J. Math. Study
doi: 10.4208/jms.v49n4.16.04

Vol. **49**, No. 4, pp. 393-428
December 2016

A Survey of the Matrix Exponential Formulae with Some Applications

Baoying Zheng[1],Lin Zhang[1], Minhyung Cho[2,*]and Junde Wu[3]

[1] *Institute of Mathematics, Hangzhou Dianzi University, Hangzhou 310018, China*
[2] *Department of Applied Mathematics, Kumoh National Institute of Technology, Kyungbuk 39177, Korea*
[3] *School of Mathematical Sciences, Zhejiang University, Hangzhou 310027, China*

Received June 27, 2016; Accepted August 25, 2016

Abstract. The matrix exponential formulae is a very important tool in diverse fields of mathematical physics, in particular, it is also useful in studying quantum information theory. In this paper, we survey results related to matrix exponential thoroughly. The purpose of the article is pedagogical.

AMS subject classifications: 40C05

Key words: Matrix, exponential formula, quantum information.

1 Introduction and preliminary

The matrix exponential formulae have applications in diverse fields in mathematical physics. For example, one can apply it to prove the famous strong subadditivity of the von Neumann entropy, and the Bessis-Moussa-Villani conjecture [16]. We use these results to study also some interesting quantum informational problems [24–29]. In this paper, we present a detailed survey, where some emphasis is put on the Lie-Trotter-Suzuki product formulae, Thompson formula, Wasin-So formula, Stahl's theorem, Peierls- Bogoliubov inequality, Golden-Thompson inequality, reverse inequality to Golden-Thompson type inequalities, trace inequality in quantum information theory, Itzykson-Zuber integral formula, etc. We just collect similar results from various places and put together. Although there are no new results in this article, but some detailed and elementary proofs of partial results are still included for completeness. Firstly, we state some fundamental knowledge of matrix exponential map, which is based on [11].

*Corresponding author. *Email address:* `mignon@kumoh.ac.kr` (M. Cho)

1.1 One-parameter groups of linear transformations

In this section, we show how one-parameter groups of linear transformation of a vector space can be described using the exponential map on matrices. Let V be a finite dimensional vector space, $\mathrm{End}(V)$ denote the algebra of linear maps from V to itself, and $\mathrm{GL}(V)$ denote the group of invertible linear maps from V to itself. The usual name for $\mathrm{GL}(V)$ is the *general linear group* of V. If $V=\mathbb{K}^n$, then $\mathrm{End}(V)=M_n(\mathbb{K})$, the $n\times n$ matrices over field $\mathbb{K}$, where $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$, and $\mathrm{GL}(V)=\mathrm{GL}_n(\mathbb{K})$, the matrices with non-vanishing determinants. Let $(V,\|*\|)$ be a normed space, where the norm is defined by

$$\|A\|\stackrel{\text{def}}{=}\sup\left\{\frac{\|Av\|}{\|v\|}:0\neq v\in V\right\},\quad A\in\mathrm{End}(V). \tag{1.1}$$

Definition 1.1. A *one-parameter group of linear transformations* of V is a continuous homomorphism

$$M:\mathbb{R}\longrightarrow\mathrm{GL}(V).$$

That is, $M(t)$ is a collection of linear maps such that

(i) $M(0)=\mathbb{1}_V$,

(ii) $M(s)M(t)=M(s+t)\quad\forall s,t\in\mathbb{R}$,

(iii) $M(t)$ depends continuously on t.

Remark 1.1. (a) For $A\in\mathrm{End}(V)$ and $r>0$, set

$$B_r(A)\stackrel{\text{def}}{=}\Big\{X\in\mathrm{End}(V):\|X-A\|<r\Big\}.$$

The Neumann formula:

$$(\mathbb{1}_V-A)^{-1}=\sum_{n=0}^{\infty}A^n \tag{1.2}$$

valid for $A\in B_1(0)$ shows that

$$(B_r(\mathbb{1}_V))^{-1}\subseteq B_\alpha(\mathbb{1}_V) \tag{1.3}$$

with $\alpha=\frac{r}{1-r}$. Similarly, the formula

$$(\mathbb{1}_V+A)(\mathbb{1}_V+B)=\mathbb{1}_V+A+B+AB \tag{1.4}$$

shows

$$B_r(\mathbb{1}_V)B_s(\mathbb{1}_V)\subseteq B_{r+s+rs}(\mathbb{1}_V).$$

(b) For $A\in\mathrm{L}(V)$, define

$$\exp(A)\stackrel{\text{def}}{=}\sum_{n=0}^{\infty}\frac{A^n}{n!}. \tag{1.5}$$

Since $\| A^n \| \leqslant \| A \|^n$, we see, by the standard estimates in the exponential series, that the series defining $\exp(A)$ converges absolutely for all A and uniformly on $B_r(0)$. Hence exp defines a smooth, in fact analytic, map from $\mathrm{End}(V)$ to itself. We will see shortly that in fact $\exp(A) \in \mathrm{GL}(V)$.

Proposition 1.1. ([11])*If $A,B \in \mathrm{End}(V)$ such that $[A,B] = AB - BA = 0$, then*

$$\exp(A+B) = \exp(A)\exp(B). \tag{1.6}$$

Proof. Computing formally, we have

$$\begin{aligned}\exp(A)\exp(B) &= \left(\sum_{i=0}^{\infty}\frac{A^i}{i!}\right)\left(\sum_{j=0}^{\infty}\frac{B^j}{j!}\right) = \sum_{i,j=0}^{\infty}\frac{A^iB^j}{i!j!} \\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\left(\sum_{i+j=n}\frac{n!}{i!j!}A^iB^j\right) = \sum_{n=0}^{\infty}\frac{1}{n!}\left(\sum_{k=0}^{n}\binom{n}{k}A^kB^{n-k}\right).\end{aligned}$$

If A and B commute, the familiar binomial formula applies and says

$$(A+B)^n = \sum_{k=0}^{n}\binom{n}{k}A^kB^{n-k}.$$

Substituting this in our formula for $\exp(A)\exp(B)$, and noting that all manipulations are valid because the series converge absolutely, we see the proposition follows. ☐

Corollary 1.1. ([11]) *For any $A \in \mathrm{End}(V)$, the map $t \mapsto \exp(tA)$ is a one-parameter group of linear transformations on V. In particular,*

$$\exp(A) \in \mathrm{GL}(V) \quad \text{and} \quad (\exp(A))^{-1} = \exp(-A).$$

Proof. Since for any real numbers s and t, the matrices sA and tA commute with one another, this corollary follows immediately from the above Proposition 1.1 ☐

In what follows, Theorem 1.1 is the converse of Corollary 1.1. Since for $v \in V$, we have

$$\exp(tA)v = v + tAv + \sum_{n=2}^{\infty}\frac{t^nA^nv}{n!}.$$

Therefore, we have

$$Av = \lim_{t\to 0}\frac{M(t)v - v}{t} = \frac{\mathrm{d}}{\mathrm{d}t}(M(t)v)\big|_{t=0}. \tag{1.7}$$

Proposition 1.2. ([11]) *For sufficiently small $r > 0$, the map* exp *takes $B_r(0)$ bijectively onto an open neighborhood of $\mathbb{1}_V$ in $\mathrm{GL}(V)$. One has $\exp(B_r(0)) \subseteq B_s(\mathbb{1}_V)$, where $s = e^r - 1$.*

Proof. Let $D\exp_A$ be the differential of exp at A. It is a linear map from $\mathrm{End}(V)$ to $\mathrm{End}(V)$ defined by

$$D\exp_A(B) = \lim_{t\to 0}\frac{\exp(A+tB)-\exp(A)}{t}.$$

From the definition of exp, it is easy to compute that

$$D\exp_0(B) = B.$$

That is $D\exp_0$ is the identity map on $\mathrm{End}(V)$. In particular $D\exp_0$ is invertible. Therefore the first statement of the proposition follows from the *Inverse Function Theorem*. The inclusion $\exp(B_r(0)) \subseteq B_s(\mathbb{1}_V)$ follows from the obvious termwise estimation of $\exp(A) - \mathbb{1}_V$. ∎

Remark 1.2. If one defines

$$\log(\mathbb{1}_V - A) = -\sum_{n=1}^{\infty}\frac{A^n}{n}, \tag{1.8}$$

then just as for real numbers, one sees this series converges absolutely for $\|A\| < 1$. Further, for all $B \in B_1(\mathbb{1}_V)$, one has

$$\exp(\log B) = B. \tag{1.9}$$

This formula is known in the scalar case, and this implies that in fact it is an identity in absolutely convergent power series, whence it follows in the matrix case.

Proposition 1.3. ([11]) *Choose an $r < \log 2$, and let $T \in \exp(B_r(0))$, say $T = \exp(A)$. Then the transformation $S = \exp(\frac{A}{2})$ is a square root of T; that is, $S^2 = T$. Moreover, S is the unique square root of T contained in $\exp(B_r(0))$.*

Proof. That $S^2 = T$ follows directly from Proposition 1.1. It is only necessary to prove the uniqueness of S. From Proposition 1.2, we see that our restriction on r implies $\exp(B_r(0)) \subseteq B_1(\mathbb{1}_V)$. Hence it will suffice to show that if A, B are distinct linear maps of norm less than 1, then $(\mathbb{1}_V + A)^2 \neq (\mathbb{1}_V + B)^2$. Suppose the contrary. Then expanding the squares, canceling the $\mathbb{1}_V$'s and transposing, we find the equation:

$$2(A-B) = B^2 - A^2 = B(B-A) + (B-A)A.$$

Taking norms yields

$$2\|A-B\| \leqslant \|B\|\,\|B-A\| + \|B-A\|\,\|A\| = \Big(\|B\| + \|A\|\Big)\|B-A\|.$$

This implies either $\|A-B\| = 0$, which is false since $A \neq B$, or $\|A\| + \|B\| \geqslant 2$, which is false since both $\|A\|$ and $\|B\|$ are less than 1. This contradiction establishes the uniqueness of S. ∎

Theorem 1.1. ([11]) *Every one-parameter group M of linear transformations of V has the form:*

$$M(t)=\exp(tA) \tag{1.10}$$

for some $A\in\mathrm{End}(V)$.

The transformation A is called the *infinitesimal generator* of the group $t\mapsto\exp(tA)$.

Proof. Let $t\mapsto M(t)$ be a continuous one-parameter group in $\mathrm{GL}(V)$. Since $M(0)=\mathbb{1}_V$, if we specify $r>0$, we may, by continuity, find an $\varepsilon>0$ such that $M(t)\in\exp(B_r(0))$ for $|t|\leqslant\varepsilon$. We take $r<\log 2$. Write

$$M(\varepsilon)=\exp(A_1)$$

for appropriate $A_1\in B_r(0)$. If we set

$$A=\frac{1}{\varepsilon}A_1,$$

then $M(\varepsilon)=\exp(\varepsilon A)$. The transformations $M(\varepsilon/2)$ and $\exp(\varepsilon A/2)$ are the both square roots of $M(t)$ lying in $\exp(B_r(0))$. By Proposition 1.3 we conclude

$$M(\varepsilon/2)=\exp(\varepsilon A/2).$$

An obvious induction using Proposition 1.3 shows that

$$M(2^{-n}\varepsilon)=\exp(2^{-n}\varepsilon A)$$

for all positive integers n. Taking mth powers, we conclude

$$M(m2^{-n}\varepsilon)=\exp(m2^{-n}\varepsilon A)$$

for all integers m and n. Since the numbers $m2^{-n}\varepsilon$ are dense in $\mathbb{R}$, the Theorem follows by continuity. $\square$

Proposition 1.4. ([24]) *For every $T\in\mathrm{GL}(\mathbb{C}^d)$, there exists $M\in\mathrm{End}(\mathbb{C}^d)$ such that $T=e^M$.*

Proof. It is easy to show that if T is a diagonalizable matrix, then the conclusion is true. For a general case, we prove it in the following two steps:

Case 1. There is a sequence of diagonalizable matrices T_k satisfying that

(i) $\lim_k T_k=T$,

(ii) If $T_k=e^{M_k}$, then there is a constant $c>0$ such that $\|M_k\|\leqslant c$ holds for every k.

Now we show that the *existence* of T_k. Consider the Jordan decomposition of T for $T=PJP^{-1}$. Let t_i be the diagonal entries of J. Note that T is an invertible matrix, so $t_i\neq 0$ for every $1\leqslant i\leqslant d$. Let

$$T_k:=P(J+\Lambda_k)P^{-1},$$

where $\Lambda_k:=\mathrm{diag}(\lambda_1^k,\dots,\lambda_d^k)$. Then T_k meets the conditions (i) and (ii) in **Case 1** if

(a) $\lim_k \lambda_i^k = 0$ for $i = 1,\ldots,d$;

(b) $t_i + \lambda_i^k$ are all different when i runs from 1 to d for every given k. Thus T_k has d different eigenvalues $t_i + \lambda_i^k$, and of course T_k is diagonalizable;

(c) there is a constant c such that $\left|\ln(t_i + \lambda_i^k)\right| \leqslant c$ for every k and i. Note that if (b) is true, then $\|M_k\| = \max_i \left|\ln(t_i + \lambda_i^k)\right|$.

