NORTH-HOLLAND

Eigenvalue-Constrained Faces

A. S. Lewis
Department of Combinatorics and Optimization
University of Waterloo
Waterloo, Ontario, Canada N2L 3G1

Submitted by Richard A. Brualdi

ABSTRACT

We characterize the exposed faces of convex sets $\mathscr{C}$ of symmetric matrices, invariant under orthogonal similarity ($U^T \mathscr{C} U = \mathscr{C}$ for all orthogonal U). Such sets $\mathscr{C}$ are exactly those determined by eigenvalue constraints: typical examples are the positive semidefinite cone and the unit balls of common matrix norms. The set $\mathscr{D}$ of all diagonal matrices in $\mathscr{C}$ is known to be convex if and only if $\mathscr{C}$ is, and $\mathscr{D}$ is invariant under the group of permutations (acting on diagonal entries). We show how any exposed face of $\mathscr{C}$ is naturally associated with an exposed face of $\mathscr{D}$, by relating the stabilizer groups of the two faces. © 1998 Elsevier Science Inc.

1. INTRODUCTION

The beautiful facial structure of the cone of real symmetric $n \times n$ positive semidefinite matrices has been well understood for many years [19, 2]. This structure is strikingly analogous to the facial structure of the positive orthant in $\mathbf{R}^n$. This paper aims to explain the foundations of this analogy, and thereby to understand the exposed faces of general convex sets of matrices satisfying eigenvalue constraints.

* Research partially supported by the Natural Sciences and Engineering Research Council of Canada. E-mail: `aslewis@orion.uwaterloo.ca`; WWW: `http://orion.uwaterloo.ca/~aslewis`.

LINEAR ALGEBRA AND ITS APPLICATIONS 269:159–181 (1998)

655 Avenue of the Americas, New York, NY 10010

0024-3795/98/$19.00
PII S0024-3795(97)00063-3

Faces of the positive orthant in $\mathbf{R}^n$ play a crucial role in linear programming. Analogously, in the newer area of semidefinite programming, a number of recent papers have examined the role of the facial structure of the semidefinite cone (for example, [1, 13, 14]). We hope to clarify the analogies, both between linear and semidefinite programming, and in more general eigenvalue optimization problems.

Let $\mathscr{S}(n)$ denote the space of real symmetric $n \times n$ matrices, and let $\mathscr{S}(n)_+$ denote the positive semidefinite cone. A typical (exposed) face of this cone is

$$\mathscr{F} = \left\{ \begin{pmatrix} W & 0 \\ 0 & 0 \end{pmatrix} \middle| W \in \mathscr{S}(m)_+ \right\},$$

where $0 \leqslant m \leqslant n$. In fact, these faces, together with their rotations $V^T\mathscr{F}V$ (for orthogonal V), comprise all the nonempty faces of the cone.

The Euclidean space $\mathbf{R}^n$ (whose elements we always regard as column vectors) behaves very similarly. A typical exposed face of the positive orthant, $\mathbf{R}^n_+$, is

$$E = \left\{ \begin{pmatrix} w \\ 0 \end{pmatrix} \middle| w \in \mathbf{R}^m_+ \right\},$$

where $0 \leqslant m \leqslant n$. These faces, together with their permutations PE (for permutation matrices P), comprise all the nonempty faces of $\mathbf{R}^n_+$.

The diagonal matrices in $\mathscr{S}(n)_+$ are naturally identified with $\mathbf{R}^n_+$. More generally, the diagonal matrices in the face $\mathscr{F}$ are identified with the face E. Furthermore, the stabilizers of the two faces are related. Specifically, the orthogonal V for which $V^T\mathscr{F}V = \mathscr{F}$ are those with the block structure

$$V = \begin{pmatrix} Q & 0 \\ 0 & R \end{pmatrix}, \qquad \text{where } Q \text{ is } m \times m;$$

likewise, the permutation matrices P for which $PE = E$ are those with the same block structure.

Knowing the relationship between the stabilizers of the two faces, and between E and the diagonal matrices in $\mathscr{F}$, would enable us to "compute" $\mathscr{F}$ from E. Hence we could describe all the faces of $\mathscr{S}(n)_+$ in terms of those of $\mathbf{R}^n_+$. It is this description we wish to generalize.

Define the *eigenvalue* map $\lambda : \mathscr{S}(n) \to \mathbf{R}^n$ by letting $\lambda_i(X)$ be the ith largest eigenvalue of the symmetric matrix X (counted by multiplicity). It is easy to see that a subset $\mathscr{K}$ of $\mathscr{S}(n)$ is invariant under orthogonal similarity

($U^T\mathscr{H}U = \mathscr{H}$ for all orthogonal U) if and only if there s a subset C of $\mathbf{R}^n$ which is permutation-invariant ($PC = C$ for all permutation matrices P) satisfying

$$\mathscr{H} = \lambda^{-1}(C) = \{X \in \mathscr{S}(n) \mid \lambda(X) \in C\}.$$

The condition that C is permutation-invariant is in some sense superfluous for this observation, but is crucial for our later development. For example, with this assumption the matrix set $\mathscr{H}$ is closed and convex if and only if C is closed and convex (see [7, 10]). If we define the *diagonal* map $\mathrm{Diag}: \mathbf{R}^n \to \mathscr{S}(n)$ by letting $\mathrm{Diag}\, x$ be a diagonal matrix with diagonal entries $x_1, x_2, \dots, x_n$, then the set of diagonal matrices in $\mathscr{H}$ is just $\mathrm{Diag}\, C$. Our main result describes the exposed faces of $\mathscr{H}$ in terms of those of C.

Let E be a proper, exposed face of the closed, convex, permutation-invariant set C: in other words, E is the intersection of C with a supporting hyperplane. The *stabilizer* of E is the group of permutation matrices satisfying $PE = E$: we show this subgroup consists of those P with a certain block-diagonal structure associated with the face E (after a suitable reordering of the basis). Now consider the group of orthogonal matrices with the same block-diagonal structure: we denote this group by $\mathscr{O}(n)_{\sim_E}$. Our main result (Theorem 5.1) is that the matrix set

$$\mathscr{F} = \left\{U^T(\mathrm{Diag}\, x)U \middle| U \in \mathscr{O}(n)_{\sim_E},\ x \in E\right\}$$

is an exposed face of $\lambda^{-1}(C)$, as is any rotation $V^T\mathscr{F}V$ (for orthogonal V); furthermore, every exposed face of $\lambda^{-1}(C)$ may be constructed in this manner from some exposed face of C. The case $C = \mathbf{R}^n_+$ gives the example of the semidefinite cone. Other interesting examples arise from choosing C to be the l_1 or l_∞ unit ball: Section 6 has the details.

We develop the theory for real symmetric matrices: the Hermitian versions of our results are entirely analogous, in the usual way.

We might expect an analogous set of results for the unit balls of unitarily invariant matrix norms: an extensive study of the facial structure of such balls may be found in [4, 3, 5]. In the above discussion, we replace the space $\mathscr{S}(n)$ by the space of all complex $n \times n$ (or more generally $n \times m$) matrices, the the map λ is replaced by the analogous *singular-value* map (cf. [8]). Orthogonal similarity transformations are replaced by transformations $X \mapsto UXV$ with unitary U and V, and the group of permutations is enlarged to allow coordinate sign changes. More generally, we might expect the present results to extend to the Lie algebraic framework described in [9]. We defer further discussion.

In the present framework, we might also expect a completely analogous set of results for faces, rather than exposed faces. Our wide application of duality arguments apparently facilitates the exposed case. Again, we defer further discussion.

2. INVARIANT CONVEX SETS AND EXPOSED FACES

This section concerns the facial structure of a nonempty, convex subset C of a Euclidean space E. We consider the implications of the invariance of C under a compact subgroup G of the general linear group $\mathrm{GL}(E)$.

