1988

On an SSOR Matari Relationship and its Consequences

S. Galanis

A. Hadjidimos

D. Noutsos

Report Number:
88-815

On an SSOR Matari Relationship
and it Consequences

S. Galanis*
A. Hadjidimos+,*
D. Noutsos*,+

Computer Sciences Department
Purdue University
CSD-TR-815
October 1988

Abstract

A matrix relationship connecting the Jacobi and the SSOR matrices associated with a k-cyclic consistently ordered matrix A is presented. Next the equivalence of the SSOR method and a certain monoparametric k-step one for the solution of the linear algebraic system $Ax = b$ is established. The aforementioned equivalence can be exploited to derive regions of convergence, optimum parameters involved, etc. of the two iterative methods above. This is done by studying the simplest of the two methods that is the monoparametric k-step one. To show how the idea works the case $k = 2$ is very briefly discussed.

1980 Mathematics Subject Classifications AMS (MOS): Primary 65F10. C.R. Categories: 5.14
* Department of Mathematics, University of Ioannina, GR-451 10 Ioannina, Greece.
+ Visiting, Department of Computer Science, Purdue University, West Lafayette, Indiana 47907. (The work of these two authors was supported in part by NSF Grant CCR-8619817.)

1. Introduction

Let

$$A = I - T \quad , \tag{1.1}$$

where I is the $n \times n$ unit matrix and $T \in \mathbb{C}^{n,n}$ a weakly k-cyclic consistently ordered (c.o.) (cf. [10], [14] or [2]) one of the form

$$T := \begin{bmatrix} 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & T_{1k} \\ T_{21} & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & T_{32} & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & T_{k,k-1} & 0 \end{bmatrix} \tag{1.2}$$

and write

$$T := L + U \quad , \tag{1.3}$$

where L and U are strictly lower and strictly upper (block) triangular matrices. Form then the corresponding block Symmetric Successive Overrelaxation (SSOR) matrix associated with A as

$$S_\omega = U_\omega L_\omega \quad , \tag{1.4}$$

where

$$L_\omega = (I - \omega L)^{-1} \left[(1-\omega) I + \omega U\right] \quad , \tag{1.5a}$$

$$U_\omega = (I - \omega U)^{-1} \left[(1-\omega) I + \omega L\right] \quad . \tag{1.5b}$$

It is known that the sets of the eigenvalues $\mu \in \sigma(T)$ (the spectrum of T) and $0 \neq \lambda \in \sigma(S_\omega)$ are connected through the relationship

$$(2-\omega)^2 \omega^k \mu^k [\lambda + (1-\omega)]^{k-2} \lambda = [\lambda - (1-\omega)^2]^k \tag{1.6}$$

(see [11] and [4]). So, our main objective in this paper is to show that the matrix analogue of (1.6) also holds, namely

$$(2-\omega)^2 \omega^k T^k [S_\omega + (1-\omega)I]^{k-2} S_\omega = [S_\omega - (1-\omega)^2 I]^k \quad . \tag{1.7}$$

The idea of deriving matrix analogues from relationships connecting the eigenvalues of the Jacobi and the SOR (1.5a) matrices for k-cyclic c.o. ones goes back to Young and Kincaid [15] who dealt with the case $k = 2$ (cf. [13]). Recently, Galanis, Hadjidimos and Noutsos [2] have extended the previous result to any $k \geq 2$, (cf. [2]) and then in [3] to the case of the $(k - \ell,\ell)$– generalized c.o. matrices (cf. [12] , [7] or [14]).

The proof of the identity (1.7) and the background material for it will be given in Section 2. The exploitation of (1.7) to show that for the solution of the system $x = Tx + b$ the SSOR method is equivalent to a certain monoparametric k-step one is discussed in Section 3. The consequences of the aforementioned equivalence with regard to the determination of regions of convergence, optimal parameters etc. is the main topic of a forthcoming paper. Here, in Section 4, we restrict ourselves to applying the theory developed to the case $k = 2$.

2. The Relationship $(2-\omega)^2\, \omega^k\, T^k\, [S_\omega + (1-\omega)I]^{k-2}\, S_\omega = [S_\omega - (1-\omega)^2 I]^k$

The statement of our main result is as follows:

THEOREM: Let T, in (1.2), be the block Jacobi matrix, which is k-cyclic and consistently ordered (c.o.), and S_ω, in (1.4) - (1.5), be the associated SSOR matrix. Then for any $\omega \in \mathbb{C}$, T and S_ω satisfy the functional relationship (1.7).■

To proceed on to the proof of the Theorem some special cases have to be examined first and some background material has to be developed.

Remark 1: First we observe that for $\omega = 0$ and 2 (1.7) is trivially satisfied. Next for $\omega = 1$, (1.7) becomes

$$T^k\, S_1^{k-1} = S_1^k \quad . \tag{2.1}$$

From the special forms of T in (1.2) - (1.3) we have

$$U^j = 0,\ j > 1,\ L^j = 0,\ j > k-1,\ UL^jU = 0,\ j < k-1 \quad .$$

Consequently

$$\begin{aligned} T^k\, S_1 &= (L+U)^k\, (I-U)^{-1}\, L(I-L)^{-1}\, U = \\ &= \Big(\sum_{i=0}^{k-1} L^i\, U\, L^{k-1-i}\Big)(I+U)\, L\, \Big(\sum_{i=1}^{k-1} L^i\Big)\, U = \\ &= \Big(\sum_{i=0}^{k-1} L^i\, U\, L^{k-1-i}\Big)\Big(\sum_{i=1}^{k-1} L^i\, U + U\, L^{k-1}\, U\Big) = \end{aligned}$$

$$= \sum_{i=1}^{k-1} L^i\, U L^{k-1}\, U + U L^{k-1}\, U L^{k-1}\, U$$

and

$$S_1^2 = \Big(\sum_{i=1}^{k-1} L^i\, U + U L^{k-1}\, U\Big)^2 =$$

$$= \sum_{i=1}^{k-1} L^i\, U L^{k-1}\, U + U L^{k-1}\, U L^{k-1}\, U$$

from which (2.1) follows. So we restrict to $\omega \in \mathbb{C} \setminus \{0, 1, 2\}$.■

