Proceedings of ALGORITMY 2009
pp. 429–439

ON BLOCK JACOBI ANNIHILATORS *

VJERAN HARI†

Abstract. The paper reveals the structure of the block Jacobi annihilator associated with one step of the general block Jacobi-type process of the form $\mathbf{A}^{(k+1)} = [\mathbf{P}^{(k)}]^* \mathbf{A}^{(k)} \mathbf{Q}^{(k)}$, $k \geq 0$. Here $\mathbf{P}^{(k)}$ and $\mathbf{Q}^{(k)}$ are nonsingular elementary block-matrices which differ from the identity in four blocks: two diagonal and the two corresponding off-diagonal blocks. In the case of unitary $\mathbf{P}^{(k)}$ and $\mathbf{Q}^{(k)}$, the block Jacobi annihilator is up to a permutational similarity a direct sum of an identity matrix, of a zero matrix and of a unitary matrix. The block Jacobi annihilators are building blocks of the block Jacobi operators, which are used in proving the global convergence of block Jacobi-type processes.

Key words. eigenvalues, singular values, block Jacobi-type method, global convergence

AMS subject classifications. 65F15

1. Introduction.

Let $\mathbf{A}$ be a square matrix of order n and let

$$\mathbf{A} = \begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mm} \end{bmatrix} \begin{matrix} n_1 \\ \vdots \\ n_m \end{matrix}, \tag{1.1}$$

be the *matrix block-partition* where the diagonal blocks are square. The matrix block-partition (1.1) is determined by the partition $\pi = (n_1, \ldots, n_m)$ of n, where $n_i \geq 1$ for all $1 \leq i \leq m$ and $n_1 + \cdots + n_m = n$.

Block Jacobi-type methods are iterative processes of the form

$$\mathbf{A}^{(k+1)} = [\mathbf{P}^{(k)}]^* \mathbf{A}^{(k)} \mathbf{Q}^{(k)}, \quad k \geq 0, \tag{1.2}$$

where $\mathbf{P}^{(k)}$, $\mathbf{Q}^{(k)}$ are *elementary block matrices*, $[\mathbf{P}^{(k)}]^*$ is the Hermitian transpose of $\mathbf{P}^{(k)}$ and $\mathbf{A}^{(0)} = \mathbf{A}$ is the initial matrix. Generally, elementary block matrix $\mathbf{E}$ is a nonsingular $n \times n$ matrix of the form

$$\mathbf{E}_{ij} = \begin{bmatrix} I & & & & \\ & E_{ii} & & E_{ij} & \\ & & I & & \\ & E_{ji} & & E_{jj} & \\ & & & & I \end{bmatrix} \begin{matrix} \\ \}n_i \\ \\ \}n_j \\ \\ \end{matrix}, \quad i < j, \quad \text{or} \quad \mathbf{E}_{ii} = \begin{bmatrix} I & & \\ & E_{ii} & \\ & & I \end{bmatrix} \}n_i, \quad i = j,$$

where $\mathbf{E}$ carries the same partition as $\mathbf{A}$. All elements of $\mathbf{E}$, except possibly for the elements in the blocks E_{ii}, E_{ij}, E_{ji} and E_{jj}, are as in the identity matrix I_n. Indices i, j are the *pivot indices*, (i, j) is the *pivot pair* and

$$\hat{\mathbf{E}} = \begin{bmatrix} E_{ii} & E_{ij} \\ E_{ji} & E_{jj} \end{bmatrix} \text{ if } i < j \qquad \text{or} \qquad \hat{\mathbf{E}} = E_{ii} \text{ if } i = j$$

is the *pivot submatrix* or the (i, j) – *restriction* of $\mathbf{E}$. We will write $\mathbf{E}_{ij} = \mathcal{E}(i, j; \hat{\mathbf{E}})$, where $\mathcal{E} = \mathcal{E}_\pi$ is the mapping which constructs the $n \times n$ matrix $\mathbf{E}_{ij}$ from the input data i, j and $\hat{\mathbf{E}}$.

*This work was supported by the Croatian Ministry of Science, Education and Sports, grant 037-0372783-3042
†Department of Mathematics, University of Zagreb, Croatia (hari@math.hr).

Each iteration of the block Jacobi-type method (1.2) is associated with pivot indices $i = i(k)$, $j = j(k)$, $i \le j$, $\mathbf{P}^{(k)} = \mathcal{E}(i(k), j(k); \hat{\mathbf{P}}^{(k)})$ and $\mathbf{Q}^{(k)} = \mathcal{E}((i(k), j(k); \hat{\mathbf{Q}}^{(k)})$. When the emphasis is on pivot indices, we shall write $\mathbf{P}_{i(k)j(k)}$ instead of $\mathbf{P}^{(k)}$or sometimes $\mathbf{P}_{ij}$ when k is clear from the context. If $n_1 = n_2 = \cdots = n_m = 1$, we speak of a *non-block Jacobi-type* method or simply of a Jacobi-type method.

The process (1.2) is defined if at each step k one knows the pivot pair (i, j) and the *algorithm* which determines the pivot submatrices $\hat{\mathbf{P}}^{(k)}$ and $\hat{\mathbf{Q}}^{(k)}$ from the elements of $\mathbf{A}^{(k)}$.

The way of choosing the pivot pairs is referred to as *pivot strategy*, but since each pivot pair addresses one block, one can use the term *block pivot strategy*. For simplicity we call it briefly *strategy*. Let $\mathbb{N}_0 = \{0, 1, 2, \ldots, \}$ and $\mathbb{P}_m = \{(s,t) : 1 \le s \le t \le m\}$. Then $\mathrm{card}(\mathbb{P}_m) = m(m+1)/2$ is the cardinality of $\mathbb{P}_m$. We can define pivot strategies as functions from $\mathbb{N}_0$ to $\mathbb{P}_m$. Each strategy $\mathrm{I} : \mathbb{N}_0 \mapsto \mathbb{P}_m$ satisfies $\mathrm{I}(k) = (i(k), j(k)),\ k \ge 0$. If I is a periodic function, then I is called *periodic strategy*. Let I be a periodic strategy with period M. If $M \ge \mathrm{card}(\mathbb{P}_m)$ ($M = \mathrm{card}(\mathbb{P}_m)$) and $\{\mathrm{I}(k) : k = 0, 1, \ldots, M-1\} = \mathbb{P}_m$, then I is called *quasi-cyclic* (*cyclic*) strategy.

For any square matrix $X = (x_{ij})$, the function Off (X), Off $^2(X) = \|X - \mathrm{diag}(X)\|_F^2$ is referred to as *departure from the diagonal form* or the *off-norm* of X. Here $\| \cdot \|_F$ is the Frobenius norm while $\mathrm{diag}(X)$ is the diagonal part of X. When used with the iteration matrix $\mathbf{A}^{(k)}$ generated by the iteration (1.2), it measures how far the process has advanced. So, in proving the global convergence of Jacobi-type methods, it is important to find some sufficient conditions for the convergence of $\mathbf{A}^{(k)}$ to diagonal form, i.e. for Off $(\mathbf{A}^{(k)}) \to 0$ as $k \to \infty$.

As has been shown in [4], one important tool for proving the convergence to diagonal form of the block Jacobi-type process (1.2) is the theory of *block Jacobi operators*. In [6] and [5] Henrici and Zimmermann introduced *Jacobi operators* as tool for proving the global and asymptotic convergence of the column-cyclic Jacobi method for symmetric matrices. Later, this tool has been generalized to work for complex Hermitian matrices [3], and for proving convergence to diagonal form of general Jacobi-type processes [1], [2], [3]. Each Jacobi operator is a product of M Jacobi annihilators, where M is the period of the pivot strategy. In [4] Jacobi annihilators and operators have been generalized to cope with the block Jacobi-type processes. Therefore, we call them here *block Jacobi annihilators* and *operators*.

