Optimal Convergence of the Simultaneous Iteration Method for Normal Matrices

by
A. Sidi

Technical Report #725
April 1992
Optimal Convergence of the Simultaneous Iteration Method for Normal Matrices

Avram Sidi
Computer Science Department
Technion - Israel Institute of Technology
Haifa 32000, Israel.

Abstract
Simultaneous iteration methods are extensions of the power method that were devised for approximating several dominant eigenvalues of a matrix and the corresponding eigenvectors. The convergence analysis of these methods has been given for both hermitian and nonhermitian matrices. In this paper we concentrate on normal matrices which include also the hermitian matrices, and provide a rigorous analysis for a common version of simultaneous iteration methods. We improve on some of the known results and derive some new ones as well. We also present a detailed analysis concerning the formation of spurious eigenvalue approximations.
1 Introduction

Simultaneous iteration methods are extensions of the power method that were devised for approximating several dominant eigenvalues of a matrix and the corresponding eigenvectors. In these methods power iterations are carried out with a number of trial vectors, this number being equal to the number of dominant eigenvalues and their eigenvectors which we desire to approximate. The first simultaneous iteration method, called Treppeniteration, was developed by Bauer [Ba] and is discussed in Wilkinson [W]. Improvements of this method were subsequently proposed by Jennings [J1], Clint and Jennings [CJ1,CJ2], Stewart [St1, St2], Rutishauser [R1,R2], Vandergraft [Vand], Jennings and Stewart [JSt], and Stewart and Jennings [StJ]. See also the books by Jennings [J2], Parlett [Pa], and Golub and Van Loan [GVan].

A rigorous convergence analysis of the simultaneous iteration method for hermitian matrices is given in [St1], while [St2] provides the analysis for nonhermitian matrices. A less rigorous approach to the analysis of this method for hermitian matrices can be found in [J2]. We shall describe the results of [St1] and [J2] for hermitian matrices in the next section. We only mention that the result of [St1], although correct, is not the best possible, and that the result of [J2] that pertains only to the eigenvalues, although the best possible, is not optimal and has been obtained by heuristic arguments.

The purpose of the present work is to provide a rigorous and asymptotically optimal convergence analysis for the method of simultaneous iteration as applied to normal matrices. (Recall that a hermitian matrix is also a normal matrix, but the opposite is not necessarily true). We restrict our attention to normal matrices because of the fact that results for nonnormal matrices, although valid for normal matrices too, are not the best possible for the latter.

The plan of this work is as follows: In Section 2 we give a brief description of the method of simultaneous iteration and mention the relevant results of [St1], [R1], and [J2]. In Section 3 we give convergence results for eigenvalue approximations for both simple and multiple eigenvalues under certain sufficient conditions. The main results of Section 3 are Theorems 3.2 and 3.3, and they provide precise rates of convergence. In Section 4 we give convergence results for eigenvector approximations when the corresponding eigenvalue is simple, the main result of this section being Theorem 4.1. In Section 5 we consider the cases in which the sufficient conditions that lead to the results of Sections 3 and 4 are not satisfied. We show that the method of simultaneous iteration converges for these cases too, provided the matrix A is normal. (If A is not normal, convergence does
not necessarily take place). We prove that the method produces spurious eigenvalue approximations as well as approximations to actual eigenvalues, and we provide rates of convergence for actual eigenvalues. The main result of Section 5 is Theorem 5.1. All the necessary mathematics for obtaining the results of Sections 3-5 is put in the appendix so as to make the reading of the paper easy. We believe that the results of the appendix and the techniques used for obtaining them are of interest in themselves.

The techniques of analysis in the present work are based on those developed in Sidi [Si1] for the analysis of other similar extensions of the power method proposed in Sidi and Bridger [SiBr] and on those of Sidi [Si2] that were developed for the analysis of Padé approximants for meromorphic functions. These techniques are purely algebraic/analytic in nature and do not utilize projection operators.

2 Mathematical Description of the Method

Let \( A \) be an \( N \times N \) matrix and denote its eigenvalues and corresponding eigenvectors by \( \mu_i \) and \( v_i \), respectively, \( i = 1, 2, \ldots, N \). We are assuming here that \( A \) is diagonalizable. We order the eigenvalues such that

\[
|\mu_1| \geq |\mu_2| \geq |\mu_3| \geq \cdots.
\]

(2.1)

Suppose that we are interested in \( (\mu_i, v_i), i = 1, \ldots, k \), for some preassigned \( k \). Pick \( k \) linearly independent vectors \( x_1, x_2, \ldots, x_k \) in \( \mathbb{C}^N \), and form the \( N \times k \) matrix \( U \) as

\[
U = [x_1|x_2|\cdots|x_k].
\]

(2.2)

Then the approximations to the eigenvalues \( \mu_1, \ldots, \mu_k \) of \( A \) are taken to be the roots \( \mu_1(n), \ldots, \mu_k(n) \) of the \( k \times k \) generalized eigenvalue problem

\[
[(A^nU)^*(A^nU)]\xi_j(n) = \mu_j(n)[(A^nU)^*(A^nU)]\xi_j(n), \quad j = 1, \ldots, k,
\]

(2.3)

under certain conditions. The eigenvector \( \xi_j(n) \) that corresponds to \( \mu_j(n) \) is used in constructing the approximation to (a scalar multiple of) \( v_j \), which we denote by \( v_j(n) \) and that is given by

\[
v_j(n) = (A^nU)\xi_j(n), \quad j = 1, \ldots, k.
\]

(2.4)

We note that in the relevant literature the \( \mu_j(n) \) and \( v_j(n) \) are called Ritz values and Ritz vectors, respectively.
We note that various formulations of the method of simultaneous iteration used for hermitian and nonhermitian matrices turn out to be equivalent to (2.2)-(2.4) mathematically, although algorithmically they may differ from each other. This equivalence is in the sense that at their nth stage these formulations produce precisely \( \mu_j(n) \) and \( v_j(n) \) mentioned above as approximations to \( \mu_j \) and \( v_j \), respectively, \( j = 1, \ldots, k \). We shall demonstrate this for one of the formulations that appears also in [J2, p.305] and [StJ], and that is known as lopsided iteration. The steps of this method are as follows:

**Step 0.** Set up the \( N \times k \) matrix \( U_0 = [x_1|x_2|\cdots|x_k] \).

**Step 1.** For \( n = 0, 1, \ldots \), do until convergence

(i) Compute \( V_n = AU_n \)

(ii) Compute \( G = U_n^*U_n \) and \( H = U_n^*V_n \) and \( B = G^{-1}H \).

(iii) Find diagonal \( \Theta_n \) and nonsingular \( P_n \) such that \( B = P_n\Theta_nP_n^{-1}, \) i.e., solve the eigenproblem for \( B \). (The jth diagonal element of \( \Theta_n \) and the jth column of \( U_nP_n \) are the approximations to \( \mu_j \) and \( v_j \), respectively.)

(iv) Compute \( U_{n+1} = V_nP_n \)

Now for \( n = 0 \) the eigenproblem \( Bp = \theta p \) is equivalent to the generalized eigenproblem \( U_0^*AU_0\xi = \mu U_0^*U_0\xi \), with \( \theta = \mu \) and \( p = \xi \). Since also \( U_0p \) is the eigenvector approximation corresponding to \( \theta \), we see that the equivalence of the \( n = 0 \) stage of lopsided iteration with the \( n = 0 \) case of (2.2)-(2.4) is complete.

Next, for \( n = 1 \) the eigenproblem \( Bp = \theta p \) is equivalent to the generalized eigenproblem \( (AU_0)^*A(U_0)\xi = \mu(U_0)^*(AU_0)\xi \), with \( \theta = \mu \) and \( \xi = P_0p \). Since also \( U_1p = AU_0P_0p = AU_0\xi \) is the eigenvector approximation corresponding to \( \theta \), we see that the equivalence of the \( n = 1 \) stage of lopsided iteration with the \( n = 1 \) case of (2.2)-(2.4) is also complete.