The construction of λ_i^k satisfying (a)–(c) is described as follows: For any given k, let $\lambda_1^k = \frac{t_1}{k}$, and λ_j^k be one of $\frac{t_j}{k}, \frac{t_j}{k+1}, \ldots, \frac{t_j}{k+j}$ such that $t_i + \lambda_i^k \neq t_j + \lambda_j^k$ whenever $i < j$. Apparently (a) and (b) are satisfied. To check (c), we have

$$\left|\ln(t_i + \lambda_i^k)\right| = \left|\ln(t_i) + \ln(1 + \lambda_i^k / t_i)\right| \leqslant |\ln(t_i)| + \left|\ln(1 + \lambda_i^k / t_i)\right|,$$

taking $c = \max_i |\ln(t_i)| + \ln 2$ is enough. That is $\|M_k\| = \max_i \left|\ln(t_i + \lambda_i^k)\right| \leqslant c$ for all k.

Case 2. When **Case 1** holds, since the exponential function is a smooth and continuous function, so the image of the *compact* set, $\exp(B(0,c))$ must be closed, where $B(0,c)$ is the closed ball with radius c, thus the limit T of e^{M_k} is also in $\exp(B(0,c))$. This means that there exists $M \in B(0,c)$ such that $T = e^M$. ☐

The second proof. We note also that for arbitrary $T \in \mathrm{GL}(\mathbb{C}^d)$, there is a $M \in \mathrm{End}(\mathbb{C}^d)$ such that $T = e^M$. In fact, a matrix $A \in \mathrm{End}(\mathbb{C}^d)$ is called *nilpotent* if $A^k = 0$ for some positive integer k. A nilpotent matrix has trace zero, since zero is its only eigenvalue. A matrix $X \in \mathrm{End}(\mathbb{C}^d)$ is called *unipotent* if $X - \mathbb{1}$ is nilpotent. Note that a unipotent transformation is nonsingular and has determinant 1, since 1 is its only eigenvalue.

Let $A \in \mathrm{End}(\mathbb{C}^d)$ be nilpotent. Then $A^d = 0$ and for $z \in \mathbb{C}$, we have

$$e^{zA} = \mathbb{1} + Y,$$

where

$$Y = zA + \frac{z^2}{2!}A^2 + \cdots + \frac{z^{n-1}}{(n-1)!}A^{n-1}$$

is also nilpotent. Hence the matrix e^{zA} is unipotent and $z \mapsto e^{zA}$ is a *regular* homomorphism from the additive group $\mathbb{C}$ to $\mathrm{GL}(\mathbb{C}^d)$. Conversely, if $X = \mathbb{1} + Y \in \mathrm{GL}(\mathbb{C}^d)$ is unipotent, then $Y^d = 0$ and we define

$$\ln X = \sum_{k=1}^{n-1} (-1)^{k+1} \frac{1}{k} Y^k.$$

By the fact that any equation involving power series in a complex variable x that holds as an identity of absolutely convergent series when $|x| < r$ also holds as an identity of matrix power series in a matrix variable X, and these series converge absolutely in the matrix norm when $\|X\| < r$, we have

$$\exp(\ln(\mathbb{1} + A)) = \mathbb{1} + A.$$

Thus the exponential function is a *bijective* polynomial map from the nilpotent elements in $\mathrm{End}(\mathbb{C}^d)$ onto the unipotent elements in $\mathrm{GL}(\mathbb{C}^d)$, with inverse $X \mapsto \ln X$. $\square$

1.2 Properties of the exponential map

The map exp is the basic link between the linear structure on $\mathrm{End}(V)$ and the multiplicative structure on $\mathrm{GL}(V)$. We will describe some salient properties of this link.

Choose $r \in (0,\frac{1}{2}]$ such that exp is one-to-one on $B_r(0)$. Choose $r_1 < r$ so that if $A,B \in B_{r_1}(0)$, then $\exp(A)\exp(B)$ is contained in $\exp(B_r(0))$. Then we can write

$$\exp(A)\exp(B) = \exp(C) \tag{1.11}$$

for some $C \in B_r(0)$. The Inverse Function Theorem guarantees that C is a smooth (in fact analytic) function of A and B. There is a beautiful formula, the Campbell-Hausdorff formula, which expresses C as a universal power series in A and B. To develop this completely would take too long. We will just give the first two terms in the expression for C. These suffice for most purposes.

Proposition 1.5. ([11]) *Suppose A,B,C have norm at most $\frac{1}{2}$ and satisfy Eq.* (1.11)*. Then we have*

$$C = A + B + \frac{1}{2}[A,B] + S, \tag{1.12}$$

where the remainder term S satisfies

$$\|S\| \leqslant 65(\|A\| + \|B\|)^3. \tag{1.13}$$

Proof. We have

$$\exp(C) = \mathbb{1}_V + C + R_1(C), \tag{1.14}$$

where the remainder $R_1(C)$ is

$$R_1(C) = \sum_{n=2}^{\infty} \frac{C^n}{n!}$$

and satisfies the obvious estimate

$$\|R_1(C)\| \leqslant \|C^2\| \left(\sum_{n=2}^{\infty} \frac{\|C\|^{n-2}}{n!} \right) \leqslant \|C\|^2$$

when $\|C\| \leqslant 1$, hence certainty when $\|C\| \leqslant \frac{1}{2}$. Similarly, we have

$$\exp(A)\exp(B) = \mathbb{1}_V + A + B + R_1(A,B), \tag{1.15}$$

where by rearrangement of the double sum

$$R_1(A,B)=\sum_{n=2}^{\infty}\frac{1}{n!}\left(\sum_{k=0}^{n}\binom{n}{k}A^kB^{n-k}\right).$$

Hence we have the estimate

$$\|R_1(A,B)\|\leqslant(\|A\|+\|B\|)^2\left(\sum_{n=2}^{\infty}\frac{(\|A\|+\|B\|)^{n-2}}{n!}\right)\leqslant(\|A\|+\|B\|)^2$$

when $\|A\|+\|B\|\leqslant 1$. Comparison Eq. (1.14) and Eq. (1.15), we see that Eq. (1.11) implies

$$C=A+B+R_1(A,B)-R_1(C). \tag{1.16}$$

Hence

$$\|C\|\leqslant\|A\|+\|B\|+(\|A\|+\|B\|)^2+\|C\|^2\leqslant 2(\|A\|+\|B\|)+\frac{1}{2}\|C\|$$

when A,B and C all have norm at most $\frac{1}{2}$. Thus

$$\|C\|\leqslant 4(\|A\|+\|B\|). \tag{1.17}$$

Returning to Eq. (1.16), we further find

$$\begin{aligned}\|C-(A+B)\|&\leqslant\|R_1(A,B)\|+\|R_1(C)\|\\&\leqslant(\|A\|+\|B\|)^2+[4(\|A\|+\|B\|)]^2\\&=17(\|A\|+\|B\|)^2.\end{aligned} \tag{1.18}$$

We now refine these estimates to second order. In analogy with Eq. (1.14), we have

$$\exp(C)=\mathbb{1}_V+C+\frac{C^2}{2}+R_2(C), \tag{1.19}$$

where

$$R_2(C)=\sum_{n=3}^{\infty}\frac{C^n}{n!}$$

is easily estimated by

$$\|R_2(C)\|\leqslant\frac{1}{3}\|C\|^3 \tag{1.20}$$

when $\|C\|\leqslant 1$. If we substitute Eq. (1.12) for C in Eq. (1.19), we obtain

$$\begin{aligned}\exp(C)&=\mathbb{1}_V+A+B+\frac{1}{2}[A,B]+S+\frac{1}{2}C^2+R_2(C)\\&=\mathbb{1}_V+A+B+\frac{1}{2}[A,B]+\frac{1}{2}(A+B)^2+T\\&=\mathbb{1}_V+A+B+\frac{1}{2}(A^2+2AB+B^2)+T,\end{aligned} \tag{1.21}$$

where

$$T=S+\frac{1}{2}\Big(C^2-(A+B)^2\Big)+R_2(C).$$

On the other hand, we have

$$\exp(A)\exp(B)=\mathbb{1}_V+A+B+\frac{1}{2}(A^2+2AB+B^2)+R_2(A,B), \tag{1.22}$$

where

$$R_2(A,B)=\sum_{n=3}^{\infty}\frac{1}{n!}\left(\sum_{k=0}^{\infty}\binom{n}{k}A^kB^{n-k}\right)$$

satisfies

$$\|R_2(A,B)\|\leqslant\frac{1}{3}(\|A\|+\|B\|)^3$$

when $\|A\|+\|B\|\leqslant 1$. Comparison of Eq. (1.21) and Eq. (1.22) in the light of Eq. (1.11) yields

$$S=R_2(A,B)+\frac{1}{2}\big((A+B)^2-C^2\big)-R_2(C).$$

Taking norms, we find

$$\begin{aligned}\|S\|&\leqslant \|R_2(A,B)\|+\|R_2(C)\|+\frac{1}{2}\Big(\|(A+B)(A+B-C)+(A+B-C)C\|\Big)\\&\leqslant\frac{1}{3}(\|A\|+\|B\|)^3+\frac{1}{2}(\|A\|+\|B\|+\|C\|)\|A+B-C\|+\frac{1}{3}\|C\|^3\\&\leqslant\frac{1}{3}(\|A\|+\|B\|)^3+\frac{5}{2}(\|A\|+\|B\|)\cdot 17(\|A\|+\|B\|)^2\\&\qquad+\frac{1}{3}(4(\|A\|+\|B\|))^3\leqslant 65(\|A\|+\|B\|)^3,\end{aligned} \tag{1.23}$$

as was to be shown. □

2 The Lie-Trotter-Suzuki product formulae

We will derive two main consequences of Proposition 1.5 in this section. These relate group operations in $\mathrm{GL}(V)$ to the linear operations in $\mathrm{End}(V)$, and are crucial ingredients in the proof of the main theorem that relates Lie algebras to Lie groups. Theorem 2.1 relates group multiplication in $\mathrm{GL}(V)$ to addition in $\mathrm{End}(V)$.

Theorem 2.1 (The Lie-Trotter product formula, [10])**.** *For any two matrices A and B,*

$$\lim_{n\to\infty}\left[\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right)\right]^n=\exp(A+B). \tag{2.1}$$

Proof. For n large enough, $\frac{A}{n}$ and $\frac{B}{n}$ will be close enough to the origin that formula Eq. (1.12) applies. We then have

$$\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right)=\exp(C_n),$$

where by estimate Eq. (1.18)

$$\left\|C_n-\frac{A+B}{n}\right\|\leqslant\frac{17}{n^2}(\|A\|+\|B\|).$$

Hence as $n\to\infty$, we have $nC_n\to A+B$. Since $\exp(nC_n)=[\exp(C_n)]^n$, Eq. (2.1) follows. $\square$

The second proof is given by Bhatia [4].

Proof. For any two matrices X,Y, and for $n=1,2,\ldots$, we have

$$X^n-Y^n=\sum_{j=0}^{n-1}X^{n-1-j}(X-Y)Y^j. \tag{2.2}$$

Using this, we obtain

$$\|X^n-Y^n\|\leqslant nM^{n-1}\|X-Y\|, \tag{2.3}$$

where $M\stackrel{\text{def}}{=}\max\{\|X\|,\|Y\|\}$. Now let

$$X_n=\exp\left(\frac{A+B}{n}\right),\quad Y_n=\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right).$$

Then

$$\exp\left(\frac{\|A\|+\|B\|}{n}\right)\geqslant\max\{\|X_m\|,\|Y_m\|\}.$$

From the power series expansion for the exponential function, we see that

$$\begin{aligned}X_n-Y_n=&\,\mathbb{1}+\frac{A+B}{n}+\frac{1}{2}\left(\frac{A+B}{n}\right)^2+\cdots\\&-\left\{\left[\mathbb{1}+\frac{A}{n}+\frac{1}{2}\left(\frac{A}{n}\right)^2+\cdots\right]\left[\mathbb{1}+\frac{B}{n}+\frac{1}{2}\left(\frac{B}{n}\right)^2+\cdots\right]\right\}\\=&\,O\left(\frac{1}{n^2}\right)\end{aligned} \tag{2.4}$$

for large n. Hence, using the inequality Eq. (2.3), we see that

$$\|X_n^n-Y_n^n\|\leqslant n\exp(\|A\|+\|B\|)O\left(\frac{1}{n^2}\right).$$

This goes to zero as $n\to\infty$. But $X_n^n=\exp(A+B)$ for all n. Hence,

$$\lim_{n\to\infty} Y_n^n=\exp(A+B).$$

This proves the theorem. □

Remark 2.1. For any two matrices X and Y, we have

$$\otimes^n X-\otimes^n Y=\sum_{j=0}^{n-1}(\otimes^{n-1-j}X)\otimes(X-Y)\otimes(\otimes^j Y). \tag{2.5}$$

Hence,

$$\|\otimes^n X-\otimes^n Y\|\leqslant nM^{n-1}\|X-Y\|, \tag{2.6}$$

where $M\overset{\text{def}}{=}\max\{\|X\|,\|Y\|\}$.

