# SOME REMARKS ON THE GEOMETRY OF SOME SURFACES OF MATRICES ASSOCIATED WITH TOEPLITZ EIGENPROBLEMS

KENNETH R. DRIESSEL*

## CONTENTS



**Introduction.** I am interested in the following problem: Given a Toeplitz matrix, find its singular values. This problem is quite important. For example, consider a time-invariant, causal linear system. Such a system can be represented by means of a Volterra integral operator $T$ having the following form:

$$Tx(t) := \int_{-\infty}^{t} h(t-s)x(s)ds.$$

*Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN 55455 and Department of Mathematics , Idaho State University, Campus Box 8085, Pocatello, ID 83209-8085

(See, e.g., Naylor and Sell 1982, section 2.8 "System types".) Associated with such an operator we typically have the follow "deconvolution" problem: Given $T$ and $y$, find $x$ which satisfies $Tx = y$. When the deconvolution problem is discretized, we usually obtain a "Toeplitz least-squares inversion" problem: Given Toeplitz matrix $T \in \mathbb{R}^{m\times n}$ and $y \in \mathbb{R}^m$, find $x \in \mathbb{R}^n$ which satisfies $Tx = y$. Recall that a matrix is *Toeplitz* if it has constant diagonals; that is, $T(i,j)$ only depends on the difference $i-j : T(i,j) = h(i-j)$. For example, when $m = 3$ and $n = 2$, we have

$$T = \begin{pmatrix} h_0 & h_{-1} \\ h_1 & h_0 \\ h_2 & h_1 \end{pmatrix} .$$

There are standard algorithms for solving equations of the form $Ax = b$ in the least squares sense. These algorithms generally find an orthogonal-triangular factorization of the coefficient matrix: $A = QR$. (See, e.g., Golub and Van Loan 1983 and 1989.) These algorithms generally require $O(mn^2)$ floating point operations to find solutions. Toeplitz inversion problems are often quite large annd ill-conditioned. (For example, in seismic oil exploration, $m$ and $n$ often exceed 1000.) Special algorithms have been developed to take advantage of the Toeplitz structure of the coefficient matrix. (See, e.g., Golub and Van Loan 1983 section 5.7 "Toeplitz systems" and the references at the end of that section for early work; for more recent work, see Cybenko (1987), Cybenko and Barry (1990), Nagy (1991), Nagy and Plemmons (1991), Bojanczyk, Brent and de Hoog (1986), Chun, Kailath and Lev-Ari (1987), Sweet (1984), Sweet (1991), Heinig and Rost (1984) and references in these reports.)

When solving $Ax = b$, it is is often useful to have the singular value decomposition of the coefficient matrix: $A = U\Sigma V^T$ where $U$ and $V$ are unitary and $\Sigma$ is diagonal. This decomposition is especially useful when $A$ is ill-conditioned. (See, e.g., Golub and Van Loan 1983 section 6.5 "Rank deficiency II: The singular value decomposition".) There are standard algorithms for computing this decomposition. (see, e.g., Golub and Van Loan 1983 section 8.3 "Once again: the singular value decomposition".) These algorithms generally require $O(mn^2)$ floating point operations ("flops"). The matrix $A$ is first reduced to a bidiagonal matrix by means of orthogonal transformations. (This non-iterative reduction step requires $O(mn^2)$ flops.) The computation is then finished by means of an iterative procedure that preserves singular values, preserves the bidiagonal structure and tends toward a diagonal matrix. In particular this iteration is efficient because it preserves the bidiagonal structure. I expect that there is an analogous iterative procedure for finding the singular values of a Toeplitz matrix. In particular I want to find an iterative procedure that preserves singular values, preserves an appropriate Toeplitz-like structure and tends toward a matrix which has easily computable singular values. In this report we study the geometry of some surfaces which I believe may be the fundamental objects related to my goal. In particular, we review the (apparently) well-known properties of iso-singular surfaces. We also need to pick a geometrical object which embodies a useful representation of "Toeplitz-like" structure. I believe that the notion of "displacement rank" is the appropriate concept. The importance of this notion has been emphasized by T. Kailath and his co-workers. (See, e.g., Kailath, Kung and Morf 1979a,b and Chun and Kailath 1991. See also Heinig and Rost (1984.) In particular, we study the (apparently) well-known properties of surfaces of matrices with given rank. We also study the associated surfaces of matrices with given displacement rank. In order to preserve a particular set of singular values and preserve a particular displacement rank, an iterative procedure needs to stay in the intersection of these two surfaces. We also study this intersection.

Another way to solve $Ax = b$ in the least squares sense is to solve the "normal" equation: $A^*Ax = A^*b$. Here the coefficient matrix $A^*A$ is hermitian. Consequently, we also study the appropriate surfaces in the space of hermitian matrices-namely, surfaces of hermitian matrices with a particular set of eigenvalues, surfaces of hermitian matrices with a particular inertia and surfaces of hermitian matrices with a particular displacement inertia. On an infinitesimal scale, the paths preserving structure (i.e., staying in the surface) correspond to directions which preserve the structure (i.e., vectors tangent to the surface). Consequently, we also compute the spaces tangent to the surfaces of interest.

I shall now describe more precisely the contents of this report and its motivation. Let $m$ be a positive integer. I use $\mathrm{Herm}(m)$ to denote the set of hermitian matrices; in set-theoretic symbols,

$$\mathrm{Herm}(m) := \{X \in \mathbb{C}^{m\times m} : X^* = X\},$$

where $X^* := \overline{X}^T$ denotes the conjugate transpose of $X$ and $\mathbb{C}^{m\times m}$ denotes the set of $m\times m$ complex matrices. Let $p$ and $n$ be nonnegative integers satisfying $p+n \leq m$. A hermitian matrix $A$ has *inertia* $(p, n, z)$ where $z := m - (p+n)$ if $A$ has $p$ positive eigenvalues, $n$ negative eigenvalues and $z$ zero eigenvalues. Let $Z_m$ denote the $m \times m$ (*lower*) *shift* matrix defined by

$$Z_m(i,j) := \delta(i, j+1)$$

where $\delta$ is the Kronecker delta. (I write $Z$ in place of $Z_m$ when $m$ is easily determined from the context.) For example,

$$Z_3 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} .$$

Let $\lrcorner$ be the linear map defined by

$$\lrcorner := \mathbb{C}^{m\times m} \to \mathbb{C}^{m\times m} : X \to X - Z_m X Z_m^* .$$

Note that this linear map preserves $\mathrm{Herm}(m)$; that is, $X \in \mathrm{Herm}(m)$ implies $\lrcorner X \in \mathrm{Herm}(m)$. (In fact, since $\lrcorner$ is invertible, $\lrcorner\,\mathrm{Herm}(m) = \mathrm{Herm}(m)$.) A hermitian matrix $X$ has *displacement inertia* $(p, n, z)$ if $\lrcorner X$ has inertia $(p, n, z)$. T. Kailath appears to be one of the first to emphasize the importance of displacement structure of matrices. (Again see Kailath, Kung and Morf 1979a,b.) I want to review a few of the major results in this area in order to illustrate the significance of this concept. Note first that hermitian Toeplitz matrices usually have displacement inertia $(1, 1, m-2)$. For example, if

$$A := \begin{pmatrix} t_0 & t_1 & t_2 \\ \bar{t}_1 & t_0 & t_1 \\ \bar{t}_2 & \bar{t}_1 & t_0 \end{pmatrix}$$

then

$$A - ZAZ^* = \begin{pmatrix} t_0 & t_1 & t_2 \\ \bar{t}_1 & t_0 & t_1 \\ \bar{t}_2 & \bar{t}_1 & t_0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & t_0 & t_1 \\ 0 & \bar{t}_1 & t_0 \end{pmatrix}$$

$$= \begin{pmatrix} t_0 & t_1 & t_2 \\ \bar{t}_1 & 0 & 0 \\ \bar{t}_2 & 0 & 0 \end{pmatrix} = \begin{pmatrix} t_0/2 \\ \bar{t}_1 \\ \bar{t}_2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} t_0/2 & t_1 & t_2 \end{pmatrix}$$

$$= te_1^* + e_1 t^* = (1/2)((t+e_1)(t+e_1)^* - (t-e_1)(t-e_1)^*)$$

where $t := (t_0/2, \bar{t}_1, \bar{t}_2)^T$ and $e_1 := (1\ 0\ 0)^T$. Hermitian matrices with low displacement inertia (i.e., small $p$ and small $n$) are regarded as being "near Toeplitz". The following result shows that displacement inertia is preserved (loosely speaking) under inversion. I learned its proof from Tiberiu Constantinescu (Institute of Mathematics of the Romanian Academy of Sciences). As far as I know this proof has never been published. However, the result is closely related to the inversion theorem for displacement rank that appears in Kailath, Kung and Morf (1979a,b) and which I review below. Let $E := E_m$ denote the $m \times m$ *exchange* matrix defined by

$$E_m(i,j) := \delta(i, m+1-j).$$

For example,

$$E_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

(This terminology and notation appears in Golub and Van Loan 1983 section 5.7 "Toeplitz systems".) Note that $E^{-1} = E^* = E$.

THEOREM. *Let $A$ be an invertible hermitian matrix. Then the displacement inertia of $EA^{-1}E$ is the same as the displacement inertia of $A$.*

*Proof.* Recall that inertia is preserved under congruence; that is, if $R$ is an invertible matrix then the inertia of $RXR^*$ is the same as the inertia of $X$. The following congruence calculations often occur in connection with discussions of "Schur complements". (See, e.g., Horn and Johnson 1985 p. 472.) Consider the following block matrix

$$\begin{pmatrix} A & B \\ B^* & C \end{pmatrix} .$$

Let $X := -A^{-1}B$. Then

$$\begin{pmatrix} I & 0 \\ X^* & I \end{pmatrix} \begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & C - B^*A^{-1}B \end{pmatrix} .$$

Let $Y := -C^{-1}B$. Then

$$\begin{pmatrix} I & Y^* \\ 0 & I \end{pmatrix} \begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \begin{pmatrix} I & 0 \\ Y & I \end{pmatrix} = \begin{pmatrix} A - BC^{-1}B^* & 0 \\ 0 & C \end{pmatrix} .$$

In particular, we have that the following matrices are congruents

$$\begin{pmatrix} A & Z \\ Z^* & A^{-1} \end{pmatrix} \simeq \begin{pmatrix} A & 0 \\ 0 & A^{-1} - Z^*A^{-1}Z \end{pmatrix} \simeq \begin{pmatrix} A - ZAZ^* & 0 \\ 0 & A^{-1} \end{pmatrix} .$$

Since the inertia of $A$ is the same as the inertia of $A^{-1}$ we have

$$\text{inertia}(A - ZAZ^*) = \text{inertia}(A^{-1} - Z^*A^{-1}A) .$$

Note that $EZE = Z^*$. Hence

$$\begin{aligned} \text{inertia}(A^{-1} - Z^*A^{-1}Z) &= \text{inertia}\, E(A^{-1} - EZEA^{-1}EZ^*E)E \\ &= \text{inertia}(EA^{-1}E - ZEA^{-1}EZ^*) . \qquad \square \end{aligned}$$

Displacement structure can also be defined for matrices that are not square. Let $\mathbb{C}^{m\times n}$ denote the space of $m \times n$ complex matrices. Let $\lrcorner$ be the linear map defined as follows:

$$\lrcorner := \mathbb{C}^{m\times n} \to \mathbb{C}^{m\times n} : X \to X - Z_m X Z_n^* .$$

A matrix $X \in \mathbb{C}^{m\times n}$ has *displacement rank* $k$ if $\lrcorner X$ has rank $k$. Note that Toeplitz matrices usually have displacement rank 2. For example, if

$$A := \begin{pmatrix} t_0 & t_1 & t_2 \\ t_{-1} & t_0 & t_1 \\ t_{-2} & t_{-1} & t_0 \end{pmatrix}$$

then

$$A - ZAZ^* = \begin{pmatrix} t_0 & t_1 & t_2 \\ t_{-1} & 0 & 0 \\ t_{-2} & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ t_{-1} \\ t_{-2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} t_0 & t_1 & t_2 \end{pmatrix}$$

Matrices with low displacement rank are regarded as being "near Toeplitz". The following result shows that displacement rank is preserved (loosely speaking) under inversion. The theorem is from Kailath, Kung and Morf (1979a,b). I include a proof which has the style of the proof of the previous theorem.

THEOREM. *Let $A \in \mathbb{C}^{n\times n}$ be an invertible matrix. Then the displacement rank of $EA^{-1}E$ is the same as the displacement rank of $A$.*

*Proof.* Recall that rank is preserved under equivalence; that is, if $R$ and $S$ are invertible matrices, then the rank of $RXS$ is the same as the rank of $X$. Consider the following block matrix:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

We have

$$\begin{pmatrix} I & 0 \\ -CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & -A^{-1}B \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix} \begin{pmatrix} I & -A^{-1}B \\ 0 & I \end{pmatrix}$$

$$= \begin{pmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{pmatrix}$$

We also have

$$\begin{pmatrix} I & BD^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & 0 \\ -D^{-1}C & I \end{pmatrix} = \begin{pmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} I & 0 \\ -D^{-1}C & I \end{pmatrix}$$

$$= \begin{pmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{pmatrix} .$$

In particular, we then have the following equivalences:

$$\begin{pmatrix} A & Z \\ Z^* & A^{-1} \end{pmatrix} \simeq \begin{pmatrix} A & 0 \\ 0 & A^{-1} - Z^* A^{-1} Z \end{pmatrix} \simeq \begin{pmatrix} A - ZAZ^* & 0 \\ 0 & A^{-1} \end{pmatrix} .$$

Since $\operatorname{rank} A = \operatorname{rank} A^{-1} = n$, we have

$$\begin{aligned} \operatorname{rank}(A - ZAZ^*) &= \operatorname{rank}(A^{-1} - Z^* A^{-1} Z) \\ &= \operatorname{rank} E(A^{-1} - EZEAAEZ^*E)E \\ &= \operatorname{rank}(EA^{-1}E - ZEAAEZ^*) . \qquad \square \end{aligned}$$

Other versions of displacement structure can be defined by using other matrices in place of the shift matrix $Z$. Theorems analogous to the last two can often be proved for these other versions. See, e.g., Chun and Kailath (1991) and Heinig and Rost (1984). The following inequality shows that if $A$ has small displacement rank then so does its pseudo-inverse $A^+$:

$$\operatorname{rank}(A^+ - Z^* A^+ Z) \leq 2 \operatorname{rank}(A - ZAZ^*) .$$

This result is due to Comon (1992).

This report is divided into a main part (which in turn is divided into a number of sections) and one appendix. In the section entitled "Orbits and Quotients" I present a review of some of the basic results about Lie groups and homogeneous spaces which are used later. (For an alternative, more elementary and more complete exposition of these results see Driessel 1991.) In particular I consider the orbits of points under the action of a Lie group. Under fairly general conditions such orbits are homogeneous spaces.

In the next four sections I consider some spaces of rectangular matrices. I usually only consider generic cases of such spaces. For any $n$-tuple $\sigma := (\sigma_1, \sigma_2, \ldots, \sigma_n) \in \mathbb{R}^n$ of nonnegative real numbers let $\operatorname{sing}^{-1}(\sigma)$ denote the subspace of $\mathbb{C}^{m\times n}$, where $n \leq m$, consisting of the $m \times n$ matrices with these numbers as singular values; in symbols,

$$\operatorname{sing}^{-1}(\sigma) := \{A \in \mathbb{C}^{m\times n} : \ \text{singular-spectrum} \ \ A = \sigma\} .$$

In the section on "Matrices With Fixed Singular Values", I show that this space is homogeneous. I also determine its associated tangent spaces. For any positive integer $k \leq n$ let $\operatorname{rank}^{-1}(k)$ denote the subspace of $\mathbb{C}^{m\times n}$ consisting of the matrices with rank $k$; in symbols,

$$\operatorname{rank}^{-1}(k) := \{A \in \mathbb{C}^{m\times n} : \operatorname{rank} A = k\} .$$

In the section on "Matrices With Fixed Rank" I show that this space is homogeneous. I also determine its associated tangent spaces. Let $\text{dis-rank}^{-1}(k)$ denote the subspace of $\mathbb{C}^{m\times n}$ consisting of the matrices with displacement rank $k$; in symbols,

$$\text{dis-rank}^{-1}(k) := \{A \in \mathbb{C}^{m\times n} : \text{displacement-rank}\, A = k\}.$$

In the section on "Matrices With Fixed Displacement Rank" I show that this space is homogeneous. I also determine its associated tangent spaces. In the section on "Matrices With Fixed Singular Values and Fixed Displacement Rank" I consider spaces of the form $\text{sing}^{-1}(\sigma) \cap \text{dis-rank}^{-1}(k)$. I show that the spaces $\text{sing}^{-1}(\sigma)$ and $\text{dis-rank}^{-1}(k)$ usually intersect transversely.