For all other values of \( n \) this equivalence can now be shown by using induction.

Finally, the \( \mu_j(n), j = 1, \ldots, k \), are the solution of the polynomial equation \( D_n(\mu) = 0 \), where

\[
D_n(\mu) = \det[\mu(A^nU)^*(A^nU) - (A^nU)^*A(A^nU)] \equiv \det C. \tag{2.5}
\]
If we write
\[ C = \begin{bmatrix}
  c_{11} & c_{12} & \cdots & c_{1k} \\
  c_{21} & c_{22} & \cdots & c_{2k} \\
  \vdots & \vdots & \ddots & \vdots \\
  c_{k1} & c_{k2} & \cdots & c_{kk}
\end{bmatrix} \]
then the element \( c_{ij}^{(n)}(\mu) \) is given by

\[ c_{ij}^{(n)}(\mu) = \mu(A^nx)^*(A^nx) - (A^nx)^*A(A^nx). \]  \hspace{1cm} (2.6)

Our analysis of the simultaneous iteration method will be based exclusively on a detailed study of the polynomial \( D_n(\mu) \). This study will start with the \( c_{ij}^{(n)}(\mu) \).

Before we go on, we mention briefly the results obtained in [St1] and mentioned also in [Pa] and [GVan] and that of [J2]. Provided \(|\mu_k| > |\mu_{k+1}|\), [St1] shows that, with proper ordering of the \( \mu_j(n), j = 1, \ldots, k \),

\[ \mu_j(n) - \mu_j = O_\epsilon \left( \left| \frac{\mu_{k+1}}{\mu_j} \right|^n \right) \quad \text{as} \quad n \to \infty \]  \hspace{1cm} (2.8)

and, with proper normalization,

\[ v_j(n) - v_j = O_\epsilon \left( \left| \frac{\mu_{k+1}}{\mu_j} \right|^n \right) \quad \text{as} \quad n \to \infty, \]  \hspace{1cm} (2.9)

provided also that \( \mu_j \) is a simple eigenvalue. Under the condition \(|\mu_k| > |\mu_{k+1}|\), [J2] heuristically improves the result in (2.8) to read, roughly speaking,

\[ \mu_j(n) - \mu_j = O_\epsilon \left( \left| \frac{\mu_{k+1}}{\mu_j} \right|^{2n} \right) \quad \text{as} \quad n \to \infty. \]  \hspace{1cm} (2.10)

The notation \( O_\epsilon(r^n) \) in (2.8)-(2.10) stands for \( O((|r| + \epsilon)^n) \) for any \( \epsilon > 0 \) so that, e.g., (2.8) and (2.9) are equivalent to

\[ \limsup_{n \to \infty} |\mu_j(n) - \mu_j|^\frac{1}{n} \leq \left| \frac{\mu_{k+1}}{\mu_j} \right| \quad \text{and} \quad \limsup_{n \to \infty} ||v_j(n) - v_j||^\frac{1}{n} \leq \left| \frac{\mu_{k+1}}{\mu_j} \right|. \]

In the present work we prove in a rigorous manner that (2.10) is true in general, and that when \( \mu_j \) is a simple eigenvalue, the \( O_\epsilon \) in (2.10) and (2.9) is to be replaced by \( O \) only. For this case we also give a very precise spectral analysis of \( v_j(n) \) in terms of the \( v_i \).

The case in which \( |\mu_k| = |\mu_{k+1}| \) is not treated in any of the works mentioned above. This is the case that produces the spurious eigenvalue approximations as was mentioned in the previous section. The paper by Parlett and Poole [PaPo] considers the convergence of the subspace \( U_n \).
spanned by the columns of the matrix $A^nU$ as $n \to \infty$ under the condition $|\mu_k| = |\mu_{k+1}|$ as well as $|\mu_k| > \mu_{k+1}$. Although the convergence of $U_n$ is an important issue in simultaneous iteration, the results of [PaPo] for $|\mu_k| = |\mu_{k+1}|$ do not seem to have a direct relation to those produced in the present work.

3 Eigenvalue Analysis for Normal Matrices

Let us now assume that the matrix $A$ is normal. This implies that $A$ has $N$ eigenvectors $v_1, v_2, ..., v_N$ that form an orthonormal set in the sense

$$v_i^*v_j = \delta_{ij}. \quad (3.1)$$

Thus every vector in $C^N$ can be expressed as a linear combination of the $v_i$. In particular, the vectors $x_1, ..., x_k$ in (2.1) can be expressed in the form

$$x_r = \sum_{i=1}^{N} \alpha_{ri} v_i, \quad r = 1, 2, ..., k. \quad (3.2)$$

From (3.2) and $Av_i = \mu_i v_i, \quad i = 1, ..., N$, we next have

$$A^n x_r = \sum_{i=1}^{N} \alpha_{ri} \mu_i^n v_i. \quad (3.3)$$

As a result of (3.1) and (3.3) we have the following expansion for $c_{rs}^{(n)}(\mu)$ in (2.7):

$$c_{rs}^{(n)}(\mu) = \sum_{j=1}^{N} \alpha_{rj} \alpha_{sj} |\mu_j|^{2n}(\mu - \mu_j). \quad (3.4)$$

We assume throughout that the $\mu_j$ have been ordered as in (2.1).

Theorem 3.1 below provides the complete expansion of $D_n(\mu)$ without any conditions, and forms the heart of the analysis that follows.

**Theorem 3.1:** Let $D_n(\mu)$ be given as in (2.5)-(2.7), with $c_{rs}$ as in (3.4). Define the scalars $Z_{j_1,j_2,...,j_k}$ by

$$Z_{j_1,j_2,...,j_k} = \begin{vmatrix} \alpha_{1j_1} & \alpha_{2j_1} & \cdots & \alpha_{kj_1} \\ \alpha_{1j_2} & \alpha_{2j_2} & \cdots & \alpha_{kj_2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1j_k} & \alpha_{2j_k} & \cdots & \alpha_{kj_k} \end{vmatrix}, \quad (3.5)$$
\(j_1, j_2, \ldots, j_k\) being positive integers. Then \(D_n(\mu)\) has the expansion

\[
D_n(\mu) = \sum_{1 \leq j_1 < j_2 \cdots < j_k \leq N} |Z_{j_1, j_2, \ldots, j_k}|^2 \left[ \prod_{p=1}^{k} (\mu - \mu_{j_p}) \right] \left[ \prod_{p=1}^{k} |\mu_{j_p}|^{2n} \right].
\]  

(3.6)

**Proof:** Substituting (3.4) in \(D_n(\mu) = \det C\), cf. (2.5) and (2.6), we see that \(D_n(\mu)\) is identical in form to \(H_n(\sigma)\) given in Lemma A.2 with \(\sigma_j = \mu_j\), \(\zeta_j = |\mu_j|^2\), and \(\beta_{rj} = \alpha_{rj}\) there, while \(\alpha_{rj}\) retains its meaning. Consequently, Lemma A.2 applies, and (A.14), the result of this lemma, now becomes (3.6). \(\square\)

Note that since \(N\) is finite, the expansion in (3.6) is true whether the ordering in (2.1) holds or not. This ordering becomes very crucial in the following theorems as they are of asymptotic nature as \(n \to \infty\).

**Theorem 3.2:** In Theorem 3.1 let

\[
|\mu_1| \geq \cdots \geq |\mu_k| > |\mu_{k+1}| \geq \cdots,
\]

and assume that

\[
Z_{1,2,\ldots,k} \neq 0.
\]

Then

\[
D_n(\mu) = |Z_{1,2,\ldots,k}|^2 \left[ \prod_{j=1}^{k} |\mu_j|^{2n} \right] \left[ \prod_{j=1}^{k} (\mu - \mu_j) + O \left( \left| \frac{\mu_{k+1}}{\mu_k} \right|^{2n} \right) \right],
\]

and hence \(D_n(\mu)\) has precisely \(k\) zeros that tend to \(\mu_1, \mu_2, \ldots, \mu_k\) as \(n \to \infty\).

