A Metric Variant of Frobenius's Theorem and Some Other Remarks on Positive Matrices

H. H. Schaefer
Universität Tübingen
Mathematisches Institut
7400 Tübingen, Germany

Submitted by Hans Schneider

ABSTRACT

The theory of positive (=nonnegative) finite square matrices continues, three quarters of a century after the pioneering and well-known papers of Perron and Frobenius [4], to present a multitude of different aspects. This is evidenced, for example, by the recent papers [1] and [2], as well as by the vast literature concerned with extensions to operators on infinite dimensional spaces (see [5]). Supposing A to be a positive $n\times n$ matrix with spectral radius $r(A)=1$, the main purpose of this note is to display the role of $\lambda=1$ as a root of the minimal polynomial of A (or equivalently, of certain norm conditions on A), for the lattice structure of the space M spanned by the unimodular eigenvectors of A as well as for the permutational character of A on M. Proposition 1 can thus be viewed as a variant of Frobenius's theorem on the peripheral spectrum of indecomposable square matrices, and we hope that the proof of Proposition 2 will clarify to what extent indecomposability is responsible for the main results available in that special case. The remaining remarks (Propositions 3 and 4) are concerned with the spectral characterization of permutation matrices and with finite groups of positive matrices. Some of that material is undoubtedly known, but we give simple, transparent proofs.

1. MAIN RESULTS

We follow essentially the notation and terminology of [5, Chapter 1]. In particular, if $x=(\xi_i)$ and $y=(\eta_i)$ are vectors of $\mathbb{C}^n$ ($n\in\mathbb{N}$), then $x\leqslant y$ means that x, y are real and $\xi_i\leqslant\eta_i$ for $i=1,\dots,n$; $x<y$ is used to denote the relations $x\leqslant y$ *and* $x\neq y$. An $n\times n$ matrix A is, accordingly, positive (= nonnegative) if all its entries are $\geqslant 0$; in the usual way A is considered an element of $L(\mathbb{C}^n)$. For a given norm τ on $\mathbb{C}^n$, we denote by $\|A\|_\tau$ the corresponding matrix norm. A norm τ on $\mathbb{C}^n$ is called *strictly monotone* on a linear subspace $M\subset\mathbb{C}^n$ if $x, y\in M$ and $0\leqslant x<y$ implies $\tau(x)<\tau(y)$. Finally, $\sigma(A)$ denotes the spectrum of A, $r(A)$ the spectral radius of A, and Γ the circle group $\{z\in\mathbb{C}:|z|=1\}$.

© Elsevier Science Publishing Co., Inc., 1982
52 Vanderbilt Ave., New York, NY 10017 0024-3795/82/010175+8$02.75

If K is a subset of $\sigma(A)$, by the *principal A-subspace L_K associated with K* we understand the largest A-invariant subspace L of $\mathbb{C}^n$ such that $\sigma(A|L)=K$; if B is a basis of $\mathbb{C}^n$ with respect to which A assumes its Jordan normal form, then L_K is spanned by precisely those vectors in B for which the corresponding Jordan block has a diagonal entry $\lambda \in K$.

PROPOSITION 1. *Let $A \geq 0$, $r(A)=1$, and denote by M the principal A-subspace of $\mathbb{C}^n$ associated with $\sigma(A)\cap\Gamma$. The following assertions are mutually equivalent:*

(a) *M has a basis, consisting of vectors >0, on which A acts as a permutation.*

(b) *A is a contraction, $\|A\|_\tau=1$, for some norm τ on $\mathbb{C}^n$ which is strictly monotone on M.*

(c) *A is a contraction for some norm on $\mathbb{C}^n$.*

(d) *The set of powers $\{A^m : m \in \mathbb{N}\}$ is bounded in the euclidean metric of $\mathbb{C}^{nn}$.*

(e) *Every unimodular eigenvalue of A is a simple root of the minimal polynomial of A.*

(f) *$\lambda=1$ is a simple root of the minimal polynomial of A.*

Before giving the proof, we wish to state several remarks and consequences.