The block Jacobi operators are made up of block Jacobi annihilators. Hence, the latter are the building blocks for the whole theory. Here, we reveal the structure of the block Jacobi annihilators. This structure is used in [4] for estimating the norms of certain block Jacobi operators .

2. Block Jacobi Annihilators. To an arbitrary $p \times q$ matrix X we can associate the column-vector col (X) and the row-vector row (X) as follows

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1q} \\ \vdots & \ddots & \vdots \\ x_{p1} & \cdots & x_{pq} \end{bmatrix} \mapsto \begin{cases} \mathrm{col}\,(X) = [x_{11}, x_{21}, \ldots, x_{p1}, \ldots, x_{1q}, x_{2q}, \ldots, x_{pq}]^T, \\ \mathrm{row}\,(X) = [x_{11}, x_{12}, \ldots, x_{1q}, \ldots, x_{p1}, x_{p2}, \ldots, x_{pq}]. \end{cases}$$

Thus, col (X) (row (X)) is the column- (row-) vector obtained by using the column-(row-) wise ordering of the elements of X. Here, generally, Z^T stands for the transpose of Z.

Let $\pi = (n_1, \ldots, n_m)$ be a partition of n, and let $\mathbf{A} = (A_{st})$ be the corresponding block matrix partition (1.1). For $2 \le i, j \le m$, let

$$\mathbf{r}_i = [\mathrm{row}\,(A_{i1})\ \mathrm{row}\,(A_{i2})\ \ldots\ \mathrm{row}\,(A_{i,i-1})], \qquad \mathbf{c}_j = \begin{bmatrix} \mathrm{col}\,(A_{1j}) \\ \mathrm{col}\,(A_{2j}) \\ \vdots \\ \mathrm{col}\,(A_{j-1,j}) \end{bmatrix}.$$

With $\mathbf{A}$ and π, we associate the column-vector

$$\mathbf{a} = \mathrm{vec}(\mathbf{A}) = \left[\mathbf{c}_2^T, \mathbf{c}_3^T, \ldots, \mathbf{c}_m^T,\ \mathbf{r}_2, \mathbf{r}_3, \ldots, \mathbf{r}_m\right]^T. \tag{2.1}$$

Thus, $\mathrm{vec} = \mathrm{vec}_\pi :\ \mathbf{C}^{n\times n} \to \mathbf{C}^{2K}$ where

$$K = K_\pi = N - \sum_{t=1}^{m} \frac{n_t(n_t-1)}{2}, \quad N = \frac{n(n-1)}{2}. \tag{2.2}$$

Let $\mathbf{C}_\circ^{n\times n} = \mathbf{C}_{\circ,\pi}^{n\times n} \subseteq \mathbf{C}^{n\times n}$ be the linear subspace of n by n matrices whose all diagonal blocks from the block partition (1.1) equal to zero. It is immediate to see that the restriction of vec to this subspace is a regular linear operator from $\mathbf{C}_\circ^{n\times n}$ to $\mathbf{C}^{2K}$ and it is isometry if $\mathbf{C}_\circ^{n\times n}$ is equipped with the Frobenius norm and $\mathbf{C}^{2K}$ with the Euclidean vector norm. We denote this restriction by $\mathrm{vec}_\circ$.

We shall also use the vector space $\mathbf{C}^K$ of K-vectors for which holds $\mathbf{C}^{2K} = \mathbf{C}^K \oplus \mathbf{C}^K$. Here $\oplus$ denotes the orthogonal sum of vector spaces. If $\mathbf{a}$ is as in the relation (2.1), then

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix}, \quad \mathbf{a}_1 = \left[\mathbf{c}_2^T, \ldots, \mathbf{c}_m^T\right]^T, \quad \mathbf{a}_2 = \left[\mathbf{r}_2, \ldots, \mathbf{r}_m\right]^T, \quad \mathbf{a}_1,\ \mathbf{a}_2 \in \mathbf{C}^K.$$

Let $i < j$ and let $\mathbf{P} = \mathcal{E}(i,j;\hat{\mathbf{P}})$, $\mathbf{Q} = \mathcal{E}(i,j;\hat{\mathbf{Q}})$ be elementary block matrices with

$$\hat{\mathbf{P}} = \begin{bmatrix} P_{ii} & P_{ij} \\ P_{ji} & P_{jj} \end{bmatrix} \begin{matrix} n_i \\ n_j \end{matrix}, \qquad \hat{\mathbf{Q}} = \begin{bmatrix} Q_{ii} & Q_{ij} \\ Q_{ji} & Q_{jj} \end{bmatrix} \begin{matrix} n_i \\ n_j \end{matrix}. \tag{2.3}$$

The block partition in (2.3) is inherited from π. Let

$$\mathbf{A}' = \mathbf{P}^*\mathbf{A}\mathbf{Q}, \quad \mathbf{A} \in \mathbf{C}^{n\times n}. \tag{2.4}$$

If $\hat{\mathbf{P}}$ and $\hat{\mathbf{Q}}$ are chosen such that the (i,j)-restriction of $\mathbf{A}'$ is diagonal, then the relation (2.4) induces in $\mathbf{C}^{2K}$ the equation

$$\mathbf{a}' = \Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})\mathbf{a}, \qquad \mathbf{a} = \mathrm{vec}(\mathbf{A}), \quad \mathbf{a}' = \mathrm{vec}(\mathbf{A}'). \tag{2.5}$$

We call the $2K$ by $2K$ matrix $\Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$ *Jacobi annihilator* or *Jacobi factor* associated with the elementary block matrices $\mathbf{P}$ and $\mathbf{Q}$. It is evident that the diagonal elements of $\Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$ which correspond to those elements of $\mathbf{a}$ which are not affected by the transformation are one, and those which correspond to the elements of $\mathbf{a}'$ which are annihilated, are zero.

Next, we extend the definition of the Jacobi annihilator/factor. Let $\mathbf{P}$ and $\mathbf{Q}$ be any elementary block matrices, i.e. we only require that $\hat{\mathbf{P}}$ and $\hat{\mathbf{Q}}$ from the relation (2.3) are regular. This means that the (i,j)-restriction of $\mathbf{A}'$ from the relation (2.4) does not have be diagonal. We define the Jacobi factor $\Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$ by help of the relation (2.5), but require that $\mathbf{A}'$ is obtained from $\mathbf{A}$ in the following way

$$\mathbf{A}' = \mathcal{Z}_{ij}(\tilde{\mathbf{A}}), \quad \tilde{\mathbf{A}} = \mathbf{P}^*\mathbf{A}\mathbf{Q}, \quad \mathbf{A} \in \mathbf{C}^{n\times n}.$$

Here, $\mathcal{Z}_{ij}$ maps the (i,j)-restriction of the argument matrix to zero. Thus, if $\mathbf{X}$ carries the block matrix partition defined by π, then the (i,j)-restriction $\mathcal{Z}_{ij}(\mathbf{X})$ is zero.

The structure of $\Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$ in the special case $n_1 = \cdots = n_m = 1$, $\hat{\mathbf{P}}$, $\hat{\mathbf{Q}}$ unitary, is described in [3] .

The partition π and the pair of elementary block matrices $\mathbf{P} = \mathcal{E}(i,j;\hat{\mathbf{P}})$, $\mathbf{Q} = \mathcal{E}(i,j;\hat{\mathbf{Q}})$ uniquely determine the Jacobi factor $\Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$. To see that, we describe with more details how $\Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$ acts on an arbitrary vector $\mathbf{x} \in \mathbf{C}^{2K}$.

