JOURNAL OF APPROXIMATION THEORY 22, 119–149 (1978)

The Location of Eigenvalues and Eigenvectors of Complex Matrices

M. GURARI*

Communicated by Oved Shisha

Received May 3, 1976

1. INTRODUCTION AND SUMMARY OF RESULTS

It is of considerable interest, both for theoretical purposes and for the applications, to obtain information about the location of the eigenvalues and eigenvectors of a matrix. There is an extensive literature on the location of the eigenvalues. (See, for example, the survey by Householder [11].) Somewhat less is known about the location of eigenvectors.

One important result is the Perron theorem which states that a positive matrix has a positive eigenvector belonging to a positive eigenvalue (Perron [17]; Frobenius [6, 7]). This has given rise to a considerable literature (see Brauer [4], Seneta [20]). Of course, many theorems on the effect of perturbations give information on the eigenvectors of a matrix close to a given matrix (Kato [14]).

Many results on the eigenvalues and eigenvectors of matrices can be extended to infinite-dimensional spaces. We may mention, for example, the Jentzsch theorem on integral operators with positive kernel [12], and its generalizations. (See Ostrowski [16], Krein and Rutman [15].)

In this paper we present several results on the location of the eigenvalues and eigenvectors of complex matrices, together with some extensions to infinite-dimensional sequence spaces. For example, we can obtain a result of the form (Theorem 12):

Let $C = (C_{jk})$, $0 \leqslant j, k \leqslant N$, be a matrix such that

$$C_{jk} = r_{jk} \exp(i\theta_{jk}), \qquad r_{jk} \geqslant 0, \qquad -\pi < \theta_{jk} \leqslant \pi,$$

* At the time of Mark Gurari's death on May 8, 1952, he was in the Department of Theoretical Physics at the University of Liverpool. A manuscript in German on the above subject was found among his papers. We have prepared this paper from his manuscript, extended and simplified some results, and put the material in relation to other published work. In the manuscript the author refers to a uniqueness theorem related to Theorem 12, and to analogs for integral equations. Unfortunately, these results are apparently lost. *Paul C. Rosenbloom, Department of Mathematics, Teachers College, Columbia University, New York, New York 10027.*

0021-9045/78/0222-0119$02.00/0

Copyright © 1978 by Academic Press, Inc.
All rights of reproduction in any form reserved.

640/22/2-2

and set

$$a = \max_{j \geqslant 1, k \geqslant 0} (r_{jk}/r_{0k}), \qquad \theta = \max_{j,k \geqslant 0} | \theta_{jk} |,$$

$$b = \max_{j \geqslant 0} \left(\sum_{k=1}^{N} r_{jk}/r_{j0} \right).$$

If $0 < \theta \leqslant \pi/8$ and

$$a < F(b, \theta), \tag{1}$$

then C has an eigenvector z in the set

$$S(3\theta): z_0 = 1, \qquad | \arg z_j | < 3\theta \qquad \text{for} \quad 0 < j \leqslant N.$$

Our proof gives $F(b, \theta) = \sin(4\theta) \cos(\theta/2)/b$. We have not tried to obtain a very sharp result, but have been concerned in getting an F which is explicit and easy to compute, and which lends itself to extension to the infinite-dimensional case.

If $a = 0$, then $C_{jk} = 0$ for $jk \neq 0$, so that we may call C a *border matrix*. Condition (1) says that C has a *dominant border*. While there is a considerable literature on matrices with a dominant *main diagonal*, little seems to have been done on matrices of the above type.

If C is a border matrix, then we may easily show that:

(a) C has the eigenvalue $\lambda = 0$ with multiplicity $N - 1$, and the corresponding $(N - 1)$-dimensional eigenspace defined by

$$z_0 = 0, \qquad \sum_{1}^{N} C_{0k} z_k = 0.$$

(b) The roots of the quadratic equation

$$\lambda^2 - C_{00}\lambda - d = 0, \qquad d = \sum_{k=1}^{N} C_{0k} C_{k0},$$

are also eigenvalues, and have the eigenvectors

$$z_0 = \lambda, \qquad z_j = C_{j0} \qquad \text{for} \quad j > 0.$$

A border matrix is also a matrix of rank at most 2, but for our purposes the above representation in a particular coordinate system seems more convenient.

By applying known results in perturbation theory (see Kato [14], Rosenbloom[18]), we can also obtain sufficient conditions for the uniqueness of an eigenvalue in a specified region, and information on the location of the corresponding eigenvector. For instance, we obtain a result of the form: If

$$\left| \sum_{k=1}^{N} C_{jk} C_{k0} \right| \leqslant | C_{00} | \, | C_{j0} | \, \delta, \qquad \text{for} \quad j \geqslant 0,$$

and

$$C_{0j} C_{j0} / C_{00}^2 \geqslant 0 \qquad \text{for all } j, \text{ and } d \neq 0,$$

and

$$\delta < B(d/C_{00}^2),$$

then C has a unique eigenvalue in the half-plane $R(\lambda/C_{00}) \geqslant \frac{1}{2}$ and a corresponding eigenvector x in the region

$$| \arg(x_j/C_{j0}) | \leqslant \alpha \qquad \text{for} \quad j \geqslant 0,$$

where

$$\sin \alpha = B_1(d/C_{00}^2) \, \delta, \qquad 0 \leqslant \alpha < \pi/2.$$

Here B and B_1 are explicitly computable functions of d/C_{00}^2.

Of course, we can apply perturbation theory to obtain similar results for nearly positive matrices. Combining perturbation theory with the results of Ostrowski [16] (see also Birkhoff [2], Hopf [10]), we obtain results of the type:

If $C_{jk} = r_{jk} \exp(i\theta_{jk})$ and $r_{jk} > 0$, $| \theta_{jk} | \leqslant \theta < \pi/2$ for $j, k \geqslant 0$, and $0 < \gamma < \pi/2$, and

$$2 \sin(\theta/2) \leqslant B_2 \sin \gamma,$$

then C has an eigenvector in $S(\gamma)$.

Here B_2 is an easily computable function of the r_{jk}.

If $R = (r_{jk})$ is positive, and λ_R is the positive eigenvalue of R, then there is a certain constant $N_R < 1$ such that for $N_R \lambda_R < r < \lambda_R$ and θ sufficiently small, the matrix C has a unique eigenvalue in $| \lambda | \geqslant r$. There will be an eigenvector of C, belonging to this eigenvalue, in the region $S(\gamma)$. Also C has no other eigenvector in $S(\gamma)$. For N_R we can use Ostrowski's sharpening of Birkhoff's bound. Again all bounds are computable from the data R, r, and γ.

If R is nonnegative but some power R^m is positive, so that R belongs to the class of power-positive matrices studied by Brauer [3], then we can obtain similar results.

In many applications we are dealing with large matrices, or matrices depending on parameters, or families of matrices. It is then important to find comparatively simple functions of the elements, in terms of which we can obtain the desired information about the eigenvalues and eigenvectors. Thus the practical significance of our results is perhaps the identification of such computable functions of the data, and the orders of magnitude of the bounds we obtain.

We note, finally, that the approaches of Kantorovich *et al.* [13], and Krein and Rutman [15], may lend themselves to extensions to infinite-dimensional spaces.

2. The Perron Theorem

We begin by recalling the Perron theorem.[1] A matrix $C = (C_{jk})$, $1 \leqslant j, k \leqslant n$, is called *nonnegative*, $C \geqslant 0$, if $C_{jk} \geqslant 0$ for all j, k, and is called *positive*, $C > 0$, if $C_{jk} > 0$ for all j, k. We defined similarly the concepts of nonnegative and positive vectors.

Perron's Theorem [17]. *If $C > 0$, then C has a positive eigenvector belonging to a positive eigenvalue λ.*

(a) *If μ is any other eigenvalue, then $|\mu| < \lambda$.*

(b) *The eigenspace of λ is one-dimensional.*

(c) *There is no other eigenvalue which has a nonnegative eigenvector.*

We shall denote this eigenvalue by λ_C. It can also be characterized in terms of a variational problem. For any $x > 0$, let

$$\tau(x) = \min_j \left(\sum_k C_{jk} x_k / x_j \right). \tag{2}$$

Then the maximum of $\tau(x)$ for $x > 0$ and

$$\| x \| = \max_j | x_j | = 1$$

is attained, and this maximum is λ_C.

Frobenius [6, 7] extended Perron's theorem to certain classes of nonnegative matrices and characterized those nonnegative matrices which have more than one eigenvalue of maximum modulus. The extension is especially simple for the class of matrices introduced by Brauer [3]. He calls a matrix C *power-positive* if some power C^m is positive.

[1] *Editor's note*: In the original manuscript, the Perron theorem is rediscovered. The proof is similar to the one given in Bellman [1], and ascribed to unpublished work of Bohnenblust. We have presented Gurari's argument in a way which brings out some additional points of interest.

