

Matrix Analysis with some Applications



Dénes Petz

Department of Mathematical Analysis
Budapest University of Technology and Economics

Homepage: http://www.math.bme.hu/~petz

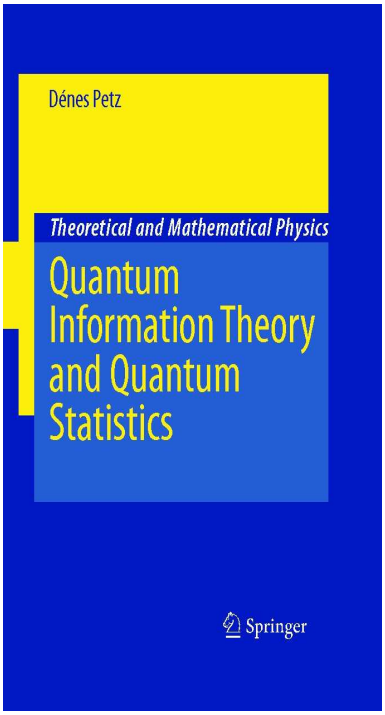


Az elsőéves matematikus és fizikus hallgatók tanulnak lineáris algebrát. Azt gondolom, hogy ez a tárgy szükséges a megértéshez. (Ha valakinek hiányossága van az alapozással, akkor belenézhet *Freud Róbert* jól érthető könyvébe.) Ebben a jegyzetben majdnem kizárólag komplex elemű mátrixokról van szó. *Rózsa Pál* könyvében valós mátrixok is vannak (és a mátrixok gyakrabban le vannak rajzolva, mint itt).

Ha egy olvasó tanult funkcionálanalízist, az kétségtelenül kedvező, mert számára a Hilbert-tér formalizmus nem új. A jegyzetben a véges dimenziós Hilbert-tér szerepel csak, a fogalmak megtalálhatók. (Ha valakit több minden érdekel az operátorok és a Hilbert-terek témakörében, akkor javasolni tudom *Kérchy László* könyveit és az én *Lineáris anaízis* könyvemet.)

A tartalomjegyzék megmutatja a témaköröket. Minden fejezet vége tartalmaz feladatokat, azok általában nem triviálisak, de nem is különösen nehezek. A nehezebbeknél van egy kis útmutatás is.

A mátrixanalízisnek alkalmazásai előfordulnak. Néha a valószínűségszámításban, statisztikában, de leginkább a kvantumelmétettel kapcsolatban. Jóval több minden szerepel az alábbi könyvben:



Contents

1 Matrices and operators **5**

1.1 Basics on matrices 5
1.2 Hilbert space 6
1.3 Notes and remarks 14
1.4 Exercises 14

2 Square matrices **17**

2.1 Jordan canonical form 17
2.2 Spectrum and eigenvalues 19
2.3 Trace and determinant 23
2.4 Notes and remarks 26
2.5 Exercises 26

3 Tensor products **28**

3.1 Algebraic tensor product 28
3.2 Tensor product of linear mappings 30
3.3 Symmetric and antisymmetric tensor powers 32
3.4 Notes and remarks 35
3.5 Exercises 35

4 Positive matrices **37**

4.1 Positivity and square root 37
4.2 Relation to tensor product 41
4.3 Partial ordering 42
4.4 Hadamard product 44
4.5 Lattice of projections 44
4.6 Kernel functions 46
4.7 Notes and remarks 48
4.8 Exercises 48

5 Functional calculus for matrices **51**

5.1 The exponential function . . . 51
5.2 Other functions . . . 56
5.3 Derivation . . . 61
5.4 Notes and remarks . . . 67
5.5 Exercises . . . 67

6 Block matrices **70**
6.1 Direct sum of Hilbert spaces . . . 70
6.2 Positivity and factorization . . . 71
6.3 Cramér-Rao inequality . . . 76
6.4 Notes and remarks . . . 78
6.5 Exercises . . . 78

7 Geometric mean **80**
7.1 Motivation by a Riemannian manifold . . . 80
7.2 Geometric mean of two matrices . . . 82
7.3 Geometric mean for more matrices . . . 84
7.4 Notes and remarks . . . 86
7.5 Exercises . . . 87

8 Convexity **89**
8.1 Convex sets . . . 89
8.2 Convex functionals . . . 90
8.3 Matrix convex functions . . . 96
8.4 Notes and remarks . . . 99
8.5 Exercises . . . 99

9 Matrix monotone functions **101**
9.1 Examples . . . 101
9.2 Characterization and properties . . . 102
9.3 Matrix means . . . 105
9.4 Quasi-entropies . . . 110
9.5 Notes and remarks . . . 112
9.6 Exercises . . . 112

10 Matrices in quantum theory **114**
10.1 Axioms . . . 114
10.2 State Transformations . . . 123
10.3 Bipartite systems . . . 131
10.4 Dense coding and teleportation . . . 139

10.5 Notes and remarks 143
10.6 Exercises 144

Index **148**

Bibliography **152**

Chapter 1

Matrices and operators

1.1 Basics on matrices

For $n, m \in \mathbb{N}$, $\mathbb{M}_{n\times m} = \mathbb{M}_{n\times m}(\mathbb{C})$ denotes the space of all $n \times m$ complex matrices. A matrix $M \in \mathbb{M}_{n\times m}$ is a mapping $\{1, 2, \ldots, n\} \times \{1, 2, \ldots, m\} \to \mathbb{C}$. It is represented as an array with n rows and m columns:

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1m} \\ m_{21} & m_{22} & \cdots & m_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nm} \end{bmatrix}$$

m_{ij} is the intersection of the ith row and the jth column. If the matrix is denoted by M, then this entry is denoted by M_{ij}. If $n = m$, then we write $\mathbb{M}_n$ instead of $\mathbb{M}_{n\times n}$. A simple example is the **identity matrix** $I_n \in \mathbb{M}_n$ defined as $m_{ij} = \delta_{i,j}$, or

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

$\mathbb{M}_{n\times m}$ is complex vector space of dimension nm. The linear operations are defined as follows.

$$[\lambda A]_{ij} := \lambda A_{ij}, \qquad [A + B]_{ij} = A_{ij} + B_{ij}.$$

(Here λ is a complex number and $A, B \in \mathbb{M}_{n\times m}$.)

Example 1.1 For $i, j = 1, \ldots, n$ let $E(ij)$ be the $n \times n$ matrix such that (i, j)-entry equals to one and all other entries equal to zero. Then $E(ij)$ are called **matrix-units** and form a basis of $\mathbb{M}_n$.

$$A = \sum_{i,j=1}^{n} A_{ij} E(ij).$$

Furthermore,

$$I_n = \sum_{i=1}^{n} E(ii)\,.$$

is a representation of the identity matrix. □

If $A \in \mathbb{M}_{n\times m}$ and $B \in \mathbb{M}_{m\times k}$, then the **product** AB of A and B is defined by

$$[AB]_{ij} = \sum_{\ell=1}^{m} A_{i\ell}B_{\ell j}.$$

if $1 \le i \le n$ and $1 \le j \le k$. Hence $AB \in \mathbb{M}_{n\times k}$.

For $A, B \in \mathbb{M}_n$, the product AB is defined, so $\mathbb{M}_n$ becomes an algebra. The most significant feature of matrices is the non-commutativity of the product: $AB \neq BA$. For example,

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

In the matrix algebra $\mathbb{M}_n$, the identity matrix I_n behaves as a unit: $I_nA = AI_n = A$ for every $A \in \mathbb{M}_n$. The matrix $A \in \mathbb{M}_n$ is **invertible** if there is a $B \in \mathbb{M}_n$ such that $AB = BA = I_n$. More generally, $A \in \mathbb{M}_{n\times m}$ is invertible if there is a $B \in \mathbb{M}_{m\times n}$ such that $AB = I_n$ and $BA = I_m$. This B is called the **inverse** of A, in notation A^{-1}.

The **transpose** A^t of the matrix $A \in \mathbb{M}_{n\times m}$ is an $m \times n$ matrix,

$$[A^t]_{ij} = A_{ji} \qquad (1 \le i \le m, 1 \le j \le n).$$

It is easy to see if the product AB is defined, then $(AB)^t = B^tA^t$. The **adjoint matrix** A^* is the complex conjugate of the transpose A^t. The space $\mathbb{M}_n$ is a *-algebra:

$$\begin{aligned} &(AB)C = A(BC), \quad (A+B)C = AC + BC, \quad A(B+C) = AB + AC, \\ &(A+B)^* = A^* + B^*, \quad (\lambda A)^* = \bar{\lambda}A^*, \quad (A^*)^* = A, \quad (AB)^* = B^*A^*. \end{aligned}$$

Let $A \in \mathbb{M}_n$. The **trace** of A is the sum of the diagonal entries:

$$\operatorname{Tr} A := \sum_{i=1}^{n} A_{ii}. \tag{1.1}$$

It is easy to show that $\operatorname{Tr} AB = \operatorname{Tr} BA$.

The **determinant** of $A \in \mathbb{M}_n$ is slightly more complicated.

$$\det A := \sum_{\pi} (-1)^{\sigma(\pi)} A_{1\pi(1)}A_{2\pi(2)} \dots A_{n\pi(n)} \tag{1.2}$$

where the sum is over all permutations π of the set $\{1, 2, \dots, n\}$ and $\sigma(\pi)$ is the parity of the permutation π. It can be proven that

$$\det(AB) = (\det A)(\det B). \tag{1.3}$$

More details about the determinant appear later.

1.2 Hilbert space

Let $\mathcal{H}$ be a complex vector space. A functional $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ of two variables is called **inner product** if

(1) $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle \quad (x,y,z\in\mathcal{H})$,

(2) $\langle \lambda x,y\rangle = \overline{\lambda}\langle x,y\rangle \quad (\lambda\in\mathbb{C},\ x,y\in\mathcal{H})$,

(3) $\langle x,y\rangle = \overline{\langle y,x\rangle} \quad (x,y\in\mathcal{H})$,

(4) $\langle x,x\rangle \geq 0$ for every $x\in\mathcal{H}$ and $\langle x,x\rangle = 0$ only for $x=0$.

These conditions imply the **Schwarz inequality**

$$|\langle x,y\rangle|^2 \leq \langle x,x\rangle\,\langle y,y\rangle. \tag{1.4}$$

The inner product determines a **norm**

$$\|x\| := \sqrt{\langle x,x\rangle} \tag{1.5}$$

which has the properties

$$\|x+y\| \leq \|x\| + \|y\| \qquad \text{and} \qquad |\langle x,y\rangle| \leq \|x\|\cdot\|y\|\,.$$

$\|x\|$ is interpreted as the length of the vector x. A further requirement in the definition of a Hilbert space is that every Cauchy sequence must be convergent, that is, the space is **complete**. (In the finite dimensional case, the completeness always holds.)



David Hilbert (1862-1943)

Hilbert had an essential role in the creation of functional analysis. For him a "Hilbert space" was a kind of infinite sequence. Actually the abstract definition was in the formulation of von Neumann.

The linear space $\mathbb{C}^n$ of all n-tuples of complex numbers becomes a Hilbert space with the inner product

$$\langle x, y\rangle = \sum_{i=1}^{n} \overline{x}_i y_i = [\overline{x}_1, \overline{x}_2, \dots, \overline{x}_n] \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix},$$

where $\overline{z}$ denotes the complex conjugate of the complex number $z \in \mathbb{C}$. Another example is the space of square integrable complex-valued function on the real Euclidean space $\mathbb{R}^n$. If f and g are such functions then

$$\langle f, g\rangle = \int_{\mathbb{R}^n} \overline{f(x)}\, g(x)\, dx$$

gives the inner product. The latter space is denoted by $L^2(\mathbb{R}^n)$ and it is infinite dimensional contrary to the n-dimensional space $\mathbb{C}^n$. Below we are mostly satisfied with finite dimensional spaces.

If $\langle x, y\rangle = 0$ for vectors x and y of a Hilbert space, then x and y are called **orthogonal**, in notation $x \perp y$. When $H \subset \mathcal{H}$, then $H^\perp := \{x \in \mathcal{H} : x \perp h \text{ for every } h \in H\}$. For any subset $H \subset \mathcal{H}$ the orthogonal complement $H^\perp$ is a closed subspace.

A family $\{e_i\}$ of vectors is called **orthonormal** if $\langle e_i, e_i\rangle = 1$ and $\langle e_i, e_j\rangle = 0$ if $i \neq j$. A maximal orthonormal system is called **basis** or orthonormal basis. The cardinality of a basis is called the dimension of the Hilbert space. (The cardinality of any two bases is the same.)

In the space $\mathbb{C}^n$, the standard orthonormal basis consists of the vectors

$$\delta_1 = (1, 0, \dots, 0), \quad \delta_2 = (0, 1, 0, \dots, 0), \quad \dots, \quad \delta_n = (0, 0, \dots, 0, 1), \tag{1.6}$$

each vector has 0 coordinate $n - 1$ times and one coordinate equals 1.

Example 1.2 The space $\mathbb{M}_n$ of matrices becomes Hilbert space with the inner product

$$\langle A, B\rangle = \operatorname{Tr} A^* B \tag{1.7}$$

which is called **Hilbert–Schmidt inner product**. The matrix units $E(ij)$ $(1 \leq i, j \leq n)$ form an orthormal basis.

It follows that the **Hilbert–Schmidt norm**

$$\|A\|_2 := \sqrt{\langle A, A\rangle} = \sqrt{\operatorname{Tr} A^* A} = \left(\sum_{i,j=1}^{n} |A_{ij}|^2 \right)^{1/2} \tag{1.8}$$

is a norm for the matrices. □

Assume that in an n dimensional Hilbert space linearly independent vectors $\{v_1, v_2, \dots, v_n\}$ are given. By the **Gram-Schmidt procedure** an orthonormal basis can be obtained by linear combinations:

$$e_1 := \frac{1}{\|v_1\|} v_1,$$

$$
\begin{aligned}
e_2 &:= \frac{1}{\|w_2\|}w_2 \quad \text{with} \quad w_2 := v_2 - \langle e_1, v_2\rangle e_1,\\
e_3 &:= \frac{1}{\|w_3\|}w_3 \quad \text{with} \quad w_3 := v_3 - \langle e_1, v_3\rangle e_1 - \langle e_2, v_3\rangle e_2,\\
&\vdots\\
e_n &:= \frac{1}{\|w_n\|}w_n \quad \text{with} \quad w_n := v_n - \langle e_1, v_n\rangle e_1 - \ldots - \langle e_{n-1}, v_n\rangle e_{n-1}.
\end{aligned}
$$

The next theorem tells that any vector has a unique **Fourier expansion**.

Theorem 1.1 *Let $e_1, e_2, \ldots$ be a basis in a Hilbert space $\mathcal{H}$. Then for any vector $x \in \mathcal{H}$ the expansion*

$$x = \sum_n \langle e_n, x\rangle e_n$$

holds. Moreover,

$$\|x\|^2 = \sum_n |\langle e_n, x\rangle|^2$$

Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. A mapping $A : \mathcal{H} \to \mathcal{K}$ is called linear if it preserves linear combination:

$$A(\lambda f + \mu g) = \lambda Af + \mu Ag \qquad (f, g \in \mathcal{H}, \quad \lambda, \mu \in \mathbb{C}).$$

The **kernel** and the **range** of A are

$$\ker A := \{x \in \mathcal{H} : Ax = 0\}, \qquad \operatorname{ran} A := \{Ax \in \mathcal{K} : x \in \mathcal{H}\}.$$

The dimension formula familiar in linear algebra is

$$\dim \mathcal{H} = \dim (\ker A) + \dim (\operatorname{ran} A). \tag{1.9}$$

Let $e_1, e_2, \ldots, e_n$ be a basis of the Hilbert space $\mathcal{H}$ and $f_1, f_2, \ldots, f_m$ be a basis of $\mathcal{K}$. The linear mapping $A : \mathcal{H} \to \mathcal{K}$ is determined by the vectors Ae_k, $k = 1, 2, \ldots, n$. Furthermore, the vector Ae_k is determined by its coordinates:

$$Ae_k = c_{1,k}f_1 + c_{2,k}f_2 + \ldots + c_{n,k}f_m.$$

The numbers $c_{i,j}$, $1 \le i \le m, 1 \le j \le n$, form an $m \times n$ matrix, it is called the **matrix** of the linear transformation A with respect to the bases $(e_1, e_2, \ldots, e_n)$ and $(f_1, f_2, \ldots, f_m)$. If we want to distinguish the linear operator A from its matrix, then the latter one will be denoted by $[A]$. We have

$$[A]_{ij} = \langle f_i, Ae_j\rangle \qquad (1 \le i \le m, \quad 1 \le j \le n).$$

Note that the order of the basis vectors is important. We shall mostly consider linear operators of a Hilbert space into itself. Then only one basis is needed and the matrix of the operator has the form of a square. So a linear transformation and a basis yield a matrix. If an $n \times n$ matrix is given, then it can be always considered as a linear transformation of the space $\mathbb{C}^n$ endowed with the standard basis (1.6).

The inner product of the vectors $|x\rangle$ and $|y\rangle$ will be often denoted as $\langle x|y\rangle$, this notation, sometimes called **bra and ket**, it is popular in physics. On the other hand, $|x\rangle\langle y|$ is a linear operator which acts on the vector $|z\rangle$ as

$$(|x\rangle\langle y|)\,|z\rangle := |x\rangle\,\langle y|z\rangle \equiv \langle y|z\rangle\,|x\rangle.$$

Therefore,

$$|x\rangle\langle y| = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{bmatrix} [\overline{y}_1, \overline{y}_2, \ldots, \overline{y}_n]$$

is conjugate linear in $|y\rangle$, while $\langle x|y\rangle$ is linear.

Example 1.3 Fix a natural number n and let $\mathcal{H}$ be the space of polynomials of at most n degree. Assume that the variable of these polynomials is t and the coefficients are complex numbers. The typical elements are

$$p(t) = \sum_{i=0}^{n} u_i t^i \qquad \text{and} \qquad q(t) = \sum_{i=0}^{n} v_i t^i.$$

If their inner product is defined as

$$\langle p(t), q(t)\rangle := \sum_{i=0}^{n} \overline{u}_i v_i,$$

then $\{1, t, t^2, \ldots, t^n\}$ is an orthonormal basis.

The differentiation is a linear operator:

$$\sum_{k=0}^{n} u_k t^k \mapsto \sum_{k=0}^{n} k u_k t^{k-1}$$

In the above basis, its matrix is

$$\begin{bmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 2 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & 0 & \ldots & 0 & n \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{bmatrix}. \tag{1.10}$$

This is an **upper triangular matrix**, the (i, j) entry is 0 if $i > j$. □

Let $\mathcal{H}_1, \mathcal{H}_2$ and $\mathcal{H}_3$ be Hilbert spaces and we fix a basis in each of them. If $B : \mathcal{H}_1 \to \mathcal{H}_2$ and $A : \mathcal{H}_2 \to \mathcal{H}_3$ are linear mappings, then the composition

$$f \mapsto A(Bf) \in \mathcal{H}_3 \qquad (F \in \mathcal{H}_1)$$

is linear as well and it is denoted by AB. The matrix $[AB]$ of the composition AB can be computed from the the matrices $[A]$ and $[B]$ as follows

$$[AB]_{ij} = \sum_k [A]_{ik}[B]_{kj}. \tag{1.11}$$

The right-hand-side is defined to be the product $[A]\,[B]$ of the matrices $[A]$ and $[B]$. Then $[AB] = [A]\,[B]$ holds. It is obvious that for a $k \times m$ matrix $[A]$ and an $m \times n$ matrix $[B]$, their product $[A]\,[B]$ is a $k \times n$ matrix.

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces and we fix a basis in each of them. If $A, B : \mathcal{H}_1 \to \mathcal{H}_2$ are linear mappings, then their linear combination

$$(\lambda A + \mu B)f \mapsto \lambda(Af) + \mu(Bf)$$

is a linear mapping and

$$[\lambda A + \mu B]_{ij} = \lambda[A]_{ij} + \mu[B]_{ij}. \tag{1.12}$$

Let $\mathcal{H}$ be a Hilbert space the linear operators $\mathcal{H} \to \mathcal{H}$ form an algebra. This algebra $B(\mathcal{H})$ has a unit, the identity operator denoted by I and the product is non-commutative. Assume that $\mathcal{H}$ is n dimensional and fix a basis. Then to each linear operator $A \in B(\mathcal{H})$ an $n \times n$ matrix A is associated. The correspondence $A \mapsto [A]$ is an algebraic isomorphism from $B(\mathcal{H})$ to the algebra $M_n(\mathbb{C})$ of $n \times n$ matrices. This isomorphism shows that the theory of linear operators on an n dimensional Hilbert space is the same as the theory of $n \times n$ matrices.

Theorem 1.2 (Riesz-Fischer theorem) *Let $\phi : \mathcal{H} \to \mathbb{C}$ be a linear mapping on a finite-dimensional Hilbert space $\mathcal{H}$. Then there is a unique vector $v \in \mathcal{H}$ such that $\phi(x) = \langle v, x\rangle$ for every vector $x \in \mathcal{H}$.*

Proof: Let $e_1, e_2, \ldots, e_n$ be an orthonormal basis in $\mathcal{H}$. Then we need a vector $v \in \mathcal{H}$ such that $\phi(e_i) = \langle v, e_i\rangle$. So

$$v = \sum_i \overline{\phi(e_i)} e_i$$

will satisfy the condition. □

The linear mappings $\phi : \mathcal{H} \to \mathbb{C}$ are called functionals. If the Hilbert space is not finite dimensional, then in the previous theorem the condition $|\phi(x)| \leq c\|x\|$ should be added, where c is a positive number.

The **operator norm** of a linear operator $A : \mathcal{H} \to \mathcal{K}$ is defined as

$$\|A\| := \sup\{\|Ax\| : x \in \mathcal{H}, \|x\| = 1\}.$$

It can be shown that $\|A\|$ is finite. In addition to the common properties $\|A + B\| \leq \|A\| + \|B\|$ and $\|\lambda A\| = |\lambda|\|A\|$, the submultiplicativity

$$\|AB\| \leq \|A\|\,\|B\|$$

also holds.

If $\|A\| \leq 1$, then the operator A is called **contraction**.

The set of linear operators $\mathcal{H} \to \mathcal{H}$ is denoted by $B(\mathcal{H})$. The convergence $A_n \to A$ means $\|A - A_n\| \to 0$. In the case of finite dimensional Hilbert space the norm here can be the operator norm, but also the Hilbert-Schmidt norm. The operator norm of a matrix is not expressed explicitly by the matrix entries.

Example 1.4 Let $A \in B(\mathcal{H})$ and $\|A\| < 1$. Then $I - A$ is invertible and

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n.$$

Since

$$(I - A)\sum_{n=0}^{N} A^n = I - A^{N+1} \qquad \text{and} \qquad \|A^{N+1}\| \le \|A\|^{N+1},$$

we can see that the limit of the first equation is

$$(I - A)\sum_{n=0}^{\infty} A^n = I.$$

This shows the statement which is called **Neumann series**. □

Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. If $T : \mathcal{H} \to \mathcal{K}$ is a linear operator, then its **adjoint** $T^* : \mathcal{K} \to \mathcal{H}$ is determined by the formula

$$\langle x, Ty\rangle_{\mathcal{K}} = \langle T^*x, y\rangle_{\mathcal{H}} \qquad (x \in \mathcal{K}, y \in \mathcal{H}). \tag{1.13}$$

The operator $T \in B(\mathcal{H})$ is called **self-adjoint** if $T^* = T$. The operator T is self-adjoint if and only if $\langle x, Tx\rangle$ is a real number for every vector $x \in \mathcal{H}$. For self-adjoint operators and matrices the notations $B(\mathcal{H})^{sa}$ and $\mathbb{M}_n^{sa}$ are used.

Theorem 1.3 *The properties of the adjoint:*

(1) $(A + B)^* = A^* + B^*$, $(\lambda A)^* = \overline{\lambda} A^* \qquad (\lambda \in \mathbb{C})$,

(2) $(A^*)^* = A$, $(AB)^* = B^*A^*$,

(3) $(A^{-1})^* = (A^*)^{-1}$ *if* A *is invertible,*

(4) $\|A\| = \|A^*\|$, $\|A^*A\| = \|A\|^2$.

Example 1.5 Let $A : \mathcal{H} \to \mathcal{H}$ be a linear mapping and $e_1, e_2, \dots, e_n$ be a basis in the Hilbert space $\mathcal{H}$. The (i, j) element of the matrix of A is $\langle e_i, Ae_j\rangle$. Since

$$\langle e_i, Ae_j\rangle = \overline{\langle e_j, A^*e_i\rangle},$$

this is the complex conjugate of the (j, i) element of the matrix of A^*. □

Theorem 1.4 (Projection theorem) *Let* $\mathcal{M}$ *be a closed subspace of a Hilbert space* $\mathcal{H}$. *Any vector* $x \in \mathcal{H}$ *can be written in a unique way in the form* $x = x_0 + y$, *where* $x_0 \in \mathcal{M}$ *and* $y \perp \mathcal{M}$.

The mapping $P : x \mapsto x_0$ defined in the context of the previous theorem is called **orthogonal projection** onto the subspace $\mathcal{M}$. This mapping is linear:

$$P(\lambda x + \mu y) = \lambda Px + \mu Py.$$

Moreover, $P^2 = P = P^*$. The converse is also true: If $P^2 = P = P^*$, then P is an orthogonal projection (onto its range).

Example 1.6 The matrix $A \in \mathbb{M}_n$ is self-adjoint if $A_{ji} = \overline{A_{ij}}$. A particular example is the **Toeplitz matrix**:

$$\begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_{n-1} & a_n \\ \overline{a_2} & a_1 & a_2 & \dots & a_{n-2} & a_{n-1} \\ \overline{a_3} & \overline{a_2} & a_1 & \dots & a_{n-3} & a_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & a_3 \\ \overline{a_{n-1}} & \overline{a_{n-2}} & \overline{a_{n-3}} & \dots & a_1 & a_2 \\ \overline{a_n} & \overline{a_{n-1}} & \overline{a_{n-2}} & \dots & \overline{a_2} & a_1 \end{bmatrix}, \tag{1.14}$$

where $a_1 \in \mathbb{R}$. □

The operator $U \in B(\mathcal{H})$ is called **unitary** if U^* is the inverse of U. Then $U^*U = I$ and

$$\langle x, y\rangle = \langle U^*Ux, y\rangle = \langle Ux, Uy\rangle$$

for any vectors $x, y \in \mathcal{H}$. Therefore the unitary operators preserve the inner product. In particular, orthogonal unit vectors are mapped into orthogonal unit vectors.

Example 1.7 The **permutation matrices** are simple unitaries. Let π be a permutation of the set $\{1, 2, \dots, n\}$. The $A_{i,\pi(i)}$ entries of $A \in \mathbb{M}(\mathbb{C})$ are 1 and all others are 0. Every line and every column contain exactly one 1 entry. If such a matrix A is applied to a vector, it permutes the coordinates, this is the reason of the terminology. □

The operator $A \in B(\mathcal{H})$ is called **normal** if $AA^* = A^*A$. It follows immediately that

$$\|Ax\| = \|A^*x\| \tag{1.15}$$

for any vector $x \in \mathcal{H}$. Self-adjoint and unitary operators are normal.

The operators we need are mostly linear, but sometimes **conjugate-linear** operators appear. $\Lambda : \mathcal{H} \to \mathcal{K}$ is conjugate-linear if

$$\Lambda(\lambda x + \mu y) = \overline{\lambda}x + \overline{\mu}y$$

for any complex numbers λ and μ and for any vectors $x, y \in \mathcal{H}$. The adjoint Λ^* of the conjugate-linear operator Λ is determined by the equation

$$\langle x, \Lambda y\rangle_{\mathcal{K}} = \langle y, \Lambda^* x\rangle_{\mathcal{H}} \qquad (x \in \mathcal{K}, y \in \mathcal{H}). \tag{1.16}$$

A mapping $\phi : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ is called **complex bilinear form** if ϕ is linear in the second variable and conjugate linear in the first variables. The inner product is a particular example.

Theorem 1.5 *On a finite dimensional Hilbert space there is a one-to-one correspondence*

$$\phi(x,y) = \langle Ax, y\rangle$$

between the complex bilinear forms $\phi : \mathcal{H}\times\mathcal{H} \to \mathbb{C}$ *and the linear operators* $A : \mathcal{H} \to \mathcal{H}$.

Proof: Fix $x \in \mathcal{H}$. Then $y \mapsto \phi(x,y)$ is a linear functional. Due to the Riesz-Fischer theorem $\phi(x,y) = \langle z, y\rangle$ for a vector $z \in \mathcal{H}$. We set $Ax = z$. □

The **polarization identity**

$$4\phi(x,y) = \phi(x+y,x+y) + \mathrm{i}\phi(x+\mathrm{i}y,x+\mathrm{i}y) - \phi(x-y,x-y) - \mathrm{i}\phi(x-\mathrm{i}y,x-\mathrm{i}y) \quad (1.17)$$

shows that a complex bilinear form ϕ is determined by its so-called quadratic form $x \mapsto \phi(x,x)$.

1.3 Notes and remarks

The origins of mathematical matrices lie with the study of systems of simultaneous linear equations. Takakazu Shinsuke Seki Japanese mathematician was the first person to study determinants in 1683. Gottfried Leibnitz (1646-1716), one of the founders of calculus, used determinants in 1693 and Gabriel Cramer (1704-1752) presented his determinant-based formula for solving systems of linear equations (today known as Cramer's Rule) in 1750. A modern matrix method of solution outlined by Carl Friedrich Gauss (1777-1855) is known as Gaussian elimination. The term "matrix" was introduced in 1850 by James Joseph Sylvester (1814-1897).

1.4 Exercises

1. Show that in the Schwarz inequalty (1.4) the equality occurs if and only if x and y are linearly dependent.

2. Show that
$$\|x-y\|^2 + \|x+y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad (1.18)$$
for the norm in a Hilbert space. (This is called **parallelogram law**.)

3. Show the polarization identity (1.17).

4. Show that an orthonormal family of vectors is linearly independent.

5. Show that the vectors $|x_1\rangle, |x_2,\rangle, \ldots, |x_n\rangle$ form an orthonormal basis in an n-dimensional Hilbert space if and only if
$$\sum_i |x_i\rangle\langle x_i| = I.$$

6. Show that Gram-Schmidt procedure constructs an orthonormal basis $e_1, e_2, \ldots, e_n$. Show that e_k is the linear combination of $v_1, v_2, \ldots, v_k$ $(1 \le k \le n)$.



James Joseph Sylvester (1814-1897)

Sylvester was an English mathematician. He had results in matrix and determinant theories. He introduced the term "matrix" in 1850.

7. Show that the upper triangular matrices form an algebra.

8. Show the following properties:

$$(|u\rangle\langle v|)^* = |v\rangle\langle u|, \quad (|u_1\rangle\langle v_1|)(|u_2\rangle\langle v_2|) = \langle v_1, u_2\rangle |u_1\rangle\langle v_2|,$$
$$A(|u\rangle\langle v|) = |Au\rangle\langle v|, \quad (|u\rangle\langle v|)A = |u\rangle\langle A^*v| \quad \text{for all } A \in B(\mathcal{H}).$$

9. Let $A, B \in B(\mathcal{H})$. Show that $\|AB\| \leq \|A\| \, \|B\|$.

10. Let $\mathcal{H}$ be an n-dimensional Hilbert space. For $A \in B(\mathcal{H})$ let $\|A\|_2 := \sqrt{\operatorname{Tr} A^*A}$. Show that $\|A + B\|_2 \leq \|A\|_2 + \|B\|_2$. Is it true that $\|AB\|_2 \leq \|A\|_2 \times \|B\|_2$?

11. Find constants $c(n)$ and $d(n)$ such that

$$c(n)\|A\| \leq \|A\|_2 \leq d(n)\|A\|$$

for every matrix $A \in \mathbb{M}_n(\mathbb{C})$.

12. Show that $\|A^*A\| = \|A\|^2$ for every $A \in B(\mathcal{H})$.

13. Let $\mathcal{H}$ be an n-dimensional Hilbert space. Show that given an operator $A \in B(\mathcal{H})$ we can chose an orthonormal basis such that the matrix of A is upper triangular.

14. Let $A, B \in \mathbb{M}_n$ be invertible matrices. Show that $A + B$ is invertible if and only if $A^{-1} + B^{-1}$ is invertible, moreover

$$(A + B)^{-1} = A^{-1} - A^{-1}(A^{-1} + B^{-1})^{-1}A^{-1}.$$

15. Let $A \in \mathbb{M}_n$ be self-adjoint. Show that

$$U = (I - \mathrm{i}A)(I + \mathrm{i}A)^{-1}$$

is a unitary. (U is the **Cayley transform** of A.)

Chapter 2

Square matrices

The $n \times n$ matrices $\mathbb{M}_n$ can be identified with the linear operators $B(\mathcal{H})$ where the Hilbert space $\mathcal{H}$ is n-dimensional. To make a precise identification an orthonormal basis should be fixed in $\mathcal{H}$.

2.1 Jordan canonical form

A **Jordan block** is a matrix

$$J_k(a) = \begin{bmatrix} a & 1 & 0 & \cdots & 0 \\ 0 & a & 1 & \cdots & 0 \\ 0 & 0 & a & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a \end{bmatrix}, \tag{2.1}$$

where $a \in \mathbb{C}$. This is an upper triangular matrix $J_k(a) \in \mathbb{M}_k$. We use also the notation $J_k := J_k(0)$. Then

$$J_k(a) = aI_k + J_k \tag{2.2}$$

and the sum consists of commuting matrices.

Example 2.1 The matrix J_k is

$$(J_k)_{ij} = \begin{cases} 1 & \text{if} \quad j = i+1, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore

$$(J_k)_{ij}(J_k)_{jk} = \begin{cases} 1 & \text{if} \quad j = i+1 \text{ and } k = i+2, \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$(J_k^2)_{ij} = \begin{cases} 1 & \text{if} \quad j = i+2, \\ 0 & \text{otherwise.} \end{cases}$$

We observe that taking the powers of J_k the line of the 1 entries is going upper, in particular $J_k^k = 0$. The matrices $\{J_k^m : 0 \le m \le k-1\}$ are linearly independent.

If $a \neq 0$, then $\det J_k(a) \neq 0$ and $J_k(a)$ is invertible. We can search for the inverse by the equation

$$(aI_k + J_k)\left(\sum_{j=0}^{k-1} c_j J_k^j\right) = I_k.$$

Rewriting this equation we get

$$ac_0 I_k + \sum_{j=1}^{k-1} (ac_j + c_{j-1}) J_k^j = I_k.$$

The solution is

$$c_j = -(-a)^{-j-1} \qquad (0 \le j \le k-1).$$

In particular,

$$\begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}^{-1} = \begin{bmatrix} a^{-1} & -a^{-2} & a^{-3} \\ 0 & a^{-1} & -a^{-2} \\ 0 & 0 & a^{-1} \end{bmatrix}.$$

Computation with a Jordan block is convenient. □

The **Jordan canonical form theorem** is the following.

Theorem 2.1 *Given a matrix $X \in \mathbb{M}_n$, there is an invertible matrix $S \in \mathbb{M}_n$ such that*

$$X = S \begin{bmatrix} J_{k_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{k_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{k_m}(\lambda_m) \end{bmatrix} S^{-1} = SJS^{-1},$$

where $k_1 + k_2 + \ldots + k_m = n$. The Jordan matrix J is uniquely determined (up to the permutation of the Jordan blocks in the diagonal.)

Note that the numbers $\lambda_1, \lambda_2, \ldots, \lambda_m$ are not necessarily different. Example 2.1 showed that it is rather easy to handle a Jordan block. If the Jordan canonical decomposition is known, then the inverse can be obtained. The theorem is about complex matrices, but for $X \in \mathbb{M}_n(\mathbb{R})$, S and J is in $\mathbb{M}_n(\mathbb{R})$ as well.

Example 2.2 An essential application is concerning the determinant. Since $\det X = \det(SJS^{-1}) = \det J$, it is enough to compute the determinant of the upper-triangular Jordan matrix J. Therefore

$$\det X = \prod_{j=1}^{m} \lambda_j^{k_j}. \tag{2.3}$$

The **characteristic polynomial** of $X \in \mathbb{M}_n$ is defined as

$$p(x) := \det(xI_n - X)$$

From the computation (2.3) we have

$$p(x) = \prod_{j=1}^{m} (x - \lambda_j)^{k_j}.$$

The numbers λ_j are roots of the characteristic polynomial. □

The powers of a matrix $X \in \mathbb{M}_n$ is well-defined, therefore for a polynomial $p(x) = \sum_{k=0}^{m} c_k x^k$ the matrix $p(X)$ is defined as well. If q is a polynomial, then it is **annihilating** for a matrix $X \in \mathbb{M}_n$ if $q(X) = 0$. If p is the characteristic polynomial of $X \in \mathbb{M}_n$, then $p(X) = 0$.

2.2 Spectrum and eigenvalues

Let $\mathcal{H}$ be a Hilbert space. For $A \in B(\mathcal{H})$ and $\lambda \in \mathbb{C}$, we say that λ is an **eigenvalue** of A if there is a non-zero vector $v \in \mathcal{H}$ such that $Av = \lambda v$. Such a vector v is called an **eigenvector** of A for the eigenvalue λ. If $\mathcal{H}$ is finite-dimensional, then $\lambda \in \mathbb{C}$ is an eigenvalue of A if and only if $A - \lambda I$ is not invertible.

Generally, the **spectrum** $\sigma(A)$ of $A \in B(\mathcal{H})$ consists of the numbers $\lambda \in \mathbb{C}$ such that $A - \lambda I$ is not invertible. Therefore in the finite-dimensional case the spectrum is the set of eigenvalues.

Example 2.3 In the history of matrix theory the particular matrix

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} \tag{2.4}$$

has importance. Its eigenvalues were computed by **Joseph Louis Lagrange** in 1759. He found that the eigenvalues are $2\cos j\pi/(n+1)$ $(j = 1, 2, \dots, n)$. □

The matrix (2.4) is **tridiagonal**. This means that $A_{ij} = 0$ if $|i - j| > 1$.

Example 2.4 Let $\lambda \in \mathbb{R}$ and consider the matrix

$$J_3(\lambda) = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}. \tag{2.5}$$

Now λ is the only eigenvalue and $(y, 0, 0)$ is the only eigenvector. The situation is similar in the $k \times k$ generalization $J_k(\lambda)$: λ is the eigenvalue of $SJ_k(\lambda)S^{-1}$ for an arbitrary invertable S and there is one eigenvector (up to constant multiple). This has the consequence that the characteristic polynomial gives the eigenvalues without the multiplicity.

If X has the Jordan form as in Theorem 2.1, then all λ_j's are eigenvalues. Therefore the roots of the characteristic polynomial are eigenvalues.

For the above $J_3(\lambda)$ we can see that

$$J_3(\lambda)(0, 0, 1) = (0, 1, \lambda), \qquad J_3(\lambda)^2(0, 0, 1) = (1, 2\lambda, \lambda^2),$$

therefore $(0, 0, 1)$ and these two vectors linearly span the whole space $\mathbb{C}^3$. The vector $(0, 0, 1)$ is called **cyclic vector**.

When we want formal matrix product, then

$$J_3(\lambda) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \lambda \end{bmatrix}$$

is the correct notation, or $J_3(\lambda)(0, 0, 1)^t = (0, 1, \lambda)^t$ is written by the use of transform.

Assume that a matrix $X \in \mathbb{M}_n$ has a cyclic vector $v \in \mathbb{C}^n$ which means that the set $\{v, Xv, X^2v, \dots, X^{n-1}v\}$ spans $\mathbb{C}^n$. Then $X = SJ_n(\lambda)S^{-1}$ with some invertible matrix S, the Jordan canonical form is very simple. □

Theorem 2.2 *Assume that $A \in B(\mathcal{H})$ is normal. Then there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ and $u_1, \ldots, u_n \in \mathcal{H}$ such that $\{u_1, \ldots, u_n\}$ is an orthonormal basis of $\mathcal{H}$ and $Au_i = \lambda_i u_i$ for all $1 \leq i \leq n$.*

Proof: Let us prove by induction on $n = \dim \mathcal{H}$. The case $n = 1$ trivially holds. Suppose the assertion for dimension $n-1$. Assume that $\dim \mathcal{H} = n$ and $A \in B(\mathcal{H})$ is normal. Choose a root λ_1 of $\det(\lambda I - A) = 0$. As explained before the theorem, λ_1 is an eigenvalue of A so that there is an eigenvector u_1 with $Au_1 = \lambda_1 u_1$. One may assume that u_1 is a unit vector, i.e., $\|u_1\| = 1$. Since A is normal, we have

$$\begin{aligned}(A-\lambda_1 I)^*(A-\lambda_1 I) &= (A^*-\overline{\lambda}_1 I)(A-\lambda_1 I)\\ &= A^*A - \overline{\lambda}_1 A - \lambda_1 A^* + \lambda_1\overline{\lambda}_1 I\\ &= AA^* - \overline{\lambda}_1 A - \lambda_1 A^* + \lambda_1\overline{\lambda}_1 I\\ &= (A-\lambda_1 I)(A-\lambda_1 I)^*,\end{aligned}$$

that is, $A - \lambda_1 I$ is also normal. Therefore,

$$\|(A^*-\overline{\lambda}_1 I)u_1\| = \|(A-\lambda_1 I)^* u_1\| = \|(A-\lambda_1 I)u_1\| = 0$$

so that $A^*u_1 = \overline{\lambda}_1 u_1$. Let $\mathcal{H}_1 := \{u_1\}^\perp$, the orthogonal complement of $\{u_1\}$. If $x \in \mathcal{H}_1$ then

$$\begin{aligned}\langle Ax, u_1\rangle = \langle x, A^*u_1\rangle = \langle x, \overline{\lambda}_1 u_1\rangle = \overline{\lambda}_1\langle x, u_1\rangle = 0,\\ \langle A^*x, u_1\rangle = \langle x, Au_1\rangle = \langle x, \lambda_1 u_1\rangle = \lambda_1\langle x, u_1\rangle = 0\end{aligned}$$

so that $Ax, A^*x \in \mathcal{H}_1$. Hence we have $A\mathcal{H}_1 \subset \mathcal{H}_1$ and $A^*\mathcal{H}_1 \subset \mathcal{H}_1$. So one can define $A_1 := A|_{\mathcal{H}_1} \in B(\mathcal{H}_1)$. Then $A_1^* = A^*|_{\mathcal{H}_1}$, which implies that A_1 is also normal. Since $\dim \mathcal{H}_1 = n-1$, the induction hypothesis can be applied to obtain $\lambda_2, \ldots, \lambda_n \in \mathbb{C}$ and $u_2, \ldots, u_n \in \mathcal{H}_1$ such that $\{u_2, \ldots, u_n\}$ is an orthonormal basis of $\mathcal{H}_1$ and $A_1 u_i = \lambda_i u_i$ for all $i = 2, \ldots, n$. Then $\{u_1, u_2, \ldots, u_n\}$ is an orthonormal basis of $\mathcal{H}$ and $Au_i = \lambda_i u_i$ for all $i = 1, 2, \ldots, n$. Thus the assertion holds for dimension n as well. □

It is an important consequence that the matrix of a normal operator is diagonal in an appropriate orthonormal basis and the trace is the sum of the eigenvalues.

Theorem 2.3 *Assume that $A \in B(\mathcal{H})$ is self-adjoint. If $Av = \lambda v$ and $Aw = \mu w$ with non-zero eigenvectors v, w and the eigenvalues λ and μ are different, then $v \perp w$ and $\lambda, \mu \in \mathbb{R}$.*

Proof: First we show that the eigenvalues are real:

$$\lambda\langle v, v\rangle = \langle v, \lambda v\rangle = \langle v, Av\rangle = \langle Av, v\rangle = \langle \lambda v, v\rangle = \overline{\lambda}\langle v, v\rangle.$$

The $\langle v, w\rangle = 0$ orthogonality comes similarly:

$$\mu\langle v, w\rangle = \langle v, \mu w\rangle = \langle v, Aw\rangle = \langle Av, w\rangle = \langle \lambda v, w\rangle = \lambda\langle v, w\rangle.$$

□

If A is a self-adjoint operator on an n-dimensional Hilbert space, then from the eigenvectors we can find an orthonormal basis $v_1, v_2, \ldots, v_n$. If $Av_i = \lambda_i v_i$, then

$$A = \sum_{i=1}^{n} \lambda_i |v_i\rangle\langle v_i| \tag{2.6}$$

which is called the **Schmidt decomposition**. The Schmidt decomposition is unique if all the eigenvalues are different, otherwise not. Another useful decomposition is the **spectral decomposition**. Assume that the self-adjoint operator A has the eigenvalues $\mu_1 > \mu_2 > \ldots > \mu_k$. Then

$$A = \sum_{j=1}^{k} \mu_j P_j, \tag{2.7}$$

where P_j is the orthogonal projection onto the subspace spanned by the eigenvectors with eigenvalue μ_j. From the Schmidt decomposition (2.6),

$$P_j = \sum_i |v_i\rangle\langle v_i|,$$

where the summation is over all i such that $\lambda_i = \mu_j$. This decomposition is always unique. Actually, the Schmidt decomposition and the spectral decomposition exist for all normal operators.

