On the Convergence of the Cyclic Jacobi Method*

Larry Nazareth

Applied Mathematics Division
Argonne National Laboratory
Argonne, Illinois

Submitted by Peter Henrici

ABSTRACT

In this paper we develop a class of cyclic orderings and prove that the cyclic Jacobi method using any cyclic ordering from this class converges to the eigensystem of a symmetric matrix.

1. INTRODUCTION

The convergence of the cyclic Jacobi method for computing the eigensystem of a symmetric matrix has never been decisively settled. Forsythe and Henrici [1] established convergence to a diagonal matrix for cyclic orderings by rows or by columns. Hansen [2] has shown that of the possible cyclic orderings, some are equivalent, and convergence for a given cyclic ordering implies convergence for all cyclic orderings equivalent to it. A good discussion of Jacobi methods may be found in Wilkinson [3].

In this paper we assume that the reader is familiar with Jacobi methods and has some acquaintance with the works cited above. Here we identify a fairly broad class of cyclic orderings for which we prove convergence of the cyclic Jacobi method. Our discussion is organized as follows. In Sec. 2, we introduce briefly the cyclic Jacobi method and some basic notation. In Sec. 3 we develop the class of cyclic orderings, and in Sec. 4 we prove convergence. Our aim is to address certain theoretical questions concerning the cyclic Jacobi process. In practice, of course, the threshold cyclic Jacobi

*This research was initially supported at The University of California, Berkeley by the Office of Naval Research, Contract N00017-A-0200-1017, and completed under the auspices of the Atomic Energy Commission.
process would be implemented, and the convergence of this method for an
arbitrary cyclic ordering is well known (see Wilkinson [3]).

2. THE CYCLIC JACOBI METHOD

We deal throughout with real matrices. A Jacobi Method seeks to reduce
a symmetric $n \times n$ matrix $A$ to diagonal form through a series of similarity
transformations using plane rotations. A plane rotation corresponding to a
pair of indices $(i, j)$, $1 < i < j < n$, is an orthogonal matrix of the form

$$ U = (u_{pq}), $$

where

$$
\begin{align*}
   u_{pp} &= 1, \\
   u_{ii} &= u_{ij} = \cos \phi, \\
   u_{ij} &= -u_{ii} = \sin \phi, \\
   u_{pq} &= 0 \quad \forall \text{ other pairs.}
\end{align*}
$$

$\phi$ is a real angle called the angle of rotation.

In a Jacobi method we define a sequence of matrices $A_k = (a_{pq}^k)$ by

$$ A_0 = A, \quad A_{k+1} = U_k A_k U_k^T, \quad k = 0, 1, 2, \ldots, \quad (1) $$

where each $U_k$ is a plane rotation corresponding to a pair $(i_k, j_k)$, with angle
of rotation $\phi_k$. $(i_k, j_k)$ will be termed the \textit{rotated element} henceforth. The
eigenvalues of $A_k$ are the same as those of $A$, and we want $A_k$ to tend in the
limit to a diagonal matrix corresponding to the set of eigenvalues.

To simplify the notation let us drop the subscript on $(i_k, j_k)$ and $\phi_k$.

From (1) we have

$$
\begin{align*}
   a_{ti}^{(k+1)} &= a_{ti}^{(k)} \cos \phi + a_{ti}^{(k)} \sin \phi, \\
   a_{ti}^{(k+1)} &= -a_{ti}^{(k)} \sin \phi + a_{ti}^{(k)} \cos \phi, \quad t \neq i, j, \\
   a_{ij}^{(k+1)} &= a_{ij}^{(k)} \cos^2 \phi + 2a_{ij}^{(k)} \sin \phi \cos \phi + a_{ij}^{(k)} \sin^2 \phi, \\
   a_{ij}^{(k+1)} &= a_{ij}^{(k)} \sin^2 \phi - 2a_{ij}^{(k)} \sin \phi \cos \phi + a_{ij}^{(k)} \cos^2 \phi, \\
   a_{ij}^{(k+1)} &= (a_{ij}^{(k)} a_{ij}^{(k)} \cos \phi \sin \phi + a_{ij}^{(k)} (\cos^2 \phi - \sin^2 \phi). \quad (2)
\end{align*}
$$
The cyclic Jacobi method chooses pairs \((i, j)\) in some cyclic ordering (e.g. by rows), and the corresponding angle of rotation \(\phi\), according to some rule. Perhaps the most natural rule, and the one we shall consider throughout this paper, seeks \(\phi\) such that \(a_{ij}^{(k+1)} = 0\).

