isid/ms/2004/21
September 10, 2004
http://www.isid.ac.in/~ statmath/eprints

Generalized Lyapunov Equations and Positive Definite Functions

Rajendra Bhatia
Driss Drissi

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

Generalized Lyapunov Equations and Positive Definite Functions

Rajendra Bhatia
Indian Statistical Institute
New Delhi - 110 016, INDIA
rbh@isid.ac.in

and

Driss Drissi
Department of Mathematics
Kuwait University
Safat, Kuwait
drissi_math@yahoo.com

Abstract: We establish the positive definiteness of some functions and of some matrices that arise as solutions of generalized Lyapunov equations.

Mathematics Subject Classification (2000) : 42A82, 47A62, 15A24, 15A48.

Key words : Positive definite matrix, positive definite function, Fourier transform, Bochner's theorem, Lyapunov matrix equation, operator means

December 1, 2004

1 Introduction

Let A be a positive definite matrix and consider the following matrix equations:

(L1) $AX + XA = B$.

(L2) $A^2X + XA^2 + tAXA = B$.

(L3) $A^3X + XA^3 + t(A^2XA + AXA^2) = B$.

(L4) $A^4X + XA^4 + t(A^3XA + AXA^3) + 6A^2XA^2 = B$.

(L5) $A^4X + XA^4 + 4(A^3XA + AXA^3) + tA^2XA^2 = B$.

The equation (L1) is known as the *Lyapunov equation* and has been studied extensively.

We may choose an orthonormal basis for the underlying space in which A is diagonal, $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Then, in component form, (L1) may be written as

$$(\lambda_i + \lambda_j)x_{ij} = b_{ij}$$

and solved as

$$x_{ij} = \frac{b_{ij}}{\lambda_i + \lambda_j}. \tag{1}$$

The matrix C with entries

$$c_{ij} = \frac{1}{\lambda_i + \lambda_j} \tag{2}$$

is called the *Cauchy matrix.* It is positive definite. One way of seeing this is by writing

$$\frac{1}{\lambda_i + \lambda_j} = \int_0^\infty e^{-t(\lambda_i + \lambda_j)} dt, \tag{3}$$

and thus realising C as the Gram matrix associated with the vectors $\{e^{-t\lambda_i}\}$ in the Hilbert space $L_2(0, \infty)$. The solution (1) can be written in matrix form as

$$X = C \circ B$$

where $\circ$ denotes the entrywise product (*Schur product*) of two matrices. So, if B is positive semidefinite, then X is also positive semidefinite. This is an important fact in Lyapunov's theory of stability of dynamical systems [13]. Some generalizations of Lyapunov's equation where the solution is positive definite, and their importance, are discussed in [5, Ch.10].

The analogous question about the equation (L2) was considered, and partly answered, by M. K. Kwong [12]. We want to know for what real numbers t, the equation (L2) has a positive definite solution X whenever B is positive definite. Our arguments above show that this is the same as asking when the matrix Z with entries

$$z_{ij} = \frac{1}{\lambda_i^2 + \lambda_j^2 + t\lambda_i\lambda_j} \tag{4}$$

is positive definite (for any given positive numbers $\lambda_1, \ldots, \lambda_n$). Clearly the restriction $t > -2$ is necessary. By a somewhat intricate topological argument Kwong [12] showed that the condition $-2 < t \leq 2$ is sufficient to guarantee that the matrix (4) is positive definite for all orders n. (A simple proof of this statement is given in Section 2 below.) He showed also that when $n = 2$, the matrix (4) is positive definite for all $t > -2$, but further restrictions on t are needed for higher dimensions. The problem was solved completely in a paper of Bhatia and Parthasarathy [4, Thm 5.1]. We have :

Theorem 1.1 (Kwong, Bhatia-Parthasarathy) The $n \times n$ matrices Z defined by (4) are positive definite for all n and all positive numbers $\lambda_1, \ldots, \lambda_n$ if and only if $-2 < t \leq 2$.

The method of Bhatia and Parthasarathy [4] consists of showing that matrices such as (4) are congruent to others related to difference kernels, and then checking whether these kernels are positive definite. In the process this leads to interesting examples of positive definite functions which may be useful in other contexts. See, in particular, the monograph of F. Hiai and H. Kosaki [9] which shows numerous applications to operator theory discovered by these authors in a series of papers [7,8,11]. Our aim in this paper is to apply these ideas to the equations (L3)-(L5). We prove the following three theorems.