Since the analysis leading to the proof of the Theorem will be based on elementary graph theory, as this was done in [3], we adopt some simplifications in the notations to facilitate our study. Thus:

Remark 2: If in (1.7) instead of S_ω we had $S'_\omega = L_\omega U_\omega$ the analysis would be much easier. That (1.7) can be recovered from

$$(2-\omega)^2\, \omega^k\, T^k\, [S'_\omega + (1-\omega)I]^{k-2}\, S'_\omega = [S'_\omega - (1-\omega)^2 I]^k \tag{2.2}$$

is proven by using the matrix U_ω (1.5b) and simple similarity transformations. (2.2) becomes

$$(2-\omega)^2\, \omega^k\, U_\omega\, T^k\, [S'_\omega + (1-\omega)I]^{k-2}\, S'_\omega\, U_\omega^{-1} = U_\omega [S'_\omega - (1-\omega)^2 I]^k\, U_\omega^{-1}$$

or equivalently

$$(2-\omega)^2\, \omega^k\, U_\omega\, T^k\, U_\omega^{-1}\, [S_\omega + (1-\omega)I]^{k-2}\, S_\omega = [S_\omega - (1-\omega)^2 I]^k \quad . \tag{2.3}$$

In view of the known identity

$$\omega^k\, T^k\, U_\omega = [U_\omega + (\omega - 1)I]^k$$

(cf. [2] and [3]) one has

$$\omega^k\, T^k = [U_\omega + (\omega-1)I]^k\, U_\omega^{-1} = U_\omega^{-1}\, [U_\omega + (\omega-1)I]^k$$

or

$$\omega^k\, U_\omega\, T^k = [U_\omega + (\omega-1)I]^k \quad ,$$

meaning that

$$U_\omega\, T^k\, U_\omega^{-1} = T^k \quad . \tag{2.4}$$

Using (2.4) into (2.3), (1.7) follows.■

The material needed from graph theory is very elementary. Thus we consider the graph $G(X)$ of a block matrix X to be defined in the usual way (cf. Varga [10] and

Harary [6]) and adopt the notations

$$\tilde{T} = \omega\, T,\ \tilde{L} = \omega\, L,\ \tilde{U} = \omega\, U \quad . \tag{2.5}$$

In $G(\tilde{T})$, for a given $i = 2\ (1)\ k$, we have

$$\bigcup_{j=1}^{k} (i,\ j) = (i,\ i-1) \text{ as in Figure 1}$$

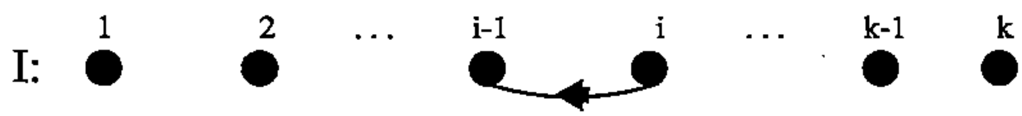

Figure 1

while for $i = 1$

$$\bigcup_{j=1}^{k} (i,\ j) = (1,\ k) \text{ as in Figure 2} \quad .$$

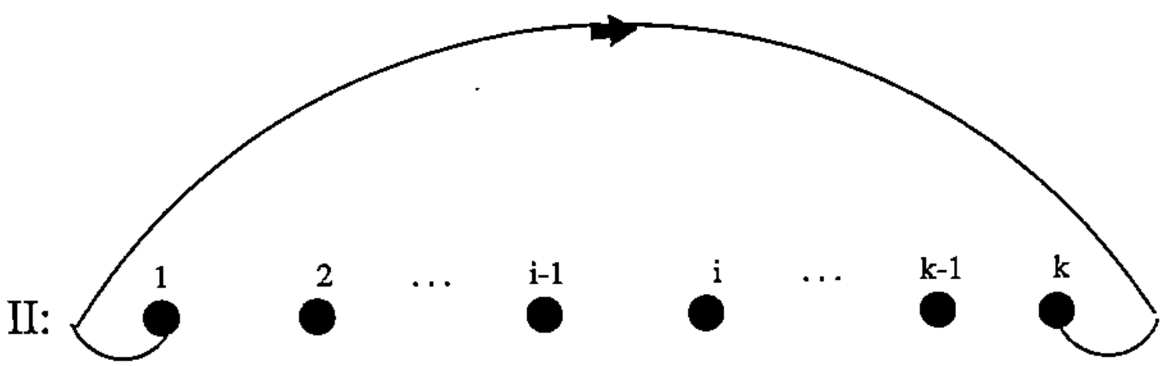

Figure 2

In the case of a type II edge we shall say that we have a "folding" (edge). Edges of type I are associate with the matrix $\tilde{L}$ while those of type II with $\tilde{U}$. Consequently

$$G(\tilde{L}) = \bigcup_{i=2}^{k} (i,\ i-1),\ G(\tilde{U}) = (1,\ k),\ G(\tilde{T}) = G(\tilde{L}) \cup G(\tilde{U}) \quad .$$

Obviously $G(\widetilde{T}^k)$ consists of only one cycle (closed path from any node to itself) of length k. This cycle contains precisely one folding.

Here we consider the "multiple-arrowed" paths that is paths associated with nonidentically zero products of $\widetilde{L}$ and $\widetilde{U}$ matrix factors having different scalar coefficients. More specifically: A path $\overrightarrow{P_i P_j}$ will be considered as a single-, double- or triple-arrowed path and will be denoted by $(\overrightarrow{i, j})$ or $\overrightarrow{P_i P_j}$, $(\overrightarrow{\rightarrow i, j})$ or $(\overrightarrow{\rightarrow P_i P_j})$ or $(\overrightarrow{\rightarrow\rightarrow i, j})$ or $(\overrightarrow{\rightarrow\rightarrow P_i P_j})$, iff it is associated with a product of the aforementioned form whose scalar coefficient is $(1-\omega)(2-\omega)$, $(2-\omega)$ or $(1-\omega)^2$ respectively.

