SIAM J. MATRIX ANAL. APPL.
Vol. 25, No. 2, pp. 495–506

LOW RANK PERTURBATION OF JORDAN STRUCTURE*

JULIO MORO† AND FROILÁN M. DOPICO†

Abstract. Let A be a matrix and λ_0 be one of its eigenvalues having g elementary Jordan blocks in the Jordan canonical form of A. We show that for most matrices B satisfying $\operatorname{rank}(B) \leq g$, the Jordan blocks of $A+B$ with eigenvalue λ_0 are just the $g - \operatorname{rank}(B)$ smallest Jordan blocks of A with eigenvalue λ_0. The set of matrices for which this behavior does not happen is explicitly characterized through a scalar determinantal equation involving B and some of the λ_0-eigenvectors of A. Thus, except for a set of zero Lebesgue measure, a low rank perturbation $A+B$ of A destroys for each of its eigenvalues exactly the $\operatorname{rank}(B)$ largest Jordan blocks of A, while the rest remain unchanged.

Key words. Jordan canonical form, matrix spectral perturbation theory

AMS subject classifications. 15A18, 15A21

DOI. 10.1137/S0895479802417118

1. Introduction.

It is well known [1, 4] that the multiple eigenvalues of a matrix split typically under perturbation into simple, distinct eigenvalues. If A is the unperturbed matrix, then each Jordan block of dimension k of A gives rise to a so-called *ring* or *cycle* [4, section II.1.2] of k different simple eigenvalues of the perturbed matrix, say $A+B$. This typical behavior takes place for sufficiently small B provided a certain genericity condition is satisfied by the perturbation (see [12, 5, 7] for more details).

In this paper we study a class of perturbations B which are only able to break some, but not all, of the Jordan blocks of A, namely perturbations with low rank. To be more precise, let λ_0 be an eigenvalue of A with geometric multiplicity g, i.e., $g = \dim\ker(A - \lambda_0 I)$, where ker denotes the null space and I is the identity matrix. By "low" rank we will mean in what follows that the rank of B satisfies

$$\operatorname{rank}(B) \leq g. \tag{1.1}$$

It is easy to check that this kind of perturbation cannot break all g Jordan blocks: using the elementary facts that $\operatorname{rank}(A+B-\lambda_0 I) \leq \operatorname{rank}(A-\lambda_0 I) + \operatorname{rank}(B)$ and $\operatorname{rank}(A-\lambda_0 I) = \operatorname{rank}(A+B-\lambda_0 I - B) \leq \operatorname{rank}(A+B-\lambda_0 I) + \operatorname{rank}(B)$, one easily gets

$$g - \operatorname{rank}(B) \leq \dim\ker(A+B-\lambda_0 I) \leq g + \operatorname{rank}(B). \tag{1.2}$$

Since every Jordan block corresponds to one independent eigenvector, the previous inequality implies that the perturbation B can destroy at most $\operatorname{rank}(B)$ of the Jordan blocks of A and can create at most $\operatorname{rank}(B)$ new Jordan blocks associated with each eigenvalue of A. This constraint still allows for a great deal of freedom as to the number and dimensions of the Jordan blocks of $A+B$. The purpose of this paper is to find out which is the most usual behavior in this respect.

*Received by the editors October 31, 2002; accepted for publication (in revised form) by I. C. F. Ipsen March 7, 2003; published electronically September 17, 2003. This research was partially supported by the Ministerio de Ciencia y Tecnología of Spain through grant BFM 2000-0008.
http://www.siam.org/journals/simax/25-2/41711.html

†Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911-Leganés, Spain (jmoro@math.uc3m.es, dopico@math.uc3m.es).

The following naive argument sheds light on the question: for most B's, the equality $\operatorname{rank}(A + B - \lambda_0 I) = \operatorname{rank}(A - \lambda_0 I) + \operatorname{rank}(B)$ holds, and consequently $\dim\ker(A + B - \lambda_0 I) = g - \operatorname{rank}(B)$. Hence, in most cases $A + B$ will have *exactly* $\operatorname{rank}(B)$ fewer Jordan blocks with eigenvalue λ_0 than A. Furthermore, the larger the size of a Jordan block, the more algebraic conditions are needed to ensure its existence, so the largest Jordan blocks should be more sensitive to perturbation than the smaller ones. According to this argument, the *generic* behavior one would expect for most perturbations B is that, *for each eigenvalue* λ_0 *of* A *satisfying* (1.1), *precisely the* $\operatorname{rank}(B)$ *largest Jordan blocks of* A *corresponding to that eigenvalue are destroyed in the Jordan form of* $A+B$*, and the other Jordan blocks of* A *persist as Jordan blocks of* $A+B$.

Of course this hand-waving argument does not always hold true, as shown in the following examples. An appropriately chosen "nontypical" rank one perturbation can increase the size of the Jordan blocks corresponding to $\lambda_0 = 1$ in

$$A+B = \left[\begin{array}{cc|cc|c} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

or it may increase the number of Jordan blocks associated with λ_0, as in

$$A+B = \left[\begin{array}{cc|cc|c} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{cc|cc|c} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

However, we will see that, in both cases, very special structures of the perturbation are needed to produce these unusual behaviors.