(i) If $B=\{x_\rho, y_\sigma,\ldots,z_\tau\}$ is the basis of M referred to in (a) and the permutation of B effected by A is written as the union $(x_1\ldots x_r)(y_1\ldots y_s)\cdots(z_1\ldots z_t)$ of independent cycles, then the number $k \geq 1$ of these independent cycles is the dimension of the fixed space of A, and $\sigma(A)\cap\Gamma$ is the union of k (not necessarily distinct) groups of roots of unity, of respective orders $r, s,\ldots,t$.

(ii) Simple examples such as

$$A=\begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix}$$

show that, in general, a matrix $A \geq 0$ satisfying (a) above is not a contraction with respect to any norm τ strictly monotone on all of $\mathbb{C}^n$. The point of assertion (b) is the following: If any one of the assertions of Proposition 1 holds, then M is a lattice ordered subspace (though not, in general, a sublattice) of $\mathbb{C}^n$ and $\varepsilon x = Ax$, $|\varepsilon|=1$, always implies $|x|=A|x|$ with respect to this ordering (cf. Example below).

(iii) The subsequent proof will show that a suitable set of generators of extreme rays of the positive cone of M provides a basis B of the desired kind;

by this condition B is determined uniquely (except for numeration and some normalization).

Proof of Proposition 1. We begin by recalling the following well-known facts. If M and N denote the principal A-subspaces of $\mathbb{C}^n$ associated with $\sigma(A)\cap\Gamma$ and $\sigma(A)\setminus\Gamma$, respectively, then $\mathbb{C}^n = M+N$ is a direct sum reducing A. Let P denote the projection $\mathbb{C}^n \to M$ with kernel N, and set $U := AP$, $R = A(I-P)$; A commutes with P and we have

$$A = U + R, \tag{$*$}$$

where $RU = UR = 0$ and $\lim_{k\to\infty} R^k = 0$; thus the asymptotic behavior of A^m as $m\to\infty$ is essentially that of U.

In addition, we will make use of the following results on positive matrices: The spectral radius $r(A)$ is an eigenvalue with an eigenvector >0, and if $r(A)\varepsilon\in\sigma(A)$, $|\varepsilon|=1$, then $r(A)\varepsilon^k\in\sigma(A)$ for all $k\in\mathbb{Z}$ (see, for example, [5, I.2.3] and [5, I.2.7]).

(a)$\Rightarrow$(b): If the basis B of M referred to in (a) is written as in remark (i) above and if p is the least common multiple of the integers $r, s, \dots, t$, then it follows that $A^p|M$ is the identity map of M, and hence that $U^p = P$. Since $P = \lim_{k\to\infty} U^{pk} = \lim_{k\to\infty} A^{pk}$, it follows that $P \geqslant 0$ and that $U_0 := U|M$ is an order isomorphism of M (for $U_0^{p-1} = U_0^{-1}$). [From $P\geqslant 0$ it follows that M is lattice ordered; in fact, for $0 \leqslant x, y\in M$, $\sup_M(x,y)$ is given by $P(x\vee y)$, where $x\vee y$ denotes the usual supremum in $\mathbb{R}^n$.] Because $U^p = P$, U is a periodic matrix, and from $(*)$ we conclude that the set $\{A^m : m\in\mathbb{N}\}$ is bounded.