Let us now examine how $G(S'_\omega)$ is derived. It is

$$
\begin{aligned}
S'_\omega &= L_\omega U_\omega = (I-\widetilde{L})^{-1}\,[(1-\omega)I+\widetilde{U}]\,(I-\widetilde{U})^{-1}\,[(1-\omega)I+\widetilde{L}] = \\
&= (I+\widetilde{L}+\widetilde{L}^2+\cdots+\widetilde{L}^{k-1})\,[(1-\omega)I+(2-\omega)\widetilde{U}]\,[(1-\omega)I+\widetilde{L}] = \\
&= (I+\widetilde{L}+\widetilde{L}^2+\cdots+\widetilde{L}^{k-1})\,[(1-\omega)^2 I+(1-\omega)\widetilde{L}+(1-\omega)(2-\omega)\widetilde{U}+(2-\omega)\widetilde{U}\widetilde{L}] = \qquad (2.6)\\
&= (1-\omega)^2 I+(1-\omega)(2-\omega)(\widetilde{L}+\widetilde{L}^2+\cdots+\widetilde{L}^{k-1})+ \\
&\quad +(1-\omega)(2-\omega)(I+\widetilde{L}+\widetilde{L}^2+\cdots+\widetilde{L}^{k-1})\widetilde{U}+(2-\omega)(I+\widetilde{L}+\cdots+\widetilde{L}^{k-1})\widetilde{U}\widetilde{L}\ .
\end{aligned}
$$

(2.6) gives

$$
G(S'_\omega) = \bigcup_{i=1}^{k}\,[(\overrightarrow{\rightarrow\rightarrow i, i})\cup(\bigcup_{j=1}^{i-1}(\overrightarrow{i, j}))\cup(\overrightarrow{i, k})\cup(\overrightarrow{\rightarrow i, k-1})] \qquad (2.7)
$$

where $(\overrightarrow{1, 0}) = \varnothing$. The subgraph of $G(S'_\omega)$ which contains the edges whose origin is the node i is illustrated in Figure 3. We notice that in view of (2.6) and (2.7)

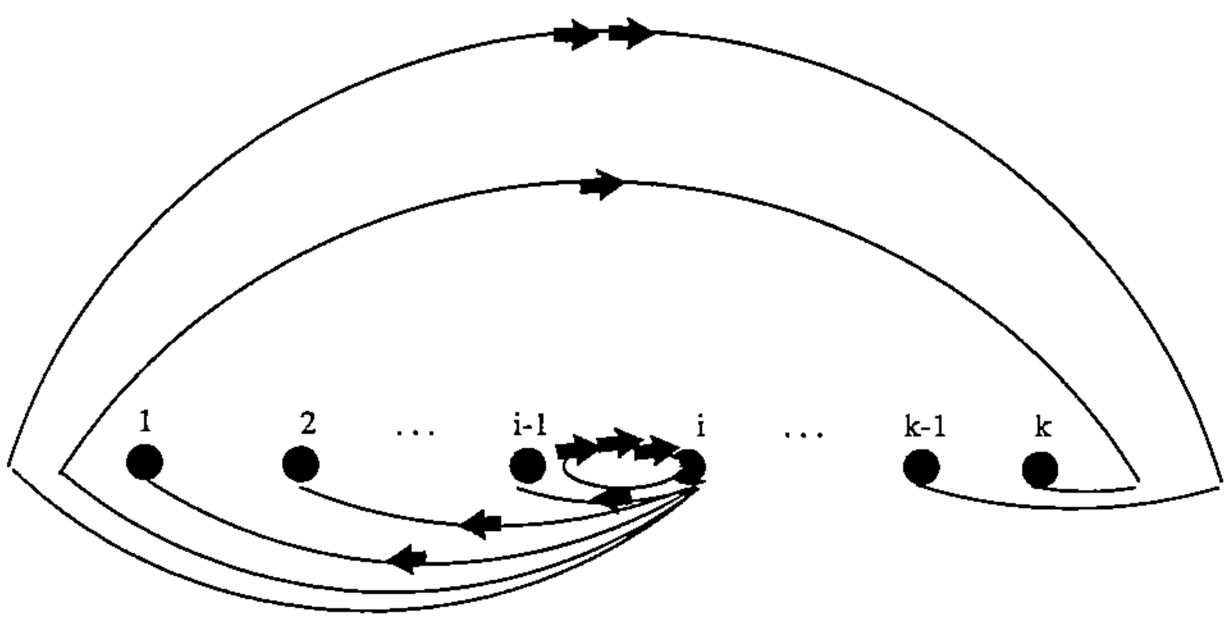

Figure 3

$$S'_\omega - (1-\omega)^2 I = (1-\omega)(2-\omega)(\tilde{L} + \tilde{L}^2 + \cdots + \tilde{L}^{k-1}) +$$
$$+ (1-\omega)(2-\omega)(I + \tilde{L} + \tilde{L}^2 + \cdots + \tilde{L}^{k-1})\tilde{U} + (2-\omega)(I + \tilde{L} + \tilde{L}^2 + \cdots + \tilde{L}^{k-1})\,\tilde{U}\,\tilde{L} \quad (2.8)$$

and

$$G(S'_\omega - (1-\omega)^2 I) = \bigcup_{i=1}^{k}[(\bigcup_{j=1}^{i-1} \overrightarrow{(i,\ j)}) \cup \overrightarrow{(i,\ k)} \cup \stackrel{\rightarrow\rightarrow}{(i,\ k-1)}] \quad . \qquad (2.9)$$

Its subgraph corresponding to the previous i is illustrated in Figure 4. In the same way we have that