The main contribution of this paper is to obtain, for any matrix A and each eigenvalue λ_0, a simple, explicit characterization of the set of perturbations B for which the previously described typical behavior occurs. The necessary and sufficient condition for this is simply that a single scalar quantity, denoted by C_0, is not equal to zero. The scalar C_0 is defined through a sum of determinants of matrices involving B and some of the λ_0-eigenvectors of A. As a trivial consequence, the set of perturbations B for which the generic behavior does not happen, i.e., those fulfilling $C_0 = 0$, is an algebraic manifold of zero Lebesgue measure in the set of $n \times n$ complex matrices of given rank. This precise mathematical formulation allows us to term properly the expected behavior described above as *generic*.

The problem we address here was solved when the perturbation B has rank equal to one by Savchenko [9]. In fact, Savchenko conjectured without proof in [9] the generic behavior for perturbations of arbitrary rank. This conjecture motivated our work, leading first to the partial answer given in [8, section 3.2.1] and ultimately to the present paper. Recently, Savchenko [10] has found an independent (and different) proof of the results we present here. Both in [9] and in [10], the proofs rely on functional analytic techniques based on spectral resolvents. Our approach, based only on elementary linear algebra results, is probably better suited for the matrix analysis community. However, the approach in [9, 10] might be more amenable to extend results of this nature to infinite-dimensional operators.

An important point to be made is that all theorems below are valid for perturbations B *of any size*, i.e., they are by no means first-order perturbation results. This makes especially surprising the prominent role of the scalar C_0, a quantity which closely resembles the quantities defining the genericity conditions in first-order eigenvalue perturbation theory [5, 7]. In this respect, the results we present below are related to previous contributions in the context of first-order perturbation theory, dealing with perturbations restricted to some nongeneric manifold. Some preliminary results for nongeneric perturbations may be found in [7, section 3] as an extension of Lidskii's [5] classical results for generic perturbations, but the first systematic description of a class of structured perturbations was obtained by Ma and Edelman [6] for upper k-Hessenberg perturbations of Jordan blocks. More recently, Jeannerod [3] has extended Lidskii's results by obtaining explicit formulas for both the leading exponents and leading coefficients of the Puiseux expansions of the eigenvalues of *analytic* perturbations $J + B(\varepsilon)$ of a Jordan matrix J, provided the powers of ε in the perturbation matrix $B(\varepsilon)$ conform in a certain way to the Jordan structure given by J. However, in both cases [6, 3] the particular structure of the perturbations to the Jordan blocks is not preserved by undoing the change of basis leading to the Jordan form. Hence, not much information is provided for nongeneric perturbations of *arbitrary* matrices. The rank of the perturbation, on the other hand, does not change by undoing the Jordan change of basis. Therefore, to our knowledge, this work is a first contribution in this respect.

Another remarkable feature of the characterization via the scalar C_0 is that, taking into account the properties of the Jordan canonical form (see, for instance, [2, pp. 126–127]), the generic behavior will take place if and only if several equations involving the ranks of different powers of $A + B - \lambda_0 I$ and $A - \lambda_0 I$ are fulfilled. Surprisingly, in the case of low rank perturbations, this set of equations is equivalent to the single condition $C_0 \neq 0$, where C_0 does not involve explicitly any power, either of $A - \lambda_0 I$ or of $A + B - \lambda_0 I$.

Finally, although in this paper we only pay attention to which Jordan blocks are destroyed under a low rank perturbation, and which ones are preserved for each eigenvalue of A, another question which naturally arises is, What happens with the eigenvalues of the destroyed blocks? As stated before, classical first-order eigenvalue perturbation results answer the question for small perturbations: for each destroyed Jordan block of dimension k, a ring of k different simple eigenvalues of $A+B$ appears, and there are explicit formulas for the first-order corrections [5, 7]. For perturbations of arbitrary size, however, the information available is much more limited, and reduces to fairly general (and usually pessimistic) bounds on the variation of the eigenvalues [11].

The paper is organized as follows. In the second section, after setting the appropriate notation, we study in Theorem 2.1 the algebraic multiplicity, as an eigenvalue of $A+B$, of each eigenvalue λ_0 of A for which condition (1.1) holds. This multiplicity turns out to depend crucially on C_0, and $C_0 \neq 0$ is the necessary and sufficient condition for the algebraic multiplicity of λ_0 to be compatible with the predicted generic behavior, i.e., the Jordan blocks of $A+B$ with eigenvalue λ_0 are just the $g - \operatorname{rank}(B)$ smallest Jordan blocks of A with eigenvalue λ_0, where g is the number of λ_0-Jordan blocks of A. However, the algebraic and geometric multiplicity of an eigenvalue do not determine by themselves the corresponding part of the Jordan structure. In the third section, we prove in Theorem 3.1 that $C_0 \neq 0$ ensures the generic behavior by explicitly constructing the corresponding Jordan chains of $A+B$ starting from those of A. This will show that $C_0 \neq 0$ is a necessary and sufficient condition for the generic behavior, a fact we summarize in a final, concluding theorem.