Let $\mathbf{x} \in \mathbf{C}^{2K}$ and let $\mathbf{X} \in \mathbf{C}_\circ^{n\times n}$ be the unique matrix $\mathbf{X} = \mathrm{vec}_\circ^{-1}(\mathbf{x})$. The matrix $\mathbf{X}$ carries the partition (X_{st}) induced by π and its diagonal blocks are zero. Let $\mathbf{X}$ be transformed into $\mathbf{Y} \in \mathbf{C}_\circ^{n\times n}$ by the following rule

$$\begin{aligned}
&Y_{ir} = P_{ii}^* X_{ir} + P_{ji}^* X_{jr}, \quad 1 \le r \le m\\
&Y_{jr} = P_{ij}^* X_{ir} + P_{jj}^* X_{jr}, \quad 1 \le r \le m\\
&Y_{ri} = X_{ri} Q_{ii} + X_{rj} Q_{ji}, \quad 1 \le r \le m\\
&Y_{rj} = X_{ri} Q_{ij} + X_{rj} Q_{jj}, \quad 1 \le r \le m\\
&Y_{ij} = O, \quad Y_{ji} = O, \quad Y_{ii} = O, \quad Y_{jj} = O\\
&Y_{st} = X_{st} \quad \text{whenever} \quad \{s,t\} \cap \{i,j\} = \emptyset\,,
\end{aligned} \tag{2.6}$$

where the lines in (2.6) are to be read (i.e. performed) sequentially from top to bottom. The transformation (2.6) is a composition of the linear transformation $\mathbf{X} \mapsto \tilde{\mathbf{X}} = \mathbf{P}^*\mathbf{X}\mathbf{Q}$ and the linear transformation $\tilde{\mathbf{X}} \mapsto \mathbf{Y} = \mathcal{Z}_{ij}(\tilde{\mathbf{X}})$ which simply sets the blocks $\tilde{X}_{ii}, \tilde{X}_{ij}, \tilde{X}_{ji}, \tilde{X}_{jj}$ of $\tilde{\mathbf{X}}$ to zero. Let $\mathbf{y} = \mathrm{vec}(\mathbf{Y}) = \mathrm{vec}_\circ(\mathbf{Y})$. Then $\Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$ is the unique matrix which satisfies $\mathbf{y} = \Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})\mathbf{x}$ for all $\mathbf{x} \in \mathbf{C}^{2K}$. Since $\mathrm{vec}_\circ$ and $\mathrm{vec}_\circ^{-1}$ are linear transformations, the mapping $\mathbf{x} \mapsto \mathbf{y}$ is linear. So, the matrix $\Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$ exists.

To show the uniqueness of $\Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$, suppose that

$$\Im_{ij}(\hat{\mathbf{P}}', \hat{\mathbf{Q}}') \ne \Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}}) \quad \text{and} \quad \Im_{ij}(\hat{\mathbf{P}}', \hat{\mathbf{Q}}')x = \Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})x \quad \text{for all } x.$$

This would imply

$$\begin{aligned}
&O = (P_{ii} - P_{ii}')^* X_{ir} + (P_{ji} - P_{ji}')^* X_{jr}, \quad \text{for } 1 \le r \le m\\
&O = (P_{ij} - P_{ij}')^* X_{ir} + (P_{jj} - P_{jj}')^* X_{jr}, \quad \text{for } 1 \le r \le m\\
&O = X_{ri}(Q_{ii} - Q_{ii}') + X_{rj}(Q_{ji} - Q_{ji}'), \quad \text{for } 1 \le r \le m\\
&O = X_{ri}(Q_{ij} - Q_{ij}') + X_{rj}(Q_{jj} - Q_{jj}'), \quad \text{for } 1 \le r \le m
\end{aligned} \tag{2.7}$$

If for example, $P_{ii} \ne P_{ii}'$, then $e_\alpha^T P_{ii} e_\beta \ne e_\alpha^T P_{ii}' e_\beta$ for some α and β, where e_α and e_β are the columns of I_{n_i}. Choosing an x such that $X_{ir} e_\beta = e_\beta$ and $X_{jr} e_\beta = 0$, we obtain that the first equation in (2.7) is violated. The arguments for other cases are similar.

Finally, let us consider the case $i = j$. In this case the relation (2.6) simplifies to just three lines. The second, the forth and the fifth line can be removed, while the first and the third lines are simplified to $Y_{ir} = P_{ii}^* X_{ir}$, $1 \le r \le m$ and $Y_{ri} = X_{ri} Q_{ii}$, $1 \le r \le m$. Hence, $\mathbf{X}$ and $\mathbf{Y}$ are linked solely by the transformation $\mathbf{Y} = \mathbf{P}^*\mathbf{X}\mathbf{Q}$, i.e. the operator $\mathcal{Z}_{ii}$ is not needed.

2.1. The structure of $\Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$. Here we find the structure of the matrix $\Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$. We first consider the non-trivial case: $i < j$.

Let $\pi = (n_1, \ldots, n_m)$ be a partition of n and let

$$\Sigma = \Sigma_\pi = (s_1, s_2, \ldots, s_m), \quad s_r = n_1 + \cdots + n_r,\ 1 \le r \le m.$$

Obviously, we have $s_r = s_{r-1} + n_r$, $1 \le r \le m$, provided that $s_0 = 0$. Let

$$b(r) = (s_{r-1} + 1, \ldots, s_r), \quad 1 \le r \le m,$$

denote the “block index” corresponding to the r’th block-column and block-row of $\mathbf{X}$ and $\mathbf{Y}$ from the relation (2.6).

Let $\mathbf{x}, \mathbf{y} \in \mathbf{C}^{2K}$ with $\mathbf{x}$ arbitrary and $\mathbf{y}$ satisfying $\mathbf{y} = \Im_{ij}\mathbf{x}$. The most natural block-partition of $\mathbf{x}$ and $\mathbf{y}$ is defined by the following partition of $2K$,

$$\begin{aligned}
\nu = \nu(2K) &= (\nu_1, \ldots, \nu_{N_m}, \nu_{N_m+1}, \ldots, \nu_{2N_m}) \qquad (2.8)\\
&= (n_1n_2\,,\ n_1n_3\,,\ n_2n_3\,,\ \ldots\,,\ n_1n_m\,,\ n_2n_m\,,\ \ldots\,,\ n_{m-1}n_m,\\
&\quad\ n_1n_2\,,\ n_1n_3\,,\ n_2n_3\,,\ \ldots\,,\ n_1n_m\,,\ n_2n_m\,,\ \ldots\,,\ n_{m-1}n_m),
\end{aligned}$$

where $N_m = m(m-1)/2$. In addition, let

$$\rho = \rho(2K, \nu) = (\rho_1, \rho_2, \ldots, \rho_{N_m}, \ldots, \rho_{2N_m}),$$

where ρ_t is the sum of the first t elements of ν. Note that

$$n_1 n_2 + n_1 n_3 + n_2 n_3 + \cdots + n_1 n_m + n_2 n_m + \cdots + n_{m-1} n_m = K,$$

so we have $\rho_{t+N_m} = K + \rho_t$, $1 \le t \le N_m$. Let

$$\mathsf{b}(t) = (\rho_{t-1} + 1, \ldots, \rho_t), \quad 1 \le t \le 2N_m,$$

denote the "block index" corresponding to the t'th block of the vectors $\mathbf{x}$ and $\mathbf{y}$. Let $\Im$ be any Jacobi factor of order $2K$. For the blocks of $\Im$ and $\mathbf{x}$, $\mathbf{y}$ as well, we shall use the following notation

$$y_t = \mathbf{y}(\mathsf{b}(t)), \quad x_t = \mathbf{x}(\mathsf{b}(t)), \quad \Im_{t_1,t_2} = \Im(\mathsf{b}(t_1), \mathsf{b}(t_2)), \quad 1 \le t, t_1, t_2 \le 2N_m.$$

Thus,

$$\mathbf{y} = \begin{bmatrix} y_1, \\ y_2 \\ \vdots \\ y_{2N_m} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1, \\ x_2 \\ \vdots \\ x_{2N_m} \end{bmatrix}, \quad \Im = \begin{bmatrix} \Im_{1,1} & \Im_{1,2} & \cdots & \Im_{1,2N_m} \\ \Im_{2,1} & \Im_{2,2} & \cdots & \Im_{2,2N_m} \\ \vdots & \vdots & \ddots & \vdots \\ \Im_{2N_m,1} & \Im_{2N_m,2} & \cdots & \Im_{2N_m,2N_m} \end{bmatrix}.$$

Let

$$\tau(i,j) = \begin{cases} (j-1)(j-2)/2 + i, & 1 \le i < j \le m, \\ \tau(j,i) + N_m, & 1 \le j < i \le m \end{cases} \tag{2.9}$$

be the function which counts how many steps are needed within one cycle to reach the stage when X_{ij} and X_{ji} become the pivot blocks under the column-cyclic strategy. Since we have assumed that the step counter starts with zero, then within the current cycle, at step $k = \tau(i,j) - 1$, $i < j$, X_{ij} and X_{ji} are the pivot blocks. If $i > j$, then X_{ij} and X_{ji} are again the pivot blocks, but in the next cycle.