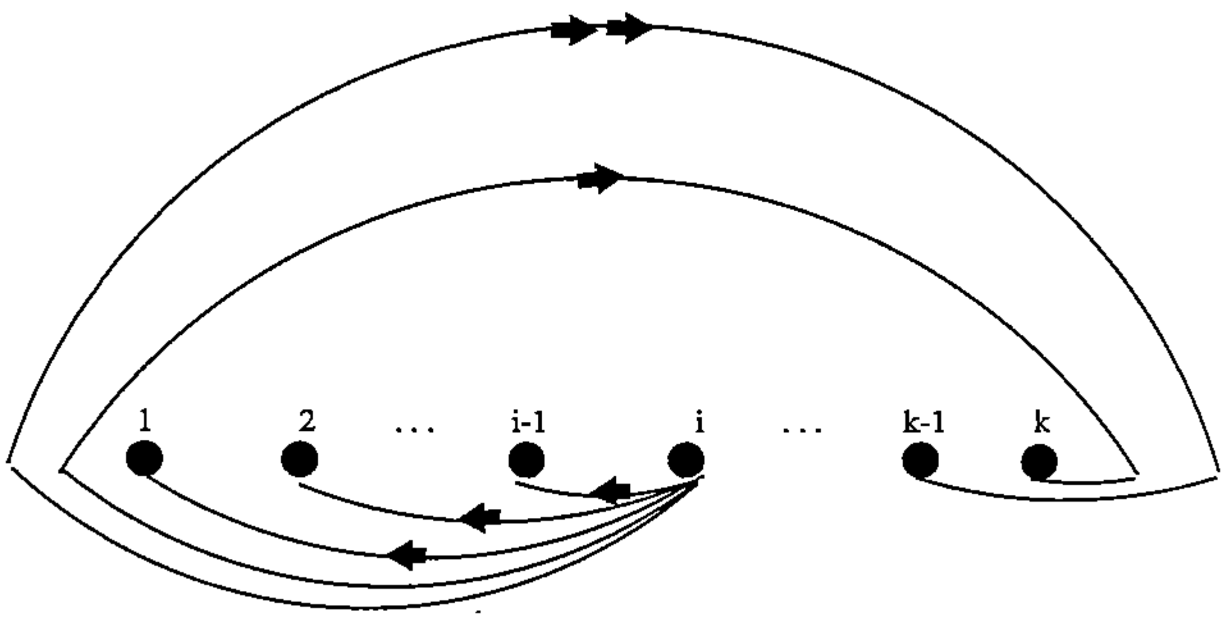

Figure 4

$$S'_\omega + (1-\omega)I = (1-\omega)(2-\omega)(I + \tilde{L} + \tilde{L}^2 + \cdots + \tilde{L}^{k-1}) +$$
$$+ (1-\omega)(2-\omega)(I + \tilde{L} + \tilde{L}^2 + \cdots + \qquad (2.10)$$
$$+ \tilde{L}^{k-1})\tilde{U} + (2-\omega)(I + \tilde{L} + \tilde{L}^2 + \cdots + \tilde{L}^{k-1})\tilde{U}\tilde{L}$$

and

$$G(S'_\omega + (1-\omega)I) = \bigcup_{i=1}^{k} [(\bigcup_{j=1}^{i} \overrightarrow{(i,\ j)}) \cup \overrightarrow{(i,\ k)} \cup \stackrel{\rightarrow\rightarrow}{(i,\ k-1)}] \quad , \qquad (2.11)$$

whose subgraph corresponding to the same i is illustrated in Figure 5.

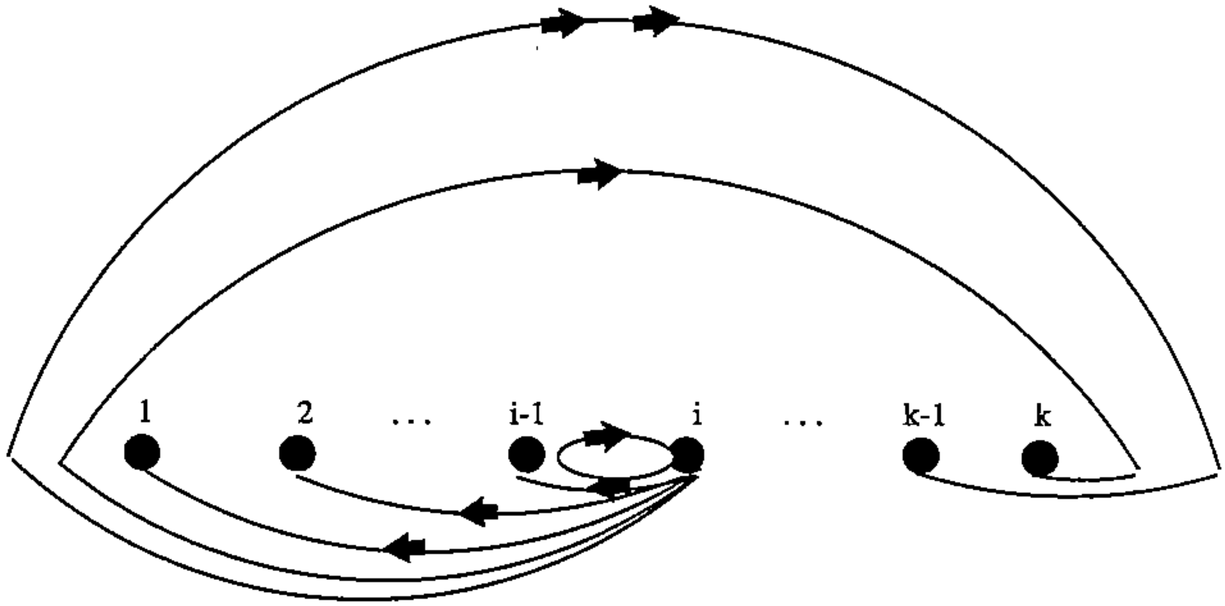

Figure 5

To be able to go on with our analysis we ddefine the matrices

$$B := [S'_\omega - (1-\omega)^2 I]^k \,,\; C := (2-\omega)^2\, \tilde{T}^k\, [S'_\omega + (1-\omega)I]^{k-2}\, S'_\omega \quad . \qquad (2.12)$$

Our main effort will be put in the statement and proof of the following two lemmas from which the proof of the Theorem is immediately obtainedd.

Lemma 1: Let $G(B)$ and $G(C)$ be the graphs of the $k \times k$ block matrices defined in (2.12). If in $G(B)$ there exists a path from i to j containing $m+1$ foldings then there exists an identical path in $G(C)$ and vice versa.