If $\lambda_i \geq 0$ in (2.6), then we can set $|x_i\rangle := \sqrt{\lambda_i}|v_i\rangle$ and we have

$$A = \sum_{i=1}^{n} |x_i\rangle\langle x_i|.$$

If the orthogonality of the vectors $|x_i\rangle$ is not assumed, then there are several similar decompositions, but they are connected by a unitary. The next lemma and its proof is a good exercise for the bra and ket formalism. (The result and the proof is due to **Schrödinger** [51].)

Lemma 2.1 *If*

$$A = \sum_{j=1}^{n} |x_j\rangle\langle x_j| = \sum_{i=1}^{n} |y_i\rangle\langle y_i|,$$

then there exists a unitary matrix $(U_{ij})_{i,j=1}^{n}$ *such that*

$$\sum_{j=1}^{n} U_{ij}|x_j\rangle = |y_i\rangle. \tag{2.8}$$

Proof: Assume first that the vectors $|x_j\rangle$ are orthogonal. Typically they are not unit vectors and several of them can be 0. Assume that $|x_1\rangle, |x_2\rangle, \ldots, |x_k\rangle$ are not 0 and $|x_{k+1}\rangle = \ldots = |x_n\rangle = 0$. Then the vectors $|y_i\rangle$ are in the linear span of $\{|x_j\rangle : 1 \leq j \leq k\}$, therefore

$$|y_i\rangle = \sum_{j=1}^{k} \frac{\langle x_j|y_i\rangle}{\langle x_j|x_j\rangle}|x_j\rangle$$

is the orthogonal expansion. We can define (U_{ij}) by the formula

$$U_{ij} = \frac{\langle x_j|y_i\rangle}{\langle x_j|x_j\rangle} \qquad (1 \le i \le n, 1 \le j \le k).$$

We easily compute that

$$\begin{aligned}\sum_{i=1}^{k} U_{it}U_{iu}^* &= \sum_{i=1}^{k} \frac{\langle x_t|y_i\rangle}{\langle x_t|x_t\rangle}\frac{\langle y_i|x_u\rangle}{\langle x_u|x_u\rangle} \\ &= \frac{\langle x_t|A|x_u\rangle}{\langle x_u|x_u\rangle\langle x_t|x_t\rangle} = \delta_{t,u},\end{aligned}$$

and this relation shows that the k column vectors of the matrix (U_{ij}) are orthonormal. If $k < n$, then we can append further columns to get a $n \times n$ unitary, see Exercise 13. (One can see in (2.8) that if $|x_j\rangle = 0$, then U_{ij} does not play any role.)

In the general situation

$$A = \sum_{j=1}^{n} |z_j\rangle\langle z_j| = \sum_{i=1}^{n} |y_i\rangle\langle y_i|$$

we can make a unitary U from an orthogonal family to $|y_i\rangle$'s and and a unitary V from the same orthogonal family to $|z_i\rangle$'s and UV^* goes from $|z_i\rangle$'s to $|y_i\rangle$'s. □

Example 2.5 Let $A \in B(\mathcal{H})$ be a self-adjoint operator with eigenvalues $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_n$ (counted with multiplicity). Then

$$\lambda_1 = \max\{\langle v, Av\rangle \,:\, v \in \mathcal{H},\, \|v\| = 1\}. \tag{2.9}$$

We can take the Schmidt decomposition (2.6). Assume that

$$\max\{\langle v, Av\rangle \,:\, v \in \mathcal{H},\, \|v\| = 1\} = \langle w, Aw\rangle$$

for a unit vector w. This vector has the expansion

$$w = \sum_{i=1}^{n} c_i|v_i\rangle$$

and we have

$$\langle w, Aw\rangle = \sum_{i=1}^{n} |c_i|^2\lambda_i \le \lambda_1.$$

The equality holds if and only if $\lambda_i < \lambda_1$ implies $c_i = 0$. The maximizer should be an eigenvector with eigenvalue λ_1.

Similarly,

$$\lambda_n = \min\{\langle v, Av\rangle \,:\, v \in \mathcal{H},\, \|v\| = 1\}. \tag{2.10}$$

The formulas (2.9) and (2.10) will be extended below. □

Theorem 2.4 (Poincaré's inequality) *Let $A \in B(\mathcal{H})$ be a self-adjoint operator with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ (counted with multiplicity) and let $\mathcal{K}$ be a k-dimensional subspace of $\mathcal{H}$. Then there are unit vectors $x, y \in \mathcal{K}$ such that*

$$\langle x, Ax\rangle \leq \lambda_k \qquad \textit{and} \qquad \langle y, Ay\rangle \geq \lambda_k.$$

Proof: Let $v_k, \ldots, v_n$ be orthonormal eigenvectors corresponding to the eigenvalues $\lambda_k, \ldots, \lambda_n$. They span a subspace $\mathcal{M}$ of dimension $n-k+1$ which must have intersection with $\mathcal{K}$. Take a unit vector $x \in \mathcal{K}, \mathcal{M}$ which has the expansion

$$x = \sum_{i=k}^{n} c_i v_i$$

and it has the required property:

$$\langle x, Ax\rangle = \sum_{i=k}^{n} |c_i|^2 \lambda_i \leq \lambda_k \sum_{i=k}^{n} |c_i|^2 = \lambda_k.$$

To find the other vector y, the same argument can be used with the matrix $-A$. □

The next result is a **minimax principle**.

Theorem 2.5 *Let $A \in B(\mathcal{H})$ be a self-adjoint operator with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ (counted with multiplicity). Then*

$$\lambda_k = \min\Big\{ \max\{\langle v, Av\rangle \,:\, v \in \mathcal{K},\, \|v\| = 1\} \,:\, \mathcal{K} \subset \mathcal{H},\, \dim \mathcal{K} = n+1-k \Big\}.$$

Proof: Let $v_k, \ldots, v_n$ be orthonormal eigenvectors corresponding to the eigenvalues $\lambda_k, \ldots, \lambda_n$. They span a subspace $\mathcal{K}$ of dimension $n+1-k$. According to (2.9) we have

$$\lambda_k = \max\{\langle v, Av\rangle \,:\, v \in \mathcal{K}\}$$

and it follows that in the statement of the theorem $\geq$ is true.

To complete the proof we have to show that for any subspace $\mathcal{K}$ of dimension $n+1-k$ there is a unit vector v such that $\lambda_k \leq \langle v, Av\rangle$, or $-\lambda_k \geq \langle v, (-A)v\rangle$. The decreasing eigenvalues of $-A$ are $-\lambda_n \geq -\lambda_{n-1}, \ldots, -\lambda_1$ where the ℓth is $-\lambda_{n+1-\ell}$. The existence of a unit vector v is guaranted by the Poincaré's inequality and we take ℓ with the property $n+1-\ell = k$. □

2.3 Trace and determinant

When $\{e_1, \ldots, e_n\}$ is an orthonormal basis of $\mathcal{H}$, the **trace** $\operatorname{Tr} A$ of $A \in B(\mathcal{H})$ is defined as

$$\operatorname{Tr} A := \sum_{i=1}^{n} \langle e_i, Ae_i\rangle.$$

It will be shown that the definition is independent of the choice of an orthonormal basis and $\operatorname{Tr} AB = \operatorname{Tr} BA$ for all $A, B \in B(\mathcal{H})$. We have

$$\operatorname{Tr} AB \;=\; \sum_{i=1}^{n} \langle e_i, ABe_i\rangle = \sum_{i=1}^{n} \langle A^* e_i, Be_i\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{\langle e_j, A^* e_i\rangle} \langle e_j, Be_i\rangle$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{n}\overline{\langle e_i, B^*e_j\rangle}\langle e_i, Ae_j\rangle = \sum_{j=1}^{n}\langle e_j, BAe_j\rangle = \operatorname{Tr} BA.$$

Now, let $\{f_1, \ldots, f_n\}$ be another orthonormal basis of $\mathcal{H}$. Then a unitary U is defined by $Ue_i = f_i$, $1 \le i \le n$, and we have

$$\sum_{i=1}^{n}\langle f_i, Af_i\rangle = \sum_{i=1}^{n}\langle Ue_i, AUe_i\rangle = \operatorname{Tr} U^*AU = \operatorname{Tr} AUU^* = \operatorname{Tr} A,$$

which says that the definition of $\operatorname{Tr} A$ is actually independent of the choice of an orthonormal basis.

When $A \in \mathbb{M}_n$, the trace of A is nothing but the sum of the principal diagonal entries of A:

$$\operatorname{Tr} A = A_{11} + A_{22} + \ldots + A_{nn}.$$

If A is normal, then the trace is the sum of the eigenvalues.

Computation of the trace is very simple, the case of the determinant (1.2) is very different. In the terms of the Jordan canonical form described in Theorem 2.1, we have

$$\operatorname{Tr} X = \sum_{j=1}^{m} k_j\lambda_j \qquad \text{and} \qquad \det X = \prod_{j=1}^{m}\lambda_j^{k_j}.$$

Example 2.6 On the linear space $\mathbb{M}_n$ we can define a linear mapping $\alpha : \mathbb{M}_n \to \mathbb{M}_n$ as $\alpha(A) = VAV^*$, where $V \in \mathbb{M}_n$ is a fixed matrix. We are interested in $\det\alpha$.

Let $V = SJS^{-1}$ be the canonical Jordan decomposition and set

$$\alpha_1(A) = S^{-1}A(S^{-1})^*, \qquad \alpha_2(B) = JBJ^*, \qquad \alpha_3(C) = SCS^*.$$

Then $\alpha = \alpha_3 \circ \alpha_2 \circ \alpha_1$ and $\det\alpha = \det\alpha_3 \times \det\alpha_2 \times \det\alpha_1$. Since $\alpha_1 = \alpha_3^{-1}$, we have $\det\alpha = \det\alpha_2$, so only the Jordan block part has influence to the determinant.

The following example helps to understand the situation. Let

$$J = \begin{bmatrix}\lambda_1 & x\\ 0 & \lambda_2\end{bmatrix}, \quad A_1 = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \quad A_2 = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}, \quad A_3 = \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}, \quad A_4 = \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}.$$

Then $\{A_1, A_2, A_3, A_4\}$ is a basis in $\mathbb{M}_2$. If $\alpha(A) = JAJ^*$, then from the data

$$\alpha(A_1) = \lambda_1^2A_1, \qquad \alpha(A_2) = \lambda_1xA_1 + \lambda_1\lambda_2A_2,$$

$$\alpha(A_3) = \lambda_1xA_1 + \lambda_1\lambda_2A_3, \qquad \alpha(A_4) = x^2A_1 + x\lambda_2A_2 + x\lambda_2A_3 + \lambda_2^2A_4$$

we can observe that the matrix of α is upper triangular:

$$\begin{bmatrix}\lambda_1^2 & x\lambda_1 & x\lambda_1 & x^2\\ 0 & \lambda_1\lambda_2 & 0 & x\lambda_2\\ 0 & 0 & \lambda_1\lambda_2 & x\lambda_2\\ 0 & 0 & 0 & \lambda_2^2\end{bmatrix},$$

So its determinat is the product of the diagonal entries:

$$\lambda_1^2(\lambda_1\lambda_2)(\lambda_1\lambda_2)\lambda_2^2 = \lambda_1^4\lambda_2^4 = (\det V)^4.$$

Now let $J \in \mathbb{M}_n$ and assume that only the entries J_{ii} and $J_{i,i+1}$ can be non-zero. In $\mathbb{M}_n$ we choose the basis of the matrix units,

$$E(1,1), E(1,2), \dots, E(1,n), E(2,1), \dots, E(2,n), \dots, E(3,1), \dots, E(n,n).$$

We want to see that the matrix of α is upper triangular.

From the computation

$$JE(j,k))J^* = J_{j-1,j}\overline{J_{k-1,k}}\ E(j-1,k-1) + J_{j-1,j}\overline{J_{k,k}}\ E(j-1,k)$$

$$+J_{jj}\overline{J_{k-1,k}}\ E(j,k-1) + J_{jj}\overline{J_{k,k}}\ E(j,k)$$

we can see that the matrix of the mapping $A \mapsto JAJ^*$ is upper triangular. (In the lexicographical order of the matrix units $E(j-1,k-1), E(j-1,k), E(j,k-1)$ are before $E(j,k)$.) The determinant is the product of the diagonal entries:

$$\prod_{j,k=1}^{n} J_{jj}\overline{J_{kk}} = \prod_{k=1}^{m}(\det J)\overline{J_{kk}}^n = (\det J)^n\overline{\det J}^n$$

It follows that the determinant of $\alpha(A) = VAV^*$ is $(\det V)^n\overline{\det V}^n$, since the determinant of V equals to the determinant of its block matrix. If $\beta(A) = VAV^t$, then the argument is similar, $\det\beta = (\det V)^{2n}$, only the conjugate is missing.

Next we deal with the space $\mathcal{M}$ of real symmetric $n \times n$ matrices. Set $\gamma : \mathcal{M} \to \mathcal{M}$, $\gamma(A) = VAV^t$. The canonical Jordan decomposition holds also for real matrices and it gives again that the Jordan block J of V determines the determinant of γ.

To have a matrix of $A \mapsto JAJ^t$, we need a basis in $\mathcal{M}$. We can take

$$\{E(j,k) + E(k,j) : 1 \le j \le k \le n\}$$

Similarly to the above argument, one can see that the matrix is upper triangular. So we need the diagonal entries. $J(E(j,k) + E(k,j))J^*$ can be computed from the above formula and the coefficient of $E(j,k) + E(k,j)$ is $J_{kk}J_{jj}$. The determinant is

$$\prod_{1\le j\le k\le n} J_{kk}J_{jj} = (\det J)^{n+1} = (\det V)^{n+1}$$

□

Theorem 2.6 *The determinant of a positive matrix $A \in \mathbb{M}_n$ does not exceed the product of the diagonal entries:*

$$\det A \le \prod_{i=1}^{n} A_{ii}$$

This is a consequence of the concavity of the log function, see Example 8.7 (or Example 4.5).

If $A \in \mathbb{M}_n$ and $1 \le i, j \le n$, then in the next theorems $[A]^{ij}$ denotes the $(n-1)\times(n-1)$ matrix which is obtained from A by striking out the ith row and the jth column.

Theorem 2.7 *Let $A \in \mathbb{M}_n$ and $1 \leq j \leq n$. Then*

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det([A]^{ij}).$$

The determinant has an important role in the computation of the **inverse**. The next result is called Cramer's rule.

Theorem 2.8 *Let $A \in \mathbb{M}_n$ be invertible. Then*

$$(A^{-1})_{ki} = (-1)^{i+k} \frac{\det([A]^{ik})}{\det A}$$

for $1 \leq i, k \leq n$.

2.4 Notes and remarks

Computation of the determinant of concreate special matrices had a huge literature. Theorem 2.6 is the Hadamard inequality from 1893.

Example 2.6 is related to the Hunyadi-Scholtz determinant theorem from 1870's.

2.5 Exercises

1. Show that

$$\begin{bmatrix} \lambda + z & x - \mathrm{i}y \\ x + \mathrm{i}y & \lambda - z \end{bmatrix}^{-1} = \frac{1}{\lambda^2 - x^2 - y^2 - z^2} \begin{bmatrix} \lambda - z & -x + \mathrm{i}y \\ -x - \mathrm{i}y & \lambda + z \end{bmatrix}$$

   for real parameters λ, x, y, z.

2. Let $m \leq n$, $A \in \mathbb{M}_n, B \in \mathbb{M}_m, Y \in \mathbb{M}_{n \times m}$ and $Z \in \mathbb{M}_{m \times n}$. Assume that A and B are invertible. Show that $A + YBZ$ is invertible if and only if $B^{-1} + ZA^{-1}Y$ is invertible. Moreover,

$$(A + YBZ)^{-1} = A^{-1} - A^{-1}Y(B^{-1} + ZA^{-1}Y)^{-1}ZA^{-1}.$$

3. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of the matrix $A \in \mathbb{M}_n(\mathbb{C})$. Show that A is normal if and only if

$$\sum_{i=1}^{n} |\lambda_i|^2 = \sum_{i,j=1}^{n} |A_{ij}|^2.$$

4. Show that $A \in \mathbb{M}_n$ is normal if an only if $A^* = AU$ for a unitary $U \in \mathbb{M}_n$.

5. Give an example such that $A^2 = A$, but A is not an orthogonal projection.

6. $A \in \mathbb{M}_n$ is called idempotent if $A^2 = A$. Show that each eigenvalue of a idempotent matrix is either 0 or 1.

7. Compute the eigenvalues and eigenvectors of the **Pauli matrices**:

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma_2 = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}, \qquad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.11}$$

8. Show that the Pauli matrices (2.11) are orthogonal to each other (with respect to the Hilbert–Schmidt inner product). What are the matrices which orthogonal to all Pauli matrices?

9. Let λ be an eigenvalue of a unitary operator. Show that $|\lambda| = 1$.

10. Let A be an $n \times n$ matrix and let $k \geq 1$ be an integer. Assume that $A_{ij} = 0$ if $j \geq i + k$. Show that A^{n-k} is the 0 matrix.

11. Show that $|\det U| = 1$ for a unitary U.

12. Let $U \in \mathbb{M}_n$ and $u_1, \ldots, u_n$ be n column vectors of U, i.e., $U = [u_1\, u_2\, \ldots\, u_n]$. Prove that U is a unitary matrix if and only if $\{u_1, \ldots, u_n\}$ is an orthonormal basis of $\mathbb{C}^n$.

13. Let the $U = [u_1\, u_2\, \ldots\, u_n] \in \mathbb{M}_n$ matrix be described by column vectors. Assume that that $\{u_1, \ldots, u_k\}$ are given and orthonormal in $\mathbb{C}^n$. Show that $u_{k+1}, \ldots, u_n$ can be chosen in such a way that U will be a unitary matrix.

14. Compute $\det(\lambda I - A)$ when A is the tridiagonal matrix (2.4).

15. Let $U \in B(\mathcal{H})$ be a unitary. Show that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} U^n x$$

exists for every vector $x \in \mathcal{H}$. (Hint: Consider the subspaces $\{x \in \mathcal{H} : Ux = x\}$ and $\{Ux - x : x \in \mathcal{H}\}$.) What is the limit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} U^n \,?$$

(This is the **ergodic theorem**.)

Chapter 3

Tensor products

Let $\mathcal{H}$ be the linear space of polynomials in the variable x and with degree less than n. A natural basis consists of the powers $1, x, x^2, \dots, x^n$. Similarly, let $\mathcal{K}$ be the space of polynomials in y of degree less than m. Its basis is $1, y, y^2, \dots, y^m$. The tensor product of these two spaces is the space of polynomials of two variables with basis $x^i y^j$, $i \leq n, j \leq m$. This simple example contains the essential ideas.

3.1 Algebraic tensor product

Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. Their **algebraic tensor product** consists of the formal finite sums

$$\sum_{i,j} x_i \otimes y_j \qquad (x_i \in \mathcal{H}, y_i \in \mathcal{K}).$$

Computing with these sums, one should use the following rules:

$$\begin{aligned} (x_1 + x_2) \otimes y &= x_1 \otimes y + x_2 \otimes y, \quad (\lambda x) \otimes y = \lambda(x \otimes y)\,, \\ x \otimes (y_1 + y_2) &= x \otimes y_1 + x \otimes y_2, \quad x \otimes (\lambda y) = \lambda(x \otimes y)\,. \end{aligned} \tag{3.1}$$

The inner product is defined as

$$\Big\langle \sum_{i,j} x_i \otimes y_j, \sum_{k,l} z_k \otimes w_l \Big\rangle = \sum_{i,j,k,l} \langle x_i, z_k \rangle \langle y_j, w_l \rangle.$$

When $\mathcal{H}$ and $\mathcal{K}$ are finite dimensional spaces, then we arrived at the **tensor product** Hilbert space $\mathcal{H} \otimes \mathcal{K}$, otherwise the algebraic tensor product must be completed in order to get a Banach space.

Example 3.1 $L^2[0,1]$ is the Hilbert space of the square integrable functions on $[0,1]$. If $f, g \in L^2[0,1]$, then the elementary tensor $f \otimes g$ can be interpreted as a function of two variables, $f(x)g(y)$ defined on $[0,1] \times [0,1]$. The computational rules (3.1) are obvious in this approach. □

The tensor product of finitely many Hilbert spaces is defined similarly.

If $e_1, e_2, \dots$ and $f_1, f_2, \dots$ are bases in $\mathcal{H}$ and $\mathcal{K}$, respectively, then $\{e_i \otimes f_j : i, j\}$ is a basis in the tensor product space. This basis is called **product basis**. An arbitrary

vector $x \in \mathcal{H} \otimes \mathcal{K}$ admits an expansion

$$x = \sum_{i,j} c_{ij}\, e_i \otimes f_j \tag{3.2}$$

for some coefficients c_{ij}, $\sum_{i,j} |c_{ij}|^2 = \|x\|^2$. This kind of expansion is general, but sometimes it is not the best.

Lemma 3.1 *Any unit vector $x \in \mathcal{H} \otimes \mathcal{K}$ can be written in the form*

$$x = \sum_k \sqrt{p_k}\, g_k \otimes h_k, \tag{3.3}$$

where the vectors $g_k \in \mathcal{H}$ and $h_k \in \mathcal{K}$ are orthonormal and (p_k) is a probability distribution.

Proof: We can define a conjugate-linear mapping $\Lambda : \mathcal{H} \to \mathcal{K}$ as

$$\langle \Lambda\alpha, \beta\rangle = \langle \Psi, \alpha \otimes \beta\rangle$$

for every vector $\alpha \in \mathcal{H}$ and $\beta \in \mathcal{K}$. In the computation we can use the bases $(e_i)_i$ in $\mathcal{H}$ and $(f_j)_j$ in $\mathcal{K}$. If x has the expansion (3.2), then

$$\langle \Lambda e_i, f_j\rangle = \overline{c_{ij}}$$

and the adjoint Λ^* is determined by

$$\langle \Lambda^* f_j, e_i\rangle = \overline{c_{ij}}.$$

(Concerning the adjoint of a conjugate-linear mapping, see (1.16).)

One can compute that the partial trace of the matrix $|x\rangle\langle x|$ is $D := \Lambda^*\Lambda$. It is enough to check that

$$\langle x|e_k\rangle\langle e_\ell|x\rangle = \mathrm{Tr}\, \Lambda^*\Lambda |e_k\rangle\langle e_\ell|$$

for every k and ℓ.

Choose now the orthogonal unit vectors g_k such that they are eigenvectors of D with corresponding non-zero eigenvalues p_k, $Dg_k = p_k g_k$. Then

$$h_k := \frac{1}{\sqrt{p_k}} |\Lambda g_k\rangle$$

is a family of pairwise orthogonal unit vectors. Now

$$\langle x, g_k \otimes h_\ell\rangle = \langle \Lambda g_k, h_\ell\rangle = \frac{1}{\sqrt{p_\ell}} \langle \Lambda g_k, \Lambda g_\ell\rangle = \frac{1}{\sqrt{p_\ell}} \langle g_\ell, \Lambda^*\Lambda g_k\rangle = \delta_{k,\ell}\sqrt{p_\ell}$$

and we arrived at the orthogonal expansion (3.3). □

The product basis tells us that

$$\dim(\mathcal{H} \otimes \mathcal{K}) = \dim(\mathcal{H}) \times \dim(\mathcal{H}).$$

Example 3.2 In the quantum formalism the orthonormal basis in the two dimensional Hilbert space $\mathcal{H}$ is denoted as $|\uparrow\rangle, |\downarrow\rangle$. Instead of $|\uparrow\rangle \otimes |\downarrow\rangle$, the notation $|\uparrow\downarrow\rangle$ is used. Therefore the product basis is

$$|\uparrow\uparrow\rangle, \quad |\uparrow\downarrow\rangle, \quad |\downarrow\uparrow\rangle, \quad |\downarrow\downarrow\rangle.$$

Sometimes $\downarrow$ is replaced 0 and $\uparrow$ by 1.

Another basis

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \quad \frac{\mathrm{i}}{\sqrt{2}}(|10\rangle - |01\rangle), \quad \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \tag{3.4}$$

is often used, it is called **Bell basis**. □

Example 3.3 In the Hilbert space $L^2(\mathbb{R}^2)$ we can get a basis if the space is considered as $L^2(\mathbb{R}) \otimes L^2(\mathbb{R})$. In the space $L^2(\mathbb{R})$ the Hermite functions

$$\varphi_n(x) = \exp(-x^2/2)H_n(x)$$

form a good basis, where $H_n(x)$ is the appropriately normalized Hermite polynomial. Therefore, the two variable Hermite functions

$$\varphi_{nm}(x,y) := e^{-(x^2+y^2)/2}H_n(x)H_m(y) \qquad (n,m = 0,1,\ldots) \tag{3.5}$$

form a basis in $L^2(\mathbb{R}^2)$. □

3.2 Tensor product of linear mappings

The tensor product of linear transformations can be defined as well. If $A : V_1 \to W_1$ és $B : V_2 \to W_2$ are linear transformations, then there is a unique linear transformation $A \otimes B : V_1 \otimes V_2 \to W_1 \otimes W_2$ such that

$$(A \otimes B)(v_1 \otimes v_2) = Av_1 \otimes Bv_2 \qquad (v_1 \in V_1,\ v_2 \in V_2).$$

Since the linear mappings (between finite dimensional Hilbert spaces) are identified with matrices, the tensor product of matrices appears as well.

Example 3.4 Let $\{e_1, e_2, e_3\}$ be a basis in $\mathcal{H}$ and $\{f_1, f_2\}$ be a basis in $\mathcal{K}$. If $[A_{ij}]$ is the matrix of $A \in B(\mathcal{H}_1)$ and $[B_{kl}]$ is the matrix of $B \in B(\mathcal{H}_2)$, then

$$(A \otimes B)(e_j \otimes f_l) = \sum_{i,k} A_{ij}B_{kl}e_i \otimes f_k\,.$$

It is useful to order the tensor product bases lexicographically: $e_1 \otimes f_1, e_1 \otimes f_2, e_2 \otimes f_1, e_2 \otimes f_2, e_3 \otimes f_1, e_3 \otimes f_2$. Fixing this ordering, we can write down the matrix of $A \otimes B$ and we have

$$\begin{bmatrix} A_{11}B_{11} & A_{11}B_{12} & A_{12}B_{11} & A_{12}B_{12} & A_{13}B_{11} & A_{13}B_{12} \\ A_{11}B_{21} & A_{11}B_{22} & A_{12}B_{21} & A_{12}B_{22} & A_{13}B_{21} & A_{13}B_{22} \\ A_{21}B_{11} & A_{21}B_{12} & A_{22}B_{11} & A_{22}B_{12} & A_{23}B_{11} & A_{23}B_{12} \\ A_{21}B_{21} & A_{21}B_{22} & A_{22}B_{21} & A_{22}B_{22} & A_{23}B_{21} & A_{23}B_{22} \\ A_{31}B_{11} & A_{31}B_{12} & A_{32}B_{11} & A_{32}B_{12} & A_{33}B_{11} & A_{33}B_{12} \\ A_{31}B_{21} & A_{31}B_{22} & A_{32}B_{21} & A_{32}B_{22} & A_{33}B_{21} & A_{33}B_{22} \end{bmatrix}.$$

In the block matrix formalism we have

$$A \otimes B = \begin{bmatrix} A_{11}B & A_{12}B & A_{13}B \\ A_{21}B & A_{22}B & A_{23}B \\ A_{31}B & A_{32}B & A_{33}B \end{bmatrix}, \tag{3.6}$$

see Chapter 6.

The tensor product of matrices is also called **Kronecker product**. □

Example 3.5 When $A \in \mathbb{M}_n$ and $B \in \mathbb{M}_m$, the matrix

$$I_m \otimes A + B \otimes I_n \in \mathbb{M}_{nm}$$

is called the **Kronecker sum** of A and B.

If u is an eigenvector of A with eigenvalue λ and v is an eigenvector of B with eigenvalue μ, then

$$(I_m \otimes A + B \otimes I_n)(u \otimes v) = \lambda(u \otimes v) + \mu(u \otimes v) = (\lambda + \mu)(u \otimes v).$$

So $u \otimes v$ is an eigenvector of the Kronecker sum with eigenvalue $\lambda + \mu$. □

The computation rules of the tensor product of Hilbert spaces imply straightforward properties of the tensor product of matrices (or linear mappings).

Theorem 3.1 *The following rules hold:*

(1) $(A_1 + A_2) \otimes B = A_1 \otimes B + A_2 \otimes B$,

(2) $B \otimes (A_1 + A_2) = B \otimes A_1 + B \otimes A_2$,

(3) $(\lambda A) \otimes B = A \otimes (\lambda B) = \lambda(A \otimes B) \qquad (\lambda \in \mathbb{C})$,

(4) $(A \otimes B)(C \otimes D) = AC \otimes BD$,

(5) $(A \otimes B)^* = A^* \otimes B^*$,

(6) $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$,

(6) $\|A \otimes B\| = \|A\| \, \|B\|$.

For example, the tensor product of self-adjoint matrices is self-adjoint, the tensor product of unitaries is unitary.

The linear mapping $\mathbb{M}_n \otimes \mathbb{M}_n \to \mathbb{M}_n$ defined as

$$\mathrm{Tr}_2 : A \otimes B \mapsto (\mathrm{Tr}\, B)A$$

is called **partial trace**. The other partial trace is

$$\mathrm{Tr}_1 : A \otimes B \mapsto (\mathrm{Tr}\, A)B.$$

Example 3.6 Assume that $A \in \mathbb{M}_n$ and $B \in \mathbb{M}_m$. Then $A \otimes B$ is an $nm \times nm$-matrix. Let $C \in \mathbb{M}_{nm}$. How can we decide if it has the form of $A \otimes B$ for some $A \in \mathbb{M}_n$ and $B \in \mathbb{M}_m$?

First we study how to recognize A and B from $A \otimes B$. (Of course, A and B are not uniquely determined, since $(\lambda A) \otimes (\lambda^{-1} B) = A \otimes B$.) If we take the trace of all entries of (3.6), then we get

$$\begin{bmatrix} A_{11}\operatorname{Tr} B & A_{12}\operatorname{Tr} B & A_{13}\operatorname{Tr} B \\ A_{21}\operatorname{Tr} B & A_{22}\operatorname{Tr} B & A_{23}\operatorname{Tr} B \\ A_{31}\operatorname{Tr} B & A_{32}\operatorname{Tr} B & A_{33}\operatorname{Tr} B \end{bmatrix} = \operatorname{Tr} B \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = (\operatorname{Tr} B)A.$$

The sum of the diagonal entries is

$$A_{11}B + A_{12}B + A_{13}B = (\operatorname{Tr} A)B.$$

If $X = A \otimes B$, then

$$(\operatorname{Tr} X)X = (\operatorname{Tr}_2 X) \otimes (\operatorname{Tr}_1 X).$$

For example, the matrix

$$X := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

in $\mathbb{M}_2 \otimes \mathbb{M}_2$ is not a tensor product. Indeed,

$$\operatorname{Tr}_1 X = \operatorname{Tr}_2 X = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and their tensor product is the identity in $\mathbb{M}_4$. □

Let $\mathcal{H}$ be a Hilbert space. The k-fold tensor product $\mathcal{H} \otimes \ldots \otimes \mathcal{H}$ is called the kth tensor power of $\mathcal{H}$, in notation $\mathcal{H}^{\otimes k}$. When $A \in B(\mathcal{H})$, then $A^{(1)} \otimes A^{(2)} \ldots \otimes A^{(k)}$ is a linear operator on $\mathcal{H}^{\otimes k}$ and it is denoted by $A^{\otimes k}$.

3.3 Symmetric and antisymmetric tensor powers

$\mathcal{H}^{\otimes k}$ has two important subspaces, the symmetric and the antisymmetric ones. If $v_1, v_2, \ldots, v_k \in \mathcal{H}$ are vectors, then their **antisymmetric** tensor-product is the linear combination

$$v_1 \wedge v_2 \wedge \ldots \wedge v_k := \frac{1}{\sqrt{k!}} \sum_{\pi} (-1)^{\sigma(\pi)} v_{\pi(1)} \otimes v_{\pi(2)} \otimes \ldots \otimes v_{\pi(k)} \tag{3.7}$$

where the summation is over all permutations π of the set $\{1, 2, \ldots, k\}$ and $\sigma(\pi)$ is the number of inversions in π. The terminology "antisymmetric" comes from the property that an antisymmetric tensor changes its sign if two elements are exchanged. In particular, $v_1 \wedge v_2 \wedge \ldots \wedge v_k = 0$ if $v_i = v_j$ for different i and j.

The computational rules for the antisymmetric tensors are similar to (3.1):

$$\lambda(v_1 \wedge v_2 \wedge \ldots \wedge v_k) = v_1 \wedge v_2 \wedge \ldots \wedge v_{\ell-1} \wedge (\lambda v_\ell) \wedge v_{\ell+1} \wedge \ldots \wedge v_k$$

for every ℓ and

$$\begin{aligned}(v_1 \wedge v_2 \wedge \ldots \wedge v_{\ell-1} \wedge v \wedge v_{\ell+1} \wedge \ldots \wedge v_k) + \\ + (v_1 \wedge v_2 \wedge \ldots \wedge v_{\ell-1} \wedge v' \wedge v_{\ell+1} \wedge \ldots \wedge v_k) = \\ = \ v_1 \wedge v_2 \wedge \ldots \wedge v_{\ell-1} \wedge (v + v') \wedge v_{\ell+1} \wedge \ldots \wedge v_k \, .\end{aligned}$$

Lemma 3.2 *The inner product of $v_1 \wedge v_2 \wedge \ldots \wedge v_k$ and $w_1 \wedge w_2 \wedge \ldots \wedge w_k$ is the determinant of the $k \times k$ matrix whose (i,j) entry is $\langle v_i, w_j \rangle$.*

Proof: The inner product is

$$\begin{aligned}&\frac{1}{k!} \sum_\pi \sum_\kappa (-1)^{\sigma(\pi)} (-1)^{\sigma(\kappa)} \langle v_{\pi(1)}, w_{\kappa(1)} \rangle \langle v_{\pi(2)}, w_{\kappa(2)} \rangle \ldots \langle v_{\pi(k)}, w_{\kappa(k)} \rangle \\ &= \frac{1}{k!} \sum_\pi \sum_\kappa (-1)^{\sigma(\pi)} (-1)^{\sigma(\kappa)} \langle v_1, w_{\kappa\pi^{-1}(1)} \rangle \langle v_2, w_{\kappa\pi^{-1}(2)} \rangle \ldots \langle v_k, w_{\kappa\pi^{-1}(k)} \rangle \\ &= \frac{1}{k!} \sum_\pi \sum_\kappa (-1)^{\sigma(\kappa\pi^{-1})} \langle v_1, w_{\kappa\pi^{-1}(1)} \rangle \langle v_2, w_{\kappa\pi^{-1}(2)} \rangle \ldots \langle v_k, w_{\kappa\pi^{-1}(k)} \rangle \\ &= \sum_\pi (-1)^{\sigma(\pi)} \langle v_1, w_{\pi(1)} \rangle \langle v_2, w_{\pi(2)} \rangle \ldots \langle v_k, w_{\pi(k)} \rangle\end{aligned}$$

This is the determinant. □

The subspace spanned by the vectors $v_1 \wedge v_2 \wedge \ldots \wedge v_k$ is called the kth antisymmetric tensor power of $\mathcal{H}$, in notation $\wedge^k \mathcal{H}$. So $\wedge^k \mathcal{H} \subset \otimes^k \mathcal{H}$ and this subspace can be defined also in a different way.

Example 3.7 A transposition is a permutation of $1, 2, \ldots, n$ which exchanges the place of two entries. For a transposition κ, there is a unitary $U_\kappa : \otimes^k \mathcal{H} \to \otimes^k \mathcal{H}$ such that

$$U_\kappa(v_1 \otimes v_2 \otimes \ldots \otimes v_n) = v_{\kappa(1)} \otimes v_{\kappa(2)} \otimes \ldots \otimes v_{\kappa(n)}.$$

Then

$$\wedge^k \mathcal{H} = \{x \in \otimes^k \mathcal{H} : U_\kappa x = -x \text{ for every } \kappa\}. \tag{3.8}$$

The terminology "antisymmetric " comes from this description. □

If $A \in B(\mathcal{H})$, then the transformation $\otimes^k A$ leaves the subspace $\wedge^k \mathcal{H}$ invariant. Its restriction is denoted by $\wedge^k A$ which is equivalently defined as

$$\wedge^k A(v_1 \wedge v_2 \wedge \ldots \wedge v_k) = Av_1 \wedge Av_2 \wedge \ldots \wedge Av_k. \tag{3.9}$$

If $e_1, e_2, \ldots, e_n$ is a basis in $\mathcal{H}$, then

$$\{e_{i(1)} \wedge e_{i(2)} \wedge \ldots \wedge e_{i(k)} \ : \ 1 \le i(1) < i(2) < \ldots < i(k)) \le n\} \tag{3.10}$$

is a basis in $\wedge^k \mathcal{H}$. It follows that the dimension of $\wedge^k \mathcal{H}$ is

$$\binom{n}{k} \quad \text{if} \quad k \le n,$$

otherwise for $k > n$ the power $\wedge^k\mathcal{H}$ has dimension 0. Consequently, $\wedge^n\mathcal{H}$ has dimension 1 and for any operator $A \in B(\mathcal{H})$, we have

$$\wedge^n A = \lambda \times \text{identity} \tag{3.11}$$

Theorem 3.2 *For $A \in \mathbb{M}_n$, the constant λ in (3.11) is* $\det A$.

Proof: If $e_1, e_2, \ldots, e_n$ is a basis in $\mathcal{H}$, then in the space $\wedge^n V$ the vector $e_1 \wedge e_2 \wedge \ldots \wedge e_n$ forms a basis. We should compute $(\wedge^n A)(e_1 \wedge e_2 \wedge \ldots \wedge e_n)$.

$$\begin{aligned}
&(\wedge^n A)(e_1 \wedge e_2 \wedge \ldots \wedge e_n) = (Ae_1) \wedge (Ae_2) \wedge \ldots \wedge (Ae_n) = \\
&= \Big(\sum_{i(1)=1}^{n} A_{i(1),1} e_{i(1)}\Big) \wedge \Big(\sum_{i(2)=1}^{n} A_{i(2),2} e_{i(2)}\Big) \wedge \ldots \wedge \Big(\sum_{i(n)=1}^{n} A_{i(n),n} e_{i(n),n}\Big) = \\
&= \sum_{i(1),i(2),\ldots,i(n)=1}^{n} A_{i(1),1} A_{i(2),2} \ldots A_{i(n),n} e_{i(1)} \wedge \ldots \wedge e_{i(n)} = \\
&= \sum_{\pi} A_{\pi(1),1} A_{\pi(2),2} \ldots A_{\pi(n),n} e_{\pi(1)} \wedge \ldots \wedge e_{\pi(n)} = \\
&= \sum_{\pi} A_{\pi(1),1} A_{\pi(2),2} \ldots A_{\pi(n),n} (-1)^{\sigma(\pi)} e_1 \wedge \ldots \wedge e_n \,.
\end{aligned}$$

Here we used that $e_{i(1)} \wedge \ldots \wedge e_{i(n)}$ can be non-zero if the vectors $e_{i(1)}, \ldots, e_{i(n)}$ are all different, in other words, this is a permutation of $e_1, e_2, \ldots, e_n$. □

The symmetric tensor product of the vectors $v_1, v_2, \ldots, v_k \in \mathcal{H}$ is

$$v_1 \vee v_2 \vee \ldots \vee v_k := \frac{1}{\sqrt{k!}} \sum_{\pi} v_{\pi(1)} \otimes v_{\pi(2)} \otimes \ldots \otimes v_{\pi(k)} \,,$$

where the summation is over all permutations π of the set $\{1, 2, \ldots, k\}$ again. The linear span of the symmetric tensors is the symmetric tensor power $\vee^k\mathcal{H}$. Similarly to (3.8), we have

$$\vee^k\mathcal{H} = \{x \in \otimes^k\mathcal{H} : U_\kappa x = x \text{ for every } \kappa\}. \tag{3.12}$$

It follows immediately, that $\vee^k\mathcal{H} \perp \wedge^k\mathcal{H}$. Let $u \in \vee^k\mathcal{H}$ and $v \in \wedge^k\mathcal{H}$. Then

$$\langle u, v\rangle = \langle U_\kappa u, -U_\kappa v\rangle = -\langle u, v\rangle$$

and $\langle u, v\rangle = 0$.

If $e_1, e_2, \ldots, e_n$ is a basis in $\mathcal{H}$, then $\vee^k\mathcal{H}$ has the basis

$$\{e_{i(1)} \vee e_{i(2)} \vee \ldots \vee e_{i(k)} \ : \ 1 \le i(1) \le i(2) \le \ldots \le i(k) \le n\}. \tag{3.13}$$

Similarly to the proof of Lemma 3.2 we have

$$\langle v_1 \vee v_2 \vee \ldots \vee v_k, w_1 \vee w_2 \vee \ldots \vee w_k\rangle = \sum_{\pi} \langle v_1, w_{\pi(1)}\rangle \langle v_2, w_{\pi(2)}\rangle \ldots \langle v_k, w_{\pi(k)}\rangle \tag{3.14}$$

The right-hand-side is similar to a determinant, but the sign is not changing. This is called the **permanent** of the $k \times k$ matrix whose (i, j) entry is $\langle v_i, w_j\rangle$. So

$$\operatorname{per} A = \sum_{\pi} A_{1,\pi(1)} A_{2,\pi(2)} \ldots A_{n,\pi(n)}.$$

3.4 Notes and remarks

Note that the Kronecker sum is often denoted by $A \oplus B$ in the literature, but in this book $\oplus$ is the notation for the direct sum.

3.5 Exercises

1. Let
$$|\beta_0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \in \mathbb{C}^2 \otimes \mathbb{C}^2$$
and
$$|\beta_i\rangle = (\sigma_i \otimes I_2)|\beta_0\rangle \qquad (i = 1, 2, 3)$$
by means of the Pauli matrices σ_i. Show that $\{|\beta_i\rangle : 0 \leq i \leq 3\}$ is the Bell basis.

2. Show that the vectors of the Bell basis are eigenvectors of the matrices $\sigma_i \otimes \sigma_i$, $1 \leq i \leq 3$.

3. Show the identity
$$|\psi\rangle \otimes |\beta_0\rangle = \frac{1}{2} \sum_{k=0}^{3} |\beta_k\rangle \otimes \sigma_k |\psi\rangle \tag{3.15}$$
in $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$, where $|\psi\rangle \in \mathbb{C}^2$ and $|\beta_i\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ is defined above.

4. Write the so-called **Dirac matrices** in the form of elementary tensor (of two 2×2 matrices):
$$\gamma_1 = \begin{bmatrix} 0 & 0 & 0 & -\mathrm{i} \\ 0 & 0 & -\mathrm{i} & 0 \\ 0 & -\mathrm{i} & 0 & 0 \\ -\mathrm{i} & 0 & 0 & 0 \end{bmatrix}, \qquad \gamma_2 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix},$$
$$\gamma_3 = \begin{bmatrix} 0 & 0 & -\mathrm{i} & 0 \\ 0 & 0 & 0 & \mathrm{i} \\ \mathrm{i} & 0 & 0 & 0 \\ 0 & -\mathrm{i} & 0 & 0 \end{bmatrix}, \qquad \gamma_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

5. Give the dimension of $\vee^k \mathcal{H}$ if $\dim(\mathcal{H}) = n$.

6. Let $A \in B(\mathcal{H})$ and $B \in B(\mathcal{H})$ be operators on the finite dimensional spaces $\mathcal{H}$ and $\mathcal{K}$. Show that
$$\det(A \otimes B) = (\det A)^m (\det B)^n,$$
where $n = \dim \mathcal{H}$ and $m = \dim \mathcal{K}$. (Hint: The determinant is the product of the eigenvalues.)

7. Show that $\|A \otimes B\| = \|A\| \cdot \|B\|$.

8. Use Theorem 3.2 to prove that $\det(AB) = \det A \times \det B$. (Hint: Show that $\wedge^k(AB) = (\wedge^k A)(\wedge^k B)$.)

9. Let $x^n + c_1 x^{n-1} + \ldots + c_n$ be the characteristic polynomial of $A \in \mathbb{M}_n$. Show that $c_k = \mathrm{Tr}\ \wedge^k A$.

10. Show that
$$\mathcal{H} \otimes \mathcal{H} = (\mathcal{H} \vee \mathcal{H}) \oplus (\mathcal{H} \wedge \mathcal{H})$$
for a Hilbert space $\mathcal{H}$.

11. Show that
$$|\mathrm{per}\,(AB)|^2 \leq \mathrm{per}\,(AA^*)\mathrm{per}\,(B^*B).$$

12. Let $A \in \mathbb{M}_n$ and $B \in \mathbb{M}_m$. Show that
$$\mathrm{Tr}\,(I_m \otimes A + B \otimes I_n) = m\mathrm{Tr}\,A + n\mathrm{Tr}\,B.$$

13. For a vector $f \in \mathcal{H}$ the linear operator $a^+(f) : \vee^k\mathcal{H} \to \vee^{k+1}\mathcal{H}$ is defined as
$$a^+(f)\, v_1 \vee v_2 \vee \ldots \vee v_k = f \vee v_1 \vee v_2 \vee \ldots \vee v_k. \tag{3.16}$$
Compute the adjoint of $a^+(f)$ which is denoted by $a(f)$.

