SRC TR 87-173

Generalized Eulerian Numbers and the Topology of the Hessenberg Variety of a Matrix

by

F. De Mari and M. A. Shayman

GENERALIZED EULERIAN NUMBERS AND THE TOPOLOGY OF THE HESSENBERG VARIETY OF A MATRIX*

Filippo De Mari

Department of Mathematics

Washington University

St. Louis, Missouri 63130

Mark A. Shayman

Electrical Engineering Department and Systems Research Center

University of Maryland

College Park, Maryland 20742

July 8, 1987

* Research partially supported by the National Science Foundation under Grant ECS-8696108.

Abstract: Let $A \in g\ell(n, \mathbb{C})$ and let p be a positive integer. The Hessenberg variety of degree p for A is the subvariety Hess (p, A) of the complete flag manifold consisting of those flags $S_1 \subset ... \subset S_{n-1}$ in $\mathbb{C}^n$ which satisfy the condition $AS_i \subset S_{i+p}$, for all i. We show that if A has distinct eigenvalues, then Hess (p, A) is smooth and connected. The odd Betti numbers of Hess (p, A) vanish, while the even Betti numbers are given by a natural generalization of the Eulerian numbers.

I. Introduction

In this paper, we describe some of the basic topological properties of the Hessenberg varieties of a matrix. These are certain subsets of the flag manifold which are closely related to Hessenberg form for a matrix, and especially to the use of Hessenberg form in the efficient implementation of the QR-algorithm for matrix eigenvalue problems.

A. Mathematical Problem

Let n be a positive integer, and let $K = (k_1, ..., k_d)$ be a d-tuple of positive integers satisfying $0 < k_1 < ... < k_d < k_{d+1} = n$. Let F denote either the field of real numbers or the field of complex numbers. Let Flag (K, F^n) denote the (partial) flag manifold consisting of d-tuples $(S_1, ..., S_d)$, where S_i is a k_i-dimensional subspace of F^n and $S_1 \subset ... \subset S_d$. If $K = K_o = (1, 2, ..., n-1)$, then Flag (K_0, F^n) is referred to as a complete flag manifold and is denoted by Flag (F^n).

Let $A \in g\ell(n, F)$, a linear operator on F^n, and let p be a nonnegative integer. We say that $(S_1, ..., S_d)$ is a degree p Hessenberg flag for A if and only if

$$A(S_i) \subset S_{i+p}, \qquad i = 1, ..., d-p,$$

and we denote by Hess (p, K, A) the subset of Flag (K, F^n) consisting of all such flags. We refer to Hess (p, K, A) as the Hessenberg variety of degree p for A. Note that

$$\text{Hess } (0, K, A) \subset \text{ Hess } (1, K, A) \subset \cdots \subset \text{ Hess } (d, K, A) = \text{ Flag } (K, F^n)$$

and if $A \in GL(n, F)$, Hess $(0, K, A)$ consists of those flags which are fixed by A - i.e., for which $A(S_i) = S_i$ for all i. the notion of Hessenberg flags was introduced for the case $K = K_0, p = 1$ by Ammar and Martin [1] in connection with a geometric approach to the QR-algorithm. (See also Ammar [2].) This connection is briefly described below.

Most of this paper is concerned with the case where $K = K_0$ and where A has distinct eigenvalues all of which are in F. Thus, it should be regarded essentially as a study of the Hessenberg varieties of the complete complex flag manifold Flag $(\mathbb{C}^n)$, under generic assumptions on A. In particular, we focus our attention on the three issues of smoothness, connectedness and homology. Under the aforementioned assumptions, we show that for $p \geq 1$, Hess (p, K_0, A) is a smooth connected subvariety of Flag (F^n) of dimension $p(2n-p-1)/2$ over F.

In the case of a complex matrix with distinct eigenvalues, we characterize the Betti numbers of Hess (p, K_0, A) in terms of a combinatorial property of the elements of $\sum(n)$, the symmetric group of permutations on n letters. For $\sigma \in \sum(n)$, we define the p-th Eulerian dimension of σ to be

$$E_p(\sigma) \triangleq \text{ card } \{(i,j) \mid 1 \leq i,j \leq n, \ 1 \leq i-j \leq p, \ \sigma(i) < \sigma(j)\}.$$

In the special case where $p = 1, E_1(\sigma)$ simply counts the number of "falls" in the permutation σ, a fall being a value of i for which $\sigma(i) > \sigma(i+1)$. In the special case $p = n-1, E_{n-1}(\sigma)$ coincides with the length function on the Weyl group $\sum(n)$ [3]. We show that the odd Betti numbers of Hess (p, K_0, A) vanish, while the 2k-th Betti number is equal to

$$A_p(n, k+1) \triangleq \text{ card } \{\sigma \in \Sigma(n) \mid E_p(\sigma) = k\}, \qquad k = 0, ..., p(2n-p-1)/2$$

In the special case where $p = 1$, these Betti numbers coincide with the well-known Eulerian numbers, while in the special case where $p = n-1$, these numbers are also well-known, and are sometimes referred to as the Mahonian numbers [4]. We refer to the numbers $\{A_p(n,k)\}$ as generalized Eulerian numbers of degree p. Several generalizations of the Eulerian numbers have been considered in the combinatorics literature. (See e.g. [5][6][7][8][9].) However, we have not been able to locate any references to the particular generalization $\{A_p(n,k)\}$ which arises in our study of the Hessenberg varieties.

If A is nonsingular and has distinct eigenvalues, the variety of fixed flags, Hess $(0, K_0, A)$, is uninteresting topologically, consisting of a finite set of points. Thus, it is somewhat surprising that, for such a matrix, Hess $(p, K_0, A), p \geq 1$, has such a rich topological structure.

B. Relationship to Hessenberg Matrices and the QR-Algorithm

In the remainder of this section, we briefly describe the relationship of the Hessenberg varieties and Hessenberg matrices, with particular application to the QR-algorithm. An $n \times n$ matrix B is said to be in (upper) Hessenberg-form if $b_{ij} = 0$ whenever $i - j > 1$. Let $G \triangleq GL(n, F)$, and let $U \triangleq U(F^n)$ be the subgroup of G consisting of those elements which are upper-triangular. Let $\pi : G \to G/U \cong$ Flag (F^n) be the natural projection; we will write $< g >$ for $\pi(g), g \in G$. If $T \in G$, then $< T > = (S_1, ..., S_{n-1})$ where S_i is the subspace spanned by the first i columns of T.

Let $A \in g\ell(n,F)$. Let $T \in G$, and let B be the matrix for A relative to the ordered basis given by the columns of T. Since $AT = TB$, it follows that B is a Hessenberg matrix if and only if $< T > \in$ Hess $(1, K_0, A)$. Thus, the degree one Hessenberg flags for A correspond to those ordered bases relative to which A has a Hessenberg representation. More generally, the flags in Hess (p, K_0, A) correspond to generalized Hessenberg representations for which B satisfies the condition that $B_{ij} = 0$ whenever $i - j > p$, while the flags in Hess (p, K, A) correspond to block generalized Hessenberg representations.

The QR-algorithm is the most commonly used method for finding the eigenvalues and invariant subspaces of a matrix. Ammar and Martin [1] and Shub and Vasquez [10] have described how the QR-algorithm may be interpreted as a linear-induced dynamical system on Flag (F^n). The QR-algorithm applied to a given $A \in GL(n,F)$ generates a sequence of matrices $\{A_i\}_{i\geq 0}$, all similar to A, obtained by successively performing QR-factorizations. A_0 is a representation of A relative to a chosen initial basis P-i.e., $A_0 = P^{-1}AP.A_0$ is factored into the product Q_0R_0 of an orthogonal (if $F = \mathbb{R}$) or unitary (if $F = \mathbb{C}$) matrix Q_0 and an upper-triangular matrix R_0 with positive diagonal entries. Then one defines $A_1 = R_0Q_0$. Inductively, if Q_iR_i is the QR-factorization of A_i, then A_{i+1} is defined to be R_iQ_i.