Each Jacobi factor *differs from the identity matrix only in certain principal submatrices obtained at the intersection of two block-rows and block-columns*. We shall now indicate their position and their structure. These submatrices can be nicely expressed by help of the Kronecker matrix product, denoted by $\otimes$.

THEOREM 2.1. *Let $\pi = (n_1, \ldots, n_m)$ be a partition of n such that $n \ge m \ge 2$ and let K be as in the relation (2.2).*

Let $1 \le i < j \le m$. Let $\mathbf{P}_{ij}$ and $\mathbf{Q}_{ij}$ be elementary block matrices and let $\Im = \Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$ be the associated Jacobi annihilator. Then $\Im$ differs from the identity matrix I_{2K} in exactly $2m-2$ principal submatrices. Using the function τ from the relation (2.9), these submatrices can be written in the following form:

$$\Im_{\tau(i,j),\tau(i,j)} = O, \qquad \Im_{\tau(j,i),\tau(j,i)} = O,$$

$$\begin{bmatrix} \Im_{\tau(i,r),\tau(i,r)} & \Im_{\tau(i,r),\tau(j,r)} \\ \Im_{\tau(j,r),\tau(i,r)} & \Im_{\tau(j,r),\tau(j,r)} \end{bmatrix} = \begin{cases} \begin{bmatrix} P_{ii}^* \otimes I_{n_r} & P_{ji}^* \otimes I_{n_r} \\ P_{ji}^* \otimes I_{n_r} & P_{jj}^* \otimes I_{n_r} \end{bmatrix} = \hat{\mathbf{P}}^* \otimes I_{n_r},\ 1 \le r \le i-1 \\ \begin{bmatrix} I_{n_r} \otimes P_{ii}^* & S\,(P_{ji}^* \otimes I_{n_r}) \\ \tilde{S}\,(I_{n_r} \otimes P_{ij}^*) & P_{jj}^* \otimes I_{n_r} \end{bmatrix},\ i+1 \le r \le j-1 \\ \begin{bmatrix} I_{n_r} \otimes P_{ii} & I_{n_r} \otimes P_{ij} \\ I_{n_r} \otimes P_{ji} & I_{n_r} \otimes P_{jj} \end{bmatrix}^*, \quad j+1 \le r \le m \end{cases}$$

$$\begin{bmatrix} \Im_{\tau(r,i),\tau(r,i)} & \Im_{\tau(r,i),\tau(r,j)} \\ \Im_{\tau(r,j),\tau(r,i)} & \Im_{\tau(r,j),\tau(r,j)} \end{bmatrix} = \begin{cases} \begin{bmatrix} Q_{ii}^T \otimes I_{n_r} & Q_{ji}^T \otimes I_{n_r} \\ Q_{ij}^T \otimes I_{n_r} & Q_{jj}^T \otimes I_{n_r} \end{bmatrix} = \hat{\mathbf{Q}}^T \otimes I_{n_r},\ 1 \le r \le i-1 \\ \begin{bmatrix} I_{n_r} \otimes Q_{ii}^T & S\,(Q_{ji}^T \otimes I_{n_r}) \\ \tilde{S}\,(I_{n_r} \otimes Q_{ij}^T) & Q_{jj}^T \otimes I_{n_r} \end{bmatrix},\ i+1 \le r \le j-1 \\ \begin{bmatrix} I_{n_r} \otimes Q_{ii} & I_{n_r} \otimes Q_{ij} \\ I_{n_r} \otimes Q_{ji} & I_{n_r} \otimes Q_{jj} \end{bmatrix}^T, \quad j+1 \le r \le m, \end{cases}$$

where

$$S = \begin{bmatrix} I_{n_i} \otimes e_1^T \\ \vdots \\ I_{n_i} \otimes e_{n_r}^T \end{bmatrix} = [I_{n_r} \otimes \tilde{e}_1 \ \ \ldots \ \ I_{n_r} \otimes \tilde{e}_{n_i}], \quad \tilde{S} = \begin{bmatrix} I_{n_r} \otimes \hat{e}_1^T \\ \vdots \\ I_{n_r} \otimes \hat{e}_{n_j}^T \end{bmatrix} = [I_{n_j} \otimes e_1 \ \ \ldots \ \ I_{n_j} \otimes e_{n_r}].$$

Here, e_i, $\tilde{e}_i$ and $\hat{e}_i$ denote the ith column of I_{n_r}, I_{n_i} and I_{n_j}, respectively.

Let $i = j$. Then $\Im = \Im_{ii}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$ differs from the identity matrix I_{2K} in exactly $2m-2$ principal submatrices, which can be written in the form:

$$\Im_{\tau(i,r),\tau(i,r)} = \begin{cases} P_{ii}^* \otimes I_{n_r} = \hat{\mathbf{P}}^* \otimes I_{n_r}, & 1 \le r \le i-1 \\ I_{n_r} \otimes P_{ii}^* = I_{n_r} \otimes \hat{\mathbf{P}}^*, & i+1 \le r \le m \end{cases}$$

$$\Im_{\tau(r,i),\tau(r,i)} = \begin{cases} Q_{ii}^T \otimes I_{n_r} = \hat{\mathbf{Q}}^T \otimes I_{n_r}, & 1 \le r \le i-1 \\ I_{n_r} \otimes Q_{ii}^T = I_{n_r} \otimes \hat{\mathbf{Q}}^T, & i+1 \le r \le m \end{cases}.$$

Proof. Let us fix i, j, $i < j$ and denote $\Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}})$ simply by $\Im$. The matrix $\Im$ has order $2K$ which is given by (2.2).

Obviously, the block rows of $\Im$, with subscripts $\tau(i,j)$ and $\tau(j,i) = \tau(i,j) + N_m$ have to be zero. Thus,

$$\Im_{\tau(i,j),t} = O, \quad \Im_{\tau(j,i),t} = O, \quad 1 \le t \le 2N_m.$$

All information on the non-trivial principal submatrices of $\Im$ can be extracted from the relation (2.6). So, let us consider the first two equations from (2.6). They should be combined with the fact that the pivot blocks are annihilated. Then they split into the three relations

$$\text{(A1)} \quad \left.\begin{aligned} Y_{ir} &= P_{ii}^* X_{ir} + P_{ji}^* X_{jr} \\ Y_{jr} &= P_{ij}^* X_{ir} + P_{jj}^* X_{jr} \end{aligned}\right\} \ 1 \le r \le i-1$$

$$\text{(A2)} \quad \left.\begin{aligned} Y_{ir} &= P_{ii}^* X_{ir} + P_{ji}^* X_{jr} \\ Y_{jr} &= P_{ij}^* X_{ir} + P_{jj}^* X_{jr} \end{aligned}\right\} \ i+1 \le r \le j-1$$

$$\text{(A3)} \quad \left.\begin{aligned} Y_{ir} &= P_{ii}^* X_{ir} + P_{ji}^* X_{jr} \\ Y_{jr} &= P_{ij}^* X_{ir} + P_{jj}^* X_{jr} \end{aligned}\right\} \ j+1 \le r \le m$$

We consider how these relations define $\Im$.

