Convexity and Log Convexity for the Spectral Radius

Roger D. Nussbaum*
Department of Mathematics
Rutgers University
New Brunswick, New Jersey 08903

Submitted by Shmuel Friedland

ABSTRACT

The starting point of this paper is a theorem by J. F. C. Kingman which asserts that if the entries of a nonnegative matrix are log convex functions of a variable then so is the spectral radius of the matrix. A related result of J. Cohen asserts that the spectral radius of a nonnegative matrix is a convex function of the diagonal elements. The first section of this paper gives a new, unified proof of these results and also analyzes exactly when one has strict convexity. The second section gives some very simple proofs of results of Friedland and Karlin concerning "min-max" characterizations of the spectral radius of nonnegative matrices. These arguments also yield, as will be shown in another paper, min-max characterizations of the principal eigenvalue of second order elliptic boundary value problems on bounded domains. The third section considers the cone K of nonnegative vectors in R^n and continuous maps $f: K \to K$ which are homogeneous of degree one and preserve the partial order induced by K. The (cone) spectral radius of such maps is defined and a direct generalization of Kingman's theorem to a subclass of such nonlinear maps is given. The final section of this paper treats a problem that arises in population biology. If K_0 denotes the interior of K and f is as above, when can one say that f has a unique eigenvector (to within normalization) in K_0? A subtle point to be noted is that f may have other eigenvectors in the boundary of K. If $u \in K_0$ is an eigenvector of f, $|u| = 1$, and $g(x) = f(x)/|f(x)|$, when can one say that for any $x \in K_0$, $g^p(x)$, the pth iterate of g acting on x, converges geometrically to u? The fourth section provides answers to these questions that are adequate for many of the population biology problems.

*Partially supported by NSF MCS 82-101316 and as a visiting member of the Courant Institute, 1983–84.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/86/$3.50

INTRODUCTION

The spectral radius $r(A)$ of a square matrix A is the maximum of $\{ |\lambda| : \lambda$ an eigenvalue of $A\}$. If $A = (a_{ij})$ is a "nonnegative matrix" (so $a_{ij} \geqslant 0$ for $1 \leqslant i, j \leqslant n$), Cohen [7, 8] has shown that $r(A)$ is a convex function of the diagonal elements of A, and Kingman [17] has proved that if the entries of A are log convex functions of a parameter t, then $r(A)$ is also a log convex function of t. Friedland [12] has also given related results concerning the convex dependence of $r(A)$ on various parameters. In the first section of this paper we shall present a simple and unified approach to refinements of the Cohen, Kingman and Friedland theorems. In particular we shall obtain necessary and sufficient conditions for strict convexity or strict log convexity to hold in our theorems. With the partial exception of some results in [12], such necessary and sufficient conditions are inaccessible by previous methods.

The second section of this paper presents a very simple approach to a "minimax" variational formula (obtained by Friedland) for the spectral radius of a nonnegative matrix A. We also give a simple proof of an earlier, closely related theorem of Friedland and Karlin [13]. We should remark that at least part of Friedland's theorem is a consequence of an earlier, more general result of Donsker and Varadhan [11]. However, Friedland's result is sharper, for it explicitly gives the saddlepoint at which the minimax is achieved.

Our real interest in the results of Section 2 is that the proofs can be generalized to the context of second order elliptic eigenvalue problems like

$$\Lambda(u) = \sum a_{ij} u_{x_i x_j} + \sum b_i u_{x_i} + cu = \lambda u \qquad \text{on } \Omega,$$

$$\alpha u + \beta \cdot \nabla u = 0 \qquad \text{on } \partial\Omega. \qquad (0.1)$$

Here Λ is assumed uniformly elliptic on a smooth, bounded domain Ω; $a_{ij}(x)$, $b_i(x)$, and $c(x)$ are Hölder-continuous; $\alpha(x) \geqslant 0$ on $\partial\Omega$; and $\beta(x)$ is an outward-pointing vector on $\partial\Omega$ or $\beta(x) \equiv 0$ [in which case $\alpha(x)$ is assumed positive on $\partial\Omega$]. We shall prove in [26] that variants of arguments like those in Section 2 can be used to give a variational characterization of the principal eigenvalue of (0.1).

The third section of this paper is concerned with nonlinear generalizations of the Cohen and Kingman theorems. Let $K = \{x \in \mathbb{R}^n : x_i \geqslant 0 \text{ for all } i\}$ and $K_0 = \{x \in \mathbb{R}^n : x_i > 0 \text{ for all } i\}$ (this notation is maintained throughout this paper), and suppose $f: K \to K$ is a continuous map which is homogeneous of degree one $[f(\lambda x) = \lambda f(x)$ for $x \in K$ and $\lambda > 0]$ and order-preserving (with respect to the partial ordering induced by $x \leqslant y$ if $y - x \in K$). One can define eigenvectors and eigenvalues in the usual way for such maps and define $r(f) = \sup\{\lambda \geqslant 0 : \lambda$ is an eigenvalue of $f\}$. We call $r(f)$ the spectral radius of f. Our operators will depend in a natural way on certain parameters; if these parameters are, in turn, log convex functions of a variable t, we shall prove (directly generalizing Kingman's theorem) that the generalized spectral radius is a log convex function of t. We shall also give a direct generalization in this framework of Cohen's theorem.

One point should be emphasized here: The approach which we give to the linear questions of Section 1 generalizes directly to the nonlinear context of Section 3, but other approaches to the linear theory do not seem to generalize to this framework.

Eigenvectors of f in K_0 are equivalent to fixed points of $g(x) = f(x)/|f(x)|$ in K_0 (where $|u|$ denotes a suitable norm). It is natural to ask whether g has a unique fixed point u in K_0 and whether, for any $x \in K_0$, $g^p(x)$, the pth iterate of g acting on x, converges to u. In Section 4 we consider such questions and obtain theorems which reduce in the linear case to the Perron-Frobenius theorem and to a theorem of Birkhoff [3]. Corollaries 4.6 and 4.7 below are very special cases of our results which nevertheless convey the flavor of our theorems.

One point should be strongly emphasized about the results in Section 4: If there exists an integer p such that $f^p(K - \{0\}) \subset K_0$, then the arguments given in Section 4 can be simplified enormously. However, precisely this condition fails in many examples of interest to us, e.g., in Corollaries 4.6 and 4.7.

We should also remark that the results of Section 4 are of interest in studying so-called "two-sex models" in population biology, and the particular class of functions f we emphasize is motivated by examples in the population-biology literature.

1. CONVEXITY AND LOG CONVEXITY FOR THE SPECTRAL RADIUS

Our prerequisites for this section comprise only some elementary facts from the theory of nonnegative matrices (see [29, Chapter 1]). If $A = (a_{ij})$ and $B = (b_{ij})$ are $n \times n$ matrices, we shall write $A \geqslant B$ if $a_{ij} \geqslant b_{ij}$ for $1 \leqslant i, j \leqslant n$, and $A > B$ if $a_{ij} > b_{ij}$ for $i \leqslant i, j \leqslant n$. Analogously, if $x, y \in \mathbb{R}^n$, we shall write $x \geqslant y$ if $x_i \geqslant y_i$ for $1 \leqslant i \leqslant n$ and $x > y$ if $x_i > y_i$ for $1 \leqslant i \leqslant n$; we set $K = \{x \in \mathbb{R}^n\colon\ x \geqslant 0\}$. An $n \times n$, nonnegative matrix A is called "irreducible" if for each pair of integers (i, j) with $1 \leqslant i, j \leqslant n$, there exists an integer $m = m(i, j)$ such that the entry in row i and column j of A^m is positive.

The Perron-Frobenius theorem [29, pp. 20, 25] asserts that if $A \geqslant 0$ (A a square matrix) and $r = r(A)$, there exists $u \in K - \{0\}$ such that $Au = ru$. Furthermore, one can easily prove directly or obtain from [29]

LEMMA 1.1 (See Theorem 1.6 in [29]). *Suppose that A is a nonnegative, irreducible square matrix and that there exists $u \in K - \{0\}$ and a real number ρ such that*

$$Au \leqslant \rho u. \tag{1.1}$$

Then one has $\rho > 0$, $u > 0$, and $r(A) \leqslant \rho$. Furthermore, $r(A) < \rho$ unless equality holds in (1.1).

As an immediate consequence of Lemma 1.1 we have the following simple but useful observation.

LEMMA 1.2. *If A and B are $n \times n$ irreducible, nonnegative matrices such that $A \leqslant B$ and $A \neq B$, then $r(A) < r(B)$.*

Proof. Lemma 1.1 and the Perron-Frobenius theorem imply that there exists $v > 0$ such that

$$Bv = r(B)v.$$

Because $A \leqslant B$, it follows that

$$Av \leqslant r(B)v,$$

and because $A \neq B$ and $v > 0$, equality cannot hold in the previous inequality. Lemma 1.1 now implies that

$$r(A) < r(B). \qquad \blacksquare$$

If $A = (a_{ij})$ is a matrix with nonnegative off-diagonal elements (so $a_{ij} \geqslant 0$ for all $i \neq j$), we shall say, following notation in [10], that A is *essentially nonnegative*. Seneta calls such matrices *ML*-matrices; see [29, p. 40]. If A is essentially nonnegative, the Perron-Frobenius theorem implies that A has a real eigenvalue $\lambda_1 = \lambda_1(A)$ with corresponding eigenvector in K and

$$\operatorname{Re} \lambda \leqslant \lambda_1(A)$$

for every other eigenvalue λ of A; $\lambda_1(A)$ is the *principal eigenvalue of* A. If I denotes the identity matrix and $A + \alpha I \geqslant 0$, the Perron-Frobenius theorem implies

$$\lambda_1(A) = r(A + \alpha I) - \alpha.$$

Throughout this paper, $r(A)$ and $\lambda_1(A)$ will denote the spectral radius and principal eigenvalue respectively of a matrix A.

If $A = (a_{ij})$ is an essentially nonnegative matrix, define a nonnegative matrix $B = (b_{ij})$ by $b_{ij} = a_{ij}$ for $i \neq j$ and $b_{ii} = 0$, and say that A is irreducible if B is irreducible. We leave it to the reader to check that this definition agrees with the previous one when A is nonnegative. Equivalently, A is irreducible if $A + \alpha I$ is irreducible whenever $A + \alpha I \geqslant 0$.

We need also to recall some definitions. If B is a matrix such that $b_{ij} = 0$ for $i \neq j$, then B will be called a diagonal matrix and we shall write $B = \operatorname{diag}(b_{ii})$; B is a *positive diagonal matrix* if B is diagonal and $b_{ii} > 0$ for $1 \leqslant i \leqslant n$.

If U is a convex subset of $\mathbb{R}^n$ and $f: U \to \mathbb{R}$ is a real-valued function, f is called *convex* if for all vectors x and y in U and all real numbers t such that $0 \leqslant t \leqslant 1$ one has

$$f((1-t)x + ty) \leqslant (1-t)f(x) + tf(y).$$

If $-f(x)$ is convex, $f(x)$ is called concave. If $f(x)$ is positive on U and $\log f(x)$ is convex, $f(x)$ is called log convex. It is known [1, 17] that the sum or product of log convex functions is log convex.

Our first theorem generalizes both Kingman's theorem [17] and Cohen's theorem [6–8]. The theorem also gives necessary and sufficient conditions for strict convexity or strict log convexity and generalizes earlier partial results of this type due to S. Friedland [12]. Our motivation for the following proof comes from an analogous, unpublished trick [21] which has been used by P. L. Lions in studying eigenvalues of second-order elliptic partial differential equations.

THEOREM 1.1. *For* $0 \leqslant t \leqslant 1$, *assume that* $F(t) = (f_{ij}(t))$ *is an* $n \times n$ *nonnegative, irreducible matrix and that* $G(t) = \operatorname{diag}(g_{ii}(t))$ *is an* $n \times n$ *diagonal matrix. For* $1 \leqslant i, j \leqslant n$ *assume that* $f_{ij}(t)$ *is either identically zero or a log convex function of* t *and that* $g_{ii}(t)$ *is a convex function for* $1 \leqslant i \leqslant n$. *Define* $F(0) = A = (a_{ij})$, $F(1) = B = (b_{ij})$, $G(0) = \operatorname{diag}(c_{ii})$, *and*

$G(1) = \mathrm{diag}(d_{ii})$, *and define a matrix* $M(t) = (m_{ij}(t))$ *by*

$$m_{ij}(t) = \begin{cases} (1-t)c_{ii} + td_{ii} + a_{ii}^{1-t}b_{ii}^{t} & \text{for} \quad i = j \\ a_{ij}^{1-t}b_{ij}^{t} & \text{for} \quad i \neq j. \end{cases}$$

Then one has

$$\lambda_1(F(t)+G(t)) \leqslant \lambda_1(M(t)), \qquad 0 \leqslant t \leqslant 1, \tag{1.2}$$

and equality holds in (1.2) *for some t with* $0 < t < 1$ *if and only if*

$$F(t)+G(t) = M(t)$$

for all t with $0 \leqslant t \leqslant 1$. *Furthermore,*

$$\lambda_1(M(t)) \leqslant \max_{1 \leqslant i \leqslant n} \left\{ \left[(1-t)c_{ii} + td_{ii}\right] + \left[\lambda_1(A) - c_{ii}\right]^{1-t}\left[\lambda_1(B) - d_{ii}\right]^{t} \right\} \tag{1.3}$$

and

$$\lambda_1(M(t)) \leqslant (1-t)\lambda_1(M(0)) + t\lambda_1(M(1)). \tag{1.4}$$

Equality occurs in (1.4) *for some t with* $0 < t < 1$ *if and only if there exist a real number c and a positive diagonal matrix E such that*

$$G(1) - G(0) = cI \tag{1.5}$$

and

$$B = E^{-1}AE, \tag{1.6}$$

and then equality holds in (1.4) *for all t with* $0 \leqslant t \leqslant 1$. *If* $G(0) = G(1) = 0$, *one has*

$$\lambda_1(M(t)) \leqslant \lambda_1(A)^{1-t}\lambda_1(B)^{t}, \tag{1.7}$$

and (if $G(0) = G(1) = 0$*) equality holds in* (1.7) *for some t with* $0 < t < 1$ *if and only if there exist a positive constant k and a positive diagonal matrix E* *such that*

$$B = kE^{-1}AE, \tag{1.8}$$

and in this case equality holds in (1.7) *for all* t *with* $0 \leqslant t \leqslant 1$. *The inequalities* (1.2), (1.3), (1.4) *and (for* $G(0)=G(1)=0$*)* (1.7) *are valid even if* $F(t)$ *is not irreducible for* $0 \leqslant t \leqslant 1$.

Proof. By using the comments preceding Theorem 1 and by adding a multiple of the identity to $G(t)$, we can assume that $G(t)$ is nonnegative for $0 \leqslant t \leqslant 1$ and work with the spectral radius instead of λ_1. The convexity assumptions on $f_{ij}(t)$ and $g_{ii}(t)$ imply that, for $0 \leqslant t \leqslant 1$,

$$F(t)+G(t) \leqslant M(t). \tag{1.9}$$

Convexity of the entries of $F(t)+G(t)$ and of $M(t)$ and the fact that $M(0)=F(0)+G(0)$ and $M(1)=F(1)+G(1)$ imply that if equality occurs in (1.9) for some t_0 with $0<t_0<1$, then it occurs for all t with $0 \leqslant t \leqslant 1$. It follows that

$$r(F(t)+G(t)) \leqslant r(M(t)), \qquad 0 \leqslant t \leqslant 1, \tag{1.10}$$

and because $F(t)+G(t)$ is irreducible, if equality occurs in (1.10) for some t_0 with $0<t_0<1$, Lemma 1.2 implies that

$$F(t_0)+G(t_0)=M(t_0)$$

and hence

$$F(t)+G(t)=M(t)$$

for all t with $0 \leqslant t \leqslant 1$.

If $F(t)$ is not irreducible, it can still be approximated by the irreducible matrix $F_\varepsilon(t) \equiv (f_{ij}(t)+\varepsilon)$, $\varepsilon>0$; one then obtains the inequality (1.2) by taking the limit as $\varepsilon \to 0^+$ and using the fact that the map $N \to r(N)$ is continuous on the set of $n \times n$ matrices N.

Define $r_0 = r(M(0))$ and $r_1 = r(M(1))$, and select $u>0$ and $v>0$ such that

$$M(0)u = r_0 u \quad \text{and} \quad M(1)v = r_1 v.$$

For a fixed t, $0 < t < 1$, define $w = u^{1-t}v^t$, i.e., $w_i = u_i^{1-t}v_i^t$ for $1 \leqslant i \leqslant n$. The ith component of $M(t)w$ satisfies

$$(M(t)w)_i = [(1-t)c_{ii} + td_{ii}]\,w_i + \sum_{j=1}^{n} (a_{ij}u_j)^{1-t}(b_{ij}v_j)^t,$$

and Hölder's inequality gives

$$(M(t)w)_i \leqslant [(1-t)c_{ii} + td_{ii}]\,w_i + \Big(\sum_j a_{ij}u_j\Big)^{1-t}\Big(\sum_j b_{ij}v_j\Big)^t. \quad (1.11)$$

According to Hölder's inequality, equality holds in (1.11) if and only if there exists $\gamma_i > 0$ such that

$$b_{ij}v_j = \gamma_i a_{ij}u_j \qquad \text{for} \quad 1 \leqslant j \leqslant n. \quad (1.12)$$

Since $\sum_j a_{ij}u_j = (r_0 - c_{ii})u_i$ and similarly for B, (1.11) gives

$$(M(t)w)_i \leqslant \Big\{(1-t)c_{ii} + td_{ii} + (r_0 - c_{ii})^{1-t}(r_1 - d_{ii})^t\Big\}\,w_i. \quad (1.13)$$

If one defines ρ_t by

$$\rho_t = \max_{1 \leqslant i \leqslant n} \Big\{[(1-t)c_{ii} + td_{ii}] + (r_0 - c_{ii})^{1-t}(r_1 - d_{ii})^t\Big\},$$

the inequality (1.13) gives

$$M(t)w \leqslant \rho_t w, \quad (1.14)$$

so Lemma 1.1 implies

$$r(M(t)) \leqslant \rho_t,$$

which is (1.3). The inequality between geometric and arithmetic means yields

$$(r_0 - c_{ii})^{1-t}(r_1 - d_{ii})^t \leqslant (1-t)(r_0 - c_{ii}) + t(r_1 - d_{ii}) \quad (1.15)$$

with equality if and only if

$$r_0 - c_{ii} = r_1 - d_{ii}. \quad (1.16)$$

Substituting (1.15) in the formula for ρ_t gives

$$\rho_t \leqslant (1-t)r_0 + tr_1$$

and

$$M(t))w \leqslant \left[(1-t)r_0 + tr_1\right]w. \tag{1.17}$$

Furthermore, equality holds in (1.17) if and only if (1.12) and (1.16) are valid for $1 \leqslant i \leqslant n$. Lemma 1.1 implies that

$$r(M(t)) \leqslant (1-t)r_0 + tr_1, \tag{1.18}$$

and equality holds in (1.18) if and only if (1.12) and (1.16) are valid for $1 \leqslant i \leqslant n$. The inequality (1.18) proves (1.4).

