POSITIVE DEFINITE FUNCTIONS AND OPERATOR INEQUALITIES

RAJENDRA BHATIA AND K. R. PARTHASARATHY

Abstract

We construct several examples of positive definite functions, and use the positive definite matrices arising from them to derive several inequalities for norms of operators.

1. *Introduction*

This paper has two goals. The first is to present, and to advance, a technique for proving the positive definiteness of several classes of matrices related to important problems in operator theory. The second is to show how this positivity can be used to prove inequalities for norms of operators.

The matrices that we discuss include the *divided-difference matrices* or the *Loewner matrices*. These are matrices whose entries are defined as

$$a_{ij} = \frac{f(\lambda_i) - f(\lambda_j)}{\lambda_i - \lambda_j}, \tag{1.1}$$

where $\lambda_1, \ldots, \lambda_n$ are distinct points in an interval I, f is a differentiable function on I, and it is understood that $a_{ii} = f'(\lambda_i)$. It is a fundamental fact in Loewner's theory of operator monotone functions that f is operator monotone (see Section 4 for the definition) if and only if all matrices given by (1.1) associated with f are positive definite. Following the seminal paper of Löwner in 1934, there have been several expositions of this theory. However, no direct proof of the positivity of these matrices (for the standard examples of operator monotone functions) seems to have been found. Our method provides such a proof.

More examples of the efficacy of this technique are provided by giving new and simple proofs of the classical Heinz inequalities and their more recent generalisations that have aroused a lot of interest. Some new inequalities are established, and a decade-old problem related to the Lyapunov Equation is solved.

We summarise the essence of the technique in Section 2, and then illustrate its use in the subsequent sections.

2. *Preliminaries*

All matrices in this paper are $n \times n$ complex matrices. The entries of a matrix X are denoted by x_{ij}. A positive (semi)definite matrix will simply be called positive.

Received 20 August 1998; revised 23 April 1999.

1991 *Mathematics Subject Classification* 42A82, 47A63, 15A45, 15A60.

The second author is grateful to the Indian National Science Academy for financial support in the form of a C. V. Raman Research Professorship.

The symbol $|||\cdot|||$ stands for a unitarily invariant norm on the space of matrices; see [**3**, Chapter IV] for basic properties of such norms. The usual operator norm $\|\cdot\|$ is an example of such a norm.

The *Schur product* or the *Hadamard product* of two matrices X and Y will be denoted as $X \circ Y$. This is the matrix whose (i, j) entry is $x_{ij}y_{ij}$. A well-known theorem of Schur says that if X and Y are positive, then so is $X \circ Y$.

If X is positive, then for any matrix Y we have

$$|||X \circ Y||| \leqslant \max_i x_{ii} |||Y||| \tag{2.1}$$

for every unitarily invariant norm [**19**, p. 343].

We shall say that two matrices X and Y are *congruent* if $Y = T^*XT$ for some nonsingular matrix T. Congruence is an equivalence relation. If X is positive, then every matrix Y congruent to X is also positive.

We shall repeatedly use the connection between positive matrices, positive definite functions on $\mathbb{R}$ and positive definite kernels; see, for example, [**18**, pp. 400–402]. We shall also use basic facts about Fourier transforms. Let f be a function in $L^1(\mathbb{R})$. The Fourier transform of f is the function $\hat{f}$ defined as

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-i\xi x}\, dx.$$

By a well-known theorem of Bochner, if $f \in L^1(\mathbb{R})$, then $f(x) \geqslant 0$ for almost all x if and only if $\hat{f}$ is positive definite; see [**14**, p. 70].

Our method involves three ideas. The matrices that we study here have entries of the form $x_{ij} = f(\lambda_i, \lambda_j)$. First, we find a congruence that converts such a matrix to one whose entries have the form $x_{ij} = g(\lambda_i - \lambda_j)$. These matrices are positive if the function g is positive definite. The second step in our argument is calculating the Fourier transform $\hat{g}(\xi)$. If we find that $\hat{g}(\xi) \geqslant 0$, then we can conclude that g is positive definite. This establishes the positivity of the matrix X. Now we can use X as a Schur multiplier, and obtain norm inequalities using the inequality (2.1).

Germs of some of these ideas can be found in several recent papers; see, in particular, the papers by Corach, Porta and Recht [**5**, **6**], Horn [**17**], Mathias [**26**] and Zhan [**30**]. In a paper written while our work was in progress, Kosaki [**20**] has taken an approach closely related to ours. This has been elaborated further in two papers by Hiai and Kosaki [**15**, **16**].

When writing Fourier transforms, we shall ignore constant factors, since the only property we use is that of being positive almost everywhere.

For simplicity, we state and prove all our results for $n \times n$ matrices. Many of them are valid for operators in Hilbert space. The extensions are often (but not always) routine. Also, the passage from positive definite matrices to positive semidefinite ones is a routine matter via standard continuity arguments; we shall not make an explicit mention of it.

3. *The Heinz inequalities*

In [**13**] Heinz proved several inequalities for fractional powers of positive operators, and used them to obtain many results on the perturbation of spectral families of self-adjoint operators. Some of these have been generalised in different directions and given different proofs. We consider here two of these basic inequalities.

Let A, B be positive matrices, and let ν be any real number, $0 \leqslant \nu \leqslant 1$. Then for all X and all unitarily invariant norms,

$$|||A^{\nu}XB^{1-\nu} + A^{1-\nu}XB^{\nu}||| \leqslant |||AX + XB|||, \tag{3.1}$$

$$|||A^{\nu}XB^{1-\nu} - A^{1-\nu}XB^{\nu}||| \leqslant |2\nu - 1| \, |||AX - XB|||. \tag{3.2}$$

To appreciate the power of these inequalities, note that the very special case $X = I$ and $\nu = \frac{1}{2}$ of (3.1) gives

$$|||A^{1/2}B^{1/2}||| \leqslant \tfrac{1}{2}|||A + B|||,$$

a noncommutative arithmetic-geometric mean inequality. See the discussion in [**3**, Section IX.4] for connections between such results, and for different proofs.

