*The Jordan forms of AB and BA#*



**Citation:** Lippert, Ross A. and Gilbert Strang. "The Jordan Forms of AB and BA*." Electronic Journal of Linear Algebra 18 (2009) : 281-288.

**As Published:** http://www.math.technion.ac.il/iic/ela/ela-articles/articles/vol18_pp281-288.pdf

**Publisher:** International Linear Algebra Society

**Persistent URL:** http://hdl.handle.net/1721.1/63178

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 281-288, June 2009

ELA
http://math.technion.ac.il/iic/ela

# THE JORDAN FORMS OF $AB$ AND $BA$*

ROSS A. LIPPERT† AND GILBERT STRANG‡

**Abstract.** The relationship between the Jordan forms of the matrix products $AB$ and $BA$ for some given $A$ and $B$ was first described by Harley Flanders in 1951. Their non-zero eigenvalues and non-singular Jordan structures are the same, but their singular Jordan block sizes can differ by 1. We present an elementary proof that owes its simplicity to a novel use of the Weyr characteristic.

**Key words.** Jordan form, Weyr characteristic, eigenvalues

**AMS subject classifications.** 15A21, 15A18

**1. Introduction.** Suppose $A$ and $B$ are $n \times n$ complex matrices, and suppose $A$ is invertible. Then $AB = A(BA)A^{-1}$. The matrices $AB$ and $BA$ are similar. They have the same eigenvalues with the same multiplicities, and more than that, they have the same Jordan form. This conclusion is equally true if $B$ is invertible.

If both $A$ and $B$ are singular (and square), a limiting argument involving $A + \epsilon I$ is useful. In this case $AB$ and $BA$ still have the same eigenvalues with the same multiplicities. What the argument does not prove (because it is not true) is that $AB$ is similar to $BA$. Their Jordan forms may be different, in the sizes of the blocks associated with the eigenvalue $\lambda = 0$. This paper studies that difference in the block sizes.

*The block sizes can increase or decrease by* 1. This is illustrated by an example in which $AB$ has Jordan blocks of sizes 2 and 1 while $BA$ has three 1 by 1 blocks. We could begin with Jordan matrices $A$ and $B$:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The product $AB$ is zero. The product $BA$ also has a triple zero eigenvalue but the

* Received by the editors May 5, 2009. Accepted for publication May 28, 2009. Handling Editor: Roger A. Horn.

† 123 West 92 Street #1, New York, NY 10025, USA (ross.lippert@gmail.com).

‡ MIT Department of Mathematics, 77 Massachusetts Avenue, Building Room 2-240, Cambridge, MA 02139, USA (gs@math.mit.edu).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 281-288, June 2009

ELA
http://math.technion.ac.il/iic/ela



rank is 1. In fact, $BA$ is in Jordan form:

$$BA = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

A different 3 by 3 example illustrates another possibility:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

with

$$AB = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad BA = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Those examples show all the possible differences for $n = 3$, when $AB$ is nilpotent. More generally, we want to find every possible pair of Jordan forms for $AB$ and $BA$, for any $n \times m$ matrix $A$ and $m \times n$ matrix $B$ over an algebraically closed field. The solution to this problem, generalized to matrices over an arbitrary field, was given over 50 years ago by Harley Flanders [3], with subsequent generalizations and specializations [4, 6]. In this article, we give a novel elementary proof by using the *Weyr characteristic*.

**2. The Weyr Characteristic.** There are two dual descriptions of the Jordan block sizes for a specific eigenvalue. We can list the block dimensions $\sigma_i$ in decreasing order, giving the row lengths in Figure 2.1. This is the Segre characteristic. We can

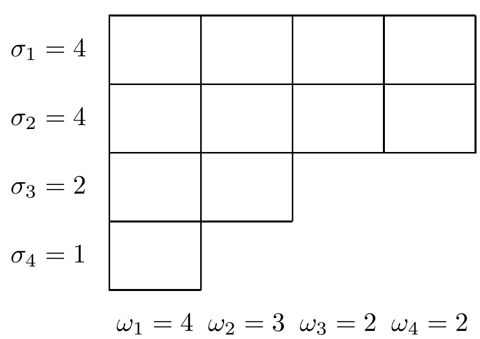


FIG. 2.1. *A tableau representing the Jordan structure* $J_4 \oplus J_4 \oplus J_2 \oplus J_1$.