The *stabilizer* of C in G is the subgroup $G_C = \{g \in G \mid gC = C\}$. It is easy to check that the stabilizer of any closed set is closed. We say C is *invariant under* G when $G_C = G$. For a point x in E, we write G_x for $G_{\{x\}}$. The following slight refinement of a well-known result is fundamental for us.

THEOREM 2.1 (Fixed point). *If a nonempty, convex set C is invariant under a compact group G, then so is its relative interior* $\mathrm{ri}\, C$, *which therefore contains a point which is invariant under G.*

Proof. For any transformation g in G, [15, Theorem 6.6] implies $g\, \mathrm{ri}\, C = \mathrm{ri}\, gC = \mathrm{ri}\, C$, as required. The result follows by applying the fixed point theorem in [12, p. 130]. ∎

DEFINITION 2.2. A convex set C is *relatively invariant* under a group G if every group element fixing a point in $\mathrm{ri}\, C$ leaves C invariant: $G_x \subset G_C$ for all x in $\mathrm{ri}\, C$.

PROPOSITION 2.3. *A closed, convex set C is invariant under a linear transformation g (that is, $gC = C$) if and only if its relative interior is invariant under g.*

Proof. If C is invariant, then, as before, $g\, \mathrm{ri}\, C = \mathrm{ri}\, gC = \mathrm{ri}\, C$. If, on the other hand, $\mathrm{ri}\, C$ is invariant, then since $\mathrm{cl}\, \mathrm{ri}\, C = C$ by [15, Theorem 6.3],

$$gC = g\, \mathrm{cl}\, \mathrm{ri}\, C \subset \mathrm{cl}(g\, \mathrm{ri}\, C) = \mathrm{cl}\, \mathrm{ri}\, C = C,$$

by continuity.

It remains to show $gC \supset C$. Fix a point x in ri C (so gx also lies in ri C), and let L be the affine span of C (or, equivalently, of ri C). Notice that

$$gL = g[x + \mathbf{R}_+(\operatorname{ri} C - x)] = gx + \mathbf{R}_+(g \operatorname{ri} C - gx)$$
$$= gx + \mathbf{R}_+(\operatorname{ri} C - gx) = L.$$

Thus the restricted map $g|_L : L \to L$ is invertible. Since $g|_L \operatorname{int}_L C = \operatorname{int}_L C$, we deduce

$$C = \operatorname{cl} \operatorname{int}_L C = \operatorname{cl} g|_L^{-1} \operatorname{int}_L C \supset g|_L^{-1} \operatorname{cl} \operatorname{int}_L C = g|_L^{-1} C,$$

by continuity. Thus $gC = g|_L C \supset C$, as required. ■

COROLLARY 2.4. *A closed, convex set C is relatively invariant under a group G if and only if its relative interior is relatively invariant.*

Proof. By the previous proposition, C and ri C have the same stabilizer, and the result follows, since ri ri C = ri C.

A *face* of C is a convex subset $F \subset C$ such that points x and y in C must lie in F whenever $\lambda x + (1 - \lambda)y$ lies in F for some real λ in $(0, 1)$. In this paper our primary interest is a special class of faces—exposed faces. The *indicator function* of C is defined by

$$\delta_C(x) = \begin{cases} 0 & \text{if } \; x \in C, \\ +\infty & \text{otherwise,} \end{cases}$$

and the *support function* is defined by $\delta_C^*(y) = \sup_{x \in C} \langle x, y \rangle$ for any element y of E (where $\langle \cdot, \cdot \rangle$ denotes the inner product for E). An *exposed face* of C is a set of the form

$$F = \{x \in C \,|\, \langle x, y \rangle = \delta_C^*(y)\} \tag{2.5}$$

for some y in E. (Any exposed face is a face.) If C is closed then $F = \partial \delta_C^*(y)$, where ∂ denotes the convex subdifferential (see [15]). Notice that the empty set is allowed: in fact it is not difficult to see that the empty set is an exposed face of a nonempty closed convex set C if and only if C is unbounded.

For results analogous to the following for cones, see [17, Theorem 3.3(a)] and [18, Lemma 3.7].

THEOREM 2.6 (Equivalent faces). *Suppose a convex set C is invariant under a linear transformation g, and consider a nonempty (exposed) face F of C. Then gF is also a(n) (exposed) face of C, and the following properties are equivalent:*

(i) $gF = F$;
(ii) $gF \subset F$;
(iii) $F \cap g \operatorname{ri} F \neq \emptyset$.

Proof. By restricting to the subspace spanned by C we may without loss of generality assume g is invertible. To see gF is a face, suppose $x \in F$, $y, z \in C$, and $gx = \frac{1}{2}(y + z)$. Because $x = \frac{1}{2}(g^{-1}y + g^{-1}z)$ and F is a face we deduce $g^{-1}y$ and $g^{-1}z$ belong to F so y and z lie in gF, as required.

If F is an exposed face, we can describe it by Equation (2.5), and then it is easy to check

$$gF = \{x \in C | \langle x, u \rangle = \delta_C^*(u)\},$$

where $u = (g^{-1})^*(y)$ (* denoting the adjoint). Thus gF is also an exposed face.

The implication (i) $\Rightarrow$ (iii) is immediate. To see (iii) $\Rightarrow$ (ii), suppose the point x lies in ri F and satisfies $gx \in F$. Noting $gx \in \operatorname{ri} gF$, we see that for any point y in gF there is a real $\epsilon > 0$ with $gx + \epsilon(gx - y)$ in gF. Since gx belongs to the face F, so must y.

Finally, to prove (ii) $\Rightarrow$ (i), note that gF and F are faces of the same dimension (since g is invertible), so the result follows (cf. [15, Theorem 18.1]). ∎

For any compact subgroup G of GL(E), there is an inner product $(\cdot, \cdot)$ which is G-invariant: $(gx, gy) = (x, y)$ for all elements x and y of E and elements g of G (see [12, p. 131]). With respect to this inner product, G is a subgroup of the orthogonal group $O(E)$: we thus lose no essential generality in always assuming this.

COROLLARY 2.7 (Stabilizers). *If a convex set C is invariant under a group G, then any face F of C is relatively invariant. If furthermore G is compact and F is nonempty and closed, then there is a point x in* ri F *satisfying $G_x = G_F$.*

Proof. The first part follows from the equivalent faces theorem above. By the fixed point theorem 2.1 applied to the face F and the (compact) group G_F, there is a point x in ri F with $(G_F)_x = G_F$, whence $G_x \supset G_F$. But the equivalent faces theorem shows $G_x \subset G_F$. ∎

The next result, an apparently rather innocuous refinement of the exposed faces definition (2.5), is central. Just as the stabilizers corollary above shows that any nonempty closed face contains a point whose stabilizer coincides with that of the face, so this result shows that any exposed face can be exposed by a vector whose stabilizer coincides with that of the face.

THEOREM 2.8 (Exposing vectors). *If a closed convex subset C of the Euclidean space E is invariant under a closed subgroup G of the orthogonal group $O(E)$, then we can write any exposed face F of C in the form*

$$F = \{x \in C \,|\, \langle x, y\rangle = \delta_C^*(y)\}, \tag{2.9}$$

for some vector $y \in E$ with the same stabilizer as F (that is, $G_y = G_F$).

Proof. Suppose first that F is nonempty. Let K be the (nonempty) set of exposing vectors for F:

$$K = \{y \in E \,|\, (2.9) \text{ holds}\}.$$

We begin by showing that K is convex. For this, let us assume $0 \in \operatorname{int} C$: the general case follows after a translation and restriction to a subspace.

