BOUNDS ON CHARACTERISTIC VALUES

EDWARD W. BARANKIN

In a previous paper [1][1] the author obtained several upper bounds for the characteristic roots of a finite matrix. The purpose of this note is to establish the corresponding results for characteristic values of certain linear transformations in certain infinite spaces.

1. **Infinite matrices.** We consider a non-null infinite matrix, $A = (a_{st})$, s, $t = 1$, 2, $\cdots$, of complex numbers. (The present considerations pertain also to the finite case when $a_{st} = 0$ for $s > n$ and $t > n$.) We shall say that A is *absolutely summable* (abbreviated a.s.) when the double series $\sum_{s,t} |a_{st}|$ converges. In general, if $\sum_{s,t} |a_{st}|^p$ converges for a positive number p, A will be said to be *absolutely summable* (p). If A is a.s.(p), then $\sum_t |a_{st}|^p$ converges for all s, $\sum_s |a_{st}|^p$ converges for all t, and

$$\sum_{s,t} |a_{st}|^p = \sum_s \sum_t |a_{st}|^p = \sum_t \sum_s |a_{st}|^p.$$

We define

$$R_s^{(p)} = \sum_t |a_{st}|^p, \qquad T_t^{(p)} = \sum_s |a_{st}|^p,$$

and, for brevity, $R_s = R_s^{(1)}$, $T_t = T_t^{(1)}$. Clearly if A is a.s.(p), it is a.s.(q) for every $q > p$. Observe also that $R_s^{(p)} = 0$ or $T_t^{(p)} = 0$ imply, respectively, the vanishing of all elements in the sth row or of all elements in the tth column of A.

Let A be a.s.(p), $p \leq 1$, and l denote the space of bounded sequences of complex numbers $\{x_r\}$. Throughout, we shall view A as a transformation in l. As such, A carries l into a proper sub-manifold. For, let $\{x_r\} \in l$, and $|x_r| < M$, $r = 1$, 2, $\cdots$. Then the transform of $\{x_r\}$ has the components

$$y_s = \sum_t a_{st} x_t, \qquad\qquad s = 1, 2, \cdots .$$

Since $p \leq 1$, we have

$$|y_s|^p \leq \left(\sum_t |a_{st}| \cdot |x_t| \right)^p \leq M^p \left(\sum_t |a_{st}| \right)^p \leq M^p \sum_t |a_{st}|^p,$$

and therefore $\sum_s |y_s|^p$ is convergent. An immediate consequence of this is the following.

LEMMA. *If A is a.s.(p), $p \leq 1$, and $\{x_r\} \in l$ is a characteristic vector of A belonging to a nonzero characteristic value, then $\sum |x_r|^p$ is convergent.*

It will be assumed in the details of all proofs to follow that the characteristic value $\lambda \neq 0$; the truth of the theorems is evident for $\lambda = 0$.

THEOREM 1. *If A is a.s., and has a characteristic value λ, then:*

$$\text{(i)} \quad |\lambda|^2 \leq \max_{(s)} R_s T_s.$$

$$\text{(ii)} \quad |\lambda| \leq \max_{(s)} R_s.$$

$$\text{(iii)} \quad |\lambda| \leq \max_{(s)} T_s.$$

The bound (i) is stated separately because of its simplicity; it is a particular case of the next theorem. The proofs of (ii) and (iii) are completed by the addition to the proof for the finite case (cf. [1]) of the remark that, by virtue of the lemma above, the l.u.b. of the components of a characteristic vector is attained.

THEOREM 2. *If A is a.s.(p), $p \leq 1$, and has a characteristic value λ, then for any number q in the interval $p \leq q \leq 2 - p$,*

$$|\lambda|^2 \leq \max_{(s)} R_s^{(q)} T_s^{(2-q)}.$$

Let $\{x_s\}$ be a characteristic vector of A belonging to λ, so that $\lambda x_s = \sum_t a_{st} x_t$, $s = 1, 2, \cdots$. From these equations we derive the inequalities

$$\text{(1)} \qquad |\lambda| |x_s| \leq \sum_t |a_{st}| |x_t|, \qquad s = 1, 2, \cdots,$$

which we write in the form:

$$|\lambda| |x_s| \leq \sum_t |a_{st}|^{q/2} \cdot |a_{st}|^{1-q/2} |x_t|.$$

Application of the Schwarz inequality to the right-hand member, as indicated by the decomposition, yields

$$|\lambda| |x_s| \leq \left\{ \left(\sum_t |a_{st}|^q \right) \left(\sum_t |a_{st}|^{2-q} |x_t|^2 \right) \right\}^{1/2}$$

$$= (R_s^{(q)})^{1/2} \left(\sum_t |a_{st}|^{2-q} |x_t|^2 \right)^{1/2};$$

or, squaring this,

(2) $$|\lambda|^2| x_s|^2 \leqq R_s^{(q)} \sum_t | a_{st}|^{2-q}| x_t|^2, \qquad s = 1, 2, \cdots.$$

