The Electronic Journal of Linear Algebra.
A publication of the International Linear Algebra Society.
Volume 5, pp. 1-10, January 1999.
ISSN 1081-3810. http://math.technion.ac.il/iic/ela

ELA

AN EXTENSION OF A RESULT OF LEWIS*

TIN-YAU TAM†

Abstract. A result of Lewis on the extreme properties of the inner product of two vectors in a Cartan subspace of a semisimple Lie algebra is extended. The framework used is an Eaton triple which has a reduced triple. Applications are made for determining the minimizers and maximizers of the distance function considered by Chu and Driessel with spectral constraint.

Key words. Eaton triple, reduced triple, finite reflection group

AMS subject classifications. 15A42

1. Main results.

The purpose of this note is to establish an extension of a result of Lewis [7, Theorem 3.2] and make some applications.

THEOREM 1.1. (Lewis [7]) *Let $\mathfrak{g}$ be a real semisimple Lie algebra with Cartan decomposition $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ where the analytic group of $\mathfrak{k}$ is $K \subset G$. For $x \in \mathfrak{p}$, let x_0 denote the unique element of the singleton set $Ad(K)x \cap \mathfrak{a}_+$ where $\mathfrak{a}_+$ is a closed fundamental Weyl chamber. For $x, y \in \mathfrak{p}$, $(x, y) \le (x_0, y_0)$, where $(\cdot,\cdot)$ denotes the Killing form, with equality holding if and only if there is $k \in K$ such that both $Ad(k)x$ and $Ad(k)y$ are in $\mathfrak{a}_+$.*

Lewis' result generalizes some well-known results including von Neumann's result [10] and a result of Fan [2] and Theobald [15], which are corresponding to the real simple Lie algebra $\mathfrak{su}_{p,q}$ and the reductive Lie algebra $\mathfrak{gl}_n(\mathbb{R})$. Here is a framework for the extension which only requires basic knowledge of linear algebra. Let G be a closed subgroup of the orthogonal group on a real Euclidean 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 $x, y \in F$.

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 $H := \{g|_W : g \in G,\ gW = W\} \subset O(W)$ [14]. For $x \in V$, let x_0 denote the unique element of the singleton set $Gx \cap F$. It is known that H is a finite reflection group [11]. Let us recall some rudiments of finite reflection groups [4]. 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, i.e., $s_\alpha \lambda := \lambda - 2(\lambda, \alpha)/(\alpha, \alpha)\, \alpha$, $\lambda \in V$. 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 $\alpha \in \Phi$.

*Received by the editors on 3 August 1998. Accepted for publication on 1 January 1999. Handling Editor: Danny Hershkowitz.

†Department of Mathematics, 218 Parker Hall Auburn University, Auburn, AL 36849-5310, USA (tamtiny@mail.auburn.edu, http://www.auburn.edu/~tamtiny). Dedicated to Hans Schneider on the occasions of his 70th birthday and retirement.

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 [4, 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 [4, p.8] a unique *simple system* Δ, i.e., 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 $F := \{\lambda \in V : (\lambda, \alpha) \geq 0, \text{ for all } \alpha \in \Delta\}$, i.e., $F = C^-$ 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 [11]). 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 $\Delta = \{\alpha_1, \ldots, \alpha_n\}$, i.e., $\dim V_1 = n$. If $\{\lambda_1, \ldots, \lambda_n\}$ denotes the basis of $V_1 = V_0^\perp$ dual to the basis $\{\beta_i = 2\alpha_i/(\alpha_i, \alpha_i) : i = 1, \ldots, n\}$, i.e., $(\lambda_i, \beta_j) = \delta_{ij}$, then $F = \{\sum_{i=1}^n c_i\lambda_i : c_i \geq 0\} \oplus V_0$. Thus the interior $\operatorname{Int} F = C$ of F is the nonempty set $\{\sum_{i=1}^n c_i\lambda_i : c_i > 0\} \oplus V_0$. There is a unique element $\omega \in G$ sending Φ^+ to Φ^- and thus sending F to $-F$. Moreover, the length [4, p.12] of ω is the longest one [4, p.15-16]. So we call it the *longest element*.