14. For $A \in B(\mathcal{H})$ let $\mathcal{F}(A) : \vee^k\mathcal{H} \to \vee^k\mathcal{H}$ be defined as
$$\mathcal{F}(A)\, v_1 \vee v_2 \vee \ldots \vee v_k = \sum_{i=1}^{k} v_1 \vee v_2 \vee \ldots \vee v_{i-1} \vee Av_i \vee v_{i+1} \vee \ldots \vee v_k.$$
Show that
$$\mathcal{F}(|f\rangle\langle g|) = a^+(f)a(g)$$
for $f, g \in \mathcal{H}$. (Recall that a and a^+ are defined in the previous exercise.)

Chapter 4

Positive matrices

Mostly the statements and definitions are formulated in the Hilbert space setting. The Hilbert space is always assumed to be finite dimensional, so instead of operators one can consider a matrices.

4.1 Positivity and square root

Let $\mathcal{H}$ be a (complex) Hilbert space and $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator. T is called **positive** (or positive semidefinite) if $\langle x, Tx\rangle \geq 0$ for every vector $x \in \mathcal{H}$, in notation $T \geq 0$. It follows from the definition that a positive operator is self-adjoint: Since $\langle x, Tx\rangle \in \mathbb{R}$,

$$\langle x, Tx\rangle = \overline{\langle x, Tx\rangle} = \langle Tx, x\rangle.$$

(The argument used the complex structure. In the real case any scalar product is real. So in the real-valued case the positivity of T needs the requirement of the self-adjoint property as well.)

If T_1 and T_2 are positive operators, then $T_1 + T_2$ is positive as well.

Theorem 4.1 *Let $T \in B(\mathcal{H})$ be an operator. The following conditions are equivalent.*

(1) T is positive.

(2) $T = T^*$ and the spectrum of T lies in $\mathbb{R}^+$.

(3) T is of the form A^*A for some operator $A \in B(\mathcal{H})$.

An operator T is positive if and only if UTU^* is positive for a unitary U. The positivity of an operator is equivalent to the positivity of its matrix (with respect to any orthonormal basis).

Theorem 4.2 *Let T be a positive operator. Then there is a unique positive operator B such that $B^2 = T$. If the self-adjoint operator A commutes with T, then it commutes with B as well.*

Proof: We restrict ourselves to the finite dimensional case. In this case it is enough to find the eigenvalues and the eigenvectors. If $Bx = \lambda x$, then x is an eigenvector of T with eigenvalue λ^2. This determines B uniquely, T and B have the same eigenvectors.

$AB = BA$ holds if for any eigenvector x of B the vector Ax is an eigenvector, too. If $TA = AT$, then this follows. □

B is called the **square root** of T, $T^{1/2}$ is a notation. It follows from the theorem that the product of commuting positive operators T and A is positive. Indeed,

$$TA = T^{1/2}T^{1/2}A^{1/2}A^{1/2} = T^{1/2}A^{1/2}A^{1/2}T^{1/2} = (A^{1/2}T^{1/2})^*A^{1/2}T^{1/2}.$$

Example 4.1 For each $A \in B(\mathcal{H})$, we have $A^*A \geq 0$. So, define $|A| := (A^*A)^{1/2}$ that is called the **absolute value** of A. The mapping

$$|A|x \mapsto Ax$$

is norm preserving:

$$\| |A|x\|^2 = \langle |A|x, |A|x\rangle = \langle x, |A|^2x\rangle = \langle x, A^*Ax\rangle = \langle Ax, Ax\rangle = \|Ax\|^2$$

It can be extended to a unitary U. So $A = U|A|$ and this is called **polar decomposition**.

The definition $|A| := (A^*A)^{1/2}$ makes sense also if $A : \mathcal{H}_1 \to \mathcal{H}_2$. Then $|A| \in B(\mathcal{H}_1)$. The above argument tells that $|A|x \mapsto Ax$ is norm preserving, but it is not sure that it can be extended to a unitary. If $\dim \mathcal{H}_1 \leq \dim \mathcal{H}_2$, then $|A|x \mapsto Ax$ can be extended to an isometry $V : \mathcal{H}_1 \to \mathcal{H}_2$. Then $A = V|A|$, where $V^*V = I$.

The eigenvalues of $|A|$ are called the **singular values** of A. □

Example 4.2 Let T be a positive operator acting on a finite dimensional Hilbert space such that $\|T\| \leq 1$. We want to show that there is unitary operator U such that

$$T = \frac{1}{2}(U + U^*).$$

We can choose an orthonormal basis $e_1, e_2, \ldots, e_n$ consisting of eigenvectors of T and in this basis the matrix of T is diagonal, say, $\mathrm{Diag}(t_1, t_2, \ldots, t_n)$, $0 \leq t_j \leq 1$ from the positivity. For any $1 \leq j \leq n$ we can find a real number θ_j such that

$$t_j = \frac{1}{2}(e^{\mathrm{i}\theta_j} + e^{-\mathrm{i}\theta_j}).$$

Then the unitary operator U with matrix $\mathrm{Diag}(\exp(\mathrm{i}\theta_1)), \ldots, \exp(\mathrm{i}\theta_n))$ will have the desired property. □

If T acts on a finite dimensional Hilbert space which has an orthonormal basis $e_1, e_2, \ldots, e_n$, then T is uniquely determined by its matrix

$$[\langle e_i, Te_j\rangle]_{i,j=1}^n.$$

T is positive if and only if its matrix is positive (semi-definite).

Example 4.3 Let

$$A = \begin{bmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_n \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

Then

$$[A^*A]_{i,j} = \overline{\lambda}_i \lambda_j \qquad (1 \le i, j \le n)$$

and this matrix is positive. Every positive matrix is the sum of matrices of this form. (The minimum number of the summands is the rank of the matrix.) □

Example 4.4 Take numbers $\lambda_1, \lambda_2, \dots, \lambda_n > 0$ and set

$$A_{ij} = \frac{1}{\lambda_i + \lambda_j} \qquad (1 \le i, j \le n). \tag{4.1}$$

A is called a **Cauchy matrix**. We have

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

and the matrix

$$A(t)_{ij} := e^{-t\lambda_i} e^{-t\lambda_j}$$

is positive for every $t \in \mathbb{R}$ due to Example 4.3. Therefore

$$A = \int_0^\infty A(t)\, dt$$

is positive as well. □

Theorem 4.3 *Let $T \in B(\mathcal{H})$ be an invertible self-adjoint operator and $e_1, e_2, \dots, e_n$ be a basis in the Hilbert space $\mathcal{H}$. T is positive if and only if for any $1 \le k \le n$ the determinant of the $k \times k$ matrix*

$$[\langle e_i, Te_j \rangle]_{ij=1}^k$$

is positive (that is, ≥ 0).

An invertible positive matrix is called **positive definite**. Such matrices appear in probability theory in the concept of **Gaussian distribution**. The work with Gaussian distributions in probability theory requires the experience with matrices. (This is in the next example, but also in Example 6.1.)

Example 4.5 Let M be a positive definite $n \times n$ real matrix and $\mathbf{x} = (x_1, x_2, \dots, x_n)$. Then

$$f_M(\mathbf{x}) := \sqrt{\frac{\operatorname{Det} M}{(2\pi)^n}} \exp\Big(-\tfrac{1}{2}\langle \mathbf{x}, M\mathbf{x} \rangle \Big) \tag{4.2}$$

is a multivariate Gaussian probability distribution (with 0 expectation, see, for example, III.6 in [18]). The matrix M will be called the **quadratic matrix** of the Gaussian distribution.

For an $n \times n$ matrix B, the relation

$$\int \langle \mathbf{x}, B\mathbf{x} \rangle f_M(\mathbf{x})\, d\mathbf{x} = \operatorname{Tr} BM^{-1} \tag{4.3}$$

holds.

We first note that if (4.3) is true for a matrix M, then

$$\begin{aligned} \int \langle B\mathbf{x}, \mathbf{x} \rangle f_{U^*MU}(\mathbf{x})\, d\mathbf{x} &= \int \langle BU^*\mathbf{x}, U^*\mathbf{x} \rangle f_M(\mathbf{x})\, d\mathbf{x} \\ &= \operatorname{Tr} (UBU^*)M^{-1} \\ &= \operatorname{Tr} B(U^*MU)^{-1} \end{aligned}$$

for a unitary U, since the Lebesgue measure on $\mathbb{R}^n$ is invariant under unitary transformation. This means that (4.3) holds also for U^*MU. Therefore to check (4.3), we may assume that M is diagonal. Another reduction concerns B: We may assume that B is a matrix unit E_{ij}. Then the n variable integral reduces to integrals on $\mathbb{R}$ and the known integrals

$$\int_{\mathbb{R}} t \exp\Big(-\frac{1}{2}\lambda t^2\Big)\, dt = 0 \quad \text{and} \quad \int_{\mathbb{R}} t^2 \exp\Big(-\frac{1}{2}\lambda t^2\Big)\, dt = \frac{\sqrt{2\pi}}{\lambda}$$

can be used.

Formula (4.3) has an important consequence. When the joint distribution of the random variables $(\xi_1, \xi_2, \ldots, \xi_n)$ is given by (4.2), then the **covariance matrix** is M^{-1}.

The **Boltzmann entropy** of a probability density $f(\mathbf{x})$ is defined as

$$h(f) := -\int f(\mathbf{x}) \log f(\mathbf{x})\, d\mathbf{x} \tag{4.4}$$

if the integral exists. For a Gaussian f_M we have

$$h(f_M) = \frac{n}{2} \log(2\pi e) - \frac{1}{2} \log \operatorname{Det} M.$$

Assume that f_M is the joint distribution of the (number-valued) random variables $\xi_1, \xi_2, \ldots, \xi_n$. Their joint Boltzmann entropy is

$$h(\xi_1, \xi_2, \ldots, \xi_n) = \frac{n}{2} \log(2\pi e) + \log \operatorname{Det} M^{-1}$$

and the Boltzmann entropy of ξ_i is

$$h(\xi_i) = \frac{1}{2} \log(2\pi e) + \frac{1}{2} \log(M^{-1})_{ii}.$$

The subadditivity of the Boltzmann entropy is the inequality

$$h(\xi_1, \xi_2, \ldots, \xi_n) \leq h(\xi_1) + h(\xi_2) + \ldots + h(\xi_n)$$

which is

$$\log \operatorname{Det} A \leq \sum_{i=1}^{m} \log A_{ii}$$

in our particular Gaussian case, $A = M^{-1}$. What we obtained is the **Hadamard inequality**

$$\operatorname{Det} A \leq \prod_{i=1}^{m} A_{ii}$$

for a positive definite matrix A, cf. Theorem 2.6. □

Example 4.6 If the matrix $X \in \mathbb{M}_n$ can be written in the form

$$X = S\mathrm{Diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)S^{-1},$$

with $\lambda_1, \lambda_2, \ldots, \lambda_n > 0$, then X is called **weakly positive**. Such a matrix has n linearly independent eigenvectors with strictly positive eigenvalues. If the eigenvectors are orthogonal, then the matrix is positive definite. Since

$$X = \left(S\mathrm{Diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})S^*\right)\left((S^*)^{-1}\mathrm{Diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})S^{-1}\right),$$

this is the product of two positive definite matrices.

Although this X is not positive, but the eigenvalues are strictly positive. Therefore we can define the square root as

$$X^{1/2} = S\mathrm{Diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})S^{-1}.$$

(See also 5.6). □

4.2 Relation to tensor product

If $0 \leq A \in \mathbb{M}_n$ and $0 \leq B \in \mathbb{M}_m$, then $0 \leq A \otimes B$. More generally if $0 \leq A_i \in \mathbb{M}_n$ and $0 \leq B_i \in \mathbb{M}_m$, then

$$\sum_{i=1}^{k} A_i \otimes B_i$$

is positive. These matrices in $\mathbb{M}_n \otimes \mathbb{M}_m$ are called **separable positive matrices**. Is it true that every positive matrix in $\mathbb{M}_n \otimes \mathbb{M}_m$ is separable? A counterexample follows.

Example 4.7 Let $\mathbb{M}_4 = \mathbb{M}_2 \otimes \mathbb{M}_2$ and

$$D := \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

D is a rank 1 positive operator, it is a projection. If $D = \sum_i D_i$, then $D_i = \lambda_i D$. If D is separable, then it is a tensor product. If D is a tensor product, then up to a constant factor it equals to $(\mathrm{Tr}_2 D) \otimes (\mathrm{Tr}_1 D)$. We have

$$\mathrm{Tr}_1 D = \mathrm{Tr}_2 D = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Their tensor product has rank 4 and it cannot be λD. It follows that this D is not separable. □

In quantum theory the non-separable positive operators are called **entangled**. The positive operator D is **maximally entangled** if it has minimal rank (it means rank 1) and the partial traces have maximal rank. The matrix D in the previous example is maximally entangled.

It is interesting that there is no procedure to decide if a positive operator in a tensor product space is separable or entangled.

4.3 Partial ordering

Let $A, B \in B(\mathcal{H})$ be self-adjoint operators. The partial ordering $A \leq B$ holds if $B - A$ is positive, or equivalently

$$\langle x, Ax \rangle \leq \langle x, Bx \rangle$$

for all vectors x. From this formulation one can easily see that $A \leq B$ implies $XAX^* \leq XBX^*$ for every operator X.

Example 4.8 Assume that for the orthogonal projections P and Q the inequality $P \leq Q$ holds. If $Px = x$ for a unit vector x, then $\langle x, Px \rangle \leq \langle x, Qx \rangle \leq 1$ shows that $\langle x, Qx \rangle = 1$. Therefore the relation

$$\|x - Qx\|^2 = \langle x - Qx, x - Qx \rangle = \langle x, x \rangle - \langle x, Qx \rangle = 0$$

gives that $Qx = x$. The range of Q contains the range of P. □

Let A_n be a sequence of operators on a finite dimensional Hilbert space. Fix a basis and let $[A_n]$ be the matrix of A_n. Similarly, the matrix of the operator A is $[A]$. Let the Hilbert space m-dimensional, so the matrices are $m \times m$. Recall that the following conditions are equivalent:

(1) $\|A - A_n\| \to 0$.

(2) $A_n x \to Ax$ for every vector x.

(3) $\langle x, A_n y \rangle \to \langle x, Ay \rangle$ for every vectors x and y.

(4) $\langle x, A_n x \rangle \to \langle x, Ax \rangle$ for every vector x.

(5) $\mathrm{Tr}\,(A - A_n)^*(A - A_n) \to 0$

(6) $[A_n]_{ij} \to [A]_{ij}$ for every $1 \leq i, j \leq m$.

These conditions describe in several ways the **convergence** of a sequence of operators or matrices.

Theorem 4.4 *Let A_n be an increasing sequence of operators with an upper bound: $A_1 \leq A_2 \leq \ldots \leq B$. Then there is an operator $A \leq B$ such that $A_n \to A$.*

Proof: Let $\phi_n(x,y) := \langle x, A_n y\rangle$ be a sequence of complex bilinear functionals. $\lim_n \phi_n(x,x)$ is a bounded increasing real sequence and it is convergent. Due to the polarization identity $\phi_n(x,y)$ is convergent as well and the limit gives a complex bilinear functional ϕ. If the corresponding operator is denoted by A, then

$$\langle x, A_n y\rangle \to \langle x, Ay\rangle$$

for every vectors x and y. This is the convergence $A_n \to A$. The condition $\langle x, Ax\rangle \le \langle x, Bx\rangle$ means $A \le B$. □

Example 4.9 Assume that $0 \le A \le I$ for an operator A. Define a sequence T_n of operators by recursion. Let $T_1 = 0$ and

$$T_{n+1} = T_n + \frac{1}{2}(A - T_n^2) \qquad (n \in \mathbb{N})\,.$$

T_n is a polynomial of A with real coefficients. So these operators commute with each other. Since

$$I - T_{n+1} = \frac{1}{2}(I - T_n)^2 + \frac{1}{2}(I - A)\,,$$

induction shows that $T_n \le I$.

We show that $T_1 \le T_2 \le T_3 \le \dots$ by mathematical induction. In the recursion

$$T_{n+1} - T_n = \frac{1}{2}\left((I - T_{n-1})(T_n - T_{n-1}) + (I - T_n)(T_n - T_{n-1})\right)$$

$I - T_{n-1} \ge 0$ and $T_n - T_{n-1} \ge 0$ due to the assumption. Since they commute their product is positive. Similarly $(I - T_n)(T_n - T_{n-1}) \ge 0$. It follows that the right-hand-side is positive.

Theorem 4.4 tells that T_n converges to an operator B. The limit of the recursion formula yields

$$B = B + \frac{1}{2}(A - B^2)\,,$$

therefore $A = B^2$. □

Theorem 4.5 *Assume that $0 < A, B \in \mathbb{M}_n$ are invertible matrices and $A \le B$. Then $B^{-1} \le A^{-1}$*

Proof: The condition $A \le B$ is equivalent to $B^{-1/2}AB^{-1/2} \le I$ and the statement $B^{-1} \le A^{-1}$ is equivalent to $I \le B^{1/2}A^{-1}B^{1/2}$. If $X = B^{-1/2}AB^{-1/2}$, then we have to show that $X \le I$ implies $X^{-1} \ge I$. The condition $X \le I$ means that all eigenvalues of X are in the interval $(0,1]$. This implies that all eigenvalues of X^{-1} are in $[1,\infty)$. □

Assume that $A \le B$. It follows from (2.9) that the largest eigenvalue of A is smaller than the largest eigenvalue of B. Let $\lambda(A) = (\lambda_1(A), \dots, \lambda_n(A))$ denote the vector of the eigenvalues of A in decreasing order (with counting multiplicities).

The next result is called **Weyl's monotonicity theorem**.

Theorem 4.6 *If $A \le B$, then $\lambda_k(A) \le \lambda_k(B)$ for all k.*

This is a consequence of the minimax principle, Theorem 2.5.

4.4 Hadamard product

Theorem 4.7 *(**Schur theorem**) Let A and B be positive $n \times n$ matrices. Then*

$$C_{ij} = A_{ij}B_{ij} \qquad (1 \le i, j \le n)$$

determines a positive matrix.

Proof: If $A_{ij} = \overline{\lambda}_i\lambda_j$ and $B_{ij} = \overline{\mu}_i\mu_j$, then $C_{ij} = \overline{\lambda_i\mu_i}\lambda_j\mu_j$ and C is positive due to Example 4.3. The general case is reduced to this one. □

The matrix C of the previous theorem is called the **Hadamard** (or Schur) **product** of the matrices A and B. In notation, $C = A \circ B$.

Corollary 4.1 *Assume that $0 \le A \le B$ and $0 \le C \le D$. Then $A \circ C \le B \circ D$.*

Proof: The equation

$$B \circ D = A \circ C + (B - A) \circ C + (D - C) \circ A + (B - A) \circ (D - C)$$

implies the statement. □

Example 4.10 Let $A \in \mathbb{M}_n$ be a positive matrix. The mapping $S_A : B \mapsto A \circ B$ sends a positive matrix to a positive matrix, therefore it is a positive mapping. A linear mapping $\alpha : \mathbb{M}_n \to \mathbb{M}_n$ is called **completely positive** if it has the form

$$\alpha(B) = \sum_{i=1}^{k} V_i^* A V_i$$

for some matrices V_i. The sum of completely positive mappings is completely positive.

We want to show that S_A is completely positive. S_A is additive in A, hence it is enough to show the case $A_{ij} = \overline{\lambda}_i\lambda_j$. Then

$$S_A(B) = \mathrm{Diag}(\overline{\lambda}_1, \overline{\lambda}_2, \dots, \overline{\lambda}_n)\, B\, \mathrm{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

and S_A is completely positive. □

Example 4.11 Take numbers $t, \lambda_1, \lambda_2, \dots, \lambda_n > 0$ and set

$$A_{ij} = \frac{1}{(\lambda_i + \lambda_j)^t} \qquad (1 \le i, j \le n). \tag{4.5}$$

For $t = 1$ this A is a Cauchy matrix which is positive, see Example 4.4. If t is an integer, the Schur theorem gives the positivity of the Hadamard powers. Actually, $A > 0$ for every $t > 0$. It is easy to prove for a 2×2 matrix. □

4.5 Lattice of projections

Let $\mathcal{K}$ be a closed subspace of a Hilbert space $\mathcal{H}$. Any vector $x \in \mathcal{H}$ can be written in the form $x_0 + x_1$, where $x_0 \in \mathcal{K}$ and $x_1 \perp \mathcal{K}$. The linear mapping $P : x \mapsto x_0$ is

called (orthogonal) **projection** onto $\mathcal{K}$. The orthogonal projection P has the properties $P = P^2 = P^*$. If an operator $P \in B(\mathcal{H})$ satisfies $P = P^2 = P^*$, then it is an (orthogonal) projection (onto its range). Instead of orthogonal projection the expression **ortho-projection** is also used.

If P and Q are projections, then the relation $P \leq Q$ means that the range of P is contained in the range of Q. An equivalent algebraic formulation is $PQ = P$. The largest projection is the identity I and the smallest one is 0. Therefore $0 \leq P \leq I$ for any projection P.

If P is a projection, then $I - P$ is a projection as well and it is often denoted by $P^\perp$, since the range of $I - P$ is the orthogonal complement of the range of P.

Example 4.12 Let P and Q be projections. The relation $P \perp Q$ means that the range of P is orthogonal to the range of Q. An equivalent algebraic formulation is $PQ = 0$. Since the orthogonality relation is symmetric, $PQ = 0$ if and only if $QP = 0$. (We can arrive at this statement by taking adjoint as well.)

We show that $P \perp Q$ if and only if $P+Q$ is a projection as well. $P+Q$ is self-adjoint and it is a projection if

$$(P+Q)^2 = P^2 + PQ + QP + Q^2 = P + Q + PQ + QP = P + Q$$

or equivalently

$$PQ + QP = 0.$$

This is true if $P \perp Q$. On the other hand, the condition $PQ + QP = 0$ implies that $PQP + QP^2 = PQP + QP = 0$ and QP must be self-adjoint. We can conclude that $PQ = 0$ which is the orthogonality. □

Assume that P and Q are projections on the same Hilbert space. Among the projections which are smaller than P and Q there is a largest, it is the orthogonal projection onto the intersection of the ranges of P and Q. This has the notation $P \wedge Q$.

Theorem 4.8 *Assume that P and Q are ortho-projections. Then*

$$P \wedge Q = \lim_{n\to\infty} (PQP)^n = \lim_{n\to\infty} (QPQ)^n.$$

Proof: The operator $A := PQP$ is a positive contraction. Therefore the sequence A^n is monotone decreasing and Theorem 4.4 implies that A^n has a limit R. The operator R is selfadjoint. Since $(A^n)^2 \to R^2$ we have $R = R^2$, in other words, R is an ortho-projection. If $Px = x$ and $Qx = x$ for a vector x, then $Ax = x$ and it follows that $Rx = x$. This means that $R \geq P \wedge Q$.

From the inequality $(PQP)^n \leq Q$, $R \leq Q$ follows. Taking the limit of the relation $(QPQ)^n = Q(PQP)^{n-1}Q$, we have

$$\lim_{n\to\infty} (QPQ)^n = QRQ = R.$$

Similarly to the above argument we have that $R \leq P$.

It has been proved that $R \leq P, Q$ and $R \geq P \wedge Q$. So $R = P \wedge Q$ is the only possibility. □

Corollary 4.2 *Assume that P and Q are ortho-projections and $0 \leq H \leq P, Q$. Then $H \leq P \wedge Q$.*

Proof: One can show that $PHP = H, QHQ = H$. This implies $H \leq (PQP)^n$ and the limit $n \to \infty$ gives the result. □

Let P and Q be ortho-projections. If the ortho-projection R has the property $R \geq P, Q$, then the image of R contains the images of P and Q. The smallest such R projects to the linear space generated by the images of P and Q. This ortho-projection is denoted by $P \vee Q$. The set of ortho-projections becomes a lattice with the operations $\wedge$ and $\vee$.

4.6 Kernel functions

Let $\mathcal{X}$ be a nonempty set. A function $\psi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ is often called a **kernel**. A kernel $\psi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ is called **positive definite** if

$$\sum_{j,k=1}^{n} c_j \overline{c_k} \psi(x_j, x_k) \geq 0$$

for all finite sets $\{c_1, c_2, \ldots, c_n\} \subset \mathbb{C}$ and $\{x_1, x_2, \ldots, x_n\} \subset \mathcal{X}$.

Example 4.13 It follows from the Schur theorem that the product of positive definite kernels is a positive definite kernel as well.

If $\psi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ is positive definite, then

$$e^{\psi} = \sum_{n=0}^{\infty} \frac{1}{n!} \psi^m$$

and $\bar{\psi}(x, y) = f(x)\psi(x, y)\overline{f(y)}$ are positive definite for any function $f : \mathcal{X} \to \mathbb{C}$. □

The function $\psi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ is called **conditionally negative definite** kernel if $\psi(x, y) = \overline{\psi(y, x)}$ and

$$\sum_{j,k=1}^{n} c_j \overline{c_k} \psi(x_j, x_k) \leq 0$$

for all finite sets $\{c_1, c_2, \ldots, c_n\} \subset \mathbb{C}$ and $\{x_1, x_2, \ldots, x_n\} \subset \mathcal{X}$ when $\sum_{j=1}^{n} c_j = 0$.

The above properties of a kernel depend on the matrices

$$\begin{bmatrix} \psi(x_1, x_1) & \psi(x_1, x_2) & \ldots & \psi(x_1, x_n) \\ \psi(x_2, x_1) & \psi(x_2, x_2) & \ldots & \psi(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ \psi(x_n, x_1) & \psi(x_n, x_2) & \ldots & \psi(x_n, x_n) \end{bmatrix}.$$

If a kernel f is positive definite, then $-f$ is conditionally negative definite, but the converse is not true.

Lemma 4.1 *Assume that the function $\psi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ has the property $\psi(x, y) = \overline{\psi(y, x)}$ and fix $x_0 \in \mathcal{X}$. Then*

$$\varphi(x, y) := -\psi(x, y) + \psi(x, x_0) + \psi(x_0, y) - \psi(x_0, x_0)$$

is positive definite if and only if ψ is conditionally negative definite.

The proof is rather straightforward, but an interesting particular case is below.

Example 4.14 Assume that $f : \mathbb{R}^+ \to \mathbb{R}$ is a C^1-function with the property $f(0) = f'(0) = 0$. Let $\psi : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ be defined as

$$\psi(x,y) = \begin{cases} \dfrac{f(x)-f(y)}{x-y} & \text{if} \quad x \neq y, \\ \\ f'(x) & \text{if} \quad x = y. \end{cases}$$

(This is the so-called kernel of divided difference.) Assume that this is conditionally negative definite. Now we apply the lemma with $x_0 = \varepsilon$:

$$-\frac{f(x)-f(y)}{x-y} + \frac{f(x)-f(\varepsilon)}{x-\varepsilon} + \frac{f(\varepsilon)-f(y)}{\varepsilon-y} - f'(\varepsilon)$$

is positive definite and from the limit $\varepsilon \to 0$, we have the positive definite kernel

$$-\frac{f(x)-f(y)}{x-y} + \frac{f(x)}{x} + \frac{f(y)}{y} = -\frac{f(x)y^2 - f(y)x^2}{x(x-y)y}.$$

The multiplication by $xy/(f(x)f(y))$ gives a positive definite kernel

$$\frac{\frac{x^2}{f(x)} - \frac{y^2}{f(y)}}{x-y}$$

which is again a divided difference of the function $g(x) := x^2/f(x)$. □

Theorem 4.9 *(**Schoenberg theorem**) Let $\mathcal{X}$ be a nonempty set and let $\psi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ be a kernel. Then ψ is conditionally negative definite if and only if $\exp(-t\psi)$ is positive definite for every $t > 0$.*

Proof: If $\exp(-t\psi)$ is positive definite, then $1 - \exp(-t\psi)$ is negative definite and so is

$$\psi = \lim_{t \to 0} \frac{1}{t}(1 - \exp(-t\psi)).$$

Assume now that ψ is conditionally negative definite. Take $x_0 \in \mathcal{X}$ and set

$$\varphi(x,y) := -\psi(x,y) + \psi(x,x_0) + \psi(x_0,x) - \psi(x_0,x_0)$$

which is positive definite due to the previous lemma. Then

$$e^{-\psi(x,y)} = e^{\varphi(x,y)} e^{-\psi(x,x_0)} \overline{e^{-\psi(y,x_0)}} e^{-\psi(x_0,x_0)}$$

is positive definite. This was $t = 1$, for general $t > 0$ the argument is similar. □

The kernel functions are a kind of generalization of matrices. If $A \in \mathbb{M}_n$, then the corresponding kernel function has $\mathcal{X} := \{1, 2, \ldots, n\}$ and

$$\psi_A(i,j) = A_{ij} \qquad (1 \leq i, j \leq n).$$

Therefore the results of this section have matrix consequences.

4.7 Notes and remarks

Weakly positive matrices were introduced by **Eugene P. Wigner** in 1963. He showed that if the product of two or three weakly positive matrices is self-adjoint, then it is positive definite.



Eugene Paul Wigner (1902–1995)

Wigner was born in Budapest, he attended a high-school together with John von Neumann whose Hungarian name is *Neumann János*. They became good friends, both of them studied first chemistry in university. Wigner received the Nobel Prize in Physics in 1963.

Example 4.11 is related to the concept of *infinite divisibility*, there is a good overview in the paper R. Bhatia and H. Kosaki, Mean matrices and infinite divisibility, Linear Algebra Appl. **424**(2007), 36–54.

The lattice of ortho-projections has applications in quantum theory. An important results says that if ψ is a positive real valued function on the ortho-projections of $\mathbb{M}_n$ with $n \geq 3$ and $\psi(P+Q) = \psi(P) + \psi(Q)$ for all orthogonal ortho-projections, then ψ has a linear extension to $\mathbb{M}_n$. This is the Gleason theorem, see R. Cooke, M. Keane and W. Moran: An elementary proof of Gleason's theorem. Math. Proc. Cambridge Philos. Soc. 98 1985), 117–128.

Positive and conditionally negative definite kernel function are well discussed in the book C. Berg, J.P.R. Christensen and P. Ressel: Harmonic analysis on semigroups. Theory of positive definite and related functions. Graduate Texts in Mathematics, 100. Springer-Verlag, New York, 1984. (It is remarkable that the conditionally negative definite is called there negative definite.)

4.8 Exercises

1. Give an example of $A \in \mathbb{M}_n(\mathbb{C})$ such that the spectrum of A is in $\mathbb{R}^+$ and A is not positive.

2. Let $A \in \mathbb{M}_n(\mathbb{C})$. Show that A is positive if and only if X^*AX is positive for every $X \in \mathbb{M}_n(\mathbb{C})$.

3. Let $A \in B(\mathcal{H})$. Prove the equivalence of the following assertions: (i) $\|A\| \leq 1$, (ii) $A^*A \leq I$, and (iii) $AA^* \leq I$.

4. Let $A \in \mathbb{M}_n(\mathbb{C})$. Show that A is positive if and only if $\operatorname{Tr} XA$ is positive for every positive $X \in \mathbb{M}_n(\mathbb{C})$.

5. Let $\|A\| \leq 1$. Show that there are unitaries U and V such that

$$A = \frac{1}{2}(U + V).$$

(Hint: Use Example 4.2.)

6. Show that a matrix is weakly positive if and only if it is the product of two positive definite matrices.

7. Let $V : \mathbb{C}^n \to \mathbb{C}^n \otimes \mathbb{C}^n$ be defined as $Ve_i = e_i \otimes e_i$. Show that

$$V^*(A \otimes B)V = A \circ B \tag{4.6}$$

for $A, B \in \mathbb{M}_n(\mathbb{C})$. Conclude the **Schur theorem**.

8. Let A be a positive 2×2 matrix with real entries. Show that

$$B_{ij} = (A_{ij})^t$$

is a positive matrix for every $t > 0$. Give a counter example that the similar statement is not true for 3×3 matrices.

9. Assume that P and Q are projections. Show that $P \leq Q$ is equivalent to $PQ = P$.

10. Assume that $P_1, P_2, \ldots, P_n$ are projections and $P_1 + P_2 + \ldots + P_n = I$. Show that the projections are pairwise orthogonal.

11. Let $U|A|$ be the polar decomposition of $A \in \mathbb{M}_n$. Show that M is normal if and only if $U|A| = |A|U$.

12. Let $P \in \mathbb{M}_n$ be idempotent, $P^2 = P$. Show that P is an ortho-projection if and only if $\|P\| \leq 1$.

13. Show that the kernels $\psi(x, y) = \cos(x - y), \quad \cos(x^2 - y^2), \quad (1 + |x - y|)^{-1}$ are positive semidefinite on $\mathbb{R} \times \mathbb{R}$.

14. Assume that the kernel $\psi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ is positive definite and $\psi(x, x) > 0$ for every $x \in \mathcal{X}$. Show that

$$\bar{\psi}(x, y) = \frac{\psi(x, y)}{\psi(x, x)\psi(y, y)}$$

is a positive definite kernel.

15. Assume that the kernel $\psi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ is negative definite and $\psi(x, x) \geq 0$ for every $x \in \mathcal{X}$. Show that
$$\log(1 + \psi(x, y))$$
is a negative definite kernel.

16. Show that the kernel $\psi(x, y) = (\sin(x - y))^2$ is negative semidefinite on $\mathbb{R} \times \mathbb{R}$.

Chapter 5

Functional calculus for matrices

Let $A \in \mathbb{M}_n(\mathbb{C})$ and $p(x) := \sum_i c_i x^i$ be a polynomial. It is quite obvious that by $p(A)$ one means the matrix $\sum_i c_i A^i$.

5.1 The exponential function

The Taylor expansion of the exponential function can be used to define e^A for a matrix $A \in \mathbb{M}_n(\mathbb{C})$:

$$e^A := \sum_{n=0}^{\infty} \frac{A^n}{n!}. \tag{5.1}$$

(The right-hand-side is an absolutely convergent series.)

Theorem 5.1 *If $AB = BA$, then*

$$e^{A+B} = e^A e^B.$$

Proof: We can compute the product $e^A e^B$ by multiplying term by term the series:

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

Therefore,

$$e^A e^B = \sum_{k=0}^{\infty} \frac{1}{k!} C_k,$$

where

$$C_k := \sum_{m+n=k} \frac{k!}{m!n!} A^m B^n.$$

Due to the commutation relation $AB = BA$, the binomial formula holds and $C_k = (A+B)^k$. We conclude

$$e^A e^B = \sum_{k=0}^{\infty} \frac{1}{k!} (A+B)^k$$

which is the statement.

We can give another proof by differentiation. It follows from the expansion (5.1) that the derivative of the matrix-valued function $t \mapsto e^{tA}$ defined on $\mathbb{R}$ is $e^{tA}A$. Therefore,

$$\frac{\partial}{\partial t} e^{tA} e^{C-tA} = e^{tA} A e^{C-tA} - e^{tA} A e^{C-tA} = 0$$

if $AC = CA$. Therefore, the function $t \mapsto e^{tA} e^{C-tA}$ is constant. In particular,

$$e^{A} e^{C-A} = e^{C}.$$

If we put $A + B$ in place of C, we get the statement. □

Example 5.1 In case of 2×2 matrices, the use of the **Pauli matrices**

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma_2 = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}, \qquad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

is efficient, together with I they form an orthogonal system.

Let $A \in \mathbb{M}_2^{sa}$ be such that

$$A = c_1\sigma_1 + c_2\sigma_2 + c_3\sigma_3, \qquad c_1^2 + c_2^2 + c_3^2 = 1$$

in the representation with Pauli matrices. It is simple to check that $A^2 = I$. Therefore, for even powers $A^{2n} = I$, but for odd powers $A^{2n+1} = A$. Choose $c \in \mathbb{R}$ and combine the two facts with the knowledge of the relation of the exponential to sine and cosine:

$$e^{\mathrm{i}cA} = \sum_{n=0}^{\infty} \frac{\mathrm{i}^n c^n A^n}{n!} = \sum_{n=0}^{\infty} \frac{(-1)^n c^{2n} A^{2n}}{(2n)!} + \mathrm{i} \sum_{n=0}^{\infty} \frac{(-1)^n c^{2n+1} A^{2n+1}}{(2n+1)!} = (\cos c) I + \mathrm{i} (\sin c) A$$

A general matrix has the form $C = c_0 I + cA$ and

$$e^{\mathrm{i}C} = e^{\mathrm{i}c_0} (\cos c) I + \mathrm{i} e^{\mathrm{i}c_0} (\sin c) A.$$

(e^C is similar, see Exercise 3.) □

The next theorem gives the so-called **Lie-Trotter formula**.

Theorem 5.2 *Let $A, B \in \mathbb{M}_n(\mathbb{C})$. Then*

$$e^{A+B} = \lim_{n \to \infty} \left(e^{A/n} e^{B/n} \right)^n$$

Proof: First we observe that the identity

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

implies the norm estimate

$$\|X^n - Y^n\| \leq n t^{n-1} \|X - Y\| \tag{5.2}$$

for the submultiplicative operator norm when the constant t is chosen such that $\|X\|, \|Y\| \leq t$.

Now we choose $X_n := \exp((A+B)/n)$ and $Y_n := \exp(A/n)\exp(B/n)$. From the above estimate we have

$$\|X_n^n - Y_n^n\| \leq nu\|X_n - Y_n\|, \tag{5.3}$$

if we can find a constant u such that $\|X_n\|^{n-1}, \|Y_n\|^{n-1} \leq u$. Since

$$\|X_n\|^{n-1} \leq (\exp((\|A\| + \|B\|)/n))^{n-1} \leq \exp(\|A\| + \|B\|)$$

and

$$\|Y_n\|^{n-1} \leq (\exp(\|A\|/n))^{n-1} \times (\exp(\|B\|/n))^{n-1} \leq \exp\|A\| \times \exp\|B\|,$$

$u = \exp(\|A\| + \|B\|)$ can be chosen to have the estimate (5.3).

The theorem follows from (5.3) if we show that $n\|X_n - Y_n\| \to 0$. The power series expansion of the exponential function yields

$$X_n = I + \frac{A+B}{n} + \frac{1}{2}\left(\frac{A+B}{n}\right)^2 + \dots$$

and

$$Y_n = \left(I + \frac{A}{n} + \frac{1}{2}\left(\frac{A}{n}\right)^2 + \dots\right) \times \left(I + \frac{B}{n} + \frac{1}{2}\left(\frac{B}{n}\right)^2 + \dots\right).$$

If $X_n - Y_n$ is computed by multiplying the two series in Y_n, one can observe that all constant terms and all terms containing $1/n$ cancel. Therefore

$$\|X_n - Y_n\| \leq \frac{c}{n^2}$$

for some positive constant c. □

If A and B are self-adjoint matrices, then it can be better to reach e^{A+B} as the limit of self-adjoint matrices.

Corollary 5.1

$$e^{A+B} = \lim_{n\to\infty}\left(e^{\frac{A}{2n}}\, e^{\frac{B}{n}}\, e^{\frac{A}{2n}}\right)^n$$

Proof: We have

$$\left(e^{\frac{A}{2n}}\, e^{\frac{B}{n}}\, e^{\frac{A}{2n}}\right)^n = e^{-\frac{A}{2n}}\left(e^{A/n}e^{B/n}\right)^n e^{\frac{A}{2n}}$$

and the limit $n \to \infty$ gives the result. □

Theorem 5.3 *For matrices $A, B \in \mathbb{M}_n$ the Taylor expansion of the function $\mathbb{R} \ni t \mapsto e^{A+tB}$ is*

$$\sum_{k=0}^{\infty} t^k A_k(1)\,,$$

where $A_0(s) = e^{sA}$ and

$$A_k(s) = \int_0^s dt_1 \int_0^{t_1} dt_2 \dots \int_0^{t_{k-1}} dt_k e^{(s-t_1)A} B e^{(t_1-t_2)A} B \dots B e^{t_k A}$$

for $s \in \mathbb{R}$.

Proof: To make differentiation easier we write

$$A_k(s) = \int_0^s e^{(s-t_1)A} B A_{k-1}(t_1)\, dt_1 = e^{sA} \int_0^s e^{-t_1 A} B A_{k-1}(t_1)\, dt_1$$

for $k \geq 1$. It follows that

$$\begin{aligned} \frac{d}{ds} A_k(s) &= A e^{sA} \int_0^s e^{-t_1 A} B A_{k-1}(t_1)\, dt_1 + e^{sA} \frac{d}{ds} \int_0^s e^{-t_1 A} B A_{k-1}(t_1)\, dt_1 \\ &= A A_k(s) + B A_{k-1}(s). \end{aligned}$$

Therefore

$$F(s) := \sum_{k=0}^{\infty} A_k(s)$$

satisfies the differential equation

$$F'(s) = (A+B)F(s), \qquad F(0) = I.$$

Therefore $F(s) = e^{s(A+B)}$. If $s = 1$ and we write tB in place of B, then we get the expansion of e^{A+tB}. □

Corollary 5.2

$$\frac{\partial}{\partial t} e^{A+tB}\Big|_{t=0} = \int_0^1 e^{uA} B e^{(1-u)A}\, du.$$

The following result is Lieb's extension of the Golden-Thompson inequality.

Theorem 5.4 *(Golden-Thompson-Lieb) Let A, B and C be self-adjoint matrices. Then*

$$\operatorname{Tr} e^{A+B+C} \leq \int_0^{\infty} \operatorname{Tr} e^A (t + e^{-C})^{-1} e^B (t + e^{-C})^{-1}\, dt\,.$$

When $C = 0$, the we have

$$\operatorname{Tr} e^{A+B} \leq \operatorname{Tr} e^A e^B \tag{5.4}$$

which is the original Golden-Thompson inequality. If $BC = CB$, then in the right-hand-side, the integral

$$\int_0^{\infty} (t + e^{-C})^{-2}\, dt$$

appears. This equals to e^C and we have $\operatorname{Tr} e^{A+B+C} \leq \operatorname{Tr} e^A e^B e^C$. Without the assumption $BC = CB$, this inequality is not true.

An example of the application of the Golden-Thompson-Lieb inequality is the **strong subadditivity** of the von Neumann entropy (defined by (5.5)).

Example 5.2 The **partial trace** $\operatorname{Tr}_1 : \mathbb{M}_k \otimes \mathbb{M}_m \to \mathbb{M}_m$ is a linear mapping which is defined by the formula $\operatorname{Tr}_1(A \otimes B) = (\operatorname{Tr} A)B$ on elementary tensors. It is called partial trace, since trace of the first tensor factor was taken. $\operatorname{Tr}_2 : \mathbb{M}_k \otimes \mathbb{M}_m \to \mathbb{M}_k$ is similarly defined. We shall need this concept for three-fold tensor product.

Recall that $D \in \mathbb{M}_k$ is a density matrix if $0 \le D$ and $\operatorname{Tr} D = 1$. The von Neumann entropy of a density matrix D is $S(D) = -\operatorname{Tr} D \log D$ (see also in (5.26)).

Let D_{123} be a density matrix in $\mathbb{M}_k \otimes \mathbb{M}_l \otimes \mathbb{M}_m$. The reduced density matrices are defined by the partial traces

$$D_{12} := \operatorname{Tr}_1 D_{123} \in \mathbb{M}_k \otimes \mathbb{M}_l, \quad D_2 := \operatorname{Tr}_{13} D_{123} \in \mathbb{M}_l \quad \text{and} \quad D_{23} := \operatorname{Tr}_1 D_{123} \in \mathbb{M}_k$$

The **strong subadditivity** is the inequality

$$S(D_{123}) + S(D_2) \le S(D_{12}) + S(D_{23}). \tag{5.5}$$

which is equivalent to

$$\operatorname{Tr} D_{123} \left(\log D_{123} - (\log D_{12} - \log D_2 + \log D_{23})\right) \ge 0.$$

The operator

$$\exp(\log D_{12} - \log D_2 + \log D_{23})$$

is positive and can be written as λD for a density matrix D. Actually,

$$\lambda = \operatorname{Tr} \exp(\log D_{12} - \log D_2 + \log D_{23}).$$

We have

$$\begin{aligned} &S(D_{12}) + S(D_{23}) - S(D_{123}) - S(D_2) \\ &\qquad = \operatorname{Tr} D_{123} \left(\log D_{123} - (\log D_{12} - \log D_2 + \log D_{23})\right) \\ &\qquad = S(D_{123} \| \lambda D) = S(D_{123} \| D) - \log \lambda \end{aligned} \tag{5.6}$$

Here $S(X\|Y) := \operatorname{Tr} X(\log X - \log Y)$ is the **relative entropy**. If X and Y are density matrices, then $S(X\|Y) \ge 0$. Therefore, $\lambda \le 1$ implies the positivity of the left-hand-side (and the strong subadditivity). Due to Theorem 5.4, we have

$$\operatorname{Tr} \exp(\log D_{12} - \log D_2 + \log D_{23})) \le \int_0^\infty \operatorname{Tr} D_{12}(tI + D_2)^{-1} D_{23} (tI + D_2)^{-1} \, dt$$

Applying the partial traces we have

$$\operatorname{Tr} D_{12}(tI + D_2)^{-1} D_{23}(tI + D_2)^{-1} = \operatorname{Tr} D_2 (tI + D_2)^{-1} D_2 (tI + D_2)^{-1}$$

and that can be integrated out. Hence

$$\int_0^\infty \operatorname{Tr} D_{12}(tI + D_2)^{-1} D_{23}(tI + D_2)^{-1} \, dt = \operatorname{Tr} D_2 = 1.$$

and $\lambda \le 1$ is obtained and the strong subadditivity is proven.

