With respect to Example 1.4 and $p = q = n$, the symmetric matrix $((\lambda_k, \lambda_j))$ is:

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

Direct observation leads to the conclusions that $f^D_{\lambda_n}(A) = 2f_{\lambda_1}(A) = 2s_1(A)$ and

$$f^D_{\lambda_k}(A) = \max\{f_{\lambda_1}(A), 2f_{\lambda_n}(A)/k\} = \max\{s_1(A), (\sum_{i=1}^n s_i(A))/k\}$$

whenever $k \leq n - 1$. So the dual of the Ky Fan k-norm is, following [1, p. 90], $\max\{s_1(A), (\sum_{i=1}^n s_i(A))/k\}$, $k = 1, \dots, n$.

The irreducible root systems associated with the finite reflection groups are well known [9]. One can readily compute the dual norms of f_{λ_j}, $j = 1, \ldots, n$. We need the following lemma to prove our next result which is a slight extension of the previous theorem.

Lemma 2.9. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F) such that H is essential relative to W, $H \cong H_1 \times \cdots \times H_r$ is a decomposition of H into its irreducible components, where $H_i = H|_{W_i}$, with $H_0 = \{id\}$, $i = 1, \ldots, r$, and $W = W_1 + \cdots + W_r$ is an orthogonal direct sum so that $F = F_1 + \cdots + F_r$. Then (V_j, G_j, F_j) is an Eaton triple with reduced triple (W_j, H_j, F_j) where $V_j := GF_j$ and $G_j := G|_{V_j}$, $j = 1, \ldots, r$.*

Proof. We first establish that $V_j = GF_j$ is a subspace of V. By (A1),

$$\begin{aligned} V &= GF = G(F_1 + \cdots + F_r) = GF_1 + \cdots + GF_r \\ &\subset \operatorname{span} GF_1 + \cdots + \operatorname{span} GF_r \subset V. \end{aligned}$$

Since $GF_k \perp GF_l$ if $k \neq l$, we conclude that $\operatorname{span} GF_i = GF_i$, and thus $V_i = GF_i$ is a subspace.

Clearly $G_j = G|_{V_j}$ is well defined, that is, G leaves V_j invariant for all j since $GV_j = GGF_j = GF_j = V_j$. Moreover G_j is a group. Thus $V_j = GF_j = G_jF_j$ since $F_j \subset V_j$. Hence (A1) is satisfied for (V_j, G_j, F_j). To see (A2), let $x, y \in F_j \subset F$. Since G_j contains the identity,

$$\max_{g\in G}(x, gy) \geq \max_{g\in G_j}(x, gy) \geq (x, y) = \max_{g\in G}(x, gy).$$

The last equality is due to (A2) for (V, G, F). Therefore, $(x, y) = \max_{g\in G_j}(x, gy)$, and (A2) is satisfied. Since (W_j, H_j, F_j) is an irreducible Eaton triple for $j = 1, \ldots, r$ [24, p. 105], we have $W_j = \operatorname{span} F_j \subset V_j$. We need to show that $H_j = \{g|_{W_j} : g \in G_j,\ gW_j = W_j\}$. Now

$$\begin{aligned} H_j = H|_{W_j} &= \{h \in H : hW_j = W_j,\ h|_{W_k} = id,\ k \neq j\} \\ &= \{h = g|_W : g \in G,\ gW = W,\ h|_{W_j}W_j = W_j,\ h|_{W_k} = id,\ k \neq j\} \\ &= \left\{ h = (g_1 \times \cdots \times g_r)|_{W_1+\cdots+W_r} : \begin{array}{l} g_j \in G_j,\ g_jW_j = hW_j = W_j, \\ h|_{W_k} = id,\ k \neq j \end{array} \right\} \\ &= \{g_j|_{W_j} : g_j \in G_j,\ g_jW_j = W_j\}. \end{aligned}$$

Hence (W_j, H_j, F_j) is a reduced triple of (V_j, G_j, F_j). □

Proposition 2.10. *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 $H_0 = \{id\}$, $i = 1, ..., r$, and $W = W_1 + \cdots + W_r$ is an orthogonal direct sum so that $F = F_1 + \cdots + F_r$. Let $\{\alpha_{j,1}, \ldots, \alpha_{j,n_j}\}$ be the simple roots for H_j, $j = 1, \ldots, r$. Suppose $f_{\lambda_{j,k}} : V \to \mathbb{R}$, $1 \leq j \leq r$, $1 \leq k \leq n_j$ defined by $f_{\lambda_{j,k}}(x) = (\lambda_{j,k}, \psi_F(x))$ is a norm. Then*

$$f^D_{\lambda_{j,k}}(x) = \max_{t=1,\ldots,n_j} (\psi_{F_j}(x), \lambda_{j,k})/(\lambda_{j,k}, \lambda_{j,t}), \qquad \textit{for all } x \in V.$$

Proof. Apply Lemma 2.9 to (V, G, F). Then apply Theorem 2.7 on each Eaton triple (V_j, G_j, F_j) with reduced triple (W_j, H_j, F_j). $\square$

We remark that the assumption that H is essential relative to W in Theorem 2.10 is necessary for each $f_{\lambda_{j,k}}$ being a norm. See Theorem 2.2.