Thus, from (2),
\[
\tan 2\phi = \frac{2a_{ij}^{(k)}}{a_{ii}^{(k)} - a_{jj}^{(k)}}
\]
and, in the event that \(a_{ij}^{(k)} = a_{ii}^{(k)}\), we take \(\phi = (\pi/4) \text{sgn} a_{ij}^{(k)}\). \footnote{sgn \(x = -1\) if \(x < 0\), \(0\) if \(x = 0\), \(1\) if \(x > 0\).} Many other variants of this rule are possible.

The following results may be shown (see Wilkinson [3], Forsythe and Henrici [2]):

**Fact 1:** \(\phi\) may always be chosen to satisfy \(|\phi| < \pi/4\)

**Fact 2:**
\[
(a_{ii}^{(k)})^2 + (a_{jj}^{(k)})^2 = (a_{ii}^{(k+1)})^2 + (a_{jj}^{(k+1)})^2 \quad \forall t \neq i \text{ or } j.
\]

**Fact 3:** \(\forall t \neq i \text{ or } j\), if \(\max[|a_{ii}^{(k)}|, |a_{ij}^{(k)}|] < \epsilon\), then \(\max[|a_{ii}^{(k+1)}|, |a_{ij}^{(k+1)}|] < \sqrt{2} \epsilon\), and if \(\max[|a_{ii}^{(k)}|, |a_{ij}^{(k)}|] > \epsilon\), then \(\max[|a_{ii}^{(k+1)}|, |a_{ij}^{(k+1)}|] > \epsilon/\sqrt{2}\).

**Fact 4:** Defining \(S_k = \sum_{1 \leq p < q \leq n} (a_{pq}^{(k)})^2\), we have \(S_{k+1} = S_k - (a_{ii}^{(k)})^2\).

**Fact 5:** The rotated element tends to zero as \(k \to \infty\).

**Lemma 1** (Forsythe and Henrici). Suppose \(|a_{ii}^{(k)}| < \epsilon/C\) and \(|a_{ij}^{(k)}| > \epsilon\), where \(C = (1 + \sqrt{2})\). Then \(|a_{ij}^{(k+1)}| > \epsilon/C\).

**Terminology.**

1. \(|a_{ij}^{(k)}|\) will be called the weight on pair \((i, j)\).
2. If \((i, j)\) is the rotated element, we shall say that \((i, t)\) and \((j, t)\) are coupled (where \(t \neq i \text{ or } j\)).

3. **The Class of Cyclic Patterns.**

Hansen [2] has pinpointed difficulties associated with convergence proofs for the general cyclic Jacobi method. He demonstrated that given any real
number \( R \) and any cyclic ordering that contains a subsequence \((p,q), (r,t), (p,r), (q,t), (p,t), (q,r)\), one can find a matrix whose sums of squares of off-diagonal elements is \( R \) and for which there is an arbitrarily small decrease in the sums of squares of off-diagonal elements during one complete cycle of iterations.

Here we develop a class of cyclic orderings for which we have been able to demonstrate convergence. (This choice and the proofs were motivated by some earlier work, Nazareth [4, Part II], which dealt with the analysis of an algorithm for unconstrained minimization using conjugate directions.)

The class of cyclic orderings \( \mathcal{C} \) is most conveniently developed using a recursive definition. This is a little unfortunate for the reader unfamiliar with recursion. We therefore give a detailed illustration after the formal definition.

2.1. Procedure \( P \)

When the following procedure is called with its input parameter \( G \) set equal to the set of indices \( \{1, 2, \ldots, n\} \), it returns with a cyclic ordering in the output parameter \( S \). The class \( \mathcal{C} \) consists of all possible cyclic orderings that this procedure could return.