also list the column lengths $\omega_1, \omega_2, \ldots$ (they automatically come in decreasing order).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 281-288, June 2009

ELA
http://math.technion.ac.il/iic/ela



This is the Weyr characteristic. By convention, we define $\sigma_i$ and $\omega_i$ for all $i > 0$ by setting them to 0 for sufficiently large $i$. If we consider $\{\sigma_i\}$ and $\{\omega_j\}$ to be partitions of their common sum $n$, then they are *conjugate partitions*: $\sigma_i$ counts the number of $j$'s for which $\omega_j \geq i$ and vice versa. The relationship between conjugate partitions $\{\sigma_i\}$ and $\{\omega_i\}$ is compactly summarized by $\omega_{\sigma_i} \geq i > \omega_{\sigma_i+1}$ (or by $\sigma_{\omega_i} \geq i > \sigma_{\omega_i+1}$), the first inequality making sense only when $\sigma_i > 0$. Tying the two descriptions to linear algebra is the *nullity index* $\nu_j$:

$$\nu_j(A) = \dim Null(A^j) = \text{ dimension of the nullspace of } A^j \quad (\text{with } \nu_0(A) = 0).$$

Thus $\nu_j$ counts the number of generalized eigenvectors for $\lambda = 0$ with *height* $j$ or less. In the example in Figure 2.1, $\nu_0, \ldots, \nu_5$ are $0, 4, 7, 9, 11$. Then $\omega_j = \nu_j - \nu_{j-1}$ counts the number of Jordan blocks of size $i$ or greater for $\lambda = 0$. Further exposition of the Weyr characteristic can be found in [5] and some geometric applications in [1, 2].

*Our main theorem is captured in the statement that* $\omega_i(BA) \geq \omega_{i+1}(AB)$. Reversing $A$ and $B$ gives a parallel inequality that we re-index as $\omega_{i-1}(AB) \geq \omega_i(BA)$. This observation, although in different terms, was central to the original proof by Flanders [3].

THEOREM 2.1. *Let $\mathbb{F}$ be an algebraically closed field. Given $A, B^t \in \mathbb{F}^{n\times m}$, the non-singular Jordan blocks of $AB$ and $BA$ have matching sizes, i.e., their Weyr characteristics are equal:*

$$\omega_i(AB - \lambda I) = \omega_i(BA - \lambda I) \quad \text{for } \lambda \neq 0 \text{ and all } i. \tag{2.1}$$

*For the eigenvalue $\lambda = 0$, the Jordan forms of $AB$ and $BA$ have Weyr characteristics that satisfy*

$$\omega_{i-1}(AB) \geq \omega_i(BA) \geq \omega_{i+1}(AB) \quad \text{for all } i, \tag{2.2}$$

*which is equivalent to*

$$|\sigma_i(AB) - \sigma_i(BA)| \leq 1 \quad \text{for all } i. \tag{2.3}$$

*If $P \in \mathbb{F}^{n\times n}$ and $Q \in \mathbb{F}^{m\times m}$ satisfy $\omega_i(P - \lambda I) = \omega_i(Q - \lambda I)$ for $\lambda \neq 0$ and $\omega_{i-1}(P) \leq \omega_i(Q) \leq \omega_{i+1}(P)$, then there exist $A, B^t \in \mathbb{F}^{n\times m}$ such that $P = AB$ and $Q = BA$.*

The equivalence of (2.2) and (2.3) is purely a combinatorial property of conjugate partitions (see Lemma 3.2).