3. Dual elements and facial structure

Motivated by [29] we introduce the notion of the φ-dual elements to $A \in V$, where $\varphi\colon V \to \mathbb{R}$ is a norm (not necessarily G-invariant): An element $K \in V$ with $\varphi(K) \leq 1$ is said to be a *φ-dual element* to A if $\varphi^D(A) = (A, K)$. It is clear that if $A \neq 0$, 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,

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

If $A \neq 0$, 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 $B_\varphi = \{A \in V : \varphi(A) \leq 1\}$ the unit ball in V associated with the norm φ. Clearly $\mathbb{D}_V(0 : \varphi) = B_\varphi$. A norm $\varphi\colon V \to \mathbb{R}$ 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 $A \neq 0$ and φ is strictly convex.

Given a norm φ on $V \neq \{0\}$, 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 $B, C \in B_\varphi$ [29]. In other words, every closed line segment in B_φ with a relative interior point in $\mathbb{F}$ has both endpoints in $\mathbb{F}$ [19, 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 [19, Corollary 18.11]. A nontrivial face $\mathbb{F}$ is called an *maximal face* of B_φ if there are no other nontrivial 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 [21, Theorem 2] is merely a particular case.

Example 3.2. Let φ be a norm (not necessarily G-invariant) on V. Note that 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 [29, Theorem 4.1] and the idea of the proof is from [21, 29]. Also see [21, 28]. 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 we have $\bigcap_{H\in\mathbb{F}} \mathbb{D}_V(H : \varphi^D) \neq \phi$. Hence*

(1) *there exists $A \in V$ such that $\mathbb{F} \subset \mathbb{D}_V(A : \varphi)$, and*

(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 [16, 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, since $\varphi^D(K) = 1$,

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

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) This 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. We state [29, Theorem 4.1] as a corollary.

Corollary 3.4. (Zietak, 1993) *Let B_φ be the unit ball with respect to a unitarily invariant norm φ, and let a convex subset $\mathbb{F}$ be a proper face of B_φ. Then there exists a matrix A, $A \neq 0$, such that $\mathbb{F} \subseteq \mathbb{D}_V(A : \varphi)$.*

4. Characterization of the dual elements

In this section we generalize some results of Zietak. The following is a generalization of [29, Theorem 3.1].

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 $\varphi\colon V \to \mathbb{R}$:

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

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 $(x, \varphi(x))$ [19, p. 215]. We will make use of the following well known result so we state it explicitly.

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 $\partial\varphi^D(A) = \mathbb{D}_V(A, \varphi)$.*

The Supporting Hyperplane Theorem [20, p. 58] states that given a closed convex subset C of V with a boundary point x, there exists $w \in V$ such that $(w, x) \geq (w, y)$ for all $y \in C$. If we let C be the closed unit ball with respect to a norm φ and $\varphi(x) = 1$, then the set of all w such that $(w, x) = 1$ which satisfy the Supporting Hyperplane Theorem is precisely the set of φ^D-dual elements of x. To see this, let $w \in \mathbb{D}_V(x, \varphi^D)$. Then

$$(w, x) = 1 \geq \varphi^D(w) = \max_{\varphi(y) \leq 1} (w, y) \geq (w, y)$$

for all $y \in C$. Conversely, if w satisfies the Supporting Hyperplane Theorem and $(w, x) = 1$, then $\varphi(x) = 1 = (w, x)$, and $\varphi^D(w) = \max_{\varphi(y) \leq 1}(w, y) \leq (w, x) = 1$.

Example 4.3. Consider the space $\mathbb{R}^2$ and let φ be the 1-norm. Then φ^D is the max norm. Let C be the unit ball with respect to φ and take $x = (1, 0)$. Then

$$\begin{aligned}\mathbb{D}_{\mathbb{R}^2}(x, \varphi^D) &= \{z : \varphi(x) = (x, z),\ \varphi^D(z) \leq 1\} \\ &= \{z : z_1 = 1,\ |z_2| \leq 1\} = \{(1, z_2) : |z_2| \leq 1\}.\end{aligned}$$

It is easy to see that any $w \in \mathbb{D}_{\mathbb{R}^2}(x, \varphi^D)$ satisfies the Supporting Hyperplane Theorem and that $(w, x) = 1$. In this case there are infinitely many such w (and thus many hyperplanes), since x is a "corner" of the unit ball.