Theorem 1.2 If $t > -1$, then for all n and all positive numbers $\lambda_1, \ldots, \lambda_n$ the $n \times n$ matrices Y with entries

$$y_{ij} = \frac{1}{\lambda_i^3 + \lambda_j^3 + t(\lambda_i^2 \lambda_j + \lambda_i \lambda_j^2)} \tag{5}$$

are positive definite. The restriction $t > -1$ is necessary for any such matrix to be positive definite.

Theorem 1.3 If $t \geq 4$, then for all n and all positive numbers $\lambda_1, \ldots, \lambda_n$ the $n \times n$ matrices W with entries

$$w_{ij} = \frac{1}{\lambda_i^4 + \lambda_j^4 + t(\lambda_i^3 \lambda_j + \lambda_i \lambda_j^3) + 6\lambda_i^2 \lambda_j^2} \tag{6}$$

are positive definite. The restriction $t \geq 4$ is necessary in general, while the restriction $t > -4$ is necessary for any such matrix to be positive definite.

Theorem 1.4 If $-10 < t \leq 6$, then for all n and all positive numbers $\lambda_1, \ldots, \lambda_n$ the $n \times n$ matrices V with entries

$$v_{ij} = \frac{1}{\lambda_i^4 + \lambda_j^4 + 4(\lambda_i^3 \lambda_j + \lambda_i \lambda_j^3) + t\lambda_i^2 \lambda_j^2} \tag{7}$$

are positive definite. The restriction $t > -10$ is necessary for any such matrix to be positive definite, while the restriction $t \leq 6$ is necessary in general.

Note the differences in the nature of restrictions on t in the three theorems. The restrictions $t > -1$, $t > -4$, and $t > -10$, respectively, are necessary even in the case $n = 1$. For $n \geq 2$, in the absence of these restrictions the matrices in question have nonpositive diagonal entries and can not be positive definite.

Some parts of these statements can be proved by rather simple arguments. These are shown in Section 2. Complete proofs need more elaborate arguments given in Section 4. This separation brings out more interesting differences among the three results.

For brevity we use the term *positive matrix* to mean a positive semidefinite matrix in the rest of the paper.

2 The easier parts

Schur's Theorem says that the Schur product $A \circ B$ of two positive matrices is positive. Two matrices X and Y are said to be *congruent* if $Y = T^{\star}XT$ for some nonsingular matrix T. If X is positive, then so is every matrix congruent to it.

Let z_{ij} be the numbers defined by the equation (4). For all real numbers t we have

$$z_{ij} = \frac{1}{(\lambda_i + \lambda_j)^2} \quad \frac{1}{1 - \frac{(2-t)\lambda_i\lambda_j}{(\lambda_i+\lambda_j)^2}}.$$

Since $4\lambda_i\lambda_j/(\lambda_i + \lambda_j)^2 \leq 1$, we can expand this as an infinite series

$$z_{ij} = \frac{1}{(\lambda_i + \lambda_j)^2} \sum_{n=0}^{\infty} (2-t)^n \frac{\lambda_i^n \lambda_j^n}{(\lambda_i + \lambda_j)^{2n}}, \tag{8}$$

which is convergent provided $|2-t| < 4$. The matrix with entries $1/(\lambda_i+\lambda_j)^2$ is positive being the Schur product of the Cauchy matrix with itself. The matrix with entries $\lambda_i\lambda_j/(\lambda_i + \lambda_j)^2$ is congruent to this matrix (via a congruence by the diagonal matrix $T = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$). Each n-fold Schur product of this matrix with itself is again positive. Hence, if $t < 2$, the matrix Z is positive, being a sum of several positive matrices. This proves one half of Theorem 1.1 first proved by Kwong [12]. (The case $t = 2$ follows by a limits argument.) It is interesting, and perhaps surprising, that the sufficient condition $-2 < t \leq 2$ turns out to be necessary as well.