2. Counting algebraic multiplicities. Throughout this section we follow the notation in [7]: let A be an arbitrary $n \times n$ complex matrix and

$$\left[\begin{array}{c|c} J & \\ \hline & \widehat{J} \end{array}\right] = \left[\begin{array}{c} Q \\ \hline \widehat{Q} \end{array}\right] A \left[\begin{array}{c|c} P & \widehat{P} \end{array}\right] \tag{2.1}$$

be a Jordan decomposition of A, so

$$\left[\begin{array}{c} Q \\ \hline \widehat{Q} \end{array}\right] \left[\begin{array}{c|c} P & \widehat{P} \end{array}\right] = I. \tag{2.2}$$

The matrix J contains all Jordan blocks associated with the eigenvalue of interest λ_0, while $\widehat{J}$ is the part of the Jordan form containing the other eigenvalues. Let

$$J = \Gamma_1^1 \oplus \cdots \oplus \Gamma_1^{r_1} \oplus \cdots \oplus \Gamma_q^1 \oplus \cdots \oplus \Gamma_q^{r_q}, \tag{2.3}$$

where, for $j = 1, \ldots, q$,

$$\Gamma_j^1 = \cdots = \Gamma_j^{r_j} = \begin{bmatrix} \lambda_0 & 1 & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ & & & & \lambda_0 \end{bmatrix}$$

is a Jordan block of dimension n_j repeated r_j times and ordered so that

$$n_1 > n_2 > \cdots > n_q.$$

The n_j are called the *partial multiplicities* for λ_0. The eigenvalue λ_0 is semisimple (nondefective) if $q = n_1 = 1$ and nonderogatory if $q = r_1 = 1$. Set

$$a = \sum_{j=1}^{q} r_j n_j \qquad \text{and} \qquad g = \sum_{j=1}^{q} r_j, \tag{2.4}$$

i.e., we denote by a the *algebraic* multiplicity of λ_0 as an eigenvalue of A, and by g its geometric multiplicity.

We further partition

$$P = \left[\begin{array}{c|c|c|c|c|c|c} P_1^1 & \ldots & P_1^{r_1} & \ldots & P_q^1 & \ldots & P_q^{r_q} \end{array}\right] \tag{2.5}$$

conformally with (2.3). The columns of each P_j^k form a right Jordan chain of A with length n_j corresponding to λ_0. The lth column of P_j^k is a right Jordan vector of order l. In particular, if we denote by x_j^k the first column of P_j^k, each x_j^k is a right eigenvector of A associated with λ_0. Analogously, we split

$$Q = \left[\begin{array}{c} Q_1^1 \\ \hline \vdots \\ \hline Q_1^{r_1} \\ \hline \vdots \\ \hline Q_q^1 \\ \hline \vdots \\ \hline Q_q^{r_q} \end{array}\right],$$

also conformally with (2.3). The rows of each Q_j^k form a left Jordan chain of A of length n_j corresponding to λ_0. The lth row, counting from below, of Q_j^k is a left Jordan vector of order l. Hence, if we denote by y_j^k the last (i.e., n_jth) row of Q_j^k, each y_j^k is a left eigenvector corresponding to λ_0. With these eigenvectors we build up matrices

$$L_j = \begin{bmatrix} y_j^1 \\ \vdots \\ y_j^{r_j} \end{bmatrix}, \qquad R_j = [x_j^1, \dots, x_j^{r_j}]$$

for $j = 1, \dots, q$,

$$W_i = \begin{bmatrix} L_1 \\ \vdots \\ L_i \end{bmatrix}, \qquad Z_i = [R_1, \dots, R_i]$$

for $i = 1, \dots, q$, and we define square matrices Φ_i of dimension

$$f_i = \sum_{j=1}^{i} r_j$$

by

$$\Phi_i = W_i B Z_i, \qquad i = 1, \dots, q. \tag{2.6}$$

Note that, due to the cumulative definitions of W_i and Z_i, every $\Phi_{i-1},\ i = 2, \dots, q$, is the upper left block of Φ_i.

Take, for instance, the unperturbed matrix

$$A = \left[\begin{array}{ccc|cc|cc|c} 0 & 1 & 0 & & & & & \\ 0 & 0 & 1 & & & & & \\ 0 & 0 & 0 & & & & & \\ \hline & & & 0 & 1 & & & \\ & & & 0 & 0 & & & \\ \hline & & & & & 0 & 1 & \\ & & & & & 0 & 0 & \\ \hline & & & & & & & 2 \end{array}\right], \tag{2.7}$$

and set $\lambda_0 = 0$, i.e., $g = 3$, $a = 7$, $n_1 = 3$, $n_2 = 2$, $r_1 = 1$, $r_2 = 2$. Then, since the right Jordan vectors of A are columns of the identity matrix, any given perturbation matrix

$$B = \left[\begin{array}{ccc|cc|cc|c} * & * & * & * & * & * & * & * \\ * & * & * & * & * & * & * & * \\ \blacksquare & * & * & \clubsuit & * & \spadesuit & * & * \\ \hline * & * & * & * & * & * & * & * \\ \clubsuit & * & * & \clubsuit & * & \heartsuit & * & * \\ \hline * & * & * & * & * & * & * & * \\ \spadesuit & * & * & \heartsuit & * & \spadesuit & * & * \\ \hline * & * & * & * & * & * & * & * \end{array}\right] \tag{2.8}$$

gives rise to the two matrices

$$\Phi_1 = \left[\ \blacksquare\ \right], \qquad \Phi_2 = \left[\begin{array}{c|cc} \blacksquare & \clubsuit & \spadesuit \\ \hline \clubsuit & \clubsuit & \heartsuit \\ \spadesuit & \heartsuit & \spadesuit \end{array}\right]$$

with dimensions $f_1 = 1$ and $f_2 = 3$.