We claim the set $K_1 = \{y \in K \,|\, \delta_C^*(y) = 1\}$ is convex. To see this, choose elements v and w in K_1, and real α in $(0, 1)$. If a point x lies in F, then $\langle x, v\rangle = 1 = \langle x, w\rangle$, and hence

$$\begin{aligned}\delta_C^*(\alpha v + (1-\alpha)w) &\leqslant \alpha\delta_C^*(v) + (1-\alpha)\delta_C^*(w)\\ &= 1\\ &= \langle x, \alpha v + (1-\alpha)w\rangle\\ &\leqslant \delta_C^*(\alpha v + (1-\alpha)w).\end{aligned}$$

Since F is nonempty, we deduce $\delta_C^*(\alpha v + (1-\alpha)w) = 1$, and

$$F \subset \{x \in C \,|\, \langle x, \alpha v + (1-\alpha)w\rangle = 1\}.$$

To see the reverse inclusion, if x belongs to the right hand side, then

$$1 = \alpha\langle x, v\rangle + (1-\alpha)\langle x, w\rangle \leqslant \alpha\delta_C^*(v) + (1-\alpha)\delta_C^*(w) = 1,$$

whence $\langle x, v\rangle = \delta_C^*(v)$, and so x lies in F. Thus $\alpha v + (1-\alpha)w$ belongs to K_1, as required.

If $F = C$ then $K = \{0\}$. If F is a nonempty, proper, exposed face of C, then $K = \bigcup_{\beta > 0} \beta K_1$, which is convex. In either case, K is convex.

We next observe that K is invariant under the stabilizer G_F: for a linear transformation g in G_F and a vector y in K,

$$\begin{aligned}\{x \in C | \langle x, gy\rangle = \delta_C^*(gy)\} &= \{x \in C | \langle g^T x, y\rangle = \delta_C^*(y)\}\\ &= g\{z \in C | \langle z, y\rangle = \delta_C^*(y)\}\\ &= gF = F,\end{aligned}$$

whence $gy \in K$. By the fixed point theorem 2.1 applied to the convex set K and the group G_F, we can fix an exposing vector $y \in K$ with $(G_F)_y = G_F$. Thus $G_y \supset G_F$. But if g belongs to G_y, then

$$\begin{aligned}F &= \{x \in C | \langle x, gy\rangle = \delta_C^*(y)\}\\ &= \{x \in C | \langle g^T x, y\rangle = \delta_C^*(y)\}\\ &= g\{z \in C | \langle z, y\rangle = \delta_C^*(y)\}\\ &= gF,\end{aligned}$$

and hence g belongs to G_F. Thus $G_y = G_F$.

Finally we return to the case where the exposed face F is empty. If the convex set C is empty, then we can choose $y = 0$. Otherwise, C must be unbounded, and hence its recession cone $(0+)C$ is nontrivial [15, Theorem 8.4]. It is easily checked that $(0+)C$ is invariant under G. The case $C = E$ is easy (again choose $y = 0$), so we can assume C is a proper subset of E, in which case so is $(0+)C$. By the fixed point theorem, there is a point y in $\mathrm{ri}[(0+)C]$ with $G_y = G = G_\varnothing = G_F$, and y must be nonzero. Fix any point x in C. Since $x + \alpha y$ lies in C for all $\alpha \geqslant 0$, we have

$$\delta_C^*(y) \geqslant \sup_{\alpha \geqslant 0} \langle x + \alpha y, y\rangle = +\infty,$$

and so Equation (2.9) holds, as required. ∎

3. PERMUTATIONS AND ORTHOGONAL SIMILARITY

We concentrate on two particular Euclidean spaces: $\mathbf{R}^n$, with the usual inner product, and the space $\mathcal{S}(n)$ of $n \times n$, real symmetric matrices, with the trace inner product $\langle X, Y \rangle = \operatorname{tr} XY$. The diagonal map $\operatorname{Diag} : \mathbf{R}^n \to \mathcal{S}(n)$ gives an isomorphism between $\mathbf{R}^n$ and the subspace of diagonal matrices.

On $\mathbf{R}^n$ we are interested in the group of $n \times n$ permutation matrices $\mathscr{P}(n)$: we consider elements x of $\mathbf{R}^n$ as column vectors, and permutation matrices P as linear transformations, $x \mapsto Px$.

On $\mathcal{S}(n)$ we are interested in the group of orthogonal similarity transformations. Let $\mathscr{O}(n)$ be the group of $n \times n$ orthogonal matrices, and define the *adjoint representation* $\operatorname{Ad} : \mathscr{O}(n) \to O(\mathcal{S}(n))$ by $(\operatorname{Ad} U)X = U^T X U$, for orthogonal U and symmetric X. Then we are interested in the group $\operatorname{Ad} \mathscr{O}(n)$.

The results in the previous section depend heavily on properties of stabilizers, in the relevant group, of certain elements of the underlying Euclidean space. In both of the above cases, these stabilizers are associated with equivalence relations on $\{1, 2, \dots, n\}$. For such a relation $\sim$, let $\mathscr{P}(n)_\sim$ denote the subgroup of matrices representing permutations which leave the equivalence classes of $\sim$ invariant. If we reorder the basis so that the equivalence classes of $\sim$ are blocks of consecutive integers, then $\mathscr{P}(n)_\sim$ consists of the permutation matrices having the corresponding block-diagonal structure. Now let $\mathscr{O}(n)_\sim$ denote the subgroup of orthogonal matrices with the same block-diagonal structure.

More formally, define a subspace $\mathbf{R}^n_\sim = \{x \in \mathbf{R}^n \mid x_i = x_j \text{ if } i \sim j\}$. Then

$$\mathscr{P}(n)_\sim = \left\{P \in \mathscr{P}(n) \middle| Px = x \text{ for all } x \text{ in } \mathbf{R}^n_\sim\right\},$$

$$\mathscr{O}(n)_\sim = \left\{U \in \mathscr{O}(n) \middle| Ux = x \text{ for all } x \text{ in } \mathbf{R}^n_\sim\right\}.$$

For example, if $n = 5$ and the equivalence classes of $\sim$ are the blocks $\{1, 2\}$ and $\{3, 4, 5\}$, then

$$\mathscr{P}(n)_\sim = \left\{\begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix} \middle| R \in \mathscr{P}(2),\ Q \in \mathscr{P}(3)\right\},$$

$$\mathscr{O}(n)_\sim = \left\{\begin{pmatrix} V & 0 \\ 0 & W \end{pmatrix} \middle| V \in \mathscr{O}(2),\ W \in \mathscr{O}(3)\right\}.$$

The following result is clear.

LEMMA 3.1. *For two equivalence relations* $\sim$ *and* $\bowtie$ *on* $\{1, 2, \dots, n\}$, *the following are equivalent*:

(i) $\sim$ *and* $\bowtie$ *are identical*;
(ii) $\mathscr{P}(n)_{\sim} = \mathscr{P}(n)_{\bowtie}$;
(iii) $\mathscr{O}(n)_{\sim} = \mathscr{O}(n)_{\bowtie}$.

We can identify the stabilizer in $\mathscr{P}(n)$ of a vector x in $\mathbf{R}^n$ through an associated equivalence relation $\sim_x$, defined by $i \sim_x j$ if and only if $x_i = x_j$. The next result is easy to check.

PROPOSITION 3.2. *The stabilizer of any vector* x *in* $\mathbf{R}^n$ *satisfies*

$$\mathscr{P}(n)_x = \mathscr{P}(n)_{\sim_x}.$$

We can also express stabilizers of elements of $\mathscr{S}(n)$ in the group $\operatorname{Ad} \mathscr{O}(n)$ easily, with this notation.