Theorem 2.2. ([10]) *For any two matrices A and B,*

$$\lim_{n\to\infty}\left[\exp\left(\frac{A}{2n}\right)\exp\left(\frac{B}{n}\right)\exp\left(\frac{A}{2n}\right)\right]^n=\exp(A+B). \tag{2.7}$$

Proof. Since

$$\begin{aligned}&\left[\exp\left(\frac{A}{2n}\right)\exp\left(\frac{B}{n}\right)\exp\left(\frac{A}{2n}\right)\right]^n\\ =&\exp\left(-\frac{A}{2n}\right)\left[\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right)\right]^n\exp\left(\frac{A}{2n}\right),\end{aligned}$$

it follows from taking the limit $n\to\infty$ that

$$\begin{aligned}&\lim_{n\to\infty}\left[\exp\left(\frac{A}{2n}\right)\exp\left(\frac{B}{n}\right)\exp\left(\frac{A}{2n}\right)\right]^n\\ =&\lim_{n\to\infty}\exp\left(-\frac{A}{2n}\right)\lim_{n\to\infty}\left[\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right)\right]^n\lim_{n\to\infty}\exp\left(\frac{A}{2n}\right)\\ =&\exp(A+B).\end{aligned}$$

Thus the desired conclusion is proved. □

As is well-known, an exponential function e^x is expressed by

$$e^x=\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n, \tag{2.8}$$

or

$$e^x=\sum_{k=0}^{\infty}\frac{x^k}{k!}. \tag{2.9}$$

The above two formulae give methods to calculate e^x numerically. The second expression is more convenient for such a purpose, because the convergence of the latter is better than the former.

Clearly

$$e^x = \left(\exp\left(\frac{x}{n}\right)\right)^n = \left[\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{x}{n}\right)^k\right]^n.$$

Let

$$e_{n,m}(x) = \left[\sum_{k=0}^{m}\frac{1}{k!}\left(\frac{x}{n}\right)^k\right]^n.$$

This indicates that

$$e^x = \lim_{n\to\infty} e_{n,m}(x) = \lim_{m\to\infty} e_{n,m}(x). \tag{2.10}$$

The case $m=1$ corresponds to the first expression and $n=1$ to the second one.

Is there a much more rapidly convergent expression for e^x? To answer this question, we try to unify or combine the above two formulae as follows. It is easy to evaluate (the upper bound of) the remainder $r_{n,m}(x)$ defined by $r_{n,m}(x) = e^x - e_{n,m}(x)$. In fact, using the generalized mean value theorem or Taylor's theorem, Suzuki obtained the following result:

Proposition 2.1. ([20]) *With the above notation, it holds that*

$$|e^x - e_{n,m}(x)| \leqslant \frac{1}{n^m}\frac{|x|^{m+1}}{(m+1)!}e^{|x|}.$$

Proof. Let $a = \exp(x/n) = \sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{x}{n}\right)^k$ and $b = \sum_{k=0}^{m}\frac{1}{k!}\left(\frac{x}{n}\right)^k$. Then

$$\begin{aligned}
|e^x - e_{n,m}(x)| &= |a^n - b^n| \leqslant n|a-b|(\max\{|a|,|b|\})^{n-1} \\
&\leqslant n|a-b|\left(\exp\left(\frac{|x|}{n}\right)\right)^{n-1}.
\end{aligned} \tag{2.11}$$

In what follows, we give an upper bound for $|a-b|$. Apparently,

$$\begin{aligned}
|a-b| &= \left|\sum_{k=m+1}^{\infty}\frac{1}{k!}\left(\frac{x}{n}\right)^k\right| = \frac{1}{(m+1)!}\left(\frac{|x|}{n}\right)^{m+1}\left|\sum_{k=m+1}^{\infty}\frac{(m+1)!}{k!}\left(\frac{x}{n}\right)^{k-m-1}\right| \\
&\leqslant \frac{1}{(m+1)!}\left(\frac{|x|}{n}\right)^{m+1}\sum_{k=m+1}^{\infty}\frac{(m+1)!}{k!}\left(\frac{|x|}{n}\right)^{k-m-1} \\
&\leqslant \frac{1}{(m+1)!}\left(\frac{|x|}{n}\right)^{m+1}\exp\left(\frac{|x|}{n}\right).
\end{aligned} \tag{2.12}$$

The desired inequality is proven. ∎

Proposition 2.2. ([20]) *For any set of operators* $\{A_j:1,\ldots,k\}$ *in a Banach algebra, we have*

$$\left\|\exp\left(\sum_{j=1}^{k}A_j\right)-\left(\prod_{j=1}^{k}\exp\left(\frac{1}{n}A_j\right)\right)^n\right\|\leqslant\frac{1}{n}\left(\sum_{j=1}^{k}\|A_j\|\right)^2\exp\left(\sum_{j=1}^{k}\|A_j\|\right). \tag{2.13}$$

Proof. It will be instructive to give here a brief proof of the present Proposition. If we set

$$X=\exp\left(\frac{1}{n}\sum_{j=1}^{k}A_j\right)\ \text{ and }\ Y=\prod_{j=1}^{k}\exp\left(\frac{1}{n}A_j\right), \tag{2.14}$$

hence

$$\|X\|,\|Y\|\leqslant\exp\left(\frac{1}{n}\sum_{j=1}^{k}\|A_j\|\right). \tag{2.15}$$

This implies that

$$\|X\|^{n-1-i}\|Y\|^i\leqslant\exp\left(\frac{n-1}{n}\sum_{j=1}^{k}\|A_j\|\right),\quad i=0,\ldots,n-1.$$

Since $e^x=\sum_{m=0}^{\infty}\frac{1}{m!}x^m$, it follows that, for $x>0$,

$$\begin{aligned}2(e^x-x-1)&=\sum_{m=2}^{\infty}\frac{2}{m!}x^m=x^2\sum_{m=2}^{\infty}\frac{2}{m!}x^{m-2}\\&\leqslant x^2\sum_{m=2}^{\infty}\frac{1}{(m-2)!}x^{m-2}=x^2e^x.\end{aligned} \tag{2.16}$$

Then we obtain

$$\begin{aligned}\|X^n-Y^n\|&\leqslant\|X-Y\|\left(\|X\|^{n-1}+\|X\|^{n-2}\|Y\|+\cdots+\|X\|\|Y\|^{n-2}+\|Y\|^{n-1}\right)\\&\leqslant n\|X-Y\|\exp\left(\frac{n-1}{n}\sum_{j=1}^{k}\|A_j\|\right).\end{aligned}$$

Next it is easy to show that

$$\begin{aligned}\|X-Y\|&\leqslant 2\left[\exp\left(\frac{1}{n}\sum_{j=1}^{k}\|A_j\|\right)-\left(1+\frac{1}{n}\sum_{j=1}^{k}\|A_j\|\right)\right]\\&\leqslant\frac{1}{n^2}\left(\sum_{j=1}^{k}\|A_j\|\right)^2\exp\left(\frac{1}{n}\sum_{j=1}^{k}\|A_j\|\right),\end{aligned} \tag{2.17}$$

where we have used the mean value theorem, implying the desired inequality. □

This yields the following theorem:

Theorem 2.3 (A generalized Lie-Trotter-Suzuki product formula, [4])**.** *For any k matrices $A_1,\ldots,A_k$,*

$$\lim_{n\to\infty}\left[\exp\left(\frac{A_1}{n}\right)\cdots\exp\left(\frac{A_k}{n}\right)\right]^n=\exp(A_1+\cdots+A_k). \tag{2.18}$$

Corollary 2.1. *If A,B are positive define matrices, then*

$$\lim_{n\to\infty}\left(A^{\frac{1}{2n}}B^{\frac{1}{n}}A^{\frac{1}{2n}}\right)^n=\exp(\log A+\log B). \tag{2.19}$$

Proof. In fact, replacing A,B with $\log A,\log B$, respectively, in Eq. (2.7) gives the desired conclusion. $\square$

Proposition 2.3. *For any two matrices A and B,*

$$\lim_{t\to 0}\left[\exp\left(\frac{tA}{2}\right)\exp(tB)\exp\left(\frac{tA}{2}\right)\right]^{\frac{1}{t}}=\exp(A+B). \tag{2.20}$$

Proposition 2.4. ([20]) *For any two matrices A and B, define a function f as follows:*

$$f(n):=\mathrm{Tr}\left(\left(e^{A/n}e^{B/n}\right)^n\right). \tag{2.21}$$

Then the function f is an even function of n.

Proof. We have

$$\begin{aligned}f(-n)=&\ \mathrm{Tr}\left(\left(e^{-A/n}e^{-B/n}\right)^{-n}\right)=\mathrm{Tr}\left(\left((e^{-A/n}e^{-B/n})^{-1}\right)^n\right)\\=&\ \mathrm{Tr}\left(\left(e^{B/n}e^{A/n}\right)^n\right)=\mathrm{Tr}\left(\left(e^{A/n}e^{B/n}\right)^n\right)=f(n),\end{aligned}$$

completing the proof. $\square$

Proposition 2.5. ([20]) *For any matrices A and B, we have*

$$\left\|e^{A+B}-\left(e^{A/2n}e^{B/n}e^{A/2n}\right)^n\right\|\leqslant\frac{1}{3n^2}(\|A\|+\|B\|)^3\exp(\|A\|+\|B\|). \tag{2.22}$$

Proof. By noting that

$$\begin{aligned}&\left\|e^{A+B}-\left(e^{A/2n}e^{B/n}e^{A/2n}\right)^n\right\|\\\leqslant&\ 2\left[\exp\left(\frac{\|A\|+\|B\|}{n}\right)-\left(1+\frac{\|A\|+\|B\|}{n}+\frac{(\|A\|+\|B\|)^2}{2n^2}\right)\right]\\\leqslant&\ \frac{1}{3n^3}(\|A\|+\|B\|)^3\exp\left(\frac{\|A\|+\|B\|}{n}\right).\end{aligned} \tag{2.23}$$

This indicates that the desired conclusion. $\square$

The following lemma is the direct consequence of the above proposition.

Proposition 2.6. ([20]) *For any matrices A and B, we have*

$$\left|\mathrm{Tr}\left(e^{A+B}\right)-\mathrm{Tr}\left(\left(e^{A/n}e^{B/n}\right)^n\right)\right|\leqslant\frac{d}{3n^2}(\|A\|+\|B\|)^3\exp(\|A\|+\|B\|),\tag{2.24}$$

where d denotes the dimensionality of the matrices A and B.

Note that $\mathrm{Tr}\left(\left(e^{A/n}e^{B/n}\right)^n\right)=\mathrm{Tr}\left(\left(e^{A/2n}e^{B/n}e^{A/2n}\right)^n\right)$.

Proposition 2.7. ([20]) *The symmetrized function $f_s(n)$, defined by*

$$f_s(n)=\mathrm{Tr}\left(\left[\exp\left(\frac{A_1}{2n}\right)\cdots\exp\left(\frac{A_{k-1}}{2n}\right)\exp\left(\frac{A_k}{n}\right)\exp\left(\frac{A_{k-1}}{2n}\right)\cdots\exp\left(\frac{A_1}{2n}\right)\right]^n\right),$$

is an even function of n.

Proposition 2.8. ([20]) *For any set of matrices $\{A_j:j=1,\ldots,k\}$, we have*

$$\begin{aligned}&\left\|\exp\left(\sum_{j=1}^k A_j\right)-\left[\exp\left(\frac{A_1}{2n}\right)\cdots\exp\left(\frac{A_{k-1}}{2n}\right)\exp\left(\frac{A_k}{n}\right)\exp\left(\frac{A_{k-1}}{2n}\right)\cdots\exp\left(\frac{A_1}{2n}\right)\right]^n\right\|\\&\leqslant\frac{1}{3n^2}\left(\sum_{j=1}^k\|A_j\|\right)^2\exp\left(\sum_{j=1}^k\|A_j\|\right)\end{aligned}$$

and

$$\left|\mathrm{Tr}\left(\exp\left(\sum_{j=1}^k A_j\right)\right)-f_s(n)\right|\leqslant\frac{d}{3n^2}\left(\sum_{j=1}^k\|A_j\|\right)^2\exp\left(\sum_{j=1}^k\|A_j\|\right).$$

Proposition 2.9. ([20]) *For any matrices A and B, we have*

$$\left\|e^{A+B}-\left(e^{A/n}e^{B/n}\right)^n\right\|\leqslant\frac{1}{2n}\|[A,B]\|\exp(\|A\|+\|B\|).$$

Proof. If we set

$$F(x)=e^{x(A+B)}-e^{xA}e^{xB},$$

then we obtain the following expression:

$$F(x)=\int_0^x\mathrm{d}t\int_0^t\mathrm{d}s\left(e^{tA}e^{(t-s)B}[B,A]e^{sB}e^{(x-t)(A+B)}\right)\tag{2.25}$$

using Kubo's identity,

$$[A,e^{tB}]=\int_0^t e^{(t-s)B}[A,B]e^{sB}\mathrm{d}s.$$

Therefore, we obtain the following inequality:

$$\|F(x)\|\leqslant\frac{x^2}{2}\|[A,B]\|e^{x(\|A\|+\|B\|)}.$$

Letting $x=1/n$, gives the desired inequality. $\square$

3 Thompson formula

In 1986, Thompson obtained the following result:

Theorem 3.1. ([21]) *Let A,B be two $n\times n$ Hermitian matrices. Then there exist two $n\times n$ unitary matrices U and V such that*

$$\exp(\mathrm{i}A)\exp(\mathrm{i}B)=\exp\left(\mathrm{i}UAU^\dagger+\mathrm{i}VBV^\dagger\right). \tag{3.1}$$

Later, in 2012, Antezana *et al.* have made extensions of this result to compact operators as well as to operators in an embeddable II_1 factor [1], and use it to study the optimal path in unitary group [2].