As announced in the introduction, we want to determine the most likely Jordan structure for the eigenvalue λ_0 of a low rank perturbation $A + B$ of A, where by low we mean that B and λ_0 satisfy (1.1). Let n_s be the smallest one among the sizes of the rank(B) largest Jordan blocks of A associated with λ_0, i.e., $s \in \{1, \ldots, q\}$ is the index such that

$$\operatorname{rank}(B) \equiv \rho = f_{s-1} + \beta, \quad 0 < \beta \le r_s, \tag{2.9}$$

where we have set $f_0 = 0$ for convenience. In the 8×8 example above, if we consider perturbations with $\operatorname{rank}(B) = \rho = 2$, then $\rho = f_1 + \beta$ with $\beta = 1 < r_2 = 2$, i.e., $s = 2$ since the two largest Jordan blocks of A are the single 3×3 block, together with either one of the two 2×2 blocks.

We have already seen in formula (1.2) that the geometric multiplicity of λ_0 can decrease at most by ρ under the perturbation B. The following result shows how much the algebraic multiplicity usually decreases. If only the ρ largest Jordan blocks of A with eigenvalue λ_0 disappear, then the algebraic multiplicity of λ_0 in $A + B$ is

$$\widetilde{a} = (r_s - \beta) n_s + r_{s+1} n_{s+1} + \cdots + r_q n_q. \tag{2.10}$$

It is shown in Theorem 2.1 that the algebraic multiplicity of λ_0 in $A + B$ is always larger than or equal to $\widetilde{a}$, and the necessary and sufficient condition for equality is $C_0 \neq 0$.

THEOREM 2.1. *Let A be an $n \times n$ matrix with Jordan form* (2.1), *i.e., having an eigenvalue λ_0 with Jordan blocks of dimensions $n_1 > n_2 > \cdots > n_q$ repeated $r_1, r_2, \ldots, r_q$ times and algebraic and geometric multiplicities a and g given by* (2.4). *Let B be an $n \times n$ matrix with rank given by* (2.9), *and let the matrices Φ_i, $i = 1, \ldots, q$, be given by* (2.6). *Then the characteristic polynomial of $A + B$ is of the form*

$$p(\lambda) = (\lambda - \lambda_0)^{\widetilde{a}}\, t(\lambda - \lambda_0),$$

where $\widetilde{a}$ is given by (2.10) *and $t(\lambda - \lambda_0)$ is a monic polynomial of degree $n - \widetilde{a}$. Moreover, the constant coefficient of $t(\cdot)$ is*

$$t(0) = (-1)^{\rho+n-a}\, C_0\, \det(\widehat{Q}\, A\, \widehat{P} - \lambda_0 I), \tag{2.11}$$

where $\widehat{Q}$, $\widehat{P}$ *are as in* (2.1) *and* C_0 *is the sum of all principal minors of* Φ_s *corresponding to submatrices of dimension* ρ *containing the upper left block* Φ_{s-1} *of* Φ_s*. (If* $s = 1$, all *principal minors of dimension* ρ *are to be considered.) If, in particular,* $\rho = f_s$ *(i.e., if* $\beta = r_s$*) for some* $s \in \{1, \ldots, q\}$*, then* C_0 *is simply* $\det \Phi_s$*.*

Proof. We begin by writing the characteristic polynomial of $A + B$ as

$$p(\lambda) = \det((\lambda - \lambda_0)I - \operatorname{diag}(J - \lambda_0 I, \widehat{J} - \lambda_0 I) - \widetilde{B}),$$

where

$$\widetilde{B} = \left[\begin{array}{c} Q \\ \hline \widehat{Q} \end{array}\right] B \left[\begin{array}{c|c} P & \widehat{P} \end{array}\right].$$

For the sake of simplicity we define $\tilde{\lambda} \equiv \lambda - \lambda_0$ and $p_0(\tilde{\lambda}) \equiv p(\lambda)$, so the coefficient of $\tilde{\lambda}^{n-k}$ in $p_0(\tilde{\lambda})$ is $(-1)^k$ times the sum of all k-dimensional principal minors of $\operatorname{diag}(J-\lambda_0 I, \widehat{J} - \lambda_0 I) + \widetilde{B}$ [2, p. 42]. Notice that all principal minors whose corresponding submatrices have more than $\rho = \operatorname{rank}(B)$ rows (equivalently, columns) containing *only* elements of $\widetilde{B}$ are zero, since $\operatorname{rank}(\widetilde{B}) = \operatorname{rank}(B)$. This simple observation is the key to proving the theorem.