The Jordan block sizes are hence restricted to change by at most 1 for $\lambda = 0$. Taking Figure 2.1 as the Jordan structure of $AB$ at $\lambda = 0$, Figure 2.2 is an admissible modification (by + and $-$) for $BA$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 281-288, June 2009

ELA
http://math.technion.ac.il/iic/ela



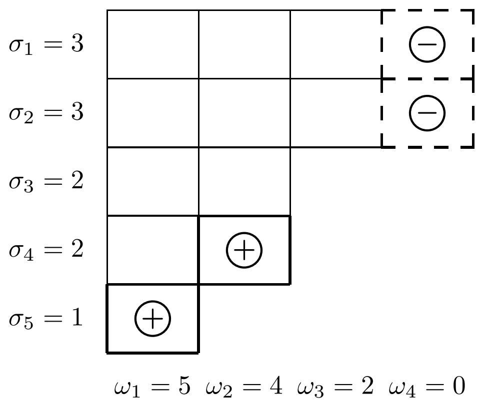


FIG. 2.2. *If $AB$ is nilpotent with Jordan structure $J_4 \oplus J_4 \oplus J_2 \oplus J_1$, then a permitted $BA$ structure is $J_3 \oplus J_3 \oplus J_2 \oplus J_2 \oplus J_1$.*

**3. Main results.** Our results are ultimately derived from the associativity of matrix multiplication. A typical example is $B(AB\cdots AB) = (BA\cdots BA)B$.

THEOREM 3.1. *If $A$ and $B^t$ are $n \times m$ matrices over a field $\mathbb{F}$, then for all $i > 0$*

$$\omega_i(AB - \lambda I) = \omega_i(BA - \lambda I) \quad \text{for } \lambda \in \mathbb{F} - \{0\}$$
$$\omega_i(BA) \geq \omega_{i+1}(AB) \quad (\text{for } \lambda = 0).$$

*Proof.* (For $\lambda \neq 0$) For any polynomial $p(x)$, $p(BA)B = Bp(AB)$. Thus $p(AB)v = 0$ implies $p(BA)Bv = 0$. Since $Bv = 0$ implies $p(AB)v = p(0)v$, we have $\dim Null(p(AB)) = \dim Null(p(BA))$ when $p(0) \neq 0$. Hence $\nu_i(AB - \lambda I) = \nu_i(BA - \lambda I)$ when $\lambda \neq 0$.

(For $\lambda = 0$) We define the following nullspaces for $i \geq 0$:

$$\mathcal{R}_i = \{v \in \mathbb{F}^n : B(AB)^i v = 0\}$$
$$\mathcal{R}'_i = \{v \in \mathbb{F}^n : (AB)^i v = 0\}$$
$$\mathcal{L}_i = \{v \in \mathbb{F}^m : v^t(BA)^i = 0\}$$
$$\mathcal{L}'_i = \{v \in \mathbb{F}^m : v^t(BA)^i B = 0\}$$

We see that, $\mathcal{R}_i \subset \mathcal{R}'_{i+1}$ and $\mathcal{L}_i \subset \mathcal{L}'_{i+1}$, and $\dim\{\mathcal{R}_{i+1}\} - \dim\{\mathcal{R}_i\} = \dim\{\mathcal{L}'_{i+1}\} - \dim\{\mathcal{L}'_i\}$.

Let $v_1, \dots, v_k \in \mathcal{R}'_{i+2}$ be a set of vectors that are linearly independent modulo $\mathcal{R}_{i+1}$. Thus $\sum_{i=1}^k c_i v_i \in \mathcal{R}_{i+1}$ only if $c_1 = \cdots = c_k = 0$. Then the vectors

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 281-288, June 2009

ELA
http://math.technion.ac.il/iic/ela



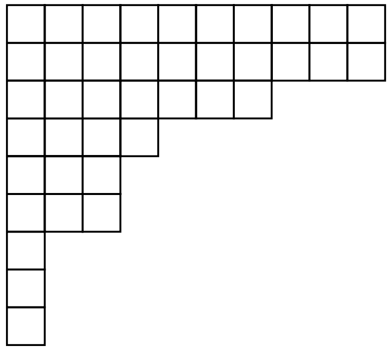

FIG. 3.1. *A tableau representing the Jordan structure* $\sigma_i = (10, 10, 7, 4, 3, 3, 1, 1, 1, 0, \ldots)$*, with Weyr characteristic* $\omega_i = (9, 6, 6, 4, 3, 3, 3, 2, 2, 2, 0, \ldots)$.