If the equality holds in (5.5), then $\exp(\log D_{12} - \log D_2 + \log D_{23})$ is a density matrix and

$$S(D_{123} \| \exp(\log D_{12} - \log D_2 + \log D_{23})) = 0$$

implies

$$\log D_{123} = \log D_{12} - \log D_2 + \log D_{23}. \tag{5.7}$$

This is the necessary and sufficient condition for the equality. □

A function $f : \mathbb{R}^+ \to \mathbb{R}$ is the **Laplace transform** of a measure μ on $[0, \infty) \subset \mathbb{R}$ if

$$f(t) = \int_0^\infty e^{-tx}\, d\mu(x) \qquad (t \in \mathbb{R}^+).$$

According to the **Bernstein theorem** such a measure μ exists if and only if the nth derivative of f has the sign $(-1)^n$ on the whole $\mathbb{R}^+$ and for every $n \in \mathbb{N}$. (Such a function is often called completely monotone.)

Statement (ii) in the next theorem is the Bessis-Moussa-Villani conjecture from 1975. The theorem is due to Lieb and Seiringer. It gives equivalent conditions, property (i) has a very simple formulation.

Theorem 5.5 *Let $A, B \in \mathbb{M}_n^{sa}$ and let $t \in \mathbb{R}$. The following statements are equivalent:*

(i) The polynomial $t \mapsto \mathrm{Tr}\,(A + tB)^p$ has only positive coefficients for every A and $B \geq 0$ and all $p \in \mathbb{N}$.

(ii) For every A and $B \geq 0$, the function $t \mapsto \mathrm{Tr}\,\exp(A - tB)$ is the Laplace transform of a positive measure supported in $[0, \infty)$.

(iii) For every $A > 0$, $B \geq 0$ and all $p \geq 0$, the function $t \mapsto \mathrm{Tr}\,(A + tB)^{-p}$ is the Laplace transform of a positive measure supported in $[0, \infty)$.

Proof: (i)⇒(ii): We have

$$\mathrm{Tr}\,\exp(A - tB) = e^{-\|A\|} \sum_{k=0}^{\infty} \frac{1}{k!} \mathrm{Tr}\,(A + \|A\| I - tB)^k \tag{5.8}$$

and it follows from Bernstein's theorem and (i) that the right-hand-side is the Laplace transform of a positive measure supported in $[0, \infty)$.

(ii)⇒(iii): This follows from taking the trace of the matrix equation

$$(A + tB)^{-p} = \frac{1}{\Gamma(p)} \int_0^\infty \exp\left[-u(A + tB)\right] u^{p-1} du\ . \tag{5.9}$$

(iii)⇒(i): It suffices to assume (iii) only for $p \in \mathbb{N}$. For invertible A we observe that the r-th derivative of $\mathrm{Tr}\,(A_0 + tB_0)^{-p}$ at $t = 0$ is related to the coefficient of t^r in $\mathrm{Tr}\,(A + tB)^p$ as given by (5.31) with A, A_0, B, B_0 related as in Lemma 5.1. The left side of (5.31) has the sign $(-1)^r$ because it is the derivative of a function that is the Laplace transform of a positive measure supported in $[0, \infty)$. Thus the right-hand-side has the correct sign as stated in item (i). The case of non-invertible A follows from continuity argument. □

5.2 Other functions

All reasonable functions can be approximated by polynomials. Therefore, it is basic to compute $p(X)$ for a matrix $X \in \mathbb{M}_n$ and for a polynomial p. The canonical Jordan

decomposition

$$X = S \begin{bmatrix} J_{k_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{k_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{k_m}(\lambda_m) \end{bmatrix} S^{-1} = SJS^{-1},$$

gives that

$$p(X) = S \begin{bmatrix} p(J_{k_1}(\lambda_1)) & 0 & \cdots & 0 \\ 0 & p(J_{k_2}(\lambda_2)) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p(J_{k_m}(\lambda_m)) \end{bmatrix} S^{-1} = Sp(J)S^{-1}.$$

The crucial point is the computation of $(J_k(\lambda))^m$. Since $J_k(\lambda) = \lambda I_n + J_k(0) = \lambda I_n + J_k$ is the sum of commuting matrices, to compute the mth power, we can use the binomial formula:

$$(J_k(\lambda))^m = \lambda^m I_n + \sum_{j=1}^{m} \binom{m}{j} \lambda^{m-j} J_k^j$$

The powers of J_k are known, see Example 2.1. Let $m > 3$, then the example

$$J_4(\lambda)^m = \begin{bmatrix} \lambda^m & m\lambda^{m-1} & \dfrac{m(m-1)\lambda^{m-2}}{2!} & \dfrac{m(m-1)(m-2)\lambda^{m-3}}{3!} \\ 0 & \lambda^m & m\lambda^{m-1} & \dfrac{m(m-1)\lambda^{m-2}}{2!} \\ 0 & 0 & \lambda^m & m\lambda^{m-1} \\ 0 & 0 & 0 & \lambda^m \end{bmatrix}.$$

shows the point. In another formulation,

$$p(J_4(\lambda)) = \begin{bmatrix} p(\lambda) & p'(\lambda) & \dfrac{p''(\lambda)}{2!} & \dfrac{p^{(3)}(\lambda)}{3!} \\ 0 & p(\lambda) & p'(\lambda) & \dfrac{p''(\lambda)}{2!} \\ 0 & 0 & p(\lambda) & p'(\lambda) \\ 0 & 0 & 0 & p(\lambda) \end{bmatrix}.$$

which is actually correct for all polynomials and for every smooth function. We conclude that if the canonical Jordan form is known for $X \in \mathbb{M}_n$, then $f(X)$ is computable. In particular, the above argument gives the following result.

Theorem 5.6 *For $X \in \mathbb{M}_n$ the relation*

$$\det e^X = \exp(\operatorname{Tr} X)$$

holds between trace and determinant.

The matrix $A \in \mathbb{M}_n$ is **diagonizable** if

$$A = S\,\mathrm{Diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)S^{-1}$$

with an invertible matrix S. Observe that this condition means that in the Jordan canonical form all Jordan blocks are 1×1 and the numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A. In this case

$$f(A) = S\,\mathrm{Diag}(f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n))S^{-1} \tag{5.10}$$

when the complex-valued function f is defined on the set of eigenvalues of A.

Example 5.3 We consider the matrix

$$X = \begin{bmatrix} 1+z & x-y\mathrm{i} \\ x+y\mathrm{i} & 1-z \end{bmatrix} \equiv \begin{bmatrix} 1+z & w \\ \bar{w} & 1-z \end{bmatrix}$$

when $x, y, x \in \mathbb{R}$. From the characteristic polynomial we have the eigenvalues

$$\lambda_1 = 1 + R \qquad \text{and} \qquad \lambda_2 = 1 - R,$$

where $R = \sqrt{x^2 + y^2 + z^2}$. If $R < 1$, then X is positive and invertible. The eigenvectors are

$$u_1 = (R + z, \bar{w}) \qquad \text{and} \qquad u_2 = (R - z, -\bar{w}).$$

Set

$$\Delta = \begin{bmatrix} 1+R & 0 \\ 0 & 1-R \end{bmatrix}, \qquad S = \begin{bmatrix} R+z & R-z \\ \bar{w} & -\bar{w} \end{bmatrix}.$$

We can check that $XS = S\Delta$, hence

$$X = S\Delta S^{-1}.$$

To compute S^{-1} we use the formula

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Hence

$$S^{-1} = \frac{1}{2\bar{w}R} \begin{bmatrix} \bar{w} & R-z \\ \bar{w} & -R-z \end{bmatrix}.$$

It follows that

$$X^t = a_t \begin{bmatrix} b_t + z & w \\ \bar{w} & b_t - z \end{bmatrix},$$

where

$$a_t = \frac{(1+R)^t - (1-R)^t}{2R}, \qquad b_t = R\,\frac{(1+R)^t + (1-R)^t}{(1+R)^t - (1-R)^t}.$$

The matrix $X/2$ is a density matrix and has applications in quantum theory. □

The function used are typically continuous. By C^n-function on an interval of the real line we mean n-times differentiable function whose nth derivative is still continuous.

Remember that self-adjoint matrices are diagonalizable and they have a spectral decomposition. Let $A = \sum_i \lambda_i P_i$ be the spectral decomposition of the self-adjoint $A \in \mathbb{M}_n(\mathbb{C})$. ($\lambda_i$ are the different eigenvalues and P_i are the corresponding eigenprojections, the rank of P_i is the multiplicity of λ_i.) Then

$$f(A) = \sum_i f(\lambda_i) P_i \,. \tag{5.11}$$

Usually we assume that f is continuous on an interval containing the eigenvalues of A.

Example 5.4 Consider

$$f_+(t) := \max\{t, 0\} \quad \text{and} \quad f_-(t) := \max\{-t, 0\} \quad \text{for} \quad t \in \mathbb{R}.$$

For each $A \in B(\mathcal{H})^{sa}$ define

$$A_+ := f_+(A) \quad \text{and} \quad A_- := f_-(A).$$

Since $f_+(t), f_-(t) \geq 0$, $f_+(t) - f_-(t) = t$ and $f_+(t) f_-(t) = 0$, we have

$$A_+, A_- \geq 0, \quad A = A_+ - A_-, \quad A_+ A_- = 0.$$

These A_+ and A_- are called the **positive part** and the **negative part** of A, respectively, and $A = A_+ + A_-$ is called the **Jordan decomposition** of A. □

Theorem 5.7 *If f_k and g_k are functions $(\alpha, \beta) \to \mathbb{R}$ such that for some $c_k \in \mathbb{R}$*

$$\sum_k c_k f_k(x) g_k(y) \geq 0$$

for every $x, y \in (\alpha, \beta)$, then

$$\sum_k c_k \mathrm{Tr}\, f_k(A) g_k(B) \geq 0$$

whenever A, B are self-adjoint matrices with spectrum in (α, β).

Proof: Let $A = \sum_i \lambda_i p_i$ and $B = \sum_j \mu_j q_j$ be the spectral decompositions. Then

$$\begin{aligned} \sum_k c_k \mathrm{Tr}\, f_k(A) g_k(B) &= \sum_k \sum_{i,j} c_k \mathrm{Tr}\, p_i f_k(A) g_k(B) q_j \\ &= \sum_{i,j} \mathrm{Tr}\, p_i q_j \sum_k c_k f_k(\lambda_i) g_k(\mu_j) \geq 0 \end{aligned}$$

due to the hypothesis. □

Example 5.5 In order to show the application of the previous theorem, assume that f is convex. Then

$$f(x) - f(y) - (x - y) f'(y) \geq 0$$

and

$$\operatorname{Tr} f(A) \geq \operatorname{Tr} f(B) + \operatorname{Tr} (A-B)f'(B)\,. \tag{5.12}$$

Replacing f by $\eta(t) = -t\log t$ we have

$$-\operatorname{Tr} A\log A \geq -\operatorname{Tr} B\log B - \operatorname{Tr} (A-B) - \operatorname{Tr} (A-B)\log B$$

or equivalently

$$\operatorname{Tr} A(\log A - \log B) \geq \operatorname{Tr} (A-B). \tag{5.13}$$

The left-hand-side is the **relative entropy** of the positive matrices A and B. If $\operatorname{Tr} A = \operatorname{Tr} B = 1$, then the lower bound is 0.

Concerning the relative entropy we can have a better estimate. If $\operatorname{Tr} A = \operatorname{Tr} B = 1$, then all eigenvalues are in $[0,1]$. Analysis tells us that for some $\xi \in (x,y)$

$$-\eta(x) + \eta(y) + (x-y)\eta'(y) = -\frac{1}{2}(x-y)^2\eta''(\xi) \geq \frac{1}{2}(x-y)^2 \tag{5.14}$$

when $x, y \in [0,1]$. According to Theorem 5.7 we have

$$\operatorname{Tr} A(\log A - \log B) \geq \tfrac{1}{2}\operatorname{Tr} (A-B)^2. \tag{5.15}$$

The **Streater inequality** (5.15) has the consequence that $A = B$ if the relative entropy is 0. □

Let f be holomorphic inside and on a positively oriented simple contour Γ in the complex plane and let A be an $n \times n$ matrix such that its eigenvalues are inside of Γ. Then

$$f(A) := \frac{1}{2\pi i}\int_\Gamma f(z)(zI - A)^{-1}\,dz \tag{5.16}$$

is defined by a contour integral. When A is self-adjoint, then (5.11) makes sense and it is an exercise to show that it gives the same result as (5.16).

Example 5.6 We can define the square root function on the set

$$\mathbb{C}^+ := \{Re^{\mathrm{i}\varphi} \in \mathbb{C} : R > 0,\ -\pi/2 < \varphi < \pi/2\}$$

as $\sqrt{Re^{\mathrm{i}\varphi}} := \sqrt{R}e^{\mathrm{i}\varphi/2}$ and this is a holomorphic function on $\mathbb{C}^+$.

When $X = S\operatorname{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n)\, S^{-1} \in \mathbb{M}_n$ is a weakly positive matrix, then $\lambda_1, \lambda_2, \dots, \lambda_n > 0$ and to use (5.16) we can take a positively oriented simple contour Γ in $\mathbb{C}^+$ such that the eigenvalues are inside. Then

$$\begin{aligned}
\sqrt{X} &= \frac{1}{2\pi i}\int_\Gamma \sqrt{z}(zI - X)^{-1}\,dz \\
&= S\left(\frac{1}{2\pi i}\int_\Gamma \sqrt{z}\operatorname{Diag}(1/(z-\lambda_1), 1/(z-\lambda_2), \dots, 1/(z-\lambda_n))\,dz\right) S^{-1} \\
&= S\operatorname{Diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n})\, S^{-1}.
\end{aligned}$$

□

5.3 Derivation

This section contains derivatives of number-valued and matrix-valued functions. From the latter one number-valued can be obtained by trace, for example.

Example 5.7 Assume that $A \in \mathbb{M}_n$ is invertible. Then $A + tT$ is invertible as well for $T \in \mathbb{M}_n$ and for small real number t. The identity

$$(A+tT)^{-1} - A^{-1} = (A+tT)^{-1}(A - (A+tT))A^{-1} = -t(A+tT)^{-1}TA^{-1},$$

gives

$$\lim_{t\to 0} \frac{1}{t}\Big((A+tT)^{-1} - A^{-1}\Big) = -A^{-1}TA^{-1}.$$

The derivative is computed at $t = 0$, but if $A + tT$ is invertible, then

$$\frac{d}{dt}(A+tT)^{-1} = -(A+tT)^{-1}T(A+tT)^{-1} \tag{5.17}$$

by similar computation. We can continue the derivation:

$$\frac{d^2}{dt^2}(A+tT)^{-1} = 2(A+tT)^{-1}T(A+tT)^{-1}T(A+tT)^{-1} \tag{5.18}$$

$$\frac{d^3}{dt^3}(A+tT)^{-1} = -6(A+tT)^{-1}T(A+tT)^{-1}T(A+tT)^{-1}T(A+tT)^{-1} \tag{5.19}$$

So the Taylor expansion is

$$\begin{aligned}(A+tT)^{-1} &= A^{-1} - tA^{-1}TA^{-1} + t^2A^{-1}TA^{-1}TA^{-1} - t^3A^{-1}TA^{-1}TA^{-1}TA^{-1} + \dots \\ &= \sum_{n=0}^{\infty}(-t)^n A^{-1/2}(A^{-1/2}TA^{-1/2})^n A^{-1/2}. \end{aligned} \tag{5.20}$$

Since

$$(A+tT)^{-1} = A^{-1/2}(I + tA^{-1/2}TA^{-1/2})^{-1}A^{-1/2}$$

we can get the Taylor expansion also from the Neumann series of $(I + tA^{-1/2}TA^{-1/2})^{-1}$, see Example 1.4. □

Example 5.8 Assume that $A \in \mathbb{M}_n$ is positive invertible. Then $A + tT$ is positive invertible as well for $T \in \mathbb{M}_n^{sa}$ and for a small real number t. Therefore $\log(A + tT)$ is defined and it is expressed as

$$\log(A+tT) = \int_0^\infty (x+1)^{-1}I - (x+A+tT)^{-1}\,dx.$$

This is a convenient formula for the derivation (with respect to $t \in \mathbb{R}$):

$$\frac{d}{dt}\log(A+tT) = \int_0^\infty (x+A+tT)^{-1}T(x+A+tT)^{-1}\,dx$$

from the derivative of the inverse. The derivation can be continued and we have the Taylor expansion

$$\begin{aligned}\log(A+tT) &= \log A + t\int_0^\infty (x+A)^{-1}T(x+A)^{-1}\,dx \\ &\qquad -t^2\int_0^\infty (x+A)^{-1}T(x+A)^{-1}T(x+A)^{-1}\,dx + \dots \\ &= \log A - \sum_{n=1}^\infty (-t)^n \int_0^\infty (x+A)^{-1/2}((x+A)^{-1/2}T(x+A)^{-1/2})^n(x+A)^{-1/2}\,dx\end{aligned}$$

□

Theorem 5.8 *Let $A, B \in \mathbb{M}_n(\mathbb{C})$ be self-adjoint matrices and $t \in \mathbb{R}$. Assume that $f : (\alpha, \beta) \to \mathbb{R}$ is a continuously differentiable function defined on an interval and assume that the eigenvalues of $A + tB$ are in (α, β) for small $t - t_0$. Then*

$$\frac{d}{dt}\mathrm{Tr}\, f(A+tB)\Big|_{t=t_0} = \mathrm{Tr}\,(Bf'(A+t_0B))\,.$$

Proof: One can verify the formula for a polynomial f by an easy direct computation: $\mathrm{Tr}\,(A+tB)^n$ is a polynomial of the real variable t. We are interested in the coefficient of t which is

$$\mathrm{Tr}\,(A^{n-1}B + A^{n-2}BA + \dots + ABA^{n-2} + BA^{n-1}) = n\mathrm{Tr}\, A^{n-1}B.$$

We have the result for polynomials and the formula can be extended to a more general f by means of polynomial approximation. □

Example 5.9 Let $f : (\alpha, \beta) \to \mathbb{R}$ be a continuous increasing function and assume that the spectrum of the self-adjoint matrices B and C lies in (α, β). We use the previous theorem to show that

$$A \le C \quad \text{implies} \quad \mathrm{Tr}\, f(A) \le \mathrm{Tr}\, f(C). \tag{5.21}$$

We may assume that f is smooth and it is enough to show that the derivative of $\mathrm{Tr}\, f(A+tB)$ is positive when $B \ge 0$. (To observe (5.21), one takes $B = C - A$.) The derivative is $\mathrm{Tr}\,(Bf'(A+tB))$ and this is the trace of the product of two positive operators. Therefore, it is positive. □

For a holomorphic function f, we can compute the derivative of $f(A+tB)$ on the basis of (5.16), where Γ is a positively oriented simple contour satisfying the properties required above. The derivation is reduced to the differentiation of the resolvent $(zI-(A+tB))^{-1}$ and we obtain

$$X := \frac{d}{dt}f(A+tB)\Big|_{t=0} = \frac{1}{2\pi\mathrm{i}}\int_\Gamma f(z)(zI-A)^{-1}B(zI-A)^{-1}\,dz\,. \tag{5.22}$$

When A is self-adjoint, then it is not a restriction to assume that it is diagonal, $A = \mathrm{Diag}(t_1, t_2, \dots, t_n)$, and we compute the entries of the matrix (5.22) using the Frobenius formula

$$f[t_i, t_j] := \frac{f(t_i) - f(t_j)}{t_i - t_j} = \frac{1}{2\pi\mathrm{i}}\int_\Gamma \frac{f(z)}{(z-t_i)(z-t_j)}\,dz\,.$$

Therefore,

$$X_{ij} = \frac{1}{2\pi \mathrm{i}} \int_\Gamma f(z) \frac{1}{z - t_i} B_{ij} \frac{1}{z - t_j}\, dz = \frac{f(t_i) - f(t_j)}{t_i - t_j} B_{ij}.$$

A C^1 function can be approximated by polynomials, hence we have the following result.

Theorem 5.9 *Assume that $f : (\alpha, \beta) \to \mathbb{R}$ is C^1 function and $A = \mathrm{Diag}(t_1, t_2, \dots, t_n)$ with $\alpha < t_i < \beta$ $(1 \le i \le n)$. If $B = B^*$, then the derivative $t \mapsto f(A + tB)$ is an Hadamard product:*

$$\frac{d}{dt} f(A + tB)\Big|_{t=0} = D \circ B, \tag{5.23}$$

where D is the divided difference matrix,

$$D_{ij} = \begin{cases} \dfrac{f(t_i) - f(t_j)}{t_i - t_j} & \text{if} \quad t_i - t_j \neq 0, \\ \\ f'(t_i) & \text{if} \quad t_i - t_j = 0. \end{cases} \tag{5.24}$$

Let $f : (\alpha, \beta) \to \mathbb{R}$ be a continuous function. It is called **matrix monotone** if

$$A \le C \quad \text{implies} \quad f(A) \le f(C) \tag{5.25}$$

when the spectrum of the self-adjoint matrices B and C lies in (α, β).

Theorem 4.5 tells us that $f(x) = -1/x$ is matrix monotone function. Matrix monotonicity means that $f(A + tB)$ is an increasing function when $B \ge 0$. The increasing property is equivalent to the positivity of the derivative. We use the previous theorem to show that the function $f(x) = \sqrt{x}$ is matrix monotone.

Example 5.10 Assume that $A > 0$ is diagonal: $A = \mathrm{Diag}(t_1, t_2, \dots, t_n)$. Then derivative of the function $\sqrt{A + tB}$ is $D \circ B$, where

$$D_{ij} = \begin{cases} \dfrac{1}{\sqrt{t_i} + \sqrt{t_j}} & \text{if} \quad t_i - t_j \neq 0, \\ \\ \dfrac{1}{2\sqrt{t_i}} & \text{if} \quad t_i - t_j = 0. \end{cases}$$

This is a Cauchy matrix, see Example 4.4 and it is positive. If B is positive, then so is the Hadamard product. We have shown that the derivative is positive, hence $f(x) = \sqrt{x}$ is matrix monotone.

The idea of another proof is in Exercise 18. □

A subset $K \subset \mathbb{M}_n$ is **convex** if for $A, B \in K$ and for a real number $0 < \lambda < 1$

$$\lambda A + (1 - \lambda) B \in K.$$

The functional $F : K \to \mathbb{R}$ is convex if for $A, B \in K$ and for a real number $0 < \lambda < 1$ the inequality

$$F(\lambda A + (1 - \lambda) B) \le \lambda F(A) + (1 - \lambda) F(B)$$

holds. This inequality is equivalent to the convexity of the function

$$G : [0, 1] \to \mathbb{R}, \quad G(\lambda) := F(B + \lambda(A - B)).$$

It is well-known in analysis that the convexity is related to the second derivative.

Theorem 5.10 *Let K be the set of self-adjoint $n \times n$ matrices with spectrum in the interval (α, β). Assume that the function $f : (\alpha, \beta) \to \mathbb{R}$ is a convex C^2 function. Then the functional $A \mapsto \operatorname{Tr} f(A)$ is convex on K.*

Proof: The stated convexity is equivalent with the convexity of the numerical functions

$$t \mapsto \operatorname{Tr} f(tX_1 + (1-t)X_2) = \operatorname{Tr}(X_2 + t(X_1 - X_2)) \qquad (t \in [0,1]).$$

It is enough to prove that the second derivative of $t \mapsto \operatorname{Tr} f(A+tB)$ is positive at $t = 0$.

The first derivative of the functional $t \mapsto \operatorname{Tr} f(A+tB)$ is $\operatorname{Tr} f'(A+tB)B$. To compute the second derivative we differentiate $f'(A+tB)$. We can assume that A is diagonal and we differentiate at $t = 0$. We have to use (5.23) and get

$$\Big[\frac{d}{dt} f'(A+tB)\Big|_{t=0}\Big]_{ij} = \frac{f'(t_i) - f'(t_j)}{t_i - t_j} B_{ij}\,,$$

Therefore,

$$\begin{aligned}
\frac{d^2}{dt^2} \operatorname{Tr} f(A+tB)\Big|_{t=0} &= \operatorname{Tr}\Big[\frac{d}{dt} f'(A+tB)\Big|_{t=0}\Big] B \\
&= \sum_{ik} \Big[\frac{d}{dt} f'(A+tB)\Big|_{t=0}\Big]_{ik} B_{ki} \\
&= \sum_{ik} \frac{f'(t_i) - f'(t_k)}{t_i - t_k} B_{ik} B_{ki} \\
&= \sum_{ik} f''(s_{ik}) |B_{ik}|^2,
\end{aligned}$$

where s_{ik} is between t_i and t_k. The convexity of f means $f''(s_{ik}) \geq 0$. Therefore we conclude the positivity. □

Note that another, less analytic, proof is sketched in Exercise 14.

Example 5.11 The function

$$\eta(x) = \begin{cases} -x \log x & \text{if} \quad 0 < x, \\ 0 & \text{if} \quad x = 0 \end{cases}$$

is continuous and concave on $\mathbb{R}^+$. For a positive matrix $D \geq 0$

$$S(D) := \operatorname{Tr} \eta(D) \tag{5.26}$$

is called **von Neumann entropy**. It follows from the previous theorem that $S(D)$ is a concave function of D. If we are very rigorous, then we cannot apply the theorem, since η is not differentiable at 0. Therefore we should apply the theorem to $f(x) := \eta(x + \varepsilon)$, where $\varepsilon > 0$ and take the limit $\varepsilon \to 0$. □

Example 5.12 Let a self-adjoint matrix H be fixed. The state of a quantum system is described by a density matrix D which has the properties $D \geq 0$ and $\operatorname{Tr} D = 1$. The equilibrium state is minimizing the energy

$$F(D) = \operatorname{Tr} DH - \frac{1}{\beta} S(D),$$

where β is a positive number. To find the minimizer, we solve the equation

$$\frac{\partial}{\partial t}F(D+tX)\Big|_{t=0}=0$$

for self-adjoint matrices X with the property $\operatorname{Tr} X=0$. The equation is

$$\operatorname{Tr} X\left(H+\frac{1}{\beta}\log D+\frac{1}{\beta}I\right)=0$$

and

$$H+\frac{1}{\beta}\log D+\frac{1}{\beta}I$$

must be cI. Hence the minimizer is

$$D=\frac{e^{-\beta H}}{\operatorname{Tr} e^{-\beta H}} \tag{5.27}$$

which is called **Gibbs state**. □

Example 5.13 Next we restrict ourselves to the self-adjoint case $A,B\in\mathbb{M}_n(\mathbb{C})^{sa}$ in the analysis of (5.22).

The space $M_n(\mathbb{C})^{sa}$ can be decomposed as $\mathcal{M}_A\oplus\mathcal{M}_A^\perp$, where $\mathcal{M}_A:=\{C\in\mathbb{M}_n(\mathbb{C})^{sa}: CA=AC\}$ is the commutant of A and $\mathcal{M}_A^\perp$ is its orthogonal complement. When the operator $\mathcal{L}_A: X\mapsto \mathrm{i}(AX-XA)\equiv \mathrm{i}[A,X]$ is considered, $\mathcal{M}_A$ is exactly the kernel of $\mathcal{L}_A$, while $\mathcal{M}_A^\perp$ is its range.

When $B\in\mathcal{M}_A$, then

$$\frac{1}{2\pi\mathrm{i}}\int_\Gamma f(z)(zI-A)^{-1}B(zI-A)^{-1}\,dz=\frac{B}{2\pi\mathrm{i}}\int_\Gamma f(z)(zI-A)^{-2}\,dz=Bf'(A)$$

and we have

$$\frac{d}{dt}f(A+tB)\Big|_{t=0}=Bf'(A)\,. \tag{5.28}$$

When $B=\mathrm{i}[A,X]\in\mathcal{M}_A^\perp$, then we use the identity

$$(zI-A)^{-1}[A,X](zI-A)^{-1}=[(zI-A)^{-1},X]$$

and we conclude

$$\frac{d}{dt}f(A+t\mathrm{i}[A,X])\Big|_{t=0}=\mathrm{i}[f(A),X]\,. \tag{5.29}$$

To compute the derivative in an arbitrary direction B we should decompose B as $B_1\oplus B_2$ with $B_1\in\mathcal{M}_A$ and $B_2\in\mathcal{M}_A^\perp$. Then

$$\frac{d}{dt}f(A+tB)\Big|_{t=0}=B_1f'(A)+\mathrm{i}[f(A),X]\,, \tag{5.30}$$

where X is the solution of the equation $B_2=\mathrm{i}[A,X]$. □

Lemma 5.1 *Let $A_0,B_0\in\mathbb{M}_n^{sa}$ and assume $A_0>0$. Define $A=A_0^{-1}$ and $B=A_0^{-1/2}B_0A_0^{-1/2}$, and let $t\in\mathbb{R}$. For all $p,r\in\mathbb{N}$*

$$\frac{d^r}{dt^r}\operatorname{Tr}(A_0+tB_0)^{-p}\Big|_{t=0}=\frac{p}{p+r}(-1)^r\,\frac{d^r}{dt^r}\operatorname{Tr}(A+tB)^{p+r}\Big|_{t=0}\,. \tag{5.31}$$

Proof: By induction it is easy to show that

$$\frac{d^r}{dt^r}(A+tB)^{p+r} = r! \sum_{\substack{0\le i_1,\ldots,i_{r+1}\le p \\ \sum_j i_j = p}} (A+tB)^{i_1} B \ldots B (A+tB)^{i_{r+1}} \ .$$

By taking the trace at $t=0$ we obtain

$$I_1 \equiv \frac{d^r}{dt^r} \mathrm{Tr}\,(A+tB)^{p+r}\Big|_{t=0} = r! \sum_{\substack{0\le i_1,\ldots,i_{r+1}\le p \\ \sum_j i_j = p}} \mathrm{Tr}\, A^{i_1} B \cdots B A^{i_{r+1}} \ .$$

Moreover, by similar arguments,

$$\frac{d^r}{dt^r}(A_0+tB_0)^{-p} = (-1)^r r! \sum_{\substack{1\le i_1,\ldots,i_{r+1}\le p \\ \sum_j i_j = p+r}} (A_0+tB_0)^{-i_1} B_0 \cdots B_0 (A_0+tB_0)^{-i_{r+1}} \ .$$

By taking the trace at $t=0$ and using cyclicity, we get

$$I_2 \equiv \frac{d^r}{dt^r} \mathrm{Tr}\,(A_0+tB_0)^{-p}\Big|_{t=0} = (-1)^r r! \sum_{\substack{0\le i_1,\ldots,i_{r+1}\le p-1 \\ \sum_j i_j = p-1}} \mathrm{Tr}\, A\, A^{i_1} B \cdots B A^{i_{r+1}} \ .$$

We have to show that

$$I_2 = \frac{p}{p+r}(-1)^r I_1 \ .$$

To see this we rewrite I_1 in the following way. Define $p+r$ matrices M_j by

$$M_j = \begin{cases} B & \text{for } 1 \le j \le r \\ A & \text{for } r+1 \le j \le r+p \ . \end{cases}$$

Let $\mathcal{S}_n$ denote the permutation group. Then

$$I_1 = \frac{1}{p!} \sum_{\pi \in \mathcal{S}_{p+r}} \mathrm{Tr} \prod_{j=1}^{p+r} M_{\pi(j)} \ .$$

Because of the cyclicity of the trace we can always arrange the product such that M_{p+r} has the first position in the trace. Since there are $p+r$ possible locations for M_{p+r} to appear in the product above, and all products are equally weighted, we get

$$I_1 = \frac{p+r}{p!} \sum_{\pi \in \mathcal{S}_{p+r-1}} \mathrm{Tr}\, A \prod_{j=1}^{p+r-1} M_{\pi(j)} \ .$$

On the other hand,

$$I_2 = (-1)^r \frac{1}{(p-1)!} \sum_{\pi \in \mathcal{S}_{p+r-1}} \mathrm{Tr}\, A \prod_{j=1}^{p+r-1} M_{\pi(j)} \ ,$$

so we arrive at the desired equality. □

5.4 Notes and remarks

The first proof of (5.5) is due to E.H. Lieb and M.B. Ruskai, see the book [40]. The presented proof is from J. Pitrik.

The Bessis, Moussa and Villani conjecture (or BMV conjecture) was published in the paper D. Bessis, P. Moussa and M. Villani: Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics, J. Math. Phys. **16**, 2318–2325 (1975). Theorem 5.5 is from E. H. Lieb and R. Seiringer: Equivalent forms of the Bessis-Moussa-Villani conjecture, J. Statist. Phys. **115**, 185–190 (2004).

The contour integral representation (5.16) was found by Henri Poincaré in 1899.

5.5 Exercises

1. Prove the Golden-Thompson inequality using the trace inequality

$$\operatorname{Tr}(CD)^n \le \operatorname{Tr} C^n D^n \qquad (n \in \mathbb{N}) \tag{5.32}$$

for $C, D \ge 0$.

2. Let A and B be self-adjoint matrices. Show that

$$|\operatorname{Tr} e^{A+\mathrm{i}B}| \le \operatorname{Tr} e^A . \tag{5.33}$$

3. Let

$$C = c_0 I + c(c_1\sigma_1 + c_2\sigma_2 + c_3\sigma_3) \quad \text{with} \quad c_1^2 + c_2^2 + c_3^2 = 1,$$

where σ_1, σ_2, σ_3 are the Pauli matrices and c_0, c_1, c_2, $c_3 \in \mathbb{R}$. Show that

$$e^C = e^{c_0}\left((\cosh c)I + (\sinh c)(c_1\sigma_1 + c_2\sigma_2 + c_3\sigma_3)\right).$$

4. Let $A, B \in \mathbb{M}_n$ and $\|A\|, \|B\| \le \lambda$. Prove that

$$\|e^A - e^B\| \le \|A - B\| e^\lambda .$$

(Hint: Show first that $\|A^m - B^m\| \le m\lambda^{m-1}\|A - B\|$.)

5. Assume that

$$e^A = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}.$$

What is the matrix A?

6. Let $A \in \mathbb{M}_3$ have eigenvalues λ, λ, μ with $\lambda \ne \mu$. Show that

$$e^{tA} = e^{\lambda t}(I + t(A - \lambda I)) + \frac{e^{\mu t} - e^{\lambda t}}{(\mu - \lambda)^2}(A - \lambda I)^2 - \frac{te^{\lambda t}}{\mu - \lambda}(A - \lambda I)^2.$$

7. Assume that $A \in \mathbb{M}_3$ has different eigenvalues λ, μ, ν. Show that

$$e^{tA} = e^{\lambda t}\frac{(A - \mu I)(A - \nu I)}{(\lambda - \mu)(\lambda - \nu)} + e^{\mu t}\frac{(A - \lambda I)(A - \nu I)}{(\mu - \lambda)(\mu - \nu)} + e^{\nu t}\frac{(A - \lambda I)(A - \mu I)}{(\nu - \lambda)(\nu - \mu)}.$$

8. Assume that the $n \times n$ matrix A is diagonalizable and let $f(t) = t^m$ with $m \in \mathbb{N}$. Show that (5.10) and (5.16) are the same matrices.

9. Prove Corollary 5.2 directly in the case $B = AX - XA$.

10. Let $0 < D \in \mathbb{M}_n$ be a fixed invertible positive matrix. Show that the inverse of the linear mapping

$$\mathbb{J}_D : \mathbb{M}_n \to \mathbb{M}_n, \qquad \mathbb{J}_D(B) = \tfrac{1}{2}(DB + BD) \tag{5.34}$$

is the mapping

$$\mathbb{J}_D^{-1}(A) = \int_0^\infty e^{-tD/2} A e^{-tD/2}\, dt \tag{5.35}$$

.

11. Let $0 < D \in \mathbb{M}_n$ be a fixed invertible positive matrix. Show that the inverse of the linear mapping

$$\mathbb{J}_D : \mathbb{M}_n \to \mathbb{M}_n, \qquad J_D(B) = \int_0^1 D^t B D^{1-t}\, dt \tag{5.36}$$

is the mapping

$$\mathbb{J}_D^{-1}(A) = \int_0^\infty (D+t)^{-1} A (D+t)^{-1}\, dt. \tag{5.37}$$

12. Prove (5.23) directly for the case $f(t) = t^n$, $n \in \mathbb{N}$.

13. Let $f : [\alpha, \beta] \to \mathbb{R}$ be a convex function. Show that

$$\operatorname{Tr} f(B) \geq \sum_i f(\operatorname{Tr} B\, p_i)\,. \tag{5.38}$$

for a pairwise orthogonal family (p_i) of minimal projections with $\sum_i p_i = I$ and for a self-ajoint matrix B with spectrum in $[\alpha, \beta]$. (Hint: Use the spectral decomposition of B.)

14. Prove Theorem 5.10 using formula (5.38). (Hint: Take the spectral decomposition of $B = \lambda B_1 + (1-\lambda)B_2$ and show $\lambda \operatorname{Tr} f(B_1) + (1-\lambda)\operatorname{Tr} f(B_2) \geq \operatorname{Tr} f(B)$.)

15. Show that

$$\frac{d^2}{dt^2} \log(A + tK)\Big|_{t=0} = -2 \int_0^\infty (A+s)^{-1} K (A+s)^{-1} K (A+s)^{-1}\, ds. \tag{5.39}$$

16. Show that

$$\begin{aligned} \partial^2 \log A(X_1, X_2) = \quad & -\int_0^\infty (A+s)^{-1} X_1 (A+s)^{-1} X_2 (A+s)^{-1}\, ds \\ & -\int_0^\infty (A+s)^{-1} X_2 (A+s)^{-1} X_1 (A+s)^{-1}\, ds \end{aligned}$$

for a positive invertible variable A.

17. Show that

$$\partial^2 A^{-1}(X_1, X_2) = A^{-1}X_1A^{-1}X_2A^{-1} + A^{-1}X_2A^{-1}X_1A^{-1}$$

for an invertible variable A.

18. Differentiate the equation

$$\sqrt{A+tB}\,\sqrt{A+tB} = A + tB$$

and show that for positive A and B

$$\frac{d}{dt}\sqrt{A+tB}\Big|_{t=0} \geq 0.$$

19. For a real number $0 < \alpha \neq 1$ the **Rényi entropy** is defined as

$$S_\alpha(D) = \frac{1}{1-\alpha}\log \operatorname{Tr} D^\alpha \tag{5.40}$$

for a positive matrix D such that $\operatorname{Tr} D = 1$. Show that $S_\alpha(D)$ is a decreasing function of α. What is the limit $\lim_{\alpha\to 1} S_\alpha(D)$? Show that $S_\alpha(D)$ is a concave functional of D for $0 < \alpha < 1$.

20. Fix a positive invertible matrix $D \in \mathbb{M}_n$ and set a linear mapping $\mathbb{M}_n \to \mathbb{M}_n$ by $\mathbb{J}_D(A) = DAD$. Consider the differential equation

$$\frac{\partial}{\partial t}D(t) = \mathbb{J}_{D(t)}T, \qquad D(0) = \rho_0, \tag{5.41}$$

where ρ_0 is positive invertible and T is self-adjoint in $\mathbb{M}_n$. Show that $D(t) = (\rho_0^{-1} - tT)^{-1}$ is the solution of the equation.

21. When $f(x) = x^k$ with $k \in \mathbb{N}$, verify that

$$f^{[n]}[x_1, x_2, \dots, x_{n+1}] = \sum_{\substack{u_1,u_2,\dots,u_{n+1}\geq 0\\ u_1+u_2+\dots+u_{n+1}=k-n}} x_1^{u_1}x_2^{u_2}\cdots x_n^{u_n}x_{n+1}^{u_{n+1}}.$$

Chapter 6

Block matrices

A block matrix means a matrix whose entries are not numbers but matrices. For example,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

when

$$A = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 3 \\ 6 \end{bmatrix}, \quad C = \begin{bmatrix} 7 & 8 \end{bmatrix}, \quad D = \begin{bmatrix} 9 \end{bmatrix}.$$

6.1 Direct sum of Hilbert spaces

If $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces, then $\mathcal{H}_1 \oplus \mathcal{H}_2$ consists of all the pairs (f_1, f_2), where $f_1 \in \mathcal{H}_1$ and $f_2 \in \mathcal{H}_2$. The linear combinations of a pair is computed entry-wise and the inner product is defined as

$$\langle (f_1, f_2), (g_1, g_2) \rangle := \langle f_1, g_1 \rangle + \langle f_2, g_2 \rangle.$$

It follows that the subspaces $\{(f_1, 0) : f_1 \in \mathcal{H}_1\}$ and $\{(0, f_2) : f_2 \in \mathcal{H}_2\}$ are orthogonal and span the direct sum $\mathcal{H}_1 \oplus \mathcal{H}_2$.

Assume that $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$, $\mathcal{K} = \mathcal{K}_1 \oplus \mathcal{K}_2$ and $A : \mathcal{H} \to \mathcal{K}$ is a linear operator. A general element of $\mathcal{H}$ has the form $(f_1, f_2) = (f_1, 0) + (0, f_2)$. We have $A(f_1, 0) = (g_1, g_2)$ and $A(0, f_2) = (g_1', g_2')$ for some $g_1, g_1' \in \mathcal{K}_1$ and $g_2, g_2' \in \mathcal{K}_2$. The linear mapping A is determined uniquely by the following 4 linear mappings:

$$A_{i1} : f_1 \mapsto g_i, \quad A_{i1} : \mathcal{H}_1 \to \mathcal{K}_i \qquad (1 \le i \le 2)$$

and

$$A_{i2} : f_2 \mapsto g_i', \quad A_{i2} : \mathcal{H}_2 \to \mathcal{K}_i \qquad (1 \le i \le 2).$$

We write A in the form

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

The advantage of this notation is the formula

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} A_{11} f_1 + A_{12} f_2 \\ A_{21} f_1 + A_{22} f_2 \end{pmatrix}.$$

(The right-hand side is $A(f_1, f_2)$ written in the form of a coloum vector.)

Assume that $e_1^i, e_2^i, \dots, e_{m(i)}^i$ is a basis in $\mathcal{H}_i$ and $f_1^j, f_2^j, \dots, f_{n(j)}^j$ is a basis in $\mathcal{K}_j$, $1 \le i, j \le 2$. The linear operators $A_{ij} : \mathcal{H}_j \to \mathcal{K}_i$ have a matrix $[A_{ij}]$ with respect to these bases. Since

$$\{(e_t^1, 0) : 1 \le t \le m(1)\} \cup \{0, e_u^2) : 1 \le u \le m(2)\}$$

is a basis in $\mathcal{H}$ and similarly

$$\{(f_t^1, 0) : 1 \le t \le n(1)\} \cup \{(0, f_u^2) : 1 \le u \le m(2)\}$$

is a basis in $\mathcal{K}$, the operator A has an $(m(1) + m(2)) \times (n(1) + n(2))$ matrix which is expressed by the $n(i) \times m(j)$ matrices $[A_{ij}]$ as

$$[A] = \begin{bmatrix} [A_{11}] & [A_{12}] \\ [A_{21}] & [A_{22}] \end{bmatrix}.$$

This is a 2×2 matrix with matrix entries and it is called **block matrix**.

The computation with block matrices is similar to that of ordinary matrices.

$$\begin{bmatrix} [A_{11}] & [A_{12}] \\ [A_{21}] & [A_{22}] \end{bmatrix}^* = \begin{bmatrix} [A_{11}]^* & [A_{21}]^* \\ [A_{12}]^* & [A_{22}]^* \end{bmatrix},$$

$$\begin{bmatrix} [A_{11}] & [A_{12}] \\ [A_{21}] & [A_{22}] \end{bmatrix} + \begin{bmatrix} [B_{11}] & [B_{12}] \\ [B_{21}] & [B_{22}] \end{bmatrix} = \begin{bmatrix} [A_{11}] + [B_{11}] & [A_{12}] + [B_{12}] \\ [A_{21}] + [B_{21}] & [A_{22}] + [B_{22}] \end{bmatrix}$$

and

$$\begin{bmatrix} [A_{11}] & [A_{12}] \\ [A_{21}] & [A_{22}] \end{bmatrix} \times \begin{bmatrix} [B_{11}] & [B_{12}] \\ [B_{21}] & [B_{22}] \end{bmatrix} = \begin{bmatrix} [A_{11}] \cdot [B_{11}] + [A_{12}] \cdot [B_{21}] & [A_{11}] \cdot [B_{12}] + [A_{12}] \cdot [B_{22}] \\ [A_{21}] \cdot [B_{11}] + [A_{22}] \cdot [B_{21}] & [A_{21}] \cdot [B_{12}] + [A_{22}] \cdot [B_{22}] \end{bmatrix}.$$

6.2 Positivity and factorization

If we do not emphasize that the entries of the block matrix are matrices, we can write

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

This matrix is self-adjoint if and only if $A = A^*$, $B^* = C$ and $D = D^*$. (These conditions include that A and D are square matrices.)

