On a Matrix Identity Connecting Iteration Operators Associated with a p-Cyclic Matrix

Apostolos Hadjidimos

Dimitrios Noutsoss

Report Number:
91-058
ON A MATRIX IDENTITY CONNECTING ITERATION OPERATORS ASSOCIATED WITH A P-CYCLIC MATRIX

Apostolos Hadjidimos
Dimitrios Noutsos

CSD-TR-91-058
August 1991
On a Matrix Identity Connecting Iteration Operators Associated with a $p$-Cyclic Matrix

Apostolos Hadjidimos* and Dimitrios Noutsos†

Computer Science Department
Purdue University
West Lafayette, IN 47907

August 16, 1991

Abstract

The Successive Overrelaxation (SOR) and the Symmetric (S)SOR iteration matrices are connected with the Jacobi iteration matrix in case these operators are associated with a $(q, p-q)$-Generalized Consistently Ordered matrix through certain matrix identities. The validity of these identities has been proved in the last couple of years. Very recently an analogous matrix identity was shown to hold for the Modified (MSOR and Jacobi iteration matrices in the very particular cases $(p, q) = (2, 1), (3, 1)$ and $(3, 2)$. It is the main objective of this paper to extend the validity of this identity to cover an entire class of pairs $(p, q)$. The identity in question is not only of theoretical interest but of practical importance too since it can be used to show the equivalence of the MSOR and a class of $p$-step iterative methods for the solution of a linear system whose matrix coefficient possesses the $(q, p-q)$-generalized consistently ordered property.

Key words and phrases: $p$-cyclic matrices, Jacobi, Successive Overrelaxation (SOR) and Modified (M)SOR iteration matrices.

Abbreviated Title: Matrix identities for $p$-cyclic matrices.


*The work of this author was supported in part by NSF grant CCR-8619817 and by AFOSR grant 8810243.
†Department of Mathematics, University of Ioannina, GR-451 10 Ioannina, Greece.
1 Introduction and Preliminaries

Let us consider the matrix $A \in \mathbb{C}^{n \times n}$ and let us suppose that it is partitioned into $p \times p$ blocks and has the form

$$A = \begin{bmatrix}
A_{11} & 0 & \cdots & 0 & A_{1,p-q+1} & 0 & \cdots & 0 \\
0 & A_{22} & \cdots & 0 & 0 & A_{2,p-q+2} & \cdots & 0 \\
\vdots & & & & \vdots & & & \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & A_{q}
\end{bmatrix}, \quad (1.1)
$$

where its diagonal blocks are square and non-singular and where the integer $q$ is relatively prime to $p$. The matrix $A$ in (1.1) belongs to the class of block $p$-cyclic matrices [12] or more specifically to that of the generalized consistently ordered (GCO) $(q,p-q)$-matrices [14]. The Jacobi iteration matrix $T$ associated with $A$ above has the form

$$T = \begin{bmatrix}
0 & 0 & \cdots & 0 & T_{1,p-q+1} & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 0 & T_{2,p-q+2} & \cdots & 0 \\
\vdots & & & & \vdots & & & \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & T_{q}
\end{bmatrix}, \quad (1.2)
$$

where the non-identically zero blocks of it are given by

$$T_{ij} = \begin{cases}
-A_{ii}^{-1} A_{i,p-q+i}, & i = 1(1)q \\
-A_{ii}^{-1} A_{i,i-q}, & i = q + 1(1)p;
\end{cases} \quad (1.3)
$$

this matrix can also be written as

$$T = L + U \quad (1.4)
$$

where $L$ and $U$ are strictly lower and strictly upper triangular matrices respectively. Furthermore let us consider the Modified Successive Overrelaxation (MSOR) iteration matrix $L_\Omega$ associated with $A$ which, in view of (1.2)-(1.4), can be written as

$$L_\Omega := (I - \Omega L)^{-1} (I - \Omega + \Omega U) \quad (1.5)
$$

(see [9] and for $p = 2$ see also [14]). In (1.5) $\Omega := (\omega_1 I_1, \omega_2 I_2, \ldots, \omega_p I_p)$ where $I_j, j = 1(1)p$, are unit matrices of orders the respective ones of $A_{jj}$ and $\omega_j, j = 1(1)p$, are scalars. Note that for $\Omega = \omega I$ the MSOR matrix operator reduces to that of the SOR one, that is to
\( L_\omega := (I - \omega L)^{-1} \left( (1 - \omega)I + \omega U \right). \)  \hspace{1cm} (1.6)

It is known [9] that the sets of eigenvalues \( \mu \in \sigma(T) \) and \( \lambda \in \sigma(L_\omega) \) are connected through the relationship

\[
\prod_{j=1}^{p} (\lambda + \omega_j - 1) = \prod_{j=1}^{p} \omega_j \mu^{p-\eta}.
\]  \hspace{1cm} (1.7)

From (1.7) the well-known eigenvalue relationships of Young's [13] (SOR, \( p = 2, q = 1 \)), Varga's [11] (SOR, \( p \geq 2, q = 1 \)), Verner and Bernal's [10] (SOR, \( p \geq 2, p - 1 \geq q \geq 1 \)) and Young and Kingaid's [15] (MSOR, \( p = 2, q = 1 \)) are readily recovered.

Our main objective in this work is to derive the matrix analogue of the relationship (1.7). More specifically to show that the identity

\[
\prod_{j=1}^{p} (L_\omega + (\omega_j - 1)I) = (\Omega T)^p L_\omega^{p-\eta}
\]  \hspace{1cm} (1.8)
always holds. For this we note that for \( \Omega = \omega I \) (SOR) the corresponding to (1.8) identity was obtained in [2] while for the MSOR operator for \( p = 2, q = 1 \) was obtained by Young and Kincaid [15] and for \((p,q) = (3,1),(3,2)\) was obtained by Hadjidimos and Yeyios [4]. For the proof of (1.8) our main tools will be elementary graph theory (see, e.g., [12], [5], etc.) and simple combinatorics in the way these tools were effectively used in similar previous works (see [2], [3]). Apart from the theoretical interest the identity (1.8) presents, it is of practical importance too. More specifically, as was shown in [4], where the observation made by Hubner [7] ([6]) should be taken into consideration, the problem of the determination of "good" or "optimal" relaxation factors \( \omega_j, j = 1(1)p, \) for the solution of the linear system \( Ax = b, \) with \( A \) in (1.1) nonsingular, using the MSOR method is equivalent to that of the determination of a set of \( p \) parameters of a \( p \)-step iterative method "equivalent" to the MSOR one. The determination of the parameters in the latter problem may turn out to be simpler as this was in the case \( p = 2 \) [4], (see also [8], [1], [2], [3] for similar SOR and SSOR problems).