Since q lies in the interval $[p, 2-p]$, the numbers q and $2-q$ are not less than p, and therefore $R_s^{(q)}$ and

$$\sum_t | a_{st}|^{2-q}| x_t|^2 \qquad (\leqq M^2 R_s^{(2-q)})$$

exist for each s.

Now if $R_k^{(q)} = 0$, the inequality (2) for $s = k$ gives $x_k = 0$. Hence, in *each* of the inequalities (2), the term $|a_{sk}|^{2-q} \cdot |x_k|^2$, in the sum on the right, vanishes. That is, if all the matrix elements of the kth row vanish, the elements of the kth *column* do not appear in the inequalities (2). Let us denote by $\tilde{A} = (\tilde{a}_{mn})$ the matrix obtained from A by deleting the null-rows and the corresponding columns, and re-indexing the elements that remain, in the order in which they remain. Also, let $\{\tilde{x}_n\}$ be the re-indexed characteristic vector after dropping those zero-components which correspond to null-rows. We can depict the correspondence of the elements of $\tilde{A}$ with those of A by the notation $\tilde{a}_{mn} = a_{s_m s_n}$. We remark that

$$\sum_m | \tilde{a}_{mn}|^{2-q} = T_{s_n}^{(2-q)}.$$

This is so since the elements of the s_nth column of A which are deleted to leave the nth column of $\tilde{A}$ are those which lie in deleted rows, and such elements are all 0. Thus

$$\sum_m | \tilde{a}_{mn}|^{2-q} \text{ differs from } \sum_r | a_{rs_n}|^{2-q} = T_{s_n}^{(2-q)}$$

only by 0-terms.

With the new notation, (2) is completely represented by the system

(3) $$|\lambda|^2| \tilde{x}_m|^2 \leqq R_{s_m}^{(q)} \sum_n | \tilde{a}_{mn}|^{2-q}| \tilde{x}_n|^2, \qquad m = 1, 2, \cdots,$$

where now $R_{s_m}^{(q)} \neq 0$ for each m.

We may therefore divide the mth inequality by $R_{s_m}^{(q)}$, and sum over m, to obtain

(4) $$\sum_m \frac{|\lambda|^2}{R_{s_m}^{(q)}}| \tilde{x}_m|^2 \leqq \sum_{m,n} | \tilde{a}_{mn}|^{2-q}| \tilde{x}_n|^2 = \sum_n T_{s_n}^{(2-q)}| \tilde{x}_n|^2.$$

This last summation is valid since

$$\sum_n T_{s_n}^{(2-q)} \mid \bar{x}_n \mid^2 \leq M^2 \sum_n T_{s_n}^{(2-q)} \leq M^2 \sum_t T_t^{(2-q)},$$

the last member here being convergent for $2-q \leq p$. The inequality (4) may be put in the form

$$\sum_m \left(\frac{\mid \lambda \mid^2}{R_{s_m}^{(q)}} - T_{s_m}^{(2-q)} \right) \mid \bar{x}_m \mid^2 \leq 0.$$

Hence it appears that for some s_0, since $\{x_s\} \neq \{0\}$,

$$\frac{\mid \lambda \mid^2}{R_{s_0}^{(q)}} - T_{s_0}^{(2-q)} \leq 0,$$

or

$$\mid \lambda \mid^2 \leq R_{s_0}^{(q)} T_{s_0}^{(2-q)} \leq \max_{(s)} R_s^{(q)} T_s^{(2-q)}.$$

This proves the theorem.

We state without detailed proof the following generalization of the last result.

