The geometry of spectral interlacing

Ricardo S. Leite and Carlos Tomei

August 16, 2021

Abstract

We provide a detailed description of the maps associated with spectral interlacing, for rank one perturbations and bordering of symmetric and Hermitian matrices. The arguments rely on standard techniques of nonlinear analysis.

Keywords: Spectral interlacing, convex polytopes, degree theory.

MSC-class: 15A29, 15A42, 15B57

1 Introduction

Usually when one thinks of geometry and interlacing, one is reminded of the min-max characterization of eigenvalues of Hermitian matrices. In this text, we have something different in mind: how does spectrum vary with the perturbation? We present a detailed description of this process. There are two familiar scenarios in which the spectra of two real, symmetric (resp., Hermitian) matrices $S$ and $T$ interlace. The first is when $S - T$ is a rank one perturbation. In the second, $S$ is obtained from $T$ by bordering, i.e., $T$ is obtained by removing the last row and column of $S$. We start with the real case.

Let $S$ be a real, $n \times n$ symmetric matrix with simple spectrum $\lambda_1 < \ldots < \lambda_n$ yielding an ordered $n$-uple $\lambda \in \mathbb{R}^n$. Fix a normalized eigenbasis $Q$ of $S$ and dispose the eigenvectors along the columns of the matrix $Q$. Let $O_Q \in \mathbb{R}^n$ be the positive orthant associated with $Q$, the set of vectors of the form $Qp$, where $p \in \mathbb{R}^n$ has nonnegative entries. Given a sequence of signs $s = (s_1, \ldots, s_k) \in \{1, -1\}^n$, define the sign matrix $\Sigma = \text{diag}(s_1, \ldots, s_n)$. Sequences of signs are associated to the $2^n$ signed orthants in $\mathbb{R}^n$. Vectors $v, w \in \mathbb{R}^n$ are equivalent, $v \sim w$, if and only if there is $\Sigma$ such that $v = \Sigma w$. The global $n$-fold

$$\text{abs} : \mathbb{R}^n \to O_Q \simeq \mathbb{R}^n/\sim, \quad (v_1, \ldots, v_n) \mapsto (|v_1|, \ldots, |v_n|)$$

1
has a simple geometric interpretation. Let \( S(r) \subset \mathbb{R}^n \) be the Euclidean sphere of radius \( r \) centered at the origin. An ordered \( n \)-uple of \( \mathbb{R}^n \) is a vector whose entries are in nondecreasing order. Define

\[ \tilde{F}_S : \mathcal{D} = \mathcal{O}_Q \to \mathbb{R}^n \]

\[ v \mapsto \sigma_o(S + v \otimes v) \]

where the ordered spectrum \( \sigma_o(T) \) of a matrix \( T \) consists of the ordered \( n \)-uple whose entries are the eigenvalues of \( T \). Here \( v \otimes v = vv^T \). Define the restriction \( \hat{F}_S : \mathcal{D} \cap S(r) \to \mathbb{R}^n \). We frequently replace the counterdomain of a function by its image: we say \( \tilde{H} : X \to Y \) induces a function \( H : X \to \text{Image}(F) \). Finally set

\[ \mathcal{P}_F = [\lambda_1, \lambda_2] \times [\lambda_2, \lambda_3] \times \ldots \times [\lambda_n, \infty) , \mathcal{P}'_F = \mathcal{P}_F \cap \{ \mu \in \mathbb{R}^n, \sum \mu_i = r + \sum \lambda_i \} . \]

**Theorem 1.1**

(i) The image of the maps \( \tilde{F}_S \) and \( \hat{F}_S \) are \( \mathcal{P}_F \) and \( \mathcal{P}'_F \).

(ii) The induced maps \( F_S : \mathcal{D} \to \mathcal{P}_F \) and \( F'_S : \mathcal{D} \cap S(r) \to \mathcal{P}'_F \) are homeomorphisms which restrict to diffeomorphisms between the interiors of domain and image.

(iii) \( F_S \circ \text{abs}(v) = F_S \circ \text{abs}(w) \) if and only if \( v \) and \( w \) are equivalent.

Said differently, \( F_S \circ \text{abs} : \mathbb{R}^n \to \mathcal{P} \) is a global \( n \)-fold. Matrices \( S + v \otimes v \) and \( S + w \otimes w \) are equal if and only if \( v = \pm w \). In particular, the theorem counts the number of matrices yielding a given ordered sequence \( \mu \) interlacing with \( \lambda \).

We relate the theorem and interlacing in a precise sense. The general rank one symmetric perturbation matrix is of the form \( cv \otimes v \) for a real unit vector \( v \) and \( c \in \mathbb{R} \). The sign of \( c \) specifies if the perturbation of \( S \) pushes the spectrum to the right (the case \( c > 0 \)) or to the left (\( c < 0 \)). Our choice of domain and image corresponds to the case \( c \geq 0 \): minor alterations handle \( c \leq 0 \). For \( c \geq 0 \), from the standard interlacing theorem, the ordered \( n \)-uples \( \lambda \) and \( \mu \) associated with the matrices \( S \) and \( S + (cv) \otimes (cv) \) satisfy

\[ \lambda_1 \leq \mu_1 \leq \lambda_2 \leq \ldots \leq \mu_{n-1} \leq \lambda_n \leq \mu_n , \]

which is equivalent to \( \mu \in \mathcal{P}_F \). From the above theorem, all possible interlacing \( n \)-uples \( \mu \) are obtained by performing a unique positive semidefinite rank one perturbation \( v \otimes v \) of \( S \) for some \( v \in \mathcal{D} \).