Proof: The present lemma can be proved in a way quite analogous to the one in the first part of Lemma 4 in [3], where a similar case is treated. So, to avoid duplication we refer the interested reader to the reference just given.■

Lemma 2: Consider the expansions of B and C in (2.12) in terms of products of $\tilde{L}$ and $\tilde{U}$ by using (2.5), (2.6), (2.8) and (2.10). Then the path in Lemma 1 comes from the graphs of identical terms of B and C whose scalar coefficients (s.c.'s), after all like terms are summed up, are identical products of integer numbers and powers of $(1-\omega)$ and $(2-\omega)$.

Proof: Assume that there exists a path from i to j containing $m+1$ foldings, identically the same, in both graphs $G(B)$ and $G(C)$. This path corresponds to the matrix prodduct

$$\tilde{L}^{i-1}\,\tilde{U}\,\tilde{L}^{k-1}\,\tilde{U}\,\tilde{L}^{k-1}\,\cdots\,\tilde{U}\,\tilde{L}^{k-j} \tag{2.13}$$

which contains $m+1$ $\tilde{U}$'s. We compare the s.c.'s of (2.13) in both matrices in (2.12). By considering three cases:

Case I: $j < k-1$: We begin with the matrix B. It is noted that any one of the possible paths connecting i with j and containing $m+1$ foldings is a sequence of k paths of $G(S'_\omega - (1-\omega)^2 I)$. Let t be all the nodes which are passed on the way from i to j after $m+1$ foldings not counting i. It is checked that $t = k(m+1)+i-j$. The $m+1$ foldings are taken from the single-arrowed foldings ending at k or from the double-arrowed foldings ending at $k-1$ (see (2.7)). Let m_2 be the number of the double-arrowed foldings. Obviously the different paths with $m+1$ foldings are $\begin{bmatrix} m+1 \\ m_2 \end{bmatrix}$. The other $k-m-2$ paths (in fact $k-m-1$ except the last path ending at the given j) will be taken as single-arrowed paths ending at $t-m-1-m_2$ nodes (that is t nodes except the $m+1$ nodes k except the m_2 nodes $k-1$ and the node j). This gives $\begin{bmatrix} t-m-m_2-2 \\ k-m-2 \end{bmatrix}$ different paths. After the previous obeservations and taking into consideration the arrowed nature of all different paths we have that the s.c. of the matrix in (2.13) will be

$$(2-\omega)^k \sum_{m_2=0}^{m+1} \begin{bmatrix} t-m-m_2-2 \\ k-m-2 \end{bmatrix} \begin{bmatrix} m+1 \\ m_2 \end{bmatrix} (1-\omega)^{k-m_2} \quad . \tag{2.14}$$

Now we work in exactly the same way on the matrix C. The path which corresponds to $G(\tilde{T}^k)$ is a cycle connecting i with itself and containing one folding. The corresponding s.c. here is 1. Any path connecting i with j and containing now m foldings can be obtained from $k-2$ successive paths of $G(S'_\omega + (1-\omega)I)$ followed by one path of $G(S'_\omega)$. All the nodes which are passed on the way from i to j after m foldings are $t-k$. We have to distinguish two subcases: a) The last path corresponding to $G(S'_\omega)$ is not a triple-arrowed path: Let then m_2 be the number of the double-arrowed foldings. So the number of different paths is $\begin{bmatrix} m \\ m_2 \end{bmatrix}$. Let p be the number of the single-arrowed paths of the form $(\overrightarrow{i, i})$. From (2.12) these paths can be obtained from the $k-1-p$ nodes ($k-2$ except the p paths plus the first node). The "distribution" of the above p paths to $k-1-p$ nodes can be made in as many ways as the number of combinations with repetitions of $k-1-p$ chosen p. This equals

$\begin{bmatrix} (k-1-p)+p-1 \\ p \end{bmatrix} = \begin{bmatrix} k-2 \\ p \end{bmatrix}$. Finally the $k-1-p-m-1$ single-arrowed paths (the above $k-1-p$ paths except the m foldings, which are given, and the first node i) will be obtained from the $t-k-m-m_2-1$ nodes ($t-k$ nodes except the m nodes k except the m_2 nodes $k-1$ except the last node which corresponds to $G(S'_\omega)$). This gives $\begin{bmatrix} t-k-m-m_2-1 \\ k-m-p-2 \end{bmatrix}$ different paths. By combining the results so far and having always in mind the arrowed nature of the paths involved the s.c. of (2.13) corresponding to the present subcase is

$$(2-\omega)^2 \sum_{m_2=0}^{m} \sum_{p=0}^{k-2-m_2} \begin{bmatrix} t-k-m-m_2-1 \\ k-m-p-2 \end{bmatrix} \begin{bmatrix} k-2 \\ p \end{bmatrix} \begin{bmatrix} m \\ m_2 \end{bmatrix} (2-\omega)^{k-1} (1-\omega)^{k-1-m_2}$$
$$= (2-\omega)^{k+1} \sum_{m_2=0}^{m} \begin{bmatrix} t-m-m_2-3 \\ k-m-p-2 \end{bmatrix} \begin{bmatrix} m \\ m_2 \end{bmatrix} (1-\omega)^{k-1-m_2} \quad . \tag{2.15}$$

b) The last path corresponding to $G(S'_\omega)$ is a triple-arrowed one: The previous analysis is applied with the difference that $p = 0(1)k-m-3$. (In fact $k-m-2$ except the path ending at the last node j). Hence the s.c. is given by

$$(2-\omega)^k \sum_{m_2=0}^{m} \begin{bmatrix} t-m-m_2-3 \\ k-m-3 \end{bmatrix} \begin{bmatrix} m \\ m_2 \end{bmatrix} (1-\omega)^{k-m_2} \quad . \tag{2.16}$$

Obviously the s.c. in (2.13) is the sum of the two quantities in (2.15) and (2.16) which after a lengthy manipulation involving properties of sums and combinatorics can be proved it is equal to the quantity in (2.14). This concludes the proof of the lemma in this case.

