EIGENVALUE INEQUALITIES FOR PRODUCTS OF MATRIX EXPONENTIALS

Joel E. Cohen
Rockefeller University, New York, USA

Shmuel Friedland
Hebrew University, Jerusalem, Israel

Tosio Kato
University of California, Berkeley, California, USA

Frank P. Kelly
University of Cambridge, Cambridge, UK

RR-83-23
September 1983

Reprinted from *Linear Algebra and Its Applications*, volume 45 (1982)

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
Laxenburg, Austria

Reprinted with permission from *Linear Algebra and Its Applications* 45:55–95.

Printed by Novographic, Vienna, Austria

FOREWORD

In IIASA's work on demographic population dynamics and stochastic optimization, certain purely mathematical problems emerge that have contributed to an increase in the understanding of the problems of population evolution. The work of Joel E. Cohen, who spent the summer of 1980 at IIASA, is an example of such results.

ANDRZEJ WIERZBICKI
Program Leader
Systems and Decision Sciences Program

Eigenvalue Inequalities for Products of Matrix Exponentials

Joel E. Cohen
Rockefeller University
New York, New York

Shmuel Friedland
Hebrew University
Jerusalem, Israel

Tosio Kato
University of California
Berkeley, California

and

Frank P. Kelly
University of Cambridge
Cambridge, England

Submitted by Ingram Olkin

ABSTRACT

Motivated by models from stochastic population biology and statistical mechanics, we prove new inequalities of the form $(*)$ $\varphi(e^Ae^B) \geqslant \varphi(e^{A+B})$, where A and B are $n \times n$ complex matrices, $1 < n < \infty$, and φ is a real-valued continuous function of the eigenvalues of its matrix argument. For example, if A is essentially nonnegative, B is diagonal real, and φ is the spectral radius, then $(*)$ holds; if in addition A is irreducible and B has at least two different diagonal elements, then the inequality $(*)$ is strict. The proof uses Kingman's theorem on the log-convexity of the spectral radius, Lie's product formula, and perturbation theory. We conclude with conjectures.

1. INTRODUCTION

Let A and B be $n \times n$ matrices over the field of complex numbers, where n is a fixed integer, $1 < n < \infty$. Let $\varphi(A)$ be a real-valued continuous function of the eigenvalues of A. If $\varphi(A)$ is finite when all elements of A are finite, φ

52 Vanderbilt Ave., New York, NY 10017 0024-3795/82/040055+41$02.75

will be called a *spectral function*. For example, $\varphi(A)$ might be the *spectral radius* of A, which is the maximum of the magnitudes of the eigenvalues of A. Whenever $\log\varphi$ is considered, we shall always assume, without a further explicit statement, that $\varphi > 0$. To emphasize that $\varphi(A)$ depends only on the eigenvalues of A, we assume that any spectral function φ satisfies

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

In this paper, we give conditions on A, B, and φ that imply

$$\varphi(e^A e^B) \geqslant \varphi(e^{A+B}). \tag{1}$$

Our main new results are given in Theorems 1 to 7 below. We also state some conjectures.

Before proceeding to the mathematics, we review the scientific reasons for interest in (1). Under distinct conditions on A, B, and φ, the inequality (1) arises in statistical mechanics and population biology. Products of matrix exponentials under other special assumptions arise also in quantum mechanics [32].

In statistical mechanics, Golden [12] proved that if A and B are Hermitian and nonnegative definite and $\varphi = \text{trace}$, then (1) holds. Independently, Thompson [26] proved (1) if A and B are Hermitian and $\varphi = \text{trace}$, without any requirement that A and B be nonnegative definite. Golden [12] observed that (1) can be used to obtain lower bounds for the Helmholtz free-energy function by appropriate partitioning of the Hamiltonian. Thompson [26] showed that (1) improves a convexity property that has been used to obtain an upper bound for the partition function of an antiferromagnetic chain.

Thompson [27, p. 476] proved (1) for Hermitian matrices A and B and for any continuous real-valued matrix function $\varphi(X)$ satisfying

$$\varphi(XY) = \varphi(YX) \qquad \text{for} \quad Y \text{ positive definite} \tag{2}$$

and

$$\varphi([XX^*]^s) \geqslant |\varphi(X^{2s})|, \qquad s = 1, 2, \ldots, \tag{3}$$

where X^* is the conjugate transpose of the matrix X. All spectral functions satisfy (2). Thompson [27, pp. 477–478] observed that many spectral functions satisfy (3). For example [30], if the eigenvalues $a_1, a_2, \ldots, a_n$ of an arbitrary $n \times n$ complex matrix are ordered so that $|a_1| \geqslant |a_2| \geqslant \cdots \geqslant |a_n|$, then, for $k = 1, \ldots, n$, the function

$$\varphi_k(A) = \sum_{i=1}^{k} |a_i| \tag{4}$$

satisfies (3) for every real positive s. A special case is $r(A) = \varphi_1(A)$.

In Section 2, we obtain inequalities analogous to (1) for arbitrary complex matrices A and B and spectral functions φ that satisfy (3). We apply our first main result, Theorem 1, to several special cases, including that of reversible Markov chains.

The main results of Section 3 are motivated by a problem in population dynamics. Suppose a homogeneous continuous-time population of size $z(t)$, $t \geqslant 0$, grows according to

$$\frac{dz}{dt} = s(t)z(t), \qquad z(0) = 1, \tag{5}$$

where $s(t)$ is the piecewise constant sample path of a continuous-time homogeneous Markov chain, with $n \times n$ intensity matrix Q, taking values in the set $\{s_1, \ldots, s_n\}$ of n real numbers s_i. The random process $z(t)$ is an example of a multiplicative functional [2, p. 98] or a random evolution [13]. If r is the spectral radius, $S = \operatorname{diag}(s_1, \ldots, s_n)$, and $E_i(z(t))$ is the expectation of $z(t)$ given that $s(0) = s_i$, then [5]

$$\lim_{t \to \infty} t^{-1} \log \max_i E_i(z(t)) = \log r(e^{Q+S}). \tag{6}$$

This random evolution $z(t)$ in continuous time can be approximated by a random evolution $y(t)$ in discrete time. Suppose the instantaneous growth rate $s(t)$ governed by the continuous-time chain is observed at $t = 0, 1, 2, \ldots$. The sequence of states occupied would be described by a discrete-time Markov chain with one-step transition probability matrix $P = e^{Q}$. It would be plausible to suppose that if the discrete process were in state i at some integral time t, then

$$y(t+1) = e^{s_i} y(t), \qquad t = 0, 1, \ldots. \tag{7}$$

Denote the expectation of this discrete approximation $y(t)$ given that $s(0)=s_i$ by $E_i(y(t))$, $t=0,1,\dots$. Then [6]

$$\lim_{t\to\infty} t^{-1}\log\max_i E_i(y(t)) = \log r(e^Q e^S). \tag{8}$$

This formula was derived by Cohen [6] as a special case of a formula for the large-time expectation of a Markovian product of random matrices. It can also be derived as a special case of a formula of LeBras [19, p. 441]. When Q is irreducible, $E_i(z(t))$ and $E_i(y(t))$ are independent of i and $\max_i$ can be dropped from (6) and (8).

In numerical examples [6, p. 249], the long-run rate of growth of the average population $E(y(t))$ in the discrete approximation is greater than or equal to the long-run rate of growth of the average population $E(z(t))$ in the continuous-time model. To rationalize this observation, we prove in Theorem 2 of Section 3 that (1) holds when A is an essentially nonnegative matrix (as Q is), B is a diagonal real matrix (as S is), and $\varphi = r$.

In population genetics [3] the stability of equilibrial gene frequencies in organisms that migrate among multiple niches depends on $r(PD)$, where P is a nonnegative row-stochastic $n\times n$ matrix and D is an $n\times n$ diagonal nonnegative matrix. For those special cases where PD takes the form $e^Q e^S$, (1) gives a lower bound on $r(PD)$.

In Section 4, we observe that sufficient conditions for (1) are that, for any positive integer m, $\varphi(A^m)=[\varphi(A)]^m$ and

$$F(t)=\log\varphi(e^{At}e^{Bt}) \text{ is convex}, \qquad t\in[0,\infty). \tag{9}$$

We then show that (9) holds if A and B are Hermitian and φ is the product or sum of the k largest eigenvalues, $k=1,\dots,n$; or if A is the intensity matrix of a reversible Markov chain, B is diagonal real, and φ is the product or sum of the k largest eigenvalues, $k=1,\dots,n$.

Finally, in Section 5, we state conjectures and open problems.

2. INEQUALITIES FOR COMPLEX MATRICES

THEOREM 1. *If A and B are $n\times n$ complex matrices and φ is a spectral function that satisfies* (3), *then*

$$\varphi(e^{(A+A^*)/2}e^{(B+B^*)/2}) \geq |\varphi(e^{A+B})|. \tag{10}$$

Proof. For any complex $n \times n$ matrix M, let $\mathrm{sp}(M)$, the spectrum of M, be the set of n eigenvalues of M, each repeated according to its multiplicity. Since

$$\mathrm{sp}(AB) = \mathrm{sp}(BA) \tag{11}$$

[18, p. 104, Exercise 12], (2) is guaranteed. Let $X = AB$. Then $X^* = B^*A^*$ and $XX^* = ABB^*A^*$. Substituting into (3) gives

$$\varphi\left([ABB^*A^*]^s\right) \geqslant \left|\varphi\left([AB]^{2s}\right)\right|. \tag{12}$$

Setting $s = 2^{k-1}$ for a positive integer k and using (11) on the left in (12) gives

$$\varphi\left([BB^*A^*A]^{2^{k-1}}\right) \geqslant \left|\varphi\left([AB]^{2^k}\right)\right|. \tag{13}$$

By first taking the absolute value of the left member of (13) and then applying (3) and then (11), we have $\varphi([BB^*A^*A]^{2^{k-1}}) \leqslant \left|\varphi\left[\left([BB^*A^*A]^2\right)^{2^{k-2}}\right]\right| \leqslant \varphi([BB^*A^*A(BB^*A^*A)^*]^{2^{k-2}}) = \varphi([BB^*A^*AA^*ABB^*]^{2^{k-2}}) = \varphi([(A^*A)^2(BB^*)^2]^{2^{k-2}})$. Combining this inequality with (13) gives

$$\varphi\left(\left[(A^*A)^2(BB^*)^2\right]^{2^{k-2}}\right) \geqslant \left|\varphi\left([AB]^{2^k}\right)\right|. \tag{14}$$

Repeated application of the steps from (13) to (14), applied to the left member of (14), yields

$$\varphi\left([A^*A]^{2^{k-1}}[BB^*]^{2^{k-1}}\right) \geqslant \left|\varphi\left([AB]^{2^k}\right)\right|. \tag{15}$$

Now replace A by $\exp(2^{-k}A)$ and B by $\exp(2^{-k}B)$ in (15). Since $M = e^A$ implies $M^* = e^{A^*}$,

$$\varphi\left(\left[e^{2^{-k}A^*}e^{2^{-k}A}\right]^{2^k/2}\left[e^{2^{-k}B}e^{2^{-k}B^*}\right]^{2^k/2}\right) \geqslant \left|\varphi\left(\left[e^{2^{-k}A}e^{2^{-k}B}\right]^{2^k}\right)\right|. \tag{16}$$

For any complex $n \times n$ matrices A and B,

$$\lim_{s \to \infty} \left(e^{A/s} e^{B/s}\right)^{s} = e^{A+B}. \tag{17}$$

[We discuss below the provenance of (17).] Let $k \to \infty$ in (16). Now the limit of products is the product of limits and φ is continuous. Thus (16) and (17) imply (10). ■

This proof is very similar to Thompson's [27, Lemma 6, p. 476].

Reed and Simon [22, p. 295] and Davies [7, p. 90] attribute (17) to Lie but give no exact source. Butler and Friedman [4, (12), p. 289] state (17) without proof and with no explicit restrictions on A and B. Marvin H. Friedman (conversation, 30 September 1980) said he and Butler came upon (17) by themselves. They were told by Ed Salpeter, Cornell University, that it had been published previously, though Salpeter gave them no source. Golden [11, (2.14), p. 1284] states (17) independently and proves it by a method that assumes complex square A and B without further restrictions. Equation (17) is generalized by Trotter [29], who does not mention Lie, or Butler and Friedman [4], or Golden [11]. Equation (17), in the matrix case, is attributed to Trotter [29] by Bellman [1, p. 181], Thompson [27, p. 476], and many others. Since, for matrices, (17) probably dates back at least to Lie, the risk of doing a historical injustice could probably be reduced by referring to (17), in matrix applications, as the exponential product formula or Lie's product formula.

