DETERMINATION OF THE EXTREME VALUES OF THE SPECTRUM OF A BOUNDED SELF-ADJOINT OPERATOR[1]

WILLIAM KARUSH

1. **Introduction.** Let A be a bounded self-adjoint operator on a Hilbert space $\mathfrak{R}$ with spectral family of projections $F(\lambda)$. Let λ_1' be the supremum of the spectrum $\mathfrak{S}(A)$ of A. We consider an iterative procedure for the determination of λ_1' and, in case λ_1' is a characteristic number, for the determination of a characteristic vector belonging to λ_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

$$\mu(x) = \frac{(x, Ax)}{(x, x)}, \qquad x \neq 0 \text{ in } \mathfrak{R}, \tag{1}$$

we have

$$\lambda_1' = \sup_{x \neq 0} \mu(x), \qquad x \text{ in } \mathfrak{R}. \tag{2}$$

The iteration is determined as follows. Select an integer $s > 1$ and an initial vector x^0 in $\mathfrak{R}$. In the space $\mathcal{A}(x^0)$ spanned by the vectors $x^0, Ax^0, \cdots, A^{s-1}x^0$ determine a vector x^1 which maximizes $\mu(z)$ 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 $\mu(x^0) \leqq \mu(x^1)$. In the next step we similarly choose $x^2 = x^1 + \eta$ as the vector in the space $\mathcal{A}(x^1)$ spanned by $x^1, Ax^1, \cdots, A^{s-1}x^1$ 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.[2]

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 λ_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 λ_1' if λ_1' is a characteristic number, but that u^i converges weakly to 0 if λ_1' is in the continuous spectrum. If λ_1' is an isolated point of the spectrum $\mathfrak{S}(A)$ (and thus necessarily a characteristic number), and if the initial vector x^0 has a non-null projection on the characteristic manifold belonging to λ_1', 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 $\mathfrak{R}$ is interpreted as the invariant closed linear manifold determined by $x^0, Ax^0, A^2x^0, \cdots$.

By minimizing instead of maximizing μ 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 $\mathfrak{R}$ 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[3] 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 $\mathfrak{R}$, not necessarily separable, whose elements we call vectors. We let $F(\lambda)$ (continuous on the right) denote the spectral family of A and $\mathfrak{S}(A)$ the spectrum of A. Thus

$$A = \int_{-\infty}^{\infty} \lambda dF(\lambda) = \int_{\mathfrak{S}(A)} \lambda dF(\lambda).$$

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

$$Ay = \lambda y$$

for some $y \neq 0$ in $\mathfrak{R}$; y is a characteristic vector of A belonging to λ. The characteristic vectors belonging to λ determine a closed linear manifold, the characteristic manifold, which we designate by $\mathfrak{L}(\lambda)$. 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

$$\lambda_1' = \sup \mathfrak{S}(A).$$

We have

$$\mu(x) \leqq \lambda_1', \qquad x \neq 0 \text{ in } \mathfrak{R}. \tag{3}$$

[3] W. Karush, *An iterative method for finding characteristic vectors of a symmetric matrix,* Pacific Journal of Mathematics vol. 1 (1951) pp. 233–248.

The proof is like that of (4) below. (We shall not require the stronger result (2).) Suppose now that λ_1' is a characteristic number. Let $\mathcal{B}'$ be the set $\mathcal{S}(A)$ with λ_1' deleted, and put $\lambda_2' = \sup \mathcal{B}'$. Then we show that

$$(4) \qquad \mu(x) \leqq \lambda_2', \qquad x \neq 0 \text{ in } \mathfrak{R} \text{ and orthogonal to } \mathcal{L}(\lambda_1').$$

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

$$\mu(x) = (Ax, x) = \int_{\mathcal{S}(A)} \lambda d \,|F(\lambda)x|^2 = \int_{\mathcal{B}'} + \int_{\lambda_1'}.$$