Now let $x\to\|x\|$ be any strictly monotone norm on $\mathbb{C}^n$ (for example, the euclidean norm). The norm τ defined by

$$\tau(x) := \sup_{k\geqslant 0} \|A^k x\| \qquad (x\in\mathbb{C}^n)$$

is strictly monotone on $M = P(\mathbb{C}^n)$: Let $0<x<y$, and suppose that $\tau(x)=\tau(y)$. Since U is an order isomorphism on M, for each $k\in\mathbb{N}$ we have

$$A^k x = U^k x < U^k y = A^k y$$

and hence $\|A^k x\| < \|A^k y\|$. Thus there must exist an infinite sequence $\{n_k\}$ in $\mathbb{N}$ for which

$$\tau(x) = \lim_{k\to\infty} \|A^{n_k} x\| = \lim_{k\to\infty} \|A^{n_k} y\| = \tau(y).$$

Without loss of generality we may assume that all n_k belong to the same residue class $[p_0]$ mod p. Then we have $\lim_{k\to\infty} A^{n_k} = U^{p_0}$ and hence $\|U^{p_0}x\| = \|U^{p_0}y\|$, which is contradictory.

Thus τ is a norm which is strictly monotone on M. Since, clearly, A is a τ-contraction, assertion (b) is proved.

The implications (b)$\Rightarrow$(c)$\Rightarrow$(d)$\Rightarrow$(e)$\Rightarrow$(f) are trivial [for (d)$\Rightarrow$(e), consider C. Neumann's series for A].

(f)$\Rightarrow$(a): It is well known (and easy to see, for instance from the Jordan normal form of A) that (f) is equivalent to $(\lambda - A)^{-1}$ having a first order pole at $\lambda = 1$. Now for each eigenvalue $\varepsilon \in \Gamma$ of A, a brief look at C. Neumann's series for $(\lambda - A)^{-1}$ $(|\lambda| > 1)$ shows that for $\lambda = |\lambda|\varepsilon$,

$$|(\lambda-\varepsilon)(\lambda-A)^{-1}| \leq (|\lambda|-1)(|\lambda|-A)^{-1}$$

holds by the positivity of A; therefore, ε is a first order pole of $(\lambda - A)^{-1}$ [which is again assertion (e)]. Thus for all $p \in \mathbb{N}$, $(\lambda - A^p)^{-1}$ has a first order pole at $\lambda = 1$. On the other hand, since every $\alpha \in \sigma(A) \cap \Gamma$ is a root of unity, there exists a smallest integer $p \geq 1$ such that $\sigma(A^p) \cap \Gamma = \{1\}$. From the decomposition $(*)$ above we obtain $A^p = U^p + R^p$ and $\sigma(U^p|M) = \sigma(A^p|M) = \{1\}$; since by the preceding, $(\lambda - U^p)^{-1}$ has a first order pole at $\lambda = 1$, we obtain $U^p|M = \mathrm{id}\, M$ and thus $U^p = P$.

As in (a)$\Rightarrow$(b), we conclude that $P \geq 0$ and hence that $U = AP \geq 0$. Thus $U|M = A|M$ is an order isomorphism of M and, as observed above, M is a (complex) archimedean vector lattice under the ordering induced by $\mathbb{C}^n$. Now every real (respectively, complex) archimedean vector lattice M of finite dimension k is isomorphic to $\mathbb{R}^k$ (respectively, $\mathbb{C}^k$) under its usual ordering (cf. [5, II.3.9, Corollary 1]). Thus the positive cone of M contains a set E of exactly k extreme rays, and obviously A must map E bijectively onto itself, i.e., A must permute E. Clearly, then, there exists a basis of M satisfying (a).

This completes the proof of Proposition 1. ∎

Example. An example where the space M is not a sublattice of $\mathbb{C}^n$ is furnished by the matrix

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

with eigenvalues $\pm 1, 0$. M is the space of all vectors $(\alpha, \alpha+\beta, \beta)$ $(\alpha, \beta \in \mathbb{C})$ and a two dimensional vector lattice under the ordering induced by $\mathbb{C}^3$; the extreme positive rays of M are generated by the vectors $(1,1,0)$ and $(0,1,1)$.

Proposition 1 now gives fairly easy access to the main conclusions of Frobenius's classical result on irreducible (=indecomposable) positive matrices plus some supplementary information.