Corollary 1 (Thompson [27, p. 476]). *If A and B are $n \times n$ Hermitian matrices and φ is a spectral function that satisfies* (3), *then* (1) *holds.*

Proof. If A is Hermitian, $A = (A + A^*)/2$, so (10) implies (1). ■

Corollary 2. *If A and B are $n \times n$ complex matrices, A is skew-Hermitian, and φ is a spectral function that satisfies* (3), *then*

$$\varphi\left(e^{(B+B^*)/2}\right) \geq \left|\varphi\left(e^{A+B}\right)\right|. \tag{18}$$

If, in addition, B is Hermitian,

$$\varphi\left(e^{B}\right) \geq \left|\varphi\left(e^{A+B}\right)\right|. \tag{19}$$

Proof. If A is skew-Hermitian, then $A^* = -A$, so $(A^*+A)/2=0$. Then use (10). ■

COROLLARY 3. *Under the assumptions of Theorem* 1,

$$\varphi\left(e^{(A+A^*)/2}e^{(B+B^*)/2}\right) \geqslant \varphi\left(e^{(A+B+A^*+B^*)/2}\right) \geqslant \left|\varphi\left(e^{A+B}\right)\right|. \tag{20}$$

Proof. Since $(A+A^*)/2$ is symmetric for any A, Corollary 1 justifies replacing A by $(A+A^*)/2$ and B by $(B+B^*)/2$ in (1), giving the left-hand inequality in (20). Now if $B=0$ in (10), we have, for any complex A,

$$\varphi\left(e^{(A+A^*)/2}\right) \geqslant \left|\varphi\left(e^{A}\right)\right|. \tag{21}$$

Replacing A in (21) by $A+B$ gives the right-hand inequality in (20). ■

If A is a complex $n\times n$ matrix that is *normal*, i.e. $AA^* = A^*A$, then (21) is a direct consequence of (3). For with $X=e^{A/2}$, $s=1$, (3) becomes $\varphi(e^{A/2}e^{A^*/2}) \geqslant |\varphi(e^{A})|$ and $AA^* = A^*A$ implies that $e^{A/2}e^{A^*/2} = e^{(A+A^*)/2}$.

Let A be an $n\times n$ real matrix. Define A to be *essentially nonnegative* if $a_{ij}\geqslant 0$ for all $i\neq j$. Define A to be *quasisymmetric* if there exist real $n\times n$ matrices H and D, H symmetric, D diagonal and nonsingular, such that

$$A = D^{-1}HD. \tag{22}$$

In the theory of n-state homogeneous continuous-time Markoff chains, an *intensity matrix* $Q=(q_{ij})$ is defined to be an $n\times n$ essentially nonnegative matrix such that

$$\sum_{j=1}^{n} q_{ij} = 0, \qquad i=1,\dots,n. \tag{23}$$

An intensity matrix Q is defined to be *reversible* if there exist n positive numbers π_i, $i=1,\dots,n$, such that

$$\pi_i q_{ij} = \pi_j q_{ji}, \qquad i,j=1,\dots,n. \tag{24}$$

LEMMA 1. *Let Q be an $n\times n$ intensity matrix. Then Q is reversible if and only if Q is quasisymmetric.*

Proof. Let Q be reversible. The following proof that Q is quasisymmetric is due to Whittle [31]. If $P=\operatorname{diag}(\pi_i)$, with all $\pi_i>0$, then $(PQ)_{ij}=\pi_i q_{ij}$, while $[(PQ)^T]_{ij}=\pi_j q_{ji}$. Thus (24) is equivalent to

$$PQ=(PQ)^T=Q^TP. \tag{25}$$

If $M=PQ$, (25) says M is symmetric. Therefore $P^{-1/2}MP^{-1/2}=H$ is also symmetric. But $Q=P^{-1/2}HP^{+1/2}$, so Q is quasisymmetric.

Now suppose $Q=D^{-1}HD$, H symmetric, D diagonal nonsingular. For $i,j=1,\dots,n$, $q_{ij}=d_i^{-1}h_{ij}d_j=[h_{ij}/(d_i d_j)]d_j^2$. Thus $Q=CS$, where C is a symmetric matrix with elements $c_{ij}=h_{ij}/(d_i d_j)$ and S is diagonal with diagonal elements $s_i=d_i^2>0$. So $C=QS^{-1}=C^T=S^{-1}Q^T$ implies $SQ=Q^TS$ or $s_i q_{ij}=s_j q_{ji}$, which is reversibility. ■

COROLLARY 4. *If A is a quasisymmetric matrix or a reversible intensity matrix, B is a diagonal real matrix, and φ is a spectral function that satisfies (3), then (1) holds.*

Proof. If A is reversible intensity matrix, A is also quasisymmetric. Therefore $A=D^{-1}HD$ for some real H and D, H symmetric, D diagonal and nonsingular. But $B=D^{-1}BD$, since diagonal matrices commute. So $e^A=D^{-1}e^HD$ and $e^B=D^{-1}e^BD$. Thus $\varphi(e^Ae^B)=\varphi(D^{-1}e^He^BD)=\varphi(e^He^B)\geqslant\varphi(e^{H+B})$ (by Corollary 1) $=\varphi(D^{-1}e^{H+B}D)=\varphi(e^{D^{-1}(H+B)D})=\varphi(e^{A+B})$. ■

COROLLARY 5. *Let A be an $n\times n$ complex matrix with spectrum* $\operatorname{sp}(A)=\{\lambda_1(A),\dots,\lambda_n(A)\}$ *labeled so that*

$$\operatorname{Re}\lambda_1(A)\geqslant\cdots\geqslant\operatorname{Re}\lambda_n(A). \tag{26}$$

Then

$$\lambda_1\left(\frac{A+A^*}{2}\right)\geqslant\operatorname{Re}\lambda_1(A). \tag{27}$$

This result is attributed to Hirsch by Marshall and Olkin [20, p. 238].

Proof. In (21), take $\varphi=r$, the spectral radius. Since $(A+A^*)/2$ is Hermitian, its spectrum is real. Hence $r(e^{(A+A^*)/2})=\exp\lambda_1[(A+A)/2]$.

Also $r(e^A) = \max_{i=1,\ldots,n}|\exp \lambda_i(A)| = \exp[\operatorname{Re} \lambda_1(A)]$. Since $r(e^{(A+A^*)/2}) \geq r(e^A)$ by (21), taking logarithms of both sides yields (27). ■

For any matrix A, write $A \geq 0$ and say A is nonnegative if every element of A is real and nonnegative; write $A > 0$ and say A is *positive* if $A \geq 0$ and no element of A is 0.

Define an $n \times n$ matrix A to be *irreducible* if, for each $i, j = 1, \ldots, n$, there is a positive integer k such that $(A^k)_{ij} \neq 0$.

COROLLARY 6. *If A is an $n \times n$ matrix and $A \geq 0$, then*

$$r\left(\frac{A + A^*}{2}\right) \geq r(A). \tag{28}$$

Suppose, in addition, that A is irreducible. Then equality holds in (28) *if and only if, for some n-vector $u > 0$ such that $u^T u = 1$,*

$$Au = r(A)u \tag{29}$$

and

$$A^T u = r(A)u. \tag{30}$$

Proof. By the Perron-Frobenius theorem [18, 25], $r(A) = \operatorname{Re} \lambda_1(A)$. Then (28) follows from (27). Now suppose A is irreducible. If (29) and (30) hold, then $(A + A^T)u = 2r(A)u$, so u is a positive eigenvector of the nonnegative irreducible matrix $A + A^T$. Thus $r(A + A^T) = 2r(A)$, and equality holds in (28). Conversely, assume (29) and equality in (28). Now $r(A + A^*) = \max\{x^T(A + A^*)x : x \text{ is a real } n\text{-vector and } x^T x = 1\}$ and the maximum is attained at the n-vector v such that $(A + A^*)v = r(A + A^*)v$ [18, pp. 109–110]. But for u given by (29), $u^T(A + A^*)u = u^T(Au) + (u^T A^T)u = 2r(A)u^T u = r(A + A^*)$, so $u = v$. Therefore $(A + A^T)u = Au + A^T u = r(A)u + A^T u = 2r(A)u$, which implies (30). ■

3. INEQUALITIES FOR ESSENTIALLY NONNEGATIVE MATRICES

The major results of this section depend on a simple but powerful result of Kingman [16]. Define a function $f(t)$ to be *log-convex* for t in some interval if and only if $f(t) > 0$ and $\log f(t)$ is convex for t in the interval. Kingman's theorem is this: If $A(t)$ is an $n\times n$ nonnegative matrix function of a parameter t on some interval such that $r(A(t))>0$ on the interval and, for $i,j=1,\dots,n$, either $a_{ij}(t)$ vanishes or $a_{ij}(t)$ is log-convex on the interval, then $\log r(A(t))$ is a convex function of t on the interval.

Define an $n\times n$ matrix $B=(b_{ij})$ to be *real diagonal*, and write $B=\operatorname{diag}(b_1,\dots,b_n)$ if $b_{ij}=0$ when $i\neq j$ and $b_{ii}=b_i$ with all b_i real.

Define an $n\times n$ matrix B to be a *scalar* matrix if there is a (real or complex) scalar b such that $B=bI$, where I is the $n\times n$ identity matrix.

THEOREM 2. *If A is an $n\times n$ essentially nonnegative matrix, B is an $n\times n$ real diagonal matrix, and $\varphi=r$, the spectral radius, then* (1) *holds. The inequality* (1) *is strict if A is also irreducible and B is also not a scalar matrix.*

The proof depends on Lemmas 2 to 5. For two complex vectors u and v, we denote $(u,v)=\sum_{i=1}^{n}u_i\bar{v}_i$ and $\|u\|=(u,u)^{1/2}$. For any $n\times n$ complex matrix A, define $\|A\|=[r(A^*A)]^{1/2}$.

LEMMA 2. *Let $A\geq 0$ be an irreducible $n\times n$ matrix. Then there exists an $n\times n$ real diagonal matrix $S=\operatorname{diag}(s_1,\dots,s_n)$ with $s_i>0$, $i=1,\dots,n$, and an n-vector $w>0$ such that $A_1=SAS^{-1}$ satisfies*

$$A_1w=r(A)w,\ A_1^Tw=r(A)w. \tag{31}$$

Moreover, $\|A_1\|=\|A_1^T\|=r(A_1)=r(A_1^T)$.

Proof. Since A is irreducible, a theorem of Frobenius [10, vol. 2, p. 53] implies that there exist n-vectors u and v such that

$$\begin{aligned} Au=r(A)u,&\qquad A^Tv=r(A)v,\\ u>0,\quad v>0,&\qquad u^Tv=1.\end{aligned} \tag{32}$$

For $i=1,\dots,n$, let $s_i=(v_i/u_i)^{1/2}$. Since $s_i>0$, $SAS^{-1}\geq 0$ is irreducible. Again by Frobenius's theorem, there exists an n-vector $w>0$ such that $SAS^{-1}w=r(SAS^{-1})w=r(A)w$. In fact, with $w=Su=S^{-1}v$, both parts of (31) hold because of (32).

To see that $\|A_1\|=r(A_1)$, multiply $A_1w=r(A)w$ on the left by A_1^T and use (31) to get $A_1^TA_1w=r^2(A)w=r^2(A_1)w$. ■

LEMMA 3. *In Lemma 2, assume further that A^TA is irreducible and $r(A)=1$. (In case $r(A)\neq 1$, replace A by $A/r(A)$.) Then, for any real*

n-vector x, $\|A_1x\| = \|x\|$ if and only if $A_1x = x = cw$ for some real scalar c, where w satisfies (31).

Proof. If A^TA is irreducible, then so is $A_1^TA_1$, which is symmetric and has $\|A_1\| = 1$. But $\|A_1x\| = \|x\|$ if and only if $(A_1^TA_1x, x) = (A_1x, A_1x) = (x, x)$, which is true if and only if $A_1^TA_1x = x$. The lemma follows because all eigenvectors of $A_1^TA_1$ corresponding to eigenvalue 1 must be of the form cw for scalar c. ■

LEMMA 4. *Let A be a nonnegative $n \times n$ matrix, C a real diagonal $n \times n$ matrix. Then $r(Ae^{tC}Ae^{-tC}) \geqslant r(A)^2$ for all real numbers t. The inequality is strict, except for $t = 0$, if both A^2 and A^TA are irreducible and C is not a scalar matrix.*

Proof. It suffices to prove the strict inequality; the general case then follows by continuity.

The elements of $A(t) = Ae^{tC}Ae^{-tC}$ are log-convex in t or identically 0 for all t. According to Kingman's theorem, it follows that $r(A(t))$ is convex in t. Since $A(t)$ is irreducible with A^2 (because $e^{\pm tC}$ are positive diagonal), $r(A(t))$ is a simple eigenvalue of $A(t)$. Since $A(t)$ is analytic in t, it follows that $r(A(t))$ is analytic in t. To prove the strict inequality, therefore, it suffices to show that

$$\left[\frac{dr(A(t))}{dt}\right]_{t=0} = 0 \tag{33}$$

and, in addition, that $r(A(t))$ is not constant in t.

A perturbation theorem of Kato [15, Equation II-(2.36), p. 81] gives

$$\left[\frac{dr(A(t))}{dt}\right]_{t=0} = \left(\left[\frac{dA(t)}{dt}\right]_{t=0} u, v\right) = ((ACA - A^2C)u, v) = 0,$$

where u and v are respectively the eigenvectors of A and A^T for the principal eigenvalue $r(A)$, normalized by $(u, v) = 1$. Note that $(ACAu, v) = r(A)^2(Cu, v) = (A^2Cu, v)$.