The second integral vanishes by the orthogonality of x; the first integral is dominated by $\lambda_2' |x|^2 = \lambda_2'$, as desired. We shall not require the sharper result $\lambda_2' = \sup \mu(x)$ (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, $A^2x^0, \cdots$. We denote this manifold by $\mathcal{K}$; symbolically

$$(5) \qquad \mathcal{K} = (x^0, Ax^0, A^2x^0, \cdots).$$

$\mathcal{K}$ is a Hilbert space which is invariant under A, that is, $A\mathcal{K} \subset \mathcal{K}$. We denote by B the operator A with domain restricted to $\mathcal{K}$, that is, B is a bounded self-adjoint operator on $\mathcal{K}$ such that $Bx = Ax$ for x in $\mathcal{K}$. 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

$$(6) \qquad E(\lambda)x = F(\lambda)x, \qquad x \text{ in } \mathcal{K};$$

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 $\mathcal{S}(B) \subset \mathcal{S}(A)$. We let

$$(7) \qquad \lambda_1 = \sup \mathcal{S}(B).$$

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 λ be such a number and let y_1, y_2 be independent characteristic vectors of B belonging to λ. 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 $A^j x^0 = B^j x^0$, $j = 0, 1, 2, \cdots$. Hence, by (5), y is orthogonal to $\mathcal{K}$. But y is in $\mathcal{K}$. Hence $y = 0$; contradiction. The next lemma shows how the characteristic vectors of B are determined.

LEMMA 1. *Let λ 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 λ 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 λ.*

Suppose λ is characteristic for B with characteristic vector $y\neq 0$ in $\mathcal{H}$. If $y(\lambda)=0$, then $\mathcal{L}(\lambda)$ is orthogonal to x^0 and hence, by (5), to $\mathcal{H}$. But y is in $\mathcal{L}(\lambda)$. 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 λ is a characteristic number of B and $y(\lambda)$ is a characteristic vector. The uniqueness follows from the earlier remark that λ is simple for B.

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

$$\mathcal{A}(x) = (x, Ax, \cdots, A^{s-1}x),$$

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

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

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

$$j = 1, 2, \cdots, s-1,$$

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

$$p_0(\lambda) = 1, \qquad p_1(\lambda) = \lambda - \mu_0,$$
$$p_{j+1}(\lambda) = p_j(\lambda)(\lambda - \mu_j) - t_j^2 p_{j-1}(\lambda), \qquad j = 1, 2, \cdots, s-1,$$

then

$$\xi_j = p_j(A)x. \tag{9}$$

The roots of $p_j(\lambda)$ are simple and real; let ν_j denote the maximum root. Then ν_j is the maximum of $\mu(x)$ for x in $(xAx, \cdots, A^{j-1}x)$. In particular, letting $\nu=\nu_s$,

$$\begin{gathered} \nu = \max \mu(z), \qquad z \neq 0 \text{ in } \mathcal{A}(x), \\ \nu_1 \leqq \nu_2 \leqq \cdots \leqq \nu_s = \nu. \end{gathered} \tag{10}$$

Also

$$(11)\qquad p_j(\lambda')p_j(\lambda) \geqq 0, \qquad |p_j(\lambda')| \geqq |p_j(\lambda)|, \qquad \lambda' \geqq \lambda \geqq \nu_j.$$

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

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

It is given by

$$(13)\qquad x^* = x + \sum_{j=1}^{s-1} \frac{p_j(\nu)}{\tau_j^2}\xi_j$$

where

$$(14)\qquad \tau_j = t_1 t_2 \cdots t_j = \frac{|\xi_j|}{|x|}\cdot$$

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

$$(15)\qquad x^{i+1} = x^i + \sum_{j=1}^{s-1} \frac{\overset{i}{p}_j(\overset{i}{\nu})}{(\tau_j^i)^2}\overset{i}{\xi}_j, \qquad i = 0, 1, 2, \cdots,$$

where the superscript "i" on the right has the obvious interpretation. It is clear that all vectors arising in this construction lie in the invariant space $\mathcal{H}$ of (5).

