DETERMINATION OF THE EXTREME VALUES OF THE SPECTRUM OF A BOUNDED SELF-ADJOINT OPERATOR

WILLIAM KARUSH

1. Introduction. Let $A$ be a bounded self-adjoint operator on a Hilbert space $\mathcal{H}$ with spectral family of projections $F(\lambda)$. Let $\lambda'_1$ be the supremum of the spectrum $\mathcal{S}(A)$ of $A$. We consider an iterative procedure for the determination of $\lambda'_1$ and, in case $\lambda'_1$ is a characteristic number, for the determination of a characteristic vector belonging to $\lambda'_1$. The procedure involves the solution of a characteristic value problem of finite dimension at each step of the iteration, this dimension being the same for each step and fixed at a convenient value at the outset.

The method is based upon the observation that with

$$
\mu(x) = \frac{(x, Ax)}{(x, x)}, \quad x \neq 0 \text{ in } \mathcal{H},
$$

we have

$$
\lambda'_1 = \sup_{x \neq 0} \mu(x), \quad x \text{ in } \mathcal{H}.
$$

The iteration is determined as follows. Select an integer $s > 1$ and an initial vector $x^0$ in $\mathcal{H}$. In the space $\mathcal{A}(x^0)$ spanned by the vectors $x^0, Ax^0, \ldots, A^{s-1}x^0$ determine a vector $x^1$ which maximizes $\mu(x)$ for $x$ in $\mathcal{A}(x^0)$; we shall show that $x^1$ is unique, apart from a scalar factor, and that it may be chosen in the form $x^1 = x^0 + \eta$, with $(x^0, \eta) = 0$. Clearly $\mu(x^0) \leq \mu(x^1)$. In the next step we similarly choose $x^2 = x^1 + \eta$ as the vector in the space $\mathcal{A}(x^0)$ spanned by $x^1, Ax^1, \ldots, A^{s-1}x^1$ which maximizes $\mu(x)$ in this space. In this way we construct a sequence $\{x^i\}$ with nondecreasing values $\mu(x^i)$. The determination of each $x^i$ involves solving for a characteristic vector of a self-adjoint matrix of order at most $s$.

Now suppose $x^0$ is such that $F(\lambda)x^0 \neq x^0$ for $\lambda < \lambda'_1$. Then we shall prove that the numbers $\mu(x^i)$ converge to $\lambda'_1$; and, further, that the unit vectors $u^i = x^i/|x^i|$ converge weakly to a characteristic vector...

Received by the editors December 22, 1950.

1 The preparation of this paper was sponsored (in part) by the Office of Naval Research.

2 An iterative procedure different from ours is given by $x^{i+1} = Ax^i$. For a discussion of this method see R. Wavre, L'itération directe des opérateurs hermitiens et deux théories qui en dépendent, Comment Math. Helv. vol. 15 (1942–1943) pp. 299–317.
belonging to \( \lambda_i' \) if \( \lambda_i' \) is a characteristic number, but that \( u^i \) converges weakly to 0 if \( \lambda_i' \) is in the continuous spectrum. If \( \lambda_i' \) is an isolated point of the spectrum \( S(A) \) (and thus necessarily a characteristic number), and if the initial vector \( x^0 \) has a non-null projection on the characteristic manifold belonging to \( \lambda_i' \), then the vectors \( x^i \) will converge (strongly) to a characteristic vector in this manifold. More generally, if no assumptions are made on \( x^0 \), then these convergence properties hold if \( R \) is interpreted as the invariant closed linear manifold determined by \( x^0, Ax^0, A^2x^0, \ldots \).

By minimizing instead of maximizing \( \mu \) at each step we obtain entirely analogous results for the infimum of the spectrum of \( A \). We shall limit our discussion to the maximizing procedure. Further we shall suppose that \( R \) is a real Hilbert space, not necessarily separable; there is no difficulty in extending the treatment to the complex case.

This paper is a generalization of an earlier one\(^8\) and some of the results obtained there will be used here.