The following analogy may be of interest. For a real, symmetric matrix \( S \) with eigenvalues \( \{\lambda_1, \ldots, \lambda_n\} \), the Schur-Horn theorem states that the map

\[ SH : \{Q^T SQ, Q \text{ orthogonal}\} \to \text{conv}\{ (\lambda_{\pi(1)}, \ldots, \lambda_{\pi(n)}) , \pi \in S_n \} \]
is a surjective map. Here, conv $X$ is the convex hull of the points of $X$, and $S_n$ is the permutation group acting in the set of indices $\{1, 2, \ldots, n\}$. The fact that the image of $SH$ is contained in the set above follows from min-max estimates, and was proved by Schur (12). Surjectivity was proved by Horn (4) much later. In a sense, the usual interlacing property is analogous to Schur’s result, while we are concerned with the counterpart to Horn’s. The fact that dimensions of domain and image are the same allow us to obtain a more detailed result. The Schur-Horn theorem in the context of rank one matrices is trivial: for example, if $\|v\| = r$, the diagonal entries of $D + v \otimes v^*$ belong to the simplex $(\lambda_1, \ldots, \lambda_n) + r \text{conv}\{(e_1, \ldots, e_n)\}$.

Our approach is geometric: we follow the basic steps used to draw the graph of a function in one variable. We identify the boundary $\partial D$ of the domain of $\tilde{F}_S$ and its image $\tilde{F}_S(\partial D)$, the points where the $\tilde{F}_S$ is not smooth and the points where its differential is not invertible. From Proposition 2.1 all such points lie in $\partial D$. We then show that $\partial D$ is taken to $\partial \mathcal{F}_F$ injectively (including, informally, points at infinity: $\tilde{F}_S$ is a proper map). Standard topological arguments then imply injectivity in $D$. Along the argument, we explain how faces of $D$ are creased by the map $\tilde{F}_S$ so as to obtain the faces of the parallelotope $\partial \mathcal{F}_F$.

The simple geometry of the map $\tilde{F}_S$ has implications to numerics. Consider the following inverse problem (6) — given a symmetric matrix $S$ with ordered spectrum $\lambda$ and an interlacing ordered $n$-uple $\mu$, find a rank one perturbation $cv \otimes v$ such that the spectrum of $S + cv \otimes v$ is $\mu$. The problem now, interpreted as inverting the function $\tilde{F}_S$, is granted to be solvable by numerical continuation starting from any point in the interior of $D$: there are no critical values in the interior of $D$ and, in principle, there are no obstruction to continuation. Recently, Maciazek and Smilansky (10) considered analogous inverse problems and pointed out the relevance of discrete information provided by strings of signs. We believe our presentation sheds some light on the issue.

We now describe our results for the second scenario, in the real case. Take $S, Q, Q$ and $O_Q$ as before. Define the $(n + 1) \times (n + 1)$ bordered matrix

$$B_S(v, c) = \begin{pmatrix} S & v^T \\ v & 2c \end{pmatrix},$$

the map

$$\tilde{G}_S : D = O_Q \times \mathbb{R} \to \mathbb{R}^{n+1}$$

$$(v, c) \mapsto \sigma_o(B_S(v, c))$$

and its restriction $\tilde{G}_S : D \cap S(r) \to \mathbb{R}^{n+1}$. Set

$$\mathcal{P}_G = (-\infty, \lambda_1] \times [\lambda_1, \lambda_2] \times [\lambda_2, \lambda_3] \times \ldots \times [\lambda_n, \infty),$$

$$\mathcal{P}_G^r = \mathcal{P}_G \cap \{ \mu \in \mathbb{R}^{n+1}, \sum_i \mu_i = r + \sum_i \lambda_i \}.$$
Theorem 1.2 The image of the maps $\tilde{G}_S$ and $\tilde{G}_S^r$ are respectively $P_G$ and $P_G^r$. The induced maps $G_S : D \to P_G$ and $G_S^r : D \to P_G^r$ are homeomorphisms which restrict to diffeomorphisms between interiors of domain and image. Moreover, $v$ and $w$ are equivalent if and only if $\tilde{G}_S \circ \text{abs}(v) = \tilde{G}_S \circ \text{abs}(w)$.

Again, $\tilde{G}_S \circ \text{abs} : \mathbb{R}^n \times \mathbb{R} \to P_G$ is an $n$-fold. Clearly, matrices $B_S(v, c)$ and $B_S(w, d)$ are equal if and only if $v = w, c = d$. The proof requires minor alterations of the previous one. The comments related to the computation of the inverse of $F_S$ also apply to $G_S$.

Complex versions of the above results follow without difficulty. Let $H$ be an $n \times n$ Hermitian matrix with ordered simple spectrum $\lambda$. Again, fixed normal eigenvectors $w_k$ for which $H w_k = \lambda_k w_k$, and consider the unitary matrix $W$ having $w_k$ as columns. Let $U_W$ consist of the linear combinations $u = \sum_k p_k w_k$, for $p_k \geq 0$. For $y, z \in \mathbb{C}^n$, set $y \sim z$ if and only if $y = \Sigma z$ for some diagonal matrix of phases $\Sigma = \text{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n})$, $\theta_k \in [0, 2\pi)$. We consider the projection

$$\text{mod} : \mathbb{C}^n \to (\mathbb{C}^n / \sim) \simeq U_W, \quad (v_1, \ldots, v_n) \mapsto (|v_1|, \ldots, |v_n|)$$

and the maps

$$\tilde{F}_H : U_W \to \mathbb{R}^n, \quad v \mapsto \sigma_0(H + vv^*)$$

$$\tilde{G}_H : U_W \times [0, \infty) \to \mathbb{R}^n, \quad (v, c) \mapsto \sigma_0(B_H(v, c))$$

where

$$B_H(v, c) = \begin{pmatrix} H & v^* \\ v & 2c \end{pmatrix}.$$  

Corollary 1.3 The induced maps $F_H : U_W \to P_G$ and $G_H : U_W \times [0, \infty) \to P_G$ are homeomorphisms and (real) diffeomorphisms between the interiors of both sets. Vectors $y, z \in \mathbb{C}^n$ (resp. pairs $(y, c)$ and $(z, c)$) satisfy $F_H \circ \text{mod}(y) = F_H \circ \text{mod}(z)$ (resp. $G_H(y, c) \circ \text{mod} = G_H(z, c) \circ \text{mod}$) if and only if $y \sim z$.