It remains to show that $r(A(t))$ is not constant. We may assume, without loss of generality, that $r(A) = 1$. Suppose that $r(A(t)) = r(A_1(t))$ were constant, where $A_1(t) = A_1e^{tC}A_1e^{-tC} = SA(t)S^{-1}$ with S and A_1 as in Lemma 2. Then $r(A_1(t)) = r(A_1)^2 = 1$ for all real t. Since $r(A_1(t))$ is an eigenvalue of $A_1(t)$, which is analytic in t, it follows from perturbation theory that $A_1(it)$

also has an eigenvalue 1 for all real t. Let $u_1 = u_1(t)$ be an associated eigenvector normalized to norm 1:

$$A_1 e^{itC} A_1 e^{-itC} u_1 = u_1, \qquad \|u_1\| = 1. \tag{34}$$

Set $u_2 = e^{-itC}u_1$, $u_3 = A_1 u_2$, $u_4 = e^{itC}u_3$, so that (34) gives $u_1 = A_1 u_4$; here all the u_k depend on t. Since $\|A_1\| = \|e^{\pm itC}\| = 1$, we have

$$1 = \|u_1\| \leqslant \|u_4\| \leqslant \|u_3\| \leqslant \|u_2\| \leqslant \|u_1\|. \tag{35}$$

Hence we must have equality everywhere. In particular $\|u_3\| = \|A_1 u_2\| = \|u_2\|$. By Lemma 3 this implies that $u_3 = A_1 u_2 = u_2 = c_2(t)w$. Similarly, we have $\|u_1\| = \|A_1 u_4\| = \|u_4\|$, which implies $u_1 = A_1 u_4 = u_4 = c_1(t)w$. [Here $c_1(t), c_2(t)$ are nonzero scalars.] It follows that $c_1(t)w = e^{itC}c_2(t)w$, which is obviously impossible if C is not a scalar matrix. ■

If, in Lemma 4, one replaces the assumption that A^2 and A^TA are irreducible by the stronger assumption that $A > 0$, it is easier to show $r(A(t))$ is not a constant. Assume $A > 0$ and C is not a scalar matrix. Then $e^{tC}Ae^{-tC}$ has at least one element larger than ae^{ct}, where $a > 0$, $c > 0$, say in row i and column j. Hence $[A(t)]_{ij} \geqslant ae^{ct}$. Therefore $r(A(t)) \geqslant ae^{ct}$. So $r(A(t))$ cannot be a constant for $-\infty < t < \infty$.

LEMMA 5. *If A and B are $n \times n$ nonnegative real matrices and B is diagonal, then*

$$r(AB) \leqslant [r(A^2B^2)]^{1/2}. \tag{36}$$

(The positive root is always intended.) If, in addition, A^2 and A^TA are irreducible and B is not a scalar matrix and B is nonsingular, then the inequality in (36) *is strict.*

Proof. It is sufficient to prove (36) assuming that B is nonsingular diagonal. For if B is singular, i.e. some $b_i = 0$, we can choose a sequence $\{B_k\}$ of nonsingular diagonal matrices B_k such that $B_k \to B$, as $k \to \infty$, and (36) will then hold by the continuity of r. So if B is nonsingular diagonal, then $r(A^2B^2) = r(AB^2A) = r([AB]B[AB]B^{-1}) \geqslant [r(AB)]^2$; the inequality follows from Lemma 4, with $B = e^C$, $t = 1$, and AB here replacing A in Lemma 4.

Strict inequality in (36) follows similarly from the conditions that assure the strict inequality in Lemma 4. ■

Define an $n \times n$ matrix A to be *primitive* if $A \geqslant 0$ and there is a positive integer k such that $A^k > 0$. A primitive matrix is irreducible, but not necessarily conversely.

If A is primitive but not positive, then strict inequality in (36) need not hold. For example, let $0 < a < 1$, $b > 0$, and

$$A = \begin{pmatrix} 0 & ab & 1-a \\ 0 & (1-a)b & a \\ 1 & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/b & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then $A^3 > 0$, while $r(AB) = r(A^2B^2) = 1$.

A slight modification of this example shows that Remark 3.2 of Friedland and Karlin [9, p. 471] is false. Take $a = \frac{1}{2}$, $b = 1$ in A and B above, and define $D = \operatorname{diag}(d, 1, d^{-1})$, $0 < d < 1$. Being doubly stochastic, A has left and right eigenvectors $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and $(1,1,1)^T$ corresponding to $r(A) = 1$. Moreover $r(AD) = 1$, since

$$DA\begin{pmatrix} d \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} d \\ 1 \\ 1 \end{pmatrix}.$$

Thus $r(DA) = r(AD) = 1 = d^{1/3}1^{1/3}d^{-1/3}$, which is equality in Equation (1.8) of Friedland and Karlin [9] even though D is not a scalar matrix, contrary to their Remark 3.2. The conclusions of Remark 3.2 are true if the $n \times n$ matrix $M \geqslant 0$ there is assumed to be irreducible and to have positive diagonal. The proof follows that of Theorem 3.2 in [9, p. 471].

Under the assumptions of Lemma 5, it need not be true that $r(A^3B^3) \geqslant [r(A^2B^2)]^{3/2}$. For example, if

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

then $r(A^3B^3) = 8^{1/2}$ while $r(A^2B^2) = 4$.

Proof of Theorem 2. Assume that A is irreducible and B is not a scalar matrix. Then $e^A > 0$. By Lemma 5, replacing B there by e^B here, for $k = 0, 1, 2, \ldots,$

$$\left[r\left(e^{A/2^k}e^{B/2^k}\right)\right]^{2^k} > \left[r\left(e^{A/2^{k+1}}e^{B/2^{k+1}}\right)\right]^{2^{k+1}}. \tag{37}$$

For any $n\times n$ matrix $M\geqslant 0$, $r(M^k)=[r(M)]^k$. Comparing the left side of (37) for $k=0$ with the limit of the right side of (37) in the limit as $k\to\infty$, and using (17), gives $r(e^Ae^B)>r(e^{A+B})$.

If B is a scalar matrix, then $r(e^Ae^B)=r(e^{A+B})$. If A is reducible, then $r(e^Ae^B)\geqslant r(e^{A+B})$ follows from (36) as does (37). ■

S. R. S. Varadhan (personal communication, 6 May 1981) pointed out that if Q is an $n\times n$ intensity matrix and S is an $n\times n$ real diagonal matrix, then

$$r(e^Qe^S)\geqslant r(e^{Q+S}) \tag{38}$$

follows from Lemma 3.1 of Donsker and Varadhan [8, p. 33]. Here is his argument. Let $p=(p_1,\dots,p_n)$ be an arbitrary probability distribution on the integers $1,\dots,n$, so that $p_i\geqslant 0$, $p_1+\cdots+p_n=1$. Then (6) and (8) above combine, respectively, with Equations (1.16) and (1.9) of Donsker and Varadhan [8, pp. 6, 4] to yield

$$\log r(e^{Q+S})=\sup_p\left[\sum_{i=1}^n s_ip_i-I(p)\right], \tag{39}$$

$$\log r(e^Qe^S)=\sup_p\left[\sum_{i=1}^n s_ip_i-I_1(p)\right]. \tag{40}$$

The functions $I(p)$ and $I_1(p)$ need not be defined here explicitly. But, for every p, according to their Lemma 3.1,

$$I_1(p)\leqslant I(p). \tag{41}$$

(Take $h=1$ in their notation.) Then (38) follows immediately from using the inequality (41) in (39) and (40).

We now show that Theorem 2 sharpens a special case of Theorem 3.1 of Friedland and Karlin [9, p. 462].

LEMMA 6. *Let A be an essentially nonnegative $n\times n$ matrix with eigenvalues $\{\lambda_i\}_1^n$ ordered by (26), so that $\lambda_1=r(A)$. Suppose there exist n-vectors u and v such that (32) holds. Then for any $n\times n$ real diagonal matrix $D=\operatorname{diag}(d_1,\dots,d_n)$, if δ_1 is the necessarily real eigenvalue of $A+D$*

with largest real part,

$$\delta_1 \geqslant \lambda_1 + \sum_{i=1}^{n} u_i v_i d_i. \tag{42}$$

Proof. If $A=(a_{ij})$ and $\varepsilon>0$, define $A(\varepsilon)$ by

$$(A(\varepsilon))_{ij} = a_{ij} + \varepsilon u_i v_j, \qquad i,j=1,\dots,n.$$

So if a is a sufficiently large positive scalar,

$$A(\varepsilon)+aI>0 \qquad \text{for all} \quad \varepsilon>0,$$

$$[A(\varepsilon)+aI]u=(\lambda_1+\varepsilon+a)u,$$

$$[A^T(\varepsilon)+aI]v=(\lambda_1+\varepsilon+a)v.$$

Corollary 3.1 of Friedland and Karlin [9, p. 471] implies that, for any n-vector $x=(x_i)>0$,

$$\sum_{i=1}^{n} \frac{u_i v_i \{[A(\varepsilon)+aI]x\}_i}{x_i} \geqslant \lambda_1+\varepsilon+a.$$

Now choose a large enough so that, for any $\varepsilon>0$,

$$A(\varepsilon)+aI+D>0.$$

Then by the Perron-Frobenius theorem, there exists an n-vector $y>0$ such that

$$[A(\varepsilon)+aI+D]y=[\delta_1(\varepsilon)+a]y$$

where $\delta_1(\varepsilon)$ is the (necessarily real) eigenvalue of $A(\varepsilon)+D$ with largest real part. Consequently

$$\sum_{i=1}^{n} \frac{u_i v_i \{[A(\varepsilon)+aI+D]y\}_i}{y_i} = \sum_{i=1}^{n} u_i v_i[\delta_1(\varepsilon)+a] = \delta_1(\varepsilon)+a;$$

but also

$$\sum_{i=1}^{n} \frac{u_i v_i \{[A(\varepsilon)+aI+D]y\}_i}{y_i} = \sum_{i=1}^{n} \frac{u_i v_i [A(\varepsilon)y]_i}{y_i} + a + \sum_{i=1}^{n} u_i v_i d_i$$

$$\geqslant \lambda_1 + \varepsilon + a + \sum_{i=1}^{n} u_i v_i d_i .$$

Hence

$$\delta_1(\varepsilon) \geqslant \lambda_1 + \varepsilon + \sum_{i=1}^{n} u_i v_i d_i .$$

As $\varepsilon \to 0$, $\delta_1(\varepsilon) \to \delta_1$, and we get (42). ■

COROLLARY 7. *Let A be an essentially nonnegative $n \times n$ matrix such that there exist n-vectors u and v that satisfy* (32). *Then for any $n \times n$ real diagonal matrix D*

$$r(e^A e^D) \geqslant r(e^{A+D}) \geqslant r(e^A) \prod_{i=1}^{n} e^{d_i u_i v_i}. \tag{43}$$

Proof. The left inequality in (43) follows from Theorem 2. Defining (again) δ_1 as the eigenvalue of $A+D$ with largest real part, and using Lemma 6,

$$r(e^{A+D}) = e^{\delta_1} \geqslant e^{\lambda_1 + \Sigma u_i v_i d_i} = r(e^A) \prod_{i=1}^{n} e^{d_i u_i v_i}.$$ ■

Equation (1.8′) of Friedland and Karlin implies only that

$$r(e^A e^D) \geqslant r(e^A) \prod_{i=1}^{n} e^{d_i u_i v_i}.$$

We now present another line of argument leading to the weak inequality asserted in Theorem 2.