$ABv_1, \ldots, ABv_k \in \mathcal{R}'_{i+1}$ are linearly independent modulo $\mathcal{R}_i$. Thus, $\dim\{\mathcal{R}'_{i+1}/\mathcal{R}_i\} \geq \dim\{\mathcal{R}'_{i+2}/\mathcal{R}_{i+1}\}$. If $v_1, \ldots, v_k \in \mathcal{L}'_{i+2}$ is a set of vectors, linearly independent modulo $\mathcal{L}_{i+1}$, then the vectors $(BA)^t v_1, \ldots, (BA)^t v_k \in \mathcal{L}'_{i+1}$ are linearly independent modulo $\mathcal{L}_i$. Thus, $\dim\{\mathcal{L}'_{i+1}/\mathcal{L}_i\} \geq \dim\{\mathcal{L}'_{i+2}/\mathcal{L}_{i+1}\}$. Notice that

$$\dim\{\mathcal{R}'_{i+2}/\mathcal{R}_{i+1}\} = \nu_{i+2}(AB) - \dim\{\mathcal{R}_{i+1}\}$$
$$\dim\{\mathcal{L}'_{i+2}/\mathcal{L}_{i+1}\} = \dim\{\mathcal{L}'_{i+2}\} - \nu_{i+1}(BA).$$

Then $\dim\{\mathcal{R}'_{i+1}/\mathcal{R}_i\} \geq \dim\{\mathcal{R}'_{i+2}/\mathcal{R}_{i+1}\}$ implies

$$\dim\{\mathcal{R}_{i+2}\} - \dim\{\mathcal{R}_{i+1}\} \geq \nu_{i+2}(AB) - \nu_{i+1}(AB)$$

and $\dim\{\mathcal{L}'_{i+1}/\mathcal{L}_i\} \geq \dim\{\mathcal{L}'_{i+2}/\mathcal{L}_{i+1}\}$ implies

$$\nu_{i+1}(BA) - \nu_i(BA) \geq \dim\{\mathcal{L}'_{i+2}\} - \dim\{\mathcal{L}'_{i+1}\}.$$

Therefore, $\omega_{i+1}(BA) \geq \omega_{i+2}(AB)$, since $\omega_{i+1} = \nu_{i+1} - \nu_i$. ∎

The first part of Theorem 3.1 says that the Jordan structures of $AB$ and $BA$ for $\lambda \neq 0$ are identical, if $\mathbb{F}$ is algebraically closed. For a general field, the results can be adapted to show that the elementary divisors of $AB$ and $BA$, that do not have zero as a root, are the same. An illustration is helpful in understanding the constraints implied by the second part, $\omega_{i-1}(AB) \geq \omega_i(BA) \geq \omega_{i+1}(AB)$. Suppose the tableau in Figure 3.1 represents the Jordan form of $AB$ at $\lambda = 0$. Theorem 3.1 constrains the tableau of the Jordan form of $BA$ at $\lambda = 0$ to be that of $AB$ plus or minus the areas covered by the circles of Figure 3.2.

The constraints on Weyr characteristics are equivalent to constraining the block sizes of the Jordan forms of $AB$ and $BA$ to differ by no more than 1. Although this

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 281-288, June 2009

ELA
http://math.technion.ac.il/iic/ela



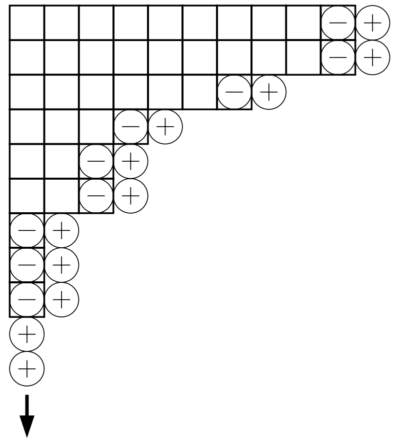

FIG. 3.2. *Given $AB$ (boxes), Theorem 3.1 imposes these constraints on the Weyr characteristic of $BA$ (a circle can be added or subtracted from each row of the tableau):* $\omega_1 \geq 6, 9 \geq \omega_2 \geq 6, 6 \geq \omega_3 \geq 4, 6 \geq \omega_4 \geq 3, 4 \geq \omega_5 \geq 3, \omega_6 = 3, 3 \geq \omega_7 \geq 2, 3 \geq \omega_8 \geq 2, \omega_9 = 2, 2 \geq \omega_9 \geq 0, 2 \geq \omega_{10} \geq 0$.