To find the lowest power of $\tilde{\lambda}$ in $p_0(\tilde{\lambda})$ we can just look for the largest possible dimension of a principal submatrix of $\operatorname{diag}(J - \lambda_0 I, \widehat{J} - \lambda_0 I) + \widetilde{B}$ containing at most ρ rows with only elements of $\widetilde{B}$. If we denote by $k_{\max}$ the maximal dimension we are looking for, then

$$p_0(\tilde{\lambda}) = \tilde{\lambda}^{n-k_{\max}}\, t(\tilde{\lambda}),$$

with t a monic polynomial of degree $k_{\max}$. Notice first that, since we are looking for the *largest* dimension, we can restrict ourselves to principal submatrices containing *exactly* ρ rows with only elements of $\widetilde{B}$: if the principal submatrix contains less than ρ rows with only elements of $\widetilde{B}$, then one may always construct a new principal submatrix of larger dimension by including a new row with only elements of $\widetilde{B}$ (and the corresponding column). For instance, any row in the position of a bottom row of a Jordan block of $J - \lambda_0 I$ contains only elements of $\widetilde{B}$, and since there are g of them, with $\rho \le g$, at least one of these bottom rows can be used to increase the dimension.

To determine $k_{\max}$, let $\alpha \subset \{1, 2, \ldots, n\}$ be any index set and denote by $(\operatorname{diag}(J-\lambda_0 I, \widehat{J} - \lambda_0 I) + \widetilde{B})(\alpha, \alpha)$ the principal submatrix of $\operatorname{diag}(J - \lambda_0 I, \widehat{J} - \lambda_0 I) + \widetilde{B}$ that lies in the rows and columns indexed by α. By definition, this principal submatrix contains all the diagonal elements in the positions indexed by α. Since the eigenvalues of $\widehat{J}$ are all different from λ_0, the diagonal elements in the positions $a+1, a+2, \ldots, n$ are *not* elements of $\widetilde{B}$, so the corresponding indices can be always included in α without increasing the number of rows with only elements of $\widetilde{B}$. Hence, any set α of the maximal size $k_{\max}$ containing exactly ρ rows with only elements of $\widetilde{B}$ must be of the form

$$\alpha = \{i_1, \ldots, i_j, a+1, a+2, \ldots, n\} \quad \text{with} \quad 1 \le i_1 < i_2 < \cdots < i_j \le a.$$

Furthermore, the rows $i_1, \ldots, i_j$ intersect with a certain number, say l, of the g Jordan blocks in $J - \lambda_0 I$. Take any of these Jordan blocks and denote by i_b the largest index corresponding to a row in α intersecting with that particular Jordan block. Then, the i_bth row contributes to the principal submatrix only with elements of $\widetilde{B}$, either because it is the bottom row of the Jordan block or because i_b+1 does not belong to α, and thus the element in the position (i_b, i_b+1), where $J-\lambda_0 I$ has a superdiagonal 1, is not in the submatrix. This imposes the restriction $l \leq \rho$ on l. Hence, no choice for α can give rise to a larger dimension than taking $l=\rho$ and choosing the indices $i_1 < \cdots < i_j$ to cover *all* rows of a set of ρ complete *largest* Jordan blocks of $J-\lambda_0 I$. Actually, any of these choices is admissible, since each contains exactly ρ rows with only elements of $\widetilde{B}$, namely one bottom row for each of the ρ Jordan blocks chosen from $J-\lambda_0 I$. The number of possible choices is $r_s!/(\beta!(r_s-\beta)!)$, which is simply one when $\rho = f_s$. Hence, we have shown that

$$k_{\max} = r_1 n_1 + \cdots + r_{s-1} n_{s-1} + \beta n_s + n - a,$$

and consequently $\widetilde{a} = n - k_{\max}$ with $\widetilde{a}$ given by (2.10).

Now we prove (2.11). Recall that $t(0)$ is $(-1)^{k_{\max}}$ times the sum of all $k_{\max}$-dimensional principal minors of $\operatorname{diag}(J-\lambda_0 I, \widehat{J}-\lambda_0 I)+\widetilde{B}$. Moreover, the only nonzero $k_{\max}$-dimensional principal minors correspond to the submatrices described in the previous paragraph. Consider one of these minors and call it M. Set $h = k_{\max} - (n-a) - \rho$ and denote by $1 = j_1 < j_2 < \cdots < j_h$ the indices of rows of the principal submatrix corresponding to M, where $J-\lambda_0 I$ has superdiagonal 1's. The j_kth row of this submatrix is the sum of two rows: one is the (j_k+1)st row e_{j_k+1} of the identity matrix, the other is a piece of a row of $\widetilde{B}$. Using this fact, we can expand M as a sum of 2^h determinants whose j_kth row, with $1 \leq k \leq h$, is either e_{j_k+1} or a row with only elements of $\widetilde{B}$. With the exception of the determinant with all the vectors $e_{j_1+1}, e_{j_2+1}, \ldots, e_{j_h+1}$, the rest of these determinants are zero because each contains more than ρ rows with elements of $\widetilde{B}$. A similar argument on the last $n-a$ rows of the submatrix corresponding to M allows us to replace every element of $\widetilde{B}$ in these rows by zero without changing the value of M. The cofactor expansion of the remaining determinant along the rows $1 = j_1 < j_2 < \cdots < j_h$ leads to a value for M equal to $(-1)^h \det(\widehat{J}-\lambda_0 I)$ times a minor of Φ_s corresponding to a principal submatrix of dimension ρ containing the upper left block Φ_{s-1}. Extending this argument to all nonzero $k_{\max}$-dimensional principal minors of $\operatorname{diag}(J-\lambda_0 I, \widehat{J}-\lambda_0 I)+\widetilde{B}$ leads to (2.11). $\square$