In the last four sections I consider some spaces of hermitian matrices. Let $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{R}^m$ be an ordered $m$-tuple of real numbers and let $\text{spec}^{-1}(\lambda)$ denote the subspaces of $\text{Herm}(m)$ consisting of the matrices with these numbers as eigenvalues; in symbols,

$$\text{spec}^{-1}(\lambda) := \{A \in \text{Herm}(m) : \text{spectrum}\, A = \lambda\}.$$

In the section on "Hermitian Matrices With Fixed Eigenvalues" I show that this space is homogeneous. I also determine its associated tangent spaces. Let $p$ and $n$ be nonnegative integers satisfying $0 < p + n \leq m$ and let $\text{inert}^{-1}(p, n)$ denote the subspace of $\text{Herm}(m)$ consisting of the matrices with inertia $(p, n, z)$ where $z := m - (p + n)$; in symbols,

$$\text{inert}^{-1}(p, n) := \{A \in \text{Herm}(m) : \text{inertia}\, A = (p, n, z)\}.$$

In the section on "Hermitian Matrices With Fixed Inertia" I show that this space is homogeneous. I also determine its associated tangent spaces. Let $\text{dis-inert}^{-1}(p, n)$ denote the subspace of $\text{Herm}(m)$ consisting of the matrices with displacement inertia $(p, n, z)$; in symbols,

$$\text{dis-inert}(p, n) := \{A \in \text{Herm}(m) : \text{displacement-inertia}\, A = (p, n, z)\}.$$

In the section on "Hermitian Matrices With Fixed Displacement Inertia" I show that this space is homogeneous. I also determine its associated tangent spaces. In the section on "Hermitian Matrices With Fixed Eigenvalues and Fixed Displacement Inertia" I consider spaces of the form $\text{spec}^{-1}(\lambda) \cap \text{dis-inert}^{-1}(p, n)$. I show that the spaces $\text{spec}^{-1}(\lambda)$ and $\text{dis-inert}^{-1}(p, n)$ usually intersect transversely.

In the appendix "Tridiagonal Hermitian Matrices With Fixed Eigenvalues" I consider spaces of the form $\text{spec}^{-1}(\lambda) \cap \text{Tridiag}(m)$ where $\text{Tridiag}(m)$ is the space consisting of tridiagonal hermitian matrices. I explicitly describe the tangent spaces associated with this intersection space. The main result in this appendix is a prototype of the kind of result that I would like to discover about the intersection space $\text{spec}^{-1}(\lambda) \cap \text{dis-inert}^{-1}(p, n)$.

I have sometimes repeated material in order to make this report more readable. In particular, most of the sections can be read somewhat independently of the others. However, most of them depend on the results presented in the section on "Orbits and Quotients".

In my report Driessel (1991) I treated questions like the ones treated here. However, that earlier report deals with matrices over the real numbers rather than over the complex numbers.

**Orbits and quotients.** In this section we shall review some results about topological groups. In particular, we shall consider the relationship between orbits and quotient spaces. (I mainly follow Varadarajan 1984 which may be consulted for further related theory and details. I also presented this material in detail in Driessel 1991.)

Let $G$ be a group and let $M$ be a set. Then a map $\alpha : G \times M \to M$ is an *action* of $G$ on $M$ (or a *transformation group*) if it satisfies the following conditions for all $g, h \in G$ and $m \in M$:

(1)
$$e * m = m \ , \quad \text{and}$$
$$(gh) * m = g * (h * m)$$

where $e$ is the identity element of $G$ and $h * m := \alpha(h, m)$.

For $m \in M$ let the *orbit* of $m$ be the set of all translates of $m$ under the group action; in set theoretic symbols,

$$\mathrm{Orbit}(m) := G * m := \{g * m : g \in G\}.$$

There is a close relationship between orbits in $M$ and certain quotients of $G$. Let $H$ be a subgroup of $G$. Then $G/H$ denotes the collection of all (left) cosets of $G$ modulo $H$; in symbols,

$$G/H := \{kH : k \in G\}.$$

This collection is called the *quotient* of $G$ modulo $H$. Let $m \in M$. Then the *isotropy* (or *stability*) subgroup determined by $m$ is the subgroup of $G$ consisting of the transformations that fix $m$; in symbols,

$$\mathrm{Fix}(m) := \{g \in G : g * m = m\}.$$

PROPOSITION. *Relation between isotropy subgroups and orbits.*

*Let $m \in M$. Then the following relation is a bijection:*

$$G/\,\mathrm{Fix}(m) \to G * m : g\,\mathrm{Fix}(m) \to g * m.$$

*Proof.* We have the following equivalent conditions for $g, h \in G : g\,\mathrm{Fix}(m) = h\,\mathrm{Fix}(m); h^{-1}g \in \mathrm{Fix}(m); (h^{-1}g) * m = m;\ g * m = h * m.$ □

Proposition. *Relation between isotropy groups.*

*Let $p, q \in M$. If $p$ is in the orbit of $q$ then $p$ and $q$ have the same orbits and their isotropy groups are conjugate.*

*Proof.* Assume $p \in \mathrm{Orbit}(q) = G * q$. Then there exists an element $g \in G$ satisfying $p = g * q$. We then have the following equivalent conditions: $m \in G * p$; $\exists h \in G, m = h * p$; $\exists h \in G, m = (hg) * q; m \in G * q$. We also have the following equivalent conditions: $h \in \mathrm{Fix}(p); h * p = p; hg * q = g * q; g^{-1}hg * q = q; g^{-1}hg \in \mathrm{Fix}(q); h \in q\,\mathrm{Fix}(q)g^{-1}$. □

We now add some topology to our discussion. Let $G$ be a topological group. Let $U$ be a subset of $G$; then $U$ is *symmetric* if $U = U^{-1}$. We shall need the following result in one of our later proofs.

Proposition. *Symmetric neighborhoods of the identity.*

*Let $U$ be an open neighborhood of the identity $e$ in $G$. Then there exists an open symmetric neighborhood $W$ of $e$ satisfying $WW \subseteq U$.*

*Proof.* Since the product map $G \times G \to G : (g, h) \to gh$ is continuous and $ee = e$, there exist open neighborhoods $V_1$ and $V_2$ of $e$ such that $V_1 V_2 \subseteq U$. Take $W := V_1 \cap V_1^{-1} \cap V_2 \cap V_2^{-1}$. Then $WW \subseteq V_1 V_2 \subseteq U$. □

Let $G$ be a topological group, let $H$ be a subgroup of $G$ and let $T$ denote the natural map of $G$ onto $G/H$; in symbols,

$$T := G \to G/H : g \to gH.$$

(I use ":=" because $T$ denotes the map defined by the given expression.) We topologize the quotient by taking the largest topology on it that makes $T$ continuous; in other words, the topology on $G/H$ is the following collection of sets:

$$\{U \subseteq G/H : T^{-1}(U) \quad \text{is open in} \quad G\}.$$

The space $G/H$ with this topology is called the *quotient* (or *coset*) space of $G$ modulo $H$. Spaces homeomorphic to such quotient spaces are called *homogeneous* spaces.

*Remarks.* A group action is *transitive* on $N \subseteq M$ if $\forall\, p, q \in N, \exists g \in G,\ g * p = q$. Note that the following group action is transitive on $G/H$:

$$\beta := G \times G/H \to G/H : (g, kH) \to gkH.$$

Also note that for any $g \in G$, the following map is a homeomorphism:

$$\beta_g := G/H \to G/H : kH \to gkH.$$

Hence, from a topological point of view, all the points of $G/H$ look the same. Thus the word "homogeneous" is appropriate. Felix Klein emphasized the importance of homogeneous spaces in his Erlangen program. $\square$

PROPOSITION. *The quotient space.*

*Let $G$ be a topological group, let $H$ be a subgroup of $G$, let $T$ be the natural map of $G$ onto $G/H$ and let $G/H$ have the quotient topology. Then $T$ is continuous and open. Furthermore, if $G$ is Hausdorff and $H$ is closed then $G/H$ is Hausdorff.*

*Proof.*

*Claim.* The map $T$ is continuous.

Let $U$ be an open subset of $G/H$. Then $T^{-1}(U)$ is open in $G$ by the definition of the quotient topology.

*Claim.* For every subset $U$ of $G, T^{-1}(T(U)) = UH$.

Note $T(U) = \{Tu : u \in U\} = \{uH : u \in U\}$. If $Tg \in T(U)$ then $gH = uH$ for some $u \subseteq U$ and hence $g \in uH \subseteq UH$. On the other hand, if $g = uh$ where $u \in U$ and $h \in H$ then $Tg = gH = uhH = uH \in T(U)$.

*Claim.* The map $T$ is open.

Let $U \subseteq G$ be open. We want to see that $T(U)$ is open in $G/H$, i.e., that $T^{-1}(T(U)) = UH$ is open in $G$. Now $UH = \cup\{Uh : h \in H\}$ and each $Uh$ is open since the map $G \to G : x \to xh$ is a homeomorphism. Thus $UH$ is the union of open sets and hence is open.

*Claim.* If $H$ is closed then $G/H$ is Hausdorff.

Let $xH \neq yH$ where $x, y \in G$. Then $x \notin yH$ and $yH$ is closed. Now there exists a neighborhood of $U$ of $e$ such that $Ux \cap yH = \emptyset$ (e.g., take $U := G - yHx^{-1}$ and note that $e = xx^{-1} \notin yHx^{-1}$ and $yHx^{-1}$ is closed). Let $W$ be a symmetric open neighborhood of $e$ satisfying $WW \subseteq U$. Note $xH \subseteq WxH, yH \subseteq WyH$ and that $WxH = \cup\{Wxh : h \in H\}$ and $WyH = \cup\{Wyh : h \in H\}$ are open sets. To finish the proof of the claim, we need only see that $W(xH) \cap W(yH) = \emptyset$. Suppose not; i.e., suppose there exists $w_1, w_2 \in W$ and $h_1, h_2 \in H$ satisfying $w_1xh_1 = w_2yh_2$. Then $w_2^{-1}w_1x = yh_2h_1^{-1} \in Ux \cap yH$ which is a contradiction. $\square$

We now add some manifold structure to our discussion. Let $M$ and $N$ be smooth manifolds and let $f : M \to N$ be a smooth map from $M$ into $N$. Then $f$ is an *immersion* if, for all $x \in$M, the derivative map $df(x) : Tan.M.x \to Tan.N.f(x)$ from the space tangent to $M$ at $x$ into the space tangent to $N$ at $f(x)$ is injective. Such a map $f$ is an *imbedding* if it is a one-to-one immersion; it is a *regular imbedding* if it is an imbedding and if $f$ is a homeomorphism of $M$ onto $f(M)$ where $f(M)$ has the subspace (or relative) topology determined by $N$. If $M$ is a subset of $N$, then $M$ is a *submanifold* of $N$ if the inclusion map is an imbedding and $M$ is a *regular submanifold* of $N$ if the inclusion map is regular.

Let $G$ be a Lie group, let $M$ be a smooth manifold and let $\alpha : G \times M \to M$ be a smooth action of $G$ on $M$. In general, orbits need not be regular submanifolds of $M$. For example, consider an irrational Kronecker flow on the 2-torus $T^2 := S^1 \times S^1$; i.e., consider the following differential equations

$$\dot{x}_1 = r_1 \qquad \pmod 1$$
$$\dot{x}_2 = r_2 \qquad \pmod 1$$

where $r_1/r_2$ is an irrational number. Any orbit of this flow is a dense subset of the torus. (For a detailed discussion of such flows see, for example, Arnold 1973 section 24.3 "Phase curves ... on the torus".) The following theorem provides necessary and sufficient conditions for regularity of orbits.

THEOREM. *Canonical homeomorphism between orbits and quotients.*

*Let $M$ be a smooth manifold, and let $G$ be a Lie group acting smoothly on $M$, and let $m \in M$. Then the orbit $G * m$ of $m$ is a smooth submanifold of $M$ with the following dimension:*

$$\dim(G * m) = \dim G - \dim \mathrm{Fix}(m).$$

*The orbit of $m$ is a regular submanifold of $M$ if and only if it is locally closed in $M$.*

For a proof of this theorem see Varadarajan (1984) section 2.6 "Closed Lie subgroups and homogeneous spaces. Orbits and spaces of orbits" or see Driessel (1991). A similar result appears in Montgomery and Zippin (1955) section 2.13 "Relation of homogeneous spaces and coset spaces"; they attribute the result to Arens (1946).

**Matrices with fixed singular values.** Let $\sigma := (\sigma_1, \sigma_2, \ldots, \sigma_n)$ be an ordered $n$-tuple of nonnegative real numbers satisfying $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ and let $\mathrm{sing}^{-1}(\sigma)$ denote the space (with the relative topology) of all $m \times n$ complex matrices having these numbers as singular values; we assume that $m \geq n$; in symbols,

$$\mathrm{sing}^{-1}(\sigma) := \{A \in \mathbb{C}^{m\times n} : \quad \text{singular-spectrum } (A) = \sigma\}.$$

We call $\mathrm{sing}^{-1}(\sigma)$ the "isosingular" space determined by $\sigma$. We want to see that this space is homogeneous; i.e., that it is homeomorphic to the quotient of a Lie group.

We first want to see that $\mathrm{sing}^{-1}(\sigma)$ is the orbit of a certain group action on $\mathbb{C}^{m\times n}$. Let $U(n)$ denote the group of unitary $n \times n$ complex matrices. Define the smooth action $\alpha$ by

$$\alpha := U(m) \times U(n) \times \mathbb{C}^{m\times n} \to \mathbb{C}^{m\times n} : (U, V, A) \to UAV^*.$$

Two matrices $A$ and $B$ in $\mathbb{C}^{m\times n}$ are *unitarily equivalent* if there exists $U \in U(m)$ and $V \in U(n)$ such that $B = UAV^*$. Recall that any matrix $A$ with singular spectrum $\sigma$ is unitarily equivalent to the $m \times n$ diagonal matrix with diagonal entries $\sigma_1, \sigma_2, \ldots, \sigma_n$; i.e., there exists $U \in U(m)$ and $V \in U(n)$ such that $UAV^* = \Sigma$ where $\Sigma := \operatorname{diag}(\sigma)$. (See, for example, Horn and Johnson 1985 section 7.3: "Polar and singular value decomposition".) Thus $\operatorname{sing}^{-1}(\sigma)$ is the orbit of $\Sigma$ under the group action given above. We also have that the following relation is a bijection:

$$(U(m) \times U(n))/\operatorname{Fix}(\Sigma) \to \operatorname{sing}^{-1}(\sigma) : (U,V)\operatorname{Fix}(\Sigma) \to U\Sigma V^*.$$

PROPOSITION. *The $m \times n$ complex matrices with a given singular spectrum form a homogeneous space.*

*Proof.* In view of the remarks preceding this proposition, we need only check the hypotheses of the theorem concerning the homeomorphism of orbits and quotient spaces. Note that $U(m)\times U(n)$ is a compact topological space (since it is closed and bounded). We finish the proof by noting that $\operatorname{sing}^{-1}(\sigma)$ is compact since it is the image of $U(m) \times U(n)$ under the following smooth map:

$$U(m) \times U(n) \to \operatorname{orbit}(\Sigma) : (U,V) \to U\Sigma V^*. \qquad \square$$

We want to determine $\operatorname{Fix}(\Sigma)$ is a generic case.