2. An invariant subspace. We are dealing with a bounded self-adjoint operator \( A \) on a real Hilbert space \( R \), not necessarily separable, whose elements we call vectors. We let \( F(\lambda) \) (continuous on the right) denote the spectral family of \( A \) and \( S(A) \) the spectrum of \( A \). Thus

\[
A = \int_{-\infty}^{\infty} \lambda dF(\lambda) = \int_{\mathbb{R}} \lambda dF(\lambda).
\]

A characteristic number (that is, a member of the point spectrum of \( A \)) is a real number \( \lambda \) such that

\[
Ay = \lambda y
\]

for some \( y \neq 0 \) in \( R \); \( y \) is a characteristic vector of \( A \) belonging to \( \lambda \). The characteristic vectors belonging to \( \lambda \) determine a closed linear manifold, the characteristic manifold, which we designate by \( L(\lambda) \). This manifold is the projection manifold of \( F(\lambda+) - F(\lambda-) = F(\lambda) - F(\lambda-) \), where \( F(\lambda+) \) and \( F(\lambda-) \) denote, respectively, right- and left-hand limits.

We define \( \mu(x) \) as in (1), and as before let

\[
\lambda_i' = \sup S(A).
\]

We have

\[
\mu(x) \leq \lambda_i', \quad x \neq 0 \text{ in } R.
\]

The proof is like that of (4) below. (We shall not require the stronger result (2).) Suppose now that $\lambda'_I$ is a characteristic number. Let $B'$ be the set $S(A)$ with $\lambda'_I$ deleted, and put $\lambda'_I = \sup B'$. Then we show that

$$
(4) \quad \mu(x) \leq \lambda'_I, \quad x \neq 0 \text{ in } \mathbb{R} \text{ and orthogonal to } \mathcal{L}(\lambda'_I).
$$

For let $x$ be as described; we may assume $|x| = 1$. Then

$$
\mu(x) = (Ax, x) = \int_{S(A)} \lambda'_I |F(\lambda)x|^2 = \int_{B'} + \int_{\lambda'_I}.
$$

The second integral vanishes by the orthogonality of $x$; the first integral is dominated by $\lambda'_I |x|^2 = \lambda'_I$, as desired. We shall not require the sharper result $\lambda'_I = \sup \mu(x)$ (as in (4)).

In the main we shall deal with the closed linear manifold determined by an initial non-null vector $x^0$ and the powers $Ax^0, A^2x^0, \ldots$. We denote this manifold by $\mathcal{C}$; symbolically

$$
(5) \quad \mathcal{C} = (x^0, Ax^0, A^2x^0, \ldots).
$$

$\mathcal{C}$ is a Hilbert space which is invariant under $A$, that is, $A\mathcal{C} \subset \mathcal{C}$. We denote by $B$ the operator $A$ with domain restricted to $\mathcal{C}$, that is, $B$ is a bounded self-adjoint operator on $\mathcal{C}$ such that $Bx = Ax$ for $x$ in $\mathcal{C}$. It is not difficult to see that the spectral family $E(\lambda)$ of $B$ is obtained from the spectral family $F(\lambda)$ of $A$ by

$$
(6) \quad E(\lambda)x = F(\lambda)x, \quad x \in \mathcal{C};
$$

we omit the proof.

Clearly if a number is in the resolvent set of $A$ then it is in the resolvent set of $B$. Thus $S(B) \subset S(A)$. We let

$$
(7) \quad \lambda_1 = \sup S(B).
$$

The point spectrum of $B$ is obviously a subset of the point spectrum of $A$. Further, the characteristic numbers of $B$ are all simple. For let $\lambda$ be such a number and let $y_1, y_2$ be independent characteristic vectors of $B$ belonging to $\lambda$. Then some non-null linear combination $y$ of these vectors is orthogonal to $x^0$. From $By = \lambda y$ it follows readily that $y$ is orthogonal to all powers $A^jx^0 = B^jx^0, j = 0, 1, 2, \ldots$. Hence, by (5), $y$ is orthogonal to $\mathcal{C}$. But $y$ is in $\mathcal{C}$. Hence $y = 0$; contradiction. The next lemma shows how the characteristic vectors of $B$ are determined.