Case II: $j = k-1$: a) If we assume that the last path is of the form $(k, \overrightarrow{k-1})$ then the theory of the previous Case I applies and gives the s.c. (2.14) for (2.13) for both matrices B and C. b) If the last path is a double-arrowed folding then the previous theory applies. The only difference is that in (2.14) m foldings are considered instead of $m+1$. This is because the last folding is given. However, the nodes at which a folding ends can be taken as being $m+1$. Thus (2.14) gives

$$(2-\omega)^k \sum_{m_2=0}^{m} \begin{bmatrix} t-m-m_2-2 \\ k-m-1 \end{bmatrix} \begin{bmatrix} m \\ m_2 \end{bmatrix} (1-\omega)^{k-m_2-1} \quad . \tag{2.17}$$

Of course the s.c. of (2.13) in both matrices B and C is the sum of the two expressions in (2.14) and (2.17).

Case III: $i = k$: In this case the last path is a single-arrowed folding and the previous analysis holds with m replacing $m+1$. Thus, from (2.14) the s.c. of (2.13) in both matrices B and C is given by

$$(2-\omega)^k \sum_{m_2=0}^{m} \begin{pmatrix} t-m-m_2-1 \\ k-m-1 \end{pmatrix} \begin{pmatrix} m \\ m_2 \end{pmatrix} (1-\omega)^{k-m_2} \quad . \tag{2.18}$$

This concludes the proof of the present lemma.■

We come now to the proof of the Theorem.

PROOF OF THE THEOREM: Lemmas 1 and 2 effectively prove that $B = C$ in (2.12) or equivalently that (2.2) holds for any $\omega \in \mathbb{C} \setminus \{0, 1, 2\}$. However, as was proved in Remark 2, (2.2) is equivalent to (1.7) except perhaps for $\omega = 1$. Since this is covered by Remark 1 the validity of (1.7) is established and the Theorem is proved.■

3. Equivalence of the SSOR to a Monoparametric k-Step Method.

To show that the SSOR method

$$x^{(m)} = S_\omega x^{(m-1)} + \omega(2-\omega)(I-\omega U)^{-1}\,(I-\omega L)^{-1}\, b \quad , \tag{3.1}$$

used for the solution of $(I-T)x = b$, is equivalent to a certain monparametric k-step method we proceed in a way analogous to that in [2] and [3]. Let then $x^{(m-k)}$ be the $(m-k)^{\text{th}}$ iteration of (3.1) with $m = 1, 2, \cdots$. From (1.7) we have

$$[S_\omega - (1-\omega)^2 I]^k\, x^{(m-k)} = (2-\omega)^2\omega^k T^k\, [S_\omega + (1-\omega)I]^{k-2}\, S_\omega x^{(m-k)}$$

or equivalently

$$\begin{aligned} &\sum_{j=0}^{k} \begin{pmatrix} k \\ j \end{pmatrix} (-1)^j\,(1-\omega)^{2j}\, S_\omega^{k-j}\, x^{(m-k)} = \\ &= (2-\omega)^2\omega^k T^k \sum_{j=0}^{k-2} \begin{pmatrix} k-2 \\ j \end{pmatrix} (1-\omega)^j\, S_\omega^{k-1-j}\, x^{(m-k)} \quad . \end{aligned} \tag{3.2}$$

By successively applying (3.1) it can be obtained that

$$S_\omega^j x^{(m-k)} = x^{(m-k+j)} - \omega(2-\omega) \left[\sum_{p=0}^{j-1} S_\omega^p \right] (I-\omega U)^{-1}(I-\omega L)^{-1}\, b \quad .$$

Substituting into (3.2) we take after some manipulation

$$x^{(m)} = \sum_{j=1}^{k-1} (1-\omega)^{j-1} \left[\begin{pmatrix} k-2 \\ j-1 \end{pmatrix} (2-\omega)^2\omega^k T^k + \begin{pmatrix} k \\ j \end{pmatrix} (\omega-1)^{j+1} I \right] x^{(m-j)} +$$

$$+ (-1)^{k+1}(1-\omega)^{2k} x^{(m-k)} +$$

$$+ \omega(2-\omega)\left[\sum_{j=0}^{k} \begin{pmatrix} k \\ j \end{pmatrix} (-1)^j (1-\omega)^{2j} \left(\sum_{p=0}^{k-j-1} S_\omega^p \right) - \right.$$

$$\left. - \sum_{j=0}^{k-2} \begin{pmatrix} k-2 \\ j \end{pmatrix} (1-\omega)^j (2-\omega)^2 \omega^k T^k \left(\sum_{p=0}^{k-j-2} S_\omega^p \right) \right] (I-\omega U)^{-1} (I-\omega L)^{-1} b \ .$$

The constant term of the right hand side above can be simplified after some algebra takes place and under the assumption $1 \notin \sigma(T^k)$. The scheme obtained is the following

$$\begin{aligned} x^{(m)} = & \sum_{j=1}^{k-1} (1-\omega)^{j-1} \left[\begin{pmatrix} k-2 \\ j-1 \end{pmatrix} (2-\omega)^2 \omega^k T^k + \begin{pmatrix} k \\ j \end{pmatrix} (\omega - 1)^{j+1} I \right] x^{(m-j)} + \\ & + (-1)^{k+1} (1-\omega)^{2k} x^{(m-k)} + \omega^k (2-\omega)^k \left(\sum_{j=0}^{k-1} T^j \right) b, \quad m = 1, 2, 3, \ldots, \end{aligned} \tag{3.3}$$

where $x^{(j)} \in \mathbb{C}^n$, $j = 0(-1) - k + 1$ arbitrary.

In the sense explained, (3.1) and (3.3) are equivalent and the study of (3.1) can be made by studying (3.3) and vice versa.

One may also observe that because of the special cyclic nature of T in (1.2), (3.3) can be split into k simpler and of smaller dimensions k-step iterative methods provided that all the vectors involved are partitioned in accordance with T. Each of these k k-step methods has the same convergence rates as (3.3). So the solution of any one of these simpler methods provides us with the corresponding vector component of the solution x of $x = Tx + b$ and from the latter equation all the other components of x and therefore x itself are readily obtained.