THEOREM 3. *Let A, λ be as in the preceding theorem. Let (q_{st}), s, $t = 1, 2, \cdots$, be a matrix of real numbers, only a finite number of which lie outside the interval $(p, 2-p)$. Then*

$$\mid \lambda \mid^2 \leq \max_{(s)} \left\{ \left(\sum_t \mid a_{st} \mid^{q_{st}} \right) \left(\sum_r \mid a_{rs} \mid^{2-q_{rs}} \right) \right\}.$$

The method of proof is clear. One starts with the inequalities (1) decomposed as follows: $\mid \lambda \mid \mid x_s \mid \leq \sum_t \mid a_{st} \mid^{q_{st}/2} \cdot \mid a_{st} \mid^{1-q_{st}/2} \mid x_t \mid.$

In the direction of a still more comprehensive class of bounds, we indicate the possibility of substituting the general Hölder inequality for the Schwarz inequality above.

2. **Integral equations.** Analogues of all the results of §1 hold for characteristic values of integral equations. Here we shall state and prove only the correspondent of Theorem 1, to fix the procedure. We consider the integral equation

$$(5) \qquad\qquad \lambda f(s) = \int_{G_t} K(s, t) f(t) dt,$$

where G_t is a fixed set on the t-axis, and the kernel $K(s, t)$ is summable on $G_s \times G_t$, G_s being the set on the s-axis corresponding to G_t on the t-axis (such corresponding sets will be designated in this way through-

out). We are concerned with characteristic functions $f(s)$ of (5) in the space of bounded measurable functions on G_s. In this case we have, corresponding to the lemma of §1, that *a characteristic function of K, belonging to a nonzero characteristic value, is integrable on G_s.*

If, when these integrals exist, we set

$$R(s) = \int_{G_t} |K(s, t)|\, dt, \qquad T(s) = \int_{G_t} |K(t, s)|\, dt,$$

we have the following theorem.

THEOREM 4. *Let the kernel K be as above, and λ a characteristic value of K. Then:*

(i) $|\lambda|^2 \leq \sup_{(s \in G_s)} R(s) T(s)$,

(ii) $|\lambda| \leq \sup_{(s \in G_s)} R(s)$,

(iii) $|\lambda| \leq \sup_{(s \in G_s)} T(s)$.

From the summability of K it follows that $R(s)$ and $T(s)$ exist almost everywhere on G_s; on the set E_s, say. Then, from (5),

$$|\lambda| \cdot |f(s)| \leq \int_{E_t} |K(s, t)| \cdot |f(t)|\, dt$$

$$= \int_{E_t} |K(s, t)|^{1/2} \cdot |K(s, t)|^{1/2} |f(t)|\, dt.$$

For $s \in E_s$ we may apply the Schwarz inequality to the extreme right member, and on squaring the result obtain

$$|\lambda|^2 |f(s)|^2 \leq R(s) \int_{E_t} |K(s, t)| |f(t)|^2 dt.$$

The integral on the right is finite since $f(t)$ is bounded. Let H_s be the subset of E_s on which $R(s) > 0$. We observe that $f(t) = 0$ for $t \in E_t - H_t$. Therefore, we may write, for $s \in H_s$,

$$\frac{|\lambda|^2}{R(s)} |f(s)|^2 \leq \int_{H_t} |K(s, t)| |f(t)|^2 dt.$$

We integrate both sides of this inequality over H_s:

$$\int_{H_s} \frac{|\lambda|^2}{R(s)} |f(s)|^2 ds \leq \int_{H_s} \int_{H_t} |K(s, t)| \cdot |f(t)|^2 dt \cdot ds.$$

By Fubini's theorem the order of integration in the right-hand mem-

ber can be reversed, the two repeated integrals being in fact equal to

$$\iint_{H_s \times H_t} |K(s, t)| \, |f(t)|^2 ds \, dt \leq (\sup |f(t)|)^2 \iint_{G_s \times G_t} |K(s, t)| \, ds \, dt,$$

which is evidently finite. Hence, we have, on further interchanging the letters s and t on the right,

$$\int_{H_s} \frac{|\lambda|^2}{R(s)} |f(s)|^2 ds \leq \int_{H_s} |f(s)|^2 \left\{ \int_{H_t} |K(t, s)| \, dt \right\} ds$$

$$\leq \int_{H_s} T(s) |f(s)|^2 ds.$$

This may be put into the form

$$\int_{H_s} \left(\frac{|\lambda|^2}{R(s)} - T(s) \right) |f(s)|^2 ds \leq 0.$$

$f(s)$ is non-null on H_s, and therefore, for some s_0,

$$\frac{|\lambda|^2}{R(s_0)} - T(s_0) \leq 0,$$

or

$$|\lambda|^2 \leq R(s_0) \cdot T(s_0) \leq \sup_{s \in G_s} R(s) T(s),$$

which is the desired result.