2 Main Result and Preliminary Analysis

The statement of our main result is given in the following theorem.

Theorem 2.1: Let \( T \) in (1.2) be the block Jacobi and \( L_\omega \) in (1.5) be the block MSOR iteration matrices associated with \( A \) in (1.1). Then for any diagonal matrix \( \Omega := \text{diag}(\omega_1 I_1, \omega_2 I_2, \ldots, \omega_p I_p), \) \( \omega_j \in \mathbb{C}, j = 1(1)p, \) with \( I_j \) the unit matrix of the order of \( A_{jj}, \) the matrices \( T \) and \( L_\omega \) satisfy (1.8). \( \Box \)
The proof of Theorem 2.1 will be given in Section 3 where a number of other auxiliary statements will also be given and proved. In this section we shall develop the necessary background material on which these proofs are based. Note that (1.8) is trivially satisfied for $\Omega = 0$; so we restrict to $\Omega \neq 0$. Note also that $\omega_j T_{jk}$, $j = 1(1)p$, are of exactly the same form as the blocks $T_{jk}$ in (1.3). So, in view of (1.2) and (1.4), $\Omega T$, $\Omega L$ and $\Omega U$ will be denoted from now on by $T$, $L$ and $U$ respectively. With this simplified notation (1.8) is written as

$$\prod_{j=1}^{p}(L_n + (\omega_j - 1)I) = TP_L^{p-q},$$

(2.1)

where

$$L_{\Omega} := (I - L)^{-1}(I - \Omega + U).$$

(2.2)

Also in the present work only the case $q < p - q$ will be analyzed and studied. The corresponding analysis when $q > p - q$ is very similar and is therefore omitted while for $q = p - q$, actually $p = 2$, $q = 1$, it is given in [14].

Assuming that we are referring to the set $P := \{1, 2, \ldots, p\}$ of nodes we begin our analysis by noting that in the directed graph $G(T)$ of $T$ (see [12] or [5]) for any given node $i = q + 1(1)p$ we have

$$\bigcup_{j=1}^{p}(i,j) = (i, i - q),$$

as in Figure 1, while for any $i = 1(1)q$ it will be

$$\bigcup_{j=1}^{p}(i,j) = (i, i + p - q),$$

(2.3)

(2.4)

as in Figure 2. To distinguish the edges of type I which are associated with the matrix $L$ from those of type II associated with $U$ we shall refer to an edge of type II as a "folding" edge. From (2.3) and (2.4) we have that

$$G(L) = \bigcup_{i=q+1}^{p}(i, i - q), \quad G(U) = \bigcup_{i=1}^{p}(i, i + p - q)$$

$$G(T) = G(L) \cup G(U).$$

(2.5)

Let us denote by $B$ and $C$ the two members of (2.1). Specifically

$$B := \prod_{\ell=0}^{p}(L_{\Omega} + (\omega - 1)I), \quad C := TP_L^{p-q}.$$ 

(2.6)

$B$ can be written in terms of powers of $L_{\Omega}$ as follows

$$B = \sum_{\ell=0}^{p}(-1)^{s}L_{\Omega}^{p-\ell},$$

(2.7)
with

$$\sigma_0 = 1, \quad \sigma_\ell = \sum_{k=1}^{\ell} (1 - \omega_{r_k}), \quad \ell = 1(1)p,$$

where \(1 \leq r_1 < r_2 \ldots < r_\ell \leq p\) and the summation in (2.8) extends over all possible combinations of the \(p\) nodes (points) in \(P\) chosen \(\ell\). Note that in both \(B\) and \(C\) various powers of \(L_0\) are involved. Since our intention is to express, indirectly, \(B\) and \(C\) in terms of products of \(L\) and \(U\) the expansion of \(L_0\) in \(L's\) and \(U's\) is needed. For this we have

$$L_0 = (I - L)^{-1}(I - \Omega + U) = L_{0,1} + L_{0,2}$$

where

$$L_{0,1} := (I + L + L^2 + \ldots + L^\ell)(I - \Omega)$$
$$L_{0,2} := (I + L + L^2 + \ldots + L^\ell)U$$

and \(\ell = [(p - 1)/q]\), with \([x]\) denoting the integral part of the real number \(x\).

Let us now examine how \(G(L_0)\) is derived from \(G(L)\) and \(G(U)\) considering all scalar coefficients \(1 - \omega_j\), \(j = 1(1)p\), even if one (or more) of them is zero. Obviously \(G(I - \Omega) = \bigcup_{i=1}^{p} (i,i)\) consists of \(p\) closed paths of length zero each, which will be called identity paths (see Figure 3). Since the scalar coefficient associated with the \(i\)th identity path is \(1 - \omega_i\) we may consider the graph of \(L_{0,1}\) in (2.9)-(2.10) as being the "weighted" graph of \(I + L + L^2 + \ldots + L^\ell\). All the weighted paths in \(G(L_{0,1})\), which are of lengths 0, 1, 2, \ldots, \(\ell\) and are coming from the terms \(I - \Omega\), \(L(I - \Omega)\), \(L^2(I - \Omega)\), \ldots, \(L^\ell(I - \Omega)\), respectively, will be indicated by a double arrow and will be represented by a single edge of type I or III, whichever applies. Hence

$$G(L_{0,1}) = \bigcup_{i=1}^{p} \left[ \bigcup_{j=0}^{\ell} (i,i - jq) \right].$$

Thus for a particular set of indices \(i, i - jq \in P\) the weighted path \((i,i - jq)\) in (2.11) is of length \(j\), has weight \(1 - \omega_{i-jq}\) and is the graph of the nonidentically zero \((i,i - jq)\)th block of the term \(L^j(I - \Omega)\) of \(L_{0,1}\).