The corresponding dynamical system on Flag (F^n) is obtained as follows: Setting $P_i \triangleq Q_0 \cdot \cdots \cdot Q_i$, it follows easily that

$$A_{i+1} = (PP_i)^{-1}A(PP_i) \tag{1}$$

$$A(PP_i) = (PP_{i+1})R_{i+1}. \tag{2}$$

The first equation shows that the sequence $\{A_i\}_{i\geq 1}$ is completely determined by the sequence $\{PP_i\}_{i\geq 0}$. Thus, the QR-algorithm can be viewed as generating a sequence of orthogonal (or unitary) matrices. However, it is demonstrated in [1] and [10] that the convergence properties of the algorithm can be characterized by the behaviour of the sequence $\{< PP_i >\}_{i\geq 0}$ in Flag (F^n). (Note that $< PP_i >$ determines PP_i up to right multiplication by an element of $U \cap O(n)$ if $F = \mathbb{R}$ or $U \cap U(n)$ if $F = \mathbb{C}$. $U \cap O(n)$ consists of the 2^n matrices $\{\text{diag}(\pm 1, \ldots, \pm 1)\}$, while $U \cap U(n)$ is the maximal torus $\{\text{diag}\ (e^{i\theta_1}, \cdots, e^{i\theta_n})\}$.)

Let $\overline{A}$ denote the diffeomorphism of $G/U \cong$ Flag (F^n) induced by A - i.e., $\overline{A}(< T >) \triangleq$

$< AT >$. It follows from (2) that

$$< PP_i >= \overline{A}^i < PP_o >$$

Since $\overline{A} < P >=< AP >=< PP_0R_0 >=< PP_0 >$, we obtain

$$< PP_i >= \overline{A}^{i+1} < P > .$$

<u>This shows that the QR-algorithm starting at the initial representation</u> $A_0 = P^{-1}AP$ <u>for</u> A <u>corresponds to the linear-induced dynamical system</u> $\overline{A} : G/U \rightarrow G/U$ <u>initialized at the flag</u> $< P >$.

For a differentiable dynamical system on a compact manifold, there is a close relationship between the properties of the dynamical system and the topology (particularly the homology) of the underlying manifold [11]. Thus, Shub and Vasquez [10] are able to relate the behavior of the QR-algorithm to topological properties (e.g., the Bruhat decomposition) of Flag (F^n). However, in actual practice the QR-algorithm is rarely applied directly to a given matrix A. Instead, A is first reduced to Hessenberg form using a finite sequence of elementary orthogonal (unitary) transformations (Householder transformations) [12]. In other words, P is chosen such that $A_0 = P^{-1}AP$ is in Hessenberg form. Thus $< P >$ belongs to the subvariety Hess $(1, K_0, A)$. This initial reduction greatly improves the efficiency of the QR-algorithm since a QR-step for a Hessenberg matrix requires $O(n^2)$ arithmetic operations as compared with $O(n^3)$ operations for a full matrix. Moreover, if one starts with A_0 in Hessenberg form, all the matrices A_i produced by the algorithm are in Hessenberg form as well.

If $< T >\in$ Hess $(1, K_0, A)$, then $AT = TB$ with B a Hessenberg matrix. Since $\overline{A} < T >=< AT >$ and $A(AT) = (AT)B$, it follows that Hess $(1, K_0, A)$ is an invariant subset for the dynamical system $\overline{A} : G/U \rightarrow G/U$. Thus, the efficient implementation of the QR-algorithm by initial reduction of A to Hessenberg form corresponds to the restriction of the dynamical system $\overline{A} : G/U \rightarrow G/U$ to the invariant subset Hess $(1, K_0, A)$. Consequently, the properties of the QR-algorithm as applied to Hessenberg matrices would be expected to reflect the topological properties of the subvariety Hess $(1, K_0, A)$. This is the primary motivation for our investigation of the topology of Hess $(1, K_0, A)$ and of its natural generalizations $\{\text{Hess}(p, K_0, A)\}$.

II. Preliminaries on the Flag Manifold

Let n be a positive integer, and let $K = (k_1, ..., k_d)$ be a d-tuple of positive integers satisfying $0 < k_1 < ... < k_d < d_{d+1} = n$. Let F denote either the field of real numbers or the field of complex numbers. Let Flag (K, F^n) denote the (partial) flag manifold consisting of d-tupes $(S_1, \ldots, S_d)$ where S_i is a k_i-dimensional subspace of F^n and $S_1 \subset \ldots \subset S_d$. Flag (K, F^n) can be given the structure of a smooth manifold as follows: Let $\{e_1, ..., e_n\}$ be the standard basis vectors for F^n and let $W_j = sp\{e_1, ..., e_j\}$ (where "sp" denotes span). Let $GL(n, F)$ be the general linear group of F^n and let $P(K, F^n)$ denote the parabolic subgroup consisting of those matrices in $GL(n, F)$ which are upper-triangular as matrices of blocks when the rows and columns are partitioned according to the partition $(k_1, k_2 - k_1, \cdots, k_d - k_{d-1}, n - k_d)$ of n. The group $GL(n, F)$ acts transitively on Flag (k, F^n) by $(g, S_1, \cdots, S_d)) \rightarrow (g(S_1), \cdots, g(S_d))$, and the stabilizer of the flag $(W_{k_1}, \cdots, W_{k_d})$ is the subgroup $P(K, F^n)$. Consequently, Flag (K, F^n) can be identified with the homogeneous space $GL(n, F)/P(K, F^n)$. We observe that if $K = K_0 = (1, 2, ..., n-1)$, then the complete flag manifold Flag $(F^n) \triangleq$ Flag (K_0, F^n) can be identified with $GL(n, F)/U(F^n)$, where $U(F^n)$ denotes the subgroup of $GL(n, F)$ consisting of upper-triangular matrices.

We now describe in more detail the analytic structure of the complete flag manifold Flag $(F^n) = GL(n, F)/U(F^n)$. Let $\sum(n)$ denote the symmetric group of $n \times n$ permutation matrices and let $L^+(F^n)$ denote the strict lower-triangular subgroup consisting of lower-triangular matrices with ones along the main diagonal. It is easy to see that the sets $ch(\sigma) = \{< \sigma X > | X \in L^+(F^n)\}, \sigma \in \sum(n)$ (where $< \cdot >$ denotes the left cosets modulo $U(F^n)$ for an element in $GL(n, F)$) give rise to a system of analytic charts for Flag (F^n), of dimension $n(n-1)/2$ over F. In fact, under the identification Flag $(F^n) = GL(n, F)/U(F^n)$ each flag $(S_1, ..., S_{n-1})$ has a representative $\sigma X, \sigma \in \sum(n)$ and $X \in L^+(F^n)$. Moreover, for a fixed $\sigma \in \sum(n)$, X is unique.

Next we recall the Bruhat decomposition of the flag manifold Flag (K, F^n). We choose, and fix, a basis $\{e_1, ..., e_n\}$ for F^n. The Bruhat decomposition of Flag (K, F^n) will be constructed relative to the choice of this basis. For any ℓ-dimensional linear subspace S of F^n, the signature of S is defined as the set $\text{sig}(S) = \{\beta_1, ..., \beta_\ell\}$, where $\beta_1 < \cdots < \beta_\ell$ are the "jump points" of S, i.e.,

$$S \cap \mathrm{sp}\{e_1, \cdots, e_{\beta_i - 1}\} \neq S \cap sp\{e_1, ..., e_{\beta_i}\}, \qquad i = 1, ..., \ell.$$

It is easily seen that $S_1 \subset S_2 \subset F^n$ implies $\mathrm{sig}(S_1) \subset \mathrm{sig}(S_2)$. In particular, for any flag $S = (S_1, \cdots, S_d) \in$ Flag (K, F^n), there is an increasing sequence of signatures

$$\mathrm{sig}(S) = (\mathrm{sig}(S_1), \cdots, \ \mathrm{sig}(S_d))$$

with

$$(i) \qquad \mathrm{sig}(S_1) \subset \cdots \subset \mathrm{sig}(S_d) \subset \{1, \cdots, n\}$$

$$(ii) \qquad \mathrm{card}\ \mathrm{sig}(S_j) = k_j \text{ for } j = 1, \cdots, d.$$

Any such sequence $s = (s_1, \cdots, s_d)$ of subsets $s_j \subset \{1, ..., n\}$ satisfying i) and ii) is called a flag symbol. Let $\mathcal{S}(K, n)$ denote the set of all such flag symbols corresponding to K. Then $\mathcal{S}(K, n)$ has

$$\binom{n}{k_1, k_2 - k_1, \cdots, k_d - k_{d-1}, n - k_d} = \binom{n}{k_1} \binom{n - k_1}{k_2 - k_1} \cdots \binom{n - k_{d-1}}{k_d - k_{d-1}}$$

elements, where

$$\binom{n}{a_1, \cdots, a_d} = \frac{n!}{a_1! \cdots a_d!}$$

is the multinomial coefficient.