We will present two examples requiring some basic knowledge of Lie theory [5]. Let $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ be a Cartan decomposition of a real semisimple Lie algebra $\mathfrak{g}$. Denote the Killing form of $\mathfrak{g}$ by $B(\cdot, \cdot)$. The Killing form is positive definite on $\mathfrak{p}$ but negative definite on $\mathfrak{k}$. Let K be an analytic subgroup of $\mathfrak{k}$ in the analytic group G of $\mathfrak{g}$. Now $\operatorname{Ad}(K)$ is a subgroup of the orthogonal group on $\mathfrak{p}$ with respect to the restriction of the Killing form on $\mathfrak{p}$ since the Killing form is invariant under $\operatorname{Ad}(K)$. Among the Abelian subalgebras of $\mathfrak{g}$ that are contained in $\mathfrak{p}$, choose a maximal one $\mathfrak{a}$ (*maximal Abelian subalgebra* in $\mathfrak{p}$). For $\alpha \in \mathfrak{a}^*$ (the dual space of $\mathfrak{a}$), set $\mathfrak{g}_\alpha = \{x \in \mathfrak{g} : [h, x] = \alpha(h)x \text{ for all } h \in \mathfrak{a}\}$. If $0 \neq \alpha \in \mathfrak{a}^*$ and $\mathfrak{g}_\alpha \neq 0$, then α is called a (*restricted*) *root* [5, p.313] of the pair $(\mathfrak{g}, \mathfrak{a})$. The set of roots will be denoted Σ. We have the orthogonal direct sum $\mathfrak{g} = \mathfrak{g}_0 + \sum_{\alpha\in\Sigma} \mathfrak{g}_\alpha$ known as *restricted-root space decomposition* [5, p.313]. We view $\mathfrak{a}$ as a Euclidean space by taking the inner product to be the restriction of B to $\mathfrak{a}$. The map $\mathfrak{a}^* \to \mathfrak{a}$ that assigns to each $\lambda \in \mathfrak{a}^*$ the unique element x_λ of $\mathfrak{a}$ satisfying $\lambda(x) = B(x, x_\lambda)$ for all $x \in \mathfrak{a}$ is a vector space isomorphism. We use this isomorphism to identify $\mathfrak{a}^*$ with $\mathfrak{a}$, allowing us, in particular, to view Σ as a subset of $\mathfrak{a}$. The set $\Phi = \{\alpha \in \Sigma : \frac{1}{2}\alpha \notin \Sigma\}$ generates a finite reflection group W, i.e., W is generated by the reflections s_α $(\alpha \in \Sigma)$, which is called the *Weyl group* of $(\mathfrak{g}, \mathfrak{a})$, and is a root system of W. It is called the *reduced root system* of the pair $(\mathfrak{g}, \mathfrak{a})$. Now fix a simple system Δ for the root system Φ. Then Δ determines a fundamental domain $\mathfrak{a}_+$ for the action of W on $\mathfrak{a}$. We now describe another way to view the Weyl group W. Use juxtaposition to represent the adjoint action of G on $\mathfrak{g}$, i.e., $gx = \mathrm{Ad}\,(g)x$, $g \in G$, $x \in \mathfrak{g}$. Set $N_K(\mathfrak{a}) = \{k \in K : k\mathfrak{a} \subset \mathfrak{a}\}$ (the normalizer of $\mathfrak{a}$ in K) and $Z_K(\mathfrak{a}) = \{k \in K : kx = x \text{ for all } x \in \mathfrak{a}\}$ (the centralizer of $\mathfrak{a}$ in K). Then the action of K on $\mathfrak{g}$ induces an action of the group $N_K(\mathfrak{a})/Z_K(\mathfrak{a})$ on $\mathfrak{a}$, i.e., $[k]x = kx$ for $[k] \in N_K(\mathfrak{a})/Z_K(\mathfrak{a})$. There exists an isomorphism $\psi : W \to N_K(\mathfrak{a})/Z_K(\mathfrak{a})$ that is compatible with the two actions on $\mathfrak{a}$, or more precisely, for which $wx = \psi(w)x$, $w \in W$, $x \in \mathfrak{a}$ [5, p.325, p.394]. We use the isomorphism ψ to identify these two groups (in the literature, the Weyl group is usually defined to be $N_K(\mathfrak{a})/Z_K(\mathfrak{a})$). Note in particular that, given $x \in \mathfrak{a}$, we have $Wx = N_K(\mathfrak{a})x \subset Kx$. Since $\mathrm{Ad}\,(k)$ is an automorphism of $\mathfrak{g}$, $N_K(\mathfrak{a}) = \{k \in K : k\mathfrak{a} = \mathfrak{a}\}$. Thus $W = \{\mathrm{Ad}\,(k)|_{\mathfrak{a}} : k \in K, k\mathfrak{a} = \mathfrak{a}\}$. Obviously $\mathfrak{a} = \operatorname{span} \mathfrak{a}_+$. A theorem of Cartan asserts that $\mathrm{Ad}\,(K)x \cap \mathfrak{a} \neq \phi$ [5, p.320] for any $x \in \mathfrak{p}$. Since W acts simply transitively on the (open) Weyl chambers in $\mathfrak{a}$, $\mathrm{Ad}\,(K)x \cap \mathfrak{a}_+ \neq \phi$ which is (A1) for $(\mathfrak{p}, \mathrm{Ad}\,(K), \mathfrak{a}_+)$. Indeed $|\mathrm{Ad}\,(K)x \cap \mathfrak{a}_+| = 1$. For verification of (A2), see [7].

Example 1.2. (real semisimple Lie algebras) With the above notation. Now $(\mathfrak{p}, \mathrm{Ad}\,(K), \mathfrak{a}_+)$ is an Eaton triple with a reduced triple $(\mathfrak{a}, W, \mathfrak{a}_+)$. It is also true for real reductive Lie algebras [3].