We now consider the case of equality in (1.18). If one sums Equation (1.12) over j, one obtains

$$(r_1 - d_{ii})v_i = \gamma_i(r_0 - c_{ii})u_i. \tag{1.19}$$

The irreducibility of A implies $r_0 - c_{ii} > 0$, so (1.16) and (1.19) yield

$$\gamma_i = \frac{v_i}{u_i}.$$

Thus, if we define $E = \text{diag}(u_i/v_i)$, then Equation (1.12) implies

$$B = E^{-1}AE \tag{1.20}$$

and (1.16) gives

$$G(1) - G(0) = (r_1 - r_0)I = cI. \tag{1.21}$$

Conversely, suppose (1.20) holds for a positive diagonal matrix $E = \text{diag}(e_i)$ and (1.21) is valid for some real number c. Select $u > 0$ to be a positive eigenvector of A with corresponding eigenvalue r_0. Define a vector $v > 0$ by $v_i = e_i^{-1}u_i$ and $r_1 = r_0 + c$. If, for any t with $0 < t < 1$, a positive vector w is defined by $w_i = u_i^{1-t}v_i^t$, an easy calculation shows

$$M(t)w = \left[(1-t)r_0 + tr_1\right]w,$$

so

$$r(M(t)) = (1-t)r_0 + tr_1.$$

It remains to consider the case $G(1) = G(0) = 0$. In this case (1.14) and Lemma 1.1 imply

$$r(M_t) \leqslant r_0^{1-t} r_1^t \tag{1.22}$$

which is inequality (1.7). Furthermore, equality holds in (1.22) for some t with $0 < t < 1$ if and only if (1.12) is satisfied for $1 \leqslant i \leqslant n$. Summing Equation (1.12) over j gives

$$r_1 v_i = \gamma_i r_0 u_i, \tag{1.23}$$

and Equations (1.12) and (1.23) give

$$B = kE^{-1}AE, \tag{1.24}$$

where $E = \mathrm{diag}(u_i/v_i)$ and $k = r_1/r_0$.

Conversely, suppose that $G(1) = G(0) = 0$ and that Equation (1.24) is satisfied for some $k > 0$ and some positive diagonal matrix $E = \mathrm{diag}(e_i)$. Select a positive vector $u > 0$ so that $Au = r_0 u$, and define $v > 0$ by $u_i = v_i e_i$ and $r_1 > 0$ by $r_1 = kr_0$. If, for $0 < t < 1$, one then defines $w = u^{1-t}v^t$, a direct calculation gives

$$M(t)w = \left(r_0^{1-t} r_1^t\right)w.$$

It remains to verify that (1.3), (1.4), and (1.7) are valid even if $F(t)$ is not irreducible, but this follows by the same argument used to prove the inequality (1.2) when $F(t)$ is not irreducible. ■

REMARK 1.1. The proof of Theorem 1.1 actually shows that the matrix E in the statement of the theorem must be of the form $E = \mathrm{diag}(u_i/v_i)$, where $u > 0$ is a positive eigenvector of A and $v > 0$ is a positive eigenvector of B. Thus the matrix E is determined uniquely to within positive scalar multiples.

Theorem 1.1 immediately gives Corollary 1.1 below, which is Cohen's theorem [6–8]. The strict convexity in Corollary 1.1 was first obtained in the stated generality by Friedland (see Theorem 4.1 in [12]).

COROLLARY 1.1 (See [7, 8] and [12]). *Assume that* A *is an* $n \times n$, *nonnegative irreducible matrix and that* $C = \operatorname{diag}(c_{ii})$ *and* $D = \operatorname{diag}(d_{ii})$ *are diagonal matrices. For* $0 \leqslant t \leqslant 1$ *define* $G(t) = (1-t)C + tD$. *Then one has for* $0 \leqslant t \leqslant 1$

$$\lambda_1(G(t)+A) \leqslant (1-t)\lambda_1(C+A) + t\lambda_1(D+A), \tag{1.25}$$

and equality holds in (1.25) *for some* t *with* $0 < t < 1$ *if and only if* $D - C$ *is a scalar multiple of the identity.*

Theorem 1.1 also yields Kingman's theorem as an immediate consequence and moreover gives necessary and sufficient conditions for strict convexity. The latter information appears to be new and inaccessible by other proofs.

COROLLARY 1.2 (See [17]). *For* $0 \leqslant t \leqslant 1$ *assume that* $F(t) = (f_{ij}(t))$ *is an* $n \times n$, *nonnegative irreducible matrix, and suppose that for* $1 \leqslant i, j \leqslant n$, $f_{ij}(t)$ *is either identically zero or positive and a log convex function of* t. *It then follows that* $r(F(t))$ *is a log convex function of* t *for* $0 \leqslant t \leqslant 1$. *If* $F(0) = (a_{ij})$, $F(1) = (b_{ij})$, *and a matrix* $M(t) = (m_{ij}(t))$ *is defined by*

$$m_{ij}(t) = a_{ij}^{1-t} b_{ij}^{t},$$

then one has

$$r(F(t)) \leqslant r(M(t)), \tag{1.26}$$

and equality occurs in (1.26) *for some* t *with* $0 < t < 1$ *if and only if* $F(t) = M(t)$ *for all* t *with* $0 \leqslant t \leqslant 1$. *If* $r(A) = r_0$ *and* $r(B) = r_1$, *then*

$$r(M(t)) \leqslant r_0^{1-t} r_1^{t}, \tag{1.27}$$

and equality occurs in (1.27) *for some* t *with* $0 < t < 1$ *if and only if there exists a constant* $k > 0$ *and a positive diagonal matrix* E *such that*

$$B = kE^{-1}AE.$$

REMARK 1.2. In Theorem 4.2 of [12], Friedland studies a special case of Kingman's theorem. Let $H = (h_{ij})$ be a fixed $n \times n$, nonnegative irreducible matrix, and for diagonal matrices G consider

$$R(G) \equiv \log r(e^{G}H).$$

Corollary 1.2 implies that $R(G)$ is a convex function of G. Specifically, if $C = \mathrm{diag}(c_{ii})$ and $D = \mathrm{diag}(d_{ii})$, define $F(t) = e^{(1-t)C + tD}H$ and note that in the notation of Corollary 1.2, $F(t) = M(t)$. Furthermore, if $A = F(0)$, $r_0 = r(A)$, $B = F(1)$, and $r_1 = r(B)$, one obtains from Corollary 1.2 that

$$\log r(F(t)) \leqslant (1-t)\log r(A) + t\log r(B), \tag{1.28}$$

which gives the convexity of $R(G)$. Furthermore, equality holds in (1.28) for some t with $0 < t < 1$ if and only if there is a positive diagonal matrix $E = \mathrm{diag}(e_i)$ and $k > 0$ such that

$$B = \left(e^{d_{ii}}h_{ij}\right) = kE^{-1}AE = \left(ke_i^{-1}e^{c_{ii}}h_{ij}e_j\right). \tag{1.29}$$

If the diagonal entries of H are positive, (1.29) implies that $d_{ii} - c_{ii} = \log k$ for $1 \leqslant i \leqslant n$, or $D - C$ is a scalar multiple of the identity. Conversely, if $D - C$ is a scalar multiple of I, Equation (1.28) clearly becomes an equality for $0 \leqslant t \leqslant 1$. Thus we obtain Friedland's necessary and sufficient conditions for strict convexity in the case he considers.

REMARK 1.3. In some work on matrix theory it is useful to have versions of Corollary 1.2 in which the parameter t is a vector and lies in a convex set C in $\mathbb{R}^n$; see [10], for example. Such versions can be derived from Corollary 1.2, but it may be worthwhile to describe such a result explicitly. Thus, suppose C is a convex subset of $\mathbb{R}^n$ and that, for $1 \leqslant i, j \leqslant n$, $g_{ij}: C \to \mathbb{R}$ is either identically zero or positive and a log convex function of $t \in C$. Define $G(t) = (g_{ij}(t))$, and assume $G(t)$ is irreducible for all $t \in C$. Suppose $t^{(1)}, t^{(2)}, \ldots, t^{(m)}$ are m distinct points in C, define $P_m = \{\theta \in \mathbb{R}^m \mid \theta \geqslant 0$ and $\sum_{i=1}^m \theta_i = 1\}$, and for $\theta \in P_m$ define $H(\theta) = G(\sum_{i=1}^m \theta_i t^{(i)})$. For $1 \leqslant k \leqslant m$ define $A^{(k)} \equiv (a_{ij}^{(k)}) = G(t^{(k)})$, and for $\theta \in P_m$ define $B(\theta) = (b_{ij}(\theta))$ by

$$b_{ij}(\theta) = \prod_{k=1}^{m} \left(a_{ij}^{(k)}\right)^{\theta_k}.$$

Then one has for all $\theta \in P_m$

$$r(H(\theta)) \leqslant r(B(\theta)), \tag{1.30a}$$

and equality holds in (1.30) for some $\theta > 0$, $\theta \in P_m$, if and only if equality holds for all $\theta \in P_m$. If $r_k = r(A^{(k)})$ for $1 \leqslant k \leqslant m$, one has for all $\theta \in P_m$

$$r(B(\theta)) \leqslant \prod_{k=1}^{m} r_k^{\theta_k}. \tag{1.30b}$$

Equality holds in Equation (1.30b) for some $\theta \in P_m$, $\theta > 0$, if and only if there exist positive diagonal matrices E_k for $2 \leqslant k \leqslant m$ and positive constants c_k for $2 \leqslant k \leqslant m$ such that

$$A^{(k)} = c_k E_k^{-1} A^{(1)} E_k.$$

It is perhaps easiest to prove the above result by arguing in analogy with Theorem 1.1. Specifically, select $v^{(k)} > 0$ so

$$A^{(k)} v^{(k)} = r_k v^{(k)};$$

and given $\theta \in P_m$, $\theta > 0$, define $w > 0$, $w \in \mathbb{R}^n$, by

$$w_i = \prod_{k=1}^{m} \left(v_i^{(k)}\right)^{\theta_i},$$

and observe how $B(\theta)$ acts on w. We leave the details to the reader.

We shall now turn to a closely related question posed by Friedland [12]. If H is a fixed $n \times n$, nonnegative matrix, Friedland asks if the map $D \to r(DH)$ is a convex function on the set of positive diagonal matrices. If $H \neq I$ is a permutation matrix, the map $D \to r(DH)$ is not convex (see [12, Section 6]); however, Friedland proves (see Theorem 4.3 in [12]) that if H is invertible and $-H^{-1}$ is essentially nonnegative, then $D \to r(DH)$ is convex on the set of positive diagonal matrices.

Our next theorem will contain Friedland's theorem and "almost" contain Theorem 1.1. To state the theorem we need a definition.

Definition 1.1. Let C be a convex subset of a vector space E, and $f: C \to \mathbb{R}$ a real-valued function. The map f is called "quasiconvex" if for every real number α, the set $\{x \in C \mid f(x) < \alpha\}$ is convex (possibly empty); f is "quasiconcave" if $-f(x)$ is quasiconvex.

Theorem 1.2. *Let notation and assumptions be as in the statement of Theorem 1.1, except do not assume that $F(t)$ is irreducible and suppose that $G(t)$ is a positive diagonal matrix. Let H be an $n \times n$, nonnegative, irreducible, invertible matrix, and suppose that $-H^{-1}$ is essentially nonnegative. Then one has*

$$r(HF(t) + HG(t)) \leqslant r(HM(t)), \qquad 0 \leqslant t \leqslant 1, \tag{1.31}$$

and equality holds in (1.31) *for some t with* $0 < t < 1$ *if and only if*

$$F(t)+G(t)=M(t)$$

for all t with $0 \leqslant t \leqslant 1$. *Furthermore, if* $R(t) \equiv r(HM(t))$, *then* $R(t)$ *is a quasiconvex function of* t, $0 \leqslant t \leqslant 1$; *and this is true even if H is not irreducible. If* $F(t)$ *is identically zero, then* $R(t)$ *is a convex function of* t, *and this is also true even if H is not irreducible.*

If $R(0)=R(1)=\alpha$, *then there exists a* t_0, $0 < t_0 < 1$, *such that* $R(t_0)=\alpha$ *if and only if* $G(0)=G(1)$ *and there exists a positive diagonal matrix E such that*

$$E^{-1}F(0)E = F(1) \tag{1.32a}$$

and

$$E^{-1}HE = H. \tag{1.32b}$$

Proof. It is known (and not hard to prove) that if H is a nonnegative, invertible matrix and $-H^{-1}$ is essentially nonnegative, then H^{-1} can be written in the form

$$H^{-1}=\xi - K,$$

where $K \geqslant 0$ and $\xi > r(K)$. We shall assume from the start that H is irreducible. [If H is not irreducible, one can obtain the desired quasiconvexity or convexity by approximating H by $H_n=(\xi - K_n)^{-1}$, where K_n is a sequence of strictly positive matrices which approach K in norm.]

Because H is nonnegative and irreducible and $G(t)$ is a *positive* diagonal matrix, $HF(t)+HG(t)$ and $HM(t)$ are irreducible and nonnegative. It was observed before that

$$F(t)+G(t) \leqslant M(t),$$

so

$$HF(t)+HG(t) \leqslant HM(t).$$

It follows that the inequality (1.31) holds for $0 \leqslant t \leqslant 1$, and Lemma 1.2

implies that if equality holds for some t_0, $0 \leqslant t_0 \leqslant 1$, then

$$F(t)+G(t)=M(t)$$

for all t, $0 \leqslant t \leqslant 1$.

To prove that $R(t)$ is quasiconvex we must prove that if $R(t_0) \leqslant \alpha$ and $R(t_1) \leqslant \alpha$, then $R(t) \leqslant \alpha$ for $t_0 \leqslant t \leqslant t_1$. It is an easy exercise to see (by reparametrizing) that it suffices to prove this when $t_0 = 0$ and $t_1 = 1$. If u is a positive eigenvector for $HM(0)$ and v is a positive eigenvector for $HM(1)$, one has

$$Cu + Au \leqslant \alpha(\xi - K)u \tag{1.33a}$$

and

$$Dv + Bv \leqslant \alpha(\xi - K)v. \tag{1.33b}$$

For fixed t, $0 < t < 1$, define w by $w_i = u_i^{1-t}v_i^t$ and define $\tilde{M}(t) = (a_{ij}^{1-t}b_{ij}^t)$. By Hölder's inequality one has

$$\begin{aligned}([\alpha K + \tilde{M}(t)]w)_i &= \sum_j (\alpha k_{ij}u_j)^{1-t}(\alpha k_{ij}v_j)^t + \sum_j (a_{ij}u_j)^{1-t}(b_{ij}u_j)^t \\ &\leqslant \left(\sum_j \alpha k_{ij}u_j + \sum_j a_{ij}u_j\right)^{1-t}\left(\sum_j \alpha k_{ij}v_j + \sum b_{ij}v_j\right)^t.\end{aligned} \tag{1.34}$$

Equality holds in (1.34) if and only if there exists a positive constant λ_i such that

$$\lambda_i k_{ij}u_j = k_{ij}v_j \qquad \text{for} \quad 1 \leqslant j \leqslant n \tag{1.35}$$

and

$$\lambda_i a_{ij}u_j = b_{ij}v_j \qquad \text{for} \quad 1 \leqslant j \leqslant n. \tag{1.36}$$

Using (1.33) in (1.34), one obtains

$$([\alpha K + \tilde{M}(t)]w)_i \leqslant (\alpha\xi - c_{ii})^{1-t}(\alpha\xi - d_{ii})^t w_i,$$

and the previous inequality implies

$$\alpha\xi w_i - (\alpha K w)_i - (\tilde{M}(t)w)_i \geqslant \left[\alpha\xi - (\alpha\xi - c_{ii})^{1-t}(\alpha\xi - d_{ii})^t\right] w_i$$
$$\geqslant \left[(1-t)c_{ii} + t d_{ii}\right] w_i. \tag{1.37}$$

Strict inequality holds in (1.37) unless

$$c_{ii} = d_{ii}. \tag{1.38}$$

The inequality (1.37) implies that

$$HM(t)w \leqslant \alpha w. \tag{1.39}$$

Furthermore, a little thought shows that our arguments imply equality holds in (1.39) if and only if equality holds in (1.33a) and (1.33b), and in (1.35), (1.36), and (1.38) for $1 \leqslant i \leqslant n$. Lemma 1.1 thus implies that

$$R(t) = r(HM(t)) \leqslant \alpha, \tag{1.40}$$

which proves quasiconvexity.

Now assume that $R(0) = R(1) = \alpha$ [so u and v in (1.33) are eigenvectors] and that $R(t) = \alpha$ for some t, $0 < t < 1$. By adding Equations (1.35) and (1.36) we obtain

$$\lambda_i\left(\sum_{j=1}^{n} \alpha k_{ij} u_j + \sum_{j=1}^{n} a_{ij} u_j\right) = \lambda_i(\alpha\xi - c_{ii})u_i$$
$$= \left(\sum_{j=1}^{n} \alpha k_{ij} v_j + \sum_{j=1}^{n} b_{ij} v_j\right)$$
$$= (\alpha\xi - d_{ii})v_i.$$

If $\alpha\xi - c_{ii} > 0$, Equation (1.38) implies $\lambda_i = v_i/u_i$. However, Equation (1.33) implies $\alpha\xi - d_{ii} > 0$ unless $b_{ij} = k_{ij} = 0$ for $1 \leqslant j \leqslant n$. Thus, even if $\alpha\xi - c_{ii} = 0$, Equations (1.35) and (1.36) are still satisfied if one defines $\lambda_i = v_i/u_i$. With this convention, if we define $E = \mathrm{diag}(u_i/v_i)$, we obtain

$$E^{-1}AE = B,$$

which is Equation (1.31), and

$$E^{-1}KE = K.$$

The latter equation implies

$$E^{-1}(I-K)E = I-K$$

and

$$E^{-1}HE = H.$$

Conversely, if $R(0)=R(1)=\alpha$, $G(0)=G(1)$, and Equations (1.31) and (1.32) are satisfied, let u be a positive eigenvector of $HM(0)$. If v is defined by $\operatorname{diag}(u_i/v_i)=E$, one verifies that v is a positive eigenvector of $HM(1)$. If w is defined by $w_i = u_i^{1-t}v_i^t$, one works backwards through the above inequalities to verify that

$$HM(t)w = \alpha w, \qquad 0 \leqslant t \leqslant 1.$$

We leave the details to the reader.

It remains to show that $R(t)$ is a convex function of t if $F(t)\equiv 0$. Initially, we do not assume $F(t)\equiv 0$. If

$$r(HM(0)) = r_0,$$
$$r(HM(1)) = r_1,$$

then select s, $0<s<1$, and define $\lambda_0 = (1/r_0)[(1-s)r_0 + sr_1]$ and $\lambda_1 = (1/r_1)[(1-s)r_0+sr_1]$. Define $\tilde{A}=\lambda_0 A$, $\tilde{C}=\lambda_0 C$, $\tilde{B}=\lambda_1 B$, and $\tilde{D}=\lambda_1 D$. We then have

$$r(H(\tilde{A}+\tilde{C})) = \alpha = (1-s)r_0 + sr_1 = r(H(\tilde{B}+\tilde{D})),$$

and quasiconvexity implies that for $0\leqslant t\leqslant 1$,

$$r\left(H\left[(1-t)\tilde{C} + t\tilde{D} + \lambda_0^{1-t}\lambda_1^t \tilde{M}(t)\right]\right) \leqslant (1-s)r_0 + sr_1,$$

where $\tilde{M}(t) = (a_{ij}^{1-t}b_{ij}^t)$. If we take $t = s/\lambda_1$ [so $1-t = (1-s)/\lambda_0$], a calculation gives

$$r\left(H\left[(1-s)C + sD + \lambda_0^{1-t}\lambda_1^t\tilde{M}(t)\right]\right) \leqslant (1-s)r_0 + sr_1.$$

If $F(t)\equiv 0$, the previous inequality becomes

$$r(H[(1-s)C+sD])\leqslant(1-s)r(HC)+sr(HD), \tag{1.41}$$

where C and D are positive diagonal matrices, and this is the desired convexity result. ■

REMARK 1.4. Theorem 1.2 immediately implies that $t\to r(HF(t)+HG(t))$ is quasiconvex.

