Reprinted from JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS Vol. 183, No. 1, April 1, 1994


# Iterated Exponentiation, Matrix–Matrix Exponentiation, and Entropy

IGNACIO BARRADAS AND JOEL E. COHEN

*Centro de Investigación en Matemáticas, A.C., Apartado Postal 402, 36000, Guanajuato, Guanajuato, México, and Rockefeller University, 1230 York Avenue, New York, New York 10021-6399*

*Submitted by William F. Ames*

Received March 20, 1992

On $[0, \infty)$, $\log(x^{x^{x^x}})$ is strictly convex. Matrix–matrix exponentiation $A^B$ is defined when $A$ is normal and nonsingular. Von Neumann's quantum-mechanical entropy $S(A)$ of a density matrix $A$ can be written $S(A) = -\log[\det(A^A)]$. Using convexity, an obvious generalization of $S(A)$, namely $-\log[\det(A^{A^{A^A}})]$, is shown to satisfy the same monotonicity inequality as $S(A)$. Matrix–matrix exponentiation is used to generalize several results about iterated exponentiation of scalars.


## 1. INTRODUCTION

A positive-valued function $f$ on some real interval is called log-convex if $\log f$ is convex; $f$ is called log-concave if $\log f$ is concave. For example, on the positive half-line, $f(x) = x$ is log-concave and $f(x) = x^x$ is log-convex, but $f(x) = x^{x^x}$ is neither log-concave nor log-convex on $(0, 1)$. The first major result of this paper (Theorem 2.1) is that on $[0, \infty)$, $\log(x^{x^{x^x}})$ is strictly convex (see Fig. 1). (Strict convexity of a real-valued function $f$ means that if $\alpha + \beta = 1$ and $\alpha > 0$, $\beta > 0$, then $f(\alpha a + \beta b) < \alpha f(a) + \beta f(b)$ for all $a$, $b$, $a \neq b$, in the domain of $f$. Strict concavity means the inequality is reversed.) This utterly elementary result does not seem to have been noticed before, yet is unexpectedly painful to prove. The proof given here is by brute force; a better proof would be welcome. Figure 1 and additional numerical studies suggest that the sixth iterated exponential, like the second and fourth, is log-convex on $[0, 1]$ but that the eighth iterated exponential is not.

Matrix–matrix exponentiation $A^B$ is defined for certain matrices $A$ and $B$ (Definition 3.1), and properties of $A^B$ analogous to those of scalar–scalar exponentiation $a^b$ are described. Von Neumann's quantum-mechanical entropy can be written using matrix–matrix exponentiation in a form



0022-247X/94 $6.00

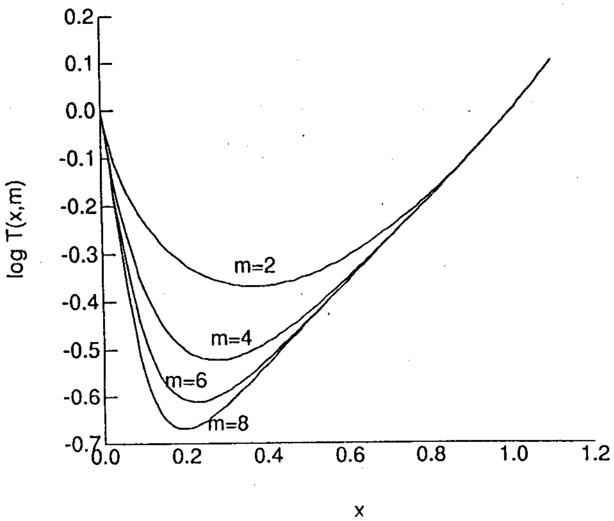


FIG. 1. Log $T(x, m)$ for $m=2, 4, 6, 8$, where $T(x, m)$ is the $m$-fold iterated exponential of $x$; e.g., $T(x, 2)=x^x$.

that suggests a natural generalization. The convexity of $\log(x^{x^{x^x}})$ makes it possible to prove a monotonicity inequality for one of these generalizations (Theorem 4.1) analogous to a well-known monotonicity inequality for the usual entropy. Finally, matrix–matrix exponentiation is used to generalize several results about iterated exponentiation of scalars.

## 2. The Logarithm of the Fourth Hyperpower Is Convex

THEOREM 2.1. *On* $[0, \infty)$, $\log(x^{x^{x^x}})$ *is strictly convex.*

*Proof.* Because $\log(x)$ appears many times in the formulas that follow, we introduce

$$L \equiv \log(x).$$

Thus, for example, $L^2=[\log(x)]^2$, $dL/dx=1/x$, and $x=e^L$.