For a 2×2 matrix, it is very easy to check the positivity:

$$\begin{bmatrix} a & b \\ \bar{b} & c \end{bmatrix} \ge 0 \quad \text{if} \quad a \ge 0 \quad \text{and} \quad b\bar{b} \le ac. \tag{6.1}$$

If the entries are matrices, then the condition for positivity is similar but it is a bit more complicated. It is obvious that a diagonal block matrix

$$\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}.$$

is positive if and only if the diagonal entries A and D are positive.

Theorem 6.1 *Assume that A is invertible. The self-adjoint block matrix*

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \tag{6.2}$$

positive if and only if A is positive and

$$B^*A^{-1}B \leq C.$$

Proof: First assume that $A = I$. The positivity of

$$\begin{bmatrix} I & B \\ B^* & C \end{bmatrix}$$

is equivalent to the condition

$$\langle (f_1, f_2), \begin{bmatrix} I & B \\ B^* & C \end{bmatrix} (f_1, f_2) \rangle \geq 0$$

for every vector f_1 and f_2. Computation gives that this condition is

$$\langle f_1, f_1 \rangle + \langle f_2, Cf_2 \rangle \geq -2\mathrm{Re}\, \langle Bf_2, f_1 \rangle.$$

If we replace f_1 by $e^{\mathrm{i}\varphi} f_1$ with real φ, then the left-hand-side does not change, while the right-hand-side becomes $2|\langle Bf_2, f_1 \rangle|$ for an appropriate φ. Choosing $f_1 = Bf_2$, we obtain the condition

$$\langle f_2, Cf_2 \rangle \geq \langle f_2, B^*Bf_2 \rangle$$

for every f_2. This means that positivity implies the condition $C \geq B^*B$. The converse is also true, since the right-hand side of the equation

$$\begin{bmatrix} I & B \\ B^* & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ B^* & 0 \end{bmatrix} \begin{bmatrix} I & B \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & C - B^*B \end{bmatrix}$$

is the sum of two positive block matrices.

For a general positive invertible A, the positivity of (6.2) is equivalent to the positivity of the block matrix

$$\begin{bmatrix} A^{-1/2} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \begin{bmatrix} A^{-1/2} & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} I & A^{-1/2}B \\ B^*A^{-1/2} & C \end{bmatrix}.$$

This gives the condition $C \geq B^*A^{-1}B$. □

Theorem 6.1 has applications in different areas, see for example the Cramér-Rao inequality, Section 6.3.

Theorem 6.2 *For an invertible A, we have the so-called* **Schur factorization**

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix} \cdot \begin{bmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{bmatrix} \cdot \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}. \tag{6.3}$$

The proof is simply the computation of the product on the right-hand-side.

Since

$$\begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix}^{-1} = \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix}$$

is invertible, the positivity of the left-hand-side of (6.3) is equivalent to the positivity of the middle factor of the right-hand-side. This fact gives a second proof of Theorem 6.1.

In the Schur factorization the first factor is lower triangular, the second factor is block diagonal and the third one is upper triangular. This structure allows and easy computation of the determinant and the inverse.

Theorem 6.3 *The determinant can be computed as follows.*

$$\operatorname{Det} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \operatorname{Det} A \operatorname{Det} (D - CA^{-1}B).$$

If

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

then $D - CA^{-1}B$ is called the **Schur complement** of A in M, in notation M/A. Hence the determinant formula becomes $\operatorname{Det} M = \operatorname{Det} A \times \operatorname{Det} (M/A)$.

Theorem 6.4 *Let*

$$M = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix}$$

be a positive invertible matrix. Then

$$\sup \left\{ X \geq 0 : \begin{bmatrix} X & 0 \\ 0 & 0 \end{bmatrix} \leq \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \right\}$$

is $A - BC^{-1}B^* = M/C$.

Proof: The condition

$$\begin{bmatrix} A - X & B \\ B^* & C \end{bmatrix} \geq 0$$

is equivalent to

$$A - X \geq BC^{-1}B^*$$

and this gives the result. □

It follows from the factorization that for an invertible block matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

both A and $D - CA^{-1}B$ must be invertible. This implies that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} \times \begin{bmatrix} A^{-1} & 0 \\ 0 & (D - CA^{-1}B)^{-1} \end{bmatrix} \times \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix}$$

After multiplication on the right-hand-side, we have the following.

$$\begin{aligned}\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} &= \begin{bmatrix} A^{-1} + A^{-1}BW^{-1}CA^{-1} & -A^{-1}BW^{-1} \\ -W^{-1}CA^{-1} & W^{-1} \end{bmatrix} \\ &= \begin{bmatrix} V^{-1} & -V^{-1}BD^{-1} \\ -D^{-1}CV^{-1} & D^{-1} + D^{-1}CV^{-1}BD^{-1} \end{bmatrix} \end{aligned} \tag{6.4}$$

where $W = M/A := D - CA^{-1}B$ and $V = M/D := A - BD^{-1}C$.

Example 6.1 Let $X_1, X_2, \ldots, X_{m+k}$ be random variables with (Gaussian) joint probability distribution

$$f_M(\mathbf{z}) := \sqrt{\frac{\operatorname{Det} M}{(2\pi)^{m+k}}} \exp\Big(-\tfrac{1}{2}\langle \mathbf{z}, M\mathbf{z}\rangle\Big), \tag{6.5}$$

where $\mathbf{z} = (z_1, z_2, \ldots, z_{n+k})$ and M is a positive definite $(m+k)\times(m+k)$ matrix, see Example 4.5. We want to compute the distribution of the random variables $X_1, X_2, \ldots, X_m$.

Let

$$M = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix}$$

be written in the form of a block matrix, A is $m \times m$ and D is $k \times k$. Let $\mathbf{z} = (\mathbf{x}_1, \mathbf{x}_2)$, where $\mathbf{x}_1 \in \mathbb{R}^m$ and $\mathbf{x}_2 \in \mathbb{R}^k$. Then the marginal of the Gaussian probability distribution

$$f_M(\mathbf{x}_1, \mathbf{x}_2) = \sqrt{\frac{\operatorname{Det} M}{(2\pi)^{m+k}}} \exp\Big(-\tfrac{1}{2}\langle (\mathbf{x}_1, \mathbf{x}_2), M(\mathbf{x}_1, \mathbf{x}_2)\rangle\Big)$$

on $\mathbb{R}^m$ is the distribution

$$f_1(\mathbf{x}_1) = \sqrt{\frac{\operatorname{Det} M}{(2\pi)^m \operatorname{Det} D}} \exp\Big(-\tfrac{1}{2}\langle \mathbf{x}_1, (A - BD^{-1}B^*)\mathbf{x}_1\rangle\Big). \tag{6.6}$$

We have

$$\begin{aligned}\langle (\mathbf{x}_1, \mathbf{x}_2), M(\mathbf{x}_1, \mathbf{x}_2)\rangle &= \langle A\mathbf{x}_1 + B\mathbf{x}_2, \mathbf{x}_1\rangle + \langle B^*\mathbf{x}_1 + D\mathbf{x}_2, \mathbf{x}_2\rangle \\ &= \langle A\mathbf{x}_1, \mathbf{x}_1\rangle + \langle B\mathbf{x}_2, \mathbf{x}_1\rangle + \langle B^*\mathbf{x}_1, \mathbf{x}_2\rangle + \langle D\mathbf{x}_2, \mathbf{x}_2\rangle \\ &= \langle A\mathbf{x}_1, \mathbf{x}_1\rangle + 2\langle B^*\mathbf{x}_1, \mathbf{x}_2\rangle + \langle D\mathbf{x}_2, \mathbf{x}_2\rangle \\ &= \langle A\mathbf{x}_1, \mathbf{x}_1\rangle + \langle D(\mathbf{x}_2 + W\mathbf{x}_1), (\mathbf{x}_2 + W\mathbf{x}_1)\rangle - \langle DW\mathbf{x}_1, W\mathbf{x}_1\rangle,\end{aligned}$$

where $W = D^{-1}B^*$. We integrate on $\mathbb{R}^k$ as

$$\begin{aligned}\int \exp\Big(-\tfrac{1}{2}(\mathbf{x}_1, \mathbf{x}_2)M(\mathbf{x}_1, \mathbf{x}_2)^t\Big)\, d\mathbf{x}_2 &= \exp\Big(-\tfrac{1}{2}(\langle A\mathbf{x}_1, \mathbf{x}_1\rangle - \langle DW\mathbf{x}_1, W\mathbf{x}_1\rangle)\Big) \\ &\quad \times \int \exp\Big(-\tfrac{1}{2}\langle D(\mathbf{x}_2 + W\mathbf{x}_1), (\mathbf{x}_2 + W\mathbf{x}_1)\rangle\Big)\, d\mathbf{x}_2 \\ &= \exp\Big(-\frac{1}{2}\langle (A - BD^{-1}B^*)\mathbf{x}_1, \mathbf{x}_1\rangle\Big) \sqrt{\frac{(2\pi)^k}{\operatorname{Det} D}}\end{aligned}$$

and obtain (6.6).

This computation gives a proof of Theorem 6.3 as well. If we know that $f_1(\mathbf{x}_1)$ is Gaussian, then its quadratic matrix can be obtained from formula (6.4). The covariance of $X_1, X_2, \ldots, X_{m+k}$ is M^{-1}. Therefore, the covariance of $X_1, X_2, \ldots, X_m$ is $(A - BD^{-1}B^*)^{-1}$. It follows that the quadratic matrix is the inverse: $A - BD^{-1}B^* \equiv M/D$. □

Theorem 6.5 *Let A be a positive $n\times n$ block matrix with $k\times k$ entries. Then A is the sum of block matrices B of the form $[B]_{ij} = X_i^* X_j$ for some $k \times k$ matrices $X_1, X_2, \ldots, X_n$.*

Proof: A can be written as C^*C for some

$$C = \begin{bmatrix} C_{11} & C_{12} & \ldots & C_{1n} \\ C_{21} & C_{22} & \ldots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \ldots & C_{nn} \end{bmatrix}.$$

Let B_i be the block matrix such that its ith raw is the same as in C and all other elements are 0. Then $C = B_1 + B_2 + \ldots + B_n$ and for $t \neq i$ we have $B_t^* B_i = 0$. Therefore,

$$A = (B_1 + B_2 + \ldots + B_n)^*(B_1 + B_2 + \ldots + B_n) = B_1^* B_1 + B_2^* B_2 + \ldots + B_n^* B_n.$$

The (i, j) entry of $B_t^* B_t$ is $C_{ti}^* C_{tj}$, hence this matrix is of the required form. □

As an application of the block matrix technique, we consider the following result, called **UL-factorization**.

Theorem 6.6 *Let X be an $n \times n$ invertible positive matrix. Then there is a unique upper triangle matrix T with positive diagonal such that $X = TT^*$.*

Proof: The proof can be done by mathematical induction for n. For $n = 1$ the statement is clear. We assume that the factorization is true for $(n-1)\times(n-1)$ matrices and write X in the form

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix}, \tag{6.7}$$

where A is an (invertible) $(n-1)\times(n-1)$ matrix and C is a number. If

$$T = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix}$$

is written in a similar form, then

$$TT^* = \begin{bmatrix} T_{11}T_{11}^* + T_{12}T_{12}^* & T_{12}T_{22}^* \\ T_{22}T_{12}^* & T_{22}T_{22}^* \end{bmatrix}$$

The condition $X = TT^*$ leads to the equations

$$\begin{aligned} T_{11}T_{11}^* + T_{12}T_{12}^* &= A, \\ T_{12}T_{22}^* &= B, \\ T_{22}T_{22}^* &= C. \end{aligned}$$

If T_{22} is positive (number), then $T_{22} = \sqrt{C}$ is the unique solution, moreover

$$T_{12} = BC^{-1/2} \quad \text{and} \quad T_{11}T_{11}^* = A - BC^{-1}B^*.$$

From the positivity of (6.7), we have $A - BC^{-1}B^* \geq 0$. The induction hypothesis gives that the latter can be written in the form of $T_{11}T_{11}^*$ with an upper triangular T_{11}. Therefore T is upper triangular, too. □

6.3 Cramér-Rao inequality

The Cramér-Rao inequality is a basic result in statistics. Assume that $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$ are unknown parameters of a probability distribution $f_\theta(x)$. Some random variables ξ_1 and ξ_2 are used to estimate θ_1 and θ_2. The unbiased conditions are

$$E_\theta(\xi_i) = \int f_\theta(x)\xi_i(x)\,dx = \theta_i \qquad (i = 1, 2)$$

which means that the expectation value of ξ_i is the parameter θ_i. Differentiating the above equations with respect to θ_1 and θ_2, we have

$$\int [\partial_i f_\theta(x)]\xi_i(x)\,dx = \delta(i, j) \qquad (i, j = 1, 2).$$

The Cramér-Rao inequality is

$$C := \begin{bmatrix} E_\theta(\xi_1^2) & E_\theta(\xi_1\xi_2) \\ E_\theta(\xi_1\xi_2) & E_\theta(\xi_1^2) \end{bmatrix} \le \begin{bmatrix} \int \frac{(\partial_1 f_\theta)^2}{f_\theta}\,dx & \int \frac{(\partial_1 f_\theta)(\partial_2 f_\theta)}{f_\theta}\,dx \\ \int \frac{(\partial_1 f_\theta)(\partial_2 f_\theta)}{f_\theta}\,dx & \int \frac{(\partial_1 f_\theta)^2}{f_\theta}\,dx \end{bmatrix}^{-1} =: F^{-1},$$

where C is called the **covariance** and F is the **Fisher information matrix**. Note that in the literature the identity

$$\int \frac{(\partial_1 f_\theta)^2}{f_\theta}\,dx = \int (\partial_1 \log f_\theta)^2 f_\theta\,dx$$

appears.

The dyadic matrix

$$\begin{bmatrix} \xi_1(f_\theta)^{1/2} \\ \xi_2(f_\theta)^{1/2} \\ (\partial_1 f_\theta)(f_\theta)^{-1/2} \\ (\partial_1 f_\theta)(f_\theta)^{-1/2} \end{bmatrix} \begin{bmatrix} \xi_1(f_\theta)^{1/2} & \xi_2(f_\theta)^{1/2} & (\partial_1 f_\theta)(f_\theta)^{-1/2} & (\partial_1 f_\theta)(f_\theta)^{-1/2} \end{bmatrix}$$

is positive and the integration entrywise yields

$$\begin{bmatrix} C & I_2 \\ I_2 & F \end{bmatrix} \ge 0.$$

(Of course, it should be assumed that all integrals are finite.) According to Theorem 6.1, the positivity of the matrix gives $C \ge F^{-1}$.

In the quantum setting the probability distribution is replaced by a positive matrix of trace 1, called **density matrix**. The random variables ξ_i are replaced by selfadjoint matrices. Therefore, the role of matrices is deeper in the quantum formalism.

Let $\mathcal{M} := \{D_\theta : \theta \in G\}$ be a smooth m-dimensional **manifold** of $n \times n$ density matrices. Formally $G \subset \mathbb{R}^m$ is an open set including 0. If $\theta \in G$, then $\theta = (\theta_1, \theta_2, \dots, \theta_m)$. The **Riemannian structure** on $\mathcal{M}$ is given by the inner product

$$\gamma_D(A, B) = \operatorname{Tr} A\mathbb{J}_D^{-1}(B) \tag{6.8}$$

of the tangent vectors A and B at the foot point $D \in \mathcal{M}$, where $\mathbb{J}_D : \mathbb{M}_n \to \mathbb{M}_n$ is a positive mapping when $\mathbb{M}_n$ is regarded as a Hilbert space with the Hilbert-Schmidt inner product. (This means $\operatorname{Tr} A\mathbb{J}_D(A)^* \geq 0$.)

Assume that a collection $A = (A_1, \dots, A_m)$ of self-adjoint matrices is used to estimate the true value of θ. The expectation value of A_i with respect to the density matrix D is $\operatorname{Tr} DA_i$. A is an **unbiased estimator** if

$$\operatorname{Tr} D_\theta A_i = \theta_i \qquad (1 \leq i \leq n). \tag{6.9}$$

(In many cases an unbiased estimator $A = (A_1, \dots, A_m)$ does not exist, therefore a weaker condition is more useful.)

The (generalized) **covariance matrix** of the estimator A is a positive definite matrix C,

$$C_{ij}(D) = \operatorname{Tr}(D - (\operatorname{Tr} DA_i)I)\mathbb{J}_D(D - (\operatorname{Tr} DA_j)I).$$

The **Fisher information** matrix of the estimator A is a positive definite matrix F,

$$F_{ij}(D) = \operatorname{Tr} L_i\mathbb{J}_D(L_j), \quad \text{where} \quad L_i = \mathbb{J}_D^{-1}(\partial_i D_\theta).$$

Both C and F depend on the actual state D, when we formulate on the manifold they are parametrized by θ.

The next theorem is the quantum version of the **Cramér-Rao inequality**. The point is that the right-hand-side does not depend on the estimators.

Theorem 6.7 *Let $A = (A_1, \dots, A_m)$ be an unbiased estimator of θ. Then for the above defined matrices the inequality*

$$C(D_\theta) \geq F(D_\theta)^{-1}$$

holds.

Proof: In the proof the block-matrix method is used and we restrict ourselves for $m = 2$ for the sake of simplicity and assume that $\theta = 0$. The matrices A_1, A_2, L_1, L_2 are considered as vectors and from the inner product $\langle A, B\rangle = \operatorname{Tr} A\mathbb{J}_D(B)^*$ we have the positive matrix

$$X := \begin{bmatrix} \operatorname{Tr} A_1\mathbb{J}_D(A_1) & \operatorname{Tr} A_1\mathbb{J}_D(A_2) & \operatorname{Tr} A_1\mathbb{J}_D(L_1) & \operatorname{Tr} A_1\mathbb{J}_D(L_2) \\ \operatorname{Tr} A_2\mathbb{J}_D(A_1) & \operatorname{Tr} A_2\mathbb{J}_D(A_2) & \operatorname{Tr} A_2\mathbb{J}_D(L_1) & \operatorname{Tr} A_2\mathbb{J}_D(L_2) \\ \operatorname{Tr} L_1\mathbb{J}_D(A_1) & \operatorname{Tr} L_1\mathbb{J}_D(A_2) & \operatorname{Tr} L_1\mathbb{J}_D(L_1) & \operatorname{Tr} L_1\mathbb{J}_D(L_2) \\ \operatorname{Tr} L_2\mathbb{J}_D(A_1) & \operatorname{Tr} L_2\mathbb{J}_D(A_2) & \operatorname{Tr} L_2\mathbb{J}_D(L_1) & \operatorname{Tr} L_2\mathbb{J}_D(L_2) \end{bmatrix}.$$

From the condition (6.9), we have

$$\operatorname{Tr} A_i\mathbb{J}_D(L_i) = \frac{\partial}{\partial\theta_i}\operatorname{Tr} D_\theta A_i = 1$$

for $i = 1, 2$ and

$$\operatorname{Tr} A_i\mathbb{J}_D(L_j) = \frac{\partial}{\partial\theta_j}\operatorname{Tr} D_\theta A_i = 0$$

if $i \neq j$. Hence the matrix X has the form

$$\begin{bmatrix} C(0) & I_2 \\ I_2 & I(0) \end{bmatrix}, \tag{6.10}$$

where

$$C(0) = \begin{bmatrix} \operatorname{Tr} A_1 \mathbb{J}_D(A_1) & \operatorname{Tr} A_1 \mathbb{J}_D(A_2) \\ \operatorname{Tr} A_2 \mathbb{J}_D(A_1) & \operatorname{Tr} A_2 \mathbb{J}_D(A_2) \end{bmatrix}, \qquad F(0) = \begin{bmatrix} \operatorname{Tr} L_1 \mathbb{J}_D(L_1) & \operatorname{Tr} L_1 \mathbb{J}_D(L_2) \\ \operatorname{Tr} L_2 \mathbb{J}_D(L_1) & \operatorname{Tr} L_2 \mathbb{J}_D(L_2) \end{bmatrix}.$$

The positivity of (6.10) implies the statement of the theorem. □

If we analyze the off-diagonal part of X, then a generalization can be obtained. The **bias** of the estimator is defined as

$$b(\theta) = (b_1(\theta), b_2(\theta)) = (\operatorname{Tr} D_\theta(A_1 - \theta_1), \operatorname{Tr} D_\theta(A_2 - \theta_2)).$$

(Note that for an *unbiased estimator* we have $b(\theta) = 0$.) From the bias vector we form a **bias matrix**

$$B_{ij}(\theta) := \partial_{\theta_i} b_j(\theta) = \partial_{\theta_i} \operatorname{Tr} D_\theta A_j - \partial_{\theta_i} \theta_j$$

Then

$$\begin{bmatrix} \operatorname{Tr} A_1 \mathbb{J}_D(L_1) & \operatorname{Tr} A_1 \mathbb{J}_D(L_2) \\ \operatorname{Tr} A_2 \mathbb{J}_D(L_1) & \operatorname{Tr} A_2 \mathbb{J}_D(L_2) \end{bmatrix} = \begin{bmatrix} 1 + B_{11}(\theta) & B_{12}(\theta) \\ B_{21}(\theta) & 1 + B_{22}(\theta) \end{bmatrix} = I_2 + B$$

and the block-matrix (6.10) becomes

$$\begin{bmatrix} C(0) & I_2 + B \\ I_2 + B^* & F(0) \end{bmatrix}, \tag{6.11}$$

and we have

$$C(0) \geq (I + B(\theta)) I(0)^{-1} (I + B(\theta)^*) \tag{6.12}$$

as the generalization of the previous theorem. For a **locally unbiased estimator** at $\theta = 0$, we have $B(0) = 0$.

6.4 Notes and remarks

6.5 Exercises

1. Show that $\|(f_1, f_2)\|^2 = \|f_1\|^2 + \|f_2\|^2$.

2. Give the analogue of Theorem 6.1 when C is assumed to be invertible.

3. Let $0 \leq A \leq I$. Find the matrices B and C such that

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix}$$

is a projection.

4. Let
$$M = \begin{bmatrix} A & B \\ B & A \end{bmatrix}$$
and assume that A and B are self-adjoint. Show that M is positive if and only if $-A \leq B \leq A$.

5. Give the analogue of the factorization (6.3) when D is assumed to be invertible.

6. Show that the self-adjoint invertible matrix
$$\begin{bmatrix} A & B & C \\ B^* & D & 0 \\ C^* & 0 & E \end{bmatrix}$$
has inverse in the form
$$\begin{bmatrix} Q^{-1} & -P & -R \\ -P^* & D^{-1}(I + B^*P) & D^{-1}B^*R \\ -R^* & R^*BD^{-1} & E^{-1}(I + C^*R \end{bmatrix},$$
where
$$Q = A - BD^{-1}B^* - CE^{-1}C^*, \qquad P = Q^{-1}BD^{-1}, \qquad R = Q^{-1}CE^{-1}.$$

7. Find the determinant and the inverse of the block matrix
$$\begin{bmatrix} A & 0 \\ a & 1 \end{bmatrix}.$$

8. Let A be an invertible matrix. Show that
$$\mathrm{Det} \begin{bmatrix} A & b \\ c & d \end{bmatrix} = (c - cA^{-1}b)\mathrm{Det}\, A.$$

Chapter 7

Geometric mean

The inequality

$$\sqrt{ab} \leq \frac{a+b}{2}$$

is well-known for the geometric and arithmetic mean of positive numbers. In this chapter the geometric mean will be generalized for positive matrices.

7.1 Motivation by a Riemannian manifold

The positive definite matrices might be considered as the variance of multivariate normal distributions and the information geometry of Gaussians yields a natural Riemannian metric. Those distributions (with 0 expectation) are given by a positive definite matrix $A \in \mathbb{M}_n$ in the form

$$f_A(x) := \frac{1}{\sqrt{(2\pi)^n \det A}} \exp\,(-\langle A^{-1}x, x\rangle/2) \qquad (x \in \mathbb{C}^n). \tag{7.1}$$

The set $\mathcal{P}$ of positive definite matrices can be considered as an open subset of an Euclidian space $\mathbb{R}^{n^2}$ and they form a manifold. The tangent vectors at a footpoint $A \in \mathcal{P}$ are the self-adjoint matrices $\mathbb{M}_n^{sa}$.

A standard way to construct an information geometry is to start with an **information potential function** and to introduce the Riemannian metric by the Hessian of the potential. The information potential is the **Boltzmann entropy**

$$S(f_A) := -\int f_A(x) \log f_A(x)\, dx = C + \frac{1}{2}\mathrm{Tr}\,\log A \qquad (C \text{ is a constant}). \tag{7.2}$$

The Hessian is

$$\frac{\partial^2}{\partial s \partial t} S(f_{A+tH_1+sH_2})\Big|_{t=s=0} = \mathrm{Tr}\, A^{-1}H_1A^{-1}H_2$$

and the inner product on the tangent space at A is

$$g_A(H_1, H_2) = \mathrm{Tr}\, A^{-1}H_1A^{-1}H_2\,. \tag{7.3}$$

We note here that this geometry has many symmetries, each similarity transformation of the matrices becomes a symmetry. Namely,

$$g_{S^{-1}AS^{-1}}(S^{-1}H_1S^{-1}, S^{-1}H_2S^{-1}) = g_A(H_1, H_2). \tag{7.4}$$

A differentiable function $\gamma : [0,1] \to \mathcal{P}$ is called **curve**, its tangent vector at t is $\gamma'(t)$ and the length of the curve is

$$\int_0^1 \sqrt{g_{\gamma(t)}(\gamma'(t),\gamma'(t))}\,dt.$$

Given $A, B \in \mathcal{P}$ the curve

$$\gamma(t) = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2} \qquad (0 \le t \le 1) \tag{7.5}$$

connects these two points: $\gamma(0) = A$, $\gamma(1) = B$. This is the shortest curve connecting the two points, it is called **geodesic**.

Lemma 7.1 *The geodesic connecting $A, B \in \mathcal{P}$ is (7.5) and the geodesic distance is*

$$\delta(A,B) = \|\log(A^{-1/2}BA^{-1/2})\|_2\,,$$

where $\|\cdot\|_2$ stands for the Hilbert–Schmidt norm.

Proof: Due to the property (7.4) we may assume that $A = I$, then $\gamma(t) = B^t$. Let $\ell(t)$ be a curve such that $\ell(0) = \ell(1) = 0$. This will be used for the perturbation of the curve $\gamma(t)$ in the form $\gamma(t) + \varepsilon\ell(t)$.

We want to differentiate the length

$$\int_0^1 \sqrt{g_{\gamma(t)+\varepsilon\ell(t)}(\gamma'(t)+\varepsilon\ell'(t),\gamma'(t)+\varepsilon\ell'(t))}\,dt$$

with respect to ε at $\varepsilon = 0$. This is

$$\int_0^1 \frac{1}{2}\Big(g_{\gamma(t)}(\gamma'(t),\gamma'(t)\Big)^{-1/2} \frac{\partial}{\partial\varepsilon} g_{\gamma(t)+\varepsilon\ell(t)}(\gamma'(t)+\varepsilon\ell'(t),\gamma'(t)+\varepsilon\ell'(t))\,dt$$

$$= \frac{1}{2\sqrt{\mathrm{Tr}\,(\log B)^2}} \int_0^1 \frac{\partial}{\partial\varepsilon} \mathrm{Tr}\,(B^t+\varepsilon\ell(t))^{-1}(B^t\log B+\varepsilon\ell'(t))(B^t+\varepsilon\ell(t))^{-1}(B^t\log B+\varepsilon\ell'(t))\,dt$$

$$= \frac{1}{\sqrt{\mathrm{Tr}\,(\log B)^2}} \int_0^1 \mathrm{Tr}\,(-B^t(\log B)^2\ell(t) + B^{-t}(\log B)\ell'(t))\,dt.$$

Since we want to remove $\ell'(t)$, we integrate by part the second term:

$$\int_0^1 \mathrm{Tr}\,B^{-t}(\log B)\ell'(t)\,dt = \Big[\mathrm{Tr}\,B^{-t}(\log B)\ell(t)\Big]_0^1 + \int_0^1 \mathrm{Tr}\,B^{-t}(\log B)^2\ell(t)\,dt$$

Since $\ell(0) = \ell(1) = 0$, the first term wanishes here and the derivative at $\varepsilon = 0$ is 0 for every perturbation $\ell(t)$. On the other hand

$$g_{\gamma(t)}(\gamma'(t),\gamma'(t)) = \mathrm{Tr}\,B^{-t}B^t(\log B)B^{-t}B^t\log B = \mathrm{Tr}\,(\log B)^2$$

does not depend on t, we can conclude that $\gamma(t) = B^t$ is the geodesic curve between I and B. The distance is

$$\int_0^1 \sqrt{\mathrm{Tr}\,(\log B)^2}\,dt = \sqrt{\mathrm{Tr}\,(\log B)^2}.$$

The lemma is proved. □

The midpoint of the curve (7.5) will be called **geometric mean** of $A, B \in \mathcal{P}$.

7.2 Geometric mean of two matrices

Let $A, B \geq 0$ and assume that A is invertible. We want to study the positivity of the matrix

$$\begin{bmatrix} A & X \\ X & B \end{bmatrix}. \tag{7.6}$$

for a positive X. The positivity of the block matrix implies

$$B \geq XA^{-1}X,$$

see Theorem 6.1. From the matrix monotonicity of the square root function (Example 5.10), we obtain $(A^{-1/2}BA^{-1/2})^{1/2} \geq A^{-1/2}XA^{-1/2}$, or

$$A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2} \geq X.$$

The left-hand-side is defined to be the **geometric mean** of A and B, in notation $A\#B$. It is easy to see that for $X = A\#B$, the block matrix (7.6) is positive. Therefore, $A\#B$ is the largest positive matrix X such that (7.6) is positive. This can be the definition for non-invertible A. An equivalent possibility is

$$A\#B := \lim_{\varepsilon \to +0} (A + \varepsilon I)\#B.$$

If $AB = BA$, then $A\#B = A^{1/2}B^{1/2} (= (AB)^{1/2})$. The inequality between geometric and arithmetic means holds also for matrices, see Exercise 1.

Example 7.1 The partial ordering $\leq$ of operators has a geometric interpretation for projections. The relation $P \leq Q$ is equivalent to $\operatorname{Rng} P \subset \operatorname{Rng} Q$, that is P projects to a smaller subspace than Q. This implies that any two projections P and Q have a largest lower bound denoted by $P \wedge Q$. This operator is the orthogonal projection to the (closed) subspace $\operatorname{Rng} P \cap \operatorname{Rng} Q$.

We want to show that $P\#Q = P \wedge Q$. First we show that the block matrix

$$\begin{bmatrix} P & P \wedge Q \\ P \wedge Q & Q \end{bmatrix}$$

is positive. This is equivalent to the relation

$$\begin{bmatrix} P + \varepsilon P^\perp & P \wedge Q \\ P \wedge Q & Q \end{bmatrix} \geq 0 \tag{7.7}$$

for every constant $\varepsilon > 0$. Since

$$(P \wedge Q)(P + \varepsilon P^\perp)^{-1}(P \wedge Q) = P \wedge Q$$

is smaller than Q, the positivity (7.7) is true due to Theorem 6.1. We conclude that $P\#Q \geq P \wedge Q$.

The positivity of

$$\begin{bmatrix} P + \varepsilon P^\perp & X \\ X & Q \end{bmatrix}$$

gives the condition

$$Q \geq X(P + \varepsilon^{-1}P^{\perp})X = XPX + \varepsilon^{-1}XP^{\perp}X.$$

Since $\varepsilon > 0$ is arbitrary, $XP^{\perp}X = 0$. The latter condition gives $X = XP$. Therefore, $Q \geq X^2$. Symmetrically, $P \geq X^2$ and Corollary 4.2 tells us that $P \wedge Q \geq X^2$ and so $P \wedge Q \geq X$. □

Theorem 7.1 *Assume that A_1, A_2, A_3, A_4 are positive matrices and $A_1 \leq A_2, A_3 \leq A_4$. Then $A_1 \# A_3 \leq A_2 \# A_4$.*

Proof: The statement is equivalent to the positivity of the block matrix

$$\begin{bmatrix} A_2 & A_1\#A_3 \\ A_1\#A_3 & A_4 \end{bmatrix}.$$

This is a sum of positive matrices:

$$\begin{bmatrix} A_1 & A_1\#A_3 \\ A_1\#A_3 & A_3 \end{bmatrix} + \begin{bmatrix} A_2 - A_1 & 0 \\ 0 & A_4 - A_3 \end{bmatrix}$$

The proof is complete. □

Theorem 7.2 (Löwner theorem) *Assume that for the matrices A and B the inequalities $0 \leq A \leq B$ hold and $0 < t < 1$ is a real number. Then $A^t \leq B^t$.*

Proof: Due to the continuity, it is enough to show the case $t = k/2^n$, that is, t is a dyadic rational number. We use Theorem 7.1 to deduce from the inequalities $A \leq B$ and $I \leq I$ the inequality

$$A^{1/2} = A\#I \leq B\#I = B^{1/2}.$$

A second application of Theorem 7.1 gives similarly $A^{3/4} \leq B^{3/4}$. The procedure can be continued to cover all dyadic rational powers. Arbitrary $t \in [0, 1]$ can be the limit of dyadic numbers. □

Theorem 7.3 *The geometric mean of matrices is jointly concave, that is,*

$$\frac{A_1 + A_3}{2} \# \frac{A_2 + A_4}{2} \geq \frac{A_1\#A_2 + A_3\#A_4}{2}.$$

Proof: The block matrices

$$\begin{bmatrix} A_1 & A_1\#A_2 \\ A_1\#A_2 & A_2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A_3 & A_3\#A_4 \\ A_4\#A_3 & A_4 \end{bmatrix}$$

are positive and so is there arithmetic mean,

$$\begin{bmatrix} \frac{1}{2}(A_1 + A_3) & \frac{1}{2}(A_1\#A_2 + A_3\#A_4) \\ \frac{1}{2}(A_1\#A_2 + A_3\#A_4) & \frac{1}{2}(A_2 + A_4) \end{bmatrix}.$$

Therefore the off-diagonal entry is smaller than the geometric mean of the diagonal entries. □

The next theorem of Ando [4] is the generalization of Example 7.1. For the sake of simplicity the formulation is in block-matrices.

Theorem 7.4 *The geometric mean of the orthogonal projection*

$$P = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$$

and the positive invertible matrix

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

is

$$\begin{bmatrix} (R_{11} - R_{12}R_{22}^{-1}R_{21})^{-1/2} & 0 \\ 0 & 0 \end{bmatrix},$$

or $P\#R = (PR^{-1}P)^{-1/2}$.

Proof: We have already P and R in block-matrix form. Due to (7.6) we are looking for matrices

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$$

such that

$$\begin{bmatrix} P & X \\ X & R \end{bmatrix} = \begin{bmatrix} I & 0 & X_{11} & X_{12} \\ 0 & 0 & X_{21} & X_{22} \\ X_{11} & X_{12} & R_{11} & R_{12} \\ X_{21} & X_{22} & R_{21} & R_{22} \end{bmatrix}$$

should be positive. From the positivity $X_{12} = X_{21} = X_{22} = 0$ follows and the necessary and sufficient condition is

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \geq \begin{bmatrix} X_{11} & 0 \\ 0 & 0 \end{bmatrix} R^{-1} \begin{bmatrix} X_{11} & 0 \\ 0 & 0 \end{bmatrix},$$

or

$$I \geq X_{11}(R^{-1})_{11}X_{11}.$$

It was shown at the beginning of the section that this is equivalent to

$$X_{11} \leq \Big((R^{-1})_{11}\Big)^{-1/2}.$$

The inverse of a block-matrix is described in (6.4) and the proof is complete. □

7.3 Geometric mean for more matrices

The arithmetic mean is simpler, than the geometric mean: for (positive) matrices A and B it is

$$\mathbf{A}(A_1, A_2, \ldots, A_n) := \frac{A_1 + A_2 + \ldots + A_n}{n}.$$

Only the linear structure plays role. The arithmetic mean is a good example to show how to move from the means of two variables to three variables.

Suppose we have a device which can compute the mean of two matrices. How to compute the mean of three? Assume that we aim to obtain the mean of A, B and C. We can make a new device

$$W : (A, B, C) \mapsto (\mathbf{A}(A, B), \mathbf{A}(A, C), \mathbf{A}(B, C)) \tag{7.8}$$

which applied to (A, B, C) many times gives the mean of A, B and C:

$$W^n(A, B, C) \to \mathbf{A}(A, B, C) \quad \text{as} \quad n \to \infty. \tag{7.9}$$

Indeed, $W^n(A, B, C)$ is a convex combination of A, B and C,

$$W^n(A, B, C) = (A_n, B_n, C_n) = \lambda_1^{(n)} A + \lambda_2^{(n)} B + \lambda_3^{(n)} C.$$

One can compute the coefficients $\lambda_i^{(n)}$ explicitly and show that $\lambda_i^{(n)} \to 1/3$. The idea is shown by a picture and will be extended to the geometric mean.

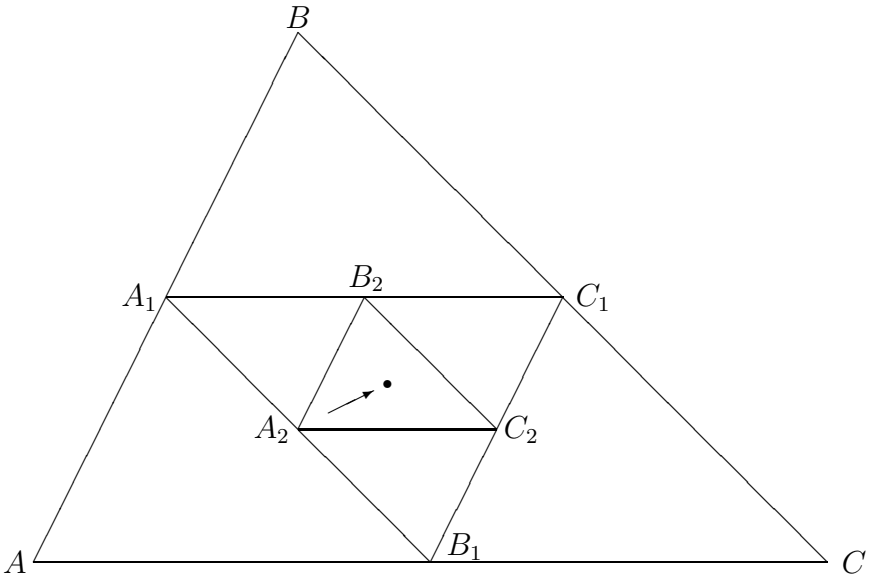


Figure 7.1: The triangles Δ_0, Δ_1 and Δ_2.

Theorem 7.5 *Let $A, B, C \in \mathbb{M}_n$ be positive definite matrices and set a recursion as*

$$A_0 = A, \qquad B_0 = B, \qquad C_0 = C,$$

$$A_{n+1} = A_n \# B_n, \qquad B_{n+1} = A_n \# C_n, \qquad C_{n+1} = B_n \# C_n.$$

Then the limits

$$\mathbf{G}_3(A, B, C) := \lim_n A_n = \lim_n B_n = \lim_n C_n \tag{7.10}$$

exist.

Proof: First we assume that $A \leq B \leq C$.

From the monotonicity property of the geometric mean, see Theorem 7.1, we obtain that $A_n \leq B_n \leq C_n$. It follows that the sequence (A_n) is increasing and (C_n) is decreasing. Therefore, the limits

$$L := \lim_{n\to\infty} A_n \qquad \text{and} \qquad U = \lim_{n\to\infty} C_n$$

exist. We claim that $L = U$.

Assume that $L \neq U$. By continuity, $B_n \to L\#U =: M$, where $L \leq M \leq U$. Since

$$B_n \# C_n = C_{n+1},$$

the limit $n \to \infty$ gives $M\#U = U$, therefore $M = U$. This contradicts to $M \neq U$.

The general case can be reduced to the case of ordered triplet. If A, B, C are arbitrary, we can find numbers λ and μ such that $A \leq \lambda B \leq \mu C$ and use the formula

$$(\alpha X)\#(\beta Y) = \sqrt{\alpha\beta}(X\#Y) \tag{7.11}$$

for positive numbers α and β.

Let

$$A'_1 = A, \qquad B'_1 = \lambda B, \qquad C'_1 = \mu C,$$

and

$$A'_{n+1} = A'_n \# B'_n, \qquad B'_{n+1} = A'_n \# C'_n, \qquad C'_{n+1} = B'_n \# C'_n.$$

It is clear that for the numbers

$$a := 1, \qquad b := \lambda \qquad \text{and} \qquad c := \mu$$

the recursion provides a convergent sequence (a_n, b_n, c_n) of triplets.

$$(\lambda\mu)^{1/3} = \lim_n a_n = \lim_n b_n = \lim_n c_n.$$

Since

$$A_n = A'_n/a_n, \qquad B_n = B'_n/b_n \qquad \text{and} \qquad C_n = C'_n/c_n$$

due to property (7.11) of the geometric mean, the limits stated in the theorem must exist and equal $\mathbf{G}(A', B', C')/(\lambda\mu)^{1/3}$. □

The geometric mean of the positive definite matrices $A, B, C \in \mathbb{M}_n$ is defined as $\mathbf{G}_3(A, B, C)$ in (7.10). Explicit formula is not known and the same procedure can be used to make definition of the geometric mean of n matrices.

7.4 Notes and remarks

The geometric mean of operators first appeared in the paper of Wieslaw Pusz and Stanislav L. Woronowicz (Functional calculus for sesquilinear forms and the purification map, Rep. Math. Phys., (1975), 159–170.) and the detailed study was in the paper Tsuyoshi Ando and Fumio Kubo [31]. The geometric mean for more matrices is from the paper [5]. A popularization of the subject is the paper Rajendra Bhatia and John

Holbrook: Noncommutative geometric means. Math. Intelligencer **28**(2006), no. 1, 32–39.

Lajos Molnár proved that if a bijection $\alpha : \mathbb{M}_n^+ \to \mathbb{M}_n^+$ preserves the geometric mean, then for $n > 2$ $\alpha(A) = SAS^{-1}$ for a linear or conjugate linear mapping S (Maps preserving the geometric mean of positive operators. Proc. Amer. Math. Soc. **137**(2009), 1763–1770.)

7.5 Exercises

1. Show that for positive invertible matrices A and B the inequalities

$$2(A^{-1} + B^{-1})^{-1} \leq A\#B \leq \frac{1}{2}(A + B)$$

hold. (Hint: Reduce the general case to $A = I$.)

2. Show that

$$A\#B = \frac{1}{\pi}\int_0^1 \frac{(tA^{-1} + (1-t)B^{-1})^{-1}}{\sqrt{t(1-t)}}\, dt.$$

3. Let $A, B > 0$. Show that $A\#B = A$ implies $A = B$.

4. Let $A \geq 0$ and P be a projection of rank 1. Show that $A\#P = \sqrt{\mathrm{Tr}\, AP}P$.

5. Argue that the natural map

$$(A, B) \longmapsto \exp\Big(\frac{\log A + \log B}{2}\Big)$$

would not be a good definition for geometric mean.

6. Let A and B be positive matrices and assume that there is a unitary U such that $A^{1/2}UB^{1/2} \geq 0$. Show that $A\#B = A^{1/2}UB^{1/2}$.

7. Let A and B be positive definite matrices. Set $A_0 := A$, $B_0 := B$ and define recurrently

$$A_n = \frac{A_{n-1} + B_{n-1}}{2} \quad \text{and} \quad B_n = 2(A_{n-1}^{-1} + B_{n-1}^{-1})^{-1} \qquad (n = 1, 2, \ldots).$$

Show that

$$\lim_{n\to\infty} A_n = \lim_{n\to\infty} B_n = A\#B.$$

8. Show that

$$\mathrm{Det}\,(A\#B) = \sqrt{\mathrm{Det}\, A\, \mathrm{Det}\, B}.$$

9. Assume that A and B are invertible positive matrices. Show that

$$(A\#B)^{-1} = A^{-1}\#B^{-1}.$$

10. Let
$$A := \begin{bmatrix} 3/2 & 0 \\ 0 & 3/4 \end{bmatrix} \quad \text{and} \quad B := \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}.$$
Show that $A \geq B \geq 0$ and for $p > 1$ the inequality $A^p \geq B^p$ does not hold.

11. Show that
$$\operatorname{Det}\Big(\mathbf{G}(A,B,C)\Big) = \Big(\operatorname{Det} A \operatorname{Det} B \operatorname{Det} C\Big)^{1/3}.$$

12. Show that
$$\mathbf{G}(\alpha A, \beta B, \gamma C) = (\alpha\beta\gamma)^{1/3}\mathbf{G}(A,B,C)$$
for positive numbers α, β, γ.

13. Show that $A_1 \geq A_2,\ B_1 \geq B_2,\ C_1 \geq C_2$ imply
$$\mathbf{G}(A_1,B_1,C_1) \geq \mathbf{G}(A_2,B_2,C_2).$$

14. Show that
$$\mathbf{G}(A,B,C) = \mathbf{G}(A^{-1},B^{-1},C^{-1})^{-1}.$$

15. Show that
$$3(A^{-1}+B^{-1}+C^{-1})^{-1} \ \leq\ \mathbf{G}(A,B,C)\ \leq\ \frac{1}{3}(A+B+C).$$

Chapter 8

Convexity

Convex sets are subsets of linear spaces. Number-valued convex functions defined on an interval are treated in analysis. Here the functionals are typically matrix-valued or they are defined on a subset of matrices. Convexity is defined by inequalities and it has many applications in different areas.