LEMMA 7. *For a positive integer k, let $A_1, \dots, A_k$ be $n \times n$ nonnegative commuting matrices, i.e. $A_i A_j = A_j A_i$, and call their product $C = A_1 A_2 \cdots A_k$. Let $b_1, \dots, b_k$ be nonnegative scalars, and call their sum $b = \sum_{i=1}^{k} b_i$.*

Let $D_1,\dots,D_k$ be $n\times n$ diagonal matrices. Then

$$\log r\left(A_1 e^{b_1 D_1}\cdots A_k e^{b_k D_k}\right)\leqslant \sum_{i=1}^{k}\frac{b_i}{b}\log r\left(Ce^{bD_i}\right).$$

In particular, if all $D_i=D$, then

$$r\left(A_1 e^{b_1 D}\cdots A_k e^{b_k D}\right)\leqslant r\left(Ce^{bD}\right). \tag{44}$$

Proof. Kingman's [16] theorem implies that $\log r(A_1e^{D_1}\cdots A_ke^{D_k})$ is a convex function of the kn diagonal elements of $D_1,\dots,D_k$. Thus, for fixed diagonal matrices D_i, $i=1,\dots,k$, define

$$g(b_1,\dots,b_k)=\log r\left(A_1 e^{b_1 D_1}\cdots A_k e^{b_k D_k}\right).$$

By Kingman's theorem g is a convex function of $(b_1,\dots,b_k)$, whether or not $b_i\geqslant 0$. Now suppose $b_i\geqslant 0$, $i=1,\dots,k$. Let e_i be the k-tuple with every element equal to 0 except the ith and with the ith element equal to 1, for $i=1,\dots,k$. Then clearly

$$(b_1,\dots,b_k)=\sum_{i=1}^{k}\frac{b_i}{b}be_i.$$

By the convexity of g,

$$\begin{aligned}
g(b_1,\dots,b_k)&\leqslant\sum_{i=1}^{k}\frac{b_i}{b}g(be_i)\\
&=\sum\frac{b_i}{b}\log r\left(A_1\cdots A_ie^{bD_i}A_{i+1}\cdots A_k\right)\\
&=\sum\frac{b_i}{b}\log r\left(A_{i+1}A_{i+2}\cdots A_kA_1\cdots A_ie^{bD_i}\right)\\
&=\sum\frac{b_i}{b}\log r\left(Ce^{bD_i}\right).
\end{aligned}$$

■

THEOREM 3. *Let A be an essentially nonnegative $n\times n$ matrix, and B be a diagonal real $n\times n$ matrix. Let $a_i\geqslant 0$, $b_i\geqslant 0$, $i=1,\dots,k$, and $a=\Sigma_i a_i$,*

$b=\Sigma_i b_i$. *Then*

$$r\left(e^{a_1 A}e^{b_1 B}\cdots e^{a_k A}e^{b_k B}\right)\leqslant r\left(e^{aA}e^{bB}\right). \tag{45}$$

Proof. Set $A_i=e^{a_i A}$, $i=1,\ldots,k$, in Lemma 7, so that $C=e^{aA}$. ∎

COROLLARY 8. *If A is an essentially nonnegative $n\times n$ matrix and B is an $n\times n$ diagonal real matrix, then*

$$r\left(e^{A+B}\right)\leqslant r\left(e^{A}e^{B}\right).$$

Proof. With $a_i=b_i=1/k$ in (45), Theorem 3 implies that $r[(e^{A/k}e^{B/k})^k]\leqslant r(e^A e^B)$. Let $k\to\infty$ and apply (17). ∎

Recall that Theorem 2 and Corollary 8 imply

$$\lim_{t\to\infty} t^{-1}\log\max_i E_i(z(t))\leqslant \lim_{t\to\infty} t^{-1}\log\max_i E_i(y(t)),$$

where $z(t)$ is a continuous-time random evolution and $y(t)$ is its discrete-time approximation, as defined in Section 1. We now show that, provided that the initial state of the random evolutions is distributed according to the equilibrium distribution of the governing Markov chain, we have $E(z(t))\leqslant E(y(t))$ for $t=0,1,2,\ldots,$and we give sufficient conditions for strict inequality. As before, these inequalities for random evolutions follow from more general inequalities for essentially nonnegative matrices.

We say that a real-valued function f is *strictly log-convex* if $f>0$ and $\log f$ is strictly convex.

LEMMA 7A. *For $d_1\geqslant 0,\ldots,d_m\geqslant 0$, $\Sigma_i d_i>0$, and real $t,c_1,\ldots,c_m$, let $f(t)=\Sigma_{i=1}^m d_i e^{tc_i}$. Then $\log f(t)$ is convex in t and is not strictly convex in t if and only if there exists c such that, whenever $d_i>0$, we have $c_i=c$.*

Proof. $\log f$ is convex if and only if $f''f-(f')^2\geqslant 0$, which follows from the Cauchy-Buniakowsky-Schwarz inequality. Necessary and sufficient conditions for the CBS inequality to be an equality are, in this application, just that $c_i=c$ whenever $d_i>0$. ∎

LEMMA 7B. *Let A be an essentially nonnegative $n\times n$ matrix and B be a diagonal real $n\times n$ matrix. Let x and y be nonnegative n-vectors. Then for*

$k \geqslant 1$ and $a_1 \geqslant 0, \ldots, a_k \geqslant 0$, $a_{k+1} \geqslant 0$, and real $b_1, \ldots, b_k$, define the real-valued function

$$h(b_1, \ldots, b_k) = x^T e^{a_1 A} e^{b_1 B} \cdots e^{a_k A} e^{b_k B} e^{a_{k+1} A} y.$$

If there exists $(b_1, \ldots, b_k)$ at which $h > 0$, then $h > 0$ for all $(b_1, \ldots, b_k)$. Provided $h > 0$,

$$g = \log h$$

is convex in $(b_1, \ldots, b_k)$. If, in addition, A is irreducible, B is not a scalar matrix, $x > 0$, $y > 0$, and $a_2 \cdots a_k > 0$ (interpret $a_2 \cdots a_k = 1$ if $k = 1$), then $h > 0$ and g is strictly convex in $(b_1, \ldots, b_k)$.

Proof. Since the diagonal elements of $e^{b_i B}$, $i = 1, \ldots, k$, are all positive, regardless of b_i, if some $(b_1, \ldots, b_k)$ makes $h(b_1, \ldots, b_k)$ positive, no other $(b_1, \ldots, b_k)$ could make $h(b_1, \ldots, b_k) = 0$.

Now assume A is irreducible, B is not a scalar matrix, $x > 0$, $y > 0$, and $a_2 \cdots a_k > 0$. Then $x^T e^{a_1 A} > 0$ for all $a_1 \geqslant 0$, and $e^{a_{k+1} A} y > 0$ for all $a_{k+1} \geqslant 0$, and $e^{a_j A} > 0$ for $j = 2, \ldots, k$. So $h > 0$.

To show $g = \log h$ is strictly convex in $(b_1, \ldots, b_k)$, it suffices to show that, for every w_i and x_i, $i = 1, \ldots, k$, such that $\Sigma_i |w_i| \neq 0$, if

$$b_i = w_i t + x_i,$$

then g is a strictly convex function of the real variable t. Let

$$h(w_1 t + x_1, \ldots, w_k t + x_k) = h(t) = \sum_{j=1}^{n^k} d_j e^{c_j t},$$

where d_j and c_j, $j = 1, \ldots, n^k$, are functions of x, y, A, B, a_i, w_i, and x_i, and all $d_j > 0$. We must show that for no real c do we have $c_j = c$, $j = 1, \ldots, n^k$.

Suppose $B = \operatorname{diag}(b_{11}, \ldots, b_{nn})$ and $b_{11} \neq b_{22}$. Then we can order the coefficients c_j so that $c_1 = b_{11} \Sigma_{i=1}^k w_i$ and $c_2 = b_{22} \Sigma_{i=1}^k w_i$. [To see this for c_1, consider the summand $x_1 (e^{a_1 A})_{11} (e^{b_1 B})_{11} \cdots (e^{a_{k+1} A})_{11} y_1$.] So if $\Sigma_{i=1}^k w_i \neq 0$, then $c_1 \neq c_2$. By Lemma 7A, $h(t)$ is strictly log-convex.

If $\Sigma_{i=1}^{k} w_i = 0$, then since $\Sigma_i |w_i| \neq 0$, we may assume that $w_l \neq 0$, say, where l is fixed and $1 \leqslant l \leqslant k$. Now $h(t)$ contains the summand

$$x_1(e^{a_1 A})_{11}(e^{b_1 B})_{11} \cdots (e^{a_{l-1} A})_{11}(e^{b_{l-1} B})_{11}(e^{a_l A})_{12}(e^{b_l B})_{22}$$
$$\times (e^{a_{l+1} A})_{21}(e^{b_{l+1} B})_{11} \cdots (e^{a_{k+1} A})_{11} y_1$$

which leads to a coefficient, say c_3, such that

$$\begin{aligned} c_3 &= w_1 b_{11} + \cdots + w_{l-1} b_{11} + w_l b_{22} + w_{l+1} b_{11} + \cdots + w_k b_{11} \\ &= w_l b_{22} + \left(\Sigma_{i=1}^{k} w_i - w_l\right) b_{11} \\ &= w_l (b_{22} - b_{11}). \end{aligned}$$

Another term of $h(t)$ similarly leads to a coefficient, say c_4, where

$$\begin{aligned} c_4 &= w_1 b_{22} + \cdots + w_{l-1} b_{22} + w_l b_{11} + w_{l+1} b_{22} + \cdots + w_k b_{22} \\ &= w_l (b_{11} - b_{22}). \end{aligned}$$

Thus $c_3 \neq c_4$. Again Lemma 7A shows that $h(t)$ is strictly log-convex.

The log-convexity of $g = \log h$ holds by continuity if the assumptions that imply strict log-convexity are dropped. ∎

THEOREM 3A. *Let A and B be real $n \times n$ matrices, A essentially nonnegative and B diagonal. Let $\lambda = r(e^A)$; then $\lambda > 0$. Let $u \geqslant 0$ and $v \geqslant 0$ be n-vectors such that $v^T e^A = \lambda v^T$ and $e^A u = \lambda u$. Then, for $t = 0, 1, 2, \ldots$ and $k = 1, 2, \ldots,$*

$$v^T \left(e^{A/(2k)} e^{B/(2k)}\right)^{2kt} u \leqslant v^T \left(e^{A/k} e^{B/k}\right)^{kt} u$$

and

$$v^T e^{(A+B)t} u \leqslant v^T \left(e^A e^B\right)^t u.$$

If, in addition, $t \geqslant 1$, A is irreducible, and B is not a scalar matrix, then both inequalities are strict.

Proof. Assume A irreducible and B not scalar. Then $u > 0$ and $v > 0$, and both u and v are unique to within scalars. (If A is not irreducible, none of these facts need hold.)

Define, for $t \geqslant 1$,

$$h(b_1,\ldots,b_{2k}) = \exp[g(b_1,\ldots,b_{2k})]$$

$$= v^T\left[e^{A/(2k)}e^{b_1B}e^{A/(2k)}e^{b_2B}\cdots e^{A/(2k)}e^{b_{2k}B}\right]^t u.$$

By Lemma 7B, g is strictly convex, and therefore so is h. Let α and β be $(2k)$-tuples defined by

$$\alpha = (0,1/k,0,1/k,\ldots,0,1/k),$$

$$\beta = (1/k,0,1/k,0,\ldots,1/k,0).$$

Then

$$h(\alpha) = h(\beta) = v^T\left[e^{A/k}e^{B/k}\right]^{kt}u,$$

while

$$h\left(\frac{\alpha+\beta}{2}\right) = v^T\left[e^{A/(2k)}e^{B/(2k)}\right]^{2kt}u.$$

Because h is strictly convex,

$$h\left(\frac{\alpha+\beta}{2}\right) < \frac{h(\alpha)+h(\beta)}{2},$$

which is the first inequality to be proved. Hence

$$v^T\left[e^{A/(2^k)}e^{B/(2^k)}\right]^{2^k t}u$$

is a strictly decreasing function of $k = 1,2,\ldots$. Again (17) gives, as $k \to \infty$,

$$v^T e^{(A+B)t}u < v^T\left[e^A e^B\right]^t u.$$

The weak inequalities when A is reducible or B is a scalar matrix or $t = 0$ follow by continuity. ■

COROLLARY 8A. *Let $z(t)$, $t \geq 0$, be the continuous-time random evolution governed by (5), with $n \times n$ intensity matrix Q. Let π be an equilibrium vector of Q, i.e. $\pi^T Q = 0$, $\pi \geq 0$, and $\Sigma_{i=1}^n \pi_i = 1$. Assume $P[s(0) = s_i] = \pi_i$, $i = 1,\ldots,n$. Let $E_\pi(z(t))$ be the expectation of z at t conditional on these initial conditions. Similarly, let $E_\pi(y(t))$ be the expectation at $t = 0,1,2,\ldots$ of the discrete approximation y with the same initial conditions. Then*

$$E_\pi(z(t)) \leq E_\pi(y(t)), \quad t = 0,1,2,\ldots.$$

If $t \geq 1$, Q is irreducible, and $S = \mathrm{diag}\,(s_i)$ is not a scalar matrix, then the inequality is strict.

Proof. From [13] and Cohen [5, p. 346], it is immediate that for $t \geq 0$, $E_\pi(z(t)) = \pi^T e^{(Q+S)t}\mathbf{1}$, where $\mathbf{1}$ is an n-vector with all elements equal to 1. Direct calculation along the lines shown in [19] gives, for $t = 0,1,2,\ldots, E_\pi(y(t)) = \pi^T(e^Q e^S)^t \mathbf{1}$. The desired inequalities, weak and strict, follow from the corresponding cases of Theorem 3A. ∎

We conclude this section with one more application of Kingman's [16] theorem.