**Proof:** The proof is accomplished by applying Lemma A.3 to the expansion in (3.6). \(\square\)

**Note:** The condition given in (3.8) is equivalent to the requirement that the orthogonal projections \(x_r^{(k)} = \sum_{i=1}^{k} a_{r,i} v_i\) of the respective vectors \(x_r\), \(r = 1, \ldots, k\), be linearly independent. This can be seen by observing that the Gram determinant of the set \(\{x_r^{(k)}\}_{r=1}^{k}\) is simply \(|Z_{1,2,\ldots,k}|^2\), which is nonzero if and only if \(\{x_r^{(k)}\}_{r=1}^{k}\) is a linearly independent set.

We next go on to analyze the convergence rates of the zeros of \(D_n(\mu)\) to the appropriate \(\mu_1, \mu_2, \ldots, \mu_k\) for \(n \to \infty\). As can be seen from the statement of Theorem 3.3 below, different rates of convergence are obtained for the different \(\mu_j\). These rates do not depend on the multiplicities of the \(\mu_j\), however.
Theorem 3.3: Assume that the conditions of Theorem 3.2 hold, and denote the k zeros of $D_n(\mu)$ by $\mu_1(n), \mu_2(n), \ldots, \mu_k(n)$ ordered such that $\mu_j(n) \to \mu_j$ as $n \to \infty$, $j = 1, \ldots, k$.

1. Let $\mu_s$ be a simple eigenvalue for some $s \in \{1, \ldots, k\}$, i.e., $\mu_s \neq \mu_j$ for $1 \leq j \leq k$, $j \neq s$. Then

$$\mu_s(n) = \mu_s + O \left( \left( \frac{\mu_{k+1}}{\mu_s} \right)^{2n} \right) \text{ as } n \to \infty. \quad (3.10)$$

This result can be refined in an optimal way as follows: Let $r$ be that integer for which

$$|\mu_{k+1}| = \cdots = |\mu_{k+r}| > |\mu_{k+r+1}|, \quad (3.11)$$

and define

$$Z_{1,2,\ldots,s-1,s+1,\ldots,k,k+i} = Z_{\{s,k+i\}}. \quad (3.12)$$

Then

$$\mu_s(n) \sim \mu_s - \left[ \sum_{i=1}^{r} Z_{\{s,k+i\}} \frac{Z_{1,2,\ldots,k}}{Z_{1,2,\ldots,k}}^2 (\mu_s - \mu_{k+1}) \right] \left( \frac{\mu_{k+1}}{\mu_s} \right)^{2n} \text{ as } n \to \infty. \quad (3.13)$$

2. Let $\mu_s$ be an eigenvalue of multiplicity $\omega > 1$. Then there are precisely $\omega$ zeros of $D_n(\mu)$ that tend to $\mu_s$. Denote these zeros by $\mu_{s,1}(n), \ldots, \mu_{s,\omega}(n)$. Then a slightly weaker form of (3.10) holds, and this form reads

$$\limsup_{n \to \infty} |\mu_{s,l}(n) - \mu_s|^2 \leq \left( \frac{\mu_{k+1}}{\mu_s} \right)^2, \quad 1 \leq l \leq \omega. \quad (3.14)$$

Proof: The proof of this theorem is accomplished by applying Lemma A.5 to the expansion in (3.6). As before, we let $\zeta_j = |\mu_j|^2$, $\sigma_j = \mu_j$, and $\beta_{rj} = \sigma_{rj}$ in Lemma A.5. \(\square\)

We note that in case $\mu_s$ is a simple eigenvalue the result in (3.13) is the best possible asymptotically in the sense that the coefficient of $|\mu_{k+1}/\mu_s|^{2n}$ in the error is given in full detail. This can be used to deduce the following interesting result: If $r = 1$, $\mu_s$ and $\mu_{k+1}$ are real and $\mu_s > 0$, then $\mu_s(n)$ tends to $\mu_s$ from below when $n$ is sufficiently large. The same is true also when $r > 1$, $\mu_s$ is real and positive, and $\mu_{k+i}$, $1 \leq i \leq r$, are real or come in complex conjugate pairs.

4 Eigenvector Analysis for Normal Matrices

We now analyze the convergence behavior of the eigenvector approximations $v_s(n)$, cf. (2.4), for $n \to \infty$ when $\mu_s$ is a simple eigenvalue. If we denote the $i$th component of the $k$-dimensional eigenvector $\xi_s(n)$ by $\xi_{si}(n)$, we then have

$$v_s(n) = (A^nU)\xi_s(n) = \sum_{i=1}^{k} (A^n x_i)\xi_{si}(n). \quad (4.1)$$
Invoking (3.3) in (4.1), we obtain the spectral decomposition

\[ v_s(n) = \sum_{j=1}^{N} \delta_{sj}(n) \mu_j^n v_j, \]  

(4.2)

where

\[ \delta_{sj}(n) = \sum_{i=1}^{k} \alpha_{ij} \xi_{si}(n). \]

(4.3)

It is obvious that in order to determine the behavior of \( v_s(n) \) for \( n \to \infty \) we need to analyze the behavior of \( \delta_{sj}(n), j = 1, 2, \ldots \), for \( n \to \infty \).

**Theorem 4.1:** Let \( \mu_s \) be a simple eigenvalue, i.e., \( \mu_s \neq \mu_j, 1 \leq j \leq k, j \neq s \). Then, with proper normalization of the \( \xi_{si}(n) \), the \( \delta_{sj}(n) \) satisfy

\[ \delta_{sj}(n) = \begin{cases} 
O(1) & \text{at most, } j \geq k + 1, \\
O(1) & \text{precisely, } j = s, \\
O \left( \mu_{k+1}^{2n} \mu_j^{-2n} \right) & \text{at most, } 1 \leq j \leq k, j \neq s.
\end{cases} \]  

(4.4)

Consequently, the vectors \( v_s(n) \), when normalized properly, satisfy

\[ v_s(n) = v_s + \sum_{j=1}^{k} \epsilon_{sj}(n) v_j, \]  

(4.5)

with

\[ \epsilon_{sj}(n) = \begin{cases} 
O(\mu_s^n) & \text{as } n \to \infty, j \geq k + 1, \\
O(\mu_{k+1}^{2n} \mu_j^{-2n}) & \text{as } n \to \infty, 1 \leq j \leq k, j \neq s.
\end{cases} \]  

(4.6)

which implies that

\[ v_s(n) = v_s + O \left( \left[ \frac{\mu_{k+1}}{\mu_r} \right]^n \right) \text{ as } n \to \infty. \]  

(4.7)

In addition, if \( \mu_r \) and \( \mu_s \) are two simple eigenvalues of \( A \), \( 1 \leq r, s \leq k \), and \( v_r(n) \) and \( v_s(n) \) are normalized such that \( v_j(n)^*v_j(n) = 1, j = r, s \), then

\[ v_r(n)^*v_s(n) = O \left( \left| \frac{\mu_{k+1}}{\mu_r} \right|^n \right) \text{ as } n \to \infty. \]  

(4.8)

**Proof:** First, we recall that the \( \xi_{si}(n) \) are the solution of the homogeneous system of equation

\[ \sum_{r=1}^{k} \xi_{ri}(n) \xi_{si}(n) = 0, \ 1 \leq r \leq k. \]  

(4.9)
Noting the similarity of (4.9) to (A.44) and of (4.3) to (A.45), and after making the substitutions $\zeta_j = |\mu_j|^2$, $\sigma_j = \mu_j$, and $\beta_j = \overline{\alpha_j}$, while retaining $\alpha_{rj}$ in Lemma A.6, we see that the latter applies, and we obtain (4.4). The results in (4.5)-(4.7) are now obtained by substituting (4.4) in (4.3), while that in (4.8) follows from (4.5) and (4.6) recalling that the $v_j$ form an orthonormal set of vectors.