For any flag symbol $s = (s_1, \cdots, s_d) \in \mathcal{S}(K, n)$, the Bruhat cell of the flag manifold is defined as the set

$$B_s = \{S \in \ \mathrm{Flag}(K, F^n) \mid sig(S) = s\}$$

It is an easy exercise to prove that B_s is in fact an F-analytic submanifold of Flag (K, F^n) and it is diffeomorphic to an affine space F^n, where g is the dimension of B_s.

In the particular case in which $K = K_0 = (1, 2, \cdots, n-1)$, there is a natural bijection between $\mathcal{S}(K_0, n)$ and the symmetric group $\sum(n)$ given by:

$$\begin{aligned} s \ &= \ (s_1, \cdots, s_{n-1}) \mapsto \sigma = (\sigma(1), \cdots, \sigma(n)), \\ &\text{where } \{\sigma(i)\} = s_i - s_{i-1}, \qquad i = 1, \cdots n, \\ s_0 \ &\triangleq \phi \text{ and } s_n \ \triangleq \{1, \cdots, n\}. \end{aligned}$$

Thus, the Bruhat cells of Flag (F^n) are parametrized by $\sum(n)$. In fact, under the identification Flag $(F^n) = GL(n,F)/U(F^n)$ a Bruhat cell $B_\sigma, \sigma \in \sum(n)$, is given by

$$B_\sigma = \{< X > \in GL(n,F)/U(F^n) \mid X = u\sigma, \text{ some } u \in U(F^n)\}.$$

To see this, let us consider an element of the form $u\sigma$, with $u \in U(F^n)$ and $\sigma \in \sum(n)$. By performing column operations, first on the first column, next on the first two columns and so forth, we can transform it into an $n \times n$ matrix that has ones in the positions prescribed by σ and zeros on the right and below each such one. This operation corresponds to right multiplication of $u\sigma$ by an element in $U(F^n)$. Thus, in the left coset $< u\sigma >$ there is a representative of the form described above. From this, one readily sees that the flag corresponding to such a coset has signature σ. Also one sees that the dimension of B_σ is given by the length of the permutation σ, namely by

$$\ell(\sigma) = \text{ card } \{(i,j) \mid 1 \le i < j \le n \ \& \ \sigma(i) > \sigma(j)\}.$$

We remark that the Bruhat decomposition can be carried out in the broader context of the so-called generalized flag manifolds G/B, where G is a reductive Lie group and B a Borel subgroup (see e.g. [3]). In fact, J. Tits [13] has shown that the Bruhat decomposition is a formal consequence of the axioms for a BN-pair structure on a group. This decomposition is an immediate consequence of the (Bruhat) decomposition of G into the disjoint union of double cosets BwB, where w is an element in the Weyl group W of G.

III. Hessenberg Flags

Let F denote the field of real or complex numbers, and let $A \in g\ell(n,F)$ be a linear operator. Given a d-tuple $K = (k_1, \cdots, k_d)$ of positive integers satisfying the condition $0 < k_1 < \cdots < k_d < k_{d+1} = n$, we form the product

$$M = \prod_{i=1}^{d} G(k_i, F^n)$$

of the Grassmannians $G(k_i, F^n)$ of k_i-dimensional subspaces of F^n, and we define for $p = 0, 1, ..., d-1$

$$M(p,K,A) = \{(S_1, \cdots, S_d) \in M \mid AS_i \subset S_{i+p}, i = 1, \cdots, d\}$$

where we set conventionally $S_\ell = F^n$ if $\ell \geq d+1$. Also, with the above convention understood, it is natural to set

$$M(d, K, A) = M.$$

We observe immediately that if $A = I$, the identity operator, then

$$M(1, K, I) \;=\; \text{Flag } (K, F^n).$$

Also, for a fixed "signature" K, we have inclusions:

$$M(0, K, A) \subset M(1, K, A) \subset \cdots \subset M(d-1, K, A) \subset M(d, K, A) = M.$$

(III.1) Definition. Given $A \in g\ell(n, F), K = (k_1, \cdots, k_d)$ and $p \in \{0, 1, \cdots, d-1\}$ as above, we define the Hessenberg flags of degree p as the elements of

$$\text{Hess } (p, K, A) = M(p, K, A) \cap M(1, K, I).$$

In other words, a Hessenberg flag (of degree p) is a flag $S_1 \subset \cdots \subset S_d$ such that $AS_i \subset S_{i+p}$. Again we observe that we have inclusions:

$$\text{Hess } (0, K, A) \subset \text{ Hess } (1, K, A) \subset \cdots \subset \text{ Hess}(d-1, K, A) \subset \text{ Hess } (d, K, A)$$

where Hess $(d, K, A) =$ Flag (K, F^n).

Remark. It is immediate to see that for $A \in g\ell(n, F)$ and $\lambda \in F$ we have

$$\text{Hess } (p, K, A) = \text{ Hess } (p, K, A - \lambda I), \qquad\qquad p = 0, \cdots, d-1.$$

Thus, we can assume A to be nonsingular, i.e. $A \in GL(n, F)$. In particular we observe that for $A \in GL(n, F)$, Hess $(0, K, A)$ consists of those flags that are fixed by A.

For $p = 1, \cdots, d-1$, the subset $M(p, K, I)$ of M can be identified with a product of flag manifolds as follows: For $i = 1, \cdots, p$ define

$$K(i) = (k_i, k_{i+p}, \cdots, k_{i+r(i)p})$$

where $r(i) \triangleq \max\{r \mid r \text{ a positive integer with } i + rp \leq d\}$. Then

$$M(p, K, I) \simeq \prod_{i=1}^{p} \text{ Flag } (K(i), F^n), \qquad\qquad p = 1, \cdots, d-1$$

under the mapping

$$(S_1,\cdots,S_d) \to ((S_1,S_{1+p},...,S_{1+r(1)p}),(S_2,S_{2+p},\cdots,S_{2+r(2)p}),$$

$$\cdots,(S_p,S_{p+p},\cdots,S_{p+r(p)p})).$$

Let now $A \in GL(n,F)$ and $1 \le p \le d-1$. We define a map $\hat{A}_p : M(p,K,I) \to M(p,K,A)$ by setting

$$\hat{A}_p((S_1,\cdots,S_d)) = (S_1,\cdots,S_p,AS_{p+1},\cdots,AS_{2p},A^2S_{2p+1},\cdots$$

$$\cdots,A^2S_{3p},\cdots,A^{r(1)}S_{r(1)p+1},\cdots,A^{r(1)}S_d).$$

Clearly $\hat{A}_p$ is well-defined and it admits an inverse $(\hat{A}_p)^{-1} : M(p,K,A) \to M(p,K,I)$ given by $(\hat{A}_p)^{-1} = (\widehat{A^{-1}})_p$. Therefore we have that

$$M(p,K,A) \simeq M(p,K,I), \qquad p = 1,\cdots,d-1.$$

The above isomorphism enables us to identify $M(p,K,A)$ with an algebraic subset of M isomorphic to a product of flag manifolds, for $p = 1,\cdots,d-1$. Thus, since Hess $(0,K,A)$ consists of the flags fixed by A, we have the following:

(III.2) Proposition: For $A \in GL(n,F)$ and $p = 0,\cdots,d-1$, the set Hess (p,K,A) is a projective algebraic variety.

From now on we will focus our attention on more particular cases, as stated in the following assumptions:

(A1) The flags we consider are the complete flags, that is $K = K_0 = (1,2,\cdots,n-1)$.

(A2) The operator A is assumed diagonalizable and with spectrum in F; so we choose once and for all a basis of F^n on which A is diagonal.

In view of assumption (A1), we will simplify the notation by writing Hess (p,K_0,A) = Hess (p,A). It is clear that in the case $F = \mathbb{R}$, assumption (A2) is not generic, so that most of the following results are of greater interest in the case $F = \mathbb{C}$. On the other hand we remark that if $Q \in GL(n,F)$, the analytic isomorphism $\psi((S_1,...,S_{n-1})) = (Q^{-1}S_1,\cdots,Q^{-1}S_{n-1})$, carries Hess (p,A) one-to-one and onto Hess $(p,Q^{-1}AQ)$ so that there is no further loss of generality in fixing a basis and assuming A to be diagonal.