PROPOSITION 3.3. *The stabilizer of a diagonal matrix* Diag x (*for a vector* x *in* $\mathbf{R}^n$) *satisfies*

$$[\operatorname{Ad} \mathscr{O}(n)]_{\operatorname{Diag} x} = \operatorname{Ad}[\mathscr{O}(n)_{\sim_x}].$$

Proof. By definition, an orthogonal matrix U has $\operatorname{Ad} U$ in $[\operatorname{Ad} \mathscr{O}(n)]_{\operatorname{Diag} x}$ if and only if $U^T(\operatorname{Diag} x)U = \operatorname{Diag} x$. Suppose that U lies in $\mathscr{O}(n)_{\sim_x}$. Without loss of generality, suppose that the equivalence classes of $\sim_x$ are blocks (by reordering the basis, if necessary). Then U and Diag x both have the corresponding block-diagonal structure, the diagonal blocks of U being orthogonal matrices, and those of Diag x being multiples of identity matrices. It follows that $U^T(\operatorname{Diag} x)U = \operatorname{Diag} x$.

Conversely, if $U^T(\operatorname{Diag} x)U = \operatorname{Diag} x$ for some orthogonal matrix U, then the ith column of U, denoted u^i, is an eigenvector for the matrix Diag x, with corresponding eigenvalue x_i. But for any $j \not\sim_x i$, the standard unit vector e^j is an eigenvector with distinct corresponding eigenvalue x_j. The (j, i) entry in U is $(e^j)^T u^i$, and is hence zero. It follows that U belongs to $\mathscr{O}(n)_{\sim_x}$. ∎

The stabilizer in $\operatorname{Ad} \mathscr{O}(n)$ of a general symmetric matrix X is now easily computed by diagonalizing X.

PROPOSITION 3.4. *Vectors x and y in $\mathbf{R}^n$ satisfy*

$$[\operatorname{Ad} \mathscr{O}(n)]_{\operatorname{Diag} x} = [\operatorname{Ad} \mathscr{O}(n)]_{\operatorname{Diag} y}$$

if and only if $\mathscr{P}(n)_x = \mathscr{P}(n)_y$.

Proof. This follows from Lemma 3.1 and Propositions 3.2 and 3.3. ■

LEMMA 3.5. *Suppose that vectors x, y, and z in $\mathbf{R}^n$ and an orthogonal matrix U satisfy $U^T(\operatorname{Diag} x)U = \operatorname{Diag} x$ and $U^T(\operatorname{Diag} y)U = \operatorname{Diag} z$. Then there is a permutation matrix P satisfying $Px = x$ and $Py = z$.*

Proof. Reorder the basis so that $\operatorname{Diag} x$ is the block-diagonal matrix $\bigoplus_i \alpha_i I_i$, where the scalars α_i are distinct, and I_i is an $r_i \times r_i$ identity matrix for each index i. By Proposition 3.3, there are matrices U_i in $\mathscr{O}(r_i)$ with $U = \bigoplus_i U_i$. If we decompose the vectors y and z into the same size blocks, $y = \bigoplus_i y^i$ and $z = \bigoplus_i z^i$, with y^i and z^i in $\mathbf{R}^{r_i}$ for each i, then by assumption, $U_i^T(\operatorname{Diag} y^i)U_i = \operatorname{Diag} z^i$ for each i. By considering eigenvalues, it is clear that there are permutation matrices P_i with $P_i y^i = z^i$, for each i. The required permutation matrix can then be chosen as $P = \bigoplus_i P_i$. ■

4. A CONVEX CHAIN RULE

Given a matrix X in $\mathscr{S}(n)$, let $\lambda_1(X) \geqslant \lambda_2(X) \geqslant \cdots \geqslant \lambda_n(X)$ denote the eigenvalues of X (counted by multiplicity). In this way we define a function $\lambda : \mathscr{S}(n) \to \mathbf{R}^n$. We say a function $f : \mathbf{R}^n \to (-\infty, +\infty]$ is *permutation-invariant* if $f(Px) = f(x)$ for every point x in $\mathbf{R}^n$ and every permutation matrix P. If such a function is convex, then so is the composite function $f \circ \lambda$ [where $(f \circ \lambda)(X) = f(\lambda(X))$ for matrices X in $\mathscr{S}(n)$], by [11, Theorem 4.3], and, whether or not f is convex, the Fenchel conjugate is given by

$$(f \circ \lambda)^* = f^* \circ \lambda \tag{4.1}$$

(see [10, Theorem 2.3]).

Given a vector x in $\mathbf{R}^n$, let $\bar{x}$ denote the vector with the same components arranged in nonincreasing order.

The following chain rule for subgradients of the function $f \circ \lambda$ is central to our development.

THEOREM 4.2 (Chain rule). *Let the function* $f: \mathbf{R}^n \to (-\infty, +\infty]$ *be permutation-invariant. Then for any vector* y *in* $\mathbf{R}^n$,

$$\partial(f \circ y)(\mathrm{Diag}\, y) = \mathrm{Ad}[\mathscr{O}(n)_{\sim_y}]\, \mathrm{Diag}\, \partial f(y).$$

Proof. By [11, Theorem 4.6], $\partial(f \circ \lambda)(\mathrm{Diag}\, y) = (\mathrm{Ad}\, \mathscr{Q})\, \mathrm{Diag}\, \partial f(\bar{y})$, where the set $\mathscr{Q}$ consists of those orthogonal U satisfying $\mathrm{Diag}\, y = U^T(\mathrm{Diag}\, \bar{y})U$, and [11, Corollary 4.8] shows $\partial f(y) = \mathscr{R}\, \partial f(\bar{y})$, where the set $\mathscr{R}$ consists of those permutations matrices P satisfying $y = P\bar{y}$. Hence

$$\mathrm{Diag}\, \partial f(\bar{y}) = (\mathrm{Ad}\, \mathscr{R})\, \mathrm{Diag}\, \partial f(y),$$

so, by Proposition 3.3, it suffices to prove

$$\mathrm{Ad}(\mathscr{R}\mathscr{Q}) = [\mathrm{Ad}\, \mathscr{O}(n)]_{\mathrm{Diag}\, y}.$$

For matrices U in $\mathscr{Q}$ and P in $\mathscr{R}$,

$$(\mathrm{Ad}\, PU)\, \mathrm{Diag}\, y = (PU)^T(\mathrm{Diag}\, y)(PU) = U^T(\mathrm{Diag}\, \bar{y})U = \mathrm{Diag}\, y,$$

so

$$\mathrm{Ad}(\mathscr{R}\mathscr{Q}) \subset [\mathrm{Ad}\, \mathscr{O}(n)]_{\mathrm{Diag}\, y}.$$

To see the opposite inclusion, suppose an orthogonal matrix U satisfies $U^T(\mathrm{Diag}\, y)U = \mathrm{Diag}\, y$. Fix a matrix P in $\mathscr{R}$, so $\mathrm{Diag}\, y = P(\mathrm{Diag}\, \bar{y})P^T$. Hence $P^T U \in \mathscr{Q}$, so U belongs to $\mathscr{R}\mathscr{Q}$, as required. ■

Our next step is to apply this chain rule to study sets in $\mathbf{R}^n$ via their indicator functions. We say that a subset C of $\mathbf{R}^n$ is *permutation-invariant* if $PC = C$ for every permutation matrix P. Suppose in addition that C is closed and convex. Then the matrix set

$$\lambda^{-1}(C) = \{X \in \mathscr{S}(n) \mid \lambda(X) \in C\}$$

is also closed and convex: to see this, note that its indicator function satisfies

$$\delta_{\lambda^{-1}(C)} = \delta_C \circ \lambda, \tag{4.3}$$

so we can apply the results in [10] (see also [7]). We also see from this equation and the conjugacy formula (4.1) that the support functions satisfy

$$\delta^*_{\lambda^{-1}(C)} = \delta^*_C \circ \lambda. \tag{4.4}$$

By assumption, the set C is invariant under the group $\mathscr{P}(n)$, and it is easy to see that the matrix set $\lambda^{-1}(C)$ is invariant under the group Ad $\mathscr{O}(n)$. Our ultimate aim is to characterize the exposed faces of $\lambda^{-1}(C)$ in terms of those of C.