4. The Case k = 2

In this section the convergence properties of the 2-cyclic SSOR method (3.1) are studied by means of its equivalent 2-step method (3.3). It should be said that some of them are already known.

Restricting ourselves to real ω and therefore to $\omega \in (0,2)$, which is a necessary condition for the SSOR method to converge, assuming that $1 \notin \sigma(T^2)$, and setting

$$\omega' = (2-\omega)\omega \in (0, 1] \ , \tag{4.1}$$

(3.3) becomes

$$x^{(m)} = [\omega^2 T^2 + 2(1-\omega)] x^{(m-1)} - (1-\omega)^2 x^{(m-2)} + \omega^2 (I+T) b \ , \tag{4.2}$$

where to simplify the notation primes have been dropped from ω. Following Parsons [9], who was able to get rid of the univalence restriction required in [11], we have that the generating function of (4.2) is given by

$$h(z) = \frac{(1-(1-\omega)z)^2}{\omega^2 z} \quad . \tag{4.3}$$

Consequently (4.2) converges iff $\sigma(T^2) \subset \mathbb{C} \setminus \Omega(\bar{\Delta})$, where $\bar{\Delta}$ is the unit disc and $\Omega(z)$ the mapping defined by (4.3). On the other hand $\sigma(T^2) \subset \mathbb{C} \setminus \Omega(\Delta_\eta)$ and has at least one of its elements on the boundary of $\Omega(\Delta_\eta)$, with Δ_η being the open disc of radius $\eta > 1$, implies optimum convergence for one member of the family of schemes (4.2) or equivalently for the SSOR method with $\rho(S_\omega) = 1/\eta$.

By writing $z = \eta(\cos\theta + i\sin\theta)$, $\eta > 1$ and $\theta \in [0, 2\pi)$, (4.3) gives

$$Re h(z) = \frac{1}{\omega^2 \eta} \left[(1 + (1-\omega)^2 \eta^2) \cos\theta - 2(1-\omega)\eta\right], \tag{4.4}$$

$$Im h(z) = -\frac{1}{\omega^2 \eta} \left[1 - (1-\omega)^2 \eta^2\right] \sin\theta \quad .$$

Obviously the point $(x,y) = (Re h(z), Im h(z))$ lies on an ellipse symmetric about the real axis

$$\frac{(x-c)^2}{a^2} + \frac{y^2}{b^2} = 1 \quad , \tag{4.5}$$

whose center c, "real" semiaxis a and "imaginary" semiaxis b are given by the expressions

$$c = -\frac{2(1-\omega)}{\omega^2}, \quad a = \frac{1 + (1-\omega)^2 \eta^2}{\omega^2 \eta}, \quad b = \frac{|\, 1 - (1-\omega)^2 \eta^2 \,|}{\omega^2 \eta} \quad . \tag{4.6}$$

We fix the right vertex of the above ellipse in such a way as to have

$$a + c = Re h(z) \mid_{\theta = 0} = \frac{1}{\omega^2 \eta} (1 - (1-\omega)\eta)^2 = \rho \in [0, 1) \tag{4.7}$$

Then a simple analysis shows that for a given ρ and ω decreasing continuously in $(0, 1]$, $1/\eta$ increases continuously in $[\rho, 1)$, c decreases continuously in $(-\infty, 0]$ and a, b increase continuously in $[\rho, +\infty)$. Thus one obtains a family of ellipses $\mathcal{E}_{\rho,\omega}$, each member of which is tangent to the line $x = \rho$ at $(\rho,0)$ and is strictly contained (except for the aforementioned point) within any other member of the same family corresponding to a smaller ω. The limiting ellipse $\lim\limits_{\omega \to 0^+} \mathcal{E}_{\rho,\omega}$ is easily found out to be the parabola

$$P_\rho :\equiv y^2 = -4\rho x + 4\rho^2 \quad ; \tag{4.8}$$

which has its vertex at $(\rho,0)$ and passes through the points $(0, \pm 2\rho)$. In the very special case $\rho = 0$, $b = 0$, the ellipses $\mathcal{E}_{0,\omega}$ degenerate into straight-line segments whose end-points are $(-\frac{4(1-\omega)}{\omega^2}, 0)$ and (0,0), while the degenerate parabola P_0 is the left real semi-line with origin (0,0). Consequently if it so happens and $\sigma(T^2) \subset \text{int}(P_\rho)$ for a given $\rho \in [0,1)$, then there exists a unique ellipse $\mathcal{E}_{\rho,\omega}$ corresponding to a unique $\omega \in (0,1]$ such that $\sigma(T^2) \subset \mathcal{E}_{\rho,\omega}$ with at least one element of $\sigma(T^2)$ on the boundary of $\mathcal{E}_{\rho,\omega}$. The smallest of the two η's from (4.7) provides us with the "optimum" scheme (4.2) and therefore with the "optimum" SSOR method for which $\rho(S_\omega) = 1/\eta$.

Allowing now ρ in (4.7) to vary in $[0,1)$ it is obvious that the family of ellipses discussed in the previous paragraph becomes a two-parametric one and the limiting parabola in (4.8) as $\rho \to 1^-$ becomes

$$P_1 :\equiv y^2 = -4x + 4 \quad . \tag{4.9}$$

It is easy for one to conclude that if $\sigma(T^2) \in \text{int}(P_1)$ then and only then one can find convergent SSOR methods. In addition it becomes obvious, from the previous analysis, that in such a case there will exist infinitely many "optimum" ellipses $\mathcal{E}_{\rho,\omega}$, with the property "$\sigma(T^2) \subset \mathcal{E}_{\rho,\omega}$ and has at least one element on the boundary of $\mathcal{E}_{\rho,\omega}$", which correspond to all the values of $\rho \in (\hat{\rho}, 1)$. $\hat{\rho}$ is the unique value of ρ for which $\sigma(T^2)$ is contained in the closure of the interior of P_ρ and has at least one element on it. The determination of that ω which corresponds to the smallest "optimum" $1/\eta$ and which in turn corresponds to the overall optimum ω is given in [5].