We now proceed to a description of local algebraic equations for Hess (p, A). First of all we realize the flag manifold Flag (F^n) as the homogeneous space $GL(n,F)/U(F^n)$, as described in section II. By its very definition, Hess (p, A) is contained in Flag (F^n). Hence, if $< h > = < \sigma X >$, with $\sigma \in \sum(n)$ and $X \in L^+(F))$, we have that $< h >$ is a Hessenberg flag (of degree p), if and only if $A\sigma X = \sigma X R$, i.e.

$$(\sigma^{-1}A\sigma)X = XR \tag{3}$$

where R is an $n \times n$ matrix in the Hessenberg form

$$R = \begin{bmatrix} r_{11} & \cdots & \cdots & \cdots & \cdots & r_{1n} \\ \vdots & & & & & \vdots \\ r_{p+11} & & & & & \vdots \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ 0 & \cdots & 0 & r_{nn-p} & \cdots & r_{nn} \end{bmatrix}$$

From (3) we deduce algebraic conditions on the entries of X. If $A = diag(\lambda_1, \cdots, \lambda_n)$, we have

$$(\sigma^{-1}A\sigma)X = \begin{bmatrix} \lambda_{\sigma(1)} & 0 & \cdots & \cdots & \cdots & 0 \\ \lambda_{\sigma(2)}x_{21} & \lambda_{\sigma(2)} & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ \lambda_{\sigma(n)}x_{n1} & \cdots & \cdots & \cdots & \lambda_{\sigma(n)}x_{nn-1} & \lambda_{\sigma(n)} \end{bmatrix}$$

so that from (3) we easily deduce that

$$\begin{cases} r_{ij} = 0 & \text{if } j > i \\ r_{ii} = \lambda_{\sigma(i)}, & i = 1, \cdots, n. \end{cases}$$

Hence we may rewrite the matrix equation (3) explicitly as:

$$\begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ x_{21} & 1 & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ x_{n1} & \cdots & \cdots & \cdots & x_{nn-1} & 1 \end{bmatrix} \begin{bmatrix} \lambda_{\sigma(1)} & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ r_{21} & \lambda_{\sigma(2)} & \ddots & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & \vdots \\ r_{p+11} & & \ddots & \ddots & \ddots & & \vdots \\ 0 & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & r_{nn-p} & \cdots & r_{nn-1} & \lambda_{\sigma(n)} \end{bmatrix}$$

$$
(4) \qquad = \begin{bmatrix}
\lambda_{\sigma(1)} & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
\lambda_{\sigma(2)}x_{21} & \lambda_{\sigma(2)} & \ddots & & & & \vdots \\
\vdots & \ddots & \ddots & \ddots & & & \vdots \\
\vdots & & \ddots & \ddots & \ddots & & 0 \\
\lambda_{\sigma(n)}x_{n1} & \cdots & \cdots & \cdots & \cdots & \lambda_{\sigma(n)}x_{nn-1} & \lambda_{\sigma(n)}
\end{bmatrix}
$$

We shall regard (4) as n systems of n linear equations as follows:

$$X \cdot (\beta^{th} \text{ column of } R) = (\beta^{th} \text{ column of } (\sigma^{-1}A\sigma)X)$$

for $\beta = 1, \cdots, n$. We observe that fixing the β^{th} column of $(\sigma^{-1}A\sigma)X$, say ${}^t(0, \cdots, 0, \lambda_{\sigma(\beta)}, \lambda_{\sigma(\beta+1)}x_{\beta+1\beta}, \cdots, \lambda_{\sigma(n)}x_{n\beta})$, we automatically have for the solution vector - i.e., the β^{th} column of R - that its first β entries are $(0, \cdots, 0, \lambda_{\sigma(\beta)})$. By using Cramer's rule, we can obtain explicit expressions for the remaining r_{ij}'s, whereas the solutions corresponding to the zeros in the left bottom corner of R give us the desired algebraic equations. Thus we will obtain equations of the type:

$$f_{\alpha\beta} = 0, \qquad \alpha - \beta > p$$

where $f_{\alpha\beta}$ is a polynomial in the x_{ij}'s, $1 \leq j < i \leq n$.

More precisely, since $\det X = 1$, Cramer's rule applied to the β^{th} system gives

$$
f_{\alpha\beta} = \det \left[\begin{array}{ccccc|c|ccccc}
1 & 0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\
x_{21} & 1 & \ddots & & \vdots & \vdots & \vdots & & & & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots & \lambda_{\sigma(\beta)} & \vdots & & & & \vdots \\
\vdots & & \ddots & \ddots & 0 & \vdots & \vdots & & & & \vdots \\
\vdots & & & \ddots & 1 & \vdots & \vdots & & & & \vdots \\
x_{\alpha 1} & & & & x_{\alpha\ \alpha-1} & \lambda_{\sigma(\alpha)}x_{\alpha\beta} & 0 & & & & \vdots \\
\vdots & & & & \vdots & \vdots & 1 & \ddots & & & \vdots \\
\vdots & & & & \vdots & \vdots & x_{\alpha+2\ \alpha+1} & \ddots & \ddots & & \vdots \\
\vdots & & & & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\
x_{n1} & \cdots & \cdots & \cdots & x_{n\alpha-1} & \lambda_{\sigma(n)}x_{n\beta} & x_{n\ \alpha+1} & \cdots & \cdots & x_{n\ n-1} & 1
\end{array}\right]
$$

which we rewrite as $f_{\alpha\beta} = \det\theta_{\alpha\beta}$. We now compute this determinant by performing column operations. We subtract from the α^{th} column of $\theta_{\alpha\beta}$ one of the first $(\alpha - 1)$ columns suitably multiplied by a constant, in order to put $\theta_{\alpha\beta}$ in upper-triangular form. Its determinant will then be the value of the α^{th} entry of the transformed column. Explicitly we have:

(III.3) Proposition. For $\alpha - \beta > p$ we have:

$$f_{\alpha\beta} = (\lambda_{\sigma(\alpha)} - \lambda_{\beta)})x_{\alpha\beta} +$$

$$+ \sum_{t=1}^{\alpha-\beta-1} \sum_{\alpha > \gamma_1 > \cdots > \gamma_t > \beta} (-1)^t (\lambda_{\sigma(\gamma t)} - \lambda_{\sigma(\beta)}) x_{\alpha\gamma_1} x_{\gamma_1\gamma_2} \cdot \cdots \cdot x_{\gamma_t\beta}.$$

Proof: We have already described the k th step of the chosen algorithm, so we prove the result by induction. To simplify the formulas we put

$$\begin{cases} \Lambda(\gamma, \beta) = \lambda_{\sigma(\gamma)} - \lambda_{\sigma(\beta)} \\ X(\alpha, \gamma_1, \ldots, \gamma_t, \beta) = x_{\alpha\gamma_1} x_{\gamma_1\gamma_2} \quad \cdot \cdots \cdot x_{\gamma_{t-1}\gamma_t} x_{\gamma_t\beta} \end{cases}$$

and we rewrite the statement:

$$f_{\alpha\beta} = \Lambda(\alpha, \beta) x_{\alpha\beta} + \sum_{t=1}^{\alpha-\beta-1} \sum_{\alpha > \gamma_1 > \cdots > \gamma_t > \beta} (-1)^t \Lambda(\gamma_t, \beta) X(\alpha, \gamma_1, \cdots, \gamma_t, \beta)$$

The first step ($k = 0$) consists of subtracting from the α^{th} column of $\theta_{\alpha\beta}$ its β^{th} column, multiplied by $\lambda_{\sigma(\beta)}$, so that the transformed column, say v_0, is

$$\begin{aligned} v_0 &=^t (0, \cdots, 0, (\lambda_{\sigma(\beta+1)} - \lambda_{\sigma(\beta)}) x_{\beta+1\beta}, \ldots, (\lambda_{\sigma(n)} - \lambda_{\sigma(\beta)}) x_{n\beta}) \\ &=^t (0, \cdots, 0, \Lambda(\beta+1, \beta) x_{\beta+1,\beta}, \cdots, \Lambda(n, \beta) x_{n\beta}). \end{aligned}$$

Suppose inductively that after k transformations ($k < \alpha - \beta$) the obtained column v_{k-1} has entries $(v_{k-1})_\omega$ given by

$$(v_{k-1})_\omega = \begin{cases} \Lambda(\omega, \beta) x_{\omega\beta} + \sum_{t=1}^{k-1} \sum_{\beta+k > \gamma_1 > \cdots \gamma_t > \beta} (-1)^t \Lambda(\gamma_t, \beta) X(\alpha, \gamma_1, \cdots, \gamma_t) & \text{if } \omega \geq \beta + k \\ 0 & \text{if } \omega < \beta + k \end{cases}$$