If $F(t)\equiv 0$ and if equality holds in (1.41) for some s with $0<s<1$, then in the notation of the above proof and with $t=s/\lambda_1$

$$r(H\tilde{C})=r(H[(1-t)\tilde{C}+t\tilde{D}])=r(H\tilde{D}).$$

Our theorem for strict convexity implies $\tilde{C}=\tilde{D}$ or $D=\lambda C$, $\lambda>0$. Conversely, if $D=\lambda C$ for some $\lambda>0$, one clearly has equality in (1.41) for $0\leqslant s\leqslant 1$. Thus we obtain Friedland's necessary and sufficient condition for strict convexity (see Theorem 4.3 in [12]).

2. MINIMAX FORMULAS FOR THE SPECTRAL RADIUS

If A is an $n\times n$ matrix, write A^T for the transpose of A, and denote by P the set of probability vectors in $\mathbb{R}^n$, $P=\{\alpha\in\mathbb{R}^n\,|\,\alpha\geqslant 0,\ \Sigma\alpha_j=1\}$, and by P_0 the set $\{\alpha\in P\,|\,\alpha_i>0 \text{ for } 1\leqslant i\leqslant n\}$. If A is a given $n\times n$, nonnegative matrix, we shall consider in this section variants of the function $f(\alpha,x)$ given by

$$f(\alpha,x)=\sum_{i=1}^{n}\alpha_i\log\left(\frac{(Ax)_i}{x_i}\right) \tag{2.1}$$

for $(\alpha,x)\in P\times P_0$ or for $(\alpha,x)\in\{(\alpha,x):\alpha\in P,\ x>0\}$. For a given $\alpha\in P$ one can consider the map

$$x\to f(\alpha,x)=g(x) \tag{2.2}$$

and consider g as defined on $\{x>0\}$ or on P_0. Since g is homogeneous of degree zero, any critical point x of $g\,|\,P_0$ actually satisfies $\nabla g(x)=0$.

If A is an $n \times n$ nonnegative matrix and $r = r(A)$, there exist nonnegative (nonzero) vectors u and v such that

$$Au = ru \tag{2.3}$$

and

$$A^T v = rv. \tag{2.4}$$

If A is also irreducible, then necessarily $u > 0$ and $v > 0$, and u and v can be chosen so

$$\sum_{i=1}^{n} u_i v_i = 1. \tag{2.5}$$

However, even if A is not irreducible, it may happen that $u > 0$, and then (2.5) can also be satisfied.

Our first theorem presents a simple proof of a result which was first proved by Friedland and Karlin (see [13, Section 3]) and which plays a central role in [12].

THEOREM 2.1 (Friedland and Karlin [13]). *Let $A = (a_{ij})$ be a nonnegative, irreducible matrix such that $a_{ii} > 0$ for $1 \leqslant i \leqslant n$. If $\alpha \in P_0$ and $g(x)$ is defined by Equation* (2.2), *then $g \mid P_0$ has a unique critical point ξ and*

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

In particular, if u and v satisfy Equations (2.3), (2.4), *and* (2.5) *and if α is defined by*

$$\alpha = (u_1 v_1, u_2 v_2, \dots, u_n v_n), \tag{2.6}$$

then

$$\sum_{i=1}^{n} u_i v_i \log\left(\frac{(Ax)_i}{x_i} \right) \geqslant \sum_{i=1}^{n} u_i v_i \log\left(\frac{(Au)_i}{u_i} \right) = \log r(A) \tag{2.7}$$

for all vectors $x > 0$.

Proof. As is observed in [12, 13], if α is given by (2.6), the function $g(x)$ has a critical point at $x = u$ (this is an easy calculation), and Equation (2.7) then follows from the first part of the theorem.

Thus we concentrate on proving the first part of the theorem. Select $\alpha \in P_0$, and let g be given by Equation (2.2). We first claim that g achieves its minimum on P_0. To prove this it suffices to prove that if $z \in P$ and $z \notin P_0$, then $\lim_{x \to z, x > 0} g(x) = \infty$. We know that $(Ax)_i / x_i \geqslant a_{ii} > 0$ for any $x \in P_0$, so we shall be done if we can find an index i such that $z_i = 0$ and $(Az)_i > 0$. Let $S = \{i \mid 1 \leqslant i \leqslant n,\ z_i = 0\}$; we claim that $(Az)_i > 0$ for some $i \in S$. If not, for each $k \notin S$ multiply row k of A by $\lambda_k > 0$ such that $\lambda_k (Az)_k = z_k$. This gives a new matrix B which is also irreducible but which has a nonnegative eigenvector z which is not strictly positive, thus contradicting Lemma 1.1. It follows that f has a minimum on P_0, say at ξ.

Given vectors $x > 0$ and $y > 0$ and a real number t, define a positive vector $w = x^{1-t} y^t$ by $w_i = x_i^{1-t} y_i^t$. We next claim that for $0 \leqslant t \leqslant 1$

$$g(x^{1-t} y^t) \leqslant (1-t) g(x) + t g(y) \tag{2.8}$$

and that equality holds in (2.8) for some t with $0 < t < 1$ if and only if there exists a positive real λ such that $y = \lambda x$. To see this, observe that for a fixed t with $0 < t < 1$ Hölder's inequality implies

$$\begin{aligned}(A(x^{1-t} y^t))_i &= \sum_{j=1}^{n} (a_{ij} x_j)^{1-t} (a_{ij} y_j)^t \\ &\leqslant \left(\sum_{j=1}^{n} a_{ij} x_j \right)^{1-t} \left(\sum_{j=1}^{n} a_{ij} y_j \right)^t. \end{aligned} \tag{2.9}$$

Equality holds in (2.9) if and only if there exists a positive number λ_i such that

$$a_{ij} y_j = \lambda_i a_{ij} x_j, \qquad 1 \leqslant j \leqslant n. \tag{2.10}$$

Because $\alpha > 0$, (2.9) and (2.10) imply that

$$\begin{aligned}\sum_{i=1}^{n} \alpha_i \left[\log (A(x^{1-t} y^t))_i - (1-t) \log x_i - t \log y_i \right] &= g(x^{1-t} y^t) \\ &\leqslant (1-t) g(x) + t g(y), \end{aligned} \tag{2.11}$$

and that equality holds in (2.11) if and only if Equation (2.10) is satisfied for $1 \leqslant i \leqslant n$.

Since we assume $a_{ii} > 0$, taking $j = i$ in (2.10) gives $\lambda_i = y_i/x_i$, and (2.10) then implies

$$\lambda_j \equiv \frac{y_j}{x_j} = \frac{y_i}{x_i} \equiv \lambda_i \qquad \text{if} \quad a_{ij} > 0. \tag{2.12}$$

If i and j are arbitrary indices, the irreducibility of A implies that there are indices $j_1, j_2, \ldots, j_p$ such that

$$a_{ij_1} a_{j_1 j_2} a_{j_2 j_3} \cdots a_{j_p j} > 0,$$

and we obtain from this fact and (2.12) that

$$\lambda_i = \lambda_{j_1} = \lambda_{j_2} = \cdots = \lambda_{j_p} = \lambda_j, \tag{2.13}$$

i.e., $y = \lambda x$ for some $\lambda > 0$.

The above observations show that the point $\xi \in P_0$ where g achieves its minimum is unique. For suppose g achieves its minimum also at $\bar{\xi} \in P_0$ and $\bar{\xi} \neq \xi$. Then one has

$$g(\xi) = \min_{x > 0} g(x),$$

but $g((\xi\bar{\xi})^{1/2}) < g(\xi)$, a contradiction.

We know that $\nabla g(\xi) = 0$, but it remains to show that if $x \in P_0$ and $x \neq \xi$, then $\nabla g(x) \neq 0$. If $\nabla g(x) = 0$ and we defined $x_t = \xi^t x^{1-t}$, we would have

$$\lim_{t \to 0^+} \frac{g(x) - g(x_t)}{\|x - x_t\|} = 0.$$

However, Equation (2.8) implies that

$$g(x) - g(x_t) \geqslant t\,[g(x) - g(\xi)],$$

and the differentiability of $t \to x_t$ implies that

$$\|x - x_t\| \leqslant Mt$$

for some constant M. Thus we have

$$\lim_{t \to 0^+} \frac{g(x) - x(x_t)}{\|x - x_t\|} \geqslant \frac{1}{M}[g(x) - g(\xi)] > 0,$$

and the assumption that $\nabla g(x) = 0$ was wrong. ∎

REMARK 2.1. Suppose that $\psi: \mathbb{R} \to \mathbb{R}$ is a map such that $\lim_{t \to -\infty} \psi(t) \equiv \psi(-\infty)$ exists either as a real number or in the sense that $\lim_{t \to -\infty} \psi(t) = +\infty$ or $\lim_{t \to -\infty} \psi(t) = -\infty$. If A is an arbitrary $n \times n$ nonnegative matrix, $\alpha \in P$, $x > 0$, and ψ is as above, we shall want to make sense of

$$h(\alpha, x) = \sum_{i=1}^{n} \alpha_i \psi\left(\log\left(\frac{(Ax)_i}{x_i}\right)\right)$$

even if $(Ax)_i = 0$ for some i. If $\psi(-\infty)$ is finite and $(Ax)_i = 0$, define $\alpha_i \psi(\log[(Ax)_i/x_i]) = \alpha_i \psi(-\infty)$. If $\psi(-\infty) = \pm\infty$, $(Ax)_i = 0$ and $\alpha_i > 0$, define $h(\alpha, x) = \psi(-\infty)$; but if $\alpha_i = 0$, define $\alpha_i \psi(\log[(Ax)_i/x_i]) = 0$, no matter what the value of $(Ax)_i$. These conventions give a well-defined value of $h(\alpha, x)$ and, in particular [taking $\psi(t) = t$], of $f(\alpha, x)$. An examination of the proof of Theorem 2.1 shows that the inequality (2.8) is still true in this generality.

It turns out that the inequality (2.7) is true under less restrictive assumptions. Since the argument we shall give also has the virtue of extending to the case of partial differential equations, we present it here.

THEOREM 2.2. *Let A be a nonnegative matrix such that A has an eigenvector $u > 0$, $Au = ru$. (It is then necessarily true that $r = r(A)$.) Let v be a nonnegative eigenvector of A^T such that v satisfies Equations* (2.4) *and* (2.5). *The inequality* (2.7) *is then satisfied for all $x > 0$.*

Proof. If A is identically zero, (2.7) is trivially satisfied (both sides are $-\infty$). If A is not identically zero, then $r > 0$ and the assumption on A implies $Ax > 0$ for $x > 0$. If $\alpha_i \equiv u_i v_i$ and $g(x)$ is defined by

$$g(x) = \sum_{i=1}^{n} \alpha_i \log\left(\frac{(Ax)_i}{x_i}\right),$$

it follows that $g(x)$ is finite for vectors $x > 0$. As before, a calculation gives

$$\nabla g(u) = 0 \quad \text{and} \quad g(u) = \log r.$$

Assume there exists $\xi > 0$ such that $g(\xi) < g(u)$, and for $0 \leqslant t \leqslant 1$ define $u_t = \xi^t u^{1-t} \equiv (\xi_1^{1-t} u_1^{1-t}, \dots, \xi_n^{1-t} u_n^{1-t})$. We know that g satisfies the inequality (2.8) (see Remark 2.1), so exactly the argument at the end of the proof of Theorem 2.1 shows that $\nabla g(u) \neq 0$, a contradiction. ■

REMARK 2.2. The hypotheses of Theorem 2.2 may be satisfied for matrices A which are not irreducible and for which A^T does not have an eigenvector $v > 0$, e.g., for

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

As an immediate consequence of Theorem 2.2, we obtain the following corrected variant of Theorem 3.3 in [12]. Notice that we do *not* claim that Equation (2.14b) below is true for arbitrary nonnegative matrices A (compare Theorem 3.3 in [12]). In fact, we show in Remark 2.3 below that for general A Equation (2.14b) is false.

THEOREM 2.3 (Compare Friedland [12]). *Assume that A is an $n \times n$, nonnegative matrix which has an eigenvector $u > 0$, so $Au = ru$ (necessarily $r = r(A)$). Let $v \geqslant 0$ satisfy Equations* (2.4) *and* (2.5). *If $\psi : \mathbb{R} \to \mathbb{R}$ is a continuous, convex function which is nondecreasing on* $[\log r, \infty)$, *and if we define $\phi(t) = \psi(\log t)$, then*

$$\inf_{x>0} \sum_{i=1}^{n} u_i v_i \phi\left(\frac{(Ax)_i}{x_i} \right) = \phi(r) \tag{2.14a}$$

and

$$\sup_{\alpha \in P} \inf_{x>0} \sum \alpha_i \phi\left(\frac{(Ax)_i}{x_i} \right) = \phi(r). \tag{2.14b}$$

Proof. We proved in Theorem 2.2 that

$$\inf_{x>0} \sum_{i=1} u_i v_i \log\left(\frac{(Ax)_i}{x_i} \right) = \log r, \tag{2.15}$$

and because ψ is nondecreasing on $[\log r, \infty)$ and convex, Equation (2.15) implies for any $x > 0$

$$\sum_{i=1}^{n} u_i v_i \psi\left(\log\left(\frac{(Ax)_i}{x_i}\right)\right) \geqslant \psi\left(\sum_{i=1}^{n} u_i v_i \log\left(\frac{(Ax)_i}{x_i}\right)\right) \geqslant \phi(r). \quad (2.16)$$

Since one obtains equality in (2.16) for $x = u$, we have proved Equation (2.14a).

Equation (2.14a) immediately implies

$$\sup_{\alpha \in P} \inf_{x > 0} \sum_{i=1}^{n} \alpha_i \phi\left(\frac{(Ax)_i}{x_i}\right) \geqslant \phi(r)$$

by taking $\alpha = (u_i v_i, \dots, u_n v_n)$. On the other hand, for any $\alpha \in P$ one has

$$\inf_{x > 0} \sum \alpha_i \phi\left(\frac{(Ax)_i}{x_i}\right) \leqslant \sum \alpha_i \phi\left(\frac{(Au)_i}{u_i}\right) = \phi(r),$$

and the preceding two inequalities give (2.14b). ∎

REMARK 2.3. It is tempting to conjecture that if ψ satisfies the assumptions of Theorem 2.3 and A is an arbitrary nonnegative matrix, then Equation (2.14b) is satisfied. However, if

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix},$$

so $r = r(A) = 2$ and ψ is any convex function on $\mathbb{R}$ such that ψ is increasing on $[\log 2, \infty)$ but $\psi(0) = \phi(1) > \psi(\log 2) = \phi(2)$, one can easily check that

$$\inf_{x > 0} \inf_{\alpha \in P} \sum \alpha_j \phi\left(\frac{(Ax)_j}{x_j}\right) = \sup_{\alpha \in P} \inf_{x > 0} \sum \alpha_j \phi\left(\frac{(Ax)_j}{x_j}\right) = \psi(0) > \psi(\log r).$$

A classical (and fairly easy) result of Wielandt [32] asserts that for any nonnegative matrix A,

$$\inf_{x > 0} \max_{1 \leqslant i \leqslant n} \left(\frac{(Ax)_i}{x_i}\right) = r(A) \equiv r.$$

If $\psi:\mathbb{R}\to\mathbb{R}$ is continuous and nondecreasing on $[\log r,\infty)$, one assigns an arbitrary value for $\phi(-\infty)$, and one defines $\phi(t)=\psi(\log t)$ for $t\geqslant 0$, then one can derive easily that for any vector $x>0$ one has

$$\max_{1\leqslant i\leqslant n}\phi\left(\frac{(Ax)_i}{x_i}\right)\geqslant\phi(r(A)).$$

Because

$$\max_{1\leqslant i\leqslant n}\phi\left(\frac{(Ax)_i}{x_i}\right)=\max_{\alpha\in P}\sum_{i=1}^{n}\alpha_i\phi\left(\frac{(Ax)_i}{x_i}\right),$$

one concludes that

$$\inf_{x>0}\max_{\alpha\in P}\sum\alpha_i\phi\left(\frac{(Ax)_i}{x_i}\right)\geqslant\phi(r(A)).$$

If $Au=ru$ for some $u>0$ (e.g., if A is irreducible), and if one takes $x=u$ in the previous inequality, one sees that

$$\inf_{x>0}\max_{\alpha\in P}\sum\alpha_i\phi\left(\frac{(Ax)_i}{x_i}\right)=\phi(r(A)). \tag{2.17}$$

If A is an arbitrary nonnegative matrix, define J to be a matrix all of whose entries are 1 and $A_\varepsilon\equiv A+\varepsilon J$. *Assume in addition that* $\psi:\mathbb{R}\to\mathbb{R}$ *is nondecreasing (not just on* $[\log r,\infty)$*) and that* $\psi(-\infty)=\lim_{t\to-\infty}\psi(t)$. A simple monotonicity argument then gives

$$\begin{aligned}\phi(r(A))&\leqslant\inf_{x>0}\max_{\alpha\in P}\sum\alpha_i\phi\left(\frac{(Ax)_i}{x_i}\right)\\&\leqslant\inf_{x>0}\max_{\alpha\in P}\sum\alpha_i\phi\left(\frac{(A_\varepsilon x)_i}{x_i}\right)=\phi(r(A_\varepsilon)),\end{aligned}$$

and taking the limit as $\varepsilon\to0^+$, one finds that Equation (2.17) holds for general $A\geqslant0$ if ψ is nondecreasing on $\mathbb{R}$.

Theorem 2.3 and the above remarks yield

PROPOSITION 2.1. *Let A be an $n\times n$ nonnegative matrix with $r=r(A)$. Let $\psi:\mathbb{R}\to\mathbb{R}$ be a continuous function which is nondecreasing on* $[\log r,\infty)$, *and define*

$$\psi(-\infty) = \liminf_{t \to -\infty} \psi(t) \qquad \textit{(possibly} \quad \psi(-\infty) = \pm\infty).$$

If $\phi(t) = \psi(\log t)$,

$$\inf_{x>0} \sup_{\alpha \in P} \sum \alpha_i \phi\left(\frac{(Ax)_i}{x_i}\right) \geqslant \phi(r), \tag{2.18}$$

and if A *possesses an eigenvector* $u > 0$ *or* ψ *is nondecreasing on* $\mathbb{R}$, *one has*

$$\inf_{x>0} \max_{\alpha \in P} \sum \alpha_i \phi\left(\frac{(Ax)_i}{x_i}\right) = \phi(r), \tag{2.19}$$

but in general strict inequality can hold in (2.18). *If* ψ *is convex on* R *and nondecreasing on* $[\log r, \infty)$ *and* A *has an eigenvector* $u > 0$, *then*

$$\inf_{x>0} \sup_{\alpha \in P} \sum \alpha_i \phi\left(\frac{(Ax)_i}{x_i}\right) = \sup_{\alpha \in P} \inf_{x>0} \sum \alpha_i \phi\left(\frac{(Ax)_i}{x_i}\right) = \phi(r). \tag{2.20}$$

Friedland remarks (see [12, p. 306]) that for $\psi(t) = t$ Equation (2.14b) can be proved easily for all $A \geqslant 0$ once one knows it for $A > 0$, because the left-hand side of (2.14b) is a continuous function of A. In fact, this point is nontrivial, and we would like to sketch a proof of Equation (2.14b) for arbitrary $A \geqslant 0$ and for arbitrary nondecreasing convex functions ψ. We shall use a theorem of Sion [30] (the same theorem was also used in [11]); Friedland has shown the author a purely matrix-theoretic (but also nontrivial) proof for the case $\psi(t) = t$.

