Reprinted from

M. Hochster C. Huneke J.D. Sally
Editors

Commutative Algebra

Proceedings of a Microprogram Held

June 15-July 2, 1987

Springer-Verlag
New York Berlin Heidelberg
London Paris Tokyo

Rank Varieties of Matrices

DAVID EISENBUD AND DAVID SALTMAN

Abstract. In this paper we extend work of Gerstenhaber [5], Kostant [9], Kraft-Procesi [10], Tanisaki [15], and others on orbit closures in the nilpotent cone of matrices by studying varieties of square matrices defined by conditions on the ranks of powers of the matrices, or more generally on the ranks of polynomial functions of them. We show that the irreducible components of such varieties are always Gorenstein with rational singularities (in particular they are normal). We compute their tangent spaces, and also their limits under deformations of the defining polynomial functions. We also study generators for the ideals of such varieties, and we compute the singular loci of the hypersurfaces in the space of $n \times n$ matrices given by the vanishing of a single coefficient of the characteristic polynomial.

Introduction.

Let V be an n-dimensional vectorspace over the complex numbers (or, with suitable scheme-theoretic interpretations of what is to come, over any field of characteristic 0), and let $X = \operatorname{End} V$ be the space of linear transformations of V into itself.

By analogy with the algebraic subsets of the affine line, it is natural to consider sets of transformations $A \in X$ defined by the vanishing of a collection of polynomial functions $p_i(A)$. Since the rank of a matrix in case $n > 1$ can take on values other than 0 and n, it is natural to extend consideration to sets of matrices of the form

$$\{A \in X \mid \operatorname{rank} p_i(A) \leq r_i, \ i = 1, \dots\}$$

for various polynomials $p_i(t)$ of one variable, and numbers r_i. In this paper we are concerned with sets defined in this way; we will call them *rank sets.*

A rank set is clearly invariant under the action of $\mathrm{PGL}(V)$ on X by conjugation. On the other hand, among the rank sets are the closures of the orbits of this action. Indeed, as shown by Gerstenhaber [5], the closure of the orbit of a transformation A is the set of those transformations B such that for every $i = 1, 2, \dots$ and for every $\lambda \in \mathbb{C}$ (or just every λ which is an eigenvalue of A),

$$\operatorname{rank}(B - \lambda)^i \leq \operatorname{rank}(A - \lambda)^i.$$

This case, and in particular the case of a nilpotent orbit closure to which it quickly reduces, has been studied by many authors from both topological and algebraic points of view; in particular, the results of this paper were already known for orbit closures.

At the opposite extreme of the rank sets are the determinantal varieties, each given by a single condition of the form rank $A \leq r$. Here there are many orbits, of many different dimensions, and infinitely many of these have maximal dimension.

Both these examples of rank sets are irreducible, but in the general case, irreducibility cannot be expected: a set of the form

$$\{A \mid \operatorname{rank}[(A-\lambda)^i(A-\mu)^j]\} \leq r,$$

with $\lambda \neq \mu$, will consist in the union, over pairs of positive integers r_1, r_2 whose sum is r, of the set of A with rank $(A-\lambda)^i \leq r_1$ and $\operatorname{rank}(A-\mu)^j \leq r_2$. We are thus led to consider first the rank varieties defined in terms of polynomials $p_i(t)$, each of which is a power of an irreducible polynomial, and we accordingly make the following definition:

For any sequence of complex numbers $\lambda = (\lambda_1, \dots, \lambda_s)$, and for any doubly indexed set of integers $r(i,j)$, $1 \leq i \leq s$, $1 \leq j$ let $X_{r,\lambda}$ be the subset of $X = \operatorname{End}(V)$ given by

$$X_{r,\lambda} = \{A \in X \mid \operatorname{rank}(A-\lambda_i)^j \leq r(i,j) \text{ for all } 1 \leq i \leq s, 1 \leq j\}.$$

Extending the example above we easily derive:

PROPOSITION. *Every rank set is a union of sets of the form* $X_{r,\lambda}$.

PROOF SKETCH: Since the intersection of two sets of the form $X_{r,\lambda}$ is again of that form, it is enough to treat the case of a rank set Y defined by a single condition rank $p(A) \leq u$. We may factor $p(t)$ into powers of irreducibles $p(t) = (t-\lambda_1)^{i_1}(t-\lambda_2)^{i_2}\dots(t-\lambda_s)^{i_s}$. Given a tranformation $A \in X$, we may split V into eigenspaces of A with distinct eigenvalues, and from this splitting we see that the corank of $p(A)$ is simply the sum of the coranks of the transformations $(A-\lambda_j)^{i_j}$. Thus Y is the union, over all sequences $(u_1, \dots, u_s)$ summing to u, of the sets $X_{r,\lambda}$ with $r(j,k) = n$, the dimension of V, for $k < i_j$, while $r(j,k) = u_j$ for $k \geq i_j$. □

This being so, we focus our attention on the $X_{r,\lambda}$. Of course the simplest of these are the ones given by conditions relative to a single eigenvalue, that is, those for which s is 1, and, after translating by a scalar matrix, we may even assume that the eigenvalue in question is 0. For convenience, we define, for any integer valued function $r(j)$,

$$X_r = \{A \in X \,|\, \operatorname{rank} A^j \leq r(j) \text{ for all } 1 \leq j\}.$$

Of course for general r there may not be any matrix A with rank $A^j = r(j)$ for all j. But it is not hard to show that such an A exists iff r is a decreasing concave non-negative integral function with $r(0) = n$; we will call such a function a rank function. Given an arbitrary non-negative function, r with $r(0) \geq n$, it is easy to see that there is a unique maximal rank function r' with $r'(j) \leq r(j)$ for every j, and $X_r = X_{r'}$.

The main results of Sections 1 and 2 of this paper may now be summarized as follows:

THEOREM 1. *The varieties $X_{r,\lambda}$ are Gorenstein, and are normal with rational singularities. Their limits under deformations of the λ_i are always varieties of the same form. Each $X_{r,\lambda}$ is a generically transverse intersection of translates by scalar matrices of varieties of the form $X_{r'}$, for suitable $r'(i)$.*

In some respects the hardest part of this is the normality statement, in whose proof we closely follow and augment ideas of DeConcini and Procesi [**3**]. The fundamental notions in their arguments are that of the variety of pairs of transformations whose composition is nilpotent, and that of a rather natural complete intersection of which it is the quotient, whose points correspond roughly to the transformations which are the possible products of the transformations in the original pair. In Section 1 we make the necessary changes to extend these ideas to the non-nilpotent situation, proving the normality of the varieties X_r. As in the argument of DeConcini and Procesi [**3**], the hypothesis of characteristic 0 enters to prove that the X_r are actually the desired quotients.

We return to the geometry of the $X_{r,\lambda}$ in Section 2. We construct a canonical resolution of singularities of the varieties $X_{r,\lambda}$. We then use the Grauert-Riemenschneider vanishing theorem to derive the normality of these varieties from that of the X_r and to establish the remaining properties given in the Theorem; of course this requires the characteristic 0 hypothesis again. Nevertheless, it seems reasonable to hope that the theorem remains true over an arbitrary algebraically closed field. We also derive an interesting representation for the tangent spaces to X_r at some smooth points.

In the third section of the paper we study the equations of the varieties $X_{r,\lambda}$. By the Theorem, it is enough to treat the X_r. Choosing a basis for V, we may identify $\operatorname{End} V$ with the set of $n \times n$ matrices. We give generators up to radical for the ideals of polynomials in the entries of these matrices which vanish on the varieties X_r. In a preliminary version of this paper we conjectured, extending previous conjectures of DeConcini-Procesi [**3**] and Tanisaki [**15**] which dealt with the nilpotent case, that these generators suffice to generate the ideals of the varieties themselves. After having read our preliminary version, J. Weyman [**16**] proved our conjectures in this direction.

The basic building blocks for the ideals defining the equations of the X_r are ideals identified by Tanisaki that we call $I(\lambda_d^t)$, for various t and d. $I(\lambda_d^t)$ is the ideal generated by the coefficients of x^{t-d} in the t-th exterior power of the matrix

$$x - A_{\text{gen}},$$

where A_{gen} is the generic matrix, whose entries are the coordinate functions on X, and x is an auxilliary variable; thus the generators of $I(\lambda_d^t)$ are polynomials of degree d in the entries of a matrix. By way of familiar examples, $I(\lambda_d^n)$ is generated by the single polynomial which is the d-th coeffient of the characteristic polynomial of the generic matrix, while $I(\lambda_d^d)$ is the ideal generated by all $d \times d$ minors of the generic matrix.

In the final section, solving the problem from which this paper originally arose, we identify for all k the singular locus S_k of the hypersurface in X defined by the vanishing of the coefficient of x^{n-k} in the characteristic polynomial

$$\det(x - A_{\text{gen}}):$$

THEOREM 4.10. *The matrix A is a singular point of the k-th coefficient of the characteristic polynomial iff writing $d := \operatorname{rank} A^n$, we have $d \leq k-2$ and $\operatorname{rank} A^{k-d+1} \leq d$.*

We also give an irredundant decomposition of S_k into irreducible components: these turn out to be rank varieties (Corollary 4.14).

These singular loci are related in an interesting way to the reduced varieties defined by the ideals $I(\lambda_k^t)$ introduced above. The variety defined by the vanishing of the k-th coefficient of the characteristic polynomial is $V(\lambda_k^n)$, and we have:

THEOREM 4.15. *The singular locus of* $V(\lambda_k^n)$ *is* $V(\lambda_{k-1}^{n-1})$.

This result is particularly suggestive in view of the well-known identification of the singular locus of a determinantal variety; in that case one has $\operatorname{Sing} V(\lambda_k^k) = V(\lambda_{k-1}^{k-1})$. One might at first conjecture that $\operatorname{Sing} V(\lambda_k^t) = V(\lambda_{k-1}^{t-1})$ in general, but examples show that this is false; however, it remains plausible that this conjecture becomes true if one replaces the two varieties by their normalizations, or even by the disjoint union of their irreducible components.

1. Normality.

1A) Linear Algebra. Let $r : Z^+ \to Z^+$ be a rank function. Our goal is to study $X_r = \{A \in M_n(F) \mid \operatorname{rank}(A^i) \le r(i)\}$. To begin with, we consider the open subset $Y_r = \{A \in M_n(F) \mid \operatorname{rank}(A^i) = r(i)\}$. The stable value of r is zero if and only if Y_r consists of nilpotent matrices. The stable rank of an $A : U \to U$ is the stable value of its rank function.

Suppose $A : U \to U$ is arbitrary. If U has dimension n define the stable space, V, of A to be $A^n(U)$, a subspace of U. Observe that V is the direct sum of the eigenspaces of A associated to nonzero eigenvalues. It follows that $A(V) = V$ and A restricted to V is an isomorphism. Of course, the dimension of V is the stable rank of A. For any A, we define A' to be the induced linear transformation $U/V \to U/V$. Let s be the stable rank of A. Note that A' is nilpotent with rank function $r - s$, where r is the rank function of A. In general, if r is a rank function, define r' to be $r - s$, where s is the stable value of r. For any Y_r, choose U' to be a space with dimension $\dim(U) - s$, s being the stable value of r. Consider $Y_{r'}$ as a subset of $\operatorname{End}(U')$.

Let us state two elementary properties of the Y_r's.

LEMMA 1.1. Y_r *is irreducible and* $\operatorname{codim}_{\operatorname{End}}(U)(Y_r) = \operatorname{codim}_{\operatorname{End}}(U')(Y_{r'})$.

PROOF (outline): Let s be the stable value of r and consider the Grassmann variety G_s of s dimensional subspaces of U. Define $W \subseteq \operatorname{End}(U) \times G_s$ to be the locally closed subvariety of pairs (A, V) such that $A(V) = V$. Let $p : W \to \operatorname{End}(U)$ and $q : W \to G_s$ be the restrictions of the projection maps. Define $W_r \subseteq W$ to be the inverse image under p of Y_r. The map p restricted to W_r is an isomorphism. The inverse map sends A to (A, V), where V is the stable space of A.

Next consider the induced map $q_r : W_r \to G_s$. This is a surjection. If $V \in G_s$, $q_r^{-1}(V)$ consists of $A : U \to U$ with stable space V such that $A' : U/V \to U/V$ has rank function $r' = r - s$. We write $q_r^{-1}(V)$ as $Y_r(V)$. There is a regular map $t : Y_r(V) \to \mathrm{End}(U/V)$ defined by $t(A) = A'$. The image of t is precisely $Y_{r'}$. For any $C \in Y_{r'}$, $t^{-1}(C)$ consist of maps, A, whose matrix (with respect to a suitable basis) in block form is:

$$\begin{pmatrix} C & 0 \\ D & E \end{pmatrix}$$

where E is the matrix of A restricted to V, and so is nonsingular. It follows that all the $t^{-1}(C)$ are irreducible of dimension ns. By e.g., Shafarevich [**13**, p. 61], $Y_r(V)$ is irreducible of dimension $\dim(Y_{r'})+ns$ and again W_r is irreducible of dimension $\dim(Y_{r'})+ns+s(n-s)$. Thus $\mathrm{codim}_{\mathrm{End}}(U)(Y_r) = n^2 - \dim(Y_r) = n^2 - \dim(W_r) = n^2 - 2ns + s^2 - \dim(Y_{r'}) = \mathrm{codim}_{\mathrm{End}}(U')(Y_{r'})$. □

Our next subject is a theory for pairs of linear maps that parallels the one above for single ones. Formally speaking, a *pair* will be $\alpha = (A : U \to V; B : V \to U)$ where U, V are vector spaces and A and B are linear maps. For fixed U, V, then, a pair α is an element of $\mathrm{Hom}(U, V) \times \mathrm{Hom}(V, U)$ and we will write $\alpha = (A, B)$ for convenience. $\alpha = (A : U \to V; B : V \to U)$ and $\beta = (C : S \to T; D : T \to S)$ are isomorphic if there are linear isomorphisms $\varphi : U \to S$ and $\tau : V \to T$ such that $C = \tau A \varphi^{-1}$ and $D = \varphi B \tau^{-1}$. If we fix U and V, then these same equations define an action of $GL(U) \times GL(V)$ on $\mathrm{Hom}(U, V) \times \mathrm{Hom}(V, U)$, that is, on all pairs defined using U, V. If α and β are as above, we define the direct sum $\alpha \oplus \beta$ to be the pair $(A \oplus C : (U \oplus S) \to (V \oplus T);\ B \oplus D : (V \oplus T) \to (U \oplus S))$. We say α is a *nilpotent pair* if AB (and therefore BA) is nilpotent.