We now subtract from v_{k-1} the $(\beta+k)^{th}$ column of $\theta_{\alpha\beta}$ multiplied by $(v_{k-1})_{\beta+k}$. Hence for $\omega \geq \beta+k+1$ we have:

$$
\begin{aligned}
(v_k)_\omega &= \Lambda(\omega,\beta) + \sum_{t=1}^{k-1} \sum_{\beta+k>\gamma_1>\cdots>\gamma_t>\beta} (-1)^t \Lambda(\gamma_t,\beta) X(\omega,\gamma_1,\cdots,\gamma_t,\beta) \\
&\quad - \Lambda(\beta+k,\beta) X(\omega,\beta+k,\beta) \\
&\quad - \sum_{t=1}^{k-1} \sum_{\beta+k>\gamma_1>\cdots>\gamma_t>\beta} (-1)^t \Lambda(\gamma_t,\beta) X(\omega,\beta+k,\gamma_1,\cdots,\gamma_t,\beta) \\
&= \Lambda(\omega,\beta) x_{\omega\beta} + \sum_{t=1}^{k} \sum_{\beta+k+1>\cdots>\gamma_t>\beta} (-1)^t \Lambda(\gamma_t,\beta) X(\omega,\gamma_1,\cdots,\gamma_t,\beta)
\end{aligned}
$$

and $(v_k)_\omega = 0$ for $\omega < \beta+k+1$. This proves the induction argument. So after $(\alpha-\beta)$ transformations we get $v_{\alpha-\beta-1}$ with

$$
(v_{\alpha-\beta-1})_\omega = \Lambda(\omega,\beta) x_{\omega\beta} + \sum_{t=1}^{\alpha-\beta-1} \sum_{\alpha>\gamma_1>\cdots>\gamma_t>\beta} (-1)^t \Lambda(\gamma_t,\beta) X(\omega,\gamma_1,\cdots,\gamma_t,\beta)
$$

for $\omega \geq \alpha$. Setting $\omega = \alpha$ gives the asserted result.

QED

We recall that for each chart $ch(\sigma), \sigma \in \sum(n)$, of Flag (F^n), the corresponding local equations defining Hess (p,A) have indices (α,β) equal to the indices of the zeros of a matrix in Hessenberg form (of degree p), i.e., $\alpha-\beta > p$.

(III.4) Theorem: If $A \in GL(n,F)$ has distinct eigenvalues, all of which are in F, then Hess (p,A) is a smooth submanifold of Flag (F^n) of dimension $p(2n-p-1)/2$ (as a real manifold if $F = \mathbb{R}$ and as a complex manifold if $F = \mathbb{C}$).

Proof: We shall use the Jacobian criterion to show that Hess (p,A) is smooth. Fix a chart $ch(\sigma)$ of Flag (F^n) and consider the Jacobian

$$
(J_{(\alpha,\beta)(i,j)}) = \left(\frac{\partial f_{\alpha\beta}}{\partial x_{ij}}\right), \qquad \alpha-\beta > p, \qquad 1 \leq j < i \leq n.
$$

We order the row indices (α,β) and the column indices (i,j) as follows:

$$
(\alpha,\beta) < (\gamma,\delta) \Leftrightarrow \text{ either } \alpha-\beta < \gamma-\delta \text{ or } \alpha-\beta = \gamma-\delta \ \&\ \beta < \delta.
$$

Next we look at the square submatrix of $(J_{(\alpha,\beta)(i,j)})$ obtained by choosing its last $(n-p-1)(n-p)/2$ columns, namely those with indices (i,j) that satisfy $i-j>p$. We claim that:

i) $J_{(\alpha,\beta)(\mu,\nu)} = 0$ if $\alpha-\beta<\mu-\nu$

ii) For $r=p+1,\cdots,n-1$, the square matrix $(J(r)_{(\alpha,\beta)(i,j)}) \triangleq (J_{(\alpha,\beta)(i,j)})_{\alpha-\beta=i-j=r}$ is diagonal with diagonal entries $\Lambda(r+1,1),\Lambda(r+2,2),\cdots,\Lambda(n,n-r)$.

If i) and ii) hold, then the chosen square submatrix of $(J_{(\alpha,\beta)(i,j)})$ is lower-triangular as a matrix of blocks, when the rows and columns are partitioned according to the partition $(n-p-1,n-p-2,\cdots,1)$ of $(n-p-1)(n-p)/2$. In particular, since its diagonal blocks are themselves diagonal, its determinant is equal to

$$\prod_{\substack{p+1<\alpha\leq n\\ \alpha-\beta>p}} \Lambda(\alpha,\beta) = \prod_{\substack{p+1<\alpha\leq n\\ \alpha-\beta>p}} (\lambda_{\sigma(\alpha)}-\lambda_{\sigma(\beta)}) \neq 0$$

independently of the point in $ch(\sigma)$. Thus the mapping defined by $\{f_{\alpha\beta}\}_{\alpha\beta}$ on $ch(\sigma)$ is a submersion and its fiber $\{f_{\alpha\beta}^{-1}(\{0\})\}_{\alpha\beta}$ is an embedded submanifold of $ch(\sigma)$. Moreover the codimension of Hess $(p,A)\cap ch(\sigma) = \{f_{\alpha\beta}^{-1}(\{0\})\}_{\alpha\beta}$ in $ch(\sigma)$ hence in Flag (F^n), is equal to the rank of the mapping. In other words the dimension (over F) of Hess (p,A) is equal to the number of nonzero subdiagonal entries of a matrix in Hessenberg form (of degree p), namely

$$(n-1)+\cdots+(n-p) = p(2n-p-1)/2.$$

Let us now prove i) and ii). We have

$$f_{\alpha\beta} = \Lambda(\alpha,\beta)x_{\alpha\beta} + \sum_{t=1}^{\alpha-\beta-1} \sum_{\alpha>\gamma_1>\cdots,\gamma_t,\beta} (-1)^t \Lambda(\gamma_t,\beta) X(\alpha,\gamma_1,\cdots,\gamma_t,\beta)$$

so that each factor $x_{\gamma\delta}$ appearing in

(5) $$X(\alpha,\gamma_1,\cdots,\beta) = x_{\alpha\gamma_1}x_{\gamma_1,\gamma_2}\cdot\cdots\cdot x_{\gamma_t\beta}, \qquad \alpha>\gamma_1>\cdots>\gamma_t>\beta$$

is such that $\gamma-\delta\leq\alpha-\beta$. Therefore if (μ,ν) is a pair of indices with $\alpha-\beta<\mu-\nu$, we have $\frac{\partial f_{\alpha\beta}}{\partial x_{\mu\nu}}=0$, proving i). Suppose now that (μ,ν) is such that $\alpha-\beta=\mu-\nu$. If $\nu<\beta$, none of the factors in (5) is equal to $x_{\mu\nu}$ and if $\nu>\beta$ then $\mu>\alpha$ and again none of the factors in (5) is equal to $x_{\mu\nu}$. Thus, if $\mu - \nu = \alpha - \beta$ then $\frac{\partial f_{\alpha\beta}}{\partial x_{\mu\nu}} = 0$, unless $(\alpha,\beta) = (\mu,\nu)$, in which case $\frac{\partial f_{\alpha\beta}}{\partial x_{\mu\nu}} = \Lambda(\alpha,\beta)$. This proves ii).

QED

Example. Let $n = 5$ and $p = 2$. For a fixed $\sigma \in \sum(5)$ we have

$$f_{41} = \Lambda(4,1)x_{41} - \{\Lambda(3,1)x_{43}x_{31} + \Lambda(2,1)x_{42}x_{21}\} + \Lambda(2,1)x_{43}x_{32}x_{21}$$
$$f_{52} = \Lambda(5,2)x_{52} - \{\Lambda(4,2)x_{54}x_{42} + \Lambda(3,2)x_{53}x_{32}\} + \Lambda(3,2)x_{54}x_{43}x_{32}$$

$$\begin{aligned} f_{51} = \Lambda(5,1)x_{51} &- \{\Lambda(4,1)x_{54}x_{41} + \Lambda(3,1)x_{53}x_{31} + \Lambda(2,1)x_{52}x_{21}\} \\ &+ \{\Lambda(3,1)x_{54}x_{43}x_{31} + \Lambda(2,1)x_{54}x_{42}x_{21} + \Lambda(2,1)x_{53}x_{32}x_{21}\} \\ &- \Lambda(2,1)x_{54}x_{43}x_{32}x_{21}. \end{aligned}$$

The 3×3, full rank, submatrix of the Jacobian is then

$$\begin{bmatrix} \Lambda(4,1) & 0 & \vline & 0 \\ 0 & \Lambda(5,2) & \vline & 0 \\ -\Lambda(4,1)x_{54} & -\Lambda(2,1)x_{21} & \vline & \Lambda(5,1) \end{bmatrix}$$
$$(4,1) \qquad (5,2) \qquad (5,1)$$

IV. Connectedness

In this section we prove the connectedness of Hess $(p, A), p \geq 1$, in the case in which A has distinct eigenvalues, all of which are in F. In fact we can prove a somewhat stronger connectedness result which implies the previous one. In order to accomplish this, we need to consider the Riccati flow induced by a linear operator $B \in GL(n, F)$ on Flag (F^n) namely the flow

$$\phi_B(t, (S_1, \cdots, S_{n-1})) = (e^{Bt}(S_1), \cdots, e^{Bt}(S_{n-1})), \qquad t \in \mathbb{R}$$

whose features have been investigated in [1][10][14][15]. In particular we will use the fact that if B has eigenvalues with distinct real parts, then the set of the flags fixed by B is the positive limit set for the differential equation associated with the flow ϕ_B.