Example 1: Let \(p = 5\) and \(q = 2\).

a) For \(i = 4\) and \(j = 1\) the path in \(G(L_{0,1})\) is \((4,2)\), has length 1, weight \(1 - \omega_2\) and comes from the term \(L(I - \Omega)\) of \(L_{0,1}\) whose block in the position \((4,2)\) is \((1 - \omega_2)T_{42}\).

b) For \(i = 5\) and \(j = 2\) the path \((5,1)\) has length 2, weight \(1 - \omega_1\) and comes from the term \(L^2(I - \Omega)\) whose block in the \((5,1)\) position is \((1 - \omega_1)T_{51}T_{31}\).
On the other hand, all the paths in \( G(\Omega, \tau) \), which are of lengths 1, 2, \ldots, \( t + 1 \) and are coming from the terms \( U, LU, L^2 U, \ldots, L^t U \) respectively, will have a folding as their last edge, will be indicated by a single arrow and will be represented in \( G(\Omega, \tau) \) by a single folding. Therefore

\[
G(\Omega, \tau) = \bigcup_{i=1}^{p} (i, i + p - ([\frac{i-1}{q}] + 1)q)
\]

and hence

\[
G(\Omega) = G(\Omega, 1) \cup G(\Omega, \tau)
\]

\[
= \bigcup_{i=1}^{p} \left\{ \bigcup_{j=0}^{\left[\frac{i-1}{q}\right]} (i, i - jq) \bigcup (i, i + p - ([\frac{i-1}{q}] + 1)q) \right\}.
\]

The subgraph of \( G(\Omega) \) that contains only the paths that have origin the node \( i \in P \), have lengths 0, 1, 2, \ldots, \( \left[\frac{i-1}{q}\right], \left[\frac{i-1}{q}\right] + 1 \) and are coming from the terms \( I - \Omega, L(I - \Omega), L^2(I - \Omega), \ldots, L^{\left[\frac{i-1}{q}\right]}(I - \Omega), L^{\left[\frac{i-1}{q}\right]+1}U \) of \( \Omega \), is illustrated in Figure 4.

**Example 2:** For \( p = 5 \) and \( q = 2 \) the subgraph of \( G(\Omega) \) whose paths have the node \( i = 5 \) as their origin is the union of the paths \((5, 5), (5, 3), (5, 1), \) and \((5, 4)\). These paths have lengths 0, 1, 2 and 3, weights \( 1 - \omega_5, 1 - \omega_3, 1 - \omega_1, 1 \), and are coming from the terms \( I - \Omega, L(I - \Omega), L^2(I - \Omega) \) and \( L^3 U \) and represent the blocks \((1 - \omega_5)T_{53}, (1 - \omega_3)T_{53}T_{31}, (1 - \omega_1)T_{53}T_{31}, \) and \( T_{53}T_{31} \) of \( \Omega \).

To prove Theorem 2.1 we will show that \( B_{ij} = C_{ij} \) for every pair \( i, j \in P \). For this it is assumed that \( B \) and \( C \) in (2.6) have been expanded in terms of nonidentically zero products of \( L \)'s and \( U \)'s and all the like terms, if any, have been summed up. The graph of each term in either \( B \) or \( C \) consists of the union of one or more paths. Each path consists of consecutive subpaths and represents the graph of a nonidentically zero block of the term in question. Our objective will be then accomplished if we show that all paths in \( G(B) \) and \( G(C) \) from \( i \) to \( j \) with \( m \) foldings \( (0 \leq m \leq p) \) coincide and are associated with equal overall weights. Obviously, any two “different” paths in \( G(B_{ij}) \) and \( G(C_{ij}) \) with a particular number of \( m \) foldings are “identical” except for possible double-arrowed identity subpaths and the weights of type I subpaths. These identity subpaths can be practically anywhere between any two subpaths of such a path. All the above “identical” paths will be considered as one with which an overall weight will be associated. This overall weight will be equal to the sum of all the weights associated with all the “identical” paths. The determination of this weight constitutes the basic key to the proof of our main result. For this a number of propositions will be stated and proved in the next section before the proof of the main result is given. However, in order to make some of the points of the previous discussion clear to the reader we will give an example.

**Example 3:** Suppose that \( p = 3, q = 1 \) and let us try to find the contribution to the formation of the \((i, j)\)th block of \( \Omega^2 \) of the terms of the expansion of \( \Omega^2 \) whose graphs have \( m \) foldings in the following two cases:

a) When \( i = 2, j = 3, m = 2 \) where the last subpath in anyone of those graphs, with identity
subpaths being ignored, is a folding and

b) When \( i = 2, j = 1, m = 1 \) where the last subpath in these graphs is not a folding.

a) Based on the given information we shall try to find, ignoring weights, the term of \((L_3^2)_{23}\) specified as a product of nonidentically zero blocks. For this we have that the first factor in the product in question must be \(T_{21}\), the only nonidentically zero block in the second block row of \(L\) and \(U\) (or \(T\)), while the last factor must be \(T_{13}\). To find out which factors, if any, should be placed between \(T_{21}\) and \(T_{13}\) we must keep in mind the number of foldings \(m = 2\) and also the fact that the only nonidentically zero block of \(T\) whose graph is a folding is the block \(T_{13}\). This leads us to the conclusion that the only product satisfying these constraints is \(T_{21} T_{13} T_{32} T_{21} T_{13}\). It is obvious now that this product is the block in the position \((2,3)\) of \(LLL = LU L^2 U\). It remains then to find the weight associated with \(T_{21} T_{13} T_{32} T_{21} T_{13}\) or the sum of the weights of all the terms in the expansion of \((L_3^2)_{23}\) of the form \(LU L^2 U\). Having in mind that this product comes from a product of three factors of \(L\) two of which are coming from \(L_{1,2}\) due to the presence of the two \(U\)'s, and that the last subpath in the graphs we are considering is a folding it is implied that in the product in question the third factor must come from \(L_{1,1}\) unless, of course, it is followed by the factor \(L - \Omega\), from \(L_{1,1}\). All these requirements are satisfied only by the following six products: 

\[
(I - \Omega)(LU)(L^2 U), (L(I - \Omega)U(L^2 U), (LU)(I - \Omega)(L^2 U), (LU)(L(I - \Omega))(LU), (LU)(L^2(I - \Omega))(LU) \text{ and } (LU)(L^2 U)(I - \Omega) \text{. (Note: It is very easy to find these products provided one inserts the one factor } (I - \Omega), \text{ coming from the one term of } L_{1,1}, \text{ in all six possible positions of } LUL^2 U \text{ and finds those products that give acceptable products of three factors, one from } L_{1,1} \text{ and two from } L_{1,2}. \)