By (9) we have the alternative formula

$$(16)\qquad x^{i+1} = \left[I + \sum_{j=1}^{s-1} \frac{\overset{i}{p}_j(\overset{i}{\nu})}{(\tau_j^i)^2}\overset{i}{p}_j(B)\right]x^i,$$

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

$$(17)\qquad \mu^{i+1} \geqq \mu^i, \qquad \text{where } \mu^i = \mu(x^i).$$

From (15), (14) and the orthogonality of the ξ_j^i for each i,

$$(18)\qquad |x^{i+1}|^2 = |x^i|^2\left\{1 + \sum_{j=1}^{s-1}\left[\frac{\overset{i}{p}_j(\overset{i}{\nu})}{\tau_j^i}\right]^2\right\}.$$

We consider now the case when for some (first) value k, $\mathcal{A}(x^k)$ has dimension less than s. Then $\mathcal{A}(x^k)$ is invariant under A, and the vector x^{k+1} maximizing $\mu(z)$ 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 $\mathcal{A}(x^{k+1})=(x^{k+1})$, it follows that the sequence $\{x^i\}$ has the constant value x^{k+1} for $i\geqq k+1$. 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 $\mathcal{A}(x^i)$ 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*

$$u^i = \frac{x^i}{|x^i|}\cdot$$

Then

$$\lim_{i\to\infty}\mu(x^i) = \lambda_1,$$

where λ_1 is given by (7). *Further, if λ_1 is a characteristic number of B, then u^i converges weakly* (*in* $\mathfrak{R}$) *to a characteristic vector of B belonging to λ_1; if λ_1 is not a characteristic number of B, then u^i converges weakly to* 0.

PROOF. By (3), interpreted for $\mathcal{H}$, $\mu(x)\leqq\lambda_1$ for $x\neq 0$ in $\mathcal{H}$. Thus, from (17), the numbers μ^i have a limit $\bar{\mu}\leqq\lambda_1$. By definition $p_2^i(\lambda)=(\lambda-\mu^i)(\lambda-\mu(\xi^i))-|\xi^i|^2/|x^i|^2$, where

$$\xi^i = Bx^i - \mu^i x^i. \tag{19}$$

From (10) and (11) we have $p_2^i(\nu^i)\geqq 0$. Since $\mu^{i+1}=\nu^i$ by (12), we conclude that

$$\frac{|\xi^i|^2}{|x^i|^2} \leqq (\mu^{i+1}-\mu^i)(\mu^{i+1}-\mu(\xi^i)).$$

Since the second factor on the right is bounded, the sequence on the left tends to 0; in fact

$$\sum (t^i)^2 < \infty, \qquad t^i = |\xi^i|/|x^i|\cdot \tag{20}$$

From (19),

$$\lim_{i\to\infty}(Bu^i - \mu^i u^i) = 0. \tag{21}$$

Suppose $\bar{\mu}<\lambda_1$. Choose $\bar{\lambda}$ so that $\bar{\mu}<\bar{\lambda}<\lambda_1$. Let $\Delta=I-E(\bar{\lambda}-)$. 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 $x=B^k x^0$, $k=0, 1, 2, \cdots$, and thus $\Delta x=0$ for x in $\mathcal{H}$, by (5). Hence $E(\bar{\lambda}-)=I$, so that $\sup \mathfrak{S}(B)\leqq\bar{\lambda}$, that is, $\lambda_1\leqq\bar{\lambda}$, a contradiction.