Composing these maps with mod, we see that going from real to complex does not enlarge the images of the maps $F_S$ and $G_S$. Moreover, preimages of ordered $n$-uples consist of tori whose dimensions are easily computed.

Similar results may be obtained from the sophisticated machinery of symplectic geometry. The Schur-Horn theorem for Hermitian matrices is a consequence of the celebrated theorem about the convexity of the image of moment maps of torus actions by Atiyah ([1]) and Guillemin-Sternberg ([3]). The real case then follows by an argument by Duistermaat ([2]). The results in this paper are a special case of the characterization of the possible spectra of the sum $A + B$ of two Hermitian matrices, where only the spectrum of $A$ and $B$ is known ([3], [8], [9]). Concretely, $A = D$ has known spectrum $\lambda$, and, for $|v| = r$, $B = vv^*$ has
one eigenvalue equal to $r$, all others being zero. The convexity of the image now follows from a very general theorem of Kirwan ([7], [13]). We take what Thompson ([14]) calls a low road in linear algebra, but gain some information which does not follow directly from rote application of these more general results.

Tomei acknowledges support from CNPq, CAPES and FAPERJ.

2 Real rank one perturbations – Theorem 1.1

Without loss, we may take $S = D$, a diagonal matrix with simple spectrum $\lambda_1 < \ldots < \lambda_n$. Simplifying notation further, we consider

$$\tilde{F} = \tilde{F}_D : D = O_I \to \mathbb{R}^n$$

$$v \mapsto \sigma_o(D + v \otimes v)$$

To prove Theorem 1.1 among other facts, we must show that the image of $\tilde{F}$ is

$$\mathcal{P}_F = [\lambda_1, \lambda_2] \times [\lambda_2, \lambda_3] \times \ldots \times [\lambda_n, \infty).$$

Most of the argument consists in proving that $\tilde{F}$ restricts to a homeomorphism between the boundaries $\partial D$ and $\partial \mathcal{P}_F$. The set $\partial D$ consists of $n$ faces

$$E_i = \{v \in \mathbb{R}^n , v_i = 0\}, \quad i = 1, \ldots, n$$

of the positive octant $D = O_I$. Set

$$L_i = [\lambda_1, \lambda_2] \times \ldots \times [\lambda_{i-1}, \lambda_i] , \quad R_i = [\lambda_i, \lambda_{i+1}] \times \ldots \times [\lambda_n, \infty).$$

The parallelotope $\mathcal{P}_F \subset \mathbb{R}^n$ has $2n - 1$ faces. The eigenvalue $\lambda_1$ is associated with the face $\{\lambda_1\} \times R_1$. Each eigenvalue $\lambda_i, i > 1$ is associated with two faces

$$L_{i-1} \times \{\lambda_i\} \times R_i \quad \text{and} \quad L_i \times \{\lambda_i\} \times R_{i+1}.$$ 

As we shall see, each face $E_i \subset \partial D, i > 1$ is creased by $\tilde{F}$ over two faces of $\partial \mathcal{P}_F$.

Let $\text{int} X$ denote the interior of a set $X$. Consider the following subsets of $\mathcal{D}$.

(a) The boundary $\partial \mathcal{D} = \partial O_I$.

(b) The set $\mathcal{D}_d$ of points in which $\tilde{F}(v)$ has a double eigenvalue.

(c) The critical set $\mathcal{C}$, consisting of points in $\text{int} \mathcal{D}$ in which $\tilde{F}$ is differentiable, but its Jacobian is not invertible.

(d) The regular points, which are the points of $\mathcal{D}$ not in the three sets above.
Proposition 2.1  

(i) \( \partial \tilde{F}(\mathcal{D}) \subset \tilde{F}(\partial \mathcal{D} \cup \mathcal{D}_d \cup \mathcal{C}) \).

(ii) \( \mathcal{D}_d \subset \partial \mathcal{D} \). Thus, at points in \( \text{int} \mathcal{D} \), \( \tilde{F} \) is differentiable.

(iii) \( \mathcal{C} = \emptyset \): points in the interior of \( \mathcal{D} \) are regular.

(iv) The matrices \( D \) and \( D + v \otimes v \) share an eigenvalue \( \lambda \) if and only if \( v \in E_i \).

In particular, \( \partial \tilde{F}(\mathcal{D}) \subset \tilde{F}(\partial \mathcal{D}) \).

Recall the expression for the derivative of a simple eigenvalue \( \lambda \): if \( T w_i = \lambda_i w_i \) for a normalized \( w_i \in \mathbb{R}^n \), \( \lambda_i = \langle Tw_i, w_i \rangle \).

Proof:  At regular points, \( \tilde{F} \) is a local diffeomorphism: this settles (i).