8.1 Convex sets

Let V be a vector space (over the real numbers). If $u, v \in V$, then they are called the endpoints of the line-segment

$$[u, v] := \{\lambda u + (1 - \lambda)v \,:\, \lambda \in \mathbb{R},\, 0 \leq \lambda \leq 1\}.$$

A subset $\mathcal{A} \subset V$ is **convex** if for $u, v \in \mathcal{A}$ the line-segment $[u, v]$ is contained in $\mathcal{A}$. It is easy to check that a set $\mathcal{A} \subset V$ is convex if and only if

$$\sum_{i=1}^{n} \lambda_i v_i \in \mathcal{A}.$$

for every finite subset $\{v_1, v_2, \ldots, v_n\} \subset \mathcal{A}$ and for every family of real positive numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ with sum 1, $\sum_j \lambda_i = 1$. The intersection of convex sets is a convex set.

Example 8.1 Let V be a vector space and $\|\cdot\| : V \to \mathbb{R}^+$ be a norm. Then

$$\{v \in V \,:\, \|v\| \leq 1\}$$

is a convex set. Indeed, if $\|u\|, \|v\| \leq 1$, then

$$\|\lambda u + (1 - \lambda)v\| \leq \|\lambda u\| + \|(1 - \lambda)v\| = \lambda\|u\| + (1 - \lambda)\|v\| \leq 1.$$

□

Example 8.2 In the vector space $\mathbb{M}_n$ the self-adjoint matrices and the positive matrices form a convex set. Let (a, b) be a real intervall. Then

$$\{A \in \mathbb{M}_n^{sa} \,:\, \sigma(A) \subset (a, b)\}$$

is a convex set. This follows from the fact that $\sigma(A) \subset (a, b)$ is equivalent to the property $aI_n \leq A \leq bI_n$. □

Example 8.3 Set

$$\mathcal{S}_n := \{D \in \mathbb{M}_n^{sa} \,:\, D \geq 0 \quad \text{and} \quad \operatorname{Tr} D = 1\}.$$

This is a convex set, in quantum theory it is called state space.

If $n = 2$, then a popular parametrization of the matrices in $\mathcal{S}_2$ is

$$\frac{1}{2}\begin{bmatrix} 1-\lambda_3 & \lambda_1 - \mathrm{i}\lambda_2 \\ \lambda_1 + \mathrm{i}\lambda_2 & 1-\lambda_3 \end{bmatrix} = \frac{1}{2}(I + \lambda_1\sigma_1 + \lambda_2\sigma_2 + \lambda_3\sigma_3),$$

where $\sigma_1, \sigma_2, \sigma_3$ are the Pauli matrices, see (2.11), and the necessary and sufficient condition to be in $\mathcal{S}_2$ is

$$\lambda_1^2 + \lambda_2^2 + \lambda_3^2 \leq 1.$$

This shows that the convex set $\mathcal{S}_2$ can be viewed as the unit ball in $\mathbb{R}^3$. If $n > 2$, then the geometric picture of $\mathcal{S}_n$ is not so clear. □

If $\mathcal{A}$ is a subset of the vectorspace V, then its **convex hull** is the smallest convex set containg $\mathcal{A}$, it is denoted by $\operatorname{co}\mathcal{A}$.

$$\operatorname{co}\mathcal{A} = \left\{ \sum_{i=1}^{n} \lambda_i v_i \,:\, v_i \in \mathcal{A},\, \lambda_i \geq 0,\, 1 \leq i \leq n,\, \sum_{i=1}^{n} \lambda_i = 1,\, n \in \mathbb{N} \right\}.$$

Let $\mathcal{A} \subset V$ be a convex set. The vector $v \in \mathcal{A}$ is an **extreme point** of $\mathcal{A}$ if the conditions

$$v_1, v_2 \in \mathcal{A}, \quad 0 < \lambda < 1, \quad \lambda v_1 + (1-\lambda)v_2 = v$$

imply that $v_1 = v_2 = v$.

In the convex set $\mathcal{S}_2$ the extreme points corresponds to the parameters satisfying $\lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 1$. (If $\mathcal{S}_2$ is viewed as a ball in $\mathbb{R}^3$, then the extreme points are in the boundary of the ball.)

8.2 Convex functionals

Let $J \subset \mathbb{R}$ be an interval. A function $f : J \to \mathbb{R}$ is said to be **convex** if

$$f(ta + (1-t)b) \leq tf(a) + (1-t)f(b) \tag{8.1}$$

for all $a, b \in J$ and $0 \leq t \leq 1$. This inequality is equivalent to the positivity of the **second divided difference**

$$\begin{aligned} f[a,b,c] &= \frac{f(a)}{(a-b)(a-c)} + \frac{f(b)}{(b-a)(b-c)} + \frac{f(c)}{(c-a)(c-b)} \\ &= \frac{1}{c-b}\Big(\frac{f(c)-f(a)}{c-a} - \frac{f(b)-f(a)}{b-a}\Big) \end{aligned} \tag{8.2}$$

for every different $a, b, c \in J$. If $f \in C^2(J)$, then for $x \in J$ we have

$$\lim_{a,b,c \to x} f[a,b,c] = f''(x).$$

Hence the convexity is equivalent to the positivity of the second derivative. For a convex function f the **Jensen inequality**

$$f\Big(\sum_i t_i a_i\Big) \le \sum_i t_i f(a_i) \tag{8.3}$$

holds whenever $a_i \in J$ and for real numbers $t_i \ge 0$ and $\sum_i t_i = 1$. This inequality has an integral form

$$f\left(\int g(x)\,d\mu(x)\right) \le \int f\circ g(x)\,d\mu(x). \tag{8.4}$$

For a discrete measure μ this is exactly the Jensen inequality, but it holds for any normalized (=probabilistic) measure μ and for a bounded function g.

Definition (8.1) makes sense if J is a convex subset of a vector space and f is a real functional defined on it. So let V be a finite dimensional vector space and $\mathcal{A} \subset V$ be a convex subset. The functional $F : \mathcal{A} \to \mathbb{R} \cup \{+\infty\}$ is called **convex** if

$$F(\lambda x + (1-\lambda)y) \le \lambda F(x) + (1-\lambda)F(y) \tag{8.5}$$

for every $x, y \in \mathcal{A}$ and real number $0 < \lambda < 1$. Let $[u, v] \subset \mathcal{A}$ be a line-segment and define the function

$$F_{[u,v]}(\lambda) = F(\lambda u + (1-\lambda)v)$$

on the interval $[0, 1]$. F is convex if and only if all function $F_{[u,v]} : [0, 1] \to \mathbb{R}$ are convex when $u, v \in \mathcal{A}$.

Convexity makes sense if the functional F is matrix-valued. If $F : \mathcal{A} \to M_n(\mathbb{C})^{sa}$, then (8.5) defines convexity. For a continuous function the convexity (8.5) follows from the particular case

$$F\left(\frac{x+y}{2}\right) \le \frac{F(x)+F(y)}{2}. \tag{8.6}$$

A functional F is **concave** if $-F$ is convex.

Example 8.4 We show that the functional

$$A \mapsto \log \operatorname{Tr} e^A$$

is convex on the self-adjoint matrices, cf. Example 8.5.

The statement is equivalent to the convexity of the function

$$f(t) = \log \operatorname{Tr}\left(e^{A+tB}\right) \qquad (t \in \mathbb{R}) \tag{8.7}$$

for every $A, B \in \mathbb{M}_n^{sa}$. To show this we prove that $f''(0) \ge 0$. It follows from Theorem 5.8 that

$$f'(t) = \frac{\operatorname{Tr} e^{A+tB}B}{\operatorname{Tr} e^{A+tB}}\,.$$

In the computation of the second derivative we use the identity

$$e^{A+tB} = e^A + t\int_0^1 e^{uA}Be^{(1-u)(A+tB)}\,du\,. \tag{8.8}$$

In order to write $f''(0)$ in a convenient form we introduce the inner product

$$\langle X, Y\rangle_{Bo} := \int_0^1 \operatorname{Tr} e^{tA} X^* e^{(1-t)A} Y \, dt. \tag{8.9}$$

(This is frequently termed Bogoliubov inner product.) Now

$$f''(0) = \frac{\langle I, I\rangle_{Bo}\langle B, B\rangle_{Bo} - \langle I, B\rangle_{Bo}^2}{(\operatorname{Tr} e^A)^2}$$

which is positive due to the Schwarz inequality. □

Let V be a finite dimensional vector space with dual V^*. Assume that the duality is given by a bilinear pairing $\langle \cdot, \cdot \rangle$. For a convex function $F : V \to \mathbb{R} \cup \{+\infty\}$ the **conjugate convex function** $F^* : V^* \to \mathbb{R} \cup \{+\infty\}$ is given by the formula

$$F^*(v^*) = \sup\{\langle v, v^*\rangle - F(v) \,:\, v \in V\}.$$

F^* is sometimes called the **Legendre transform** of F. F^* is the supremum of continuous linear functionals, therefore it is convex and lower semi-continuous. The following result is basic in convex analysis.

Theorem 8.1 *If $F : V \to \mathbb{R} \cup \{+\infty\}$ is a lower semi-continuous convex functional, then $F^{**} = F$.*

Example 8.5 The negative von Neumann entropy $-S(D) = -\operatorname{Tr}\eta(D) = \operatorname{Tr} D \log D$ is continuous and convex on the density matrices. Let

$$F(X) = \begin{cases} \operatorname{Tr} X \log X & \text{if } X \geq 0 \text{ and } \operatorname{Tr} X = 1, \\ +\infty & \text{otherwise.} \end{cases}$$

This is a lower semicontinuous convex functional on the linear space of all self-adjoint matrices. The duality is $\langle X, H\rangle = \operatorname{Tr} XH$. The conjugate functional is

$$\begin{aligned} F^*(H) &= \sup\{\operatorname{Tr} XH - F(X) \,:\, X \in \mathbb{M}_n^{sa}\} \\ &= -\inf\{-\operatorname{Tr} XH - S(D) \,:\, D \in \mathbb{M}_n^{sa}, D \geq 0, \operatorname{Tr} D = 1\}. \end{aligned}$$

According to Example 5.12 the minimizer is $D = e^H/\operatorname{Tr} e^H$, therefore

$$F^*(H) = \log \operatorname{Tr} e^H.$$

This is a continuous convex function of $H \in \mathbb{M}_n^{sa}$. □

Example 8.6 Fix a **density matrix** $\rho = e^H$ and consider the functional F

$$F(X) = \begin{cases} \operatorname{Tr} X(\log X - H) & \text{if } X \geq 0 \text{ and } \operatorname{Tr} X = 1 \\ +\infty & \text{otherwise.} \end{cases}$$

defined on self-adjoint matrices. F is essentially the **relative entropy** with respect to D: $S(X\|D) := \operatorname{Tr} X(\log X - \log D)$.

The duality is $\langle X, B\rangle = \operatorname{Tr} XB$ if X and B are self-adjoint matrices. We want to show that the functional $B \mapsto \log \operatorname{Tr} e^{H+B}$ is the Legendre transform or the conjugate function of F:

$$\log \operatorname{Tr} e^{B+H} = \max\{\operatorname{Tr} XB - S(X\|e^H) : X \text{ is positive, } \operatorname{Tr} X = 1\}\,. \tag{8.10}$$

Introduce the notation

$$f(X) = \operatorname{Tr} XB - S(X||e^H)$$

for a density matrix X. When $P_1, \dots, P_n$ are projections of rank one with $\sum_{i=1}^n P_i = I$, we write

$$f\Big(\sum_{i=1}^n \lambda_i P_i\Big) = \sum_{i=1}^n (\lambda_i \operatorname{Tr} P_i B + \lambda_i \operatorname{Tr} P_i H - \lambda_i \log \lambda_i)\,,$$

where $\lambda_i \geq 0$, $\sum_{i=1}^n \lambda_i = 1$. Since

$$\frac{\partial}{\partial \lambda_i} f\Big(\sum_{i=1}^n \lambda_i P_i\Big)\Big|_{\lambda_i = 0} = +\infty\,,$$

we see that $f(X)$ attains its maximum at a positive matrix X_0, $\operatorname{Tr} X_0 = 1$. Then for any self-adjoint Z, $\operatorname{Tr} Z = 0$, we have

$$0 = \frac{d}{dt} f(X_0 + tZ)\Big|_{t=0} = \operatorname{Tr}\, Z(B + H - \log X_0)\,,$$

so that $B + H - \log X_0 = cI$ with $c \in \mathbb{R}$. Therefore $X_0 = e^{B+H}/\operatorname{Tr} e^{B+H}$ and $f(X_0) = \log \operatorname{Tr}\, e^{B+H}$ by simple computation.

On the other hand, if X is positive invertible with $\operatorname{Tr} X = 1$, then

$$S(X||e^H) = \max\{\operatorname{Tr} XB - \log \operatorname{Tr} e^{H+B} : B \text{ is self-adjoint}\} \tag{8.11}$$

due to the duality theorem. □

Theorem 8.2 *Let $\alpha : \mathcal{M}_n \to \mathcal{M}_m$ be a positive unital linear mapping and $f : \mathbb{R} \to \mathbb{R}$ be a convex function. Then*

$$\operatorname{Tr} f(\alpha(A)) \leq \operatorname{Tr} \alpha(f(A))$$

for every $A \in \mathbb{M}_n^{sa}$.

Proof: Take the spectral decompositions

$$A = \sum_j \nu_j Q_j \qquad \text{and} \qquad \alpha(A) = \sum_i \mu_i P_i.$$

So we have

$$\mu_i = \operatorname{Tr}(\alpha(A)P_i)/\operatorname{Tr} P_i = \sum_j \nu_j \operatorname{Tr}(\alpha(Q_j)P_i)/\operatorname{Tr} P_i$$

whereas the convexity of f yields

$$f(\mu_i) \le \sum_j f(\nu_j) \mathrm{Tr}\,(\alpha(Q_j)P_i)/\mathrm{Tr}\,P_i\,.$$

Therefore,

$$\mathrm{Tr}\,f(\alpha(A)) = \sum_i f(\mu_i)\mathrm{Tr}\,P_i \le \sum_{i,j} f(\nu_j)\mathrm{Tr}\,(\alpha(Q_j)P_i) = \mathrm{Tr}\,\alpha(f(A))\,,$$

which was to be proven. □

It was stated in Theorem 5.10 that for a convex function $f : (\alpha, \beta) \to \mathbb{R}$, the functional $A \mapsto \mathrm{Tr}\,f(A)$ is convex. It is rather surprising that in the convexity of this functional the number coefficient $0 < t < 1$ can be replaced by a matrix.

Theorem 8.3 *Let $f : (\alpha, \beta) \to \mathbb{R}$ be a convex function and $C_i, A_i \in \mathbb{M}_n$ be such that*

$$\sigma(A_i) \subset (\alpha, \beta) \qquad \textit{and} \qquad \sum_{i=1}^k C_i C_i^* = I.$$

Then

$$\mathrm{Tr}\,f\left(\sum_{i=1}^k C_i A_i C_i^*\right) \le \sum_{i=1}^k \mathrm{Tr}\,C_i f(A_i) C_i^*.$$

Proof: We prove only the case

$$\mathrm{Tr}\,f(CAC^* + DBD^*) \le \mathrm{Tr}\,Cf(A)C^* + \mathrm{Tr}\,Df(B)D^*,$$

when $CC^* + DD^* = I$. (The more general version can be treated similarly.)

Set $F := CAC^* + DBD^*$ and consider the spectral decomposition of A and B as integrals:

$$X = \sum_i \mu_i^X P_i^X = \int \lambda dE^X(\lambda)$$

where μ_i^X are eigenvalues, P_i^X are eigenprojections and the measure E^X is defined on the Borel subsets S of $\mathbb{R}$ as

$$E^X(S) = \sum\{P_i^X : \mu_i^X \in S\},$$

$X = A, B$.

Assume that $A, B, C, D \in \mathbb{M}_n$ and for $\xi \in \mathbb{R}^n$ we define a measure μ_ξ:

$$\mu_\xi(S) = \langle (CE^A(S)C^* + DE^B(S)D^*)\xi, \xi\rangle = \langle (E^A(S)C^*\xi, C^*\xi\rangle + \langle (E^A(S)D^*\xi, D^*\xi\rangle.$$

The reason of the definition of this measure is the formula

$$\langle F\xi, \xi\rangle = \int \lambda d\mu_\xi(\lambda).$$

If ξ is a unit eigenvector of F (and $f(F)$), then

$$\langle f(CAC^* + DBD^)\xi, \xi\rangle \;\; = \;\; \langle f(F)\xi, \xi\rangle = f(\langle F\xi, \xi\rangle) = f\left(\int \lambda d\mu_\xi(\lambda)\right)$$

$$\begin{aligned} &\leq \int f(\lambda) d\mu_\xi(\lambda) \\ &= \langle (C^* f(A) C + D^* f(B) D)\xi, \xi \rangle. \end{aligned}$$

The statement follows if the inequalities for an orthonormal basis of eigenvectors of F are summarized. □

Example 8.7 The log function is concave. If $A \in \mathbb{M}_n$ is positive and the we set the projections $P_i := E(ii)$, then from the previous theorem we have

$$\mathrm{Tr}\, \log \sum_{i=1}^n P_i A P_i \geq \sum_{i=1}^n \mathrm{Tr}\, P_i (\log A) P_i.$$

This means

$$\sum_{i=1}^n \log A_{ii} \geq \mathrm{Tr}\, \log A$$

and the exponential is

$$\prod_{i=1}^n A_{ii} \geq \exp(\mathrm{Tr}\, \log A) = \mathrm{Det}\, A.$$

This is the well-known **Hadamard inequality** for the determinant. □

When $F(A, B)$ is a real valued function of two matrix variables, then F is called **jointly concave** if

$$F(\lambda A_1 + (1-\lambda)A_2, \lambda B_1 + (1-\lambda)B_2) \geq \lambda F(A_1, B_1) + (1-\lambda) F(A_2, B_2)$$

for $0 < \lambda < 1$. The function $F(A, B)$ is jointly concave if and only if the function

$$A \oplus B \mapsto F(A, B)$$

is concave. In this way the joint convexity and concavity are conveniently studied.

The geometric mean of positive matrices is jointly concave, see Theorem 7.3. A similar proof gives the following example.

Example 8.8 Assume that $X, B \in M_n(\mathbb{C})$ and B is positive invertible. Then

$$(X, B) \mapsto X^* B^{-1} X$$

is jointly convex. This means that

$$\left(\frac{X_1 + X_2}{2} \right)^* \left(\frac{B_1 + B_2}{2} \right)^{-1} \left(\frac{X_1 + X_2}{2} \right) \leq \frac{1}{2} \left(X_1^* B_1^{-1} X_1 + X_2^* B_2^{-1} X_2 \right).$$

In the proof we use the block matrix method, see Theorem 6.1. We have

$$\begin{bmatrix} B_i & X_i \\ X_i^* & X_i^* B_i^{-1} X_i \end{bmatrix} \geq 0 \qquad (i = 1, 2).$$

So is their sum:

$$\frac{1}{2} \begin{bmatrix} B_1 + B_2 & X_1 + X_2 \\ X_1^* + X_2^* & X_1^* B_1^{-1} X_1 + X_2^* B_2^{-1} X_2 \end{bmatrix} \geq 0$$

This implies our statement due to the positivity theorem. □

8.3 Matrix convex functions

Let $J \subset \mathbb{R}$ be an interval. A function $f : J \to \mathbb{R}$ is said to be **matrix convex** if

$$f(tA + (1-t)B) \leq tf(A) + (1-t)f(B) \tag{8.12}$$

for all self-adjoint matrices A and B whose spectra are in J and for all numbers $0 \leq t \leq 1$. Actually in case of a continuous function f it is enough to have the inequality for $t = 1/2$. We can say that a function f is matrix convex if the functional $A \mapsto f(A)$ is convex. f is matrix concave if $-f$ is matrix convex.

Example 8.9 The function $f(t) = t^2$ is matrix convex on the whole real line. This follows from the obvious inequality

$$\left(\frac{A+B}{2}\right)^2 \leq \frac{A^2+B^2}{2}.$$

□

Example 8.10 The function $f(x) = (x+t)^{-1}$ is matrix convex on $[0,\infty)$ when $t > 0$. It is enough to show that

$$\left(\frac{A+B}{2}\right)^{-1} \leq \frac{A^{-1}+B^{-1}}{2} \tag{8.13}$$

which is equivalent with

$$\left(\frac{B^{-1/2}AB^{-1/2}+I}{2}\right)^{-1} \leq \frac{(B^{-1/2}AB^{-1/2})^{-1}+I}{2}.$$

This holds, since

$$\left(\frac{X+I}{2}\right)^{-1} \leq \frac{X^{-1}+I}{2}$$

is true for an invertible matrix $X \geq 0$.

Note that the inequality (8.13) is equivalent to the relation of arithmetic and harmonic means. □

The classical result is about matrix convex functions on the interval $(-1,1)$. They have the integral decomposition

$$f(x) = \beta_0 + \beta_1 x + \frac{1}{2}\beta_2 \int_{-1}^{1} x^2(1-\alpha x)^{-1}\, d\mu(\alpha), \tag{8.14}$$

where μ is a probability measure and $\beta_2 \geq 0$. (In particular, f must be an analytic function.) An integral representation is used also in the next example.

Example 8.11 The function $f(x) = x^t$ is matrix concave on $[0,\infty)$ when $0 < t < 1$. The integral representation

$$x^t = \frac{\sin \pi t}{\pi} \int_0^\infty \left(1 - s(x+s)^{-1}\right) s^{t-1}\, ds$$

is used. It follows from Example 8.10 that $1 - s(x+s)^{-1}s^{t-1}$ is matrix concave, therefore the integral is matrix concave as well. □

Since self-adjoint operators may be approximated by self-adjoint matrices, (8.12) holds for operators when it holds for matrices. The point in the next theorem that in the convex combination $tA + (1-t)B$ the numbers t and $1-t$ can be replaced by matrices.

Theorem 8.4 *Let $f : (\alpha, \beta) \to \mathbb{R}$ be a matrix convex function and $C_i, A_i = A_i^* \in \mathbb{M}_n$ be such that*

$$\sigma(A_i) \subset (\alpha, \beta) \qquad \text{and} \qquad \sum_{i=1}^{k} C_i C_i^* = I.$$

Then

$$f\left(\sum_{i=1}^{k} C_i A_i C_i^*\right) \leq \sum_{i=1}^{k} C_i f(A_i) C_i^*. \tag{8.15}$$

Proof: The essential idea is in the case

$$f(CAC^* + DBD^*) \leq Cf(A)C^* + Df(B)D^*,$$

when $CC^* + DD^* = I$.

The condition $CC^* + DD^* = I$ implies that we can find a unitary block matrix

$$U := \begin{bmatrix} C & D \\ X & Y \end{bmatrix}$$

when the entries X and Y are chosen properly. Then

$$U \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} U^* = \begin{bmatrix} CAC^* + DBD^* & CAX^* + DBY^* \\ XAC^* + YBD^* & XAX^* + YBY^* \end{bmatrix}.$$

It is easy to check that

$$\frac{1}{2} V \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} V + \frac{1}{2} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$$

for

$$V = \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix}.$$

It follows that the matrix

$$Z := \frac{1}{2} VU \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} U^* V + \frac{1}{2} U \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} U^*$$

is diagonal, $Z_{11} = CAC^* + DBD^*$ and $f(Z)_{11} = f(CAC^* + DBD^*)$.

Next we use the matrix convexity of the function f:

$$\begin{aligned} f(Z) &\leq \frac{1}{2} f\left(VU \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} U^* V\right) + \frac{1}{2} f\left(U \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} U^*\right) \\ &= \frac{1}{2} VU f\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\right) U^* V + \frac{1}{2} U f\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\right) U^* \\ &= \frac{1}{2} VU \begin{bmatrix} f(A) & 0 \\ 0 & f(B) \end{bmatrix} U^* V + \frac{1}{2} U \begin{bmatrix} f(A) & 0 \\ 0 & f(B) \end{bmatrix} U^* \end{aligned}$$

The right-hand-side is diagonal with $Cf(A)C^*+Df(B)D^*$ as $(1,1)$ element. The inequality implies the inequality between the $(1,1)$ elements and this is exactly the inequality (8.15). □

In the proof of (8.15) for $n\times n$ matrices, the ordinary convexity was used for $(2n)\times(2n)$ matrices. This is an important trick. The theorem is due to Hansen and Pedersen [20].

Corollary 8.1 *Let f be an matrix convex function on an interval J such that $0 \in J$. If $\|V\| \le 1$ and $f(0) \le 0$, then*

$$f(V^*AV) \le V^*f(A)V \tag{8.16}$$

if the spectrum of $A = A^*$ lies in J.

Proof: Choose $B = 0$ and W such that $V^*V + W^*W = I$. Then

$$f(V^*AV + W^*BW) \le V^*f(A)V + W^*f(B)W$$

holds and gives our statement. □

Example 8.12 From the previous corollary we can deduce that if $f : \mathbb{R}^+ \to \mathbb{R}$ is a matrix convex function and $f(0) \le 0$, then $f(x)/x$ is matrix monotone on the interval $(0,\infty)$.

Assume that $0 < A \le B$. Then $B^{-1/2}A^{1/2} =: V$ is a contraction, since

$$\|V\|^2 = \|VV^*\| = \|B^{-1/2}AB^{-1/2}\| \le \|B^{-1/2}BB^{-1/2}\| = 1.$$

Therefore the corollary gives

$$f(A) = f(V^*BV) \le V^*f(B)V = A^{1/2}B^{-1/2}f(B)B^{-1/2}A^{1/2}$$

which is equivalent to $A^{-1}f(A) \le B^{-1}f(B)$. □

Example 8.13 Heuristically we can say that Theorem 8.4 replaces all the numbers in the the Jensen inequality $f(\sum_i t_i a_i) \le \sum_i t_i f(a_i)$ by matrices. Therefore

$$f\left(\sum_i a_i A_i\right) \le \sum_i f(a_i)A_i \tag{8.17}$$

holds for a matrix convex function f if $\sum_i A_i = I$ for the positive matrices $A_i \in \mathbb{M}_n$ and for the numbers $a_i \in (\alpha,\beta)$.

We want to show that the property (8.17) is equivalent to the matrix convexity

$$f(tA + (1-t)B) \le tf(A) + (1-t)f(B).$$

Let

$$A = \sum_i \lambda_i P_i \qquad \text{and} \qquad B = \sum_j \mu_j Q_j$$

be the spectral decompositions. Then

$$\sum_i tP_i + \sum_j (1-t)Q_j = I$$

and from (8.17) we obtain

$$\begin{aligned} f(tA+(1-t)B) &= f\left(\sum_i t\lambda_i P_i + \sum_j (1-t)\mu_j Q_j\right) \\ &\leq \sum_i f(\lambda_i) tP_i + \sum_j f(\mu_j)(1-t)Q_j \\ &= tf(A) + (1-t)f(B). \end{aligned}$$

This inequality was the aim. □

8.4 Notes and remarks

The integral representation (8.14) was obtained by Julius Bendat and Seymur Sherman [8]. Theorem 8.4 is from the paper of Frank Hansen and Gert G. Pedersen [20].

8.5 Exercises

1. Show that the extreme points of the set

$$\mathcal{S}_n := \{D \in \mathbb{M}_n^{sa} \,:\, D \geq 0 \quad \text{and} \quad \operatorname{Tr} D = 1\}.$$

   are the orthogonal projections of trace 1. Show that for $n > 2$ not all points in the boundary are extreme.

2. Let the block matrix

$$M = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix}$$

   be positive and $f : \mathbb{R}^+ \to \mathbb{R}$ be a convex function. Show that

$$\operatorname{Tr} f(M) \geq \operatorname{Tr} f(A) + \operatorname{Tr} f(C).$$

3. Show that for $A, B \in \mathbb{M}_n^{sa}$ the inequality

$$\log \operatorname{Tr} e^{A+B} \geq \log \operatorname{Tr} e^A + \frac{\operatorname{Tr} Be^A}{\operatorname{Tr} e^A}$$

   holds. (Hint: Use the function (8.7).)

4. Let the block matrix

$$M = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix}$$

   be positive and invertible. Show that

$$\det M \leq \det A \cdot \det C.$$

5. Show that for $A, B \in \mathbb{M}_n^{sa}$ the inequality

$$|\log \mathrm{Tr}\, e^{A+B} - \log \mathrm{Tr}\, e^{A}| \leq \|B\|$$

holds. (Hint: Use the function (8.7).)

6. Is it true that the function

$$\eta_\alpha(x) = \frac{x^\alpha - x}{1-\alpha} \qquad (x > 0) \tag{8.18}$$

is matrix concave if $\alpha \in (0, 2)$?

7. Let A be a positive invertible operator on a finite dimensional Hilbert space $\mathcal{H}$ and $\xi \in \mathcal{H}$. Show that

$$(\xi, A) \mapsto \langle \xi, A^{-1}\xi \rangle$$

is jointly convex.

Chapter 9

Matrix monotone functions

Let $J \subset \mathbb{R}$ be an interval. A function $f : J \to \mathbb{R}$ is said to be **matrix monotone** for $n \times n$ matrices if $f(A) \leq f(B)$ whenever A and B are self-adjoint $n \times n$ matrices, $A \leq B$ and their eigenvalues are in J. If a function is matrix monotone for every size n, then it is called matrix monotone. (One can see by an approximation argument that if a function is matrix monotone for every matrix size, then $A \leq B$ implies $f(A) \leq f(B)$ also for operators of infinite dimensional Hilbert spaces.)

9.1 Examples

Example 9.1 Let $t > 0$ be a parameter. The function $f(x) = -(t + x)^{-1}$ is matrix monotone on $[0, \infty)$.

Let A and B positive matrices of the same order. Then $A_t := tI + A$ and $B_t := tI + B$ are invertible, and

$$A_t \leq B_t \iff B_t^{-1/2} A_t B_t^{-1/2} \leq I \iff \|B_t^{-1/2} A_t B_t^{-1/2}\| \leq 1 \iff \|A_t^{1/2} B_t^{-1/2}\| \leq 1.$$

Since the adjoint preserves the operator norm, the latest condition is equivalent to $\|B_t^{-1/2} A_t^{1/2}\| \leq 1$ which implies that $B_t^{-1} \leq A_t^{-1}$. □

Example 9.2 The function $f(x) = \log x$ is matrix monotone on $(0, \infty)$.

This follows from the formula

$$\log x = \int_0^\infty \frac{1}{1+t} - \frac{1}{x+t}\, dt\,.$$

which is easy to verify. The integrand

$$f_t(x) := \frac{1}{1+t} - \frac{1}{x+t}$$

is matrix monotone according to the previous example. It follows that

$$\sum_{i=1}^n c_i f_{t(i)}(x)$$

is matrix monotone for any $t(i)$ and positive $c_i \in \mathbb{R}$. The integral is the limit of such functions, therefore it is a matrix monotone function as well.

There are several other ways to show the matrix monotonicity of the logarithm. □

Example 9.3 To show that the square root function is matrix monotone, consider the function

$$F(t) := \sqrt{A + tX}$$

defined for $t \in [0, 1]$ and for fixed positive matrices A and X. If F is increasing, then $F(0) = \sqrt{A} \leq \sqrt{A + X} = F(1)$.

In order to show that F is increasing, it is enough to see that the eigenvalues of $F'(t)$ are positive. Differentiating the equality $F(t)F(t) = A + tX$, we get

$$F'(t)F(t) + F(t)F'(t) = X.$$

As the limit of self-adjoint matrices, F' is self-adjoint and let $F'(t) = \sum_i \lambda_i E_i$ be its spectral decomposition. (Of course, both the eigenvalues and the projections depend on the value of t.) Then

$$\sum_i \lambda_i (E_i F(t) + F(t) E_i) = X$$

and after multiplication by E_j from the left and from the right, we have for the trace

$$2\lambda_j \operatorname{Tr} E_j F(t) E_j = \operatorname{Tr} E_j X E_j.$$

Since both traces are positive, λ_j must be positive as well.

Another approach is based on Theorem 7.1. Assume that $A \leq B$. Since $I \leq I$, $\sqrt{A} = A\#I \leq B\#I = \sqrt{B}$. Repeating this idea one can see that $A^t \leq B^t$ if $0 < t < 1$ is a dyadic rational number, $k/2^n$. □

The previous example contained an important idea. To decide about the matrix monotonicity of a function f, one has to investigate the derivative of $f(A + tX)$.

9.2 Characterization and properties

Theorem 9.1 *A smooth function $f : (a, b) \to \mathbb{R}$ is matrix monotone for $n \times n$ matrices if and only if the divided difference matrix $D \in \mathbb{M}_n$ defined as*

$$D_{ij} = \begin{cases} \dfrac{f(t_i) - f(t_j)}{t_i - t_j} & \text{if} \quad t_i - t_j \neq 0, \\ f'(t_i) & \text{if} \quad t_i - t_j = 0, \end{cases} \tag{9.1}$$

is positive semi-definite for $t_1, t_2, \ldots, t_n \in (a, b)$.

Proof: Let A be a self-adjoint and B be a positive semi-definite $n \times n$ matrix. When f is matrix monotone, the function $t \mapsto f(A + tB)$ is increasing function of the real variable t. Therefore, the derivative, which is a matrix, must be positive semi-definite. To compute the derivative, we use formula (5.16) of Theorem 5.9. The Schur theorem implies that the derivative is positive if the divided difference matrix is positive.

To show the converse, take a matrix B such that all entries are 1. Then positivity of the derivative $D \circ B = D$ is the positivity of D. □

The assumption about the smooth property in the previous theorem is not essential. At the beginning of the theory Löwner proved that if the function $f : (a, b) \to \mathbb{R}$ has the property that $A \leq B$ for $A, B \in \mathbb{M}_2$ implies $f(A) \leq f(B)$, then f must be a C^1 function.

The previous theorem can be reformulated in terms of a positive definite kernel. The divided difference

$$\psi(x,y) = \begin{cases} \dfrac{f(x)-f(y)}{x-y} & \text{if} \quad x \neq y, \\ \\ f'(x) & \text{if} \quad x = y \end{cases}$$

is a $(a, b) \times (a, b) \to \mathbb{R}$ kernel function. f is matrix monotone if and only if ψ is a positive definite kernel.

Example 9.4 The function $f(x) := \exp x$ is not matrix monotone, since the divided difference matrix

$$\begin{pmatrix} \exp x & \dfrac{\exp x - \exp y}{x-y} \\ \dfrac{\exp y - \exp x}{y-x} & \exp y \end{pmatrix}$$

does not have positive determinant (for $x = 0$ and for large y). □

Theorem 9.2 *(***Löwner theorem***) Matrix monotone functions on $\mathbb{R}^+$ have a special integral representation*

$$f(x) = f(0) + \beta x + \int_0^\infty \frac{\lambda x}{\lambda + x}\, d\mu(\lambda)\,, \tag{9.2}$$

where μ is a measure such that

$$\int_0^\infty \frac{\lambda}{\lambda + 1}\, d\mu(\lambda)$$

is finite and $\beta \geq 0$.

Since the integrand

$$\frac{\lambda x}{\lambda + x} = \lambda - \frac{\lambda^2}{\lambda + x}$$

is a matrix monotone function of x, see Example 9.1, one part of the Löwner theorem is straightforward. It follows from the theorem that a matrix monotone function is matrix concave.

Theorem 9.3 *If $f : \mathbb{R}^+ \to \mathbb{R}$ is matrix monotone, then $xf(x)$ is matrix convex.*

Proof: Let $\lambda > 0$. First we check the function $f(x) = -(x + \lambda)^{-1}$. Then

$$xf(x) = -\frac{x}{\lambda + x} = -1 + \frac{\lambda}{\lambda + x}$$

and it is well-known that $x \mapsto (x + \lambda)^{-1}$ is matrix convex.

For a general matrix monotone f, we use the integral decomposition (9.2) and the statement follows from the previous special case. □

Theorem 9.4 *If $f : (0,\infty) \to (0,\infty)$, then the following conditions are equivalent:*

(1) f is matrix monotone;

(2) $x/f(x)$ is matrix monotone;

(3) f is matrix concave.

Proof: For $\varepsilon > 0$ the function$f_\varepsilon(x) := f(x+\varepsilon)$ is defined on $[0,\infty)$. If the statement is proved for this function, then the limit $\varepsilon \to 0$ gives the result. So we assume $f : [0,\infty) \to (0,\infty)$.

Recall that (1) $\Rightarrow$ (3) was already remarked above.

The implication (3) $\Rightarrow$ (2) is based on Example 8.12. It says that $-f(x)/x$ is matrix monotone. Therefore $x/f(x)$ is matrix monotone as well.

(2) $\Rightarrow$ (1): $x/f(x)$ is matrix monotone on $[0,\infty)$, then it follows from the Löwvner representation that divided by x we have

$$\frac{\beta}{x} + \int_0^\infty \frac{\lambda}{\lambda + x}\, d\mu(\lambda)\,,$$

This multiplied with -1 is the matrix monotone $-1/f(x)$. Therefore $f(x)$ is matrix monotone as well. □

It was proved that the matrix monotonicity is equivalent to the positive definiteness of the divided difference kernel. Concavity has somewhat similar property.

Theorem 9.5 *Let $f : [0,\infty) \to [0,\infty)$ be a smooth function. If the divided difference kernel function is conditionally negative definite, then f is matrix convex.*

Proof: The Example 4.14 and Theorem 9.1 give that $g(x) = x^2/f(x)$ is matrix monotone. Then $x/g(x) = f(x)/x$ matrix monotone due to Theorem 9.4. Multiplying by x we get a matrix convex function, Theorem 9.3. □

It is not always easy to decide if a function is matrix monotone. An efficient method is based on holomorphic extension. The set $\mathbb{C}_+ := \{a + \mathrm{i}b : a, b \in \mathbb{R} \quad \text{and} \quad b > 0\}$ is called upper half-plain. A function $\mathbb{R}^+ \to \mathbb{R}$ is matrix monotone if and only if it has a holomorphic extension to the upper half-plain such that its range is in the closure of $\mathbb{C}_+$ [11]. (Such functions are studied in the next section.) It is surprising that a matrix monotone function is very smooth and connected with functions of a complex variable.

Example 9.5 The representation

$$x^t = \frac{\sin \pi t}{\pi} \int_0^\infty \frac{\lambda^{t-1} x}{\lambda + x}\, d\lambda \tag{9.3}$$

shows that $f(x) = x^t$ is matrix monotone when $0 < t < 1$. In other words,

$$0 \le A \le B \quad \text{imply} \quad A^t \le B^t,$$

which is often called **Löwner-Heinz inequality**.

We can arrive at the same conclusion by holomorphic extension. If

$$a + \mathrm{i}b = Re^{\varphi \mathrm{i}} \quad \text{with} \quad 0 \le \varphi \le \pi,$$

then $a + \mathrm{i}b \mapsto R^t\, e^{t\varphi \mathrm{i}}$ is holomorphic and it maps $\mathbb{C}_+$ into itself when $0 \le t \le 1$. This shows that $f(x) = x^t$ is matrix monotone for these values of the parameter but not for any other value. □

Example 9.6 Let

$$f_p(x) = p(1-p)\frac{(x-1)^2}{(x^p-1)(x^{1-p}-1)}\,. \tag{9.4}$$

The integral representation

$$\frac{1}{f_p(x)} = \frac{\sin p\pi}{\pi}\int_0^\infty d\lambda\, \lambda^{p-1}\int_0^1 ds \int_0^1 dt \frac{1}{x((1-t)\lambda + (1-s)) + (t\lambda + s)}, \tag{9.5}$$

shows that $1/f_p$ is matrix monotone decreasing when $0 < p < 1$, since so is the integrand as a function of all variables. It follows that $f_p(x)$ is matrix monotone. □

Theorem 9.6 *A function $f : [0,\infty) \to [0,\infty)$ is matrix monotone decreasing if and only if*

$$f(x) = \alpha + \int_0^\infty \frac{1}{\lambda + x}\, d\mu(\lambda)$$

with a constant $\alpha \ge 0$ and a Borel measure μ such that the integrals

$$\int_0^\infty \frac{1}{(1+\lambda)^2}\, d\mu(\lambda) \text{ and } \int_0^\infty \frac{\lambda}{(1+\lambda)^2}\, d\mu(\lambda)$$

are finite.

9.3 Matrix means

The means of numbers is a popular subject. If we move from 1×1 matrices to $n \times n$ matrices, then the arithmetic mean does not require any theory. Historically the harmonic mean was the first essential subject for matrices.

Example 9.7 It is well-known in electricity that if two resistors with resistance a and b are connected parallelly, the the total resistance q is the solution of the equation

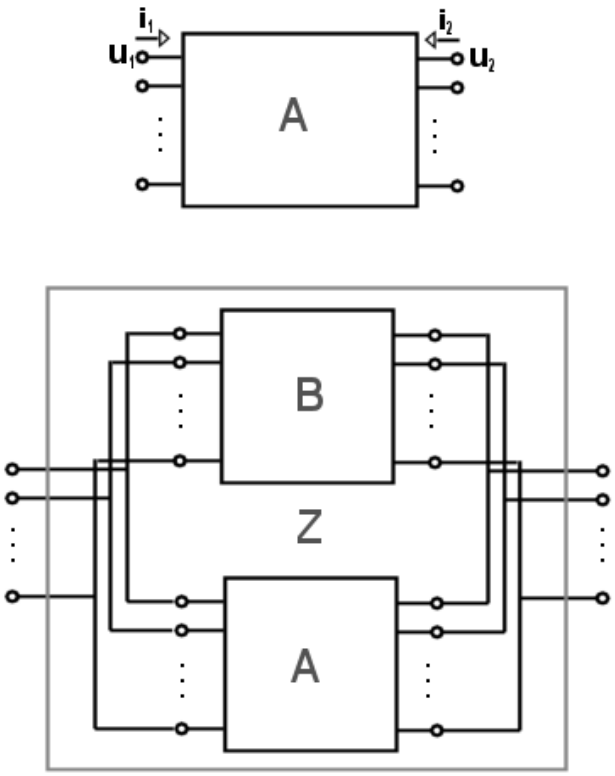
$$\frac{1}{g} = \frac{1}{a} + \frac{1}{b}.$$

Then

$$q = (a^{-1} + b^{-1})^{-1} = \frac{ab}{a+b}$$

is the harmonic mean up to a factor 2. More generally, one can consider n-point network, where the voltage and current vectors are connected by a positive matrix. The **parallel sum**

$$A : B = (A^{-1} + B^{-1})^{-1}$$



Upper part: An n-point network with the input and output voltage vectors. Below: Two parallelly connected networks

of two positive definite matrices represents the combined resistance of two n-port networks connected in parallel.

One can check that

$$A : B = A - A(A + B)^{-1}A.$$

Therefore $A : B$ is the **Schur complement** of $A + B$ in the block matrix

$$\begin{bmatrix} A & A \\ A & A + B \end{bmatrix},$$

see Theorem 6.4. □

On the basis of the previous example, the **harmonic mean** of the positive definite matrices A and B is defined as

$$H(A, B) := 2(A^{-1} + B^{-1})^{-1}. \tag{9.6}$$

Matrix monotone functions on $\mathbb{R}^+$ may be used to define **positive matrix means**. A theory of means of positive matrices was developed by Kubo and Ando [31]. Their theory has many interesting applications. Here we do not go into the details concerning matrix means but we do confine ourselves to the essentials. Matrix means are binary operations on matrices, they satisfy the following conditions.

(1) $M(A,A) = A$ for every A,

(2) $M(A,B) = M(B,A)$ for every A and B,

(3) if $A \le B$, then $A \le M(A,B) \le B$,

(4) if $A \le A'$ and $B \le B'$, then $M(A,B) \le M(A',B')$,

(5) M is continuous.

Note that the above conditions are not independent, (1) and (4) imply (3).

An important further requirement is the **transformer inequality**:

(6) $C\,M(A,B)\,C^* \le M(CAC^*, CBC^*)$

for all not necessary self-adjoint operator C.

The key issue of the theory is that operator means are in a 1-to-1 correspondence with operator monotone functions satisfying conditions $f(1) = 1$ and $tf(t^{-1}) = f(t)$. Given an operator monotone function f, the corresponding mean is

$$M_f(A,B) = A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2} \tag{9.7}$$

when A is invertible. (When A is not invertible, take a sequence A_n of invertible operators approximating A and let $M_f(A,B) = \lim_n M_f(A_n,B)$.) It follows from the definition (9.7) of means that

$$\text{if } f \le g, \text{ then } M_f(A,B) \le M_g(A,B). \tag{9.8}$$

An important example is the **geometric mean**

$$A\#B = A^{1/2}(A^{-1/2} B A^{-1/2})^{1/2} A^{1/2} \tag{9.9}$$

which corresponds to $f(x) = \sqrt{x}$. The geometric mean $A\#B$ is the unique positive solution of the equation

$$XA^{-1}X = B \tag{9.10}$$

and therefore $(A\#B)^{-1} = A^{-1}\#B^{-1}$.