LEMMA 7C. *Let A and D_i, $i = 1,\ldots,k$, be $n \times n$ matrices, A nonnegative and not nilpotent, D_i all diagonal real. Let $D = (D_1 + \cdots + D_k)/k$. Then*

$$r(Ae^D)^k \leq r(Ae^{D_1}Ae^{D_2}\cdots Ae^{D_k}).$$

Proof. Let σ be the permutation of the numbers $1,\ldots,k$ defined by $\sigma(i) = i+1$ for $i = 1,\ldots,k-1$, and $\sigma(k) = 1$. Define

$$f(D_1,\ldots,D_k) = \log r(Ae^{D_1}Ae^{D_2}\cdots Ae^{D_k}).$$

Then, by (11), for $i = 0,1,2,\ldots,$ $f(D_1,\ldots,D_k) = f(D_{\sigma^i(1)},\ldots,D_{\sigma^i(k)})$. As remarked earlier, Kingman's theorem implies that f is convex in its arguments. Hence

$$f(D_1,\ldots,D_k) = k^{-1}\Sigma_{i=0}^{k-1} f(D_{\sigma^i(1)},\ldots,D_{\sigma^i(k)}) \geq f(D,\ldots,D). \quad ∎$$

LEMMA 7D. *Let A and D_i, $i = 1,\ldots,k$ be $n \times n$ matrices, A essentially nonnegative, D_i all diagonal real. Let $D = D_1 + \cdots + D_k$, a_i be nonnegative* *scalars,* $i=1,\ldots,k$, *and* $a=\Sigma_i a_i$. *Then*

$$r\left(e^{aA+D}\right)\leqslant r\left(e^{a_1A}e^{D_1}e^{a_2A}e^{D_2}\cdots e^{a_kA}e^{D_k}\right).$$

Proof. By continuity, it suffices to prove the desired inequality when $a_1,\ldots,a_k$ are rational numbers. Suppose $a_j=m_j/N$, $j=1,\ldots,k$, and let m be any positive integer. Let $C=e^{A/(Nm)}$. Then

$$\begin{aligned} e^{a_1A}e^{D_1}\cdots e^{a_kA}e^{D_k} &= C^{mm_1}e^{D_1}C^{mm_2}e^{D_2}\cdots C^{mm_k}e^{D_k} \\ &= \left[Ce^{0}\right]^{mm_1-1}Ce^{D_1}\left[Ce^{0}\right]^{mm_2-1}Ce^{D_2}\cdots\left[Ce^{0}\right]^{mm_k-1}Ce^{D_k}. \end{aligned}$$

Now, applying Lemma 7C with C here replacing A there, and with $M=k+m(m_1+\cdots+m_k)$ (here M is an integer), we have

$$\begin{aligned} r\left(C^{mm_1}e^{D_1}\cdots C^{mm_k}e^{D_k}\right) &\geqslant r\left(\left[Ce^{D/M}\right]^M\right) \\ &= r\left(\left[e^{A/(Nm)}e^{D/M}\right]^M\right). \end{aligned}$$

Let $m\to\infty$. Then $M\to\infty$ and

$$\left[e^{A/(Nm)}e^{D/M}\right]^M\to e^{A(m_1+\cdots+m_k)/N+D}=e^{aA+D},$$

so $r([e^{A/(Nm)}e^{D/M}]^M)\to r(e^{aA+D})$. Since $C=e^{A/(Nm)}$, $C^{mm_j}=e^{Am_j/N}$, which is independent of m, so the other side of the inequality does not change as $m\to\infty$. ■

THEOREM 4 (The mixing inequality). *Let A be an essentially nonnegative $n\times n$ matrix and B a diagonal real $n\times n$ matrix. For $k\geqslant 1$, $a_i\geqslant 0$, $b_i\geqslant 0$, $i=1,\ldots,k$, $a=\Sigma_i a_i$, $b=\Sigma_i b_i$,*

$$r\left(e^{aA+bB}\right)\leqslant r\left(e^{a_1A}e^{b_1B}\cdots e^{a_kA}e^{b_kB}\right)\leqslant r\left(e^{aA}e^{bB}\right).$$

Proof. The left inequality follows from Lemma 7D with $D_j=b_jB$, $j=1,\ldots,k$. The right inequality restates Theorem 3. ■

4. LOG-CONVEXITY OF SPECTRAL FUNCTIONS

A spectral function φ is defined to be *homogeneous* if, for any $n \times n$ complex matrix A and any positive integer m, $\varphi(A^m) = [\varphi(A)]^m$. Since $|\lambda_i(A^m)| = |\lambda_i(A)|^m$, $\prod_{i=1}^{k}|\lambda_i(A)|$ is a homogeneous spectral function for $k = 1, \ldots, n$.

THEOREM 5. *If A and B are $n \times n$ complex matrices and φ is a homogeneous spectral function, then* (9) *implies* (1), *and strict convexity in* (9) *implies strict inequality in* (1).

The proof depends on Lemmas 8 and 9.

LEMMA 8. *The real-valued function $f(t)$, $t > 0$, is convex (respectively, strictly convex) in t if and only if $tf(1/t)$, $t > 0$, is a convex (respectively, strictly convex) function of t.*

This lemma generalizes and provides a converse to Exercise 7 of [28, p. 77].

Proof. Suppose $f(t)$, $t > 0$, is a convex function of t. Then for $0 < a < 1$, $x > 0$, $y > 0$, and $z = ax + (1-a)y$,

$$\frac{ax}{z} f\left(\frac{1}{x}\right) + \frac{(1-a)y}{z} f\left(\frac{1}{y}\right) \geqslant f\left(\frac{a}{z} + \frac{1-a}{z}\right) = f\left(\frac{1}{z}\right). \tag{46}$$

Hence, multiplying both sides of (46) by $z > 0$,

$$axf(1/x) + (1-a)yf(1/y) \geqslant zf(1/z),$$

and so $h(t) = tf(1/t)$, $t > 0$, is a convex function of t. If f is strictly convex, the inequality is strict and so h is strictly convex. The converse statements follow since $th(1/t) = f(t)$. ■

LEMMA 9. *Let $f: [0, \infty] \to [-\infty, +\infty]$ satisfy*

$$\lim_{t \to \infty} f(t) \equiv f(\infty) < \infty, \tag{47}$$

where $f(\infty)$ may be finite in magnitude or equal to $-\infty$. Let $0 \leqslant t_1 < t_2 < \infty$. If f is convex, then $f(t_1) \geqslant f(t_2)$. If f is strictly convex, then $f(t_1) > f(t_2)$. In *both cases* $f(\infty)=\inf_{t\geqslant 0} f(t)$. *These assertions remain true if* $\lim_{t\to\infty} f(t)$ *is replaced by* $\lim_{m\to\infty} f(m)$ *for integral* m.

Proof. Assume f is convex and $0\leqslant t_1<t_2<t<\infty$. Then

$$f(t_2)\leqslant \frac{t-t_2}{t-t_1}f(t_1)+\frac{t_2-t_1}{t-t_1}f(t).$$

Letting $t\to\infty$, we deduce that $f(t_2)\leqslant f(t_1)$. Thus $f(t)$ is a nonincreasing function for t in $[0,\infty)$, and by (47) for t in $[0,\infty]$.

Suppose now that $0\leqslant t_1<t_2<t<\infty$ as before, and $f(t_1)=f(t_2)$. Since f is convex, $f(t)\geqslant f(t_2)$, but since f is nonincreasing, $f(t)\leqslant f(t_2)$. Thus $f(t)=f(t_2)$, and so f is constant for all $t\geqslant t_2$. Thus if f is strictly convex, then $f(t_1)>f(t_2)$. ∎

Proof of Theorem 5. Since $F(t)=\log\varphi(e^{At}e^{Bt})$ is convex on $[0,\infty)$, Lemma 8 implies $tF(1/t)=\log[\varphi(e^{A/t}e^{B/t})]^t$ is convex; and if $F(t)$ is strictly convex, then so is $tF(1/t)$. But if $tF(1/t)$ is convex, so is $\exp[tF(1/t)]=[\varphi(e^{A/t}e^{B/t})]^t=f(t)$; and if $tF(1/t)$ is strictly convex, so is $f(t)$. (Since $F(t)$ is twice differentiable, prove this by taking second derivatives of $\exp[tF(1/t)]$.) When all elements of A and B are finite, $\varphi(e^{A+B})$ is finite. Since φ is homogeneous, $f(t)=\varphi[(e^{A/t}e^{B/t})^t]$ for integral t. By (17) and the continuity of φ,

$$\varphi(e^{A+B})=\lim_{t\to\infty}\varphi\left[(e^{A/t}e^{B/t})^t\right], \tag{48}$$

where t moves along the integers on the right.

By Lemma 9, with $t_1=1$, $t_2=\infty$, (1) follows, with strict inequality when $f(t)$ is strictly convex. ∎

Theorem 5 makes it desirable to find log-convex homogeneous spectral functions φ in order to prove inequalities like (1). Theorem 6 establishes a large class of log-convex spectral functions, some of which are homogeneous. Define $\mathcal{A}$ to be a commutative set of $n\times n$ matrices if and only if, for all A_1, A_2 in $\mathcal{A}$, $A_1A_2=A_2A_1$. Define a set $\mathcal{A}$ of $n\times n$ matrices to be convex if and only if, for $0\leqslant a\leqslant 1$ and A_1, A_2 in $\mathcal{A}$, $aA_1+(1-a)A_2$ is in $\mathcal{A}$.

THEOREM 6. *Let* $\mathcal{A}$ *and* $\mathcal{B}$ *be two commutative convex sets of Hermitian* $n\times n$ *matrices. With the ordering of eigenvalues given in* (26), *for* $k=$

$1,2,\dots,n$,

$$\log \prod_{i=1}^{k} \lambda_i(e^A e^B) \quad \textit{and} \quad \log \sum_{i=1}^{k} \lambda_i(e^A e^B)$$

are convex functions of A in $\mathcal{A}$ and B in $\mathcal{B}$ jointly.

Proof. In view of the continuity of the functions involved, it suffices to prove, for A_1, A_2 in $\mathcal{A}$ and B_1, B_2 in $\mathcal{B}$, that

$$\sum_{i=1}^{k} \lambda_i(e^{(A_1+A_2)/2}e^{(B_1+B_2)/2}) \leqslant \left[\sum_{i=1}^{k} \lambda_i(e^{A_1}e^{B_1})\right]^{1/2}\left[\sum_{i=1}^{k} \lambda_i(e^{A_2}e^{B_2})\right]^{1/2},$$

and similarly with Σ replaced by Π. Let $X = e^{(A_1+A_2)/2}e^{(B_1+B_2)/2}$, $X_j = e^{A_j B_j}$, $Y_j = e^{A_j/2}e^{B_j/2}$, $j = 1,2$. Then

$$\begin{aligned}\sum_{i=1}^{k} \lambda_i(X) &= \sum \lambda_i(Y_1 Y_2^*) \leqslant \sum \sigma_i(Y_1 Y_2^*)\\ &\leqslant \sum \sigma_i(Y_1)\sigma_i(Y_2)\\ &\leqslant \left[\sum \sigma_i^2(Y_1)\right]^{1/2}\left[\sum \sigma_i^2(Y_2)\right]^{1/2}\\ &= \left[\sum \lambda_i(Y_1^* Y_1)\right]^{1/2}\left[\sum \lambda_i(Y_2^* Y_2)\right]^{1/2}\\ &= \left[\sum \lambda_i(X_1)\right]^{1/2}\left[\sum \lambda_i(X_2)\right]^{1/2},\end{aligned}$$

where the first inequality is due to Weyl [30] (see Theorem 9.E.1.a in [20, p. 232]), the second inequality is due to Horn [14] (see Theorems 5.A.2.b and 9.H.1 in [20, pp. 117, 246]), and the third is the Cauchy-Schwarz inequality (see e.g. Theorem 16.D.1.e in [20, p. 459]). This proves the theorem for Σ.

Similarly,

$$\begin{aligned}\prod_{i=1}^{k} \lambda_i(X) &= \prod \lambda_i(Y_1 Y_2^*) \leqslant \prod \sigma_i(Y_1 Y_2^*)\\ &\leqslant \prod \sigma_i(Y_1)\sigma_i(Y_2^*)\\ &= \left[\prod \lambda_i(X_1)\right]^{1/2}\left[\prod \lambda_i(X_2)\right]^{1/2},\end{aligned}$$

where the first inequality is due to Weyl [30] (see Theorem 9.E.1 in [20, p.

231]) and the second inequality is due to Horn [14] (again see Theorem 9.H.1. in [20, p. 246]). ■

COROLLARY 9. *Let A and B be $n \times n$ Hermitian matrices, with the ordering of eigenvalues given in* (26). *Then, for $k = 1, 2, \ldots, n$,*

$$f_k(t,\tau) = \log \prod_{i=1}^{k} \lambda_i(e^{At}e^{B\tau}), \tag{49}$$

$$g_k(t,\tau) = \log \sum_{i=1}^{k} \lambda_i(e^{At}e^{B\tau}) \tag{50}$$

are convex functions of the finite real pair (t,τ). For $k = n$, $f_k(t,\tau)$ is linear in (t,τ). The eigenvalues on the right in (49) *and* (50) *are positive.*

Proof. For any fixed Hermitian $n \times n$ matrix A, $\mathcal{Q} = \{At: t \text{ is real}\}$ is a commutative convex set of Hermitian $n \times n$ matrices. Apply Theorem 6. To prove linearity for $k = n$ note that $f_n(t,\tau) = \log\det(e^{At}e^{B\tau}) = \log\det(e^{At}) + \log\det(e^{B\tau}) = t\operatorname{Tr}(A) + \tau\operatorname{Tr}(B)$ by Jacobi's identity. Finally to prove positivity of the eigenvalues, for $i = 1,\ldots,n$, $\lambda_i(e^{At}e^{B\tau}) = \lambda_i(e^{At/2}e^{B\tau}e^{At/2}) = \lambda_i([e^{B\tau/2}e^{At/2}]^*[e^{B\tau/2}e^{At/2}]) \geqslant 0$. If any one of these eigenvalues were 0, then we would have $0 = \det(e^{At}e^{B\tau}) = \det(e^{At})\det(e^{B\tau}) = \exp[t\operatorname{Tr}(A) + \tau\operatorname{Tr}(B)] > 0$, a contradiction. ■

Log-convex functions of one parameter are obtained by setting $t = \tau$ in (49) and (50). The functions $\exp f_k(\tau, t)$, being homogeneous and log-convex, satisfy the assumptions of Theorem 5.