PROPOSITION 2. *Let $A \geqslant 0$ be irreducible, and suppose that $r(A)=1$. Then $\sigma(A)\cap\Gamma$ is the group Γ_r of all rth roots of unity for some r, $1\leqslant r\leqslant n$, and each $\lambda\in\Gamma_r$ has algebraic and geometric multiplicity one. Moreover, the principal A-subspace M associated with $\sigma(A)\cap\Gamma$ is a sublattice of $\mathbf{C}^n$ whose extreme positive rays are cyclically permuted by A.*

Proof. There exists a positive fixed vector $y\neq 0$ of the transpose ${}^tA$ and, since ${}^tA$ is irreducible if (and only if) A is, y has all its coordinates >0. Therefore, $x\to\langle |x|, y\rangle$ is a strictly monotone norm on $\mathbf{C}^n$ for which A is obviously a contraction, so (b) of Proposition 1 is satisfied. Moreover, the fixed space $F:=\{x\colon Ax=x\}$ must be one dimensional. For if $x=Ax$ then $|x|\leqslant A|x|$ $(A\geqslant 0)$, and $\langle |x|,y\rangle\leqslant\langle A|x|,y\rangle=\langle |x|,y\rangle$ implies that $|x|=A|x|$; now if F is a linear sublattice of $\mathbf{C}^n$ each of whose positive nonzero vectors has all coordinates >0, then $\dim F\leqslant 1$.

Thus $\dim F=1$, and this implies that the permutation effected by A on the basis B of M [see remark (i) following Proposition 1] has precisely one independent cycle $(x_1\cdots x_r)$; for each cycle gives rise to a fixed vector $\sum_{\rho=1}^{r}x_\rho$, and fixed vectors stemming from distinct cycles are clearly linearly independent. By remark (iii), we may and will assume that the basis vectors $x_1,\ldots,x_r$ of M generate extreme positive rays in M. Now if $Ax_\rho=x_{\rho+1}$ (ρ mod r) and ε is a primitive rth root of unity, the vectors

$$u_k=\sum_{\rho=1}^{r}\bar{\varepsilon}^{k\rho}x_\rho \qquad (k=0,\ldots,r-1)$$

satisfy $Au_k=\varepsilon^k u_k$. Since $\det(\bar{\varepsilon}^{k\rho})$ is a van der Monde determinant $\neq 0$, the vectors u_k $(k=0,\ldots,r-1)$ are linearly independent (in particular, $\neq 0$) and thus constitute a basis of M. This shows that the group $\Gamma_r=\{\varepsilon^k\colon k=0,\ldots,r-1\}$ is the set of all unimodular eigenvalues of A, each of geometric and algebraic multiplicity one.

It remains to show that M is a sublattice of $\mathbf{C}^n$. In the proof of Proposition 1, (a)⇒(b), we have observed that the projection P (of $\mathbf{C}^n$ onto M, with kernel the principal A-subspace associated with $\sigma(A)\setminus\Gamma$) is positive, and that for $0\leqslant x, y\in M$ the lattice supremum $\sup_M(x,y)$ is given by $P(x\vee y)$. Now $P=\sum_{\rho=1}^{r}y_\rho\otimes x_\rho$ for suitable vectors y_ρ satisfying $\langle x_\rho, y_\sigma\rangle=\delta_{\rho\sigma}$. Since by assumption the x_ρ are extreme positive rays in M, $P\geqslant 0$ implies that $y_\rho\geqslant 0$

$(\rho=1,\dots,r)$. On the other hand, from $\Sigma_1^r y_\rho \otimes x_{\rho+1} = \Sigma_1^r y_\rho \otimes Ax_\rho = \Sigma_1^r {}^tAy_\rho \otimes x_\rho$ we conclude that ${}^tAy_\rho = y_{\rho-1}$ for all $\rho \bmod r$. Therefore, $\Sigma_1^r y_\rho$ is fixed under ${}^tA$ and hence has all coordinates >0. But then $x>0$ implies $Px>0$. Since $\sup_M(x,y) \geqslant x \vee y$ in $\mathbb{C}^n$ and since $P(\sup_M(x,y) - x\vee y)=0$, we conclude that $x \vee y = \sup_M(x,y)$ for all $0 \leqslant x, y \in M$. ∎