To see what happens when x is not a "corner" of the unit ball, let us consider the same x on the unit ball C^D with respect to φ^D. Since $\varphi^D(x) = 1$, it must be that $(w, x) = 1$ for $w \in \mathbb{D}_{\mathbb{R}^2}(x, \varphi)$. Hence $w_1 = 1$. Also $\varphi(w) \leq 1$, so $w_2 = 0$. Thus $w = x$ is the only element of $\mathbb{R}^2$ satisfying the Supporting Hyperplane Theorem such that $(w, x) = 1$, and hence there is only one supporting hyperplane at x.

Proposition 4.4. (1) *Let V be a real inner product space and let $x \in V$ and $g \in O(V)$. Let $\varphi\colon V \to \mathbb{R}$ 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 $\varphi\colon V \to \mathbb{R}$ be a G-invariant convex function. Then $\partial\varphi(x) = g\partial\varphi(\psi_F(x))$, where $g\psi_F(x) = x$, $x \in V$.*

Proof. The second part follows immediately from the first, which follows from a well-known result in Convex Analysis [19, Theorem 23.9], stating that for a convex function $f\colon Y \to \mathbb{R}$ on a finite dimensional normed space Y and a linear mapping $A\colon X \to Y$ from X from a finite dimensional space X into Y,

$$\partial(f \circ A) = A^*((\partial f)(x)), \quad \text{for all } x \in X.$$

Alternatively, one may get the second part in the following way:

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

Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). It is clear that the restriction $\hat{\varphi}\colon W \to \mathbb{R}$ of a G-invariant norm $\varphi\colon V \to \mathbb{R}$ on W is also an H-invariant norm. On the other hand, if $\hat{\varphi}\colon W \to \mathbb{R}$ is an H-invariant norm, then one can define $\varphi\colon V \to \mathbb{R}$ by $\varphi(x) = \hat{\varphi}(\psi_F(x))$ which is a G-invariant norm. This is a generalization [24] of von Neumann's well-known result on the one-to-one correspondence between unitarily invariant norms and symmetric gauge functions [26] as well as the result of Davis [3]. This is also true for G-invariant convex functions [24]. Given $\gamma \in F$, how do we obtain $\mathbb{D}_W(\gamma : \hat{\varphi})$ from $\mathbb{D}_V(\gamma : \varphi)$? We now give the answer in the following proposition.

Lemma 4.5. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi\colon V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat{\varphi}\colon W \to \mathbb{R}$ 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 $x \in W$,

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

In other words, $\widehat{(\varphi^D)} \geq (\hat{\varphi})^D$. Moreover we may assume that $x \in F$. Let $u \in V$ such that $\varphi(u) \leq 1$ and $\widehat{(\varphi^D)}(x) = (u, x)$. By (A2), $(u, x) \leq (gu, x)$, 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 $\varphi(w) \leq 1$. So $\widehat{(\varphi^D)}(x) \leq (\hat{\varphi})^D(x)$ and thus $\widehat{(\varphi^D)} = (\hat{\varphi})^D$. □

Proposition 4.6. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi\colon V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat{\varphi}\colon W \to \mathbb{R}$ 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 $\psi_F(X) \in \mathbb{D}_W(\psi_F(A) : \hat{\varphi})$.*

Proof. (1) $X \in \mathbb{D}_W(\gamma : \hat{\varphi}) \;\Leftrightarrow\; X \in W,\ \hat{\varphi}(X) \le 1,\ (\gamma, X) = \hat{\varphi}^D(\gamma)$
$\Leftrightarrow\; X \in W,\ \varphi(X) \le 1,\ (\gamma, X) = \widehat{(\varphi^D)}(\gamma)$ since $\widehat{(\varphi^D)} = (\hat{\varphi})^D$
$\Leftrightarrow\; X \in W,\ \varphi(X) \le 1,\ (\gamma, X) = \varphi^D(\gamma) \;\Leftrightarrow\; X \in W,\ X \in \mathbb{D}_V(\gamma : \varphi)$.

(2) If $X \in \mathbb{D}_V(A : \varphi)$, then $\varphi(X) = 1$ and

$$\begin{aligned}\varphi^D(A) = (X, A) &\le (\psi_F(X), \psi_F(A)) \le \hat{\varphi}^D(\psi_F(A)) \\ &= \widehat{\varphi^D}(\psi_F(A)) = \varphi^D(\psi_F(A)) = \varphi^D(A).\end{aligned}$$

Thus $\psi_F(X) \in \mathbb{D}_W(\psi_F(A) : \hat{\varphi})$ since $\hat{\varphi}(\psi_F(X)) = \varphi(X) = 1$. □

Proposition 4.6 is an extension of [28, Theorem 4.1] and its first part 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 $\mathbb{D}_W(\gamma : \hat{\varphi})$. To this end we need to recall some basics about Clarke's generalized gradient.