We now draw some further consequences of Corollary 9.

LEMMA 10. *Let $g: [0,\infty) \to (-\infty,\infty)$ be a convex function such that $g(0) = 0$. Then for $t > 0$, $g(t)/t$ is a nondecreasing function of t.*

Proof. Let $0 < t_1 < t_2$. Then $(t_1/t_2)g(t_2) = (t_1/t_2)g(t_2) + (1 - t_1/t_2)g(0) \geqslant g((t_1/t_2)t_2 + (1 - t_1/t_2)0) = g(t_1)$. ■

COROLLARY 10. *Let A and B be Hermitian $n \times n$ matrices. Then for $t > 0$ and $k = 1,\ldots,n$, $[f_k(t,t)]^{1/t}$ and $[g_k(t,t)]^{1/t}$ defined by* (49) *and* (50) *are increasing functions of t. In particular $r(e^{At}e^{Bt})^{1/t}$ is an increasing function of $t > 0$.*

LEMMA 11. *Let A and B be $n \times n$ complex matrices and $f(t) = \log r(e^{At}e^{Bt})^{1/t}$, $t > 0$. Order eigenvalues by* (26). *Then the following limits exist and*

$$f(0) = \lim_{t \downarrow 0} f(t) = \operatorname{Re} \lambda_1(A+B). \tag{51}$$

$$f(\infty) = \lim_{t \to +\infty} f(t) \leqslant \operatorname{Re}[\lambda_1(A) + \lambda_1(B)]. \tag{52}$$

If A and B are Hermitian, the inequality in (52) *holds if and only if the eigenspaces U and V corresponding respectively to $\lambda_1(A)$ and $\lambda_1(B)$,*

$$U = \{x \colon Ax = \lambda_1(A)x\},$$
$$V = \{x \colon Bx = \lambda_1(B)x\},$$

are mutually orthogonal.

Proof. The exponential product formula (17) implies that

$$f(0) = \log r(e^{A+B}) = \log e^{\operatorname{Re}\lambda_1(A+B)} = \operatorname{Re} \lambda_1(A+B).$$

This proves (51).

For any matrix norm $\|\cdot\|$, it is well known (e.g. [15]) that $r(C) = \lim_{m \to \infty} \|C^m\|^{1/m}$ for any $n \times n$ matrix C. Hence, taking $t \to \infty$, $r(e^{At}e^{Bt})^{1/t} \leqslant \|e^{At}e^{Bt}\|^{1/t} \leqslant \|e^{At}\|^{1/t}\|e^{Bt}\|^{1/t} \to r(e^A)r(e^B) = e^{\operatorname{Re}\lambda_1(A) + \operatorname{Re}\lambda_1(B)}$. Thus $\limsup_{t \to \infty} f(t) \leqslant \operatorname{Re}[\lambda_1(A) + \lambda_1(B)]$.

To see that $\lim_{t \to \infty} f(t)$ exists, recall [10] that

$$e^{At} = \sum_{\mu_j \in \operatorname{sp}^*(A)} e^{\mu_j t} \sum_{k=0}^{d(\mu_j, A)-1} \frac{t^k Z_{jk}(A)}{k!}, \tag{53}$$

and similarly for e^{Bt}, where $\operatorname{sp}^*(A)$ is the set of *distinct* eigenvalues of A, $d(\mu_j, A)$ is the multiplicity of μ_j in the minimal polynomial of A, and $Z_{jk}(A)$ is the kth component of A on μ_j. So

$$e^{At}e^{Bt} = e^{(\mu+\nu)t}t^K C(t)$$

when $\mu = \lambda_1(A)$, $\nu = \lambda_1(B)$ in the ordering (26), $K = d(\mu, A) - 1 + d(\nu, A) - 1 \geqslant 0$, and $\lim_{t \to \infty} C(t) = C \neq 0$. Thus

$$\lim_{t \to \infty} r(e^{At}e^{Bt})e^{-\operatorname{Re}(\mu+\nu)t}t^{-K} = r(C) \geqslant 0,$$

and since $\lim_{t\to\infty} t^{K/t} = 1$,

$$\lim_{t\to\infty} f(t)e^{-\mathrm{Re}(\mu+\nu)} = \lim_{t\to\infty} r^{1/t}(C) = 0 \text{ or } 1,$$

so $\lim_{t\to\infty} f(t)$ exists. This establishes (52).

If A and B are Hermitian, let

$$Ax_i = \lambda_i(A)x_i, \qquad By_i = \lambda_i(B)y_i,$$

$$(x_i, x_j) = (y_i, y_j) = \delta_{ij}, \qquad i, j = 1, \ldots, n.$$

Again decomposing into components by (53)

$$e^{At}e^{Bt} = \sum_{i,j} e^{[\lambda_i(A)+\lambda_j(B)]t}(x_i^* y_j)x_i y_j^*. \tag{54}$$

The term with coefficient $e^{[\lambda_1(A)+\lambda_1(B)]t}$ will appear on the right in (54) if and only if $x_1^* y_1 \neq 0$, i.e. if and only if U and V are not mutually orthogonal. ■

COROLLARY 11. *Let A and B be Hermitian $n \times n$ matrices and $f(t) = \log r(e^{At}e^{Bt})^{1/t}$, $t > 0$. Then ordering eigenvalues by* (26), $f(0) = \lambda_1(A+B) \leqslant f(1) = \log r(e^A e^B) \leqslant f(\infty) \leqslant \lambda_1(A) + \lambda_1(B)$.

That $f(t)$ is nondecreasing parametrizes the classical inequality for Hermitian A and B:

$$\lambda_1(A+B) \leqslant \lambda_1(A) + \lambda_1(B).$$

As another application of Corollary 10, we give a different proof of a special case of Corollary 1, by means of a lemma of independent interest.

LEMMA 12. *Let A and B be Hermitian $n \times n$ matrices, and $0 \leqslant a_i \leqslant a'$ for $i = 1, \ldots, k$ a positive integer, where a' and a_i are real scalars. Let $a = \sum_{i=1}^k a_i$. Then*

$$\| e^{a_1 A}e^{a_1 B} \cdots e^{a_k A}e^{a_k B} \| \leqslant r(e^{2a'A}e^{2a'B})^{a/(2a')}. \tag{55}$$

Proof. Using the submultiplicative property of norms at the first inequality and Corollary 10 at the second, we have

$$\|e^{a_1A}e^{a_1B}\cdots e^{a_kA}e^{a_kB}\| \leqslant \|e^{a_1A}e^{a_1B}\|\cdots\|e^{a_kA}e^{a_kB}\|$$

$$= \prod_{i=1}^{k}\left[r\left(e^{2a_iA}e^{2a_iB}\right)^{1/(2a_i)}\right]^{a_i}$$

$$\leqslant \prod_{i=1}^{k}\left[r\left(e^{2a'A}e^{2a'B}\right)^{1/(2a')}\right]^{a_i} = r\left(e^{2a'A}e^{2a'B}\right)^{a/(2a')}.$$

∎

Corollary 12. *For positive integral k, and Hermitian $n\times n$ matrices A and B,*

$$r\left[\left(e^{A/k}e^{B/k}\right)^k\right] \leqslant r\left(e^{2A/k}e^{2B/k}\right)^{k/2} \leqslant r\left(e^Ae^B\right).$$

As $k\to\infty$, we have $r(e^{A+B})\leqslant r(e^Ae^B)$.

Corollary 13. *If A and B are $n\times n$ Hermitian matrices, then $\log\operatorname{Tr}(e^{At}e^{Bt})$ is convex in the real variable t.*

Corollary 13 is a special case of Corollary 9. Gert Roepstorff (personal communications) found several independent proofs of Corollary 13. We give two of his proofs.

First alternate proof of Corollary 13. For any $n\times n$ Hermitian matrix H and any n-vector v, $g(t)=(v,e^{Ht}v)$ is log-convex in t. To see this, compute $(d^2/dt^2)\log g(t) = [g''g-(g')^2]/g^2 = [(v,\,H^2e^{Ht}v)(v,\,e^{Ht}v)-(v,\,He^{Ht}v)^2]/g^2$. Define $u=e^{Ht/2}v$. Since $H=H^*$, we may write $g^2(d^2/dt^2)\log g(t)=(Hu,\,Hu)(u,u)-(u,\,Hu)^2\geqslant 0$ by Schwarz's inequality. Thus $g(t)$ is log-convex.

Now let the eigenvalues (*not* the diagonal elements) of the Hermitian matrix B be b_i, and let $Bv_i=b_iv_i$, $i=1,\ldots,n$. Then

$$\operatorname{Tr}(e^{At}e^{Bt}) = \sum_{i=1}^{n}\left(v_i,\,e^{(A+b_iI)t}v_i\right). \tag{56}$$

Each summand has the form of $g(t)$ and is therefore log-convex. The sum of log-convex functions of t is log-convex in t [16], so (56) is log-convex in t. ■

Second alternate proof of Corollary 13. Let a be any real number. Using (56) and the assumption that $A=A^*$, we compute

$$\frac{d^2}{dt^2}\left[e^{at}\operatorname{Tr}(e^{At}e^{Bt})\right]=\sum_{i=1}^{n}\left(v_i,\left[A+(b_i+a)I\right]^2 e^{[A+(b_i+a)I]t}v_i\right)$$
$$=\sum_{i=1}^{n}(u_i,u_i)\geqslant 0,$$

where $u_i=[A+(b_i+a)I]e^{[A+(b_i+a)I]t/2}v_i$. Thus for every real a, $e^{at}\operatorname{Tr}(e^{At}e^{Bt})$ is convex in t. A theorem of Montel [21, pp. 32–33] shows that this implies $\log\operatorname{Tr}(e^{At}e^{Bt})$ is convex in t. ■

From the argument to prove Corollary 4, it is evident that Corollaries 9 and 13 hold when A is a reversible intensity matrix and B is a real diagonal matrix. In particular, under these assumptions, $\log\operatorname{Tr}(e^{At}e^{Bt})$ is convex in t. This fact also follows immediately from the observations that log-convex functions are closed under addition and multiplication [16] and that $\log(e^{At})_{jj}$ is convex, $j=1,\dots,n$. Kingman [17, pp. 1–2] established that the diagonal elements of the transition-probability matrix of a reversible Markoff chain are log-convex without the restriction that n must be finite.

The argument of Corollary 4 can also be used in a converse sense to establish this proposition: If G and H are $n\times n$ Hermitian matrices, there exist $n\times n$ matrices A and B, A Hermitian and B diagonal real, such that, for any pair (t,τ) of real variables,

$$\operatorname{sp}(e^{Gt}e^{H\tau})=\operatorname{sp}(e^{At}e^{B\tau}).$$

Consequently, if φ is a spectral function, any property proved about $\varphi(e^{At}e^{B\tau})$ when A is Hermitian or quasi-Hermitian and B is diagonal real or complex is also true about $\varphi(e^{Gt}e^{H\tau})$ when G and H are Hermitian.

To prove the proposition, let B be a diagonal real $n\times n$ matrix with $B=\operatorname{diag}(\lambda_1(H),\dots,\lambda_n(H))$. There exists an $n\times n$ complex matrix U such that $UU^*=I$ and $H=UBU^*$. Thus, using (11),

$$\operatorname{sp}(e^{Gt}e^{H\tau})=\operatorname{sp}(e^{Gt}Ue^{B\tau}U^*)=\operatorname{sp}(U^*e^{Gt}Ue^{B\tau})=\operatorname{sp}(e^{At}e^{B\tau}),$$

where $A=U^*GU$ and $A=A^*$. ■

COROLLARY 14.

(a) *Let* A *be a* 2×2 *essentially nonnegative matrix*, B *a* 2×2 *real diagonal matrix, and* $\varphi = r$, *the spectral radius. Then* (9) *holds.*

(b) *If, in addition,* A *is irreducible and* B *is not a scalar matrix, the convexity in* (9) *is strict.*

Proof. It entails no loss of generality to assume $r(A)=\mathrm{Tr}(A+B)=0$, since real scalar matrices may be added to A and to B without affecting the convexity of $F(t)=\log r(e^{At}e^{Bt})$.

If either A is reducible or B is a scalar matrix, then an elementary calculation shows that $F(t)$ is directly proportional to t.