PROPOSITION. *Let $m \geq n$ be positive integers. Let $\sigma := (\sigma_1, \sigma_2, \ldots, \sigma_n) \in \mathbb{R}^n$ and let $\Sigma := \operatorname{diag}(\sigma)$ be the $m \times n$ diagonal matrix with diagonal entries $\sigma_1, \sigma_2, \ldots, \sigma_n$. If the $\sigma_i$ are distinct positive numbers then $\operatorname{Fix}(\Sigma)$ consists of the pairs $(U,V) \in U(m) \times U(n)$ satisfying the following conditions:*

*(i) there exist $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{C}$ such that $V = \operatorname{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and $|\alpha_i| = 1$*

*(ii) there exists $U_{22} \in U(m-n)$ such that*

$$U = \begin{pmatrix} V & 0 \\ 0 & U_{22} \end{pmatrix} .$$

*Proof.* Let $D$ be the diagonal $n \times n$ matrix with diagonal entries $\sigma_1, \sigma_2, \ldots, \sigma_n$; that is, $D := \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$. Then

$$\Sigma = \begin{pmatrix} D \\ 0 \end{pmatrix} .$$

Let $U \in U(m), V \in U(n)$ satisfy $U\Sigma V^* = \Sigma$. Then we have

$$D^2 = \Sigma^*\Sigma = (U\Sigma V^*)^*(U\Sigma V^*) = V\Sigma^2 V^*.$$

Hence $V$ commutes with $D^2$. It follows that $V$ is a diagonal matrix. Let $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{C}$ satisfy $V = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$. Since $VV^* = I$ we get $|\alpha_i| = 1$. Now define $U_{11} \in \mathbb{C}^{n\times n}, U_{12} \in \mathbb{C}^{n\times(m-n)}, U_{21} \in \mathbb{C}^{(m-n)\times n}$ and $U_{22} \in \mathbb{C}^{(m-n)\times(m-n)}$ by

$$\begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} := U \ .$$

We have

$$\begin{pmatrix} U_{11}D \\ U_{12}D \end{pmatrix} = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \begin{pmatrix} D \\ 0 \end{pmatrix} = \begin{pmatrix} D \\ 0 \end{pmatrix} V = \begin{pmatrix} D\,V \\ 0 \end{pmatrix} .$$

Hence $U_{11} = DVD^{-1} = V$ and $U_{21} = 0$. We also have

$$\begin{pmatrix} I_n & 0 \\ 0 & I_{m-n} \end{pmatrix} = I_m = UU^* = \begin{pmatrix} V^* & 0 \\ U_{12}^* & U_{22}^* \end{pmatrix} \begin{pmatrix} V & U_{12} \\ 0 & U_{22} \end{pmatrix}$$

$$= \begin{pmatrix} I_n & V^*U_{12} \\ U_{12}^*\ V & U_{12}^*U_{12} + U_{22}^*\ U_{22} \end{pmatrix} .$$

Hence $U_{12} = 0$ and $U_{22} \in U(m-n)$. □

COROLLARY. *Let $m \geq n$ be positive integers. If $\sigma := (\sigma_1, \sigma_2, \ldots, \sigma_n)$ where the $\sigma_i$ are distinct positive real numbers then* $\dim \operatorname{sing}^{-1}(\sigma) = (2m-1)n$.

*Proof.* We have

$$\dim\ \operatorname{sing}^{-1}(\sigma) = \dim(U(m) \times U(n)) - \dim \operatorname{Fix}(\Sigma)$$

$$= m^2 + n^2 - ((m-n)^2 + n). \qquad □$$

Let $B \in \operatorname{sing}^{-1}(\sigma)$. We want to compute the space tangent to $\operatorname{sing}^{-1}(\sigma)$ at $B$. We denote this tangent space by $\operatorname{Tan}.\operatorname{sing}^{-1}(\sigma).B$. Consider the function:

$$F := U(m) \times U(n) \to \operatorname{sing}^{-1}(\sigma) : (U, V) \to U\Sigma V^*.$$

Note that this function maps $U(m) \times U(n)$ smoothly onto $\operatorname{sing}^{-1}(\sigma)$. Assume $B = F(U,V) = U\Sigma V^*$. Then the derivative of $f$ at $(U,V)$ is a linear map of the following type:

$$dF(U,V) : \operatorname{Tan}.(U(m) \times U(n)).(U,V) \to \operatorname{Tan}.\operatorname{sing}^{-1}(\sigma).B.$$

Since $F$ is onto we guess $dF(U,V)$ will also be onto. Later we shall need to check to see if our guess is correct. In this way we can hope to compute $\operatorname{Tan}.\operatorname{sing}^{-1}(\sigma).B$. We can compute the derivative as follows. First we extend the definition of $F$:

$$F := \mathbb{C}^{m\times m} \times \mathbb{C}^{n\times n} \to \mathbb{C}^{m\times n} : (X, Y) \to X\Sigma Y^*.$$

It is easy to compute the derivative of this extended map. We obtain

$$dF(U,V).(X,Y) = X\Sigma V^* + U\Sigma Y^*$$

by means of the following calculation:

$$d((X,Y) \to X\Sigma Y^*).(U,V).(X,Y)$$

{by the partial derivative formula}

$$= (X \to X\Sigma V^*).U.X + d(Y \to U\Sigma Y^*).V.Y$$

{by the formula for the derivative of a linear map}

$$= X\Sigma V^* + U\Sigma Y^*.$$

Now we can identify $\mathrm{Tan}\,.U(m).U$ with the set

$$\mathrm{Skew}(m) * U := \{KU : K \quad \text{is skew-hermitian}\},$$

and consequently we can identify $\mathrm{Tan}\,.(U(m) \times U(n)).(U,V)$ with $(\mathrm{Skew}(m) * U) \times (\mathrm{Skew}(n) * V)$. Now for $J \in \mathrm{Skew}(m)$ and $K \in \mathrm{Skew}(n)$ we have

$$dF(U,V).(JU,KV) = JU\Sigma V^* + U\Sigma(KV)^* = JB - BK.$$

Thus we expect

$$\mathrm{Tan}\,.\,\mathrm{sing}^{-1}(\sigma).B = \{JB - BK : J \in \mathrm{Skew}(m), K \in \mathrm{Skew}(n)\}.$$

We need to check that $dF(U,V)$ is onto. In the generic case, we compute the dimension of the kernel of the map $dF(I,I)$. Note

$$dF(I,I) = \mathrm{Skew}(m) \times \mathrm{Skew}(n) \to \mathbb{C}^{m\times n} : (J,K) \to J\Sigma - \Sigma K.$$

We have

$$0 = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix} \begin{pmatrix} D \\ 0 \end{pmatrix} - \begin{pmatrix} D \\ 0 \end{pmatrix} K = \begin{pmatrix} J_{11}D - DK \\ J_{21}D \end{pmatrix}$$

iff $J_{11}D = DK$ and $J_{21} = 0$. Note that $J_{11}D = DK$ implies that $J_{11}$ and $K$ are diagonal matrices. For example, when $n = 2$, we have

$$\begin{pmatrix} j_{11}d_1 & j_{12}d_2 \\ -\bar{j}_{12}d_1 & j_{22}d_2 \end{pmatrix} = \begin{pmatrix} d_1k_{11} & d_1k_{12} \\ -d_2\bar{k}_{12} & d_2k_{22} \end{pmatrix}$$

which implies

$$\begin{pmatrix} d_1 & d_2 \\ d_2 & d_1 \end{pmatrix} \begin{pmatrix} j_{12} \\ k_{12} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} .$$

Thus the dimension of the kernel is $n$ (determined by the possible choices for $J_{11}$) plus $(m-n)^2$ (determined by the possible choices of $J_{22}$). This dimension is the same as the dimension of $\mathrm{Fix}(\Sigma)$ which checks. Finally note that $dF(U,V).(JU,KV) = U(U^*JU\Sigma - \Sigma V^*KV)V^*$ and hence $\dim \mathrm{kernel}\, dF(U,V) = \dim \mathrm{kernel}\, dF(I,I)$.

*Remark.* The following example shows that the given check is necessary. Let $f := \mathbb{R} \to \mathbb{R} : x \to x^3$. Then $f$ is onto but $df(0)$ is not onto. □

**Matrices with fixed rank.** Let $k$ be a positive integer satisfying $k \le \min(m,n)$ and let $\operatorname{rank}^{-1}(k)$ denote the space (with the relative topology) of all $m \times n$ complex matrices having rank $k$; in symbols,

$$\operatorname{rank}^{-1}(k) := \{A \in \mathbb{C}^{m\times n} : \operatorname{rank} A = k\}.$$

We want to see that the space $\operatorname{rank}^{-1}(k)$ is homogeneous; i.e., that it is homeomorphic to the quotient of a topological group.

We first want to see that $\operatorname{rank}^{-1}(k)$ is the orbit of a certain group action on $\mathbb{C}^{m\times n}$. Let $Gl(m)$ denote the group of nonsingular $m \times m$ complex matrices. Define the smooth action $\alpha$ by

$$\alpha := Gl(m) \times Gl(n) \times \mathbb{C}^{m\times n} \to \mathbb{C}^{m\times n} : (U,V,A) \to UAV^{-1}.$$

Recall that two matrices $A$ and $B$ are *equivalent* if there exists $U \in Gl(m)$ and $V \in Gl(n)$ such that $B = UAV^{-1}$. (See, for example, Birkhoff and MacLane 1953 section VIII.9 "General equivalence and canonical forms".) Also recall that any matrix $A$ with rank $k$ is equivalent to the following matrix:

$$J_k := \begin{array}{cc} \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} & \begin{array}{c} k \\ m-k \end{array} \\ \begin{array}{cc} k & n-k \end{array} & \end{array}$$

where $I_k$ is the $k \times k$ identity matrix. Thus $\operatorname{rank}^{-1}(k)$ is the orbit of $J_k$ under the group action given above. We also have the following relation is a bijection:

$$(Gl(m) \times Gl(n))/\operatorname{Fix}(J_k) \to \operatorname{rank}^{-1}(k) : (U,V)\operatorname{Fix}(J_k) \to UJ_kV^{-1}.$$

PROPOSITION. *The $m \times n$ complex matrices with a given rank form a homogeneous space.*

*Proof.* In view of the remarks preceding this proposition, we need only check the hypotheses of the theorem concerning homeomorphism of orbits and quotient spaces. We finish the proof of the proposition by means of the following lemma.

LEMMA. *The space of $m \times n$ matrices with fixed rank is locally compact.*

We shall use the following result (which appears, e.g., in Horn and Johnson 1985 section 7.3 "Polar and singular value decomposition") which is an easy consequence of the Hoffman–Wielandt estimate for perturbation of eigenvalues of symmetric matrices (which appears, e.g., in Horn and Johnson 1985 section 6.3 "Perturbation theorems").

Proposition. *Hoffman–Wielandt estimate for perturbation of singular values. Let $A, A+B \in \mathbb{C}^{m\times n}$ where $m \geq n$. Then*

$$\sum_{i=1}^{n}(\sigma_i(A+B)-\sigma_i(A))^2 \leq \|B\|^2$$

*where $\sigma_i(X)$ denotes the ith singular value of $X \in \mathbb{C}^{m\times n}$ in the order*

$$\sigma_1(X) \geq \sigma_2(X) \geq \cdots \geq \sigma_n(X) \geq 0$$

*and $\|X\|$ is the Frobenius norm of $X$.*

*Proof.* (of the lemma) Let $A \in \operatorname{rank}^{-1}(k)$ have singular values $\sigma(A)$ ordered as above. Note that $\sigma_k(A) > 0$ and $\sigma_{k+1}(A) = \cdots = \sigma_n(A) = 0$. Let $\varepsilon := \sigma_k(A)/2$. Note that the closed ball $B_\varepsilon(A) := \{B \in \mathbb{C}^{m\times n} : \|B-A\| \leq \varepsilon\}$ with radius $\varepsilon$ and centered at $A$ is compact. We shall see that the set $B := B_\varepsilon(A) \cap \operatorname{rank}^{-1}(k)$ is closed. Consider any sequence $X_k \in B$ that converges to $X \in B_\varepsilon(A)$. By the Hoffman–Wielandt estimates for singular values, we have $\sigma_i(X_k) \to \sigma_i(X)$. Hence $\sigma_{k+1}(X) = \cdots = \sigma_n(X) = 0$ and $\sigma_1(X) \geq \sigma_2(X) \geq \sigma_i \geq \sigma_k(X) > 0$. We conclude that rank $X = k$. □

We want to compute $\operatorname{Fix}(J_k)$. We shall see that

$$\operatorname{Fix}(J_k) = \left\{ \left( \begin{pmatrix} U & V \\ 0 & W \end{pmatrix}, \begin{pmatrix} U & 0 \\ X & Y \end{pmatrix} \right) : U \in Gl(k), V \in \mathbb{C}^{k\times(m-k)}, \right.$$

$$\left. W \in Gl(m-k), X \in \mathbb{C}^{(n-k)\times k}, Y \in Gl(n-k) \right\}.$$

Let

$$P := \begin{array}{cc} \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} & \begin{array}{c} k \\ m-k \end{array} \\ \begin{array}{cc} k & m-k \end{array} & \end{array} \quad \text{and} \quad Q := \begin{array}{cc} \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} & \begin{array}{c} k \\ n-k \end{array} \\ \begin{array}{cc} k & n-k \end{array} & \end{array} .$$

We have the following equivalent conditions:

$$PJ_kQ^{-1} = J_k;$$

$$\begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix};$$

$$\begin{pmatrix} P_{11} & 0 \\ P_{21} & 0 \end{pmatrix} = \begin{pmatrix} Q_{11} & Q_{12} \\ 0 & 0 \end{pmatrix};$$

$$P_{11} = Q_{11}, P_{21} = 0 \quad \text{and} \quad Q_{12} = 0.$$

Using the characterization of $\mathrm{Fix}(J_k)$ we obtain

$$\dim \operatorname{rank}^{-1}(k) = 2k(m+n-k).$$

In particular, we have the following calculation:

$$\begin{aligned}
&\dim \operatorname{rank}^{-1}(k) = \dim(Gl(m) \times Gl(n)) - \dim(\mathrm{Fix}(J_k)) \\
&= 2m^2 + 2n^2 - (2k^2 + 2k(m-k) + 2(m-k)^2 + 2(n-k)k + 2(n-k)^2) \\
&= 2(m^2 + n^2 - (k^2 + m(m-k) + n(n-k))) \\
&= 2k(m+n-k).
\end{aligned}$$

Let $B \in \operatorname{rank}^{-1}(k)$. We want to compute the space tangent to $\operatorname{rank}^{-1}(k)$ at $B$. We denote this space by $\mathrm{Tan}\,.\operatorname{rank}^{-1}(k).B$. Consider the following function:

$$F := Gl(m) \times Gl(n) \to \mathbb{C}^{m\times n} : (U,V) \to UJ_kV^{-1}.$$

Note that this function maps $Gl(m) \times Gl(n)$ smoothly onto $\operatorname{rank}^{-1}(k)$. Assume $B = F(U,V) = UJ_kV^{-1}$. Then the derivative of $F$ at $(U,V)$ is a linear map of the following type:

$$dF(U,V) : \mathrm{Tan}\,.(Gl(m) \times Gl(n)).(U,V) \to \mathrm{Tan}\,.\operatorname{rank}^{-1}(k).B.$$

Since $F$ is onto, we guess $dF(U,V)$ will also be onto. In this way we can hope to compute $\mathrm{Tan}\,.\operatorname{rank}^{-1}(k).B$. It is easy to compute the derivative of $F$. We obtain

$$dF(U,V).(X,Y) = XJ_kV^{-1} - UJ_kV^{-1}YV^{-1}$$

by means of the following calculation:

$$d((X,Y) \to XJ_kY^{-1}).(U,V).(X,Y)$$

{by the partial derivative formula}

$$= d(X \to XJ_kV^{-1}).U.X + d(Y \to UJ_kY^{-1}).V.Y$$

{by the derivative of a linear map and the formula $d(Y \to Y^{-1}).V.Y = -V^{-1}YV^{-1}$}

$$= XJ_kV^{-1} - UJ_kV^{-1}YV^{-1}.$$

Now we can identify $\mathrm{Tan}\,.Gl(m).U$ with the set $\mathbb{C}^{m\times m}U$ and consequently $\mathrm{Tan}\,.(Gl(m) \times Gl(n)).(U,V)$ with the set $(\mathbb{C}^{m\times m}U) \times (\mathbb{C}^{n\times m}V)$. Now for $X \in \mathbb{C}^{m\times m}$ and $Y \in \mathbb{C}^{n\times n}$ we have

$$dF(U,V).(XU,YV) = XUJ_kV^{-1} - UJ_kV^{-1}YVV^{-1} = XB - BY.$$

Thus we expect

$$\operatorname{Tan}.\operatorname{rank}^{-1}(k).B = \{XB - BY : X \in \mathbb{C}^{m\times m}, Y \in \mathbb{C}^{n\times n}\}.$$

We need to check that $dF(U,V)$ is onto. We compute the dimension of the kernel of the map $dF(I,I)$. The general case will follow from homogeneity. Note

$$dF(I,I) = \mathbb{C}^{m\times m} \times \mathbb{C}^{n\times n} \to \mathbb{C}^{m\times n} : (X,Y) \to XJ_k - J_kY.$$

We have

$$0 = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix} = \begin{pmatrix} X_{11} - Y_{11} & Y_{12} \\ X_{21} & 0 \end{pmatrix},$$

$$\text{iff } X_{11} = Y_{11}, Y_{12} = 0 \quad \text{and} \quad X_{21} = 0.$$

Thus the dimension of the kernel is $2k(m-k)$ (which is determined by the possible choices of $X_{12}$) + $2(m-k)^2$ (which is determined by the possible choices of $X_{22}$) + $2k^2$ (which is determined by the possible choices of $Y_{11}$) + $2(n-k)k$ (which is determined by the possible choices of $Y_{21}$) + $(n-k)^2$ (which is determined by the possible choices of $Y_{22}$). This dimension is the same as the dimension of $\operatorname{Fix}(J_k)$ which checks.