Henceforth assume $x \in (0, \infty)$. Let $g(x) \equiv x^x$, $f(x) \equiv x^{x^{g(x)}}$. Then $d^2 \log f(x)/dx^2>0$ if and only if

$$f(x)f''(x)-(f'(x))^2>0, \tag{2.1.1}$$

which is the inequality we aim to prove. By elementary calculus

$$f'(x)=f(x)\, e^{Lg(x)}\left[\frac{1}{x}+L\left(\frac{g(x)}{x}+Lg'(x)\right)\right],$$



$$\begin{aligned}
f''(x)=f(x)\, e^{Lg(x)}\Bigg[&-\frac{1}{x^2}+\frac{g(x)}{x^2}+\frac{Lg'(x)}{x}\\
&+L\left(\frac{xg'(x)-g(x)}{x^2}+\frac{g'(x)}{x}+Lg''(x)\right)\\
&+\left(\frac{1}{x}+L\left[\frac{g(x)}{x}+Lg'(x)\right]\right)\\
&\times\left\{e^{Lg(x)}\left(\frac{1}{x}+L\left(\frac{g(x)}{x}+Lg'(x)\right)\right)+\frac{g(x)}{x}+Lg'(x)\right\}\Bigg],
\end{aligned}$$

$$g'(x)=(1+L)\,g(x),$$

$$g''(x)=\left[(1+L)^2+\frac{1}{x}\right]g(x).$$

Substituting these expressions into (2.1.1) gives

$$\begin{aligned}
&-\frac{1}{x^2}+g(x)\left[\frac{2}{x^2}-\frac{L}{x^2}+\frac{4(1+L)\,L}{x}+(1+L)^2\,L^2+\frac{L^2}{x}\right]\\
&\quad+g^2(x)\left[\frac{L}{x^2}+\frac{2(1+L)\,L^2}{x}+(1+L)^2\,L^3\right]>0.
\end{aligned} \tag{2.1.2}$$

If we multiply (2.1.2) by $x^2$ and move the first term to the other side of the inequality, we obtain

$$\begin{aligned}
h(x)\equiv g(x)\Bigg[&-L(1-g(x))+2+xL^3\left\{\frac{4}{L^2}+2x+\frac{x+5}{L}+xL\right.\\
&\left.+g(x)\left(2+x+xL^2+\frac{2}{L}+2xL\right)\right\}\Bigg]>1.
\end{aligned} \tag{2.1.3}$$

Now consider $x \in (0, 1)$. Here $0<g(x)<1$, $g'(x)<0$ on $(0, 1/e)$, and $g'(x)>0$ on $(1/e, 1)$. Also, for any natural numbers $m$ and $n$, $x^m L^n$ is a negative function of $x$ for odd $n$, is a positive function of $x$ for even $n$, and in either case has a unique extreme point where $x=\exp(-n/m)$.

Thus, for example,

$$\begin{aligned}
\min_{x\in[a,b]}\{xL^3\}&=b\log^3 b &&\text{if } 0\leqslant a\leqslant b\leqslant e^{-3},\\
&=a\log^3 a &&\text{if } e^{-3}\leqslant a\leqslant b\leqslant 1,\\
\max_{x\in[a,b]}\{xL^2\}&=b\log^2 b &&\text{if } 0\leqslant a\leqslant b\leqslant e^{-2},\\
&=a\log^2 a &&\text{if } e^{-2}\leqslant a\leqslant b\leqslant 1.
\end{aligned} \tag{2.1.4}$$

For any interval $[a, b] \subset (0, 1/e]$ and for $x \in [a, b]$,

$$\begin{aligned}
h(x) &\geqslant \min_{x \in [a,b]} \{h(x)\} \\
&\geqslant g(b)\left[1 - g(a) + 2 + \min_{x \in [a,b]} \{xL^3\} \cdot \max_{x \in [a,b]} \left\{0, \frac{4}{L^2} + 2x + \frac{x+5}{L} + xL \right.\right. \\
&\quad \left.\left. + g(x)\left(2 + x + xL^2 + \frac{2}{L} + 2xL\right)\right\}\right] \\
&\geqslant g(b)\left[1 - g(a) + 2 + \min_{x \in [a,b]} \{xL^3\} \cdot \max_{x \in [a,b]} \left\{0, \frac{4}{\log^2 b} + 2b \right.\right. \\
&\quad + g(a)\left(2 + b + \max_{x \in [a,b]} \{xL^2\}\right) + \frac{a+5}{\log a} \\
&\quad \left.\left. + a \log a + g(b)\left(\frac{2}{\log a} + 2a \log a\right)\right\}\right] \\
&\equiv y(a, b). \qquad (2.1.5)
\end{aligned}$$