Having obtained the above products we can readily find their blocks in their \((2,3)\) position. These are \((T_{32} T_{21} T_{13} T_{21} T_{13}), (T_{21} T_{13} T_{32} T_{21} T_{13}), (T_{21} T_{13} T_{21} T_{13}((1 - \omega_3)T_{32} T_{21} T_{13}), (T_{21} T_{13} T_{32} T_{21} T_{13}), (T_{21} T_{13} T_{32} T_{21} T_{13})((1 - \omega_3)T_{32} T_{21} T_{13}),\) respectively. The weighted graphs of these six products are given in Figures 6a-6f. Note that except for weights and identity paths all six paths are identical and all the previous six products are equal to the \(T_{21} T_{13} T_{32} T_{21} T_{13}\) term each multiplied by the associated weight \((1 - \omega_2), (1 - \omega_1), (1 - \omega_2), (1 - \omega_1)\) and \((1 - \omega_3)\), respectively. Thus we have that

\[
(L_3^2)_{23}^{\text{II}} = [2(1 - \omega_1) + 2(1 - \omega_2) + 2(1 - \omega_3)] T_{21} T_{13} T_{32} T_{21} T_{13}, \tag{2.14}
\]

where, in the superscripts, 2 denotes the number of foldings and II that the last edge is a folding.

b) Based on a similar reasoning as in the previous case it is easy to see that

\[
(L_3^2)_{21}^{\text{II}} = \alpha_{21}^{\text{II}} T_{21} T_{13} T_{32} T_{21}, \tag{2.15}
\]
where the weight \( \alpha_{21}^{11} \) is to be determined. Obviously, the product in the right hand side of (2.15) can only come from the block in the (2,1) position of \( LUL^2 \). Since only one folding is involved two out of the three factors that form this product will come from \( L_0,1 \). Hence two \((I - \Omega)\) factors must be placed among the \( L's \) and \( U's \) of \( LUL^2 \) so that to produce acceptable products involving two terms from \( L_0,1 \) and one from \( L_0,2 \). It can be checked that only the following five possibilities exist:

\[
(I - \Omega)(LU)(L^2(I - \Omega)), (L(I - \Omega))U(L^2(I - \Omega)), (LU)(I - \Omega)(L^2(I - \Omega)), (LU)(L(I - \Omega))(L(I - \Omega)) \text{ and } (L^2)(I - \Omega)(I - \Omega).
\]

The (2,1) blocks of these products are \((1 - \omega_2)T_2(T_1T_1)(T_3T_1(1 - \omega_1)I_1), (T_2(1 - \omega_1)I_1)T_3(T_2T_1(1 - \omega_1)I_1), (T_2T_1)((1 - \omega_3)I_3)T_3T_1(1 - \omega_1)I_1), (T_3T_1)(T_3T_1(1 - \omega_2)I_2)(T_3(1 - \omega_1)I_1) \text{ and } (T_2T_1)(T_3T_2T_1(1 - \omega_1)I_1)((1 - \omega_1)I_1), \text{ respectively.}

The weighted graphs of these blocks are given in Figures 7a-7f. If the previous graphs are analyzed in the way suggested in the previous case (a) we will have the equivalent graphs given in Figures 8a-8e. All five graphs are then identical, in the sense already explained, and represent the term \( T_2T_1T_3T_2T_2T_1 \) multiplied by the associated weights \((1 - \omega_2)(1 - \omega_1), (1 - \omega_1)^2, (1 - \omega_3)(1 - \omega_1), (1 - \omega_2)(1 - \omega_1) \text{ and } (1 - \omega_1)^2, \) respectively. Hence

\[
\alpha_{21}^{11} = (1 - \omega_1)[2(1 - \omega_1) + 2(1 - \omega_2) + (1 - \omega_3)].
\]  

3 Proof of Theorem 2.1

We begin this section with the statement and proof of three lemmas which constitute the basis for the subsequent analysis. In these lemmas we give the theoretical formulation and proof of what we tried to make clear with Example 3 in the previous section.

Lemma 3.1: Let \( s \in P \). Then in \( G(L^*_n) \) all “identical” paths from \( i \) to \( j \) \((i, j \in P)\) with \( m \) foldings, \( 1 \leq m \leq s \), (in the sense explained in Example 3) whose last edge (ignoring identity subpaths) is a folding, can be considered as one path with an overall (scalar) weight \( \alpha_{ij}^{m, III} \) given by

\[
\alpha_{ij}^{m, III} = \sum_{\sum q_k = s - m} \prod_{k=1}^{P} \left( \frac{q_k + j_k - 1}{j_k - 1} \right) (1 - \omega_k)^{q_k}.
\]

In (3.1) \( q_k \) denotes the number of double-arrowed edges, including identity subpaths, that a path from \( i \) to \( j \) with \( m \) foldings consists of which have \( k \in P \) as an ending node. On the other hand \( j_k \) denotes the maximum number of times the node \( k \) can be an ending node of an edge, including identity subpaths, in any one of the “identical” paths considered. (Note: If one (or more) \( q_k \) is zero the factor \((1 - \omega_k)^0\) is defined to be 1 even if \( \omega_k = 1 \). Also, \( \begin{pmatrix} -1 \\ -1 \end{pmatrix} \) or \( \begin{pmatrix} 0 \\ 0 \end{pmatrix} \) are defined to be 1.)