We prove (ii). A double eigenvalue \( \lambda_i \) of \( D + v \otimes v \) admits a (nonzero) eigenvector \( w \) for which \( w_1 = 0 \) (consider the appropriate linear combination of two eigenvectors associated with \( \lambda_i \)). Since \( (D - \lambda_i)w = -(v \otimes v)w = -\langle v, w \rangle v \), equating first coordinates either \( v_1 = 0 \) or \( \langle v, w \rangle = 0 \). In the first case, \( v \in E_1 \subset \partial \mathcal{D} \) and we are done. Otherwise, \( (D - \lambda_i)w = 0 \) and \( w \) must be a canonical vector, \( w = e_i \), so that, as \( \langle v, w \rangle = 0 \), we must have \( v_i = 0 \) so that \( v \in E_i \).

To prove (iii), let \( T = D + v \otimes v \), \( Tw_i = \lambda_i w_i \), \( i = 1, \ldots, n \), \( \|w_i\| = 1 \). The Jacobian of \( \tilde{F} \) at a point \( v \) is

\[
J(v)\dot{v} = \left( \langle w_1, \dot{T} w_1 \rangle, \ldots, \langle w_n, \dot{T} w_n \rangle \right),
\]

where \( \dot{v} \in \mathbb{R}^n \) and \( \dot{T} = \dot{v} \otimes v + v \otimes \dot{v} \). Let \( V \) be the vector space of such matrices. Write the linear transformation \( J(v) \) as a composition,

\[
J(v)\dot{v} = 2 \left( \langle w_1, v \rangle \langle w_1, \dot{v} \rangle, \ldots, \langle w_n, v \rangle \langle w_n, \dot{v} \rangle \right)
= 2 \text{ diag}(\langle w_1, v \rangle, \ldots, \langle w_n, v \rangle) (\langle w_1, \dot{v} \rangle, \ldots, \langle w_n, \dot{v} \rangle)^T.
\]

A point \( v \) in the interior of \( \mathcal{D} \) is critical if and only if \( J(v) \) is not invertible. Clearly \( \dot{v} \mapsto (\langle w_1, \dot{v} \rangle, \ldots, \langle w_n, \dot{v} \rangle) \) is invertible, as the vectors \( \{w_i\} \) are linearly independent. Suppose by contradiction that, for some \( i \), we have \( \langle w_i, v \rangle = 0 \). Equation \( (D + v \otimes v)w_i = \lambda_i w_i \) becomes \( (D - \lambda_i)w_i = 0 \), so that \( w_i = te_i, t \neq 0 \). Now, \( \langle v, w_i \rangle = 0 \) implies \( v_i = 0 \), and again \( v \in E_i \subset \partial \mathcal{D} \).

To prove (iv), take a common eigenvalue \( \lambda \) and eigenvectors \( e_i, y \neq 0 \), so that \( De_i = \lambda_i e_i \) and \( Dy + \langle v, y \rangle v = \lambda_i y \) and \( (D - \lambda_i)(y - e_i) = -\langle v, y \rangle v \). The \( i \)-entry of both sides of the last equation is zero. If \( v_i = 0 \), we are done. Suppose \( \langle v, y \rangle = 0 \): \( y = te_i \in E_i, t \neq 0 \), and then \( v_i = 0 \). The converse is trivial.

One might worry about the possibility that the ‘boundary at infinity’ of the domain gives rise to a piece of boundary of the image. The next result implies that this may not happen.

Lemma 2.2 The map \( F \) is proper.
Proof: With the Frobenius norm, for a matrix \( T = D + v \otimes v \) with eigenvalues \( \{\mu_i\} \), we have \( \|D + v \otimes v\|^2 = \sum_i \mu_i^2 \). 

We still cannot say that \( \hat{F}(E_i) \) (or equivalently \( \hat{F}(\partial D) \)) lies in \( \partial \hat{F}(\partial D) \): when ordering the spectrum, \( \lambda_i \) must be either in position \( i - 1 \) or \( i \). The study of \( \partial \hat{F}(E_i) \) requires the understanding of the map \( F_D \) for the \((n - 1) \times (n - 1)\) matrix \( \hat{D} \), obtained from \( D \) by removing the eigenvalue \( \lambda_i \). Said differently, the proof that \( \hat{F} \) takes \( D \) to \( \partial \mathcal{P}_F \) homeomorphically in dimension \( n \) naturally uses Theorem [1.1] in dimension \( n - 1 \). It is the ordering of the eigenvalues that leads to creasing the faces of \( D \) as to cover the faces of \( \mathcal{P}_F \).

We use a simple topological lemma, which we restrict to the case of interest. Given \( \hat{F} : D \to \mathbb{R}^n \), a preimage of \( \mu \in \mathbb{R}^n \) is a vector \( v \in D \) such that \( \hat{F}(v) = \mu \).

**Lemma 2.3** The number of preimages is constant (and finite) for points in the same connected component of \( \mathbb{R}^n \setminus F(\partial D) \).

Part of the proof of Theorem [1.1] will be showing that points in \( \mathbb{R}^n \setminus \mathcal{P}_F \) have zero preimages (so that \( F : D \to \mathcal{P}_F \) is well defined) and points in \( \text{int} \mathcal{P}_F \) have exactly one preimage (so that \( F : \text{int} D \to \text{int} \mathcal{P}_F \) is a diffeomorphism).

**Proof:** Take \( \mu \in \mathbb{R}^n \setminus F(\partial D) \). By connectivity, it suffices to show that, for a small open neighborhood \( U \) of \( \mu \), points \( \tilde{\mu} \in U \) have the same number of preimages. If \( \mu \) has infinite preimages, by properness they have to accumulate at some preimage \( v_* \). Clearly preimages in \( \text{int} D \) are isolated, by the inverse function theorem, thus \( v_* \in \partial D \), contradicting \( \hat{F}(v_*) = \mu \in \mathbb{R}^n \setminus F(\partial D) \).