In the same way we can write the numbers y_{ij} defined by (5) as

$$y_{ij} = \frac{1}{(\lambda_i + \lambda_j)^3} \quad \frac{1}{1 - \frac{(3-t)(\lambda_i^2\lambda_j+\lambda_i\lambda_j^2)}{(\lambda_i+\lambda_j)^3}}.$$

Since $4(\lambda_i^2\lambda_j + \lambda_i\lambda_j^2)/(\lambda_i + \lambda_j)^3 \leq 1$ this can be expanded as an infinite series

$$y_{ij} = \frac{1}{(\lambda_i + \lambda_j)^3} \sum_{n=0}^{\infty} (3-t)^n \left(\frac{\lambda_i^2\lambda_j + \lambda_i\lambda_j^2}{(\lambda_i + \lambda_j)^3} \right)^n, \tag{9}$$

which is convergent provided $|3 - t| < 4$; i.e., provided $-1 < t < 7$. Note that

$$\frac{\lambda_i^2\lambda_j + \lambda_i\lambda_j^2}{(\lambda_i + \lambda_j)^3} = \frac{\lambda_i\lambda_j}{(\lambda_i + \lambda_j)^2},$$

and we have observed that the matrix with these numbers as its ij entries is positive. By the same arguments as we used for the matrix Z we can conclude that the matrix Y is positive for $-1 < t \leq 3$. This is where the similarity with Z ends. The matrices Y turn out to be positive for $t > 3$ as well.

More interesting things happen when we apply these considerations to the matrix W defined by (6). Since

$$8(\lambda_i^3\lambda_j + \lambda_i\lambda_j^3) \leq (\lambda_i + \lambda_j)^4, \tag{10}$$

we have

$$w_{ij} = \frac{1}{(\lambda_i + \lambda_j)^4} \sum_{n=0}^{\infty} (4-t)^n \left(\frac{\lambda_i^3\lambda_j + \lambda_i\lambda_j^3}{(\lambda_i + \lambda_j)^4} \right)^n, \tag{11}$$

provided $|4-t| < 8$; i.e., provided $-4 < t < 12$. To proceed further we need to know whether the matrix G with entries

$$g_{ij} = \frac{\lambda_i^3\lambda_j + \lambda_i\lambda_j^3}{(\lambda_i + \lambda_j)^4} \tag{12}$$

is positive. From the inequality (10) it follows that any 2×2 matrix G with entries given by (12) is positive. Thus from the identity (11) we can conclude that for $-4 < t \leq 4$ the matrix W is positive in the special case $n = 2$. We will see in Section 4 that the matrix G is not always positive in general for all n. So this argument does not carry us further. We have to rely on other arguments to decide whether W is positive in all dimensions. It turns out that this is *not* always the case for $t < 4$ but is always the case for $t \geq 4$. In other words the matrix W is positive for $t \geq 4$, and for arbitrary n this restriction is necessary. For $n = 2$, the matrix W is positive for all $-4 < t \leq 4$ as well.

Finally let us consider the matrix V defined by (7). Since $16\lambda_i^2\lambda_j^2 \leq (\lambda_i + \lambda_j)^4$ we can write

$$v_{ij} = \frac{1}{(\lambda_i + \lambda_j)^4} \sum_{n=0}^{\infty} (6-t)^n \left(\frac{\lambda_i^2\lambda_j^2}{(\lambda_i + \lambda_j)^4} \right)^n, \tag{13}$$

provided $|6-t| < 16$; i.e., provided $-10 < t < 22$. Our arguments using the Cauchy matrix, congruence, and Schur products allow us to deduce from this series expansion that V is positive if $-10 < t \leq 6$. (This fact has been noted by Kwong [12] as well.) It turns out that in this case again the condition $-10 < t \leq 6$ is necessary to ensure the positivity of all matrices V for all orders n.

3 Some Fourier transforms

Fourier transforms of some functions are needed for our arguments. First we indicate how they are calculated using contour integrations. See [14, p.116] for the theory behind the method that we use. For a function $f \in L_1(\mathbb{R})$ we use the notation $\hat{f}$ for the function

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{i\xi x} dx. \tag{14}$$

If f is an even function, then

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) \cos \xi x \, dx. \tag{15}$$

Let f_1 be the function

$$f_1(x) = \frac{1}{\cosh x + t}, \quad -1 < t < 1. \tag{16}$$

The integral (15) corresponding to this can be calculated by the method of residues. Let

$$\varphi(z) = \frac{e^{i\xi z}}{\cosh z + t}. \tag{17}$$

If $z = x + iy$, then $\cosh z = \cosh x \cos y + i \sinh x \sin y$. So the poles of φ are at the points $z = i \arccos(-t) = i(\pi \pm \arccos t)$. Choose the rectangular contour $[-R, R, R+2\pi i, -R+2\pi i]$, where $R > 0$. There are two poles of φ inside this rectangle: $z_1 = i(\pi + \arccos t)$, $z_2 = i(\pi - \arccos t)$. Using the periodicity of the exponential function we get by the method of residues

$$\int_{-\infty}^{\infty} \frac{e^{i\xi x}}{\cosh x + t} dx = \frac{2\pi i}{1 - e^{-\xi 2\pi}} \left[\frac{e^{i\xi z_1}}{\sinh z_1} + \frac{e^{i\xi z_2}}{\sinh z_2} \right].$$