Remark: One may verify (A2) by Kostant's theorem. If $x, y \in \mathfrak{p}$, let $x_0 \in \mathrm{Ad}\,(K)x \cap \mathfrak{a}_+$ and let $k' \in K$ such that $\mathrm{Ad}\,(k')x = x_0$, then for any $k \in K$, $B(x, \mathrm{Ad}\,(k)y) = B(\mathrm{Ad}\,(k')x, \mathrm{Ad}\,(k')\mathrm{Ad}\,(k)y)$ since the Killing form is $\mathrm{Ad}\,(K)$-invariant. Thus we have $B(x, \mathrm{Ad}\,(k)y) = B(x_0, \mathrm{Ad}\,(k'k)y) = B(x_0, \pi(y'))$, where $\pi : \mathfrak{p} \to \mathfrak{a}$ is the orthogonal projection with respect to the Killing form in $\mathfrak{p}$, where $y' = \mathrm{Ad}\,(k'k)y$. By Kostant's convexity theorem [6] $\pi(y')$ is in the convex hull of Wy_0. Now let $\pi(y') = \sum_{w \in W} \alpha_w w y_0$ where $\alpha_w \geq 0$ for all $w \in W$, and $\sum_{w \in W} \alpha_w = 1$. Thus $B(x_0, \pi(y')) = B(x_0, \sum_{w \in W} \alpha_w w y_0) = \sum_{w \in W} \alpha_w (x_0, w y_0) \leq \sum_{w \in W} \alpha_w (x_0, y_0) = (x_0, y_0)$ by Lemma 3.2 of [6].

Similarly we have the following example.

Example 1.3. (compact connected Lie groups) Let G be a (real) compact connected Lie group and let $(\cdot, \cdot)$ be a bi-invariant inner product on $\mathfrak{g}$. Now $\mathrm{Ad}\,(G)$ is a subgroup of the orthogonal group on $\mathfrak{g}$ [5, p.196]. Let $\mathfrak{t}_+$ be a fixed (closed) fundamental chamber of the Lie algebra $\mathfrak{t}$ of a maximal torus T of G. Now $(\mathfrak{g}, \mathrm{Ad}\,(G), \mathfrak{t}_+)$ is an Eaton triple with reduced triple $(\mathfrak{t}, W, \mathfrak{t}_+)$ where the Weyl group W of G is often defined as $N(T)/T$ where $N(T)$ is the normalizer of T in G [5, p.201]. We remark that Theorem 1.1 is also true for compact connected Lie groups.

The following lemma is a slight extension of Proposition 2.1 in [7] which is stated for Weyl groups. Now we add the lower bound and the proof is similar.

Lemma 1.4. *Let H be a finite reflection group acting on a real Euclidean space W. For any $x, y \in W$, $(x_0, \omega y_0) \leq (x, y) \leq (x_0, y_0)$, where $\omega \in H$ is the longest element. The upper (lower) bound is achieved if and only if there exists $h \in H$ such* *that both hx and hy lie in a common closed chamber F ($hx \in F$ and $hy \in -F$).*

The following is an extension of Theorem 1.1 except we add an assumption for the upper (lower) bound attainment.

THEOREM 1.5. *Let (V, G, F) be an Eaton triple with a reduced triple (W, H, F). Then for any $x, y \in V$, we have $(x_0, -(-y_0)_0) \leq (x, y) \leq (x_0, y_0)$ and $\omega y = -(-y_0)_0$ where $\omega \in H$ is the longest element. Thus, if G is connected, then $\{(x, gy) : g \in G\}$ is the interval $[(x_0, -(-y_0)_0), (x_0, y_0)]$. If, in addition, x_0 or $y_0 \in Int_W F$ which denotes the relative interior of F in W, then the upper (lower) bound is achieved if and only if there exists $g \in G$ such that both gx and gy are in F ($gx \in F$ and $gy \in -F$).*

Proof. Let $g \in G$ such that $gx = x_0$ according to (A1). So $(x, y) = (gx, gy) = (x_0, gy) \leq \max_{g \in G}(x_0, gy) = (x_0, y_0)$ by (A2). We notice that the inequality is true without using reduced triple.

For the lower bound, notice that $(x, y) = (x_0, gy) = (x_0, \pi(gy))$ where $\pi : V \to W$ is the orthogonal projection onto W. According to a result of Niezgoda [11, Theorem 3.2], $\pi(gy) = \sum_{h \in H} \alpha_h h y_0$ where $\sum_{h \in H} \alpha_h = 1$, $\alpha_h \geq 0$ for all $h \in H$ and H is indeed a finite reflection group. Now $(x, y) = \sum_{h \in H} \alpha_h (x_0, h y_0) \geq \sum_{h \in H} \alpha_h (x_0, \omega y) = (x_0, \omega y)$ by Lemma 1.4.

