PACIFIC JOURNAL OF MATHEMATIC
Vol. 26, No. 3, 1968

ERGODIC PROPERTIES OF NONNEGATIVE MATRICES-II

D. VERE-JONES

This paper develops some properties of matrices which have nonnegative elements and act as bounded operators on one of the sequence spaces l_p or $l_p(\mu)$, where μ is a measure on the integers. Its chief aim is to relate the operator properties of such matrices to the matrix properties (the value of the convergence parameter R, R--recurrence and R-positivity) described in detail in Part I. The relationships between the convergence norm (equal to $1/R$), the spectral radius, and the operator norm are discussed, and conditions are set up for the convergence norm to lie in the point spectrum. It is shown that the correspondence between matrix and operator properties depends very much on the choice of the underlying space. Other topics considered are the problem of characterizing the operator norm in matrix terms, and a general theorem on the structure of a positive contraction operator on l_p.

The notation, terminology and results of Part I of this paper will be assumed without further comment. Reference to equations and theorems from that part will be in the form I (13); I, Theorem 4.1. etc.

The symbol $l_p (1 \leqq p < \infty)$ will be used throughout to denote the (complex) Banach space, the elements of which are sequences of complex numbers $\{x_k\}$ satisfying $|x_k|^p < \infty$, with norm

$$||x||_p = (\Sigma_k |x_k|^p)^{1/p} ,$$

termwise addition, and scalar multiplication. More generally, if μ is a fixed finite or sigma-finite measure on the integers, $l_p(\mu)$ $(1 \leqq p < \infty)$ will denote the Banach space of sequences satisfying $\Sigma |x_k|^p \mu_k < \infty$ and with norm $\{\Sigma |x_k|^p \mu_k\}^{1/p}$. By l_∞ (respectively $l_\infty(\mu)$) we shall denote the Banach space of bounded sequences (respectively sequences satisfying $\sup_k |x_k/\mu_k| < \infty$) with norm $||x||_\infty = \sup_k |x_k|$ (respectively $\sup_k |x_k/\mu_k|$). It is well known that any bounded linear operator on an l_p space $(1 \leqq p < \infty)$ is determined by an infinite dimensional matrix, through the equations

$$(Ax)_k = \sum_i x_i a_{ik}$$

(note that we have chosen the matrix to act on the left). The Banach adjoint A^*, acting on the dual space l_q $(1/p + 1/q = 1)$, is then represented by the transposed matrix:

$$(A^*y)_k = \sum_j a_{kj} y_j .$$

The operator norm is defined by

$$|| A || = \sup_{||x||=1} \frac{|| Ax ||}{|| x ||}$$

and we shall leave it for the context to make clear the underlying space on which the operator is acting, and with reference to which the operator norm is defined. We have always $|| A || = || A^* ||$. A is said to be a *contraction* if $|| A || \leqq 1$.

A thorough discussion of these ideas is contained in Part I of the monograph by Dunford and Schwartz [3], to which we also refer the reader for the notions of resolvent operator, resolvent set, spectrum, spectral radius, point spectrum, and mean ergodic theorem.

An operator on an l_p space $(1 \leqq p < \infty)$ is *positive* if the elements in its matrix representation are nonnegative. More generally, a positive operator can be defined as one which leaves invariant a suitably defined positive cone. It is well-known that for a very wide class of positive operators, which include all positive operators on l_p spaces $(1 \leqq p < \infty)$ as defined above, the spectral radius is a point in the spectrum—i.e., the positive real axis contains a point in the spectrum of maximum modulus. For a discussion of this proposition see (for example) Bonsall [1] and Schaeffer [13]. It may be regarded as a rather weak generalization of the Perron-Frobenius theory to positive operators. From the same point of view, the present paper is an attempt to utilize the theory developed in Part I to investigate some ways of strengthening this result.

2. Spectral properties of positive operators on l_p spaces.

Suppose that the operator A on l_p can be represented by an irreducible matrix with nonnegative elements and convergence parameter $R > 0$. (We should remark that the choice of l_p, rather than any of the other spaces $l_p(\mu)$, is a matter of convenience only. All of the results in this section hold equally for an operator on $l_p(\mu)$, as can be seen if only by mapping $l_p(\mu)$ back onto l_p by the one-to-one isometric isomorphism which takes $\{y_j\}$ in $l_p(\mu)$ into $\{y_j \mu_j^{1/p}\}$ in l_p. The essential point is that throughout this section it is the operator, and by implication the underlying space, which is assumed given). We start by examining the relationships which hold between the convergence norm $1/R$ of A and its spectral radius.

THEOREM 2.1. *If A is a bounded linear operator on $l_p(1 \leqq p \leqq \infty)$ with spectral radius ρ, and if A can be represented by an irreducible, nonnegative matrix with convergence parameter R, then*

$$1/R \leqq \rho \tag{1}$$

where equality and strict inequality are both possible. In the case of strict inequality, the interval $1/R \leqq \lambda \leqq \rho$ *of the positive axis is entirely contained in the spectrum of* A.

Proof. Since $1/R$ is a singularity of the function $(1/2)A_{ij}(1/2) = \Sigma a_{ij}^{(n)}/z^{n+1}$, which can be identified with the $i - j$ element in the matrix representation of the resolvent operator of A (Equation I. (6)), it follows that, a fortiori, $1/R$ is a singularity of the resolvent operator itself. The converse is not true, however. To show this we give an example of an irreducible matrix defining a bounded linear operator whose spectral radius is not a singularity of the resolvent elements. Let A be defined by the matrix

$$A = \begin{pmatrix} f_1 & f_2 & f_3 & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot \end{pmatrix}$$

which can be regarded as a modified shift operator. If $\Sigma f_i \leqq 1$, the matrix defines a bounded linear operator on l_1 whose norm and spectral radius are both equal to unity (this is most easily seen by considering the adjoint: if e is the vector all of whose components are equal to unity, we have $\|(A^*)^n e\| = 1$ for all n, so that the spectral radius is at least equal to unity; yet the matrix defines a contraction). On the other hand, it is easily verified that

$$T_{11}(z) = \frac{1}{1 - F(z)} \tag{2}$$

(where $F(z) = \Sigma f_i z^i$), so that if $\Sigma f_i < 1$, and $F(z)$ is analytic in a disc $|z| < 1 + \alpha\,(\alpha > 0)$, then unity is not a singularity of the resolvent elements. Slight modifications yield similar counterexamples for the l_p spaces with $1 < p < \infty$ (for details see [17], § 5.3).