Suppose $\alpha = (A : U \to V; B : V \to U)$ is a pair, and let V'' be the stable space of $AB : V \to V$. Set $U'' = B(V'')$. If n is greater than the dimension of U and V, then $U'' = B(AB)^n(V) = (BA)^n B(V) \subseteq (BA)^n(U)$. On the other hand, $U'' \supseteq (BA)^n BA(U) = (BA)^{n+1}(U)$. But $n \geq \dim(U)$ so $(BA)^n(U) = (BA)^{n+1}(U)$ and U'' is the stable space of BA. Now B is injective on V'', and so B induces an isomorphism from V'' to U''. Dually, A restricts to an isomorphism from U'' to V''. If $U' = U/U''$, and

$V' = V/V''$, then A and B induce $A' : U' \to V'$ and $B' : V' \to U'$. If we define $\alpha' = (A' : U' \to V'; B' : U' \to V')$ then α' is a nilpotent pair. We call α' the reduced pair of α.

We next must recall the classification theory of nilpotent pairs, for which Kraft-Procesi [**10**] is a good source. Fix U and V, and let α be a nilpotent pair. Then α can be written uniquely (up to order and isomorphism) as a direct sum of indecomposable nilpotent pairs. There is a one to one correspondence between isomorphism classes of these indecomposables and strings of alternating symbols "a" and "b". For example, consider the string $\sigma = abab\dots ba$ of length $2n+1$. To this string we associate the indecomposable pair described as follows. U has basis $a_1, \dots, a_{n+1}$ and V has basis $b_1, \dots, b_n$. A is defined by $A(a_i) = b_i$ for $i \le n$, $A(a_{n+1}) = 0$, and $B(b_i) = a_{i+1}$. Such an A and B look like:

$$A: \begin{array}{ccc} a_1 & \longrightarrow & b_1 \\ \vdots & & \vdots \\ a_n & \longrightarrow & b_n \\ a_{n+1} & \longrightarrow & 0 \end{array} \qquad\qquad B: \begin{array}{ccc} b_1 & & a_1 \\ & \searrow & \\ \vdots & \vdots & \vdots \\ b_n & & a_n \\ & \searrow & \\ & & a_{n+1} \end{array}$$

We denote the pair we just defined by $\alpha(\sigma)$. If σ' is the string $bab\dots ab$ of length $2n+1$, we define the pair $\alpha(\sigma')$ associated to σ' just as above but with U, V; A, B; and the a's and b's switched.

We define the indecomposable nilpotent pairs associated with even length strings as follows. Let σ be the string $abab\cdots ab$ of length $2n$. Let U have basis $a_1, \dots, a_n$ and V have basis $b_1, \dots, b_n$. Define $A(a_i) = b_{i+1}$ for $i < n$, $A(a_n) = 0$, and $B(b_i) = a_i$. Such A and B look like:

$$A: \begin{array}{ccc} a_1 & & b_1 \\ & \searrow & \\ \vdots & \vdots & \vdots \\ a_n & & b_n \\ & \searrow & \\ & & 0 \end{array} \qquad\qquad B: \begin{array}{ccc} b_1 & \longrightarrow & a_1 \\ \vdots & & \vdots \\ b_n & \longrightarrow & a_n \end{array}$$

This defines a nilpotent pair we again denote by $\alpha(\sigma)$. By reversing a's, b's; A, B and U, V we define $\alpha(\sigma')$, for the string $\sigma' = baba\dots ba$. Note that our definition here is reversed from the one in Kraft-Procesi [**10**]. All indecomposable nilpotent pairs are isomorphic to $\alpha(\sigma)$ for some string σ, and different strings give nonisomorphic pairs.

Let Σ be the set of strings as above and let $\wedge$ denote the set of formal sums of elements of Σ with nonnegative integer coefficients. To any nilpotent pair α, the decomposition of α into indecomposables defines a unique $\eta(\alpha) \in \wedge$ where for each string σ, the coefficient of σ in $\eta(\alpha)$ is the number of times $\alpha(\sigma)$ appears in α. Obviously, we have a one to one correspondence between isomorphism classes of nilpotent pairs and $\wedge$. Also, $\eta(\alpha \oplus \alpha') = \eta(\alpha) + \eta(\alpha')$. If α is an arbitrary pair, let α' be the reduced nilpotent pair, and define $\eta(\alpha) = \eta(\alpha')$.

For any string $\sigma \in \Sigma$, we can define $\sigma(A, B)$ to be the linear map gotten by substituting A for a and B for b in the string and then computing the composition. For example, if $\sigma = ababa$, $\sigma(A, B)$ is $ABABA$. If $\alpha = (A, B)$ is a pair, we write $\sigma(A, B)$ as $\sigma(\alpha)$.

Suppose α is a pair, not necessarily nilpotent. Then α defines a rank function $r_\alpha : \Sigma \to Z^+$ as follows. If $\sigma \in \Sigma$, define $r_\alpha(\sigma) = \text{rank}(\sigma(\alpha))$. If σ, σ' are strings, and σ is strictly longer than σ', then $r_\alpha(\sigma) \leq r_\alpha(\sigma')$. There is an integer N such that if σ is longer than N, $r_\alpha(\sigma)$ has constant value, and this stable value of r_α. is the stable rank of α. A function $r : \Sigma \to Z^+$ is called a rank function if $r = r_\alpha$ for some α.

THEOREM 1.2.

a) *Let α, β be two pairs. Then $\eta(\alpha) = \eta(\beta)$ if and only if $r_\alpha = r_\beta$.*
b) *Let α be a pair and set $\eta(\alpha) = \eta$. Let $V_\eta \subseteq \text{Hom}(U, V) \times \text{Hom}(V, U)$ be all pairs β with $\eta(\beta) = \eta$. Then V_η is a locally closed subvariety of $\text{Hom}(U, V) \times \text{Hom}(V, U)$.*

PROOF: Part a) implies that V_η is defined by rank equations, so b) follows from a). As for a), suppose $\eta(\alpha) = \eta(\beta)$. If α' and β' are the respective reduced pairs, then α' and β' are isomorphic. In particular, $r_{\alpha'} = r_{\beta'}$, and α' and β' are defined on spaces of the same dimension. Thus α, β have equal stable rank, s, and $r_\alpha = r_{\alpha'} + s = r_{\beta'}, + s = r_\beta$.

Conversely, suppose $r_\alpha = r_\beta$. We must show that $\eta(\alpha) = \eta(\beta)$. As r_α and r_β have equal stable values, α and β have equal stable rank s. Letting α', β' be the reduced pairs again, we have $r_{\alpha'} = r_{\beta'}$. Thus we may assume α, β are both nilpotent pairs. Let σ be a string appearing in $\eta(\alpha)$ of maximal length m. It suffices by induction to show that σ appears in β.

To finish our argument, we consider the behavior of the rank function of an indecomposable nilpotent pair. Recall that $\alpha(\sigma)$ is the indecomposable associated with σ, and $\tau(\alpha(\sigma))$ is the linear map derived by plugging the pair of maps of $\alpha(\sigma)$ into the pattern τ and composing. One can compute that $\sigma(\alpha(\sigma))$ is the zero map, and if σ' is the string derived from σ by removing the rightmost symbol, $\sigma'(\alpha(\sigma)) = 0$ also. If m is the length of σ, then $\tau(\alpha(\sigma)) = 0$ for any string τ of length greater than or equal to m. However, if σ'' is the string with the leftmost symbol removed, $\sigma''(\alpha(\sigma))$ is not zero.

The fact that m is the length of a maximal string in $\eta(\alpha)$ is thus equivalent to the fact that $r_\alpha(\sigma) = 0$ for both strings of length m but $r_\alpha(\sigma') \neq 0$ for some σ' of length $m-1$. Thus m is also the maximal length of a string in β. Let σ have this maximal length m and let σ'' be the string with the leftmost symbol removed. The fact that σ appears in $\eta(\alpha)$ is equivalent to $r_\alpha(\sigma'') \neq 0$, and so if σ appears in $\eta(\alpha)$ it appears in $\eta(\beta)$. The theorem is proved.

Since we have shown V_η is defined by rank equations, it makes sense to change our notation and to define, for a rank function $r = r_\alpha : \Sigma \to Z^+$, V_r to be the subvariety of $\mathrm{Hom}(U,V) \times \mathrm{Hom}(V,U)$ consisting of pairs β with $r_\beta = r$. We have seen that if $\eta = \eta(\alpha)$, then $V_r = V_\eta$. We study the variety V_r by reducing to the case of nilpotent pairs, as follows.

If $r = r_\alpha$ is a rank function, let s be the stable value of r and set $r' = r - s$. We know that r' is the rank function of the reduced (nilpotent) pair α'. Fix u and v as the dimensions of U and V respectively. Choose spaces U', V' of dimension $u-s$ and $v-s$ respectively, and consider $V_{r'} \subseteq \mathrm{Hom}(U',V') \times \mathrm{Hom}(V',U')$. Since α' is a nilpotent pair, $V_{r'}$ is the orbit of α' under $GL(U') \times GL(V')$ and so $V_{r'}$ is irreducible. In order to state the next result, set $P = \mathrm{Hom}(U,V) \times \mathrm{Hom}(V,U)$ and $P' = \mathrm{Hom}(U',V') \times \mathrm{Hom}(V',U')$.

PROPOSITION 1.3. *For any rank function $r = r_\alpha$, V_r is irreducible and* $\mathrm{codim}_{P'}(V_r) = \mathrm{codim}_{P'}(V_{r'})$.

PROOF: Let s be the stable rank of α and let $G_s(U)$ and $G_s(V)$ be the Grassman varieties of subspaces of U respectively V of dimension s. Let $W \subseteq \mathrm{Hom}(U,V) \times \mathrm{Hom}(V,U) \times G_s(U) \times G_s(V)$ be the locally closed subvariety of points (A, B, U'', V'') such that $A(U'') = V''$ and $B(V'') = U''$. There are regular maps $p : W \to G_s(U) \times G_s(V)$ and $q : W \to \mathrm{Hom}(U,V) \times \mathrm{Hom}(V,U)$ induced by the respective projections. Define W_r to be $q^{-1}(V_r)$. Let q_r be the restriction of q to W_r. If $\beta = (A,B) \in V_r$, let U'' and V'' be the stable spaces of β in U, V. U'' and V'' necessarily have dimension s and we may define $f : V_r \to W_r$ by setting $f(\beta) = (A, B, U'', V'')$. The regular map f is the inverse to q_r and so $W_r \cong V_r$.

The restriction, $p_{r'}$, of p to W_r is a surjection. If $(U'', V'') \in G_s(U) \times G_s(V)$, then $p_r(U'', V'')$ consists of all pairs $\beta = (A, B)$ such that $A(U'') = V''$, $B(V'') = U''$, and the reduced pair $(A' : U/U'' \to V/V''; B' : V/V'' \to U/U'')$ lies in $V_{r'}$. Thus there is a surjective regular map $f : p_r^{-1}(U'', V'') \to V_{r'} \subseteq \mathrm{Hom}(U/U'', V/V'') \times \mathrm{Hom}(V/V'', U/U'')$. If $\beta' \in V_r$, then $f^{-1}(\beta')$ consists of all pairs of linear maps whose matrices (with respect to a suitable basis) look like:

$$\begin{pmatrix} A' & 0 \\ B & D \end{pmatrix}, \quad \begin{pmatrix} B' & 0 \\ E & G \end{pmatrix}$$

where D and G are nonsingular. Thus all $f^{-1}(\beta')$ are irreducible and isomorphic, of dimension $us + vs$. Hence $p^{-1}(U'', V'')$ is irreducible of dimension $\dim(V_{r'}) + us + vs$ and W_r is irreducible of dimension $\dim(V_{r'}) + us + vs + s(u-s) + s(v-s)$. Thus $\mathrm{codim}(V_r) = 2uv - us - vs - s(u-s) - s(v-s) - \dim(V_{r'}) = 2(u-s)(v-s) - \dim(V_{r'}) = \mathrm{codim}(V_{r'})$.

If $\alpha = (A : U \to V; B : V \to U)$ is a pair define $\pi(\alpha) = AB \in \mathrm{End}(V)$ and $\rho(\alpha) = BA \in \mathrm{End}(U)$. If α is a nilpotent pair, we can describe the Jordan normal form of $\pi(\alpha)$ and $\rho(\alpha)$ as follows. Let Π be the set of positive integers, and Ω the set of formal sums of elements of Π with nonnegative coefficients. We think of Ω as the set of all possible Jordan normal forms of a nilpotent matrix. Define a map $\mu_a : \Sigma \to \Pi \cup \{0\}$ by counting the number of a's in any string in Σ. μ_a induces a map, we also call μ_a, from Λ to Ω, where μ_a maps the string "b" to 0. We make an analogous definition of μ_b. It is observed in Kraft-Procesi **[10]** that $\mu_a(\eta(\alpha))$ is the Jordan block decomposition of $\pi(\alpha)$, and $\mu_b(\eta(\alpha))$ is the block decomposition of $\rho(\alpha)$.

Keep the spaces U, V fixed. Let r be a rank function $r : \Sigma \to Z^+$ with stable value s. To r is associated an element $\eta \in \Lambda$ such that $V_r = V_\eta$. Let $\theta = \mu_a(\eta)$. Let t be the rank function $t : Z^+ \to Z^+$ associated to θ. If $t = t' + s$, then t is the rank function of any $C : V \to V$ such that the reduction C' has form θ. Using reduction it is easy to see that the restriction of π to V_r maps onto $Y_t \subseteq \mathrm{End}(V)$.