**Lemma 1.** Let $\lambda$ be a characteristic number of $A$ and $y(\lambda)$ be the projection of $x^0$ on the characteristic manifold $\mathcal{L}(\lambda)$ of $A$. Then $\lambda$ is a
characteristic number of $B$ if and only if $y(\lambda) \neq 0$. If $y(\lambda) \neq 0$, then this vector is the unique characteristic vector of $B$ (apart from a scalar factor) belonging to $\lambda$.

Suppose $\lambda$ is characteristic for $B$ with characteristic vector $y \neq 0$ in $\mathfrak{C}$. If $y(\lambda) = 0$, then $L(\lambda)$ is orthogonal to $x^0$ and hence, by (5), to $\mathfrak{C}$. But $y$ is in $L(\lambda)$. Hence $y = 0$; contradiction.

Now suppose $y(\lambda) \neq 0$. By (6) and the definition of $y(\lambda)$,

$$[E(\lambda) - E(\lambda-)]x^0 = [F(\lambda) - F(\lambda-)]x^0 = y(\lambda).$$

Thus $\lambda$ is a characteristic number of $B$ and $y(\lambda)$ is a characteristic vector. The uniqueness follows from the earlier remark that $\lambda$ is simple for $B$.

3. The iteration procedure. Consider a fixed integer $s > 1$ and a fixed vector $x \neq 0$ in $\mathfrak{R}$. Define $\mathfrak{A}(x)$ as the finite-dimensional space

$$\mathfrak{A}(x) = (x, Ax, \ldots, A^{s-1}x),$$

that is, the space spanned by the vectors indicated on the right. Suppose that the dimension of $\mathfrak{A}(x)$ is $s$. The following statements in this paragraph were established in §§2 and 3 of the previously cited paper by the author. The vectors $\xi_0, \xi_1, \ldots, \xi_{s-1}$ defined recursively by

$$\xi_0 = x, \quad \xi_1 = A\xi_0 - \mu_0\xi_0, \quad \mu_0 = \mu(x),$$

$$(8) \quad \xi_{j+1} = A\xi_j - \mu_j\xi_j - t_j^2\xi_{j-1}, \quad \mu_j = \mu(\xi_j), \quad t_j = \frac{\|\xi_j\|}{\|\xi_{j-1}\|}, \quad j = 1, 2, \ldots, s - 1,$$

form an orthogonal basis for $\mathfrak{A}(x)$. If the polynomials $p_j(\lambda)$ are defined by

$$p_0(\lambda) = 1, \quad p_1(\lambda) = \lambda - \mu_0,$$

$$p_{j+1}(\lambda) = p_j(\lambda)(\lambda - \mu_j) - t_j^2p_{j-1}(\lambda), \quad j = 1, 2, \ldots, s - 1,$$

then

$$(9) \quad \xi_j = p_j(A)x.$$  

The roots of $p_j(\lambda)$ are simple and real; let $\nu_j$ denote the maximum root. Then $\nu_j$ is the maximum of $\mu(x)$ for $x$ in $(x, Ax, \ldots, A^{s-1}x)$. In particular, letting $\nu = \nu_n$,

$$\nu = \max \mu(x), \quad x \neq 0 \text{ in } \mathfrak{A}(x),$$

$$\nu_1 \leq \nu_2 \leq \ldots \leq \nu_n = \nu.$$
Also

(11) \( p_i(\lambda') p_j(\lambda) \geq 0 \), \( |p_i(\lambda')| \geq |p_j(\lambda)| \), \( \lambda' \leq \lambda \leq \nu_j \).

Finally, there is a unique vector \( x^* \) in \( \mathcal{A}(x) \) of the form \( x + \eta \) with \( (x, \eta) = 0 \) for which

(12) \( \mu(x^*) = \nu \).

It is given by

(13) \[ x^* = x + \sum_{i=1}^{r-1} \frac{p_j(v)}{\tau_j} \xi_i \]

where

(14) \[ \tau_j = \frac{t_1 t_2 \cdots t_j}{|x|} \]

Consider now the proposed iteration scheme. We construct a sequence of vectors \( x^0, x^1, x^2, \cdots \) by choosing \( x^{i+1} \) as that vector in \( \mathcal{A}(x^i) \) of the form \( x^i + \eta \), \( (x^i, \eta) = 0 \), which maximizes \( \mu(z) \), \( z \in \mathcal{A}(x^i) \).