Let Y be a subset of V. A function $f\colon Y \to \mathbb{R}$ is said to be *Lipschitz* [2, p. 25] on Y with Lipschitz constant K if for some $K \ge 0$,

$$|f(y) - f(y')| \le K\|y - y'\|, \qquad y, y' \in Y, \tag{1}$$

where the norm is induced by the inner product. We say that f is *Lipschitz near* x if for some $\epsilon > 0$, f satisfies the Lipschitz condition (1) on the set $x + \epsilon B$, where B is the open unit ball. The Lipschitz condition clearly implies continuity.

Let f be Lipschitz near a given $x \in V$ and let $0 \ne v \in V$. The *Clarke directional derivative* [2, p. 25] of f at x in the direction v is defined as

$$f^o(x; v) := \limsup_{y \to x\ t \downarrow 0} \frac{f(y + tv) - f(y)}{t}, \tag{2}$$

where $y \in V$ and $t > 0$. The *Clarke generalized gradient* of f at x, denoted by $\partial f(x)$, is defined as

$$\partial f(x) := \{\xi \in V : f^o(x; v) \ge (\xi, v) \text{ for all } v \in V\}. \tag{3}$$

When f is smooth (continuously differentiable) at x, $\partial f(x)$ coincides with the usual gradient $\nabla f(x)$, that is, $\partial f(x) = \{\nabla f(x)\}$. It is known [2, p. 27–28] that $\partial f(x)$ is a nonempty compact convex subset of V and

$$f^o(x; v) = \max_{\xi \in \partial f(x)} (\xi, v),$$

that is, $f^o(x; \cdot)$ is the support function of $\partial f(x)$.

Remark 4.7. The definition of $\partial f(x)$ in [2, p. 27] is given as a subset of V^*, the dual space of V since the setting there is for a general Banach space. Riesz's representation theorem for V tells us that linear functionals on V are uniquely represented by vectors in V. Thus we have (3) since we are primarily concerned with finite-dimensional inner product spaces.

To prove our next result we need the following lemma due to Tam and Hill [25].

Lemma 4.8. (Tam and Hill, 2016) *Let $U \subset W$ be open and H-invariant. Suppose that the function $f: U \to \mathbb{R}$ is H-invariant and locally Lipschitz around $\mu \in F$. Then*
$$\partial(f \circ \psi_F)(\mu) = \operatorname{conv}(\{k\gamma : k \in G_\mu, \gamma \in \partial f(\mu)\}).$$

Theorem 4.9. *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*

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

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

Proof. We first notice that $\hat{\varphi}: W \to \mathbb{R}$ is convex and Lipschitz, and $\varphi = \hat{\varphi} \circ \psi_F$. Since φ is convex and Lipschitz, the subdifferentials of φ and $\hat{\varphi}$ coincide with the Clarke's generalized gradients of φ and $\hat{\varphi}$ [2, Proposition 2.2.7] respectively, and thus we have

$$\begin{aligned}
\partial\varphi(A) &= g\,\partial\varphi(\psi_F(A)) \quad \text{by Proposition 4.4, where } g \in G_A \\
&= g\,\partial_C\varphi(\psi_F(A)) = g\,\partial_C(\hat{\varphi} \circ \psi_F)(\psi_F(A)) \quad \text{since } \hat{\varphi} \circ \psi_F = \varphi \\
&= g \operatorname{conv}(G_{\psi_F(A)} \partial_C \hat{\varphi}(\psi_F(A))) \quad \text{by Lemma 4.8} \\
&= \operatorname{conv}(G_A \partial\hat{\varphi}(\psi_F(A))) \quad \text{since } g\,G_{\psi_F(A)} = G_A. \qquad \square
\end{aligned}$$

Proposition 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. Then*
$$\partial\varphi(A) = \operatorname{conv}(G_A \partial\hat{\varphi}(\psi_F(A))),$$
where $G_A := \{g \in G : A = g\psi_F(A)\}$. Hence

$$\mathbb{D}_V(A : \varphi) = \operatorname{conv}(G_A \mathbb{D}_W(\psi_F(A) : \hat{\varphi})).$$

In particular, when $\gamma \in F$, we have

$$\mathbb{D}_V(\gamma : \varphi) = \operatorname{conv}(G_\gamma \mathbb{D}_W(\gamma : \hat{\varphi})) \quad \text{and} \quad G_\gamma = \{g \in G : g\gamma = \gamma\}.$$