We note that the result in (4.8) is not new, and can be found, e.g., in [Pa].

5 Formation of Spurious Eigenvalue Approximations

An important requirement that makes the results of Theorems 3.2, 3.3, and 4.1 is the condition $|\mu_k| > |\mu_{k+1}|$ in (3.7). In the absence of this condition, i.e., when $|\mu_k| = |\mu_{k+1}|$, the proofs of these theorems are no longer valid and hence we expect some modification to take place in their statements. Actually, this modification can be considerable. In this section we provide new versions of Theorem 3.2 and Theorem 3.3. We recall that Theorem 3.2 concerns the convergence of $D_n(\mu)$ and Theorem 3.3 concerns the convergence of zeros of $D_n(\mu)$ to $k$ largest eigenvalues of the matrix $A$ whether the latter are simple or multiple.

One of the consequences of the absence of the condition $|\mu_k| > |\mu_{k+1}|$ is that some of the zeros of $D_n(\mu)$ are approximations to the first largest eigenvalues of $A$, while the rest are spurious approximations that depend on the $\alpha_{rj}$ and have nothing to do with the spectrum of $A$. Furthermore, the spurious approximations do not share any of the favorable convergence properties of the actual eigenvalue approximations.

Theorem 5.1: In Theorem 3.1 let

$$|\mu_1| \geq \cdots \geq |\mu_t| > |\mu_{t+1}| = \cdots = |\mu_{t+r}| > |\mu_{t+r+1}| \geq \cdots,$$

for some $t \geq 1$ and $r \geq 2$, and let

$$t + 1 \leq k < t + r.$$  \hspace{1cm} (5.2)

For some integers $i_1, i_2, \ldots, i_{k-t}$ satisfying

$$t + 1 \leq i_1 < i_2 < \cdots < i_{k-t} \leq t + r$$  \hspace{1cm} (5.3)

assume also

$$Z_{i_1, i_2, \ldots, i_{k-t}} \neq 0.$$  \hspace{1cm} (5.4)
Define

\[ S(\mu) = \sum_{t+1 \leq j_{t+1} < \cdots < j_k \leq t+r} |Z_{1,2,\ldots,t,j_{t+1},\ldots,j_k}|^2 \prod_{j=t+1}^{t+r} (\mu - \mu_{j_p}) \]  

Then \( S(\mu) \) is a polynomial of degree exactly \( q = k - t \), whose zeros we denote by \( \mu_1', \ldots, \mu_q' \). Consequently, for \( D_n(\mu) \) we have

\[ d_n D_n(\mu) = \prod_{j=1}^t (\mu - \mu_j) \prod_{j=1}^q (\mu - \mu_j') + O(\epsilon^2) \quad \text{as} \quad n \to \infty, \]  

where \( d_n \) is an appropriate constant and

\[ \epsilon_t = \max \left( \frac{|\mu_{t+1}|}{\mu_t}, \frac{|\mu_{t+r+1}|}{\mu_{t+1}} \right). \]  

Consequently, \( D_n(\mu) \) has \( t \) zeros, \( \mu_1(n), \ldots, \mu_t(n) \), that tend to the eigenvalues \( \mu_1, \ldots, \mu_t \), respectively, and \( q = k - t \) zeros that tend to \( \mu_1', \ldots, \mu_q' \).

For \( 1 \leq s \leq t \) and provided \( \mu_s \notin \{\mu_1', \ldots, \mu_q'\} \),

\[ \limsup_{n \to \infty} |\mu_s(n) - \mu_s|^2 \leq \left| \frac{\mu_{t+1}}{\mu_s} \right|^2 \]  

If \( \mu_s \) is a simple eigenvalue, then (5.8) can be improved to read precisely

\[ \mu_s(n) = \mu_s + O \left( \left| \frac{\mu_{t+1}}{\mu_s} \right|^{2n} \right) \quad \text{as} \quad n \to \infty. \]  

**Proof:** The proof is accomplished by employing Lemmas A.7 and A.8. We leave the details to the interested reader. \( \square \)

**Remarks:**

1. The conditions given in (5.1) and (5.2) are implied by and also equivalent to \( |\mu_k| = |\mu_{k+1}| \).

2. The condition in (5.3) and (5.4) is equivalent to the projections of \( x_1, \ldots, x_k \) onto the subspace spanned by \( \{v_1, \ldots, v_t, v_i, \ldots, v_{k-1}\} \) being linearly independent.

3. In case the matrix \( A \) is not normal, then a similar but more complicated analysis shows that convergence of the sequence \( D_n(\mu), n = 0, 1, 2, \ldots \), is not guaranteed, but that there is a subsequence that converges. This is the situation described in part (1) of Lemma A.7.
Appendix

1 General Considerations

The purpose of this appendix is to provide the mathematical tools that will be useful in analyzing the polynomials $D_n(\mu)$ and their zeros in the limit as $n \to \infty$. We believe that the results that we obtain here are of interest in themselves as the techniques used in obtaining them are rather general and may apply to other problems as well.

**Lemma A.1:** Let $i_1, \ldots, i_k$ be positive integers, and assume that the scalars $v_{i_1, \ldots, i_k}$ are odd under an interchange of any two indices $i_1, \ldots, i_k$. Let $t_{i,j}, i \geq 1, 1 \leq j \leq k$, be scalars. Define

$$I_{k,N} = \sum_{i_1=1}^{N} \cdots \sum_{i_k=1}^{N} \left( \prod_{p=1}^{k} t_{i_p,i_p} \right) v_{i_1, \ldots, i_k}$$

and

$$J_{k,N} = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq N} \begin{vmatrix} t_{i_1,1} & t_{i_1,2} & \cdots & t_{i_1,k} \\ t_{i_2,1} & t_{i_2,2} & \cdots & t_{i_2,k} \\ \vdots & \vdots & \ddots & \vdots \\ t_{i_k,1} & t_{i_k,2} & \cdots & t_{i_k,k} \end{vmatrix} v_{i_1, \ldots, i_k}.$$ (A.2)

Then

$$I_{k,N} = J_{k,N}. \quad (A.3)$$

This lemma was stated and proved in [SiFSm].

Lemmas A.2-A.4 below are analogous to Lemmas 2.2 and 2.3 in [Si1].

For simplicity of notation below we shall use the shorthand notation

$$\sum_{j}^{\infty}, \sum_{j_1 < j_2 < \cdots < j_k}^{\infty}, \text{ and } \alpha_n \sim \beta_n$$

to mean, respectively,

$$\sum_{j=1}^{\infty}, \sum_{j_1=1}^{\infty} \sum_{j_2=j_1+1}^{\infty} \cdots \sum_{j_k=j_{k-1}+1}^{\infty}, \text{ and } \alpha_n \sim \beta_n \text{ as } n \to \infty.$$ (A.3)

We note that the polynomial $H_n(\sigma)$ that is defined in Lemma A.2 below and is analyzed throughout is a generalization of the polynomial $D_n(\mu)$. Consequently, all of the results that we prove for $H_n(\sigma)$, after the proper analogy is drawn, are good for $D_n(\mu)$ as well.
Lemma A.2: Let $\sigma_1, \sigma_2, \ldots,$ and $\zeta_1, \zeta_2, \ldots,$ be two sequences of complex numbers, and

$$|\zeta_1| \geq |\zeta_2| \geq |\zeta_3| \geq \cdots; \; \zeta_j \neq 0, \; j = 0, 1, \ldots,$$