*Remark.* Guillemin and Pollack (1974) suggest a quite different proof that the matrices with fixed rank form a regular submanifold of the space of $m \times n$ matrices. In particular, see Chapter 1 "Manifolds and smooth maps" Section 4 "Submersions" Exercise 13 in their book.

**Matrices with fixed displacement rank.** Let $Z_m$ denote the $m \times m$ *(lower) shift* matrix defined by

$$Z_m(i,j) := \delta(i-1,j).$$

For example,

$$Z_3 := \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Let $\lrcorner$ be the linear map defined by

$$\lrcorner := \mathbb{C}^{m\times n} \to \mathbb{C}^{m\times n} : X \to X - Z_m X Z_n^T.$$

(Note that the map $X \to Z_m X Z_n^T$ is a shift of the matrix $X$ to the lower right or southeast; consequently, we use a symbol, namely $\lrcorner$, that points in that direction.) Note that the map $\lrcorner$ is invertible. In fact the map $X \to Z_m X Z_n^T$ is nilpotent (since $Z_n^n = 0$) and hence $\lrcorner^{-1}$ is given by the following formula when $m \geq n$:

$$\lrcorner^{-1}.Y = Y + Z_m Y Z_n^T + Z_m^2 X (Z_n^2)^T + \cdots + Z_m^{n-1} X (Z_n^{n-1})^T .$$

We shall assume that $m \geq n$ thruout this section for ease of exposition.

We say that $X \in \mathbb{C}^{m\times n}$ has *displacement rank* $k$ if $\lrcorner.X$ has rank $k$. (The notion of "displacement rank" appears in Kailath, Kung and Morf (1979 a,b). We have slightly modified the definition that appears there.) Let $E_{mn} \in \mathbb{C}^{m\times n}$ be the $m \times n$ matrix defined by

$$E_{mn}(i,j) := \delta(i,j).$$

For example,

$$E_{32} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} .$$

The following theorem is a minor generalization of a theorem due to Kailath, Kung and Morf (1979 a,b).

THEOREM. *Alternate characterization of displacement rank.*

*Let $R \in \mathbb{C}^{m\times n}$. Then $R$ has displacement rank $k$ iff $k$ is the smallest integer such that*

$$R = \sum_{i=1}^{k} L_i E_{mn} U_i$$

*for some lower-triangular Toeplitz matrices $L_i \in \mathbb{C}^{m\times m}$ and some upper-triangular Toeplitz matrices $U_i \in \mathbb{C}^{n\times n}$.*

*Proof.* Let $L$ be any lower-triangular Toeplitz matrix; in other words, $L$ has the form

$$\begin{pmatrix} a_0 & 0 & 0 & \ldots & 0 \\ a_1 & a_0 & 0 & \ldots & 0 \\ a_2 & a_1 & a_0 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m-1} & a_{m-2} & a_{m-3} & \ldots & a_0 \end{pmatrix} .$$

Note that $L = a_0 I_m + a_1 Z_m + \cdots + a_{m-1} Z_m^{m-1}$; that is, $L = p(Z_m)$ where $p$ is the polynomial defined by $p(x) := a_0 + a_1 x + \cdots + a_{m-1} x^{m-1}$. Similarly, if $U$ is any $n \times n$ upper-triangular

Toeplitz matrix, then $U = q(Z_n^T)$ for some polynomial $q(x) := b_0 + b_1 x + \cdots + b_{n-1}x^{n-1}$. Hence, we have

$$\begin{aligned} \lrcorner(LE_{mn}U) &= LE_{mn}U - Z_m LE_{mn}UZ_n^T \\ &= L(E_{mn} - Z_m E_{mn} Z_n^T)U. \end{aligned}$$

Note that $E_{mn} - Z_m E_{mn} Z_n^T$ is the $m \times n$ matrix with 1 in the (1,1) position and 0's elsewhere; in particular, this matrix has rank one. We illustrate with $m = 3,\ n = 2$:

$$Z_3 E_{32} Z_2^T = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

and hence

$$E_{32} - Z_3 E_{32} Z_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} .$$

More generally if $R := \sum_{i=1}^{k} L_i E_{mn} U_i$ where the $L_i$ are $m \times m$ lower-triangular Toeplitz matrices and the $U_i$ are $n \times n$ upper-triangular Toeplitz matrices, then the matrix

$$\lrcorner . R = \sum_{i=1}^{k} L_i (E_{mn} - Z_m E_{mn} Z_n^T) U_i$$

has rank at most $k$.

Now consider the matrix $ab^T$ where $a$ and $b$ are column vectors in $\mathbb{C}^m$ and $\mathbb{C}^n$ respectively; that is, $a^T := (a_0, a_1, \ldots, a_{m-1})$ and $b^T := (b_0, b_1, \ldots, b_{m-1})$. Let $L(a) := p(Z_m)$ where $p(x) := a_0 + a_1 x + \cdots + a_{m-1}x^{m-1}$ and $U(b) := q(Z_n^T)$ where $q(x) := b_0 + b_1 x + \cdots + b_{n-1}x^{n-1}$. Note that

$$\begin{aligned} \lrcorner .(L(a)E_{mn}U(b)) &= L(a)E_{mn}U(b) - Z_m L(a) E_{mn} U(b) Z_n^T \\ &= L(a)(E_{mn} - Z_m E_{mn} Z_n^T) U(b) = ab^T. \end{aligned}$$

We illustrate with $m = 3,\ n = 2$:

$$
\begin{aligned}
&L(a)(E_{mn} - Z_m E_{mn} Z_n^T)U(b) \\
&= \begin{pmatrix} a_0 & 0 & 0 \\ a_1 & a_1 & 0 \\ a_2 & a_1 & a_0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b_0 & b_1 \\ 0 & b_0 \end{pmatrix} \\
&= \begin{pmatrix} a_0 & 0 & 0 \\ a_1 & a_0 & 0 \\ a_2 & a_1 & a_0 \end{pmatrix} \begin{pmatrix} b_0 & b_1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a_0 b_0 & a_0 b_1 \\ a_1 b_0 & a_1 b_1 \\ a_2 b_0 & a_2 b_1 \end{pmatrix} \\
&= \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} (\, b_0, b_1 \,) \ .
\end{aligned}
$$

More generally, if $S := \sum_{i=1}^{k} a_i b_1^T$ where the $a_i$ and $b_i$ are column vectors then

$$\lrcorner . \left( \sum_{i=1}^{k} L(a_i) E_{mn} U(b_i) \right) = S.$$

Hence, for any $m \times n$ matrix $S$, if $S$ has rank $k$ then $\lrcorner^{-1}.S$ has displacement rank at most $k$. □

Proposition. *The $m \times n$ complex matrices with displacement rank $k$ form a homogeneous manifold with dimension $2k(m+n-k)$.*

*Proof.* From the results on matrices with fixed rank, we know that the space $\text{rank}^{-1}(k) \subseteq \mathbb{C}^{m \times n}$ is a homogeneous manifold with dimension $2k(m+n-k)$. This proposition then follows at once from the fact that the map $\lrcorner$ is an invertible linear map. □

We shall use $\text{dis-rank}^{-1}(k)$ to denote the space of $m \times n$ matrices with displacement rank $k$; in symbols,

$$\text{dis-rank}^{-1}(k) := \{X \in \mathbb{C}^{m \times n} : \ \text{displacement-rank}\ (X) = k\}.$$

The (invertible linear) map $\lrcorner$ provides us with a smooth $1-1$ correspondence between the surface $\text{dis-rank}^{-1}(k)$ and the surface $\text{rank}^{-1}(k)$. We can use this observation to compute the spaces tangent to $\text{dis-rank}^{-1}(k)$. Let $B \in \text{dis-rank}^{-1}(k)$. Then $\lrcorner.B \in \text{rank}^{-1}(k)$ and (from the description of the spaces tangent to $\mathrm{rank}^{-1}(k)$ in the section on matrices with fixed rank) we have

$$\mathrm{Tan}\,.\,\mathrm{rank}^{-1}(k).(\lrcorner B) = \{X(\lrcorner B) - (\lrcorner B)Y : X \in \mathbb{C}^{m\times m}, Y \in \mathbb{C}^{n\times n}\}.$$

Clearly then

$$\mathrm{Tan}\,.\,\text{dis-rank}^{-1}(k).B = \lrcorner^{-1}(\mathrm{Tan}\,.\,\mathrm{rank}^{-1}(k).(\lrcorner B))$$

$$= \{\lrcorner^{-1}(X(\lrcorner B) - (\lrcorner B)Y) : X \in \mathbb{C}^{m\times m}, Y \in \mathbb{C}^{n\times n}\}.$$

Let

$$\lrcorner B = p_1 q_1^* + p_2 q_2^* + \cdots + p_k q_k^*$$

where $\{p_1, p_2, \ldots, p_k\} \subseteq \mathbb{C}^m$ is an independent set of $m$-vectors and $\{q_1, q_2, \ldots, q_k\} \subseteq \mathbb{C}^n$ is an independent set of $n$-vectors. We consider a typical element of $\mathrm{Tan}\,.\,\mathrm{rank}^{-1}(k).(\lrcorner B)$:

$$X(\lrcorner B) - (\lrcorner B)Y^* = X\left(\sum_{i=1}^{k} p_i q_i^*\right) - \left(\sum_{i=1}^{k} p_i q_i^*\right) Y^*$$

$$= \sum_{i=1}^{k}((Xp_i)q_i^* - p_i(Yq_i)^*) = \sum_{i=1}^{k}(x_i q_i^* - p_i y_i^*)$$

where $x_i := Xp_i$ and $y_i := Yq_i$. Then a typical element of $\mathrm{Tan}\,.\,\text{dis-rank}^{-1}(k).B$ has the following form:

$$\lrcorner^{-1}(X(\lrcorner B) - (\lrcorner B)Y^*) = \lrcorner^{-1}\left(\sum_{i=1}^{k}(x_i q_i^* - p_i y_i^*)\right)$$

$$= \sum_{i=1}^{k}(L(x_i)U(q_i^*) - L(p_i)U(y_i^*)).$$

We conclude that

$$\mathrm{Tan}\,.\,\text{dis-rank}^{-1}(k).B$$

$$= \left\{\sum_{i=1}^{k}(L(x_i)U(q_i^*) - L(p_i)U(y_i^*)) : x_1, x_2, \ldots, x_k \in \mathbb{C}^m, y_1, y_2, \ldots, y_k \in \mathbb{C}^n\right\}.$$

**Matrices with fixed singular values and fixed displacement rank.** Let $\sigma := (\sigma_1, \sigma_2, \ldots, \sigma_n)$ be an ordered $n$-tuple of nonnegative real numbers. We use $\mathrm{sing}^{-1}(\sigma)$ to denote the space of all $m \times n$, where $m \geq n$, complex matrices with these numbers as singular values; in symbols,

$$\text{sing}^{-1}(\sigma) := \{A \in \mathbb{C}^{m\times n} : \text{ singular-spectrum } A = \sigma\}.$$

In the section on matrices with fixed singular values, we saw that this isosingular surface $\text{sing}^{-1}(\sigma)$ is a smooth manifold with dimension $(2m-1)n$ provided the $\sigma_i$ are positive and distinct. Let $k$ be a positive integer satisfying $k \leq n$. We use $\text{dis-rank}^{-1}(k)$ to denote the space of all $m \times n$ complex matrices having displacement rank $k$; in symbols,

$$\text{dis-rank}^{-1}(k) := \{A \in \mathbb{C}^{m\times n} : \text{displacement-rank}\, A = k\}.$$

In the section on matrices with fixed rank, we saw that $\text{dis-rank}^{-1}(k)$ is a smooth manifold with dimension $2k(m+n-k)$. In this section we shall consider the space $\text{sing}^{-1}(\sigma) \cap \text{dis-rank}^{-1}(k)$ which consists of the matrices with fixed singular spectrum $\sigma$ and fixed displacement rank $k$. We shall see that this space is generally a smooth manifold with dimension $2k(m+n-k)-n$.

Let us begin by reviewing some standard results about submersions. (I mainly follow Guillemin and Pollack (1974.) Let $f : X \to Y$ be a smooth map from manifold $X$ to manifold $Y$; let $x \in X$ and let $y := f(x)$. Then the derivative of $f$ at $x$ is a linear map of the following type:

$$df(x) : \text{Tan}\,.X.x \to \text{Tan}\,.Y.y$$

where $\text{Tan}\,.X.x$ is the space tangent to $X$ at $x$. If $df(x)$ is surjective then $f$ is a *submersion* at $x$. The map $f$ is a *submersion* if $df(x)$ is surjective for all $x \in X$. A point $y \in Y$ is a *regular value* of $f$ if, for all $x \in X$, $f(x) = y$ implies $df(x)$ is surjective. The following result is very useful for showing that spaces are manifolds.