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

$$\begin{aligned}
|\varphi(y) - \varphi(y')| &\le \varphi(y - y') = \max_{\alpha \in C}(\alpha, \psi_F(y - y')) \quad \text{by Theorem 2.3} \\
&\le \max_{\alpha \in C}(\alpha, \alpha)^{1/2}(\psi_F(y - y'), \psi_F(y - y'))^{1/2} \le K(y - y', y - y')^{1/2},
\end{aligned}$$

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

$$\begin{aligned}
\mathbb{D}_V(A : \varphi) &= \partial\varphi^D(A) \quad \text{by Proposition 4.2} \\
&= \operatorname{conv}(G_A \partial\widehat{(\varphi^D)}(\psi_F(A))) \quad \text{by Theorem 2.2}
\end{aligned}$$

$$= \operatorname{conv}(G_A\mathbb{D}_W(\psi_F(A) : \widehat{((\varphi^D))}^D)) \quad \text{by Proposition 4.2}$$
$$= \operatorname{conv}(G_A\mathbb{D}_W(\psi_F(A) : (\hat{\varphi})^{DD})) \quad \text{by Lemma 4.5}$$
$$= \operatorname{conv}(G_A\mathbb{D}_W(\psi_F(A) : \hat{\varphi})) \quad \text{since } (\hat{\varphi})^{DD} = \hat{\varphi}. \qquad \square$$

We remark that the last statement of Theorem 4.10 generalizes [29, Theorem 3.2] if one considers the fact that $G_\gamma\mathbb{D}_W(\gamma, \hat{\varphi})$ is a convex set for Example 1.4 [29]. See [27, Theorem 2]. It is also the case for Example 1.2 [12, Theorem 3.12]. See Remark 3.12 of [24].

Example 4.11. 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 4.12. (1) Following Example 1.2, let $\varphi(A) = \sum_{i=1}^n |\lambda_i(A)|$, 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 [1] $\varphi^D(A) = \max_{i=1,\dots,n} |\lambda_i(A)|$, the operator norm. Let $\gamma = \operatorname{diag}(\gamma_1, \dots, \gamma_n)$, 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 $\gamma \neq 0$ and thus $\mathbb{D}_W(\gamma : \hat{\varphi}) \subset S_\varphi$) with $\max_{i=1,\dots,n} |\gamma_i| = \sum_{i=1}^n \gamma_i k_i$. So

Case 1: $|\gamma_1| \geq |\gamma_n|$ (so $\gamma_1 \geq 0$): $k_1 = 1$, and $k_2 = \cdots = k_n = 0$. So $G_\gamma\mathbb{D}_W(\gamma : \hat{\varphi}) = \mathbb{D}_W(\gamma : \hat{\varphi}) = \{\operatorname{diag}(1, \dots, 0)\}$. On the other hand, $X \in \mathbb{D}_V(\gamma : \varphi)$ means $\gamma_1 = \sum_{i=1}^n \gamma_i x_{ii}$ 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, $(s_1(X), \dots, s_n(X))$. So $x_{11} = 1$ and $x_{22} = \cdots = x_{nn} = 0$ and thus $X = \operatorname{diag}(1, 0, \dots, 0)$. So $\mathbb{D}_V(\gamma : \varphi) = \mathbb{D}_W(\gamma : \hat{\varphi})$.

Case 2: $|\gamma_1| < |\gamma_n|$ (so $\gamma_n \leq 0$): $k_n = -1$, and $k_1 = \cdots = k_{n-1} = 0$. So $G_\gamma\mathbb{D}_W(\gamma : \hat{\varphi}) = \mathbb{D}_W(\gamma : \hat{\varphi}) = \{\operatorname{diag}(0, \dots, 0, -1)\}$. 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 entire 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 $k_i \geq 0$ for all $i = 1, \dots, n$. Hence

$$\mathbb{D}_W(\gamma : \hat{\varphi}) = \{\operatorname{diag}(k_1, \dots, k_n) : k_i \geq 0,\ i = 1, \dots, n,\ \sum_{i=1}^n k_i = 1\}, \quad \text{so that}$$

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

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 $(s_1(X), \dots, s_n(X))$. So $0 \leq x_{ii}$ for all $i = 1, \dots, n$ and $\sum_{i=1}^n x_{ii} = 1$. Thus the eigenvalues of X must be nonnegative. So X must be of the form $U\operatorname{diag}(k_1, \dots, k_n)U^{-1}$ for some $U \in \mathrm{U}(n)$, where $k_i \geq 0$, for all $i = 1, \dots, n$, with $\sum_{i=1}^n k_i = 1$. Hence $\mathbb{D}_V(\gamma : \varphi) = G_\gamma\mathbb{D}_W(\gamma : \hat{\varphi})$. Indeed by [23, Theorem 11],

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

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

We have the following extension of the first part of Example 4.12. The result takes care of the regular points $\gamma\in W$, that is, the points in $\mathrm{Int}_W F$.