4 Matrix spectra

Theorem 4.1. ([13]) *Let $\sigma_i(i=1,\dots,N)$ and σ all be n-tuples with positive components in non-increasing order. The following statements are equivalent:*

- *(i) There exist matrices $A_i\in\mathrm{GL}_n(\mathbb{C})$ with singular spectra $\sigma_i=\sigma^\downarrow(A_i)$ and $\sigma=\sigma^\downarrow(A_1\cdots A_N)$.*
- *(i1) There exist Hermitian $n\times n$ matrices H_i with spectra $\lambda^\downarrow(H_i)=\log\sigma_i$ and $\lambda^\downarrow(H_1+\cdots+H_N)=\log\sigma$.*

Corollary 4.1. ([23]) *Let α,β,γ be three n-tuples of non-increasingly ordered real numbers. Then the following statements are equivalent:*

- *(i) There exist $n\times n$ Hermitian matrices H,K with $\lambda^\downarrow(H)=\alpha,\lambda^\downarrow(K)=\beta$ and $\lambda^\downarrow(H+K)=\gamma$.*
- *(ii) There exist $n\times n$ positive definite matrices A,B with $\lambda^\downarrow(A)=e^\alpha,\lambda^\downarrow(B)=e^\beta$ and $\lambda^\downarrow(AB)=e^\gamma$.*

5 Wasin-So formula

Theorem 5.1. ([23]) *Let A,B be two $n\times n$ Hermitian matrices. Then there exist two $n\times n$ unitary matrices U and V such that*

$$\exp\left(\frac{A}{2}\right)\exp(B)\exp\left(\frac{A}{2}\right)=\exp\left(UAU^\dagger+VBV^\dagger\right). \tag{5.1}$$

By taking trace on both sides on the above equation, we get

$$\mathrm{Tr}\left(e^Ae^B\right)=\mathrm{Tr}\left(e^{UAU^\dagger+VBV^\dagger}\right)=\mathrm{Tr}\left(e^{A+WBW^\dagger}\right), \tag{5.2}$$

where $W=U^\dagger V$. We conclude that for given Hermitian matrices A,B, there always exists unitary W such that

$$\mathrm{Tr}\left(e^A e^B\right)=\mathrm{Tr}\left(e^{A+WBW^\dagger}\right). \tag{5.3}$$

We can further study, based on this result, the following problem: Let $A=\mathrm{diag}(a_1,\ldots,a_n)$ with $a_1\geqslant\cdots\geqslant a_n$ and $B=\mathrm{diag}(b_1,\ldots,b_n)$ with $b_1\geqslant\cdots\geqslant b_n$. Consider the following function, defined over the unitary group $\mathrm{U}(n)$,

$$f(U):=\mathrm{Tr}\left(e^{A+UBU^\dagger}\right). \tag{5.4}$$

The following optimization problems are very interesting to be considered.

$$\max_{U\in\mathrm{U}(n)}\mathrm{Tr}\left(e^{A+UBU^\dagger}\right) \quad \text{and} \quad \min_{U\in\mathrm{U}(n)}\mathrm{Tr}\left(e^{A+UBU^\dagger}\right). \tag{5.5}$$

6 The Stahl's theorem

In 1975, Bessis, Moussa and Villani conjectured a way of rewritting the partition function of a broad class of statistical systems. In 2013, Stahl confirmed this conjecture. The precise statement can be formulated as follows:

Theorem 6.1 (Stahl's Theorem, [17])**.** *Let A and B be two $n\times n$ Hermitian matrices, where B is positive semi-definite. Then the function*

$$\mathrm{Tr}\left(e^{A-sB}\right), \quad s\geqslant 0 \tag{6.1}$$

can be represented as the Laplace transform of a non-negative measure $\mu_{A,B}$ on $[0,\infty)$. That is,

$$\mathrm{Tr}\left(e^{A-sB}\right)=\int_0^\infty e^{-st}\mathrm{d}\mu_{A,B}(t). \tag{6.2}$$

7 Peierls-Bogoliubov inequality

Theorem 7.1 (Peierls-Bogoliubov inequality, [3])**.** *For two Hermitian matrices H and K, it holds that*

$$\frac{\mathrm{Tr}\left(e^{H+K}\right)}{\mathrm{Tr}\left(e^H\right)}\geqslant\exp\left[\frac{\mathrm{Tr}\left(e^H K\right)}{\mathrm{Tr}\left(e^H\right)}\right]. \tag{7.1}$$

The equality occurs in the Peierls-Bogoliubov inequality if and only if K is a scalar matrix, i.e., $K=r\mathbb{1}$ for some real $r\in\mathbb{R}$.

The Peierls-Bogoliubov inequality is used to give reminder terms for some entropy inequalities [7,27,29].

8 Golden-Thompson inequality

Theorem 8.1 (Golden-Thompson inequality, [21])**.** *For arbitrary Hermitian matrices A and B, we have*

$$\operatorname{Tr}\left(e^{A+B}\right) \leqslant \operatorname{Tr}\left(e^A e^B\right). \tag{8.1}$$

Moreover, $\operatorname{Tr}\left(e^{A+B}\right) = \operatorname{Tr}\left(e^A e^B\right)$ *if and only if* $[A,B]=0$.

Proof. The proof of Golden-Thompson inequality based on Lie-Trotter product formula is presented as follows: Fix a natural number n and consider

$$X = \exp(2^{-n}A), \quad Y = \exp(2^{-n}B).$$

To prove Golden-Thompson inequality, it suffices to show that

$$\operatorname{Tr}\left((XY)^{2^n}\right) \leqslant \operatorname{Tr}\left(X^{2^n} Y^{2^n}\right). \tag{8.2}$$

Indeed, if this inequality holds then, taking limit as $n \to \infty$, we see that the left hand side of the above inequality converges to $\operatorname{Tr}\left(e^{A+B}\right)$ by Lie-Trotter product formula, while the right hand side equals $\operatorname{Tr}\left(e^A e^B\right)$. To prove Eq. (8.2), we need the following inequality:

$$|\operatorname{Tr}(T^m)| \leqslant \operatorname{Tr}\left(|T|^m\right) \tag{8.3}$$

for arbitrary matrix T and a positive integer m. Note that $|XY|^2 = (XY)^*(XY) = YX^2Y$. We have

$$\begin{aligned} \operatorname{Tr}\left((XY)^{2^n}\right) &\leqslant \operatorname{Tr}\left(|XY|^{2^n}\right) = \operatorname{Tr}\left(\left(YX^2Y\right)^{2^{n-1}}\right) = \operatorname{Tr}\left((X^2Y^2)^{2^{n-1}}\right) \\ &\leqslant \operatorname{Tr}\left((X^4Y^4)^{2^{n-2}}\right) \leqslant \cdots \leqslant \operatorname{Tr}\left(X^{2^n}Y^{2^n}\right) \end{aligned} \tag{8.4}$$

After n steps, we arrive at Eq. (8.2). This proves Golden-Thompson inequality. ☐

Remark 8.1. In order to show Eq. (8.3), we use the following Weyl's majorization theorem (See below). Setting $f(t) = t^m$ in Theorem 8.2:

$$|\operatorname{Tr}(T^m)| = \left|\sum_{i=1}^n \lambda_i^m\right| \leqslant \sum_{i=1}^n |\lambda_i|^m \leqslant \sum_{i=1}^n s_i^m = \operatorname{Tr}\left(|T|^m\right).$$

Theorem 8.2 (Weyl's majorization theorem, [4])**.** *Let A be an $n\times n$ matrix with singular values $s_1 \geqslant \cdots \geqslant s_n$ and eigenvalues $\lambda_1,\ldots,\lambda_n$ arranged so that $|\lambda_1| \geqslant \cdots \geqslant |\lambda_n|$. Let $f:\mathbb{R}^+ \to \mathbb{R}^+$ be a function such that $f(t)$ is convex and increasing in t. Then*

$$\sum_{i=1}^n f(|\lambda_i|) \leqslant \sum_{i=1}^n f(s_i).$$

Proposition 8.1. $\left|\mathrm{Tr}\left(X^{2n}\right)\right| \leqslant \mathrm{Tr}\left(\left(XX^{\dagger}\right)^{n}\right)$

Proposition 8.2 (Thompson's lemma, [21]). *If X and Y are Hermitian matrices, we have*

$$\mathrm{Tr}\left((XY)^{2m}\right) \leqslant \mathrm{Tr}\left(\left(X^2Y^2\right)^{m}\right) \tag{8.5}$$

for any positive integer m.

Theorem 8.3. *If A and B are Hermitian matrices, we have*

$$\mathrm{Tr}\left(e^{A+B}\right) \leqslant \mathrm{Tr}\left(\left(e^{A/n}e^{B/n}\right)^{n}\right) \tag{8.6}$$

for any nonzero integer n.

Proof. Now let

$$\alpha_j := \mathrm{Tr}\left(\left[\exp\left(\frac{A}{2^j n}\right)\exp\left(\frac{B}{2^j n}\right)\right]^{2^j n}\right) \tag{8.7}$$

for any nonzero integer n, where $j=0,1,2,\ldots$. Then Thompson's lemma yields

$$\alpha_j \geqslant \alpha_{j+1} \geqslant \cdots \geqslant \alpha_\infty.$$

Here the equality that $\alpha_\infty = \mathrm{Tr}\left(e^{A+B}\right)$ is assured by Lie-Trotter product formula. In particular, we obtain that

$$\alpha_0 \geqslant \alpha_\infty.$$

The case $n=1$ is the Golden-Thompson-Symanzik inequality. $\square$

The well-known Golden-Thompson inequality and its saturation is used in studying some mathematical problems in quantum informational theory [28]. It is also useful in random matrix theory [8, 22]. More extensions can be found in [9].

9 Reverse inequality to Golden-Thompson type inequalities

Recently Bourin and Seo [6] considered the comparison between e^{A+B} and e^Ae^B, where A and B are $n\times n$ Hermitian matrices. They have obtained the following result:

Theorem 9.1. ([6]) *Let A and B be $n\times n$ Hermitian matrices. Then*

$$\mathrm{Tr}\left(e^Ae^B\right) \leqslant S(\kappa_A)\,\mathrm{Tr}\left(e^{A+B}\right), \tag{9.1}$$

where κ_A is the condition number of e^A and

$$S(r) := \frac{(r-1)r^{\frac{1}{r-1}}}{e\log r}$$

is the Specht ratio of the reverse arithmetic-geometric mean inequality. It is a sharp reverse result to the Golden-Thompson inequality. This can be extended to each eigenvalue. Equivalently there exists a unitary V such that

$$\exp\left(\frac{A}{2}\right)\exp(B)\exp\left(\frac{A}{2}\right)\leqslant S(\kappa_A)V\exp(A+B)V^{\dagger}. \tag{9.2}$$

There also exists a unitary W such that

$$W\exp(A+B)W^{\dagger}\leqslant S(\kappa_A)\exp\left(\frac{A}{2}\right)\exp(B)\exp\left(\frac{A}{2}\right). \tag{9.3}$$

Proof. Note that $\exp\left(\frac{A}{2}\right)\exp(B)\exp\left(\frac{A}{2}\right)$ is unitarily equivalent to $\exp\left(\frac{B}{2}\right)\exp(A)\exp\left(\frac{B}{2}\right)$ and for $p>1$,

$$\exp\left(\frac{B}{2}\right)\exp(A)\exp\left(\frac{B}{2}\right)=\left[\exp\left(\frac{B}{2p}\right)\right]^p\left[\exp\left(\frac{A}{p}\right)\right]^p\left[\exp\left(\frac{B}{2p}\right)\right]^p.$$

Then we need the following result, which is a reverse inequality of Araki's inequality:

$$\frac{1}{K(z,p)}U(AZA)^pU^{\dagger}\leqslant A^pZ^pA^p\leqslant K(z,p)V(AZA)^pV^{\dagger}. \tag{9.4}$$