Substitute the values of z_1 and z_2. Use the relations $\sinh iz = i \sin z$ and $\sin(\pi \pm \arccos t) = \mp\sqrt{1-t^2}$. A little calculation leads to the result that

$$\hat{f}_1(\xi) = \frac{2\pi \sinh(\xi \arccos t)}{\sqrt{1-t^2} \, \sinh \xi\pi}. \tag{18}$$

is the Fourier transform of the function (16).

Now consider the function

$$f_2(x) = \frac{1}{\cosh x + t}, \quad t > 1. \tag{19}$$

In this case the function φ defined as in (17) has poles at $z = \pm \operatorname{arccosh} t + ik\pi$, where k is an odd integer. For large R, the rectangular contour we chose above has two of these poles inside it. These are $z_1 = \operatorname{arccosh} t + i\pi$ and $z_2 = -\operatorname{arccosh} t + i\pi$. Again by the method of residues

$$\int_{-\infty}^{\infty} \frac{e^{i\xi x}}{\cosh x + t} dx = \frac{2\pi i}{1 - e^{-\xi 2\pi}} \left[\frac{e^{i\xi z_1}}{\sinh z_1} + \frac{e^{i\xi z_2}}{\sinh z_2} \right].$$

Substitute the values of z_1 and z_2. For $t \geq 1$ we have $\operatorname{arccosh} t = \log(t + \sqrt{t^2 - 1})$ and $\sinh(\operatorname{arccosh} t) = \sqrt{t^2-1}$. A little calculation shows that the Fourier transform of the function f_2 in (19) is

$$\hat{f}_2(\xi) = \frac{2\pi \, \sin(\xi \operatorname{arccosh} t)}{\sqrt{t^2-1} \, \sinh \xi\pi}. \tag{20}$$

Now consider the function

$$f_3(x) = \frac{1}{\cosh x + \sigma}. \tag{21}$$

where σ is a complex number. For σ varying in the plane slit along $(-\infty, 1]$, f and its Fourier transform are analytic in σ. Hence we have

$$\hat{f}_3(\xi) = \begin{cases} \frac{2\pi \sinh(\xi \arccos \sigma)}{\sqrt{1-\sigma^2} \sinh \xi\pi} & \text{for } |\sigma| < 1 \\ \frac{2\pi \sin(\xi \operatorname{arccosh} \sigma)}{\sqrt{\sigma^2-1} \sinh \xi\pi} & \text{for } |\sigma| > 1, \sigma \notin (-\infty, 1] \end{cases} \tag{22}$$

If $\sigma \notin \mathbb{R}$, both formulas make sense and give the same value for $\hat{f}(\xi)$. The special case when σ is purely imaginary will be needed in our calculation. Let

$$f_4(x) = \frac{1}{\cosh x + ia}, \quad a \in \mathbb{R}. \tag{23}$$

Use the formula (22) and the relation $\arccos(ia) = \frac{\pi}{2} - i \operatorname{arcsinh} a$. A little calculation shows

$$\hat{f}_4(\xi) = \frac{\pi}{\sqrt{1+a^2}} \left[\frac{\cos(\xi \operatorname{arcsinh} a)}{\cosh \frac{\xi\pi}{2}} - i \frac{\sin(\xi \operatorname{arcsinh} a)}{\sinh \frac{\xi\pi}{2}} \right]. \tag{24}$$

(When $\sigma = 1$, we can obtain an expression for $\hat{f}_3(\xi)$ by taking limits in either of the formulas (22). We get $\hat{f}(\xi) = 2\pi\xi / \sinh \xi\pi$ in this case.)

4 Proofs of Theorems 1.2 - 1.4

Following Bhatia and Parthasarathy [1] we convert the question of positive definiteness of the matrices under consideration to that of positive definiteness of some functions on the real line.