Notice that the longest element ω sends vectors in F to $-F$. Now $-(-y_0)_0 \in -F$ obviously and $(-y_0) \in G(-y)$ implying $-(-y_0)_0 \in Gy$. So $-(-y_0)_0 \in (-F) \cap Gy = (-F) \cap G y_0$ and thus $\omega y_0 = -(-y_0)_0$ for any $y \in W$. So we have $(x_0, -(-y_0)_0) \leq (x, y) \leq (x_0, y_0)$.

Now we are going to handle the attainment of the upper bound. Suppose x_0 or $y_0 \in \mathrm{Int}_W F$. For definiteness, let $x_0 \in \mathrm{Int}_W F$. Now, as before $(x, y) = (x_0, gy) = (x_0, \pi(gy))$ where $gx = x_0$, $\pi(gy) = \sum_{h \in H} \alpha_h h y_0$, where $\sum_{h \in H} \alpha_h = 1$, $\alpha_h \geq 0$ for all $h \in H$, and $\pi : V \to W$ is the orthogonal projection onto W. If $(x, y) = (x_0, y_0)$, then $\sum_{h \in H} \alpha_h (x_0, h y_0) = (x_0, y_0) = \sum_{h \in H} \alpha_h (x_0, y_0)$. If $\alpha_h \neq 0$, then $(x_0, h y_0) = (x_0, y_0)$, because otherwise $\sum_{h \in H} \alpha_h (x_0, y_0) < \sum_{h \in H} \alpha_h (x_0, y_0)$ by (A2), a contradiction. In consequence, there exists $g_h \in H$ such that both $g_h x_0$ and $g_h h y_0$ are in F by Lemma 1.4. By [4, Theorem 1.12, p.22], $g_h = id$ and $g_h h = h$ is a product of simple reflections fixing y_0. So either $\alpha_h = 0$ or $h y_0 = y_0$ and hence $\pi(gy) = y_0$. However gy and y_0 have the same norm (induced by the inner product) since g is an element of the orthogonal group on V; thus $gy = y_0$. So $gx = x_0$ and $gy = y_0$.

The proof of the attainment of the lower bound is similar. $\square$ We remark that it is not known whether the condition x_0 or $y_0 \in \mathrm{Int}_W F$ can be removed. For the semisimple Lie algebra case [7], no such assumption is made (same for compact connected Lie groups) and is mainly due to the richer structure of the Lie framework.

Applying Theorem 1.5 on the reductive Lie algebras $\mathfrak{gl}_n(\mathbb{F})$, $\mathbb{F} = \mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$ respectively, we have the following result.

COROLLARY 1.6. (Fan [2], Theobald [15]) *Let A and B be real symmetric (Hermitian, quaternionic Hermitian) matrices with eigenvalues $\alpha_1 \geq \cdots \geq \alpha_n$ and $\beta_1 \geq \cdots \geq \beta_n$, respectively. The set*

$$\{trUAU^*B : U \in SU_n(\mathbb{F})\}$$

is the interval $[a,b]$ *where*

$$a = \sum_{i=1}^{n} \alpha_i \beta_{n-i+1},$$
$$b = \sum_{i=1}^{n} \alpha_i \beta_i.$$

Proof. Just take $F = \{\mathrm{diag}(a_1,\ldots,a_n) : a_1 \geq \cdots \geq a_n\}$ and H is then the symmetric group S_n. The longest element sends the elements $\mathrm{diag}(a_1,\ldots,a_n)$ in F to its reversal $\mathrm{diag}(a_n,\ldots,a_1)$. □

Let $I_{n,n} = (-I_n) \oplus I_n$. The group $G = SO(n,n)$ is the group of matrices in $SL(2n,\mathbb{R})$ which leaves invariant the quadratic form $-x_1^2 - \cdots - x_n^2 + x_{n+1}^2 + \cdots + x_{2n}^2$. In other words, $SO(n,n) = \{A \in SL(2n,\mathbb{R}) : A^T I_{n,n} A = I_{n,n}\}$. The group $SO(n,n)$ has two components (usually $SO_0(n,n)$ refers to the identity component) and hence is not connected. It is also noncompact. It is well known that [5]

$$\begin{aligned}
\mathfrak{so}_{n,n} &= \{\begin{pmatrix} X_1 & Y \\ Y^T & X_2 \end{pmatrix} : X_1^T = -X_1,\ X_2^T = X_2,\ Y \in \mathbb{R}_{n\times n}\},\\
K &= SO(n) \times SO(n),\\
\mathfrak{k} &= \mathfrak{so}(n) \oplus \mathfrak{so}(n),\ i.e.,\ Y = 0,\\
\mathfrak{p} &= \{\begin{pmatrix} 0 & Y \\ Y^T & 0 \end{pmatrix}, Y \in \mathbb{R}_{n\times n}\},\ i.e.,\ X_1 = X_2 = 0,\\
\mathfrak{a} &= \oplus_{1\leq j\leq n} \mathbb{R}(E_{j,n+j} + E_{n+j,j}),
\end{aligned}$$

where $E_{i,j}$ is the $2n \times 2n$ matrix and 1 at the (i,j) position is the only nonzero entry. The Killing form is

$$B(\begin{pmatrix} 0 & X \\ X^T & 0 \end{pmatrix}, \begin{pmatrix} 0 & Y \\ Y^T & 0 \end{pmatrix}) = 4(n-1)\mathrm{tr}\, XY^T.$$