Here we give yet another proof of these two inequalities that brings out a new connection between them.

It is enough to prove these inequalities for the special case $A = B$; the general case follows from the special one by a much used 2×2 block matrix argument; see, for example, [**3**, p. 264].

Assume $A = B$ in (3.1). Since the norms that we are considering are unitarily invariant, we may assume that A is diagonal with positive diagonal entries $\lambda_1, \ldots, \lambda_n$. Let Y be the matrix with entries y_{ij} given as

$$y_{ij} = \frac{\lambda_i^{\nu}\lambda_j^{1-\nu} + \lambda_i^{1-\nu}\lambda_j^{\nu}}{\lambda_i + \lambda_j}. \tag{3.3}$$

Note that

$$A^{\nu}XA^{1-\nu} + A^{1-\nu}XA^{\nu} = Y \circ (AX + XA).$$

The matrix Y is Hermitian, and all its diagonal entries are 1. So if we show that for $0 \leqslant \nu \leqslant 1$, Y is a positive matrix, then the inequality (3.1) would follow from (2.1).

The proof is easy for three special values of ν. When $\nu = 0$ or 1, all entries of Y are equal to 1. So Y is positive. When $\nu = \frac{1}{2}$, $y_{ij} = 2(\lambda_i\lambda_j)^{1/2}/(\lambda_i + \lambda_j)$. Let D be the diagonal matrix with diagonal entries $\lambda_1^{1/2}, \ldots, \lambda_n^{1/2}$, and let C be the matrix with entries

$$c_{ij} = \frac{1}{\lambda_i + \lambda_j}. \tag{3.4}$$

The matrix C given by (3.4) is called the Cauchy matrix and is known to be positive. Indeed, if $g_i = e^{-\lambda_i t}$, $1 \leqslant i \leqslant n$, then the inner product between g_i and g_j in the space $L_2([0, \infty))$ is

$$\langle g_i, g_j \rangle = \int_0^{\infty} e^{-(\lambda_i + \lambda_j)t} \, dt = \frac{1}{\lambda_i + \lambda_j}. \tag{3.5}$$

So the matrix C is a Gram matrix and, hence, is positive. Since $Y = 2DCD$, Y is congruent to C and is, therefore, positive too.

Now let ν be any real number in $(0, 1)$. Note that

$$y_{ij} = \lambda_i^{1-\nu} \left(\frac{\lambda_i^{2\nu-1} + \lambda_j^{2\nu-1}}{\lambda_i + \lambda_j} \right) \lambda_j^{1-\nu}.$$

Thus, applying a congruence, we see that the positivity of Y for $0 < \nu < 1$ is equivalent to that of the matrix Z with entries

$$z_{ij} = \frac{\lambda_i^{\alpha} + \lambda_j^{\alpha}}{\lambda_i + \lambda_j}, \quad -1 < \alpha < 1. \tag{3.6}$$

Since $\lambda_i > 0$, we can put $\lambda_i = e^{x_i}$ for some $x_i \in \mathbb{R}$. Thus to show that the matrix Z is positive, it suffices to show that the kernel

$$K(x, y) = \frac{e^{\alpha x} + e^{\alpha y}}{e^x + e^y}, \quad -1 < \alpha < 1, \tag{3.7}$$

is positive definite. Note that

$$\begin{aligned} K(x, y) &= \frac{e^{\alpha x/2}}{e^{x/2}} \left(\frac{e^{\alpha(x-y)/2} + e^{\alpha(y-x)/2}}{e^{(x-y)/2} + e^{(y-x)/2}} \right) \frac{e^{\alpha y/2}}{e^{y/2}} \\ &= \frac{e^{\alpha x/2}}{e^{x/2}} \left(\frac{\cosh \alpha(x-y)/2}{\cosh(x-y)/2} \right) \frac{e^{\alpha y/2}}{e^{y/2}}. \end{aligned}$$

So $K(x, y)$ is positive definite if and only if the kernel

$$L(x, y) = \frac{\cosh \alpha(x-y)/2}{\cosh(x-y)/2}, \quad -1 < \alpha < 1, \tag{3.8}$$

is positive definite. This follows from the following theorem.

THEOREM 3.1. *For $-1 < \alpha < 1$, the function*

$$f(x) = \frac{\cosh \alpha x}{\cosh x} \tag{3.9}$$

is a positive definite function on $\mathbb{R}$.

Proof. The function f is even. Its Fourier cosine transform can be seen to be

$$\hat{f}(\xi) = \frac{\cos(\alpha\pi)/2 \ \cosh(\xi\pi)/2}{\cosh \xi\pi + \cos \alpha\pi}; \tag{3.10}$$

see [**11**, p. 1192]. For $-1 < \alpha < 1$, $\cos \alpha\pi/2$ is positive. For all $\xi \neq 0$, $\cosh \xi\pi > 1$. So the denominator in (3.10) is also positive. Thus $\hat{f}(\xi) \geqslant 0$. Hence, by Bochner's Theorem, f is positive definite.