First we introduce some important notation. For a nonempty, exposed face E of C we know, by the stabilizers corollary 2.7, that there is a point x in ri E with $\mathscr{P}(n)_E = \mathscr{P}(n)_x$. Define an equivalence relation $\sim_E$ on $\{1, 2, \ldots, n\}$ by $\sim_E\ =\ \sim_x$. This relation is well defined, since if the point y satisfies $\mathscr{P}(n)_E = \mathscr{P}(n)_y$, then

$$\mathscr{P}(n)_{\sim_x} = \mathscr{P}(n)_x = \mathscr{P}(n)_E = \mathscr{P}(n)_y = \mathscr{P}(n)_{\sim_y},$$

by Proposition 3.2, and hence $\sim_x\ =\ \sim_y$, by Lemma 3.1. If E is empty, then we define $\sim_E$ by $u \sim_E v$ for all points u and v.

The equivalence relation $\sim_E$ describes the stabilizer of the face E:

$$\mathscr{P}(n)_E = \mathscr{P}(n)_{\sim_E}. \tag{4.5}$$

This is the notation we shall use to characterize the exposed faces of $\lambda^{-1}(C)$. The first step is a direct application of the chain rule (Theorem 4.2).

THEOREM 4.6 (Facial chain rule). *Suppose the subset C of $\mathbf{R}^n$ is closed, convex, and permutation-invariant, with an exposed face E. Then the set* Ad$[\mathscr{O}(n)_{\sim_E}]$ Diag E *is an exposed face of the matrix set $\lambda^{-1}(C)$.*

Proof. When C is empty the result is trivial, so assume C is nonempty. By the exposing vectors theorem 2.8, there is a vector y in $\mathbf{R}^n$ with $E = \partial\delta^*_C(y)$ and $\mathscr{P}(n)_E = \mathscr{P}(n)_y$. Now by the chain rule (Theorem 4.2), Equations (4.4) and (4.5), and Lemma 3.1,

$$\begin{aligned}\partial\delta^*_{\lambda^{-1}(C)}(\operatorname{Diag} y) &= \partial(\delta^*_C \circ \lambda)(\operatorname{Diag} y)\\ &= \operatorname{Ad}\left[\mathscr{O}(n)_{\sim_y}\right] \operatorname{Diag} \partial\delta_C(y)\\ &= \operatorname{Ad}\left[\mathscr{O}(n)_{\sim_E}\right] \operatorname{Diag} E,\end{aligned}$$

so the result follows. ■

In the above result, we see immediately by the equivalent faces theorem 2.6 that any rotation

$$V^T\left(\mathrm{Ad}\left[\mathscr{O}(n)_{\sim_E}\right] \mathrm{Diag}\, E\right)V \qquad \text{(with } V \text{ orthogonal)}$$

is also an exposed face. Our main result, proved in the next section, states that this construction in fact gives *all* the exposed faces of $\lambda^{-1}(C)$.

5. EXPOSED FACES OF MATRIX SETS

We are now ready to prove our main result. Given a closed, convex, permutation-invariant subset C of $\mathbf{R}^n$, we characterize the exposed faces of the matrix set $\lambda^{-1}(C)$ in terms of the exposed faces of C. Just as C is invariant under the permutation group $\mathscr{P}(n)$, so $\lambda^{-1}(C)$ is invariant under the group of orthogonal similarity transformations $\mathrm{Ad}\,\mathscr{O}(n)$. The idea of our proof is to relate the stabilizers of exposed faces of $\lambda^{-1}(C)$ in $\mathrm{Ad}\,\mathscr{O}(n)$ to the stabilizers of exposed faces of C in $\mathscr{P}(n)$.

For a matrix set $\mathscr{F} \subset \mathscr{S}(n)$, we define

$$\mathrm{Diag}^{-1} \mathscr{F} = \{x \in \mathbf{R}^n \mid \mathrm{Diag}\, x \in \mathscr{F}\}.$$

THEOREM 5.1 (Exposed faces). *Suppose the subset C of $\mathbf{R}^n$ is closed, convex, and permutation-invariant. Then the following properties of a nonempty subset $\mathscr{F}$ of $\lambda^{-1}(C)$ are equivalent:*

(i) *$\mathscr{F}$ is an exposed face of $\lambda^{-1}(C)$;*

(ii) *$\mathscr{F} = V^T(\mathrm{Ad}[\mathscr{O}(n)_{\sim_E}]\mathrm{Diag}\, E)V$ for some orthogonal V and some exposed face E of C;*

(iii) *$\mathscr{F}$ is convex, is relatively $\mathrm{Ad}\,\mathscr{O}(n)$-invariant, and satisfies*

$$\mathrm{Diag}^{-1}\left[V(\mathrm{ri}\,\mathscr{F})V^T\right] \neq \emptyset,$$

$$\mathrm{Diag}^{-1}(V\mathscr{F}V^T) = E,$$

for some orthogonal V and some exposed face E of C.

If property (ii) (or equivalently (iii)) holds, then the stablizers of the faces E and $\mathscr{F}$ are related by

$$[\mathrm{Ad}\,\mathscr{O}(n)]_{\mathscr{F}} = \mathrm{Ad}\left[V^T\mathscr{O}(n)_{\sim_E} V\right]. \tag{5.2}$$

Proof. (ii) $\Rightarrow$ (i): This follows immediately from the facial chain rule (Theorem 4.6) and the equivalent faces Theorem 2.6.

(iii) $\Rightarrow$ (ii): We restrict to the case $V = I$: the general case is a straightforward consequence. Define a matrix set $\mathscr{H} = \mathrm{Ad}[\mathscr{O}(n)_{\sim_E}]\,\mathrm{Diag}\, E$. This set is a face of $\lambda^{-1}(C)$, by the facial chain rule (Theorem 4.6): we wish to show $\mathscr{H} = \mathscr{F}$.

By the stabilizers corollary 2.7, we can choose a point x in ri E with $\mathscr{P}(n)_x = \mathscr{P}(n)_E$. Since $\mathrm{Diag}^{-1}\,\mathrm{ri}\,\mathscr{F}$ is nonempty, we have $\mathrm{Diag}^{-1}\,\mathrm{ri}\,\mathscr{F} = \mathrm{ri}\,\mathrm{Diag}^{-1}\mathscr{F} = \mathrm{ri}\, E$. Thus $\mathrm{Diag}\, x \in \mathscr{H} \cap \mathrm{ri}\,\mathscr{F}$, and since $\mathscr{F}$ is convex and $\mathscr{H}$ is a face, we deduce $\mathscr{F} \subset \mathscr{H}$.

On the other hand, clearly $\mathrm{Diag}\, E \subset \mathscr{F}$, and since $\mathscr{F}$ is relatively invariant, by Proposition 3.3 and Lemma 3.1 we obtain

$$[\mathrm{Ad}\,\mathscr{O}(n)]_{\mathscr{F}} \supset [\mathrm{Ad}\,\mathscr{O}(n)]_{\mathrm{Diag}\,x} = \mathrm{Ad}\big[\mathscr{O}(n)_{\sim_x}\big] = \mathrm{Ad}\big[\mathscr{O}(n)_{\sim_E}\big].$$

Thus $\mathscr{H} \subset \mathscr{F}$, whence $\mathscr{H} = \mathscr{F}$.