Put $\lambda_i = e^{x_i}$, $x_i \in \mathbb{R}$. Then the expression (5) is equal to

$$\begin{aligned} y_{ij} &= \frac{1}{e^{3x_i} + e^{3x_j} + t\left(e^{2x_i + x_j} + e^{x_i + 2x_j}\right)} \\ &= \frac{1}{e^{3x_i/2}} \left(\frac{1}{e^{3(x_i - x_j)/2} + e^{3(x_j - x_i)/2} + t\left(e^{(x_i - x_j)/2} + e^{(x_j - x_i)/2}\right)} \right) \frac{1}{e^{3x_j/2}}. \end{aligned}$$

Thus the matrix Y is congruent to one with entries

$$\frac{1}{\cosh\left(\frac{3(x_i - x_j)}{2}\right) + t \cosh\left(\frac{x_i - x_j}{2}\right)}.$$

To say that all such matrices are positive definite for $t > -1$ is to say that the function

$$f(x) = \frac{1}{\cosh 3x + t \cosh x}$$

on the real line is positive definite for $t > -1$. Our argument in Section 2 shows that this is certainly the case for $-1 < t \le 3$. Using the identity $\cosh 3x = 4\cosh^3 x - 3\cosh x$ we have

$$f(x) = \frac{1}{4\cosh^3 x + (t-3)\cosh x}.$$

The positive definiteness of this function for $t \ge 3$ is equivalent to the following assertion.

Proposition 4.1 For all real numbers a, the function

$$f(x) = \frac{1}{\cosh^3 x + a^2 \cosh x} \tag{25}$$

is positive definite.

Proof. By a theorem of Bochner f is positive definite if and only if its Fourier transform $\hat{f}(\xi) \geq 0$. The function f in (25) can be expressed also as

$$f(x) = \frac{1}{a^2} \left[\frac{1}{\cosh\, x} - \text{Re}\, \frac{1}{\cosh\, x + ia} \right].$$

The Fourier transform of this can be found using (24); it is

$$\hat{f}(\xi) = \frac{\pi}{a^2 \cosh \frac{\xi\pi}{2}} \left[1 - \frac{\cos(\xi \operatorname{arcsinh}\, a)}{\sqrt{1+a^2}} \right].$$

It is clear that $\hat{f}(\xi) \geq 0$. ■

As explained above this proves Theorem 1.2.

Using the same argument we write the expression (6) as

$$w_{ij} = \frac{1}{e^{2x_i}} \left(\frac{1}{e^{2(x_i - x_j)} + e^{2(x_j - x_i)} + t\left(e^{x_i - x_j} + e^{x_j - x_i}\right) + 6} \right) \frac{1}{e^{2x_j}}.$$

The assertion of Theorem 1.3 is equivalent to the statement that the function

$$f(x) = \frac{1}{\cosh\, 2x + t\, \cosh\, x + 3}$$

is positive definite for $t \geq 4$ but not for $-4 < t < 4$. Using the identity $\cosh\, 2x = 2\cosh^2 x - 1$, we see that this is equivalent to the following :

Proposition 4.2 The function

$$f(x) = \frac{1}{\cosh^2\, x + t\, \cosh\, x + 1} \tag{26}$$

is positive definite for all $t \geq 2$ but not for any $t < 2$.

Proof. Let $t > 2$. Let α, β be the roots of the polynomial $x^2 + tx + 1$. Then α, β are negative real numbers and their product is 1. We can write the function f in (26) as

$$f(x) = \frac{1}{\sqrt{t^2 - 4}} \left[\frac{1}{\cosh\, x + a} - \frac{1}{\cosh\, x + b} \right] \tag{27}$$

where $0 < a < 1$, and $b = \frac{1}{a} > 1$. (The numbers a, b are the negatives of α and β.) The Fourier transform of the expression in the square brackets in (27) is calculated using (22). This is equal to

$$\begin{aligned} &\frac{2\pi}{\sinh \xi\pi} \left[\frac{\sinh(\xi \arccos\, a)}{\sqrt{1-a^2}} - \frac{\sin(\xi \operatorname{arccosh}\, b)}{\sqrt{b^2-1}} \right] \\ &= \frac{2\pi}{\sinh\, \xi\pi} \frac{1}{\sqrt{1-a^2}} \left[\sinh(\xi \arccos\, a) - a\, \sin(\xi \operatorname{arccosh} \frac{1}{a}) \right]. \end{aligned} \tag{28}$$

We want to show that this is nonnegative for all ξ. As a function of ξ this expression is even. So it suffices to show that for $\xi \geq 0$ the expression inside the square brackets on the right hand side of (28) is nonnegative.