**PROCEDURE** \( P(G, S) \).

*Comment:* The input parameter \( G \) is a list of consecutive indices \( k, k + 1, \ldots, m \).

*Comment:* The output parameter \( S \) is a list of pairs formed by the procedure using the input \( G \).

**LOCAL VARIABLES** \( G_1, G_2, H, L, T \).

*Comment:* Each call to the procedure results in these local variables being defined for that call.

**Step A:** If \( G \) contains only a single member, then set \( S \) equal to the null list and return.

**Step B:** Otherwise, split \( G \) into two non-empty sets of consecutive indices

\[
G_1 = \{k, k + 1, \ldots, l - 1, l\} \quad \text{and} \quad G_2 = \{l + 1, \ldots, m - 1, m\}.
\]

**Step C:** Form a list of pairs \( L \) as follows: either (i) pick any member of the first group \( G_1 \) and pair it with every member of the second group \( G_2 \) taken in any order (repeat until all members of \( G_1 \) are exhausted), or (ii) carry our (i) with \( G_1 \) and \( G_2 \) interchanged.

**Step D:** Call \( P(G_1, H) \).
Comment: This is a recursive call to the procedure $P$ with input parameter set to $G_1$ and it returns with a list of pairs in $H$.

Step E: Call $P(G_2, T)$.

Comment: This is another recursive call to the procedure $P$ with input parameter set equal to $G_2$, and this call returns with a list of pairs in $T$.

Step F: Form the output list by concatenating $H, L$ and $T$, and return this in $S$. Thus $S = HLT$.

RETURN.

3.2. Example

Figure 1 illustrates in detail the development of a typical member of the class of cyclic orderings when $n = 7$.

Thus the first call to the procedure $P$ is made with the input parameter $G$ set equal to the indices $\{1,2,\ldots,7\}$. This is split at step $B$ into $\{1,2,3,4\}$ and $\{5,6,7\}$. The list $L$ formed at step $C$ might then be $(2,6), (2,5), (2,7), (3,5), (3,7), (3,6), (1,5), (1,6), (1,7), (4,6), (4,5), (4,7)$, corresponding to the first option (i). The recursive call at step $D$ is then made with input parameter $G_1 = \{1,2,3,4\}$. This returns after three further levels of recursion with the output parameter $H$ set equal to, say, $(1,2), (2,4), (1,3), (2,3), (3,4), (2,6), (2,5), \ldots (4,5), (4,7), (5,6), (6,7), (5,7)$. Similarly the recursive call at step $E$ could return with the output parameter $T$ set equal to $(5,6), (6,7), (5,7)$. These are then concatenated to give the
ordering of Fig. 1. This is a typical member of the class of ordering defined by the above procedure for $n = 7$.

Suppose we write out, as in Fig. 2, the set of all pairs for $n = 7$. The cyclic pattern defined in the example is given by taking each pair of Fig. 2 in the order determined by the number associated with it.

Consider the first partition of the example, i.e., $G_1 = \{1,2,3,4\}$ and $G_2 = \{5,6,7\}$. Referring to Fig. 2, all pairs with both members in $G_1$ are contained in triangle $T_1$. All pairs with both members in $G_2$ are contained in triangle $T_2$. All pairs with one member in $G_1$ and the other in $G_2$ are given by rectangle $R$. Forming the list $L$ corresponds to doing Step C(i), i.e., corresponds to picking one row of rectangle $R$ and taking all pairs in it in any order, then doing this in sequence until all rows of $R$ are exhausted. The process is then repeated recursively within triangles $T_1$ and $T_2$. All pairs taken from $T_1$ come before those taken from $R$, and all pairs taken from $T_2$ come after those taken from $R$. At the next level of recursion another rectangle identified within $T_1$, and this time Step C(ii) is used, with selection by columns. The procedure terminates when no further pairs can be formed.

REMARKS.