REMARK 3.1. Another proof of Theorem 3.1 goes as follows. It has the advantage of avoiding the complicated-looking transform (3.10). Use the familiar factoring

$$\cosh x = \prod_{k=0}^{\infty} \left(1 + \frac{4x^2}{(2k+1)^2\pi^2} \right), \tag{3.11}$$

to write

$$\begin{aligned} \frac{\cosh \alpha x/2}{\cosh x/2} &= \prod_{k=0}^{\infty} \frac{1 + \alpha^2 x^2/(2k+1)^2\pi^2}{1 + x^2/(2k+1)^2\pi^2} \\ &= \prod_{k=0}^{\infty} \left[\frac{1-\alpha^2}{1 + x^2/(2k+1)^2\pi^2} + \alpha^2 \right]. \end{aligned} \tag{3.12}$$

To prove that the product is positive definite, it suffices to prove that each factor is positive definite. Since $-1 < \alpha < 1$, it suffices to prove that for each k, the function

$$g(x) = \frac{1}{1 + x^2/(2k+1)^2\pi^2}$$

is positive definite. The function

$$h(x) = \frac{1}{1+x^2}$$

is an easily recognised positive definite function, being the Fourier transform of the positive function $e^{-|\xi|}$. Thus for each a, the function $1/(1+a^2x^2)$ is also positive definite. So $g(x)$ is positive definite.

REMARK 3.2. The Schur product is used to derive the arithmetic-geometric mean inequality in [**5**], [**17**] and [**26**]. In [**6**], the authors derive a norm inequality by appealing to Theorem 3.1 and to Schur's Theorem.

REMARK 3.3. We have used the positivity of the matrix given by (3.3) to prove the Heinz inequality (3.1). The argument can be turned around. Mathias [**26**] has given a proof of the positivity of the matrix given by (3.6) for $0 < \alpha < 1$, using a norm inequality from [**4**], also given in [**3**, Theorem IX.4.8].

Now we consider the inequality (3.2). It is enough to prove this for $A = B$, and it is enough to consider the case $\frac{1}{2} < v < 1$.

Let W be the matrix whose entries are

$$\begin{aligned} w_{ij} &= \frac{\lambda_i^v \lambda_j^{1-v} - \lambda_i^{1-v}\lambda_j^v}{\lambda_i - \lambda_j}, \quad i \neq j, \\ w_{ii} &= 2v - 1 \quad \text{for all } i. \end{aligned} \tag{3.13}$$

By the arguments that we used to prove (3.1), the inequality (3.2) will be proved if we can show that W is positive whenever the λ_i are positive.

Let D be the diagonal matrix with entries $\lambda_1^{v-1}, \ldots, \lambda_n^{v-1}$ down its diagonal, and let $V = DWD$. Then V has entries $v_{ij} = (\lambda_i^{2v-1} - \lambda_j^{2v-1})/(\lambda_i - \lambda_j)$, $i \neq j$, and $v_{ii} = (2v-1)\lambda_i^{2v-2}$ for all i. We want to prove that this matrix is positive for all v, $\frac{1}{2} < v < 1$. Put $\alpha = 2v - 1$. We then need to prove that the matrix V with entries

$$\begin{aligned} v_{ij} &= \frac{\lambda_i^\alpha - \lambda_j^\alpha}{\lambda_i - \lambda_j}, \quad i \neq j, \\ v_{ii} &= \alpha\lambda_i^{\alpha-1} \quad \text{for all } i, \end{aligned} \tag{3.14}$$

is positive for all α, $0 < \alpha < 1$.

As before, one can see that the matrix V given by (3.14) is positive if and only if the kernel $K_1(x, y)$ defined as

$$\begin{aligned} K_1(x, y) &= \frac{e^{\alpha x} - e^{\alpha y}}{e^x - e^y}, \quad x \neq y, \\ K_1(x, x) &= \alpha e^{(\alpha-1)x}, \end{aligned} \tag{3.15}$$

is positive definite for $0 < \alpha < 1$. In turn, this is positive definite if and only if the kernel

$$\begin{aligned} L_1(x, y) &= \frac{\sinh \alpha(x-y)/2}{\sinh(x-y)/2}, \quad x \neq y, \\ L_1(x, x) &= \alpha, \end{aligned} \tag{3.16}$$

is positive definite for $0 < \alpha < 1$ (see the passage from (3.7) to (3.8)). This positive definiteness is a consequence of the following theorem.

THEOREM 3.2. *For $0 < \alpha < 1$, the function*

$$f_1(x) = \frac{\sinh \alpha x}{\sinh x} \tag{3.17}$$

is a positive definite function on $\mathbb{R}$.

Proof. The function f_1 is even. Its Fourier cosine transform is

$$\hat{f}_1(\xi) = \frac{\sin \alpha\pi}{\cosh \xi\pi + \cos \alpha\pi}; \tag{3.18}$$

see [**11**, p. 1192]. For $0 < \alpha < 1$, the numerator in (3.18) is positive; the denominator is also positive, since $\cosh \xi\pi > 1$ for all $\xi \neq 0$. Thus $\hat{f}_1(\xi) \geqslant 0$ for all ξ. Hence, by Bochner's Theorem, f_1 is a positive definite function.

REMARK 3.4. Once again, we could avoid the computation of the transform (3.18) if we use the factoring

$$\frac{\sinh x}{x} = \prod_{k=1}^{\infty} \left(1 + \frac{x^2}{k^2\pi^2}\right).$$

Then we can write

$$\frac{\sinh \alpha x}{\sinh x} = \alpha \prod_{k=1}^{\infty} \frac{1 + \alpha^2 x^2/k^2\pi^2}{1 + x^2/k^2\pi^2}.$$

Each factor in this product is of the form

$$\frac{1 + b^2x^2}{1 + a^2x^2} = \frac{b^2}{a^2} + \frac{1 - b^2/a^2}{1 + a^2x^2}, \quad 0 \leqslant b < a,$$

and is, therefore, positive definite.

REMARK 3.5. The positivity of the matrix V given by (3.14) is equivalent to the operator monotonicity of the function $f(t) = t^\alpha$, $0 < \alpha < 1$, on the positive half line. This is a consequence of Loewner's Thoerem [**3**, Theorem V.3.4]. We shall discuss this in greater detail in Section 4.