Assume for the moment that for each \( i \), \( \mathcal{A}(x^i) \) has dimension \( s \). By (13) we have the explicit formula

(15) \[ x^{i+1} = x^i + \sum_{j=1}^{r-1} \frac{p_j(v)}{(\tau_j)^2} \xi_j, \quad i = 0, 1, 2, \cdots, \]

where the superscript "\(^{*}\)" on the right has the obvious interpretation.

It is clear that all vectors arising in this construction lie in the invariant space \( \mathcal{X} \) of (5).

By (9) we have the alternative formula

(16) \[ x^{i+1} = \left[ I + \sum_{j=1}^{r-1} \frac{p_j(v)}{(\tau_j)^2} p_j(B) \right] x^i, \]

where \( I \) is the identity operator. By the extremum property of \( x^{i+1} \) in \( \mathcal{A}(x^i) \),

(17) \( \mu^{i+1} \geq \mu^i \), \quad where \( \mu^i = \mu(x^i) \).

From (15), (14) and the orthogonality of the \( \xi_j \) for each \( i \),

(18) \[ |x^{i+1}|^2 = |x^i|^2 \left\{ 1 + \sum_{j=1}^{r-1} \left[ \frac{p_j(v)}{\tau_j} \right]^2 \right\}. \]

We consider now the case when for some (first) value \( k \), \( \mathcal{A}(x^k) \) has
dimension less than \( s \). Then \( A(x^k) \) is invariant under \( A \), and the
vector \( x^{k+1} \) maximizing \( \mu(x) \) in this subspace is a characteristic vector
of \( A \). (This vector is given by (15) with "s" replaced by the dimension
of the subspace.) Since \( A(x^{k+1}) = (x^{k+1}) \), it follows that the sequence
\( \{ x^i \} \) has the constant value \( x^{k+1} \) for \( i \geq k+1 \). The iteration now becomes trivial, and the forthcoming proofs can be readily simplified
to apply to this case. To save space we shall henceforth assume that for each \( i \) the dimension of \( A(x^i) \) is \( s \), and leave to the reader
the appropriate modifications for the other case.

4. Convergence theorem. We establish the following convergence theorem.

**Theorem 1.** For a given integer \( s > 1 \) and initial vector \( x^0 \neq 0 \), let
\( \{ x^i \} \) be the sequence determined above. Let
\[
\mu^i = \frac{x^i}{|x^i|}.
\]
Then
\[
\lim_{i \to \infty} \mu(x^i) = \lambda_1,
\]
where \( \lambda_1 \) is given by (7). Further, if \( \lambda_1 \) is a characteristic number of \( B \),
then \( u^i \) converges weakly (in \( \mathbb{R} \)) to a characteristic vector of \( B \) belonging
to \( \lambda_1 \); if \( \lambda_1 \) is not a characteristic number of \( B \), then \( u^i \) converges weakly
to 0.

**Proof.** By (3), interpreted for \( \mathcal{C} \), \( \mu(x) \leq \lambda_1 \) for \( x \neq 0 \) in \( \mathcal{C} \). Thus,
from (17), the numbers \( \mu^i \) have a limit \( \mu \leq \lambda_1 \). By definition \( p_2'(\lambda) \)
\[
= (\lambda - \mu^2)(\lambda - \mu(\xi^i)) - |\xi^i|^2/|x^i|^2,
\]
where
\[
\xi^i = Bx^i - \mu^i x^i.
\]
From (10) and (11) we have \( p_2'(\nu^i) \geq 0 \). Since \( \mu^{i+1} = \nu^i \) by (12), we
conclude that
\[
\frac{|\xi^i|^2}{|x^i|^2} \leq (\mu^{i+1} - \mu^i)(\mu^{i+1} - \mu(\xi^i)).
\]
Since the second factor on the right is bounded, the sequence on the
left tends to 0; in fact
\[
\sum (\nu^i)^2 < \infty, \quad \nu^i = |\xi^i|/|x^i|.
\]
From (19),
\[
\lim_{i \to \infty} (Bu^i - \mu^i u^i) = 0.
\]
Suppose \( \mu < \lambda_1 \). Choose \( \hat{\lambda} \) so that \( \mu < \hat{\lambda} < \lambda_1 \). Let \( \Delta = I - E(\hat{\lambda} - \cdot) \).
We note first that \( \Delta x^0 \neq 0 \), i.e., \( \Delta u^0 \neq 0 \). For, suppose \( \Delta x^0 = 0 \). Then from \( \Delta B = B \Delta \) would follow \( \Delta x = 0 \) for \( x = B x^0 \), \( k = 0, 1, 2, \ldots \), and thus \( \Delta x = 0 \) for \( x \) in \( \mathfrak{C} \), by (5). Hence \( E(\hat{\lambda} - \cdot) = I \), so that \( \sup \mathfrak{S} (B) \leq \hat{\lambda} \), that is, \( \lambda_1 \leq \hat{\lambda} \), a contradiction.