This difficulty seems first to have been pointed out by Putnam [10]; the counterexample given there is not irreducible, however.

The final statement of the theorem is rather deeper in character, and is one of the few places where we have been able to apply results from the general theory of operators leaving invariant a positive cone. We make use of the following lemma, due to Schaeffer [13].

LEMMA 2.1. *Suppose that the positive cones in the Banach space* X *and its dual are both normal*[1], *and that* A *is a positive operator on* X *with spectral radius* $\rho > 0$; *then the resolvent operator* $R(\lambda; A)$ *is positive if and only if* λ *is real,* $\lambda > \rho$. (Note that the resolvent

(1) A cone K is *normal* if for $x \in K, y \in K$, $\|x + y\| \geqq \|x\|$, or more generally $\|x + y\| \geqq \alpha \|x\|$ for some $\alpha > 0$.

operator is defined only for λ in the resolvent set of A).

We now observe that the elements of the matrix representation of $R(\lambda; A)$ are real and positive whenever $\lambda \geqq 1/R$. This would lead to a contradiction with the necessary and sufficient condition of the lemma if it was possible to have $1/R < \lambda < \rho$, and λ a point in the resolvent set. Thus all points in the open interval $(1/R, \rho)$ must lie in the spectrum; since the spectrum is a closed set, the end-points of the interval must lie in the spectrum also.

This concludes the proof of the theorem.

We pass on next to a discussion of the spectral character of the point 1/R and its relationship to the classification of the matrix as R-positive, R-null, or R-transient. It might be conjectured that R-positivity was always associated with the situation where $1/R$ belonged to the point spectrum, but in fact this is not so and no completely general statements of this kind seem possible.

A few examples will illustrate the sorts of difficulty that may arise. In § 5 of [14] it is shown that if A is the irreducible part of a subcritical branching-process matrix, then under mild conditions the matrix is R-positive ($R > 1$) and every point in the range $1 < r \leqq R$ is associated with an r-invariant probability distribution. Since the matrix is substochastic, it defines a contraction operator on l_1. In this case, therefore, we have an example of an R-positive matrix for which the interval ($1/R \leqq \lambda < 1$) lies in the point spectrum. On the other hand, for a supercritical branching process, the reduced matrix of transition probabilities is again R-positive, with $R > 1$, and again each point in the range $1 < r \leqq R$ is associated with an r-invariant vector, but in this case the invariant vectors are no longer summable, so that the corresponding points $\lambda = 1/r$ are no longer in the point spectrum.

Similar difficulties arise in the case of R-null and R-transient matrices. The semi-infinite random walk considered in § 6 of [14] furnishes an example of an R-transient substochastic matrix for which $R > 1, \rho = 1$, and $1/R$ is in the point spectrum of the operator. A variety of further examples, which can be constructed along the lines of the matrix in § 2, show that the same sorts of situation arise also for operators on l_p with $p > 1$.

Such examples make it clear that the best we can hope for is to find reasonably simple conditions on the operator or the matrix that will ensure a better correspondence between matrix and operator properties. One natural condition which suggests itself is that the operator have bounded iterates. Even here, no real progress can be made unless we assume also that the convergence norm and the spectral radius coincide. Under these two assumptions, however, a much more satisfactory situation prevails. We start by examining the R-positive case.

THEOREM 2.2. *Suppose that A is a positive operator on $l_p(1 \leqq p < \infty)$, with spectral radius equal to unity and bounded iterates, and that its matrix representation is irreducible with period $d \geqq 1$. Then the two statements*

(i) *unity is in the point spectrum of both A and A^*;*

(ii) *the convergence parameter R equals unity, and the matrix is R-positive;*

are equivalent, and either implies that

(iii) *each of the points $e^{2\pi i m/d}$ $(m = 1, 2, \cdots d)$ is in the point spectrum of both A and A^*;*

(iv) *each of these eigenvalues has index unity and is associated with unique eigenvectors for both A and A^*;*

(v) *A has no further eigenvalues on the boundary of the unit disc.*

Proof. We prove first (i) and (ii) are equivalent, then establish the remaining properties.

The main difficulty in proving (i) $\rightarrow$ (ii) lies in the fact that we cannot assume the eigenvectors to be nonnegative. (i) supplies us only with the weaker information that unity is associated with left and right superinvariant vectors. To get over this difficulty we shall make use of the following property of superinvariant vectors.

LEMMA 2.2. *If T is a nonnegative matrix, $\{u_k\}$ is a left superinvariant vector, and $u_i = 1$, then for all j and for all n,*

$$\sum_k u_k t_{kj}^{(n)} \geqq \sum_{r=1}^{n} l_{ij}^{(r)} . \tag{3}$$

The lemma can be proved by induction in a similar manner to I, Lemma 4.1, and we omit details. A similar result holds for right superinvariant vectors, with the obvious changes.

Returning to the main proof, let $\{\alpha_k\}$, $\{\beta_k\}$ denote the eigenvectors for A, A^*, respectively, referred to in (i), and put $u_k = |\alpha_k|$, $t_{kj}^{(n)} = a_{kj}^{(n)}$ in the lemma above. From the assumption that A has bounded iterates, the left hand side of (3) remains bounded as $n \rightarrow \infty$. Hence the sum $L_{ij}(1)$ is convergent. Further, multiplying (3) by $v_j = |\beta_j|$ and summing over j, we obtain

$$\sum_j v_j \left\{ \sum_{d=1}^{n} l_{ij}^{(r)} \right\} \leqq \| \alpha \|_p . \| \beta \|_q . \| T^n \| . \tag{4}$$

Letting $n \to \infty$ in this expression, we see that the sum $\Sigma_j L_{ij}(1) v_j$ is convergent. But the vector $\{L_{ij}(1)\}$ (i fixed) is a left subinvariant vector, and $\{v_k\}$ is a right superinvariant vector, so that I, Lemma 5.2 and I, Criterion III (following I, Lemma 5.3) apply, and we deduce that the matrix representing A is R-positive, with $R = 1$.