The constant

$$K(z,p):=\frac{z^p-z}{(p-1)(z-1)}\left(\left(1-\frac{1}{p}\right)\frac{z^p-1}{z^p-z}\right)^p$$

and its inverse are optimal. By this result, it follows that

$$\exp\left(\frac{B}{2}\right)\exp(A)\exp\left(\frac{B}{2}\right)\leqslant K(\kappa_A^{\frac{1}{p}},p)V_p\left[\exp\left(\frac{B}{2p}\right)\exp\left(\frac{A}{p}\right)\exp\left(\frac{B}{2p}\right)\right]^pV_p^{\dagger}$$

for some unitary $V_p(p>1)$. Now $\lim_{p\to\infty}K(\kappa_A^{\frac{1}{p}},p)=S(\kappa_A)$, via Lie-Trotter product formula Eq. (2.7) (Theorem 2.2), we can get that Eq. (9.2) holds. Next by taking trace on both sides gives Eq. (9.1). The first inequality of Eq. (9.4) implies

$$\begin{aligned}&\left[\exp\left(\frac{B}{2p}\right)\exp\left(\frac{A}{p}\right)\exp\left(\frac{B}{2p}\right)\right]^p\\&\leqslant K(\kappa_A^{\frac{1}{p}},p)U_p\exp\left(\frac{B}{2}\right)\exp(A)\exp\left(\frac{B}{2}\right)U_p^{\dagger}\end{aligned} \tag{9.5}$$

for some unitaries $U_p(p>1)$. Hence, the remaining part of proof goes similarly as above. Hence Eq. (9.3). This proves all results. We are done. □

Remark 9.1. In summary, we have

$$\frac{1}{S(\kappa_A)}U\exp(A+B)U^\dagger \leqslant \exp\left(\frac{A}{2}\right)\exp(B)\exp\left(\frac{A}{2}\right) \leqslant S(\kappa_A)V\exp(A+B)V^\dagger \tag{9.6}$$

for some unitaries U and V. Moreover,

$$\frac{1}{S(\kappa)}\mathrm{Tr}\left(e^{A+B}\right) \leqslant \mathrm{Tr}\left(e^A e^B\right) \leqslant S(\kappa)\mathrm{Tr}\left(e^{A+B}\right). \tag{9.7}$$

Of course, we see from Eq. (9.7) by the symmetry of A and B, that

$$\max\left\{\frac{1}{S(\kappa_A)},\frac{1}{S(\kappa_B)}\right\} \leqslant \frac{\mathrm{Tr}\left(e^A e^B\right)}{\mathrm{Tr}\left(e^{A+B}\right)} \leqslant \min\{S(\kappa_A),S(\kappa_B)\} \tag{9.8}$$

Later, Forrester and Thompson [8] investigated this trace inequality of matrix exponential from the view of random matrix theory, they have obtained in the 2×2 case that

$$\frac{\mathbb{E}\,\mathrm{Tr}\left(e^A e^B\right)}{\mathbb{E}\,\mathrm{Tr}\left(e^{A+B}\right)} = \frac{4}{3}. \tag{9.9}$$

There is a naturally question: what is the analytical value of $\mathbb{E}\left[\frac{\mathrm{Tr}\left(e^A e^B\right)}{\mathrm{Tr}\left(e^{A+B}\right)}\right]$ (e.g. in the 2×2 case)?

Remark 9.2. There is an reverse inequality for the arithmetic and geometric means of positive numbers: Let $x_1 \geqslant \cdots \geqslant x_n > 0$ and set $r = x_1/x_n$. Then

$$\left(\sqrt[n]{\prod_{j=1}^n x_j} \leqslant\right) \frac{1}{n}\sum_{j=1}^n x_j \leqslant S(r)\sqrt[n]{\prod_{j=1}^n x_j}, \tag{9.10}$$

where $S(r) := \frac{(r-1)r^{\frac{1}{r-1}}}{e\log r}$ is called the Specht ratio at r, in particular $S(1) = 1$. Specht's inequality is a ratio type reverse inequality of the classical arithmetic-geometric mean inequality.

10 The bipolar decomposition

Recently, Bhatia have shown the following theorem:

Theorem 10.1 (The bipolar decomposition, [5]). *Every $n\times n$ invertible complex matrix M can be factored as*

$$M = e^L e^{\mathrm{i}C_L} e^{\mathrm{i}C_R} e^R$$

where R and C_L are real symmetric matrices, *and L and C_R are* real skew-symmetric matrices.

Proof. Firstly, we employ the Mostow decomposition theorem (see Theorem 10.3), $M = Ue^{\mathrm{i}C_R}e^R$, where U is unitary, R is real and symmetric, and C_R is real and skew-symmetric. Secondly, we use Theorem 10.4 to decompose $U = e^L e^{\mathrm{i}T} C_L$ where L is real skew-symmetric matrix, and C_L is real symmetric matrix. Putting them together completes the proof. $\square$

In order to get this result, we need to make some observations. If A and B are two positive definite matrices, denoted by $A,B>0$, then the equation

$$XA^{-1}X = B \tag{10.1}$$

has a unique positive definite solution given by

$$X = \sqrt{A}\sqrt{A^{-\frac{1}{2}}BA^{-\frac{1}{2}}}\sqrt{A} := g(A,B) > 0. \tag{10.2}$$

This is called the *geometric mean* of A and B. Thus

$$X = g(A,B) \Longleftrightarrow XA^{-1}X = B.$$

Clearly X is also the solution of $XB^{-1}X = A$, and hence $g(A,B) = g(B,A)$. Also we have $g(A^{-1},B^{-1}) = g(A,B)^{-1}$. In what follows, we show that

$$g(\overline{A},\overline{B}) = \overline{g(A,B)}. \tag{10.3}$$

A consequence of this is that $g(A,\overline{A})$ is a real matrix, and $g(A^{-1},\overline{A})$ is a circular matrix. Indeed, The matrix $X = g(A,B)$ is the unique positive definite solution of $XA^{-1}X = B$. Taking complex conjugates changes this equation to $\overline{X}(\overline{A})^{-1}\overline{X} = \overline{B}$. So, $\overline{X} = g(\overline{A},\overline{B})$. This completes the proof. If $Y = g(A,\overline{A})$, then $Y = g(\overline{A},A) = \overline{Y}$, thus Y is real. Similarly, if $Y = g(A^{-1},\overline{A})$, then we have $\overline{Y} = g((\overline{A})^{-1},A) = g(A,(\overline{A})^{-1})$. Thus

$$(\overline{Y})^{-1} = g(A,(\overline{A})^{-1})^{-1} = g(A^{-1},\overline{A}) = Y,$$

that is, $Y\overline{Y} = \mathbb{1}$, it is circular. We summarize the above discussion into a proposition.

Proposition 10.1. *Given $A,B>0$. It holds that*

(i) $g(A,B) = g(B,A)$.

(ii) $g(A^{-1},B^{-1}) = g(A,B)^{-1}$.

(iii) $g(\overline{A},\overline{B}) = \overline{g(A,B)}$.

(iv) $g(A,\overline{A})$ *is a real matrix.*

(v) $g(A^{-1},\overline{A})$ *is a circular matrix.*

Theorem 10.2. ([5]) *(i). Every complex positive definite matrix M can be factored as*

$$M = e^D e^{\mathrm{i}C} e^D, \tag{10.4}$$

where D is a real symmetric matrix, and C a real skew-symmetric matrix. Both C and D are uniquely determined.

(ii). Every complex positive definite matrix N can also be factored as

$$N = e^{\mathrm{i}S} e^T e^{\mathrm{i}S}, \tag{10.5}$$

where S and T are real skew-symmetric and symmetric matrices, respectively. Both are uniquely determined by these conditions.

Proof. (i). Since $g(M,\overline{M})$ is a real positive definite matrix, there exists a unique real symmetric matrix D such that $g(M,\overline{M}) = e^{2D}$. Let $Y = e^{-D} M e^{-D}$. Then Y is positive definite and

$$Y\overline{Y} = e^{-D} M e^{-2D} \overline{M} e^{-D}.$$

Clearly $g(M,\overline{M}) = e^{2D}$ can be expressed as $\overline{M} = e^{2D} M^{-1} e^{2D}$. In view of this, we see that $Y\overline{Y} = \mathbb{1}$. Thus Y is positive definite and circular. So $Y = e^{\mathrm{i}C}$, where C is a real skew-symmetric matrix. Hence the result.

The uniqueness can be obtained by

$$M^{-1} = e^{-D} e^{-\mathrm{i}C} e^{-D}, \quad \overline{M} = e^D e^{-\mathrm{i}C} e^D.$$

In fact, $\overline{M} = e^{2D} M^{-1} e^{2D}$, and therefore, D is uniquely determined. Then it is easy to see that C is unique also.

(ii). We express the circular matrix $g(N^{-1},\overline{N})$ as $g(N^{-1},\overline{N}) = e^{-2\mathrm{i}S}$, where S is real and skew-symmetric. Let $X = e^{-\mathrm{i}S} N e^{-\mathrm{i}S}$. Then X is positive definite, and $\overline{X} = e^{\mathrm{i}S} \overline{N} e^{\mathrm{i}S}$. Thus we have $\overline{N} = e^{-2\mathrm{i}S} N e^{-2\mathrm{i}S}$. So X is a real positive definite matrix. Hence there exists a real symmetric matrix T such that $X = e^T$. Then $N = e^{\mathrm{i}S} e^T e^{\mathrm{i}S}$. It is easily to check the uniqueness of S and T. We are done. $\square$

Theorem 10.3 (The Mostow decomposition theorem, [5])**.** *(i). Let M be an invertible complex matrix. Then M can be factored as*

$$M = U e^{\mathrm{i}C} e^R, \tag{10.6}$$

where U is unitary, R is real and symmetric, and C is real and skew-symmetric. Such a factorization is unique.

(ii). Let N be an invertible complex matrix. Then N can be factored as

$$N = V e^S e^{\mathrm{i}T}, \tag{10.7}$$

where V is unitary, S real symmetric, and T real skew-symmetric. Such a factorization is unique.

Proof. (i). Let $A=M^{\dagger}M$. Since $g(A,\overline{A})$ is a real positive definite matrix, there exists a unique real symmetric matrix R such that $g(A,\overline{A})=e^{2R}$. Let $Y=e^{-R}Ae^{-R}$. Thus Y is positive definite and circular. Let C be a real skew-symmetric matrix such that $Y=e^{2\mathrm{i}C}$. Then let $U=Me^{-R}e^{-\mathrm{i}C}$. Since e^{-R} and $e^{-\mathrm{i}C}$ both are Hermitian, we have

$$\begin{aligned} U^{\dagger}U &= e^{-\mathrm{i}C}e^{-R}M^{\dagger}Me^{-R}e^{-\mathrm{i}C}=e^{-\mathrm{i}C}e^{-R}Ae^{-R}e^{-\mathrm{i}C} \\ &= e^{-\mathrm{i}C}Ye^{-\mathrm{i}C}=e^{-\mathrm{i}C}e^{2\mathrm{i}C}e^{-\mathrm{i}C}=\mathbb{1}. \end{aligned} \tag{10.8}$$

Thus U is unitary, and hence the result. To arrive at the uniqueness of this decomposition, we see that

$$A:=M^{\dagger}M=e^{R}e^{2\mathrm{i}C}e^{R}. \tag{10.9}$$

Since R and C are real, it follows that

$$\overline{A}=e^{R}e^{-2\mathrm{i}C}e^{R}=e^{R}\left(e^{-R}Ae^{-R}\right)^{-1}e^{R}=e^{2R}A^{-1}e^{2R}. \tag{10.10}$$

Thus $g(A,\overline{A})=e^{2R}$. The uniqueness is proved.

(ii). We denote $B:=N^{\dagger}N$. Then $g(B^{-1},\overline{B})$ is a circular matrix, which can be expressed as $g(B^{-1},\overline{B})=e^{-2\mathrm{i}T}$, where T is real and skew-symmetric. Let $X=e^{-\mathrm{i}T}Be^{-\mathrm{i}T}$. Let S be the real symmetric matrix such that $X={}^{2S}$. Let $V=Ne^{-\mathrm{i}T}e^{-S}$. Then

$$V^{\dagger}V=e^{-S}e^{-\mathrm{i}T}Be^{-\mathrm{i}T}e^{-S}e^{-S}Xe^{-S}=\mathbb{1}. \tag{10.11}$$

So V is unitary, and the decomposition holds. It can be checked that this decomposition is unique. $\square$

Theorem 10.4. ([5]) *(i). Every complex unitary matrix U can be factored as*

$$U=e^{L}e^{\mathrm{i}T}, \tag{10.12}$$

where L is real skew-symmetric matrix, and T is real symmetric matrix.

(ii). Every complex unitary matrix V can be factored as

$$V=e^{\mathrm{i}S}e^{R}, \tag{10.13}$$

where S is real symmetric matrix, and R is real skew-symmetric matrix.