In example (2.7)–(2.8) above, with a perturbation B with

$$\operatorname{rank}(B) = \rho = 2,$$

the quantity C_0 is given by

$$C_0 = \det \left[\begin{array}{c|c} \blacksquare & \clubsuit \\ \hline \clubsuit & \clubsuit \end{array}\right] + \det \left[\begin{array}{c|c} \blacksquare & \spadesuit \\ \hline \spadesuit & \spadesuit \end{array}\right].$$

According to Theorem 2.1, any perturbation with $C_0 \neq 0$ is such that $\lambda_0 = 0$ is an eigenvalue of $A+B$ with algebraic multiplicity two and, according to (1.2), geometric multiplicity at least one. Hence, the Jordan form of $A+B$ can either have just one 2×2, or have two 1×1 Jordan blocks corresponding to λ_0. We shall prove in the next section that $C_0 \neq 0$ actually implies the first possibility.

3. Building Jordan chains. In this section we prove that the genericity condition $C_0 \neq 0$ actually implies that the $\operatorname{rank}(B)$ largest Jordan blocks of A disappear for each eigenvalue, and the rest of the Jordan blocks of A remain as Jordan blocks of $A+B$. If $\operatorname{rank}(B)$ is given by (2.9), we will construct, for the eigenvalue λ_0 of $A+B$,

$r_s - \beta$ Jordan chains of length n_s and r_k chains of length n_k for $k = s+1, \dots, q$. Due to Theorem 2.1, these are the only Jordan chains of $A+B$ for λ_0, since $C_0 \neq 0$ implies that the algebraic multiplicity of λ_0 is given by (2.10). Although the construction is more involved for perturbations of arbitrary rank, the crucial step in the proof is the recursive formula (3.5), a multidimensional analogue of the one employed by Savchenko [9] for the case of rank one perturbations.

In order to give a concise proof of the results in this section we need to introduce some further notation. Recall that each column of the matrix P in decomposition (2.1) is a Jordan vector of A associated with λ_0. Furthermore, the set of columns of each P_j^k, $j = 1, \dots, q$, $k = 1, \dots, r_j$, in (2.5) forms a right Jordan chain with length n_j of A associated with λ_0, and the lth column of P_j^k is a right Jordan vector of order l.

For each $l \in \{1, \dots, n_s\}$ we consider all right Jordan vectors of order l of A associated with λ_0 and denote by X_l (resp., Y_l) the submatrix of P containing all right Jordan vectors of order l corresponding to the ρ largest (resp., the $g-\rho$ smallest) Jordan blocks in J. Both the columns of X_l and of Y_l are assumed to appear in the same relative order as in P. Notice that whenever $\beta < r_s$ in (2.9), the ρ largest Jordan blocks in J are not uniquely determined: we need to further specify which β of the r_s Jordan blocks of size n_s contribute to the X_l, and this fixes which blocks contribute to the Y_l. We do this with the aid of the genericity condition $C_0 \neq 0$: recall that C_0 is the sum of all ρ-dimensional principal minors of Φ_s containing Φ_{s-1}, where $\Phi_s = W_s B Z_s$ and the columns of Z_s are right eigenvectors, i.e., right Jordan vectors of order 1. If $C_0 \neq 0$, then one or more of these principal minors of Φ_s must be different from zero. Let γ be the set of indices corresponding to the ρ rows and columns of Φ_s in any of the nonzero principal minors, and denote, as before, by $\Phi_s(\gamma, \gamma)$ the corresponding principal submatrix of Φ_s. Then γ must be of the form

$$\gamma = \{1, \dots, f_{s-1}, i_1, i_2, \dots, i_\beta\}, \qquad f_{s-1} < i_1 < \cdots < i_\beta \leq f_s, \tag{3.1}$$

and we define X_1 as the $n \times \rho$ submatrix of Z_s containing the columns indexed by γ. The $r_s - \beta$ remaining columns of Z_s are assigned to Y_1. Once X_1 (and therefore Y_1) is fixed, the columns of the remaining X_l (resp., Y_l) are chosen from the same set of Jordan blocks as the eigenvectors in X_1 (resp., Y_1). This implies that equations (3.3) below are satisfied.

In the example (2.7)–(2.8), with rank $(B) = 2$, there are only two principal minors of Φ_2 containing Φ_1, namely

$$\begin{aligned} \Phi_2(\{1,2\},\{1,2\}) &= \det \left[\begin{array}{c|c} \blacksquare & \clubsuit \\ \hline \clubsuit & \clubsuit \end{array}\right], \\ \Phi_2(\{1,3\},\{1,3\}) &= \det \left[\begin{array}{c|c} \blacksquare & \spadesuit \\ \hline \spadesuit & \spadesuit \end{array}\right]. \end{aligned} \tag{3.2}$$

If the first (resp., the second) minor is different from zero, then the two columns of $X_1 \in \mathbb{C}^{8\times 2}$ are the first and second (resp., first and third) columns of $Z_2 \in \mathbb{C}^{8\times 3}$, which are the first and fourth (resp., the first and sixth) columns of $P \in \mathbb{C}^{8\times 7}$. In that case, Y_1 reduces to the third (resp., second) column of Z_2.