First, we consider the relations labeled (A3). If e_l denotes the lth column of I_{n_r}, then

$$\left.\begin{aligned} Y_{ir} e_l &= P_{ii}^* X_{ir} e_l + P_{ji}^* X_{jr} e_l, \quad 1 \le l \le n_r \\ Y_{jr} e_l &= P_{ij}^* X_{ir} e_l + P_{jj}^* X_{jr} e_l, \quad 1 \le l \le n_r \end{aligned}\right\} \ j+1 \le r \le m,$$

shows the rule how the l'th column of Y_{ir} (Y_{jr}) is obtained from the l'th columns of X_{ir} and X_{jr}. Note that $x_{\tau(i,r)} = [(X_{ir}e_1)^T, \ldots, (X_{ir}e_{n_r})^T]^T$ and similar for $y_{\tau(i,r)}$. Hence, for each $j+1 \le r \le m$, we have

$$\begin{bmatrix} \Im_{\tau(i,r),\tau(i,r)} & \Im_{\tau(i,r),\tau(j,r)} \\ \Im_{\tau(j,r),\tau(i,r)} & \Im_{\tau(j,r),\tau(j,r)} \end{bmatrix} = \left[\begin{array}{ccc|ccc} P_{ii}^* & & & P_{ji}^* & & \\ & \ddots & & & \ddots & \\ & & P_{ii}^* & & & P_{ji}^* \\ \hline P_{ij}^* & & & P_{jj}^* & & \\ & \ddots & & & \ddots & \\ & & P_{ij}^* & & & P_{jj}^* \end{array}\right] = \begin{bmatrix} I_{n_r} \otimes P_{ii} & I_{n_r} \otimes P_{ij} \\ I_{n_r} \otimes P_{ji} & I_{n_r} \otimes P_{jj} \end{bmatrix}^*$$

Let us now consider the relation (A1). If $\tilde{e}_l$ denotes the lth column of I_{n_i} and $\hat{e}_l$ the lth column of I_{n_j}, then for all $1 \le r \le i-1$

$$\left.\begin{aligned} \tilde{e}_l^T Y_{ir} &= \tilde{e}_l^T P_{ii}^* X_{ir} + \tilde{e}_l^T P_{ji}^* X_{jr}, \quad 1 \le l \le n_i \\ \hat{e}_l^T Y_{jr} &= \hat{e}_l^T P_{ij}^* X_{ir} + \hat{e}_l^T P_{jj}^* X_{jr}, \quad 1 \le l \le n_j \end{aligned}\right\} \; 1 \le r \le i-1$$

shows the rule how the l'th row of Y_{ir} (Y_{jr}) is obtained from the matrices X_{ir} and X_{jr}. Transposing the equations yields

$$\left.\begin{aligned} Y_{ir}^T \tilde{e}_l &= X_{ir}^T \bar{P}_{ii} \tilde{e}_l + X_{jr}^T \bar{P}_{ji} \tilde{e}_l, \quad 1 \le l \le n_i \\ Y_{jr}^T \hat{e}_l &= X_{ir}^T \bar{P}_{ij} \hat{e}_l + X_{jr}^T \bar{P}_{jj} \hat{e}_l, \quad 1 \le l \le n_j \end{aligned}\right\} \; 1 \le r \le i-1$$

Since $i > r$ and $j > r$, we have $x_{\tau(i,r)} = [\tilde{e}_1^T X_{ir}, \ldots, \tilde{e}_{n_i}^T X_{ir}]^T$ and similar for $y_{\tau(i,r)}$. Note also that

$$X_{ir}^T \bar{P}_{ii} e_l = \sum_{s=1}^{n_i} (\bar{P}_{ii})_{sl} (X_{ir}^T \tilde{e}_s), \quad 1 \le l \le n_i,$$

and similar holds for $\tilde{e}_l^T P_{ji}^* X_{jr}$. This implies

$$\Im_{\tau(i,r),\tau(i,r)} = \begin{bmatrix} (\bar{P}_{ii})_{11} I_{n_r} & (\bar{P}_{ii})_{21} I_{n_r} & \cdots & (\bar{P}_{ii})_{n_i,1} I_{n_r} \\ (\bar{P}_{ii})_{12} I_{n_r} & (\bar{P}_{ii})_{22} I_{n_r} & \cdots & (\bar{P}_{ii})_{n_i,2} I_{n_r} \\ \vdots & \vdots & \ddots & \vdots \\ (\bar{P}_{ii})_{1,n_1} I_{n_r} & (\bar{P}_{ii})_{2,n_i} I_{n_r} & \cdots & (\bar{P}_{ii})_{n_i n_i} I_{n_r} \end{bmatrix} \begin{matrix} n_r \\ n_r \\ \vdots \\ n_r \end{matrix} = P_{ii}^* \otimes I_{n_r}.$$

and similarly

$$\Im_{\tau(i,r),\tau(j,r)} = P_{ji}^* \otimes I_{n_r}.$$

In a similar way we obtain $\Im_{\tau(j,r),\tau(i,r)} = P_{ji}^* \otimes I_{n_r}$, $\Im_{\tau(j,r),\tau(j,r)} = P_{jj}^* \otimes I_{n_r}$. Thus,

$$\begin{bmatrix} \Im_{\tau(i,r),\tau(i,r)} & \Im_{\tau(i,r),\tau(j,r)} \\ \Im_{\tau(j,r),\tau(i,r)} & \Im_{\tau(j,r),\tau(j,r)} \end{bmatrix} = \begin{bmatrix} P_{ii}^* \otimes I_{n_r} & P_{ji}^* \otimes I_{n_r} \\ P_{ji}^* \otimes I_{n_r} & P_{jj}^* \otimes I_{n_r} \end{bmatrix} = \hat{\mathbf{P}}^* \otimes I_{n_r},\ 1 \le r \le i-1.$$

Let us now consider the relation (A2). If e_l denotes the lth column of I_{n_r} and $\hat{e}_l$ the lth column of I_{n_j}, then for all $i+1 \le r \le j-1$

$$\left.\begin{aligned} Y_{ir} e_l &= P_{ii}^* X_{ir} e_l + P_{ji}^* X_{jr} e_l, \quad 1 \le l \le n_r \\ \hat{e}_l^T Y_{jr} &= \hat{e}_l^T P_{ij}^* X_{ir} + \hat{e}_l^T P_{jj}^* X_{jr}, \quad 1 \le l \le n_j \end{aligned}\right\} \quad i+1 \le r \le j-1, \qquad (2.10)$$

shows the rule how the l'th column (row) of Y_{ir} (Y_{jr}) is obtained from the columns (rows) of X_{ir} and X_{jr}. This implies $\Im_{\tau(i,r),\tau(i,r)} = \mathrm{diag}(P_{ii}^*, \ldots, P_{ii}^*) = I_{n_r} \otimes P_{ii}^*$.