Based on the analysis so far one may tackle and solve a number of problems, the solutions to some of which are known, in connection with the convergence of the 2-cyclic SSOR method (3.1) ($\rho(S_\omega) < 1$). For example: i) Determine in a $(\rho(T), \omega)$-plane the region R_2 so that the SSOR converges for any $(\rho(T), \omega) \in R_2$. ii) For $\sigma(T^2)$ real nonnegative determine as in (i) the corresponding region R_2^+ as well as the optimal SSOR parameter $\hat{\omega}$. iii) For $\sigma(T^2)$ nonpositive determine as in (ii) R_2^- and $\hat{\omega}$.

i) Since convergence is guaranteed iff $\sigma(T^2) \subset P_1$ and no other information on $\sigma(T^2)$ is given it suffices to have, for all possible T, $\rho(T^2) < 1$ or $\rho(T) < 1$. On the other hand as was proved before for each $\rho = \rho(T^2) (< 1)$, in (4.7), convergence is ensured for all $\omega \in (0,1]$. Since the latter value for ω is in fact ω', from (4.1), the values for ω for which convergence is guaranteed are those satifsying $0 < (2-\omega)\omega \le 1$. Therefore

$$R_2 := \{(\rho(T), \omega) : 0 \le \rho(T) < 1,\ 0 < \omega < 2\} \quad . \tag{4.10}$$

ii) From the requirement $\sigma(T^2) \subset P_1$ and in view of our further assumption it is now sufficient and necessary to have $\rho(T^2) < 1$ leading to the same conclusion as before. Consequently

$$R_2^+ := R_2 \quad . \tag{4.11}$$

To determine the optimum ω for a given $\rho(T)$ we observe that since $\rho = \rho(T^2)$ is given, only one family of ellipses $\mathcal{E}_{\omega,\rho^2(T)}$ is associated with the specific ρ. So, the optimum ω corresponds to the smallest of them and hence $\omega = 1$. This gives $\hat{\omega} = 1$ for the SSOR method as well.

iii) Now we are in the degenerate case of the straight-line segments $\mathcal{E}_{\omega,0}$. So, convergence is guaranteed for all $(-\rho^2(T), 0) \in P_1$ that is for any $\rho(T)$. Obviously for a given $\rho(T) < 1$ we have convergence provided c, in (4.6), satisfies $c \le 0$ or $\omega \in (0,1]$, leading to the conclusion above. For $\rho(T) \ge 1$ convergence is guaranteed for any $\rho \in (0,1)$ provided c satsifies, from (4.6), $c = \frac{-2(1-\omega)}{\omega^2} < \frac{1-\rho^2(T)}{2}$ or $\rho^2(T) < \frac{(2-\omega)^2}{\omega^2}$. Going back to the original ω for the SSOR method one obtains $\rho(T) < \frac{1+(1-\omega)^2}{\omega(2-\omega)}$. Thus

$$R_2^- := \{(\rho(T),\omega) : \omega \in (0,2),\ 0 \le \rho(T) < \frac{1+(1-\omega)^2}{\omega(2-\omega)}\} \tag{4.12}$$

(see [4]). Again for a given $\rho(T)$ the optimum ω corresponds to the smallest degenerate ellipse that is to the straight-line segment $[-\rho^2(T),0]$. It is $c = -\frac{\rho^2(T)}{2} = -\frac{2(1-\omega)}{\omega^2}$ or $\omega = 2/(1+(1+\rho^2(T))^{1/2})$. From the latter two optimum values $\hat{\omega}$ for the SSOR are obtained, namely

$$\hat{\omega}_1 = 1 - \frac{\rho(T)}{1+(1+\rho^2(T))^{1/2}}\ ,\ \hat{\omega}_2 = 1 + \frac{\rho(T)}{1+(1+\rho^2(T))^{1/2}} \quad . \tag{4.13}$$

References

1. Berman, A. and R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, 1979.

2. Galanis, S., A. Hadjidimos and D. Noutsos, "On the Equivalence of the k-Step Iterative Euler Methods and Successive Overrelaxation (SOR) Methods for k-Cyclic Mactrices", *Math. Comput. Simulation*, **30** (1988), 213-230.

3. Galanis, S., A. Hadjidimos and D. Noutsos, "On the Relationship Between the Jacobi and the Successive Overrelaxation (SOR) Matrices of the k-Cyclic Matrix, *Comp. & Math. with Appls.*, (to appear).

4. Hadjidimos, A. and M. Neumann, "Precise Domains of Convergence for the Block SSOR Method Associated with p-Cyclic Matrices", submitted for publication.

5. Hadjidimos, A. and D. Noutsos, "The Determination of the Optimum Real Relaxation Factor for the 2-Cyclic SSOR Method", (to appear).

6. Harary, F., *Graph Theory*, Addison-Wesley, Reading, Massachusetts, 1969.

7. Nichols, N.K. and L. Fox "Generalized Consistent Ordering and the Optimum Successive Over-relaxation Factor", *Numer. Math.*, **13** (1969), 425-433.

8. Niethammer, W. and R.S. Varga, "The Analysis of k-Step Iterative Methods for Linear Systems from Summability Theory", *Numer. Math.*, **41** (1983), 177-206.

9. Parsons, B.N., "General k-Part Stationary Iterative Solutions to Linear Systems", *SIAM J. Numer. Anal.*, **24** (1987), 188-198.

10. Varga, R.S., *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey, 1962.

11. Varga, R.S., W. Niethammer and D.-Y. Cai, "p-Cyclic Matrices and the Symmetric Successive Overrelaxation Method", *Linear Algebra Appl.*, **58** (1984), 425-439.

12. Verner, J.H. and M.J.M. Bernal, "On the Generalizations of the Theory of Consistently Orderings for Successive Overrelaxation Methods", *Numer. Math.*, **12** (1968), 215-222.

13. Young, D.M., *Iterative Methods for Solving Partial Difference Equations of Elliptic Type*, Doctoral Thesis, Harvard University, 1950.

14. Young, D.M., *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.

15. Young, D.M. and D.R. Kincaid, "Norms of the Successive Overrelaxation Methods", Report TNN-94, Computation Center, University of Texas, Austin, Texas, 1969.