To prove the converse, we show that the left and right invariant vectors (whose existence follows from the assumption of R-positivity) belong to l_q and l_p, respectively. Denote these vectors by $\{\alpha_k\}$, $\{\beta_k\}$ respectively, and let $\{y_k\}$ be any vector in l_q. Then, from the assumption of bounded iterates,

$$\sum_k l_{ik}^{(n)} | y_k | \leqq \| T^n \| \, \| y \|_q \leqq K . \| y \|_q < \infty$$

where K is independent of n. Taking $C - 1$ averages, letting $n \to \infty$, and using the fact (I, Theorem D) that $C - 1 \lim_{n\to\infty} a_{ik}^{(n)} = \alpha_k \beta_i / \Sigma \alpha_j \beta_j$, we obtain from Fatou's lemma

$$\sum_k \alpha_k | y_k | < \infty . \tag{5}$$

Since $\{y_k\}$ is an arbitrary l_q vector, it follows from (5) by a standard argument that $\{\alpha_k\}$ is an l_p vector.

If $1 < p < \infty$ a dual argument shows $\{\beta_k\}$ is an l_q-vector. For $p = 1, q = \infty$ we have

$$\sup_k \left[\lim_{n\to\infty} \frac{1}{n} \sum_{r=1}^{n} a_{ik}^{(r)} \right] \leqq \limsup_{n\to\infty} \left\{ \sup_k \frac{1}{n} \sum_{r=1}^{n} a_{ik}^{(r)} \right\} \leqq \limsup_{n\to\infty} \| A^n \| < \infty$$

which, letting $n \to \infty$, shows directly that $\sup \beta_k < \infty$, so that $\{\beta_k\}$ is an l_∞ vector.

This completes the proof that (i) and (ii) are equivalent. The remaining properties are almost immediate corollaries of the results established in Part I, in particular I, Theorem 7.1. Note that if $\{\beta_k\}$ is the right eigenvector described above, and $\{x_k\}$ is any left eigenvector (used in the strict sense implying that the vectors belong to the appropriate spaces) then the condition $\Sigma | x_k | \beta_k < \infty$ is satisfied. This remark takes care of (iii), (v), and the second statement in (iv). The fact that the eigenvalues have index unity is a consequence of bounded iterates. If $(I - T)^2 \boldsymbol{x} = 0$ for some vector $\boldsymbol{x}$, then

$$T^n \boldsymbol{x} = \boldsymbol{x} - n(I - T)\boldsymbol{x} .$$

Dividing this equation by n and letting $n \to \infty$ it follows also that $(I - T)\boldsymbol{x} = 0$. (cf. [3], Lemma VIII. 8.1). (In fact the property can be deduced from weaker assumptions by making use of the positivity.) This completes the proof of the theorem.

There are various directions in which Theorem 2.2 can be strengthened and complemented, and we pass on now to a consideration of some of these.

The first point to note is that it is sufficient in (i) to have unity in the point spectrum of A, rather than in the point spectra of both A and A^*. To see this, we consider separately the cases $p = 1$ and $p > 1$. In the case $p = 1$, the argument from (4) on carries through as before if we put $v_j = 1$ in (4), since this vector both belongs to l_∞ and is subinvariant. For $p > 1$ we can obtain the same result from the mean ergodic theorem ([3], Theorem VIII. 5.1) which applies here because the spaces $l_p(1 < p < \infty)$ are reflexive, so that the unit ball is weakly sequentially compact. In this case, if unity is in the spectrum of either A or A^*, the ergodic limit to which the operators converge cannot be zero, so that the sequences $a_{ij}^{(n)}$ cannot tend to zero, and the matrix for A is necessarily 1-positive.

If the matrix for A is 1-null or 1-transient, similar arguments show that no points on the boundary of the unit disc can belong to the point spectrum. Thus if $p = 1$ and $\{\alpha_k\}$ is an eigenvector for A associated with an eigenvalue of unit modulus, we can put $\beta_k = 1$ in I, Theorem 6.1 to justify taking $C - 1$ limits under the summation sign in the expression $a_j e^{-n\theta} = \Sigma_k \alpha_k t_{kj}^{(n)}$. This leads to a contradiction if we assume that the matrix representing A is null-recurrent or transient. If $p > 1$, a similar contradiction follows directly from the mean ergodic theorem.

These remarks establish the following result.

THEOREM 2.3. *Under the assumptions of Theorem* 2.2, *the matrix representing A is* 1-*positive if and only if unity is in the point spectrum of A. If the matrix representing A is either* 1-*null or* 1-*transient, A has no eigenvalues of unit modulus.*

Theorems 2.2 and 2.3 form an extension of the work of Šidák [15] and Holmes [4], who establish similar results for the special case that A is an operator associated with a Markov chain (see the next section for further discussion).

An interesting unresolved problem is to find further conditions on the matrix which are necessary and sufficient to ensure that the operator it represents has unity not only in its point spectrum, but as a pole of the resolvent operator. It is shown in [18] that for $p = 1$ "geometric ergodicity" (i.e., the convergence radius is a pole of each of the resolvent elements $(1/2)A_{ij}(1/2)$) is a necessary but not sufficient condition. Whether geometric ergodicity is a sufficient condition when $p > 1$ remains an open question. More easily established is the property, which holds also for more general classes of positive operators (see [8]), that if the spectral radius is a pole, and the matrix irreducible, then the pole is necessarily simple.

The results contained in part 1 also have implications for the ergodic behaviour of operator iterates. Suppose that the conditions of Theorem 2.2 hold, so that in particular the operator iterates are uniformly bounded. Since the coordinate vectors are dense in the l_p spaces ($1 \leqq d < \infty$), and also in the space c_0 of bounded sequences, the element-wise convergence described by I, Theorem D can be extended to weak convergence (c_0-weak convergence in the case $p = 1$) by a double application of the Banach-Steinhaus theorem ([3], Theorem II-3.6).