Because of (2.1.4), for suitable values of $a$ and $b$, $y(a, b)$ can be evaluated numerically using only elementary functions. Table I lists a finite number of intervals $[a, b]$ with union containing $[e^{-6.5}, e^{-1}]$ such that $y(a, b) > 1$. This proves (2.1.3) on $[e^{-6.5}, e^{-1}]$.

For any interval $[a, b] \subset [1/e, 1]$ and for $x \in [a, b]$,

$$\begin{aligned}
h(x) &\geqslant \min_{x \in [a,b]} \{h(x)\} \\
&\geqslant g(a)[-\log b(1 - g(b)) + 2 + \min\{0, \min_{x \in [a,b]} \{4xL + 2x^2L^3 \\
&\quad + g(x)(2xL^3 + x^2L^3 + x^2L^5)\} + \min_{x \in [a,b]} \{x^2L^2 + 5xL^2 + x^2L^4 \\
&\quad + g(x)(2xL^2 + 2x^2L^4)\}\}] \\
&\geqslant g(a)[-\log b(1 - g(b)) + 2 + \min\{0, 4a \log a + 2a^2 \log^3 a \\
&\quad + g(b)(2a \log^3 a + a^2 \log^3 a + a^2 \log^5 a) \\
&\quad + g(a)(2b \log^2 b + 2b^2 \log^4 b) \\
&\quad + b^2 \log^2 b + 5b \log^2 b + b^2 \log^4 b\}] \\
&\equiv z(a, b). \qquad (2.1.6)
\end{aligned}$$

Table II lists a finite number of intervals $[a, b]$ with union containing $[e^{-1}, 1]$ such that $z(a, b) > 1$. This proves (2.1.3) on $[e^{-1}, 1]$.



TABLE I

Evaluations of $y(a, b)$ Defined in (2.1.5)

| $\log a$ | $\log b$ | $y(a, b)$ |
|---|---|---|
| −6.5 | −5.5 | 1.2017 |
| −5.5 | −5 | 1.1629 |
| −5 | −4.5 | 1.0653 |
| −4.5 | −4 | 1.0154 |
| −4 | −3.5 | 1.0383 |
| −3.5 | −3 | 1.1440 |
| −3 | −2.5 | 1.2818 |
| −2.5 | −2 | 1.3840 |
| −2 | −1.5 | 1.2973 |
| −1.5 | −1.2 | 1.5002 |
| −1.2 | −1 | 1.5027 |

We now deal with $x \in (0, e^{-6.5}]$. From (2.1.3), multiplying through the curly brackets { } by $xL^3$ and then dropping all positive terms except the constant 2 gives

$$h(x) > g(x)[2 + 4xL + 2x^2L^3 + g(x)(2xL^3 + x^2L^3 + x^2L^5)]. \quad (2.1.7)$$

Since $g(x) < 1$ and each term in the sum $2xL^3 + x^2L^3 + x^2L^5$ is negative, we may continue (2.1.7) by replacing the inner $g(x)$ by 1, giving

$$h(x) > g(x)[2 + 4xL + 2xL^3 + 3x^2L^3 + x^2L^5] \equiv g(x)[2 + u(x)]. \quad (2.1.8)$$

Now for $x \in (0, e^{-6.5}]$, each term of $u(x)$ is negative and decreasing as $x$ increases. Moreover, $u(e^{-6.5}) = -0.8929 > -0.9$ and $g(e^{-6.5}) = 0.9902 > 0.99$, so

$$h(x) > 0.99[2 - 0.9] = 1.089 > 1.$$

For $x > 1$, $h(x) > 1$ by inspection because $g(x) > 1$ and $L > 0$. ∎

TABLE II

Evaluations of $z(a, b)$ Defined in (2.1.6)

| $a$ | $b$ | $z(a, b)$ |
|---|---|---|
| −1 | −0.8 | 1.3562 |
| −0.8 | −0.6 | 1.0725 |
| −0.6 | −0.5 | 1.1277 |
| −0.5 | −0.4 | 1.0516 |
| −0.4 | −0.3 | 1.0175 |
| −0.3 | −0.2 | 1.0487 |
| −0.2 | 0 | 1.1169 |

## 3. Matrix–Matrix Exponentiation

Here matrices, denoted by capital letters, are assumed to have complex elements and to be $n \times n$, $1 \leqslant n < \infty$. Complex scalars are denoted by lower-case letters.