PREIMAGE THEOREM. *Let $f : X \to Y$ be a smooth map between manifolds and let $y \in Y$ be a regular value of $f$. Then the preimage $f^{-1}(y) := \{x \in X : f(x) = y\}$ is a submanifold of $X$ with*

$$\dim f^{-1}(y) = \dim X - \dim Y.$$

*In particular, $f^{-1}(y)$ is a manifold.*

*Proof.* See, e.g. Guillemin and Pollack (1974). □

The preimage theorem provides another way to see that $\text{sing}^{-1}(\sigma)$ is a smooth manifold with dimension $2mn-n$ provided the $\sigma_i$ are positive and distinct. We consider the following function:

$$f := \mathbb{C}^{m\times n} \to \mathbb{R}^n : X \to (\text{trace}(X^*X), \text{trace}(X^*X)^2, \ldots, \text{trace}(X^*X)^N).$$

It follows easily from the following result that $\operatorname{sing}^{-1}(\sigma) = f^{-1}(\mu_1, \mu_2, \ldots, \mu_n)$ where

$$\mu_k := \operatorname{trace}(X^*X)^k = \sigma_1^{2k} + \sigma_2^{2k} + \cdots + \sigma_n^{2k}.$$

PROPOSITION. *Let $A, B \in \mathbb{C}^{n\times n}$. Then $A$ and $B$ have the same eigenvalues iff, for all $k = 1, 2, \ldots, n, \operatorname{trace} A^k = \operatorname{trace} B^k$ .*

*Proof.* See Horn and Johnson (1985) section 1.2 "The characteristic polynomial" Problem 12. $\square$

We need to compute the derivative of $f$ in order to apply the preimage theorem. We have

$$d(X \to X^k).A.X = A^{k-1}X + A^{k-2}XA + A^{k-3}XA^2 + \cdots + XA^{k-1} \quad \text{and}$$

$$d(X \to X^*X).A.X = A^*X + X^*A.$$

Using the chain rule, we get

$$\begin{aligned} d(X \to \operatorname{trace}(X^*X)^k).A.X &= \operatorname{trace}(d(X \to (X^*X)^k).A.X) \\ &= k\operatorname{trace}((A^*A)^{k-1}(A^*X + X^*A)) \end{aligned}$$

and hence

$$df.A.X = (\operatorname{trace}(A^*X+X^*A), 2\operatorname{trace}((A^*A)(A^*X+X^*A)), \ldots, n\operatorname{trace}((A^*A)^{n-1}(A^*X+X^*A))).$$

We want to see that $df(A)$ is surjective if $A$ has distinct nonzero singular values. In other words, we want to see that for all $b \in \mathbb{R}^n$, there exists $X \in \mathbb{C}^{m\times n}$ such that

$$(*) \qquad \begin{aligned} &\operatorname{trace}(A^*X + X^*A) = b_1 \ , \\ &\operatorname{trace}(A^*A)(A^*X + X^*A)) = b_2 \ , \\ &\operatorname{trace}(A^*A)^{n-1}(A^*X + X^*A)) = b_n \ . \end{aligned}$$

Let $A = U\Sigma V$, where $U \in U(m), V \in U(n), \Sigma := \operatorname{diag}(\sigma)$, be a singular value decomposition of $A$. Then we have $A^*A = V^*\Sigma^*\Sigma V$ and hence

$$\begin{aligned} \operatorname{trace}(A^*A)^k(A^*X + X^*A) &= \operatorname{trace} V(A^*A)^kV^*V(A^*X + X^*A)V^* \\ &= \operatorname{trace}(\Sigma^*\Sigma)(\Sigma^*VXV^* + VX^*V^*\Sigma). \end{aligned}$$

Thus we may assume without loss of generality that $A = \mathrm{diag}(\sigma)$. We take $X := \mathrm{diag}(x_1, \ldots, x_n) \in \mathbb{R}^{m\times n}$. Then the equations $(*)$ reduce the following ones

$$\begin{aligned} 2(\sigma_1 x_1 + \sigma_2 x_2 + \cdots + \sigma_n x_n) &= b_1 \\ 2(\sigma_1^3 x_1 + \sigma_2^3 x_2 + \cdots + \sigma_n^3 x_n) &= b_2 \\ &\vdots \\ 2(\sigma_1^{2n-1} x_1 + \sigma_2^{2n-1} x_2 + \cdots + \sigma_n^{2n-1} x_n) &= b_n \end{aligned}$$

or

$$2\begin{pmatrix} 1 & 1 & \ldots & 1 \\ \sigma_1^2 & \sigma_2^2 & \ldots & \sigma_n^2 \\ \vdots & \vdots & & \vdots \\ \sigma_1^{2(n-1)} & \sigma_2^{2(n-1)} & \ldots & \sigma_n^{2(n-1)} \end{pmatrix} \begin{pmatrix} \sigma_1 x_1 \\ \sigma_2 x_2 \\ \vdots \\ \sigma_n x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

The coefficient matrix is a vandermonde matrix. If the $\sigma_i$ are distinct then its determinant is nonzero. Thus $df.A$ is surjective if $A$ has distinct, nonzero singular values $\sigma$. We conclude (using the preimage theorem) that $\mathrm{sing}^{-1}(\sigma)$ is a smooth manifold with dimension $2mn - n = (2m-1)n$.

We now turn to the following result.

PROPOSITION. *Let $k$ be a positive integer satisfying $1 \le k \le n$ where $n \le m$. Then, for almost all $\sigma \in \mathbb{R}^n$, $\mathrm{sing}(\sigma) \cap \text{dis-rank}^{-1}(k)$ is a manifold with dimension $2k(m+n-k) - n$.*

As usual, "almost all" means "except for a set of measure 0". Recall the following result.

THEOREM. *(Sard) Let $f : X \to Y$ be a smooth map between manifolds. Then almost every point in $Y$ is a regular value of $f$.*

*Proof.* (of Sard's theorem) See, e.g., Guillemin and Pollack (1974), Spivak (1965), Hirsch (1976) or Milnor (1965). □

*Proof.* (of the proposition) We apply Sard's theorem and the preimage theorem. In particular, we take $X := \text{dis-rank}^{-1}(k)$, $Y := \mathbb{R}^n$ and define $f$ by

$$f(A) := (\mu_1(A), \mu_2(A), \ldots, \mu_n(A))$$

where $\mu_k(A) := \mathrm{trace}(A^*A)^k$. □

**Hermitian matrices with fixed eigenvalues.** Let $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_n)$ be an ordered $n$-tuple of real numbers satisfying $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ and let $\operatorname{spec}^{-1}(\lambda)$ denote the space (with the relative topology) of all hermitian $n \times n$ complex matrices with these numbers as eigenvalues; in symbols,

$$\operatorname{spec}^{-1}(\lambda) := \{A \in \operatorname{Herm}(n) : \quad \text{spectrum } A = \lambda\}$$

where $\operatorname{Herm}(n)$ is the space of all $n \times n$ hermitian matrices; in symbols,

$$\operatorname{Herm}(n) := \{A \in \mathbb{C}^{n \times n} : \ A^* = A\}.$$

We call $\operatorname{spec}^{-1}(\lambda)$ the "isospectral" space determined by $\lambda$. We want to see that this space is homogeneous; i.e., that it is homeomorphic to the quotient space of a Lie group.

We first want to see that $\operatorname{spec}^{-1}(\lambda)$ is the orbit of a certain group action on $\operatorname{Herm}(n)$. Let $U(n)$ denote the group of unitary matrices; in symbols,

$$U(n) := \{Q \in \mathbb{C}^{n \times n} : QQ^* = I\}.$$

Define the smooth action $\alpha$ by

$$\alpha := U(n) \times \operatorname{Herm}(n) \to \operatorname{Herm}(n) : (Q, A) \to QAQ^*.$$

Recall that two matrices $A, B \in \mathbb{C}^{n \times n}$ are *unitarily similar* if there exists $U \in Gl(n)$ such that $B = UAU^{-1}$. Also recall that any hermitian matrix $A$ with spectrum $\lambda$ is unitarily similar to the diagonal matrix with diagonal matrices $\lambda_1, \lambda_2, \ldots, \lambda_n$; in other words, there exists $Q \in U(n)$ such that $QAQ^* = \Lambda$ where $\Lambda := \operatorname{diag}(\lambda)$. (See, for example, Birkhoff and MacLane 1953 section IX.12 "Unitary and hermitian matrices" or Strang 1980 Chapter 5" Eigenvalues and eigenvectors" or Horn and Johnson 1985 section 2.5 "Normal matrices.") Thus $\operatorname{spec}^{-1}(\lambda)$ is the orbit of $\Lambda$ under the action given above. We also have that the following relation is bijection:

$$U(n)/\operatorname{Fix}(\Lambda) \to \operatorname{spec}^{-1}(\lambda) : Q\operatorname{Fix}(\Lambda) \to Q\Lambda Q^*.$$

PROPOSITION. *The hermitian matrices with a given spectrum form a homogeneous space.*

*Proof.* In view of the remarks preceding this proposition, we need only check the hypotheses of the theorem concerning the homeomorphism of orbits and quotient spaces. Note that $U(n)$ is a compact topological group (since it is closed and bounded). We finish the proof by noting that $\operatorname{spec}^{-1}(\lambda)$ is compact since it is the image of $U(n)$ under the smooth map $U(n) \to \operatorname{orbit}(\Lambda) : Q \to Q\Lambda Q^*$. $\square$

We want to determine $\operatorname{Fix}(\Lambda)$ in a generic case.

PROPOSITION. *Let* $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{R}^n$ *and let* $\Lambda := \operatorname{diag}(\lambda)$ *be the* $n \times n$ *diagonal matrix with diagonal entries* $\lambda_1, \lambda_2, \ldots, \lambda_n$. *If the* $\lambda_i$ *are distinct real numbers then*

$$\operatorname{Fix}(\Lambda) = \{\operatorname{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n) : \alpha_i \in \mathbb{C}, |\alpha_i| = 1\}.$$

*Proof.* Let $Q \in U(n)$ satisfy $Q\Lambda Q^* = \Lambda$; i.e., $Q\Lambda = \Lambda Q$ or $Q$ commutes with $\Lambda$. It follows that $Q$ is a diagonal matrix: $Q = \operatorname{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)$ for some $\alpha_i \in \mathbb{C}$. Since $QQ^* = I$, we get $|\alpha_i| = 1$. $\square$

COROLLARY. *Let* $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{R}^n$. *If the* $\lambda_i$ *are distinct then*

$$\dim \operatorname{spec}^{-1}(\lambda) = n^2 - n.$$

*Proof.* We have

$$\begin{aligned} \dim \operatorname{spec}^{-1}(\lambda) &= \dim U(n) - \dim \operatorname{Fix}(\Lambda) \\ &= n^2 - n. \qquad \square \end{aligned}$$

Let $B \in \operatorname{spec}^{-1}(\lambda)$. We want to compute the space tangent to $\operatorname{spec}^{-1}(\lambda)$ at $B$. We denote this tangent space by $\operatorname{Tan}.\operatorname{spec}^{-1}(\lambda).B$. Consider the following function:

$$F := U(n) \to \operatorname{Herm}(n) : Q \to Q\Lambda Q^*.$$

Note that this function maps $U(n)$ smoothly onto $\operatorname{spec}^{-1}(\lambda)$. Assume $B = F(Q) = Q\Lambda Q^*$. Then the derivative of $F$ at $Q$ is a linear map of the following type:

$$dF(Q) : \operatorname{Tan}.U(n).Q \to \operatorname{Tan}.\operatorname{spec}^{-1}(\lambda).B.$$

Since $F$ is onto, we guess $dF(Q)$ will also be onto. In this way we can hope to compute $\operatorname{Tan}.\operatorname{spec}^{-1}(\lambda).B$. We can easily compute the derivative map. We extend the definition of $F$:

$$F := \mathbb{C}^{n \times n} \to \operatorname{Herm}(n) : X \to X\Lambda X^*.$$

It is easy to compute the derivative of this map. We obtain (e.g., by modifying the derivative calculation in the section on matrices with fixed singular values):

$$dF(Q).X = X\Lambda Q^* + Q\Lambda X^*.$$

Now we can identify $\operatorname{Tan}.U(n).Q$ with the set $\operatorname{Skew}(n) * Q := \{KQ : K \text{ is skew-hermitian}\}$. For $K \in \operatorname{Skew}(n)$, we have

$$dF(Q).(KQ) = KQ\Lambda Q^* + Q\Lambda(KQ)^* = KB - BK.$$

Thus

$$\mathrm{Tan}\,.\,\mathrm{spec}^{-1}(\lambda).B = \{[K,B] : K \in \mathrm{Skew}(n)\}.$$

Recall that the Lie bracket is defined by $[X,Y] := XY - YX$.

As a check we compute the dimension of the kernel of the map $dF(I)$ which is given by

$$dF(I) : \mathrm{Skew}(n) \to \mathrm{Herm}(n) : K \to [K, \Lambda].$$

We do the computation in the generic case when $\Lambda$ has distinct diagonal entries. We have

$$0 = K\Lambda - \Lambda K$$

iff $K$ is a diagonal matrix. In other words $K$ is in the kernel iff $K$ is a diagonal matrix with purely imaginary diagonal entries. Hence the dimension of the kernel is $n$. This equals the dimension of $\mathrm{Fix}(\Lambda)$ and our check is successful. The dimension of the kernel of $dF(Q)$ equals the dimension of the kernel of $dF(I)$ by homogeneity.

**Hermitian matrices with fixed inertia.** Let $m$ be a positive integer and let $p$ and $n$ be nonnegative integers satisfying $p+n \leq m$. A hermitian matrix $A$ has *inertia* $(p,n,z)$ where $z := m - (p+n)$ if $A$ has $p$ positive eigenvalues, $n$ negative eigenvalues and $z$ zero eigenvalues. Let $\mathrm{inert}^{-1}(p,n)$ denote the space (with the relative topology) of all $m \times m$ hermitian matrices with inertia $(p,n,z)$; in symbols,

$$\mathrm{inert}^{-1}(p,n) := \{A \in \mathrm{Herm}(m) : \mathrm{inertia}(A) = (p,n,z)\}.$$

We want to see that the space $\mathrm{inert}^{-1}(p,n)$ is homogeneous; i.e. that it is homeomorphic to the quotient of a Lie group.

We first want to see that $\mathrm{inert}^{-1}(p,n)$ is the orbit of a certain group action on $\mathrm{Herm}(m)$. Let $Gl(m)$ denote the group of non singular $m \times m$ complex matrices. Define the smooth action $\alpha$ by

$$\alpha := Gl(m) \times \mathrm{Herm}(m) \to \mathrm{Herm}(m) : (C,A) \to CAC^*.$$

Recall that two hermitian matrices $A$ and $B$ are *congruent* if there exists $C \in Gl(m)$ such that $B = CAC^*$. Also recall that congruent matrices have the same inertia; in fact, any hermitian matrix $A$ with inertia $(p,n,z)$ is congruent to the diagonal matrix

$$\Lambda_{pn} := \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1, 0, \ldots 0)$$

which has diagonal entries consisting of $p$ positive ones, $n$ negative ones and $z := m - (p+n)$ zeros. (This result is called Sylvester's law of inertia. See, for example, Horn and Johnson 1985 section 4.5 " Congruence and simultaneous diagonalization of hermitian and symmetric matrices" or Strang 1980 section 6.3 "Semidefinite and indefinite matrices;

$Ax = \lambda x$".) Thus $\mathrm{inert}^{-1}(p,n)$ is the orbit of $\Lambda_{pn}$ under the action of $Gl(n)$ given above. We also have that the following relation is a bijection:

$$Gl(m)/\operatorname{Fix}(\Lambda_{pn}) \to \mathrm{inert}^{-1}(p,n) : C \operatorname{Fix}(\Lambda_{pn}) \to C\Lambda_{pn}C^*.$$

PROPOSITION. *The hermitian matrices with given inertia form a homogeneous space.*

*Proof.* In view of the remarks preceding this proposition, we need only check the hypotheses of the theorem concerning homeomorphism of orbits and quotients. We finish the proof by means of the following lemma:

LEMMA. *The space of $m \times m$ hermitian matrices with fixed inertia is locally compact.*

We shall use the following theorem (which appears, e.g., in Horn and Johnson 1985 section 6.3 "Perturbation theorems"):

PROPOSITION. *Hoffman–Wielandt estimate for perturbation of eigenvalues.*
*Let $A$ and $A+X$ be hermitian $m \times m$ matrices. Then*

$$\sum_{i=1}^{m} (\lambda_i(A+X) - \lambda_i(A))^2 \leq \|X\|^2$$

*where $\lambda_i(Y)$ denotes the ith eigenvalue of $Y \in \mathrm{Herm}(m)$ in the order*

$$\lambda_1(Y) \geq \lambda_2(Y) \geq \cdots \geq \lambda_n(Y),$$

*and $\|Y\|$ denotes the Frobenius norm of $Y$.*

*Proof (of the lemma).* Let $A \in \mathrm{inert}^{-1}(p,q)$ have eigenvalues $\lambda_i(A)$ as follows:

$$\lambda_1(A) \geq \cdots \geq \lambda_p(A) > \lambda_{p+1}(A) = \cdots = \lambda_{p+z}(A) = 0 > \lambda_{p+z+1}(A) \geq \cdots \geq \lambda_m(A).$$

Let $\varepsilon := \min(\lambda_p(A), -\lambda_{p+z+1}(A))/2$. Note that the closed ball

$$B_\varepsilon(A) := \{Y \in \mathrm{Herm}(m) : \|Y - A\| \leq \varepsilon\}$$

with radius $\varepsilon$ and centered at $A$ is compact. We shall see that the set $B := B_\varepsilon(A) \cap \mathrm{inert}^{-1}(p,n)$ is compact. In particular, we need to see that $B$ is closed. Consider any sequence $Y_k \in B$ that converges to $Y \in B_\varepsilon(A)$. By the Hoffman–Wielandt estimate for perturbation of eigenvalues, we have $\lambda_i(Y_k) \to \lambda_i(Y)$. Hence

$$\lambda_1(Y) \geq \cdots \geq \lambda_p(Y) \geq \varepsilon > 0,$$

$$\lambda_{p+1}(Y) = \cdots = \lambda_{p+z}(Y) = 0, \quad \text{and}$$

$$0 > \varepsilon \geq \lambda_{p+z+1}(Y) \geq \cdots \geq \lambda_m(Y).$$

We conclude that $Y$ has inertia $(p, n, z)$. □

We want to compute Fix $\Lambda_{pq}$. Let $I_{pn}$ be the $(p+n) \times (p+n)$ diagonal matrix which has diagonal entries consisting of $p$ positive ones and $n$ negative ones; in other words, $I_{pn}$ is the direct sum of $I_p$ and $-I_n$ : $I_{pn} := I_p \oplus (-I_n)$ or

$$I_{pn} := \begin{pmatrix} I_p & 0 \\ 0 & -I_n \end{pmatrix} .$$

Let

$$U_{pn} := \{C \in Gl(p+n) : C I_{pn} C^* = I_{pn}\}.$$

Note that $U_{pn}$ is a subgroup of $Gl(u+p)$. We compute its Lie algebra. (Compare, for example, Varadarajan 1984 pp. 48–49.)