(i) $\Rightarrow$ (iii): Clearly $\mathscr{F}$ is convex and relatively invariant, by the stabilizers corollary 2.7. Furthermore, by the stabilizers corollary and the exposing vectors theorem (2.8), there are symmetric matrices, X in ri $\mathscr{F}$, and Y with $\mathscr{F} = \partial\delta^*_{\lambda^{-1}(C)}(Y)$ satisfying

$$[\mathrm{Ad}\,\mathscr{O}(n)]_X = [\mathrm{Ad}\,\mathscr{O}(n)]_Y = [\mathrm{Ad}\,\mathscr{O}(n)]_{\mathscr{F}}.$$

Since X belongs to $\partial(\delta^*_C \circ \lambda)(Y)$, there is an orthogonal V with

$$X = V^T\,\mathrm{Diag}\,\lambda(X)\,V \quad \text{and} \quad Y = V^T\,\mathrm{Diag}\,\lambda(Y)\,V$$

(see [10, Theorem 3.1]). We claim that if we define an exposed face of C by $E = \partial\delta^*_C(\lambda(Y))$, then property (iii) holds.

Note first that the set $\mathrm{Diag}^{-1}[V(\mathrm{ri}\,\mathscr{F})V^T]$ is nonempty, since it contains the vector $\lambda(X)$. It remains to show $\mathrm{Diag}^{-1}(V\mathscr{F}V^T) = E$. For a vector x in $\mathbf{R}^n$, we have

$$\begin{aligned} x \in \mathrm{Diag}^{-1}(V\mathscr{F}V^T) \quad &\Leftrightarrow \quad V^T(\mathrm{Diag}\,x)V \in \partial(\delta^*_C \circ \lambda)(V^T\,\mathrm{Diag}\,\lambda(Y)V) \\ &\Leftrightarrow \quad \mathrm{Diag}\,x \in \partial(\delta^*_C \circ \lambda)(\mathrm{Diag}\,\lambda(Y)) \\ &\Leftrightarrow \quad \mathrm{Diag}\,x \in \mathrm{Ad}\big[\mathscr{O}(n)_{\sim_{\lambda(Y)}}\big]\,\mathrm{Diag}\,E, \end{aligned}$$

by the chain rule (Theorem 4.2). We deduce immediately $E \subset \mathrm{Diag}^{-1}(V\mathscr{F}V^T)$. Conversely, for a point x in $\mathrm{Diag}^{-1}(V\mathscr{F}V^T)$, we see from the above that there is a matrix U in $\mathscr{O}(n)_{\sim_{\lambda(Y)}}$ with $U^T(\mathrm{Diag}\, x)U$ in Diag E. Hence by Lemma 3.5 there is a matrix P in $\mathscr{P}(n)_{\lambda(Y)}$ with Px in E. But from its definition it is clear that the face E is invariant under $\mathscr{P}(n)_{\lambda(Y)}$. Thus x lies in E, as required.

It remains to prove the stabilizer characterization (5.2). The inclusion

$$[\mathrm{Ad}\, \mathscr{O}(n)]_{\mathscr{F}} \supset \mathrm{Ad}\big[V^T \mathscr{O}(n)_{\sim_E} V\big]$$

follows immediately from property (ii). To see the reverse inclusion, choose the symmetric matrix Y as above. Then

$$[\mathrm{Ad}\, \mathscr{O}(n)]_{\mathscr{F}} = [\mathrm{Ad}\, \mathscr{O}(n)]_Y = [\mathrm{Ad}\, \mathscr{O}(n)]_{V^T[\mathrm{Diag}\, \lambda(Y)]V},$$

so any matrix U in $\mathscr{O}(n)$ with Ad U in $[\mathrm{Ad}\, \mathscr{O}(n)]_{\mathscr{F}}$ satisfies

$$U^T V^T \,\mathrm{Diag}\, \lambda(Y)\, VU = V^T \mathrm{Diag}\, \lambda(Y)\, V.$$

Thus

$$\mathrm{Ad}(VUV^T) \in [\mathrm{Ad}\, \mathscr{O}(n)]_{\mathrm{Diag}\, \lambda(Y)} = \mathrm{Ad}\, \mathscr{O}(n)_{\sim_{\lambda(Y)}},$$

whence by Proposition 3.3,

$$VUV^T \in \mathscr{O}(n)_{\sim_{\lambda(Y)}} \subset \mathscr{O}(n)_{\sim_E},$$

since $E = \partial \delta_C^*(\lambda(Y))$ (see [10, Theorem 3.1]). ∎

In the above result, the matrix set $\lambda^{-1}(C)$ is unbounded (and hence has the empty set as an exposed face) if and only if the set C is unbounded. We therefore immediately deduce the following corollary.

COROLLARY 5.3. *Suppose the subset C of $\mathbf{R}^n$ is closed, convex, and permutation-invariant. Then a subset $\mathscr{F}$ of the matrix set $\lambda^{-1}(C)$ is an exposed face if and only if*

$$\mathscr{F} = V^T\big(\mathrm{Ad}\big[\mathscr{O}(n)_{\sim_E}\big]\, \mathrm{Diag}\, E\big)V$$

for some orthogonal matrix V and some exposed face E of C. In this case

$$[\mathrm{Ad}\,\mathscr{O}(n)]_{\mathscr{F}} = \mathrm{Ad}\big[V^T\mathscr{O}(n)_{\sim_E} V\big].$$

The statement of property (iii) in the exposed faces theorem 5.1 suggests that the equivalence (i) $\Leftrightarrow$ (iii) is more naturally stated in terms of *relative interiors* of exposed faces. The following example shows the necessity of the rather technical phrasing in property (iii).

EXAMPLE 5.4. In the framework of the exposed faces theorem 5.1, let $n = 2$, $C = \mathbf{R}^2_+$, and $\mathscr{F} = \mathbf{R}_+ A$, where

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}.$$

Clearly C is closed, convex, and permutation-invariant, and $\mathscr{F}$ is convex. Since A has distinct eigenvalues, we can check that the stabilizer $[\mathrm{Ad}\,\mathscr{O}(2)]_A$ consists just of the identity, and hence $\mathscr{F}$ is relatively $\mathrm{Ad}\,\mathscr{O}(2)$-invariant. Now $\mathrm{Diag}^{-1}\,\mathscr{F}$ consists just of the origin, and is thus an exposed face of C. However, $\mathscr{F}$ is certainly not a face of $\lambda^{-1}(C)$, the cone of 2×2 positive semidefinite matrices: for example,

$$A = \frac{1}{2}\begin{pmatrix} 4 & -4 \\ -4 & 4 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}.$$

The exposed faces theorem 5.1 does not apply, because $\mathrm{Diag}^{-1}\,\mathrm{ri}\,\mathscr{F}$ is empty.

Motivated by this, the next result is a less clumsy restatement of the equivalence of properties (i) and (iii) in the exposed faces theorem 5.1. We say that a convex set is *relatively open* if it is open with respect to its affine span.