(A.4)

and assume that there can be only a finite number of $\zeta_j$'s having the same modulus. Let $H_n(\sigma)$ be a polynomial in $\sigma$ of degree $k$ defined by

$$H_n(\sigma) = \begin{vmatrix}
    u_{11}(\sigma) & u_{12}(\sigma) & \cdots & u_{1k}(\sigma) \\
    u_{21}(\sigma) & u_{22}(\sigma) & \cdots & u_{2k}(\sigma) \\
    \vdots & \vdots & \ddots & \vdots \\
    u_{k1}(\sigma) & u_{k2}(\sigma) & \cdots & u_{kk}(\sigma)
\end{vmatrix},$$

(A.5)

where

$$u_{r3}(\sigma) \sim \sum_{j=1}^{\infty} \beta_{rj} \alpha_{kj} \zeta_j^n (\sigma - \sigma_j) \quad \text{as} \; n \to \infty.$$ 

(A.6)

Define the scalars

$$Z_{j_1,j_2,\ldots,j_k}(\alpha) = \begin{vmatrix}
    \alpha_{1j_1} & \alpha_{2j_1} & \cdots & \alpha_{kj_1} \\
    \alpha_{1j_2} & \alpha_{2j_2} & \cdots & \alpha_{kj_2} \\
    \vdots & \vdots & \ddots & \vdots \\
    \alpha_{1j_k} & \alpha_{2j_k} & \cdots & \alpha_{kj_k}
\end{vmatrix},$$

(A.7)

with $j_1, \ldots, j_k$ being positive integers. Define the scalars $Z_{j_1,j_2,\ldots,j_k}(\beta)$ similarly, and set

$$W_{j_1,j_2,\ldots,j_k} = Z_{j_1,j_2,\ldots,j_k}(\alpha)Z_{j_1,j_2,\ldots,j_k}(\beta).$$

(A.8)

Then we have

$$H_n(\sigma) \sim \sum_{j_1<j_2<\cdots<j_k} W_{j_1,j_2,\ldots,j_k} \left[ \prod_{p=1}^{k} (\sigma - \sigma_{j_p}) \right] \left( \prod_{p=1}^{k} \zeta_{j_p}^n \right).$$

(A.9)

(If the summation in (A.6) is finite, and $\sim$ is replaced by $=$, then the multiple sum in (A.9) is finite, and $\sim$ is replaced by $=$ there too.)

Proof: Substituting (A.6) in (A.5), we obtain

$$H_n(\sigma) \sim \begin{vmatrix}
    \sum_{j_1} \beta_{1j_1} \alpha_{1j_1} \zeta_{j_1}^n (\sigma - \sigma_{j_1}) & \sum_{j_1} \beta_{1j_2} \alpha_{2j_1} \zeta_{j_1}^n (\sigma - \sigma_{j_1}) & \cdots & \sum_{j_1} \beta_{1j_k} \alpha_{kj_1} \zeta_{j_1}^n (\sigma - \sigma_{j_1}) \\
    \sum_{j_2} \beta_{2j_1} \alpha_{1j_2} \zeta_{j_2}^n (\sigma - \sigma_{j_2}) & \sum_{j_2} \beta_{2j_2} \alpha_{2j_2} \zeta_{j_2}^n (\sigma - \sigma_{j_2}) & \cdots & \sum_{j_2} \beta_{2j_k} \alpha_{kj_2} \zeta_{j_2}^n (\sigma - \sigma_{j_2}) \\
    \vdots & \vdots & \ddots & \vdots \\
    \sum_{j_k} \beta_{kj_1} \alpha_{1j_k} \zeta_{j_k}^n (\sigma - \sigma_{j_k}) & \sum_{j_k} \beta_{kj_2} \alpha_{2j_k} \zeta_{j_k}^n (\sigma - \sigma_{j_k}) & \cdots & \sum_{j_k} \beta_{kj_k} \alpha_{kj_k} \zeta_{j_k}^n (\sigma - \sigma_{j_k})
\end{vmatrix}.$$ 

(A.10)
Using the multilinearity property of determinants, and removing common factors from each row, we can express (A.10) in the form

\[ H_n(\sigma) \sim \sum_{j_1} \sum_{j_2} \cdots \sum_{j_k} \left( \prod_{p=1}^k \beta_{p,j_p} \right) \left( \prod_{p=1}^k \zeta^\sigma_{p,j_p} \right) \left( \prod_{p=1}^k (\sigma - \sigma_{j_p}) \right) Z_{j_1,j_2,\ldots,j_k}(\alpha). \]  

(A.11)

Since the product \( \left( \prod_{p=1}^k \zeta^\sigma_{p,j_p} \right) \left( \prod_{p=1}^k (\sigma - \sigma_{j_p}) \right) \) in (A.11) is odd under an interchange of any two of the indices \( j_1, \ldots, j_k \), Lemma A.1 applies, and (A.9) follows. \( \square \)

2 Analysis of the Zeros of \( H_n(\sigma) \) When \( |\zeta_k| > |\zeta_{k+1}| \)

We now start the analysis of the zeros of \( H_n(\sigma) \). We show that, under appropriate conditions, the zeros of \( H_n(\sigma) \) tend to \( \sigma_1, \sigma_2, \ldots, \sigma_k \) as \( n \to \infty \). We also provide the precise rates of convergence.

Lemma A.3 If in Lemma A.2 we also assume that

\[ |\zeta_k| > |\zeta_{k+1}| \]  

(A.12)

and

\[ W_{1,2,\ldots,k} \neq 0, \]  

(A.13)

then, for \( \sigma \neq \sigma_i, 1 \leq i \leq k \), the dominant behavior of \( H_n(\sigma) \) is given by

\[ H_n(\sigma) = W_{1,2,\ldots,k} \left( \prod_{j=1}^k \zeta^\sigma_j \right) \left( \prod_{j=1}^k (\sigma - \sigma_j) + O \left( \left| \frac{\zeta_{k+1}}{\zeta_k} \right|^n \right) \right) \]  

as \( n \to \infty \).  

(A.14)

This implies that \( H_n(\sigma) \) has precisely \( k \) zeros that tend to \( \sigma_1, \sigma_2, \ldots, \sigma_k \) as \( n \to \infty \).

Proof: By (A.12) and (A.13) the dominant term in (A.9) is that for which \( j_1 = 1, j_2 = 2, \ldots, j_k = k \), this term being of order \( |\zeta_1 \zeta_2 \cdots \zeta_k|^n \). The next dominant terms are of order \( |\zeta_1 \zeta_2 \cdots \zeta_{k-1} \zeta_{k+1}|^n \) by the assumption that there can be only a finite number of \( \zeta_j \)'s having modulus equal to \( |\zeta_{k+1}| \). By this (A.14) now follows. \( \square \)

Definition: Let \( s \) be an integer in \( \{1, 2, \ldots, k\} \). We shall say that \( \sigma_s \) has multiplicity \( \omega \) if the set \( \{\sigma_1, \sigma_2, \ldots, \sigma_k\} \) contains precisely \( \omega \) elements that are equal to \( \sigma_s \), including \( \sigma_s \) itself. If \( \omega = 1 \), we shall say that \( \sigma_s \) is simple, otherwise, we say that it is multiple. In case \( \sigma_s \) has multiplicity \( \omega > 1 \), we shall also assume that \( \sigma_s = \sigma_{s+1} = \cdots = \sigma_{s+\omega-1} \) and \( |\zeta_s| = |\zeta_{s+1}| = \cdots = |\zeta_{s+\omega-1}| \), as happens in eigenvalue problems.
Now by Lemma A.3 the zeros $\sigma_1(n), \ldots, \sigma_k(n)$ of $H_n(\sigma)$ tend to $\sigma_1, \ldots, \sigma_k$ as $n \to \infty$, whether the latter are simple or multiple. The rates of convergence, however, depend entirely on the $\sigma_j$ but not on their multiplicities as we show in Lemma A.4 below. Before going into this lemma, however, we wish to present a perturbation lemma concerning the zeros of certain polynomials, which is of interest in itself. We note that Lemma 2.5 in [Si2] is similar in spirit to the perturbation lemma we are about to state, although its results are rather different.