To begin, let $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ denote the two-point compactification of $\mathbb{R}$. If Z is a convex subset of a topological vector space and $g: Z \to \overline{\mathbb{R}}$, define g to be quasiconvex or quasiconcave by using Definition 1.1. Define g to be lower semicontinuous if $\{z \in Z \mid g(z) > c\}$ is open for all real numbers c and define g to be upper semicontinuous if $-g$ is lower semicontinuous. If $g: Z \to \mathbb{R}$ we use the same definition of upper or lower semicontinuity. Sion [30, Corollary 3.3] has proved the following lemma in the case that the map $f(x, y)$ below is real-valued.

LEMMA 2.1. *Assume that* X *and* Y *are convex subsets of topological vector spaces* E *and* F *respectively and that at least one of* X *or* Y *is compact. Let* $f: X \times Y \to \overline{\mathbb{R}} \equiv \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ *be a map such that* (1) *for each* $x \in X$, *the map* $y \to f(x, y)$ *is quasiconvex and lower semicontinuous and* (2) *for each $y \in Y$ the map $x \to f(x, y)$ is quasiconcave and upper semicontinuous. Then one has*

$$\sup_x \inf_y f(x, y) = \inf_y \sup_x f(x, y).$$

Proof. Sion has proved this lemma if f is real-valued, so it suffices to reduce to this case. For a positive integer N define a retraction $\rho_N : \overline{\mathbb{R}} \to [-N, N]$ by $\rho_N(t) = N$ if $t \geqslant N$, $\rho_N(t) = t$ if $-N \leqslant t \leqslant N$, and $\rho_N(t) = -N$ if $t \leqslant N$. If Z is a convex subset of a topological vector space and $g: Z \to \overline{\mathbb{R}}$ is lower semicontinuous (respectively upper semicontinuous) and quasiconvex (respectively quasiconcave), then it is a straightforward exercise which we leave to the reader to prove that $\rho_N \circ g$ is lower semicontinuous and quasiconvex (respectively upper semicontinuous and quasiconcave). If we define $f_N(x, y)$ by $f_N(x, y) = \rho_N(f(x, y))$ and apply the previous comment, Sion's theorem implies

$$\sup_x \inf_y f_N(x, y) = \inf_y \sup_x f_N(x, y). \tag{2.21}$$

For any function $f(x, y)$ one has

$$\sup_x \inf_y f(x, y) \leqslant \inf_y \sup_x f(x, y), \tag{2.22}$$

and we shall use (2.21) to obtain equality in (2.22). If the right-hand side of (2.22) is $-\infty$, we are done, so assume

$$\inf_y \sup_x f(x, y) = \sigma$$

where $-\infty < \sigma < \infty$. Choose $N > -\sigma$, and observe that for any $y \in Y$ one has then

$$\sup_x f(x, y) \geqslant \sup_x f_N(x, y) \geqslant \min(\sigma, N). \tag{2.23}$$

Using Equations (2.21) and (2.23) one obtains

$$\inf_y \sup_x f_N(x, y) = \sup_x \inf_y f_N(x, y) \geqslant \min(\sigma, N),$$

and (since $-N < \sigma$) one derives from the previous equation that

$$\sup_x \inf_y f(x, y) \geqslant \min(\sigma, N).$$

The desired result follows by letting $N \to \infty$. ■

Now let $\psi:\mathbb{R}\to\mathbb{R}$ be a nondecreasing convex function, and extend $\psi:\overline{\mathbb{R}}\to\overline{\mathbb{R}}$ by $\psi(\pm\infty)=\lim_{x\to\pm\infty}\psi(x)$. If $z\in\mathbb{R}^n$, define $e^z=(e^{z_1},e^{z_2},\dots,e^{z_n})$, and for a given nonnegative matrix A and $(\alpha,z)\in P\times\mathbb{R}^n$ define $h(\alpha,z)$ by

$$h(\alpha,z)=\sum\alpha_i\phi\left(\frac{(Ae^z)_i}{e^{z_i}}\right), \tag{2.24}$$

where $\phi(t)\equiv\psi(\log t)$ and $h(\alpha,z)$ is defined as in Remark 2.1.

THEOREM 2.4. *Assume $\psi:\mathbb{R}\to\mathbb{R}$ is a nondecreasing convex function, and define $\psi(\pm\infty)=\lim_{t\to\pm\infty}\psi(t)$ and $\phi(t)=\psi(\log t)$ for $t\geqslant 0$. If A is an $n\times n$, nonnegative matrix, then*

$$\sup_{\alpha\in P}\inf_{x>0}\sum_{i=1}^{n}\alpha_i\phi\left(\frac{(Ax)_i}{x_i}\right)=\inf_{x>0}\sup_{\alpha\in P}\sum_{i=1}^{n}\alpha_i\phi\left(\frac{(Ax)_i}{x_i}\right)=\phi(r(A)). \tag{2.25}$$

Proof. It suffices to prove that

$$\sup_{\alpha\in P}\inf_{z\in\mathbb{R}^n}h(\alpha,z)=\inf_{z\in\mathbb{R}^n}\sup_{\alpha\in P}h(\alpha,z), \tag{2.26}$$

where h is given by Equation (2.24). Proposition 2.1 implies that the right-hand side of (2.26) equals $\phi(r(A))$.

The inequality (2.9) is valid for $A\geqslant 0$, and one easily concludes that for $0<t<1$ and $z,w\in\mathbb{R}^n$ and $\alpha\in P$,

$$h(\alpha,(1-t)z+tw)\leqslant(1-t)h(\alpha,z)+th(\alpha,w),$$

so $z\to h(\alpha,z)$ is certainly quasiconvex. The facts that $\alpha\to h(\alpha,z)$ is quasiconcave and upper semicontinuous and $z\to h(\alpha,z)$ is lower semicontinuous are straightforward but tedious, and we leave them to the reader. Some care is necessary because of possible $-\infty$ values. The conclusion of Theorem 2.4 follows now from Lemma 2.1. ∎

The cases of greatest interest in Theorem 2.4 are $\psi(t)=t$ and $\phi(t)=\log t$ or $\psi(t)=e^t$ and $\phi(t)=t$.

REMARK 2.4. *If ψ is convex and nonincreasing on $\mathbb{R}$, one obtains*

$$\psi(\log r(A)) = \inf_{\alpha \in P} \sup_{x > 0} \sum_{i=1}^{n} \alpha_i \psi\left(\log\left(\frac{(Ax)_i}{x_i}\right)\right)$$

$$= \sup_{x > 0} \inf_{\alpha \in P} \sum_{i=1}^{n} \alpha_i \psi\left(\log\left(\frac{(Ax)_i}{x_i}\right)\right),$$

because then the function $h(\alpha, z)$ in Equation (2.24) *is quasiconcave in z.*

3. CONVEXITY THEOREMS FOR THE SPECTRAL RADIUS OF NONLINEAR, HOMOGENEOUS CONE MAPS

If X is a real Banach space and C is a subset in X, we shall say that C is a cone (with vertex at 0) if C is closed and convex and (1) if $x \in C$ and t is any nonnegative real, then $tx \in C$, and (2) if $x \in C - \{0\}$, then $-x \notin C$. The cone C induces a partial order on X by $x \leqslant y$ if and only if $y - x \in C$. If $f: C \to (C$ and $f(0) = 0$, by an eigenvector of f is meant a vector $x \in C - \{0\}$ such that $f(x) = \lambda x$ for some real number λ (the eigenvalue), and by a nonzero fixed point of f is meant an eigenvector with eigenvalue $\lambda = 1$. There is an enormous literature concerning the existence of fixed points or eigenvalues of nonlinear maps of cones: see [18], [19, Section 9], [22, Section 1], [23, Section 5], [25], [4], [27] and [31], [33], and [34] for example.

One can also ask somewhat more delicate questions than just the existence of eigenvectors. If $f: C \to C$ is continuous, define f to be *homogeneous of degree one* if for all real numbers t and all $x \in C$, $f(tx) = tf(x)$, and define f to be *order-preserving* (*with respect to* C) if for all $x, y \in C$ such that $x \leqslant y$ one has $f(x) \leqslant f(y)$. Krein and Rutman [19] use the term "monotonic" instead of "order-preserving." If $f: C \to C$ is continuous and the image of any bounded set in C under f has compact closure, f is called a *compact map*. If $f: C \to C$ is a compact map (or, more generally, a "condensing map" [23]), homogeneous of degree 1 and order-preserving, one can define the spectral radius of f (with respect to C), $r(f)$. If f has an eigenvector in C, then

$$r(f) = \sup\{\lambda \mid f(x) = \lambda x \text{ for some } x \in C,\ x \neq 0\};$$

and for general compact $f: C \to C$ as above, $r(f)$ is defined by approximating f by compact maps f_n of the same type such that f_n has an eigenvector in C and defining

$$r(f) = \lim_{n \to \infty} r(f_n).$$

(The existence of appropriate f_n follows by results in Section 9 of [19] or, for more general noncompact f, from results in Section 5 of [23].) Having defined $r(f)$, one can ask how $r(f_t)$ varies with a parameter t if f_t depends in an appropriate way on t and $f_t: C \to C$. We shall prove in this section direct analogues of Cohen's theorem and Kingman's theorem for nonlinear maps and also of the "min-max" and "max-min" formulas for the spectral radius.

If x and y are elements of C, define x and y to be *comparable* if there exist positive numbers α and β such that

$$\alpha x \leqslant y \leqslant \beta x.$$

If $u \in C - \{0\}$, define $C(u) = \{x \in C \mid x \text{ is comparable to } u\}$. If $f: C \to C$ is compact, homogeneous of degree 1, and order-preserving and f has an eigenvector in $C(u)$, one can ask whether this eigenvector is unique (to within scalar multiples). If the eigenvector v in $C(u)$ is unique and if $f(C(u)) \subset C(u)$, then one can define $g: C(u) \to C(u)$ by $g(x) = f(x)/\|f(x)\|$, and one can ask whether $\lim_{n \to \infty} g^n(x) = v/\|v\|$ for any $x \in C(u)$ (where g^n denotes the iteration of g with itself n times). We shall consider such questions in Section 4 below.

In this section and the next section we shall usually restrict ourselves for simplicity to $X = \mathbb{R}^n$, and to the cone $K = \{x \in \mathbb{R}^n \mid x \geqslant 0\}$. We wish to emphasize, however, that much of what we shall prove has extensions to the general framework just described. This is particularly so if $X = C(M)$, the continuous real-valued functions on a compact space M, or $X = L^p(\Omega, \mu)$, $1 \leqslant p < \infty$, where Ω is a σ-finite measure space with measure μ; C is taken to be the cone of nonnegative functions in either case. In fact, it will be useful if the reader considers $\mathbb{R}^n$ as $C(M)$, where $M = \{1, 2, \dots, n\}$, i.e., continuous real-valued functions on M. Thus $\mathbb{R}^n$ is an algebra, and if $u, v \in \mathbb{R}^n$, then $uv = w$, where $w_i = u_i v_i$, $1 \leqslant i \leqslant n$. If t is a nonnegative real and $u \in K$, then $u^t = w$, where $w_i \equiv u_i^t$; and if $u, v \in K$ and $0 \leqslant t \leqslant 1$, then $u^{1-t}v^t = w$, where $w_i = u_i^{1-t}v_i^t$.

We shall always write K for the cone of nonnegative vectors in $\mathbb{R}^n$, and $K_0 = \{x \in \mathbb{R}^n : x_i > 0 \text{ for } 1 \leqslant i \leqslant n\}$, i.e., the vectors in K which are comparable to $(1, 1, \dots, 1) = u$. If $x \in \mathbb{R}^n$, we shall always write $|x| \equiv \sum_{i=1}^n |x_i|$, $P = \{x \in K \mid |x| = 1\}$ and $P_0 = \{x \in K_0 \mid |x| = 1\}$. We shall say $x \in K$ is a positive vector if $x > 0$, i.e., if $x \in K_0$. To save repetition, we make the following definition.

DEFINITION 3.1. We shall say that a continuous map $f: K \to K$ satisfies hypothesis H1 if f is homogeneous of degree one and order-preserving.

Our first lemma is well known, but we sketch a proof for completeness.

LEMMA 3.1. *If $f: P \to K$ is a continuous map, f has an eigenvector.*

Proof. If $f(x)=0$ for some $x \in P$, x is an eigenvector with eigenvalue 0. If $f(x) \neq 0$ for all $x \in P$, the map $g(x) = f(x)/|f(x)|$ is a continuous map of P into itself. The Brouwer fixed-point theorem implies that $g(x)$ has a fixed point, which is an eigenvector of f. ■

Note that if $f^m(P) \subset K_0$ for some integer $m \geqslant 1$, then the eigenvectors of f all lie in K_0.

With the aid of Lemma 3.1 we can define the spectral radius $r(f)$ of a map f satisfying H1.

DEFINITION 3.2. If f satisfies H1, then $r(f)$, the spectral radius of f, is given by

$$r(f) = \sup\{\lambda \mid \lambda x = f(x) \text{ for some } x \in K - \{0\}\}.$$

Properly we should speak of the "cone spectral radius of f" and write $r_K(f)$ to indicate dependence on K. However, for simplicity we shall use the previous notation. For linear maps in Banach space, the idea of the cone spectral radius was introduced by Bonsall [35], although he used the term "partial spectral radius."

In order to establish the basic properties of the spectral radius we need two more lemmas. The next lemma is a special case of Theorem 9.1 in [19]; other generalizations can be found in Section 5 of [23].

LEMMA 3.2 (Krein and Rutman [19]). *Assume that $f: K \to K$ satisfies H1 and that there exist $u \in K - \{0\}$ and $\lambda \in \mathbb{R}$ such that $f(u) \geqslant \lambda u$. Then there exists $x \in K - \{0\}$ and $\lambda' \geqslant \lambda$ such that $f(x) = \lambda' x$.*

The following simple lemma will also be useful.

LEMMA 3.3. *Assume that C is a cone in a real Banach space X and that $f: C \to C$ is a continuous map which is homogeneous of degree one and order-preserving. Suppose that there exists $x \in C - \{0\}$ and a real number λ such that*

$$f(x) \leqslant \lambda x.$$

If $f(y) = \mu y$ for some $y \in C - \{0\}$ and if there exists a real number $\delta > 0$ such that $\delta y \leqslant x$, then $\mu \leqslant \lambda$. In particular, if the interior C_0 of C is *nonempty and* $x \in C_0$, *then every eigenvalue* μ *of* f *satisfies* $\mu \leqslant \lambda$.

Proof. Because f is order-preserving and homogeneous, the fact that $\delta y \leqslant x$ implies

$$\delta\mu^n y = f^n(\delta y) \leqslant f^n(x) \leqslant \lambda^n x, \tag{3.1}$$

where f^n denotes the composition of f with itself n times. If μ were strictly greater than λ, Equation (3.1) would imply

$$-\delta y = \lim_{n \to \infty} \left[\left(\frac{\lambda^n}{\mu^n} \right) x - \delta y \right] \in C,$$

a contradiction. ■

Our next theorem collects the basic properties of the spectral radius.

THEOREM 3.1.

(1) *If* $f\colon K \to K$ *satisfies* H1 *and if there exists* $y \in K_0$ *such that* $f(y) \leqslant \lambda y$, *then* $r(f) \leqslant \lambda$. *If* $g\colon K \to K$ *also satisfies* H1 *and* $f(x) \leqslant g(x)$ *for all* $x \in K$, *then* $r(f) \leqslant r(g)$.

(2) *Suppose that* $f\colon K \to K$ *satisfies* H1 *and that* $\{f_m : m \geqslant 1\}$ *is a sequence of maps* $f_m\colon K \to K$ *all of which satisfy* H1 *and*

$$\lim_{m \to \infty} \sup_{x \in P} |f_m(x) - f(x)| = 0. \tag{3.2}$$

Then one has $\lim_{m \to \infty} r(f_m) = r(f)$.

(3) *If* $f\colon K \to K$ *satisfies* H1, *one has*

$$r(f) = \inf_{x > 0} \max_{1 \leqslant i \leqslant n} \frac{(f(x))i}{x_i} = \inf_{x > 0} \sup_{\alpha \in P} \sum_{i=1}^{n} \alpha_i \left(\frac{(f(x))i}{x_i} \right), \tag{3.3}$$

where $(f(x))_i$ *denotes the* ith *component of* $f(x)$ *and* α_i *the* ith *component of* α.

(4) *Assume that* f *satisfies* H1 *and that for all* $x, y \in K$ *and all real numbers* θ *with* $0 \leqslant \theta \leqslant 1$ *one has*

$$f(x^{1-\theta} y^{\theta}) \leqslant f(x)^{1-\theta} f(y)^{\theta}. \tag{3.4}$$

Let $\psi:\mathbb{R}\to\mathbb{R}$ be a nondecreasing, convex function, and define $\psi(-\infty)=\lim_{t\to-\infty}\psi(t)$ (possibly $\psi(-\infty)=-\infty$) and $\phi(t)=\psi(\log t)$ for $t\geqslant 0$, with $\phi(0)=\psi(-\infty)$. Then one has

$$\phi(r(f))=\inf_{x>0}\sup_{\alpha\in P}\sum_{i=1}^{n}\alpha_i\phi\left(\frac{(f(x))_i}{x_i}\right)=\sup_{\alpha\in P}\inf_{x>0}\sum_{i=1}^{n}\alpha_i\phi\left(\frac{(f(x))_i}{x_i}\right). \tag{3.5}$$

Proof. (1): The first part of (1) follows immediately from Lemma 3.3. By the compactness of P, there exists $u\in P$ such that $f(u)=r_1u$, where $r_1=r(f)$. The assumption on g implies that

$$f(u)=r_1u\leqslant g(u),$$

so Lemma 3.2 implies that g has an eigenvalue greater than or equal to r_1, i.e., $r(g)\geqslant r_1$.