Note that all matrices $X_l \in \mathbb{C}^{n\times \rho}$, $l = 1, \dots, n_s$, have the same dimensions, while $Y_l \in \mathbb{C}^{n \times d_l}$, with d_l the number of Jordan blocks of dimension larger than or equal to l among the $g - \rho$ smallest Jordan blocks contributing to Y_1. Hence,

$d_1 = g - \rho \geq d_2 \geq \cdots \geq d_{n_s}$. The fact that both the X_l and the Y_l are constituted by consecutive pieces of Jordan chains is reflected by the conditions

$$(3.3) \qquad (A - \lambda_0 I)X_l = X_{l-1}, \qquad (A - \lambda_0 I)Y_l = Y_{l-1}^{(l)}, \quad l = 1, \dots, n_s,$$

where $Y_{l-1}^{(l)}$ is the leftmost $n \times d_l$ submatrix of Y_{l-1}, and both X_0 and Y_0 are defined to be zero. Notice that if $\beta = r_s$, then $d_{n_s} = 0$ and Y_{n_s} is an empty matrix, so the second equation in (3.3) makes sense only for $l = 1, \dots, n_{s+1}$.

After all these conventions we are in the position to obtain the main result of this section.

THEOREM 3.1. *Let A, B, λ_0, and C_0 be as in the statement of Theorem* 2.1. *If $C_0 \neq 0$, then the Jordan blocks of $A + B$ with eigenvalue λ_0 are just the $g - \operatorname{rank}(B)$ smallest Jordan blocks of A with eigenvalue λ_0. More precisely, if the rank of B is given by* (2.9)*, then the Jordan structure of $A + B$ with eigenvalue λ_0 consists of $r_s - \beta$ Jordan blocks of dimension n_s and r_k Jordan blocks of dimension n_k for $k = s + 1, \dots, q$.*

Proof. As commented in the beginning of this section, it suffices to explicitly construct Jordan chains of the appropriate length for $A+B$. This amounts to constructing matrices $\widetilde{Y}_l \in \mathbb{C}^{n \times d_l}$ for $l = 1, \dots, n_s$ such that

$$(3.4) \qquad (A + B - \lambda_0 I)\widetilde{Y}_l = \widetilde{Y}_{l-1}^{(l)}, \qquad l = 1, \dots, n_s,$$

where $\widetilde{Y}_{l-1}^{(l)}$ is the leftmost $n \times d_l$ submatrix of $\widetilde{Y}_{l-1}$ and $\widetilde{Y}_0 = 0$. We must also prove that the columns of $[\widetilde{Y}_1, \widetilde{Y}_2, \dots, \widetilde{Y}_{n_s}]$ are linearly independent.

We will construct these matrices recursively through the formula

$$(3.5) \qquad \widetilde{Y}_l = Y_l - \sum_{i=1}^{l} X_i C_{l-i+1}^{(l)}, \qquad l = 1, \dots, n_s,$$

where, at the lth step, the $\rho \times d_l$ matrix $C_l^{(l)}$ is chosen in such a way that

$$(3.6) \qquad B\widetilde{Y}_l = 0,$$

and we denote by $C_j^{(l)}$ for $j < l$, the leftmost $\rho \times d_l$ submatrix of the $\rho \times d_j$ matrix $C_j^{(j)}$ already chosen at the jth step. The fact that condition (3.6) uniquely determines the matrix $C_l^{(l)}$ at each step is a consequence of our previous choice of the last β columns of the matrix X_1: since B has rank ρ, one can write $B = \mathcal{U}\mathcal{V}^*$ with $\mathcal{U}$, $\mathcal{V} \in \mathbb{C}^{n \times \rho}$ of full rank and, accordingly, rewrite (3.6) as

$$\mathcal{V}^* X_1 C_l^{(l)} = \mathcal{V}^* \left(Y_l - \sum_{i=2}^{l} X_i C_{l-i+1}^{(l)} \right),$$

where the right-hand side is already known. Hence, the solution $C_l^{(l)}$ is unique provided the square matrix $\mathcal{V}^* X_1$ is nonsingular. Now, recall that the ρ-dimensional principal submatrix $\Phi_s(\gamma, \gamma)$ of Φ_s indexed by the set γ in (3.1) is nonsingular, and $\Phi_s(\gamma, \gamma) = W_s(\gamma) B X_1$, where $W_s(\gamma)$ is the $\rho \times n$ submatrix of W_s containing the rows indexed by γ. Hence, $\Phi_s(\gamma, \gamma)$ is the product of two $\rho \times \rho$ matrices, $W_s(\gamma)\mathcal{U}$ and $\mathcal{V}^* X_1$, each of them nonsingular as well.