REMARK 3.6. Furuta [**10**] has observed that several norm inequalities for operators are equivalent to the operator monotonicity of $f(t) = t^\alpha$, $0 < \alpha < 1$, on $\mathbb{R}_+$. Our analysis shows the equivalence of (3.2) to this list. In this context, see also [**9**].

REMARK 3.7. Our analysis shows that (3.1) and (3.2) are as intimately related as the functions cosh and sinh. There is another sense in which these two inequalities are real and imaginary parts of one statement. Let X be a Hermitian and A a positive matrix. Let $0 < v < 1$, and let

$$S = A^v X A^{1-v}, \quad T = vAX + (1-v)XA.$$

If we use the notations $\operatorname{Re} Z$ and $\operatorname{Im} Z$ to mean

$$\operatorname{Re} Z = \tfrac{1}{2}(Z + Z^*), \quad \operatorname{Im} Z = \tfrac{1}{2i}(Z - Z^*)$$

for every matrix Z, then we see that

$$\begin{aligned}
\operatorname{Re} S &= \tfrac{1}{2}(A^{v}XA^{1-v} + A^{1-v}XA^{v}),\\
\operatorname{Im} S &= \tfrac{1}{2i}(A^{v}XA^{1-v} - A^{1-v}XA^{v}),\\
\operatorname{Re} T &= \tfrac{1}{2}(AX + XA),\\
\operatorname{Im} T &= \tfrac{1}{2i}(2v-1)(AX - XA).
\end{aligned}$$

So the inequalities (3.1) and (3.2) say that

$$|||\operatorname{Re} S||| \leqslant |||\operatorname{Re} T|||, \quad |||\operatorname{Im} S||| \leqslant |||\operatorname{Im} T|||. \tag{3.19}$$

Remark 3.8. For the Hilbert–Schmidt norm $\|Z\|_2 = (\operatorname{tr} Z^*Z)^{1/2}$, we have

$$\|Z\|_2^2 = \|\operatorname{Re} Z\|_2^2 + \|\operatorname{Im} Z\|_2^2.$$

In this case, the two inequalities in (3.19) can be combined to say $\|S\|_2 \leqslant \|T\|_2$. This need not be true for other norms.

Remark 3.9. It is easy to see that if X is any matrix and A any positive matrix, then

$$\|A^{v}XA^{1-v}\|_2 \leqslant \|vAX + (1-v)XA\|_2, \tag{3.20}$$

for $0 < v < 1$. A corresponding statement for other unitarily invariant norms need not be true. Choose

$$A = \begin{pmatrix} 8 & 0 \\ 0 & 1 \end{pmatrix}, \quad v = \tfrac{1}{3},$$

to obtain an easy counter-example to (3.20) when the norm $\|\cdot\|_2$ is replaced by $|||\cdot|||$. When $v = \frac{1}{2}$, the inequality (3.20) does hold for all unitarily invariant norms. This is a special case of (3.1).

Remark 3.10. In a recent paper [**20**], Kosaki has given other proofs of (3.1), (3.2) and related inequalities. He uses integral transforms, and the trigonometric and hyperbolic functions enter his calculations, but in a way different from ours.

4. *Operator monotone functions*

Let f be a real-valued function on an interval I. Then f is said to be operator monotone if it satisfies the following property: whenever A and B are two Hermitian matrices of the same size, with all their eigenvalues in I and such that $A \geqslant B$, then $f(A) \geqslant f(B)$. (Here $A \geqslant B$ means that $A - B$ is positive.)

A rich theory of operator monotone functions was developed by Löwner [**24**]. Among the several papers and books where such functions are studied are [**1**, **2**, **3**, **7**, **8**, **12**, **13**, **19**, **22**, **23**, **25**].

Given any n points $\lambda_1, \ldots, \lambda_n$ in I, consider the matrix of divided differences of f: this is the matrix with (i, j) entries

$$\begin{cases} \dfrac{f(\lambda_i) - f(\lambda_j)}{\lambda_i - \lambda_j} & \text{if } i \neq j, \\ f'(\lambda_i) & \text{if } i = j. \end{cases}$$

One of the first steps in Loewner's theory is to show that f is operator monotone on I if and only if all such matrices (for all choices of $\lambda_1, \ldots, \lambda_n$ in I) are positive [**3**, Theorem V.3.4]. This is a natural generalisation of the fact that a function with a positive derivative is monotone.

Other characterisations of operator monotone functions can be derived from this: they have analytic continuations that map the upper half plane into itself [**3**, Theorem V.4.7], and they have integral representations of a special type [**3**, pp. 134–145].

It is somewhat curious that the operator monotonicity of special functions f is proved in all the sources cited above by appealing to the latter characterisations, or by special arguments for each function. No proof based on the first characterisation (positivity of the divided-difference matrix) seems to be known. Our approach in Section 3 readily leads to such proofs for all the standard examples; see below.

EXAMPLE 4.1. For $0 < \alpha < 1$, the function $f(t) = t^\alpha$ is operator monotone on $(0, \infty)$.

In this case, the matrix of divided differences is the matrix V given by (3.14). We have seen that this is a positive matrix.

EXAMPLE 4.2. The function $f(t) = \log t$ is operator monotone on $(0, \infty)$.