Let \( \mu \) have preimages \( v_1, \ldots, v_k \). From the inverse function theorem, for every sufficiently small open ball \( B \) centered in \( \mu \), there are open disjoint sets \( V_i, i = 1, \ldots, k \), each containing \( v_i \), for which \( \hat{F} : V_i \to B \) is a diffeomorphism. Thus, points in \( B \) have at least \( k \) preimages. If, for balls \( B_n \) of radius \( 1/n \) there are points \( \tilde{\mu}_n \) with (at least) \( k + 1 \) preimages, one preimage \( w_n \in D \) is outside \( \cup_{i=1}^k V_i \). By properness of \( \hat{F} \), they accumulate at \( w_* \notin \cup_{i=1}^k V_i \). But then \( \hat{F}(w_*) = \mu \), contradicting the fact that \( \mu \) has exactly \( k \) preimages, all in \( \cup_{i=1}^k V_i \). 

**Proof of Theorem [1.1]:** The argument is by induction. The case \( n = 2 \) contains the gist of the proof. For \( v \in E_1 \), we must have \( v = (0, c) \), so that

\[
T = D + v \otimes v = \begin{pmatrix}
\lambda_1 & 0 \\
0 & \lambda_2 + c^2
\end{pmatrix}.
\]

As \( \lambda_1 < \lambda_2 \), we also have \( \lambda_1 < \lambda_2 + c^2 \), so that

\[
\hat{F}(T) = (\lambda_1, \lambda_2 + c^2) \in \{\lambda_1\} \times [\lambda_2, \infty) \in \partial \mathcal{P}.
\]

If instead \( v \in E_2, v = (c, 0) \) and

\[
T = D + v \otimes v = \begin{pmatrix}
\lambda_1 + c^2 & 0 \\
0 & \lambda_2
\end{pmatrix}.
\]
There are two possibilities. If $\lambda_1 + c^2 \leq \lambda_2$,

$$\hat{F}(T) = (\lambda_1 + c^2, \lambda_2) \in [\lambda_1, \lambda_2] \times \{\lambda_2\} \in \partial P_F .$$

Otherwise,

$$\hat{F}(T) = (\lambda_2, \lambda_1 + c^2) \in \{\lambda_2\} \times [\lambda_2, \infty) \in \partial P_F .$$

As $\hat{F} : D \rightarrow \mathbb{R}^n$ is proper, the restriction $F : \partial D \rightarrow \partial P_F$ is a homeomorphism. This completes the first part of the argument for $n = 2$.

We begin the second part by showing that $F : \text{int} D \rightarrow \text{int} P_F$ is a diffeomorphism. As $F(\partial D) = \partial P_F$, Lemma [2.3] applied to $F : D \rightarrow \mathbb{R}^n$ implies that points in $\text{int} P_F$ have the same number of preimages, as do points in $\mathbb{R}^n \setminus P_F$. As some points in $\mathbb{R}^n \setminus P_F$ do not have their coordinates in nondecreasing order (as do all points of $F(D)$), no point in $\mathbb{R}^n \setminus P_F$ is in $F(D)$. Thus, the restriction $F : D \rightarrow \hat{P}_F$ is well defined.

The existence of a regular point (any point of $\text{int} D$) implies that $F(D)$ contains an open ball, which necessarily contains some point of $\text{int} P_F$. Again, from Lemma [2.3] all points of $\text{int} P_F$ belong to $F(D)$: $F : D \rightarrow \hat{P}_F$ is surjective.

We now prove injectivity of $F : \text{int} D \rightarrow \text{int} P_F$. As $F : D \rightarrow \mathbb{R}^n$ is a proper map, there is a well defined degree $\text{deg}(\hat{F}, \mu)$ for points $\mu \notin F(\partial D)$ (in particular, $\mu \in \text{int} P_F$) (an excellent reference for degree theory is [11]), given by

$$\text{deg}(\hat{F}, \mu) = \sum_{v \in \hat{F}^{-1}(\mu)} \text{sgn} \det J(v) .$$

The case of interest is especially simple: points $v \in \hat{F}^{-1}(\mu)$ belong to $\text{int} D$, which in turn contains no critical points, from Proposition [2.1](ii). Thus all (necessarily nonzero) determinants of Jacobians have the same sign, so that the number of preimages of a point $\mu$ equals $|\text{deg}(F, \mu)|$. Finally, as $F : \partial D \rightarrow \partial P_F$ of is a homeomorphism, we must have $|\text{deg}(F, \mu)| = 1$. Thus, $\mu$ has a single preimage and $F : \text{int} D \rightarrow \text{int} P_F$ is injective, hence a diffeomorphism (as all points are regular), which in turn gives that $F : D \rightarrow \hat{P}_F$ is a homeomorphism. This completes the proof of the case $n = 2$.