This result is not a corollary of the mean ergodic theorem, because, taking as a particular example the case when the matrix representing A is aperiodic, it is convergence of the iterates themselves which is asserted, and not convergence of the $C - 1$ averages. However, the fact that the mean ergodic theorem asserts *strong* convergence of the $C - 1$ averages leads one to hope that when the matrix is irreducible and aperiodic, the $C - 1$ averaging can be dropped without upsetting the strong convergence. We have not been able to obtain such an extension in general, but we can prove that strong convergence holds whenever the conditions of the Theorem 2.2 are satisfied and in addition A is a contraction and its matrix is 1-positive. For the case $p = 1$, this follows from I, Theorem 6.2, and was proved there as a corollary. The case $p > 1$ can be obtained from a recent result of Mrs. S.C. Moy's [7] concerning associated operators for Markov chains. It will be convenient to leave the proof of this assertion until after the discussion of associated operators in the next section. The results concerning the convergence of operator iterates are then summarized in the following theorem. (For convenience we restrict attention to the aperiodic case; for periodic matrices corresponding results can be formulated in terms of convergence as $n \longrightarrow \infty$ along a suitable residue class, and are readily deduced from the theorems in I, § 7.)

THEOREM 2.4. *Suppose that A satisfies the conditions of Theorem 2.2, and that the matrix representing A is aperiodic. Then as $n \longrightarrow \infty$ the operator iterates A^n converge weakly (c_0-weakly if $p = 1$) to their ergodic limit and their $C - 1$ averages converge strongly to the same limit. If, in addition, the operator is a contraction and its matrix is* 1*-positive, the iterates themselves tend strongly to their ergodic limit.*

To end this section we add a few further remarks about the case when A is a contraction operator. This condition imposes remarkably stringent requirements on the left and right eigenvectors (for eigenvalue unity) when they exist. Suppose that A is 1-positive, so that the ergodic limit is nonzero; in fact its elements are given by (using I, Theorem D)

$$\lambda_{ij} = \beta_i \alpha_j / \Sigma \alpha_k \beta_k \tag{6}$$

where $\{\alpha_k\}$, $\{\beta_k\}$ are the left and right invariant vectors.

From the fact that A is a contraction it follows readily that the ergodic limit Λ is also a contraction. Then, substituting from (6),

$$1 \geqq \| \Lambda \| = \sup_{\|x\|_p=1} \| \Lambda x \|_p = \sup_{\|x\|_p=1} \frac{\| \alpha \|_p}{\Sigma \alpha_j \beta_j} | \Sigma x_j \beta_j | .$$

Now the supremum on the extreme right is attained when $\{x_j\}$ is proportional to $\{\beta_j^{q-1}\}$, and $\Sigma_j x_j \beta_j = \| \beta \|_q$. Thus the above inequality leads to

$$\Sigma \alpha_j \beta_j \geqq \| \alpha \|_p . \| \beta \|_q .$$

This shows that in fact both sides must be equal and we have the special case of equality in Hölder's inequality. It follows that the left and right eigenvectors must be linked by the equations

$$\beta_j = \alpha_j^{q-1} \qquad (j = 1, 2, \cdots) \tag{7}$$

so that the right eigenvector is completely determined by the left eigenvector (and vice versa if $1 < p < \infty$).

This result may also be regarded as an extension of a well-known theorem of Riesz and Nagy ([11]; see also [12]) to the effect that if A is a Hilbert contraction, and A^H its Hilbert-adjoint, then the eigenvectors of A and A^H coincide. The properties of positivity and irreducibility are irrelevant to this theorem, as is the sequential nature of the underlying space, and in fact the result we have just proved has a quite general statement as follows (cf. [17], Lemma 4.7).

THEOREM 2.5. *Let A be a contraction operator on a Banach space X whose unit ball is smooth*[2]*; then for every eigenvector $\boldsymbol{x}$ of A, associated with an eigenvalue of unit modulus, there is an eigenvector $\boldsymbol{y}$ of A^*, associated with the same eigenvalue, and given by the unique (up to constant factors) norm-determining element for $\boldsymbol{x}$ in X^*.*

[2] The unit ball in a Banach space is said to be smooth if there is a unique supporting hyperplane at every point, i.e., if for every $\boldsymbol{x}$ in X, there is a unique linear functional $\boldsymbol{y}$ in X^* such that $\| \boldsymbol{y} \| = 1$ and $| (\boldsymbol{y}, \boldsymbol{x}) | = \| \boldsymbol{x} \|$. The symbol $(\boldsymbol{y}, \boldsymbol{x})$ is used here, and in the proof above, to denote the value taken by the linear functional $\boldsymbol{y}$ at the vector $\boldsymbol{x}$.

Proof. Since A is a contraction, we have

$$\begin{aligned}\|\boldsymbol{x}\| = (\boldsymbol{y}, \boldsymbol{x}) = (\boldsymbol{y}, \lambda A\boldsymbol{x}) &= (\lambda^{-1}A^*\boldsymbol{y}, \boldsymbol{x}) \\ &\leqq \|\lambda^{-1}A^*\boldsymbol{y}\|\cdot\|\boldsymbol{x}\| \\ &\leqq |\lambda^{-1}|\cdot\|A^*\|\cdot\|\boldsymbol{y}\|\cdot\|\boldsymbol{x}\| \\ &\leqq \|\boldsymbol{x}\|\end{aligned}$$

where λ is the eigenvalue (of unit modulus) for which $\boldsymbol{x}$ is an eigenvector, and we have chosen $\boldsymbol{y}$ to have unit norm. It follows from this chain of equalities and inequalities that in fact equality must hold at every stage, so that

$$(\lambda^{-1}A^*\boldsymbol{y}, \boldsymbol{x}) = 1$$

and

$$\|\lambda^{-1}A^*\boldsymbol{y}\| = 1 .$$

Now these two equations assert that $\lambda^{-1}A^*\boldsymbol{y}$ is a norm-determining element for $\boldsymbol{x}$. We have assumed, however, that to every element in X there is a *unique* norm-determining element in X^*, and so $\boldsymbol{y} = \lambda A^*\boldsymbol{y}$. This proves the theorem.

3. Associated operators. In the previous section, the operator, rather than the matrix, was taken as the given quantity, so that there was no question of altering the underlying space on which the matrix was assumed to act. In this section we take up the opposite point of view; given a nonnegative matrix, on which spaces does it induce a bounded operator, and what are the spectral properties of the operators so induced ? Many of the difficulties described in the previous section stem from an inappropriate choice of the underlying space; conversely, the main result of the present section is that (for an irreducible matrix) it is always possible to find a σ-finite measure on the integers, say $\mu_p(\,.\,)$, such that, after suitable normalization, the matrix defines a bounded linear operator on the space $l_p(\mu_p)$ with norm, spectral radius and convergence norm all equal to unity.