Proposition. *The tangent space to $U_{pn}$ at the identity is given by*

$$\operatorname{Tan} .U_{pn}.I = \{K \in \mathbb{C}^{(p+n)\times(p+n)} : K = -I_{pn} K^* I_{pn}\}.$$

*Proof.* $\subseteq$: Consider any curve $C : \mathbb{R} \to U_{pn}$ satisfying $C(0) = I$. Then for all $t \in \mathbb{R}$ we have

$$I_{pn} = C(t) I_{pn}\ C(t)^*.$$

Hence

$$\begin{aligned} 0 &= C'(t) I_{pn} C(t)^* + C(t) I_{pn} C'(t)^*|_{t=0} \\ &= K I_{pn} + I_{pn} K^* \end{aligned}$$

where $K := C'(0)$.
$\supseteq$: For the other inclusion use the exponential map. □

It follows that

$$\dim \operatorname{Tan} .U_{pn}.I = (p+n)^2$$

since

$$\begin{aligned} \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} &= - \begin{pmatrix} I_p & 0 \\ 0 & -I_n \end{pmatrix} \begin{pmatrix} K_{11}^* & K_{21}^* \\ K_{12}^* & K_{22}^* \end{pmatrix} \begin{pmatrix} I_p & 0 \\ 0 & -I_n \end{pmatrix} \\ &= \begin{pmatrix} -K_{11}^* & K_{21}^* \\ K_{12}^* & -K_{22}^* \end{pmatrix} \end{aligned}$$

iff $K_{11} = -K_{11}^*, K_{22} = -K_{22}^*$ and $K_{21} = K_{12}^*$.

We can now compute Fix $\Lambda_{pn}$.

PROPOSITION. *The isotropy subgroup of $Gl(m)$ determined by $\Lambda_{pn}$ is given by*

$$\operatorname{Fix}\Lambda_{pn} = \left\{ \begin{pmatrix} U & W \\ 0 & V \end{pmatrix} \ : \ U \in U_{pn}, V \in Gl(z), W \in \mathbb{C}^{(p+n)\times z} \right\}$$

*where $z := m - (p+n)$.*

*Proof.* We have the following equivalent conditions:

$$\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \begin{pmatrix} I_{pn} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} C_{11}^* & C_{21}^* \\ C_{12}^* & C_{22}^* \end{pmatrix} = \begin{pmatrix} I_{pn} & 0 \\ 0 & 0 \end{pmatrix} ;$$

$$\begin{pmatrix} C_{11}I_{pn}C_{11}^* & C_{11}I_{pn}C_{21} \\ C_{21}I_{pn}C_{11}^* & 0 \end{pmatrix} = \begin{pmatrix} I_{pn} & 0 \\ 0 & 0 \end{pmatrix} ;$$

$$C_{11}I_{pn}C_{11}^* = I_{pn}, C_{11}I_{pn}C_{21}^* = 0 \quad \text{and} \quad C_{21}I_{pn}C_{11}^* = 0 \ ;$$

$$C_{11} \in U_{pn} \quad \text{and} \quad C_{21} = 0 \ . \qquad \square$$

It follows that

$$\dim \operatorname{inert}^{-1}(p,n) = (p+n)(p+n+2z)$$

since

$$\begin{aligned} &\dim \operatorname{inert}^{-1}(p,n) = \dim Gl(m) - \dim \operatorname{Fix}(\Lambda_{pn}) \\ &= 2(p+n+z)^2 - ((p+n)^2 + 2(p+n)z + 2z^2) \\ &= 2((p+n)^2 + 2(p+n)z + z^2) - ((p+n)^2 + 2(p+n)z + 2z^2) \\ &= (p+n)^2 + 2(p+n)z = (p+n)\ (p+n+2z). \end{aligned}$$

Let $B \in \operatorname{inert}^{-1}(p,n)$. We want to compute the space tangent to $\operatorname{inert}^{-1}(p,n)$ at $B$. We denote this tangent space by $\operatorname{Tan}.\operatorname{inert}^{-1}(p,n).B$. Consider the following function:

$$F := Gl(m) \to \operatorname{Herm}(m) : C \to C\Lambda_{pn}C^*.$$

Note that this function maps $Gl(m)$ smoothly onto $\operatorname{inert}^{-1}(p,n)$. Assume $B = C\Lambda_{pn}C^*$. Then the derivative of $F$ at $C$ is a linear map of the following type:

$$dF(C) : \operatorname{Tan}.Gl(m).C \to \operatorname{Tan}.\operatorname{inert}^{-1}(p,n).B.$$

Since $F$ is onto, we guess $dF(C)$ will also be onto. In this way we can hope to compute $\operatorname{Tan}.\operatorname{inert}^{-1}(p,n).B$. It is easy to compute the derivative. We obtain (e.g., by modifying the derivative calculation in the section on matrices with fixed singular values)

$$dF(C).X = X\Lambda_{pn}C^* + C\Lambda_{pn}X^*.$$

As before we identify $\operatorname{Tan}.Gl(m).\ C$ with the set $\mathbb{C}^{m\times m}C$. For $X \in \mathbb{C}^{m\times m}$ we have

$$dF(C).(XC) = XC\Lambda_{pn}C^* + C\Lambda_{pn}(XC)^* = XB + BX^*.$$

Thus

$$\operatorname{Tan}.\operatorname{inert}^{-1}(p,n).B = \{XB + BX^* : X \in \mathbb{C}^{m\times m}\}.$$

As a check we compute the kernel of the map $dF(I)$ which is given by

$$dF(I) = \mathbb{C}^{m\times m} \to \operatorname{Herm}(m) : X \to X\Lambda_{pn} + \Lambda_{pn}X^*.$$

We have

$$0 = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix} \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_n & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_n & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} X_{11}^* & X_{21}^* & X_{31}^* \\ X_{12}^* & X_{22}^* & X_{32}^* \\ X_{13}^* & X_{23}^* & X_{33}^* \end{pmatrix}$$

$$= \begin{pmatrix} X_{11} & -X_{12} & 0 \\ X_{21} & -X_{22} & 0 \\ X_{31} & -X_{23} & 0 \end{pmatrix} + \begin{pmatrix} X_{11}^* & X_{21}^* & X_{31}^* \\ -X_{12}^* & -X_{22}^* & -X_{32}^* \\ 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} X_{11} + X_{11}^* & X_{21}^* - X_{12} & X_{31}^* \\ X_{21} - X_{12}^* & -(X_{22} + X_{22}^*) & -X_{32}^* \\ X_{31} & -X_{32} & 0 \end{pmatrix}$$

iff

$$X_{11} + X_{11}^* = 0, X_{21}^* = X_{12}, X_{31} = 0, X_{22} + X_{22}^* = 0, X_{32} = 0.$$

Hence

$$\dim.\operatorname{Kernel}.dF(I) = p^2 + 2pn + n^2 + 2pz + 2nz + 2z^2$$

$$= (p+n)^2 + 2(p+n)z + 2z^2.$$

This equals the dimension of $\operatorname{Fix}(\Lambda_{pn})$ and our check is successful. The dimension of the kernel of $dF(C)$ equals the dimension of the kernel of $dF(I)$ by homogeneity.

**Hermitian matrices with fixed displacement inertia.** Let $Z_m$ denote the $m \times m$ *(lower) shift* matrix defined by

$$Z_m(i,j) := \delta(i-1,j).$$

Let $\lrcorner$ be the linear map defined by

$$\lrcorner := \mathbb{C}^{m\times m} \to \mathbb{C}^{m\times m} : X \to X - Z_m X Z_m^T.$$

Let $\mathrm{Herm}(m)$ denote the set of $m \times m$ hermitian matrices; in symbols,

$$\mathrm{Herm}(m) := \{X \in \mathbb{C}^{m\times m} : X^* = X\}.$$

Note that the linear map $\lrcorner$ preserves $\mathrm{Herm}(m)$; in symbols, $\lrcorner\,\mathrm{Herm}(m) = \mathrm{Herm}(m)$.

Let $p$ and $n$ be nonnegative integers satisfying $p+n \le m$. We say that $X \in \mathrm{Herm}(m)$ has *displacement inertia* $(p,n,z)$ where $z := m-(p+n)$ if $\lrcorner.X$ has inertia $(p,n,z)$. The following result is a modification of a theorem due to Kailath, Kung and Morf (1979a,b).

Proposition. *Alternative characterization of displacement inertia.*

*Let $R \in \mathrm{Herm}(m)$. Then $R$ has displacement inertia $(p,n,z)$ iff $p$ and $n$ are the smallest integers such that*

$$R = \sum_{i=1}^{p} P_i P_i^* - \sum_{i=1}^{n} N_i N_i^*$$

*for some lower triangular Toeplitz matrices $P_i, N_i \in \mathbb{C}^{m\times m}$.*

*Proof.* Let $L$ be any $m \times m$ lower triangular Toeplitz matrix; in other words, $L$ has the following form:

$$\begin{pmatrix} a_0 & 0 & 0 & \ldots & 0 \\ a_1 & a_0 & 0 & \ldots & 0 \\ a_2 & a_1 & a_0 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \\ a_{m-1} & a_{m-2} & a_{m-3} & \ldots & a_0 \end{pmatrix}.$$

Note that $L = a_0 I + a_1 Z + \cdots + a_{m-1} Z^{m-1}$ where $Z := Z_m$; that is, $L = p(Z)$ where $p$ is the polynomial defined by $p(x) := a_0 + a_1 x + \cdots + a_{m-1}x^{m-1}$. Hence we have

$$\lrcorner(LL^*) = LL^* - ZLL^*Z^T = L(I - ZZ^T)L^*.$$

Note that $I - ZZ^*$ is the $m \times m$ matrix with one in the (1,1) position and zeros elsewhere; in particular, this matrix has inertia $(1,0,m-1)$. More generally, if

$$R = \sum_{i=1}^{p} P_i P_i^* - \sum_{i=1}^{n} N_i N_i^*,$$

where the $P_i$ and $N_i$ are lower triangular Toeplitz matrices, then the matrix

$$\lrcorner .R = \sum_{i=1}^{p} P_i(I - ZZ^T)P_i^* - \sum_{i=1}^{n} N_i(I - ZZ^T)N_i^*$$

has at most $p$ positive eigenvalues at most $q$ negative eigenvalues.

Now consider the matrix $aa^*$ where $a \in \mathbb{C}^m$; that is $a := (a_0, a_1, \ldots, a_{m-1})^T$. Let $L(a) := p(Z)$ where $p(x) := a_0 + a_1 x + \cdots + a_{m-1}x^{m-1}$. Note that

$$\begin{aligned}\lrcorner(L(a)L(a)^*) &= L(a)L(a)^* - ZL(a)L(a)^*Z^T \\ &= L(a)(I - ZZ^T)L(a)^* = aa^*.\end{aligned}$$

We illustrate with $m = 3$:

$$L(a)(I - ZZ^T)L(a)^* = \begin{pmatrix} a_0 & 0 & 0 \\ a_1 & a_0 & 0 \\ a_2 & a_1 & a_0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \overline{a}_0 & \overline{a}_1 & \overline{a}_2 \\ 0 & \overline{a}_0 & \overline{a}_1 \\ 0 & 0 & \overline{a}_0 \end{pmatrix}$$

$$= \begin{pmatrix} a_1 & 0 & 0 \\ a_1 & 0 & 0 \\ a_2 & 0 & 0 \end{pmatrix} \begin{pmatrix} \overline{a}_0 & \overline{a}_1 & \overline{a}_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} \quad (\overline{a}_0, \overline{a}_1, \overline{a}_2)\ .$$

More generally, if

$$S := \sum_{i=1}^{p} q_i q_i^* - \sum_{i=1}^{n} r_i r_i^*$$

where the $q_i$ and $r_i$ are column vectors, then

$$\lrcorner \left( \sum_{i=1}^{p} L(q_i)L(q_i)^* - \sum_{i=1}^{n} L(r_i)L(r_i)^* \right) = S\ . \qquad \square$$

Proposition. *The $m \times m$ hermitian matrices with displacement inertia $(p, n, z)$ form a homogeneous space with dimension $(p + n)(p + n + 2z)$.*

*Proof.* From the results in the section on hermitian matrices with fixed inertia, we know that the space $\mathrm{inert}^{-1}(p, n) \subseteq \mathrm{Herm}(m)$ is a homogeneous space with dimension $(p + n)(p + n + 2z)$. This proposition then follows at once from the fact that the map $\lrcorner$ is an invertible linear map. $\square$

We shall use dis-inert$^{-1}(p,n)$ to denote the space of hermitian matrices with displacement inertia $(p,n)$; in symbols,

$$\text{dis-inert}^{-1}(p,n) := \{X \in \mathrm{Herm}(m) : \text{displacement-inertia}\, X = (p,n,m-(p+n))\}.$$

The (invertible, linear) map $\lrcorner$ provides us with a smooth $1-1$ correspondence between the surface dis-inert$^{-1}(p,n)$ and the surface inert$^{-1}(p,n)$. We can use this observation to compute the spaces tangent to dis-inert$^{-1}(p,n)$. Let $B \in$ dis-inert$^{-1}(p,n)$. Then $\lrcorner.B \in$ inert$^{-1}(p,n)$, and (from the description of the spaces tangent to inert$^{-1}(p,n)$ in the section on hermitian matrices with fixed inertia) we have

$$\mathrm{Tan}\,.\,\mathrm{inert}^{-1}(p,n).B = \{X(\lrcorner.B) + (\lrcorner.B)X^* : X \in \mathbb{C}^{m\times m}\}.$$

Clearly then

$$\mathrm{Tan}\,.\,\text{dis-inert}^{-1}(p,n).B = \lrcorner^{-1}(\mathrm{Tan}\,.\,\mathrm{inert}^{-1}(p,n).(\lrcorner.B))$$

$$= \{\lrcorner^{-1}(X(\lrcorner.B) + (\lrcorner.B)X^*) : X \in \mathbb{C}^{m\times m}\}.$$