equivalence "is not hard to see" [3] from Figure 3.1, it warrants a short proof. Taking $d = 1$, Lemma 3.2 establishes the equivalence of (2.2) and (2.3).

LEMMA 3.2. *Let $p_1 \geq p_2 \geq \cdots$ and $p'_1 \geq p'_2 \geq \cdots$ be partitions of $n$ and $n'$ with conjugate partitions $q_1 \geq q_2 \geq \cdots$ and $q'_1 \geq q'_2 \geq \cdots$. Let $d \in \mathbb{N}$. Then*

$$q'_i \geq q_{i+d} \text{ and } q_i \geq q'_{i+d} \text{ for all } i > 0 \text{ if and only if } |p_i - p'_i| \leq d \text{ for all } i > 0.$$

*Proof.* If $p'_i > d$, then $q'_{p'_i} \geq i > q_{p_i+1}$ by the conjugacy conditions. By hypothesis, $q_{p'_i - d} \geq q'_{p'_i} > q_{p_i+1}$ and thus $p'_i - d < p_i + 1$ since $q_j$ is monotonically decreasing in $j$. Thus $p'_i \leq p_i + d$ (trivially true when $p'_i \leq d$). By a symmetric argument (switching primed and unprimed), we have $p_i \leq p'_i + d$.

Conversely, if $q_{i+d} > 0$, then $p'_{q_{i+d}} \geq p_{q_{i+d}} - d \geq (i+d) - d = i > p'_{q'_i+1}$, the first inequality by hypothesis and the next two by the conjugacy conditions. Since $p'_j$ is monotonically decreasing, we have $q_{i+d} < q'_i + 1$, and thus $q_{i+d} \leq q'_i$ for all $i > 0$ (trivially true when $q_{i+d} = 0$). A symmetric argument gives $q'_{i+d} \leq q_i$. ∎

What remains is to show that the constraints in Theorem 3.1 are exhaustive; we can construct matrices $A, B$ that realize all the possibilities of the theorem. Here we find it easier to use the traditional Segre characteristic of block sizes $\sigma_i$:

THEOREM 3.3. *Let $\sigma_1 \geq \sigma_2 \geq \cdots$ and $\sigma'_1 \geq \sigma'_2 \geq \cdots$ be partitions of $n$ and $m$ respectively.*

*If $|\sigma_i - \sigma'_i| \leq 1$, then there exist $n \times m$ matrices $A$ and $B^t$ such that $\sigma_j(AB) = \sigma_j$ and $\sigma_j(BA) = \sigma'_j$.*

**Electronic Journal of Linear Algebra ISSN 1081-3810**
**A publication of the International Linear Algebra Society**
**Volume 18, pp. 281-288, June 2009**

ELA
http://math.technion.ac.il/iic/ela



*Proof.* For each $j$ such that $\sigma_j$ and $\sigma'_j \geq 1$, we construct $\sigma_j \times \sigma'_j$ matrices $A_j$ and $B_j^t$ such that $A_jB_j = J_{\sigma_j}(0)$ and $B_jA_j = J_{\sigma'_j}(0)$ according to these three cases:

1. $\sigma_j = \sigma'_j$: set $A_j = J_{\sigma_j}(0)$ and $B_j = I_{\sigma_j}$,
2. $\sigma_j + 1 = \sigma'_j$: set $A_j = [\,0 \quad I_{\sigma_j}\,]$ and $B_j = \begin{bmatrix} I_{\sigma_j} \\ 0 \end{bmatrix}$,
3. $\sigma_j = \sigma'_j + 1$: set $A_j = \begin{bmatrix} I_{\sigma'_j} \\ 0 \end{bmatrix}$ and $B_j = [\,0 \quad I_{\sigma'_j}\,]$.