We assume the result true for $(n - 1) \times (n - 1)$ matrices. Take a diagonal $n \times n$ matrix $D$: we consider $\hat{F}(\partial D)$. For a vector $v \in E_i$ of a face of the octant $D \subset \mathbb{R}^n$, the $i$-th column and row of the matrix $D + v \otimes v$ equal $\lambda_i e_i^2$ and $\lambda_i e_i$, so that $\lambda_i$ is one eigenvalue. The remaining eigenvalues belong to the spectrum of $\hat{D} + \hat{v} \otimes \hat{v}$, where $\hat{D}$ is obtained by removing row and column $i$ of $D$ and $\hat{v}$ is obtained from removing the $i$-th entry of $v$. To apply the inductive hypothesis, at the risk of being pedantic, identify $E_i$ with the positive orthant $D \subset \mathbb{R}^{n-1}$. The map $\hat{F} : E_i \rightarrow \mathbb{R}^{n-1}$ is then identified with a map $\hat{F}_D : \hat{D} \rightarrow \mathbb{R}^{n-1}$ which leads to a homeomorphism $F_D : \hat{D} \rightarrow \hat{P}_F$, where

$$\hat{P}_F = [\lambda_1, \lambda_2] \times \ldots \times [\lambda_{i-2}, \lambda_{i-1}] \times [\lambda_{i-1}, \lambda_{i+1}] \times [\lambda_{i+1}, \lambda_{i+2}] \times \ldots \times [\lambda_n, \infty).$$
Notice that both intervals containing $\lambda_i$ in the definition of $\mathcal{P}_F$ were removed and replaced by a single interval $[\lambda_{i-1}, \lambda_{i+1}]$. In order to compute $\tilde{F}(E_i)$, it is a matter of inserting $\lambda_i$ among the ordered eigenvalues of a matrix in $\tilde{F}_D(\hat{D})$. As in the case $n = 2$, there are two possibilities, leading to the creasing of $E_i$ by $F$ for $i > 1$, giving rise to faces of $\mathcal{P}_F$ of the form $L_{i-1} \times \{\lambda_i\} \times R_i$ or $L_i \times \{\lambda_i\} \times R_{i+1}$, where the notation was introduced in the beginning of the section.

Thus, $F : \partial D \to \partial \mathcal{P}_F$ is surjective, and injective on the restriction to each face $E_i \times \hat{D}$. We are left with showing injectivity on the union of the faces. For $v_i \in E_i$ and $v \in D$ for which $\tilde{F}(v_i) = \tilde{F}(v)$, as $\lambda_i$ is one coordinate of $\tilde{F}(v_i)$, by Proposition 2.1(iv) we must have $v \in E_i$. As the restriction of $\tilde{F}$ to $E_i$ is injective, global injectivity in $\partial \hat{D}$ follows. Thus $F : \partial D \to \partial \mathcal{P}_F$ is a homeomorphism. The first part of the argument is complete. The second part follows the case $n = 2$.

We prove the statements about $F^r_S$. Since

$$\text{tr}(S + v \otimes v) = \text{tr} D + |v|^2$$

and $F_S : D \to \mathcal{P}_F$ is a homeomorphism, we have that $F^r_S$ is also a homeomorphism. When restricting to the interior of $D$, $F^r_S$ takes one surface to another and the Jacobian at each point is easily seen to be invertible, showing that on the interior $F^r_S$ is indeed a diffeomorphism.

The proof (i) and (ii) is complete. Statement (iii) is left to the reader. ■

3 Real bordering — Theorem 1.2

Again, it suffices to consider the positive sign sequence $s = (1, \ldots, 1) \in \mathbb{R}^n$, and $S = D$, a diagonal matrix with simple eigenvalues $\lambda_1 < \ldots < \lambda_n$.

Now $D = \mathcal{O}_I \times \mathbb{R}$ has faces of the form $E_i \times \mathbb{R}$ and the parallelepiped

$$\mathcal{P}_G = (-\infty, \lambda_1] \times [\lambda_1, \lambda_2] \times [\lambda_2, \lambda_3] \times \ldots \times [\lambda_n, \infty) \subset \mathbb{R}^{n+1}$$

has $2(n + 1) - 2 = 2n$ faces. Each eigenvalue $\lambda_i$ is associated with two faces,

$$L_{i-1} \times \{\lambda_i\} \times R_i \quad \text{and} \quad L_i \times \{\lambda_i\} \times R_{i+1},$$

where $L_i = (-\infty, \lambda_1] \times \ldots \times [\lambda_{i-1}, \lambda_i]$ and $R_i = [\lambda_i, \lambda_{i+1}] \times \ldots \times [\lambda_n, \infty)$. Define the $(n + 1) \times (n + 1)$ bordered matrix

$$B(v, c) = B_D(v, c) = \begin{pmatrix} D & v^T \\ v & 2c \end{pmatrix}.$$ 

We must show that the map $\tilde{G} : D \to \mathbb{R}^{n+1}$, $(v, c) \mapsto \sigma_0(B(v, c))$ defines a homeomorphism $G : D \to \partial \mathcal{P}_G$. Again, as we shall see, $G$ creases each face of $\partial D$ so as to cover two faces of $\partial \mathcal{P}_G$. 


As before $\mathcal{D}_d$ consists of the points $(v, c) \in \mathcal{D}$ for which $B(v, c)$ has a double eigenvalue, the critical set $\mathcal{C}$ is the set of points in the interior of $\mathcal{D}$ in which $G$ is differentiable with not invertible Jacobian, and its complement in $\mathcal{D}$ is the set of regular points. The counterpart of Proposition 2.1 still holds.

**Proposition 3.1**

(i) $\partial \tilde{G}(\mathcal{D}) \subset \tilde{G}(\partial \mathcal{D} \cup \mathcal{D}_d \cup \mathcal{C})$, (ii) $\mathcal{D}_d \subset \partial \mathcal{D}$, (iii) $\mathcal{C} = \emptyset$, (iv) The matrices $D$ and $B(v, c)$ share an eigenvalue $\lambda_i$ if and only if $v \in E_i$ and $(v, c) \in \partial \mathcal{D}$. Thus, $\partial \tilde{G}(\mathcal{D}) \subset \tilde{G}(\partial \mathcal{D})$.