The matrix monotone function $f : \mathbb{R}^+ \to \mathbb{R}^+$ will be called **standard** if $f(1) = 1$ and $tf(t^{-1}) = f(t)$. Standard functions are used to define matrix means in (9.7). For the harmonic mean (9.6) the function is

$$f(x) = \frac{2x}{x+1} = 2 - \frac{2}{x+1}$$

which is matrix monotone and standard.

Example 9.8 The function

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

is matrix monotone due to the formula

$$\int_0^1 x^t \, dt = \frac{x-1}{\log x}.$$

The standard property is obvious. The matrix mean induced by the function $f(x)$ is called **logarithmic mean**. The logarithmic mean of the positive operators A and B is denoted by $L(A, B)$.

From the inequality

$$\frac{x-1}{\log x} = \int_0^1 x^t \, dt = \int_0^{1/2} (x^t + x^{1-t}) \, dt \geq \int_0^{1/2} 2\sqrt{x} \, dt = \sqrt{x}$$

of the real functions we have the matrix inequality

$$A\#B \leq L(A, B).$$

It can be proved similarly that $L(A, B) \leq (A + B)/2$. □

Example 9.9 Let $A, B \in \mathbb{M}_n$ be positive definite matrices and M be a matrix mean. The block-matrix

$$\begin{bmatrix} A & M(A,B) \\ M(A,B) & B \end{bmatrix}$$

is positive if and only if $M(A, B) \leq A\#B$. Similarly,

$$\begin{bmatrix} A^{-1} & M(A,B)^{-1} \\ M(A,B^{-1}) & B^{-1} \end{bmatrix} \geq 0.$$

if and only $M(A, B) \geq A\#B$.

If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are positive numbers, then the matrix $A \in \mathbb{M}_n$ defined as

$$A_{ij} = \frac{1}{L(\lambda_i, \lambda_j)}$$

is positive for $n = 2$ according to the above argument. However, this is true for every n due to the formula

$$\frac{1}{L(x,y)} = \int_0^1 \frac{1}{(x+t)(y+t)} \, dt. \tag{9.11}$$

From the harmonic mean we obtain the mean matrix

$$B_{ij} = \frac{2\lambda_i \lambda_j}{\lambda_i + \lambda_j}.$$

This is positive, since the Hadamard product of two positive matrices, one of them is the Cauchy matrix.

There are many examples of positive mean matrices, but the precise condition is not known. □

Example 9.10 The **Heinz mean**

$$H_t(x,y) = \frac{x^t y^{1-t} + x^{1-t} y^t}{2} \qquad (0 \leq t \leq 1) \tag{9.12}$$

approximates between the geometric and arithmetic means. The corresponding standard function

$$f_t(x) = \frac{x^t + x^{1-t}}{2}$$

is obviously matrix monotone. Therefore we can regard it as a matrix mean. □

Theorem 9.7 *If $f : \mathbb{R}^+ \to \mathbb{R}^+$ is a standard matrix monotone function, then*

$$\frac{2x}{x+1} \le f(x) \le \frac{x+1}{2}.$$

It follows from this theorem that the harmonic mean is the smallest and the arithmetic mean is the largest mean for matrices.

Theorem 9.8 *Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a standard matrix monotone function. Then f admits a canonical representation*

$$f(t) = \beta \frac{1+t}{\sqrt{2}} \exp \int_0^1 \frac{\lambda^2 - 1}{\lambda^2 + 1} \cdot \frac{1+t^2}{(\lambda+t)(1+\lambda t)} h(\lambda)\, d\lambda \tag{9.13}$$

where $h : [0,1] \to [0,1]$ is a measurable function and the real constant β satisfies $f(1) = 1$.

Example 9.11 Since the integrand is negative $h(\lambda) = 1$ gives the smallest function. From the integral

$$\int_0^1 \frac{\lambda^2 - 1}{\lambda^2 + 1} \cdot \frac{1+t^2}{(\lambda+t)(1+\lambda t)}\, d\lambda = \log \frac{2t}{(1+t)^2}$$

we can move to constant case $h(\lambda) = \gamma \in [0,1]$ and we have

$$f_\gamma(x) = \beta \frac{1+t}{\sqrt{2}} \left(\frac{2t}{(1+t)^2} \right)^\gamma = 2^{2\gamma-1} t^\gamma (1+t)^{1-2\gamma}.$$

This is a kind of interpolation between the arithmetic mean and the harmonic mean. $\gamma = 1/2$ gives the geometric mean and

$$m_{1/3}(x,y) = \left(\frac{xy^2 + x^2 y}{2} \right)^{1/3}$$

is between the arithmetic and geometric means. □

The integral representation (9.5) is not optimal. **Hansen's canonical representation** is true for any standard matrix monotone function [21].

Theorem 9.9 *If $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a standard matrix monotone function, then*

$$\frac{1}{f(t)} = \int_0^1 \frac{1+\lambda}{2} \left(\frac{1}{t+\lambda} + \frac{1}{1+t\lambda} \right) d\mu(\lambda), \tag{9.14}$$

where μ is a probability measure on $[0,1]$.

Theorem 9.10 *Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a standard matrix monotone function. Then*

$$\tilde{f}(x) := \frac{1}{2} \left((x+1) - (x-1)^2 \frac{f(0)}{f(x)} \right) \tag{9.15}$$

is standard matrix monotone as well.

9.4 Quasi-entropies

In the mathematical formalism of quantum mechanics, instead of n-tuples of numbers one works with $n \times n$ complex matrices. They form an algebra and this allows an algebraic approach. In this approach, a probability density is replaced by a positive matrix of trace 1 which is called **density matrix**. The eigenvalues of a density matrix give a probability density.

For positive definite matrices $D_1, D_2 \in \mathbb{M}_n$, for $A \in \mathbb{M}_n$ and a function $f : \mathbb{R}^+ \to \mathbb{R}$, the **quasi-entropy** is defined as

$$\begin{aligned} S_f^A(D_1\|D_2) &:= \langle AD_2^{1/2}, f(\Delta(D_1/D_2))(AD_2^{1/2})\rangle \\ &= \operatorname{Tr} D_2^{1/2}A^* f(\Delta(D_1/D_2))(AD_2^{1/2}), \end{aligned} \tag{9.16}$$

where $\langle B, C\rangle := \operatorname{Tr} B^*C$ is the so-called **Hilbert-Schmidt inner product** and $\Delta(D_1/D_2) : \mathbb{M}_n \to \mathbb{M}_n$ is a linear mapping acting on matrices:

$$\Delta(D_1/D_2)A = D_1AD_2^{-1}.$$

This concept was introduced by Petz in [43, 45].

If we set

$$\mathbf{L}_D(X) = DX\,, \quad \mathbb{R}_D(X) = XD \quad \text{and} \quad \mathbb{J}^f_{D_1,D_2} = f(\mathbf{L}_{D_1}\mathbb{R}_{D_2}^{-1})\mathbb{R}_{D_2}\,, \tag{9.17}$$

then the quasi-entropy has the form

$$S_f^A(D_1\|D_2) = \langle A, \mathbb{J}^f_{D_1,D_2}A\rangle \tag{9.18}$$

It is clear from the definition that

$$S_f^A(\lambda D_1\|\lambda D_2) = \lambda S_f^A(D_1\|D_2)$$

for positive number λ.

Let $\alpha : \mathbb{M}_n \to \mathbb{M}_m$ be a mapping between two matrix algebras. The dual $\alpha^* : \mathbb{M}_m \to \mathbb{M}_n$ with respect to the Hilbert-Schmidt inner product is positive if and only if α is positive. Moreover, α is unital if and only if α^* is trace preserving. $\alpha : \mathbb{M}_n \to \mathbb{M}_m$ is called a **Schwarz mapping** if

$$\alpha(B^*B) \geq \alpha(B^*)\alpha(B) \tag{9.19}$$

for every $B \in \mathbb{M}_n$.

The quasi-entropies are monotone and jointly convex.

Theorem 9.1 *Assume that $f : \mathbb{R}^+ \to \mathbb{R}$ is a matrix monotone function with $f(0) \geq 0$ and $\alpha : \mathbb{M}_n \to \mathbb{M}_m$ is a unital Schwarz mapping. Then*

$$S_f^A(\alpha^*(D_1)\|\alpha^*(D_2)) \geq S_f^{\alpha(A)}(D_1\|D_2) \tag{9.20}$$

holds for $A \in \mathbb{M}_n$ and for invertible density matrices D_1 and D_2 from the matrix algebra $\mathbb{M}_m$.

Proof: The proof is based on inequalities for matrix monotone and matrix concave functions. First note that

$$S_{f+c}^A(\alpha^*(D_1)\|\alpha^*(D_2)) = S_f^A(\alpha^*(D_1)\|\alpha^*(D_2)) + c\,\mathrm{Tr}\,D_1\alpha(A^*A))$$

and

$$S_{f+c}^{\alpha(A)}(D_1\|D_2) = S_f^{\alpha(A)}(D_1\|D_2) + c\,\mathrm{Tr}\,D_1(\alpha(A)^*\alpha(A))$$

for a positive constant c. Due to the Schwarz inequality (9.19), we may assume that $f(0) = 0$.

Let $\Delta := \Delta(D_1/D_2)$ and $\Delta_0 := \Delta(\alpha^*(D_1)/\alpha^*(D_2))$. The operator

$$VX\alpha^*(D_2)^{1/2} = \alpha(X)D_2^{1/2} \qquad (X \in \mathcal{M}_0) \tag{9.21}$$

is a contraction:

$$\begin{aligned} \|\alpha(X)D_2^{1/2}\|^2 &= \mathrm{Tr}\,D_2(\alpha(X)^*\alpha(X)) \\ &\le \mathrm{Tr}\,D_2(\alpha(X^*X) = \mathrm{Tr}\,\alpha^*(D_2)X^*X = \|X\alpha^*(D_2)^{1/2}\|^2 \end{aligned}$$

since the Schwarz inequality is applicable to α. A similar simple computation gives that

$$V^*\Delta V \le \Delta_0\,. \tag{9.22}$$

Since f is matrix monotone, we have $f(\Delta_0) \ge f(V^*\Delta V)$. Recall that f is matrix concave, therefore $f(V^*\Delta V) \ge V^*f(\Delta)V$ and we conclude

$$f(\Delta_0) \ge V^*f(\Delta)V\,. \tag{9.23}$$

Application to the vector $A\alpha^*(D_2)^{1/2}$ gives the statement. □

It is remarkable that for a multiplicative α we do not need the condition $f(0) \ge 0$. Moreover, $V^*\Delta V = \Delta_0$ and we do not need the matrix monotonicity of the function f. In this case the only required condition is the matrix concavity of f.

Now we apply the monotonicity (9.20) to an embedding $\alpha(X) = X \oplus X$ of $\mathbb{M}_n$ into $\mathbb{M}_n \oplus \mathbb{M}_n$. Then $\alpha^*(X_1 \oplus X_2) = X_1 + X_2$. If we take the densities $D_1 = \lambda E_1 \oplus (1-\lambda)F_1$ and $D_2 = \lambda E_2 \oplus (1-\lambda)F_2$, then we obtain the joint concavity of the quasi-entropy:

Corollary 9.1 *If $f : \mathbb{R}^+ \to \mathbb{R}$ is matrix concave, then*

$$\lambda S_f^A(E_1\|E_2) + (1-\lambda)S_f^A(F_1\|F_2) \le S_f^A(\lambda E_1 + (1-\lambda)E_2)\|\lambda F_1 + (1-\lambda)F_2)$$

holds.

Example 9.12 The concept of quasi-entropy includes some important special cases. If $f(t) = t^\alpha$, then

$$S_f^A(D_1\|D_2) = \mathrm{Tr}\,A^*D_1^\alpha AD_2^{1-\alpha}.$$

If $0 < \alpha < 1$, then f is matrix monotone. The joint concavity in (D_1, D_2) is the famous **Lieb's concavity theorem** [35].

If D_2 and D_1 are different and $A = I$, then we have a kind of relative entropy. For $f(x) = x\log x$ we have Umegaki's relative entropy $S(D_1\|D_2) = \mathrm{Tr}\,D_1(\log D_1 - \log D_2)$.

(If we want a matrix monotone function, then we can take $f(x) = \log x$ and then we get $S(D_2\|D_1)$.) Umegaki's relative entropy is the most important example, therefore the function f will be chosen to be matrix convex. This makes the probabilistic and non-commutative situation compatible as one can see in the next argument.

Let

$$f_\alpha(x) = \frac{1}{\alpha(1-\alpha)}(1 - x^\alpha).$$

This function is matrix monotone decreasing for $\alpha \in (-1, 1)$. (For $\alpha = 0$, the limit is taken and it is $-\log x$.) Then the **relative entropies of degree** α are produced:

$$S_\alpha(D_2\|D_1) := \frac{1}{\alpha(1-\alpha)}\mathrm{Tr}\,(I - D_1^\alpha D_2^{-\alpha})D_2.$$

These quantities are essential in the quantum case. □

9.5 Notes and remarks

The matrix monotonicity of the function (9.4) was recognized in [48], a proof for $p \in [-1, 2]$ is in the paper V.E. Sándor Szabó, A class of matrix monotone functions, Linear Alg. Appl. **420**(2007), 79–85.

The matrix means were defined by Fumio Kubo and Tsuyoshi Ando in [31]. Theorem 9.7 is also fro here. Different matrix means are in the book of Fumio Hiai and Hideki Kosaki [23]. there are several examples of positive mean matrices in the paper Rajendra Bhatia and Hideki Kosaki, Mean matrices and infinite divisibility, Linear Algebra Appl. **424**(2007), 36–54. (Actually the positivity of matrices $A_{ij} = m(\lambda_i, \lambda_j)^t$ are considered, $t > 0$.)

Theorem 9.8 is from the paper F. Hansen, Characterization of symmetric monotone metrics on the state space of quantum systems, Quantum Information and Computation **6**(2006), 597–605.

Quasi-entropy was introduced by Dénes Petz in 1985, see the papers D. Petz, Quasi-entropies for finite quantum systems, Rep. Math. Phys., **23**(1986), 57-65 and D. Petz, From f-divergence to quantum quasi-entropies and their use, Entropy **12**(2010), 304-325. This concept is the quantum generalization of the f-entropy of Imre Csiszár which is used in classical information theory (and statistics) [16], see also the paper F. Liese and I. Vajda, On divergences and informations in statistics and information theory, IEEE Trans. Inform. Theory **52**(2006), 4394-4412.

9.6 Exercises

1. Prove that the function $\kappa : \mathbb{R}^+ \to \mathbb{R}$, $\kappa(x) = -x\log x + (x+1)\log(x+1)$ is matrix monotone.

2. Let f be a differentiable function on the interval (a, b) such that for some $a < c < b$ the function f is matrix monotone for 2×2 matrices on the intervals $(a, c]$ and $[c, b)$. Show that f is matrix monotone for 2×2 matrices on (a,b).

3. Show that the function

$$f(x) = \frac{ax+b}{cx+d} \qquad (a,b,c,d \in \mathbb{R}, \quad ad\rangle bc)$$

is matrix monotone on any interval which does not contain $-d/c$.

4. Show that the canonical representing measure in (9.14) for the standard matrix monotone function $f(x) = (x-1)/\log x$ is the measure

$$d\mu(\lambda) = \frac{2}{(1+\lambda)^2} d\lambda\,.$$

5. The function

$$\log_\alpha(x) = \frac{x^{1-\alpha}-1}{1-\alpha} \qquad (x>0, \quad \alpha>0, \quad \alpha \neq 1) \tag{9.24}$$

is called α-logarithmic function. Is it matrix monotone?

6. Give an example of a matrix convex function such that the derivative is not matrix monotone.

7. Show that $f(z) = \tan z := \sin z/\cos z$ is in $\mathcal{P}$, where $\cos z := (e^{iz}+e^{-iz})/2$ and $\sin z := (e^{iz}-e^{-iz})/2i$.

8. Show that $f(z) = -1/z$ is in $\mathcal{P}$.

9. Show that for positive matrices $A : B = A - A(A+B)^{-1}A$.

10. Show that for positive matrices $A : B \leq A$.

11. Show that $0 < A \leq B$ imply $A \leq 2(A:B) \leq B$.

12. Show that $L(A,B) \leq (A+B)/2$.

13. Let $A, B > 0$. Show that if for a matrix mean $M_f(A,B) = A$, then $A = B$.

14. Let $f, g : \mathbb{R}^+ \to \mathbb{R}^+$ be matrix monotone functions. Show that their arithmetic and geometric means are matrix monotone as well.

15. Show that the mean matrix

$$A_{ij} = \frac{1}{H_t(\lambda_i, \lambda_j)}$$

defined by the Heinz mean is positive.

Chapter 10

Matrices in quantum theory

When quantum theory was developed in the 1920's, it had a strong influence on functional analysis. The quantum information theory appeared much later and the matrices are very important in this subject.

10.1 Axioms

The basic postulate of quantum mechanics is about the Hilbert space formalism.

(A0) To each quantum mechanical system a complex Hilbert space $\mathcal{H}$ is associated.

The (pure) physical states of the system correspond to unit vectors of the Hilbert space. This correspondence is not 1-1. When f_1 and f_2 are unit vectors, then the corresponding states identical if $f_1 = zf_2$ for a complex number z of modulus 1. Such z is often called **phase**. The **pure physical state** of the system determines a corresponding state vector up to a phase.

Example 10.1 The 2 dimensional Hilbert space $\mathbb{C}^2$ is used to describe a 2-level quantum system called **qubit**. The canonical basis vectors $(1,0)$ and $(0,1)$ are usually denoted by $|\uparrow\rangle$ and $|\downarrow\rangle$, respectively. (An alternative notation is $|1\rangle$ for $(0,1)$ and $|0\rangle$ for $(1,0)$.) Since the polarization of a photon is an important example of a qubit, the state $|\uparrow\rangle$ may have the interpretation that the "polarization is vertical" and $|\downarrow\rangle$ means that the "polarization is horizontal".

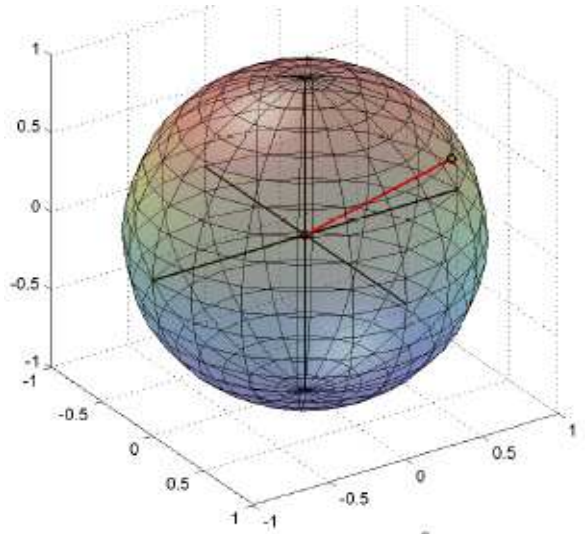
To specify a state of a qubit we need to give a real number x_1 and a complex number z such that $x_1^2 + |z|^2 = 1$. Then the state vector is

$$x_1 |\uparrow\rangle + z |\downarrow\rangle .$$

(Indeed, multiplying a unit vector $z_1 |\uparrow\rangle + z_2 |\downarrow\rangle$ by an appropriate phase, we can make the coefficient of $|\uparrow\rangle$ real and the corresponding state remains the same.)

Splitting z into real and imaginary parts as $z = x_2 + \mathrm{i}x_3$, we have the constraint $x_1^2 + x_2^2 + x_3^2 = 1$ for the parameters $(x_1, x_2, x_3) \in \mathbb{R}^3$.

Therefore, the space of all pure states of a qubit is conveniently visualized as the sphere in the three dimensional Euclidean space, it is called the **Bloch sphere**. □



Bloch ball

A 2×2 density matrix has the form $\frac{1}{2}(I + x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3)$, where $x_1^2 + x_2^2 + x_3^2 \leq 1$. The length of the vectors (x_1, x_2, x_3) is at most 1 and they form the unit ball, called Bloch ball, in the three dimensional Euclidean space. The pure states are on the surface.

Traditional quantum mechanics distinguishes between pure states and **mixed states**. Mixed states are described by **density matrices**. A density matrix or statistical operator is a positive operator of trace 1 on the Hilbert space. This means that the space has a basis consisting of eigenvectors of the statistical operator and the sum of eigenvalues is 1. (In the finite dimensional case the first condition is automatically fulfilled.) The pure states represented by unit vectors of the Hilbert space are among the density matrices under an appropriate identification. If $x = |x\rangle$ is a unit vector, then $|x\rangle\langle x|$ is a density matrix. Geometrically $|x\rangle\langle x|$ is the orthogonal projection onto the linear subspace generated by x. Note that $|x\rangle\langle x| = |y\rangle\langle y|$ if the vectors x and y differ only in a phase.

(A1) The physical states of a quantum mechanical system are described by statistical operators acting on the Hilbert space.

Example 10.2 A state of the spin (of 1/2) can be represented by the 2×2 matrix

$$\frac{1}{2}\begin{bmatrix} 1 + x_3 & x_1 - \mathrm{i}x_2 \\ x_1 + \mathrm{i}x_2 & 1 - x_3 \end{bmatrix}. \tag{10.1}$$

This is a density matrix if and only if $x_1^2 + x_2^2 + x_3^2 \leq 1$. □

The second axiom is about observables.

(A2) The observables of a quantum mechanical system are described by self-adjoint operators acting on the Hilbert space.

A **self-adjoint operator** A on a Hilbert space $\mathcal{H}$ is a linear operator $\mathcal{H} \to \mathcal{H}$ which satisfies

$$\langle Ax,\, y\rangle = \langle x,\, Ay\rangle$$

for $x, y \in \mathcal{H}$. Self-adjoint operators on a finite dimensional Hilbert space $\mathbb{C}^n$ are $n \times n$ self-adjoint matrices. A self-adjoint matrix admits a **spectral decomposition** $A = \sum_i \lambda_i E_i$, where λ_i are the different eigenvalues of A and E_i is the orthogonal projection onto the subspace spanned by the eigenvectors corresponding to the eigenvalue λ_i. Multiplicity of λ_i is exactly the rank of E_i.

Example 10.3 In case of a quantum spin (of 1/2) the matrices

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma_2 = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}, \qquad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

are used to describe the spin of direction x, y, z (with respect to a coordinate system). They are called **Pauli matrices**. Any 2×2 self-adjoint matrix is of the form

$$A_{(x_0,x)} := x_0\sigma_0 + x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3$$

if σ_0 stands for the unit matrix I. We also use the shorthand notation $x_0\sigma_0 + x \cdot \sigma$.

The density matrix (10.1) can be written as

$$\tfrac{1}{2}(\sigma_0 + x \cdot \sigma), \tag{10.2}$$

where $\|x\| \leq 1$. x is called **Bloch vector** and they form the **Bloch ball**.

Formula (10.2) makes an affine correspondence between 2×2 density matrices and the unit ball in the Euclidean 3-space. The extreme points of the ball correspond to pure states and any mixed state is the convex combination of pure states in infinitely many different ways. In higher dimension the situation is much more complicated. □

Any density matrix can be written in the form

$$\rho = \sum_i \lambda_i |x_i\rangle\langle x_i| \tag{10.3}$$

by means of unit vectors $|x_i\rangle$ and coefficients $\lambda_i \geq 0$, $\sum_i \lambda_i = 1$. Since ρ is self-adjoint such a decomposition is deduced from the spectral theorem and the vectors $|x_i\rangle$ may be chosen as pairwise orthogonal eigenvectors and λ_i are the corresponding eigenvalues. The decomposition is unique if the spectrum of ρ is non-degenerate, that is, there is no multiple eigenvalue.

Lemma 10.1 *The density matrices acting on a Hilbert space form a convex set whose extreme points are the pure states.*

Proof: Denote by Σ the set of density matrices. It is obvious that a convex combination of density matrices is positive and of trace one. Therefore Σ is a convex set.

Recall that $\rho \in \Sigma$ is an extreme point if a convex decomposition $\rho = \lambda\rho_1 + (1-\lambda)\rho_2$ with $\rho_1, \rho_2 \in \Sigma$ and $0 < \lambda < 1$ is only trivially possible, that is, $\rho_1 = \rho_2 = \rho$. The Schmidt decomposition (10.3) shows that an extreme point must be a pure state.

Let p be a pure state, $p = p^2$. We have to show that it is really an extreme point. Assume that $p = \lambda\rho_1 + (1-\lambda)\rho_2$. Then

$$p = \lambda p\rho_1 p + (1-\lambda)p\rho_2 p$$

and $\operatorname{Tr} p\rho_i p = 1$ must hold. Remember that $\operatorname{Tr} p\rho_i p = \langle p, \rho_i\rangle$, while $\langle p, p\rangle = 1$ and $\langle \rho_i, \rho_i\rangle \leq 1$. In the Schwarz inequality

$$|\langle e, f\rangle|^2 \leq \langle e, e\rangle \langle f, f\rangle$$

the equality holds if and only if $f = ce$ for some complex number c. Therefore, $\rho_i = c_i p$ must hold. Taking the trace, we get $c_i = 1$ and $\rho_1 = \rho_2 = p$. □

Quantum mechanics is not deterministic. If we prepare two identical systems in the same state, and we measure the same observable on each, then the result of the **measurement** may not be the same. This indeterminism or stochastic feature is fundamental.

(A3) Let $\mathcal{X}$ be a finite set and for $x \in \mathcal{X}$ an operator $V_x \in B(\mathcal{H})$ be given such that $\sum_x V_x^* V_x = I$. Such an indexed family of operators is a model of a measurement with values in $\mathcal{X}$. If the measurement is performed in a state ρ, then the outcome $x \in \mathcal{X}$ appears with probability $\operatorname{Tr} V_x \rho V_x^*$ and after the measurement the state of the system is

$$\frac{V_x \rho V_x^*}{\operatorname{Tr} V_x \rho V_x^*}.$$

A particular case is the measurement of an observable described by a self-adjoint operator A with spectral decomposition $\sum_i \lambda_i E_i$. In this case $\mathcal{X} = \{\lambda_i\}$ is the set of eigenvalues and $V_i = E_i$. One compute easily that the expectation of the random outcome is $\operatorname{Tr} \rho A$. The functional $A \mapsto \operatorname{Tr} \rho A$ is linear and has two important properties: 1. If $A \geq 0$, then $\operatorname{Tr} \rho A \geq 0$, 2.$\operatorname{Tr} \rho I = 1$. These properties allow to see quantum states in a different way. If $\varphi : B(\mathcal{H}) \to \mathbb{C}$ is a linear functional such that

$$\varphi(A) \geq 0 \quad \text{if} \quad A \geq 0 \quad \text{and} \quad \varphi(I) = 1, \tag{10.4}$$

then there exists a density matrix ρ_φ such that

$$\varphi(A) = \operatorname{Tr} \rho_\varphi A. \tag{10.5}$$

The functional φ associates the expectation value to the observables A.

The density matrices ρ_1 and ρ_2 are called **orthogonal** if any eigenvector of ρ_1 is orthogonal to any eigenvector of ρ_2.

Example 10.4 Let ρ_1 and ρ_2 be density matrices. They can be **distinguished with certainty** if there exists a measurement which takes the value 1 with probability 1 when the system is in the state ρ_1 and with probability 0 when the system is in the state ρ_2.

Assume that ρ_1 and ρ_2 are orthogonal and let P be the orthogonal projection onto the subspace spanned by the non-zero eigenvectors of ρ_1. Then $V_1 := P$ and $V_2 := I - P$ is a measurement and $\operatorname{Tr} V_1 \rho_1 V_1^* = 1$ and $\operatorname{Tr} V_1 \rho_2 V_1^* = 0$.

Conversely, assume that a measurement (V_i) exists such that $\operatorname{Tr} V_1 \rho_1 V_1^* = 1$ and $\operatorname{Tr} V_1 \rho_2 V_1^* = 0$. The first condition implies that $V_1^* V_1 \geq P$, where P us the support projection of ρ_1, defined above. The second condition tells is that $V_1^* V_1$ is orthogonal to the support of ρ_2. Therefore, $\rho_1 \perp \rho_2$. □

Let $e_1, e_2, \ldots, e_n$ be an orthonormal basis in a Hilbert space $\mathcal{H}$. The unit vector $\xi \in \mathcal{H}$ is **complementary** to the given basis if

$$|\langle e_i, \xi \rangle| = \frac{1}{\sqrt{n}} \qquad (1 \le i \le n). \tag{10.6}$$

The basis vectors correspond to a measurement, $|e_1\rangle\langle e_1|, \ldots, |e_n\rangle\langle e_n|$ are positive operators and their sum is I. If the pure state $|\xi\rangle\langle\xi|$ the actual state of the quantum system, then complementarity means that all outputs of the measurement appear with the same probability.

Two orthonormal bases are called **complementary** if all vectors in the first basis are complementary to the other basis.

Example 10.5 First we note that equation (10.6) is equivalent to the relation

$$\operatorname{Tr} |e_i\rangle\langle e_i| \, |\xi\rangle\langle\xi| = \frac{1}{n} \tag{10.7}$$

which is about the trace of the product of two projections.

The eigenprojections of the Pauli matrix σ_i are $(I \pm \sigma_i)/2$. We have

$$\operatorname{Tr}\Big(\frac{I \pm \sigma_i}{2}\frac{I \pm \sigma_j}{2}\Big) = \frac{1}{2}$$

for $1 \le i \ne j \le 3$. This shows that the eigenbasis of σ_i is complementary to the eigenbasis of σ_j if i and j are different. □

According to axiom (A1), a Hilbert space is associated to any quantum mechanical system. Assume that a **composite system** consists of the subsystems (1) and (2), they are described by the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. (Each subsystem could be a particle or a spin, for example.) Then we have

(A4) The composite system is described by the tensor product Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$.

When $\{e_j : j \in J\}$ is a basis in $\mathcal{H}_1$ and $\{f_i : i \in I\}$ is a basis in $\mathcal{H}_2$, then $\{e_j \otimes f_i : j \in J, i \in I\}$ is a basis of $\mathcal{H}_1 \otimes \mathcal{H}_2$. Therefore, the dimension of $\mathcal{H}_1 \otimes \mathcal{H}_2$ is $\dim \mathcal{H}_1 \times \dim \mathcal{H}_2$. If $A_i \in B(\mathcal{H}_i)$ $(i = 1, 2)$, then the action of the tensor product operator $A_1 \otimes A_2$ is determined by

$$(A_1 \otimes A_2)(\eta_1 \otimes \eta_2) = A_1\eta_1 \otimes A_2\eta_2$$

since the vectors $\eta_1 \otimes \eta_2$ span $\mathcal{H}_1 \otimes \mathcal{H}_2$.

When $A = A^*$ is an observable of the first system, then its expectation value in the vector state $\Psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$, is

$$\langle \Psi, (A \otimes I_2)\Psi \rangle \,,$$

where I_2 is the identity operator on $\mathcal{H}_2$.

Example 10.6 The Hilbert space of a composite system of two spins (of 1/2) is $\mathbb{C}^2 \otimes \mathbb{C}^2$. In this space, the vectors

$$e_1 := |\uparrow\rangle \otimes |\uparrow\rangle, \quad e_2 := |\uparrow\rangle \otimes |\downarrow\rangle, \quad e_3 := |\downarrow\rangle \otimes |\uparrow\rangle, \quad e_4 := |\downarrow\rangle \otimes |\downarrow\rangle$$

form a basis. The vector state

$$\Phi = \frac{1}{\sqrt{2}}(|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle) \tag{10.8}$$

has a surprising property. Consider the observable

$$A := \sum_{i=1}^{4} i|e_i\rangle\langle e_i|,$$

which has eigenvalues $1, 2, 3, 4$ and the corresponding eigenvectors are just the basis vectors. Measurement of this observable yields the values $1, 2, 3, 4$ with probabilities $0, 1/2, 1/2$ and 0, respectively. The 0 probability occurs when both spins are up or both are down. Therefore in the vector state Φ the spins are anti-correlated. □

We consider now the composite system $\mathcal{H}_1 \otimes \mathcal{H}_2$ in a state $\Phi \in \mathcal{H}_1 \otimes \mathcal{H}_2$. Let $A \in B(\mathcal{H}_1)$ be an observable which is localized at the first subsystem. If we want to consider A as an observable of the total system, we have to define an extension to the space $\mathcal{H}_1 \otimes \mathcal{H}_2$. The tensor product operator $A \otimes I$ will do, I is the identity operator of $\mathcal{H}_2$.

Lemma 10.2 *Assume that $\mathcal{H}_1$ and $\mathcal{H}_2$ are finite dimensional Hilbert spaces. Let $\{e_j : j \in J\}$ be a basis in $\mathcal{H}_1$ and $\{f_i : i \in I\}$ be a basis in $\mathcal{H}_2$. Assume that*

$$\Phi = \sum_{i,j} w_{ij}\, e_j \otimes f_i$$

is the expansion of a unit vector $\Phi \in \mathcal{H}_1 \otimes \mathcal{H}_2$. Set W for the matrix which is determined by the entries w_{kl}. Then W^*W is a density matrix and

$$\langle \Phi,\ (A \otimes I)\Phi \rangle = \operatorname{Tr} AW^*W\,.$$

Proof: Let E_{kl} be an operator on $\mathcal{H}_1$ which is determined by the relations $E_{kl}e_j = \delta_{lj}e_k$ $(k, l \in I)$. As a matrix, E_{kl} is called matrix unit, it is a matrix such that (k, l) entry is 1, all others are 0. Then

$$\begin{aligned}
\langle \Phi,\ (E_{kl} \otimes I)\Phi \rangle &= \left\langle \sum_{i,j} w_{ij}\, e_j \otimes f_i,\ (E_{kl} \otimes I) \sum_{t,u} w_{tu}\, e_u \otimes f_t \right\rangle = \\
&= \sum_{i,j} \sum_{t,u} \overline{w}_{ij} w_{tu} \langle e_j,\ E_{kl}e_u \rangle \langle f_i, f_t \rangle = \\
&= \sum_{i,j} \sum_{t,u} \overline{w}_{ij} w_{tu}\, \delta_{lu}\delta_{jk}\delta_{it} = \sum_{i} \overline{w}_{ik} w_{il}\,.
\end{aligned}$$

Then we arrived at the (k, l) entry of W^*W. Our computation may be summarized as

$$\langle \Phi,\ (E_{kl} \otimes I)\Phi \rangle = \operatorname{Tr} E_{kl}(W^*W) \qquad (k, l \in I).$$

Since any linear operator $A \in B(\mathcal{H}_1)$ is of the form $A = \sum_{k,l} a_{kl} E_{kl}$ ($a_{kl} \in \mathbb{C}$), taking linear combinations of the previous equations, we have

$$\langle \Phi,\, (A \otimes I)\Phi \rangle = \operatorname{Tr} A(W^*W)\,.$$

W^*W is obviously positive and

$$\operatorname{Tr} W^*W = \sum_{i,j} |w_{ij}|^2 = \|\Phi\|^2 = 1\,.$$

Therefore it is a density matrix. □

This lemma shows a natural way from state vectors to density matrices. Given a density matrix ρ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ there are density matrices $\rho_i \in B(\mathcal{H}_i)$ such that

$$\operatorname{Tr} (A \otimes I)\rho = \operatorname{Tr} A\rho_1 \qquad (A \in B(\mathcal{H}_1)) \tag{10.9}$$

and

$$\operatorname{Tr} (I \otimes B)\rho = \operatorname{Tr} B\rho_2 \qquad (B \in B(\mathcal{H}_2)). \tag{10.10}$$

ρ_1 and ρ_2 are called **reduced density matrices**. (They are the quantum analogue of marginal distributions.)

The proof of Lemma 10.2 contains the reduced density of $|\Phi\rangle\langle\Phi|$ on the first system, it is W^*W. One computes similarly the reduced density on the second subsystem, it is $(WW^*)^T$, where X^T denotes the transpose of the matrix X. Since W^*W and $(WW^*)^T$ have the same non-zero eigenvalues, the two subsystems are very strongly connected if the total system is in a pure state.

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces and let $\dim \mathcal{H}_1 = m$ and $\dim \mathcal{H}_2 = n$. It is well-known that the matrix of a linear operator on $\mathcal{H}_1 \otimes \mathcal{H}_2$ has a block-matrix form

$$U = (U_{ij})_{i,j=1}^m = \sum_{i,j=1}^m E_{ij} \otimes U_{ij},$$

relative to the lexicographically ordered product basis, where U_{ij} are $n \times n$ matrices. For example,

$$A \otimes I \;=\; (X_{ij})_{i,j=1}^m\,, \quad \text{where} \quad X_{ij} = A_{ij} I_n$$

and

$$I \otimes B \;=\; (X_{ij})_{i,j=1}^m\,, \quad \text{where} \quad X_{ij} = \delta_{ij} B.$$

Assume that

$$\rho = (\rho_{ij})_{i,j=1}^m = \sum_{i,j=1}^m E_{ij} \otimes \rho_{ij}$$

is a density matrix of the composite system written in block-matrix form. Then

$$\operatorname{Tr} (A \otimes I)\rho = \sum_{i,j} A_{ij} \operatorname{Tr} I_n \rho_{ij} = \sum_{i,j} A_{ij} \operatorname{Tr} \rho_{ij}$$

and this gives that for the first reduced density matrix ρ_1, we have

$$(\rho_1)_{ij} = \operatorname{Tr} \rho_{ij}. \tag{10.11}$$

We can compute similarly the second reduced density ρ_2. Since

$$\operatorname{Tr}(I \otimes B)\rho = \sum_i \operatorname{Tr} B\rho_{ii}$$

we obtain

$$\rho_2 = \sum_{i=1}^{m} \rho_{ii}. \tag{10.12}$$

The reduced density matrices might be expressed by the **partial traces**. The mappings $\operatorname{Tr}_2 : B(\mathcal{H}_1) \otimes B(\mathcal{H}_2) \to B(\mathcal{H}_1)$ and $\operatorname{Tr}_1 : B(\mathcal{H}_1) \otimes B(\mathcal{H}_2) \to B(\mathcal{H}_2)$ are defined as

$$\operatorname{Tr}_2(A \otimes B) = A\operatorname{Tr} B, \qquad \operatorname{Tr}_1(A \otimes B) = (\operatorname{Tr} A)B\,. \tag{10.13}$$

We have

$$\rho_1 = \operatorname{Tr}_2\rho \qquad \text{and} \qquad \rho_2 = \operatorname{Tr}_1\rho\,. \tag{10.14}$$

Axiom (A4) tells about a composite quantum system consisting of two quantum components. In case of more quantum components, the formalism is similar, but more tensor factors appear.

It may happen that the quantum system under study has a classical and a quantum component, assume that the first component is classical. Then the description by tensor product Hilbert space is still possible. A basis $(|e_i\rangle)_i$ of $\mathcal{H}_1$ can be fixed and the possible density matrices of the joint system are of the form

$$\sum_i p_i |e_i\rangle\langle e_i| \otimes \rho_i^{(2)}\,, \tag{10.15}$$

where $(p_i)_i$ is a probability distribution and $\rho_i^{(2)}$ are densities on $\mathcal{H}_2$. Then the reduced state on the first component is the probability density $(p_i)_i$ (which may be regarded as a diagonal density matrix) and $\sum_i p_i\rho_i^{(2)}$ is the second reduced density.

The next postulate of quantum mechanics tells about the **time development** of a closed quantum system. If the system is not subject to any measurement in the time interval $I \subset \mathbb{R}$ and ρ_t denotes the statistical operator at time t, then

(A5) $\qquad \rho_t = U(t,s)\rho_s U(t,s)^* \qquad (t,s \in I),$

where the **unitary propagator** $U(t,s)$ is a family of unitary operators such that

(i) $U(t,s)U(s,r) = U(t,r)$,

(ii) $(s,t) \mapsto U(s,t) \in B(\mathcal{H})$ is strongly continuous.

The first order approximation of the unitary $U(s,t)$ is the **Hamiltonian**:

$$U(t+\Delta t, t) = I - \frac{\mathrm{i}}{\hbar}H(t)\Delta t,$$

where $H(t)$ is the Hamiltonian at time t. If the Hamiltonian is time independent, then

$$U(s,t) = \exp\left(-\frac{\mathrm{i}}{\hbar}(s-t)H\right).$$

In the approach followed here the density matrices are transformed in time, this is the so-called **Schrödinger picture** of quantum mechanics. When discrete time development is considered, a single unitary U gives the transformation of the vector state in the form $\psi \mapsto U\psi$, or in the density matrix formalism $\rho \mapsto U\rho U^*$.

Example 10.7 Let $|0\rangle, |1\rangle, \ldots, |n-1\rangle$ be an orthonormal basis in an n-dimensional Hilbert space. The transformation

$$V : |i\rangle \mapsto \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \omega^{ij} |j\rangle \qquad (\omega = e^{2\pi \mathrm{i}/n}) \tag{10.16}$$

is a unitary and it is called **quantum Fourier transform**. □

When the unitary time development is viewed as a quantum algorithm in connection with quantum computation, the term **gate** is used instead of unitary.

Example 10.8 Unitary operators are also used to manipulate quantum registers and to implement quantum algorithms.

The **Hadamard gate** is the unitary operator

$$U_H := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{10.17}$$

It sends the basis vectors into uniform superposition and vice versa. The Hadamard gate can establish or destroy the superposition of a qubit. This means that the basis vector $|0\rangle$ is transformed into the vector $(|0\rangle + |1\rangle)/\sqrt{2}$ which is a superposition and superposition is created.

The **controlled-NOT gate** is a unitary acting on two qubits. The first qubit is called a control qubit, and the second qubit is the data qubit. This operator sends the basis vectors $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ of $\mathbb{C}^4$ into $|00\rangle, |01\rangle, |11\rangle, |10\rangle$. When the first character is 1, the second changes under the operation. Therefore, the matrix of the controlled-NOT gate is

$$U_{c-NOT} := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{10.18}$$

The **swap gate** moves a product vector $|i\rangle \otimes |j\rangle$ into $|j\rangle \otimes |i\rangle$. Therefore its matrix is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{10.19}$$

Quantum algorithms involve several other gates. □

Example 10.9 The unitary operators are used to transform a basis into another one. In the Hilbert space $\mathbb{C}^4 = \mathbb{C}^2 \otimes \mathbb{C}^2$ the standard basis is

$$|00\rangle, |01\rangle, |10\rangle, |11\rangle.$$

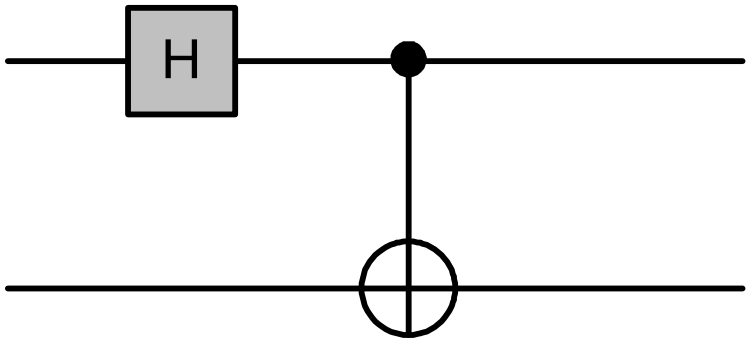
The unitary

$$(U_H \otimes I_2)U_{c-NOT} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{bmatrix}.$$

moves the standard basis into the so-called **Bell basis**:

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \quad \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \quad \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle).$$

This basis is complementary to the standard product basis.



Hadamard gate

The unitary made of the Hadamard gate and the controlled-NOT gate transforms the standard product basis into the Bell basis.

10.2 State Transformations

Assume that $\mathcal{H}$ is the Hilbert space of our quantum system which initially has a statistical operator ρ (acting on $\mathcal{H}$). When the quantum system is not closed, it is coupled to another system, called **environment**. The environment has a Hilbert space $\mathcal{H}_e$ and statistical operator ρ_e. Before interaction the total system has density $\rho_e \otimes \rho$. The dynamical change caused by the interaction is implemented by a unitary and $U(\rho_e \otimes \rho)U^*$ is the new statistical operator and the reduced density $\tilde{\rho}$ is the new statistical operator of the quantum system we are interested in. The affine change $\rho \mapsto \tilde{\rho}$ is typical for quantum mechanics and called **state transformation**. In this way the map $\rho \mapsto \tilde{\rho}$ is defined on density matrices but it can be extended by linearity to all matrices. In this way we obtain a trace preserving and positivity preserving linear transformation.

The above defined state transformation can be described in several other forms, reference to the environment could be omitted completely. Assume that ρ is an $n \times n$ matrix and ρ_e is of the form $(z_k\overline{z_l})_{kl}$ where $(z_1, z_2, \ldots, z_m)$ is a unit vector in the m dimensional space $\mathcal{H}_e$. (ρ_e is a pure state.) All operators acting on $\mathcal{H}_e \otimes \mathcal{H}$ are written in a block matrix form, they are $m \times m$ matrices with $n \times n$ matrix entries. In particular, $U = (U_{ij})_{i,j=1}^m$ and $U_{ij} \in M_n$. If U is a unitary, then U^*U is the identity and this implies that

$$\sum_i U_{ik}^* U_{il} = \delta_{kl} I_n \tag{10.20}$$

Formula (10.12) for the reduced density matrix gives

$$\begin{aligned}\tilde{\rho} &= \mathrm{Tr}_1(U(\rho_e \otimes \rho)U^*) = \sum_i (U(\rho_e \otimes \rho)U^*)_{ii} = \sum_{i,k,l} U_{ik}(\rho_e \otimes \rho)_{kl}(U^*)_{li} \\ &= \sum_{i,k,l} U_{ik}(z_k\overline{z}_l\rho)(U_{il})^* = \sum_i \Big(\sum_k z_k U_{ik}\Big)\rho\Big(\sum_l z_l U_{il}\Big)^* = \sum_i A_i \rho A_i^*\,,\end{aligned}$$

where the operators $A_i := \sum_k z_k U_{ik}$ satisfy

$$\sum_p A_p^* A_p = I \tag{10.21}$$

due to (10.20) and $\sum_k |z_k|^2 = 1$.