To obtain $\Im_{\tau(i,r),\tau(j,r)}$, we shall first find the auxiliary column-vector z and the matrix T such that $z = Tx_{\tau(j,r)}$ corresponds to the relation $Z = P_{ji}^* X_{jr}$, so that $z = \text{row}(Z)^T$. As earlier, one can find out that $T = P_{ji}^* \otimes I_{n_r}$. Note that Z has to be added to $P_{ii}^* X_{ir}$ which has been stored by columns as the column-vector $\Im_{\tau(i,r),\tau(i,r)} x_{\tau(i,r)}$. So, we have to add Z to $P_{ii}^* X_{ir}$ by columns. This means that before adding z to $\Im_{\tau(i,r),\tau(i,r)} x_{\tau(i,r)}$, we have to transform it to the column-vector $z' = \text{col}(Z)$. This can be done by the permutation matrix S of order $n_i n_r$, whose row- and column-partition has the form (here $\mathbf{e}_i$, e_i, $\tilde{e}_i$ denote the ith column of $I_{n_i n_r}$, I_{n_r}, I_{n_i}, respectively),

$$\begin{aligned} S &= [\mathbf{e}_1, \mathbf{e}_{1+n_r}, \ldots, \mathbf{e}_{1+(n_i-1)n_r}, \mathbf{e}_2, \mathbf{e}_{2+n_r}, \ldots, \mathbf{e}_{2+(n_i-1)n_r}, \ldots, \mathbf{e}_{n_r}, \mathbf{e}_{2n_r}, \ldots, \mathbf{e}_{n_i n_r}]^T \\ &= [\mathbf{e}_1, \mathbf{e}_{1+n_i}, \ldots, \mathbf{e}_{1+(n_r-1)n_i}, \mathbf{e}_2, \mathbf{e}_{2+n_i}, \ldots, \mathbf{e}_{2+(n_r-1)n_i}, \ldots, \mathbf{e}_{n_i}, \mathbf{e}_{2n_i}, \ldots, \mathbf{e}_{n_r n_i}] \\ &= \begin{bmatrix} I_{n_i} \otimes e_1^T \\ \vdots \\ I_{n_i} \otimes e_{n_r}^T \end{bmatrix} = [I_{n_r} \otimes \tilde{e}_1 \;\; I_{n_r} \otimes \tilde{e}_2 \;\; \ldots \;\; I_{n_r} \otimes \tilde{e}_{n_i}]. \end{aligned}$$

Thus, $z' = S((P_{ji}^* \otimes I_{n_r})\, x_{\tau(i,r)}) = S(P_{ji}^* \otimes I_{n_r})\, x_{\tau(i,r)}$, hence

$$\Im_{\tau(i,r),\tau(j,r)} = S\,(P_{ji}^* \otimes I_{n_r}) = \begin{bmatrix} P_{ji}^* \otimes e_1^T \\ \vdots \\ P_{ji}^* \otimes e_{n_r}^T \end{bmatrix}.$$

If we transpose the second equation in (2.10), we obtain

$$Y_{jr}^T \hat{e}_l = X_{ir}^T \bar{P}_{ij} \hat{e}_l + X_{jr}^T \bar{P}_{jj} \hat{e}_l.$$

Note that the columns of X_{jr}^T and of Y_{jr}^T are consecutively saved in $x_{\tau(j,r)}$ and $y_{\tau(j,r)}$. Therefore, the contribution to $y_{\tau(j,r)}$, coming from $P_{jj}^* X_{jr}$, is just

$$\begin{bmatrix} (\bar{P}_{jj})_{11} I_{n_r} & (\bar{P}_{jj})_{21} I_{n_r} & \cdots & (\bar{P}_{jj})_{n_j,1} I_{n_r} \\ (\bar{P}_{jj})_{12} I_{n_r} & (\bar{P}_{jj})_{22} I_{n_r} & \cdots & (\bar{P}_{jj})_{n_j,2} I_{n_r} \\ \vdots & \vdots & \ddots & \vdots \\ (\bar{P}_{jj})_{1,n_j} I_{n_r} & (\bar{P}_{jj})_{2,n_j} I_{n_r} & \cdots & (\bar{P}_{jj})_{n_j,n_j} I_{n_r} \end{bmatrix} x_{\tau(j,r)} = (P_{jj}^* \otimes I_{n_r})\, x_{\tau(j,r)}.$$

This implies

$$\Im_{\tau(j,r),\tau(j,r)} = P_{jj}^* \otimes I_{n_r}.$$

To obtain $\Im_{\tau(j,r),\tau(i,r)}$, we shall first find the auxiliary column-vector $\tilde{z}$ and the matrix $\tilde{T}$, such that $\tilde{z} = \tilde{T} x_{\tau(i,r)}$ corresponds to the relation $\tilde{Z} = P_{ij}^* X_{ir}$, so that $\tilde{z} = \text{col}(\tilde{Z})$. Since $x_{\tau(i,r)} = \text{col}(X_{ir})$, one can find as earlier, that $T = I_{n_r} \otimes P_{ij}^*$. Note that $\tilde{Z}$ has to be added to $P_{jj}^* X_{jr}$ by rows, since $P_{jj}^* X_{jr}$ is saved in the column-vector $\Im_{\tau(j,r),\tau(j,r)} x_{\tau(j,r)}$ by rows. So, before adding $\tilde{z}$ to $\Im_{\tau(j,r),\tau(j,r)} x_{\tau(j,r)}$, it has to be transformed to the column-vector $\tilde{z}' = [\text{row}(\tilde{Z})]^T$. This can be done by the permutation matrix $\tilde{S}$ of order $n_i n_r$, whose column- and row-partition has form (here $\mathbf{e}_i$, e_i, $\hat{e}_i$ denote the ith column of $I_{n_j n_r}$, I_{n_r}, I_{n_j}, respectively),

$$\begin{aligned} \tilde{S} &= [\mathbf{e}_1, \mathbf{e}_{1+n_r}, \ldots, \mathbf{e}_{1+(n_j-1)n_r}, \mathbf{e}_2, \mathbf{e}_{2+n_r}, \ldots, \mathbf{e}_{2+(n_j-1)n_r}, \ldots, \mathbf{e}_{n_r}, \mathbf{e}_{2n_r}, \ldots, \mathbf{e}_{n_j n_r}] \\ &= [\mathbf{e}_1, \mathbf{e}_{1+n_j}, \ldots, \mathbf{e}_{1+(n_r-1)n_j}, \mathbf{e}_2, \mathbf{e}_{2+n_j}, \ldots, \mathbf{e}_{2+(n_r-1)n_j}, \ldots, \mathbf{e}_{n_j}, \mathbf{e}_{2n_j}, \ldots, \mathbf{e}_{n_r n_j}]^T \\ &= \begin{bmatrix} I_{n_r} \otimes \hat{e}_1^T \\ \vdots \\ I_{n_r} \otimes \hat{e}_{n_j}^T \end{bmatrix} = [I_{n_j} \otimes e_1 \;\; I_{n_j} \otimes e_2 \;\; \ldots \;\; I_{n_j} \otimes e_{n_r}]. \end{aligned}$$

Thus, $z' = \tilde{S}((I_{n_r} \otimes P_{ij}^*)\, x_{\tau(i,r)}) = \tilde{S}(I_{n_r} \otimes P_{ij}^*)\, x_{\tau(i,r)}$, hence

$$\Im_{\tau(j,r),\tau(i,r)} = \tilde{S}\,(I_{n_r} \otimes P_{ij}^*) = \begin{bmatrix} I_{n_r} \otimes (P_{ij}\hat{e}_1)^* \\ \vdots \\ I_{n_r} \otimes (P_{ij}\hat{e}_{n_j})^* \end{bmatrix} = [P_{ij}^* \otimes e_1 \quad \ldots \quad P_{ij}^* \otimes e_{n_r})]\,.$$

The latest relations imply that for $i+1 \le r \le j-1$ holds

$$\begin{bmatrix} \Im_{\tau(i,r),\tau(i,r)} & \Im_{\tau(i,r),\tau(j,r)} \\ \Im_{\tau(j,r),\tau(i,r)} & \Im_{\tau(j,r),\tau(j,r)} \end{bmatrix} = \begin{bmatrix} I_{n_r} \otimes P_{ii}^* & S\,(P_{ji}^* \otimes I_{n_r}) \\ \tilde{S}\,(I_{n_r} \otimes P_{ij}^*) & P_{jj}^* \otimes I_{n_r} \end{bmatrix}.$$