In matrix–matrix exponentiation $A^B$, the base matrix $A$ is assumed to be normal (i.e., $AA^* = A^*A$) and nonsingular throughout, and to satisfy further restrictions as needed. It is well known [4, p. 100] that $A$ is normal if and only if $A = UDU^*$ where $D$ is a diagonal matrix $\text{diag}[\lambda_i(A)]$ with the eigenvalues $\lambda_i(A)$, $i = 1, ..., n$, of $A$ on the diagonal, and $U$ is some unitary matrix ($U^* = U^{-1}$).

For any matrix $X$, the matrix exponential $e^X$ can be defined by applying the exponential power series to the matrix argument $X$. If matrices $X$ and $Y$ satisfy $e^X = Y$, $X$ is defined to be a logarithm of $Y$. Since $e^X = e^{X + 2\pi i I}$, the matrix logarithm is not single-valued.

Let $\mathbf{N}$ be the set of normal nonsingular matrices. When $A \in \mathbf{N}$, $\log A$ may be defined in a unique way as follows. For any nonzero complex number $z = re^{i\theta}$, where $r > 0$ and $-\pi < \theta \leqslant \pi$, define $\text{Log}\, z = \log r + i\theta$. This principal value of the logarithm is unique. For $A \in \mathbf{N}$ with $A = UDU^*$, $U$ unitary, $D = \text{diag}[\lambda_i(A)]$ as above, define $\log A = U \,\text{diag}[\text{Log}\, \lambda_i(A)]\, U^*$ and $e^A = U \,\text{diag}[\exp \lambda_i(A)]\, U^*$. (For $A \in \mathbf{N}$, this definition of $e^A$ is equivalent to the power series definition of $e^A$, which is also valid for $A$ not in $\mathbf{N}$.) By this definition, $\log A$ is unique, $\log A \in \mathbf{N}$, and $e^{\log A} = A$.

Definition 3.1. If $A \in \mathbf{N}$, i.e., $A$ is normal and nonsingular, and $B$ is any (complex $n \times n$) matrix, define $A^B = e^{(\log A) B}$ and ${}^BA = e^{B(\log A)}$.

More general definitions are possible, but are not needed here.

The remainder of this section describes some basic properties of matrix–matrix exponentiation, first for normal nonsingular matrices and then for Hermitian positive definite matrices. These propositions are used to prove the results of the following sections.

Theorem 3.2. *If $A \in \mathbf{N}$ and $B$ is any matrix, $A^B$ and ${}^BA$ have the same set of eigenvalues.*

*Proof.* For any two matrices $X$ and $Y$, $XY$ and $YX$ have the same set of eigenvalues. In particular $(\log A) B$ has the same set of eigenvalues as $B(\log A)$. The spectral mapping theorem guarantees that the eigenvalues of $e^X$ are the exponentials of the eigenvalues of $X$ [6, p. 312, Theorem 6]. Therefore ${}^BA$ has the same set of eigenvalues as $A^B$. ∎

Lemma 3.3 [4, p. 52; 6, p. 420, Ex. 6]. *A set of normal matrices is a commuting set (i.e., every possible pair of matrices in it commutes) if and only if it is a simultaneously diagonalizable set, or equivalently if and only if all the matrices share a common set of orthonormal eigenvectors.*



Theorem 3.4. *If $A \in \mathbf{N}$, $B$ is normal and $AB = BA$, then* (a) $A^B \in \mathbf{N}$ *and* ${}^BA \in \mathbf{N}$, (b) $(\log A) B = B(\log A)$, *and* (c) $A^B = {}^BA$.

*Proof.* (a) If $U$ is the matrix given by the lemma, with columns that are the common set of orthonormal eigenvectors of $A$ and $B$, so that $AU = UD$ and $BU = UE$, where $D = \text{diag}[\lambda_i(A)]$ and $E = \text{diag}[\lambda_i(B)]$, then

$$(\log A) B = U(\log D) U^*UEU^* = U(\log D) EU^*, \qquad (3.4.1)$$

and $A^B = e^{(\log A)B} = UD^EU^*$ so $A^B$ is normal. Since, for any $Y$, $e^Y$ is nonsingular, with $Y = (\log A) B$ it follows that $A^B$ is nonsingular. Hence $A^B \in \mathbf{N}$. (b) Continuing (3.4.1) (since diagonal matrices like $\log D$ and $E$ commute),

$$U(\log D) EU^* = UEU^*U(\log D) U^* = B(\log A). \qquad (3.4.2)$$

(c) follows from (b).