Now the adjoint action of K on $\mathfrak{p}$ is given by

$$\begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix}^T \begin{pmatrix} 0 & S \\ S^T & 0 \end{pmatrix} \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix} = \begin{pmatrix} 0 & U^T S V \\ V^T S^T U & 0 \end{pmatrix}$$

where $U, V \in SO(n)$. We will identify $\mathfrak{p}$ with $\mathbb{R}_{n\times n}$ and thus $\mathfrak{a}$ will then be identified with real diagonal matrices. We may choose $\mathfrak{a}_+ = \{\mathrm{diag}(a_1,\ldots,a_n) : a_1 \geq \cdots \geq a_{n-1} \geq |a_n|\}$. The action of K on $\mathfrak{p}$ is then orthogonal equivalence, i.e., $H \mapsto UHV$ where $U, V \in SO(n)$. The action of the Weyl group W on $\mathfrak{a}$ is given by $\mathrm{diag}(d_1,\ldots,d_n) \mapsto \mathrm{diag}(\pm d_{\sigma(1)},\ldots,\pm d_{\sigma(n)})$, $\mathrm{diag}(d_1,\ldots,d_n) \in \mathfrak{a}$, $\sigma \in S_n$ (the symmetric group) and the number of negative signs is even. The longest element ω sends $\mathrm{diag}(a_1,\ldots,a_n) \in \mathfrak{a}_+$ to

$$\omega a = \begin{cases} \mathrm{diag}(-a_1,\ldots,-a_{n-1},a_n) & \text{if } n \text{ is odd} \\ \mathrm{diag}(-a_1,\ldots,-a_n) & \text{if } n \text{ is even.} \end{cases}$$

Applying Theorem 1.5 on the simple Lie algebra $\mathfrak{so}_{n,n}$, we have the following result.

COROLLARY 1.7. (Miranda-Thompson [9], Tam [13]) *Let A and B be $n \times n$ real matrices with singular values $\alpha_1 \geq \cdots \geq \alpha_n \geq 0$, $\beta_1 \geq \cdots \geq \beta_n \geq 0$ respectively, the set*

$$\{trUAVB : U, V \in SO(n)\}$$

is the interval $[a, b]$, where

$$a = \begin{cases} -\sum_{i=1}^{n-1} s_i + sign\ \det(AB)s_n & \text{if } n \text{ is odd} \\ -\sum_{i=1}^{n-1} s_i - sign\ \det(AB)s_n & \text{if } n \text{ is even,} \end{cases}$$

$$b = \sum_{i=1}^{n-1} s_i + sign\ \det(AB)s_n,$$

where $s_i = \alpha_i \beta_i$, $i = 1, \ldots, n$.

Proof. Recall that $\mathfrak{a}_+ = \{\operatorname{diag}(a_1, \ldots, a_n) : a_1 \geq \cdots \geq a_{n-1} \geq |a_n| \geq 0\}$. Any real $n \times n$ matrix A is special orthogonally similar to $\operatorname{diag}(a_1, \ldots, a_{n-1}, [\text{sign det } A]a_n)$ in $\mathfrak{a}_+$ where $a_1 \geq \cdots \geq a_n \geq 0$ are the singular values of A. □

2. Applications to least squares approximations with orbital constraint. In [1] Chu and Driessel considered the following two least squares problems with spectral constraints. Let $S(n)$ be the space of $n \times n$ real symmetric matrices. Given $H \in S(n)$, the *isospectral surface* $O(H)$ is defined to be the set $O(H) = \{QHQ^T : Q \in O(n)\}$. It is simply the manifold of all real symmetric matrices which are co-spectral with H, by the well-known spectral theorem.

Problem 1. Given $X \in S(n)$. Find the minimizer $L \in O(H)$ for

$$\min_{L \in O(H)} \|L - X\|^2,$$

where the norm is the Frobenius matrix norm, i.e., $\|Y\|^2 = \operatorname{tr} YY^*$.

Given $H \in \mathbb{R}_{p\times q}$, set $O'(H) = \{UHV : U \in O(p), V \in O(q)\}$. So $O'(H)$ is simply the manifold of all real matrices which have the same singular values of H, by the singular value decomposition.

Problem 2. Given $X \in \mathbb{R}_{p\times q}$. Find the minimizer $L \in O'(H)$ of

$$\min_{L \in O'(H)} \|L - X\|^2,$$

where the norm is also the Frobenius matrix norm.