We put

\[
R_i = \frac{\hat{\rho}_i(v^i)}{(v^i)^2}, \quad i = 1, 2, \ldots, s - 1, i = 0, 1, 2, \ldots.
\]

By (16)

\[
x^{i+1} = \int_{-\alpha}^{\lambda_1} \left[ 1 + R_1^i \hat{\rho}_1(\lambda) + \cdots + R_{s-1}^i \hat{\rho}_{s-1}(\lambda) \right] dE(\lambda) x^i.
\]

Hence

\[
|\Delta x^{i+1}|^2 = \int_{\lambda}^{\lambda_1} \left[ 1 + R_1^i \hat{\rho}_1(\lambda) + \cdots + R_{s-1}^i \hat{\rho}_{s-1}(\lambda) \right]^2 d|E(\lambda) x^i|^2.
\]

By (11), the quantity in brackets is an increasing function of \( \lambda \) in the range of integration. Thus

\[
|\Delta x^{i+1}|^2 \geq \left[ 1 + R_1^i \hat{\rho}_1(\lambda) + \cdots + R_{s-1}^i \hat{\rho}_{s-1}(\lambda) \right] |\Delta x^i|^2.
\]

From (18)

\[
\frac{|x^{i+1}|^2}{|x^i|^2} = 1 + R_1^i \hat{\rho}_1(v^i) + \cdots + R_{s-1}^i \hat{\rho}_{s-1}(v^i),
\]

so that, by (11),

\[
|\Delta u^{i+1}|^2 \geq \frac{1 + R_1^i \hat{\rho}_1(\lambda) + \cdots + R_{s-1}^i \hat{\rho}_{s-1}(\lambda)}{1 + R_1^i \hat{\rho}_1(v^i) + \cdots + R_{s-1}^i \hat{\rho}_{s-1}(v^i)} |\Delta u^i|^2 \geq |\Delta u^i|^2.
\]

On the other hand

\[
|Bu^i - \mu_i u^i|^2 = \int_{-\alpha}^{\lambda_1} (\lambda - \mu_i)^2 d|E(\lambda) u^i|^2
\]

\[
\geq \int_{\lambda}^{\lambda_1} (\lambda - \mu_i)^2 d|E(\lambda) u^i|^2 \geq (\hat{\lambda} - \mu_i)^2 |\Delta u^i|^2.
\]

By (21) it follows that \( |\Delta u^i| \to 0 \), a contradiction to (23) and the earlier result \( |\Delta u^0| \neq 0 \). This establishes the first conclusion of the theorem.
Suppose now that \( \lambda_1 \) is not a characteristic number of \( B \). If \( u^i \) does not converge weakly to 0 (in \( \mathcal{C} \)), then a subsequence \( u^k \) in \( \mathcal{C} \) must have a weak limit \( \bar{u} \neq 0 \) (in \( \mathcal{C} \)), since \( |u^i| = 1 \). Thus \( Bu^k - \mu^k u^k \) converges weakly to \( B\bar{u} - \lambda_1 \bar{u} \). But this sequence converges (strongly) to zero by (21). Hence \( B\bar{u} - \lambda_1 \bar{u} = 0 \), contrary to \( \lambda_1 \) not being characteristic. Hence \( u^i \) converges weakly to 0 in \( \mathcal{C} \), and hence in the original space \( \mathcal{R} \).