Theorem 3.5 (Exponential product formulas). *If $A$ is normal and nonsingular and $B$ and $C$ are complex $n \times n$ matrices, then*

$$A^{B+C} = \lim_{k \uparrow \infty} (A^{B/k} A^{C/k})^k,$$

*where $k$ runs through the positive integers. If, in addition, $\{A, B, C\}$ is a commuting set of normal matrices, then $A^{B+C} = A^B A^C$ and $A^{B+C}$ is normal and nonsingular. Analogous formulas hold for ${}^{B+C}A$.*

*Proof.* Sophus Lie's exponential product formula asserts that for any $Y$ and $Z$,

$$e^{Y+Z} = \lim_{k \uparrow \infty} (e^{Y/k} e^{Z/k})^k.$$

Various sources for this important formula are reviewed by Cohen *et al.* [2, p. 60].

Taking $Y = (\log A) B$ and $Z = (\log A) C$ yields the first claim.

Now let $\{A, B, C\}$ be a commuting set in $N$. Using the common set of orthonormal eigenvectors guaranteed by Lemma 3.3, it is easy to see that $B + C$ is normal and that $A(B + C) = (B + C) A$. Theorem 3.4(a) implies that $A^{B+C} \in \mathbf{N}$. Again taking $Y = (\log A) B$ and $Z = (\log A) C$, it follows that $YZ = ZY$, hence $e^{Y+Z} = e^Y e^Z$, which reduces to $A^{B+C} = A^B A^C$. ∎

By definition, $A$ is Hermitian if $A=A^*$ and is positive definite if all of its eigenvalues $\lambda_i(A)$, $i=1, ..., n$, are positive reals. Let $\mathbf{H}$ be the set of Hermitian positive definite matrices. Then $\mathbf{H}\subset\mathbf{N}$. When $A\in\mathbf{H}$, $\log A$ is Hermitian but $\log A$ need not be positive definite. For example, $\log A$ is a negative eigenvalue if $\lambda_i(A)<1$ for some $i=1, ..., n$. For the remainder of this section and the next, the base matrix $A$ is assumed to be in $\mathbf{H}$, i.e., Hermitian and positive definite.

THEOREM 3.6. *If $A, B\in\mathbf{H}$, then all eigenvalues of $A^B$ and ${}^BA$ are positive.*

*Proof.* Since $\log A$ is Hermitian and $B$ is Hermitian positive definite, $(\log A)\,B$ has only real eigenvalues [6, p. 180, Ex. 13]. Since the eigenvalues of $e^X$ are $e^{\lambda_i(X)}$, $i=1, ..., n$ [6, p. 312, Theorem 6], the eigenvalues of $A^B$ are all positive real. So are those of ${}^BA$ by Theorem 3.2. ∎

THEOREM 3.7. *If $A, B\in\mathbf{H}$ and $AB=BA$, then $A^B\in\mathbf{H}$, ${}^BA\in\mathbf{H}$.*

The proof is like the proof of Theorem 3.4.

THEOREM 3.8. *If $A, B\in\mathbf{H}$ and $AB=BA$, then* $\log(A^B)=(\log A)\,B$, $\log({}^BA)=B(\log A)$.

*Proof.* By Theorem 3.7, $A^B\in\mathbf{H}$. Every matrix in $\mathbf{H}$ has a unique Hermitian logarithm [6, p. 313, Ex. 5]. Using the spectral decomposition of $A$ and $B$ given in the proof of Theorem 3.4 (where now $D$ and $E$ have positive reals on their diagonals), it is routine to check that $(\log A)\,B$ is Hermitian and $e^{(\log A)B}=A^B$. A similar argument applies to $\log({}^BA)$. ∎

Theorem 3.8 need not be true for all $A, B\in\mathbf{N}$, under the present definition of $\log Y$ for $Y\in\mathbf{N}$. Even in the case of complex scalars $a$ and $b$, using the principal value of the logarithm, it is not always true that $\mathrm{Log}(e^{(\mathrm{Log}\,a)b})=(\mathrm{Log}\,a)\,b$. For example, if $a=e^{2\pi}e^{i\theta}$ $(0<\theta\leqslant\pi)$, and $b=i$, then $\mathrm{Log}(e^{(\mathrm{Log}\,a)b})=-\theta$ but $(\mathrm{Log}\,a)\,b=-\theta+2\pi i$.