The minimizer $L \in O(H)$ provides the shortest distance between the given $X \in S(n)$ and the isospectral surface $O(H)$. Notice that

$$\|L - X\|^2 = (L - X, L - X) = (L, L) + (X, X) - 2(X, L),$$

where $(X, L) := \operatorname{tr} XL$, $X, L \in S(n)$. Since $(L, L) = (H, H)$ for all $L \in O(H)$ and (X, X) are constants, finding the minimizer of Problem 1 is equivalent to finding the maximizer $L \in O(H)$ for (X,L). We can assume that $\operatorname{tr} H = \operatorname{tr} X = 0$. It is because that

$$(X,L) = (X' + \frac{1}{n}(\operatorname{tr} X)I_n, L' + \frac{1}{n}(\operatorname{tr} L)I_n) = (X',L') + \frac{1}{n}\operatorname{tr} X \operatorname{tr} H,$$

where $X' = X - \frac{1}{n}(\operatorname{tr} X)I_n$ and $L' = L - \frac{1}{n}(\operatorname{tr} L)I_n$. So the simple Lie algebra $\mathfrak{sl}_n(\mathbb{R})$ with Cartan decomposition $\mathfrak{sl}_n(\mathbb{R}) = \mathfrak{so}(n) + \mathfrak{p}$ comes into the scene. Though we have $SO(n)$ instead of $O(n)$, the orbits of H are the same.

THEOREM 2.1. *Let $\mathfrak{g}$ be a real semisimple Lie algebra with Cartan decomposition $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ where the analytic group of $\mathfrak{k}$ is $K \subset G$. For $x \in \mathfrak{p}$, let x_0 denote the unique element of the singleton set $Ad(K)x \cap \mathfrak{a}_+$ where $\mathfrak{a}_+$ is a closed fundamental Weyl chamber. Given $x, y \in \mathfrak{p}$, if $z \in Ad(K)y$, then $\|x_0 - y_0\| \le \|x - z\| \le \|x_0 + (-y_0)_0\|$ and $\omega y_0 = -(-y_0)_0$, where $\|\cdot\|$ is the norm induced by the Killing form $B(\cdot,\cdot)$ and ω is the longest element of the Weyl group of $(\mathfrak{g},\mathfrak{a})$. The lower (upper) bound is achieved if and only if there is $k \in K$ such that both $Ad(k)x$ and $Ad(k)z$ are in $\mathfrak{a}_+$ ($Ad(k)x \in \mathfrak{a}_+$ but $Ad(k)z \in -\mathfrak{a}_+$).*

Proof. Just notice that

$$\|x-z\|^2 = B(x-z,x-z) = B(x,x)+B(z,z)-2B(x,z) = B(x,x)+B(y,y)-2B(x,z),$$

where $B(x,x)$ and $B(y,y)$ are constants. Now notice that $y_0 = z_0$ and then apply Theorem 1.1. ◻

Applying Theorem 2.1 on $\mathfrak{sl}_n(\mathbb{R})$ we have the following solution to Problem 1 (Chu and Driessel made the assumption that the eigenvalues of x and y are distinct, but it is not necessary. That x has distinct eigenvalues means that x_0 is in $\operatorname{Int}_{\mathfrak{a}}\mathfrak{a}_+$ (the relative interior of $\mathfrak{a}_+$ in $\mathfrak{a}$) and x is called a *regular* element in Lie theoretic language. See [1, Theorem 4.1]).

COROLLARY 2.2. *Let X, H be $n \times n$ real symmetric matrices with eigenvalues $x_1 \ge \cdots \ge x_n$ and $h_1 \ge \cdots \ge h_n$ respectively. Then for any $Q \in SO(n)$*

$$\sum_{i=1}^n (x_i - h_i)^2 \le \|X - QHQ^T\|^2 \le \sum_{i=1}^n (x_i - h_{n-i+1})^2,$$

where $\|X\|^2 = trXX^T$. The lower (upper) bound is achieved if and only if there is a $Q \in SO(n)$ such that $QXQ^T = \operatorname{diag}(x_1,\dots,x_n)$ and $QHQ^T = \operatorname{diag}(h_1,\dots,h_n)$ ($QHQ^T = \operatorname{diag}(h_n,\dots,h_1)$).

One can deduce the solution for Problem 2 (Chu and Driessel [1] also made the assumption that the singular values of x and y are distinct, but again it is not necessary. See [1, Theorem 5.1]) by applying Theorem 2.1 on $\mathfrak{so}_{p,q}$. However, special care needs to be taken when $p = q$. Let us proceed for the case $p = q = n$.

We have the following refinement of the result of Chu and Driessel by applying Theorem 2.1 on $\mathfrak{so}_{n,n}$.