LEMMA 1.4. *$\pi : V_r \to Y_t$ is smooth.*

PROOF: If the stable value $s = 0$, then this was shown in Kraft-Procesi **[10]**. If $U' \subseteq U$ and $V' \subseteq V$, then recall that $Y_t(V')$ is the subvariety with stable space V' and define $V_r(U', V')$ to be the subvariety with stable spaces U', V'. An easy argument using reduction shows that the map $\pi : V_r(U', V') \to Y_r(V')$ is smooth.

Let $P \subseteq GL(U)$ and $Q \subseteq GL(V)$ be parabolic subgroups such that $G_s(U) \cong GL(U)/P$ and $G_s(V) \cong GL(V)/Q$. The maps $GL(U) \to G_s(U)$ and $GL(V) \to G_s(V)$ have sections defined locally, and from this it follows that the maps $f : V_r \to G_s(U) \times G_s(V)$ and $g : Y_t \to G_s(V)$ are locally trivial. That is, in the second case say, $G_s(V)$ has an open cover $G_s(V) = \cup W_i$ such that $f^{-1}(W_i) \cong Y_t(V') \times W_i$ for some $V' \in W_i$. The lemma follows.

1B) Invariant Theory. Let U be a vector space, r a rank function with $r(0) = \dim(U)$, and $X_r \subseteq \mathrm{End}(V)$ the variety $\{A \mid \mathrm{rank}(A^i) \leq r(i)$ for all $i\}$. It is the purpose of this section to show that X_r is a normal variety. Thus for this section we fix r and consider $X = X_r$.

If r has stable value 0 then X_r is the closure of $Y_r = \{A \mid \mathrm{rank}(A^i) = r(i)\}$ and Y_r is the orbit of a single nilpotent linear map. In this case Kraft and Procesi have shown that X_r is normal, and a generalization of their argument proves our theorem. In this section we will outline the proof in our more general setting, only emphasizing those parts that differ significantly from Kraft and Procesi's paper.

Notice that we do not know the radical ideal defined by X; that is, the equations we have given only define X set theoretically. Thus it is difficult to imagine how to prove normality directly. One of the beautiful ideas of Kraft-Procesi is that normality can be proved by showing that X is the quotient of a normal variety under the action of a reductive algebraic group. We follow this tack closely.

Denote by s the stable value of r. Let n be the least integer such that $r(n) = s$. Note that X_r can equally well be defined by the finite set of inequalities $\mathrm{rank}(A^i) \leq r(i)$ for $i = 1, \ldots, n$. Let $U_0, U_1, \ldots, U_n$ be vector spaces such that $U_n = U$, and U_i has dimension $r(n-i)$. Note that we have arranged it so that the dimensions of the U_i form a strictly decreasing sequence. Let M be the affine space consisting of tuples $(A_0, B_0, A_1, \ldots, A_{n-1}, B_{n-1})$ where $A_i : U_i \to U_{i+1}$ and $B_i : U_{i+1} \to U_i$ are linear maps. M is the set of all tuples of maps that fit into the diagram:

$$U_0 \underset{B_0}{\overset{A_0}{\leftrightarrows}} U_1 \underset{B_1}{\overset{A_1}{\leftrightarrows}} \ldots \underset{B_{n-1}}{\overset{A_{n-1}}{\leftrightarrows}} U_n$$

Let $Z \subseteq M$ be the closed subscheme of M defined by the equations $A_{i-1}B_{i-1} = B_i A_i$ for $i = 1, \ldots, n-1$. Let G be the group $GL(U_0) \times \ldots \times$

$GL(U_n)$ and H the subgroup $GL(U_0) \times \dots GL(U_{n-1})$. We define an action of G on M by setting $(g_0, \dots, g_n)(A_0, \dots, B_{n-1}) = (g_1 A_0 g_0^{-1}, g_0 B_0 g_1^{-1}, \dots, g_n A_{n-1} g_{n-1}^{-1}, g_{n-1} B_{n-1} g_n^{-1})$. This action induces an action of G on Z.

The definition of M and Z can be trivially generalized to any sequence of spaces $U_0', \cdots, U_t'$. We will write these varieties as $M(U_0', \dots, U_t') \supseteq Z(U_0', \dots, U_t')$.

The definition of Z can be viewed in the following manner. Let N be the variety $\mathrm{End}(U_1) \times \dots \mathrm{End}(U_{n-1})$ and define $\varphi : M \to N$ by setting:

$$\varphi(A_0, \dots, B_{n-1}) = (B_1 A_1 - A_0 B_0, \dots, B_{n-1} A_{n-1} - A_{n-2} B_{n-2}) .$$

Z is then the variety $\varphi^{-1}(0)$, where $0 \in N$ is the 0 tuple.

Next define a map $\psi : M \to \mathrm{End}(U)$ by setting $\psi(A_0, \dots, B_{n-1}) = A_{n-1} B_{n-1}$. Note that if $A = A_{n-1} B_{n-1}$ is in the image of ψ, then $A^2 = A_{n-1} B_{n-1} A_{n-1} B_{n-1} = A_{n-1} A_{n-2} B_{n-2} B_{n-1}$. By induction, $A^i = A_{n-1} \dots A_{n-i} B_{n-i} \dots B_{n-1}$. In particular, A^i factors through U_{n-i} and so $\mathrm{rank}(A^i) \leq r(i)$. Thus $A \in X$. Conversely, it is elementary to see that if $A \in X$, then there is an $\alpha = (A_0, \dots, B_{n-1}) \in Z$ such that all the A's are injections and $\psi(\alpha) = A$.

Let Z_{red} be the reduced scheme defined by Z. The action of G on Z induces an action on Z_{red}, and by restriction we get an action of H on Z_{red}. Since X is reduced by definition, there is an induced map $\psi : Z_{\mathrm{red}} \to X$. The proof of the next fact follows the proof in Kraft and Procesi with only minor changes and so we omit it.

LEMMA 1.5. *ψ induces an isomorphism $Z_{\mathrm{red}}/H \cong X$.*

We next turn to finding some smooth points on Z. Let $M^0 \subseteq M$ be the open subset defined by requiring for each $i = 1, \dots, n-1$ that either A_i or B_i has maximal rank. Arguing as in Kraft and Procesi, we observe that the tangent space map $d\varphi : T(M) \to T(N)$ is surjective over each element of M^0. Note that when one follows the proof in Kraft and Procesi, one sees that no restrictions on A_0 or B_0 are required. Set $Z^0 = Z \cap M^0$ which is open in Z, and observe that we showed above that Z_0 maps onto X. Z^0 is certainly nonempty. We have shown:

LEMMA 1.6. *Z^0 is nonsingular of dimension* $\dim(M) - \dim(N)$.

We next want to show that $Z - Z^0$ has small dimension. More specifically, we show:

PROPOSITION 1.7. *$Z - Z^0$ has codimension ≥ 2 in Z.*

We begin the proof of the above proposition by noticing that elements of Z have stable spaces that generalize the stable spaces of section 1A. Let $(A_0, \dots, B_{n-1}) = \alpha$ be an element of Z. Define α_i to be the pair (A_i, B_i). Let Σ and $\wedge$ be as in section 1A, and define $\eta_i(\alpha) = \eta(\alpha_i)$. Set $U_i'' \subseteq U_i$ to be the stable space of $B_i A_i$. Since $B_i A_i = A_{i-1} B_{i-1}$, U_i'' is also the stable space of $A_{i-1} B_{i-1}$. Thus all U_i'' have equal dimension which we call the stable rank of α. In addition, $A_i(U_i'') = U_{i+1}''$ and $B_i(U_{i+1}'') = U_i''$.

For this fixed α, define $U_i' = U_i / U_i''$. Let $M' = M(U_0', \dots, U_n')$ and $Z' = Z(U_0', \dots, U_n')$. In the obvious way, α induces a point $\alpha' = (A_0', \dots B_{n-1}')$ such that all $A_i' B_i'$ are nilpotent and α' has stable rank 0. We call α' the reduction of α.

For each $\alpha \in Z$, define $\Gamma(\alpha) \in (\wedge \times \dots \times \wedge)$ ($n-1$ times) to be the sequence $\eta_i(\alpha)$. The relations defining Z force $\pi_a(\eta_i(\alpha)) = \pi_b(\eta_{i-1}(\alpha))$. Conversely, any sequence satisfying these relations can easily be seen to be $\Gamma(\alpha)$ for some α. Note that, by definition, $\Gamma(\alpha) = \Gamma(\alpha')$. Now fix $\Gamma = \Gamma(\alpha)$, let s be the stable rank of α, set $\eta = \eta_{n-1}(\alpha)$, and set $\pi_a(\eta) = \theta$. Let $t : Z^+ \to Z^+$ be the rank function of stable value s corresponding to θ. Choose U_i' to be a vector space of dimension $\dim(U_i) - s$ and define M' and Z' as above. Set Z_Γ to be the inverse image of Γ in Z, while Z_Γ' is the same for Z'. Define $N' = \mathrm{End}(U_1') \times \dots \times \mathrm{End}(U_{n-1}')$. We claim:

PROPOSITION 1.8. *Z_Γ is smooth and irreducible of dimension* $\dim(Z_\Gamma) = \dim(M) - \dim(N) - \dim(M') + \dim(N') + \dim(Z_\Gamma')$.

PROOF: Let $Z* = Z(U_0, \dots, U_{n-1})$ and $Z*_\Gamma \subseteq Z*$ be the inverse image of Γ in $Z*$ (yes, the last component of Γ is ignored). There is a pullback diagram:

$$\begin{array}{ccc} Z_\Gamma & \longrightarrow & V_\eta \\ \downarrow & & \downarrow \\ Z*_\Gamma & \longrightarrow & Y_r \end{array}$$

The right column is smooth by 1.4, so induction shows that Z_Γ is smooth and irreducible. A similar pullback diagram holds for Z_Γ', so the dimension equation also follows by induction.

Note that U_0' is not (0), so Z_Γ' is not quite one of the Z_λ studied by Kraft and Procesi. But like the variety defined by them, Z_Γ' consists of nilpotent pairs. In fact, by adding spaces with negative subscripts, one can see that

Z'_Γ is the "upper tail" of one of the varieties Z_λ defined in their paper. The key result can now be stated (compare Kraft-Procesi [**10**, p. 236]).

PROPOSITION 1.9. *Either* $Z_\Gamma \subseteq Z_0$ *or* $\dim(Z^0) - \dim(Z_\Gamma) \geq 2$.

PROOF: Let α be such that $\Gamma = \Gamma(\alpha)$. Note first of all that $Z_\Gamma \subseteq Z_0$ if and only if each $\eta_i(\alpha)$ has the property that either all strings in $\eta_i(\alpha)$ start with a "b" or all strings in $\eta_i(\alpha)$ end with an "a". Γ which violate the above we call *defective*. By 1.6 $\dim(Z^0) = \dim(M) - \dim(N)$, so by 1.8 to prove the Proposition we must show that:

(*) If Γ is defective then $\dim(M') - \dim(N') - \dim(Z'_\Gamma) \geq 2$.

It is necessary to give a formula for the dimension of Z'_Γ. Let t be the rank function of stable value 0 associated with $\mu_a(\eta_{n-1}(\alpha))$ and b the rank function associated with $\mu_b(\eta_0(\alpha))$. The point is that if $\alpha = (A'_0, \dots, B'_{n-1}) \in Z'_\Gamma$, then t is the rank function of $A'_{n-1}B'_{n-1}) \in \operatorname{End}(U'_n)$ and b is the rank function of $B'_0A'_0 \in \operatorname{End}(U'_0)$. Using the same argument as in [**10**, Corollary p. 241] we have that:

LEMMA 1.10. *Let u_i be the dimension of U'_i. Then:*

$$\dim(Z'_\Gamma) = \left(\sum u_i u_{i+1}\right) + 1/2(\dim(Y_b) + \dim(Y_t)).$$

Now $\dim(Y_b) \leq u_0(u_0 - 1)$ since b is the rank function of a nilpotent matrix. In order to study $\dim(Y_t)$, define $t' : Z^+ \to Z^+$ as follows. For $i \leq n$, set $t'(i) = u_{n-i}$. If $n < i < u_0 + n$, set $t'(i) = u_0 - (n - i)$. Finally, set $t'(i) = 0$ if $i \geq u_0 + n$. Since $u_1 - u_0 \geq 1$, t' is a rank function. Furthermore, $t(i) \leq t'(i)$ for all i. By [**10**, part a) of the Proposition p. 229], Y_t is in the closure of $Y_{t'}$, and so $\dim(Y_t) \leq \dim(Y_{t'})$. By part b) of that same proposition, $(1/2)\dim(Y_{t'}) =$

$$\sum t'(i)(t'(i-1) - t'(i)) = \left(\sum u_i(u_{i+1} - u_i)\right) + u_0(u_0 - 1)/2$$

It follows that $(1/2)(\dim(Y_b) + \dim(Y_t))$ is less than or equal to:

$$\left(\sum u_i(u_{i+1} - u_i)\right) + u_0(u_0 - 1)$$

Of course $\dim(M') - \dim(N') =$

$$2\sum u_i u_{i+1} - \sum u_i^2 = \sum u_i u_{i+1} + \sum u_i(u_{i+1} - u_i) + u_0 u_1$$

Combining these facts we have $\dim(M')-\dim(N')-\dim(Z'_\Gamma) \geq u_0$. Thus it suffices to prove (*) in the cases $u_0 = 1$ or 0. If $u_0 = 0$, then Z'_Γ *is* a Z_λ as in Kraft-Procesi and the result of their paper applies here. If $u_0 = 1$, then we can set $U_{-1} = (0)$ and Z'_Γ, M' and N' are isomorphic to the corresponding varieties in Kraft-Procesi, so the results there apply again. Thus (*) is proved.