An inspection of the preceding proof shows the conclusion of Proposition 2 to remain valid whenever the permutation of the extreme positive rays of M effected by A has precisely one independent cycle, and whenever $\Sigma_{\rho=1}^r y_\rho$ has all coordinates >0. These requirements are evidently weaker than irreducibility of A. Also, in the more general situation of Proposition 1, each cycle $(z_1 \dots z_t)$ of that permutation gives rise to t eigenvectors

$$w_m = \sum_{\tau=1}^{t} \bar{\eta}^{m\tau} z_\tau \qquad (m=0,\dots,t-1)$$

pertaining to the eigenvalues η^m, where η denotes any primitive tth root of unity.

2. SPECTRAL CHARACTERIZATION OF PERMUTATION MATRICES

It is well known that a (row or column) stochastic matrix A is a permutation matrix iff $\sigma(A) \subset \Gamma$, Γ denoting the circle group [5, I.4.5]. This equivalence is valid more generally, as Proposition 3 below shows. We begin with the following lemma, which is very well known but seldom proved.

LEMMA. *Every $n \times n$ matrix $A \geqslant 0$ with inverse $A^{-1} \geqslant 0$ is monomial (i.e., contains exactly one entry $\neq 0$ in each row and column).*

Proof. It is easy to see that the coordinate functionals $\delta_i(x) = x_i$ $(i=1,\dots,n)$ and their multiples $\alpha\delta_i$ $(\alpha \geqslant 0)$ are the only lattice preserving linear functionals on $\mathbb{R}^n$. Now if $A \geqslant 0$ and $A^{-1} \geqslant 0$, then A is a linear lattice isomorphism of $\mathbb{R}^n$; consequently, for each $\{1,\dots,n\}$ there exist a unique $\alpha_i \geqslant 0$ and $j \in \{1,\dots,n\}$ such that $\delta_i \circ A = {}^tA\delta_i = \alpha_i \delta_j$. Hence A has precisely one entry $\neq 0$ in each row. Since the same argument applies to ${}^tA$, A is monomial. ∎

In the following we denote by e_i $(i=1,\dots,n)$ the standard basis vectors of $\mathbb{C}^n$.

PROPOSITION 3. *Let $A \geqslant 0$ be a contraction, $\|A\|_\tau \leqslant 1$, with respect to any norm τ on $\mathbb{C}^n$ for which $\tau(e_i) = \tau(e_j)$ $(i, j = 1, \ldots, n)$. Then A is a permutation matrix if and only if $\sigma(A) \subset \Gamma$.*

Proof. If $A = P_\pi$ is a permutation matrix then, clearly, $\sigma(A) \subset \Gamma$ [5, p.14]. Conversely, if $\sigma(A) \subset \Gamma$, then each eigenvalue of A is a root of unity; hence we have $\sigma(A^p) = \{1\}$ for some integer $p \geqslant 1$. Since $\|A^p\|_\tau \leqslant 1$, by Proposition 1 the minimal polynomial of A^p must be $\lambda - 1$, whence it follows that $A^p = I$. Thus $A^{-1} = A^{p-1} \geqslant 0$, and by the preceding lemma A is monomial. If $\alpha_{i\pi(i)}$ denotes the unique entry $\neq 0$ of A in the ith row, then $i \to \pi(i)$ is a permutation of $\{1, \ldots, n\}$. Thus for each i and $j = \pi(i)$ we have $Ae_j = \alpha_{ij} e_i$, which implies

$$\alpha_{ij} \tau(e_i) = \tau(Ae_j) \leqslant \tau(e_j),$$

or $\alpha_{ij} \leqslant 1$. Because $A^p = I$, we must have $\alpha_{i\pi(i)} = 1$ for each i; hence A is a permutation matrix. ■