THEOREM 1. *If $C \geqslant 0$ and $C^m > 0$, then the maximum of $\tau(x)$ for $x > 0$, $\| x \| = 1$, is attained. The maximum is a positive eigenvalue λ_C, and is attained for a positive eigenvector ξ belonging to λ_C.*

Proof. We note first that $\tau(x)$ may be characterized as the maximum of the real numbers τ such that

$$Cx - \tau x \geqslant 0.$$

Since $Cx - \tau(x)x \geqslant 0$ and $C \geqslant 0$, we infer that

$$C(Cx - \tau(x)x) \geqslant 0,$$

that is,

$$C^2x - \tau(x)\, Cx \geqslant 0,$$

so that

$$\tau(Cx) \geqslant \tau(x).$$

It follows that

$$\tau(C^m x) \geqslant \tau(x).$$

For any positive matrix A, we define

$$\gamma(A) = \min_{j,k} A_{jk}\,. \tag{3}$$

It is then trivial that if $y \geqslant 0$ and $A > 0$, and $u = Ay$, then

$$\min u_j \geqslant \gamma(A) \| y \|.$$

We note also that for any matrix A, we have

$$\| A \| = \max_{\|x\| \leqslant 1} \| A \| = \max_j \sum | A_{jk} |.$$

Now let $y = C^m x$, $z = y/\| y \|$. Then we have

$$\min y_j \geqslant \gamma(C^m) \| x \|$$

and

$$\| y \| \leqslant \| C^m \| \| x \|,$$

so that

$$\min z_j \geqslant \gamma(C^m)/\| C^m \|$$

and

$$\tau(z) = \tau(y) \geqslant \tau(x).$$

Therefore the supremum of $\tau(x)$ on the set of $x \geqslant 0$, $\| x \| = 1$, is the same as its supremum on the subset where $\min_j x_j \geqslant \gamma(C^m)/\| C^m \|$. Since τ is continuous on this subset, it attains its maximum λ there at some vector ξ. If $\epsilon = \| C\xi - \lambda\xi \| > 0$ and $y = C^m(C\xi - \lambda\xi)$, then $\min y_j \geqslant \gamma(C^m)\epsilon > 0$.

But since $y = C(C^m\xi) - \lambda C^m\xi$, this implies that $\tau(C^m\xi) > \lambda$, which contradicts the definition of λ. Hence $\epsilon = 0$ and $C\xi = \lambda\xi$. Thus all $\xi > 0$ for which $\tau(\xi) = \lambda$ are eigenvectors of C.

The function

$$\mu(C) = \gamma(C)/\| C \|$$

seems to be a natural measure of the positivity of a matrix and arises frequently in the sequel. In the course of the argument, we proved

COROLLARY 1a. *If $C \geqslant 0$ and $C^m > 0$, and ξ is the positive unit vector which maximizes τ, then* $\min_j \xi_j \geqslant \mu(C^m)$.

For the sake of completeness we prove that properties (a)–(c) in Perron's theorem hold also for nonnegative power-positive matrices. Let C' be the transpose of C, and let η be a positive unit eigenvector belonging to λ. If μ is any eigenvalue of C other than λ and z is an eigenvector belonging to μ, let $| z |$ be the vector with the components $| z_j |$, $1 \leqslant j \leqslant n$. Let $A = C^m$. We have

$$A'\eta = \lambda^m\eta \qquad \text{and} \qquad Az = \mu^m z.$$

It follows that

$$| \mu |^m \, | z | \leqslant A \, | z |$$

and

$$A \, | z | - | \mu |^m \, | z | \neq 0$$

unless z is a scalar multiple of a nonnegative vector, and then μ must also be nonnegative. Consequently, except in this case, we have

$$\eta \cdot (A \, | z | - | \mu |^m \, | z |) > 0,$$

that is,

$$(\lambda^m - | \mu |^m)(\eta \cdot | z |) > 0,$$

so that

$$| \mu | < \lambda.$$

In the exceptional case we may assume $z \geqslant 0$, $\mu \geqslant 0$. Then we obtain

$$0 = \eta \cdot (Cz - \mu z) = (\lambda - \mu)(\eta \cdot z),$$

and therefore $\mu = \lambda$.

Finally if z is any eigenvector of C belonging to λ, let

$$u = z - \frac{(\eta \cdot z)}{\eta \cdot \xi} \, \xi,$$

so that

$$Cu = \lambda u, \qquad \eta \cdot u = 0.$$

If z is not a scalar multiple of ξ, then $u \neq 0$. If u were a scalar multiple of a nonnegative vector, then $| \eta \cdot u |$ would be positive. Hence we find that

$$A | u | - \lambda^m | u | \geqslant 0, \qquad A | u | - \lambda^m | u | \neq 0,$$

and therefore

$$\eta \cdot (A | u | - \lambda^m | u |) > 0.$$

But since $\eta \cdot A | u | = \lambda^m | u |$, we have arrived at a contradiction. Therefore $u = 0$, and z is a scalar multiple of ξ.

The same argument shows that λ can be characterized by another extremal problem.

THEOREM 2. *If $C \geqslant 0$ and $C^m > 0$ and for $x > 0$*

$$\sigma(x) = \max_j \left(\sum C_{jk} x_k / x_j\right), \tag{4}$$

then

$$\lambda = \min_{x>0} \sigma(x) = \sigma(\xi).$$

Proof. Let $\mathbf{1}$ be the vector with all components equal to 1. It is sufficient to look for the minimum of σ on the set S_1 of all $x > 0$ such that $\| x \| = 1$ and $\sigma(x) \leqslant \sigma(\mathbf{1})$. The argument of Theorem 1 shows that $\sigma(C^m x) \leqslant \sigma(Cx) \leqslant \sigma(x)$, and that if $\sigma(C^m x) \leqslant \sigma(\mathbf{1})$, $x > 0$, and $\| x \| = 1$, then

$$x_j \geqslant \gamma(C^m)/\sigma(\mathbf{1}) \qquad \text{for all } j. \tag{5}$$

Thus it suffices to look for the minimum of σ on the subset S_2 of those x in S_1 which satisfy (5). Again if the minimum is attained at $v \in S_2$ and $\sigma(v)v - Cv \neq 0$, then $\sigma(C^m v) < \sigma(v)$, which contradicts the minimum property of v.

We note

COROLLARY 2a. *If $C \geqslant 0$ and $C^m > 0$, then*

$$\lambda_C \geqslant \max_j C_{jj},$$

and

$$\min_j \sum_k C_{jk} \leqslant \lambda_C \leqslant \max_j \sum_k C_{jk} = \| C \|.$$

Proof. The first estimate follows from $\lambda_C \geqslant \tau(\delta^{(j)})$, where $\delta^{(j)}$ is the vector with components δ_{jk}. The second follows from $\tau(\mathbf{1}) \leqslant \lambda_C \leqslant \sigma(\mathbf{1})$.

If τ^* and σ^* are the functions corresponding to τ and σ for the transposed matrix C', then we have the inequalities

$$\tau(x) \leqslant \sigma^*(y) \qquad \text{and} \qquad \sigma(x) \geqslant \tau^*(y)$$

for any positive x and y.

We wish to sharpen property (c) of Perron's theorem.

THEOREM 3. *If $C \geqslant 0$ and $C^m > 0$, and*

$$Cz = \mu z, \quad z_1 = 1, \quad |\arg z_j| \leqslant \pi/2 \quad \text{for } j > 1,$$

then $\mu = \lambda$ and $z > 0$.

Proof. Let η be the vector such that

$$\eta > 0, \quad \|\eta\| = 1, \qquad \text{and} \qquad C'\eta = \lambda\eta.$$

By property (b), the conclusion follows if $\mu = \lambda$. If $\mu \neq \lambda$, then from $\eta \cdot Cz = \lambda(\eta \cdot z)$ we obtain $\eta \cdot z = 0$. But

$$R(\eta \cdot z) = \eta_1 z_1 + \sum_2^n \eta_j R(z_j) > 0,$$

which is a contradiction.

Following Ostrowski [16], if $y > 0$ and x is any real vector, we define $m(x; y)$ and $M(x; y)$ as the upper and lower bounds, respectively, of m and M such that

$$my \leqslant x \leqslant My.$$

Then we have

$$\tau(x) = m(Cx; x), \qquad \sigma(x) = M(Cx; x),$$
$$\min x_j = m(\lambda; \mathbf{1}), \qquad \max x_j = M(x; \mathbf{1}).$$

Birkhoff [2] introduced into the study of positive matrices the projective metric

$$\theta(x, y) = \log(M(x; y)/m(x; y)) \qquad (x, y > 0)$$

of Hilbert [8] (see also Busemann and Kelley [5]). If C is a positive matrix, then we have

$$\max_{x,y>0} \theta(Cx, Cy) = \log T_C = \Delta_C,$$

where

$$T_C = \max_{j,k,r,s} (C_{jk}C_{rs}/C_{rk}C_{js});$$

(Ostrowski [16, p. 87]), and

$$\max \theta(Cx, Cy)/\theta(x, y) = \frac{T_C^{1/2} - 1}{T_C^{1/2} + 1} = N_C$$

(Birkhoff [2, Lemma 1; p. 221]).