Given (*), we have the Proposition 1.9. From 1.9, 1.7 immediately follows. Thus $\dim(Z) = \dim(Z_0) = \dim(M) - \dim(N)$. Since Z is defined by $\dim(N)$ equations, Z is a complete intersection. Thus Z is Cohen-Macaulay. As $Z_0 \subseteq Z$ is nonsingular and of codimension 2, Z is normal. As Z is a cone, it is also connected. By Serre's criterion (e.g., Matsumura [**11**, p. 125]), Z is reduced. Thus by 1.5 X_r is isomorphic to Z/H. It follows that X_r is normal.

2. Rational resolutions and the proof of the main theorem.

NOTATION: Throughout this section and the next we will adhere to the following notation: V will be a vectorspace of dimension n over $\mathbb{C}$, r will be a decreasing non-negative concave integral function (a "rank function") with $r(0) = n$. We will write $r(\infty)$ for the stable value of r (which is equal to $r(n)$). We write a for the partition $a(r) = (a_1(r), \ldots, 0, 0, \ldots)$ with $a_i(r) = r(i-1) - r(i)$. We write $b(r) = (b_1(r), \ldots, 0, 0, \ldots)$ for the dual partition to a, that is, $b_i(r)$ is the number of indices j with $a_j(r) \geq i$. It is easily verified that if $A \in \operatorname{End}(V)$ has rank function r, that is $\operatorname{rank} A^i = r(i)$, then

$$a_i(r) = \dim((\ker A) \cap (\operatorname{im} A^{i-1}))$$

while $b_i(r)$ is the size of the i-th (in descending order) block with eigenvalue 0 in the Jordan normal form of A.

In the last section we saw that if r is a rank function, then

$$X_r := \{A \in \operatorname{End}(V) |\, \operatorname{rank} A^k \leq r(k)\}_{\text{red}}$$

is a normal variety. In this section we will prove that it is in fact Gorenstein with rational singularities. We will also show that it fits into a flat family over $\mathbb{A}^m$ of normal varieties, whose fiber over a point $(\lambda_1, \ldots, \lambda_m)$ such that the λ_i are all distinct is

$$X_{r;\lambda_1,\ldots,\lambda_m} = \{A \in \operatorname{End}(V) |\, \operatorname{corank}(A - \lambda_i) \geq r(i-1) - r(i), i = 1, \ldots, m\}.$$

In fact we will construct a ("very weak") simultaneous resolution of singularities for this family of varieties, and we will construct the family from the resolution. We will use the resolution of X_r to give a description of the tangent space to X_r at a point corresponding to an endomorphism A such that rank $A^i = r_i$ for all i. As an application of the tangent space computation we show that the general fiber of this family can be written as the scheme-theoretic intersection of varieties like X_r.

To construct the resolution, suppose that m is the largest number such that $r(m) \neq r(m+1)$. Let $x_1, \ldots, x_m$ be coordinates on $\mathbb{A}^m$, and let $W \subset \mathbb{A}^m$ be a linear subvariety (that is, the translate of a subvectorspace). Let $F = \mathrm{Flag}(V; r_1, \ldots, r_m)$ be the variety of flags

$$(V = V_0 \supset V_1 \supset \cdots \supset V_m \supset 0) \quad \dim V_i = r(i).$$

Let

$$\mathcal{X}_{r,W} \subset \mathrm{End}(V) \times W \times F$$

be the subvariety of triples $(A, \lambda, \{V_i\})$ such that, with $\lambda_i := x_i(\lambda)$,

$$(A - \lambda_i)V_{i-1} \subset V_i \text{ for } i = 1, \ldots, m.$$

Write $\pi_1 = \pi_{1,W}$ for the projection of $\mathcal{X}_{r,W}$ to $\mathrm{End}(V) \times W$. Let $X_{r,W}$ be the reduced image of π_1 and let $X'_{r,W}$ be $\operatorname{spec} \pi_{1*}\mathcal{O}_{\mathcal{X}_{r,W}}$. Since the flag manifold is complete, π_1 is proper, so $X_{r,W}$ is a closed affine subvariety of $\mathrm{End}(V) \times W$, and $X'_{r,W}$ is also affine, and finite over $X_{r,W}$. If W is a single point $\lambda \in \mathbb{A}^m$, and the numbers λ_i are distinct, then $X_{r,W}$ is the same as the variety $X_{r;\lambda_1,\ldots,\lambda_m}$ defined previously.

THEOREM 2.1.

i) *$\mathcal{X}_{r,W}$ is smooth over W and irreducible, of dimension $n^2 - \sum_i a_i(r)^2 + \dim W$, with trivial canonical bundle, while π_1 is proper and birational, with*

$$R^i\pi_{1*}\mathcal{O}_{\mathcal{X}_{r,W}} = 0 \text{ for } i > 0.$$

ii) *$X_{r,W}$ is normal, so that $X_{r,W} = X'_{r,W}$, $\mathcal{X}_{r,W}$ is a rational resolution of singularities of $X_{r,W}$, and $X_{r,W}$ is Gorenstein.*

iii) *$X_{r,W}$ is the restriction of $X_{r,\mathbb{A}^m}$ to W, and it is flat over W.*

Taken together, statement ii) of the Theorem and the vanishing part of statement i) may be rephrased as saying that $\mathcal{X}_{r,W}$ is a rational resolution of the singularities of $X_{r,W}$; see Kempf et al [8].

We will postpone the proofs of this and our other results until after all the statements have been given.

Next we study the fibers of the family $X_{r,\mathbb{A}^m}$ over points $\lambda \in \mathbb{A}^m$ for which some of the λ_i may coincide. We partition the coordinates λ_i of λ by equality. Let p be a permutation of $\{1,\ldots,m\}$ such that

$$\begin{aligned} \lambda_{p(1)} &= \cdots = \lambda_{p(m_1)}, \\ \lambda_{p(m_1+1)} &= \cdots = \lambda_{p(m_2)}, \\ \cdots & \\ \lambda_{p(m_s+1)} &= \cdots = \lambda_{p(m_{s+1})}. \end{aligned}$$

and such that p preserves the order within each interval $m_i+1,\ldots,m_{i+1}$. Thus with $a_i = r(i-1)-r(i)$, we have $a_{p(m_i+1)} \geq \cdots \geq a_{p(m_{i+1})}$. We write

$$r(i,j) := \dim V - a_{p(m_i+1)} - \cdots - a_{p(m_i+j)} \qquad \text{for } i=1,\ldots,s \text{ and } j=1,\ldots,m_{i+1}-m_i$$

We have:

Corollary 2.2. *The scheme-theoretic fiber of $X_{r,W}$ over $\lambda \in W$ is*

$$X_{r,\lambda} = \{A \in \operatorname{End}(V) | \operatorname{rank}(A-\lambda_i)^j \leq r(i,j) \textit{ for all } i,j \textit{ as above}\}_{\text{red}}.$$

In particular, this variety is Gorenstein with rational singularities for every $\lambda \in \mathbb{A}^m$, and $\mathcal{X}_{r,\lambda}$ is a rational resolution of its singularities.

Using this description of the desingularization we can give a description of some tangent spaces to X_r:

Proposition 2.3. *If $\operatorname{rank} A^i = r(i)$ for all i then A is a smooth point of $X_r = X_{r,0}$, and the tangent space to X_r at A is naturally equal to*

$$\{\alpha \in \operatorname{End}(V) | \sum_{i+j=k-1} A^i \alpha A^j \textit{ maps } \ker A^k \textit{ into } \operatorname{im} A^k \textit{ for all } k\}.$$

A consequence is the following transversality result, which will play a role in the description to be given in the next section of the equations defining some of the varieties X_r:

COROLLARY 2.4. *For all $\lambda \in \mathbb{A}^m$, the variety $X_{r,\lambda}$ is scheme-theoretically the intersection*

$$\cap_i \left(\{A \in \operatorname{End}(V) | \operatorname{rank}(A - \lambda_i)^j \leq r(i,j) \text{ for all } j \text{ as above}\}_{\text{red}} \right).$$

PROOF OF THEOREM 2.1: We first show that $\mathcal{X}_{r,W}$ is smooth and irreducible. Write $\pi_2 : \mathcal{X}_{r,\mathbb{A}^m} \to \mathbb{A}^m \times F$ for the projection. Choose a cover of F by sets F' on which the bundles V_i are trivialized. The scheme-theoretic preimage by π_2 of $\mathbb{A}^m \times F'$ may be identified with the product of $\mathbb{A}^m \times F'$ with the set of block-upper-triangular-matrices of a suitable shape, and is thus smooth over $\mathbb{A}^m \times F'$ and irreducible. As $\mathbb{A}^m \times F$ is smooth over $\mathbb{A}^m$, we see that $\mathcal{X}_{r,W}$ is smooth for every W. Since $W \times F$ is irreducible, so is $\mathcal{X}_{r,W}$.

Next we use the adjunction formula to prove that the canonical bundle of $\mathcal{X}_{r,W}$ is trivial. Since $\mathcal{X}_{r,W}$ is a complete intersection with trivial normal bundle in $\mathcal{X} := \mathcal{X}_{r,\mathbb{A}^m}$, it suffices to treat the case $W = \mathbb{A}^m$.

Let V_F be the trivial bundle $V \times F$ on F, and let $W = (V_F = \mathcal{V}_0 \supset \cdots \supset \mathcal{V}_m \supset 0)$ be the tautological flag on F. Let $\mathcal{D} \subset \operatorname{End}(V_F)$ be the vector bundle of "upper-triangular" endomorphisms preserving $\mathcal{V}$ — that is, those A with $A\mathcal{V}_i \subset \mathcal{V}_i$ for all i — and let D be the total space of $\mathcal{D}$. We may regard $\mathcal{X}_{r,\mathbb{A}^m}$ as a subvariety of $D \times A^m$. Inside $\mathcal{D}$ is the bundle of "strictly upper-triangular" endomorphisms, that is, those A with $A\mathcal{V}_{i-1} \subset \mathcal{V}_i$ for all i; as is well-known and easy to verify by direct computation, it may be identified with the cotangent bundle $\mathcal{T}^*(F)$ to F.

On the other hand, pulling $\mathcal{D}$ back to D we get a tautological endomorphism $\mathcal{A} \in \mathcal{D}_D \subset \operatorname{End}(V_D)$ which preserves the pullback $\mathcal{V}_D = (V_D = \mathcal{V}_0 \supset \cdots \supset \mathcal{V}_m \supset 0)$ to D of the tautological flag on F. Writing $\mathcal{A}_i$ for the image of $\mathcal{A}$ in $\operatorname{End}(\mathcal{V}_{i-1}/\mathcal{V}_i)$ on D we see that $\mathcal{X}$ is the zero locus of the section

$$(\mathcal{A}_1 - \lambda_1, \ldots, \mathcal{A}_m - \lambda_m) \in H^0\left(D, \oplus_1^m (\operatorname{End}(\mathcal{V}_{i-1}/\mathcal{V}_i))_D\right).$$

Computing dimensions, we see that this section vanishes in codimension $= \operatorname{rank}(\oplus_1^m(\operatorname{End}(\mathcal{V}_{i-1}/\mathcal{V}_i))_D)$, so $\mathcal{X}$ is locally a complete intersection in D with normal bundle $\oplus_1^m(\operatorname{End}(\mathcal{V}_{i-1}/\mathcal{V}_i))_{\mathcal{X}}$.

By the adjunction formula (Hartshorne [**7**, Ch. 2]), the canonical bundle of $\mathcal{X}$ is the canonical bundle ω_D of D tensored with the highest exterior power of the normal bundle to $\mathcal{X}$ in D. But the highest exterior power of the endomorphism bundle of any bundle is trivial, so this is simply the restriction of ω_D to $\mathcal{X}$. It thus suffices to show that $\omega_D \cong \mathcal{O}_D$.

Now the cotangent bundle $\mathcal{T}^*(D)$ of D is easily seen to be the pullback from F to D of $\mathcal{D}^* \oplus \mathcal{T}^*(F)$. From the description of $\mathcal{T}^*(F)$ in terms of upper triangular matrices we get an exact sequence

$$0 \to \mathcal{T}^*(F) \to \mathcal{D} \to \oplus_1^{m+1} \operatorname{End}(\mathcal{V}_{i-1}/\mathcal{V}_i) \to 0.$$

From the triviality of the top exterior powers of the $\operatorname{End}(\mathcal{V}_{i-1}/\mathcal{V}_i)$, we thus see that the highest exterior power of $\mathcal{D}^* \oplus \mathcal{T}^*(F)$ is the same as that of $\mathcal{D}^* \oplus \mathcal{D}$. Since this is trivial we get the desired isomorphism $\omega_{\mathcal{X}} \cong \mathcal{O}_{\mathcal{X}}$.

From the vanishing theorem of Grauert-Riemenschneider [**6**] we now obtain

$$R^i\pi_{1*}\mathcal{O}_{\mathcal{X}} = R^i\pi_{1*}\omega_{\mathcal{X}} = 0 \text{ for } i > 0,$$

completing the proof of part i) of the Theorem. Using duality (see for example Elkik [4]) we also get $\omega_{X'_{rW}} = \mathcal{O}_{X'_{rW}}$, where we have written X' for $X'_{r,\mathbb{A}^m}$. Thus this last variety is Gorenstein.

ii) Since $X_{r,W}$ is the image of $\mathcal{X}_{r,W}$, it is irreducible. On the other hand π_1 is one to one over the image of the open set of those $(A, \lambda, \{V_i\})$ such that $(A - \lambda_i)V_{i-1} = V_i$ for $i = 1, \ldots, m$, so π_1 is birational. It is proper because F is a projective variety. In particular, $X'_{r,W}$ is the normalization of $X_{r,W}$.