This defines $k = \min\{\omega_1(AB), \omega_1(BA)\}$ matrix pairs $(A_j, B_j)$. Consider $\{\sigma_j\}$ as a partition for $n$ rows and $\{\sigma'_j\}$ as a partition for $m$ columns. Construct the block diagonal matrix $A = \text{diag}(A_1, \ldots, A_k, 0, \ldots, 0)$ with zeros filling any remaining lower right part. Then with partitions $\{\sigma'_j\}$ for $m$ rows and $\{\sigma_j\}$ for $n$ columns let $B = \text{diag}(B_1, \ldots, B_k, 0, \ldots, 0)$. ∎

The final construction merely stitches together a singular piece with a non-singular piece.

Corollary 3.4. *Let $P \in \mathbb{F}^{n\times n}$ and $Q \in \mathbb{F}^{m\times m}$ have Segre characteristics $\sigma_i^\lambda$ and $\sigma_i'^\lambda$ for each eigenvalue $\lambda$, i.e.*

$$P \sim \bigoplus_{\lambda\in\mathbb{F}} \bigoplus_{i>0} J_{\sigma_i^\lambda}(\lambda) \quad \text{and} \quad Q \sim \bigoplus_{\lambda\in\mathbb{F}} \bigoplus_{i>0} J_{\sigma_i'^\lambda}(\lambda).$$

*If $\sigma_i^\lambda = \sigma_i'^\lambda$ for all $\lambda \neq 0$ and $|\sigma_i^0 - \sigma_i'^0| \leq 1$, then there exist matrices $A$ and $B^t$ in $\mathbb{F}^{n\times m}$ such that $P = AB$ and $Q = BA$.*

*Proof.* If $\tilde{P} = X^{-1}PX$ and $\tilde{Q} = Y^{-1}QY$ are in canonical form with $\tilde{P} = \tilde{A}\tilde{B}$ and $\tilde{Q} = \tilde{B}\tilde{A}$, then setting $A = X\tilde{A}Y^{-1}$ and $B = Y\tilde{B}X^{-1}$, we have $P = AB$ and $Q = BA$. Hence we take $P$ and $Q$ to be in canonical form.

Let $M = \bigoplus_{\lambda\neq 0} \bigoplus_{i>0} J_{\sigma_i}(\lambda)$, i.e., $M$ is a (non-singular) $k \times k$ matrix in Jordan canonical form with Segre characteristic $\sigma_i^\lambda$, where $k = \sum_{\lambda\neq 0} \sum_i \sigma_i^\lambda$. Let $A_0$ and $B_0$ be the $A$ and $B$ matrices from Theorem 3.3 with $\sigma_i = \sigma_i^0$ and $\sigma'_i = \sigma_i'^0$. Then $A = M \oplus A_0$ and $B = I_k \oplus B_0$. ∎

**Acknowledgment.** We thank Roger Horn for pointing us to the Flanders paper and others, and for his encouragement.

## REFERENCES

[1] James W. Demmel and Alan Edelman. The dimension of matrices (matrix pencils) with given Jordan (Kronecker) canonical forms. *Linear Algebra Appl.*, 230:61–87, 1995.

[2] Alan Edelman, Erik Elmroth, and Bo Kågström. A geometric approach to perturbation theory of matrices and matrix pencils. part II: A stratification-enhanced staircase algorithm. *SIAM J. Matrix Anal. Appl.*, 20(3):667–699, 1999.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 281-288, June 2009

ELA
http://math.technion.ac.il/iic/ela



[3] Harley Flanders. Elementary divisors of AB and BA. *Proc. Amer. Math. Soc.*, 2(6):871–874, 1951.

[4] W. V. Parker and B. E. Mitchell. Elementary divisors of certain matrices. *Duke Math. J.*, 19(3):483–485, 1952.

[5] Helene Shapiro. The Weyr characteristic. *Amer. Math. Monthly*, 106(10):919–929, 1999.

[6] Robert C. Thompson. On the matrices AB and BA. *Linear Algebra Appl.*, 1:43–58, 1968.