We can use these relations to sharpen the considerations of Theorem 1. If $C^m > 0$ and $x > 0$ and $T = T(C^m)$, then

$$M(C^{m+1}x - \tau(x)\, C^m x; C^m x) \leqslant Tm(C^{m+1}x - \tau(x)\, C^m x; C^m x).$$

If $M = M(C^{m+1}x - \tau(x)\, C^m x; C^m x)$ and η is the positive eigenvector of C' as in the proof of Theorem 3, then we have

$$C^{m+1}x - \tau(x)\, C^m x \leqslant MC^m x,$$

so that, with $\lambda = \lambda_C$,

$$\lambda^m(\lambda - \tau(x))(\eta \cdot x) \leqslant M\lambda^m(\eta \cdot x),$$

and

$$\lambda - \tau(x) \leqslant M \leqslant Tm(C^{m+1}x - \tau(x)\, C^m x; C^m x).$$

We infer that

$$C^{m+1}x - \tau(x)\, C^m x \geqslant T^{-1}(\lambda - \tau(x))\, C^m x,$$

which yields

$$\tau(C^m x) \leqslant \tau(x) + T^{-1}(\lambda - \tau(x)),$$

or

$$\lambda - \tau(C^m x) \leqslant (1 - T^{-1})(\lambda - \tau(x)).$$

If $v = mq + r$, $0 \leqslant r \leqslant m$, then it follows that

$$\begin{aligned} \lambda - \tau(C^v x) &\leqslant \lambda - \tau(C^{mq} x) \\ &\leqslant (1 - T^{-1})^q(\lambda - \tau(x)), \end{aligned} \tag{6}$$

so that

$$\lim_{v \to \infty} \tau(C^v x) = \lambda.$$

Similarly, we find that

$$\sigma(C^m x) - \lambda \leqslant (1 - T^{-1})(\sigma(x) - \lambda). \tag{7}$$

In the course of the proofs of Theorems 1 and 2, we proved $\tau(Cx) \geqslant \tau(x)$, $\sigma(Cx) \leqslant \sigma(x)$, which imply that

$$\begin{aligned} \theta(C^m x; x) &\leqslant \sum_{k=0}^{m-1} \theta(C^{k+1}x; C^k x) \\ &\leqslant m\theta(Cx; x). \end{aligned}$$

But if ξ is the positive eigenvector of C such that $\| \xi \| = 1$ and $N = N(C^m)$, then

$$\begin{aligned}\theta(x; \xi) &\leqslant \theta(x; C^m x) + \theta(C^m x, \xi) \\ &\leqslant \theta(x, C^m x) + N\theta(x, \xi)\end{aligned}$$

so that

$$\theta(x, \xi) \leqslant (1 - N)^{-1} \theta(x, C^m x).$$

If x is normalized so that $m(x; \xi)\, M(x; \xi) = 1$, that is

$$\log M(x; \xi) = \theta(x, \xi)/2 = \theta/2,$$

then

$$\exp(-\theta/2)\xi \leqslant x \leqslant \exp(\theta/2)\xi,$$

and

$$\begin{aligned}\| x - \xi \| &\leqslant (\exp(\theta/2) - 1) \| \xi \| \\ &\leqslant (\sigma(x)/\tau(x))^k - 1,\end{aligned} \tag{8}$$

where

$$k = m(1 - N)^{-1}/2.$$

Thus we can estimate the distance from x to ξ in terms of the ratio $\sigma(x)/\tau(x)$.

THEOREM 4. *If $C \geqslant 0$ and $C^m > 0$ and $x > 0$, then*

$$\lim_{v \to \infty} \sigma(C^v x) = \lim_{v \to \infty} \tau(C^v x) = \lambda_C,$$

and we have inequalities (6) *and* (7) *on the rates of convergence of $\sigma(C^v x)$ and $\tau(C^v x)$ to λ_C, and* (8) *on the distance from x to the positive eigenvector.*

In the following we shall continue to denote by ξ and η the positive eigenvectors of C and C', respectively. If we set

$$\xi \cdot \eta = 1,$$

then we may still replace ξ and η by $a\xi$ and $a^{-1}\eta$, respecitvely, where a is any positive number. It will be convenient to postpone further normalization of ξ and η until later.

The transformation C' may be considered as the adjoint of C, operating on the dual space with the norm

$$\| y \|_1 = \sum | y_j |.$$

For future use we give the following modification of (8):

LEMMA 1. *If* $x > 0$ *and* $\theta(x, \xi) \leqslant \epsilon$, *then*

$$\begin{aligned}\| x - (\eta \cdot x)\xi \| &\leqslant (e^{\epsilon} - 1) \mid \eta \cdot x \mid \| \xi \| \\ &\leqslant (e^{\epsilon} - 1) \| \eta \|_1 \| \xi \| \| x \|.\end{aligned}$$

Proof. Let $z = x - (\eta \cdot x)\xi$. We have

$$\begin{aligned}M(x; \xi) &= (\eta \cdot x) + M(z; \xi), \\ m(x; \xi) &= (\eta \cdot x) + m(z; \xi),\end{aligned}$$

and $m(z; \xi) \leqslant \eta \cdot z = 0 \leqslant M(z; \xi)$.

Furthermore, we have $M(x; \xi) \leqslant e^{\epsilon} m(x; \xi)$. It follows that

$$M(z; \xi) - m(z; \xi) \leqslant M(z; \xi) - e^{\epsilon} m(z, \xi) \leqslant (e^{\epsilon} - 1)(\eta \cdot x).$$

Since

$$\| z \| \leqslant \max(M(z; \xi), -m(z; \xi)) \| \xi \|,$$

we obtain

$$\| z \| \leqslant (e^{\epsilon} - 1)(\eta \cdot x) \| \xi \|.$$

The following estimates are also useful.

LEMMA 2.

$$\begin{aligned}\| \xi \| \| \eta \|_1 &\leqslant 1/\mu(C), \\ \| \xi \| \| \eta \|_1 &\leqslant \exp((1 - N_C)^{-1} \theta(C\mathbf{1}, \mathbf{1})), \\ m(\eta; \mathbf{1}) &\geqslant \mu(C) \| \eta \|_1 .\end{aligned}$$

Proof. Since

$$1 = \eta \cdot \xi \geqslant m(\xi; \mathbf{1}) \| \eta \|_1 ,$$

the first estimate follows from Corollary 1a. An alternative estimate of $\| \xi \| / m(\xi; \mathbf{1}) = \exp(\theta(\xi; \mathbf{1})$ follows from

$$\begin{aligned}\theta(\xi; \mathbf{1}) &\leqslant \theta(\xi, C\mathbf{1}) + \theta(C\mathbf{1}; \mathbf{1}) \\ &\leqslant N_C \theta(\xi, \mathbf{1}) + \theta(C\mathbf{1}, \mathbf{1}),\end{aligned}$$

and this yields the second inequality. The third follows from

$$\lambda_C \eta = C' \eta,$$

which implies

$$\begin{aligned}\gamma(C) \| \eta \|_1 &\leqslant \lambda_C m(\eta; \mathbf{1}) \\ &\leqslant \| C \| m(\eta; \mathbf{1}).\end{aligned}$$

We note that $\theta(C\mathbf{1}, \mathbf{1}) = \log(\sigma(\mathbf{1})/\tau(\mathbf{1}))$, and $\sigma(\mathbf{1})/\tau(\mathbf{1})$ is the ratio of the bounds for λ_C given in Corollary 2a.

THEOREM 5. *If* $\eta \cdot z = 0$, *and* $C > 0$, *then*

$$\| C^v z \| \leqslant (34/\mu(C))(\lambda_C N_C)^v \| z \|.$$

Remark. Ostrowski gives bounds which imply that $\| C^v z \| = O((\lambda_C N_C)^v)$, but do not specify the constant implicit in this result.

Proof. If we set $m = m(\xi; \mathbf{1})$, then from

$$-(\| z \|/m)\xi \leqslant z \leqslant (\| z \|/m)\xi$$

we obtain

$$M(z; \xi) \leqslant \| z \|/m \leqslant (\| z \|/(\mu(C) \| \xi \|)) = a,$$

and

$$-m(z; \xi) \leqslant a.$$

If $k > 1$, then we have

$$x = ka\xi + z > 0$$

and

$$M(x; \xi) \leqslant (k + 1)a, \qquad m(x; \xi) \geqslant (k - 1)a,$$

so that

$$\theta(x; \xi) \leqslant (k + 1)/(k - 1) = t.$$

Now we obtain

$$\theta(C^v x; \xi) \leqslant N_C{}^v \theta(x; \xi) \leqslant t N_C{}^v = \epsilon.$$

From

$$C^v x = ka\lambda_C{}^v \xi + C^v z$$

and

$$\eta \cdot C^v z = 0$$

we infer, by Lemma 1, that

$$\| C^v z \| \leqslant (e^\epsilon - 1)\, ka\lambda_C{}^v \| \xi \|.$$

But $\epsilon < t$, so that we have

$$e^\epsilon - 1 \leqslant e^\epsilon \epsilon \leqslant e^t \epsilon.$$

If we now choose $k = 5$, we obtain the theorem.