Theorem 4.13. *Let (V,G,F) be an Eaton triple with reduced triple (W,H,F). Let $\varphi\colon V\to\mathbb{R}$ be a G-invariant norm and denote by $\hat\varphi\colon W\to\mathbb{R}$ the restriction of φ on W. If $\psi_F(A)\in\mathrm{Int}_W F$, then $\mathbb{D}_V(A:\varphi)=g\mathbb{D}_W(\psi_F(A):\hat\varphi)$, where $A=g\psi_F(A)$. In particular, if $\gamma\in\mathrm{Int}_W F$, 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 $\gamma\in\mathrm{Int}_W F$, 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 4.11). So $1=\varphi^D(\gamma)=(\gamma,X)$ and $1=\varphi(X)=\max_{\alpha\in C}(\alpha,\psi_F(X))$, where $C=\{\beta\in F:\varphi^D(\beta)\le 1\}$. Let $p(X)$ be the projection of X under the orthogonal projection $p\colon V\to W$ with respect to the inner product. Since $\gamma\in C\subset F\subset W$, we have $(\gamma,X)=(\gamma,p(X))=(\gamma,\psi_F(X))$. By [17, Theorem 3.2], $p(X)$ is in $\mathrm{conv}\,(H\psi_F(X))$, where H is a finite reflection group, that is, $p(X)=\sum_{i=1}^k c_ih_i\psi_F(X)$ where $h_i\in H$ and $c_i>0$ with $\sum_{i=1}^k c_i=1$, and $h_i\psi_F(X)$, $i=1,\ldots,k$, are distinct. Thus $\sum_{i=1}^k c_i(\gamma,h_i\psi_F(X))=(\gamma,\psi_F(X))$. By (A2) $(\gamma,h_i\psi_F(X))\le(\gamma,\psi_F(X))$ and thus $(\gamma,h_i\psi_F(X))=(\gamma,\psi_F(X))$ for all $i=1,\ldots,k$. Suppose $h_i\psi_F(X)\ne\psi_F(X)$. By [8, p. 22], each nonzero $\psi_F(X)-h_i\psi_F(X)$ is a nonnegative combination of the simple roots, that is, for each $i=1,\ldots,k$, there exist nonnegative numbers $d_j\ge 0$, $j=1,\ldots,n$, not all zero, such that $\psi_F(X)-h_i\psi_F(X)=\sum_{j=1}^n d_j\alpha_j$. Now, $\gamma\in\mathrm{Int}_W F$, and we have $(\gamma,\alpha_j)>0$ for all $j=1,\ldots,n$ so that $(\gamma,\psi_F(X)-h_i\psi_F(X))>0$, a contradiction. So $p(X)=\psi_F(X)$. By considering $(X,X)=(\psi_F(X),\psi_F(X))$, we have $X=\psi_F(X)$. Thus $\mathbb{D}_V(\gamma:\varphi)\subset F$. By Proposition 4.6, $\mathbb{D}_V(\gamma:\varphi)=\mathbb{D}_W(\gamma:\hat\varphi)$. □

In view of Example 4.12 and Theorem 4.13, one may guess that if $\gamma\in\mathrm{Int}_W F$, 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 4.14. Let $\varphi(A)=\sum_{i=1}^k s_i(A)$, the Ky Fan k-norm on $\mathbb{C}_{n\times n}$. Note that $\varphi^D(A)=\max\{s_1(A),(\sum_{i=1}^n s_i(A))/k\}$ [1, 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 $a_1\ge\cdots\ge a_n\ge 0$ by Theorem 4.13, $\sum_{i=1}^k a_i=1$, 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. This also shows that the corresponding Ky Fan's k-norm is not strictly convex.

5. The unit ball

Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi\colon V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat{\varphi}\colon W \to \mathbb{R}$ 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 4.11). We summarize this as:

Proposition 5.1. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi\colon V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat{\varphi}\colon W \to \mathbb{R}$ the restriction of φ on W. Then:* (1) $B_{\hat{\varphi}} = B_\varphi \cap W$, *and* (2) $B_\varphi = GB_{\hat{\varphi}}$.

We state the following theorem of Tam [23, Theorem 10] which is used in the proof of our next result.

Theorem 5.2. (Tam, 2000) *Let (W, H, F) be a reduced triple of the Eaton triple (V, G, F). If $x, y \in V$, then $(x+y)^* \prec_H x^*+y^*$, or equivalently, $x^*+y^*-(x+y)^* \in \mathrm{dual}_W F$.*

Recall that a norm $\varphi\colon V \to \mathbb{R}$ is said to be *strictly convex* if the condition 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 [28, Theorem 3.1].