Let us yet consider the third and fourth relation from (2.6). They should be combined with the fact that the pivot blocks are annihilated. Then they split into the three relations

$$\begin{array}{lll} \text{(B1)} & \left.\begin{array}{l} Y_{ri} = X_{ri}Q_{ii} + X_{jr}Q_{ji} \\ Y_{rj} = X_{ri}Q_{ij} + X_{jr}Q_{jj} \end{array}\right\} & 1 \le r \le i-1 \\ \text{(B2)} & \left.\begin{array}{l} Y_{ri} = X_{ri}Q_{ii} + X_{rj}Q_{ji} \\ Y_{rj} = X_{ri}Q_{ij} + X_{rj}Q_{jj} \end{array}\right\} & i+1 \le r \le j-1\,. \\ \text{(B3)} & \left.\begin{array}{l} Y_{ri} = X_{ri}Q_{ii} + X_{rj}Q_{ji} \\ Y_{rj} = X_{ri}Q_{ij} + X_{rj}Q_{jj} \end{array}\right\} & j+1 \le r \le m \end{array}$$

We consider how these relations define $\Im$.

First, we consider the relation (B1). If $\tilde{e}_l$ and $\hat{e}_l$ are the lth columns of I_{n_i} and I_{n_j}, respectively, then

$$\begin{array}{ll} Y_{ri}\tilde{e}_l = X_{ri}Q_{ii}\tilde{e}_l + X_{rj}Q_{ji}\tilde{e}_l, & 1 \le l \le n_i \\ Y_{rj}\hat{e}_l = X_{ri}Q_{ij}\hat{e}_l + X_{jr}Q_{jj}\hat{e}_l, & 1 \le l \le n_j, \end{array}$$

holds. Since $r < i$ and $r < j$ the columns of Y_{ri}, X_{ri} and X_{rj} take consecutive positions in the vectors $y_{\tau_{(r,i)}}$, $x_{\tau(r,i)}$, and $x_{\tau(r,j)}$. So, we can conclude as earlier,

$$\begin{bmatrix} \Im_{\tau(r,i),\tau(r,i)} & \Im_{\tau(r,i),\tau(r,j)} \\ \Im_{\tau(r,j),\tau(r,i)} & \Im_{\tau(r,j),\tau(r,j)} \end{bmatrix} = \begin{bmatrix} Q_{ii}^T \otimes I_{n_r} & Q_{ji}^T \otimes I_{n_r} \\ Q_{ij}^T \otimes I_{n_r} & Q_{jj}^T \otimes I_{n_r} \end{bmatrix} = \hat{\mathbf{Q}}^T \otimes I_{n_r},\ 1 \le r \le i-1.$$

Next, we consider the relation (B3). If e_l is the lth column of I_{n_r}, then

$$\begin{array}{l} e_l^T Y_{ri} = e_l^T X_{ri}Q_{ii} + e_l^T X_{rj}Q_{ji} \\ e_l^T Y_{rj} = e_l^T X_{ri}Q_{ij} + e_l^T X_{rj}Q_{jj} \end{array}\ , \quad 1 \le l \le n_r,$$

shows how the rows of Y_{ri} and Y_{rj} are obtained from the rows of X_{ri} and X_{rj}. Since $r > i$ and $r > j$, the transposed rows of Y_{ri}, X_{ri} and X_{rj} (that is the columns of Y_{ri}^T, X_{ri}^T and X_{rj}^T) take consecutive positions in the column vectors $y_{\tau_{(r,i)}}$, $x_{\tau(r,i)}$ and $x_{\tau(r,j)}$, respectively. Therefore, by transposing the last two relations, we obtain

$$\begin{array}{l} Y_{ri}^T e_l = Q_{ii}^T X_{ri}^T e_l + Q_{ji}^T X_{rj}^T e_l \\ Y_{rj}^T e_l = Q_{ij}^T X_{ri}^T e_l + Q_{jj}^T X_{rj}^T e_l \end{array}\ , \quad 1 \le l \le n_r.$$

From this, we obtain

$$\begin{bmatrix} \Im_{\tau(r,i),\tau(r,i)} & \Im_{\tau(r,i),\tau(r,j)} \\ \Im_{\tau(r,j),\tau(r,i)} & \Im_{\tau(r,j),\tau(r,j)} \end{bmatrix} = \begin{bmatrix} I_{n_r} \otimes Q_{ii} & I_{n_r} \otimes Q_{ij} \\ I_{n_r} \otimes Q_{ji} & I_{n_r} \otimes Q_{jj} \end{bmatrix}^T,\ j+1 \le r \le m.$$

The result for the case $i+1 \le r \le j-1$ is obtained from the relation (B2) in a similar way as we have obtained the appropriate result from the relation (A2). One can also use the argument following the relation (2.13).

Finally, for the case $i = j$, we now easily conclude that $\Im_{ii}$ is up to a permutational similarity a direct sum of an identity matrix and of the matrices

$$\Im_{\tau(i,r),\tau(i,r)} \quad \text{and} \quad \Im_{\tau(r,i),\tau(r,i)}, \quad 1 \le r \le m,\ r \ne i, \tag{2.11}$$

which are described above. This proves the theorem. □

We see that $\Im_{ij}$, $i < j$, is up to a permutational similarity a direct sum of an identity matrix, of the null matrix of order $2n_i n_j$ and of the matrices

$$\begin{bmatrix} \Im_{\tau(i,r),\tau(i,r)} & \Im_{\tau(i,r),\tau(j,r)} \\ \Im_{\tau(j,r),\tau(i,r)} & \Im_{\tau(j,r),\tau(j,r)} \end{bmatrix}, \begin{bmatrix} \Im_{\tau(r,i),\tau(r,i)} & \Im_{\tau(r,i),\tau(r,j)} \\ \Im_{\tau(r,j),\tau(r,i)} & \Im_{\tau(r,j),\tau(r,j)} \end{bmatrix}, \quad \begin{array}{c} 1 \le r \le m, \\ r \notin \{i,j\}, \end{array} \tag{2.12}$$

which are described above.

Note that $\mathbf{Y} = \mathcal{Z}_{ij}(\mathbf{P}^*\mathbf{X}\mathbf{Q})$ implies $\mathbf{Y}^* = \mathcal{Z}_{ij}(\mathbf{Q}^*\mathbf{X}^*\mathbf{P})$ and the latter yields

$$\begin{bmatrix} \bar{\mathbf{y}}_2 \\ \bar{\mathbf{y}}_1 \end{bmatrix} = \Im_{ij}(\hat{\mathbf{Q}}, \hat{\mathbf{P}}) \begin{bmatrix} \bar{\mathbf{x}}_2 \\ \bar{\mathbf{x}}_1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \mathbf{y}_2 \\ \mathbf{y}_1 \end{bmatrix} = \Im_{ij}(\hat{\bar{\mathbf{Q}}}, \hat{\bar{\mathbf{P}}}) \begin{bmatrix} \mathbf{x}_2 \\ \mathbf{x}_1 \end{bmatrix}. \tag{2.13}$$

In the vectors $[\mathbf{x}_2^T\ \mathbf{x}_1^T]^T$ and $[\mathbf{y}_2^T\ \mathbf{y}_1^T]^T$, the subscripts $\tau(r,i)$, $\tau(r,j)$ ($\tau(i,r)$, $\tau(j,r)$) address the blocks of $\mathbf{x}_2$ and $\mathbf{y}_2$ ($\mathbf{x}_1$ and $\mathbf{y}_1$). Taking into account (2.13), we can understand why in the assertion of the theorem, the right-hand sides in the second set of formulas are obtained from the right-hand sides in the first set of formulas by replacing $\hat{\mathbf{P}}$ with $\hat{\bar{\mathbf{Q}}}$. From (2.13) one can conclude that $\Im_{ij}(\hat{\mathbf{P}}, \hat{\mathbf{Q}}) = \mathbf{J}^T \Im_{ij}(\hat{\bar{\mathbf{Q}}}, \hat{\bar{\mathbf{P}}})\mathbf{J}$ for a suitable permutation (block-transposition) matrix $\mathbf{J}$.