Thus, at points in int $\mathcal{D}$, $\tilde{G}$ is differentiable.

**Proof:** The proof of item (i) is the same. For (ii), take a double eigenvalue $\rho$ and an associated eigenvector $w \in \mathbb{R}^{n+1}$ with $w_{n+1} = 0$. Expanding $(B(v, c) - \rho)w = 0$ we have that $\rho = \lambda_i$ is an eigenvalue of $D$ and $w = te_i, t \neq 0$. But then $(B(v, c) - \lambda_i)e_i = 0$ implies that $v_i = 0$, so that $v \in E_i$.

For (iii), follow the previous section:

$$J(v, c)(\dot{v}, \dot{c}) = (\langle w_1, \dot{B}w_1 \rangle, \ldots, \langle w_{n+1}, \dot{B}w_{n+1} \rangle),$$

where $B(v, c)w_i = \lambda_i w_i, i = 1, \ldots, n+1, \|w_i\| = 1$ and

$$\dot{B} = \dot{B}(v, c)(\dot{v}, \dot{c}) = \begin{pmatrix} 0 & \dot{v}^T \\ \dot{v} & 2\dot{c} \end{pmatrix}.$$

Write $B = B(v, c)$. Let $\dot{V}$ be the vector space spanned by the matrices $\dot{B}$.

Using Frobenius inner products, the Jacobian becomes

$$J(v, c)(\dot{v}, \dot{c}) = \left( \text{tr}(w_1 \otimes w_1) \dot{B}, \ldots, \text{tr}(w_{n+1} \otimes w_{n+1}) \dot{T} \right)$$

$$= (\langle w_1 \otimes w_1, \dot{B} \rangle, \ldots, \langle w_{n+1} \otimes w_{n+1}, \dot{B} \rangle).$$

Thus, a point $(v, c) \in \text{int} \mathcal{D}$ is critical if and only if some linear combination

$$\sum_{k=1}^{n+1} c_k w_k \otimes w_i = \sum_{k=1}^{n+1} d_k B_k$$

is orthogonal to some nonzero matrix $\dot{B} \in \dot{V}$. Indeed, since the spectrum of $B$ is simple (by Proposition 3.1(ii), as $(v, c) \in \text{int} \mathcal{D}$), a linear combination of the matrices $w_k \otimes w_k$ is some polynomial of $B$. The Frobenius inner product of $\dot{B} \in \dot{V}$ with an arbitrary real symmetric matrix $T$ is simply

$$\langle \dot{B}, \begin{pmatrix} * & y^T \\ y & x \end{pmatrix} \rangle = \left( \begin{pmatrix} 0 & \dot{v}^T \\ \dot{v} & 2\dot{c} \end{pmatrix}, \begin{pmatrix} * & y^T \\ y & x \end{pmatrix} \right) = 2(\langle \dot{v}, \dot{v} \rangle, (x, y)) = 2(\langle \dot{v}, \dot{v} \rangle, Te_{n+1}),$$

where $e_{n+1} = (0, \ldots, 1) \in \mathbb{R}^n$ is canonical. Matrices in $\dot{V}$ are in bijective correspondence with the vectors in $(\dot{v}, \dot{c}) \in \mathbb{R}^{n+1}$. Thus, a point $(v, c)$ corresponding
to a matrix $B = B(v, c)$ is critical if and only if there is a matrix $\hat{B}$ associated with a nonzero $(\hat{v}, \hat{c})$ such that $e_{n+1}, Be_{n+1}, \ldots, B^n e_{n+1}$ are orthogonal to $(\hat{v}, \hat{c})$.

This can happen if and only if the vectors $e_{n+1}, Be_{n+1}, \ldots, B^n e_{n+1}$ are linearly dependent, i.e., $e_{n+1}$ is not a cyclic vector of $B$. Diagonalize $B = Q_B^T D_B Q_B$, where the rows of the orthogonal matrix $Q_B$ are the eigenvectors of $B$ and $D_B$ has simple spectrum, from (ii). The vectors $e_{n+1}, Be_{n+1}, \ldots, B^n e_{n+1}$ are linearly dependent if and only if the vectors $Q_B e_{n+1}, D_B Q_B e_{n+1}, \ldots, D_B^n Q_B e_{n+1}$ are. By an argument with Vandermonde determinants, this is the case if and only if some coordinate of $Q_B e_{n+1}$, the last column of $Q_B$, equals zero. Said differently, the last coordinate of some eigenvector of $B$ is zero. Say $w = (\hat{w}, 0)$ satisfies $(B - \lambda_i)w = 0$. Then $(D - \lambda_i)\hat{w} = 0$, so that $\lambda_i$ is also an eigenvalue of $D$. Since $D$ has simple spectrum, we must have $\hat{w} = \alpha e_k$ for some $\alpha \neq 0$ and $e_k \in \mathbb{R}^n$ a canonical vector. Thus, without loss, $w = (e_k, 0)$ from the equality of entry $(n + 1)$ of $(B - \lambda_i)w = 0$, we have $B_{k,n+1} = B(v, c)_{k,n+1} = v_k = 0$. Thus $v_k \in E_k$ and $(v, c) \in \partial D$. The proof of (iii) is complete.

To prove (iv), simply expand det$(B(v, c) - \lambda I)$ along row $i$. 