The matrix of divided differences, in this case, is the matrix V with entries

$$v_{ij} = \frac{\log \lambda_i - \log \lambda_j}{\lambda_i - \lambda_j}, \quad i \neq j,$$
$$v_{ii} = \frac{1}{\lambda_i} \quad \text{for all } i. \tag{4.1}$$

Making the substitution $\lambda_i = e^{x_i}$, we see that the positivity of the matrix V is equivalent to the positive definiteness of the kernel

$$K_2(x, y) = \frac{x - y}{e^x - e^y}, \quad x \neq y.$$

Since

$$K_2(x, y) = \frac{1}{e^{x/2}} \left(\frac{(x - y)/2}{\sinh(x - y)/2} \right) \frac{1}{e^{y/2}},$$

the positive definiteness of $K_2(x, y)$ is equivalent to the positive definiteness of the function

$$f_2(x) = \frac{x}{\sinh x}.$$

This is an even function, and its Fourier cosine transform is

$$\hat{f}_2(\xi) = \frac{e^{\pi\xi}}{(1 + e^{\pi\xi})^2}$$

[**11**, p. 1185]. Since $\hat{f}_2(\xi) \geqslant 0$, the function f_2 is positive definite by Bochner's Theorem.

Once again, if we use the infinite factoring

$$\frac{x}{\sinh x} = \prod_{k=1}^{\infty} \frac{k^2\pi^2}{k^2\pi^2 + x^2},$$

then we need not calculate $\hat{f}_2(\xi)$.

Example 4.3. The function $f(t) = \tan t$ is operator monotone on $(-\pi/2, \pi/2)$. Now the matrix of divided differences is the matrix V whose entries are

$$v_{ij} = \frac{\tan\lambda_i - \tan\lambda_j}{\lambda_i - \lambda_j}, \quad i \neq j,$$
$$v_{ii} = \sec^2\lambda_i \quad \text{for all } i. \tag{4.2}$$

Using the identity

$$\tan x - \tan y = \frac{\sin(x-y)}{\cos x \cos y},$$

we can write

$$v_{ij} = \frac{1}{\cos\lambda_i}\left(\frac{\sin(\lambda_i - \lambda_j)}{\lambda_i - \lambda_j}\right)\frac{1}{\cos\lambda_j}.$$

The matrix V is thus congruent to the matrix W with entries

$$w_{ij} = \frac{\sin(\lambda_i - \lambda_j)}{\lambda_i - \lambda_j}. \tag{4.3}$$

To prove that W is positive, we have to show that the function $f(x) = (\sin x)/x$ is positive definite. This is easy: f is the Fourier transform of the characteristic function of the interval $[-1, 1]$.

Example 4.4. For completeness, we note an easy proof of the fact that the function

$$f(t) = \frac{at+b}{ct+d}, \quad ad - bc > 0,$$

is operator monotone on any interval that does not contain the point $-d/c$. Following the steps for other functions studied above, one sees that this reduces to showing that the matrix with all entries 1 is positive.

From the fact that the matrices V given by (4.1) and (4.2) are positive, we can derive some inequalities for norms of commutators, just as we did in Section 3. This is indicated below.

Let D be the diagonal matrix with entries $\lambda_1^{1/2}, \ldots, \lambda_n^{1/2}$ down its diagonal. Let $W = DVD$, where V is the matrix given by (4.1). Then W is positive. The entries of W are

$$w_{ij} = \frac{\log\lambda_i - \log\lambda_j}{\lambda_i^{1/2}\lambda_j^{-1/2} - \lambda_i^{-1/2}\lambda_j^{1/2}}, \quad i \neq j,$$
$$w_{ii} = 1 \quad \text{for all } i. \tag{4.4}$$

Using the arguments in Section 3, we obtain, from the positivity of this matrix, the inequality

$$|||(\log A)X - X(\log B)||| \leqslant |||A^{1/2}XB^{-1/2} - A^{-1/2}XB^{1/2}|||, \tag{4.5}$$

valid for positive matrices A, B and for every matrix X. From this we obtain the inequality

$$|||HX - XK||| \leqslant |||e^{H/2}Xe^{-K/2} - e^{-H/2}Xe^{K/2}|||, \tag{4.6}$$

valid for all X and for Hermitian H, K.

In the same way, from the fact that the matrix W given by (4.3) is positive, we obtain the inequality

$$|||(\sin H)X(\cos K) - (\cos H)X(\sin K)||| \leqslant |||HX - XK|||. \tag{4.7}$$

The inequalities (4.5)–(4.7) have been proved recently by Kosaki [**20**]. He remarks that since trigonometric functions are neither monotone nor convex, a proof of (4.7) using majorisation type arguments seems impossible. The proof that we have given is just such a proof. It is clear that many more inequalities could be obtained using these ideas. For example, from the positive definiteness of the function $x/(\sinh x)$, one sees that

$$|||HX - XK||| \leqslant |||(\sinh H)X(\cosh K) - (\cosh H)X(\sinh K)|||, \tag{4.8}$$

for all X and for Hermitian H, K.

The *logarithmic mean* of two positive numbers a and b is, by definition, the quantity

$$\frac{a-b}{\log a - \log b} = \int_0^1 a^t b^{1-t}\, dt. \tag{4.9}$$

We have then a refinement of the arithmetic-geometric mean inequality:

$$\sqrt{ab} \leqslant \int_0^1 a^t b^{1-t}\, dt \leqslant \tfrac{1}{2}(a+b). \tag{4.10}$$

An operator version of this has been proved recently by Hiai and Kosaki [**15**]. This says that for positive matrices A, B and for every matrix X,

$$|||A^{1/2}XB^{1/2}||| \leqslant |||\int_0^1 A^t X B^{1-t}\, dt||| \leqslant \tfrac{1}{2}|||AX + XA|||. \tag{4.11}$$