Consider, for any linear varieties $U \subset W \subset \mathbb{A}^m$, the commutative diagram (of rings, since all the varieties in question are affine):

$$\begin{array}{ccc} \mathcal{O}_{X'_{r,W}} & \longrightarrow & \mathcal{O}_{X'_{r,U}} \\ \uparrow & & \uparrow \\ \mathcal{O}_{X_{r,W}} & \overset{\alpha}{\twoheadrightarrow} & \mathcal{O}_{X_{r,U}} \end{array}$$

The horizontal map on the bottom is an epimorphism since $X_{r,U}$, being a closed subset of $\operatorname{End}(V) \times W$ by the properness of $\pi_{1,W}$, is a closed subvariety of $X_{r,W}$.

Consider first the case $U = 0$. Since $\mathcal{O}_{X_{r,0}}$ is normal by the result of Section 1, the right hand vertical map is an epimorphism. We wish to prove that the left hand vertical map is an epimorphism. This will prove part ii) for spaces W containing 0.

Since the rings on the bottom of the diagram are the reduced rings of finitely generated graded rings, they are graded, and their normalizations on the top row of the diagram are graded too. Thus we may use Nakayama's lemma to conclude that the left hand vertical map is onto if we establish that $\mathcal{O}_{X'_{r,W}}$ modulo the kernel of α, or, even better, modulo $x_1, \ldots, x_m$, is $\mathcal{O}_{X'_{r,0}}$. This is the special case $W = 0$ of the first statement of part iii) of the Theorem, applied to the X' in place of the X. We will prove it now in this form for arbitrary W:

Inductively, it suffices to show that if $W \subset W'$ are linear subvarieties of $\mathbb{A}^m$ such that W is cut out in W' by a single equation $y \in \mathcal{O}_{W'}$, then the sequence

$$0 \to \pi_1 * \mathcal{O}_{\mathcal{X}_{r,W'}} \xrightarrow{y} \pi_1 * \mathcal{O}_{\mathcal{X}_{r,W'}} \longrightarrow \pi_1 * \mathcal{O}_{\mathcal{X}_{r,W}}$$

is right exact, and this follows at once from the last statement of part i), proved above. This completes the proof of ii) for spaces W containing 0.

To prove ii) for arbitrary linear varieties $U \subset \mathbb{A}^m$, consider the diagram above with $W = \mathbb{A}^m$, so that by what we have just proved, the left-hand vertical map is onto. By the version of the first part of iii) established above, the top horizontal map is onto. Of course this implies that the right-hand vertical map is onto, proving ii) in general.

iii) We now see that $X'_{r,W} = X_{r,W}$ for each W, so that the first part of iii), the restriction statement, is already proved. As for flatness, it suffices to treat the case $W = \mathbb{A}^m$. Since $\mathcal{O}_{X_{r,\mathbb{A}^m}}$ is graded, it suffices to prove flatness locally at $0 \in W$; that is, by the "local criterion" (see for example Matsumura [**11**]), it suffices to show that $x_1, \ldots, x_m$ is a regular sequence on $\mathcal{O}_{X_{r,\mathbb{A}^m}}$. But since $\mathcal{O}_{X_{r,W}}$ is a domain for every W, this follows from the restriction statement. This concludes the proof of Theorem 2.1. □

Proof of Corollary 2.2: All but the first statement comes directly from the special case $W = \{\lambda\}$ of the Theorem. By the Theorem, the fiber $X_{r,\lambda}$ is reduced, so to prove the first statement it is enough to check that the given variety is the image of $\mathcal{X}_{r,\lambda}$ set-theoretically. This follows easily by decomposing V into eigenspaces. □

Proof of Proposition 2.3: Smoothness at A follows from Zariski's Main Theorem since we know that X_r is normal and the fiber of the desingularization $\mathcal{X}_{r,0}$ over A consists of the single point corresponding to the flag $\{A^i V\}$.

We next show that the given set of endomorphisms is contained in the tangent space by showing that it is in the image of the tangent space of $\mathcal{X}_{r,0}$. For this it is enough to show that if $\varepsilon^2 = 0$ then the set $\{(A + \varepsilon\alpha)^i V_{\mathbb{C}[\varepsilon]}\}$

of subspaces of $V_{\mathbb{C}[\varepsilon]}$ is a $\mathbb{C}[\varepsilon]$-valued point of F over $\{A^iV\}$; it will follow that the pair

$$(A+\varepsilon\alpha, \{(A+\varepsilon\alpha)^i V_{\mathbb{C}[\varepsilon]}\})$$

is a $\mathbb{C}[\varepsilon]$-valued point of $\mathcal{X}_{r,0}$.

The $\mathbb{C}[\varepsilon]$-valued points of F over $\{A^kV\}$ are the sets $\{\mathcal{V}_k\}$ of subspaces of $\mathcal{V} := V_{\mathbb{C}[\varepsilon]}$ such that for each k, $\mathcal{V}_k$ reduces mod ε to V_k, $\mathcal{V}_k \supset \mathcal{V}_{k+1}$, and $\mathcal{V}_k$ is a direct summand of $\mathcal{V}$. The subspaces $(A+\varepsilon\alpha)^k\mathcal{V}$ obviously satisfy the first two conditions, and we must show that they satisfy the third. Since $\operatorname{rank} A^k = r(k)$ the $r(k)\times r(k)$ minors of $(A+\varepsilon\alpha)^k$ generate the unit ideal of $\mathbb{C}[\varepsilon]$, so we need only show that the $(r(k)+1)-(r(k)+1)$ minors are 0, or equivalently that every exterior product

$$(A+\varepsilon\alpha)^k v_1 \wedge \cdots \wedge (A+\varepsilon\alpha)^k v_{r(k)+1}$$

vanishes. Since $\operatorname{codim}_V \ker A^k = r(k)$, we may assume that one of the v_i, say $v_{r(k)+1}$, is in $\ker A^k$. Writing $(A+\varepsilon\alpha)^k = A^k + \varepsilon B$, with $B = \sum_{i+j=k-1} A^i\alpha A^j$ we see that such an exterior product consists of a sum of terms of the form

$$(*) \qquad \varepsilon \cdot A^k v_1 \wedge \cdots \wedge Bv_j \wedge \cdots \wedge A^k v_{r(k)+1},$$

each containing one factor with B in place of A^k. Of course all but the term with $j = r(k)+1$ vanishes. By our condition on α, $Bv_{r(k)+1} \in A^kV$, so the expression in (*) vanishes for $j = r(k)+1$ as well. This proves that α is in the tangent space to X_r at A.

To show that the given set of α is exactly the tangent space to X_r, it now suffices to show that it has at least the correct dimension; that is, that its codimension in $\operatorname{End}(V)$ is at most $\sum_i a_i(r)^2$. Inductively, suppose that the set

$$E_h = \{\alpha \in \operatorname{End}(V) | \sum_{i+j=k-1} A^i\alpha A^j \text{ maps } \ker(A^k) \text{ into } \operatorname{im}(A^k) \text{ for } k < h\}$$

has codimension $\leq \sum_{i<h} a_i(r)^2$ in $\operatorname{End}(V)$. The space E_{h+1} is defined in E_h by the condition that $\sum_{i+j=h} A^i\alpha A^j$ takes $\ker A^{h+1}$ into $\operatorname{im} A^{h+1}$. But for $\alpha \in E_h$ the transformation

$$\sum_{i+j=h} A^i\alpha A^j = A^h\alpha + \left(\sum_{i+j=h-1} A^i\alpha A^j\right) A$$

already takes $\ker A^h$ into $\operatorname{im} A^h$, and both the codimension of $\ker A^h$ in $\ker A^{h+1}$ and the codimension of $\operatorname{im} A^{h+1}$ in $\operatorname{im} A^h$ are equal to $a_{h+1}(r)$. Thus $\operatorname{codim}_{E_h} E_{h+1} \leq a_{h+1}(r)^2$, and we are done. □

Proof of Corollary 2.4: The given intersection is certainly equal to $X_{r,\lambda}$ set-theoretically, so it is enough to prove that it is reduced. By the dimension formula of Theorem 2.1, the intersection has codimension equal to the sum of the codimensions of the varieties being intersected. Since each of the varieties being intersected is Cohen-Macaulay, the scheme-theoretic intersection is too, so if it were not reduced, it would be everywhere non-reduced. Thus it is enough to show that the intersection is transverse at some one point.

It is easy to write down (for example by using the Jordan normal form) an endomorphism $A \in X_{r,\lambda}$ which satisfies $\operatorname{rank} A^i = r(i,j)$ for every appropriate i, j. The condition on an element $\alpha \in \operatorname{End}(V)$ for it to belong to the the tangent space at A to one of the varieties

$$\{A \in \operatorname{End}(V) \,|\, \operatorname{rank}(A - \lambda_i)^j \leq r(i,j) \text{ for all } j \text{ as above}\}_{\text{red}}$$

being intersected is, by Corollary 2.3, a condition only on the restriction of α to the kernel of a large power of $A - \lambda_i$. Since the direct sum of these kernels, for the different i, is embedded in V, such conditions are independant. Consequently the tangent spaces at A of the varieties being intersected meet properly, and the scheme-theoretic intersection is smooth at A as required. □

3. Equations.

Throughout this section we will make use of the notations V, n, r, a, b introduced in the beginning of Section 2. We will also write A_{gen} for the generic $n \times n$ matrix, defined over the polynomial ring $\mathbb{C}\big[\{x_{ij}\}_{1\leq i,j\leq n}\big]$, which we identify with the coordinate ring of $\operatorname{End}(V)$.

The action of $\operatorname{GL}(V)$ on $\operatorname{End}(V)$ by conjugation induces an action on the coordinate ring $\mathbb{C}[x_{ij}]$ of $\operatorname{End}(V)$. Since the variety X_r is invariant, it is natural to wish for equations for X_r that are invariant under $\operatorname{GL}(V)$. However, as is well-known, the ring of $\operatorname{GL}(V)$-invariants is generated by the polynomials in x_{ij} that are the coefficients of the characteristic polynomial of the generic matrix, so if one requires that the individual equations be invariant, one is doomed to failure (except in the case of the "largest" nilpotent orbit closure X_r, that for which $r(i) \equiv n-i$, where the equations are generated by all the coefficients of the characteristic polynomial!) There seems little reason to expect in advance anything more than that the whole ideal $I(X_r)$ of functions vanishing on X_r will be invariant.

But in fact a stronger invariance does hold. One can construct matrices of functions which are conjugation invariant as matrices, in a suitable sense, and the entries of such a matrix generate a conjugation-invariant ideal; the ideals of functions vanishing on the X_r turn out to be sums of ideals constructed in this way.

Powers, exterior powers, and powers of exterior powers of the generic matrix A_{gen} are all examples of matrices F of functions which are conjugation invariant in a sufficiently strong sense: namely for some t and all endomorphisms A of V, $F(A)$ acts naturally on $\bigwedge^t V$; and for any invertible matrix B,

$$F(B^{-1}AB) = (\bigwedge^t B)^{-1} F(A)(\bigwedge^t B).$$

De Concini and Procesi [**3**] noticed that there is a very general method of making such constructions, and this method seems to provide enough invariant ideals to define all the nilpotent X_r.

Tanisaki [**15**], following the work of De Concini and Procesi, found a more compact and convenient set of invariant matrices of functions to use: Let $A \in \text{End}(V)$. Regarding $\bigwedge^t(x-A)$ as a polynomial in x whose coefficients are endomorphisms of $\bigwedge^t V$, we define $\lambda_d^t(A) \in \text{End}(\bigwedge^t V)$ to be the coefficient of x^{t-d}. (Note that if $d=0$ this is the identity map, while if $d>t$ it is the 0 map.) The index d in the notation $\lambda_d^t(A)$ is chosen in this way because if we choose a basis for V, so that we may regard A and $\lambda_d^t(A)$ as matrices, the entries of $\lambda_d^t(A)$ are polynomials of degree d in the entries of A. In particular,

$$\lambda_d^t := \lambda_d^t(A_{\text{gen}})$$

is a matrix of polynomial functions of degree d. We write $I(\lambda_d^t)$ for the ideal generated by the entries of the matrix λ_d^t, and $V(\lambda_d^t)$ for the (reduced) variety in $\mathbb{A}^{n^2}$ that they define.

For example, if $d=t$ then we get

$$\lambda_t^t = \bigwedge^t A_{\text{gen}},$$

so the $I(\lambda_t^t)$ are the usual determinantal ideals. On the other hand, λ_d^n is the 1×1 matrix whose entry is (up to sign) the d-th coefficient of the characteristic polynomial.

In the nilpotent case, where $r(\infty) = 0$, X_r is the closure of a single orbit, so a collection of functions which arise as the entries of an invariant matrix has a nice property: to check its vanishing on the whole orbit closure, it is enough to check its vanishing at a single element of the orbit. In this case, Tanisaki showed that the vanishing of certain of the λ_d^t defines the variety X_r set-theoretically (and a little more); De Concini and Procesi had previously proved the corresponding results for their matrices.

In the non-nilpotent case the X_r are no longer closures of single conjugation orbits, and it is no longer true that if an invariant matrix of functions — even an invariant function like a coefficient of the characteristic polynomial — vanishes on an element of $X_r - \bigcup_{r'<r} X_{r'}$, then it vanishes on all of X_r. For example, taking $r(i) \equiv n$, the constant fuction, so that $X_r = \mathrm{End}(V)$, the trace function vanishes on lots of invertible matrices without vanishing everywhere. But one can hope for ideals of functions with a still stronger invariance property. To describe what is desirable, we will write A', for the restriction of A to its own 0-eigenspace, so that A' is the "nilpotent part" of A. Since the membership of an endomorphism A in one of the X_r can be determined from A', what one hopes for are conjugation invariant matrices of functions in the entries of A, or collections of such matrices, whose vanishing depends only on A'. Tanisaki's construction realizes this hope in a rather strong form.