COROLLARY 5.5 (Relative interiors). *Suppose the subset C of $\mathbf{R}^n$ is closed, convex, and permutation-invariant. Then the following properties of a subset $\mathscr{E}$ of $\lambda^{-1}(C)$ are equivalent:*

(i) *$\mathscr{E}$ is the relative interior of a nonempty exposed face of $\lambda^{-1}(C)$;*

(ii) *$\mathscr{E}$ is convex, relatively open, and relatively $\mathrm{Ad}\,\mathscr{O}(n)$-invariant, and for some orthogonal matrix V, the set $\mathrm{Diag}^{-1}(V\mathscr{E}V^T)$ is the relative interior of a nonempty exposed face of C.*

Proof. (i) ⇒ (ii): If there is a nonempty exposed face $\mathscr{F}$ of $\lambda^{-1}(C)$ with relative interior $\mathscr{E}$, then $\mathscr{E}$ is certainly relatively open and convex. Furthermore, since $\mathscr{F}$ is relatively Ad $\mathscr{O}(n)$-invariant by the exposed faces theorem 4.2, so is $\mathscr{E}$ (by Proposition 2.3), and we also see that for some orthogonal V, the set $E = \mathrm{Diag}^{-1}(V\mathscr{F}V^T)$ is a nonempty exposed face of C with $\mathrm{Diag}^{-1}[V(\mathrm{ri}\,\mathscr{F})V^T]$ nonempty. Hence, by [15, Theorem 6.7], noting that the linear map $X \mapsto VXV^T$ is invertible,

$$\mathrm{Diag}^{-1}(V\mathscr{E}V^T) = \mathrm{Diag}^{-1}\big[V(\mathrm{ri}\,\mathscr{F})V^T\big] = \mathrm{Diag}^{-1}\,\mathrm{ri}(V\mathscr{F}V^T)$$

$$= \mathrm{ri}\,\mathrm{Diag}^{-1}(V\mathscr{F}V^T) = \mathrm{ri}\,E,$$

as required.

(ii) ⇒ (i): Let $\mathscr{F} = \mathrm{cl}\,\mathscr{E}$. Then the set $\mathscr{F}$ is convex, with relative interior $\mathscr{E}$, and $\mathscr{F}$ is relatively Ad $\mathscr{O}(n)$-invariant since $\mathscr{E}$ is, using Corollary 2.4. Certainly $\mathrm{Diag}^{-1}[V(\mathrm{ri}\,\mathscr{F})V^T]$ is nonempty. Furthermore, since $\mathrm{Diag}^{-1}\,\mathrm{ri}(V\mathscr{E}V^T) = \mathrm{Diag}^{-1}(V\mathscr{E}V^T)$ is nonempty, we deduce

$$\mathrm{Diag}^{-1}(V\mathscr{F}V^T) = \mathrm{Diag}^{-1}\big[V(\mathrm{cl}\,\mathscr{E})V^T\big] = \mathrm{Diag}^{-1}\,\mathrm{cl}(V\mathscr{E}V^T)$$

$$= \mathrm{cl}\,\mathrm{Diag}^{-1}(V\mathscr{E}V^T),$$

by [15, Theorem 6.7], so $\mathrm{Diag}^{-1}(V\mathscr{F}V^T)$ is an exposed face of C. Now the exposed faces theorem 4.2 shows $\mathscr{F}$ is an exposed face of $\lambda^{-1}(C)$. ■

If an exposed face of a convex set consists of just one point, then that point is called *exposed.* The following corollary, although more straightforward to obtain directly (see [10]), is a good illustration of the exposed faces theorem. The analogous result for *extreme* points is also true, without the assumption of closure (see [11]). In particular this shows that if a subset C of $\mathbf{R}^n$ is convex and permutation-invariant, then $\mathrm{Diag}\,C$ is a *diagonal* of $\lambda^{-1}(C)$ in the sense of [6]: that is, a point in $\mathrm{Diag}\,C$ lies in $\mathrm{ri}\,\mathrm{Diag}\,C$ if and only if it lies in $\mathrm{ri}\,\lambda^{-1}(C)$, and is an extreme point of $\mathrm{Diag}\,C$ only if it is an extreme point of $\lambda^{-1}(C)$.

COROLLARY 5.6 (Exposed points). *Suppose the subset C of $\mathbf{R}^n$ is closed, convex, and permutation-invariant. Then a symmetric matrix X is an exposed point of the matrix set $\lambda^{-1}(C)$ if and only if $\lambda(X)$ is an exposed point of C.*

Proof. If X is an exposed point, then, by the exposed faces theorem, there is an exposed face E of C and an orthogonal V with

$$\{X\} = V^T\big(\mathrm{Ad}[\mathscr{O}(n)_{\sim_E}]\ \mathrm{Diag}\ E\big)V.$$

Thus E must be a singleton: $E = \{x\}$ for some exposed point x of E. Furthermore we have $X = V^T(\mathrm{Diag}\ x)V$, whence $\lambda(X) = Px$ for some permutation matrix P, and since x is exposed, so is $\lambda(X)$.

Conversely, suppose $\lambda(X)$ is exposed in C. Choose an orthogonal V with $X = V^T[\mathrm{Diag}\ \lambda(X)]V$. Since

$$\{\lambda(X)\} = \mathrm{Diag}^{-1}\big(V\{X\}V^T\big) = \mathrm{Diag}^{-1}\big[V(\mathrm{ri}\{X\})V^T\big],$$

and since trivially the set $\{X\}$ is convex and relatively $\mathrm{Ad}\,\mathscr{O}(n)$-invariant, it follows by the exposed faces theorem that X is exposed. ■

In the above result, if in addition the set C is a cone, then clearly so is the matrix set $\lambda^{-1}(C)$. In this case, for a nonzero matrix X in $\mathscr{S}(n)$, the half line $\mathbf{R}_+X$ is an exposed ray of $\lambda^{-1}(C)$ if and only if $\mathbf{R}_+\lambda(X)$ is an exposed ray of C. The proof is similar.

6. EXAMPLES

In this section we illustrate our results on some examples, beginning with the positive semidefinite cone.

COROLLARY 6.1 (Positive semidefinite cone). *The exposed faces of the positive semidefinite cone, $\mathscr{S}(n)_+$, are the sets*

$$\left\{V^T\begin{pmatrix} W & 0 \\ 0 & 0\end{pmatrix}V \,\middle|\, W \in \mathscr{S}(m)_+\right\},$$

where the matrix V is orthogonal and $0 \leqslant m \leqslant n$, together with the empty set. The exposed rays are $\mathbf{R}_+uu^T$ for nonzero column vectors u in $\mathbf{R}^n$.

Proof. We apply Corollary 5.3, with the set C being the positive orthant $\mathbf{R}^n_+$. A typical exposed face of $\mathbf{R}^n_+$ is

$$E = \left\{(x,0)^T \in \mathbf{R}^n \mid x \in \mathbf{R}^m_+\right\},$$

where $0 \leqslant m \leqslant n$. The sets PE, where P is a permutation matrix (and the empty set) comprise all the exposed faces of $\mathbf{R}_+^n$.

The equivalence relation $\sim_E$ has equivalence classes $\{1, 2, \dots, m\}$ and $\{m+1, m+2, \dots, n\}$ [the blocks preserved by the stabilizer $P(n)_E$], whence

$$\mathscr{O}(n)_{\sim_E} = \left\{ \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix} \middle| P \in \mathscr{O}(m),\ Q \in \mathscr{O}(n-m) \right\}.$$

The result now follows easily. ■

This characterization may be compared with those in [19, 2]. For example, we may deduce that all the faces of $\mathscr{S}(n)_+$ are exposed. For a more general discussion of cones with this property, see [16].

We denote the unit balls for the l_1 and l_∞ norms on $\mathbf{R}^n$ by

$$B_1^n = \left\{ x \in \mathbf{R}^n \middle| \sum_1^n |x_i| \leqslant 1 \right\},$$

$$B_\infty^n = \{ x \in \mathbf{R}^n \mid |x_i| \leqslant 1 \text{ for all } i \}$$

respectively. The corresponding matrix sets are then the unit balls for the trace norm and spectral norms respectively:

$$\lambda^{-1}(B_1^n) = \left\{ X \in \mathscr{S}(n) \middle| \sum_1^n |\lambda_i(X)| \leqslant 1 \right\},$$

$$\lambda^{-1}(B_\infty^n) = \{ X \in \mathscr{S}(n) \,\|\, \lambda_1(X)| \leqslant 1 \}.$$

Corollary 6.2 (Spectral norm). *The exposed faces of the unit ball of the spectral norm on $\mathscr{S}(n)$ are the sets*

$$\left\{ V^T \begin{pmatrix} W & 0 & 0 \\ 0 & I_r & 0 \\ 0 & 0 & -I_s \end{pmatrix} V \middle| W \in \mathscr{S}(m),\ |\lambda_1(W)| \leqslant 1 \right\},$$

where the matrix V is orthogonal and the nonnegative integers m, r, and s sum to n. The exposed points are the symmetric matrices with all eigenvalues having absolute value 1.