THEOREM 3.9. *Let $A, B\in\mathbf{H}$. If $AB=BA$, then $(\log A)\,B=B(\log A)$ and $A^B={}^BA$. Conversely, if $(\log A)\,B=B(\log A)$, then $AB=BA$ and $A^B={}^BA$.*

*Proof.* If $AB=BA$, then by Lemma 3.3, $A$ and $B$ share a common set of orthonormal eigenvectors, and therefore $\log A$ and $B$ also share a common set of orthonormal eigenvectors; so by Lemma 3.3 again, $(\log A)\,B=B(\log A)$, and therefore $A^B={}^BA$. Conversely, if $(\log A)\,B=B(\log A)$, apply Lemma 3.3 to the commuting set $\{\log A, B\}$ of normal matrices to infer that $\{A=e^{\log A}, B\}$ is also a commuting set of normal matrices, i.e., $AB=BA$. In light of Theorem 3.4(c), $A^B={}^BA$. ∎



## 4. Entropy in Quantum Mechanics

Matrix–matrix exponentiation appears to have a natural role in the description of nature. In quantum mechanics (von Neumann [11]), a *density matrix* $A$ is defined to be a Hermitian positive definite matrix $A$ with trace 1; i.e.,

$$\mathrm{tr}\,A\equiv\sum_{i=1}^{n}\lambda_i(A)=1.$$

For a density matrix $A$, the spectrum $\{\lambda_i(A), i=1, ..., n\}$ may be viewed as a probability density function on $\{1, ..., n\}$. The physical interpretation of the density matrix is that a quantum mechanical system or quantity may be in one of $n$ states, indexed by $i=1, 2, ..., n$ and the probability of observing the system described by $A$ in state $i$ is $\lambda_i(A)$. Von Neumann [11] showed that the *entropy* of a gas with density matrix $A$ is, except for a constant of proportionality,

$$S(A)=-\mathrm{tr}\,A\log A.$$

Jacobi's identity states that for any matrix $A$, $\det(e^A)=e^{\mathrm{tr}(A)}$, where det means determinant [e.g., 6, p. 346]. Using this identity with the definition of matrix–matrix exponentiation yields

$$\begin{aligned} e^{S(A)}&=e^{-\mathrm{tr}\,A\log A}=\det(e^{-A\log A})\\ &=[\det({}^AA)]^{-1}=[\det(A^A)]^{-1}, \end{aligned}$$

hence

$$S(A)=-\log[\det(A^A)]=\log\frac{1}{\det(A^A)}.$$

Let $T(A, m)$ denote the iterated exponential containing $m$ copies of $A$, i.e., $T(A, 1)=A$ and $T(A, m+1)=A^{T(A,m)}$, $m=1, 2, ...$ If von Neumann's entropy is written as $S(A)=-\log[\det T(A, 2)]$, it appears as the special case when $m=2$ of an infinite sequence of "entropies,"

$$S_m(A)=-\log[\det T(A, m)],\quad m=1, 2, ....$$

$S_m(A)$ is well defined for any Hermitian positive definite $A$, not only when $A$ is a density matrix. For any density matrix $A$, it is easy to prove that $\det A<\det T(A, m)<T(\det A, m)$ and hence that $-\log T(\det A, m)<S_m(A)<-\log\det A$. These bounds on $S_m(A)$ require computing only iterated exponentials of scalars, not iterated exponentials of matrices.

A change in the density matrix of a system from $A$ to $B$ results in a change in its entropy from $S(A)$ to $S(B)$, which corresponds to a change in the number of possible microscopic states of the system proportional to $e^{S(B)-S(A)}=\det(A^A)/\det(B^B)$. When entropy increases, i.e., $S(B)-S(A)>0$, the number of compatible microscopic states of the system also increases.