It follows from an easy argument of Tanisaki that for any t and d the ideal generated by the entries of the collection of matrices $\lambda_d^t, \lambda_{d+1}^t, \ldots, \lambda_t^t$ has the strong invariance property just described. But in fact — this is perhaps the main new point in the discussion below — if $t < n$ then the vanishing of an individual λ_d^t on $A \in \mathrm{End}(V)$ depends only on A' (Theorem 3.1). Thus the λ_d^t are a natural source of equations for defining the varieties X_r for arbitrary r. We prove here that they suffice set-theoretically (Corollary 3.3), and as mentioned in the introduction, Weyman [**16**] has proved that suitable $I(\lambda_d^t)$ actually suffice to generate the ideal of X_r. (Weyman's result was proved for some special rank functions r by Strickland [**14**], and a preliminary version of this paper contained a proof in some further special cases.)

The vanishing loci of the λ_d^t are interesting varieties in their own right.

It turns out that they are unions of the X_r in all but the familiar case of λ_d^n (Corollary 3.4). In Section 4 we shall see that $V(\lambda_{d-1}^{n-1})$ arises as the singular locus of the hypersurface $V(\lambda_d^n)$.

Here is our main technical result, which give the invariance result in an explicit form:

THEOREM 3.1. *Suppose* rank $A^n = \rho$ *and let* A' *be the nilpotent part of* A. *If* $\rho < d$, *then* $\lambda_d^t(A) = 0$ *iff* $\lambda_{d-\rho}^{t-\rho}(A') = 0$. *If* $t < n$ *then in addition* $\lambda_d^t(A) = 0$ *implies* $\rho < d$.

The example of $\lambda_1^n(A)$, the trace of A, shows, as remarked above, how badly this fails for $t = n$.

We postpone the proofs until after all the results have been stated.

Theorem 3.1 allows us to analyze the varieties X_r and $V(\lambda_d^t)$ in terms of each other, at least set-theoretically. First we give a formula telling exactly which $V(\lambda_d^t)$ contain which X_r. Some of these containments follow from others, as the second statement of the result shows. Here and later in this section we will use the notations $a_i(r)$ and $b_i(r)$ introduced at the beginning of Section 2:

THEOREM 3.2. $V(\lambda_d^t) \supseteqq X_r$ *iff* $d > t - \sum_{i>n-t} b_i(r)$. *If* $e \geq d$ *and* $s \geq t$ *then* $V(\lambda_e^s) \supseteqq V(\lambda_d^t)$ *unless* $s = t = n$.

The inclusions in Theorem 3.2 are inclusions of sets — that is, reduced schemes, and the X_r are indeed defined as such. However, we could also regard $V(\lambda_d^t)$ as a scheme, defined by the ideal $I(\lambda_d^t)$ generated by the entries of the matrix λ_d^t. Thus we may ask whether the assertions of Theorem 3.2 hold scheme-theoretically — that, whether the corresponding inclusions hold among the ideals $I(\lambda_d^t)$.

It is immediate from the definitions that Theorem 3.2 holds scheme-theoretically for $d = 1$ and $d = t$, and Weyman's result implies that the first statement holds scheme-theoretically, as already mentioned. Also, we have been able to prove (using the description of the λ_d^t given in part A) that the second statement is true scheme-theoretically for $d = e$. We have checked that these statements are true in all cases with $n \leq 6$ using the computer program Macaulay of Bayer and Stillman [**1**], and we thus conjecture that they hold scheme-theoretically in all cases.

COROLLARY 3.3. $X_r = \bigcap V(\lambda_d^t)$, *the intersection being taken over all those* *t and d such that $b_{n-t+1}(r) > 0$ and*

$$t < n, \quad d = t + 1 - \sum_{i>n-t} b_i(r)$$

or

$$t = n, \quad r(\infty) + b_1(r) > d > t - \sum_{i>n-t} b_i(r).$$

Tanisaki proved the weaker result that the above formula holds if the intersection is extended over the (larger) set of indices with

$$d > t - \sum_{i>n-t} b_i(r).$$

PROOF (Tanisaki): The formula is obvious if one first diagonalizes $x - A$; that is, it suffices to consider the elementary divisors of $x - A$. □

Our result by contrast cannot be treated simply by diagonalizing the matrix $x - A$ over the ring $\mathbb{C}[x]$, because this process mixes together the ideal $I(\lambda_d^t)$ with $I(\lambda_e^t)$ for $e > d$. But by Theorem 3.2 the radical of $I(\lambda_d^t)$ contains $I(\lambda_e^t)$, and modulo this fact the results are equivalent.

Each λ_d^n is a 1×1 matrix whose entry is an irreducible polynomial, not a rank variety except for $d = n$. However for $t < n$ the irreducible components of the $V(\lambda_d^t)$ are rank varieties:

COROLLARY 3.4. *If $t < n$ then $V(\lambda_d^t)$ is the union, for $1 \leq k \leq 1 + [(d-1)/(n-t)]$, of the rank varieties*

$$\{A \mid \operatorname{rank} A^k \leq \min\bigl(n - k(n-t+1), d - 1 - (k-1)(n-t)\bigr)\}.$$

We derive:

COROLLARY 3.5. *The following conditions are equivalent if $t < n$:*

i) *$V(\lambda_d^t)$ is irreducible.*
ii) $V(\lambda_d^t) = V(\lambda_d^d) = \{A \mid \operatorname{rank} A \leq d - 1\}$.
iii) *$d = t$ or $t + d \leq n$.*

REMARK: The ideal $I(\lambda_d^t)$ is given with $({}_nC_t)^2$ generating polynomials (where as usual ${}_nC_t$ denotes the binomial coefficient $n!/(n-t)!\,t!$). It is easy to see that they are each expressible as a linear combination of the $d \times d$ minors of the generic matrix, and thus at most $({}_nC_d)^2$ of them are linearly independent. Now the ${}_nC_d$ minors of the generic matrix are linearly independent (as one sees at once from the fact that one can construct a matrix of numbers with any given minor nonzero, and all other minors zero) so if ${}_nC_t < {}_nC_d$ then $I(\lambda_d^t) \subsetneqq I(\lambda_d^d)$. In this situation, Corollary 3.5 says that the two ideals even have different radicals.

On the other hand, $t+d \leq n$ iff ${}_nC_t \geq {}_nC_d$,and in this case the Corollary says that the two ideals have the same radical; but in fact, in a number of cases that we have checked by computer (with $n \leq 6$) the two ideals are actually equal when $t+d \leq n$. It would be interesting to know whether the span of the entries of λ_d^t in the space of all forms of degree d has dimension $= \min\big(({}_nC_t)^2, ({}_nC_d)^2\big)$, the "largest possible" value, in every case.

Our next goal is the proof of Theorems 3.1 and 3.2. We first show how to compute the minors of $x - A$ if A is in Jordan normal form. (Since we are interested in the coeffiecients of the different powers of x that appear, we cannot simply diagonalize $x - A$ over $\mathbb{C}[x]$; this would of course preserve the ideal in $\mathbb{C}[x]$ generated by the minors, but would not preserve the $\mathbb{C}$-vectorspace spanned by the minors.)

Assume, then, that A is in Jordan form; that is, decompose V as a direct sum $V = \oplus V_\alpha$ in such a way that $AV_\alpha \subset V_\alpha$ and $(A - \lambda_\alpha)\big|_{V_\alpha}$ is nilpotent of index equal to the dimension of V_α for each α. Choose bases $\mathcal{B}_\alpha$ for each V_α so that $A - \lambda_\alpha$ restricted to V_α becomes upper triangular with ones on the superdiagonal and zeros elsewhere, and let $\mathcal{B} = \cup \mathcal{B}_\alpha$. For subsets I and J of $\mathcal{B}$, we let $d(I, J) = \det_{I,J}(x - A)$ be the corresponding minor of $x - A$ (or 0 if I and J do not have the same cardinality). Writing $I_\alpha = I \cap \mathcal{B}_\alpha$ and $J_\alpha = J \cap \mathcal{B}_\alpha$, we obviously have $d(I, J) = \Pi_\alpha d(I_\alpha, J_\alpha)$ if the cardinality of I_α is the same as that of J_α for each α.

If I and J have the same cardinality, we let $\varphi : I \to J$ be the unique order preserving map from I to J, and set

$$e_\alpha = \operatorname{card}\{i \in I_\alpha \mid \varphi(i) = i\}.$$

We have

LEMMA 3.6. *With notation as above, suppose that I and J are t element subsets of $\mathcal{B}$. The minor $d(I, J)$ is zero iff $\varphi I_\alpha \neq J_\alpha$ for some α or $\varphi(i) \neq i, i+1$ for some i. Else*

$$d(I, J) = \Pi(x - \lambda_\alpha)^{e_\alpha}.$$

Thus if A consists of a single Jordan block, with eigenvalue λ, then the set of possible values of $d(I,J)$ is given by

$$\{d(I,J)\} = \{(x-\lambda)^e\}$$

where $e = t$ if $t = n$, and else e takes on all values from 0 to t. If A is nilpotent, with Jordan blocks of size $b_i(r)$, then

$$\lambda_d^t(A) = 0 \text{ iff } t - d < \sum_{i>n-t} b_i(r).$$

PROOF OF LEMMA 3.6: We first derive the formula for $d(I,J)$. Since $d(I,J) = \Pi_\alpha d(I_\alpha, J_\alpha)$ we see that $d(I,J)$ can only be nonzero if I_α and J_α have the same number of elements for each α, which is equivalent to the statement that $\varphi(I_\alpha) = \varphi(J_\alpha)$ for every α. Thus we may assume that this condition is satisfied. Let i be the smallest index occurring in a block I_α, and let $I' = I - \{i\}$, $J' = J - \{\varphi(i)\}$. If $\varphi(i) \neq i$, $i+1$, then the row corresponding to i in the determinant $\det_{I,J}(x-A)$ is zero, so $d(I,J) = 0$. If $\varphi(i) = i+1$, then this row has a 1 in the i-th column and zeros elsewhere, so $d(I,J) = d(I',J')$. If $\varphi(i) = i$ then the i-th column in the determinant $\det_{I,J}(x-A)$ has $x - \lambda_a$ in the i-th row and zeros elsewhere, so $d(I,J) = (x-\lambda_\alpha)\, d(I',J')$. In either case, since $\varphi\,\big|_{I'}$ is again the unique order preserving function from I' to J', we are done by induction.

It is clear in the one block case that $d(I,J)$ cannot take on any other values than those indicated. But if $0 \le e \le t < n$ then the choice

$$\begin{aligned} I &= \{1, 2, \ldots, t\}, \\ J &= \{1, \ldots, e, e+2, \ldots, t+1\}, \end{aligned}$$

yields $d(I,J) = (x-\lambda)^e$.

Finally, in the nilpotent case, a choice of t rows I, that is, t elements of $\mathcal{B}$, excludes $n-t$ rows, and thus can exclude the elements of at most $n-t$ distinct $\mathcal{B}_\alpha$. But each $\mathcal{B}_\alpha$ which is entirely included contributes a factor of x^{b_α} to the determinants $d(I,-)$. Thus each of these determinants is divisible by the $\sum_{i>n-t} b_i(r)$ power of x, and from the one-block case it is clear that any power of x larger than this one, up to the t-th power, can be so obtained. Since by definition $\lambda_d^t(A) = 0$ iff the $(t-d)$-th power of x does not occur, we are done. □

PROOF OF THEOREM 3.2: If $t = n$ then $t - \sum_{i>n-t} b_i(r) = r(\infty)$. In this case the first statement of the theorem is elementary, while the second is vacuous.

We may thus assume $t < n$. We will first prove that the stated inclusions hold under the additional assumptions $r(\infty) = 0$, and only among the intersections of the given varieties with the set of nilpotent transformations. We will use this nilpotent case, below, to prove Theorem 3.1. But Theorem 3.1 shows that the general case of Theorem 3.2 follows from the nilpotent case of Theorem 3.2, so the proof of Theorem 3.2 will be completed when we complete the proof of Theorem 3.1.

Assume, then that $r(\infty) = 0$, and let A be a generic point of X_r, so that A is nilpotent. Of course, since the X_r are just the orbit closures in the nilpotent case, A can have any nilpotent Jordan form. The first statement of the Theorem in this case follows at once from the last statement of Lemma 3.6.

To prove the second statement of the Theorem in the nilpotent case — that is, with $V(\lambda_e^s)$ and $V(\lambda_d^t)$ replaced by their intersections with the set of nilpotent matrices — we use the fact that the X_r are orbit closures. Using the first statement of the Theorem, it is enough to show that $t \geq d > t - \sum_{i>n-t} b_i(r)$ implies $e > s - \sum_{i>n-s} b_i(r)$ in relevant cases. By induction we may further assume that e and s differ from d and t by at most 1. The cases $s = t$, $e = d + 1$ and $s = t + 1$, $e = d + 1$ are innocuous. If $s = t + 1$, $e = d$ then problems can only arise if $b_{n-t}(r) = 0$, in which case also $\sum_{i>n-t} b_i(r) = 0$, so $d > t$, a contradiction. This concludes the proof in the nilpotent case. □

PROOF OF THEOREM 3.1: First suppose that $\rho < d$. In this case we must show that $\lambda_d^t(A) \neq 0$ iff $\lambda_{d-\rho}^{t-\rho}(A') \neq 0$. Since A' is nilpotent, we may apply the last statement of Lemma 3.6 and obtain

$$\lambda_{d-\rho}^{t-\rho}(A') \neq 0$$

iff

$$d + \rho \leq (t + \rho) - \sum_{i>(n-\rho)-(t-\rho)} b_i(r),$$

that is, iff

$$d \leq t - \sum_{i>n-t} b_i(r).$$

Let A'' be the part of A with eigenvalues $\neq 0$, so that $A = A' \oplus A''$ and A'' is a $\rho \times \rho$ matrix of rank ρ. We may assume that A' is in Jordan normal form. The coefficient of x^0 in $\det(x - A'')$ is $\pm \det A''$, which is nonzero, and since by the nilpotent case of Theorem 3.2 and the above the $(t-\rho) \times (t-\rho)$ minors of $x - A'$ are the powers of x from $t - \sum_{i>n-t} b_i(r)$ to t, we see that the $t \times t$ minors of A containing all the rows and columns of A'' already have the powers of x in this range occurring with nonzero coefficient. Thus $\lambda^{t-\rho}_{d-\rho}(A') \neq 0$ implies $\lambda^t_d(A) \neq 0$. On the other hand, any $t \times t$ minor $d(I, J)$ of $x - A$ involves at least $t - \rho$ rows and columns of A', and thus has as a factor a minor of A' of size $s \geq t - \rho$. Applying the last statement of Lemma 3.6 to this minor, we see that it is divisible by x to the power

$$\sum_{i>n-\rho-s} b_i(r) \geq \sum_{i>n-t} b_i(r),$$

finishing the proof in the case $\rho < d$.