Let us see how this can be derived easily using our technique. Using (4.9) and the arguments in Section 3, one can see that the first inequality in (4.11) follows from the positivity of the matrix given by (4.1). The same arguments show that the second inequality in (4.11) would follow if we show that the matrix W with entries

$$w_{ij} = \frac{2(\lambda_i - \lambda_j)}{(\log\lambda_i - \log\lambda_j)(\lambda_i + \lambda_j)}, \quad i \neq j,$$
$$w_{ii} = 1 \quad \text{for all } i, \tag{4.12}$$

is positive. Again making the substitution $\lambda_i = e^{x_i}$, we see that

$$w_{ij} = \frac{\tanh(x_i - x_j)/2}{(x_i - x_j)/2}. \tag{4.13}$$

Thus the positivity of W would follow from the positive definiteness of the function

$$f(x) = \frac{\tanh x}{x}. \tag{4.14}$$

Since

$$\frac{\tanh x}{x} = \int_0^1 \frac{\cosh \alpha x}{\cosh x}\, d\alpha,$$

Theorem 3.1 implies that f is a positive definite function. We could also prove this directly. The function f is even, and its Fourier cosine transform is

$$\hat{f}(\xi) = \log\coth\frac{\pi\xi}{4}, \quad \xi > 0; \tag{4.15}$$

see [**11**, p. 549], [**28**, p. 36, Formula 7.37].

Further refinements and generalisations of (4.11) have been obtained in [**15**] and [**16**].

5. *Generalisations of Lyapunov's Theorem*

Consider the matrix equation

$$AX + XA = B, \tag{5.1}$$

where A is a positive matrix. This is a special case of Lyapunov's Equation. If we choose a basis in which A is diagonal with diagonal entries $\lambda_1, \ldots, \lambda_n$, then the solution X can be written as

$$x_{ij} = \frac{b_{ij}}{\lambda_i + \lambda_j}.$$

In other words, X is the Schur product of B with the Cauchy matrix given by (3.4). Thus X is positive if B is; a fact well-known, and of great importance, in the Lyapunov theory.

Equations more general than (5.1) that involve polynomial expressions in A have also been studied. The simplest such equation is

$$A^2X + XA^2 + tAXA = B. \tag{5.2}$$

If A is positive, then the solution in a basis that diagonalises A can be expressed as

$$x_{ij} = \frac{b_{ij}}{\lambda_i^2 + \lambda_j^2 + t\lambda_i\lambda_j}.$$

This is the Schur product of B with the matrix Z whose entries are

$$z_{ij} = \frac{1}{\lambda_i^2 + \lambda_j^2 + t\lambda_i\lambda_j}. \tag{5.3}$$

So if B is positive, then the solution X of the equation (5.2) would be positive if the matrix Z is positive.

It is easy to see that for $t \leqslant -2$, the matrix Z is not positive. For $t = 2$, Z is the Schur product of the Cauchy matrix (3.4) with itself, and is, therefore, positive. What happens for other values of t? This problem was studied by Kwong [**21**]. He showed by a somewhat intricate topological argument that the matrix Z is positive if $t \in (-2, 2]$, irrespective of its order n. He also showed that when $n = 2, 3, 4$, the matrix Z is positive for $t \in (-2, \infty)$, $(-2, 8)$ and $(-2, 4)$, respectively, and asked for necessary and sufficient conditions on t for Z to be positive for all n. Our next theorem says that the matrix Z is positive for all n if and only if $t \in (-2, 2]$. The proof is more transparent and simpler than the one given in [**21**] for one half of the theorem.

THEOREM 5.1. *Let λ_i be positive numbers. Then the $n \times n$ matrices Z defined in* (5.3) *are positive for all n if and only if $t \in (-2, 2]$.*

Proof. The matrix Z is positive if and only if the kernel

$$K_3(x, y) = \frac{1}{e^{2x} + e^{2y} + te^x e^y}$$

is positive definite. Since

$$K_3(x,y) = \frac{1}{e^x}\left(\frac{1}{2\cosh(x-y)+t}\right)\frac{1}{e^y},$$

this is positive definite if and only if the function

$$f_3(x) = \frac{1}{2\cosh x + t} \tag{5.4}$$

is positive definite. By Bochner's Theorem, this is so if and only if $\hat{f}_3(\xi) \geqslant 0$. The Fourier transform of the even function f_3 can be read from Formula 1 of 3.983 in [**11**, p. 538]. We have, for $\xi > 0$,

$$\hat{f}_3(\xi) = \begin{cases} \dfrac{\sinh(\xi \arccos t/2)}{(4-t^2)^{1/2}\sinh \xi\pi} & \text{for } -2 < t < 2, \\ \dfrac{\sin(\xi \arccos t/2)}{(t^2-4)^{1/2}\sinh \xi\pi} & \text{for } t > 2. \end{cases}$$

In the first case, $\hat{f}_3(\xi) \geqslant 0$ for $\xi > 0$. Since $\hat{f}_3$ is an even function, we have $\hat{f}_3(\xi) \geqslant 0$ for all $\xi \in \mathbb{R}$. In the second case, $\hat{f}_3(\xi)$ assumes negative values for ξ in a set of positive measure.

We conclude that f_3 is positive definite, and hence the matrix Z is positive for all n if $-2 < t < 2$, but not if $t > 2$. We have already noted that Z is positive for $t = 2$, but not for $t < -2$.

In the same spirit, we can give a proof of another theorem, proved by Kwong [**22**, Theorem 10] using a different technique.

THEOREM 5.2. *Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be positive numbers. Then the $n \times n$ matrix W with entries*

$$w_{ij} = \frac{\lambda_i^\nu + \lambda_j^\nu}{\lambda_i^2 + \lambda_j^2 + t\lambda_i\lambda_j} \tag{5.5}$$

is positive for $-1 < \nu < 1$ and $-2 < t < 2$.