Proof. (i). Assume that W is a unitary matrix that is also symmetric, i.e., $W^{\mathsf{T}}=W$. Then there exists a Hermitian matrix H such that $W=e^{\mathrm{i}H}$. This H is unique if its spectrum is in $(0,2\pi]$. The condition $W^{\mathsf{T}}=W$ means $H^{\mathsf{T}}=H$. Thus H is real and symmetric. Now let U be any unitary matrix. Then U^{T} is also unitary, and $U^{\mathsf{T}}U$ is unitary and symmetric. Hence there exists a real and symmetric matrix T such that $U^{\mathsf{T}}U=e^{2\mathrm{i}T}$. Let $Q=Ue^{-\mathrm{i}T}$. Then

$$Q^{\dagger}Q=\mathbb{1}, \quad Q^{\mathsf{T}}Q=\mathbb{1}.$$

So Q is unitary and $Q^\dagger = Q^\top$. Thus Q is real and orthogonal matrix. So we have $U = Qe^{\mathrm{i}T}$, where Q is a real orthogonal, and T a real symmetric matrix. Thus $\det(Q) = \pm 1$. If $\det(Q) = 1$, then there exists a real skew-symmetric matrix L such that $Q = e^L$, and we get the desired proof. If $\det(Q) = -1$, for the real symmetric matrix T, there exists a real orthogonal matrix G such that

$$GTG^\top = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Let K be the matrix given by $GKG^\top = \mathrm{diag}(\pi, 0, \ldots, 0)$. Then K is a real symmetric matrix that commutes with T. Let $J = e^{-\mathrm{i}K}$. Then $GJG^\top = \mathrm{diag}(-1, 1, \ldots, 1)$. Then the matrix J is real orthogonal, $\det(J) = -1, J^2 = \mathbb{1}$, and J commutes with $e^{\mathrm{i}T}$. From $U = Qe^{\mathrm{i}T}$, we see that

$$U = QJJe^{\mathrm{i}T} = QJe^{\mathrm{i}(K+T)}.$$

Then QJ is a real orthogonal matrix with $\det(QJ) = 1$ if $\det(Q) = -1$. So $QJ = e^L$ for some real skew-symmetric matrix L. In this case, $K+T$ in the position of T.

(ii). The proof of the second result goes similarly. That is, we start with $VV^\top$. ☐

11 The integral representation of pinching maps

Any Hermitian matrix H has a spectral decomposition $H = \sum_j \lambda_j \Pi_j$, where $\lambda_j \in \mathrm{Spec}(H) \subset \mathbb{R}$ are unique eigenvalues and Π_j are mutually orthogonal projections. The *spectral pinching map* with respect to H is

$$\mathscr{P}_H(X) = \sum_j \Pi_j X \Pi_j.$$

Theorem 11.1. ([19]) *Let M be positive definite matrix. There exists a probability measure μ on $\mathbb{R}$ such that*

$$\mathscr{P}_M(X) = \int_{-\infty}^{\infty} \mathrm{d}\mu(t) M^{\mathrm{i}t} X M^{-\mathrm{i}t} \tag{11.1}$$

for all $X \geqslant 0$.

Proof. By the spectral decomposition of M, we find

$$M^{\mathrm{i}t} = \sum_j \lambda_j^{\mathrm{i}t} \Pi_j. \tag{11.2}$$

Then

$$M^{\mathrm{i}t} X M^{-\mathrm{i}t} = \sum_{i,j} \lambda_i^{\mathrm{i}t} \lambda_j^{-\mathrm{i}t} \Pi_i X \Pi_j. \tag{11.3}$$

Denote

$$c_0 = \frac{1}{2} \min\left\{ \left| \log\lambda_i - \log\lambda_j \right| : \lambda_i, \lambda_j \in \mathrm{Spec}(M),\ \lambda_i \neq \lambda_j \right\}. \tag{11.4}$$

Let

$$\mathrm{d}\mu(t)=\frac{1-\cos(c_0t)}{\pi c_0t^2}\mathrm{d}t.$$

Note that

$$\int_{\mathbb{R}} e^{-\mathrm{i}ts}\mathrm{d}\mu(t)=\int_{\mathbb{R}} e^{-\mathrm{i}ts}\frac{1-\cos(c_0t)}{\pi c_0t^2}\mathrm{d}t=\delta(s), \tag{11.5a}$$

$$\frac{1-\cos(c_0t)}{\pi c_0t^2}=\frac{1}{2\pi}\int_{\mathbb{R}}\mathrm{d}s\left(1-\frac{s}{c_0}\right)\chi_{(-c_0,c_0)}(s)e^{\mathrm{i}ts}. \tag{11.5b}$$

It is easily to check that μ is a probability distribution on $\mathbb{R}$, i.e., nonnegative and normalized. Indeed, $\mu(t)\geqslant 0$ for any $t\in\mathbb{R}$; and

$$\begin{aligned}\int_{\mathbb{R}}\mathrm{d}\mu(t)&=\int_{\mathbb{R}}\frac{1-\cos(c_0t)}{\pi c_0t^2}\mathrm{d}t=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1-\cos(t)}{t^2}\mathrm{d}t\\&=-\frac{1}{\pi}\int_{-\infty}^{\infty}(1-\cos t)\mathrm{d}\left(\frac{1}{t}\right)\\&=-\frac{1}{\pi}\left[\frac{1-\cos(t)}{t}\Big|_{-\infty}^{\infty}-\int_{-\infty}^{\infty}\frac{1}{t}\mathrm{d}(1-\cos t)\right],\end{aligned} \tag{11.6}$$

implying

$$\int_{\mathbb{R}}\mathrm{d}\mu(t)=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin t}{t}\mathrm{d}t=1.$$

This completes the proof. $\square$

Based on the pinching inequality $\mathscr{P}_M(X)\geqslant\frac{1}{|\mathrm{Spec}(M)|}X$, combining with the fact that

$$\left|\mathrm{Spec}(M^{\otimes n})\right|\leqslant O(\mathrm{Poly}(n)),$$

where $\mathrm{Poly}(n)$ denotes any polynomial in n, the authors of [19] present a very fast track to the Golden-Thompson inequality: for any $n\in\mathbb{N}$

$$\begin{aligned}&\log\mathrm{Tr}(\exp(\log A+\log B))=\frac{1}{n}\log\mathrm{Tr}\left(\exp\left(\log A^{\otimes n}+\log B^{\otimes n}\right)\right)\\&\leqslant\frac{1}{n}\log\mathrm{Tr}\left(\exp\left(\log\mathscr{P}_{B^{\otimes n}}\left(A^{\otimes n}\right)+\log B^{\otimes n}\right)\right)+\frac{\log\mathrm{Poly}(n)}{n}\\&\frac{1}{n}\log\mathrm{Tr}\left(\mathscr{P}_{B^{\otimes n}}\left(A^{\otimes n}\right)B^{\otimes n}\right)+\frac{\log\mathrm{Poly}(n)}{n}\\&=\log\mathrm{Tr}(AB)+\frac{\log\mathrm{Poly}(n)}{n}.\end{aligned}$$

Here we used several properties of pinching map, for instance, $[\mathscr{P}_M(X),M]=0$, and $\mathrm{Tr}(\mathscr{P}_M(X)M)=\mathrm{Tr}(XM)$ for any $X\geqslant 0$. By taking the limit of $n\to\infty$, we get that

$$\log\mathrm{Tr}(\exp(\log A+\log B))\leqslant\log\mathrm{Tr}(AB)$$

because $\lim_{n\to\infty}\frac{\log\mathrm{Poly}(n)}{n}=0$. Therefore

$$\mathrm{Tr}(\exp(\log A+\log B))\leqslant\mathrm{Tr}(AB)=\mathrm{Tr}(\exp(\log A)\exp(\log B)). \tag{11.7}$$

Replacement of $(\log A,\log B)$ by (H,K) give rises the desired Golden-Thompson inequality.

12 Matrix integrals

Proposition 12.1. ([14]) Let H be an $n\times n$ real symmetric positive matrix. Then

$$\int_{\mathbb{R}^n}\mathrm{d}x e^{-\langle x|H|x\rangle}=\sqrt{\frac{\pi^n}{\det(H)}}, \tag{12.1}$$

where $\mathrm{d}x=\prod_{j=1}^n \mathrm{d}x_j$ denotes the Lebesgue volume element.

Proof. Let $|y\rangle=\sqrt{H}|x\rangle$. Then $\langle x|H|x\rangle=\langle y|y\rangle$ and $\mathrm{d}y=\det(\sqrt{H})\mathrm{d}x=\sqrt{\det(H)}\mathrm{d}x$ [25]. The mentioned integral can be rewritten as

$$\int_{\mathbb{R}^n}\mathrm{d}x e^{-\langle x|H|x\rangle}=\int_{\mathbb{R}^n}\left(\frac{1}{\sqrt{\det(H)}}\mathrm{d}y\right)e^{-\langle y,y\rangle}$$

$$=\frac{1}{\sqrt{\det(H)}}\int_{\mathbb{R}^n}\mathrm{d}y e^{-\langle y,y\rangle}=\frac{1}{\sqrt{\det(H)}}\prod_{j=1}^n\int_{-\infty}^{\infty}\mathrm{d}y_j e^{-y_j^2}=\sqrt{\frac{\pi^n}{\det(H)}}, \tag{12.2}$$

via $\int_{-\infty}^{\infty}\mathrm{d}y e^{-y^2}=\sqrt{\pi}$, implying the proof. □

Theorem 12.1. ([14]) *Let A and B be real symmetric positive $n\times n$ matrices. Then for all $\lambda\in[0,1]$,*

$$\det(\lambda A+(1-\lambda)B)\geqslant\det{}^{\lambda}(A)\det{}^{1-\lambda}(B). \tag{12.3}$$

Proof. If $\lambda=0,1$, then the result follows trivially. Let $H=\lambda A+(1-\lambda)B$, in Eq. (12.1), for $\lambda\in(0,1)$, where A,B are as mentioned in Proposition 12.1. Then

$$\sqrt{\frac{\pi^n}{\det(\lambda A+(1-\lambda)B)}}$$

$$=\int_{\mathbb{R}^n}\mathrm{d}x e^{-\langle x|\lambda A+(1-\lambda)B|x\rangle}=\int_{\mathbb{R}^n}\mathrm{d}x e^{-\lambda\langle x|A|x\rangle}e^{-(1-\lambda)\langle x|B|x\rangle}. \tag{12.4}$$

By employing Hölder's inequality

$$\left|\int_{\mathbb{R}^n} \mathrm{d}x f(x)g(x)\right| \leqslant \left(\int_{\mathbb{R}^n} \mathrm{d}x |f(x)|^p\right)^{\frac{1}{p}} \left(\int_{\mathbb{R}^n} \mathrm{d}x |g(x)|^q\right)^{\frac{1}{q}} \tag{12.5}$$

for

$$f(x)=e^{-\lambda\langle x|A|x\rangle}, \quad g(x)=e^{-(1-\lambda)\langle x|B|x\rangle},$$

and $p=\frac{1}{\lambda}, q=\frac{1}{1-\lambda}$, we get

$$\begin{aligned}\int_{\mathbb{R}^n} \mathrm{d}x e^{-\lambda\langle x|A|x\rangle} e^{-(1-\lambda)\langle x|B|x\rangle} &\leqslant \left(\int_{\mathbb{R}^n} \mathrm{d}x e^{-\langle x|A|x\rangle}\right)^{\lambda} \left(\int_{\mathbb{R}^n} \mathrm{d}x e^{-\langle x|B|x\rangle}\right)^{1-\lambda} \\ = \sqrt{\frac{\pi^n}{\det(A)}}^{\lambda} \sqrt{\frac{\pi^n}{\det(B)}}^{1-\lambda} &= \sqrt{\frac{\pi^n}{\det^{\lambda}(A)\det^{1-\lambda}(B)}}.\end{aligned} \tag{12.6}$$

Hence

$$\sqrt{\frac{\pi^n}{\det(\lambda A+(1-\lambda)B)}} \leqslant \sqrt{\frac{\pi^n}{\det^{\lambda}(A)\det^{1-\lambda}(B)}},$$

that is,

$$\det(\lambda A+(1-\lambda)B) \geqslant \det^{\lambda}(A)\det^{1-\lambda}(B).$$

This completes the proof. $\square$

The above results can be easily generalized to the complex Hermitian positive matrices.

Proposition 12.2. ([14]) *Let H be a complex Hermitian positive matrix. Then*

$$\int_{\mathbb{C}^n} \mathrm{d}z e^{-\langle z|H|z\rangle} = \frac{\pi^n}{\det(H)}, \tag{12.7}$$

where $\mathrm{d}z=\mathrm{d}x\mathrm{d}y$ *for* $z=x+\sqrt{-1}y, x,y\in\mathbb{R}^n$.