Let

$$\lrcorner.B = \sum_{i=1}^{p} q_i q_i^* - \sum_{i=1}^{n} r_i r_i^*$$

where $\{q_1,\ldots,q_p,r_1,\ldots,r_n\} \subseteq \mathbb{C}^m$ is an independent set of vectors. We consider a typical element of Tan . inert$^{-1}(p,n).(\lrcorner B)$:

$$X(\lrcorner B) + (\lrcorner B)X^* = X\left(\sum_{i=1}^{p} q_i q_i^* - \sum_{i=1}^{n} r_i r_i^*\right) + \left(\sum_{i=1}^{p} q_i q_i^* - \sum_{i=1}^{n} r_i r_i^*\right)X^*$$

$$= \sum_{i=1}^{p}(x_i q_i^* + q_i x_i^*) - \sum_{i=1}^{n}(y_i r_i^* + r_i y_i^*)$$

where $x_i := Xq_i$ and $y_i := Xr_i$. Thus a typical element of Tan . displacement-inertia$^{-1}(p,n).B$ has the following form:

$$\lrcorner^{-1}(X(\lrcorner B) + (\lrcorner B)X^*) = \lrcorner^{-1}\left(\sum_{i=1}^{p}(x_i q_i^* + q_i x_i^*) - \sum_{i=1}^{n}(y_i r_i^* + r_i y_i^*)\right)$$

$$= \sum_{i=1}^{p}(L(x_i)L(q_i)^* + L(q_i)L(x_i)^*) - \sum_{i=1}^{n}(L(y_i)L(r_i)^* + L(r_i)L(y_i)^*).$$

We conclude that

$$\operatorname{Tan}.\,\text{dis-inert}^{-1}(p,n).B$$

$$=\left\{\sum_{i=1}^{p}(L(x_i)L(q_i)^*+L(q_i)L(x_i)^*)-\sum_{i=1}^{n}(L(y_i)L(r_i)^*+L(r_i)L(y_i)^*) : x_1,\dots,x_p,y_1,\dots,y_n\in\mathbb{C}^m\right\}.$$

Note that $uv^* + vu^* = (1/2)((u+v)(u+v)^* - (u-v)(u-v)^*)$. Consequently, we have the following alternative description of this tangent space:

$$\begin{aligned}&\operatorname{Tan}.\,\text{dis-inert}^{-1}(p,n).B\\ &=\Bigg\{(1/2)\left(\sum_{i=1}^{p}L(q_i+x_i)L(q_i+x_i)^*+\sum_{i=1}^{n}L(r_i-y_i)L(r_i-y_i)^*\right)\\ &-(1/2\left(\sum_{i=1}^{p}L(q_i-x_i)L(q_i-x_i)^*)+\sum_{i=1}^{n}L(r_i+y_i)L(r_i+y_i)^*\right):\\ &\qquad x_1,x_2,\dots,x_p,y_1,1/2,\dots,y_n\in\mathbb{C}^m\Bigg\}.\end{aligned}$$

**Hermitian matrices with fixed eigenvalues and fixed displacement inertia.** Let $\lambda := (\lambda_1, \lambda_2, \dots, \lambda_m)$ be an ordered $m$-tuple of real numbers. Recall that we have been using $\text{spec}^{-1}(\lambda)$ to denote the space of all $m\times m$ hermitian matrices with these numbers as eigenvalues; in symbols,

$$\text{spec}^{-1}(\lambda) := \{A\in\operatorname{Herm}(m) : \operatorname{spectrum} A = \lambda\}.$$

In the section on hermitian matrices with fixed eigenvalues, we saw that this isospectral surface $\text{spec}^{-1}(\lambda)$ is a smooth manifold with dimension $m^2-m$ provided the eigenvalues $\lambda_i$ are distinct. Let $p$ and $n$ be nonnegative integers satisfying $p+n\le m$. Recall that we have been using $\text{dis-inert}^{-1}(p,n)$ to denote the space of all $m\times m$ hermitian matrices with inertia $(p,n,z)$ where $z := m-(p+n)$; in symbols,

$$\text{dis-inert}^{-1}(p,n) := \{A\in\operatorname{Herm}(m) : \text{displacement-inertia}\, A = (p,n,z)\}.$$

In the section on hermitian matrices with fixed displacement inertia, we saw that this surface $\text{dis-inert}^{-1}(p,n)$ is a smooth manifold with dimension $(p+n)(p+n+2z) = k(2m-k)$ where $k := p+n$. In this section we shall consider the space $\text{spec}^{-1}(\lambda)\cap\text{dis-inert}^{-1}(p,n)$ which consists of the hermitian matrices with fixed spectrum $\lambda$ and fixed inertia $(p,n)$. We shall see that this surface is generally a smooth manifold with dimension $(p+n)(p+n+2z)-m = k(2m-k)-m$. (The proof is not as satisfying as the proof that the space of tridiagonal hermitian matrices with given spectrum is generally a manifold with dimension $2(m-1)$ which appears in the appendix. Another proof of this result appears in Driessel 1987.)

We begin by reviewing some standard results about submersions. (I mainly follows Guillemin and Pollack 1974.) Let $f : X \to Y$ be a smooth map from manifold $X$ into manifold $Y$, let $x \in X$ and let $y := f(x)$. Then the derivative of $f$ at $x$ is a linear map of the following type

$$df(x) : \operatorname{Tan}.X.x \to \operatorname{Tan}.Y.y$$

where $\operatorname{Tan}.X.x$ denotes the space tangent to $X$ at $x$. If $df(x)$ is surjective then $f$ is a *submersion* at $x$. A point $y \in Y$ is a *regular value* of $f$ if, for all $x \in X, f(x) = y$ implies $df(x)$ is surjective. The following result is very useful for proving that spaces are manifolds.

Preimage Theorem. *Let $f : X \to Y$ be a smooth map between manifolds and let $y \in Y$ be a regular value of $f$. Then the preimage $f^{-1}(y) := \{x \in X : f(x) : y\}$ is a submanifold of $X$ with*

$$\dim f^{-1}(y) = \dim X - \dim Y.$$

*In particular, $f^{-1}$ is a manifold.*

*Proof.* See Guillemin and Pollack 1974. □

The preimage theorem provides another way to see that $\operatorname{spec}^{-1}(\lambda)$ is a manifold with dimension $m^2 - m$ provided the eigenvalues $\lambda_i$ are distinct. We consider the following function

$$f := \operatorname{Herm}(m) \to \mathbb{R}^m : X \to (\operatorname{trace} X, \operatorname{trace} X^2, \ldots, \operatorname{trace} X^n).$$

The following result shows that

$$\operatorname{spec}^{-1}(\lambda) = f^{-1}(\mu_1, \mu_2, \ldots, \mu_m)$$

where $\mu_k := \lambda_1^k + \lambda_2^k + \cdots + \lambda_m^k$.

Proposition. *Let $A, B \in \mathbb{C}^{m\times m}$. Then $A$ and $B$ have the same eigenvalues iff, for all $k = 1, 2, \ldots, m, \operatorname{trace} A^k = \operatorname{trace} B^k$.*

*Proof.* See Horn and Johnson (1985) section 1.2 “The characteristic polynomial” problem 12. In particular, note that if $\operatorname{spectrum} A = \lambda$ then $\operatorname{trace} A^k = \mu_k$. □

We need to compute the derivative of $f$ in order to apply the preimage theorem. We have

$$D(X \to X^k).A.X = A^{k-1}X + A^{k-2}XA + \cdots + XA^{k-1}.$$

Hence, using the chain rule, we get

$$
\begin{aligned}
d(X \to \operatorname{trace} X^k).A.X &= \operatorname{trace}(d(X \to X^k).A.X) \\
&= k \operatorname{trace}(A^{k-1}X) \quad \text{and} \\
df.A.X &= (\operatorname{trace} X, 2 \operatorname{trace} AX, \ldots, m \operatorname{trace} A^{m-1}X).
\end{aligned}
$$

We want to see that $df.A$ is surjective if $A$ has distinct eigenvalues. In other words we want to see that, for all $b \in \mathbb{R}^m$, there exists $X \in \operatorname{Herm}(m)$ such that

$$
\begin{aligned}
\operatorname{trace} X &= \beta_1, \\
\operatorname{trace} AX &= b_2, \\
&\vdots \\
\operatorname{trace} A^{m-1}X &= b_m.
\end{aligned}
$$

Since $\operatorname{trace}(A^k X)$ is invariant under unitary similarity, we may assume without loss of generality that $A = \operatorname{diag}(\lambda)$. We take $X := \operatorname{diag}(x_1, x_2, \ldots, x_m)$. Then the last set of equations reduces to the following set:

$$
\begin{pmatrix}
1 & 1 & 1 & \ldots & 1 & \\
\lambda_1 & \lambda_2 & \lambda_3 & \ldots & \lambda_m & \\
\lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \ldots & \lambda_m^2 & \\
 & \vdots & \vdots & \vdots & & \vdots \\
\lambda_1^{m-1} & \lambda_2^{m-1} & \lambda_3^{m-1} & \ldots & \lambda_m^{m-1} &
\end{pmatrix}
\begin{pmatrix}
x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_m
\end{pmatrix}
=
\begin{pmatrix}
b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m
\end{pmatrix}
$$

The coefficient matrix is the familiar vandermonde matrix which occurs in polynomial interpolation. If the $\lambda_i$ are distinct, its determinant is nonzero.

We now turn to the following result.

Proposition. *Let $p$ and $n$ be nonnegative integers satisfying $1 \leq p + n \leq m$. Then, for almost all $\lambda \in \mathbb{R}^n$, $\operatorname{spec}^{-1}(\lambda) \cap \text{dis-inert}^{-1}(p, n)$ is a manifold with dimension $(p+n)(p+n+2z) - m$ where $z := m - (p + n)$.*

As usual "almost all" means "except for a set of measure zero". Recall the following result.

THEOREM. *(Sard) Let $f : X \to Y$ be a smooth map between manifolds. Then almost every point in $Y$ is a regular value of $f$.*

*Proof.* (of Sard's theorem). See, e.g., Guillemin and Pollack (1974), Spivak (1965), Hirsch (1976) or Milnor (1965). $\square$

*Proof.* (of the proposition) We apply Sard's theorem and the preimage theorem. In particular we take $X := \text{dis-inert}^{-1}(p,n), Y := \mathbb{R}^m$ and we define $f$ by

$$f(A) := (\mu,(A), \mu_2(A), \ldots, \mu_m(A))$$

where $\mu_k(A) := \operatorname{trace} A^k$. $\square$

Recall that two submanifolds $X_1$ and $X_2$ of a manifold $Y$ intersect *transversely* if

$$\forall\ a \in X_1 \cap X_2,\ \operatorname{Tan}.X_1.a + Tan.X_2.a = \operatorname{Tan}.Y.a.$$

From the last proposition, we see (using an elementary dimension argument) that $\text{spec}^{-1}(\lambda)$ and $\text{dis-inert}^{-1}(p,n)$ intersect transversely in $\text{Herm}(m)$ for almost all spectra $\lambda$. Unfortunately, this proposition doesn't tell us specifically which spectra are the typical ones. I would like to find a generic condition on $\lambda$ which guarantees transverse intersection. I have not been able to do so. However, I shall make a conjecture.

*Conjecture.* Let $p$ and $n$ be nonnegative integers satisfying $1 \le p+n \le m$. Let $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{R}^m$. If the $\lambda_i$ are distinct and nonzero then $\text{spec}^{-1}(\lambda)$ and $\text{dis-inert}^{-1}(p,n)$ intersect transversely in $\text{Herm}(m)$.

The following proposition provides some support for this conjecture.

PROPOSITION. *Let $B \in \text{spec}^{-1}(\lambda) \cap \text{dis-inert}^{-1}(p,n)$. Then the following conditions are equivalent:*

$$(*) \qquad \operatorname{Tan}.\text{spec}^{-1}(\lambda).B + \operatorname{Tan}.\text{dis-inert}^{-1}(p,n).B = \text{Herm}(m);$$

$$(**) \qquad \forall\ Y \in \text{Herm}(m),\ [B, \lrcorner^* Y] = 0 \quad \textit{and} \quad Y(\lrcorner B) = 0 \quad \textit{implies} \quad Y = 0$$

*where* $\lrcorner^* := \text{Herm}(m) \to \text{Herm}(m) : X \to X - Z^* X Z.$

*Proof.* In this proof we shall regard $\mathbb{C}^{m\times m}$ as a vector space over the real numbers. It has dimension $2m^2$. We shall use the following inner product on $\mathbb{C}^{m\times m}$:

$$\langle,\rangle := \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m} \to \mathbb{R} : (X,Y) \to \langle \operatorname{Re} X, \operatorname{Re} Y\rangle_F + \langle ImX, ImY\rangle_F$$

where $\operatorname{Re} X$ and $ImX$ are the $m\times m$ real matrices defined by $\operatorname{Re} X + iImX = X$ and $\langle,\rangle_F$ is the Frobenius inner product on $\mathbb{R}^{m\times m}$ defined by

$$\langle,\rangle := \mathbb{R}^{m\times m} \times \mathbb{R}^{m\times m} \to \mathbb{R} : (U,V) \to \operatorname{Trace}(UV^T).$$

Note that

$$\langle X, Y\rangle = \sum_{i,j}((\operatorname{Re} x_{ij})(\operatorname{Re} y_{ij}) + (Imx_{ij})(Imy_{ij})).$$

Also note that the bilinear form $\langle X, Y\rangle$ is symmetric:

$$\langle X, Y\rangle = \langle Y, X\rangle.$$

(In particular this inner product is not the usual Frobenius inner product on $\mathbb{C}^{m\times m}$ which is defined by Trace($XY^*$).) I first used this inner product in Driessel (1987). We shall need some of the elementary results that appear there. For completeness, I shall include the ones we need.

PROPOSITION. *The adjoint of the multiplication-by-a-matrix map*

Let $X, Y, P \in \mathbb{C}^{m\times m}$. Then

(i) $\langle PX, Y\rangle = \langle X, P^*Y\rangle$ and

(ii) $\langle XP, Y\rangle = \langle X, YP^*\rangle$.

*Proof.* For part (i) we have

$$\begin{aligned}
&\langle(\operatorname{Re} P)(\operatorname{Re} X) - (ImP)(ImX), \operatorname{Re} Y\rangle + \langle(\operatorname{Re} P)(ImX) + (ImP)(\operatorname{Re} X), ImY\rangle \\
&= \langle \operatorname{Re} X, (\operatorname{Re} P)^T(\operatorname{Re} Y) + (ImP)^T(ImY)\rangle + \langle ImX, (\operatorname{Re} P)^T(ImY) - (ImP)^T(\operatorname{Re} Y)\rangle \\
&= \langle X, P^*Y\rangle.
\end{aligned}$$

The calculation for part (ii) is similar. □

COROLLARY. *The adjoint of the displacement map.*

Let $X, Y \in \mathbb{C}^{m\times m}$. Then

$$\langle \lrcorner X, Y\rangle = \langle X, \lrcorner^* Y\rangle.$$

*Proof.* We have

$$\begin{aligned}
\langle \lrcorner X, Y\rangle &= \langle X - ZXZ^*, Y\rangle = \langle X, Y\rangle - \langle ZXZ^*, Y\rangle \\
&= \langle X, Y\rangle - \langle X, Z^*YZ\rangle = \langle X, \lrcorner^* Y\rangle. \qquad \square
\end{aligned}$$

As usual we shall use the bracket notation for the Lie product of matrices:

$$[X, Y] := XY - YX.$$

Proposition. *The adjoint of the Lie-multiplication-by-a-matrix map.*

Let $X, Y, P \in \mathbb{C}^{m\times m}$. Then

$$\langle [P, X], Y\rangle = \langle X, [P^*, Y]\rangle.$$

*Proof.* We have

$$\langle PX, Y\rangle - \langle XP, Y\rangle = \langle X, P^*Y\rangle - \langle X, YP^*\rangle. \qquad \square$$

As usual we shall use $\mathrm{Skew}(m)$ to denote the space of $m \times m$ skew-hermitian matrices:

$$\mathrm{Skew}(m) := \{K \in \mathbb{C}^{m\times m} : K^* = -K\}.$$

We have the following orthogonality relation between the spaces $\mathrm{Herm}(m)$ and $\mathrm{Skew}(m)$.

Proposition. $\mathrm{Herm}(m)^\perp = \mathrm{Skew}(m)$.