(2): Let $\{f_m\mid m\geqslant 1\}$ and f be as in the statement of part (2) of the theorem, and write $\lambda=r(f)$ and $\lambda_m=r(f_m)$. By the compactness of P, there exists $u\in P$ such that $f(u)=\lambda u$, and there exist $u_m\in P$ such that $f_m(u_m)=\lambda_m u_m$. The assumptions on f_m imply that

$$f_m(u)\geqslant(\lambda-\delta_m)u,$$

where $\lim_{m\to\infty}\delta_m=0$. If $\lambda=0$, the numbers δ_m can all be taken equal to zero. Thus for m large enough, Lemma 3.2 applies and shows $r(f_m)\geqslant\lambda-\delta_m$, so

$$\liminf_{m\to\infty} r(f_m)\geqslant\lambda=r(f). \tag{3.6}$$

Because the functions f_m are uniformly bounded on P, the numbers λ_m are bounded, and by taking a subsequence λ_{m_i} we can arrange that

$$\lim_{i\to\infty}\lambda_{m_i}=\limsup_{m\to\infty}\lambda_m\equiv\lambda'.$$

By using the compactness of P and taking a further subsequence we can assume that $u_{m_i}\to v\in P$, and a simple limit argument shows

$$f(v)=\lambda' v.$$

Thus the definition of $r(f)$ implies that

$$\lambda' = \limsup_{m \to \infty} \lambda_m \leqslant r(f), \tag{3.7}$$

and Equations (3.6) and (3.7) give

$$r(f) = \lim_{m \to \infty} r(f_m).$$

(3): The second equality in Equation (3.3) is obvious, so we concentrate on proving

$$r(f) = \inf_{x > 0} \max_{1 \leqslant i \leqslant n} \frac{(f(x))_i}{x_i}.$$

If $x > 0$ (i.e., $x \in K_0$), Lemma 3.3 implies

$$\max_{1 \leqslant i \leqslant n} \frac{(f(x))_i}{x_i} \geqslant r(f),$$

so

$$r(f) \leqslant \inf_{x > 0} \max_{1 \leqslant i \leqslant n} \frac{(f(x))_i}{x_i}. \tag{3.8}$$

On the other hand, if f has an eigenvector $u \in K_0$, then by taking $x = u$ in Equation (3.3) we see that

$$r(f) \geqslant \inf_{x > 0} \max_{1 \leqslant i \leqslant n} \frac{(f(x))_i}{x_i},$$

which gives equality in this case.

To handle the possibility that f has no eigenvector in K_0, define, for $m \geqslant 1$, $f_m(x) = f(x) + (1/m)Jx$, where J will always denote a matrix with all entries equal to one. By part (2) of this theorem we have

$$r(f) = \lim_{m \to \infty} r(f_m).$$

Because $f_m(K - \{0\}) \subset K_0$, Lemma 3.1 shows f_m has an eigenvector in K_0, so the above remarks imply

$$r(f_m) \leqslant \inf_{x>0} \max_{1\leqslant i\leqslant n} \frac{(f_m(x))_i}{x_i}.$$

It is clear that

$$\inf_{x>0} \max_{1\leqslant i\leqslant n} \frac{(f_m(x))_i}{x_i} \geqslant \inf_{x>0} \max_{1\leqslant i\leqslant n} \frac{(f(x))_i}{x_i},$$

so taking limits as $m \to \infty$ gives

$$r(f) = \lim_{m\to\infty} r(f_m) \geqslant \inf_{x>0} \max_{1\leqslant i\leqslant n} \frac{(f(x))_i}{x_i}. \tag{3.9}$$

Equations (3.8) and (3.9) give the desired result.

(4): In the statement of part (4) of this theorem one encounters sums $\sum_{i=1}^n \alpha_i r_i$, where $\alpha \in P$ and $-\infty \leqslant r_i < \infty$. If $\alpha_i = 0$, $\alpha_i r_i$ is defined to be zero even if $r_i = -\infty$, but if $r_i = -\infty$ and $\alpha_i > 0$ for some i, the sum is $-\infty$.

The first part of Equation (3.5) follows from Equation (3.3) and the fact that ϕ is nondecreasing; convexity does not enter here. To prove the second part, define $e^z = (e^{z_1}, e^{z_2}, \ldots, e^{z_n})$ for $z \in R^n$, and define $h(\alpha, z)$ for $\alpha \in P$, $z \in \mathbb{R}^n$ by

$$h(\alpha, z) = \sum_{i=1}^{n} \alpha_i \phi\left(\frac{(f(e^z))_i}{e^{z_i}}\right).$$

Just as in Section 2, one easily shows that $\alpha \to h(\alpha, z)$ is upper semicontinuous and quasiconcave and that $z \to h(\alpha, z)$ is lower semicontinuous. The fact that $z \to h(\alpha, z)$ is quasiconvex follows immediately from Equation (3.4) and the assumption that ψ is convex and nondecreasing. ■

Equation 3.3 in Theorem 3.1 is closely related to a formula of Schneider and Turner (Corollary 2.10 in [34]); however, it does not seem to follow directly from that result.

Suppose C is the cone of nonnegative functions in $C(M)$, the Banach space of continuous real-valued functions on a compact space M, or in $L^p(\Omega, \mu)$, $1 \leqslant p < \infty$, where (Ω, μ) is a σ-finite measure space. Then for $u, v \in C$ and θ a real number with $0 \leqslant \theta \leqslant 1$, we have $u^{1-\theta}v^\theta \in C$. If $X = C(M)$, this is obvious, and if $X = L^p$ it follows from Hölder's inequality

$$\|u^{1-\theta}v^\theta\|_{L^p} \leqslant \|u\|_{L^p}^{1-\theta} \|v\|_{L^p}^{\theta}.$$

If $f: C \to C$ is continuous, it thus makes sense to ask whether the analogue of Equation (3.4) holds, namely, whether for all $u, v \in C$ and all θ with $0 \leqslant \theta \leqslant 1$,

$$f(u^{1-\theta}v^{\theta}) \leqslant [f(u)]^{1-\theta}[f(v)]^{\theta}. \tag{3.10}$$

DEFINITION 3.2. If $f: C \to C$ is continuous and f satisfies Equation (3.10) for all $u, v \in C$ and all θ with $0 \leqslant \theta \leqslant 1$, we shall say that f satisfies hypothesis H2.

If C is a cone in a real Banach space X, let C^* denote the set of continuous linear functionals w^* in X^* such that $w^*(x) \geqslant 0$ for all $x \in C$. Recall that if $X = C(M)$ and C is the cone of nonnegative functions, then C^* is the set of nonnegative, regular Borel measures on M; and if $X = L^p(\Omega)$, $1 \leqslant p \leqslant \infty$, then C^* is the set of nonnegative functions in $L^q(\Omega)$, where $1/p + 1/q = 1$. Using these representations and Hölder's inequality, one can easily see that for all $u, v \in C$, $w^* \in C^*$, and real θ such that $0 \leqslant \theta \leqslant 1$, one has

$$w^*(u^{1-\theta}v^{\theta}) \leqslant [w^*(u)]^{1-\theta}[w^*(v)]^{\theta}. \tag{3.11}$$

It will be useful for us to know classes of functions which satisfy Equation (3.10).

PROPOSITION 3.1. *Assume that either $X = C(M)$, the space of continuous real-valued functions on a compact space M, or $X = L^p(\Omega)$, $1 \leqslant p \leqslant \infty$, where Ω is a σ-finite measure space with respect to a measure μ. Let C be the cone of nonnegative functions in X, and $f: C \to C$ a continuous map.*

(1) *If $X = C(M)$, then f satisfies* H2 *if and only if for every $w^* \in C^*$ and every pair $u, v \in C$, the map $\theta \to w^*(f(u^{1-\theta}v^{\theta}))$ is log convex.*

(2) *IF $f: C \to C$ and $g: C \to C$ satisfy* H2, *then $f + g$ satisfies* H2; *and if f is order-preserving, $f \circ g$ satisfies* H2.

Proof. (1): Take $u, v \in C$ and $w^* \in C$. If $f: C \to C$ satisfies H2, we must show that $h(\theta) \equiv w^*(f(u^{1-\theta}v^{\theta}))$ is log convex for $0 \leqslant \theta \leqslant 1$. Take θ_0, θ_1 such that $0 \leqslant \theta_0 \leqslant \theta_1$, and for $0 \leqslant t \leqslant 1$ define $\theta_t = (1-t)\theta_0 + t\theta_1$. We must show that

$$h(\theta_t) = w^*\Big(f\big((u^{1-\theta_0}v^{\theta_0})^{1-t}(u^{1-\theta_1}v^{\theta_1})^{t}\big)\Big) \leqslant h(\theta_0)^{1-t}h(\theta_1)^{t}.$$

However, we have

$$f\big((u^{1-\theta_0}v^{\theta_0})^{1-t}(u^{1-\theta_1}v^{\theta_1})^{t}\big) \leqslant \big[f(u^{1-\theta_0}v^{\theta_0})\big]^{1-t}\big[f(u^{1-\theta_1}v^{\theta_1})\big]^{t},$$

so applying w^* and using Equation (3.11) gives the desired result. Notice that this argument applies also if $X = L^p$.

Conversely, suppose that $h(\theta)$ defined above is log convex for all choices of $u, v \in C$ and $w^* \in C^*$. If $X = C(M)$ and one chooses $w^*(z) = z(m)$ for a fixed $m \in M$, one finds that Equation (3.10) is satisfied pointwise at each m, so Equation (3.10) is satisfied.

(2): If f and g satisfy H2 and $h(z) \equiv (f+g)(z)$ for $z \in C$, then Hölder's inequality gives

$$\begin{aligned} h(u^{1-\theta}v^{\theta}) &= f(u^{1-\theta}v^{\theta}) + g(u^{1-\theta}v^{\theta}) \\ &\leqslant f(u)^{1-\theta}f(v)^{\theta} + g(u)^{1-\theta}g(v)^{\theta} \\ &\leqslant [f(u)+g(u)]^{1-\theta}[f(v)+g(v)]^{\theta} = [h(u)]^{1-\theta}[h(v)]^{1-\theta}, \end{aligned}$$

so h satisfies H2. (Pointwise almost everywhere interpretations of these inequalities must be given in the L^p case.)

If f is also order-preserving,

$$f(g(u^{1-\theta}v^{\theta})) \leqslant f\big(g(u)^{1-\theta}g(v)^{\theta}\big) \leqslant [f(g(u))]^{1-\theta}[f(g(v))]^{\theta},$$

so $f \circ g$ satisfies H2. ∎

REMARK 3.1. It is not hard to show that statement (1) of Proposition 3.1 also is true if X is one of the standard L^p spaces, but we omit the proof.

It will also be useful to specify basic properties of maps $f: K \to K$ which satisfy H1.

PROPOSITION 3.2.

(1) *If $f: K \to K$ is continuous and f is continuously differentiable on K_0, then f is order-preserving if and only if*

$$\frac{\partial f_i}{\partial x_j}(x) \geqslant 0$$

for all i and j with $1 \leqslant i, j \leqslant n$ and all $x \in K_0$. (Here $f_i(x)$ denotes the ith component of $f(x)$.)

(2) *If $f: K \to K$ and $g: K \to K$ satisfy* H1, *then $f + g$ and $f \circ g$ satisfy* H1.

Proof. (1): Suppose that

$$\left(\frac{\partial f_i}{\partial x_j}(x)\right) \geqslant 0 \qquad \text{for all} \quad x \in K_0.$$

By the continuity of f, to prove f is order-preserving on K, it suffices to prove f is order-preserving on K_0. Thus suppose $u, v \in K_0$ and $u \leqslant v$, and for $0 \leqslant t \leqslant 1$ define $u_t = (1-t)u + tv$. If $g(t) = f_i(u_t)$, it suffices to prove $g(1) \geqslant g(0)$. But the fundamental theorem of calculus gives

$$g(1) - g(0) = \int_0^1 g'(t)\, dt = \sum_{j=1}^{n} \int_0^1 \frac{\partial f_i}{\partial x_j}(u_t)(v_j - u_j)\, dt \geqslant 0.$$

Conversely, suppose that f is order-preserving. If $x \in K_0$ and e_j denotes the unit vector with 1 in the jth position and zeros elsewhere, then order-preserving implies

$$0 \leqslant \lim_{t \to 0^+} \frac{f_i(x + te_j) - f_i(x)}{t} = \frac{\partial f_i}{\partial x_j}(x).$$

Statement (2) of the proposition is trivial, and we leave it to the reader. ∎

Notice that we do *not* assume f is C^1 on K (in Proposition 3.2), because this is almost never true in our examples.

If $x \in K$, r is a real number, and $\sigma \in P$, we define the "(r, σ) mean of x," $M_{r\sigma}(x)$, as follows.

DEFINITION 3.3. If $x = (x_1, x_2, \ldots, x_n) \in K$, $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n) \in P$, and r is a real number, then for $r \neq 0$,

$$M_{r\sigma}(x) = \left(\sum_{j=1}^{n} \sigma_j x_j^r\right)^{1/r} \tag{3.12}$$

and for $r = 0$,

$$M_{0\sigma}(x) = \prod_{j=1}^{n} x_j^{\sigma_j} = \lim_{r \to 0^+} M_{r\sigma}(x). \tag{3.13}$$

If $r < 0$, some further explanation of Equation (3.12) is needed. Given $\sigma \in P$, let $B = \{ j \mid \sigma_j > 0\}$. By definition,

$$M_{r\sigma}(x) = \left(\sum_{j \in B} \sigma_j x_j^r \right)^{1/r} \tag{3.14}$$

for $x \in K$, and (for $r < 0$) the right-hand side of Equation (3.14) is interpreted as equal to zero if $x_j = 0$ for some $j \in B$.

To simplify formulas we adopt the notation $x^\sigma = \prod_{j=1}^n x_j^{\sigma_j}$ for $x \in K$, $\sigma \in P$.

With the above conventions, we leave as an exercise to the reader to prove that $x \to M_{r\sigma}(x)$ is continuous on K for every real r and every $\sigma \in P$. It is also easy to see that $M_{r\sigma}(x)$ is homogeneous of degree 1, and that if $0 \leqslant x \leqslant y$, then $M_{r\sigma}(x) \leqslant M_{r\sigma}(y)$. Furthermore, if $x, y \in K_0$ and $0 < \theta \leqslant 1$ and $r \geqslant 0$, Hölder's inequality gives

$$\begin{aligned} M_{r\sigma}(x^{1-\theta} y^\theta) &= \left(\sum_{j=1}^{n} (\sigma_j x_j^r)^{1-\theta} (\sigma_j y_j^r)^\theta \right)^{1/r} \\ &\leqslant [M_{r\sigma}(x)]^{1-\theta} [M_{r\sigma}(y)]^\theta. \end{aligned} \tag{3.15}$$

This inequality is also true for $r = 0$, by taking limits as $r \to 0^+$ (or giving a separate argument). Continuity of $M_{r\sigma}(x)$ also implies that (3.15) is true for all x and y in K.

The typical example of a function $f: K \to K$ to which the results of this section and Section 4 will apply is one such that $(f(x))_i$, the ith component of $f(x)$, is a sum of positive multiplies of functions $M_{r\sigma}(x)$ for differing nonnegative r and σ. Formally, suppose that for each i, $1 \leqslant i \leqslant n$, there exists a finite collection Γ_i of ordered pairs (r, σ) with $r \in \mathbb{R}$ and $\sigma \in P$. Suppose that $f: K \to K$ is such that $(f(x))_i$ is given by the formula

$$(f(x))_i = \sum_{(r,\sigma) \in \Gamma_i} c_{ir\sigma} M_{r\sigma}(x), \tag{3.16}$$

where $c_{ir\sigma} > 0$ for $(r, \sigma) \in \Gamma_i$.

DEFINITION 3.4. If $f: K \to K$ is defined by Equation (3.16), then we say f is a *positive sum of* (r, σ) *means* or, simply, a *positive sum of means*.

Propositions 3.1 and 3.2 and the above remarks show that functions which are positive sums of (r, σ) means satisfy H1 and H2. Therefore we have

COROLLARY 3.1. *All conclusions of Theorem* 3.1 *apply to functions which are positive sums of* (r, σ) *means with* $r \geq 0$ *for all* $(r, \sigma) \in \Gamma_i$.

We now want to give a generalization in this framework of Kingman's theorem, and to do this we must discuss parametrized families of nonlinear operators. Suppose that $f:[0,1]\times K \to K$, and for each $t \in [0,1]$ define $f_t(x) = f(t, x)$.

DEFINITION 3.5. Suppose that $f:[0,1]\times K \to K$ is continuous and that for each t with $0 \leqslant t \leqslant 1$, f_t satisfies H1. In addition, suppose that for all real numbers t_0 and t_1 in $[0,1]$, all nonnegative vectors u and v and all real θ with $0 \leqslant \theta \leqslant 1$ one has

$$f\big((1-\theta)t_0 + \theta t_1, u^{1-\theta}v^{\theta}\big) \leqslant [f(t_0, u)]^{1-\theta}[f(t_1, v)]^{\theta}. \tag{3.17}$$

Then we shall say f satisfies hypothesis H3.

If $f:[0,1]\times K \to K$ is continuous and one defines $h(t, z) = f(t, e^z)$, one can see that the inequality (3.17) is equivalent to assuming h is log convex on $[0,1]\times \mathbb{R}^n$.

The next proposition shows that there are many examples of functions satisfying H3 and that they behave nicely with respect to some standard operations.

PROPOSITION 3.3.

(1) *Assume that* $f:[0,1]\times K \to K$ *and* $g:[0,1]\times K \to K$ *and that* f *and* g *satisfy* H3. *Then* $h(t,x) \equiv f(t,x) + g(t,x)$ *satisfies* H3.

(2) *If* $f:[0,1]\times K \to K$ *satisfies* H3 *and* $g: K \to K$ *satisfies* H1 *and* H2, *then* $g(f(t,x))$ *satisfies* H3.

(3) *If* $f:[0,1]\times K \to K$ *satisfies* H3 *and* $c(t)$, $0 \leqslant t \leqslant 1$, *is a log convex, real-valued function of* t, *then* $h(t,x) = c(t)f(t,x)$ *satisfies* H3.

(4) *For* $1 \leqslant i \leqslant n$ *let* Γ_i *be a finite collection of ordered pairs* (r, σ) *with* r *nonnegative and* $\sigma \in P$, *and for* $(r, \sigma) \in \Gamma_i$ *and* $1 \leqslant i \leqslant n$ *let* $c_{ir\sigma}(t)$ *be a*

positive log convex function of t. Then if $f\colon[0,1]\times K\to K$ *is defined by*

$$\begin{aligned}(f(t,x))_i &= i\text{th component of } f(t,x)\\ &= \sum_{(r,\sigma)\in\Gamma_i} c_{ir\sigma}(t)M_{r\sigma}(x),\end{aligned}\tag{3.18}$$

$f(t,x)$ *satisfies* H3.

Proof. (1): This follows by essentially the same argument used to prove the second part of Proposition 3.1, and we leave it to the reader.

(2): If $0\leqslant t_0\leqslant t_1\leqslant 1$, x and $y\in K$, and $0\leqslant\theta\leqslant 1$, and if $t_\theta=(1-\theta)t_0+\theta t_1$, one has

$$\begin{aligned}g\big(f(t_\theta,x^{1-\theta}y^\theta)\big) &\leqslant g\big([f(t_0,x)]^{1-\theta}[f(t_1,y)]^\theta\big)\\ &\leqslant [g(f(t_0,x))]^{1-\theta}[g(f(t_1,y))]^\theta,\end{aligned}$$

which implies that $g(f(t,x))$ satisfies H3.

(3): In the notation of the preceding paragraph one has

$$\begin{aligned}h\big(t_\theta,x^{1-\theta}y^\theta\big) &= c(t_\theta)f\big(t_\theta,x^{1-\theta}y^\theta\big)\\ &\leqslant [c(t_0)]^{1-\theta}[f(t_0,x)]^{1-\theta}[c(t_1)]^\theta[f(t_1,y)]^\theta\\ &= h(t_0,x)^{1-\theta}h(t_1,y)^\theta,\end{aligned}$$

so $h(t,x)$ satisfies H3.

(4): If e_i is the unit vector with 1 in the ith position, we have already seen that $x\to M_{r\sigma}(x)e_i$ satisfies H2, so by part (3) of this proposition $x\to c_{ir\sigma}(t)M_{r\sigma}(x)e_i$ satisfies H3. Since the map $f(t,x)$ in Equation (3.18) is a sum of such maps, part (1) of this proposition implies that $f(t,x)$ satisfies H3. ∎

We now come to the principal theorem of this section, which is a direct generalization of Kingman's theorem [17].