We now check (3.4): the definition (3.5) of $\widetilde{Y}_l$, together with (3.6) and (3.3), implies that

$$(A+B-\lambda_0 I)\widetilde{Y}_l = (A-\lambda_0 I)\left(Y_l - \sum_{i=1}^{l} X_i C_{l-i+1}^{(l)}\right) = Y_{l-1}^{(l)} - \sum_{i=2}^{l} X_{i-1} C_{l-i+1}^{(l)}.$$

Shifting the dummy index to $j = i-1$, the previous expression can be rewritten as

$$(A+B-\lambda_0 I)\widetilde{Y}_l = Y_{l-1}^{(l)} - \sum_{j=1}^{l-1} X_j C_{l-j}^{(l)},$$

and, since $C_{l-j}^{(l)}$ is the $\rho \times d_l$ leftmost submatrix of $C_{l-j}^{(l-1)}$, the matrix above is just $\widetilde{Y}_{l-1}^{(l)}$, the leftmost $n \times d_l$ submatrix of $\widetilde{Y}_{l-1}$. This proves that the matrices defined by (3.5) satisfy (3.4). Finally, each $\widetilde{Y}_l$ is just the corresponding Y_l plus some linear combinations of the columns of the matrices $X_1, \ldots, X_l$. Since the columns of all X_l and Y_l are linearly independent (the columns of P are linearly independent), the columns of $\widetilde{Y}_l$ are also linearly independent. □

In the example (2.7)–(2.8) with $\operatorname{rank}(B) = 2$, we would need to construct a Jordan chain of length two. If we assume that $C_0 \neq 0$, then one of the two minors in (3.2) is nonzero. Once X_1 is chosen accordingly, the construction of a new Jordan chain of length two for $A+B$ goes as follows: if we write X_i, Y_i, $i = 1, 2$, columnwise as

$$X_i = \left[\, \xi_i^{(1)} \; \xi_i^{(2)} \,\right], \quad Y_i = [\, \eta_i \,], \quad i = 1, 2,$$

and denote $C_1^{(1)} = [\, c_{11} \; c_{12} \,]^T$, then the first matrix equation $\widetilde{Y}_1 = Y_1 - X_1 C_1^{(1)}$ in (3.5) leads to the eigenvector

$$\widetilde{\eta}_1 = \eta_1 - c_{11}\, \xi_1^{(1)} - c_{12}\, \xi_1^{(2)}$$

of $A+B$, where c_{11} and c_{12} are chosen to ensure that $B\widetilde{\eta}_1 = 0$. The second vector $\widetilde{\eta}_2$ in the new Jordan chain is found through the equation $\widetilde{Y}_2 = Y_2 - X_2 C_1^{(2)} - X_1 C_2^{(2)}$, which, if $C_2^{(2)} = [\, c_{21} \; c_{22} \,]^T$, translates into

$$\widetilde{\eta}_2 = \eta_2 - c_{11}\, \xi_2^{(1)} - c_{12}\, \xi_2^{(2)} - c_{21}\, \xi_1^{(1)} - c_{22}\, \xi_1^{(2)}$$

in vector terms. Notice that in this case $C_1^{(2)} = C_1^{(1)}$. Again, the scalars c_{21} and c_{22} are chosen in such a way that $B\widetilde{\eta}_2 = 0$.

We may summarize the discussion throughout the paper by writing the conclusion of both Theorems 2.1 and 3.1 as a final, summarizing theorem.

Concluding Theorem. *Let A be a complex $n \times n$ matrix and λ_0 an eigenvalue of A with geometric multiplicity g. Let B be a complex $n \times n$ matrix with* $\operatorname{rank}(B) \leq g$ *and C_0 be as in the statement of Theorem* 2.1*. Then the Jordan blocks of $A+B$ with eigenvalue λ_0 are just the $g - \operatorname{rank}(B)$ smallest Jordan blocks of A with eigenvalue λ_0 if and only if $C_0 \neq 0$.*

Acknowledgments. The authors wish to thank Prof. Sergey Savchenko for bringing this interesting problem to their attention and for a fruitful interchange of ideas over e-mail. The authors also thank an anonymous referee whose many detailed suggestions helped to improve the overall presentation of the results.

REFERENCES

[1] H. Baumgärtel, *Analytic Perturbation Theory for Matrices and Operators*, Birkhäuser, Basel, Boston, Stuttgart, 1985.

[2] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.

[3] C.-P. Jeannerod, *On some nongeneric perturbations of an arbitrary Jordan structure*, Linear Algebra Appl., to appear.

[4] T. Kato, *Perturbation Theory for Linear Operators*, Springer, Berlin, Heidelberg, New York, 1980.

[5] V. B. Lidskii, *Perturbation theory of non-conjugate operators*, U.S.S.R. Comput. Math. and Math. Phys., 1 (1965), pp. 73–85.

[6] Y. Ma and A. Edelman, *Nongeneric eigenvalue perturbations of Jordan blocks*, Linear Algebra Appl., 273 (1998), pp. 45–63.

[7] J. Moro, J. V. Burke, and M. L. Overton, *On the Lidskii–Vishik–Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 793–817.

[8] J. Moro and F. M. Dopico, *First order eigenvalue perturbation theory and the Newton diagram*, in Applied Mathematics and Scientific Computing, Z. Drmač et al., eds., Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002, pp. 143–175.

[9] S. Savchenko, *The typical change of the spectral properties under a rank one perturbation*, Mat. Zametki, submitted (in Russian).

[10] S. Savchenko, *On the Change of the Spectral Properties under a Generic Perturbation of Rank r*, preprint (in Russian).

[11] G. W. Stewart and J. G. Sun, *Matrix Perturbation Theory*, Academic Press, Boston, MA, 1990.

[12] M. I. Vishik and L. A. Lyusternik, *The solution of some perturbation problems for matrices and selfadjoint or non-selfadjoint differential equations* I, Russian Math. Surveys, 15 (1960), pp. 1–74 (in English); Uspekhi Mat. Nauk, 15 (1960), pp. 3–80 (in Russian).