Proof: To find \( \alpha_{ij}^{m, III} \) one has to find all the “identical” paths from \( i \) to \( j \) with \( m \) foldings together with their associated weights. These paths come from all the nonidentically zero \((i, j)\)th blocks of all terms of the expansion of \( L^*_n \), before any summing up takes place. As has been made
clear the number of foldings \( m \) indicates that in the product that forms the corresponding term in 
the expansion, with which such a path is associated, \( m \) factors will come from \( L_{\Omega,2} \) and \( s - m \) from 
\( L_{\Omega,1} \) implying that \( \sum_{k=1}^{p} q_k = s - m \). In addition, any factor from \( L_{\Omega,1} \) is either \( I - \Omega \), in which 
case it gives rise to identity double-arrowed subpaths, or \( L^w(I - \Omega) \), \( w = 1(1)4 \), when it produces 
weighted double-arrowed subpaths of length \( w \). Since the last subpath of any path of interest is, 
by assumption, a folding we must have \( j > p - q \). It is clear that to each \( k \in P \) and for given \( p \) 
and \( q \) there corresponds a unique \( j_k \) which is a constant and depends only on \( i \), \( j \), \( m \) and \( s \) while 
the \( q_k \)'s may vary but are such that \( \sum q_k = s - m \). Consider then the subset of all “identical” 
paths that correspond to a certain possible \( p \)-tuple \( \{q_1, q_2, \ldots, q_p\} \), \( \sum_{k=1}^{p} q_k = s - m \). It is obvious 
that with each node \( k \in P \) there is associated a scalar coefficient (weight) \( N_{q_k}(1 - \omega_k)^{q_k} \), with \( N_{q_k} \) 
being the number of all “identical” paths of the subset in question that have the node \( k \) as an 
ending node \( q_k \) times each. Recalling that our interest is focused only on paths that have the node 
\( k \) as an ending node of a double-arrowed path and this can only come from the graph of either 
\( (I - \Omega)_{kk} = (1 - \omega_k)j_k \) or \( (L^w(I - \Omega))_{kk} = (L^w)_{kk}(I - \Omega)_{kk} = (1 - \omega_k)(L^w)_{kk} \) it is implied that \( N_{q_k} \) is 
the number of times the above \( q_k \) identity subpaths can be “distributed” to the maximum number 
of \( j_k \) nodes \( k \) of the path. This number is equal to the number of combinations with repetitions of 
\( j_k \) chosen \( q_k \); that is
\[
N_{q_k} = \left( \frac{q_k + j_k - 1}{j_k - 1} \right), \quad k = 1(1)p.
\]
Hence the scalar coefficient (weight) associated with all “identical” paths of the subset considered 
that correspond to a certain \( p \)-tuple is given by
\[
\prod_{k=1}^{P} \left( \frac{q_k + j_k - 1}{j_k - 1} \right)(1 - \omega_k)^{q_k}.
\]
So, \( \alpha_{ij}^{m,ff} \) that is the scalar coefficient (overall weight) associated with all the “identical” paths of 
the present lemma is given by (3.1). \( \square \)

Note: In Example 3a it is \((j_1, j_2, j_3) = (2, 2, 2)\) (Figures 6a–6f) while \((q_1, q_2, q_3) = (1, 0, 0)\) 
(Figures 6b, e), \((0, 1, 0)\) (Figures 6a, d), \((0, 0, 1)\) (Figures 6c, f). \( \square \)

Lemma 3.2: Under the assumptions of Lemma 3.1, with the only difference being that the last 
subpath, except for identity subpaths, is a double-arrowed one, that is of type I, \( 0 \leq m \leq s - 1 \), 
the corresponding overall weight is given by
\[
\alpha_{ij}^{m,I} = (1 - \omega_j) \sum_{\sum q_k = s - m - 1} \prod_{k=1}^{P} \left( \frac{q_k + j_k - 1}{j_k - 1} \right)(1 - \omega_k)^{q_k}.
\]
In (3.4), \( q_k \) and \( j_k \) are the numbers defined in Lemma 3.1 except for \( q_j \), which is the number defined 
previously decreased by 1.
Proof: This time \( j \leq p - q \) and the proof duplicates the one of the previous lemma except for the following. Since the last subpath is of type I there will be the weight \( 1 - \omega_j \) associated with it and, therefore, the weighting factor \( 1 - \omega_j \) will be associated with the whole path. It is then obvious that \( q_k \) and \( j_k \) are the numbers defined in Lemma 3.1, with \( q_j \) being decreased by 1 due to the fact that the last subpath is now a double-arrowed one meaning that one identity subpath out of the previous \( q_j \) ones has somehow been fixed. \( \square \)

Note: In Example 3b it is \( (j_1, j_2, j_3) = (2, 2, 1) \) (Figures 7a-e) while \( (q_1, q_2, q_3) = (1, 0, 0) \) (Figures 8b, e), \( = (0, 1, 0) \) (Figures 8a, d), \( = (0, 0, 1) \) (Figure 8c). \( \square \)

Lemma 3.3: Under the assumptions of Lemmas 3.1 and 3.2 and the observation made regarding \( q_j \) formulas (3.1) and (3.4) are given in the following one

\[
\alpha_{ij}^m = (1 - \omega_j)^\delta \sum_{q_k = p - m - \delta} \prod_{k=1}^{p} (q_k + j_k - 1) (1 - \omega_k)^{q_k}, \tag{3.5}
\]

where \( \delta = 0 \) if \( j > p - q \) and \( \delta = 1 \) otherwise.

Proof: The proof is obvious. \( \square \)

Here we recall Lemma 1 of [2] which will be very useful in the sequel.

Lemma 3.4 ([2, Lemma 1]): \( G(T^p) \) consists of exactly one closed path (cycle) from any node \( i \in P \) to itself of length \( p \). This cycle contains \( q \) foldings no two of which are consecutive edges of it. \( \square \)

Having proved Lemma 3.3 and using Lemma 3.4 we can now determine the overall weight associated with the term of the \((i, j)\)th block of matrix \( C \) in (2.6) which comes from all the nonidentically zero \((i, j)\)th blocks of all terms of the expansion of \( C \) whose graphs are "identical" paths from \( i \) to \( j \) \((i, j \in P)\) with \( m \) foldings \((q \leq m \leq p)\). For this we have

Lemma 3.5: The overall weight associated with all "identical" paths with \( m \) foldings of the graph \( G(C_{ij}) \) is given by

\[
c_{ij}^m = (1 - \omega_j)^\delta \sum_{q_k = p - m - \delta} \prod_{k=1}^{p} (q_k + j_k - 2) (1 - \omega_k)^{q_k}, \tag{3.6}
\]

where \( q \leq m \leq p \), \( \delta = 0 \) if \( j > p - q \) and \( \delta = 1 \) otherwise, and \( j_k \geq 1 \).