A well-known sufficient condition for an increase in the entropy of a system is that the probability distribution over the states be smoothed out or averaged [7]. More precisely, let $P$ be a doubly stochastic matrix (a matrix with nonnegative elements such that the sum of each row and the sum of each column equal 1). Let $\mathrm{vec}[\lambda_i(A)]$ be the vector whose elements are the eigenvalues of $A$, and let $A=U\,\mathrm{diag}[\lambda_i(A)]\,U^*$ be the spectral decomposition of a normal matrix $A$. If $A$ is a density matrix, then $B\equiv U\,\mathrm{diag}[P\,\mathrm{vec}[\lambda_i(A)]]\,U^*$ is also a density matrix and $S_2(B)\geqslant S_2(A)$. The inequality is strict if each element of $P$ is positive, and also under much weaker conditions [7].

THEOREM 4.1. *If $A$ and $B$ are Hermitian positive definite matrices such that $A=U\,\mathrm{diag}[\lambda_i(A)]\,U^*$ and $B=U\,\mathrm{diag}[P\,\mathrm{vec}[\lambda_i(A)]]\,U^*$, then $S_4(B)\geqslant S_4(A)$, with strict inequality if all elements of $P$ are positive and $A\neq B$.*

*Proof.* Suppose two vectors with positive elements $u$ and $v$ satisfy $u=Pv$, for some doubly stochastic $P$. Let $\varphi$ be a convex real-valued function on $[0,\infty)$. Then it is well-known [7] that

$$\sum_{i=1}^{n}\phi(u_i)\leqslant\sum_{i=1}^{n}\phi(v_i),$$

and the inequality is strict if $\phi$ is strictly convex and all elements of $P$ are positive and $u\neq v$. Theorem 2.1 shows that $\phi(x)=\log T(x,4)$ is strictly convex on $[0,\infty)$. Let $\alpha_i$ denote the eigenvalues of $A$, $\beta_i$ those of $B$; $\alpha=(\alpha_i)$, $\beta=(\beta_i)$. By hypothesis $\beta=P\alpha$. Therefore

$$\sum_{i=1}^{n}\log T(\beta_i,4)\leqslant\sum_{i=1}^{n}\log T(\alpha_i,4).$$

But

$$\sum_{i=1}^{n}\log T(\alpha_i,4)=\log\prod_{i=1}^{n}T(\alpha_i,4)=\log\det T(A,4) \text{ and similarly for } B.$$

Substituting into the last inequality and multiplying by $-1$ gives

$$-\log[\det T(B,4)]\geqslant-\log[\det T(A,4)]. \quad \blacksquare$$



## 5. ITERATED EXPONENTIATION OF HERMITIAN AND NORMAL NONSINGULAR MATRICES

Recall that the Loewner ordering [4, Sec. 7.7] is defined for Hermitian matrices $A$, $B$ by $B\geqslant A$ if and only if $B-A$ is nonnegative definite; this is equivalent to asserting that every eigenvalue of $B-A$ is a nonnegative real number. If $B-A$ is positive definite (i.e., in $\mathbf{H}$), then $B>A$ in the Loewner ordering. For real $a,b$ and Hermitian $A$, $aI\leqslant A\leqslant bI$ if and only if $a\leqslant\lambda_i(A)\leqslant b$ for $i=1,\ldots,n$.

THEOREM 5.1. *Let $A$ be a Hermitian nonnegative definite matrix. Then $\lim_{m\uparrow\infty}T(A,m)$ exists; i.e., $T(A,m)$ converges to a finite limit as $m\uparrow\infty$, if and only if $e^{-e}I\leqslant A\leqslant e^{1/e}I$, and in this case $e^{-1}I\leqslant\lim_{m\uparrow\infty}T(A,m)\leqslant eI$.*

This theorem is an immediate translation, via the spectral theorem and the results of Section 3, of a well-known theorem from 1778 of Euler (see [3] and [5]) on iterated exponentiation of scalars.

Theorem 5.1 guarantees that the limiting entropy

$$S_\infty(A)=\lim_{m\uparrow\infty}-\log[\det T(A,m)]$$

exists if $e^{-e}I\leqslant A\leqslant e^{1/e}I$.

Let $T(A_1;A_2;\ldots;A_m)$ denote the iterated exponential containing $A_1,\ldots,A_m$ from bottom to top, i.e., $T(A)=A$ and $T(A_1;A_2;\ldots;A_m)=A_1^{T(A_2;\ldots;A_m)}$.