3. Perturbation of Simple Eigenvalues and Eigenvectors

In this section we give results which we need on the perturbation of simple eigenvalues and eigenvectors. While these results are implicit, in principle, in the available literature, it is hard to find there explicit quantitative results. Rosenbloom [18] and Kato [14] have obtained results of the kind we want by quite different methods. Since their estimates are expressed in terms of different data, it is difficult to compare them. We shall work out here various estimates using Kato's method, based on the analysis of the resolvent (see Hille and Phillips [9]). This approach has the advantage that it can easily be extended to eigenvalues of higher multiplicity.

Suppose that C is a linear transformation of a complex Banach space $\mathfrak{X}$ into itself, and let

$$R(\lambda; C) = (\lambda - C)^{-1}$$

be the resolvent of C. We say that an eigenvalue λ_0 of C is *simple* if it is also an eigenvalue of the conjugate transformation C^* on the conjugate space $\mathfrak{X}^*$, the null-spaces of $C - \lambda_0$ and $C^* - \lambda_0$ are one-dimensional, and λ_0 is an isolated point of the spectrum of C. It follows that $R(\lambda; C)$ has a pole of order 1 at λ_0 and that its residue there is a projection P_0 onto the null-space of $C - \lambda_0$, and

$$CP_0 = P_0C = \lambda_0 P_0 .$$

Let x_0 and x_0^* be eigenvectors of C and C^*, respectively, such that $x_0^*(x_0) = 1$. Thus P_0 can be expressed in the form

$$P_0 = x_0 \otimes x_0^*,$$

that is,

$$P_0 x = x_0^*(x)\, x_0 \qquad \text{for all} \quad x \in \mathfrak{X}.$$

If Ω is a domain with rectifiable boundary containing no eigenvalues on its boundary, then

$$P = \frac{1}{2\pi i}\int_{\partial\Omega} R(\lambda; C)\, d\lambda$$

is a projection onto the union of the eigenspaces corresponding to the portion of the spectrum of C contained in Ω. In particular, if Ω contains λ_0 and no other point of the spectrum of C, then $P = P_0$.

Suppose Ω is such a domain, and let

$$M = \max_{\lambda \in \partial\Omega} \| R(\lambda; C)\|.$$

If U is a bounded linear transformation of $\mathfrak{X}$ into itself, then we have

$$R(\lambda; C + U) = R(\lambda; C)(1 - UR(\lambda; C))^{-1}$$

and

$$R(\lambda; C + U) - R(\lambda; C) = R(\lambda; C + U)\, UR(\lambda; C),$$

by Hille and Phillips [9, p. 196–197].

Hence if

$$\| U \| \leqslant \delta < 1/M,$$

then we obtain

$$\| R(\lambda; C + U) \| \leqslant M/(1 - \delta M)$$

and

$$\| R(\lambda; C + U) - R(\lambda; C) \| \leqslant M^2\delta/(1 - \delta M).$$

If

$$P_U = \frac{1}{2\pi i} \int_{\partial\Omega} R(\lambda; C + U)\, d\lambda,$$

then

$$\| P_U - P_0 \| \leqslant BM^2\delta/(1 - \delta M),$$

where $B =$ (length of $\partial\Omega$)$/2\pi$. Consequently, if

$$\delta < 1/(M + BM^2),$$

then we obtain

$$\| P_U - P_0 \| < 1.$$

By Kato [14, p. 33], this implies that the rank of P_U is one-dimensional, so that $C + U$ has a unique eigenvalue $\lambda(U)$ in Ω, and

$$P_U x = x_U{}^*(x)\, x_U\,, \qquad x_U{}^*(x_U) = 1,$$

for all $x \in \mathfrak{X}$, where x_U and $x_U{}^*$ are the eigenvectors of $C + U$ and $(C + U)^*$, respectively. Furthermore, we have

$$(C + U) P_U = P_U(C + U) = \lambda(U) P_U\,.$$

But the formula

$$\begin{aligned}(\lambda(U) - \lambda_0)\, P_U &= \frac{1}{2\pi i} \int_{\partial\Omega} (\lambda - \lambda_0)\, R(\lambda; C + U)\, d\lambda \\ &= \frac{1}{2\pi i} \int_{\partial\Omega} (\lambda - \lambda_0)(R(\lambda; C + U) - R(\lambda; C))\, d\lambda\end{aligned}$$

implies that

$$\begin{aligned}| \lambda(U) - \lambda_0 |\, \| P(U) \| &\leqslant dBM^2\delta/(1 - \delta M) \\ &\leqslant dM(1 + BM)\delta,\end{aligned}$$

where

$$d = \max_{\lambda \in \partial\Omega} | \lambda - \lambda_0 |$$

and

$$| \lambda(U) - \lambda_0 | \leqslant dM(1 + BM)\delta$$

since

$$\| P(U) \| \geqslant 1.$$

Furthermore, we have

$$\| P_U(x_0) - P_0(x_0) \| = \| x_U{}^*(x_0)\, x_U - x_0 \| \leqslant M(1 + BM)\delta \, \| x_0 \|.$$

Similarly, we can show that

$$\| x_0{}^*(x_U)\, x_U{}^* - x_0{}^* \| \leqslant M(1 + BM)\delta \, \| x_0{}^* \|.$$

We summarize these results in

THEOREM 6. *Suppose that C is a linear transformation of $\mathfrak{X}$ into itself, that λ_0 is a simple eigenvalue of C, that x_0 and $x_0{}^*$ are eigenvectors of C and C^*, respectively, belonging to λ_0 such that $x_0{}^*(x_0) = 1$, that Ω is a domain with boundary $\partial\Omega$ of length $2\pi B$ containing λ_0 and no other points of the spectrum of C, and that*

$$M = \max_{\lambda \in \partial\Omega} \| R(\lambda; C) \|.$$

Then for

$$\| U \| \leqslant \delta < 1/(M + BM^2) = 1/K,$$

the transformation $C + U$ has a unique eigenvalue $\lambda(U)$ in Ω, which is simple. This eigenvalue satisfies

$$| \lambda(U) - \lambda_0 | \leqslant Kd\delta,$$

where

$$d = \max_{\lambda \in \partial\Omega} | \lambda - \lambda_0 |.$$

There are eigenvectors of $C + U$ and $(C + U)^*$, respectively, belonging to $\lambda(U)$ in the spheres

$$\| x - x_0 \| \leqslant K\delta \, \| x_0 \| \qquad \textit{and} \qquad \| x^* - x_0{}^* \| \leqslant K\delta \, \| x_0{}^* \|. \tag{9}$$

We remark that we can always normalize x_0 and $x_0{}^*$ so that $\| x_0{}^* \| = \| x_0 \| = \| P_0 \|^{1/2}$.

By minor modifications of the above argument, we can obtain similar results for unbounded regions Ω. For example, we have

COROLLARY 6a. *Suppose that C is a bounded linear transformation on $\mathfrak{X}$ to itself, and λ_0, x_0, and x_0^* are as above. Suppose also that there are no other points of the spectrum of C in $|\lambda| \geqslant r$, where $r < |\lambda_0|$, and that*

$$M = \max_{|\lambda|=r} \| R(\lambda; C) \|.$$

Then for

$$\| U \| \leqslant \delta < 1/(M + rM^2) = 1/K$$

the transformation $C + U$ has a unique eigenvalue $\lambda(U)$ in $|\lambda| \geqslant r$, which is simple. It satisfies

$$|\lambda(U) - \lambda_0| \leqslant (1 + 2 |\lambda_0| K)\delta.$$

There are eigenvectors of $C + U$ and $(C + U)^*$, respectively, belonging to $\lambda(U)$ and satisfying (9).

For the proof, we take Ω to be the annulus $r < |z| < R$, and let $R \to \infty$. We note that

$$\begin{aligned} R(\lambda; C) &= \sum_0^\infty C^k/\lambda^{k+1} \\ &= \lambda^{-1} + C\lambda^{-2} + O(\lambda^{-3}) \end{aligned}$$

for $|\lambda| > |\lambda_0|$. This implies that

$$P_U - P_0 = -\frac{1}{2\pi i} \int_\Gamma (R(\lambda; C + U) - R(\lambda, U))\, d\lambda$$

and

$$(\lambda(U) - \lambda_0)\, P_U = U - \frac{1}{2\pi i} \int_\Gamma (\lambda - \lambda_0)(R(\lambda; C + U) - R(\lambda, U))\, d\lambda,$$

where Γ is the circle $|\lambda| = r$. The rest of the reasoning is as before.