Proof: Since \( T^p \) in \( C = T^p L_n^{p-q} \) is block diagonal we have \( C_{ij} = (T^p)_{ii}(L_n^{p-q})_{ij} \) implying that \( G(C_{ij}) = G((T^p)_{ii}) \cup G((L_n^{p-q})_{ij}) \). However, by Lemma 3.4, \( G((T^p)_{ii}) \) is a cycle from \( i \) to \( i \) with \( q \) foldings which passes exactly once through each node \( k \in P \) \((k \neq i)\) and has a weight equal to 1. So, \( C_{ij}^m \) is nothing but the overall weight associated with all "identical" paths from \( i \) to \( j \) with \( m - q \) foldings of the graph \( G((L_n^{p-q})_{ij}) \). Bearing in mind that in this graph \( j_k \) is in fact \( j_k - 1 \), since in \( G(C_{ij}) \) each node \( k \in P \) has already been used once in the cycle \( G((T^p)_{ii}) \), a straightforward application of formula (3.5) of Lemma 3.3 gives (3.6). \( \square \)
The analysis so far and especially Lemma 3.3 applied to the matrix $B$ in (2.7)-(2.8) gives now the following result.

**Lemma 3.8:** The overall weight associated with all "identical" paths with $m$ foldings, $q \leq m \leq p$, of the graph $G(B_{ij})$ is given by the expression

$$b_{ij}^m = (1 - \omega_j)^{\delta} \sum_{\ell=0}^{p-m-\delta} (-1)^{\ell} \frac{1}{\sigma_{\ell}} \sum_{q_k=p-\ell-m-\delta}^{p} \prod_{k=1}^{p} \left( \frac{q_k + j_k - 1}{j_k - 1} \right) (1 - \omega_k)^{q_k}$$  \hspace{1cm} (3.7)

where the $\sigma_{\ell}$'s are given by (2.8).

**Proof:** In view of (2.7) a straightforward application of the result (3.5) gives

$$b_{ij}^m = \sum_{\ell=0}^{p} (-1)^{\ell} \frac{1}{\sigma_{\ell}} \sum_{q_k=p-\ell-m-\delta}^{p} \prod_{k=1}^{p} \left( \frac{q_k + j_k - 1}{j_k - 1} \right) (1 - \omega_k)^{q_k}.$$  \hspace{1cm} (3.8)

However, since we are interested in paths that have $m(\geq q)$ foldings only the powers of $\mathcal{C}_n$ with exponent $p - \ell \geq m + \delta$ must be considered. So the upper limit of the first summation, that is the maximum value of $\ell$, must be $p - m - \delta$. On the other hand since $\delta = 0$ or 1 depending on whether all "identical" paths that are considered have as their last subpath, in the sense explained, a folding or a type I one, the factor $(1 - \omega_j)^{\delta}$ is then a common factor and can be taken out of the summation. These two observations effectively show that (3.8) implies (3.7). \hspace{1cm} \Box

In the following lemma we prove the equality of the two expressions in (3.6) and (3.7). Thus we have

**Lemma 3.7:** Under the assumptions of Lemmas 3.5 and 3.6 the expressions in (3.6) and (3.7) are equal. More specifically

$$b_{ij}^m = c_{ij}^m, \quad i,j \in P, m \in \{q, \ldots, p - \delta\},$$  \hspace{1cm} (3.9)

with $\delta = 0$ if $j > p - q$ and 1 otherwise.

**Proof:** We begin with the expression (3.6) for $b_{ij}^m$ and replace $\sigma_{\ell}$, $\ell = 0(1)p - m - \delta$, using (2.8). We then have

$$b_{ij}^m = (1 - \omega_j)^{\delta} \sum_{\ell=0}^{p-m-\delta} (-1)^{\ell} \left( \sum_{r_1}^{r_2} \cdots ^{r_{\ell}} \prod_{k=1}^{p} (1 - \omega_{r_k})(1 - \omega_{r_1}) \cdots (1 - \omega_{r_{\ell}}) \right)$$

$$\times \left( \sum_{q_k=p-\ell-m-\delta}^{p} \prod_{k=1}^{p} \left( \frac{q_k + j_k - 1}{j_k - 1} \right) (1 - \omega_k)^{q_k} \right)$$  \hspace{1cm} (3.10)

$$= (1 - \omega_j)^{\delta} \sum_{\ell=0}^{p-m-\delta} (-1)^{\ell} \sum_{q_k=p-\ell-m-\delta}^{p} \prod_{k=1}^{p} \left( \frac{q_k + j_k - 1}{j_k - 1} \right) (1 - \omega_k)^{q_k+\delta_k},$$

where $\delta_k = 1$ if $k \in \{r_1, r_2, \ldots, r_{\ell}\}$ and $\delta_k = 0$ otherwise. If we set $q_k$ in the place of $q_k + \delta_k$ then the rightmost expression in (3.10) for $b_{ij}^m$ becomes successively
\[ b_{ij} = (1 - \omega_j)^{\delta} \sum_{t=0}^{p-m-\delta} (-1)^t \sum_{q_h=p-m-\delta}^{p} \prod_{k=1}^{t} (1 - \omega_k)^{q_h} \sum_{s_r=1}^{s} \prod_{k=1}^{t} \left( \frac{q_h + j_k - 1 - \delta_k}{j_k - 1} \right) \]

or

\[ b_{ij} = (1 - \omega_j)^{\delta} \sum_{q_h=p-m-\delta}^{p} \prod_{k=1}^{t} (1 - \omega_k)^{q_h} \sum_{s_r=1}^{s} \prod_{k=1}^{t} \left( \frac{q_h + j_k - 1 - \delta_k}{j_k - 1} \right) \]  \hspace{1cm} (3.11)