The function $g(a) = \arccos a - a \,\mathrm{arccosh}\frac{1}{a}$ has the following properties: $g(0) = \pi/2$, $g(1) = 0$ and $g'(a) = -\mathrm{arccosh}\frac{1}{a} < 0$. Hence $g(a) > 0$ for $0 < a < 1$. So for all $\xi \geq 0$ we have

$$\sinh(\xi \arccos a) \geq \xi \arccos a \geq a\, \xi\, \mathrm{arccosh}\frac{1}{a} \geq a\, \sin(\xi\, \mathrm{arccosh}\frac{1}{a}).$$

This proves that the expression (28) is nonnegative. So the function (26) is positive definite for all $t > 2$. By continuity this is true for $t = 2$ as well.

When $t = 0$, the function f in (26) reduces to

$$f(x) = \frac{1}{\cosh^2 x + 1} = -\mathrm{Im}\, \frac{1}{\cosh x + i}.$$

The Fourier transform of this function can be read off from (24). It is

$$\hat{f}(\xi) = \frac{\pi \sin(\xi\ \mathrm{arcsinh}\ 1)}{\sqrt{2} \sinh\frac{\xi\pi}{2}}.$$

This function is negative for some ξ, and hence f is not positive definite in this case.

More elaborate arguments are needed to decide the case $-2 < t < 2$, $t \neq 0$. In this case let

$$\alpha = \frac{-t + i\sqrt{4 - t^2}}{2}.$$

Then α and $\beta = \bar{\alpha}$ are the two roots of the polynomial $x^2 + tx + 1$. They are complex numbers of modulus 1. We have

$$\alpha = e^{i\theta} = \cos\theta + i \sin\theta, \quad 0 < \theta < \pi. \tag{29}$$

The case $\theta = \pi/2$ corresponds to $t = 0$ which we have already settled. The function f in (26) is equal to

$$f(x) = \frac{2}{\sqrt{4 - t^2}}\, \mathrm{Im}\, \frac{1}{\cosh x - \alpha}.$$

We wish to prove that if

$$g(x) = \frac{1}{\cosh x - \alpha}$$

where α is as in (29), then the imaginary part of $\hat{g}(\xi)$ is negative for some ξ.

Use the formula (22) valid by an analytic continuation for all points σ on the unit circle except -1. Our problem reduces to showing that for every point $\alpha \neq \pm 1$ on the unit circle the expression

$$\mathrm{Im}\, \frac{\sinh(\xi \arccos \alpha)}{\sqrt{1 - \alpha^2}} \tag{30}$$

is negative for some values of $\xi > 0$.

In the next few lines we write the principal values of the functions involved. A small calculation shows that for α as in (29)

$$\sqrt{1-\alpha^2} = \sqrt{\sin\,\theta + \sin^2\,\theta} - i\ \mathrm{sgn}(\sin 2\theta)\sqrt{\sin\theta - \sin^2\theta}, \tag{31}$$

and

$$\sinh(\xi\ \arccos\,\alpha) = -i\sin\left(\xi\ \log(\alpha + i\sqrt{1-\alpha^2})\right). \tag{32}$$

Let $\gamma = \alpha + i\sqrt{1-\alpha^2}$. A calculation shows

$$\gamma = \cos\,\theta + \mathrm{sgn}(\sin 2\theta)\sqrt{\sin\theta - \sin^2\theta} + i(\sin\theta + \sqrt{\sin\theta + \sin^2\theta}).$$

The conditions on θ guarantee that $\eta = \log\gamma$ is not a real number. The quantity (30) is equal to

$$\mathrm{Im}\,\frac{-i\sin(\xi\eta)}{\sqrt{1-\alpha^2}} = -\mathrm{Re}\,\frac{\sin(\xi\eta)}{\sqrt{1-\alpha^2}}. \tag{33}$$

Mapping properties of the function $\sin$ from the z-plane into the w-plane are well-known. This function maps the strip $|\mathrm{Re}\ z| < \pi/2$ bijectively onto the w-plane slit along a part of the real line $(-\infty,-1]\cup[1,\infty)$. Horizontal lines in the z-plane are mapped onto ellipses with major axes contained in the u-axis and vertical lines are mapped onto hyperbolas with vertices on the u-axis in the w-plane. From these properties it is easy to see that the quantity (33) is negative for some values of ξ.