Now suppose A is irreducible, i.e., $a_{12}>0$, $a_{21}>0$. Let $D=\mathrm{diag}(1,(a_{21}/a_{12})^{1/2})$. Then $D^{-1}AD$ is symmetric, i.e., A is quasisymmetric. Corollary 9 applies to $F(t)=\log r(e^{D^{-1}ADt}e^{Bt})$, proving part (a).

To prove strict convexity in (9) when A is irreducible and B is not a scalar matrix requires, at present, an explicit calculation, which is long but elementary. The result is

$$F(t)=\operatorname{arccosh}[p_1\cosh(b_{11}t)+p_2\cosh(b_{22}t)],$$

where $p_1=a_{22}/(a_{11}+a_{22})>0$ and $p_2=a_{11}/(a_{11}+a_{22})>0$. By Lemma 14, proved below, $F''\geqslant 0$ on $(-\infty,\infty)$ and $F''>0$ if and only if $t\neq 0$. Thus strict convexity in (9) is proved. ∎

To complete the proof of Corollary 14, we need two lemmas.

LEMMA 13. *If* f *maps a real interval into* $(1,\infty)$, *and* f'' *exists on the interval, then* arccosh f *is convex if and only if* $f(f'')-(f')^2\geqslant f''/f$ *on the interval, and* arccosh f *is strictly convex if and only if the inequality is strict.* (*Always take the* positive *value of* arccosh f.)

Proof. If $g=\operatorname{arccosh} f$, then $g''=\{f[f-1/f]f''-f(f')^2\}/(f^2-1)^{3/2}$. Since $f>1$, we have $g''\geqslant 0$ if and only if $(f-1/f)f''-(f')^2\geqslant 0$ or $f(f'')-(f')^2\geqslant f''/f$. Also $g''>0$ if and only if the latter inequalities are strict. ∎

LEMMA 14. *If* $g(t)=\operatorname{arccosh}[p_1\cosh(s_1t)+p_2\cosh(s_2t)]$, *where* $p_1,p_2\geqslant 0$, $p_1+p_2=1$, s_1,s_2 *are real and* $t\in(-\infty,+\infty)$, *then* $g''\geqslant 0$. *Moreover,* $t\neq 0$, $p_1>0$, $p_2>0$, *and* $s_1\neq s_2$ *if and only if* $g''>0$.

Proof. Let $f(t)=p_1\cosh(s_1t)+p_2\cosh(s_2t)$. Then $g=\operatorname{arccosh} f$. Since $\cosh(t)=\cosh(-t)$, it suffices to consider $t\geqslant 0$. By Lemma 13, it suffices to show $ff''-(f')^2\geqslant f''/f$ under the assumptions in the first sentence of Lemma 14, with strict inequality if and only if the assumptions of the second sentence hold.

Now $f(0)=1$ and $f'(0)=0$. Let $h_1=ff''-(f')^2$ and $h_2=f''/f$. Then $h_1(0)=f''(0)=h_2(0)$. Now $h_1'=ff'''+f'f''-2f'f''=ff'''-f'f''$, while $h_2'=(ff'''-f'f'')/f^2=h_1'/f^2$. Since $f^2\geqslant 1$, $h_2'\leqslant h_1'$ for all t. From $h_1(0)=h_2(0)$, it is immediate that $h_1\geqslant h_2$ for all t. This establishes $g''\geqslant 0$.

Under the additional assumptions $t>0$, $p_1>0$, $p_2>0$, and $s_1\neq s_2$, it follows that $f>1$, so $h_2'<h_1'$ and hence $h_2<h_1$, so $g''>0$.

Conversely, if $t=0$, then $h_1(0)=h_2(0)$ implies $g''=0$. If $p_1=0$, then $g(t)=s_2t$, so $g''=0$; similarly if $p_2=0$. If $s_1=s_2$, again $g(t)=s_2t$ and $g''=0$. So the additional conditions are necessary and sufficient for $g''>0$. ■

Lemma 14 has an immediate generalization. Let $p_i\geqslant 0$, s_i real, $i=1,\dots,n$, and $\Sigma_ip_i=1$. If $g(t)=\operatorname{arccosh}[\Sigma_ip_i\cosh(s_it)]$ and $t\in(-\infty,+\infty)$, then $g''\geqslant 0$. Moreover, $t\neq 0$, $p_i>0$, $p_j>0$, and $s_i\neq s_j$ for some $i\neq j$ if and only if $g''>0$. The proof is identical to that of Lemma 14.

If the assumption in Corollary 14 that B is diagonal be weakened to allow B to be symmetric, then $F(t)=\log r(e^{At}e^{Bt})$ need no longer be convex in t. For example, if

$$A=\begin{pmatrix}-1 & 0.001\\ 1000 & -1\end{pmatrix},\qquad B=\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix},$$

then $r(A)=0$, $\operatorname{Tr}(A+B)=0$, and $F(1)=7.23$, $F(3)=11.52$, $F(5)=15.52$. Consequently $\frac{1}{2}[F(1)+F(5)]=11.38<11.52$, so F is not convex.

If $\varphi=$ trace, while A is essentially nonnegative and B is real diagonal, then (9) is not true for all 3×3 matrices.

To conclude this section, we describe conditions under which $\log r(e^{At}e^{B\tau})$, where A and B are Hermitian $n\times n$ matrices, is a *linear* function of the pair (t,τ) of real numbers. For the rest of this section, we write $f(t,\tau)=\log r(e^{At}e^{B\tau})$ and $r(t,\tau)=r(e^{At}e^{B\tau})$.

THEOREM 7. *Let A and B be $n\times n$ Hermitian matrices. Let z be a real number such that $0\leqslant z\leqslant 1$, and let (t_1,τ_1) and (t_2,τ_2) be two pairs of real numbers such that $(t_1,\tau_1)\neq(t_2,\tau_2)$. Then*

$$f(t_1+z(t_2-t_1),\tau_1+z(\tau_2-\tau_1))=f(t_1,\tau_1)+z[f(t_2,\tau_2)-f(t_1,\tau_1)]\qquad \textit{for all } z \quad (57)$$

if and only if at least one of the following three conditions holds:

$$\text{(I)} \qquad t_1 = t_2 = 0; \tag{58}$$

$$\text{(II)} \qquad \tau_1 = \tau_2 = 0; \tag{59}$$

or (III) *there exists a vector x and real scalars a and b such that*

$$Ax = ax \quad \text{and} \quad Bx = bx. \tag{60}$$

When (60) *holds*,

$$f(t,\tau) = at + b\tau, \qquad \text{for} \quad (t,\tau) = (1-z)(t_1,\tau_1) + z(t_2,\tau_2). \tag{61}$$

The proof of Theorem 7 depends on the following lemma.

LEMMA 15. *Let P and Q be $n \times n$ complex matrices, x a complex n-vector, λ a real scalar. If*

$$PQx = \lambda x, \tag{62}$$

$$0 < \lambda = \|PQ\| = \|P\| \cdot \|Q\|, \tag{63}$$

$$\|x\| = 1, \tag{64}$$

then

$$Q^*P^*x = \lambda x, \tag{65}$$

$$P^*P(Qx) = \|P\|^2 Qx, \tag{66}$$

$$Q^*Qx = \|Q\|^2 x. \tag{67}$$

Proof. By (64), $\lambda^2 = (\lambda x, \lambda x)$, and by (62), $(\lambda x, \lambda x) = (PQx, PQx)$; hence

$$\lambda^2 = (PQx, PQx) \tag{68}$$

$$= (P^*PQx, Qx) \tag{69}$$

$$\leqslant \|P^*P\| (Qx, Qx) \tag{70}$$

$$\leqslant \|P\|^2 (Q^*Qx, x) \tag{71}$$

$$\leqslant \|P\|^2 \|Q\|^2. \tag{72}$$

But by (63),

$$\lambda^2 = \|P\|^2 \|Q\|^2 = \|PQ\|^2, \tag{73}$$

so all inequalities in (68) to (72) are equalities. By (69), $\lambda^2 = (Q^*P^*PQx, x)$, and by (63) $\lambda^2 = \|PQ\|^2 = r(Q^*P^*PQ)$. Thus $(Q^*P^*PQx, x) = r(Q^*P^*PQ)$. By the stationary property of the Rayleigh quotient, $r(Q^*P^*PQ)x = \lambda^2 x = (Q^*P^*PQx, x)$. Hence $\lambda^2 x = Q^*P^*PQx = Q^*P^*(\lambda x)$, or $\lambda x = Q^*P^*x$, which is (65). The equality between (69) and (70) implies (66), by the same argument. The equality between (71) and (72) implies (67), again by the same argument. ■

Proof of Theorem 7. Since $e^{At}e^{B\tau} = e^{-B\tau/2}[e^{+B\tau/2}e^{At}e^{+B\tau/2}]e^{+B\tau/2}$ and the matrix in brackets is Hermitian, all the eigenvalues of $e^{At}e^{B\tau}$ are positive. Also, being similar to a Hermitian matrix, $e^{At}e^{B\tau}$ is simple [18, p. 76].

Suppose (57) holds. The geometric multiplicity (the number of linearly independent eigenvectors) of $r(t, \tau)$ is an integer. Hence there is a point (t_0, τ_0) and a real $\delta > 0$ such that the neighborhood

$$N_\delta = \{(t, \tau) : |t - t_0| < \delta, |\tau - \tau_0| < \delta\} \tag{74}$$

intersects the line segment

$$(1 - z)(t_1, \tau_1) + z(t_2, \tau_2), \qquad 0 \leqslant z \leqslant 1, \tag{75}$$

and the geometric multiplicity of $r(t, \tau)$ is fixed at some value, say ν, in N_δ. By sliding (t_1, τ_1) and (t_2, τ_2) along the line segment (75) until they fall within N_δ, we shall arrange for $r(t, \tau)$ to have a fixed geometric multiplicity on the entire (contracted) segment (75).

Now let $P = \exp(B\tau_2/2)\exp(At_2/2)$ and $Q = \exp(At_1/2)\exp(B\tau_1/2)$. Then

$$r([t_1 + t_2]/2, [\tau_1 + \tau_2]/2) = r(PQ) \leqslant \|PQ\| \leqslant \|P\| \cdot \|Q\| = r^{1/2}(t_1, \tau_1) r^{1/2}(t_2, \tau_2). \tag{76}$$

With $z = \frac{1}{2}$, (57) implies that the inequalities in (76) are equalities. Hence we may apply Lemma 15. Let $x_1, \ldots, x_\nu$ be the ν linearly independent eigenvectors of PQ, as defined before (76), corresponding to $r(PQ)$. If we note that

$\|Q\|^2 = r(Q^*Q)$, (67) implies

$$e^{B\tau_1/2}e^{At_1}e^{B\tau_1/2}x_i = r(t_1,\tau_1)x_i, \qquad i=1,\dots,\nu. \tag{77}$$

Similarly, (66) implies

$$e^{At_2/2}e^{B\tau_2}e^{At_2/2}(Qx_i) = r(t_2,\tau_2)(Qx_i), \qquad i=1,\dots,\nu. \tag{78}$$

Because $r(t_1,\tau_1)$ has a fixed multiplicity ν, $x_1,\dots,x_\nu$ are *all* of the eigenvectors up to scalar multiples of Q^*Q corresponding to $r(t_1,\tau_1) = \|Q\|^2$. So (78) and

$$PQx_i = r(PQ)x_i, \qquad i=1,\dots,\nu, \tag{79}$$

hold whenever (t_2,τ_2) is replaced by a point (t_3,τ_3) such that

$$(t_3,\tau_3) = (1-z)(t_1,\tau_1) + z(t_2,\tau_2), \qquad 0 \leqslant z \leqslant 1. \tag{80}$$

But (80) and (57) imply that

$$r(t_3,\tau_3) = e^{\alpha + z\beta}, \tag{81}$$

where $r(t_1,\tau_1) = e^{\alpha}$ and $r(t_2,\tau_2) = e^{\alpha+\beta}$. With (t_2,τ_2) replaced by (t_3,τ_3), (79) becomes

$$\exp\left(\frac{B}{2}[\tau_1 + z(\tau_2-\tau_1]\right)\exp\left(\frac{A}{2}[t_1+z(t_2-t_1)]\right)e^{At_1/2}e^{B\tau_1/2}x_i = e^{\alpha+\beta z/2}x_i,$$
$$i=1,\dots,\nu \tag{82}$$

and (78) becomes

$$e^{A[t_1+z(t_2-t_1)]/2}e^{B[\tau_1+z(\tau_2-\tau_1)]}e^{A[t_1+z(t_2-t_1)]/2}(Qx_i) = e^{\alpha+z\beta}(Qx_i),$$
$$i=1,\dots,\nu. \tag{83}$$

In (82), let $z = 2y$ and multiply on the left by $e^{B\tau_1/2}$:

$$e^{B[\tau_1+y(\tau_2-\tau_1)]}e^{A[t_1+y(t_2-t_1)]}\left(e^{B\tau_1/2}x_i\right) = e^{\alpha+y\beta}\left(e^{B\tau_1/2}x_i\right),$$
$$i=1,\dots,\nu, \quad 0 \leqslant y \leqslant \tfrac{1}{2}. \tag{84}$$