The idea behind this result appears to go back to Nelson [9], who observed that the transition function $P(x, A)$ of a Markov process with subinvariant measure $\mu(\,.\,)$ defines a contraction operator on each of the spaces $L_p(\mu)$. This observation has been used for a variety of purposes by (among others) Kendall [6], Šidák [15], Holmes [4], Moy [7] and the author [18]. It extends also to the present context, with the extra complication that the measure $\mu(\,.\,)$ must be chosen separately for each p. Because of this it is rather more convenient to standardize the sequence space than the matrix, so that instead of the same matrix T acting as an operator on each of a range of spaces $l_p(\mu_p)$, we consider instead a sequence of associated operators linked to the original matrix by the similarity transformations (8), and acting on the standard sequence spaces l_p. In other words, it is convenient to map the space $l_p(\mu)$ onto the space l_p by the isometric isomorphism which takes the vector $\{x_k\}$ from $l_p(\mu)$ into the vector $\{x_k\mu_k^{1/p}\}$ in l_p. It is clear that (apart from the scale factor) the spectral properties of the operator will be unchanged by such a similarity transformation.

These considerations lead to the following formulation of the main result; in this form it can be considered as the generalization of a result attributed to Kolmogorov (see [2]) that every finite irreducible nonnegative matrix is similar to a stochastic matrix.

THEOREM 3.1. *Let T be an irreducible matrix with convergence norm R and left and right R-subinvariant vectors $\{\alpha_k\}$, $\{\beta_k\}$ respectively. Then for each $p\,(1 \leqq p \leqq \infty)$ the matrix T_p with elements*

$$(T_p)_{ij} = \left(\frac{\alpha_i}{\alpha_j}\right)^{1/q}\left(\frac{\beta_j}{\beta_i}\right)^{1/p} t_{ij}R \tag{8}$$

defines a contraction operator on the sequence space l_p.

Proof. The proof depends on an application of Hölder's inequality. If $\boldsymbol{u}$ is any l_p-vector, we have to show that the vector

$$\boldsymbol{v} = T_p\boldsymbol{u}$$

with elements

$$v_j = \sum_k u_k(\alpha_k/\alpha_j)^{1/q}(\beta_j/\beta_k)^{1/p}t_{kj}R$$

belongs to l_p, and $\|\boldsymbol{v}\|_p \leqq \|\boldsymbol{u}\|_p$. If we write $a_k = u_k(\beta_j/\beta_k)^{1/p}(t_{kj}R)^{1/p}$ $b_k = (\alpha_k/\alpha_j)^{1/q}(t_{kj}R)^{1/q}$, then from Hölder's inequality, we have

$$|v_j| \leqq \sum_k |a_k|\,|b_k| \leqq (\sum_k |a_k|^p)^{1/p}\,(\sum_k |b_k|^q)^{1/q}.$$

But $\Sigma_k |b_k|^q = (1/\alpha_j)\Sigma_k\alpha_k t_{kj}R \leqq 1$, since $\{\alpha_k\}$ is left R-subinvariant. Hence,

$$\sum_j |v_j|^p \leqq \sum_j (\sum_k |a_k|^p) = \sum_k |u_k|^p \sum_j(\beta_j/\beta_k)t_{kj}R \leqq \sum_k |u_k|^p$$

where the last step uses the fact that the vector $\{\beta_k\}$ is right R-subinvariant. This inequality proves the theorem when $1 < p < \infty$;

we leave the reader to supply the modifications required when $p = 1$ or $p = \infty$.

Further properties of the associated operators follow from the theorems in the previous section. It is readily verified that the matrices T_p all have convergence parameter equal to unity, and are 1-transient, 1-null, or 1-positive according as the original matrix T is R-transient, R-null, or R-positive. Theorem 2.1 implies that the spectral radius of T_p is equal to unity. If T is R-positive it follows from Theorem 2.2 that the operator on l_p induced by T_p has a simple eigenvalue, associated with a unique eigenvector, at each of the points $e^{2\pi i m/d}(m = 1, \cdots, d)$ where d is the period of T. If T is R-null or R-transient, then the operator induced by T_p has no eigenvalues on the boundary of the unit disc.

These remarks apply if, in particular, T is a stochastic matrix, $R = 1$, and we set $\beta_j = 1$. This situation has been discussed in detail by Šidák [15] and Holmes [4]. A slight extension of the theorem is required to cover the case that T is stochastic and geometrically transient (i.e., $R > 1$); if in the defining equations (8) we then use unity in place of R, choose any subinvariant vector $\{\alpha_k\}$, and again set $\beta_j = 1$, the previous chain of inequalities applies without modification, and shows that the operator is still a contraction. But in this case it may be possible that the spectral radius, and even the norm, is strictly less than unity (some examples to illustrate these possibilities one discussed by Šidák [16] and Holmes [4]). Of course, it is still true that in this case the operator has no eigenvalues on the boundary of the unit disc.

One other feature of the matrices T_p deserves mention: they possess left and right subinvariant vectors given by

$$\begin{aligned} u_k &= (\alpha_k\beta_k)^{1/p} \\ v_k &= (\alpha_k\beta_k)^{1/q} \end{aligned} \tag{9}$$

respectively.

These relationships hold the key to one question which is suggested by Theorem 3.1, namely, which contraction operators on l_p can arise as the result of the transformation of that theorem? Given an arbitrary contraction operator A on l_p, when does there exist some nonnegative matrix T such that A coincides with the operator derived from T by the Equations (8)?

We have not been able to find a complete answer to this question in purely operator-theoretic terms. The following result, however, identifies the *matrices* which are representable in the form (8).

THEOREM 3.2. *The necessary and sufficient conditions for a matrix S to be an associated operator on $l_p(1 \leqq p < \infty)$ are that the* *matrix be irreducible and that there exist a positive vector* $\{\mu_k\}$ *such that*

$$u_k = \mu_k^{1/p} \text{ and } v_k = \mu_k^{1/q}$$

are left and right subinvariant vectors, respectively, for S.