Now from the summation on the right hand side of (3.11) we consider only the term that corresponds to the combination of nodes \( r_s, s = 1(1)t \leq p - m - \delta \), with \( 1 \leq r_1 < r_2 < \ldots < r_t \leq p \) and for which \( q_{r_s} \geq 1, s - 1(1)t \). This term denoted by \((b_{ij})_{r_1...r_t}\) is given by

\[ (b_{ij})_{r_1...r_t} = (1 - \omega_j)^{\delta} \prod_{k=1}^{t} (1 - \omega_{r_k})^{s_{r_k}} \sum_{t=0}^{t} (-1)^t \sum_{s_{r_k}=1}^{s} \prod_{k=1}^{t} \left( \frac{q_{r_k} + j_{r_k} - 1 - \delta_{r_k}}{j_{r_k} - 1} \right) \]  \hspace{1cm} (3.12)

Considering the double summation of the right hand side of (3.12) and distinguishing the two cases of \( \delta_{r_1} = 0 \) and \( \delta_{r_1} = 1 \) we can successively transform it as follows

\[ \sum_{t=0}^{t} (-1)^t \sum_{s_{r_k}=1}^{s} \prod_{k=1}^{t} \left( \frac{q_{r_k} + j_{r_k} - 1 - \delta_{r_k}}{j_{r_k} - 1} \right) \]

\[ = \left( \frac{q_{r_1} + j_{r_1} - 1}{j_{r_1} - 1} \right) \sum_{s_{r_k}=1}^{s} \sum_{h=2}^{t} (-1)^t \sum_{s_{r_k}=1}^{s} \prod_{k=1}^{t} \left( \frac{q_{r_k} + j_{r_k} - 1 - \delta_{r_k}}{j_{r_k} - 1} \right) \]

\[ + \left( \frac{q_{r_1} + j_{r_1} - 2}{j_{r_1} - 1} \right) \sum_{s_{r_k}=1}^{s} \sum_{h=2}^{t} (-1)^t \sum_{s_{r_k}=1}^{s} \prod_{k=1}^{t} \left( \frac{q_{r_k} + j_{r_k} - 1 - \delta_{r_k}}{j_{r_k} - 1} \right) \]

\[ = \left[ \left( \frac{q_{r_1} + j_{r_1} - 1}{j_{r_1} - 1} \right) - \left( \frac{q_{r_1} + j_{r_1} - 2}{j_{r_1} - 1} \right) \right] \sum_{s_{r_k}=1}^{s} \sum_{h=2}^{t} (-1)^t \sum_{s_{r_k}=1}^{s} \prod_{k=1}^{t} \left( \frac{q_{r_k} + j_{r_k} - 1 - \delta_{r_k}}{j_{r_k} - 1} \right) \]

\[ = \left( \frac{q_{r_1} + j_{r_1} - 2}{j_{r_1} - 2} \right) \sum_{s_{r_k}=1}^{s} \sum_{h=2}^{t} (-1)^t \sum_{s_{r_k}=1}^{s} \prod_{k=1}^{t} \left( \frac{q_{r_k} + j_{r_k} - 1 - \delta_{r_k}}{j_{r_k} - 1} \right), \]  \hspace{1cm} (3.13)

It is noted that (3.13) holds for \( j_{r_1} \geq 2 \). For \( j_{r_1} = 1 \) the difference yielding \( \left( \frac{q_{r_1} + j_{r_1} - 2}{j_{r_1} - 2} \right) \) is in fact equal to \( \left( \frac{q_{r_1} - 1}{0} \right) - \left( \frac{q_{r_1} - 1}{0} \right) = 1 - 1 = 0 \), meaning that \((b_{ij})_{r_1r_2...r_t}\) = 0. Following the same analysis on the double summation of the rightmost expression of (3.13), as in the derivation of (3.13) from (3.12), we finally obtain

\[ (b_{ij})_{r_1...r_t} = (1 - \omega_j)^{\delta} \prod_{k=1}^{t} \left( \frac{q_{r_k} + j_{r_k} - 2}{j_{r_k} - 2} \right) (1 - \omega_{r_k})^{s_{r_k}}, \]  \hspace{1cm} (3.14)
\[ j_r, k = 1(1)t. \] If the expression (3.14) is considered for all the terms of \( b_{ij}^m \) that correspond to all the combinations of \( t \leq p - m - \delta \) nodes \( r_s, s = 1(1)t, \) with \( q_{r_s} \geq 1 \) and such that \( \sum_{s=1}^{t} q_{r_s} = p - m - \delta \) then (3.11) becomes

\[
b_{ij}^m = (1 - \omega_j)^{\delta} \sum_{\ell=0}^{p - m - \delta} \sum_{q_{r_\ell} = p - \ell - m - \delta} \prod_{k=1}^{t} \left( \frac{q_{r_k} + j_{r_k} - 2}{j_{r_k} - 2} \right) (1 - \omega_k)^{q_{r_k}} \tag{3.15}
\]

which is nothing but the expression (3.6) for \( c_{ij}^m. \) This concludes the proof of the present lemma. \( \square \)

In Lemma 3.6 we considered all "identical" paths in \( G(B_{ij}) \) with \( m \) foldings such that \( q \leq m \leq p. \) However in \( G(B_{ij}) \) there are also paths with a number of foldings \( m \) ranging from 0 to \( q - 1. \) For these paths there holds.

**Lemma 3.8:** Under the assumptions of Lemma 3.6 the overall weight \( b_{ij}^m \) associated with all "identical" paths of \( G(B_{ij}) \) with \( m \) foldings, where \( 0 \leq m < q, \) equals zero (!). Namely

\[
b_{ij}^m = 0 (!). \tag{3.16}
\]

**Proof:** The expression for \( b_{ij}^m \) is the same as that in (3.7) except that \( 0 \leq m < q \) instead of \( q \leq m \leq p. \) However, even in this present case the analysis in the proof of the previous Lemma 3.7 holds and can be followed step by step. The main difference is the following. Any path from \( i \) to \( j \) with \( m < q \) foldings does not pass through all the \( p \) nodes of \( P. \) (Just note that the path of smallest length which passes through all the \( p \) nodes at least once from each is the cycle of Lemma 3.4 which has \( q(> m) \) foldings!) Let that this path passes through the nodes \( r_1, r_2, \ldots, r_t. \) Then \( t < p \) and also \( j_{r_k} = 1, k = 1(1)t. \) Thus (3.7) becomes

\[
b_{ij}^m = (1 - \omega_j)^{\delta} \sum_{\ell=0}^{p - m - \delta} (-1)^{\ell} \gamma_{\ell} \sum_{q_{r_\ell} = p - \ell - m - \delta} \prod_{k=1}^{t} \left( \frac{q_{r_k} + j_{r_k} - 2}{j_{r_k} - 2} \right) (1 - \omega_k)^{q_{r_k}}. \tag{3.17}
\]