Finally, suppose \( \lambda_1 \) is characteristic for \( B \). Let \( y_1 \) be a characteristic vector of \( B \) belonging to \( \lambda_1 \). By Lemma 1, \( (x^0, y_1) \neq 0 \) and \( \lambda_1 \) is simple. We normalize \( y_1 \) so that \( |y_1| = 1 \), \( (x^0, y_1) > 0 \). Any solution of \( Bz = \lambda_1 z \) in \( \mathcal{C} \) is a multiple of \( y_1 \). By (16) and (22),

\[
(x^{i+1}, y_1) = \left[ 1 + R^i_1 p_1(\lambda_1) + \cdots + R^i_{n-1} p_{n-1}(\lambda_1) \right] (x^i, y_1),
\]

\[
(u^{i+1}, y_1) \geq \frac{1}{1 + R^i_1 p_1(\lambda_1) + \cdots + R^i_{n-1} p_{n-1}(\lambda_1)} (u^i, y_1) \geq (u^i, y_1),
\]

using (11). Hence \( (u^i, y_1) \to L > 0 \). Now suppose \( \bar{u} \) is any weak limit (in \( \mathcal{C} \)) of a subsequence \( u^k \) of \( u^i \). As in the above paragraph, \( B\bar{u} = \lambda_1 \bar{u} \). Hence \( \bar{u} = Ly_1 \). But \( L = \lim_{k \to \infty} (u^k, y_1) = (y_1, y_1) = l \). Hence \( l = L, \bar{u} = Ly_1 \), independently of \( \bar{u} \). This establishes that \( Ly_1 \) is the weak limit in \( \mathcal{C} \), and thus in \( \mathcal{R} \), of \( u^i \), and completes the proof of the theorem.

In the following corollary we relate the preceding result directly to the original operator \( A \).

**Corollary to Theorem 1.** If \( F(\lambda)x^0 \neq x^0 \) for \( \lambda < \lambda_1' \), where \( \lambda_1' = \sup S(A) \), then \( \mu^i \) converges to \( \lambda_1' \). If \( \lambda_1' \) is characteristic for \( A \) and \( x^0 \) has a non-null projection on the characteristic manifold \( \mathcal{L}_p \) of \( A \) belonging to \( \lambda_1' \), then \( u^i \) converges weakly to a characteristic vector in \( \mathcal{L}_p \).

**Proof.** Consider the first statement of the corollary. We have \( \lambda_1 \leq \lambda_1' \). Suppose \( \lambda_1 < \lambda_1' \). Then by (6)

\[
x_0 = \int_{-\infty}^{\lambda_1} dE(\lambda) x^0 = \int_{-\infty}^{\lambda_1} dF(\lambda) x^0 = F(\lambda_1) x^0,
\]

contrary to hypothesis. Hence \( \lambda_1 = \lambda_1' \), and the conclusion follows from Theorem 1. The second statement is a consequence of Lemma 1 and Theorem 1.

5. **Convergence theorem for \( \lambda_1 \) isolated.**

**Lemma 2.** Let \( \lambda_1 \) be an isolated point of \( S(B) \). Then the unit vectors \( u^i \) converge to the characteristic vector of \( B \) belonging to \( \lambda_1 \).
Proof. Since \( \lambda_1 \) is isolated, it is a characteristic number for \( B \). Let \( y_1 \) be the corresponding characteristic vector, normalized so that \( |y_1| = 1 \) and \((x^0, y_1) > 0\). In the last paragraph of the proof of Theorem 1 we showed that \( u^i \) converged weakly to \( Ly_1 \), where \( L = \lim_{i \to \infty} (u^i, y_1) \), \( i \geq 0 \). We now show that \( u^i \) converges to \( y_1 \).

Write \( u^i = y^i + z^i \) with \( y^i \) a multiple of \( y_1 \) and \( z^i \) in \( \mathcal{C} \) orthogonal to \( y_1 \). Let

\[
\lambda_2 = \sup \mathcal{B},
\]

where \( \mathcal{B} = \mathbb{S}(B) \) with \( \lambda_1 \) deleted. Then \( \lambda_2 < \lambda_1 \), and by (4), interpreted for \( \mathcal{C} \), we have \( \mu(z^i) \leq \lambda_2 \). Now \((y^i, y_1) = (u^i, y_1) - (Ly_1, y_1) = L \). Hence \( y^i = (y^i, y_1) y_1 \) converges to \( Ly_1 \).