This completes the proof of Proposition 4.2. and that, in turn, proves Theorem 1.3. ■

These arguments can be repeated *mutatis mutandis* to prove Theorem 1.4. A little calculation shows that the statement of Theorem 1.4 is equivalent to the assertion that the function

$$\frac{1}{2\cosh 2x + 8\cosh\,x + t}$$

is positive definite if and only if $-10 < t \le 6$. This is the same as saying that the function

$$f(x) = \frac{1}{\cosh\,2x + 4\cosh\,x + t}$$

is positive definite if and only if $-5 < t \le 3$. Using the identity $\cosh\,2x = 2\cosh^2 x - 1$, we see that this is equivalent to the assertion:

Proposition 4.3 The function

$$f(x) = \frac{1}{\cosh^2 x + 2\cosh\,x + t} \tag{34}$$

is positive definite if and only if $-3 < t \le 1$.

Proof. The arguments in Section 2 show that the function (34) is positive definite if $-3 < t \le 1$. Let $t > 1$ and let

$$\alpha = -1 + i\sqrt{t-1}.$$

Then α and $\beta = \bar{\alpha}$ are the two roots of the polynomial $x^2 + 2x + t$. We have $|\alpha|^2 = t > 1$. As in the proof of Proposition 4.2 our problem reduces to showing that when α is a complex number with $|\alpha| > 1$, and α is not real, then the imaginary part of the Fourier transform of the function

$$f(x) = \frac{1}{\cosh x - \alpha}$$

is negative at some points. The proof of this is similar to that given in Proposition 4.2. ■

In Section 2 we observed that when $n = 2$, the matrix G defined by (12) is positive. This matrix would have been positive for all n if the function

$$f(x) = \frac{\cosh 2x}{\cosh^4 x} = \frac{2}{\cosh^2 x} - \frac{1}{\cosh^4 x}$$

were positive definite. The Fourier transform of the function $1/\cosh^r x$ is known [9, p.138]. Using that formula we see that

$$\hat{f}(\xi) = 2 \left| \Gamma \left(1 + \frac{i\xi}{2} \right) \right|^2 \left(1 - \frac{1}{3} \left| 1 + \frac{i\xi}{2} \right|^2 \right).$$

This is negative for large ξ.

5 Operator Inequalities

Each of the positivity results in Section 2 leads to inequalities for norms of operators. Let $\|A\|$ stand for the norm of A as a linear operator on the Hilbert space $\mathbb{C}^n$, and $|||A|||$ for any unitarily invariant norm — one that has the property $|||A||| = |||UAV|||$ for any unitary matrices U, V, and is normalised so that $|||A||| = ||A||$ for any operator A of rank one. If A is positive, then for every B we have [10, p.343]

$$|||A \circ B||| \leq \max_t a_{ii} |||B|||. \tag{35}$$

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

$$|||A^\nu X B^{1-\nu} + A^{1-\nu} X B^\nu||| \leq |||AX + XB|||. \tag{36}$$

For the norm $||.||$ alone this was proved by E. Heinz [6], and used to derive many important results in perturbation theory. The general case, its history, and relations to other inequalities may be found in [1, Chapter IX].

In [4] Bhatia and Parthasarathy showed that for any positive real numbers $\lambda_1, \ldots, \lambda_n$ and for $-1 < \alpha < 1$, the matrix P with entries

$$p_{ij} = \frac{\lambda_i^\alpha + \lambda_j^\alpha}{\lambda_i + \lambda_j} \tag{37}$$

is positive and used this fact and (35) to give a proof of (36). Following this method some quaint generalizations of (36) may be obtained using the matrices in (8), (9) and (13). For example, we can prove that

$$(1+t)\,|||A^\nu XB^{1-\nu}+A^{1-\nu}XB^\nu|||\le |||AX+XB+t\left(A^{2/3}XB^{1/3}+A^{1/3}XB^{2/3}\right)|||, \quad (38)$$

for $-1<t\le 3$ and $\frac{1}{3}\le\nu\le\frac{2}{3}$.

To see this write for $-1<\alpha<1$ and $-1<t\le 3$, the identity

$$\frac{\lambda_i^\alpha+\lambda_j^\alpha}{\lambda_i^3+\lambda_j^3+t(\lambda_i^2\lambda_j+\lambda_i\lambda_j^2)}=\frac{\lambda_i^\alpha+\lambda_j^\alpha}{(\lambda_i+\lambda_j)}\frac{1}{(\lambda_i+\lambda_j)^2}\sum_{n=0}^\infty(3-t)^n\left(\frac{\lambda_i^2\lambda_j+\lambda_i\lambda_j^2}{(\lambda_i+\lambda_j)^3}\right)^n. \quad (39)$$