**Proof of Theorem 1.2:** The argument is by induction. For $n = 1$, $D = [0, \infty) \times \mathbb{R}$, so that its boundary consists of points of the form $(0, 2c), c \in \mathbb{R}$. The eigenvalues of $B(0, 2c)$ are $\{\lambda_1, 2c\}$ and we must order them. If $2c < \lambda_1$ then

$$G(0, c) = (2c, \lambda_1) \in (-\infty, \lambda_1] \times \{\lambda_1\}.$$  

If $2c > \lambda_1$ then $F(0, c) = (\lambda_1, 2c) \in \{\lambda_1\} \times [\lambda_1, \infty)$. If $2c = \lambda_1$, $G(0, c)$ lies in the common subface, a single point of double spectrum associated with the diagonal matrix $\lambda_1 I$. Again, it is the ordering which creases $\partial D$, a straight line, so as to cover both faces of $\mathcal{P}_G$. The first inductive step is complete.

Take a diagonal $n \times n$ matrix $D$: we consider $G(\partial D)$. The $i$-th face of the octant $E_i = O_Q$ consists of vectors with $i$-th entry equal to zero. The $i$-th column (and row) of the bordered matrix $B(v, c)$ consists of the vector $\lambda_i e_i$, so that $\lambda_i$ is one eigenvalue of $B(v, c)$. The remaining eigenvalues belong to the spectrum of $B_i$, obtained by removing row and column $i$ of $B(v, c)$ — from the inductive hypothesis, $G$ takes bijectively matrices of the form $B(v, c)$, for $v \in E_i$ to their ordered spectrum, consisting of points in the parallelotope

$$(-\infty, \lambda_1] \times [\lambda_1, \lambda_2] \times \ldots \times [\lambda_{i-2}, \lambda_{i-1}] \times [\lambda_{i-1}, \lambda_{i+1}] \times [\lambda_{i+1}, \lambda_{i+2}] \times \ldots \times [\lambda_n, \infty).$$

notice that both intervals containing $\lambda_i$ in the definition of $\mathcal{P}_G$ were removed and replaced by a single interval $[\lambda_{i-1}, \lambda_{i+1}]$.

As in the case $n = 1$, the ordered insertion of $\lambda_i$ among the ordered eigenvalues of the matrices $B(M)$ for $M \in E_i \times \mathbb{R}$ provides a (creasing) homeomorphism between the face $E_i \times \mathbb{R}$ of $D$ and $(L_{i-1} \times \{\lambda_i\} \times R_i) \cup (L_i \times \{\lambda_i\} \times R_{i+1})$.

Thus, $G : \partial D \to \partial \mathcal{P}_G$ is surjective, and restricts injectively to each face $E_i \times \mathbb{R}$.

We are left with showing injectivity on the union of the faces.
For $(v, c) \in E_i \times \mathbb{R}$ set $M = B(v, c)$ and consider another bordered matrix $N = B(w, d)$ with the same ordered spectrum. As $\lambda_i$ belongs to $\lambda(M)$, we must have $\lambda_i \in \lambda(N)$ and, from the previous lemma, $w_i = 0$: $(w, v) \in E_i \times \mathbb{R}$. Global injectivity now follows from injectivity of $G$ restricted to $E_i \times \mathbb{R}$.

The rest of the argument mimics the previous proof.

References

[1] M.F. ATIYAH, *Convexity and Commuting Hamiltonians*, Bull. London Math. Soc. 14 (1982), pp. 1–5.

[2] H. DUISTERMAAT, The momentum map, in Topics in Differential Geometry I and II, Colloq. Math. Soc. Janos Bolyai, 46 (1988), pp. 347–392.

[3] V. GUILLEMIN AND S. STERNBERG, *Convexity Properties of Moment Mappings*, Inv. Math. 67 (1981), pp. 491–513.

[4] A. HORN, *Doubly Stochastics Matrices and the Diagonal of a Rotation Matrix*, Amer. J. of Math. 76 (1954), pp. 620–630.

[5] A. HORN, *Eigenvalues of Sums of Hermitian Matrices*, Pac. J. Math. 12 (1962), pp. 225–241.

[6] R. HORN and C. JOHNSON, *Matrix Analysis*, 2nd ed., Cambridge U. Press, 2013.

[7] F. KIRWAN, *Convexity properties of the moment mapping, III*, Inv. Math. 77 (1984), pp. 547–552.

[8] A.A. KLYACHKO, *Stable vector bundles and Hermitian operators*, Selecta Math. (N.S.) 4 (1998), pp. 419–445.

[9] A. KNUTSON AND T. TAO, *Honeycombs and Sums of Hermitian Matrices*, Notices Amer. Math. Soc. 48 (2001), pp. 175–186.

[10] T. MACIAZEK and U. SMILANSKY, *Can one hear a matrix? Recovering a real symmetric matrix from its spectral data*, arXiv:2106.01819.

[11] E. OUTERELO AND J.M.RUIZ, *Mapping degree theory*, GTM 108 (2009), AMS.

[12] I. SCHUR, *¨Uber eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie*, S. B. Berlin Math. Ges. 22 (1923), pp. 9–20.

[13] R. SJAMAAR, *Convexity properties of the moment mapping re-examined*, Adv. Math 138 (1998), pp. 46–91.
[14] R.C. THOMPSON, *High, Low, and Quantitative Roads in Linear Algebra*, Lin. Alg. Appl. 162 (1992), pp. 23–64.

Ricardo S. Leite, Departamento de Matemática, UFES
Av. Fernando Ferrari, 514, Vitória, ES 29075-910, Brazil
Carlos Tomei, Departamento de Matemática, PUC-Rio
R. Marquês de S. Vicente 225, Rio de Janeiro, RJ 22453-900, Brazil
rsleite@ufes.br
carlos.tomei@puc-rio.br