1. Cyclic orderings by rows and by columns are two particular members of the class of cyclic orderings $\mathcal{C}$.

2. Suppose at step A of the above procedure, the set of indices $G$ were subdivided into any two disjoint subsets instead of merely being split. A cyclic ordering thus obtained would be equivalent to an ordering in $\mathcal{C}$.
applied to a suitable permuted initial matrix. Thus subdividing $G$ at step $A$ would not lead to a more general class of cyclic orderings.

4. PROOF OF CONVERGENCE

4.1. Summary

The proof of convergence rests upon the following claim. Suppose at the start of a cycle of iterations some off-diagonal element has weight $> \varepsilon$. Then, using an ordering from $\mathcal{O}$, some rotated element during a complete cycle of iterations has weight $> M(n)\varepsilon$, where $M(n)$ is a fraction dependent only on $n$.

The proof of convergence then follows fairly directly. From Fact 4, Sec. 2, $S_k$ converges. Suppose $S_k \to \delta$ and $\delta > 0$. Then some off-diagonal element of $A_k$, for all $k$, must have weight exceeding $\sqrt{\delta/N}$, $N = n(n-1)$. By the above claim, with $\varepsilon = \sqrt{\delta/N}$, some rotated element has weight $> M(n)\sqrt{\delta/N}$, and this contradicts Fact 5, namely, that the rotated elements tend to zero as $k \to \infty$.

Most of our effort goes into proving the above claim. This is done through a series of lemmas. Before each lemma we try to give some motivation for developing it.

4.2. Notation

The following notation will be used throughout the proofs and is illustrated by Fig. 3.
Given integers $h$, $l$, and $m$, with $h < l < m$, we define the following sets of pairs:

- $\alpha[h, l] = \{(h, j) : h < j < l\}$,
- $\beta[h, l, m] = \{(h, j) : l < j < l + m\}$,
- $\gamma[h, l, m] = \{(i, j) : (i < j) \& (h < i < l + m) \& (l < j < l + m)\}$;
- $X[a, b] = \{(i, j) | a < i < j < b\}$,
- $R[h, l, m] = \{(i, j) | (i < j) \& (h < i < l - 1) \& (l < j < l + m)\}$;
- $Y[h, l, m] = \alpha[h, l] \cup \beta[h, l, m] \cup \gamma[h, l, m]$, 
- $Z[h, l, m] = \beta[h, l, m] \cup \gamma[h, l, m]$.

**Lemma 2.** Starting with the matrix $A_k$, carry out $m + 1$ further iterations of the cyclic Jacobi method, which consist of successively rotating elements $(h, l), (h, l + 1), \ldots, (h, l + m)$. Let $A_{k+m+1}$ represent the final matrix obtained. Suppose $|a_{\lambda, \rho}^{(k+m+1)}| \geq \delta$ for some pair $(\lambda, \rho) \in Y[h, l, m]$. Then $|a_{\mu, \nu}^{(k)}| \geq \delta / (\sqrt{2})^m$ for some pair $(\mu, \nu) \in Y[h, l, m]$.

**Proof.** The proof follows directly from Fact 2, Sec. 2. If $|a_{ij}^{(k)}| < \epsilon / (\sqrt{2})^m$ for all pairs $(i, j) \in Y[h, l, m]$, then after $m + 1$ iterations the weight on any pair in $Y[h, l, m]$ must be $< \delta$.

Referring to Fig. 4, Lemma 1 states that under certain specified conditions, the weight on $(t, j)$ cannot all be transferred to $(i, t)$ when revising the current pair $(i, j)$. The following lemma is a generalization of this. Referring to Fig. 3 it says that under certain specified conditions the total weight on $Z[h, l, m]$ cannot all be transferred to $\alpha[h, l]$ by successively revising pairs in $\beta[h, l, m]$.
**Lemma 3.** Given a set of indices \( \{h, \ldots, l, \ldots, l+m\} \), suppose that for some pair \((\lambda, \rho) \in \mathbb{Z}[h, l, m]\),

\[
|a_{\lambda\rho}^{(k)}| > \varepsilon. \tag{7}
\]

Suppose that \(m+1\) further steps of the cyclic Jacobi method are carried out, using successive pairs \((h, l), (h, l+1), \ldots, (h, l+m)\) (i.e., successive pairs from the set \(\beta[h, l, m]\)).