Proof. Now $\langle z|H|z\rangle$ for $z=x+\sqrt{-1}y$, where $x,y\in\mathbb{R}^n$, can be realized as follows

$$\langle z|H|z\rangle = \left\langle \widehat{z}\middle|\widehat{H}\middle|\widehat{z}\right\rangle, \tag{12.8a}$$

where

$$\widehat{z}=\begin{bmatrix} x \\ y \end{bmatrix}, \quad \widehat{H}=\begin{bmatrix} \mathrm{Re}(H) & -\mathrm{Im}(H) \\ \mathrm{Im}(H) & \mathrm{Re}(H) \end{bmatrix}. \tag{12.8b}$$

Indeed, denote $H=\mathrm{Re}(H)+\sqrt{-1}\mathrm{Im}(H)$ with $\mathrm{Re}(H)$,$\mathrm{Im}(H)$ being real matrix, so $\mathrm{Re}(H)$ is real symmetric and $\mathrm{Im}(H)$ is real skew-symmetric, then

$$\left\langle \widehat{z}\middle|\widehat{H}\middle|\widehat{z}\right\rangle=\langle x,\mathrm{Re}(H)x-\mathrm{Im}(H)y\rangle+\langle y,\mathrm{Im}(H)x+\mathrm{Re}(H)y\rangle, \tag{12.9}$$

$$\begin{aligned}\langle z|H|z\rangle=&\langle x,\mathrm{Re}(H)x-\mathrm{Im}(H)y\rangle+\langle y,\mathrm{Im}(H)x+\mathrm{Re}(H)y\rangle\\&+\sqrt{-1}\left(\langle x|\mathrm{Im}(H)|x\rangle+\langle y|\mathrm{Im}(H)|y\rangle+\langle x|\mathrm{Re}(H)|y\rangle-\langle y|\mathrm{Re}(H)|x\rangle\right).\end{aligned} \tag{12.10}$$

Now that $\mathrm{Re}(H)^{\mathsf{T}}=\mathrm{Re}(H)$ and $\mathrm{Im}(H)^{\mathsf{T}}=-\mathrm{Im}(H)$, it follows that

$$\langle x|\mathrm{Im}(H)|x\rangle=\langle y|\mathrm{Im}(H)|y\rangle=0,\quad \langle x|\mathrm{Re}(H)|y\rangle=\langle y|\mathrm{Re}(H)|x\rangle. \tag{12.11}$$

That is, $\langle z|H|z\rangle=\left\langle \widehat{z}\middle|\widehat{H}\middle|\widehat{z}\right\rangle$. Thus we can rewrite the integral above into the following form:

$$\int_{\mathbb{R}^{2n}}\mathrm{d}\widehat{z}e^{-\langle\widehat{z}|\widehat{H}|\widehat{z}\rangle}=\sqrt{\frac{\pi^{2n}}{\det(\widehat{H})}}=\frac{\pi^n}{\det(H)}, \tag{12.12}$$

where we used the fact that $\sqrt{\det(\widehat{H})}=\det(H)$ [25] since $\det(H)>0$. □

Theorem 12.2. ([14]) *Let A and B be complex Hermitian positive $n\times n$ matrices. Then for all $\lambda\in[0,1]$,*

$$\det(\lambda A+(1-\lambda)B)\geqslant\det^{\lambda}(A)\det^{1-\lambda}(B). \tag{12.13}$$

Proof. Note that, for $H=\lambda A+(1-\lambda)B$, $\widehat{H}=\lambda\widehat{A}+(1-\lambda)\widehat{B}$. The proof goes similar to that of Theorem 12.1. □

13 Trace inequality in quantum information theory

Theorem 13.1. ([18,26]) *Let ρ and σ be two density matrices in $\mathrm{D}(\mathbb{C}^n)$, and Φ be a quantum channel. Then*

$$\mathrm{Tr}\left(\exp\left(\log\sigma+\Phi^{\dagger}(\log\Phi(\rho)-\log\Phi(\sigma))\right)\right)\leqslant 1. \tag{13.1}$$

14 Fréchet differential calculus

We will review very quickly some basic concepts of the *Fréchet differential calculus*, with special emphasis on matrix analysis. The proofs are given in [4].

Let $\mathcal{X}$,$\mathcal{Y}$ be *real* Banach spaces, and let $\mathrm{L}(\mathcal{X},\mathcal{Y})$ be the space of bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$.

Definition 14.1. Let $\mathcal{U}$ be an open subset of $\mathcal{X}$. A continuous map f from $\mathcal{U}$ to $\mathcal{Y}$ is said to be *differentiable* at a point u of $\mathcal{U}$ if there exists $T\in \mathrm{L}(\mathcal{X},\mathcal{Y})$ such that

$$\lim_{v\to 0}\frac{\|f(u+v)-f(u)-Tv\|}{\|v\|}=0, \tag{14.1}$$

where $v\to 0$ means that $v\xrightarrow{\|\cdot\|}0$. It is easy to see that such a T, if it exists, is unique.

If f is differentiable at u, the operator T above is called the *derivative* of f at u. We will use for it the notation $Df(u)$. This is sometimes called the *Fréchet derivative*. If f is differentiable at every point of $\mathcal{U}$, we say that it is differentiable on $\mathcal{U}$.

One can see that, if f is differentiable at u, then for every $v\in\mathcal{X}$,

$$Df(u)(v)=\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}f(u+tv). \tag{14.2}$$

This is also called the *directional derivative* of f at u in the direction v.

We will recall from elementary calculus of functions of two variables that, the existence of directional derivatives in all directions does not ensure differentiability.

Let $f(A)=e^A$. Use the formula

$$e^{A+B}-e^A=\int_0^1 e^{(1-t)A}Be^{t(A+B)}\mathrm{d}t$$

(called Dyson's expansion) to show that

$$Df(A)(B)=\int_0^1 e^{(1-t)A}Be^{tA}\mathrm{d}t.$$

Indeed, since

$$e^A=\sum_{n=0}^{\infty}\frac{1}{n!}A^n,$$

it follows from (ii) that

$$\begin{aligned}
D(e^A)(B)&=\sum_{n=0}^{\infty}\frac{1}{n!}D(A^n)(B)\\
&=\sum_{n=0}^{\infty}\frac{1}{n!}\sum_{m=0}^{n-1}\mathbb{L}_{A^m}\mathbb{R}_{A^{n-m-1}}(B)=\sum_{n=0}^{\infty}\frac{1}{n!}\sum_{m=0}^{n-1}A^mBA^{n-m-1}\\
&=\sum_{n=0}^{\infty}\frac{m!(n-m-1)!}{n!}\sum_{m=0}^{n-1}\frac{A^m}{m!}B\frac{A^{n-m-1}}{(n-m-1)!}\\
&=\sum_{m=0}^{\infty}\sum_{p=0}^{\infty}\frac{m!p!}{(m+p+1)!}\frac{A^m}{m!}B\frac{A^p}{p!}.
\end{aligned}$$

Because $\frac{m!p!}{(m+p+1)!} = \int_0^1 (1-t)^m t^p \mathrm{d}t$, the above equation becomes:

$$\begin{aligned}
D(e^A)(B) &= \sum_{m=0}^{\infty}\sum_{p=0}^{\infty}\int_0^1 (1-t)^m t^p \mathrm{d}t \frac{A^m}{m!} B \frac{A^p}{p!} \\
&= \sum_{m=0}^{\infty}\sum_{p=0}^{\infty}\int_0^1 \frac{[(1-t)A]^m}{m!} B \frac{(tA)^p}{p!} \mathrm{d}t \\
&= \int_0^1 \sum_{m=0}^{\infty}\sum_{p=0}^{\infty} \frac{[(1-t)A]^m}{m!} B \frac{(tA)^p}{p!} \mathrm{d}t \\
&= \int_0^1 \left(\sum_{m=0}^{\infty} \frac{[(1-t)A]^m}{m!}\right) B \left(\sum_{p=0}^{\infty} \frac{(tA)^p}{p!}\right) \mathrm{d}t \\
&= \int_0^1 e^{(1-t)A} B e^{tA} \mathrm{d}t.
\end{aligned}$$

Example 14.1. ([10]) Let $B(t) = e^{-t\rho} B e^{-t\rho}$ and $L = 2\int_0^\infty e^{-t\rho} B e^{-t\rho} \mathrm{d}t$. Then there is an identity

$$\frac{\mathrm{d}}{\mathrm{d}t} B(t) = -\rho B(t) - B(t)\rho,$$

which implies that

$$\begin{aligned}
\rho L + L\rho &= 2\int_0^\infty (\rho B(t) + B(t)\rho) \mathrm{d}t = -2\int_0^\infty \frac{d}{dt} B(t) \mathrm{d}t \\
&= -2\int_0^\infty \mathrm{d}B(t) = -2[B(+\infty) - B(0)] = 2B.
\end{aligned}$$

Example 14.2. ([10]) Let

$$\rho_\theta \stackrel{\text{def}}{=} \frac{e^{H+\theta B}}{\mathrm{Tr}\left(e^{H+\theta B}\right)}$$

be a Gibbsian family of states, where $H^\dagger = H, B^\dagger = B$. Assume that $\rho_0 = e^H$ is a density matrix and $\mathrm{Tr}\left(\rho_0 B\right) = 0$. In what follows, we compute the Fréchet derivative of ρ_θ at $\theta = 0$.

In fact,

$$\frac{\mathrm{d}\rho_\theta}{\mathrm{d}\theta} = \frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{1}{\mathrm{Tr}\left(e^{H+\theta B}\right)}\right) e^{H+\theta B} + \frac{1}{\mathrm{Tr}\left(e^{H+\theta B}\right)} \frac{\mathrm{d}}{\mathrm{d}\theta} e^{H+\theta B},$$

i.e.,

$$\left.\frac{\mathrm{d}\rho_\theta}{\mathrm{d}\theta}\right|_{\theta=0} = \left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=0} e^{H+\theta B}.$$

Set $f(x) = e^x$. It is seen that

$$\left.\frac{\mathrm{d}\rho_\theta}{\mathrm{d}\theta}\right|_{\theta=0} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{\theta=0} f(H+\theta B).$$

Let the spectral decomposition of H is given by $H=\sum_i h_i|h_i\rangle\langle h_i|$. Then

$$\begin{aligned}\left\langle h_i\left|\frac{\mathrm{d}\rho_\theta}{\mathrm{d}\theta}\right|_{\theta=0}\right|h_j\right\rangle &= \left\langle h_i\left|\frac{\mathrm{d}}{\mathrm{d}t}\right|_{\theta=0} f(H+\theta B)\right|h_j\right\rangle \\ &= \frac{f(h_i)-f(h_j)}{h_i-h_j}\langle h_i|B|h_j\rangle.\end{aligned}$$

Since

$$\int_0^1 a^t b^{1-t}\mathrm{d}t = \frac{a-b}{\ln a-\ln b} \quad (a,b>0),$$

it follows from putting $a=f(h_i),b=f(h_j)$ in the above equation that

$$\frac{f(h_i)-f(h_j)}{h_i-h_j} = \int_0^1 e^{th_i}e^{(1-t)h_j}\mathrm{d}t.$$

This indicates that

$$\begin{aligned}\left\langle h_i\left|\frac{\mathrm{d}\rho_\theta}{\mathrm{d}\theta}\right|_{\theta=0}\right|h_j\right\rangle &= \int_0^1 e^{th_i}e^{(1-t)h_j}\mathrm{d}t\langle h_i|B|h_j\rangle \\ &= \left\langle h_i\left|\int_0^1 e^{tH}Be^{(1-t)H}\mathrm{d}t\right|h_j\right\rangle,\end{aligned}$$

which implies that

$$\left.\frac{d\rho_\theta}{d\theta}\right|_{\theta=0} = \int_0^1 e^{tH}Be^{(1-t)H}\mathrm{d}t.$$

Example 14.3. ([10]) Let ρ be a state and $\mathcal{K}_\rho(X)\stackrel{\text{def}}{=}\int_0^1\rho^t X\rho^{1-t}\mathrm{d}t$ defined for Hermite matrices. Recall that if $\rho=\sum_i\lambda_i|\lambda_i\rangle\langle\lambda_i|$, then $\rho^{-1}=\sum_i\frac{1}{\lambda_i}|\lambda_i\rangle\langle\lambda_i|$. For a super-operator, the spectral projection is $\mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|}\mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|}$ for which its action is given by

$$\begin{aligned}&\mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|}\mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|}\mathcal{K}_\rho(X) = \left\langle\lambda_i\left|\int_0^1\rho^tX\rho^{1-t}\mathrm{d}t\right|\lambda_j\right\rangle|\lambda_i\rangle\langle\lambda_j| \\ &= \int_0^1\lambda_i^t\lambda_j^{1-t}\mathrm{d}t\langle\lambda_i|X|\lambda_j\rangle|\lambda_i\rangle\langle\lambda_j| = \frac{\lambda_i-\lambda_j}{\ln\lambda_i-\ln\lambda_j}\mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|}\mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|}(X),\end{aligned}$$

that is

$$\mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|}\mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|}\mathcal{K}_\rho = \mathcal{K}_\rho\mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|}\mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|} = \frac{\lambda_i-\lambda_j}{\ln\lambda_i-\ln\lambda_j}\mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|}\mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|}.$$

This gives that

$$\mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|}\mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|}\mathcal{K}_\rho^{-1} = \mathcal{K}_\rho^{-1}\mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|}\mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|} = \frac{\ln\lambda_i-\ln\lambda_j}{\lambda_i-\lambda_j}\mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|}\mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|}.$$

Using the integral representation of $\ln x$:

$$\ln x = \int_0^\infty \left(\frac{1}{1+t} - \frac{1}{x+t} \right) \mathrm{d}t,$$

it follows that

$$\begin{aligned}
\frac{\ln\lambda_i - \ln\lambda_j}{\lambda_i - \lambda_j} &= \frac{1}{\lambda_i - \lambda_j} \left[\int_0^\infty \left(\frac{1}{1+t} - \frac{1}{\lambda_i + t} \right) \mathrm{d}t - \int_0^\infty \left(\frac{1}{1+t} - \frac{1}{\lambda_j + t} \right) \mathrm{d}t \right] \\
&= \frac{1}{\lambda_i - \lambda_j} \left[\int_0^\infty \left(\frac{1}{\lambda_j + t} - \frac{1}{\lambda_i + t} \right) \mathrm{d}t \right] \\
&= \int_0^\infty \frac{1}{(\lambda_i + t)(\lambda_j + t)} \mathrm{d}t.
\end{aligned}$$