**Lemma A.4:** Let the polynomial $\Phi_n(\sigma)$ be given by

$$\Phi_n(\sigma) = \sum_{i=0}^{k} d_i(n)(\sigma - \hat{\sigma})^i, \text{ for some fixed } \hat{\sigma},$$

such that

$$\lim_{n \to \infty} \frac{d_i(n)}{d_0(n)} = \begin{cases} 
0 & \text{ for } 0 \leq i \leq \omega - 1, \\
\hat{d}_i & \text{ for } \omega \leq i \leq k.
\end{cases}$$

1. $\Phi_n(\sigma)$ has precisely $\omega$ zeros $\hat{\sigma}_l(n), 1 \leq l \leq \omega$, that tend to $\hat{\sigma}$ as $n \to \infty$.

2. If $\omega = 1$, then the unique zero $\hat{\sigma}_1(n)$ of $\Phi_n(\sigma)$ that tends to $\hat{\sigma}$ satisfies precisely

$$\hat{\sigma}_1(n) \sim \hat{\sigma} - \frac{d_0(n)}{d_1(n)} \text{ as } n \to \infty.$$  

Consequently, if

$$\frac{d_0(n)}{d_1(n)} = O(\epsilon_1^n) \text{ as } n \to \infty, \text{ some } \epsilon_1 \in (0,1),$$

then it follows directly from (A.17) that

$$\hat{\sigma}_1(n) = \hat{\sigma} + O(\epsilon_1^n) \text{ as } n \to \infty.$$  

3. For $\omega > 1$ assume that the $d_i(n)$ satisfy

$$\frac{d_i(n)}{d_\omega(n)} = O((\epsilon_1 \epsilon_2 \cdots \epsilon_{\omega-1})^n) \text{ as } n \to \infty, \text{ } 0 \leq i \leq \omega - 1,$$

cf. (A.18), where

$$1 > \epsilon_1 \geq \epsilon_2 \geq \cdots \geq \epsilon_{\omega-1} \geq 0.$$  

Then, the result of (A.19) is replaced by the slightly weaker result

$$\limsup_{n \to \infty} |\hat{\sigma}_l(n) - \hat{\sigma}|^{\frac{1}{n}} \leq \epsilon_1, \text{ } 1 \leq l \leq \omega.$$  

(A.22)
Proof: The proof of part (1) is obvious from
\[
\lim_{n \to \infty} \frac{\Phi_n(\sigma)}{d_\omega(n)} = (\sigma - \hat{\sigma})^\omega \left[ 1 + \sum_{i=\omega+1}^k \hat{d}_i(\sigma - \hat{\sigma})^{i-\omega} \right],
\]
(A.23)
which follows from (A.15) and (A.16).

For the proof of parts (2) and (3) we start by rewriting \( \Phi_n(\hat{\sigma}_I(n)) = 0 \) in the form
\[
-(\hat{\sigma}_I(n) - \hat{\sigma})^\omega = \frac{\sum_{i=0}^{\omega-1} [d_i(n)/d_\omega(n)](\hat{\sigma}_I(n) - \hat{\sigma})^i}{1 + \sum_{i=\omega+1}^k [d_i(n)/d_\omega(n)](\hat{\sigma}_I(n) - \hat{\sigma})^{i-\omega}}.
\]
(A.24)
Now by the assumptions in (A.16) and the fact that \( \hat{\sigma}_I(n) - \hat{\sigma} = o(1) \) as \( n \to \infty \), which follows from part (1), the denominator of (A.24) is asymptotically equivalent to 1 as \( n \to \infty \).

If \( \omega = 1 \), the numerator of (A.24) is simply \( d_0(n)/d_\omega(n) \), so that we immediately have (A.17).

If \( \omega > 1 \), we first observe that the numerator of (A.24) is \( O(\epsilon_1^n) \) as \( n \to \infty \) by (A.20), (A.21), and the fact that \( \hat{\sigma}_I(n) - \hat{\sigma} = o(1) \) as \( n \to \infty \). Therefore, (A.24) gives \( \hat{\sigma}_I(n) - \hat{\sigma} = O(\epsilon_1^{n/\omega}) \) as \( n \to \infty \). This implies that
\[
\limsup_{n \to \infty} |\sigma_I(n) - \hat{\sigma}|^\frac{1}{n} \leq \epsilon_1^{\frac{1}{\omega}}.
\]
(A.25)
Since \( 0 < \epsilon_1 < 1 \), we have \( \epsilon_1 < \epsilon_1^{1/\omega} \) so that either \( \rho \leq \epsilon_1 \) or \( \rho > \epsilon_1 \). We need to prove that it is \( \rho \leq \epsilon_1 \) that holds. Suppose to the contrary that \( \rho > \epsilon_1 \). Then given \( a > 0 \) but arbitrary otherwise,
\[
|\hat{\sigma}_I(n) - \hat{\sigma}| < (\rho + a)^n \text{ for all large } n,
\]
(A.26)
and, for some integers \( n_0 < n_1 < n_2 < \cdots \),
\[
(\rho - a)^{n_i} < |\hat{\sigma}_I(n_i) - \hat{\sigma}|, \quad i = 0, 1, 2, \ldots
\]
(A.27)
By (A.20), (A.21), and (A.26), and by the assumption that \( \rho > \epsilon_1 \), each term in the numerator of (A.24) is less than \( C\epsilon_1^n(\rho + a)^{(\omega-1)n} \) for some constant \( C > 0 \). Using this on the right hand side of (A.24), and (A.27) on its left hand side, we have
\[
(\rho - a)^{\omega n_i} < |\hat{\sigma}_I(n_i) - \hat{\sigma}|^\omega < K\epsilon_1^{n_i}(\rho + a)^{(\omega-1)n_i}, \quad i = 0, 1, 2, \ldots
\]
(A.28)
for some constant \( K > 0 \). Taking the \( n_i \)-th root of both sides of (A.28), and going to the limit, we have
\[
(\rho - a)^\omega \leq \epsilon_1(\rho + a)^{(\omega-1)}, \quad a > 0 \text{ arbitrary}.
\]
(A.29)
As a result we obtain \( \rho \leq \epsilon_1 \), in contradiction with the assumption that \( \rho > \epsilon_1 \). Thus \( \rho \leq \epsilon_1 \). This completes the proof. \( \Box \)
so that

\[ H_n(\sigma_s) \sim \left\{ \sum_{i=j}^r W_{[s:k]} \left[ \prod_{j=1}^k (\sigma_s - \sigma_j) \right] (\sigma_s - \sigma_{k+i}) \left( \frac{\zeta_{k+1}}{\zeta_s} \right)^n \right\} \left( \prod_{j=1}^k \zeta_j^n \right). \]  \hspace{1cm} (A.37)

Now by (A.14) and by the assumption that \( \sigma_s \) is simple

\[ H'_{n}(\sigma_s) = \frac{d}{d\sigma} H_n(\sigma)|_{\sigma = \sigma_s} \sim W_{1,2,\ldots,k} \left[ \prod_{j=1}^k (\sigma_s - \sigma_j) \right] \left( \prod_{j=1}^k \zeta_j^n \right) \neq 0. \]  \hspace{1cm} (A.38)

Similarly, \( H''_{n}(\sigma_s) \) for \( i \geq 2 \) are of the same order as \( H'_{n}(\sigma_s) \). Therefore, part (2) of Lemma A.4 applies with \( d_i(n) \equiv H''_{n}(\sigma_s)/i! \), \( \omega = 1 \), \( \epsilon_1 = |\zeta_{k+1}/\zeta_s| \), and \( \hat{\sigma} = \sigma_s \), and we have

\[ \sigma_s(n) - \sigma_s \sim \frac{H_n(\sigma_s)}{H'_{n}(\sigma_s)}, \]  \hspace{1cm} (A.39)

which, upon invoking (A.37) and (A.38), produces the result in (A.33).