Then there exist non-zero fractions \(K_1(m), K_2(m)\) and \(K_3(m)\) which depend only upon \(m\), and are monotonically non-increasing with \(m\) such that if

\[
|a_{hi}^k| < K_1(m)\varepsilon \quad \text{for all } (h, i) \in \alpha[h, l],
\]

then at least one of the following two statements is true:

(i) There exists a pair \((\mu, \nu) \in \gamma[h, l, m]\) such that

\[
|a_{\nu\mu}^{k+m+1}| > K_2(m)\varepsilon.
\]

(ii) Some rotated element during these \(m+1\) iterations has weight \(> K_3(m)\varepsilon\).

**Proof.** With \(h\) and \(l\) fixed, let us use induction on \(m\).

Suppose the lemma is true for all values up to \(m-1\). We must show it to be true for \(m\).

1. If \(l+m\) in (7), then using the induction hypothesis and noting that
the elements of $\gamma[h, l, l+m]$ are unaffected by rotating element $(h, m)$, we see that the lemma holds with

$$K_i(m) = K_i(m-1), \quad i = 1, 2, 3.$$ 

2. If $p = l + m$, then after rotating elements $(h, l), \ldots, (h, l + m - 1)$ the weight of some element in column $l + m$, the last column of Fig. 3, must exceed $\epsilon/\sqrt{2}^m$. This follows from the Fact 3 of Section 2.

2.1. If this element is $(h, l + m)$, then (i) holds in the statement of Lemma 3 with $K_3(m) = 1/\sqrt{2}^m$.

2.2. Suppose not. Say then that this element is $(i, l + m)$ where $h < i < l + m$.

2.2.1. After revising $(h, l + m)$, suppose the weight on $(i, l + m)$ is greater than or equal to $(1/C)\epsilon/\sqrt{2}^m$. Then (i) holds in Lemma 3 with $K_2(m) = 1/C\sqrt{2}^m$.

2.2.2. Suppose the assumption in 2.2.1 does not hold, i.e., after rotating $(h, l + m)$ suppose

$$|a_{h,l+m}^{k+m+1}| < \frac{1}{C} \frac{\epsilon}{{\sqrt{2}^m}}$$

It then follows from Lemma 1 that prior to rotating $(h, l + m)$ the weight on pair $(h, i)$ must have been $\geq (1/C)\epsilon/\sqrt{2}^m$. From Lemma 2 it follows that at the start of the process (i.e., at iteration $k$) some element in $Y[h, l, m]$ has weight $\geq [\epsilon/C(\sqrt{2})^m][1/(\sqrt{2})^m]$. This pair must be in $Z[h, l, m]$. It follows from the induction hypothesis that after revising $(h, l), \ldots, (h, l + m - 1)$ at least one of the following two statements must be true:

(i) For some $(\mu, p) \in \gamma(h, l, m - 1)$,

$$|a_{\mu p}^{(k+m)}| \geq K_2(m-1) \frac{\epsilon}{C2^m}.$$ 

(ii) Some current pair has weight $\geq K_3(m-1)\epsilon/C2^m$. Furthermore the pairs in $\gamma(h, l, m - 1)$ are unaffected when revising the pair $(h, l + m)$.

3. We see, therefore, for all cases in 1 and 2 above, that suitable values
for $K_1(m)$, $K_2(m)$ and $K_3(m)$ are

$$K_1(m) = \frac{K_1(m-1)}{C2^m},$$

$$K_2(m) = \frac{K_2(m-1)}{C2^m},$$

$$K_3(m) = \frac{K_3(m-1)}{C2^m},$$

(8)

4. To complete the argument by induction we must show that Lemma 3 is true when $m = 0$. Suppose some element in $Z[h, l, 0]$ has weight $> \varepsilon$. If this is the current pair, then the lemma holds with $K_3(0) = 1$. If not, then by setting $K_1(0) = 1/C$ and $K_2(0) = 1/C$ we see that Lemma 3 is equivalent, for this case, to Lemma 1. Therefore Lemma 3 holds for $m = 0$.