Theorem 5.3. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi\colon V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat{\varphi}\colon W \to \mathbb{R}$ 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

$$\begin{aligned}
1 = \hat{\varphi}(\psi_F(X)) &= \hat{\varphi}(\psi_F(\frac{1}{2}[X_1 + X_2])) = \hat{\varphi}(\frac{1}{2}[\psi_F(X_1 + X_2)]) \\
&= \max_{\alpha \in \hat{C}}(\alpha, \frac{1}{2}\psi_F(X_1 + X_2)) \quad \text{(by Theorem 2.3)} \\
&\le \frac{1}{2}\max_{\alpha \in \hat{C}}(\alpha, \psi_F(X_1) + \psi_F(X_2)) \quad \text{(by Theorem 5.2)} \\
&\le \frac{1}{2}\max_{\alpha \in \hat{C}}(\alpha, \psi_F(X_1)) + \frac{1}{2}\max_{\alpha \in \hat{C}}(\alpha, \psi_F(X_2)) \\
&= \frac{1}{2}\hat{\varphi}(\psi_F(X_1)) + \frac{1}{2}\hat{\varphi}(\psi_F(X_2)) = 1,
\end{aligned}$$

where $\hat{C} = \{\psi_F(A) : (\hat{\varphi})^D(A) = 1, A \in W\}$. So $\frac{1}{2}[\psi_F(X_1) + \psi_F(X_2)] \in S_{\hat{\varphi}}$. By the strict convexity of $\hat{\varphi}$, $\frac{1}{2}[\psi_F(X_1) + \psi_F(X_2)] = \psi_F(X_1) = \psi_F(X_2)$. We now proceed to show that $\psi_F(X) = \psi_F(X_1)$ by contradiction. Suppose $\psi_F(X) \neq \psi_F(X_1)$. We may assume that $\psi_F(X), \psi_F(X_1) = \psi_F(X_2) \in S_{\hat{\varphi}}$. By the strict convexity of $\hat{\varphi}$, the element $Y := \frac{1}{2}(\psi_F(X) + \psi_F(X_1)) \notin S_{\hat{\varphi}}$. For all $i = 1, \dots, n$, by Theorem 5.2,

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

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

since $\psi_F(X_1) = \psi_F(X_2)$. So by Theorem 2.4, $\hat{\varphi}(\psi_F(X)) \leq \hat{\varphi}(Y) \leq \hat{\varphi}(\psi_F(X_1))$. Thus $Y \in S_{\hat{\varphi}}$, a contradiction.

Hence $\psi_F(X) = \psi_F(X_1) = \psi_F(X_2) \in F$. 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 4.6 (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 $\varphi^D(\gamma) = (X, \gamma) \leq (\psi_F(X), \gamma) \leq \varphi^D(\gamma)$ since $\varphi(\psi_F(X)) = 1$. Thus $\psi_F(X) \in \mathbb{D}_V(\gamma : \varphi)$ which is a singleton set. So $X = \psi_F(X) \in F$. □

We state [28, Theorem 3.1] as a corollary of Theorem 5.3.

Corollary 5.4. (Zietak, 1988) *The unitarily invariant norm $\|\cdot\|_\varphi$ is strictly convex if and only if the symmetric gauge function φ is strictly convex.*

The following is an extension of a result of So [21, Theorem 1].

Theorem 5.5. *Let (V, G, F) be an Eaton triple with reduced triple (W, H, F). Let $\varphi\colon V \to \mathbb{R}$ be a G-invariant norm and denote by $\hat{\varphi}\colon W \to \mathbb{R}$ the restriction of φ on W. Then $\operatorname{Ext} B_\varphi = G \operatorname{Ext} B_{\hat{\varphi}}$, where $\operatorname{Ext} B_{\hat{\varphi}}$ denotes the set of extreme points of $B_{\hat{\varphi}}$.*

Proof. Suppose that $A \in \operatorname{Ext} B_\varphi$. Let $g \in G$ such that $A = g\psi_F(A)$. Let $\psi_F(A) = tx + (1-t)y$, 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 $\psi_F(A) \in \operatorname{Ext} B_{\hat{\varphi}}$ and thus we have the inclusion $\operatorname{Ext} B_\varphi \subset G \operatorname{Ext} B_{\hat{\varphi}}$.