Multiply (83) on the left by $e^{-A[t_1+z(t_2-t_1)]/2}$ to get

$$e^{B[\tau_1+z(\tau_2-\tau_1)]}e^{A[t_1+z(t_2-t_1)]}\left[e^{-Az(t_2-t_1)/2}e^{B\tau_1/2}x_i\right]$$
$$=e^{\alpha+z\beta}\left[e^{-Az(t_2-t_1)/2}e^{B\tau_1/2}x_i\right], \qquad i=1,\ldots,\nu. \tag{85}$$

Since all quantities in (84) are analytic functions of y, $e^{\alpha+y\beta}$ is an eigenvalue of

$$M=e^{B[\tau_1+y(\tau_2-\tau_1)]}e^{A[t_1+y(t_2-t_1)]}$$

with geometric multiplicity ν except possibly at a finite number of values of y, because we have assumed that ν is the geometric multiplicity of the eigenvalue $e^{\alpha+y\beta}$ in the neighborhood of $y=0$. But (84) asserts that the subspace U spanned by the vectors $e^{B\tau_1/2}x_i$, $i=1,\ldots,\nu$, is the eigenspace of M corresponding to $e^{\alpha+y\beta}$. Note that U does not depend on y. Therefore (85) implies

$$e^{-Az(t_2-t_1)/2}U=U. \tag{86}$$

Because $(t_1,\tau_1)\neq(t_2,\tau_2)$, it entails no loss of generality to assume $t_1\neq t_2$, since if $t_1=t_2$ we can exchange A and B and argue similarly. So with $t_1\neq t_2$, (86) implies

$$AU\subseteq U. \tag{87}$$

As U is spanned by all the eigenvectors of M corresponding to $e^{\alpha+y\beta}$ we have, using (86),

$$U=MU=e^{B[\tau_1+y(\tau_2-\tau_1)]}U.$$

Hence

$$BU\subseteq U$$

unless $\tau_1=\tau_2-\tau_1=0$, which is just (59). Assuming otherwise, we consider the action of M in the subspace U. Within U,

$$M=e^{\alpha+y\beta}I=e^{(\alpha+y\beta)I}.$$

So $\exp\{-B[\tau_1+y(\tau_2-\tau_1)]+(\alpha+y\beta)I\}$ is the inverse of $\exp\{A[t_1+y(t_2-t_1)]\}$, within the subspace U. Since this is true for all y such that $0\leqslant y\leqslant\frac{1}{2}$, A

and B must have the following properties in their action within U:

$$A(t_2 - t_1) + B(\tau_2 - \tau_1) = \beta I,$$

$$At_1 + B\tau_1 = \alpha I.$$

Recall that $t_2 \neq t_1$. If $\tau_2 \neq \tau_1$, then any eigenvector of B in U is an eigenvector of A. If $\tau_2 = \tau_1$, then any vector in U is an eigenvector of A. Thus all eigenvectors of B in U are eigenvectors of A.

Thus, assuming (57), if (58) and (59) are both false, then (60) holds and (61) follows.

Conversely, (58) or (59) imply (57) immediately. If (60) holds, $e^{At}e^{B\tau}x = e^{At}e^{b\tau}x = e^{at+b\tau}x$, so (57) holds. ∎

5. CONJECTURES AND OPEN PROBLEMS

We conclude with some conjectures, open problems, and speculations.

Conjecture 1 arose from a search for a proof of Theorem 2 that used Theorem 5.

CONJECTURE 1. Let A be an $n \times n$ essentially nonnegative matrix, B an $n \times n$ real diagonal matrix. Then $F(t) = \log r(e^{At}e^{Bt})$ is convex in the real variable t. If, in addition, A is irreducible and B is not a scalar matrix, then $F(t)$ is strictly convex in t.

Corollary 14 is the 2×2 case of this conjecture. Theorem 6 includes the analogous result (without specifying the conditions of strict convexity) for Hermitian matrices A and B. If Conjecture 1 were true, then Theorem 5 would provide another path to the conclusions of Theorem 2.

The parallel between Theorem 6's assertion about $f_1(t,t)$ and Conjecture 1, and the parallel between Corollary 1 for Hermitian matrices and Theorem 2 for nonnegative matrices, may be viewed as further instances of what Schneider [24, pp. 209–210] calls the Taussky unification problem. This problem, due to Taussky, is to find unified treatments of similar theorems for positive matrices and positive definite symmetric matrices. Informally, it appears to us that if $A \geqslant 0$ is an $n \times n$ matrix, $r(A)$ often has properties that would be expected if A were Hermitian, while the rest of $\mathrm{sp}(A)$ need not behave like the spectrum of a Hermitian matrix.

The next conjecture would provide sufficient conditions for strict inequality in Theorem 3.

CONJECTURE 2. Let $A_1,\dots,A_k$ be nonnegative irreducible $n\times n$ matrices with positive diagonal elements, for some positive integer k. Let $D_1,\dots,D_k$ be real diagonal $n\times n$ matrices with zero trace. Then

$$f(D_1,\dots,D_k)=\log r\left(A_1e^{D_1}\cdots A_ke^{D_k}\right)$$

is a strictly convex function of $(D_1,\dots,D_k)$.

Conjecture 3 generalizes Lemma 12.

CONJECTURE 3. Let A and B be $n\times n$ Hermitian matrices and $a_i\geqslant 0$, $b_i\geqslant 0$, $i=1,\dots,n$. Let $a=\Sigma_i a_i$, $b=\Sigma_i b_i$. Then

$$\|e^{a_1A}e^{b_1B}\cdots e^{a_kA}e^{b_kB}\|\leqslant r\left(e^{2aA}e^{2bB}\right)^{1/2}, \tag{88}$$

$$r\left(e^{a_1A}e^{b_1B}\cdots e^{a_kA}e^{b_kB}\right)\leqslant r\left(e^{aA}e^{bB}\right). \tag{89}$$

Lemma 12 verifies (88) in case $b_i=ca_i$, $i=1,\dots,k$, for some nonnegative scalar c, since B in Lemma 12 could then be replaced by cB. Similarly, (89) holds when $b_i=ca_i$, $i=1,\dots,k$, and $a'=\max_i a_i\leqslant a/2$. For $b_i=ca_i$ plus $r(X)\leqslant\|X\|$ for any $n\times n$ matrix X implies

$$r\left(e^{a_1A}e^{b_1B}\cdots e^{a_kA}e^{b_kB}\right)\leqslant\|e^{a_1A}e^{a_1(cB)}\cdots e^{a_kA}e^{a_k(cB)}\|.$$

Lemma 12 gives $\|e^{a_1A}e^{a_1(cB)}\cdots e^{a_kA}e^{a_k(cB)}\|\leqslant r(e^{2a'A}e^{2a'cB})^{a/(2a')}$. Then $a'\leqslant a/2$ plus Corollary 10 implies $r(e^{2a'A}e^{2a'cB})^{a/(2a')}\leqslant r(e^{aA}e^{acB})=r(e^{aA}e^{bB})$ as asserted in (89).

Finally there may be a probabilistic proof of Theorem 2 when A is an intensity matrix. (Both our proof and Varadhan's are analytical.) Interpret $\log r(e^Ae^B)$ and $\log r(e^{A+B})$ as the asymptotic growth rates of random evolutions in discrete and continuous time, respectively. The discrete-time random evolution can change states only at integer times. In the continuous-time process, the duration of a single visit to any one state is exponentially distributed. A majorization argument [20] applied to the sample paths of the discrete-time and continuous-time processes might yield the desired inequality.

A remark of Wilde [33] provides courage to attack these conjectures and problems: "Even things that are true can be proved."

We thank E. Deutsch, J. F. C. Kingman, M. B. Ruskai, L. Tarter, and C. J. Thompson for helpful comments, C. J. Stone for suggesting that we compare $E_\pi(y(t))$ *and* $E_\pi(z(t))$ *for finite t, Gert Roepstorff for permitting us to quote his proofs of Corollary 13, S. R. S. Varadhan for permitting us to quote his proof of a special case of Theorem 2, and Jean Hernandez for typing all drafts of the manuscript. The Department of Mathematics, University of Wisconsin at Madison; the Systems and Decision Sciences Area, International Institute for Applied Systems Analysis, Laxenburg, Austria; the Center for Advanced Study in the Behavioral Sciences, Stanford, California, and Mr. and Mrs. William T. Golden, Olivebridge, New York, kindly provided hospitality for J.E.C. during this work. During the completion of this work, J.E.C. was a John Simon Guggenheim Memorial Foundation Fellow, with support at the Center for Advanced Study in the Behavioral Sciences from Exxon Education Foundation and National Institute of Mental Health grant 5T32MH14581-06. The U.S. National Science Foundation supported this work through grant DEB 80-11026 to J.E.C. and grant MCS 79-02578 to T.K. The U.S.–Israel Binational Science Foundation partially supported this work through grant BSF 2242/80 to S.F.*

REFERENCES

1 R. Bellman, *Introduction to Matrix Analysis*, 2nd ed., McGraw-Hill, New York, 1970.
2 R. M. Blumenthal and R. K. Getoor, *Markov Processes and Potential Theory*, Academic, New York, 1968.
3 M. G. Bulmer, Multiple niche polymorphism, *Amer. Naturalist* 106:254–257 (1972).
4 S. T. Butler and M. H. Friedman, Partition function for a system of interacting Bose-Einstein particles, *Phys. Rev. (2)* 98:287–293 (1955).
5 J. E. Cohen, Random evolutions and the spectral radius of a nonnegative matrix, *Math. Proc. Cambridge Philos. Soc.* 86:345–350 (1979).
6 J. E. Cohen, Random evolutions in discrete and continuous time, *Stochastic Process. Appl.* 9:245–251 (1979).
7 E. B. Davies, *One-Parameter Semigroups*, Academic, New York, 1980.
8 M. D. Donsker and S. R. S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time, I, *Comm. Pure Appl. Math.* 28:1–47 (1975).
9 S. Friedland and S. Karlin, Some inequalities for the spectral radius of non-negative matrices and applications, *Duke Math. J.* 42:459–490 (1975).
10 F. R. Gantmacher, *Theory of Matrices*, 2 vols., Chelsea, New York, 1960.

11 S. Golden, Statistical theory of many-electron systems; discrete bases of representation, *Phys. Rev.* 107:1283–1290 (1957).

12 S. Golden, Lower bounds for the Helmholtz function, *Phys. Rev.* 137:B1127–B1128 (1965).

13 R. Griego and R. Hersh, Theory of random evolutions with applications to partial differential equations, *Trans. Amer. Math. Soc.* 156:405–418 (1971).

14 A. Horn, On the singular values of a product of completely continuous operators, *Proc. Nat. Acad. Sci. U.S.A.* 36:374–375 (1950).

15 T. Kato, *Perturbation Theory for Linear Operators*, 2nd ed., Springer, New York, 1976.

16 J. F. C. Kingman, A convexity property of positive matrices, *Quart. J. Math. Oxford Ser. (2)* 12:283–284 (1961).

17 J. F. C. Kingman, Markov transition probabilities, II. Completely monotonic functions, *Z. Wahrsch. Verw. Gebiete* 9:1–9 (1967).

18 P. Lancaster, *Theory of Matrices*, Academic, New York, 1969; reprinted by the author, Calgary, 1977.

19 H. Le Bras, Populations stables aléatoires, *Population (Paris)* 29:435–464 (1974).

20 A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic, New York, 1979.

21 P. Montel, Sur les fonctions convexes et les fonctions sousharmoniques, *J. Math. Pures Appl.* (9) 7:29–60 (1928).

22 M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, Academic, New York, 1972.

23 A. W. Roberts and D. E. Varberg, *Convex Functions*, Academic, New York, 1973.

24 H. Schneider, Olga Taussky-Todd's influence on matrix theory and matrix theorists, *Linear and Multilinear Algebra* 5:197–224 (1977).

25 E. Seneta, *Non-negative Matrices, an Introduction to Theory and Applications*, George Allen and Unwin, London, 1973.

26 C. J. Thompson, Inequality with applications in statistical mechanics, *J. Math. Phys.* 6:1812–1813 (1965).

27 C. J. Thompson, Inequalities and partial orders on matrix spaces, *Indiana Univ. Math. J.* 21:469–480 (1971).

28 C. J. Thompson, *Mathematical Statistical Mechanics*, Macmillan, New York, 1972; reprint, Princeton U.P., Princeton, 1979.

29 H. F. Trotter, On the product of semi-groups of operators, *Proc. Amer. Math. Soc.* 10:545–551 (1959).

30 H. Weyl, Inequalities between the two kinds of eigenvalues of a linear transformation, *Proc. Nat. Acad. Sci. U.S.A.* 35:408–411 (1949).

31 P. Whittle, Reversibility and acyclicity, in *Perspectives in Probability and Statistics: Papers in Honour of M. S. Bartlett* (J. Gani, Ed.), Academic, London, 1975, pp. 217–224.

32 R. M. Wilcox, Exponential operators and parameter differentiation in quantum physics, *J. Math. Phys.* 8:962–982 (1967).

33 O. Wilde, *The Picture of Dorian Gray*, London, 1890.

Received 1 December 1981