COROLLARY 2.3. *Let X, H be $n \times n$ real matrices with singular values $x_1 \ge \cdots \ge x_n \ge 0$ and $h_1 \ge \cdots \ge h_n \ge 0$ respectively. Then for any $U, V \in SO(n)$*

$$\sum_{i=1}^{n-1} (x_i - h_i)^2 + ([\mathit{sign}\,\det X]x_n - [\mathit{sign}\,\det H]h_n)^2$$

$$\leq \|X - UHV\|^2$$
$$\leq \begin{cases} \sum_{i=1}^{n-1}(x_i + h_i)^2 + ([sign\det X]x_n - [sign\det H]h_n)^2 & \text{if } n \text{ is odd} \\ \sum_{i=1}^{n-1}(x_i + h_i)^2 + ([sign\det X]x_n + [sign\det H]h_n)^2 & \text{if } n \text{ is even,} \end{cases}$$

where $\|X\|^2 = trXX^T$. *The lower (upper) bound is achieved if and only if there are* $U, V \in SO(n)$ *such that*

$$UXV = \mathrm{diag}(x_1, \ldots, x_{n-1}, [sign\det X]x_n)$$

and

$$UHV = \mathrm{diag}(h_1, \ldots, h_{n-1}, [sign\det H]h_n)$$

$$\left(UHV = \begin{cases} \mathrm{diag}(-h_1, \ldots, -h_{n-1}, [sign\det H]h_n) & \text{if } n \text{ is odd} \\ \mathrm{diag}(-h_1, \ldots, -h_{n-1}, -[sign\det H]h_n) & \text{if } n \text{ is even.} \end{cases}\right)$$

COROLLARY 2.4. *Let* X, H *be* $n \times n$ *real matrices with singular values* $x_1 \geq \cdots \geq x_n \geq 0$ *and* $h_1 \geq \cdots \geq h_n \geq 0$ *respectively. Then for any* $U, V \in O(n)$

$$\sum_{i=1}^{n}(x_i - h_i)^2 \leq \|X - UHV\|^2 \leq \sum_{i=1}^{n}(x_i + h_i)^2,$$

where $\|X\|^2 = trXX^T$. *The lower (upper) bound is achieved if and only if there are* $U, V \in O(n)$ *such that* $UXV = \mathrm{diag}(x_1, \ldots, x_n)$ *and* $UHV = \mathrm{diag}(h_1, \ldots, h_n)$ *(*$UHV = -\mathrm{diag}(h_1, \ldots, h_n)$*).*

Proof. Notice that we can assume that H and X are of nonnegative determinants since $L \in O'(H)$ and $\|\cdot\|$ is invariant under orthogonally similarity. Let $K = SO(n) \times SO(n)$ be the group in the discussion of $\mathfrak{so}_{n,n}$. Now $O'(H) = \mathrm{Ad}\,(K)H \cup \mathrm{Ad}\,(K)(DH)$ where $D = \mathrm{diag}(-1, 1, \ldots, 1)$. Notice that $\mathrm{Ad}\,(K)$ and $\mathrm{Ad}\,(K)(DH)$ are disjoint if and only if $\det H \neq 0$. So the problem is reduced to the optimization of $\|X - L\|$ where $L \in \mathrm{Ad}\,(K)H \cup \mathrm{Ad}\,(K)(DH)$ with $\det X \geq 0$ and $\det H \geq 0$. Applying Theorem 2.1 yields

$$\min_{L \in Ad\,(K)H} \|X - L\|^2 = \sum_{i=1}^{n}(x_i - h_i)^2,$$
$$\min_{L \in Ad\,(K)(DH)} \|X - L\|^2 = \sum_{i=1}^{n-1}(x_i - h_i)^2 + (x_n + h_n)^2$$

and

$$\max_{L \in Ad\,(K)H} \|X - L\|^2 = \begin{cases} \sum_{i=1}^{n-1}(x_i + h_i)^2 + (x_n - h_n)^2 & \text{if } n \text{ is odd} \\ \sum_{i=1}^{n}(x_i + h_i)^2 & \text{if } n \text{ is even,} \end{cases}$$

and

$$\max_{L \in Ad\,(K)(DH)} \|X - L\|^2 = \begin{cases} \sum_{i=1}^{n}(x_i + h_i)^2 & \text{if } n \text{ is odd} \\ \sum_{i=1}^{n}(x_i + h_i)^2 + (x_n - h_n)^2 & \text{if } n \text{ is even.} \end{cases}$$

The minimizers L are in $\mathrm{Ad}\,(K)H$ since

$$\min_{L\in O'(H)} \|X-L\|^2 = \min\{\min_{L\in Ad\,(K)H}\|X-L\|^2, \min_{L\in Ad\,(K)(DH)}\|X-L\|^2\} = \sum_{i=1}^n (x_i-h_i)^2.$$

The maximizers are either in $\mathrm{Ad}\,(K)H$ or $\mathrm{Ad}\,(K)(DH)$, depending n is even or odd. Nevertheless,

$$\max_{L\in O'(H)} \|X-L\|^2 = \max\{\max_{L\in Ad\,(K)H}\|X-L\|^2, \max_{L\in Ad\,(K)(DH)}\|X-L\|^2\} = \sum_{i=1}^n (x_i+h_i)^2.$$