3. FINITE GROUPS OF POSITIVE MATRICES

Our final remark concerns the characterization of finite groups G of positive $n \times n$ matrices, the group operation being matrix multiplication, and the identity of G being the $n \times n$ unit matrix I. (In the slightly more general case where the identity of G is a positive projection, the results are entirely similar; compare also [3].) In the following, we denote by Π the group of $n \times n$ permutation matrices, and by Δ the group of all diagonal matrices with diagonal entries > 0.

PROPOSITION 4. *For any group G of positive $n \times n$ matrices, the following properties are equivalent:*

(a) *G is isomorphic to a subgroup of Π.*
(b) *G is bounded (in the euclidean metric of $\mathbb{R}^{nn}$).*
(c) *For all $A \in G$, $\sigma(A) \subset \Gamma$.*
(d) *$G \cap \Delta = \{I\}$.*
(e) *G possesses a strictly positive fixed vector.*

Proof. (a)⇒(b) is clear. (b)⇒(c): If G is bounded, then $(\lambda - A)^{-1}$ and $(\lambda - A^{-1})^{-1}$ exist for all $\lambda \in \mathbb{C}$, $|\lambda| > 1$, since C. Neumann's series converges for such λ. Thus $\sigma(A) \subset \Gamma$. (c)⇒(d) is trivial.

(d)⇒(a): Each $A \in G$ is monomial (Lemma, Section 2), and hence has a representation $A = PD$ with $P \in \Pi$ and $D \in \Delta$, which is easily seen to be unique. If $A_1 = P_1 D_1$, $A_2 = P_2 D_2$ are elements of G, then $A_1 A_2 = (P_1 P_2)D$, where $D = (P_2^{-1} D_1 P_2) D_2 \in \Delta$; therefore, $A \to P$ is a homorphism of G into Π with kernel $G \cap \Delta$. Thus if $G \cap \Delta = \{I\}$ is trivial, G is isomorphic with a subgroup of Π.

(a)⇒(e): If e denotes the vector $(1,1,\dots,1)$, then $x_0 := \sum_{A \in G} Ae$ is clearly a common fixed vector for all $A \in G$. Moreover, x_0 is strictly positive (in the sense that all coordinates of x_0 are >0), since even each summand Ae has this property, A being an order isomorphism of $\mathbb{R}^n$. (e)⇒(d) is clear. ■

We observe in conclusion that, for finite G, an isomorphism of G into Π can be obtained explicitly from any vector x_0 satisfying (e) above; in fact, if $D_0 \in \Delta$ denotes the matrix with diagonal the coordinates of x_0, then $x_0 = D_0 e$ and $D_0^{-1} A D_0 e = e$. Now $A \to D_0^{-1} A D_0$ is an isomorphism of G into Π, since each $D_0^{-1} AD$ is a stochastic matrix with stochastic inverse (hence a permutation matrix).

The author is much indebted to H. Schneider for several valuable suggestions.

REFERENCES

1 G. P. Barker, and H. Schneider, Algebraic Perron-Frobenius theory, *Linear Algebra Appl.* 11:219–233 (1975).
2 S. Friedland and H. Schneider, The growth of the powers of a nonnegative matrix, *SIAM J. Alg. Disc. Meth.* 1:185–200 (1980).
3 P. Flor, On groups of non-negative matrices. *Compositio Math.* 21:376–382 (1969).
4 Frobenius, Über Matrizen aus nicht negativen Elementen, *S.-B. Preuss. Akad. Wiss. Berlin*, 1912, pp. 456–477; *Ges. Abh.*, Springer, Berlin, 1968, Vol. 3, pp. 546–567.
5 H. H. Schaefer, *Banach Lattices and Positive Operators*, Springer, New York,

Received 9 December 1980; revised 15 May 1981