(IV.1) Lemma: Let $A \in GL(n,F)$ have distinct eigenvalues all of which are in F, and let h_1 and h_2 be two flags fixed by A. Then h_1 and h_2 can be joined by a path in Hess $(1,A)$.

Proof: There exist two permutation matrices $\sigma, \tau \in \sum(n)$ such that $h_1 =<\sigma>$ and $h_2 =<\tau>$, where $<\cdot>$ denotes as usual the coset modulo $U(F^n)$. Let now Y^0 denote the Vandermonde matrix

$$Y^0 = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-1} \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \cdots & \lambda_n^{n-1} \end{bmatrix}$$

where $\lambda_1, \cdots, \lambda_n$ are the eigenvalues of A. Clearly, $<Y^0> \in$ Hess $(1,A)$ and the nonsingular matrix

$$Y_\sigma^0(u) = \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ \vdots & \vdots & & \vdots \\ u & u\lambda_{\sigma(n)} & \cdots & u\lambda_{\sigma(n)}^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{bmatrix}$$

is such that $<Y_\sigma^0(u)> \in$ Hess $(1,A)$, for $u \in (0,1]$. If we denote by e_k the k^{th} standard column vector, and we put $\alpha =^t (\alpha_1, \cdots, \alpha_n)$, the linear system

$$Y^0 \cdot \alpha = e_{\sigma(n)}$$

has a solution α, since $Y^0 \in GL(n,F)$. Moreover, by Cramer's rule

$$\alpha_n = \frac{1}{det Y^0} \cdot det \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-2} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & \lambda_{\sigma(n)} & \cdots & \lambda_{\sigma(n)}^{n-2} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-2} & 0 \end{bmatrix} \neq 0.$$

Thus, the matrix

$$B(u) = \left[\begin{array}{c|c} I_{n-1} & \begin{matrix} \alpha_1/u \\ \vdots \\ \vdots \end{matrix} \\ \hline 0 \cdots 0 & \alpha_n/u \end{array}\right]$$

belongs to $U(F^n)$ for $u \in (0,1]$, and we have

$$< Y_\sigma^0(u)B(u) > = < Y_\sigma^0(u) > \in \text{ Hess } (1,A), \qquad u \in (0,1].$$

On the other hand,

$$Y_\sigma^0(u)B(u) = \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-2} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ u\lambda_{\sigma(n)} & \cdots & & u\lambda_{\sigma(n)}^{n-2} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-2} & 0 \end{bmatrix}$$

for $u \in (0,1]$, and so we see that

$$Y_\sigma^1 \triangleq \lim_{u \to 0} Y_\sigma^0(u)B(u)$$

$$= \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-2} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-2} & 0 \end{bmatrix}$$

Thus if we set

$$\gamma^1(u) = \begin{cases} < Y_\sigma^0(u) > & \text{if } u \in (0,1] \\ < Y_\sigma^1 > & \text{if } u = 0 \end{cases}$$

we have that $\gamma^1(u), u \in [0,1]$, is a continuous path in Hess $(1,A)$ with $\gamma^1(0) = < Y_\sigma^1 >$ and $\gamma^1(1) = < Y^0 >$. Repeating this algorithm n times yields n paths $\gamma^i(u)$ in Hess $(1,A), i = 1, \cdots, n$, joining $< Y_\sigma^{i-1} >$ to $< Y_\sigma^i >$, where

$$Y_\sigma^i = \left[\begin{array}{cccc|ccc} 1 & \lambda_1 & \cdots & \lambda_1^{n-i-1} & & & \\ \vdots & \vdots & & \vdots & & & \\ \vdots & \vdots & & \vdots & e_{\sigma(n-i+1)} & \cdots & e_{\sigma(n)} \\ \vdots & \vdots & & \vdots & & & \\ 1 & \lambda_n & \cdots & \lambda_n & & & \end{array}\right]$$

(Rows $\sigma(n-i+1), \cdots, \sigma(n)$ have zero entries at column indices $1, \cdots, n-i$.) Since $Y_\sigma^n = \sigma$, we obtain a continuous path in Hess $(1,A)$ joining $h_1 = < \sigma >$ to $< Y^0 >$. Similarly we can join $h_2 = < \tau >$ to $< Y^0 >$, and the proof is complete. QED

(IV.2) Theorem. Let $A \in GL(n,F)$ have distinct eigenvalues, all of which are in F, and let $B \in GL(n,F)$ have eigenvalues with distinct real parts. Assume that A and B commute. Then any subset of Flag (F^n) which contains Hess $(1,A)$ and is positively invariant under the Riccati flow ϕ_B induced by B is connected. In particular Hess $(p,A), p \geq 1$, is connected.

Proof: Let N be a subset of Flag (F^n) containing Hess $(1,A)$ and positively invariant under the Riccati flow ϕ_B induced by B. Since the positive limit set of ϕ_B consists of those flags that are fixed by B, i.e., Hess $(0,B)$, we can join any point in N to at least one point in Hess $(0,B)$ by a continuous path in N. On the other hand, Hess $(0,B) =$ Hess $(0,A)$ and by (IV.1) any two points in Hess $(0,A)$ can be joined by a continuous path in Hess $(1,A) \subset N$. Therefore N is connected.

Finally, since A and B commute, we have that for all $t \in \mathbb{R}, Ae^{Bt} = e^{Bt}A$. Thus, if $(S_1, \cdots, S_{n-1}) \in$ Hess $(p,A), p \geq 1$, then

$$A(e^{Bt}S_i) = e^{Bt}(AS_i) \subset e^{Bt}S_{i+p}, \qquad i = 1, \cdots, n-p.$$

Thus Hess $(p,A), p \geq 1$, is ϕ_B-invariant and the proof is complete. QED

Remark: We now give a second proof of the connectedness of Hess $(p,A), p \geq 1$, which does not depend on the properties of the Riccati flow. By virtue of (IV.1), all we have to show is that each point $h \in$ Hess (VA) can be joined by a path in Hess (p,A) to a flag fixed by A.

Let $h \in ch(\sigma) \bigcap$ Hess $(p,A), \sigma \in \sum(n)$, and let $(x_{ij})_{1 \leq j < i \leq n}$ be the σ-coordinates of h in Flag (F^n). They satisfy

$$\Lambda(\alpha,\beta)x_{\alpha\beta} + \sum_{t=1}^{\alpha-\beta-1} \sum_{\alpha > \gamma_1 > \cdots > \gamma_t > \beta} (-1)^t \Lambda(\gamma_t, \beta) x_{\alpha\gamma_1} \cdot \cdots \cdot x_{\gamma_t\beta} = 0$$

for $\alpha - \beta > p$. Put now

$$y_{ij}(u) = u^{i-j}x_{ij}, \qquad 1 \leq j < i \leq n, \qquad u \in [0,1],$$

so that $(y_{ij}(0))_{1\le j<i\le n}$ are the σ-coordinates of the fixed flag $<\sigma>$. On the other hand,

$$\Lambda(\alpha,\beta)y_{\alpha\beta}(u)+\sum_{t=1}^{\alpha-\beta-1}\sum_{\alpha>\gamma_1>\cdots>\gamma_t>\beta}(-1)^t\Lambda(\gamma_t,\beta)y_{\alpha\gamma_1}(u)\cdot\cdots\cdot y_{\gamma_t\beta}(u)$$

$$=u^{\alpha-\beta}\Lambda(\alpha,\beta)x_{\alpha\beta}+\sum_{t=1}^{\alpha-\beta-1}\sum_{\alpha>\gamma_1>\cdots>\gamma_t>\beta}(-1)^t$$

$$\Lambda(\gamma_t,\beta)u^{(\alpha-\gamma_1)+\cdots+(\gamma_t-\beta)}x_{\alpha\gamma_1}\cdot\cdots\cdot x_{\gamma_t\beta}$$

$$=u^{\alpha-\beta}\{\Lambda(\alpha,\beta)x_{\alpha\beta}+\sum_{t=1}^{\alpha-\beta-1}\sum_{\alpha>\gamma_1>\cdots>\gamma_t>\beta}(-1)^t\Lambda(\gamma_t,\beta)x_{\alpha\gamma_1}\cdot\cdots\cdot x_{\gamma_t\beta}\}$$

$$=0.$$

Thus, $(y_{ij}(u))_{1\le j<i\le n}$ are the σ-coordinates of a point in Hess (p,A) for all $u\in[0,1]$. This shows that we can join a point $h\in ch(\sigma)\bigcap$ Hess (p,A) to $<\sigma>$ with a continuous path in Hess (p,A).