On the other hand, suppose $a \in \operatorname{Ext} B_{\hat{\varphi}}$. We now proceed to show that $ga \in \operatorname{Ext} B_\varphi$ 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 $p\colon V \to W$

is the orthogonal projection. By [17, Theorem 3.2], $p(x)$ is in $\mathrm{conv}\,(H\psi_F(X))$ so that $\varphi(p(x)) \le \varphi(\psi_F(x)) = \varphi(x) = 1$ 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 $a \in \mathrm{Ext}B_{\hat{\varphi}}$, $a = p(x) = p(y)$. We then have $a \in \mathrm{conv}\,(H\psi_F(x))$. However, $a \in \mathrm{Ext}B_{\hat{\varphi}}$ implies that $a = h\psi_F(x)$ for some $h \in H$ and since $a \in F$, $a = \psi_F(x)$ [9, p. 22]. Similarly $a = \psi_F(y)$. It follows that $x = p(x) = \psi_F(x)$ by considering $(\psi_F(x), \psi_F(x)) = (x, x)$ and similarly $y = p(y) = \psi_F(y)$. Hence $g^{-1}X = g^{-1}Y$ so that $X = Y$ which is the desired result. □

We now state [21, Theorem 1] as a corollary to Theorem 5.5.

Corollary 5.6. (So, 1990) *Let $B_{\|\cdot\|_\varphi}$ be the unit ball with respect to a unitarily invariant norm $\|\cdot\|_\varphi$, and let B_φ be the unit ball with respect to the symmetric gauge function φ. Then A is an extreme point of $B_{\|\cdot\|_\varphi}$ if and only if A is unitarily equivalent to* $\mathrm{diag}\, x$ *for some extreme point x of B_φ.*

Acknowledgement. Thanks are given to the anonymous referee for his/her suggestions.

References

[1] R. Bhatia: Matrix Analysis, Springer, New York et al. (1997).

[2] F. H. Clarke: Optimization and Nonsmooth Analysis, Wiley, New York et al. (1983).

[3] C. Davis: All convex invariant functions on hermitian matrices, Arch. Math. 8 (1957) 276–278.

[4] M. L. Eaton: Lectures on Topics in Probability Inequalities, CWI Tract, Stichting Mathematisch Centrum, Amsterdam (1987).

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

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

[7] R. Horn, R. Mathias: Cauchy-Schwarz inequalities associated with positive semi-definite matrices, Linear Algebra Appl. 142 (1990) 63–82.

[8] J. E. Humphreys: Introduction to Lie Algebras and Representation Theory, Springer, New York et al. (1972).

[9] J. E. Humphreys: Reflection Groups and Coxeter Groups, Cambridge University Press, Cambridge (1990).

[10] A. W. Knapp: Lie Groups Beyond an Introduction, Birkhäuser, Boston (1996).

[11] A. S. Lewis: Convex analysis on the Hermitian matrices, SIAM J. Optimization 6 (1996) 164–177.

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

[13] A. S. Lewis: Group invariance and convex matrix analysis, SIAM J. Matrix Anal. Appl. 17 (1996) 927–949.

[14] A. S. Lewis: Lidskii's theorem via nonsmooth analysis, SIAM J. Matrix Anal. Appl. 21 (1999) 379–381.

[15] A. S. Lewis: Convex analysis on Cartan subspaces, Nonlinear Anal. 42 (2000) 813–820, 2000.

[16] J. R. Munkres: Topology: a First Course, Prentice Hall, Englewood Cliffs (1975).

[17] M. Niezgoda: Group majorization and Schur type inequalities, Linear Algebra Appl. 268 (1998) 9–30.

[18] M. Niezgoda: An analytical characterization of effective and of irreducible groups inducing cone orderings, Linear Algebra Appl. 269 (1998) 105–114.

[19] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton (1970).

[20] W. Rudin: Functional Analysis, McGraw-Hill, New York et al. (1973).

[21] W. S. So: Facial structures of Schatten p-norms, Linear and Multilinear Algebra 27 (1990) 207–212.

[22] T. Y. Tam: An extension of a result of Lewis, Electronic J. Linear Algebra 5 (1999) 1–10.

[23] T. Y. Tam: Group majorization, Eaton triples and numerical range, Linear and Multilinear Algebra 47 (2000) 11–28.

[24] T. Y. Tam, W. C. Hill: On G-invariant norms, Linear Algebra Appl. 331 (2001) 101–112.

[25] T. Y. Tam, W. C. Hill: Derivatives of orbital function and an extension of Berezin-Gelfand's theorem, Special Matrices 4 (2016) 333–349.

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

[27] G. A. Watson: Characterization of the subdifferential of some matrix norms, Linear Algebra Appl. 170 (1992) 33–45.

[28] K. Zietak: On the characterization of the maximal points of the unit sphere of matrices, Linear Algebra Appl. 106 (1988) 57–75.

[29] K. Zietak: Subdifferential, faces and dual matrices, Linear Algebra Appl. 185 (1993) 125–141.