Proof. The necessity is simply a restatement of Equations (9) above. To prove sufficiency, we have to show that there exists a matrix T which can be substituted into the right hand side of (8) to yield s_{ij} on the left hand side. In fact it is sufficient to take $t_{ij} = s_{ij}$, with $\alpha_k = u_k$ and $\beta_k = v_k$ above.

Note that in this theorem we do not require $\Sigma_k \mu_k < \infty$, so that the vectors $\{u_k\}$ and $\{v_k\}$ are not necessarily elements of l_p and l_q.

As an immediate corollary to the above result, we obtain the following theorem, which answers the original question in the special case that the underlying matrix is R-positive.

THEOREM 3.3. *Suppose that the positive contraction operator* A *on* $l_p(1 \leqq p < \infty)$ *can be represented by an irreducible matrix. Then* A *is an associated operator for an* R*-positive matrix if and only if unity is in the point spectrum of* A.

Proof. It follows from Theorem 2.5 that the condition of Theorem 3.2 is satisfied whenever A is a contraction operator and unity is in its point spectrum. The converse follows from Theorem 2.2, which shows that unity is in the point spectrum of the associated operator whenever its matrix is 1-positive, i.e. whenever the original matrix is R-positive. This concludes the proof.

The sufficiency part of Theorem 3.2 can be formulated in a slightly different way, which suggests further generalizations of the preceding result. We can restate it as follows.

LEMMA 3.1. *Suppose that the matrix* T *is irreducible, and that there exists a positive vector* $\{\mu_k\}$ *and a positive number* ρ *such that*

$$\sum_k \mu_k^{1/p} t_{kj} \leqq \rho \mu_j^{1/p}$$

and

$$\sum_j t_{kj} \mu_j^{1/q} \leqq \rho \mu_k^{1/q} . \tag{10}$$

Then T *acts as a bounded linear operator on* l_p, *and* $\|T\| \leqq \rho$.

In this form it suggests the following conjecture.

Conjecture. The matrix T acts as a bounded linear operator on l_p if and only if there exists a positive vector $\{\mu_k\}$ and a positive number ρ such that Equations (10) are satisfied. Moreover, the norm of the operator can be identified with the least number ρ for which such a vector $\{\mu_k\}$ can be found.

Although we have not been able to prove this conjecture, note that it gives the correct answer in two important special cases:

(i) if $p = 1$.

(ii) if the matrix T is R-positive and irreducible, and the operator norm is equal to the convergence norm.

To conclude this section we give the proof of the assertion of Theorem 2.4, that if A is a contraction operator on $l_p(1 < p < \infty)$ and its matrix representation is aperiodic and 1-positive, then the operator iterates tend stongly to their ergodic limit. The proof depends on the following result, which is due to Mrs. S.C. Moy [7] (see also [8]).

LEMMA 3.2. *If P is an irreducible aperiodic stochastic or substochastic matrix with finite subinvariant measure $\mu(\,.\,)$, and $\boldsymbol{y} \in l_q(\mu)$, then the iterates $(P^*)^n\boldsymbol{y}$ converge in $l_q(\mu)$ norm to their ergodic limit.*

Proof. Suppose first that $\boldsymbol{y} \in l_\infty(\mu)$ (i.e. that $\boldsymbol{y}$ is a bounded vector). Since μ is finite, $\boldsymbol{y}$ is also in $l_p(\mu)$. From I, Theorem 6.1 it follows that, for each i, $\Sigma_j p_{ij}^{(n)} y_j \to \Sigma \mu_j y_j$. Since P is stochastic or substochastic, the quantities $\Sigma_j p_{ij}^{(n)} y_j$ are uniformly bounded as i and n vary. Hence, from the dominated convergence theorem,

$$\sum_i \mu_i \left| \sum_j p_{ij}^{(n)} y_j - \sum_j \mu_j y_j \right|^q \to 0 \ .$$

Thus the iterates $(P^*)^n\boldsymbol{y}$ converge in l_q-norm to their ergodic limits.

Now let $\boldsymbol{y}$ be an arbitrary vector in $l_q(\mu)$. Since the bounded vectors are dense in $l_q(\mu)$, given ε we can find a bounded $\boldsymbol{y}'$ such that $\| \boldsymbol{y} - \boldsymbol{y}' \|_q < \varepsilon$. Then, denoting the ergodic limit by Π,

$$\begin{aligned} \| (P^*)^n\boldsymbol{y} - \Pi^*\boldsymbol{y} \|_q &\leqq \| (P^*)^n\boldsymbol{y} - (P^*)^n\boldsymbol{y}' \|_q \\ &\quad + \| (P^*)^n\boldsymbol{y}' - \Pi^*\boldsymbol{y}' \|_q + \| \Pi^*\boldsymbol{y} - \Pi^*\boldsymbol{y}' \|_q \\ &\leqq 2\varepsilon + \| (P^*)^n\boldsymbol{y}' - \Pi^*\boldsymbol{y}' \|_q \ . \end{aligned}$$

The last term approaches zero as $n \to \infty$, and since ε is arbitrary it follows that the left-hand side also approaches zero. This concludes the proof of the lemma.

To apply this result, we first reinterpret the lemma in terms of associated operators. It is easy to see that any associated operator (i.e., any operator whose matrix can be represented by the Equations

(8)) is in fact an associated operator for a stochastic or substochastic matrix. Indeed, a suitable matrix is given by (using the notation of Theorem 3.1)

$$p_{ij} = (\beta_j/\beta_i)t_{ij}R \ .$$

It therefore follows from the lemma that if T_p is any associated operator on l_p, and $\boldsymbol{y}$ is any vector in l_q, the iterates $(T_p^*)^n\boldsymbol{y}$ converge in l_q-norm to their ergodic limits, provided only that the measure $\{\mu_k\} = \{\alpha_k\beta_k\}$ is finite. The latter condition will be satisfied, in particular, whenever the original matrix is R-positive.

Now suppose that A is a contraction operator on l_p, $(1 < p < \infty)$ with unity in its point-spectrum, so that the matrix is 1-positive. By Theorem 3.3, both A and A^* are then associated operators for R-positive matrices. Consequently, if we apply the above argument to A^*, with $\boldsymbol{x} \in l_p$ we see that the iterates $A^n\boldsymbol{x}$ of the adjoint of A^* (which is A itself since the l_p spaces are reflexive for $1 < p < \infty$) converge in l_p-norm to their ergodic limit. This completes the proof.