V. Betti numbers

In this section we compute the Betti numbers of Hess (p,A) with $A\in GL(n,\mathbb{C})$, an operator with distinct eigenvalues. They turn out to be generalized Eulerian numbers, in a sense that will be clearly explained at the end of this section.

The strategy is to decompose Hess (p,A) into a disjoint union of subvarieties, namely those obtained by intersecting the Bruhat cells of Flag (F^n)-relative to a basis in which A is diagonal - with Hess (p,A). These subvarieties are diffeomorphic to affine spaces $\mathbb{C}^q$, where q is in fact a quantity attached to the element $\sigma\in\sum(n)$ that indexes the corresponding Bruhat cell. (Precisely, q is the pth "Eulerian dimension" of σ.) The obtained decomposition is not cellular, but we are able to derive the Betti numbers of Hess (p,A) by means of a theorem of A.H. Durfee that gives the Betti numbers for a smooth complex projective variety that can be written as a disjoint union of smooth contractible quasiprojective subvarieties [16].

(V.1) Proposition Let $A\in GL(n,\mathbb{C})$ have distinct eigenvalues and let B_σ be a Bruhat cell of Flag $(\mathbb{C}^n)$ relative to a basis in which A is diagonal. Then $B_\sigma\bigcap$ Hess (p,A) is a subvariety of Hess (p,A) which is analytically isomorphic to a $\mathbb{C}$-affine space of complex dimension

$$E_p(\sigma)=\mathrm{card}\{(i,j)\mid 1\le i,j\le n,\ 1\le i-j\le p,\ \sigma(i)<\sigma(j)\}$$

Proof: First of all we observe that $B_\sigma \subset ch(\sigma)$ and, more precisely, B_σ is the slice of $ch(\sigma)$ obtained by setting equal to zero all the σ-coordinates X_{ij} for which $\sigma(i) > \sigma(j)$. Let us now consider the restrictions $\{\hat{f}_{\alpha\beta}\}_{\alpha-\beta>p}$ of the mappings $\{f_{\alpha\beta}\}_{\alpha-\beta>p}$ to B_σ. We claim that

$$\hat{f}_{\alpha\beta} \text{ vanishes identically } \Leftrightarrow \sigma(\alpha) > \sigma(\beta).$$

Suppose $\sigma(\alpha) > \sigma(\beta)$. Then $x_{\alpha\beta} = 0$ on B_σ. Moreover, suppose that one of the monomials $X(\alpha, \gamma_1, \cdots, \gamma_t, \beta), (\alpha > \gamma_1 > \cdots > \gamma_t > \beta)$ appearing in $f_{\alpha\beta}$ is not zero at some point in B_σ. Then necessarily $\sigma(\alpha) < \sigma(\gamma_1) < \cdots < \sigma(\gamma_t) < \sigma(\beta)$, a contradiction.

Conversely, assume that $\hat{f}_{\alpha\beta}$ vanishes identically and suppose that $x_{\alpha\beta} \neq 0$ at some point $b \in B_\sigma$, i.e. suppose that the $(\alpha, \beta)^{th}$ coordinate of b has value $b_{\alpha\beta} \neq 0$. Then necessarily $\sigma(\alpha) < \sigma(\beta)$. Hence the point c whose σ-coordinates are all zero except the $(\alpha, \beta)^{th}$ which is $b_{\alpha\beta}$ belongs to B_σ. But $\hat{f}_{\alpha\beta}(c) = b_{\alpha\beta}\Lambda(\alpha, \beta) \neq 0$, a contradiction. Therefore $x_{\alpha\beta}$ is identically zero on B_σ, which implies $\sigma(\alpha) > \sigma(\beta)$.

Next we claim that the equations

$$\hat{f}_{\alpha\beta} = 0, \ \alpha - \beta > p, \ \sigma(\alpha) < \sigma(\beta)$$

determine all the variables $x_{\alpha\beta}$ for which $\alpha - \beta > p$ and $\sigma(\alpha) < \sigma(\beta)$ as polynomial functions of the variables x_{ij} for which $i - j \leq p$ and $\sigma(i) < \sigma(j)$.

Let (α, β) be a pair of indices such that $\alpha - \beta = p + 1$ and $\sigma(\alpha) < \sigma(\beta)$. Equation $\hat{f}_{\alpha\beta} = 0$ implies that

$$x_{\alpha\beta} = \frac{-1}{\Lambda(\alpha, \beta)} \sum_{t=1}^{\alpha-\beta-1} \sum_{\alpha > \gamma_1 > \cdots > \gamma_t > \beta} (-1)^t \Lambda(\gamma_t, \beta) x_{\alpha\gamma_1} \cdot \cdots \cdot x_{\gamma_t\beta}. \tag{6}$$

All the factors $x_{\gamma\delta}$ appearing in the right hand side of (6) are such that $\gamma - \delta < \alpha - \beta = p + 1$, and vanish identically if $\sigma(\gamma) > \sigma(\delta)$. Therefore $x_{\alpha\beta}$ is a polynomial function of the variables x_{ij} for which $i - j \leq p$ and $\sigma(i) < \sigma(j)$.

Suppose inductively that for $q \geq p + 1$, all the variables $x_{\alpha\beta}$ for which $q \geq \alpha - \beta > p$ and $\sigma(\alpha) < \sigma(\beta)$ are polynomial functions of the variables x_{ij} for which $i - j \leq p$ and $\sigma(i) < \sigma(j)$. Let (α, β) be a pair of indices such that $\alpha - \beta = q + 1$ and $\sigma(\alpha) < \sigma(\beta)$. Again $\hat{f}_{\alpha\beta} = 0$ implies

(6). All the factors $x_{\gamma\delta}$ appearing in the right hand side of (6) are such that $\gamma-\delta < \alpha-\beta = q+1$ and vanish identically if $\sigma(\gamma) > \sigma(\delta)$. Since $\gamma-\delta \leq q$, we have by induction hypothesis that, in all the cases in which $\sigma(\gamma) < \sigma(\delta), x_{\gamma\delta}$ is a polynomial function of the variables x_{ij} for which $i-j \leq p$ and $\sigma(i) < \sigma(j)$. Therefore the same holds for $x_{\alpha\beta}$ and our claim is proved.

Finally, since the variables x_{ij} for which $i-j \leq p$ and $\sigma(i) < \sigma(j)$ are subject to no other constraints, we conclude that $B_\sigma \bigcap$ Hess (p, A) is analytically isomorphic to a $\mathbb{C}$-affine space of dimension

$$E_p(\sigma) = \text{ card } \{(i,j) \mid 1 \leq i,j \leq n,\ 1 \leq i-j \leq p,\ \sigma(i) < \sigma(j)\}.$$

QED

Remark. The Bruhat cells are quasiprojective subvarieties of Flag $(\mathbb{C}^n)$. Hence, their intersections with Hess (p, A) are still quasiprojective.

We now show with an example that the decomposition of Hess (p, A) into the disjoint union of the subsets $B_\sigma \bigcap$ Hess (p, A) is not cellular, in that the boundary of a subset of dimension k is not always contained in the union of the subsets of dimension $\leq (k-1)$.