*Proof.* Let $A \in \mathrm{Herm}(m)$ and $K \in \mathrm{Skew}(m)$. Then

$$\langle A, K\rangle = \langle I, A^*K\rangle = \langle I, AK\rangle \quad \text{and}$$
$$\langle A, K\rangle = \langle AK^*, I\rangle = -\langle AK, I\rangle = -\langle I, AK\rangle.$$

Hence $\langle A, K\rangle = 0$. Also note that

$$\dim \mathrm{Herm}(m) + \dim \mathrm{Skew}(m) = m^2 + m^2 = \dim \mathbb{C}^{m\times m}. \qquad \square$$

*Proof.* (of the proposition concerning $\mathrm{spec}^{-1}(\lambda)$ and $\mathrm{dis\text{-}inert}^{-1}(p, n)$ continued). Recall from the section on hermitian matrices with fixed spectra and the section on hermitian matrices with fixed displacement inertia that

$$\mathrm{Tan}\,.\,\mathrm{spec}^{-1}(\lambda).B = \{[B, K] : K \in \mathrm{Skew}(m)\} \quad \text{and}$$
$$\mathrm{Tan}\,.\,\mathrm{dis\text{-}inert}^{-1}(\lambda).B = \lrcorner^{-1}\{X(\lrcorner B) + (\lrcorner B)X^* : X \in \mathbb{C}^{m\times m}\}.$$

Also recall that $\dim(\mathrm{Tan}\,.\,\mathrm{spec}^{-1}(\lambda).B) = m^2 - m$ and $\dim(\mathrm{Tan}\,.\,\mathrm{dis\text{-}inert}^{-1}(p, n).B) = k(2m - k)$ where $k := p + n = \mathrm{rank}(\lrcorner B)$. Let

$$M := \lrcorner(\mathrm{Tan}\,.\,\mathrm{spec}^{-1}(\lambda).B) = \{\lrcorner[B, K] : K \in \mathrm{Skew}(m)\} \quad \text{and}$$
$$N := \lrcorner(\mathrm{Tan}\,.\,\mathrm{dis\text{-}inert}^{-1}(p, n).B) = \{X(\lrcorner) + (\lrcorner B)X^* : X \in \mathbb{C}^{m\times m}\}.$$

Clearly condition $(*)$ is equivalent to the following one:

$$(*') \qquad M + N = \mathrm{Herm}(m).$$

Since $(M+N)^\perp = M^\perp \cap N^\perp$, condition $(*')$ is equivalent to the following one:

$$(*'') \qquad M^\perp \cap N^\perp = 0.$$

These orthogonal complements are taken in the space of hermitian matrices. In particular,

$$M^\perp = \{Y \in \mathrm{Herm}(m) : \forall\ K \in \mathrm{Skew}(m), \langle \lrcorner[B,K], Y\rangle = 0\} \quad \text{and}$$

$$N^\perp = \{Y \in \mathrm{Herm}(m) : \forall\ X \in \mathbb{C}^{m\times m}, \langle X(\lrcorner B) + (\lrcorner B)X^*, Y\rangle = 0\}.$$

*Claim.* $M^\perp = \{Y \in \mathrm{Herm}(m) : [B, \lrcorner^* Y] = 0\}$.

Let $Y \in \mathrm{Herm}(m)$. Note that $\lrcorner^* Y = Y - Z^*YZ \in \mathrm{Herm}(m)$ and hence $[B, \lrcorner^* Y] \in \mathrm{Skew}(m)$. Thus we have the following equivalent conditions:

$$\forall\ K \in \mathrm{Skew}(m), \langle \lrcorner[B,K], Y\rangle = 0;$$

$$\forall\ K \in \mathrm{Skew}(m), \langle K, [B, \lrcorner^* Y]\rangle = 0;$$

$$[B, \lrcorner^* Y] = 0.$$

This completes the proof of the claim.

*Claim.* $N^\perp = \{Y \in \mathrm{Herm}(m) : Y(\lrcorner B) = 0\}$.

Let $A := \lrcorner B$. We have

$$\langle XA + AX^*, Y\rangle = \langle XA, Y\rangle + \langle AX^*, Y\rangle$$

$$= \langle X, YA\rangle + \langle YA, X\rangle = 2\langle X, YA\rangle.$$

Hence the following conditions are equivalent:

$$\forall\ X \in \mathbb{C}^{m\times m}, \langle XA + A^*, Y\rangle = 0;$$

$$\forall\ X \in \mathbb{C}^{m\times m}, 2\langle X, YA\rangle = 0;$$

$$YA = 0.$$

This completes the proof of the claim and the proof of the proposition. $\square$

The next proposition characterizes solutions of $AX = 0$ where $A, X \in \mathrm{Herm}(m)$.

PROPOSITION. *Let* $A := \Sigma\alpha_i a_i a_i^*, \alpha_i \in \mathbb{R}, \alpha_i \neq 0, a_i \in \mathbb{C}^m$ *be a spectral decomposition of* $A$ *and let* $X = \Sigma\gamma_i c_i c_i^*, \gamma_i \in \mathbb{R}, \gamma_i \neq 0, c_i \in \mathbb{C}^m$ *be a spectral decomposition of* $X$. *Then*

$$AX = 0 \quad \textit{iff} \quad \forall\, i,j, \quad a_i^* c_j = 0.$$

*Proof.* ($\Longrightarrow$) We have

$$
\begin{aligned}
0 = AX &= (\Sigma\alpha_i a_i a_i^*)(\Sigma\gamma_k c_k c_k^*) \\
&= \sum_{i,k} \alpha_i\gamma_k(a_i^* c_k) a_i c_k^* \\
\Longrightarrow &\quad \{\text{by the orthogonality of the } c_k\text{'s}\} \\
\forall\, j, 0 &= \sum_{i,k} \alpha_i\gamma_k(a_i^* c_k) a_i c_k^* c_j \\
&= \sum_{i,k} \alpha_i\gamma_k(a_i^* c_k) a_i \delta_{jk} \\
&= \sum_{i} \alpha_k\gamma_j(a_i^* c_j) \\
\Longrightarrow &\quad \{\text{by the independence of the } a_i\} \\
\forall\, i,j, 0 &= \alpha_i\gamma_j(a_k^* c_j) \\
\Longrightarrow &\quad \{\text{since} \quad 0 \neq \alpha_i\gamma_j\} \\
\forall\, i,j, 0 &= a_i^* c_j.
\end{aligned}
$$

The other implication is an easy verification. □

**Appendix: Tridiagonal Hermitian matrices with fixed eigenvalues.** Let $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_m)$ be an ordered $m$-tuple of real numbers. We use $\operatorname{spec}^{-1}(\lambda)$ to denote the space of all $m \times m$ hermitian matrices with these numbers as eigenvalues; in symbols,

$$\operatorname{spec}^{-1}(\lambda) := \{A \in \operatorname{Herm}(m) : \operatorname{spectrum} A = \lambda\},$$

where $\operatorname{Herm}(m)$ denotes the space of all $m \times m$ hermitian matrices. In the section on hermitian matrices with fixed eigenvalues we saw that this isospectral surface $\operatorname{spec}^{-1}(\lambda)$ is a smooth manifold with dimension $m^2 - m$ provided the eigenvalues $\lambda_i$ are distinct. We also saw that for $A \in \operatorname{spec}^{-1}(\lambda)$ the space tangent to $\operatorname{spec}^{-1}(\lambda)$ at $A$ consists of the matrices $[A, K]$ where $K$ is skew hermitian; in symbols,

$$\operatorname{Tan}.\operatorname{spec}^{-1}(\lambda).A = \{[A, K] : K \in \operatorname{skew}(m)\}.$$

Let $\mathrm{Tridiag}(m)$ denote the real linear space consisting of tridiagonal hermitian matrices. Since this space is linear, we have, for every $A \in \mathrm{Tridiag}(m), \mathrm{Tan}\,.\,\mathrm{Tridiag}(m).A = \mathrm{Tridiag}(m)$. Note that

$$\begin{aligned} &\mathrm{Tan}\,.\,\mathrm{spec}^{-1}(\lambda).A \cap \mathrm{Tan}\,.\,\mathrm{Tridiag}(m).A \\ &\qquad = \{X \in \mathrm{Tridiag}(m) : \exists K \in \mathrm{Skew}(m), X = [A, K]\}. \end{aligned}$$

The following proposition explicitly describes this intersection of tangent spaces. Using this proposition it is easy to show that $\mathrm{spec}^{-1}(\lambda)$ and $\mathrm{Tridiag}(m)$ usually intersect transversely in $\mathrm{Herm}(m)$.

PROPOSITION. *Let*

$$A := \begin{pmatrix} a_1 & b_1 & 0 & \dots & 0 & 0 \\ \overline{b}_1 & a_2 & b_2 & \dots & 0 & 0 \\ 0 & \overline{b}_2 & a_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \dots & \overline{b}_{n-1} & a_n \end{pmatrix}$$

*be a tridiagonal hermitian matrix and let $k := (0, -k_{12}, -k_{13}, \dots, -k_{1n})^*$. If the $b_i$ are non zero then there is a unique $X \in \mathrm{Tridiag}(m)$ and $K \in \mathrm{Skew}(m)$ such that $k$ is the first column of $K$, the diagonal entries of $K$ are zero and $[A, K] = X$.*

I illustrate the constructive method for determining $K$ and $X$ when $n = 4$. We have

$$A := \begin{pmatrix} a_1 & b_1 & 0 & 0 \\ \overline{b}_1 & a_2 & b_2 & 0 \\ 0 & \overline{b}_2 & a_3 & b_3 \\ 0 & 0 & \overline{b}_3 & a_4 \end{pmatrix},$$

$$X := \begin{pmatrix} x_1 & y_1 & 0 & 0 \\ \overline{y}_1 & x_2 & y_2 & 0 \\ 0 & \overline{y}_2 & x_3 & y_3 \\ 0 & 0 & \overline{y}_3 & x_4 \end{pmatrix}, \quad \text{and}$$

$$K := \begin{pmatrix} 0 & k_{12} & k_{13} & k_{14} \\ -\overline{k}_{12} & 0 & k_{23} & k_{24} \\ -\overline{k}_{13} & -\overline{k}_{23} & 0 & k_{34} \\ -\overline{k}_{14} & -\overline{k}_{24} & -\overline{k}_{34} & 0 \end{pmatrix} = (k_1, k_2, k_3, k_4)$$

where $k_1, k_2, k_3$ and $k_4$ denote the columns of $K$. We can rewrite the equation $AK - KA = X$ as follows:

$$(Ak_1, Ak_2, Ak_3, Ak_4) - (a_1k_1 + \overline{b}_1k_2, b_1k_1 + a_2k_2 + \overline{b}_2k_3, b_2k_2 + a_3k_3 + \overline{b}_3k_3, b_3k_3 + a_4k_4) = X.$$

We consider the first column. Since $k_1 = k$ is given, $Ak_1$ and $a_1k_1$ are determined. We specifically consider the (1,4) entry of the matrix equation; from it we see that $-\overline{k}_{24}$ is determined. We next specifically consider the (1,3) entry; from it we see that $-\overline{k}_{23}$ is determined. We next consider the (1,2) entry; from it we see that $\overline{y}_1$ is determined; we also use here the condition $k_{22} = 0$. Considering the (1,1) entry shows that $x_1$ is determined. Moving on to the second column we similarly successively consider the (2,4), (2,3) and (2,2) entries to determine respectively $-\overline{k}_{34}, \overline{y}_2$ and $x_2$. Clearly we can continue in this way to fully determine $K$ and $X$.

The last proposition seems to be related to an infinitesimal version of the following theorem.

THEOREM. *(Lanczos tridiagonalization)*

*Let $A$ be a $m \times m$ hermitian matrix and let $q \in \mathbb{C}^m$ be a unit vector: $q^*q = 1$. Then there is a unitary matrix $Q$ with first column equal to $q$ such that $Q^*AQ$ is tridiagonal.*

This theorem is discussed in many texts. See, e.g., Horn and Johnson (1985), Golub and Van Loan (1983) or Strang (1986).

## REFERENCES


ARENS, R., *Topologies for homeomorphism groups*, Amer. J. Math. 68, pp. 593–610 (1946).

BIRKHOFF, G. AND MACLANE, *A Survey of Modern Algebra*, Macmillan(1953).

BOJANCZYK, A.W.; BRENT, R.P. AND DE HOOG, F.R., *QR factorization of Toeplitz matrices*, Numer. Math. 49 (1986), pp. 81–94.

CHUN, J. AND KAILATH, T., *Displacement structure for Hankel, Vandermonde and related (derived) matrices*, Lin. Alg. Appl. 151 (1991), pp. 199–227.

CHUN, J.; KAILATH, T. AND LEV-ARI, H., *Fast parallel algorithms for QR and triangular factorization*, SIAM J. Sci. Stat. Comp. 8 (1987),, pp. 899–913.

COMON, P., *Displacement rank of pseudo-inverses*, IEEE 1992 International Conference on Acoustics Speech and Signal Processing, March 23–26, 1992 San Francisco, California (1992).

CYBENKO, G. AND BERRY M., *Hyperbolic Householder algorithms for factoring structured matrices*, SIAM J. Matrix Anal. Appl. 11 (1990), pp. 499–520.

CYBENKO, G., *Fast Toeplitz orthogonalization using inner products*, SIAM J. Sci. Stat. Comput. 8, (1987), pp. 734–740.

DRIESSEL, K.R., *On Isospectral Surfaces in the Space of Tridiagonal Matrices*, Technical Report, Department of Mathematical Sciences, Clemson University, Clemson, SC29634–1907 (1987).

DRIESSEL, K.R., *Some Homogeneous Spaces of Matrices*, Technical Report, Department of Mathematics, Idaho State University, Pocatello, Idaho (1991).

GOLUB, G.H. AND VAN LOAN, C.F., *Matrix Computations*, Johns Hopkins University Press (1983 and 1989).

GUILLEMIN, V. AND POLLACK, A., *Differential Topology*, Prentice–Hall, Inc. (1974).

HEINIG, G. AND ROST, K., *Algebraic Methods for Toeplitz-like Matrices and Operators*, Birkhäuser Verlag (1984).

HIRSCH, M.W., *Differential Topology*, Springer–Verlag (1976).

HORN, R.A. AND JOHNSON, C.R., *Matrix Analysis*, Cambridge University Press, Cambridge, England (1985).

KAILATH, T.; KUNG, S.-Y. AND MORF, M., *Displacement Ranks of Matrices*, AMS Bull. 1, (1979a), 769–773.

KAILATH, T.; KUNG, S.-Y. AND MORF, M., *Displacement Ranks of Matrices and Linear Equations*, J. Math. Anal. Appl. 68 (1979b), 395–407.

MILNOR, J.W., *Topology From the Differentiable Viewpoint*, University Press of Virginia, Charlottesville (1965).

MONTGOMERY, D. AND ZIPPIN, L., *Topological Transformation Groups*, John Wiley & Sons (Reprinted 1974 by Robert E. Krieger Publishing Co.) (1955).

NAGY, J.G., *Fast inverse QR factorization for Toeplitz matrices*, preprint (1991).

NAGY, J.G. AND PLEMMONS, R.J., *Some fast Toeplitz least squares algorithms*, Proc. Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations II vol. 1566, San Diego, CA (July, 1991).

NAYLOR, A.W. AND SELL, G.R., *Linear Operator Theory in Engineering and Science*, Springer–Verlag (1982).

SPIVAK, M., *Calculus on Manifolds*, Addison–Wesley Publishing Company (1965).

STRANG, G., *Introduction to Applied Mathematics*, Wellesley–Cambridge Press, Wellesley, Massachusetts (1986).

STRANG, G., *Linear Algebra and Its Applications*, Academic Press (1980).

SWEET, D.R., *Fast Toeplitz orthogonalization*, Numer. Math. 43 (1984), pp. 1–21.

SWEET, D.R., *Fast block Toeplitz orthogonalization*, Numer. Math. 58 (1991), pp. 613–629.

VARADARAJAN, V.S., *Lie Groups, Lei Algebras, and Their Representations*, Springer–Verlag (1984).

**Acknowledgement.** I did this research during the 1991-1992 academic year while visiting the Institute for Mathematics and its Applications (IMA) at the University of Minnesota. I thank the members of the IMA for their hospitality. In particular they have created a very stimulating intellectual environment. I also wish to especially thank Patricia V. Brick (IMA) for her careful typing of this report.