Another variant of the argument yields

COROLLARY 6b. *Let $a > 0$ and suppose that*

$$M_j = \sup_{R\lambda=a} |\lambda|^j \| R(\lambda; C) \|, \qquad j = 0, 1.$$

Suppose also that the bounded transformation C has the simple eigenvalue λ_0, $R\lambda_0 > a$, and that the half-plane $R\lambda \geqslant a$ contains no other points of the spectrum of C. Then for

$$\| U \| \leqslant \delta < \frac{2a}{2aM_0 + M_1^2} = \frac{1}{K}$$

the transformation $C + U$ *has a unique eigenvalue* $\lambda(U)$ *in the half-plane* $R\lambda \geqslant a$, *and it is simple. It satisfies*

$$|\lambda(U) - \lambda_0| \leqslant (2M_1^{-1} + 1) K \| C \| \delta \tag{10}$$

if $\delta \leqslant \| C \|$. *There are eigenvectors of* $C + U$ *and* $(C + U)^*$, *respectively, belonging to* $\lambda(U)$ *and satisfying* (9).

This time we take Ω to be the portion of the circle $|\lambda| < r$ in the half-plane $R\lambda > a$. Again the contribution of the circular part $|\lambda| = r$ to the integral

$$P_U - P_0 = \frac{1}{2\pi i}\int_{\partial\Omega} (R(\lambda; C + U) - R(\lambda; C))\, d\lambda$$

approaches zero, so that we obtain

$$P_U - P_0 = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} (R(\lambda; C + U) - R(\lambda; C))\, d\lambda.$$

From the identity

$$R(\lambda; C + U) - R(\lambda; C) = R(\lambda; C)(1 - UR(\lambda; C))^{-1}\, UR(\lambda; C)$$

it follows that

$$\| R(\lambda; C + U) - R(\lambda; C)\| \leqslant \frac{M_1{}^2\delta}{|\lambda|^2 (1 - M_0\delta)},$$

which yields

$$\| P_U - P_0 \| \leqslant \frac{M_1{}^2\delta}{2a(1 - M_0\delta)} \leqslant K\delta.$$

We set

$$D(\lambda) = R(\lambda; C + U) - R(\lambda; C).$$

From the identities

$$\lambda R(\lambda; C) = 1 + R(\lambda; C)C$$

and

$$\lambda R(\lambda; C + U) = 1 + R(\lambda; C + U)(C + U),$$

we derive

$$\begin{aligned} D(\lambda) &= \lambda^{-1}(D(\lambda)C + R(\lambda; C + U)U) \\ &= \lambda^{-1}D(\lambda)C + \lambda^{-2}(U + R(\lambda; C + U)(C + U)U). \end{aligned}$$

Consequently, we infer

$$\begin{aligned} (\lambda(U) - \lambda_0)\, P_U &= \frac{1}{2\pi i}\int_{\partial\Omega} (\lambda - \lambda_0)\, D(\lambda)\, d\lambda \\ &= \frac{1}{2\pi i}\int_{\partial\Omega} \lambda D(\lambda)\, d\lambda - \lambda_0(P_U - P_0) \\ &= J(C + U)\, U + (P_U - P_0)(C - \lambda_0), \end{aligned}$$

where

$$J = \frac{1}{2\pi i} \int_{\partial\Omega} \frac{R(\lambda; C + U)}{\lambda} d\lambda.$$

Again the integrand is $O(\lambda^{-2})$, is $|\lambda| \to \infty$, and J is independent of r for r sufficiently large. Hence we find that

$$J = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{R(\lambda; C + U)}{\lambda} d\lambda.$$

Now the estimate

$$\| R(\lambda; C + U)\| \leqslant \| R(\lambda; C)\|/(1 - M_0\delta)$$

yields

$$| J | \leqslant \frac{M_1}{2a(1 - M_0\delta)},$$

from which we conclude the inequality.

Remark. We always have the relation

$$M_1 \leqslant 1 + \| C \| M_0,$$

but sometimes we can obtain a much sharper bound for M_1.

4. Perturbation of Power-Positive Matrices

We wish now to apply the results of the previous section to power-positive matrices. For this purpose we need to estimate the resolvent of such a matrix. If C is power-positive, we shall denote by λ_C its largest eigenvalue and by ξ and η the positive eigenvectors of C and C^*, respectively, normalized by the conditions

$$\eta \cdot \xi = 1, \qquad \| \eta \| = \| \xi \| = \| P_0 \|^{1/2},$$

where $P_0 = \xi \otimes \eta$ is the projection defined by

$$P_0 x = (\eta \cdot x)\xi \qquad \text{for all } x.$$

Theorem 7. *If $C > 0$ and*

$$M(r) = \max_{|\lambda|=r} \| R(\lambda; C)\|,$$

then for $\lambda_C N_C < r < \lambda_C$, we have

$$M(r) \leqslant \mu(C)^{-1}((\lambda_C - r)^{-1} + B(r - \lambda_C N_C)^{-1}),$$

where

$$B = B(C) = 34(1 + \mu(C)^{-1}).$$

Proof. Let y be a given vector,

$$b = \eta \cdot y, \qquad z = (1 - P_0)y = y - b\xi,$$

and

$$(\lambda - C)x = y.$$

Then

$$x = (\lambda - \lambda_C)^{-1} b\xi + (\lambda - C)^{-1} z.$$

The estimate in Theorem 5 yields

$$\| R(\lambda; C)z \| \leqslant 34\mu(C)^{-1}(|\lambda| - \lambda_C N_C)^{-1} \| z \|.$$

Since Lemma 2 implies that

$$\| z \| \leqslant \| y \| + \| \xi \| \| \eta \| \| y \| \leqslant (1 + \mu(C)^{-1}) \| y \|,$$

we obtain the estimate for $M(r)$ stated above.

The minimum of $(\lambda_C - r)^{-1} + B(r - \lambda_C N_C)^{-1}$ is attained for $r = \alpha\lambda_C$, where $\alpha = (N_C + B^{1/2})/(1 + B^{1/2})$, from which we obtain

COROLLARY 7a. *For* $\alpha = (N_C + B^{1/2})/(1 + B^{1/2})$, *we have*

$$M(\alpha\lambda_C) \leqslant (1 + B^{1/2})^2/((1 - N_C)\,\mu(C)\,\lambda_C).$$

To deal with power-positive matrices, we use the identity

$$R(\lambda, C) = \left(\sum_{k=0}^{m-1} \lambda^k C^{m-1-k}\right) R(\lambda^m; C^m).$$

COROLLARY 7b. *If* $C \geqslant 0$ *and* $C^m > 0$, *and* $N = N(C^m)$, $B' = B(C^m)$, *and* $\lambda_C N^{1/m} < r < \lambda_C$, *then*

$$M(r) \leqslant \frac{\| C \|^m - r^m}{\| C \| - r} \frac{1}{\mu(C^m)} \frac{1}{\lambda_C{}^m - r^m} + \frac{B'}{r^m - \lambda_C{}^m N}.$$

For example, if $m = 2$, and $\alpha = (N + (B')^{1/2})/(1 + (B')^{1/2})$, then an easy computation yields

$$M(\alpha^{1/2}\lambda_C) \leqslant \frac{200 \| C \|}{(1 - N)(\mu(C^2)\,\lambda_C)^2}.$$

Application of Corollary 6a and Lemma 2 leads to

COROLLARY 7c. *If* $C > 0$ *and* α *is as in Corollary* 7a, *and*

$$\| U \| \leqslant \delta < 1/K,$$

where

$$K = M(1 + \lambda_C M),$$

$$M = \frac{100}{(1 - N_C)\,\mu(C)^2\,\lambda_C},$$

then $C + U$ has a unique eigenvalue $\lambda(U)$ in the set $|\lambda| \geqslant \alpha\lambda_C$, and $\lambda(U)$ satisfies

$$|\lambda(U) - \lambda_C| \leqslant (1 + 2\lambda_C K)\delta,$$

and this eigenvalue is simple. There are eigenvectors x and y of $C + U$ and $(C + U)^*$, respectively, in the spheres

$$\| x - \xi \| \leqslant K\delta \| \xi \|, \qquad \| y - \eta \|_1 \leqslant K\delta \| \eta \|_1 .$$

If $\delta < \mu(C)/K$, then $|\arg x_j| \leqslant \theta$, where

$$\sin\theta = K\delta/\mu(C), \qquad 0 \leqslant \theta < \pi/2.$$

We can apply Corollary 7b in a similar way to obtain a corresponding result for power-positive matrices.

We now wish to prove a generalization of Theorem 3. For this purpose we first derive a lemma.