THEOREM 5.2. *Let $A_1,A_2,\ldots\in\mathbf{N}$ be a commuting sequence of normal nonsingular matrices such that $|\log\lambda_i(A_m)|\leqslant e^{-1}$, for $i=1,\ldots,n$ and $m=1,2,\ldots$. Then*

$$M=\lim_{m\uparrow\infty}T(A_1;A_2;\ldots;A_m)$$

*exists and $|\log\lambda_i(M)|\leqslant 1$. If $A\in\mathbf{N}$ satisfies*

$$\lambda_i(A)=\exp(t_ie^{-t_i}),\qquad i=1,\ldots,n,$$

*where $|t_i|\leqslant\log 2$, $i=1,\ldots,n$, then $W=\lim_{m\uparrow\infty}T(A,m)$ exists and $\lambda_i(W)=\exp(t_i)$.*

This theorem is an immediate translation, via the spectral theorem and the results of Section 3, of theorems on scalar–scalar exponentiation due to Thron [10] and Shell [8].

THEOREM 5.3. *Let $A_1, A_2, \ldots, A_k$ be a commuting set of Hermitian positive definite matrices such that $A_i \geqslant eI$ for $i = 1, \ldots, k$. Pick any $i$ and $j$ so that $1 \leqslant i < j \leqslant k$. If $A_i \leqslant A_j$, then*

$$\begin{aligned} &T(\ldots; A_{i-1}; A_i; A_{i+1}; \ldots; A_{j-1}; A_j; A_{j+1}; \ldots) \\ &\quad \geqslant T(\ldots; A_{i-1}; A_j; A_{i+1}; \ldots; A_{j-1}; A_i; A_{j+1}; \ldots), \end{aligned}$$

*with strict inequality if $A_i < A_j$.*

In words, the "larger" tower (in the sense of the Loewner ordering) results from putting the larger of $A_i$ and $A_j$ higher in the tower. This theorem is an immediate translation, via the spectral theorem and the results of Section 3, of a theorem on scalar–scalar exponentiation due to Brunson [1].

## ACKNOWLEDGMENTS

We thank Barry W. Brunson, Arthur Knoebel, and Seth Lloyd for helpful comments. Parts of this work were done during J.E.C.'s visits to the Department of Mathematical Sciences, IBM T. J. Watson Research Center, Yorktown Heights, New York, and the Department of Mathematics, Beijing Normal University, China. J.E.C. is grateful for partial support from U.S. National Science Foundation Grants BSR 84-07461 and BSR 87-05047, and for the hospitality of Mr. and Mrs. William T. Golden during this work.

## REFERENCES

1. B. W. BRUNSON, The partial order of iterated exponentials, *Amer. Math. Monthly* **93** (1986), 779–786.
2. J. E. COHEN, S. FRIEDLAND, T. KATO, AND F. P. KELLY, Eigenvalue inequalities for products of matrix exponentials, *Linear Algebra Appl.* **45** (1982) 55–95.
3. J. M. DE VILLIERS AND P. N. ROBINSON, The interval of convergence and limiting functions of a hyperpower sequence, *Amer. Math. Monthly* **93** (1986), 13–23.
4. R. A. HORN AND C. R. JOHNSON, "Matrix Analysis," Cambridge Univ. Press, Cambridge, 1985.
5. R. A. KNOEBEL, Exponentials reiterated, *Amer. Math. Monthly* **88** (1981), 235–252.
6. P. LANCASTER AND M. TISMENETSKY, "Theory of Matrices with Applications," 2nd ed., Academic Press, New York, 1985.
7. A. W. MARSHALL AND I. OLKIN, "Inequalities: Theory of Majorization and Its Applications," Academic Press, New York, 1979.
8. D. L. SHELL, On the convergence of infinite exponentials, *Proc. Amer. Math. Soc.* **13** (1962), 678–681.
9. J. R. STEMBRIDGE, The Bruhat order and iterated exponentials, *J. Combin. Theory Ser. A* **50** (1988), 87–99.



10. W. J. THRON, Convergence of infinite exponentials with complex elements, *Proc. Amer. Math. Soc.* **8** (1957), 1040–1043.
11. J. VON NEUMANN, "Mathematische Grundlagen der Quantenmechanik," Springer, Berlin, 1932; English translation by R. T. BEYER, "Mathematical Foundations of Quantum Mechanics," Princeton Univ. Press, Princeton, 1955. Chapters V and VI reprinted as "Measurement and Reversibility" and "The Measuring Process," *in* "Quantum Theory and Measurement" (J. A. Wheeler and Wojciech H. Zurek, Eds.), Princeton Univ. Press, Princeton, 1983.

Printed by Catherine Press, Ltd., Tempelhof 41, B-8000 Brugge, Belgium