4. The ergodic decomposition of positive contraction operators on l_p-spaces. In this section we shall examine the ergodic behaviour of a simple class of operators with reducible matrix representations. Since the whole of the preceding discussion has been devoted to irreducible matrices, we start with a few comments concerning the structure of an arbitrary nonnegative matrix.

Exactly as in the stochastic case, we can say that the index i *communicates* with the index j if there is a chain of indices $k_0, k_1, \cdots, k_n$ with $k_0 = i$, $k_n = j$ and $t_{k_ik_{i+1}} > 0$ $(i = 0, 1, \cdots, n-1)$. If i communicates with j and j communicates with i, then i and j *intercommunicate*. This concept defines an equivalence relationship among the set of all indices which communicate with themselves, the disjoint equivalence classes from which we shall refer to as *irreducible classes*. Note that a matrix is irreducible if all the indices form a single irreducible class.

If C is an irreducible class, the remaining indices fall into three mutually exclusive groups relative to C: those which communicate with one (and hence all) of the indices in C; those with which one (and hence all) of the indices in C communicates; those which enjoy neither property. If we call these three groups $S(C)$, $A(C)$, $N(C)$ respectively, then

(i) columns whose indices belong to C contain zero entries in rows whose indices belong to $A(C)$, and hence nonzero entries only for rows in $S(C)$ or $N(C)$;

(ii) rows whose indices belong to C contain zero entries in columns whose indices belong to $S(C)$, and hence nonzero entries only in $A(C)$ or $N(C)$.

It follows that by suitable relabelling of indices, any nonnegative matrix can be partitioned into a sort of checker-board pattern, with square blocks, corresponding to entries with row and column indices from a particular irreducible class, centered on the main diagonal, and the other nonzero entries following into restricted groups of rows and columns. It is important to note that if C is an irreducible class, and T_C denotes the reduced matrix with row and column indices restricted to C, then

$$(T_C)^n = (T^n)_C \,. \tag{11}$$

Thus, in examining the iterates of a reducible matrix, we may, so far as the irreducible classes are concerned, restrict attention to the reduced (and now irreducible) matrices formed by restricting row and column entries to the class in question.

Under certain circumstances (for example, whenever the total number of irreducible classes is finite), it may be possible to order the irreducible classes $\{C_i\}$ so that if $i > j$, then either C_i and C_j are mutually inaccessible, or C_i is accessible from C_j. In this case the appearance of the matrix is further simplified, so that the only nonzero entries on the same level as one of the irreducible blocks lie to its left.

In general, however, any further simplification of the matrix structure will depend on special assumptions concerning the matrix and its irreducible classes.

The purpose of this final section is to investigate some of the simplifications which result from the assumption that the matrix acts as a contraction operator on one of the l_p spaces. We shall not attempt to obtain a full decomposition of the matrix; our aim is to investigate its ergodic behaviour, and to obtain a decomposition that is relevant to this problem.

If $p = 1$, the matrix corresponding to a positive contraction operator is either stochastic or substochastic; in this case it is well known that the ($C - 1$ averages of) iterate elements $t_{ij}^{(n)}$ tend to zero unless j belongs to a positive recurrent, irreducible class; that each such class is associated with an invariant probability distribution; that any invariant probability distribution can be represented as a mixture of the invariant distributions associated with these classes; and, finally, that a relabelling of the indices allows the matrix itself to be written in the form

$$T = \begin{vmatrix} T_1 & & & & & \\ & T_2 & & & & \\ & & \ddots & & & \\ & & & T_r & & \\ & & & & \ddots & \\ A_1 & A_2 & \cdots & A_r & \cdots & N \end{vmatrix} \tag{12}$$

where each of the T_i is strictly stochastic, irreducible, and positive recurrent, and N is a dissipative matrix with the property that $C - 1$ averages of its iterate elements tend to zero.

The following theorem establishes a similar decomposition for contraction operators on the reflexive l_p-spaces. The term *ergodic contraction* is used in this theorem to denote an irrreducible matrix which defines a contraction operator on l_p with unity in its point spectrum. Some special properties of such matrices have been established in Theorems 2.2, 2.5, and 3.3. "Eigenvector" is used throughout the theorem and its proof as an abbreviation for "eigenvector associated with the eigenvalue unity."

THEOREM 4.1. *Let T be the nonnegative matrix representing a positive contraction on $l_p(1 < p < \infty)$. Then T can be written in the partitioned form*

$$T = \begin{vmatrix} T_1 & & & & & \\ & T_2 & & & & \\ & & \ddots & & & \\ & & & T_r & & \\ & & & & \ddots & \\ & & & & & N \end{vmatrix} \tag{13}$$

where each of the T_i is an ergodic contraction on l_p, and N is a dissipative matrix with the property that the $C - 1$ averages of its iterate elements tend to zero.

The eigenvectors for the individual T_i, if extended by zeros for all indices outside the corresponding irreducible classes, form eigenvectors for T as a whole, and conversely every eigenvector for T can be represented as a mixture of the eigenvectors associated in this way with the irreducible classes. A similar statement holds for the transpose of T. Finally, if x is any vector in l_p, the $C - 1$ averages of the iterates $T^n x$ converge strongly to their ergodic limit, *which can be represented in a similar partitioned form, with T_i replaced by an irreducible idempotent of unit rank, and N replaced by zero.*

Proof. The significant part of the proof lies in establishing the following lemma.

LEMMA 4.1. *Suppose that T is a positive contraction operator on l_p. Then the index i is in the support of an eigenvector for T if and only if i belongs to an irreducible class, and the matrix reduced to this class is* 1*-positive.*

Proof of Lemma 4.1. Let $\boldsymbol{u}$ be any eigenvector, and $S(\boldsymbol{u})$ its support. We first note that there is no loss of generality in supposing that $\boldsymbol{u}$ is nonnegative, for in any case, the vector with components $u'_i = |u_i|$ is superinvariant, $(T\boldsymbol{u}')_i \geqq u'_i$, and from the assumption that T is a contraction there can be no indices for which the inequality is strict. Secondly, if $i \in S(\boldsymbol{u})$, where $\boldsymbol{u}$ is now nonnegative, then by Theorem 2.5 i is also in the support of a nonnegative eigenvector for the transpose of T.