This completes the inductive argument.

\[ \square \]

**Corollary.** The indices $l, l+1, \ldots, l+m$ may be permuted arbitrarily. It is clear therefore that Lemma 3 holds with the pairs in $\beta[h, l, m]$ revised in any order.

**Notation.** Denote by $L[h, l, m]$ a list of pairs formed as follows. Referring to Figure 3, take the elements of the first rows of $R = R[h, l, m]$ in any order, then the elements of the second row in any order, and so on, until all rows are exhausted.

Lemma 3, we noted, was a generalization of Lemma 1. The next lemma is a generalization of Lemma 3. Referring to Fig. 3 and using the notation developed there, Lemma 4 states that under certain specified conditions the total weight on pairs in $Z[h, l, m]$ (corresponding to triangle $T_2$ and rectangle $R$) cannot all be transferred to pairs in $X[h, l-1]$ (corresponding to triangle $T_1$) by revising pairs in rectangle $R$ taken in sequence from $L[h, l, m]$.

**Lemma 4.** Given a starting matrix $A_k$ and indices $h, \ldots, l, \ldots, l+m$, suppose that for some $(\lambda, \rho) \in Z[h, l, m]$,

$$|a_{\lambda\rho}^{(k)}| > \varepsilon.$$

Carry out $(l-h)(m+1)$ further repetitions of step $C$ using pairs selected in sequence from $L[h, l, m]$. 
Suppose that for all \((i, j) \in X[h, l-1]\)

\[ |a_{ij}^{[k]}| < K_1(m)K_2(m)^{l-h-1} \varepsilon. \]  

(9)

Then at least one of the following two statements is true:

(i) At the end of this process some pair \((\mu, \nu) \in X[l, l+m]\) has weight \(> K_2(m)^{l-h-1} \varepsilon. \)

(ii) Some rotated element during the above \((l-h)(m+1)\) iterations has weight \(> K_3(m)K_2(m)^{l-h} \varepsilon. \)

Proof. With \(l\) and \(m\) fixed, the proof is by induction on \(h\). Suppose the lemma is true for indices \(h+1, \ldots, l, \ldots, l+m\). We must show it to be true for indices \(h, \ldots, l, \ldots, l+m\). Consider rotating successively the first \(m+1\) elements of \(L[h, l, m]\), corresponding to \(\beta[h, l, m]\) in Fig. 3. By assumption all pairs in \(X[h, l-1]\) satisfy (9). Thus, since \(K_2(m) < 1\), all pairs in the set \(\alpha[h, l]\) have weight \(< K_1(m) \varepsilon. \)

Then by the corollary to Lemma 3 at least one of the following statements is true:

(a) Some element in \(\gamma[h, l,m]\) has weight \(> K_2(m) \varepsilon. \)

(b) Some rotated element has weight \(> K_3(m) \varepsilon. \)

Now, if (b) holds for indices \(h, \ldots, l, \ldots, l+m\), then statement (ii) of Lemma 4 is true, since \(K_3(m) > K_3(m)[K_2(m)]^{l-h-1} \varepsilon. \) Suppose therefore that (b) does not hold. Then (a) must hold. Furthermore, no pair in \(X[h+1,l-1]\) is affected when revising pairs in \(\beta[h, l,m]\). Then, by the induction hypothesis, one of the following two statements is true:

(i) At the end of the process some pair \((\mu, \nu) \in X[l, l+m]\) has weight \(> K_2(m)^{l-h-1} \varepsilon. \)

(ii) Some current pair has weight \(> K_3(m)K_2(m)^{l-h-2} \varepsilon. \)

Therefore in all cases Lemma 4 is true for indices \(h, \ldots, l, \ldots, l+m\).