The matrix whose ij entries are given by (39) is a sum of Schur products of several positive matrices. Hence it is positive. Multiply it on the left and the right by a diagonal matrix $\mathrm{diag}\,(\lambda_1^r,\ldots,\lambda_n^r)$ where $1<r<2$. Then put $\alpha=3-2r$. This shows that the matrix with its ij entries given by

$$(1+t)\frac{\lambda_i^{3-r}\lambda_j^r+\lambda_i^r\lambda_j^{3-r}}{\lambda_i^3+\lambda_j^3+t(\lambda_i^2\lambda_j+\lambda_i\lambda_j^2)},\quad -1<t\le 3 \quad (40)$$

is positive and all its diagonal entries are equal to 1. Using the arguments in [2] one gets from this the inequality

$$(1+t)\,|||A^rXB^{3-r}+A^{3-r}XB^r|||\le |||A^3X+XB^3+t(A^2XB+AXB^2)|||,$$

for $-1<t\le 3$, $1<r<2$. Replace A and B by their cube roots and put $\nu=r/3$. This gives the inequality (38).

The choice $t=0$ and $\nu=1/2$ reduces (38) to the arithmetic-geometric mean inequality

$$|||A^{1/2}XB^{1/2}|||\le\frac{1}{2}|||AX+XB|||,$$

first proved by Bhatia and Davis [2]; the special case $X=I$ having been proved earlier by Bhatia and Kittaneh [3].

Similar inequalities may be obtained by applying these considerations to (8) and to (13). The first of these has been noted by X. Zhan [15]. One could also use the Schur product of the matrix (37) with itself. For example, one can write

$$\frac{(\lambda_i^\alpha+\lambda_j^\alpha)^2}{\lambda_i^2+\lambda_j^2+t\lambda_i\lambda_j}=\left(\frac{\lambda_i^\alpha+\lambda_j^\alpha}{\lambda_i+\lambda_j}\right)^2\sum_{n=0}^\infty(2-t)^n\frac{\lambda_i^n\lambda_j^n}{(\lambda_i+\lambda_j)^{2n}}$$

for $-1<\alpha<1$ and $-2<t\le 2$ and get other operator inequalities.

The full power of this method has been exploited by Hiai and Kosaki in a series of interesting papers [7,8,11] and in the monograph [9] to which the reader should turn for a large collection of operator means inequalities.

The first author thanks the Department of Mathematics, Kuwait University, for their generous hospitality during a visit in January 2004 when a part of this work was done.

References :

1. R. Bhatia, *Matrix Analysis,* Springer, 1997.
2. R. Bhatia and C. Davis, *More matrix forms of the arithmetic-geometric mean inequality,* SIAM J. Matrix Anal. Appl., 14 (1993) 132-136.
3. R. Bhatia and F. Kittaneh, *On the singular values of a product of operators,* SIAM J. Matrix Anal. Appl., 11 (1990) 272-277.
4. R. Bhatia and K. R. Parthasarathy, *Positive definite functions and operator inequalities,* Bull. London Math. Soc., 32 (2000) 214-228.
5. S. K. Godunov, *Modern Aspects of Linear Algbera,* Translations of Mathematical Monographs Vol.175, American Mathematical Society, 1998.
6. E. Heinz, *Beiträge zur Störungstheorie der Spektralzerlegung,* Math. Ann., 123 (1951) 415-438.
7. F. Hiai and H. Kosaki, *Comparison of various means for operators,* J. Funct. Anal., 163 (1999) 300-323.
8. F. Hiai and H. Kosaki, *Means of matrices and comparison of their norms,* Indiana Univ. Math. J., 48 (1999) 899-936.
9. F. Hiai and H. Kosaki, *Means of Hilbert Space Operators,* Lecture Notes in Mathematics 1820, Springer, 2003.
10. R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis,* Cambridge Univ. Press, 1990.
11. H. Kosaki, *Arithmetic-geometric mean and related inequalities for operators,* J. Funct. Anal., 156 (1998) 429-451.
12. M. K. Kwong, *On the definiteness of the solutions of certain matrix equations,* Linear Algebra Appl., 108 (1988) 177-197.
13. A. M. Lyapunov, *The general problem of the stability of motion,* English translation, Taylor and Francis, 1992.
14. D. S. Mitrinovic and J. D. Keckic, *The Cauchy Method of Residues,* D. Riedel Publishing Co., 1984.
15. X. Zhan, *Inequalities for unitarily invariant norms,* SIAM J. Matrix Anal. Appl., 20 (1999) 466-470.