Theorem 3.2. *Assume that* $f\colon[0,1]\times K\to K$ *satisfies* H3 (*see Definition* 3.5), *and define* $f_t(x)=f(t,x)$. *Then the map* $t\to r(f_t)$ *is log convex. In particular* $t\to r(f_t)$ *is log convex if* f_t *is given by Equation* (3.18), *i.e.,* f_t *is a positive sum of* (r,σ) *means with coefficients* $c_{ir\sigma}(t)$ *which are log convex functions of t and* $r\geqslant 0$ *for all* $(r,\sigma)\in\Gamma_i$.

Proof. Select $\varepsilon > 0$, and define

$$g(t, x) = f(t, x) + \frac{1}{k}Jx,$$

where J is the matrix all of whose entries are 1 and $k \geqslant 1$ is an integer. Select t_0 and t_1 in $[0,1]$ and θ with $0 \leqslant \theta \leqslant 1$, and define $t_\theta = (1-\theta)t_0 + \theta t_1$. If $g_t(x) \equiv g(t, x)$, Theorem 3.1 implies that for k large enough, $r(f_{t_j}) \leqslant r(g_{t_j}) < r(f_{t_j}) + \varepsilon$ for $j = 0$ and 1, and $r(f_{t_\theta}) \leqslant r(g_{t_\theta})$. Because $g_t(K - \{0\}) \subset K_0$, Lemma 3.1 implies that there exist vectors $u \in K_0$ and $v \in K_0$ such that $g_{t_0}(u) = \rho_0 u$, $g_{t_1}(v) = \rho_1 v$, where $\rho_j = r(g_{t_j})$ for $j = 0$ and 1. Proposition 3.3, part (1), implies that g satisfies H3, so

$$g_{t_\theta}(u^{1-\theta}v^\theta) \leqslant \left[g_{t_0}(u)\right]^{1-\theta}\left[g_{t_1}(v)\right]^\theta = \rho_0^{1-\theta}\rho_1^\theta u^{1-\theta}v^\theta.$$

Theorem 3.1, part (1), therefore implies that

$$r(f_{t_\theta}) \leqslant r(g_{t_\theta}) \leqslant \left[r(f_{t_0}) + \varepsilon\right]^{1-\theta}\left[r(f_{t_1}) + \varepsilon\right]^\theta,$$

and because $\varepsilon > 0$ can be taken arbitrarily small above, we are done. ∎

The case in which all the numbers r in Equation (3.18) equal zero provides the simplest nontrivial example of Theorem 3.2. Thus assume that $G_i \subset P$ is a finite set or $1 \leqslant i \leqslant n$ and that for $\sigma \in G_i$, $c_{i\sigma}(t)$ is a log convex function of t, $0 \leqslant t \leqslant 1$. If $f_t : K \to K$ is defined by

$$(f_t(x))_i = \sum_{\sigma \in G_i} c_{i\sigma}(t)x^\sigma, \tag{3.19}$$

then $t \to r(f_t)$ is log convex (recall $x^\sigma = \prod_{j=1}^n x_j^{\sigma_j}$).

There is also a generalization of Cohen's theorem to this framework.

THEOREM 3.3. *Assume that* $f : [0,1] \times K \to K$ *satisfies* H3 *and that* $D(t) = \operatorname{diag}(d_{ii}(t))$ *is a nonnegative, diagonal matrix whose diagonal elements* $d_{ii}(t)$ *are convex functions of* t *for* $0 \leqslant t \leqslant 1$. *Then the map* $t \to r(f_t + D(t))$ *is a convex function of* t *for* $0 \leqslant t \leqslant 1$.

Proof. By using the trick of approximating $f(t, x)$ by $f(t, x) + (1/k)Jx$ (as in the proof of Theorem 3.2), we can assume that $f(t, x) \in K_0$ for $0 \leqslant t \leqslant 1$ and $x \in K - \{0\}$. Define t_0, t_1, θ, and t_θ as in the proof of Theorem

3.1. Lemma 3.1 implies that $f_{t_0}+D(t_0)$ has an eigenvector $u\in K_0$ with corresponding eigenvalue $\lambda_0=r(D(t_0)+f_{t_0})$. Similarly, $f_{t_1}+D(t_1)$ has an eigenvector $v\in K_0$ with eigenvalue $\lambda_1=r(D(t_1)+f_{t_1})$. For notational convenience write $d_{ii}(t_0)=a_i$ and $d_{ii}(t_1)=b_i$, so

$$d_{ii}(t_\theta)\leqslant(1-\theta)a_i+\theta b_i.$$

Because f satisfies H3,

$$\begin{aligned}
&d_{ii}(t_\theta)u_i^{1-\theta}v_i^\theta+\left(f\left(t_\theta,u^{1-\theta}v^\theta\right)\right)_i\\
&\quad\leqslant\left[(1-\theta)a_i+\theta b_i\right]u_i^{1-\theta}v_i^\theta+\left(f(t_0,u)\right)_i^{1-\theta}\left(\left(f(t_1,v)\right)_i^\theta\right.\\
&\quad\leqslant\left\{(1-\theta)a_i+\theta b_i+(\lambda_0-a_i)^{1-\theta}(\lambda_1-b_i)^\theta\right\}u_i^{1-\theta}v_i^\theta,
\end{aligned}\tag{3.20}$$

where the subscript i denotes the ith component of a vector. Because the geometric mean is dominated by the arithmetic mean, Equation (3.20) gives

$$D(t_\theta)u^{1-\theta}v^\theta+f_{t_\theta}(u^{1-\theta}v^\theta)\leqslant\left[(1-\theta)\lambda_0+\theta\lambda_1\right]u^{1-\theta}v^\theta,\tag{3.21}$$

and Theorem 3.1 then implies that

$$r\left(D(t_\theta)+f_{t_\theta}\right)\leqslant(1-\theta)\lambda_0+\theta\lambda_1$$

which proves the theorem. ∎

4. CONVERGENCE OF ITERATES TO A UNIQUE POSITIVE EIGENVECTOR

If A is an irreducible, nonnegative matrix, A possesses a unique (to within scalar multiples) eigenvector in K_0. The first question we shall ask is whether certain maps $f\colon K\to K$ have a unique eigenvector in K_0. Unlike the linear case, we shall not expect a unique eigenvector in K. Thus the map $f(x_1,x_2)=(\sqrt{x_1x_2},\sqrt{x_1x_2})$ clearly has a unique eigenvector $u=(1,1)$ in K_0, but also has eigenvectors $(1,0)$ and $(0,1)$, not in K_0.

A nonnegative $n\times n$ matrix A is called "primitive" if there exists an integer $p\geqslant 1$ such that $A^p>0$. If A is primitive, A has a unique eigenvector $u\in K_0$ such that $|u|=1$. However, Birkhoff [3] has shown that much more is true. For $x \in K_0$, define $g(x) = Ax/|Ax|$ and let $g^p(x)$ denote the iteration of g with itself p times. Birkhoff proved that if $x \in K_0$, then $\lim_{p\to\infty}|g^p(x) - u| = 0$ and that, in fact, the convergence is geometric. If $f: K \to K$ satisfies H1 and $f(K_0) \subset K_0$, one can define $g(x) = f(x)/|f(x)|$ and ask whether $\lim_{p\to\infty}|g^p(x) - u| = 0$ for $x \in K_0$, where $u \in K_0$ is the unique eigenvector of f, normalized so $|u| = 1$. We shall show that this is indeed true for a large class of maps f, and our results will reduce to Birkhoff's in the linear case.

The typical map f to which our results will apply is a positive sum of (r, σ) means (see Definition 3.4). One basic difficulty is that we shall to consider f^p, the composition of f with itself p times, and that in general f^p will not be a positive sum of (r, σ) means. Thus we are forced to consider more general classes which are closed under composition. Another problem to remember is that it is not, in general, true for our functions f that $f^m(K - \{0\}) \subset K_0$ for some integer m. If $f^m(K - \{0\}) \subset K_0$, many of the subsequent difficulties would vanish.

Definition 4.1. If $f: K \to K$ satisfies H1 (see Definition 3.1), f will be called *power-bounded below* if for each i, $1 \leqslant i \leqslant n$, there exists a positive constant c and a probability vector $\sigma \in P$, both dependent on i, such that

$$(f(x))_i \geqslant cx^\sigma \tag{4.1}$$

for all $x \in K$. [Here $(f(x))_i$ denotes the ith component of $f(x)$.]

We also need to define an incidence matrix for a power-bounded-below function f.

Definition 4.2. If $f: K \to K$ is power-bounded below and $A = (a_{ij})$ is an $n \times n$, nonnegative matrix, then A is called an *incidence matrix for f* (with respect to being power-bounded below) if whenever $a_{ij} > 0$, there exist a positive real c and a probability vector $\sigma \in P$ (both depending on i and j) such that σ_j, the jth component of σ, is positive and

$$(f(x))_i \geqslant cx^\sigma$$

for all $x \in K$.

Notice that an incidence matrix for a power-bounded-below map f is not unique, although there is clearly an incidence matrix A with a maximal number of nonzero entries. If incidence matrices are normalized so that their nonzero entries equal 1, then $A \geqslant B$ for every other incidence matrix B.

LEMMA 4.1.

(1) *If* $f, g: K \to K$ *satisfy* H1 (*see Definition* 3.1) *and are power-bounded below, then* $h(x) = f(g(x))$ *satisfies* H1 *and is power-bounded below. If* A *is an incidence matrix for* f (*with respect to being power-bounded below*) *and* B *is an incidence matrix for* g, *then* AB *is an incidence matrix for* h.

(2) *If* f *is as in part* 1 *and* $\phi: K \to K$ *satisfies* H1, *then* $f(x) + \phi(x)$ *is power-bounded below and has incidence matrix* A.

Proof. (1): Because g is power-bounded below, for $1 \leqslant j \leqslant n$ there exist positive constants c_j and probability vectors $\tau^{(j)} \in P$ such that for all $x \in K$

$$(g(x))_j \geqslant c_j x^{\tau^{(j)}}. \tag{4.2}$$

Because f is power-bounded below, there exists a positive constant c and a vector $\sigma \in P$ such that

$$(f(u))_i \geqslant c u^{\sigma} \tag{4.3}$$

for all $u \in K$. Taking $u = g(x)$, one obtains

$$(f(u))_i \geqslant d x^{\tau}, \tag{4.4}$$

where $d = c\prod_{j=1}^{n} c_j^{\sigma_j}$ and $\tau = \sum_{j=1}^{n} \sigma_j \tau^{(j)} \in P$. This shows that $f(g(x))$ is power-bounded below.

If the (i, k) entry of AB is positive, there exists p such that $a_{ip} b_{pk} > 0$. The fact that $b_{pk} > 0$ means that, in the notation of the preceding paragraph, $\tau^{(p)} \in P$ can be chosen so that $\tau_k^{(p)}$, the kth component of $\tau^{(p)}$, is positive; and $\sigma \in P$ can be chosen so σ_p, the pth component of σ, is positive. All of this implies that the kth component of $\tau = \sum_{j=1}^{n} \sigma_j \tau^{(j)}$ is greater than or equal to $\sigma_p \tau_k^{(p)}$, which is positive. The latter fact proves AB is an incidence matrix for h.

The second part of the proposition is obvious, and we leave it to the reader. ■

If g is defined and continuously differentiable on an open neighborhood of a point $x \in \mathbb{R}^n$ and g maps into $\mathbb{R}^n$, let $J_g(x)$ denote the Jacobian matrix of g at x, i.e.,

$$J_g(x) = \left(\frac{\partial g_i}{\partial x_j}(x) \right), \tag{4.5}$$

where $g_i(x)$ denotes the ith component of $g(x)$.

The next lemma shows that $J_g(x)$ is often an incidence matrix (with respect to being power bounded below) for a map $g: K \to K$.

LEMMA 4.2. *For $1 \leqslant i \leqslant n$ assume that Γ_i is a finite collection of ordered pairs (r, σ), where r is a real number and $\sigma \in P$. For $1 \leqslant i \leqslant n$ and $(r, \sigma) \in \Gamma_i$ assume that $c_{ir\sigma}$ is a positive real, and define $f: K \to K$ by*

$$(f(x))_i = \sum_{(r,\sigma) \in \Gamma_i} c_{ir\sigma} M_{r\sigma}(x),$$

where $M_{r\sigma}(x)$ is defined in Equations (3.12)–(3.14). Define $\Gamma_i^+ = \{(r, \sigma) \in \Gamma_i \mid r \geqslant 0\}$, assume that Γ_i^+ is nonempty for $1 \leqslant i \leqslant n$, and define $g: K \to K$ by

$$(g(x))_i = \sum_{(r,\sigma) \in \Gamma_i^+} c_{ir\sigma} M_{r\sigma}(x).$$

Then f is power-bounded below, and for any $x \in K_0$, $J_g(x)$ is an incidence matrix for f. Furthermore, if f^p denotes the composition of f with itself p times, $x \in K_0$, and $A = J_g(x)$, then A^p is an incidence matrix for f^p.

Proof. Lemma 4.1, part (2), implies that if g is power-bounded below, f is; and if A is an incidence matrix for g, it is an incidence matrix for f. Furthermore, Lemma 4.1 also implies that if A is an incidence matrix for f, then A^p is an incidence matrix for f^p.

Thus it remains to show that g is power-bounded below and if $x \in K_0$, then $J_g(x)$ is an incidence matrix for g. Because the arithmetic mean dominates the geometric mean,

$$M_{r\sigma}(x) \geqslant x^\sigma \tag{4.6}$$

for $x \in K$ and $(r, \sigma) \in \Gamma_i^+$. [Notice that the inequality (4.6) is reversed if $r < 0$.] It follows that

$$(g(x))_i \geqslant \sum_{(r,\sigma) \in \Gamma_i^+} c_{ir\sigma} x^\sigma, \tag{4.7}$$

and because Γ_i^+ is nonempty, (4.7) shows that g is power-bounded below. Furthermore, if one defines a nonnegative matrix $B = (b_{ij})$ by $b_{ij} > 0$ if there exists $(r, \sigma) \in \Gamma_i^+$ such that σ_j, the jth component of σ, is positive, then

Equation (4.7) shows that B is an incidence matrix for g. However, if $x \in K_0$ and $g_i(x)$ denotes the ith component of g, a calculation shows that

$$\frac{\partial g_i}{\partial x_j}(x) > 0$$

if there exists $\sigma \in \Gamma_i^+$ such that $\sigma_j > 0$. Thus $J_g(x)$ has the same positive entries as B and is an incidence matrix for g. ■

REMARK 4.1. For $r < 0$, $M_{r\sigma}(x)$ does not in general satisfy an inequality of the form $M_{r\sigma}(x) \geqslant cx^{\tau}$ for some $c \geqslant 0$ and $\tau \in P$, as one can easily verify for $M(x_1, x_2) = (x_1^{-1} + x_2^{-1})^{-1}$.

We shall need to introduce one more definition before proving our main theorems.

DEFINITION 4.3. If $f: K \to K$ is a continuous map, f is *superadditive* if for all $x, y \in K$,

$$f(x) + f(y) \leqslant f(x+y). \tag{4.8}$$

We are interested in superadditive functions because the map $x \to M_{r\sigma}(x)$ is often superadditive. Specifically, a classical inequality (see Theorem 24 on p. 30 in [15]) implies that if $r \leqslant 1$ and $\sigma \in P$, then for all $x, y \in K$

$$M_{r\sigma}(x) + M_{r\sigma}(y) \leqslant M_{r\sigma}(x+y), \qquad r \leqslant 1, \quad \sigma \in P. \tag{4.9}$$

The next lemma shows that the class of superadditive functions is closed under various simple operations.

LEMMA 4.3.

(1) *If $f: K \to K$ is superadditive, then f is order-preserving, and if f and g are superadditive, then $f + g$ and $f \circ g$ are superadditive.*

(2) *If f and g are superadditive and θ is a real number, $0 < \theta < 1$, then $h(u) \equiv f(u)^{1-\theta} g(u)^{\theta}$ is superadditive.*

(3) *If f is defined as in Lemma 4.2 and $(r, \sigma) \in \Gamma_i$ implies $r \leqslant 1$, then f is superadditive.*

Proof. (1): This is obvious and is left to the reader.

(2): By using Hölder's inequality and superadditivity one obtains for $x, y \in K$

$$\begin{aligned} h(x+y) &= f(x+y)^{1-\theta} g(x+y)^{\theta} \\ &\geqslant [f(x)+f(y)]^{1-\theta}[g(x)+g(y)]^{\theta} \\ &\geqslant f(x)^{1-\theta} g(x)^{\theta} + f(y)^{1-\theta} g(y)^{\theta}, \end{aligned}$$

which is the desired result.

(3): By using the inequality (4.9) one sees that f is a sum of superadditive functions and hence superadditive. ■

With these preliminaries we can prove our first theorem.

THEOREM 4.1. *Assume that $f: K \to K$ is homogeneous of degree* 1, *power-bounded below, and superadditive. In addition, suppose that f has an incidence matrix A (with respect to being power-bounded below) such that A is irreducible. Then f has a unique (to within scalar multiples) positive eigenvector $x \in K_0$.*

Note that f may have other eigenvectors in K, but these cannot lie in K_0.