Proof. The assertion of the theorem is equivalent to saying that the kernel

$$K_4(x,y) = \frac{e^{\alpha x} + e^{\alpha y}}{e^{2x} + e^{2y} + te^x e^y} \tag{5.6}$$

is positive definite for $-1 < \alpha < 1$, $-2 < t < 2$. Since

$$K_4(x,y) = \frac{e^{\alpha x/2}}{e^x}\left(\frac{2\cosh \alpha(x-y)/2}{2\cosh(x-y)+t}\right)\frac{e^{\alpha y/2}}{e^y},$$

this kernel is positive definite if and only if the function

$$f_4(x) = \frac{\cosh \alpha x}{\cosh 2x + t} \tag{5.7}$$

is positive definite for $-1 < \alpha < 1$, $-1 < t < 1$. Once again, we have an even function whose Fourier cosine transform can be read off from Formula 6 of 3.983 in [**11**, p. 539]. It is convenient to make the substitution $\delta = \arccos t$. Note that $0 < \delta < \pi$ if $-1 < t < 1$. We then have

$$\hat{f}_4(\xi) = \frac{\cos\frac{\alpha}{2}(\pi-\delta)\cosh\frac{\xi}{2}(\pi+\delta) - \cos\frac{\alpha}{2}(\pi+\delta)\cosh\frac{\xi}{2}(\pi-\delta)}{\sin\delta(\cosh \pi\xi - \cos\pi\alpha)}. \tag{5.8}$$

Since $\sin\delta > 0$ and $\cosh\pi\xi \geqslant 1 > \cos\pi\alpha$, the denominator in (5.8) is positive. The numerator is also positive. To see this, use the identities for $\cos(x \pm y)$ and $\cosh(x \pm y)$, and re-group the terms to see that the numerator is two times the quantity

$$\cos\frac{\alpha\pi}{2}\cos\frac{\alpha\delta}{2}\sinh\frac{\xi\pi}{2}\sinh\frac{\xi\delta}{2} + \sin\frac{\alpha\pi}{2}\sin\frac{\alpha\delta}{2}\cosh\frac{\xi\pi}{2}\cosh\frac{\xi\delta}{2}.$$

For $-1 < \alpha < 1$ and $0 < \delta < \pi$, this quantity is positive.

Remark 5.1. An alternate proof of Theorem 5.2 that avoids the complicated Fourier transform (5.8) goes as follows. Rewrite the expression in (5.5) as

$$w_{ij} = \frac{1}{\lambda_i + \lambda_j}\,\frac{\lambda_i^\nu + \lambda_j^\nu}{\lambda_i + \lambda_j}\sum_{n=0}^{\infty}(2-t)^n\frac{\lambda_i^n\lambda_j^n}{(\lambda_i + \lambda_j)^{2n}}.$$

The positivity of W is then a consequence of the positivity of the Cauchy matrix given by (3.4), the matrix given by (3.6) and the Schur product theorem. We are thankful to an anonymous referee for this elegant argument.

Remark 5.2. Once again, it is possible to obtain inequalities for operators from Theorems 5.1 and 5.2. Some of these have been written down by Zhan [**30**].

Remark 5.3. The idea used by Kwong [**21**, **22**] to prove positivity is extremely interesting. Our approach using congruence sheds some new light on it.

Let $\mathscr{L}$ be the set of all functions g mapping the positive half line into itself, for which the $n \times n$ matrix with (i, j) entry

$$\frac{g(\lambda_i) + g(\lambda_j)}{\lambda_i + \lambda_j}$$

is positive for all n and for all positive numbers $\lambda_1, \lambda_2, \ldots$.

The set $\mathscr{L}$ is a closed cone, and contains the functions $g(t) = 1$ and $g(t) = t$. It also contains the functions $g_s(t) = t/(t+s)$ for all $s \geqslant 0$. To see this, note that

$$\frac{g_s(\lambda_i) + g_s(\lambda_j)}{\lambda_i + \lambda_j} = \frac{1}{\lambda_i + s}\left(\frac{2\lambda_i\lambda_j + s(\lambda_i + \lambda_j)}{\lambda_i + \lambda_j}\right)\frac{1}{\lambda_j + s}.$$

So the positivity of the Cauchy matrix implies the positivity of this matrix (simply using the fact that taking positive linear combinations and applying congruences preserves positivity).

Thus $\mathscr{L}$ includes all operator monotone functions f, since such an f has an integral representation

$$f(t) = \alpha + \beta t + \int_0^\infty \frac{st}{t+s}\,d\mu(s), \tag{5.9}$$

where $\alpha, \beta \geqslant 0$ and μ is a positive measure.

This gives an alternate proof of the positivity of some of the matrices that we have considered.

Remark 5.4. Using the same ideas, it is easy to prove the positivity of some matrices whose entries are complex numbers with positive real part.

Let λ_i be complex numbers with positive real parts. The Cauchy matrix C with entries

$$c_{ij} = \frac{1}{\bar{\lambda}_i + \lambda_j} \tag{5.10}$$

is positive. (The proof is the same as we gave for the matrix given by (3.4).) The argument that we have outlined above can be repeated to show that the matrix with (i, j) entry

$$\frac{f(\bar{\lambda}_i) + f(\lambda_j)}{\bar{\lambda}_i + \lambda_j} \tag{5.11}$$

is positive for all functions f of the form (5.9).

REMARK 5.5. What we have said in Remark 5.4 presents a way of extending many of the known inequalities from positive matrices to normal matrices with all their eigenvalues in the right half plane.

6. *Infinite divisibility of the Cauchy matrix*

We say that a matrix A is entrywise positive if all of its entries a_{ij} are positive numbers. For such a matrix, we denote by $A^{(r)}$ the matrix whose entries are a_{ij}^r, $r > 0$. If a matrix A is both positive semidefinite and entrywise positive, then we say that A is infinitely divisible if $A^{(r)}$ is positive semidefinite for all $r > 0$; see [**19**, p. 456].

The techniques introduced in this paper give an interesting proof of the known fact that the Cauchy matrix C defined by (3.4) is infinitely divisible.

