The Jordan forms of $AB$ and $BA$

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://www.math.technion.ac.il/iic/ela/ela-articles/articles/vol18_pp281-288.pdf">http://www.math.technion.ac.il/iic/ela/ela-articles/articles/vol18_pp281-288.pdf</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>International Linear Algebra Society</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/63178">http://hdl.handle.net/1721.1/63178</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
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
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

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

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

\[
\begin{array}{cccc}
\sigma_1 &=& 4 \\
\sigma_2 &=& 4 \\
\sigma_3 &=& 2 \\
\sigma_4 &=& 1 \\
\omega_1 &=& 4 \\
\omega_2 &=& 3 \\
\omega_3 &=& 2 \\
\omega_4 &=& 2
\end{array}
\]

Fig. 2.1. A tableau representing the Jordan structure \( J_4 \oplus J_4 \oplus J_2 \oplus J_1 \).
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 $\{\omega_i\}$ 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 $\omega_{\sigma_i} \geq i \geq \omega_{\sigma_{i+1}}$ (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 $\nu_j$:

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

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

\begin{equation}
\omega_i(AB - \lambda I) = \omega_i(BA - \lambda I) \quad \text{for } \lambda \neq 0 \text{ and all } i.
\end{equation}

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

\begin{equation}
\omega_{i-1}(AB) \geq \omega_i(BA) \geq \omega_{i+1}(AB) \quad \text{for all } i,
\end{equation}

which is equivalent to

\begin{equation}
|\sigma_i(AB) - \sigma_i(BA)| \leq 1 \quad \text{for all } i.
\end{equation}

If $P \in F^{n \times n}$ and $Q \in F^{m \times m}$ 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 $A, B \in F^{n \times m}$ 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 $BA$. 
### 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 $F$, then for all $i > 0$

$$
\omega_i(AB - \lambda I) = \omega_i(BA - \lambda I) \quad \text{for } \lambda \in F - \{0\}
$$

$$
\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 $\dim \text{Null}(p(AB)) = \dim \text{Null}(p(BA))$ 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 $i \geq 0$:

$$
R_i = \{v \in F^n : B(AB)^i v = 0\}
$$

$$
R'_i = \{v \in F^n : (AB)^i v = 0\}
$$

$$
L_i = \{v \in F^m : v^t(AB)^i = 0\}
$$

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

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

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

---

**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 $J_3 \oplus J_3 \oplus J_2 \oplus J_2 \oplus J_1$.
Jordan forms of $AB$ and $BA$

$ABv_1, \ldots , ABv_k \in R_{i+1}'$ are linearly independent modulo $R_i$. Thus,
\[ \dim\{R_{i+1}'/R_i\} \geq \dim\{R_{i+2}'/R_{i+1}\} \]
If $v_1, \ldots , v_k \in L_{i+2}'$ is a set of vectors, linearly
independent modulo $L_{i+1}$, then the vectors $(BA)^t v_1, \ldots , (BA)^t v_k \in L_{i+1}'$ are linearly
independent modulo $L_i$. Thus,
\[ \dim\{L_{i+1}'/L_i\} \geq \dim\{L_{i+2}'/L_{i+1}\} \]
Notice that
\[ \dim\{R_{i+2}'/R_{i+1}\} = \nu_{i+2}(AB) - \dim\{R_{i+1}\} \]
\[ \dim\{L_{i+2}'/L_{i+1}\} = \dim\{L_{i+2}'\} - \nu_{i+1}(BA) \]
Then \[ \dim\{R_{i+1}'/R_i\} \geq \dim\{R_{i+2}'/R_{i+1}\} \]
implies
\[ \dim\{R_{i+2}\} - \dim\{R_{i+1}\} \geq \nu_{i+2}(AB) - \nu_{i+1}(AB) \]
and \[ \dim\{L_{i+1}'/L_i\} \geq \dim\{L_{i+2}'/L_{i+1}\} \]
implies
\[ \nu_{i+1}(BA) - \nu_i(AB) \geq \dim\{L_{i+2}'\} - \dim\{L_{i+1}'\} \]
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 $F$ 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, $\omega_{i-1}(AB) \geq \omega_i(AB) \geq \omega_{i+1}(AB)$. 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
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): $\omega_1 \geq 6, 9 \geq \omega_2 \geq 6, 6 \geq \omega_3 \geq 4, 6 \geq \omega_4 \geq 3, 4 \geq \omega_5 \geq 3, 6 = 3, 3 \geq \omega_6 \geq 2, 3 \geq \omega_7 \geq 2, 2 = \omega_8 \geq 2, 2 \geq \omega_9 \geq 0, 2 \geq \omega_{10} \geq 0$.

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

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

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

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 $\sigma_j(BA) = \sigma'_j$. 
Jordan forms of $AB$ and $BA$ 287

Proof. For each $j$ such that $\sigma_j$ and $\sigma'_j \geq 1$, we construct $\sigma_j \times \sigma'_j$ matrices $A_j$ and $B'_j$ such that $A_j B_j = J_{\sigma_j}(0)$ and $B_j A_j = J_{\sigma'_j}(0)$ 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 $A_j = [0 \ I_{\sigma_j}]$ 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 $B_j = \begin{bmatrix} 0 \ I_{\sigma'_j} \end{bmatrix}$.

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 $\sigma^\lambda_i$ and $\sigma'^\lambda_i$ for each eigenvalue $\lambda$, i.e.,

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

If $\sigma^\lambda_i = \sigma'^\lambda_i$ for all $\lambda \neq 0$ and $|\sigma_i^0 - \sigma_i'^0| \leq 1$, then there exist matrices $A$ and $B'$ 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 $A = XAY^{-1}$ and $B = YBX^{-1}$, 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 $\sigma^\lambda_i$, where $k = \sum_{\lambda \neq 0} \sum_{i} \sigma^\lambda_i$. 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