To complete the inductive argument, we need only observe that for \(l-1, l, \ldots, l+m\). Lemma 4 is equivalent to Lemma 3, and therefore the induction hypothesis holds for \(h = l-1. \)

\[ \square \]

Corollary 1. Instead of taking the rows of \(R[h, l, m]\) in increasing order when forming \(L[h, l, m]\), it is clear that Lemma 4 holds when the rows are taken in any order. This corresponds to forming a list of pairs as in step C(i) of Procedure P.
A similar result to Lemma 4 also holds when the pairs in rectangle $R[h, l, m]$ of Fig. 3 are selected in sequence from a list formed as in step C(ii) of Procedure $P$.

**Lemma 5.** Start a fresh cycle of iterations with matrix $A_k$, and suppose that

$$|a_{hp}^{(k)}| \geq \varepsilon$$

for some pair $(\lambda, \rho) \in X[1, n]$. Then during this cycle some rotated element $(p, q)$ has weight $\geq M(n)\varepsilon$, where $M(n)$ is a non-zero fraction dependent only on $n$.

**Proof.** The proof is by induction on $n$.

Suppose Lemma 5 is true for matrices of dimension up to $n - 1$. We must show it to be true for matrices of dimension $n$.

Consider the first partition used to define the cyclic pattern employed. Say it is $\{1, 2, \ldots, l-1\}$ and $\{l, \ldots, n\}$.

1. Suppose some element in $X[1, l-1]$ has weight exceeding

$$\omega = K_1(n-l)[K_2(n-l)]^{l-2}/2^l$$

where $l = l(l-1)/2$. Then by the induction hypothesis some current pair has weight $\geq M(l-1)[K_1(n-l)[K_2(n-l)]^{l-2}/2^l]\varepsilon$. Thus Lemma 5 is true with

$$M(n) \leq M(l-1) \frac{K_1(n-l)[K_2(n-l)]^{l-2}}{2^l}.$$ 

2. Suppose therefore that no element in $X[1, l-1]$ has weight $\geq \omega$ as defined by (11). After revising all elements in $X[1, l-1]$ in the sequence given by the cyclic pattern, no element in $X[1, l-1]$ has weight $\geq K_1(n-l) \cdot [K_2(n-l)]^{l-2}\varepsilon$. This follows from Fact 3 of Section 2. It follows also that the pair $(\lambda, \rho)$ in (11) must be $Z[1, l, n-l]$.

Then by Lemma 4, one of the following two statements is true:

(i) After revising all elements in $R[1, l, n-l]$ in sequence given by the cyclic pattern, some element in $X[l, n]$ has weight $\geq [K_2(n-l)]^{l-1}\varepsilon$.

(ii) Some current pair has weight $\geq K_3(n-l)[K_2(n-l)]^{l-2}\varepsilon$. 


2.1 If (ii) holds, then Lemma 5 is true with

$$M(n) = K_3(n - l)[K_2(n - l)]^{l-2}.$$  

2.2 If (i) holds, then by the induction hypothesis applied to the indices $l, \ldots, n$, some current pair has weight $M(n - l + 1)[K_2(n - l)]^{l-1}$.

3. Therefore, taking

$$M(n) = M(n - 1)K_1(n)K_3(n)[K_2(n)]^n/2^N,$$

where $N = n(n - 1)/2$ covers all cases in 1 and 2 above, we see that Lemma 5 is true for $n$ directions.

4. Trivially the lemma holds for two directions, with $M(2) = 1$, completing the proof by induction.

Using these Lemmas we prove our main theorem:

**Theorem 1.** Apply the cyclic Jacobi method (1) to the symmetric matrix $A$ using any cyclic ordering in the class $C$, and select the angles of rotation $\phi_k$ using (3). Then $A_k$ tends to a diagonal matrix as $k \to \infty$.

**Proof.** The proof follows from Lemma 5 as detailed in Sec. 4.1.

In conclusion we point out that the above results may be extended in several ways—e.g., step C of the procedure for generating the class of cyclic orderings may be generalized. Also the rule for selecting may be extended; cf. Forsythe and Henrici [1].

*My thanks go to Dr. Beresford Parlett for his guidance and advice and to the Applied Mathematics Division, Argonne National Laboratory for giving me the facilities to complete the research.*

**REFERENCES**


*Received 18 July 1974*