To establish part (ii), we let $h = \sup_{s \in E_s} |f(s)|$. Then, for $s \in E_s$,

$$|\lambda| \frac{|f(s)|}{h} \leq \int_{E_t} |K(s, t)| \frac{|f(t)|}{h} \, dt$$

$$\leq \int_{E_t} |K(s, t)| \, dt = R(s) \leq \sup_{s \in E_s} R(s).$$

A sequence of values $\{s_n\}$ can be found such that $\lim_{n \to \infty} |f(s_n)| = h$. Then, clearly,

$$\lim_{n \to \infty} |\lambda| \frac{|f(s_n)|}{h} = |\lambda| \leq \sup_{s \in G_s} R(s).$$

Finally, consider

$$|\lambda| \cdot |f(s)| \leq \int_{E_t} |K(s, t)| \cdot |f(t)| \, dt.$$

We integrate both sides of this inequality over E_s. The order of integration may be reversed, and we obtain, on interchanging the letters s and t,

$$\int_{E_s} |\lambda| \cdot |f(s)| \, ds \leqq \int_{E_s} |f(s)| \left\{ \int_{Et} |K(t, s)| \, dt \right\} ds$$

$$= \int_{E_s} T(s) |f(s)| \, ds$$

or

$$\int_{E_s} (|\lambda| - T(s)) |f(s)| \, ds \leqq 0.$$

Thus, for some s_0, $|\lambda| \leqq T(s_0) \leqq \sup_{s \in G_s} T(s)$, which is the third inequality of Theorem 4.

3. **Concerning bounds.** The literature dealing with estimation of characteristic values of finite matrices begins in 1900 with the work of Bendixson [2]. A well detailed account of results in this field up to 1939 is given by Browne [4]. In the bibliography below are listed most of the papers that have appeared since that time. One of the most effective results is the following, due to Schur [8]: *if λ_1, λ_2, $\cdots$, λ_n are the n characteristic roots of the matrix $A = (a_{st})$, s, $t = 1$, 2, $\cdots$, n, then*

$$\sum_1^n |\lambda_i|^2 \leqq \sum_{s,t=1}^n |a_{st}|^2.$$

The equality holds if and only if A is normal. The corresponding bound is established for the integral equation (5) with G_s a finite interval and K continuous on $G_s \times G_t$:

$$\sum_1^\infty |\lambda_i|^2 \leqq \int\int_{G_s \times G_t} |K(s, t)|^2 ds dt.$$

The sum on the left here is taken over all the reciprocal zeros (counting multiplicities) of the Fredholm determinant of K.

Schur's estimate and ours are best estimates, but with different ranges. There are, in fact, for each order n, matrices for which

$$n \cdot \max_{(s)} (R_s T_s) < \sum_{s,t=1}^n |a_{st}|^2,$$

the left member therefore affording a sharper bound for $\sum_1^n |\lambda_i|^2$. An

example is the matrix

$$\begin{bmatrix} 0 & 2 & 1 \\ 10 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

For normal matrices, and for normal operators in Hilbert space, Schur's global relations can be augmented by more detailed statements concerning the location of individual characteristic values. Some results in this direction will be published by the author in the near future.

BIBLIOGRAPHY

1. E. W. Barankin, *Bounds for the characteristic roots of a matrix*, Bull. Amer. Math. Soc. vol. 51 (1945) pp. 767–770.

2. I. Bendixson, *Sur les racines d'une équation fondamentale*, Acta Math. vol. 25 (1902) pp 359–365.

3. A. Brauer, *Limits for the characteristic roots of a matrix*, Duke Math. J. vol. 13 (1946) pp. 387–395.

4. E. T. Browne, *Limits to the characteristic roots of a matrix*, Amer. Math. Monthly vol 46 (1939) pp. 252–265.

5. A. B. Farnell, *Limits for the characteristic roots of a matrix*, Bull. Amer. Math. Soc. vol. 50 (1944) pp. 789–794.

6. W. V. Parker, *Limits to the characteristic roots of a matrix*, Duke Math J. vol. 10(1943) pp. 479–482.

7. ———, *The characteristic roots of matrices*, Duke Math. J. vol. 12 (1945) pp. 517–526.

8. Schur, I. *Über die charakteristischen Wurzeln einer linearen Substitution mit einer Anwendung auf die Theorie der Integralgleichungen*, Math. Ann. vol. 66 (1909) pp. 488–510.

UNIVERSITY OF CALIFORNIA