LEMMA 3. *If $C > 0$, $z \neq 0$, $|\arg z_j| \leqslant \gamma < \pi/2$ for all j, and*

$$\| Cz - \mu z \| \leqslant \epsilon,$$

then

$$|\lambda_C - \mu| \leqslant \epsilon/(\| z \|\, \mu(C) \cos\gamma) = K\epsilon/\| z \|, \tag{11}$$

and

$$\| z - (\eta \cdot z)\, \xi \| \leqslant \frac{34(1 + K)\,\epsilon}{\mu(C)\,\lambda_C(1 - N_C)} = K_1\epsilon. \tag{12}$$

Proof. Let $z_j = r_j \exp(i\theta_j)$, $|\theta_j| \leqslant \gamma < \pi/2$ for all j. Since

$$\eta \cdot (Cz - \mu z) = (\lambda_C - \mu)(\eta \cdot z)$$

and

$$\begin{aligned} R(\eta \cdot z) &= \sum \eta_j r_j \cos\theta_j \\ &\geqslant m(\eta; \mathbf{1})\, \| z \| \cos\gamma, \end{aligned}$$

we obtain

$$|\lambda_C - \mu|\, m(\eta; \mathbf{1})\, \| z \| \cos\gamma \leqslant \| \eta \|_1\, \epsilon,$$

and now Lemma 2 implies (11).

Now let $u = P_0 z = z - (\eta \cdot z)\xi$, so that

$$(C - \lambda_C)u = (C - \lambda_C)z = v.$$

Hence we have

$$u = -\sum_{k=0}^{\infty} \lambda^{-k-1} C^k v$$

and, by Theorem 5, we find that

$$\| u \| \leqslant 34 \| v \| / (\mu(C)\, \lambda_C(1 - N_C)).$$

Since

$$v = (C - \mu)z + (\mu - \lambda_C)z,$$

so that

$$\| v \| \leqslant \epsilon + K\epsilon,$$

we obtain the estimate (12).

THEOREM 8. *If $C > 0$, $\| U \| \leqslant \epsilon$, and if z is an eigenvector of $C + U$ such that*

$$\| z \| = 1 \qquad \text{and} \qquad | \arg z_0 | \leqslant \gamma < \pi/2 \quad \text{for all } j, \tag{13}$$

belonging to the eigenvalue μ, then

$$| \lambda_C - \mu | \leqslant K\epsilon$$

and

$$\| z - (\eta \cdot z)\xi \| \leqslant K_1 \epsilon,$$

where K and K_1 are as in Lemma 3.

In Corollary 7a we obtain conditions on U that $C + U$ have a *unique* eigenvalue in $| \lambda | \geqslant \alpha\lambda_C$, and then find that its eigenvector is close to ξ. In Theorem 8 we find that if $C + U$ has an eigenvector satisfying (13) and U is small, then the corresponding eigenvalue is close to λ_C and the eigenvector is close to a scalar multiple of ξ. Here U is not necessarily so small that Corollary 7a applies, and so there may very well be other eigenvalues in $| \lambda | \geqslant \alpha\lambda_C$.

In the next theorem we give a sufficient condition that $C + U$ have at least one eigenvector satisfying (13) and with a positive component. Again the condition may not be strong enough to ensure uniqueness of the corresponding eigenvalue.

THEOREM 9. *If $C > 0$ and $0 < \gamma < \pi/2$ and*

$$\epsilon < \mu(C)^{1/2} \sin \gamma / (2K_1),$$

where K_1 *is as in Lemma* 3, *then for* $\| U \| \leqslant \epsilon$, *the matrix* $C + U$ *has an eigenvector* z *in the set*

$$S(\gamma) \colon z_1 = 1, \qquad | \arg z_j | \leqslant \gamma \quad (\text{all } j).$$

Proof. For $0 \leqslant t \leqslant 1$, let $C(t) = C + tU$, and let t_0 be the least upper bound of the t in [0, 1] such that $C(\tau)$ has an eigenvector in $S(\gamma)$. Then t_0 is positive by Corollary 7c. If $t_0 < 1$, then $C(t_0)$ has an eigenvector z on the boundary of $S(\gamma)$. We may assume that $| \arg z_j | = \gamma$. Let $a = \eta \cdot z$. Then by Lemma 3, we have

$$| 1 - a\xi_1 | \leqslant K_1 \epsilon \| \xi \|$$

and

$$| z_j - a\xi_j | \leqslant K_1 \epsilon \| \xi \|.$$

Consequently, we infer that

$$\begin{aligned} | \xi_1 z_j - \xi_j | &= | \xi_1(z_j - a\xi_j) + (a\xi_1 - 1)\, \xi_j | \\ &\leqslant 2K_1 \epsilon \| \xi \| \\ &\leqslant 2K_1 \epsilon\, \mu(C)^{-1/2}\, \xi_j , \end{aligned}$$

by Lemma 2. This is impossible if ϵ satisfies the above inequality.

COROLLARY 9a. *If* $A = (C_{jk} \exp(i\theta_{jk}))$, *where* $C > 0$, *and* $| \theta_{jk} | \leqslant \theta < \pi/2$ *for all* j, k *and if*

$$2 \sin(\theta/2) < \mu(C)^{1/2} \sin \gamma / (2K_1 \| C \|),$$

then A *has an eigenvector in* $S(\gamma)$.

Proof. We set $U = A - C$ in Theorem 9.

By means of these methods, we can obtain similar results for power-positive matrices.

LEMMA 4. *If* $C \geqslant 0$, $C^m > 0$, $\| z \| = 1$, $| \arg z_j | \leqslant \gamma < \pi/2$, *and* $N = N(C^m)$, *and*

$$\| Cz - \mu z \| \leqslant \epsilon,$$

then

$$| \lambda_C - \mu | \leqslant \epsilon / (\mu(C^m) \cos \gamma) = K(C^m)\epsilon,$$

and

$$\begin{aligned} \| z - (\eta \cdot z)\, \xi \| &\leqslant \left(\sum_{0}^{m-1} \frac{\| C \|^k}{\lambda_C{}^k} \right) \frac{34(1 + K(C^m))\, \epsilon}{\lambda_C \mu(C^m)(1 - N)} \\ &= K_1(C^m)\, \epsilon. \end{aligned}$$

THEOREM 10. *If* $C \geqslant 0$, $C^m > 0$, *and* $0 < \gamma < \pi/2$, *and*

$$\epsilon < \sin \gamma / (2\mu(C^m)^{1/2} K_1(C^m)),$$

then for $\| U \| \leqslant \epsilon$, *the matrix* $C + U$ *has an eigenvector in* $S(\gamma)$.

5. Matrices with Dominant Border

In this section it will be convenient to have the indices in our vectors and matrices run from 0 to N. We shall begin by determining the eigenvalues and eigenvectors of a *border matrix* C, i.e., a matrix such that $C_{jk} = 0$ for $jk \neq 0$.

If $Cz = \lambda z$, $z \neq 0$, and $\lambda \neq 0$, then for $j > 0$ we have

$$z_j = C_{j0} z_0 / \lambda$$

so that

$$\lambda^2 z_0 = \lambda \sum C_{0k} z_k = (C_{00}\lambda + d)\, z_0\,,$$

where

$$d = \sum_1^N C_{0j} C_{j0}\,.$$

Since $z \neq 0$, we must have $z_0 \neq 0$, and therefore λ is a root of the quadratic polynomial

$$Q(\lambda) = \lambda^2 - C_{00}\lambda - d,$$

and

$$\lambda = (C_{00} \pm (C_{00}^2 + 4d)^{1/2})/2.$$

Incidentally, it is easy to prove that

$$\det(\lambda - C) = \lambda^{N-1} Q(\lambda).$$

If $\lambda = 0$, then z is in the $(N-1)$-dimensional subspace

$$z_0 = 0, \qquad \sum_1^N C_{0k} z_k = 0.$$

Similarly, we easily compute

$$R(\lambda; C)\, y = (\lambda Q(\lambda))^{-1}\, x,$$

where

$$x_0 = \lambda^2 y_0 + \lambda \sum_1^N C_{0k} y_k ,$$

$$x_j = \lambda C_{j0} y_0 + C_{j0} \sum_1^N C_{0k} y_k + y_j Q(\lambda)$$

$$= \lambda C_{j0} y_0 + C_{j0} \sum{}' C_{0k} y_k + (Q(\lambda) + C_{j0} C_{0j})\, y_j ,$$

for $j > 0$. Here $\sum'$ denotes the summation over all indices $k \neq 0, j$.

If $C_{j0} \neq 0$ for all j, then it is often convenient to normalize the matrix C. We transform by the diagonal matrix A defined by

$$A_{jj} = C_{j0} , \qquad A_{jk} = 0 \qquad \text{for} \quad j \neq k,$$

and set

$$A^{-1} C A = C_{00} \tilde{C}. \tag{14}$$

If x is an eigenvector of $\tilde{C}$ belonging to the eigenvalue μ, then Ax is an eigenvector of C belonging to the eigenvalue $C_{00}\mu$. The matrix $\tilde{C}$ is a border matrix with

$$\tilde{C}_{j0} = 1 \qquad \text{and} \qquad \tilde{C}_{0j} = C_{0j} C_{j0} / C_{00}^2 \qquad \text{for all } j.$$

Hence we first focus our attention on border matrices with $C_{j0} = 1$ and $C_{0j} \geqslant 0$ for all j. There is a unique positive eigenvalue λ_1, a unique negative eigenvalue λ_2, and an eigenvalue of multiplicity $N - 1$ at 0. Let us compute the other data needed in order to apply the results of Section 3.