The following result is important since the most important Jacobi-type methods use unitary transformations.

PROPOSITION 2.2. *If the elementary block matrices $\mathbf{P}_{ij}$ and $\mathbf{Q}_{ij}$, $i \le j$ are unitary, then the matrices defined in (2.12) and (2.11) are also unitary and*

$$\|\Im\|_2 = 1\,. \tag{2.14}$$

Proof. It suffices to prove that the matrices from the relations (2.12) and (2.11) (the latter in the case $i = j$) are unitary. Let us first consider the matrices from (2.12).

The Kronecker product of unitary matrices is unitary. So, we have to check the six types of matrices given in Theorem 2.1. For simplicity we shall denote each of of these matrices by $\hat{\Im}_r$.

For $1 \le r \le i-1$, we have either $\hat{\Im}_r = \hat{\mathbf{P}}^* \otimes I_{n_r}$ or $\hat{\Im}_r = \hat{\mathbf{Q}}^T \otimes I_{n_r}$. In the both cases the matrices $\hat{\Im}_r$ are unitary as the Kronecker product of unitary matrices.

For $j+1 \le r \le m$, we see that $\hat{\Im}_r$ is up to a permutational similarity a direct sum of n_r matrices $\hat{\mathbf{P}}^*$ or n_r matrices $\hat{\mathbf{Q}}^T$. Since $\hat{\mathbf{P}}^*$ and $\hat{\mathbf{Q}}^T$ are unitary, such is also each $\hat{\Im}_r$.

For $i+1 \le r \le j-1$, we first consider the case

$$\hat{\Im}_r = \begin{bmatrix} I_{n_r} \otimes P_{ii}^* & S\,(P_{ji}^* \otimes I_{n_r}) \\ \tilde{S}\,(I_{n_r} \otimes P_{ij}^*) & P_{jj}^* \otimes I_{n_r} \end{bmatrix}, \quad i+1 \le r \le j-1$$

To check whether $\hat{\Im}_r$ is unitary, we make the product

$$\hat{\Im}_r^* \hat{\Im}_r = \begin{bmatrix} I_{n_r} \otimes P_{ii} & (I_{n_r} \otimes P_{ij})\tilde{S}^T \\ (P_{ji} \otimes I_{n_r})S^T & P_{jj} \otimes I_{n_r} \end{bmatrix} \begin{bmatrix} I_{n_r} \otimes P_{ii}^* & S\,(P_{ji}^* \otimes I_{n_r}) \\ \tilde{S}\,(I_{n_r} \otimes P_{ij}^*) & P_{jj}^* \otimes I_{n_r} \end{bmatrix} = \begin{bmatrix} E & F \\ G & H \end{bmatrix}.$$

Next, we compute the four blocks of $\hat{\Im}_r\hat{\Im}_r^*$. Since $G = F^*$, it suffices to determine E, F and H. Assuming that the matrix product has higher priority than the Kronecker product, we have

$$\begin{aligned} E &= (I_{n_r} \otimes P_{ii})\,(I_{n_r} \otimes P_{ii}^*) + (I_{n_r} \otimes P_{ij})\tilde{S}^T\tilde{S}\,(I_{n_r} \otimes P_{ij}^*) \\ &= I_{n_r} \otimes P_{ii}P_{ii}^* + I_{n_r} \otimes P_{ij}P_{ij}^* = I_{n_r} \otimes (P_{ii}P_{ii}^* + P_{ij}P_{ij}^*) = I_{n_r} \otimes I_{n_i} = I_{n_r n_i}, \end{aligned}$$

$$\begin{aligned} F &= (I_{n_r} \otimes P_{ii})\,S\,(P_{ji}^* \otimes I_{n_r}) + (I_{n_r} \otimes P_{ij})\tilde{S}^T(P_{jj}^* \otimes I_{n_r}) \\ &= (I_{n_r} \otimes P_{ii}) \begin{bmatrix} I_{n_i} \otimes e_1^T \\ \vdots \\ I_{n_i} \otimes e_{n_r}^T \end{bmatrix} (P_{ji}^* \otimes I_{n_r}) + (I_{n_r} \otimes P_{ij}) \begin{bmatrix} I_{n_j} \otimes e_1^T \\ \vdots \\ I_{n_j} \otimes e_{n_r}^T \end{bmatrix} (P_{jj}^* \otimes I_{n_r}) \\ &= \begin{bmatrix} P_{ii} \otimes e_1^T \\ \vdots \\ P_{ii} \otimes e_{n_r}^T \end{bmatrix} (P_{ji}^* \otimes I_{n_r}) + \begin{bmatrix} P_{ij} \otimes e_1^T \\ \vdots \\ P_{ij} \otimes e_{n_r}^T \end{bmatrix} (P_{jj}^* \otimes I_{n_r}) \\ &= \begin{bmatrix} P_{ii}P_{ji}^* \otimes e_1^T \\ \vdots \\ P_{ii}P_{ji}^* \otimes e_{n_r}^T \end{bmatrix} + \begin{bmatrix} P_{ij}P_{jj}^* \otimes e_1^T \\ \vdots \\ P_{ij}P_{jj}^* \otimes e_{n_r}^T \end{bmatrix} = \begin{bmatrix} (P_{ii}P_{ji}^* + P_{ij}P_{jj}^*) \otimes e_1^T \\ \vdots \\ (P_{ii}P_{ji}^* + P_{ij}P_{jj}^*) \otimes e_{n_r}^T \end{bmatrix} = O, \end{aligned}$$

$$\begin{aligned} H &= (P_{jj} \otimes I_{n_r})\left(P_{jj}^* \otimes I_{n_r}\right) + (P_{ji} \otimes I_{n_r})\left(P_{ji}^* \otimes I_{n_r}\right) \\ &= P_{jj}P_{jj}^* \otimes I_{n_r} + P_{ji}P_{ji}^* \otimes I_{n_r} = (P_{jj}P_{jj}^* + P_{ji}P_{ji}^*) \otimes I_{n_r} = I_{n_j} \otimes I_{n_r} = I_{n_j n_r}\,. \end{aligned}$$

The proof for for the remaining case is quite similar. The role of $\hat{\mathbf{P}}^*$ is played by the unitary matrix $\hat{\mathbf{Q}}^T$.

If $i = j$, we see from Theorem 2.1 that each $\hat{\Im}_r$ is the Kronecker product of two unitary matrices. This completes the proof. □

REFERENCES

[1] V. HARI, *On the global convergence of Eberlein method for real matrices*, Numer. Math., 39 (1982), pp. 361–369.

[2] V. HARI, *On the convergence of certain cyclic Jacobi-like norm-reducing processes for real matrices*, Radovi Matem., 1 (1985), pp. 121–147.

[3] V. HARI, *On the convergence of cyclic Jacobi-like processes*, Linear Algebra and Its Appl., 81 (1986), pp. 105–127.

[4] V. HARI, *Convergence to Diagonal Form of Block Jacobi Methods*, Preprint, University of Zagreb, 2008.

[5] P. HENRICI AND K. ZIMMERMANN, *An Estimate for the Norms of Certain Cyclic Jacobi Operators*, Linear Algebra Appl., 1 (1968), pp. 289–501.

[6] K. ZIMMERMANN, *On the Convergence of the Jacobi Process for Ordinary and Generalized Eigenvalue Problems*, Ph.D. thesis, ETH Zürich 1965, Dissertation No. 4305.