Using the definition (1) of \( \mu \) we find

\[
\mu^i = \mu(u^i) = (Bu^i, u^i) = \mu(y^i) |y^i|^2 + \mu(z^i) |z^i|^2
\]

\[
= \lambda_1 (1 - |z^i|^2) + \mu(z^i) |z^i|^2.
\]

Thus

\[
\lambda_1 - \mu^i = (\lambda_1 - \mu(z^i)) |z^i|^2 \geq (\lambda_1 - \lambda_2) |z^i|^2.
\]

Since \( \mu \to \lambda_1 \) by Theorem 1, it follows that \( z^i \) converges to 0. From \( u^i = y^i + z^i \) we now deduce that \( u^i \) converges to \( Ly_1 \). Since \( |u^i| = |y_1| = 1 \), we must have \( L = 1 \). This completes the proof.

Our goal is to replace "\( u^i \)" by "\( x^i \)" in the above lemma. For this it is clearly sufficient to show that the increasing lengths \( |x^i| \) (see (18)) are bounded. To this end we introduce the next lemma. (As stated at the end of §3, we are assuming that for each \( i \) the vectors \( x^i, Ax^i, \ldots, A^{s-1}x^i \) are independent. As shown there, if this is not the case then the sequence \( \{x^i\} \) is eventually constant, and obviously the lengths \( |x^i| \) converge.)

**Lemma 3.** Let \( \lambda_1 \) be an isolated point of \( \mathbb{S}(B) \). There is a constant \( K \), independent of \( i \) and \( j \), such that for \( i \) sufficiently large,

\[
|B^j(x^i)| \leq K(\tau_i)^j, \quad j = 1, 2, \ldots, s - 1.
\]

We shall not give the details of the proof; they can be found in the proof of a similar result in the previously cited paper by the author. One first establishes, as in Lemma 1 of the earlier paper, that

\[
\lambda_1 - \mu(x) \leq \frac{1}{\mu(x) - \lambda_2} |Bx - \mu(x)x|^2
\]

for every \( x \) in \( \mathcal{C} \) with \( \mu(x) > \lambda_2 \), \( \lambda_2 \) as in (24). The lemma is then established by an argument like that of Lemma 2 of the earlier paper.
Theorem 2. Let \( \lambda_1 \) be an isolated point of \( \mathcal{S}(B) \). Then the vectors \( x^i \) of Theorem 1 converge to the characteristic vector of \( B \) belonging to \( \lambda_1 \).

Proof. We use (18). By a standard theorem on infinite products the numbers \( |x^i|^2 \) will converge if each of the series \( \sum_{j=1}^{s-1} |p^j(x^i)/\tau_j|^2 \), \( j = 1, 2, \cdots, s-1 \), converges. By Lemma 3 this will occur if each of the series \( \sum_{j=0}^{\infty} (\tau_j)^2 \) converges. By (20), this series converges for \( j = 1 \). By (8),

\[
|x_{i+1}^i| \leq K_1 |x_i^i| + (t_i^i)^2 |x_{i-1}^i|.
\]

Hence

\[
t_{i+1}^i \leq K_1 + t_i^i.
\]

Since \( t_i^i \) is bounded, it follows that there is a constant \( K_2 \) such that \( t_{i+1}^i \leq K_2 \) for all \( i \) and for \( j = 1, 2, \cdots, s-2 \). Hence by (14),

\[
\sum_i (t_{i+1}^i)^2 = \sum_i (t_j^i)^2 (t_{i+1}^i)^2 \leq K_2 \sum_i (\tau_j^i)^2.
\]

This establishes the convergence of all the series and completes the proof.

By Lemma 1 and Theorem 2 we obtain the following result.

Corollary to Theorem 2. Let \( \lambda_1' = \sup \mathcal{S}(A) \) be an isolated point of \( \mathcal{S}(A) \). If \( x^0 \) is not orthogonal to the characteristic manifold of \( A \) belonging to \( \lambda_1' \), then \( x^i \) converges to a characteristic vector in this manifold.