The positive eigenvector ξ belonging to λ_1 has the components

$$\xi_j = \xi_0 / \lambda_1 \qquad \text{for} \quad j > 0,$$

and since $\lambda_1 > 1$, we have $\| \xi \| = \xi_0$. The positive eigenvector η of C^* has the coordinates

$$\eta_j = C_{0j} \eta_0 / \lambda_1 \qquad \text{for} \quad j > 0,$$

and

$$\| \eta \|_1 = \eta_0 (1 + d/\lambda_1) = \eta_0 \lambda_1 .$$

The normalization

$$\eta \cdot \xi = 1, \qquad \| \eta \|_1 = \| \xi \|,$$

leads to

$$\eta_0 = (1 + 4d)^{-1/4}, \qquad \xi_0 = (\eta_0 + \eta_0^{-1})/2. \tag{15}$$

Finally we have

$$| \lambda Q(\lambda) |\, \| R(\lambda; C) \| = \max(| \lambda |^2 + | \lambda |\, d, | \lambda | + (d - C_{0j}) + | Q(\lambda) + C_{0j} |).$$

On the line $R\lambda = \frac{1}{2}$, the perpendicular bisector of the segment $[\lambda_2, \lambda_1]$, we have $\lambda - \lambda^2 = |\lambda|^2 \geqslant |\lambda|/2$, so that

$$|Q(\lambda)| = |\lambda|^2 + d.$$

Therefore we have

$$\frac{|\lambda| + d}{|Q(\lambda)|} \leqslant \frac{2|\lambda|^2 + d}{|\lambda|^2 + d} \leqslant 2.$$

Furthermore, we see that

$$|Q(\lambda) + C_{0j}| = |\lambda|^2 + (d - C_{0j}) \leqslant |\lambda|^2 + d.$$

It follows that

$$\begin{aligned} |\lambda| + (d - C_{0j}) + |Q(\lambda) + C_{0j}| &\leqslant |\lambda| + d + |Q(\lambda)| \\ &\leqslant 3|Q(\lambda)| \leqslant 6|\lambda Q(\lambda)|, \end{aligned}$$

and conclude that

$$\| R(\lambda; C)\| \leqslant 6 \qquad \text{for} \quad R\lambda = \tfrac{1}{2}.$$

We note that for $t > 0$, $s - 0$,

$$\begin{aligned} s^2 + sd &\leqslant s^2 + (tds^2 + t^{-1}d)/2 \\ &= ((2 + td)\, s^2 + t^{-1}d)/2. \end{aligned}$$

If we choose t as the positive solution of

$$(2 + td)t = 1, \qquad \text{i.e.,} \quad t = (-1 + (1 + d)^{1/2})/d,$$

then we obtain

$$s^2 + sd \leqslant (s^2 + d)(2 + td)/2.$$

Hence for $R\lambda = \frac{1}{2}$, we have

$$|\lambda| \, \| R(\lambda; C) \leqslant \max((1 + (1 + d)^{1/2})/2, 3) = M_1, \tag{16}$$

which is sharper than the bound $1 + 6\| C\|$ obtained from the identity

$$\lambda R(\lambda; C) = 1 + R(\lambda; C)C.$$

We can now apply Corollary 6b. A little computation yields

THEOREM 11. *If C is a nonnegative border matrix with $C_{j0} = 1$ for all j, and if*

$$\| U\| \leqslant \delta < 1/(6 + M_1^2) = 1/K,$$

then $C + U$ has a unique eigenvalue $\lambda(U)$ in the half plane $R\lambda \geqslant \frac{1}{2}$. This eigenvalue is simple, and satisfies

$$| \lambda(U) - \lambda_1 | \leqslant \varphi(\| C \|)\delta,$$

where

$$\begin{aligned} \varphi(k) &= 1 + 35k & k \leqslant 25, \\ &= (k^2/2)(1 + 12k^{-1/2}) & k > 25. \end{aligned}$$

There is an eigenvector x of $C + U$, belonging to $\lambda(U)$, such that

$$\| x - \xi \| \leqslant K\delta \| \xi \|.$$

If $\lambda_1 K\delta < 1$, then for all j we have

$$| x_j | \geqslant \frac{(1 - \lambda_1 K\delta)}{\lambda_1(1 - K\delta)} \| x \|$$

and

$$| \arg x_j | \leqslant \alpha,$$

where

$$\sin \alpha = \lambda_1 K\delta, \qquad 0 \leqslant \alpha < \pi/2.$$

There is an eigenvector y of $(C + U)^*$, belonging to $\lambda(U)$, such that

$$\| y - \eta \|_1 \leqslant K\delta \| \eta \|_1 .$$

If we apply Theorem 11 to the matrix $\tilde{C}$ defined by (14), then we obtain

COROLLARY 11a. *Suppose that C is a border matrix and that $C_{j0} \neq 0$ and $C_{j0}C_{0j}/C_{00}^2 \geqslant 0$ for all j. If V is a matrix such that*

$$\sum_0^N | V_{jk} C_{k0} | \leqslant \delta | C_{00} C_{j0} | \qquad \textit{for all } j,$$

where

$$\delta < 1/(6 + M_1{}^2) = 1/K,$$

and M_1 is defined by (16) *with*

$$d = \sum_1^N C_{0j} C_{j0} / C_{00}^2 ,$$

then $C + V$ has a unique eigenvalue $\lambda(V)$ in the half-plane $R(\lambda/C_{00}) \geqslant \frac{1}{2}$. This eigenvalue is simple. If we set

$$\lambda_1 = (1 + (1 + 4d)^{1/2})/2 \tag{17}$$

then there is an eigenvector x of $C + V$, belonging to $\lambda(V)$, such that

$$\left| \frac{x_0}{C_{00}} - 1 \right| \leqslant K\delta,$$

and

$$\left| \frac{x_j}{C_{j0}} - \frac{1}{\lambda_1} \right| \leqslant K\delta \qquad \textit{for} \quad j > 0.$$

If $\lambda_1 K\delta < 1$, *and*

$$\sin \alpha = \lambda_1 K\delta, \qquad 0 \leqslant \alpha < \pi/2,$$

then

$$| \arg(x_j/C_{j0}) | \leqslant \alpha \qquad \textit{for all } j.$$

There is an eigenvector y of $(C + V)^*$, belonging to $\lambda(V)$, such that

$$| C_{00} y_0 - 1 | + \sum_1^N \left| C_{k0} y_k - \frac{C_{0k} C_{k0}}{\lambda_1 C_{00}^2} \right| \leqslant \lambda_1 K\delta.$$

Note that x and y satisfy

$$\max_j \left| \frac{x_j}{C_{j0}} \right| \leqslant \frac{(1 + K\delta)\, \lambda_1}{(1 - \lambda_1 K\delta)} \min_j \left| \frac{x_j}{C_{j0}} \right|,$$

and

$$\lambda_1(1 - K\delta) \leqslant \sum_0^N | C_{0k} y_k | \leqslant \lambda_1(1 + K\delta).$$

By an easy limiting process, or by imitating the above argument, we can obtain extensions of this result to certain infinite-dimensional spaces.

Let l_1 be the Banach space of absolutely convergent series y with the norm

$$\| y \|_1 = \sum_0^\infty | y_j |,$$

and let $l_\infty = l_1^*$ be its conjugate space, the set of all bounded sequences x with the norm

$$\| x \|_\infty = \sup | x_j |.$$

If $\{a_n\}$, $n \geqslant 0$, is a sequence of nonzero complex numbers, and A is the diagonal transformation defined by

$$(Ax)_j = a_j x_j \qquad \text{for all } j,$$

then we may denote by Al_1 and Al_∞, respectively, the transforms of l_1 and l_∞ by A.