Proof. We apply Corollary 5.3 with the set C being B_∞^n. A typical exposed face of this set is

$$E = \left\{(x, e^r, -e^s)^T \mid x \in B_\infty^m\right\},$$

where the nonnegative integers m, r, and s sum to n, and e^r denotes the vector $(1, 1, \dots, 1)^T$ in $\mathbf{R}^r$. The sets PE, where P is a permutation matrix, comprise all the exposed faces of B_∞^n.

The equivalence relation $\sim_E$ has equivalence classes

$$\{1, 2, \dots, m\}, \qquad \{m+1, m+2, \dots, m+r\}, \quad \text{and}$$
$$\{m+r+1, m+r+2, \dots, n\},$$

whence

$$\mathscr{O}(n)_{\sim_E} = \left\{ \begin{pmatrix} P & 0 & 0 \\ 0 & Q & 0 \\ 0 & 0 & R \end{pmatrix} \middle| P \in \mathscr{O}(m),\, Q \in \mathscr{O}(r),\, R \in \mathscr{O}(s) \right\}.$$

The result now follows easily. ∎

The proof of the next result is rather similar in spirit.

COROLLARY 6.3 (Trace norm). *The exposed faces of the unit ball of the trace norm on $\mathscr{S}(n)$ are the whole unit ball and the sets*

$$\left\{ V^T \begin{pmatrix} 0 & 0 & 0 \\ 0 & X & 0 \\ 0 & 0 & -Y \end{pmatrix} V \middle| X \in \mathscr{S}(r)_+,\, Y \in \mathscr{S}(s)_+,\, \operatorname{tr} X + \operatorname{tr} Y = 1 \right\},$$

where the matrix V is orthogonal, $r, s \geqslant 0$, and $r + s \leqslant n$. The exposed points are $\pm uu^T$ for unit column vectors u in $\mathbf{R}^n$.

The remaining l_p norms on $\mathbf{R}^n$ (for $1 < p < \infty$) have unit balls for which every boundary point is exposed. The corresponding matrix norms are the *Schatten p-norms*: by the exposed points corollary 5.6, the same property must hold for the unit balls of these norms,

$$\left\{ X \in \mathscr{S}(n) \middle| \sum_{i=1}^{n} |\lambda_i(X)|^p \leqslant 1 \right\}.$$

The case $p = 2$ is the Frobenius norm.

We end with a less standard illustration.

EXAMPLE 6.4. Consider the closed, convex, permutation-invariant set

$$C = \left\{ x \in \mathbf{R}^4 \,\middle|\, \sum_{i \in I} x_i \leqslant 0 \text{ whenever } |I| = 2 \right\}.$$

A particular exposed face is

$$\begin{aligned} E &= \{x \in C \mid x_1 + x_2 = 0\} \\ &= \left\{ (\alpha, -\alpha, \beta, \gamma)^T \,\middle|\, \beta, \gamma \leqslant -|\alpha| \right\}. \end{aligned}$$

Clearly the equivalence classes of $\sim_E$ are $\{1, 2\}$ and $\{3, 4\}$, and the stabilizer of E is the group

$$\mathscr{P}(n)_E = \left\{ \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix} \,\middle|\, P, Q \in \mathscr{P}(2) \right\}.$$

Hence

$$\mathscr{O}(n)_{\sim_E} = \left\{ \begin{pmatrix} R & 0 \\ 0 & S \end{pmatrix} \,\middle|\, R, S \in \mathscr{O}(2) \right\},$$

and by Theorem 5.1, the matrix set

$$\mathscr{F} = \operatorname{Ad} \mathscr{O}(n)_{\sim_E} \operatorname{Diag} E$$

is an exposed face of the matrix set

$$\lambda^{-1}(C) = \{X \in \mathscr{S}(4) \mid \lambda_1(X) + \lambda_2(X) \leqslant 0\},$$

with stabilizer $\operatorname{Ad}[\mathscr{O}(n)_{\sim_E}]$. It is easy to express $\mathscr{F}$ more directly:

$$\mathscr{F} = \left\{ \begin{pmatrix} Y & 0 \\ 0 & Z \end{pmatrix} \,\middle|\, Y, Z \in \mathscr{S}(2),\ \operatorname{tr} Y = 0,\ \lambda_1(Z) \leqslant -\lambda_1(Y) \right\}.$$

For any orthogonal V, the set $V^T \mathscr{F} V$ is also an exposed face.

I am indebted to an anonymous referee for many helpful comments leading to substantial improvements in this paper (in particular, Theorem 2.6), and to Levent Tunçel for bringing [12] to my attention.

REFERENCES

1 F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton, Complementarity and nondeneracy in semidefinite programming, *Math. Programming*, Series B 77:111–128 (1997).

2 G. P. Barker and D. Carlson, Cones of diagonally dominant matrices, *Pacafic J. Math.* 57:15–32 (1975).

3 E. M. de Sá, Exposed faces and duality for symmetric and unitarily invariant norms, *Linear Algebra Appl.* 197–198:429–450 (1994).

4 E. M. de Sá, Faces of the unit ball of a symmetric gauge function, *Linear Algebra Appl.* 197–198:349–395 (1994).

5 E. M. de Sá, Faces of the unit ball of a unitarily invariant norm, *Linear Algebra Appl.* 197–198:451–493 (1994).

6 M. Fiedler and V. Pták, Diagonal of convex sets, *Czechoslovak Math. J.* 28:25–44 (1978).

7 P. A. Fillmore and J. P. Williams, Some convexity theorems for matrices, *Glasgow Math. J.* 12:110–117 (1971).

8 A. S. Lewis, The convex analysis of unitarily invariant matrix functions, *J. Convex Anal.* 2:173–183 (1995).

9 A. S. Lewis, Von Neumann's Lemma and a Chevalley-Type Theorem for Convex Functions on Cartan Subspaces, Technical Report CORR 95-19, University of Waterloo, 1995; *Nonlinear Analysis*, submitted for publication.

10 A. S. Lewis, Convex analysis on the Hermitian matrices, *SIAM J. Optim.* 6:164–177 (1996).

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

12 A. L. Onishchik and E. B. Vinberg, *Lie Groups and Algebraic Groups*, Springer-Verlag, Berlin, 1990.

13 G. Pataki, On Cone-LP's and Semi-definite Programs: Facial Structure, Basic Solutions and the Simplex Method, Technical Report GSIA 1995-03, Carnegie Mellon Univ., 1995.

14 M. Ramana, L. Tunçel, and H. Wolkowicz, Strong duality for semidefinite programming, *SIAM J. Optim.* 7:641–662 (1997).

15 R. T. Rockafellar, *Convex Analysis*, Princeton U.P., Princeton, N. J. 1970.

16 C. H. Sung and B. S. Tam, A study of projectionally exposed cones, *Linear Algebra Appl.* 139:225–252 (1990).

17 B. S. Tam, On the structure of the cone of positive operators, *Linear Algebra Appl.* 167:65–85 (1992).

18 B. S. Tam and H. Schneider, On the core of a cone-preserving map, *Trans. Amer. Math. Soc.* 343:479–524 (1994).

19 O. Taussky, Positive-definite matrices and their role in the study of characteristic roots of general matrices, *Adv. Math.* 2:175–186 (1968).

Received 17 October 1995; final manuscript accepted 20 March 1997