Proof of Theorem 4.1. First we prove the existence of a positive eigenvector $u \in P_0 \subset K_0$. This part of the proof only requires that f satisfy H1 and be power-bounded below. If J is the matrix all of whose entries equal 1 and $k \geqslant 1$, let $u^{(k)} \in P_0$ be a positive eigenvector for $f(x)+(1/k)Jx$, so

$$f(u^{(k)}) + \frac{1}{k} J(u^{(k)}) = \lambda_k u^{(k)}. \tag{4.10}$$

Because f is bounded on P, one obtains from (4.10) (by taking norms on both sides) that there exists a constant B (independent of $k \geqslant 1$) such that

$$\lambda_k \leqslant B.$$

Equation (4.10) implies that for any integer $p \geqslant 1$,

$$f^p(u^{(k)}) \leqslant \lambda_k^p u^{(k)}, \tag{4.11}$$

where f^p denotes composition of f with itself p times. Because A is assumed irreducible, for any pair of integers (i, j) with $1 \leqslant i, j \leqslant n$ there exists an integer $p \geqslant 1$, depending on i and j, such that the (i, j) entry of A^p is positive. Given $k \geqslant 1$, select i and j, $i \neq j$, so that $u_i^{(k)}$ is the smallest component of $u^{(k)}$ and $u_j^{(k)}$ is the largest component, and let $p = p(i, j)$ be such that the (i, j) entry of A^p is positive. Lemma 4.1 implies A^p is an incidence matrix for f^p, so there exists a positive constant $c = c_{ij}$ and a vector $\sigma \in P$ (σ dependent on i and j) such that σ_j, the jth component of σ, is positive and

$$(f^p(x))_i \geqslant cx^\sigma \tag{4.12}$$

for all $x \in K$. For notational convenience, fix k and write $v = u^{(k)}$, so v_i is the minimal component of v, and v_j the maximal component. Then (4.11) and (4.12) imply

$$cv_i^{1-\sigma_j}v_j^{\sigma_j} \leqslant \lambda_k^p v_i, \tag{4.13}$$

or

$$\frac{v_j}{v_i} \leqslant \left(\lambda_k^p c^{-1}\right)^{\sigma_j^{-1}}. \tag{4.14}$$

The number λ_k is bounded by B, and the numbers p, c, and σ_j each assume at most n^2 distinct positive values, corresponding to the n^2 ordered pairs (i, j). Thus there exists a constant M, independent of k, such that

$$\frac{\max_{1 \leqslant p \leqslant n} u_p^{(k)}}{\min_{1 \leqslant p \leqslant n} u_p^{(k)}} \leqslant M. \tag{4.15}$$

If we take a subsequence of $(\lambda_k, u^{(k)})$ such that $\lambda_k \to \lambda$ and $u^{(k)} \to u \in P$, then Equation (4.15) implies that $u \in K_0$, and continuity gives

$$f(u) = \lambda u.$$

Next we have to prove uniqueness of the positive eigenvector of f. Suppose by way of contradiction that $x, y \in K_0$ are eigenvectors of f with eigenvalues λ and μ respectively and that y is not a scalar multiple of x. Theorem 3.1 implies that $\lambda = \mu$. If we define $\bar{\delta}$ by

$$\bar{\delta} = \min_{1 \leqslant k \leqslant n} \frac{y_k}{x_k},$$

there exists an integer j such that $y_j > \bar{\delta}x_j$ and an integer i such that $y_i = \bar{\delta}x_i$. We know that f^p is superadditive for any integer $p \geqslant 1$, so if $0 < \delta \leqslant \bar{\delta}$, we have

$$(f^p(y-\delta x))_i + (f^p(\delta x))_i \leqslant (f^p(y))_i = \lambda^p y_i = \bar{\delta}\lambda^p x_i,$$

or

$$(f^p(y-\delta x))_i + \delta\lambda^p x_i \leqslant \bar{\delta}\lambda^p x_i. \tag{4.16}$$

Because A is irreducible, there exists an integer p such that the (i, j) entry of A^p is positive (where i and j are selected as above); and because A^p is an incidence matrix for f^p, there exists a positive real c and a vector $\tau \in P$ such that τ_j, the jth component of τ, is positive and

$$f^p(w) \geqslant cw^\tau \tag{4.17}$$

for all $w \in K$. By using Equation (4.17) in (4.16) with $w = y - \delta x$, we obtain

$$c(y_j - \bar{\delta}x_j)^{\tau_j}(\bar{\delta}-\delta)^{1-\tau_j}\prod_{k \neq j} x_k^{\tau_k} \leqslant (\bar{\delta}-\delta)\lambda^p x_i. \tag{4.18}$$

If $\tau_j = 1$, we obtain an immediate contradiction by taking $\delta = \bar{\delta}$. If $0 < \tau_j < 1$, the inequality (4.18) is of the form

$$\alpha(\bar{\delta}-\delta)^{1-\tau_j} \leqslant \beta(\bar{\delta}-\delta), \tag{4.19}$$

where α and β are fixed positive constants independent of δ for $0 < \delta < \bar{\delta}$, and this is impossible for $\bar{\delta} - \delta$ small. ■

The question of whether a continuous $f: K_0 \to K_0$ which satisfies H1 has an eigenvector in K_0 is central to the results of this section. It is interesting to note that, for a particular class of nonlinear maps $f: K_0 \to K_0$, precisely this question arose in a different context in work of Menon and Schneider [33].

Before applying Theorem 4.1 to the case of positive sums of (r, σ) means, it will be convenient to prove one more lemma. For real numbers $t \neq 0$ define $\phi_t: K_0 \to K_0$ by

$$\phi_t(x) = x^t = (x_1^t, \ldots, x_n^t). \tag{4.20}$$

Observe that ϕ_t is defined and continuous on K if $t > 0$ and that $\phi_s(\phi_t(x)) = \phi_t(\phi_s(x)) = x$ if $s = t^{-1}$.

LEMMA 4.4. *Suppose that* $f: K \to K$ *satisfies* H1 *and is power-bounded below with incidence matrix* A. *If* $t > 0$ *and* $h(x)$ *is defined by*

$$h(x) = \phi_{t^{-1}}(f(\phi_t(x))),$$

where $\phi_t(x)$ *is defined by Equation* (4.20), *then* h *is power-bounded below, and* A *is an incidence matrix for* h.

Proof. Suppose $a_{ij} > 0$, so there exists $c > 0$ and a vector $\sigma \in P$ such that σ_j, the jth component of σ, is positive. Then for all $x \in K$

$$(f(\phi_t(x)))_i \geqslant c(\phi_t(x))^{\sigma} = cx^{t\sigma},$$

so we obtain

$$(h(x))_i \geqslant (c^{t^{-1}})x^{\sigma}.$$

The latter equation shows h is power-bounded below and has A as an incidence matrix. ■

COROLLARY 4.1. *Assume that* $f(x)$ *and* $g(x)$ *are defined as in Lemma* 4.2 *and that for some* $x \in K_0$, *the Jacobian matrix* $J_g(x)$ *of* g *at* x *is irreducible. Then* f *has a positive eigenvector* $u \in K_0$, *and* (*to within scalar multiples*) *such an eigenvector is unique.*

Proof. If one knew that $r \leqslant 1$ whenever $(r, \sigma) \in \Gamma_i$ for some i, $1 \leqslant i \leqslant n$, Lemmas 4.2 and 4.3 would imply that f is superadditive and power-bounded below with incidence matrix $J_g(x)$, so Corollary 4.1 would follow directly from Theorem 4.1. Since one may have $r > 1$, a different approach is needed.

Select $t > 0$ such that $t \leqslant 1$ and $tr \leqslant 1$ for any real number r such that $(r, \sigma) \in \Gamma_i$ for some i and some $\sigma \in P$. Define $s = t^{-1}$ and $h(x) = \phi_s(f(\phi_t(x)))$. Lemmas 4.2 and 4.4 imply that $h(x)$ is power-bounded below and has incidence matrix $J_g(x)$ for $x \in K_0$.

We claim that h is superadditive. Because h is continuous on K (a composition of continuous functions), to prove superadditivity it suffices to show that for all $x, y \in K_0$ and $1 \leqslant i \leqslant n$,

$$(h(x+y))_i \geqslant (h(x))_i + (h(y))_i.$$

By definition one has

$$(h(x+y))_i = \left\{ \sum_{(r,\sigma)\in\Gamma_i} c_{ir\sigma}\left(\left[\sum_{j=1}^{n} \sigma_j (x_j+y_j)^{rt}\right]^{1/rt}\right)^t\right\}^s .$$

Because $rt \leqslant 1$, Equation (4.9) gives

$$\left[\sum_{j=1}^{n} \sigma_j (x_j+y_j)^{rt}\right]^{1/rt} \geqslant \left[\sum_{j=1}^{n} \sigma_j x_j^{rt}\right]^{1/rt} + \left[\sum_{j=1}^{n} \sigma_j y_j^{rt}\right]^{1/rt} . \tag{4.21}$$

If we set $\alpha_{r\sigma} = [\sum_{j=1}^{n} \sigma_j x_j^{rt}]^{1/rt}$ and $\beta_{r\sigma} = [\sum_{j=1}^{n} \sigma_j y_j^{rt}]^{1/rt}$ and use Equation (4.21) in the formula for $(h(x+y))_i$, we obtain

$$(h(x+y))_i \geqslant \left[\sum_{(r,\sigma)\in\Gamma_i} c_{ir\sigma}[\alpha_{r\sigma}+\beta_{r\sigma}]^t\right]^s . \tag{4.22}$$

If we apply Equation (4.9) again, but this time to the vectors $u = (\alpha_{r\sigma})$ and $v = (\beta_{r\sigma})$, indexed by Γ_i, we obtain

$$\begin{aligned}(h(x+y))_i &\geqslant \left[\sum_{(r,\sigma)\in\Gamma_i} c_{ir\sigma}[\alpha_{r\sigma}+\beta_{r\sigma}]^t\right]^s \\ &\geqslant \left(\sum_{(r,\sigma)\in\Gamma_i} c_{ir\sigma}\alpha_{r\sigma}^t\right)^s + \left(\sum_{(r,\sigma)\in\Gamma_i} c_{ir\sigma}\beta_{r\sigma}^t\right)^s \\ &= (h(x))_i + (h(y))_i .\end{aligned}$$

Thus h is superadditive.

Since it is clear that h is homogeneous of degree 1, Theorem 4.1 implies that h has a unique (to within multiples) eigenvector x such that $x \in K_0$. Since eigenvectors $x \in K_0$ of h are in one-to-one correspondence with eigenvectors $y \in K_0$ of f by the map $y = \phi_t(x)$, f has a unique eigenvector in K_0. ∎

It may be worthwhile to state explicitly a special case of Corollary 4.1.

COROLLARY 4.2. *For $1 \leqslant i \leqslant n$ assume that Γ_i is a finite collection of ordered pairs (r,σ) such that r is a nonnegative real and $\sigma \in P$ is a* *probability vector. For* $1 \leqslant i \leqslant n$ *and* $(r,\sigma) \in \Gamma_i$, *suppose that* $c_{ir\sigma}$ *is a positive real, and define* $f: K \to K$ *by*

$$(f(x))_i = \sum_{(r,\sigma) \in \Gamma_i} c_{ir\sigma} M_{r\sigma}(x),$$

where $M_{r\sigma}(x) = (\Sigma \sigma_j x_j^r)^{1/r}$ *if* $r > 0$ *and* $M_{r\sigma}(x) = \prod_{j=1}^n x_j^{\sigma_j}$ *if* $r = 0$. *For some positive vector* $x \in K_0$ *assume that* $J_f(x)$, *the Jacobian matrix of* f *at* x, *is irreducible. Then* f *has a unique (to within scalar multiples) eigenvector* x *such that* $x \in K_0$.

Corollary 4.1 provides no information if $f(x)$ is a positive sum of (r,σ) means such that $r < 0$ whenever $(r,\sigma) \in \Gamma_i$. The next corollary is designed to give information about precisely this case. In the statement of the corollary recall that vectors x and y in K are called comparable if there exist positive reals α and β such that

$$\alpha y \leqslant x \leqslant \beta y.$$

COROLLARY 4.3. *For* $1 \leqslant i \leqslant n$ *assume that* Γ_i *is a finite collection of ordered pairs* (r,σ) *such that* $r < 0$ *is a negative real number and* $\sigma \in P$. *In addition assume that if* (r,σ) *and* $(\bar{r},\bar{\sigma})$ *are any two elements of* Γ_i *for* $1 \leqslant i \leqslant n$, *then* σ *and* $\bar{\sigma}$ *are comparable. Define* $f: K \to K$ *by*

$$(f(x))_i = \sum_{(r,\sigma) \in \Gamma_i} c_{ir\sigma} M_{r\sigma}(x),$$

where $c_{ir\sigma}$ *is a positive real number for* $1 \leqslant i \leqslant n$ *and* $(r,\sigma) \in \Gamma_i$. *If there exists* $x \in K_0$ *such that* $J_f(x)$, *the Jacobian matrix of* f *at* x, *is irreducible, then* f *has a unique (to within scalar multiples) eigenvector* u *such that* $u \in K_0$.

Proof. Select a real number $t < 0$ such that $rt \leqslant 1$ for all r such that $(r,\sigma) \in \Gamma_i$ for some $\sigma \in P$ and some i. For $x \in K_0$ define $s = t^{-1}$ and define $h(x)$ by

$$h(x) = \phi_s(f(\phi_t(x)))$$

where ϕ_t is given by Equation (4.20). One can write

$$(h(x))_i = \left\{ \sum_{(r,\sigma)\in\Gamma_i} c_{ir\sigma} w_{r\sigma}^t \right\}^s, \qquad w_{r\sigma} = \left[\sum_{j=1}^n \sigma_j x_j^{rt} \right]^{(rt)^{-1}}. \tag{4.23}$$

We have already remarked that the map $x \to w_{r\sigma}(x)$ has a continuous extension to all of K. Similarly, the map $(w_{r\sigma}) \to (\sum_{(r,\sigma)\in\Gamma_i} c_{ir\sigma} w_{r\sigma}^t)^s$, considered as a map on vectors $(w_{r\sigma})$ (indexed by Γ_i) all of whose components are positive, has a continuous extension to vectors $(w_{r\sigma})$ with nonnegative components. Thus $x \to h(x)$ is a composition of continuous maps, and hence has a continuous extension to all of K.

Clearly h is homogeneous of degree one. We claim that h is superadditive. The proof is essentially the same as the proof of the corresponding fact in the proof of Corollary 4.1, except one must remember that exponentiation to the power t or power t^{-1} now reverses inequalities. Details are left to the reader.

In order to apply Theorem 4.1 to h it remains to show that h is power-bounded below with incidence matrix $J_f(x)$; and it is at this point that we need the special assumptions on Γ_i. For fixed integers i and k, let σ_k denote the kth component of σ for $(r,\sigma)\in\Gamma_i$. Our assumption implies that $\sigma_k = 0$ for all $(r,\sigma)\in\Gamma_i$ or $\sigma_k > 0$ for all $(r,\sigma)\in\Gamma_i$. Assuming that $\sigma_k > 0$ for all $(r,\sigma)\in\Gamma_i$, one has for $(r,\sigma)\in\Gamma_i$ and $x\in K_0$

$$w_{r\sigma}(x) = \left(\sum_{j=1}^n \sigma_j x_j^{rt} \right)^{(rt)^{-1}} \geqslant x_k \sigma_k^{(rt)^{-1}}, \tag{4.24}$$

and using Equation (4.24) in (4.23) gives

$$(h(x))_i \geqslant \left[\sum_{(r,\sigma)\in\Gamma_i} c_{ir\sigma} \sigma_k^{r^{-1}} \right]^s x_k. \tag{4.25}$$

Notice that if we did not know $\sigma_k > 0$ for *all* $(r,\sigma)\in\Gamma_i$, Equation (4.25) would give no information.

Equation (4.25) implies that h is power-bounded below. Furthermore, if one selects $(r,\sigma)\in\Gamma_i$ and puts σ in the ith row of a matrix B, then Equation (4.25) shows that B is an incidence matrix (with respect to being power-bounded below) for h. A simple calculation shows that B has precisely the same positive entries as $J_f(x)$ for any $x\in K_0$, so B is irreducible.

Theorem 4.1 now implies that h (and hence f) has a unique (to within scalar multiples) eigenvector $u\in K_0$. ■

REMARK 4.2. It is unclear exactly when a map $f(x)$ which is positive sum of (r, σ) means has a unique eigenvector u such that $u \in K_0$ and $|u| = 1$. One might conjecture that this should be the case if $J_f(x)$ is an irreducible matrix for all $x \in K_0$; however, the following example shows such a conjecture is false. For $n = 2$ and positive reals c and d, define $f: K \to K$ by $f(x_1, x_2) = (c(x_1^{-1} + x_2^{-1})^{-1}, d(x_1 + x_2))$. For $x \in K_0$, all entries of $J_f(x)$ are positive. However, if $u \in K_0$ is an eigenvector for f with eigenvalue λ and if one defines $t = u_1/u_2 > 0$, a calculation gives

$$c(1+t)^{-1} = \lambda = d(t+1),$$

or

$$(t+1)^2 = \frac{c}{d}.$$

If $d > c$, the latter equation has no solution $t \geqslant 0$, so the original equation has no eigenvector in K_0 if $d > c$. On the other hand, Corollary 4.3 applies to a slight modification of this example, namely $\bar{f}(x_1, x_2) = (c(x_1^{-1} + x_2^{-1})^{-1}, dx_1)$ [think of dx_1 as $d(x_1^{-1})^{-1}$], and Corollary 4.3 implies $\bar{f}$ has a unique positive eigenvector for any $c, d > 0$.

If $f: K \to K$ satisfies H1, $f(K_0) \subset K_0$, and $g(x) = f(x)/|f(x)|$, we next want to study when $g^p(x)$ converges to a positive eigenvector of f. The principal tool to be used is Hilbert's projective metric. Discussions of the projective metric and of generalizations and variants of it can be found in [3, 4, 14, 27, 31]. We list here the definition and basic properties of the projective metric and refer the reader to the above references for further details.

If $x, y \in K_0$, define $\alpha = \sup\{r \mid ry \leqslant x\}$ and $\beta = \inf\{s \mid x \leqslant sy\}$. Then $d(x, y)$, the Hilbert projective metric distance between x and y, is defined by

$$d(x, y) = \log\left(\frac{\beta}{\alpha}\right) \tag{4.26}$$

If $x, y \in K_0$ and λ and μ are any positive reals, then $d(\lambda y, \mu y) = d(x, y)$. If x, y, and z are any elements of K_0, then

$$d(x, z) \leqslant d(x, y) + d(y, z)$$

and

$$d(x, y) = d(y, x).$$

If one restricts d to pairs $x, y \in P_0$, then d is a metric on P_0, and P_0 with metric d is a complete metric space.

If $f: K \to K$ satisfies H1, $f(K_0) \subset K_0$, and $\alpha y \leqslant x \leqslant \beta y$ for $x, y \in K_0$, then $\alpha f(y) \leqslant f(x) \leqslant \beta f(y)$, and one obtains that

$$d(f(x), f(y)) \leqslant d(x, y). \tag{4.27}$$

Furthermore, if $g(x) = f(x)/|f(x)|$ for $x \in K_0$ and f^p denotes the pth iterate of f, $p \geqslant 1$, the properties of d imply

$$d(f^p(x), f^p(y)) = d(g^p(x), g^p(y)). \tag{4.28}$$

With these preliminaries we can state our next theorem.

THEOREM 4.2. *Assume that $f: K \to K$ is homogeneous of degree* 1, *superadditive, and power-bounded below. Suppose also that f has an incidence matrix A (with respect to being power-bounded below) such that A is primitive. If $g = f(x)/|f(x)|$ and $g^k(x)$ denotes the composition of g with itself k times, then for any $x \in K_0$ one has*

$$\lim_{k \to \infty} g^k(x) = u, \tag{4.29}$$

where $u \in P_0$ is the unique eigenvector of f in P_0. If, for a given positive constant R, $B_R(u) = \{x \in K_0 \mid d(x, u) \leqslant R\}$ [$d(x, u)$ denotes Hilbert's projective metric], then $g(B_R(u)) \subset B_R(u)$ and $f(B_R(u)) \subset B_R(u)$, and there exist constants M and c with $0 \leqslant c < 1$ (M and c depend on R and f) such that if $x \in B_R(u)$ and $k \geqslant 1$, then

$$d(g^k(x), u) = d(f^k(x), u) \leqslant Mc^k. \tag{4.30}$$

Proof. Theorem 4.1 implies that f has a unique eigenvector $u \in P_0$. Equation (4.29) will follow for a given $x \in K_0$ if we choose $R \geqslant d(x, u)$ and prove Equation (4.30).

Thus it suffices to pick $R > 0$ and prove the latter part of the theorem. Because $f(u) = \lambda u$, Equation (4.27) implies that if $x \in B_R(u)$

$$d(f(x), u) = d(f(x), f(u)) \leqslant d(x, u) = R,$$

and Equation (4.28) implies that

$$d(g(x),u) \leqslant R.$$

A simple calculation, which we leave to the reader, also shows that if $x \in B_R(u)$,

$$\max_{1 \leqslant i,j \leqslant n} \frac{x_i}{x_j} \leqslant e^R\left(\max_{1 \leqslant i,j \leqslant n} \frac{u_i}{u_j}\right). \tag{4.30a}$$

Select an integer $p \geqslant 1$ such that $A^p > 0$. We claim that there exists a real number c with $0 \leqslant c < 1$ such that for all $x, y \in B_R(u)$,

$$d(f^p(x), f^p(y)) \leqslant c^p d(x,y), \tag{4.31}$$

and by the homogeneity of f it suffices to prove (4.31) for all $x, y \in B_R(u) \cap P$.