If \( \sigma_s \) has multiplicity \( \omega > 1 \), i.e., if \( \sigma_s = \sigma_{s+1} = \cdots = \sigma_{s+\omega-1} \) and \( |\zeta_s| = |\zeta_{s+1}| = \cdots = |\zeta_{s+\omega-1}| \), then the analysis of the \( H''_{n}(\sigma_s) \) becomes more involved. First, for \( i \geq \omega \) the dominant term of \( H_n(\sigma) \) is obtained by differentiating the term with \( j_1, \ldots, j_k = 1, \ldots, k \), and setting \( \sigma = \sigma_s \) in it, and is of order \( |\zeta_1 \cdots \zeta_k|^n \) as \( n \to \infty \). In particular,

\[ H''_{n}(\omega)(\sigma_s)/\omega! \sim W_{1,2,\ldots,k} \left[ \prod_{j=1}^k \zeta_j^n \right] \left[ \prod_{j=1}^k (\sigma_s - \sigma_j) \right] \neq 0. \]  \hspace{1cm} (A.40)

For \( i = 0 \) set \( \sigma = \sigma_s \) in (A.9). We see that all the terms having any of their indices \( j_1, \ldots, j_k \) equal to \( s, s + 1, \ldots, s + \omega - 1 \) vanish. Therefore,

\[ H_n(\sigma_s) = O \left( \left( \prod_{j=1}^k \zeta_j^n \right) \left( \prod_{l=1}^\omega \zeta_{k+1}^l \right)^n \right) \text{ as } n \to \infty. \]  \hspace{1cm} (A.41)

For \( i = 1 \) differentiate (A.9) term by term and set \( \sigma = \sigma_s \) there. Now all the terms having any \( k - 1 \) of their indices \( j_1, \ldots, j_k \) equal to \( s, s + 1, \ldots, s + \omega - 1 \) vanish. Therefore,

\[ H'_{n}(\sigma_s) = O \left( \left( \prod_{j=1}^k \zeta_j^n \right) \left( \prod_{l=1}^{\omega-1} \zeta_{k+1}^l \right)^n \right) \text{ as } n \to \infty. \]  \hspace{1cm} (A.42)

For \( i \leq \omega - 1 \) differentiate (A.9) \( i \) times term by term and set \( \sigma = \sigma_s \) there. All the terms having any \( k - i \) of their indices \( j_1, \ldots, j_k \) equal to \( s, s + 1, \ldots, s + \omega - 1 \) vanish. Therefore, for \( i \leq \omega \),

\[ H''_{n}(\sigma_s) = O \left( \left( \prod_{j=1}^k \zeta_j^n \right) \left( \prod_{l=1}^{\omega-i} \zeta_{k+1}^l \right)^n \right) \text{ as } n \to \infty. \]  \hspace{1cm} (A.43)
Hence part (3) of the previous lemma applies with $d_i(n) = H^{(0)}_n(\alpha_i)/i!$, $\epsilon_i = |\zeta_{k+i}/\zeta_i|$, and $\delta = \sigma_s$, and we obtain (A.34). (The reader is urged to verify (A.41) - (A.43) for small values of $k$ and $\omega$.)

This completes the proof. □

3 A Further Result for Simple $\sigma_s$ When $|\zeta_k| > |\zeta_{k+1}|$

Lemma A.6: Assume that the conditions of Lemma A.9 hold, and denote the zeros of $H_n(\sigma)$ that tend to $\sigma_1, ..., \sigma_k$ by $\sigma_1(n), ..., \sigma_k(n)$, respectively. Let $\sigma_s$ be simple for some $s \in \{1, ..., k\}$. Then the solution of the homogeneous system of equations, cf. (A.5),

$$\sum_{i=1}^{k} y_{ri}^{(n)}(\sigma_s(n))\xi_{ni}(n) = 0, \ 1 \leq r \leq k, \quad (A.44)$$

for $n$ sufficiently large, is unique up to a multiplicative constant, and satisfies

$$\left( \prod_{j=1}^{k} \zeta_j^r \right) \sum_{i=1}^{k} \alpha_i n \xi_{ni}(n) = \begin{cases} O(1) & \text{at most, } q \geq k + 1, \\ O(1) & \text{precisely, } q = s, \\ O(\zeta_{k+1}^{-n} \zeta_q^{-n}) & \text{at most, } 1 \leq q \leq k, \ q \neq s. \end{cases} \quad (A.45)$$

provided $\sigma_i \neq 0, \ 1 \leq i \leq k$, in addition.

Proof: First, the matrix of the system in (A.44) is singular as its determinant is simply $H_n(\sigma_s(n))$, which itself is zero. By the assumption that $\sigma_s$ is simple, we see from Lemma A.5 that $\sigma_s(n) \neq \sigma_j(n)$ for $j \neq s$ for all large $n$. Now the system in (A.44) can be expressed as the generalized eigenvalue problem $E \xi = \sigma F \xi$, where the $(r, s)$ elements $e_{rs}$ and $f_{rs}$ of $E$ and $F$, respectively, have the expansions, cf. (A.6),

$$e_{rs} \sim \sum_{j=1}^{\infty} \beta_{rj} \alpha_{rj} \sigma_j \zeta_j^n \quad \text{and} \quad f_{rs} \sim \sum_{j=1}^{\infty} \beta_{rj} \alpha_{rj} \zeta_j^n.$$

Thus $\det E = (-1)^k H_n(0)$. But $H_n(0) \neq 0$ for $n$ sufficiently large by (A.14). Consequently, $1/\sigma_j(n), \ 1 \leq j \leq k$, that are naturally bounded away from zero and infinity for all large $n$, are the eigenvalues of the matrix $E^{-1}F$. Combining all this, we conclude that the matrix $E^{-1}F - \sigma_s(n)I$ and hence the matrix of the system in (A.44) have rank $k - 1$ exactly. This implies that the solution of (A.44) is unique up to a multiplicative constant, and that it can be obtained from $k - 1$ of the equations in (A.44).

Next, (A.13) implies that $Z_{1, ..., k}(\alpha) \neq 0$ and $Z_{1, ..., k}(\beta) \neq 0$ simultaneously. Now $Z_{1, ..., k}(\beta) \neq 0$ guarantees that for any $s, 1 \leq s \leq k$, there exists a $(k - 1) \times (k - 1)$ minor of the determinant representation of $Z_{1, ..., k}(\beta)$, cf. (A.7), that does not include the $s$th row and does not vanish.
Without loss of generality and for simplicity of notation, we assume that this minor is obtained by deleting the 8th row and the kth column. This amounts to

\[ Z_{1,2,...,s-1,s+1,...,k}(\beta) \neq 0. \]  

(A.46)

With this we now assume that the \( k - 1 \) equations mentioned at the end of the previous paragraph are the first \( k - 1 \) equations of (A.44). In view of this, it is easy to verify that for any scalars \( \delta_1, \delta_2, ..., \delta_k \), we have

\[
\sum_{i=1}^{k} \delta_i \xi_{si}(n) = \begin{vmatrix}
\delta_1 & \delta_2 & \ldots & \delta_k \\
u_{11} & u_{12} & \ldots & u_{1k} \\
u_{21} & u_{22} & \ldots & u_{2k} \\
\vdots & \vdots & \ddots & \vdots \\
u_{k'1} & u_{k'2} & \ldots & u_{k'k}
\end{vmatrix} \quad \text{with } k' = k - 1, \tag{A.47}
\]

where we have denoted \( u_{pq} = u_{pq}^{(n)}(\sigma_s(n)) \) and we have normalized \( \xi_{si}(n) \) appropriately. Substituting now \( \delta_i = \alpha_{i} \) and (A.6) in (A.47), and proceeding exactly as in the proof of Lemma A.3, we obtain

\[
\sum_{i=1}^{k} \alpha_{i} \xi_{si}(n) \sim \sum_{j_1 < j_2 < \ldots < j_k} Z_{j_1,...,j_k}(\beta)Z_{q,j_1,...,j_k}(\alpha) \left[ \prod_{j=1}^{k'} (\sigma_s(n) - \sigma_{j}) \right] \left( \prod_{j=1}^{k'} \zeta_{j}^{n} \right). \tag{A.48}
\]

Obviously, all those terms in (A.48) for which any one of the indices \( j_1, ..., j_k \) takes on the value \( q \) vanish on account of \( Z_{q,j_1,...,j_k}(\alpha) \) vanishing. That is to say the indices \( j_1, ..., j_k \) in (A.48) take on all values except \( q \).