Thus we can suppose that the index i lies in the support of a nonnegative eigenvector $\boldsymbol{u}$ for T, and of a nonnegative eigenvector $\boldsymbol{v}$ for its transpose. It is convenient to suppose also that $\boldsymbol{u}$ and $\boldsymbol{v}$ are normalized so that $u_i = v_i = 1$.

From the minimality lemma (I, Lemma 4.1) it follows next that for all k, $L_{ik}(1) \leqq u_k$ (note that the proof of this property does not depend on the assumption of irreducibility). Hence $\Sigma_k L_{ik}(1) v_k < \infty$, and since the vector $\{L_{ik}(1)\}$ is left subinvariant and the vector $\{v_k\}$ is right invariant

$$L_{ik}(1) = \sum_k L_{ik}(1) t_{kj}$$

for every index k in the support of v (cf. I, Lemma 5.2). But (Equation I, (13))

$$L_{ij}(1) = \sum_k L_{ik}(1) t_{kj} + t_{ij}(1 - L_{ii}(1))$$

so that either $L_{ii}(1) = 1$, or $t_{ij} = 0$ for every index j with $v_j > 0$. The latter possibility would imply $\Sigma_j t_{ij} v_j = v_i = 0$, contradicting the assertion that i is in the support of $\boldsymbol{v}$. Hence $L_{ii}(1) = 1$, which implies both that i belongs to an irreducible class (since it communicates with itself) and that the matrix reduced to this class is 1-recurrent.

Finally, the fact that this class is 1-positive follows from I, Criterion III and the inequalities $L_{ik}(1) \leqq u_k$, $F_{ki}(1) \leqq v_k$, which together imply that $\Sigma_k L_{ik}(1)F_{ki}(1) < \infty$. (Note that from (11) this summation is restricted to the irreducible class containing i).

This concludes the proof of the lemma, and we proceed to the proof of the main theorem.

Let $\mathscr{M}$ denote the set of all indices contained in the support of at least one eigenvector for T, and $\mathscr{N}$ the set of all remaining indices. From the lemma, every index in $\mathscr{M}$ belongs to an irreducible class, and since the irreducible classes are distinct, $\mathscr{M}$ can be divided up completely into a finite or denumerable set of disjoint irreducible classes, say $M_1, \cdots, M_r, \cdots$. If T_i denotes the reduced matrix, with row and column entries restricted to M_i, it follows from (10) and the lemma above that each T_i is irreducible and 1-positive.

If $\boldsymbol{u}_i'$ denotes the (reduced) left eigenvector for T_i, and $\boldsymbol{u}_i$ the same vector augmented by zeros for the indices not in M_i, then $\boldsymbol{u}_i$ is clearly superinvariant, and so, from the assumption that T is a contraction, actually invariant. Hence, using the notation of the first paragraph of § 4, the set $S(M_i)$ is empty. But a similar argument applies also to the transpose of T, and so $A(M_i)$ is also empty. (Note that the argument breaks down at this stage if $p = 1$, which accounts for the difference in the representations (12) and (13)). This establishes the representation (13). The assertion that the C $-$ 1 averages of iterate elements with row and column indices in $\mathscr{N}$ tend to zero then follows readily from the mean ergodic theorem and the definition of $\mathscr{N}$. The form of the ergodic limit follows from (11) and the results already established for irreducible matrices. Finally, the strong convergence of the C $-$ 1 averages of iterates $T^n\boldsymbol{x}$ follows from the mean ergodic theorem. The proof of Theorem 4.1 is now complete.

REFERENCES

1. F. F. Bonsall *Lectures on some fixed point theorems of functional analysis*, Tata Institute of Fundamental Research, Bombay, 1962.

2. A. Brauer, *A new proof of theorems of Perron and Frobenius on non-negative matrices. I, positive matrices*, Duke Math. J. **24** (1957), 367-78.

3. N. Dunford and J. Schwartz, *Linear Operators, Part I*, Interscience Publishers, New York, 1958.

4. P. T. Holmes, *On some spectral properties of a transition matrix*, Sankhya (Ser. A) **28** (1966), 205-214. Corrigendum Ibid **29** (1967), 106.

5. S. Karlin, *Positive operators*, J. Math. and Mech. **8** (1959), 907-37.

6. D. G. Kendall, *Unitary Dilations of Markov transition operators...*, In U. Grenander (Ed.) "Probability and Statistics" (Cramér Memorial Volume) Stockholm (Almqvist and Wiksell), 1959.

7. S. C. Moy, *λ-continuous Markov chains-II*, Trans. Am. Math. Soc. **120** (1965) 83-107.

8. ———, *Ergodic properties of expectation matrices of a branching process with countably many types*, J. Math. and Mech. **16** (1966), 1201-1225.

9. E. J. Nelson, *The adjoint Markov process*, Duke Math. J. **25** (1958), 671-90.
10. C. R. Putnam, *A note on nonnegative matrices*, Canad. J. Math. **13** (1961), 59-62.
11. F. Riesz and B. Sz-Nagy, *Über Kontraktionen des Hilbertschen Raumes*, Acta Sci. Math. Szeged. **10** (1943), 202-5.
12. F. Riesz and B. Sz-Nagy, *Leçons d'Analyse Fonctionelle*, Budapest, 1953.
13. H. Schaeffer, *On the singularities of an analytic function with values in a Banach space*, Arch. Math. **11** (1960), 40-43.
14. E. Seneta and D. Vere-Jones, *On quasistationary distributions...*, J. Applied Prob. **3** (1966), 403-34.
15. Z. Šidák, *Eigenvalues of operators in l_p spaces in denumerable Markov chains*, Czech. Math. J. **14** (1964), 438-443.
16. ———, *Some theorems and examples in the theory of operators in denumerable Markov chains*, (in Czech), Časopis pro pěstování matematiky **88** (1963), 457-478.
17. D. Vere-Jones, *Topics in the Theory of Probability*, (D. Phil. Thesis, Oxford, 1961).
18. ———, *Spectral properties of some operators associated with queueing systems*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete. **2** (1963), 12-21.
19. ———, *Ergodic properties of nonnegative matrices-I*, Pacific J. Math. (2) **22** (1967), 361-386.

Received April 6, 1967.

INSTITUTE OF ADVANCED STUDIES
AUSTRALIAN NATIONAL UNIVERSITY