Since we can follow the analysis in the proof of Lemma 3.7 we have for the factor \( \left( \frac{q_{r_1} + j_{r_1} - 2}{j_{r_1} - 2} \right) \) in the rightmost expression of (3.13) that

\[
\left( \frac{q_{r_1} + j_{r_1} - 2}{j_{r_1} - 2} \right) = \left( \frac{q_{r_1} + j_{r_1} - 1}{j_{r_1} - 1} \right) - \left( \frac{q_{r_1} + j_{r_1} - 2}{j_{r_1} - 2} \right) = \left( \frac{q_{r_1}}{0} \right) - \left( \frac{q_{r_1} - 1}{0} \right) = 0,
\]

which effectively shows (3.16). \( \square \)

Now that we have developed all the necessary tools needed the proof of the main theorem follows in a straightforward fashion.

**Proof of Theorem 2.1** To prove that \( B = C \) it suffices to prove that \( B_{ij} = C_{ij} \) for all \( i, j \in P. \) However, as we have already seen for the latter to hold it suffices to prove that \( G(B_{ij}) \equiv B(C_{ij}) \) for
all \(i, j \in P\) or equivalently \(b_{ij}^m = c_{ij}^m\), \(m = 0(1)p\), where \(b_{ij}^m\) and \(c_{ij}^m\) are the overall weights associated with all "identical" paths in \(G(B_{ij})\) and \(G(C_{ij})\) with \(m\) foldings. This must be proved for all \(i, j \in P\) and \(m = 0(1)p\). As we have already seen (Lemma 3.5) in \(G(C_{ij})\) there are no paths with \(m(q)\) foldings. On the other hand in Lemmas 3.7 and 3.8 we showed that \(b_{ij}^m = c_{ij}^m\), \(m = q(1)p\), and that \(b_{ij}^m = 0\), \(m = 0(1)q - 1\) respectively. From these results the validity of the matrix identity (2.1) (or (1.8)) follows directly which concludes the proof of our main theorem. \(\square\)

Based on the analysis so far it is easy to prove the following.

**Theorem 3.1:** Under the assumptions of Theorem 2.1 there holds

\[
(\Omega T)^p L_n^{p-q} = L_n^{p-q}(\Omega T)^p,
\]

that is the matrices \((\Omega T)^p\) and \(L_n^{p-q}\) commute.

**Proof:** The matrices \(\Omega L\), \(\Omega U\) and \(\Omega T\) are denoted by \(L\), \(U\) and \(T\), respectively, as this was done in the beginning of Section 2. So, to prove (3.18) it suffices to prove that \((T^p L_n^{p-q})_{ij} = (L_n^{p-q} T^p)_{ij}\) for all \(i, j \in P\), where \(L_n\) is given now by (2.2). However, since \(T^p\) is a block diagonal matrix it suffices to prove that \((T^p)_{ii}(L_n^{p-q})_{ij} = (L_n^{p-q})_{ij}(T^p)_{jj}\) for all \(i, j \in P\). By Lemmas 3.1–3.3 \((L_n^{p-q})_{ij}\) is a sum of \(p-q+1\) block terms with overall weights \(\alpha_{ij}^m\), \(m = 0(1)p - q\), which are given by (3.5) with \(s = p-q\). The terms corresponding to the same \(m\) have the same graphs, from \(i\) to \(j\) with \(m\) foldings, and identical scalar coefficients \(\alpha_{ij}^m\), which proves the validity of (3.18). \(\square\)

**Note:** It is noted that the identity (3.18) was observed to hold in the cases \((p, q) = (2, 1), (3, 1)\) and (3.2) in [4].

**Remark:** It is natural to ask here if (1.8) holds when \(gcd(p, q) = d > 1\). To analyze this more general case we notice that a path in any one of the graphs considered so far can only pass through two nodes \(i\) and \(j\) of \(P\) if and only if \(i = j \mod(d)\). This suggests that for the corresponding study we must consider \(d\) sets of \(p'/d\) nodes each and examine whether the theory developed so far holds for any one of these sets of nodes. However, a very simple example shows what the situation is. Let \(p = 4\) and \(q = 2\). It is then \(d = 2\) and can readily be found out that to the formation of \(B_{11}\) the term \(T_{13}T_{31}\) whose graph is of length two also contributes. However, the overall weight of the term in question is \((1 - \omega_1)(\omega_2 - \omega_1)(\omega_4 - \omega_1)\) which is not identically zero. This makes Lemma 3.8 not be valid anymore as a consequence of which Theorem 2.1 can not hold in general. \(\square\)

We conclude our analysis by stating the corresponding results for the SOR case. So, in the very special case where \(\Omega = \omega I\), that is in the case where the MSOR matrix \(L_n\) reduces to the SOR one, \(L_\omega\), we have the following valid statement.

**Corollary 3.1:** Under the assumptions of Theorem 2.1, let \(L_\omega\) be the SOR iteration matrix defined by (1.6), or by (1.5) when \(\omega_j = \omega, j = 1(1)p\). Then the following matrix identities hold

\[
(L_\omega + (\omega - 1)I)^p = \omega^p T^p L_\omega^{p-q} = \omega^p L_\omega^{p-q} T^p.
\]

\(\square\)
Note: The left matrix equality in (3.21) was obtained in [2] while the second one is new and follows directly from Theorem 3.1.

References


I: 1 2 \ldots q \ldots i-q \ldots i \ldots p

Fig. 1

II:

\begin{array}{c}
1 \\
2 \\
\cdots \\
i \\
q \\
\cdots \\
i+p-q \\
p \\
\end{array}

Fig. 2

III:

\begin{array}{c}
1 \\
2 \\
\cdots \\
i \\
p \\
\end{array}

Fig. 3

Fig. 4

\begin{array}{c}
i-(\frac{i-1}{q})q \\
i-2q \\
i-q \\
i \\
i+p-(\frac{i-1}{q})+1q \\
p \\
\end{array}