Let $p = 1, n = 4$ and $\sigma = (4,2,3,1) \in \sum(4)$. We have $E_1(\sigma) = 2$, so that $dim(B_\sigma \bigcap$ Hess $(1, A)) = 2$. Since $3 = \sigma(3) > \sigma(2) = 2, B_\sigma$ is the slice of $ch(\sigma)$ obtained by setting $x_{32} = 0$. The equations defining $ch(\sigma) \bigcap$ Hess $(1, A)$ are

$$f_{31} = \Lambda(3,1)x_{31} - \Lambda(2,1)x_{32}x_{21} = 0$$

$$f_{42} = \Lambda(4,2)x_{42} - \Lambda(3,2)x_{43}x_{32} = 0$$

$$f_{41} = \Lambda(4,1)x_{41} - \{\Lambda(3,1)x_{43}x_{31} + \Lambda(2,1)x_{42}x_{21}\} + \Lambda(2,1)x_{43}x_{32}x_{21} = 0$$

and so the equations defining $B_\sigma \bigcap$ Hess $(1, A)$ are

$$\hat{f}_{31} = \Lambda(3,1)x_{31} = 0, \hat{f}_{42} = \Lambda(4,2)x_{42} = 0$$

$$\hat{f}_{41} = \Lambda(4,1)x_{41} - \{\Lambda(3,1)x_{43}x_{31} + \Lambda(2,1)x_{42}x_{21}\} = 0$$

Thus we see that on $B_\sigma \bigcap$ Hess $(1, A), x_{31} = x_{41} = x_{42} = 0$. Therefore

$$B_\sigma \bigcap \text{ Hess } (1, A) = \{< \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ x_{21} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & x_{43} & 1 \end{bmatrix} >, x_{21}, x_{43} \in \mathbb{C}\}$$

$$\{< \begin{bmatrix} 0 & 0 & x_{43} & 1 \\ x_{21} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} >, x_{21}, x_{43} \in \mathbb{C}\}.$$

Consider now the sequence $\{f_n\}_{n \in N}$ of points in $B_\sigma \bigcap$ Hess $(1, A)$ where

$$f_n =< \begin{bmatrix} 0 & 0 & n & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} >=< \begin{bmatrix} 0 & 0 & n & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/n & 1 \\ 0 & 0 & 0 & -n \end{bmatrix} >$$

$$=< \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/n & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} > .$$

Clearly,

$$\lim_{n \to \infty} f_n =< \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} >=< (4, 2, 1, 3) > \triangleq < \tau >,$$

and $dim\ (B_\tau \bigcap$ Hess $(1, A)) = E_1(\tau) = 2$. Therefore, in the closure of a subset of dimension 2, there is a point that belongs to another subset of dimension 2.

We are now in a position to prove our main result. We will use the following recent theorem of A.H. Durfee [16].

Theorem: Let X be a smooth complex projective variety. Suppose that X is a finite disjoint union $X_1 \bigcup \cdots \bigcup X_m$, where the X_i are smooth contractible quasiprojective subvarieties. Then

$$b_k(X) = \text{ card } \{X_i \mid 2\ dim\ X_i = k\}.$$

where $b_k(X)$ denotes the k^{th} Betti number of X.

(V.2) Definition: For $n > p$ and $k = 1, \cdots, (p(2n - p - 1)/2) + 1$ we define the generalized Eulerian numbers of degree p by

$$A_p(n, k) = \text{ card } \{\sigma \in \Sigma(n) \mid E_p(\sigma) = k - 1\}$$

The reason for this terminology will be briefly discussed after the following theorem.

(V.3) Theorem: Let $A \in GL(n, \mathbb{C})$ have distinct eigenvalues. Then

$$b_{2k}(\text{ Hess } (p, A)) = A_p(n, k+1)$$
$$b_{2k+1}(\text{ Hess } (p, A)) = 0$$

I.e., the even Betti numbers of Hess (p, A) are generalized Eulerian numbers of degree p, and the odd Betti numbers vanish.

Proof: Readily follows from (V.1) and Durfee's theorem.

QED

For $p = 1$ and $\sigma \in \sum(n), E_1(\sigma)$ simply counts the number of "falls" in the permutation σ, a fall being a value of i for which $\sigma(i) > \sigma(i+1)$. Thus $A_1(n, k)$ is the number of permutations in $\sum(n)$ with $(k-1)$ falls. Equivalently, by reading the permutations backwards, $A_1(n, k)$ is the number of permutations in $\sum(n)$ with $(k-1)$ "rises", a rise being a value of i for which $\sigma(i) < \sigma(i+1)$. These numbers are the well-known Eulerian numbers, which occur in a variety of combinatorial problems. Their exact value is given by

$$A_1(n, k) = \sum_{j=0}^{k-1} (-1)^j \binom{n+1}{j} (k-j)^n$$

For the elementary properties of the Eulerian numbers, the reader is referred to L. Comtet [17] or J. Riordan [18].

Several generalizations of the Eulerian numbers have been considered in the literature; examples can be found in [5], [6], [7], [8], [9]. Our generalization, perhaps new, includes another class of interesting numbers, namely those corresponding to the case $p = n - 1$. In this latter case, the "Eulerian dimension" $E_{n-1}(\sigma)$ of a permutation $\sigma \in \sum(n)$ coincides with its length $\ell(\sigma)$, that is

$$E_{n-1}(\sigma) = \ell(\sigma) = \text{ card } \{(i, j) \mid 1 \leq i < j \leq n \ \& \ \sigma(i) > \sigma(j)\}.$$

Thus, $A_{n-1}(n, k)$ counts the number of permutations in $\sum(n)$ with length $(k-1)$. The numbers $A_{n-1}(n, k)$ are also well-known, and sometimes referred as the Mahonian numbers. In addition to being the even Betti numbers of the complete flag manifold, they are best known as the coefficients in the expansion of the polynomial

$$(1+x)(1+x+x^2) \cdot \cdots \cdot (1+x+\cdots+x^{n-1}).$$

Again we refer to L. Contet [17] or J. Riordan [18] for their elementary properties.

Acknowledgment: We would like to thank G. Ammar for suggesting the topic and J. Wolper for bringing the paper of Durfee [16] to our attention. We would also like to thank both of them for several useful discussions.

References

[1] G.S. Ammar and C.F. Martin, "The geometry of matrix eigenvalue methods," Acta Aplicandae Mathematicae, 5 (1986), 239-278.

[2] G.S. Ammar, "Geometric aspects of Hessenberg matrices," NSF Conference on Differential Geometry: The Interface Between Pure and Applied Mathematics, San Antonio, TX, April, 1986, to appear.

[3] H. Hiller, The Geometry of Coxeter Groups, Pitman, London, 1982.

[4] D. Foata, "Distributions Eulériennes et Mahoniennes sur le groupe des permutations," in Higher Combinatorics (ed. M. Aigner), Reidel, Dordrecht, 1977.

[5] L. Carlitz, "Eulerian numbers and polynomials of higher order," Duke Math. J., 27 (1960), 401-423.

[6] L. Carlitz and R. Scoville, "Generalized Eulerian numbers: combinatorial applications," Jour. für die Reine und Angew. Math., 265 (1974), 110-137.

[7] J. Dillon and D. Roselle, "Eulerian numbers of higher order," Duke Math. J., 35 (1968), 247-256.

[8] D. Lehmer, "Generalized Eulerian numbers," Jour. of Combin. Theory, Series A, 32 (1982), 195-215.

[9] D. Rawlings, "The r-major index," Jour. of Combin. Theory, Series A, 31 (1981), 175-183.

[10] M. Shub and A.T. Vasquez, "Some linearly induced Morse-Smale systems, the QR-algorithm and the Toda lattice," Contemp. Math, 64 (1986).

[11] S. Smale, "Differentiable dynamical systems," Bull. Amer. Math. Soc., 73 (1967), 747-817.

[12] G.H. Golub and C.F. Van Loan, Matrix Computations, John Hopkins Univ. Press, Baltimore, MD, 1984.

[13] J. Tits, "Théorème de Bruhat et sous-groupes paraboliques," C.R. Acad. Sci. Paris, 254 (1962), 2910-2912.

[14] M.A. Shayman, "Phase portrait of the matrix Riccati equation," SIAM J. Contr., 24 (1986), 1-65.

[15] M.A. Shayman, "Riccati equations, linear flows on the flag manifold, and the Bruhat decomposition," in Frequency Domain and State Space Methods for Linear Systems, (eds. C.I. Byrnes and A. Lindquist), North-Holland, Amsterdam, 1986.

[16] A.H. Durfee, "Algebraic varieties which are a disjoint union of subvarieties," in Geometry and Topology, Manifolds, Varieties, and Knots, (eds. C. McCrory and T. Shirfin), Dekker, New York, 1987.

[17] L. Comtet, Analyse Combinatoire, Vol. I& II, Presses Univ. de France, Paris, 1970.

[18] J. Riordan, An Introduction to Combinatorial Analysis, Wiley, New York, 1958.