Let us now analyze (A.48) for all values of \( q \). First, for \( q = s \) the dominant term there has the indices

\[ j_1, j_2, ..., j_{k'}, = 1, 2, ..., s - 1, s + 1, ..., k, \]

and, as \( n \to \infty \), is asymptotically equivalent to

\[
(-1)^{s-1} Z_{1,...,s-1,s+1,...,k}(\beta)Z_{1,2,...,k}(\alpha) \left[ \prod_{j=1}^{k} (\sigma_s - \sigma_j) \right] \left( \prod_{j=1}^{k} \zeta_j^{n} \right). 
\]

The important point here is that this term is precisely \( O(\pi^n) \) as \( n \to \infty \), where \( \pi \equiv \prod_{j=1}^{k} \zeta_j \), since the constant factors in it are all nonzero by (A.13) and (A.46). Next, for \( 1 \leq q \leq k, q \neq s \), we see that the dominant terms have indices \( j_1, ..., j_{k'} \), that take on the values \( 1, 2, ..., q - 1, q + 1, ..., k \) or
the values 1, 2, ..., k, k + i except q and s, i = 1, 2, ..., r. Invoking also (A.15), we see that all these terms are at most \( O(\pi^n \zeta_{k+1}^n / \zeta_q^n) \) as \( n \to \infty \). Finally, for \( q \geq k + 1 \) the dominant term has the same indices as in the case of \( q = s \) and is \( O(\pi^n) \) as \( n \to \infty \). This completes the proof. \( \square \)

4 Treatment of \( H_n(\sigma) \) and Its Zeros for \( |\zeta_k| = |\zeta_{k+1}| \)

The results in Lemmas A.3, A.5, and A.6 are made possible especially by the condition \( |\zeta_k| > |\zeta_{k+1}| \) in (A.12). If this condition is not satisfied, then the proofs of these results are not valid, and the question arises as to whether they can be saved or modified in a simple manner. Lemmas A.7 and A.8 below give a detailed treatment of this question regarding Lemma A.3 for \( H_n(\sigma) \) and Lemma A.5 for the \( \sigma_s \) respectively. We shall not pursue the modification of Lemma A.6.

Lemma A.7: In Lemma A.2 let

\[
|\zeta_1| \geq \cdots \geq |\zeta_t| > |\zeta_{t+1}| = \cdots = |\zeta_{t+r}| > |\zeta_{t+r+1}| \geq \cdots
\]

(A.49)

for some \( t \geq 1 \) and \( r \geq 2 \), and let

\[ t + 1 \leq k < t + r. \]

(A.50)

1. When \( \zeta_{t+1}, \ldots, \zeta_{t+r} \) are not all the same, assume that

\[
R(n; \sigma) \approx \sum_{t+1 \leq j_1 < \cdots < j_k \leq t+r} W_{1,2, \ldots, t, j_1+1, \ldots, j_k} \left[ \prod_{p=t+1}^k (\sigma - \sigma_{j_p}) \right] \left( \prod_{p=t+1}^k \frac{\zeta_{j_p}}{\zeta_{p}} \right)^n \neq 0
\]

(A.51)

for some integer \( n \). Then there exist integers \( 0 \leq n_0 \leq n_1 \leq n_2 < \cdots \), for which \( \{ R(n_i; \sigma) \}_{i=0}^\infty \) has a limit. Also, with appropriate constants \( d_n \), the subsequence \( \{ d_n H_n(\sigma) \}_{i=0}^\infty \) converges to a polynomial in \( \sigma \) of degree \( t + q \leq k \), whose zeros are \( \sigma_1, \sigma_2, \ldots, \sigma_t \) and \( \sigma_1', \ldots, \sigma_q' \), the latter being the zeros of the limit of the subsequence of polynomials \( \{ R(n_i; \sigma) \}_{i=0}^\infty \). Actually, for this subsequence

\[
d_n H_n(\sigma) = \left[ \prod_{j=1}^t (\sigma - \sigma_j) \right] \left[ \prod_{j=1}^q (\sigma - \sigma_j') \right] + O(\eta^n) \quad \text{as} \quad n \to \infty
\]

(A.52)

with

\[
\eta_n = \max \left( \frac{\zeta_{t+1}}{\zeta_t}, \frac{\zeta_{t+r+1}}{\zeta_{t+1}} \right).
\]

(A.53)

2. When \( \zeta_{t+1} = \cdots = \zeta_{t+r} \), assume that

\[
T(\sigma) \approx \sum_{t+1 \leq j_1 < \cdots < j_k \leq t+r} W_{1,2, \ldots, t, j_1+1, \ldots, j_k} \left[ \prod_{p=t+1}^k (\sigma - \sigma_{j_p}) \right] \neq 0.
\]

(A.54)
Then, with the proper normalization constants $d_n$, the sequence $\{d_nH_n(\sigma)\}_{n=0}^{\infty}$ converges to a polynomial in $\sigma$ of degree $t + q \leq k$ whose zeros are $\sigma_1, \sigma_2, \ldots, \sigma_t$ and $\sigma'_1, \ldots, \sigma'_q$, the latter being the zeros of the polynomial $T(\sigma)$. This time the whole sequence $\{d_nH_n(\sigma)\}_{n=0}^{\infty}$ satisfies (A.52) and (A.53).

**Proof:** The proof can be accomplished as that of Lemma 2.4 in [S1]. We leave the details to the reader. $\Box$

**Lemma A.8:** Assume the conditions of Lemma A.7, and let $s \in \{1, 2, \ldots, t\}$. Then, in case $\sigma_s$ is simple,

$$\sigma_s(n) = \sigma_s + O\left(\left|\frac{\zeta_{t+1}}{\zeta_s}\right|^n\right) \quad \text{as } n \to \infty,$$

(A.55)

and, in case $\sigma_s$ is multiple,

$$\limsup_{n \to \infty} |\sigma_s(n) - \sigma_s|^n \leq \left|\frac{\zeta_{t+1}}{\zeta_s}\right|,$$

(A.56)

when $\zeta_{t+1} = \cdots = \zeta_{t+r}$. When $\zeta_{t+1}, \ldots, \zeta_{t+r}$ are not all the same, (A.55) and (A.56) hold with $n$ replaced by $n_i$, where $\{d_nH_n(\sigma)\}_{n=0}^{\infty}$ is the convergent subsequence of part (1) of Lemma A.7.

**Proof:** The proof is almost identical to that of Lemma A.5. The only difference between the two is that Lemma A.7 is employed in the proof of the present lemma, whereas in the proof of Lemma A.5 we employed Lemma A.3. We also need the fact that $|\zeta_k| = |\zeta_{k+1}| = |\zeta_{t+1}|$, as follows from (A.49) and (A.50). We leave the details to the reader. $\Box$

**Acknowledgement**

This research was supported by the Fund for the Promotion of Research at the Technion.
References