Once again, let $\lambda_i > 0$ be given, and put $\lambda_i = e^{x_i}$, $x_i \in \mathbb{R}$. To show that C is infinitely divisible, we need to show that for each $r > 0$, the kernel

$$K_5(x, y) = \frac{1}{(e^x + e^y)^r} \tag{6.1}$$

is positive definite. Since

$$\begin{aligned} K_5(x, y) &= \frac{1}{e^{rx/2}} \frac{1}{[e^{(x-y)/2} + e^{(y-x)/2}]^r} \frac{1}{e^{ry/2}} \\ &= \frac{1}{e^{rx/2}} \frac{1}{[\cosh(\frac{x-y}{2})]^r} \frac{1}{e^{ry/2}}, \end{aligned}$$

the kernel $K_5(x, y)$ is positive definite if and only if the function

$$f_5(x) = (\operatorname{sech} x)^r \tag{6.2}$$

is positive definite for $r > 0$. The Fourier cosine transform of the function f_5 can be read off from Formula 7.5 in [**24**, p. 33]. We have

$$\hat{f}_5(\xi) = 2^{r-2} \frac{1}{\Gamma(r)} \left| \Gamma\left(\frac{r + i\xi}{2}\right) \right|^2. \tag{6.3}$$

This shows that f_5 is positive definite.

Other proofs of this property of the Cauchy matrix are outlined in [**19**, p. 458].

References

1. T. Ando, *Topics on operator inequalities* (Hokkaido University, Sapporo, 1978).
2. T. Ando, 'Majorizations and inequalities in matrix theory', *Linear Algebra Appl.* 199 (1994) 17–67.
3. R. Bhatia, *Matrix analysis* (Springer, New York, 1997).
4. R. Bhatia and C. Davis, 'More matrix forms of the arithmetic-geometric mean inequality', *SIAM J. Matrix Anal. Appl.* 14 (1993) 132–136.
5. G. Corach, H. Porta and L. Recht, 'An operator inequality', *Linear Algebra Appl.* 142 (1990) 153–158.
6. G. Corach, H. Porta and L. Recht, 'Jacobi fields on space of positive operators', *Linear Algebra Appl.* 179 (1993) 271–275.
7. C. Davis, 'Notions generalizing convexity for functions defined on spaces of matrices', *Convexity*, Proc. Sympos. Pure Math. 8 (Amer. Math. Soc., Providence, RI, 1963).
8. W. F. Donoghue, *Monotone matrix functions and analytic continuation* (Springer, New York, 1974).
9. J. Fujii, M. Fujii, T. Furuta and R. Nakamoto, 'Norm inequalities equivalent to Heinz inequality', *Proc. Amer. Math. Soc.* 118 (1993) 827–830.
10. T. Furuta, 'Norm inequalities equivalent to Löwner–Heinz theorem', *Rev. Math. Phys.* 1 (1989) 135–137.
11. I. S. Gradshteyn and I. M. Ryzhik, *Tables of integrals, series, and products* (5th edn, Academic Press, New York, 1994).
12. F. Hansen and G. K. Pedersen, 'Jensen's inequality for operators and Löwner's theorem', *Math. Ann.* 258 (1982) 229–241.
13. E. Heinz, 'Beiträge zur Störungstheorie der Spektralzerlegung', *Math. Ann.* 123 (1951) 415–438.
14. H. Helson, *Harmonic analysis* (2nd edn, Hindustan Book Agency, Delhi, 1995).
15. F. Hiai and H. Kosaki, 'Comparison of various means for operators', *J. Funct. Anal.* 163 (1999) 300–323.
16. F. Hiai and H. Kosaki, 'Means for matrices and comparison of their norms', Preprint 1998-23, Kyushu University, Fukuoka; *Indiana Univ. Math. J.*, to appear.
17. R. A. Horn, 'Norm bounds for Hadamard products and the arithmetic-geometric mean inequality for unitarily invariant norms', *Linear Algebra Appl.* 223/224 (1995) 355–361.
18. R. A. Horn and C. R. Johnson, *Matrix analysis* (Cambridge University Press, New York, 1985).
19. R. A. Horn and C. R. Johnson, *Topics in matrix analysis* (Cambridge University Press, New York, 1990).
20. H. Kosaki, 'Arithmetic-geometric mean and related inequalities for operators', *J. Funct. Anal.* 156 (1998) 429–451.
21. M. K. Kwong, 'On the definiteness of the solutions of certain matrix equations', *Linear Algebra Appl.* 108 (1988) 177–197.
22. M. K. Kwong, 'Some results on matrix monotone functions', *Linear Algebra Appl.* 118 (1989) 129–153.
23. C. Loewner, 'Advanced matrix theory', Mimeographed Lecture Notes, Stanford University, Winter 1957.
24. K. Löwner, 'Über monotone Matrixfunctionen', *Math. Z.* 38 (1934) 177–216.
25. A. W. Marshall and I. Olkin, *Inequalities: theory of majorization and its applications* (Academic Press, New York, 1979).
26. R. Mathias, 'An arithmetic-geometric mean inequality involving Hadamard products', *Linear Algebra Appl.* 184 (1993) 71–78.
27. A. McIntosh, 'Heinz inequalities and perturbation of spectral families', Report 79-0006, Macquarie University, North Ryde, 1979.
28. F. Oberhettinger, *Tables of Fourier transforms and Fourier transforms of distributions* (Springer, Berlin, 1990).
29. Z. Sasvári, *Positive definite and definitizable functions* (Akademie, Berlin, 1994).
30. X. Zhan, 'Inequalities for unitarily invariant norms', *SIAM J. Matrix Anal.* 20 (1999) 466–470.

Indian Statistical Institute
7, S.J.S. Sansanwal Marg
New Delhi – 110016
India