COROLLARY 11b. *Suppose that C is a border matrix and that $C_{j0} \neq 0$ and $C_{j0}C_{0j}/C_{00}^2 \geqslant 0$ for all $j \geqslant 0$, and that*

$$0 < d = \sum_1^\infty C_{0j}C_{j0}/C_{00}^2 < +\infty.$$

Let A be defined by $(Ax)_j = C_{j0}x_j$ for $j \geqslant 0$.
Then if V is a matrix such that

$$\sum_0^\infty | V_{jk}C_{k0} | \leqslant \delta | C_{00}C_{j0} | \qquad (all\ j \leqslant 0),$$

where

$$\delta < 1/K, \qquad K = 6 + M_1^2,$$

then $C + V$, as a linear transformation on Al_∞, has a unique eigenvalue $\lambda(V)$ in the half plane $R(\lambda/C_{00}) \geqslant \frac{1}{2}$. This eigenvalue is simple, and is also a simple eigenvalue of $(C + V)^*$ on $A^{-1}l_1$. If $\lambda_1 K\delta < 1$, then there is an eigenvector x of $C + V$, belonging to $\lambda(V)$ such that

$$| \arg(x_j/C_{j0})| \leqslant \alpha \qquad for\ all\ j,$$

where

$$\sin \alpha = \lambda_1 K\delta, \qquad 0 \leqslant \alpha < \pi/2.$$

We can also obtain results like Theorem 9, which may not be strong enough to imply uniqueness. Suppose that

$$C_{jk} = r_{jk} \exp(i\theta_{jk}), \qquad r_{jk} \geqslant 0,$$

and

$$| \theta_{jk} | \leqslant \theta < \pi/2 \qquad \text{for all } j, k, \tag{18}$$

and that

$$\max_{j\geqslant 1, k\geqslant 0} (r_{jk}/r_{0k}) = a, \tag{19}$$

$$\max_{j\geqslant 0} \sum_1^N r_{jk}/r_{j0} = b. \tag{20}$$

Then a and b are measures of the dominance of the border of the matrix C.

THEOREM 12. *If C satisfies Eqs.* (18)–(20), *and*

$$\begin{gathered} \theta + \gamma < \pi/2, \qquad 0 < \gamma, \\ \sin((\gamma/2) + 2\theta) < \sin(3\gamma/2) - 2ab, \end{gathered} \tag{21}$$

then C has an eigenvector z in the convex set

$$S(\gamma)\colon z_0 = 1, \qquad |\arg z_j| < \gamma \quad (all\ j).$$

Remarks. If $a = \theta = 0$, then C is a power-positive border matrix ($C^2 > 0$). If $s = \sin(\gamma/2)$, then

$$\begin{aligned}\varphi(\gamma) &= \sin(3\gamma/2) - \sin(\gamma/2)\\ &= 2s\cos\gamma = 2(s - 2s^3) > 0\end{aligned}$$

for $0 < \gamma < \pi/2$. Hence (16) is satisfied for sufficiently small θ and a. Thus (21) defines a class of matrices close to power-positive border matrices, which are sure to have an eigenvector in $S(\gamma)$. For γ close to 0, our result may be considered a perturbation of the Perron–Frobenius result, while if γ is close to $\pi/2$, the result is related to Theorem 3. Note also the maximum of φ is attained for

$$\sin(\gamma/2) = 1/6^{1/2}$$

and is $(8/27)^{1/2}$. Thus if $2ab < (8/27)^{1/2}$, then (21) is satisfied for some γ and for all sufficiently small θ.

Proof. For given γ, θ, and r_{jk} ($j, k \geqslant 0$) satisfying (21), let θ_0 be the least upper bound of the numbers such that

$$0 \leqslant \theta_0 \leqslant \theta$$

and such that every matrix C with $|\theta_{jk}| \leqslant \theta_0$ (all j, k) has an eigenvector in $S(\gamma)$. Then $\theta_0 > 0$, and if $\theta_0 < \theta$, then there is a matrix C with $|\theta_{jk}| \leqslant \theta_0$ (all j, k) having an eigenvector z on the boundary of $S(\gamma)$.

Let

$$z_k = \rho_k \exp(i\alpha_k), \qquad \rho_k \geqslant 0,$$

for $k \geqslant 0$. We may assume that $\alpha_0 = 0$, and $\alpha_j = \gamma$, $l_j > 0$, and $|\alpha_k| \leqslant \gamma$ for $k > 0$. For any $m > 0$, we have

$$|\lambda|\,\rho_m \leqslant \sum_0^N r_{mk}\rho_k \leqslant a\sum_0^N r_{0k}\rho_k\,.$$

But

$$\begin{aligned}|\lambda| \geqslant R\lambda &= R\left(\sum_0^N C_{0k}z_k\right)\\ &\geqslant \sum r_{0k}\rho_k\cos(\theta + \gamma),\end{aligned}$$

and this implies that

$$\rho m \leqslant a/\cos(\theta + \gamma)$$

for $m > 0$.

We set

$$Z_m = \sum_1^N C_{mk} z_k ,$$

so that

$$\lambda = C_{00} + Z_0 ,$$
$$\lambda z_j = C_{j0} + Z_j .$$

The estimate

$$| Z_m | \leqslant \sum_1^N r_{mk} \rho_k \leqslant abr_{m0}/\cos(\theta + \gamma)$$

yields

$$| \arg(1 + z_m/C_{m0})| \leqslant \beta,$$

where

$$\sin \beta = ab/\cos(\theta + \gamma), \qquad 0 < \beta < \pi/2.$$

Consequently, we obtain

$$\begin{aligned}\gamma = \arg z_j &= \arg((C_{j0} + Z_j)/(C_{00} + Z_0)) \\ &\leqslant \theta + \beta - (-\theta - \beta) = 2(\theta + \beta),\end{aligned}$$

or

$$\sin((\gamma/2) - \theta) \leqslant \sin \beta,$$

that is

$$2 \sin((\gamma/2) - \theta) \cos(\theta + \gamma) < 2ab.$$

But the left-hand side is

$$\sin(3\gamma/2) - \sin(2\theta + (\gamma/2)),$$

so this inequality contradicts (21).

COROLLARY 12a. *If C satisfies the conditions*

$$\begin{aligned} C_{j0} &\neq 0 && (j \geqslant 0),\\ | \arg((C_{jk}C_{k0})/(C_{00}C_{j0}))| &\leqslant \theta && (j, k \geqslant 0),\\ | C_{00} | \, | C_{jk} | &\leqslant a \, | C_{0k} | \, | C_{j0} | && (j > 0,\ k \geqslant 0),\end{aligned}$$

and

$$\sum_1^N | C_{jk}C_{k0} | \leqslant b \, | C_{00} | \, | C_{j0} | \qquad (j \geqslant 0),$$

and condition (20), *then C has an eigenvector z such that*

$$z_0 = C_{00} , \qquad | \arg(z_j/C_{j0})| < \gamma \qquad (j > 0).$$

Clearly Theorem 12 and its corollary can be extended in the obvious way to certain infinite-dimensional spaces.

REFERENCES

1. R. BELLMAN, "Introduction to Matrix Analysis," McGraw–Hill, New York, 1960.
2. G. BIRKHOFF, Extensions of Jentzsch's theorem, *Trans. Amer. Math. Soc.* **85** (1957), 219–227.
3. A. BRAUER, On the characteristic roots of power-positive matrices, *Duke Math. J.* **27** (1961), 439–446.
4. A. BRAUER, On the characteristic roots of non-negative matrices, Ref. [19, pp. 3–38].
5. H. BUSEMANN AND P. J. KELLY, "Projective Geometries and Projective Metrics," Academic Press, New York, 1953.
6. G. FROBENIUS, Über Matrizen aus positiven Elementen, *Sitzungsber. Kgl. Preuss. Akad. Wiss.* (1908), 471–476; and (1909), 514–518.
7. G. FROBENIUS, Über Matrizen aus nicht-negativen Elementen, *Sitzungsber. Kgl. Preuss. Akad. Wiss.* (1912), 456–477.
8. D. HILBERT, Neue Begründung der Bolyai-Lobatschefskyschen Geometrie, *Math. Ann.* **57** (1903), 137–150.
9. E. HILLE AND R. S. PHILLIPS, "Functional Analysis and Semi-Groups," Colloquium Publ., Vol. 31, Amer. Math. Soc., Providence, R.I., 1957.
10. E. HOPF, An inequality for positive linear operators, *J. Math. Mech.* **12** (1963), 683–692.
11. A. S. HOUSEHOLDER, Localization of the characteristic roots of matrices, Ref. [19, pp. 39–60].
12. R. JENTZSCH, Über Integralgleichungen mit positiven Kern, *J. Math.* **141** (1912), 235–244.
13. L. V. KANTOROVICH, B. Z. VULICH, AND A. G. PINSKER, Functional analysis in semi-ordered spaces (Russian), *Gos. Izdat. Tekhniko-Teoret. Lit., Moscow*, 1950.
14. T. KATO, "Perturbation Theory for Linear Operators," Springer, Berlin, 1966.
15. M. G. KREIN AND M. A. RUTMAN, Linear operators leaving invariant a cone in a Banach space, *Uspehi Mat. Nauk* **23** (1948), 3–95. English translation: Amer. Math. Soc. Translation No. 26.
16. A. M. OSTROWSKI, Positive matrices and functional analysis, Ref. [19, pp. 81–101].
17. O. PERRON, *Zm* Theorie der Matrices, *Math. Ann.* **64** (1907), 248–263.
18. P. C. ROSENBLOOM, Perturbation of linear operators in Banach spaces, *Arch. Math. (Basel)* **6** (1955), 89–101.
19. H. SCHNEIDER (Ed.), "Recent Advances in Matrix Theory," Univ. of Wisconsin Press, Madison, 1964.
20. E. SENETA, "Non-Negative Matrices, An Introduction to Theory and Applications," Wiley, New York, 1973.