We put

$$R_j^{i} = \frac{p_j^{i}(\nu^{i})}{(\tau_j^{i})^2}, \qquad i = 1, 2, \cdots, s-1,\ i = 0, 1, 2, \cdots.$$

By (16)

$$x^{i+1} = \int_{-\infty}^{\lambda_1} [1 + R_1^{i} p_1^{i}(\lambda) + \cdots + R_{s-1}^{i} p_{s-1}^{i}(\lambda)] dE(\lambda) x^{i}.$$

Hence

$$|\Delta x^{i+1}|^2 = \int_{\bar{\lambda}}^{\lambda_1} [1 + R_1^{i} p_1^{i}(\lambda) + \cdots + R_{s-1}^{i} p_{s-1}^{i}(\lambda)]^2 d|E(\lambda) x^{i}|^2.$$

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

$$|\Delta x^{i+1}|^2 \geqq [1 + R_1^{i} p_1^{i}(\bar{\lambda}) + \cdots + R_{s-1}^{i} p_{s-1}^{i}(\bar{\lambda})]^2 |\Delta x^{i}|^2.$$

From (18)

$$\frac{|x^{i+1}|^2}{|x^{i}|^2} = 1 + R_1^{i} p_1^{i}(\nu^{i}) + \cdots + R_{s-1}^{i} p_{s-1}^{i}(\nu^{i}), \tag{22}$$

so that, by (11),

$$|\Delta u^{i+1}|^2 \geqq \frac{1 + R_1^{i} p_1^{i}(\bar{\lambda}) + \cdots + R_{s-1}^{i} p_{s-1}^{i}(\bar{\lambda})}{1 + R_1^{i} p_1^{i}(\nu^{i}) + \cdots + R_{s-1}^{i} p_{s-1}^{i}(\nu^{i})} |\Delta u^{i}|^2 \geqq |\Delta u^{i}|^2. \tag{23}$$

On the other hand

$$|Bu^{i} - \mu^{i} u^{i}|^2 = \int_{-\infty}^{\lambda_1} (\lambda - \mu^{i})^2 d|E(\lambda) u^{i}|^2$$

$$\geqq \int_{\bar{\lambda}}^{\lambda_1} (\lambda - \mu^{i})^2 d|E(\lambda) u^{i}|^2 \geqq (\bar{\lambda} - \bar{\mu})^2 |\Delta u^{i}|^2.$$

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 λ_1 is not a characteristic number of B. If u^i does not converge weakly to 0 (in $\mathcal{H}$), then a subsequence u^k in $\mathcal{H}$ must have a weak limit $\bar{u}\neq 0$ (in $\mathcal{H}$), 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 λ_1 not being characteristic. Hence u^i converges weakly to 0 in $\mathcal{H}$, and hence in the original space $\mathfrak{R}$.

Finally, suppose λ_1 is characteristic for B. Let y_1 be a characteristic vector of B belonging to λ_1. By Lemma 1, $(x^0, y_1)\neq 0$ and λ_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 $\mathcal{H}$ is a multiple of y_1. By (16) and (22),

$$(\overset{i+1}{x}, y_1) = [1 + \overset{i}{R}_1\overset{i}{p}_1(\lambda_1) + \cdots + \overset{i}{R}_{s-1}\overset{i}{p}_{s-1}(\lambda_1)](\overset{i}{x}, y_1),$$

$$(\overset{i+1}{u}, y_1) \geqq \frac{1 + \overset{i}{R}_1\overset{i}{p}_1(\lambda_1) + \cdots + \overset{i}{R}_{s-1}\overset{i}{p}_{s-1}(\lambda_1)}{1 + R_1^i p_1^i(\nu^i) + \cdots + R_{s-1}^i p_{s-1}^i(\nu^i)}(\overset{i}{u}, y_1) \geqq (\overset{i}{u}, y_1),$$

using (11). Hence $(u^i, y_1)\to L>0$. Now suppose $\bar{u}$ is any weak limit (in $\mathcal{H}$) 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 $L=\lim_{k\to\infty}(u^k, y_1)=(ly_1, y_1)=l$. Hence $l=L$, $\bar{u}=Ly_1$, independently of $\bar{u}$. This establishes that Ly_1 is the weak limit in $\mathcal{H}$, and thus in $\mathfrak{R}$, 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 $\lambda_1'=\sup \mathfrak{S}(A)$, then μ^i converges to λ_1'. If λ_1' is characteristic for A and x^0 has a non-null projection on the characteristic manifold $\mathcal{L}$ of A belonging to λ_1', then u^i converges weakly to a characteristic vector in $\mathcal{L}$.*

Proof. Consider the first statement of the corollary. We have $\lambda_1\leqq\lambda_1'$. 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 λ_1 isolated.