To finish the proof we must show that if $t < n$ and $\lambda^t_d(A) = 0$ then $\rho < d$. From Lemma 3.6, whose notation we adopt, we see that because $t < n$, we can find among the entries of $\bigwedge^t(x - A)$ polynomials of degree t of the form

$$p_\beta(x) = q(x)/(x - \lambda_\beta),$$

where

$$q(x) = \Pi(x - \lambda_\alpha)^{e_\alpha}$$

the λ_α running over all the eigenvalues (or any $t+1$ of them, if there are that many) and λ_β running over all the λ_α. To say that $\lambda^t_d(A) = 0$ thus implies that the d-th elementary symmetric function in t variables vanishes when applied to any t of the roots of q. In this situation the following Lemma, applied with $n = t + 1$, shows that at least $n - d + 1$ of the eigenvalues of A (counted with multiplicity) are 0; that is, $n - \rho \geq n - d + 1$, yielding $\rho < d$ as desired. □

Write $\sigma_d(x_1, \ldots, x_t)$ for the d-th elementary symmetric function in t variables. Since $\sigma_d(x_1, \ldots, x_t)$ is the sum of all multilinear monomials of degree d, it vanishes if at least $t - d + 1$ of the variables are 0. The converse of this is of course false except for $t = d$; however, if more than t variables are involved, a sort of converse is true:

LEMMA 3.8. *If $1 \leq d \leq t < n$ are integers, and $\lambda_1, \ldots, \lambda_n$ are complex numbers such that the elementary symmetric function $\sigma_d(x_1, \ldots, x_t)$ vanishes on every set of t of the λ_i, then at least $n - d + 1$ of the λ_i are 0.*

PROOF OF LEMMA 3.8: If $d = t$ the result is obvious. If some $\lambda_i = 0$ then since

$$\sigma_d(x_1, \ldots, x_{t-1}, 0) = \sigma_d(x_1, \ldots, x_{t-1}),$$

the result follows by induction on n and t. Thus we may suppose that $\sigma_d(x_1, \ldots, x_t)$ vanishes on every set of t of the λ_i, but all the λ_i are nonzero, and we will derive a contradiction.

Note first that the sum $\sum_j \sigma_d(\lambda_{i_1}, \ldots, \hat{\lambda}_{i_j}, \ldots, \lambda_{i_{t+1}})$, leaving out one index at a time, is a positive integral multiple of $\sigma_d(\lambda_{i_1}, \ldots, \lambda_{i_{t+1}})$, so $\sigma_d(x_1, \ldots, x_{t+1})$ vanishes on every $t+1$ element subset of the λ_i. On the other hand,

$$\sigma_d(x_1, \ldots, x_{t+1}) = x_{t+1}\sigma_{d-1}(x_1, \ldots, x_t) + \sigma_d(x_1, \ldots, x_t),$$

(This formula also works for $d = 1$ if we interpret $\sigma_0(x_1, \ldots, x_t)$ as the constant function equal to 1.) Since by our hypothesis all the λ_i are nonzero, we see that $\sigma_{d-1}(x_1, \ldots, x_t)$ vanishes on every t-element subset of the λ_i. Continuing, we see that $\sigma_0(x_1, \ldots, x_t)$ vanishes on every t-element subset of the λ_i, which is the desired contradiction. □

As already remarked, Corollary 3.3 follows easily from the other results.

PROOF OF COROLLARY 3.4: The rank varieties X_r with $r(\infty) = 0$ are the minimal closed conjugation-invariant subvarieties of the set of nilpotent matrices. Since taking the nilpotent part commutes with conjugation, and the $V(\lambda_d^t)$ are invariant, it follows from Theorem 3.1 that each $V(\lambda_d^t)$ is the union of those X_r that are contained in it. It thus suffices to show that the maximal X_r contained in a given $V(\lambda_d^t)$ are among those on the list given in the Corollary.

From Theorem 3.2 we know that $X_r \subset V(\lambda_d^t)$ iff

$$d > t - \sum_{i > n-t} b_i(r) = t - \sum_{i \geq 1} (a_i(r) - (n-t))_+$$

where we have written $(s)_+$ for $\max(0, s)$. But we can make X_r larger without altering the validity of this inequality by replacing r with a rank function for which the $a_i(r)$ that are $\leq n-t$ are replaced by 0. Suppose that $a_k(r) > n-t$, but $a_{k+1}(r) = 0$, so that in particular

$$r(k) \leq n - k(n-t+1).$$

The condition given above then becomes

$$\begin{aligned} d &> t - \sum_1^k (a_i(r) - (n-t)) \\ &= t + k(n-t) - \sum_1^k a_i(r) \\ &= t + k(n-t) - (n - r(k)) \\ &= (k-1)(n-t) + r(k), \end{aligned}$$

that is,

$$r(k) < d - (k-1)(n-t).$$

But any X_r satisfying the above inequalities is contained in

$$\{A \mid \operatorname{rank} A^k \leq \min(n - k(n-t+1), d - 1 - (k-1)(n-t))\},$$

so this is the largest rank variety given our choice of k.

If the inequality on k in the Corollary is violated, then

$$d - 1 - (k-1)(n-t) < 0,$$

so the above rank variety is empty and can be dropped from the union.

□

In order to prove Corollary 3.5, we will need to know the rank of a generic member of the variety $\{A \mid \operatorname{rank} A^k \leq s\}$.

LEMMA 3.9. *A generic menber of the variety* $\{A \mid \operatorname{rank} A^k \leq s\}$ *has* rank $=$ $n - a_1$ *where* a_1 *is the smallest integer* $\geq (n-s)/k$.

PROOF OF LEMMA 3.9: If A is such a generic member, then since

$$\begin{aligned} \operatorname{rank} A^{i-1} - \operatorname{rank} A^i &= \dim(\ker A) \cap (\operatorname{im} A^{i-1}) \\ &\leq \dim \ker A, \end{aligned}$$

we have $s = \operatorname{rank} A^k \geq n - k(\dim \ker A)$, so $\dim \ker A \geq a_1$. On the other hand it is easy to construct elements of the variety with $\dim \ker A = a_1$.

□

PROOF OF COROLLARY 3.5: Let

$$V_k = \{A \mid \operatorname{rank} A^k \leq \min(n - k(n-t+1), d - 1 - (k-1)(n-t))\}.$$

The generic element of $V_1 = V(\lambda_d^d)$ is an endomorphism of stable rank $d-1$, which is therefore not an element of any of the other V_k. This proves the equivalence of i) and ii). We also see that $V(\lambda_t^d)$ is irreducible iff V_1 contains all the other V_k with $k < (d-1)/(n-t)$, and this will be true iff the generic element of V_k has rank at most $d-1$. This rank of a generic element is computed for us by Lemma 3.9, and straightforward arithmetic leads to the desired result. □

REMARK: In Section 2 we saw that the variety $X_r = X_{r,0}$ could be realized as the flat limit of a family of varieties $X_{r,\lambda}$. Further, Corollary 2.4 shows that the equations of $X_{r,\lambda}$ are known if one knows the equations of $X_{r'}$ for certain $r' > r$. This gives an effective method of calculating the ideal of X_r (for example by computer) in any particular case.

4. Varieties defined by the coefficients of the characteristic polynomial.

A classical example of the construction λ_k^t of Section 3 occurs when $t = n$ as λ_k^n is a 1×1 matrix (i.e. a polynomial) such that $\lambda_k^n(A)$ is the coefficient of x^{n-k} in the characteristic polynomial of A. In this section we will consider the singular locus (Theorem 4.10) and its components (Corollary 4.14) of the variety of matrices defined by $\lambda_k^n = 0$. For simplicity, we will often write λ_k^n as λ_k. We let x_{ij} be the variable corresponding to the i, j entry of our square matrices, and consider λ_k as a polynomial in the x_{ij}'s. Occasionally, we will work basis free and consider $A \in \mathrm{End}_F(W)$. As λ_k is PGL_n invariant, we can unambiguously write $\lambda_k(A)$. Finally, λ_1 is just the linear, trace polynomial and so its zero set is trivial. We will henceforth assume that $k > 1$.

To begin with, we mention an alternative description of λ_k. If S, $T \leq \{l, \ldots, n\}$ are subsets of equal order, and $A \in M_n(F)$, denote by $A(S;T)$ the submatrix obtained by "crossing out" the rows in S and the columns in T. A *principal* submatrix is a submatrix of the form $A(S;S)$. The following is well known.

LEMMA 4.1. *$\lambda_k(A)$ is the trace of $\Lambda^k(A)$, which is the sum of the determinants of the principal $k \times k$ submatrices of A.*

To prove 4.1, one reduces to diagonal matrices using the Zariski density of diagonalizable matrices, and then computes. It follows from 4.1 or direct computation that λ_k is a multilinear polynomial.

LEMMA 4.2. *λ_k is an irreducible polynomial.*

PROOF: Write $\lambda_k = x_{11}q + r$ where x_{11} does not appear in q or r. Of course, λ_k is irreducible if and only if q and r have no common factors. If $X = (x_{ij})$ is the generic matrix, let $X' = X(1;1)$ be the submatrix. By inspection, q is the coefficient of x^{n-k} in the characteristic polynomial of X'. By induction (and the trivial trace polynomial case), we may assume that q is irreducible. Thus it suffices to show that q does not divide r, that is, that q does not divide λ_k. It is easy to construct a diagonal A with $q(A) = 0$ but $\lambda_k(A) \neq 0$, so the lemma is proved. □

Let $W_k \subseteq M_n(F)$ be the zero set of λ_k. As λ_k is irreducible, W_k is irreducible and is scheme theoretically defined by $\lambda_k = 0$. Let $S_k = \text{Sing}\ (W_k)$ be the reduced subvariety consisting of the singular points of W_k. As F has characteristic 0, S_k is the zero set of the n^2 partial derivatives of λ_k. In order to express these partials, we introduce the following notation. If $S \subseteq \{1, 2, \ldots, n\}$ is a subset, consider the complementary set $\{1, \ldots, n\} - S$ and order this complementary set as usual. For $i \in S$, define $i' = 0_S(i)$ if i is in the i' position in $\{1, \ldots, n\} - S$ (lowest first). Note that if $X' = X(S;T)$ is a submatrix of the generic matrix, $i' = 0_S(i)$, and $j' = 0_T(j)$ then x_{ij} is the i', j' entry of X'. If $S = T$, *i.e.* if X' is principal, then $0_S(i) + 0_S(j)$ is congruent to $i + j$ modulo 2.

The proof of the next result is a calculus exercise.

LEMMA 4.3.

$$\frac{\partial \lambda_k}{\partial x_{ij}} = \sum_{S \in \Gamma} (-1)^{i+j} \det\bigl(X(S \cup \{i\}; S \cup \{j\})\bigr)$$

where Γ is the set of subsets of $\{1, 2, \ldots, n\}$ with cardinality $n - k$ and not containing i and j.

We begin our investigation of S_k by examining the restriction of the polynomials $\partial \lambda_k / \partial x_{ij}$ to upper triangular matrices. Recall from 3.8 that $\sigma_k(y_1, \ldots, y_s)$ is the k degree elementary symmetric function in $y_1, \ldots, y_s$. An immediate consequence of 4.3 is:

LEMMA 4.4. *Suppose $A \in M_n(F)$ is upper triangular with diagonal entries $a_{11}, a_{22}, \ldots, a_{nn}$. Then $\partial \lambda_k / \partial x_{jj}(A) = \sigma_{k-1}(a_{11}, a_{22}, \ldots, \hat{a}_{jj}, \ldots, a_{nn})$.*

COROLLARY 4.5. *If $A \in S_k$ and d is the stable rank of A then $d \leq k - 2$.*

PROOF: Since S_k and the stable rank are invariant under conjugation, we may assume that A is upper triangular. By 3.8 and 4.4, A has at least $n-k+2$ zeroes on its diagonal. The result follows. $\square$

To derive more information about the elements of S_k we must make use of the equations $\partial\lambda_k/\partial x_{ij} = 0$ for $i \neq j$. If A is upper triangular, half of these equations automaticaly hold.

LEMMA 4.6. *Let A be upper triangular. Then $\partial\lambda_k/\partial x_{ij}(A) = 0$ if $j > i$.*

PROOF: If $A = (a_{ij})$ then by assumption $a_{ij} = 0$ for $i > j$. Let $B = A(S;S)$ be any principal $k \times k$ submatrix with $i, j \notin S$. It suffices to show that $\det(B(S \cup \{i\};\ S \cup \{j\} = 0$. If $B' = B(S \cup \{i\};\ S \cup \{j\}) = (b_{ij})$ then a little work shows that B' is upper triangular with $b_{mm} = 0$ for $0_S(i) \leq m \leq 0_S(j)$. $\square$

Once again, S_k is invariant under conjugation and so it suffices to describe the $A \in S_k$ which are in Jordan canonical form. So for the next computation assume that A has the form:

$$\begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix} \tag{1}$$

where B,C are in Jordan canonical form, B is nilpotent, and C is nonsingular.

LEMMA 4.7. *Let A have the form (1) and suppose $i > j$. Then* $\det(A(\{i\};\{j\})) = 0$ *unless B is a single Jordan block and the i, j position in A is the lower left corner of B. Under these circumstances,* $\det(A(\{i\};\{j\})) = \det(C) \neq 0$.