Thus

$$\mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|} \mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|} \mathcal{K}_\rho^{-1} = \int_0^\infty \frac{1}{(\lambda_i + t)(\lambda_j + t)} \mathrm{d}t \mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|} \mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|}.$$

Furthermore,

$$\begin{aligned}
\mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|} \mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|} \mathcal{K}_\rho^{-1}(X) &= \int_0^\infty \frac{1}{(\lambda_i + t)(\lambda_j + t)} \mathrm{d}t \mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|} \mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|}(X) \\
&= \int_0^\infty \frac{1}{(\lambda_i + t)(\lambda_j + t)} \mathrm{d}t |\lambda_i\rangle\langle\lambda_i| X |\lambda_j\rangle\langle\lambda_j| \\
&= \int_0^\infty (\lambda_i + t)^{-1} |\lambda_i\rangle\langle\lambda_i| X (\lambda_j + t)^{-1} |\lambda_j\rangle\langle\lambda_j| \mathrm{d}t.
\end{aligned}$$

Finally,

$$\begin{aligned}
\mathcal{K}_\rho^{-1}(X) &= \sum_{i,j} \mathcal{L}_{|\lambda_i\rangle\langle\lambda_i|} \mathcal{R}_{|\lambda_j\rangle\langle\lambda_j|} \mathcal{K}_\rho^{-1}(X) \\
&= \sum_{i,j} \int_0^\infty (\lambda_i + t)^{-1} |\lambda_i\rangle\langle\lambda_i| X (\lambda_j + t)^{-1} |\lambda_j\rangle\langle\lambda_j| \mathrm{d}t \\
&= \int_0^\infty \left(\sum_i (\lambda_i + t)^{-1} |\lambda_i\rangle\langle\lambda_i| \right) X \left(\sum_j (\lambda_j + t)^{-1} |\lambda_j\rangle\langle\lambda_j| \right) \mathrm{d}t \\
&= \int_0^\infty (\rho + t)^{-1} X (\rho + t)^{-1} \mathrm{d}t.
\end{aligned}$$

In what follows, we show that

$$\mathcal{K}_\rho \left\{ C^\dagger = C : \mathrm{Tr}(\rho C) = 0 \right\} = \left\{ B^\dagger = B : \mathrm{Tr}(B) = 0 \right\}.$$

Since $\mathrm{Tr}\left(\mathcal{K}_\rho(C)\right) = \mathrm{Tr}(\rho C)$, it follows that

$$\mathcal{K}_\rho \left\{ C^\dagger = C : \mathrm{Tr}(\rho C) = 0 \right\} \subseteq \left\{ B^\dagger = B : \mathrm{Tr}(B) = 0 \right\}.$$

Now let $B\in\{B^\dagger=B:\mathrm{Tr}(B)=0\}$. Since $\mathcal{K}_\rho$ is invertible, the equation $B=\mathcal{K}_\rho(X)$ has a unique solution: $X=\mathcal{K}_\rho^{-1}(B)$. It suffice to show $\mathrm{Tr}(\rho X)=0$. Clearly

$$\begin{aligned}\mathrm{Tr}(\rho X)&=\mathrm{Tr}\left(\rho\mathcal{K}_\rho^{-1}(B)\right)=\mathrm{Tr}\left(\rho\int_0^\infty(\rho+t)^{-1}B(\rho+t)^{-1}\mathrm{d}t\right)\\&=\int_0^\infty\mathrm{Tr}\left(\rho(\rho+t)^{-2}B\right)\mathrm{d}t=\sum_i\int_0^\infty\frac{\lambda_i}{(\lambda_i+t)^2}\mathrm{d}t\,\langle\lambda_i|B|\lambda_i\rangle\\&=\sum_i\langle\lambda_i|B|\lambda_i\rangle=\mathrm{Tr}(B)=0.\end{aligned}$$

Theorem 14.1. ([4]) *Let A and B be operators whose spectra are contained in the open right half-plane and the open left half-plane, respectively. Then the solution of the equation $AX-XB=Y$ can be expressed as*

$$X=\int_0^\infty e^{-tA}Ye^{tB}\mathrm{d}t. \tag{14.3}$$

Proof. It is easy to see that the hypotheses ensure that the integral given above is convergent. If X is the operator defined by this integral, then

$$AX-XB=\int_0^\infty\left(Ae^{-tA}Ye^{tB}-e^{-tA}Ye^{tB}B\right)\mathrm{d}t=-e^{-tA}Ye^{tB}\big|_0^\infty=Y.$$

So X is indeed the solution of the equation. □

Theorem 14.2. ([4]) *Let A,B be Hermitian operators whose spectra are disjoint from each other. Let f be any function in $L^1(\mathbb{R})$ such that $\widehat{f}(\lambda)=\frac{1}{\lambda}$ whenever $\lambda\in\mathrm{Spec}(A)-\mathrm{Spec}(B)$. Then the solution of the equation $AX-XB=Y$ can be expressed as*

$$X=\int_{-\infty}^{+\infty}e^{-\mathrm{i}tA}Ye^{\mathrm{i}tB}f(t)\mathrm{d}t. \tag{14.4}$$

Proof. Let a and b be eigenvalues of A and B with eigenvectors $|a\rangle$ and $|b\rangle$, respectively. Then, using the fact that $e^{\mathrm{i}tA}$ is unitary and its adjoint is $e^{-\mathrm{i}tA}$, we see that

$$\left\langle a,Ae^{-\mathrm{i}tA}Ye^{\mathrm{i}tB}b\right\rangle=\left\langle e^{\mathrm{i}tA}Aa,Ye^{\mathrm{i}tB}b\right\rangle=e^{\mathrm{i}t(b-a)}a\langle a,Yb\rangle. \tag{14.5}$$

A similar consideration shows that

$$\left\langle a,e^{-\mathrm{i}tA}Ye^{\mathrm{i}tB}Bb\right\rangle=e^{\mathrm{i}t(b-a)}b\langle a,Yb\rangle. \tag{14.6}$$

Hence, if X is given by (14.4), we have

$$\begin{aligned}\langle a,(AX-XB)b\rangle&=(a-b)\langle a,Yb\rangle\int_{-\infty}^{+\infty}e^{\mathrm{i}t(b-a)}f(t)\mathrm{d}t\\&=(a-b)\langle a,Yb\rangle\widehat{f}(a-b)=\langle a,Yb\rangle.\end{aligned} \tag{14.7}$$

Since eigenvectors of A and B both span the whole space, this shows that $AX-XB=Y$. □

15 Itzykson-Züber integral formula

Theorem 15.1 (Itzykson-Züber [12])**.** *Given two $n\times n$ Hermitian matrices A and B. Let $a=\{a_1,\ldots,a_n\}$ and $b=\{b_1,\ldots,b_n\}$ be spectrum of A and B, respectively. It holds that*

$$\int \mathrm{d}\mu_{\mathrm{Haar}}(U)e^{\mathrm{i}\mathrm{Tr}\left(AUBU^\dagger\right)}=\mathrm{i}^{-\binom{n}{2}}\prod_{k=1}^{n-1}k!\frac{\det\left(e^{\mathrm{i}a_ib_j}\right)_{1\leqslant i,j\leqslant n}}{\Delta(a)\Delta(b)}, \tag{15.1}$$

where $\Delta(a)=\prod_{1\leqslant i<j\leqslant n}(a_j-a_i)$, and similar for $\Delta(b)$.

Theorem 15.2 (Mehta [15])**.** *Given two $n\times n$ Hermitian matrices A and B. Let $a=\{a_1,\ldots,a_n\}$ and $b=\{b_1,\ldots,b_n\}$ be spectrum of A and B, respectively. It holds that*

$$\int \mathrm{d}\mu_{\mathrm{Haar}}(U)e^{-\frac{1}{2t}\mathrm{Tr}\left((A-UBU^\dagger)^2\right)}=t^{\binom{n}{2}}\prod_{k=1}^{n}k!\frac{\det\left(e^{-\frac{1}{2t}(a_i-b_j)^2}\right)_{1\leqslant i,j\leqslant n}}{\Delta(a)\Delta(b)}, \tag{15.2}$$

where $\Delta(a)=\prod_{1\leqslant i<j\leqslant n}(a_j-a_i)$, and similar for $\Delta(b)$.

Acknowledgements

This project is supported by Research Fund, Kumoh National Institute of Technology.

References

[1] J. Antezanaa, G. Larotondaa, A. Varelaa, Thompson-type formulae, *Journal of Functional Analysis*, **262**(4)(2012), 1515-1528.
[2] J. Antezanaa, G. Larotondaa, A. Varelaa, Optimal Paths for Symmetric Actions in the Unitary Group, *Communications in Mathematical Physics*, **328**(2)(2014), 481-497.
[3] N. Bebianoa, R. Lemosb, J. da Providência, Inequalities for quantum relative entropy, *Linear Algebra and Its Applications*, **401**(2005), 159-72.
[4] R. Bhatia, Matrix Analysis, Spring-Verlag, 1997.
[5] R. Bhatia, The bipolar decomposition, *Linear Algebra and Its Applications*, **439**(10)(2013), 3031-3037.
[6] J.C. Bourin, Y. Seo, Reverse inequality to Golden-Thompson type inequalities: Comparison of e^{A+B} and e^Ae^B, *Linear Algebra and Its Applications*, **426**(2-3)(2007), 312-316.
[7] E.A. Carlen, E.H. Lieb, Remainder terms for some quantum entropy inequalities, *Journal Mathematical Physics*, **55**(2014), 042201.
[8] P.J. Forrester, C.J. Thompson, The Golden-Thompson inequality: Historical aspects and random matrix applications, *Journal of Mathematical Physics*, **55**(2014), 023503.
[9] F. Hansen, Multivariate extensions of the Golden-Thompson inequality, *Annals of Functional Analysis*, **6**(4)(2015), 301-310.
[10] F. Hiai, D. Petz, Introduction to Matrix Analysis and Applications, Spring-Verlag, 2014.
[11] R. Howe, Very Basic Lie Theory, *The American Mathematical Monthly*, **90**(9)(1983), 600-623.

[12] C. Itzykson, J.B. Züber, The planar approximation. II, *Journal of Mathematical Physics*, **21**(3)(1980), 411-421.
[13] A.A. Klyachko, Random walks on symmetric spaces and inequalities for matrix spectra, *Linear Algebra and Its Applications*, **319**(2000), 37-59.
[14] P.D. Lax, Linear Algebra and Its Applications, *John Wiley& Sons, Inc.*, Hoboken, New Jersey. Second Edition, 2007.
[15] M.L. Mehta, Random Matrices, *Elsevier Academic*, Press Third Edition, 2004.
[16] D. Petz, Matrix Analysis with some Applications, *Department of Mathematical Analysis*, Budapest University of Technology and Economics, 2011.
[17] H.R. Stahl, Proof of the BMV conjecture, *Acta Mathematica*, **211**(2013), 255-290.
[18] N. Sharma, More on a trace inequality in quantum information theory, arXiv:1512.00226.
[19] D. Sutter, M. Berta, M. Tomamichel, Multivariate trace inequalities, arXiv:1604.03023
[20] M. Suzuki, Transfer-matrix method and Monte Carlo simulation in quantum spin systems, *Physical Review B*, **31**(1985), 29-57.
[21] R.C. Thompson, Proof of a conjectured exponential formula, *Linear and Multilinear Algebra*, **19**(1986), 187-197.
[22] A.J. Tropp, User-friendly tools for random matrices: An introduction, available at http: users.cms.caltech.edu/ jtropp/notes/Tro12-User-Friendly-Tools-NIPS.pdf, Dec. 2012.
[23] W. So, The high road to an exponential formula, *Linear Algebra and Its Applications*, **379**(2004), 69-75.
[24] L. Zhang, Matrix integrals over unitary groups: An application of Schur-Weyl duality, arXiv:1408.3782.
[25] L. Zhang, Volumes of orthogonal groups and unitary groups, arXiv:1509.00537.
[26] L. Zhang, A strengthened monotonicity inequality of quantum relative entropy: A unifying approach via Rényi relative entropy, *Letters in Mathematical Physics*, **106**(4)(2016), 557-573.
[27] L. Zhang, L. Chen, K. Bu, Fidelity between a bipartite state and another one undergoing local unitary dynamics, *Quantum Information Processing*, **14**(2015), 4715-4730.
[28] L. Zhang, S.M. Fei, Quantum fidelity and relative entropy between unitary orbits, *Journal of Physics A: Mathematical and Theoretical*, **47**(5)(2014), 055301.
[29] L. Zhang, J.D. Wu, A lower bound of quantum conditional mutual information, *Journal of Physics A: Mathematical and Theoretical*, **47**(41)(2014), 415303.