□

We remark that when $p \neq q$, say $p < q$, then the action of the Weyl group on $\mathfrak{a}$ is given by $\mathrm{diag}(d_1, \ldots, d_p) \mapsto \mathrm{diag}(\pm d_{\sigma(1)}, \ldots, \pm d_{\sigma(p)})$, $\mathrm{diag}(d_1, \ldots, d_p) \in \mathfrak{a}$, $\sigma \in S_p$ (the symmetric group) and there is no restriction on the signs, under the appropriate identifications. Now $\mathfrak{a}_+ = \{\mathrm{diag}(a_1, \ldots, a_n) : a_1 \geq \cdots \geq a_n \geq 0\}$. The orbit $\mathrm{Ad}\,(K)H$ is equal to $O'(H)$. Thus we can apply Theorem 2.1 directly to arrive at the results of Chu and Driessel when $p \neq q$, i.e., Corollary 2.4 is true for real $p \times q$ matrices. When one considers $\mathfrak{su}_{p,q}$, the complex case will then be obtained and the treatment of the $p = q$ case is the same as $p \neq q$ case. Needless to say, application of Theorem 2.1 on various real simple Lie algebras yields different results, e.g., real (complex) skew symmetric matrices, complex symmetric matrices, etc.

The setting for Eaton triple with reduced triple is similar and the proof of the following is omitted.

THEOREM 2.5. *Let (V, G, F) be an Eaton triple with a reduced triple (W, H, F). Given $x, y \in V$, if $z \in Gy$, then $\|x_0 - y_0\| \leq \|x - y\| \leq \|x_0 + (-y_0)_0\|$ where $\|\cdot\|$ is induced by the inner product. If, in addition, x_0 or $y_0 \in Int_W F$ which denotes the relative interior of F in W, then the lower (upper) bound is achieved if and only if there exists $g \in G$ such that both gx and gy are in F ($gx \in F$ and $gy \in -F$).*

We remark that the inequality $\|x_0 - y_0\| \leq \|x - y\|$ is known and is true for an Eaton triple without reduced triple [12, p.85].

3. Remarks. In [8], 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 $\gamma : V \to V$ 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 $(x, y) \leq (\gamma x, \gamma y)$ with equality if and only if $x = g\gamma(x)$ and $y = g\gamma(y)$ for some $g \in G$.

By Theorem 2.4 (the proof does not use equality attainment property) of [8], 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 $\gamma : V \to F$ satisfying the above four conditions [11, p.14] except the equality case. The equality case holds if we assume that x_0 or y_0 are in $\mathrm{Int}_W F$ by Theorem 1.5.

The reduced triple is almost identical to the normal decomposition (sub)system [8, 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 the consideration of optimization and programming.

The condition for attainment in the upper bound in Theorem 1.5 is proved by using the finite reflection group structure which is an algebraic approach via a result of Niezgoda [11]. In contrast, the corresponding result [7, Theorem 3.2] of Lewis for Cartan subspace $\mathfrak{p}$ has an analytic proof (without the restriction that x_0 or $y_0 \in \mathrm{Int}_{\mathfrak{a}} \mathfrak{a}_+$ for the equality attainment).

Acknowledgement. The author gives thanks to an anonymous referee for pointing out that the assumption that x or $y \in \mathrm{Int}_W F$ needs to be made in Theorem 1.5 and for bringing [12] to the author's attention. He also thanks Professor A.S. Lewis for pointing out the relationship between Eaton triple and normal decomposition system. He thanks Professor M.T. Chu for the reprint of the reference [1].

REFERENCES

[1] M.T. Chu and K.R. Driessel. The projected gradient method for least squares matrix approximations with spectral constraints. *SIAM J. Numer. Anal.*, 27:1050-1060, 1990

[2] K. Fan. On a theorem of Weyl concerning eigenvalues of linear transformations. *Proc. Nat. Acad. Sci. U.S.A.*, 35:652-655, 1949.

[3] 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*, to appear.

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

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

[6] B. Kostant. On convexity, the Weyl group and Iwasawa decomposition. *Ann. Sci. École Norm. Sup.*, (4) 6:413-460, 1973.

[7] A.S. Lewis. Convex analysis on Cartan subspaces. *Nonlinear Anal.*, to appear.

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

[9] H. Miranda and R.C. Thompson. A supplement to the von Neumann trace inequality for singular values. *Linear Algebra Appl.*, 248: 61-66, 1994.

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

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

[12] M. Niezgoda. On Schur-Ostrowski type theorems for group majorizations. *J. Convex Anal.*, 5:81-105, 1998.

[13] T.Y. Tam. Miranda-Thompson's trace inequality and a log convexity result. *Linear Algebra Appl.*, 272:91-101, 1997.

[14] T.Y. Tam. Group majorization, Eaton triples and numerical range, *Linear and Multilinear Algebra*, to appear.

[15] C.M. Theobald. An inequality for the trace of the product of two symmetric matrices. *Math. Proc. Cambridge Philos. Soc.* 77:265, 1975.