PROOF: B has a zero row containing the bottom row of each Jordan block, and a zero column containing the leftmost column of each Jordan block. Thus if B has at least 2 Jordan blocks, $A(\{i\};\{j\})$ has a zero row and zero determinant. If B is a single Jordan block, and $\det(A(\{i\};\{j\})) = 0$, $A(\{i\};\{j\})$ must contain neither the first column nor the last row of B. Hence i, j is the position claimed. For such a B and i, j, $A(\{i\};\{j\})$ has the form:

$$\begin{pmatrix} I & 0 \\ 0 & C \end{pmatrix}$$

and the result is clear. $\square$

Given that A has form (1), the next lemma describes the principal submatrices of A. Its proof is both straightforward and elementary, and so is left to the reader.

LEMMA 4.8. *Let A be as in (1) and let D be a principal $k \times k$ submatrix of A.*

a) *D has the form:*

$$\begin{pmatrix} B' & 0 \\ 0 & C' \end{pmatrix}$$

where B', C' are in Jordan normal form, B' is nilpotent, and C' is nonsingular.

b) *A row or column of B' comes from one in B. Similarly for C and C'.*

c) *The maximum size of a Jordan block in B' is less than or equal to the maximum size of a Jordan block in B.*

We are ready to state and prove the key result, which describes the matrices in form (1) which lie in S_k.

PROPOSITION 4.9. *Suppose $A \in M_n(F)$ has the form (1), and that within B, the size of the Jordan blocks is nondecreasing as you go from top left to bottom right. Let d be the stable rank of A. The following are equivalent:*

a) *$A \in S_k$.*

b) *$d \le k-2$ and all Jordan blocks in B have size less than or equal to $k-d-1$.*

c) *All principal $k \times k$ submatrices of A have rank less than or equal to $k-2$.*

PROOF: a) $\Rightarrow$ b): Assume a). By 4.5, $d \le k-2$. Assume to the contrary that B has a Jordan block of size $r > k-d$. We can choose this block to be the one in the lower right corner of B. Let $i = n-d$, so B is an $i \times i$ matrix. Set $j = i-(k-d-1) = n-k+1$. Let $D = A(S;S)$ be a principal $k \times k$ submatrix of A with i, $j \in S$, and set $i' = 0_S(i)$, $j' = 0_S(j)$. Set $D' = D(\{i\};\{j\})$. Assume $\det(D') \neq 0$. By 4.7 and 4.8, i', j', must be the position in D in the lower left hand corner of E where $D =$

$$\begin{pmatrix} E & 0 \\ 0 & G \end{pmatrix} \tag{2}$$

and where E, G are in Jordan canonical form, E is nilpotent, and G is nonsingular. Since $j' = 1$; $1, 2, \ldots, j-1 \in S$. Since $j-1 = n-k$, $S = \{1, \ldots, j-1\}$, $G = C$, and $\det(D') = \det(C)$. Conversely if $S = \{1, \ldots, j-1\}$, E is a single Jordan block, and $\det(D(\{i'\};\{j'\}) = \det(C) \neq$

0. In other words, in the sum of 4.3 for $\partial\lambda_k/\partial x_{ij}(A)$, all terms but one are zero and that term is $\pm\det(C) \neq 0$. This contradicts our assumption that $\partial\lambda_k/\lambda x_{ij}(A) = 0$, and b) is proved.

b) $\Rightarrow$ c): Assume b). If D is any principal $k \times k$ submatrix of A, write D as in (2). Since G has size less than or equal to d, E has size greater than or equal to $k - d$. By 4.8, E contains at least 2 Jordan blocks, and so D has rank less than or equal to $k - 2$.

c) $\Rightarrow$ a): Let D be a principal $k \times k$ submatrix of A and D' a $k-1 \times k-1$ submatrix of D. Part c) says precisely that all such D' have determinant zero. In the sum of 4.3 for $\partial\lambda_k/\partial x_{ij}(A)$, all terms have the form $\pm\det(D')$ for such D'. Thus c) implies that $\partial\lambda_k/\partial x_{ij}(A) = 0$ for all i, j. That is, $A \in S_k$. □

We are ready to give our description of S_k.

THEOREM 4.10. *Suppose $A \in M_n(F)$ has stable rank d. Then $A \in S_k$ if and only if $d \leq k - 2$ and rank $(A^{k-d-1}) = d$.*

PROOF: As both sides of the equivalence are conjugation invariant, we may assume that A is in the form (1). Thus rank $(A^r) = \text{rank}(B^r) + \text{rank}(C^r)\,\text{rank}(B^r) + d$. Now part b) of 4.9 is equivalent to $d \leq k - 2$ and $k - d - 1 = \text{rank}(B^{k-d-1}) = \text{rank}(A^{k-d-1}) - d$. □

Let V_d^k be the variety $\{A | \text{rank}(A^{k-d-1}) \leq d\}$. By the observation in the introduction, $V_d^k = \{A \mid \text{rank}(A^i) \leq r(i)\}$ for some rank function r. We next explicitly determine the rank function r.

Since $V_d^k = X_r$ is irreducible, r is the rank function of a generic element in V_d^k. Intuitively, such a generic element, A, should have, in the appropriate sense, 0-eigenvalue block sizes as large as possible subject to the conditions rank $(A^n) = d$ and $A \in V_d^k$. To this end write $n - d = (k - d - 1)q + f$ where q, f are integers and $0 \leq f < k - d - 1$. Let A be a matrix in Jordan normal form with stable rank d, q many 0-eigenvalue blocks of size $k - d - 1$ and one 0-eigenvalue block of size f. Let r be the rank function of A.

LEMMA 4.11. *$V_d^k = X_r$. That is, r is the rank function of a generic point of V_d^k.*

PROOF: That $X_r \subseteq V_d^k$ is obvious. Suppose $B \in V_d^k$ has rank function s. Form the rank function $t(i) = \max(r(i), s(i))$. Note that $r(i) \leq t(i)$ for all i and that $t(k - d - 1) = d = r(k - d - 1)$. It suffices to show that $t(i) = r(i)$ for all i. We can suppose, by way of contradiction, that $t(i+1) > r(i+1)$ but $t(j) = r(j)$ for all $j \leq i$. Since $t(j) = r(j) = d$ for $j > k-d-1$, we have $i+1 < k-d-1$. Set $h = t(i) - t(i+1)$. Then $d = t(k-d-1) \geq t(i) - h(k-d-1-i) = r(i) - h(k-d-1-i)$. If $i \geq f$, this equals $n-iq-f-h(k-d-1-i)$, so $n-d-f \leq iq+h(k-d-1-i)$ implying that $h \geq q$ or $t(i+1) = t(i) - h = r(i) - h < r(i) - q = r(i+1)$, a contradiction. If $i < f$, $r(i) - h(k-d-1-i) = n-(q+1)i - h(k-d-1-i)$ so we have $n-d \leq (q+1)i + h(k-d-1-i) = (q+1)(k-d-1) - (q+1-h)(k-d-1-i)$. Thus $(q+1-h)(k-d-1-i) \leq (k-d-1-f)$. Since $i < f$, $q+1-h < 1$ or $h \geq q+1$. But $t(i+1) = t(i) - h = r(i) - h \leq r(i+1)$, a contradiction is again reached, and the lemma is proved. □

We will use the V_d^k's to give the decomposition of S_k into irreducible components. To begin, we use 4.10 to express S_k in terms of the V_d^k's.

Lemma 4.12. $S_k = \bigcup_{d=0}^{k-2} V_d^k$.

Proof: Clearly, from 4.10, S_k is contained in this union. But from 4.11 a generic point of V_d^k is in S_k and the result is proved. □

In order to purge the union in 4.12 of redundancies, we next investigate when $V_d^k \subseteq V_{d'}^k$. Write $V_d^k = X_r$ as in 4.11, and assume $V_d^k \subseteq V_{d'}^k$ with $d \neq d'$. Then $r(k-d'-1) \leq d'$. To express r, let $n-d = (k-d-1)q + f$ where q, f are integers and $0 \leq f < k-d-1$. First of all, r has stable value d so $d < d'$. If $k-d'-1 < f$ then $r(k-d'-1) = n-(k-d'-1)(q+1) = (d'-d)(q-1) + 2d' + (1-k+f)$ after substituting $n = (k-d-1)q + f + d$. Still assuming $k-d'-1 < f$, then $1-k+f > -d'$ so $(d'-d)(q-1) + 2d' + (1-k+f) > (d'-d)(q-1) + d' \geq d'$ a contradiction. Thus $k-d'-1 \geq f$ and $r(k-d'-1) = n-(k-d'-1)q - f = d + (d'-d)q$ after substituting the same expression for n. Since $r(k-d'-1) \leq d'$, we have $q = 1$. Altogether, if $d \neq d'$ and $V_d^k \subseteq V_{d'}^k$, then $d < d'$ and $q = 1$. This is part of the proof of:

Lemma 4.13. *Let d, $d \leq k-2$. The following are equivalent:*

a) $V_d^k \subseteq V_{d+1}^k$
b) $V_d^k \subseteq V_{d'}^k$ *for some* $d \neq d'$
c) $q = 1$

Proof: That a) implies b) is obvious, and that b) implies c) was argued above. Assume c). As $f < k-d-1$ we have that $d+1 \leq k-f-1$ or

$k-(d+1)-1>f$. Hence $r(k-(d+1)-1)=d+(d+1-d)q=d+1$ and so a) holds. □

With 4.13 in hand, we can give a more precise version of 4.12.

COROLLARY 4.14. *Let m be the maximum of 0 and $2k-2-n$. Then*

$$S_k=\bigcup_{d=m}^{k-2} V_d^k$$

and S_k is not the union of any subset of these V_d^k's.

PROOF: Since $q>0$, we have $q=1$ if and only if $(n-d)/(k-d-1)<2$. This last inequality is just $d<2k-2-n$. Thus it is precisely the V_d^k with $0\leq d<k-2-n$ that are redundant in 4.12, and 4.14 is proved. □

We close this section with a result relating the singular set S_k and the matrices λ_k^t studied in section 3. Recall that $\lambda_k=\lambda_k^n$ and $V(\lambda_k^t)$ is the zero set of the entries of λ_k^t. In particular, $V_k=V(\lambda_k^n)$.

THEOREM 4.15. $S_k=\mathrm{Sing}(V(\lambda_k^n))=V(\lambda_{k-1}^{n-1})$.

PROOF: According to 3.4, $V(\lambda_{k-1}^{n-1})$ is the union, for $0<s\leq k-1$, of the rank varieties $\{A \mid \mathrm{rank}(A^s)\leq \min(n-2s,k-s-1)\}$. Substituting $k-d-1$ for s, we have that $V(\lambda_{k-1}^{n-1})$ is the union, for $0\leq d\leq k-2$, of the varieties $\{A \mid \mathrm{rank}(A^{k-d-1})\leq \min(d,n-2k+2d+2)\}$. As this minimum is certainly less than or equal to d, we have by 4.12 that $S_k\supseteq V(\lambda_{k-1}^{n-1})$. Thus by 4.14 it suffices to show that if d satisfies $0\leq d\leq k-2$ and $d\geq 2k-1-n$, then $V_d^k\subseteq V(\lambda_{k-1}^{n-1})$. But $d\geq 2k-2-n$ implies that $d\leq n-2k+2d+2$, so V_d^k appears in the union for $V(\lambda_{k-1}^{n-1})$ and the result is proved. □

ACKNOWLEDGEMENTS: Both authors are grateful to the NSF for partial support during the preparation of this work and to the Mathematical Sciences Research Institute, where the work was done, for providing a congenial atmosphere in a beautiful setting.

References

1. D. Bayer and M. Stillman, *Macaulay, a computer algebra program*, Available free from the authors, for many machines including the Macintosh, IBM-PC, Sun, Vax, and others (1986).
2. W. Borho and H.-P. Kraft, *Über Bahnen und deren Deformationen bei linearen Aktionen reduktiver Gruppen*, Comment. Math. Helv. **54** (1979), 61–104.
3. C. De Concini and C. Procesi, *Symmetric functions, conjugacy classes, and the flag variety*, Invent. Math. **64** (1981), 203–219.
4. R. Elkik, *Singularités rationelles et déformations*, Invent. Math. **47** (1978), 139–147.
5. M. Gerstenhaber, *On dominance and varieties of commuting matrices*, Ann. of Math. **(2) 73** (1961), 324–348.
6. H. Grauert and O. Riemenschneider, *Verschwindungssätze für analytische Kohomologiegruppen auf komplexen Räumen*, Invent. Math. **11** (1970), 263–292.
7. R. Hartshorne, "Algebraic Geometry," Springer-Verlag, New York, 1977.
8. G. Kempf, F. Knudson, D. Mumford and B. Saint-Donat, "Toroidal embeddings, I," Springer Lecture Notes in Math., vol. 338, 1973.
9. B. Kostant, *Lie group representations on polynomial rings*, Amer. J. Math. **85** (1963), 327–404.
10. H.-P. Kraft and C. Procesi, *Closures of conjugacy classes of matrices are normal*, Invent. Math. **53** (1979), 227–247.
11. T. Matsumura, T., "Commutative Algebra," Benjamin/Cummings Publishing Co., 1980.
12. C. Procesi, C., *A formal inverse to the Cayley-Hamilton Theorem*, preprint (1986).
13. I. R. Shafarevich, "Basic Algebraic Geometry," Springer-Verlag, Berlin/Heidelberg/New York, 1977.
14. E. Strickland, *On the variety of projectors*, J. Alg. **106** (1987), 135–147.
15. T. Tanisaki, *Defining ideals of the closures of the conjugacy classes and representations of the Weyl groups*, Tohoku Math. J. **34** (1982), 575–585.
16. J. Weyman, *Equations of conjugacy classes of nilpotent matrices*, preprint (1987).

Department of Mathematics, Brandeis University, Waltham MA 02254

Department of Mathematics, University of Texas, Austin TX 78713