Lemma 2. *Let λ_1 be an isolated point of $\mathfrak{S}(B)$. Then the unit vectors u^i converge to the characteristic vector of B belonging to λ_1.*

PROOF. Since λ_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 $L = \lim_{i\to\infty} (u^i, y_1) > 0$. 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 $\mathfrak{H}$ orthogonal to y_1. Let

$$\lambda_2 = \sup \mathfrak{B}, \tag{24}$$

where $\mathfrak{B}$ is $\mathfrak{S}(B)$ with λ_1 deleted. Then $\lambda_2 < \lambda_1$, and by (4), interpreted for $\mathfrak{H}$, we have $\mu(z^i) \leqq \lambda_2$. Now $(y^i, y_1) = (u^i, y_1) \to (Ly_1, y_1) = L$. Hence $y^i = (y^i, y_1)y_1$ converges to Ly_1.

Using the definition (1) of μ we find

$$\begin{aligned} \mu^i = \mu(u^i) &= (Bu^i, u^i) = \mu(y^i) \,|\, y^i \,|^2 + \mu(z^i) \,|\, z^i \,|^2 \\ &= \lambda_1(1 - |\, z^i \,|^2) + \mu(z^i) \,|\, z^i \,|^2. \end{aligned}$$

Thus

$$\lambda_1 - \mu^i = (\lambda_1 - \mu(z^i)) \,|\, z^i \,|^2 \geqq (\lambda_1 - \lambda_2) \,|\, z^i \,|^2.$$

Since $\mu^i \to \lambda_1$ 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 $x^i, Ax^i, \cdots, A^{s-1}x^i$ 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 λ_1 be an isolated point of $\mathfrak{S}(B)$. There is a constant K, independent of i and j, such that for i sufficiently large,*

$$|\, p_j^i(\nu^i) \,| \leqq K(\tau_j^i)^2, \qquad j = 1, 2, \cdots, s-1.$$

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

$$\lambda_1 - \mu(x) \leqq \frac{1}{\mu(x) - \lambda_2} \cdot \frac{|\, Bx - \mu(x)x \,|^2}{|\, x \,|^2}$$

for every x in $\mathfrak{H}$ with $\mu(x) > \lambda_2$, λ_2 as in (24). The lemma is then established by an argument like that of Lemma 2 of the earlier paper.

THEOREM 2. *Let* λ_1 *be an isolated point of* $\mathfrak{S}(B)$. *Then the vectors* x^i *of Theorem* 1 *converge to the characteristic vector of* B *belonging to* λ_1.

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

$$| \overset{i}{\xi}_{j+1} | \leqq K_1 | \overset{i}{\xi}_j | + (\overset{i}{t}_j)^2 | \overset{i}{\xi}_{j-1} |.$$

Hence

$$\overset{i}{t}_{j+1} \leqq K_1 + \overset{i}{t}_j.$$

Since t_1^i is bounded, it follows that there is a constant K_2 such that $t_{j+1}^i \leqq K_2$ for all i and for $j=1, 2, \cdots, s-2$. Hence by (14),

$$\sum_i (\overset{i}{\tau}_{j+1})^2 = \sum_i (\overset{i}{\tau}_j)^2 \, (\overset{i}{t}_{j+1})^2 \leqq K_2^2 \sum_i (\overset{i}{\tau}_j)^2.$$

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* $\lambda_1' = \sup \mathfrak{S}(A)$ *be an isolated point of* $\mathfrak{S}(A)$. *If* x^0 *is not orthogonal to the characteristic manifold of* A *belonging to* λ_1', *then* x^i *converges to a characteristic vector in this manifold.*

UNIVERSITY OF CHICAGO AND
NATIONAL BUREAU OF STANDARDS