Before proving (4.31) recall (Lemma 4.1) that A^p is an incidence matrix for f^p. Thus by the definition of power-bounded below, for each pair of integers (i, j) with $1 \leqslant i, j \leqslant n$, there exists a positive constant b and a vector $\tau \in P$ [both depending on (i, j)] such that τ_j, the jth component of τ, is positive and

$$(f^p(w))_i \geqslant b w^\tau \tag{4.32}$$

for all $w \in K$. Because there are only finitely many pairs (i, j) with $1 \leqslant i, j \leqslant n$, we can select positive numbers δ and λ, independent of (i, j), such that $b \geqslant \delta$ and $\tau_j \geqslant \lambda$ for all the n^2 pairs (b, τ).

Now take any unequal vectors x and y in $B_R(u) \cap P$ and define numbers $\bar{\beta}$ and $\bar{\alpha}$ by

$$\bar{\beta} = \max_{1 \leqslant j \leqslant n} \frac{x_j}{y_j} \equiv \frac{x_{j_1}}{x_{j_1}} \quad \text{and} \quad \bar{\alpha} = \min_{1 \leqslant j \leqslant n} \frac{x_j}{y_j} \equiv \frac{x_{j_0}}{y_{j_0}}. \tag{4.33}$$

Select numbers $\beta \geqslant \bar{\beta}$ and $\alpha \leqslant \bar{\alpha}$, and note that superadditivity and homogeneity give

$$\beta f^p(y) - f^p(\beta y - x) \geqslant f^p(x) \geqslant \alpha f^p(y) + f^p(x - \alpha y). \tag{4.34}$$

If $1 \leqslant i \leqslant n$ and j_0 is as in Equation (4.33), there exists $b \geqslant \delta$ an $\tau \in P$ (depending on i and j_0) such that $\tau_{j_0} \geqslant \lambda$ and Equation (4.32) holds ($\delta > 0$

and $\lambda > 0$ are as in the preceding paragraph). Taking $w = \beta y - x$ in Equation (4.32) gives

$$(f^p(\beta y - x))_i \geqslant \delta(\beta y - x)^\tau \geqslant \delta\Big(\prod_{j \neq j_0} (\beta y_j - \bar{\beta} y_j)^{\tau_j}\Big)(\beta y_{j_0} - \bar{\alpha} y_{j_0})^{\tau_{j_0}}$$
$$\geqslant \delta y^\tau(\beta - \bar{\beta})^{1 - \tau_{j_0}}(\beta - \bar{\alpha})^{\tau_{j_0}} \geqslant \delta y^\tau(\beta - \bar{\beta})^{1-\lambda}(\bar{\beta} - \bar{\alpha})^\lambda. \tag{4.35}$$

There are at most n distinct vectors $\tau \in P$ corresponding to the n pairs (i, j_0), $1 \leqslant i \leqslant n$, and by using Equation (4.30a) one sees that for each such τ,

$$\min\{ y^\tau \mid y \in B_R(u) \cap P \} > 0.$$

It follows that there exists a positive constant ε such that for each of the n vectors τ corresponding to (i, j_0) and all $y \in B_R(u) \cap P$,

$$y^\tau \geqslant \frac{\varepsilon}{\delta}(f^p(y))_i. \tag{4.36}$$

Substituting (4.36) in (4.35) gives

$$f^p(\beta y - x) \geqslant \varepsilon f^p(y)(\beta - \bar{\beta})^{1-\lambda}(\bar{\beta} - \bar{\alpha})^\lambda, \tag{4.37}$$

where $\varepsilon > 0$ and λ, $0 < \lambda \leqslant 1$, are independent of x and y in $B_R(u)$. Substituting (4.37) in (4.34) gives for all $\beta \geqslant \bar{\beta}$

$$\left[\beta - \varepsilon(\beta - \bar{\beta})^{1-\lambda}(\bar{\beta} - \bar{\alpha})\right] f^p(y) \geqslant f^p(x). \tag{4.38}$$

If $\lambda = 1$, one obtains from (4.38) that

$$[\bar{\beta} - \varepsilon(\bar{\beta} - \bar{\alpha})]\, f^p(y) \geqslant f^p(x). \tag{4.39}$$

If $0 < \lambda < 1$, define β so that

$$\left(\frac{\bar{\beta} - \bar{\alpha}}{\beta - \bar{\beta}}\right)^\lambda = \frac{1}{(1-\lambda)\varepsilon},$$

and observe that for this choice of β, Equation (4.38) gives

$$[\bar{\beta} - \theta_1(\bar{\beta} - \bar{\alpha})] f^p(y) \geqslant f^p(x),$$
$$\theta_1 \equiv \frac{\lambda}{1-\lambda}[(1-\lambda)\varepsilon]^{\lambda^{-1}} > 0. \tag{4.40}$$

An exactly analogous proof shows that there exists a positive number θ_2, independent of x and y, such that

$$f^p(x) \geqslant [\bar{\alpha} + \theta_2(\bar{\beta} - \bar{\alpha})] f^p(y) \tag{4.41}$$

If we define $\theta = \min(1, \theta_1, \theta_2) > 0$, Equations (4.40) and (4.41) imply that

$$d(f^p(x), f^p(y)) = d(g^p(x), g^p(y)) \leqslant \log\left(\frac{\bar{\beta} - \theta(\bar{\beta} - \bar{\alpha})}{\bar{\alpha} + \theta(\bar{\beta} + \bar{\alpha})} \right). \tag{4.42}$$

If one writes $z = \bar{\beta}/\bar{\alpha} \geqslant 1$, the right-hand side of (4.42) can be written as

$$\log\left[\frac{(1-\theta)z + \theta}{1 + \theta(z-1)} \right],$$

and we leave it as a calculus exercise for the reader to prove that there exists a positive constant $\kappa < 1$ (depending on θ) such that for all numbers $z \geqslant 1$,

$$\log\left[\frac{(1-\theta)z + \theta}{1 + \theta(z-1)} \right] \leqslant \kappa \log z. \tag{4.43}$$

If one selects c, $0 < c < 1$, such that $c^p = \kappa$, Equations (4.42) and (4.43) imply that for all $x, y \in B_R(u)$,

$$d(f^p(x), f^p(y)) \leqslant c^p d(x, y). \tag{4.44}$$

If m is any positive integer, $0 \leqslant j < p$, and $N = mp + j$, then Equation (4.44) gives (for u the eigenvector of f in P_0)

$$d(f^N x, f^N u) \leqslant c^{mp} d(f^j x, f^j u) \leqslant c^{mp} R$$
$$\leqslant c^N \max_{0 \leqslant j < p} \frac{R}{c^j}.$$

Equation (4.30) follows by defining

$$M = \max_{0 \leqslant j < p} \frac{R}{c^j}. \qquad \blacksquare$$

We now want to apply Theorem 4.2 to the case of positive sums of (r, σ) means.

COROLLARY 4.4. *Let the notation and assumptions be as in the statement of Lemma 4.2, and assume in addition that for some $x \in K_0$, the Jacobian matrix $J_g(x)$ of g at x is a primitive matrix. If $u \in P_0$ denotes the unique positive eigenvector of f, then for any $x \in K_0$,*

$$\lim_{m \to \infty} d(g^m(x), u) = 0,$$

where $d(x, y)$ denotes Hilbert's projective metric. Furthermore, if R is a positive real and $B_R(u) = \{x \in K_0 \mid d(x, u) \leqslant R\}$, then $f(B_R(u)) \subset B_R(u)$, and there exist constants M and c with $0 < c < 1$ such that for all $x \in B_R(u)$ and integers $k \geqslant 1$,

$$d(f^k(x), u) \leqslant Mc^k. \tag{4.45}$$

Proof. Corollary 4.1 implies that f has a unique eigenvector $u \in P_0$, and the proof of Corollary 4.1 showed that f is power-bounded below with incidence matrix $A = J_g(x)$ for any $x \in K_0$.

Select $t > 0$ so that $t \leqslant 1$ and $tr \leqslant 1$ for all $(r, \sigma) \in \Gamma_i$, $1 \leqslant i \leqslant n$, and define $h(x) = \phi_{t^{-1}}(f(\phi_t(x)))$, where ϕ_t is given by Equation (4.20). Lemma 4.4 implies that h is also power-bounded below with incidence matrix A, and in the proof of Corollary 4.1 it was proved that h is superadditive.

Thus Theorem 4.2 applies to h, and since $v = \phi_{t^{-1}}(u)$ is the unique (to within scalar multiples) positive eigenvector of h, given any $R > 0$ there exists a constant M_1 and a constant c, $0 < c < 1$, such that for all $x \in B_{Rt^{-1}}(v)$ and all integers $k \geqslant 1$,

$$d(h^k(x), v) \leqslant M_1 c^k.$$

Notice that $d(\phi_t(x), \phi_t(y)) = |t| d(x, y)$ for all $x, y \in K_0$ and that $h^k = \phi_{t^{-1}} f^k \phi_t$. Thus the previous inequality implies

$$d\big(\phi_{t^{-1}} f^k(\phi_t(x)), \phi_{t^{-1}}(u)\big) \leqslant M_1 c_1^k,$$

or

$$d\big(f^k(\phi_t(x)),u\big)\leqslant (tM_1)c^k \tag{4.46}$$

for all $x\in B_{Rt^{-1}}(v)$ and $k\geqslant 1$. Because

$$\phi_t\big(B_{Rt^{-1}}(v)\big)=B_R(u),$$

the inequality (4.46) is equivalent to (4.45). The other statements of the corollary follow from (4.45). ∎

Our next corollary is a sharpening of Corollary 4.3.

COROLLARY 4.5. *Let the notation and assumptions be as in the statement of Corollary 4.3. In addition assume that there exists $x\in K_0$ such that $J_f(x)$ is a primitive matrix. Then f has a unique eigenvector $u\in K$, and for any $x\in K_0$, $d(f^m x,u)\to 0$. If R is a positive constant, there exist constants M and c, $0<c<1$, such that for all $x\in B_R(u)$ and all integers $k\geqslant 1$*

$$d\big(f^k(x),u\big)\leqslant Mc^k.$$

Proof. Select $t<0$ such that $rt\leqslant 1$ for all $(r,\sigma)\in\Gamma_i$, $1\leqslant i\leqslant n$, and define $h(x)=\phi_{t^{-1}}(f(\phi_t(x)))$ for $x\in K_0$. It was proved in the proof of Corollary 4.3 that h has a continuous extension to K, h is power-bounded below with incidence matrix $A=J_f(x)$, h is superadditive, and h is homogeneous of degree one. Thus h satisfies the hypotheses of Theorem 4.2, and the corresponding conclusions for f can be obtained just as in the proof of Corollary 4.5. The details are left to the reader. ∎

We conclude this paper by giving two simple examples of the previous results. The examples are still difficult enough to require most of the apparatus we have developed.

COROLLARY 4.6. *For $1\leqslant i\leqslant n$ let $G_i\subset P$ be a finite set of probability vectors, and for $\sigma\in G_i$, $1\leqslant i\leqslant n$, let $c_{i\sigma}$ be a positive real. Define $f\colon K\to K$ by*

$$(f(x))_i=\sum_{\sigma\in G_i} c_{i\sigma}x^{\sigma}.$$

Then if $J_f(x)$, the Jacobian matrix of f at x, is irreducible for some positive vector $x\in K_0$, f has a unique (to within scalar multiples) positive eigenvec-

tor $u \in K_0$. If $J_f(x)$ is a primitive matrix for some $x \in K_0$, then for any $y \in K_0$,

$$\lim_{k \to \infty} d(f^k(y), u) = 0$$

where $d(v, w)$ denotes Hilbert's projective metric.

The next example was shown to the author by Dan Weeks, who pointed out that such maps may occur in so-called "two-sex models" in population biology. (See [20] and [28] for background.) Let $n = 6$, and K denote the nonnegative vectors in $\mathbb{R}^6$. Define $f\colon K \to K$ by

$$f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} \phi(x_3, x_6) \\ ax_1 \\ bx_2 \\ cx_1 \\ dx_4 \\ ex_5 \end{pmatrix}. \tag{4.47}$$

In (4.47) assume that a, b, c, d, and e are positive. There are many possible choices for ϕ. One possible assumption is

$$\phi(u, v) = \sum_{j=1}^{m} c_j u^{\tau_j} v^{1-\tau_j}, \tag{4.48}$$

where $0 < \tau_j < 1$ and $c_j > 0$ for $1 \leqslant j \leqslant m$. Another possible choice is

$$\phi(u, v) = \sum_{j=1}^{m} c_j \left[\sigma_j u^{-1} + (1 - \sigma_j) v^{-1}\right]^{-1}, \tag{4.49}$$

where $0 < \sigma_j < 1$ and $c_j > 0$ for $1 \leqslant j \leqslant m$.

Corollary 4.7. *Assume that $f\colon K \to K$ is given by Equation (4.47), where $\phi(x_3, x_6)$ is given either by Equation (4.48) with $0 < \tau_j < 1$ and $c_j > 0$ for $1 \leqslant j \leqslant m$ or by Equation (4.49) with $0 < \sigma_j < 1$ and $c_j > 0$ for $1 \leqslant j \leqslant m$. Then f has an eigenvector w all of whose components are positive, and this eigenvector is unique to within scalar multiples. If $x \in K_0$, then*

$$\lim_{k \to \infty} d(f^k(x), w) = 0,$$

where d denotes Hilbert's projective metric.

Proof. Take $x \in K_0$ and write $M = J_f(x)$. A calculation shows that M is primitive and $M^{16} > 0$. Thus the conclusions of the corollary follow from Corollary 4.4 if ϕ is given by Equation (4.48), and from Corollary 4.5 if ϕ is given by Equation (4.49). [Note that in applying Corollary 4.5, linear terms like ax_1 must be written as $a(x_1^{-1})^{-1}$, so that one always has $r = -1$.] ■

I would like to thank Joel Cohen for encouraging me to establish necessary and sufficient conditions for equality in his theorem and Kingman's theorem. Thanks are also due to Norman Dancer for a helpful, early discussion about Cohen's theorem, and to Dan Weeks for showing me some of the population-biology literature.

REFERENCES

1 E. Artin, *The Gamma Function* (transl. by M. Butler of 1931 notes), Holt, Rinehart and Winston, New York, 1964.

2 F. L. Bauer, An elementary proof of the Hopf inequality for positive operators, *Numer. Math.* 7:331–337 (1965).

3 G. Birkhoff, Extensions of Jentzsch's theorem, *Trans. Amer. Math. Soc.* 85:219–227 (1957).

4 P. J. Bushell, Hilbert's metric and positive contraction mappings in a Banach space, *Arch. Rational Mech. Anal.* 52:330–338 (1973).

5 ____, On a class of Volterra and Fredholm non-linear integral equations, *Math. Proc. Cambridge Philos. Soc.* 79:329–335 (1976).

6 J. E. Cohen, Derivatives of the spectral radius as a function of nonnegative matrix elements, *Math. Proc. Cambridge Philos. Soc.* 83:183–190 (1978).

7 ____, Random evolutions and the spectral radius of a non-negative matrix, *Math. Proc. Cambridge Philos. Soc.* 83:345–350 (1979).

8 ____, Convexity of the dominant eigenvalue of an essentially nonnegative matrix, *Proc. Amer. Math. Soc.* 81:657–658 (1981).

9 ____, Ergodic theorems in demography, *Bull. Amer. Math. Soc.* 1:275–295 (1979).

10 J. E. Cohen, S. Friedland, T. Kato, and F. Kelly, Eigenvalue inequalities for products of matrix exponentials, *Linear Algebra Appl.* 45:55–95 (1982).

11 M. Donsker and S. R. S. Varadhan, On a variational formula for the principal eigenvalue for operators with maximum principle, *Proc. Nat. Acad. Sci. U.S.A.* 72:780–783 (Mar. 1975).

12 S. Friedland, Convex spectral functions, *Linear and Multilinear Algebra* 9:299–316 (1981).

13 S. Friedland and S. Karlin, Some inequalities for the spectral radius of non-negative matrices and applications, *Duke Math. J.* 42:459–490 (1975).

14 M. Golubitsky, E. Keeler, and M. Rothschild, Convergence of the age structure: Application of the projective metric, *Theoret. Population Biol.* 7:84–93 (1975).

15 G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities*, Cambridge U.P., 1934.

16 T. Kato, Superconvexity of the spectral radius and convexity of the spectral bound and the type, *Math. Z.* 180:265–273 (1982).

17 J. F. C. Kingman, A convexity property of positive matrices, *Quart. J. Math.*, 12:283–284 (1961).

18 M. A. Krasnosel'skii, *Positive Solutions of Operator Equations*, Noordhoff, Groningen, 1964.

19 M. G. Krein and M. A. Rutman, Linear operators leaving invariant a cone in a Banach space (in Russian), *Uspekhi Mat. Nauk* 3(1):23 (1948); English transl., *Amer. Math. Soc. Transl.*, No. 26.

20 K. C. Land and A. Rogers (Eds.), *Multidimensional Mathematical Demography*, Academic, New York, 1982.

21 P. L. Lions, letter, Jan. 1984.

22 R. D. Nussbaum, Periodic solutions of some nonlinear autonomous functional differential equations. II, *J. Differential Equations* 14:360–394 (1973).

23 ____, Integral equations from the theory of epidemics, in *Nonlinear Systems and Applications* (V. Lakshmikantham, Ed.), Academic, New York, 1977, pp. 235–255.

24 ____, Periodic solutions of some nonlinear integral equations, in *Dynamical Systems, Proceedings of a University of Florida International Symposium*, (A. Bednarek and L. Cesari, Eds.), Academic, New York, 1977, pp. 221–249.

25 ____, Eigenvectors of nonlinear positive operators and the linear Krein-Rutman theorem, in *Fixed Point Theory*, Springer-Verlag Lecture Notes in Math., Vol. 886, 1981, pp. 309–331.

26 ____, Min-max formulas for the principal eigenvalue of second order elliptic equations with general boundary conditions, in preparation.

27 A. J. B. Potter, Applications of Hilbert's projective metric to certain classes of non-homogeneous operators, *Quart. J. Math.* 28:93–99 (1977).

28 P. A. Samuelson, Generalizing Fisher's "reproductive value": Nonlinear, homogeneous, biparental systems, *Proc. Nat. Acad. Sci. U.S.A.* 74:5772–5775 (1977).

29 E. Seneta, *Non-negative matrices*, Wiley, New York, 1973.

30 M. Sion, On general minimax theorems, *Pacific J. Math.* 8: 171–176 (1958).

31 A. C. Thompson, On certain contraction mappings in a partially ordered vector space, *Proc. Amer. Math. Soc.* 14:438–443 (1963).

32 H. Wielandt, Unzerlegbare nicht negative Matrizen, *Math. Z.* 52:642–648 (1950).

33 M. V. Menon and H. Schneider, The spectrum of a nonlinear operator associated with a matrix, *Linear Algebra Appl.* 2:321–334 (1969).

34 H. Schneider and R. E. L. Turner, Positive eigenvectors of order preserving maps, *J. Math. Anal. Appl.* 37:506–515 (1972).

35 F. F. Bonsall, Linear operators in complete positive cones, *Proc. London Math. Soc.* 8:53–75 (1958).

Received 18 April 1984; revised 23 August 1984