Theorem 10.1 *Any state transformation $\rho \mapsto \mathcal{E}(\rho)$ can be written in the form*

$$\mathcal{E}(\rho) = \sum_p A_p \rho A_p^*,$$

where the operator coefficients satisfy (10.21)*. Conversely, all linear mappings of this form are state transformations.*

The first part of the theorem was obtained above. To prove the converse part, we need to solve the equations

$$A_i := \sum_k z_k U_{ik} \qquad (i = 1, 2, \ldots, m).$$

Choose simply $z_1 = 1$ and $z_2 = z_3 = \ldots = z_m = 0$ and the equations reduce to $U_{p1} = A_p$. This means that the first column is given from the block matrix U and we need to determine the other columns such a way that U should be a unitary. Thanks to the condition (10.21) this is possible. Condition (10.21) tells us that the first column of our block matrix determines an isometry which extends to a unitary. □

The coefficients A_p in the **operator-sum representation** are called the **operation elements** of the state transformation. The terms quantum (state) operation and channeling transformation are also often used instead of state transformation.

The state transformations form a convex subset of the set of all positive trace preserving linear transformations. (It is not known what the extreme points of this set are.)

A linear mapping $\mathcal{E}$ is called **completely positive** if $\mathcal{E} \otimes id_n$ is positivity preserving for the identical mapping $id_n : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ on any matrix algebra.

Theorem 10.2 *Let $\mathcal{E} : M_n(\mathbb{C}) \to M_k(\mathbb{C})$ be a linear mapping. Then $\mathcal{E}$ is completely positive if and only if it admits a representation*

$$\mathcal{E}(A) = \sum_u V_u A V_u^* \tag{10.22}$$

by means of some linear operators $V_u : \mathbb{C}^n \to \mathbb{C}^k$.

This result was first proven by Kraus. It follows that a state transformation is completely positive and the operator-sum representation is also called **Kraus representation**. Note that this representation is not unique.

Let $\mathcal{E} : M_n(\mathbb{C}) \to M_k(\mathbb{C})$ be a linear mapping. $\mathcal{E}$ is determined by the block-matrix $(X_{ij})_{1\le i,j\le k}$, where

$$X_{ij} = \mathcal{E}(E_{ij}) \tag{10.23}$$

(Here E_{ij} denote the matrix units.) This is the **block-matrix representation** of $\mathcal{E}$.

The next theorem is due to **Choi**.

Theorem 10.3 *Let $\mathcal{E} : M_n(\mathbb{C}) \to M_k(\mathbb{C})$ be a linear mapping. Then $\mathcal{E}$ is completely positive if and only if the representing block-matrix $(X_{ij})_{1\le i,j\le k} \in M_k(\mathbb{C}) \otimes M_n(\mathbb{C})$ is positive.*

Proof: Instead of the block-matrix formalism, we can use tensor product:

$$X = \sum_{i,j} E_{ij} \otimes \mathcal{E}(E_{ij}).$$

X is a linear function of $\mathcal{E}$, hence we can assume that $\mathcal{E}(A) = VAV^*$.

Let

$$B := \sum_{i,j} E_{ij} \otimes E_{ij}.$$

It is easy to check that $B = B^*$ and $B^2 = nB$, therefore $B \ge 0$. On the other hand,

$$X = \sum_{i,j} E_{ij} \otimes VE_{ij}V^* = (I \otimes V)B(I \otimes V)^*$$

which is positive.

Assume that the block-matrix X is positive. There are projections P_i from $\mathbb{C}^{nk} = \oplus_{\ell=1}^{n}\mathbb{C}^k$ to the ith summand ($1 \le i \le n$). They are pairwise orthogonal and

$$P_i X P_j^* = \mathcal{E}(E_{ij}).$$

We have a decomposition

$$X = \sum_{t=1}^{nk} |f_t\rangle\langle f_t|,$$

where $|f_t\rangle$ are appropriately normalized eigenvectors of X. Since P_i is a partition of unity, we have

$$|f_t\rangle = \sum_{i=1}^{n} P_i|f_t\rangle$$

and set $V_t : \mathbb{C}^n \to \mathbb{C}^k$ by

$$V_t|s\rangle = P_s|f_t\rangle.$$

($|s\rangle$ are the canonical basis vectors.) In this notation

$$X = \sum_t \sum_{i,j} P_i|f_t\rangle\langle f_t|P_j^* = \sum_{i,j} P_i \left(\sum_t V_t|i\rangle\langle j|V_t^* \right) P_j^*$$

and

$$\mathcal{E}(E_{ij}) = P_i X P_j^* = \sum_t V_t E_{ij} V_t^*.$$

Since this holds for all matrix units E_{ij}, we obtained

$$\mathcal{E}(A) = \sum_t V_t A V_t^*$$

which means the complete positivity. □

Example 10.10 Consider the transpose mapping $A \mapsto A^T$ on 2×2 matrices:

$$\begin{bmatrix} x & y \\ z & w \end{bmatrix} \mapsto \begin{bmatrix} x & z \\ y & w \end{bmatrix}.$$

The representing block-matrix is

$$X = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

This is not positive, so the transpose mapping is not completely positive. □

Example 10.11 Consider a positive trace-preserving transformation $\mathcal{E} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ such that its range consists of commuting operators. We show that $\mathcal{E}$ is automatically a state transformation.

Since a commutative subalgebra of $M_m(\mathbb{C})$ is the linear span of some pairwise orthogonal projections P_k, one can see that $\mathcal{E}$ has the form

$$\mathcal{E}(A) = \sum_k P_k \operatorname{Tr} F_k A, \tag{10.24}$$

where F_k is a positive operator in $M_n(\mathbb{C})$, it induces the coefficient of P_k as a linear functional on $M_n(\mathbb{C})$.

We want to show the positivity of the representing block-matrix:

$$\sum_{ij} E_{ij} \otimes \Big(\sum_k P_k \operatorname{Tr}(F_k E_{ij}) \Big) = \sum_k \Big(\sum_{ij} E_{ij} \otimes P_k \Big) \circ \Big(\sum_{ij} E_{ij} \operatorname{Tr}(F_k E_{ij}) \otimes I \Big),$$

where $\circ$ denotes the Hadamard (or entry-wise product) of $nm \times nm$ matrices. Recall that according to Schur's theorem the **Hadamard product** of positive matrices is positive. The first factor is

$$[P_k, P_k, \dots, P_k]^* [P_k, P_k, \dots, P_k]$$

and the second factor is $F_k \otimes I$, both are positive.

Consider the particular case of (10.24) where each P_k is of rank one and $\sum_{k=1}^r F_k = I$. Such a family of F_k's describe a measurement which associates the r-tuple $(\operatorname{Tr} \rho F_1, \operatorname{Tr} \rho F_2, \dots, \operatorname{Tr} \rho F_r)$ to the density matrix ρ. Therefore a measurement can be formulated as a state transformation with diagonal outputs. □

The Kraus representation and the block-matrix representation are convenient ways to describe a state transformation in any finite dimension. In the 2×2 case we have the possibility to expand the mappings in the basis $\sigma_0, \sigma_1, \sigma_2, \sigma_3$.

Any trace preserving mapping $\mathcal{E} : M_2(\mathbb{C}) \to M_2(\mathbb{C})$ has a matrix

$$T = \begin{bmatrix} 1 & 0 \\ t & T_3 \end{bmatrix} \tag{10.25}$$

with respect to this basis, where $T_3 \in M_3$ and

$$\mathcal{E}(w_0\sigma_0 + w \cdot \sigma) = w_0\sigma_0 + (t + T_3 w) \cdot \sigma. \tag{10.26}$$

Since $\mathcal{E}$ sends self-adjoint operators to self-adjoint operators, we may assume that T_3 is a real 3×3 matrix. It has a singular value decomposition $O_1 \Sigma O_2$, where O_1 and O_2 are orthogonal matrices and Σ is diagonal. Since any orthogonal transformation on $\mathbb{R}^3$ is induced by a unitary conjugation on $M_2(\mathbb{C})$, in the study of state transformations, we can assume that T_3 is diagonal.

The following examples of state transformations are given in terms of the T-representation:

Example 10.12 (Pauli channels) $t = 0$ and $T_3 = \mathrm{Diag}(\alpha, \beta, \gamma)$. Density matrices are sent to density matrices if and only if

$$-1 \leq \alpha, \beta, \gamma \leq 1$$

for the real parameters α, β, γ.

It is not difficult to compute the representing block-matrix, we have

$$X = \begin{bmatrix} \frac{1+\gamma}{2} & 0 & 0 & \frac{\alpha+\beta}{2} \\ 0 & \frac{1-\gamma}{2} & \frac{\alpha-\beta}{2} & 0 \\ 0 & \frac{\alpha-\beta}{2} & \frac{1-\gamma}{2} & 0 \\ \frac{\alpha+\beta}{2} & 0 & 0 & \frac{1+\gamma}{2} \end{bmatrix}. \tag{10.27}$$

X is unitarily equivalent to the matrix

$$\begin{bmatrix} \frac{1+\gamma}{2} & \frac{\alpha+\beta}{2} & 0 & 0 \\ \frac{\alpha+\beta}{2} & \frac{1+\gamma}{2} & 0 & 0 \\ 0 & 0 & \frac{1-\gamma}{2} & \frac{\alpha-\beta}{2} \\ 0 & 0 & \frac{\alpha-\beta}{2} & \frac{1-\gamma}{2} \end{bmatrix}.$$

This matrix is obviously positive if and only if

$$|1 \pm \gamma| \geq |\alpha \pm \beta|. \tag{10.28}$$

This positivity condition holds when $\alpha = \beta = \gamma = p > 0$. Hence the next example gives a channeling transformation. □

Example 10.13 (Depolarizing channel) This channel is given by the matrix T from (10.25), where $t = 0$ and $T_3 = pI$. Assume that $0 < p < 1$.

Since

$$\mathcal{E}_p(\tfrac{1}{2}\sigma_0 + w\cdot\sigma) = p(\tfrac{1}{2}\sigma_0 + w\cdot\sigma) + (1-p)\tfrac{1}{2}\sigma_0 = \tfrac{1}{2}\sigma_0 + p(w\cdot\sigma),$$

the depolarizing channel keeps the density with probability p and moves to the completely apolar state $\sigma_0/2$ with probability $1-p$.

Extension to n-level systems is rather obvious. $\mathcal{E}_{p,n} : M_n \to M_n$ is defined as

$$\mathcal{E}_{p,n}(A) = pA + (1-p)\frac{I}{n}\mathrm{Tr}\, A\,. \tag{10.29}$$

This mapping is trivially completely positive for $0 \le p \le 1$, since it is the convex combination of such mappings. In order to consider the negative values of p we study the representing block-matrix X. One can see that

$$X = p\sum_{ij} E_{ij}\otimes E_{ij} + \frac{1-p}{n} I\otimes I.$$

The matrix $\frac{1}{n}\sum_{ij} E_{ij}\otimes E_{ij}$ is a self-adjoint idempotent (that is, a projection), so its spectrum is $\{0,1\}$. Consequently, the eigenvalues of X are

$$pn + \frac{1-p}{n}, \frac{1-p}{n}\,.$$

They are positive when

$$-\frac{1}{n^2-1} \le p \le 1\,. \tag{10.30}$$

This is the necessary and sufficient condition for the complete positivity of $\mathcal{E}_{p,n}$. □

Example 10.14 (Phase-damping channel) $t = 0$ and $T_3 = \mathrm{Diag}(p, p, 2p-1)$. This channel describes decoherence, the decay of a superposition into a mixture;

$$\mathcal{E}\begin{bmatrix} a & b \\ \overline{b} & c \end{bmatrix} = (1-p)\begin{bmatrix} a & b \\ \overline{b} & c \end{bmatrix} + p\begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix}.$$

□

Example 10.15 (Fuchs channel) This channel is not unit preserving and maps σ_2 into 0:

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} \end{bmatrix}$$

It can be shown that the Fuchs channel is an extreme point in the convex set of channels $M_2(\mathbb{C}) \to M_2(\mathbb{C})$. □

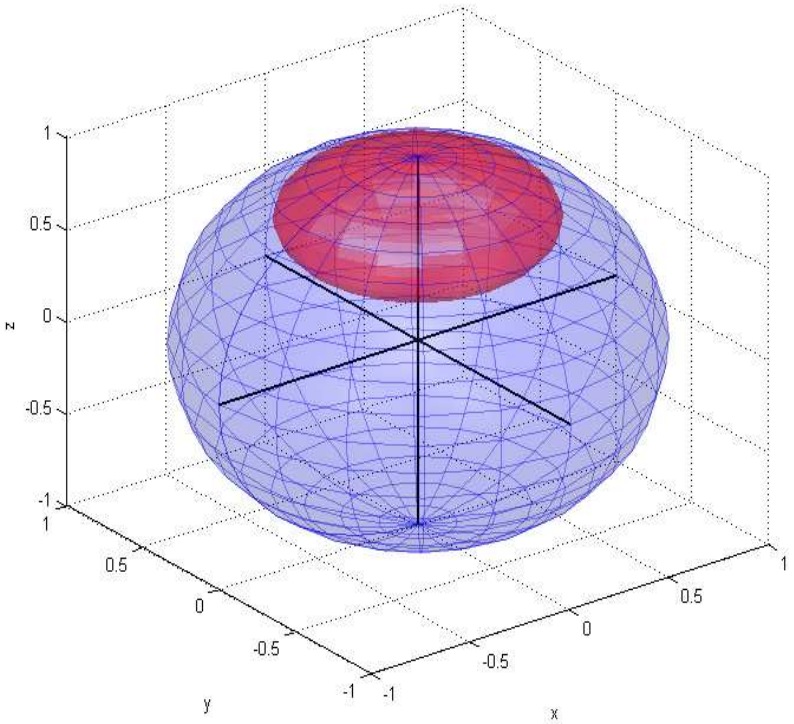
Example 10.16 (Amplitude-damping channel)

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1-p} & 0 & 0 \\ 0 & 0 & \sqrt{1-p} & 0 \\ p & 0 & 0 & 1-p \end{bmatrix}$$

or equivalently

$$\mathcal{E}\begin{bmatrix} a & b \\ \overline{b} & c \end{bmatrix} = \begin{bmatrix} a+pc & \sqrt{1-p}\,b \\ \sqrt{1-p}\,\overline{b} & (1-p)c \end{bmatrix}.$$

The Bloch ball shrinks toward the north pole. □



Amplitude-damping channel

The amplitude-damping channel shrinks the Bloch ball toward the north pole.

Example 10.17 (The Holevo–Werner channel) Set a linear mapping $\mathcal{E} : M_n \to M_n$ as

$$\mathcal{E}(D) = \frac{1}{n-1}(\mathrm{Tr}\,(D)I - D^T),$$

where D^T denotes the transpose of D. The positivity is not obvious from this form but it is easy to show that

$$\mathcal{E}(D) = \frac{1}{2(n-1)}\sum_{i,j}(E_{ij} - E_{ji})^* D(E_{ij} - E_{ji}),$$

where E_{ij} denote the matrix units. This is the Kraus representation of $\mathcal{E}$ which must be completely positive.

In the space of matrices the following matrices are linearly independent.

$$d_k = \mathrm{Diag}(1^{(1)}, 1^{(2)}, \dots, 1^{(n-k)}, -(n-k), 0, 0, \dots, 0).$$

For $k = 0$ we have the unit matrix and d_1, d_2, d_{n-1} are traceless matrices. Moreover, set

$$\begin{aligned} e_{ij} &= E_{ij} - E_{ji} \qquad (1 \le i < j \le n), \\ f_{ij} &= -\mathrm{i}\, E_{ij} + \mathrm{i}\, E_{ji} \qquad (1 \le i < j \le n). \end{aligned}$$

The matrices $\{d_k : 0 \le k \le n-1\} \cup \{e_{ij} : 1 \le i < j \le n\} \cup \{f_{ij} : 1 \le i < j \le n\}$ are pairwise orthogonal with respect to the Hilbert Schmidt inner product and up to a normalizing factor they form a basis in the Hilbert space M_n. (Actually these matrices are a kind of generalization of the Pauli matrices for $n > 2$.).

The mapping $\mathcal{E}$ is unital, hence $\mathcal{E}(d_0) = d_0$. For $0 < k < n$ we have

$$\mathcal{E}(d_k) = \frac{1}{n-1} d_k$$

and for $1 \le i < j \le n$ we have

$$\begin{aligned} \mathcal{E}(E_{ij}) &= E_{ij} \\ \mathcal{E}(F_{ij}) &= -F_{ij}. \end{aligned}$$

Hence our basis consists of eigenvectors, the spectrum of $\mathcal{E}$ is $\{1, -1, \frac{1}{n-1}\}$ with the multiplicities $n(n-1)/2 + 1$, $n(n-1)/2$, $n-2$, respectively. Although $\mathcal{E}$ is completely positive, its spectrum contains negative numbers, therefore it is not true that $\mathcal{E}$ is positive definite with respect to the Hilbert-Schmidt inner product. □

Example 10.18 (Transpose depolarizing channel) Let $\mathcal{E}^T_{p,n} : M_n \to M_n$ is defined as

$$\mathcal{E}^T_{p,n}(A) = tA^T + (1-t)\frac{I}{n}\mathrm{Tr}\, A\,, \tag{10.31}$$

where A^T is the transpose of A. In order to decide the completely positivity, we study the representing block-matrix X:

$$X = t\sum_{ij} E_{ij} \otimes E_{ji} + \frac{1-t}{n} I \otimes I.$$

The matrix $\sum_{ij} E_{ij} \otimes E_{ji}$ is a self-adjoint and its square is the identity. Therefore, its spectrum is $\{\pm 1\}$. Consequently, the eigenvalues of X are

$$-t + \frac{1-p}{n}, p + \frac{1-t}{n}\,.$$

They are positive when

$$-\frac{1}{n-1} \le t \le \frac{1}{n+1}\,. \tag{10.32}$$

This is the necessary and sufficient condition for the complete positivity of $\mathcal{E}^T_{p,n}$. The Holevo-Werner channel is a particular case. □

10.3 Bipartite systems

The physical setting to see entanglement is the **bipartite system** which corresponds to tensor product in mathematical terms. Let $B(\mathcal{H}_A)$ and $B(\mathcal{H}_B)$ be the algebras of bounded operators acting on the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. The Hilbert space of the composite system $A+B$ is $\mathcal{H}_{AB} := \mathcal{H}_A \otimes \mathcal{H}_B$. The algebra of the operators acting on $\mathcal{H}_{AB}$ is $B(\mathcal{H}_{AB}) = B(\mathcal{H}_A) \otimes B(\mathcal{H}_B)$.

Let us recall how to define ordering in a vector space V. A subset $V_+ \subset V$ is called **positive cone** if $v, w \in V_+$ implies $v + w \in V_+$ and $\lambda v \in V_+$ for any positive real λ. Given the positive cone V_+, $f \leq g$ means that $g - f \in V_+$. In the vector space $B(\mathcal{H})$ the standard positive cone is the set of all positive semidefinite matrices. This cone induces the partial ordering

$$A \leq B \iff \langle \eta, A\eta \rangle \leq \langle \eta, B\eta \rangle \quad \text{for every vector} \quad \eta\,.$$

In the product space $B(\mathcal{H}_{AB}) = B(\mathcal{H}_A) \otimes B(\mathcal{H}_B)$, we have two natural positive cones, $B(\mathcal{H}_{AB})^+$ consists of the positive semidefinite matrices acting on $\mathcal{H}_{AB} := \mathcal{H}_A \otimes \mathcal{H}_B$, and the cone $\mathcal{S}$ consists of all operators of the form

$$\sum_i A_i \otimes B_i,$$

where $A_i \in B(\mathcal{H}_A)^+$ and $B_i \in B(\mathcal{H}_B)^+$. It is obvious that $\mathcal{S} \subset B(\mathcal{H}_{AB})^+$. A state is called **separable** (or unentangled) if its density belongs to $\mathcal{S}$. The other states are the **entangled** states. Therefore the set of separable states is the convex hull of product states, or the convex hull of pure product states (since any state is the convex combination of pure states).

A pure state is separable if and only if it is a product state. Indeed, pure states are extreme points in the state space, see Lemma 10.1. If a pure state is convex combination $\sum_i p_i P_i \otimes Q_i$ of product pure states, then this convex combination must be trivial, that is, $P \otimes Q$.

Let $(e_i)_i$ be a basis for $\mathcal{H}_A$ and $(f_j)_j$ be a basis of $\mathcal{H}_B$. Then the doubly indexed family $(e_i \otimes f_j)_{i,j}$ is a basis of $\mathcal{H}_{AB}$. (Such a basis is called **product basis**.) An arbitrary vector $\Psi \in \mathcal{H}_{AB}$ admits an expansion

$$\Psi = \sum_{i,j} c_{ij}\, e_i \otimes f_j \tag{10.33}$$

for some coefficients c_{ij}, $\sum_{i,j} |c_{ij}|^2 = \|\Psi\|^2$.

If $h_i = \sum_j c_{ij}\, f_j$, then $\Psi = \sum_i e_i \otimes h_i$, however, the vectors h_i are not orthogonal. We want to see that a better choice of the representing vectors is possible.

Any unit vector $\Psi \in \mathcal{H}_{AB}$ can be written in the form

$$\Psi = \sum_k \sqrt{p_k}\, g_k \otimes h_k, \tag{10.34}$$

where the vectors $g_k \in \mathcal{H}_A$ and $h_k \in \mathcal{H}_B$ are pairwise orthogonal and normalized, moreover (p_k) is a probability distribution, see Lemma 3.1.

Expansion (10.34) is called **Schmidt decomposition**.

Let $\dim \mathcal{H}_A = \dim \mathcal{H}_B = n$. A pure state $|\Phi\rangle\langle\Phi|$ on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ is called **maximally entangled** if the following equivalent conditions hold:

- The reduced densities are maximally mixed states.
- When the vector $|\Phi\rangle$ is written in the form (10.34), then $p_k = n^{-1}$ for every $1 \leq k \leq n$.
- There is a product basis such that $|\Phi\rangle$ is complementary to it.

The density matrix of a maximally entangled state on $\mathbb{C}^n \otimes \mathbb{C}^n$ is of the form

$$\rho = \frac{1}{n} \sum_{i,j} E_{ij} \otimes E_{ij} \tag{10.35}$$

in an appropriate basis.

A common example of maximally entangled state is the **singlet state**

$$|\Phi\rangle = \frac{1}{\sqrt{2}}(|10\rangle - |01\rangle). \tag{10.36}$$

($\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^2$ which has a basis $\{|0\rangle, |1\rangle\}$.) In the singlet state there is a particular correlation between the two spins, Example 10.6.

It is worthwhile to note that formula (10.34) also shows how to purify an arbitrary density matrix $\rho = \sum_i p_i |g_i\rangle\langle g_i|$ acting on $\mathcal{H}_A$. It is enough to choose an orthonormal family (h_i) in another Hilbert space $\mathcal{H}_B$ and (3.3) gives a pure state whose reduction is ρ. In this case $|\Psi\rangle$ is called the **purification** of ρ.

Example 10.19 Entanglement is a phenomenon appearing in case of two quantum components. If the system is composite but one of the two components is classical, then the possible states are in the form

$$\rho^{cq} = \sum_i p_i |e_i\rangle\langle e_i| \otimes \rho_i^q \,, \tag{10.37}$$

where $(p_i)_i$ is a probability distribution, ρ_i^q are densities on $\mathcal{H}_q$ and $(|e_i\rangle)_i$ is a fixed basis of $\mathcal{H}_c$. (10.37) is in the convex hull of product states, therefore it is separable.

Composite systems of type classical-quantum appear often, for example, the measurement on a quantum system is described in this formalism. □

Let $\mathcal{E} : M_n(\mathbb{C}) \to M_k(\mathbb{C})$ be a linear mapping. According to Theorem 10.3 the condition

$$\sum_{i,j} E_{ij} \otimes \mathcal{E}(E_{ij}) \geq 0 \tag{10.38}$$

is equivalent to the complete positivity of $\mathcal{E}$. Therefore the following is true.

Theorem 10.4 *The linear mapping $\mathcal{E} : M_n(\mathbb{C}) \to M_k(\mathbb{C})$ is completely positive if and only if there exists a maximally entangled state $\rho \in M_n(\mathbb{C}) \otimes M_n(\mathbb{C})$ such that*

$$(id_n \otimes \mathcal{E})(\rho) \geq 0$$

holds.



Entanglement is a very special relation of two quantum systems. The same word means also a different relation in other areas. The sculpture "*Entanglement*" made by Ruth Bloch (bronze, 71 cm, 1995) might express something.

Example 10.20 Let $\mathcal{E} : M_n(\mathbb{C}) \to M_k(\mathbb{C})$ be a channel. $\mathcal{E}$ is called **entanglement breaking** if the range of $\mathrm{id} \otimes \mathcal{E}$ contains only separable states for the identical channel $\mathrm{id} : M_n(\mathbb{C}) \to M_n(\mathbb{C})$. This condition is very restrictive, an entanglement breaking channel has the form

$$\mathcal{E}(\rho) = \sum_i \rho_i \mathrm{Tr}\, S_i \rho$$

where ρ_i is a family of states and S_i are positive matrices such that $\sum_i S_i = I$. See also Exercise 34. □

Theorem 10.5 *If the state $\rho \in \mathcal{A}_{AB}$ is entangled, then there exists $W \in \mathcal{A}_{AB}^{sa}$ such that*

$$\mathrm{Tr}\; W(P \otimes Q) \geq 0$$

for all pure states $P \in \mathcal{A}_A$ and $Q \in \mathcal{A}_B$ but $\mathrm{Tr}\; W\rho < 0$.

Proof: Let $\mathcal{S}$ denote the set of separable states and assume that $\rho \notin \mathcal{S}$. let D_0 be the minimizer of a relative entropy:

$$\inf\{S(\rho \| D_s) : D_s \in \mathcal{S}\} > 0.$$

It is well-known that

$$\frac{\partial}{\partial \varepsilon} \log(X + \varepsilon K) = \int_0^\infty (X+t)^{-1} K (X+t)^{-1}\, dt,$$

and so

$$\frac{\partial}{\partial \varepsilon} S(\rho \| (1-\varepsilon)D_0 + \varepsilon \rho') \;=\; -\mathrm{Tr}\, \rho \int_0^\infty (D_0+t)^{-1}(\rho' - D_0)(D_0+t)^{-1}\, dt$$

$$\begin{aligned} &= -\mathrm{Tr}\,(\rho - D_0)\int_0^\infty (D_0+t)^{-1}\rho'(D_0+t)^{-1}\,dt \\ &= 1 - \mathrm{Tr}\,\rho\int_0^\infty (D_0+t)^{-1}\rho'(D_0+t)^{-1}\,dt, \end{aligned}$$

for any density ρ', both derivatives are taken at $\varepsilon = 0$.

Let

$$W := I - \int_0^\infty (D_0+t)^{-1}\rho(D_0+t)^{-1}\,dt$$

and we have

$$\mathrm{Tr}\,WD_s = \frac{\partial}{\partial\varepsilon}S(\rho\|(1-\varepsilon)D_0+\varepsilon D_s) = \lim_{\varepsilon\to 0}\frac{S(\rho\|(1-\varepsilon)D_0+\varepsilon D_s) - S(\rho\|D_0)}{\varepsilon} \geq 0$$

for an arbitrary $D_s \in \mathcal{S}$, since $(1-\varepsilon)D_0+\varepsilon D_s \in \mathcal{S}$ and $S(\rho\|(1-\varepsilon)D_0+\varepsilon D_s) \geq S(\rho\|D_0)$.

Due to the convexity of the relative entropy we have

$$S(\rho\|(1-\varepsilon)D_0+\varepsilon\rho) - S(\rho\|D_0) \leq -\varepsilon S(\rho\|D_0)\,.$$

Divided by $\varepsilon > 0$ and taking the limit $\varepsilon \to 0$ we arrive at

$$\mathrm{Tr}\,W\rho \leq -S(\rho\|D_0) < 0.$$

□

The operator W appearing in the previous theorem is called **entanglement witness**.

Example 10.21 Let

$$W := \sigma_1\otimes\sigma_1 + \sigma_3\otimes\sigma_3 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

For any product state $\rho_1\otimes\rho_2$, we have

$$\begin{aligned} \mathrm{Tr}\,(\rho_1\otimes\rho_2)W &= \mathrm{Tr}\,\rho_1\sigma_1 \times \mathrm{Tr}\,\rho_2\sigma_1 + \mathrm{Tr}\,\rho_1\sigma_3 \times \mathrm{Tr}\,\rho_2\sigma_3 \\ &\leq \sqrt{(\mathrm{Tr}\,\rho_1\sigma_1)^2 + (\mathrm{Tr}\,\rho_1\sigma_3)^2} \times \sqrt{(\mathrm{Tr}\,\rho_2\sigma_1)^2 + (\mathrm{Tr}\,\rho_2\sigma_3)^2} \\ &\leq 1\,, \end{aligned}$$

since

$$(\mathrm{Tr}\,\rho\sigma_1)^2 + (\mathrm{Tr}\,\rho\sigma_2)^2 + (\mathrm{Tr}\,\rho\sigma_3)^2 \leq 1$$

for any density matrix ρ. It follows that $\mathrm{Tr}\,DW \leq 1$ for any separable state D on $\mathbb{C}^4$.

Consider now the density

$$\omega := \frac{1}{6}\begin{bmatrix} 2 & 0 & 0 & 2 \\ 0 & 1 & x & 0 \\ 0 & x & 1 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}, \tag{10.39}$$

where $0 < x < 1$. Since

$$\mathrm{Tr}\,\omega W = 1 + \frac{x}{3} > 1,$$

ω must be an entangled state. □

Example 10.22 An entanglement witness determines a linear functional that may separate a state from the convex hull of product state. The separation can be done by means of non-linear functionals.

Let ρ be a state and X be an observable. The **variance**

$$\delta^2(X;\rho) := \operatorname{Tr}\rho X^2 - (\operatorname{Tr}\rho X)^2 \tag{10.40}$$

is a concave function of the variable ρ.

Let X_i be a family of observables on the system A and assume that

$$\sum_i \delta^2(X_i;\rho_A) \geq a$$

for every state ρ_A. Similarly, choose observables Y_i on the system B and let

$$\sum_i \delta^2(Y_i;\rho_B) \geq b$$

for every state ρ_B on $\mathcal{H}_B$.

Let ρ_{AB} now be a state on the composite (or bipartite) system $\mathcal{H}_A \otimes \mathcal{H}_B$. The functional

$$\psi(\rho_{AB}) := \sum_i \delta^2(X_i \otimes I + I \otimes Y_i;\rho_{AB}) \tag{10.41}$$

is the sum of convex functionals (of the variable ρ_{AB}), therefore it is convex. For a product state $\rho_{AB} = \rho_A \otimes \rho_B$ we have

$$\sum_i \delta^2(X_i \otimes I + I \otimes Y_i;\rho_{AB}) = \sum_i \delta^2(X_i;\rho_A) + \sum_i \delta^2(Y_i;\rho_B) \geq a + b.$$

It follows that $\psi(\rho_{AB}) \geq a + b$ for every separable state ρ_{AB}. If

$$\psi(\rho_{AB}) < a + b,$$

then the state ρ_{AB} must be entangled. □

Theorem 10.6 *Let ρ_{AB} be a separable state on the bipartite system $\mathcal{H}_A \otimes \mathcal{H}_B$ and let ρ_A be the reduced state. Then ρ_{AB} is more mixed than ρ_A.*

Proof: Let (r_k) be the probability vector of eigenvalues of ρ_{AB} and (q_l) is that for ρ_A. We have to show that there is a doubly stochastic matrix S which transform (q_l) into (r_k).

Let

$$\rho_{AB} = \sum_k r_k |e_k\rangle\langle e_k| = \sum_j p_j |x_j\rangle\langle x_j| \otimes |y_j\rangle\langle y_j|$$

be decompositions of a density matrix in terms of unit vectors $|e_k\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B, |x_j\rangle \in \mathcal{H}_A$ and $|y_j\rangle \in \mathcal{H}_B$. The first decomposition is the Schmidt decomposition and the second one is guaranteed by the assumed separability condition. For the reduced density ρ_A we have the Schmidt decomposition and another one:

$$\rho_A = \sum_l q_l |f_l\rangle\langle f_l| = \sum_j p_j |x_j\rangle\langle x_j|,$$

where f_j is an orthonormal family in $\mathcal{H}_A$. According to Lemma 2.1 we have two unitary matrices V and W such that

$$\begin{aligned}\sum_k V_{kj}\sqrt{p_j}|x_j\rangle\otimes|y_j\rangle &= \sqrt{r_k}|e_k\\ \sum_l W_{jl}\sqrt{q_l}|f_l\rangle &= \sqrt{p_j}|x_j\rangle.\end{aligned}$$

Combine these equations to have

$$\sum_k V_{kj}\sum_l W_{jl}\sqrt{q_l}|f_l\rangle\otimes|y_j\rangle = \sqrt{r_k}|e_k$$

and take the squared norm:

$$r_k = \sum_l \Big(\sum_{j_1,j_2}\overline{V}_{kj_1}V_{kj_2}\overline{W}_{j_1l}W_{j_2l}\langle y_{j_1},y_{j_2}\rangle\Big)q_l$$

Introduce a matrix

$$S_{kl} = \Big(\sum_{j_1,j_2}\overline{V}_{kj_1}V_{kj_2}\overline{W}_{j_1l}W_{j_2l}\langle y_{j_1},y_{j_2}\rangle\Big)$$

and verify that it is doubly stochastic. □

Separable states behaves classically in the sense that the monotonicity of the von Neumann entropy holds.

Corollary 10.1 *Let ρ_{AB} be a separable state on the bipartite system $\mathcal{H}_A\otimes\mathcal{H}_B$ and let ρ_A be the reduced state. Then $S(\rho_{AB})\geq S(\rho_A)$.*

Proof: The statement is an immediate consequence of the theorem, since the von Neumann entropy is monotone with respect to the more mixed relation. However, we give another proof.

First we observe, that for a separable state ρ_{AB}, the operator inequality

$$\rho_{AB}\leq\rho_A\otimes I_B\,. \tag{10.42}$$

holds. Indeed, for a product state the inequality is obvious and we can take convex combinations. Since log is matrix monotone, we have

$$-\log\rho_{AB}\geq-(\log\rho_A)\otimes I_B. \tag{10.43}$$

Taking the expectation values with respect to the state ρ_{AB}, we get $S(\rho_{AB})\geq S(\rho_A)$.

Both proofs show that instead of the von Neumann entropy, we can take an α-entropy as well. □

Theorem 10.7 *Let ρ_{AB} be a state on the bipartite system $\mathcal{H}_A\otimes\mathcal{H}_B$ and let ρ_A be the reduced state. If $S(\rho_{AB})<S(\rho_A)$, then there is $\varepsilon>0$ such that all states ω satisfying the condition $\|\omega-\rho_{AB}\|<\varepsilon$ are entangled.*

Proof: Due to the continuity of the von Neumann entropy $S(\omega)<S(\omega_A)$ holds in a neighborhood of ρ_{AB}. All these states are entangled. □

Theorem 10.8 *$\rho \in \mathcal{A}_{AB}$ is separable if and only if for any $k \in \mathbb{N}$ ρ has a symmetric extension to $\mathcal{A}_A^{\overset{1}{\smile}} \otimes \mathcal{A}_A^{\overset{2}{\smile}} \otimes \ldots \otimes \mathcal{A}_A^{\overset{k}{\smile}} \otimes \mathcal{A}_B$.*

Proof: For a separable state the symmetric extension is easily constructed. Assume that

$$\rho = \sum_i \lambda_i A_i \otimes B_i,$$

then

$$\sum_i \lambda_i A_i \otimes A_i \otimes \ldots \otimes A_i \otimes B_i$$

is a symmetric extension.

Conversely, let ρ_n be the assumed symmetric extension and let the state φ of the infinite product algebra be a limit points of ρ_n's. Since all ρ_n's are extensions of the given ρ, so is φ. According to the **quantum de Finetti theorem** (see Notes), φ is an integral of product state and so is its restriction, ρ. This shows that ρ is separable. □

Theorem 10.9 *Let $\rho \in \mathcal{A}_{AB}$. If there is a positive mapping $\Lambda : \mathcal{A}_B \to \mathcal{A}_B$ such that $(id \otimes \Lambda)\rho$ is not positive, then ρ is entangled.*

Proof: For a product state $\rho_{\mathcal{A}} \otimes \rho_{\mathcal{B}}$ we have

$$(\mathrm{id}_{\mathcal{A}} \otimes \Lambda)(\rho_{\mathcal{A}} \otimes \rho_{\mathcal{B}}) = \rho_{\mathcal{A}} \otimes \Lambda(\rho_{\mathcal{B}}) \geq 0\,.$$

It follows that $(\mathrm{id}_{\mathcal{A}} \otimes \Lambda)D$ is positive when D is separable. □

In place of Λ, there is no use of completely positive mapping but matrix transposition (in any basis) could be useful.

Example 10.23 Consider the state

$$\frac{1}{4}\begin{bmatrix} 1+p & 0 & 1-p & p \\ 0 & 1-p & 0 & 1-p \\ 1-p & 0 & 1-p & 0 \\ p & 1-p & 0 & 1+p \end{bmatrix},$$

where $0 \leq p \leq 1$. The partial transpose of this matrix is

$$\frac{1}{4}\begin{bmatrix} 1+p & 0 & 1-p & 0 \\ 0 & 1-p & p & 1-p \\ 1-p & p & 1-p & 0 \\ 0 & 1-p & 0 & 1+p \end{bmatrix}.$$

If this is positive, so is

$$\begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}.$$

For $1/2 < p$ this matrix is not positive, therefore Theorem 10.9 tells us that the state is entangled for these values of the parameter. □

Example 10.24 Let $|\Psi\rangle\langle\Psi| \in \mathbb{C}^2 \otimes \mathbb{C}^2$ be a maximally entangled state and τ be the tracial state. Since τ is separable

$$\rho_p := p|\Psi\rangle\langle\Psi| + (1-p)\tau \qquad (0 \le p \le 1) \tag{10.44}$$

is an interpolation between an entangled state and a separable state. ρ_p is called **Werner state**, its eigenvalues are

$$p + \frac{1-p}{4}, \frac{1-p}{4}, \frac{1-p}{4}, \frac{1-p}{4} \tag{10.45}$$

and the eigenvalues of the reduced density matrix are $(1/2, 1/2)$. (10.45) is more mixed than this pair if and only if $p \le 1/3$. Therefore, for $p > 1/3$, the state ρ_p must be entangled due to Theorem 10.6.

We can arrive at the same conclusion also from Theorem 10.9. In an appropriate basis, the matrix of ρ_p is

$$\frac{1}{4}\begin{pmatrix} 1-p & 0 & 0 & 0 \\ 0 & 1+p & 2p & 0 \\ 0 & 2p & 1+p & 0 \\ 0 & 0 & 0 & 1-p \end{pmatrix}. \tag{10.46}$$

The partial transpose of this matrix is

$$\frac{1}{4}\begin{pmatrix} 1-p & 0 & 0 & 2p \\ 0 & 1+p & 0 & 0 \\ 0 & 0 & 1+p & 0 \\ 2p & 0 & 0 & 1-p \end{pmatrix} \tag{10.47}$$

which cannot be positive when $(1-p)^2 < 4p^2$. For $p > 1/3$ this is the case and Theorem 10.9 tells us that ρ_p is entangled.

If $p = 1/3$, then ρ_p is

$$\frac{1}{6}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

and we want to show that this is separable by presenting a decomposition. We have

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Here the first summand has a decomposition

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & \varepsilon \\ \varepsilon^2 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & \varepsilon \\ \varepsilon^2 & 1 \end{pmatrix} + \begin{pmatrix} 1 & \varepsilon^2 \\ \varepsilon & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & \varepsilon^2 \\ \varepsilon & 1 \end{pmatrix},$$

where $\varepsilon := \exp(2\pi\mathrm{i}/3)$ and the second summand is

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Since separable states form a convex set, we conclude that ρ_p is separable for every $p \le 1/3$. □

10.4 Dense coding and teleportation

Quantum information and classical information are very different concepts and strictly speaking, it has no meaning to compare them. However, transmission of a single qubit can carry two bits of classical information and transmitting classical information of two bits can yield the teleportation of the state of a quantum spin. From this point of view a qubit is equivalent to two classical bits. Both protocols are based on a basis consisting of maximally entangled states on the 4 dimensional space.

The **Bell basis** of $\mathbb{C}^2 \otimes \mathbb{C}^2$ consists of the following vectors

$$\begin{aligned} |\beta_0\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \\ |\beta_1\rangle &= \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle) = (\sigma_1 \otimes I)|\beta_0\rangle, \\ |\beta_2\rangle &= \frac{\mathrm{i}}{\sqrt{2}}(|10\rangle - |01\rangle) = (\sigma_2 \otimes I)|\beta_0\rangle, \\ |\beta_3\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) = (\sigma_3 \otimes I)|\beta_0\rangle. \end{aligned}$$

All of them give maximally entangled states of the bipartite system.

Assume that Alice wants to communicate an element of the set $\{0, 1, 2, 3\}$ to Bob and both of them have a spin. Assume that the two spins are initially in the state $|\beta_0\rangle$. Alice and Bob may follow the following protocol called **dense coding**

1. If the number to communicate to Bob is k, Alice applies the unitary σ_k to her spin. After this the joint state of the two spins will be the kth vector of the Bell basis.

2. Alice sends her qubit to Bob and Bob will be in the possession of both spins.

3. Bob performs the measurement corresponding to the Bell basis, and the outcome will exactly be k.

Next we turn to the **teleportation protocol** initiated by Bennett at al. in 1993 [9]. Consider a 3-qubit system in the initial state

$$|\psi\rangle_A \otimes |\beta_0\rangle_{XB}.$$

(So the spin A is statistically independent from the other two spins.) Assume that Alice is in a possession of the qubits A and X and the spin B is at Bob's disposal. The aim is to convert Bob's spin into the state $|\psi\rangle$. Alice and Bob are separated, they could be far away from each other but they can communicate in a classical channel. How can this task be accomplished?

1. Alice measures the Bell basis on the spins A and X. The outcome of the measurement is an element of the set $\{0, 1, 2, 3\}$.

2. Alice communicates this outcome to Bob in a classical communication channel.This requires the transmission of two classical bits to distinguish among $0, 1, 2, 3$.

3. Bob applies the unitary σ_k to the state vector of spin B if the message of Alice is "k".

Then the state of spin B is the same as the state of spin A was at the beginning of the procedure.

This protocol depends on an important identity

$$|\psi\rangle_A \otimes |\beta_0\rangle_{XB} = \frac{1}{2}\sum_{k=0}^{3} |\beta_k\rangle_{AX} \otimes \sigma_k|\psi\rangle_B. \tag{10.48}$$

The measurement of Alice is described by the projections $E_i := |\beta_i\rangle\langle\beta_i| \otimes I_B$ $(0 \le i \le 3)$. The outcome k appears with probability

$$\langle \eta, (|\beta_k\rangle\langle\beta_k| \otimes I_B)\eta\rangle,$$

where η is the vector (10.48) and this is $1/4$. If the measurement gives the value k, then after the measurement the new state vector is

$$\frac{E_k\eta}{\|E_k\eta\|} = |\beta_k\rangle_{AX} \otimes \sigma_k|\psi\rangle_B.$$

When Bob applies the unitary σ_k to the state vector $\sigma_k|\psi\rangle_B$ of his spin, he really gets $|\psi\rangle_B$.

There are a few important features concerning the protocol. The actions of Alice and Bob are local, they manipulate only the spins at their disposal and they act independently of the unknown spin X. It is also important to observe that the spin A changes immediately after Alice' measurement. If this were not the case, then the procedure could be used to copy (or to clone) a quantum state which is impossible, Wooters and Zurek argued for a "no-cloning" theorem [53]. Another price the two parties have to pay for the teleportation is the entanglement. The state of AX and B becomes separable.

The identity (10.48) implies that

$$E_k\Big(|\psi\rangle\langle\psi|_A \otimes |\beta_0\rangle\langle\beta_0|_{XB}\Big)E_k = \frac{1}{4}|\beta_0\rangle\langle\beta_0|_{AX} \otimes (\sigma_k|\psi\rangle\langle\psi|_B\sigma_k)\,.$$

Both sides are linear in $|\psi\rangle\langle\psi|$ which can be replaced by an arbitrary density matrix ρ:

$$E_k\Big(\rho \otimes |\beta_0\rangle\langle\beta_0|_{XB}\Big)E_k = \frac{1}{4}|\beta_0\rangle\langle\beta_0|_{AX} \otimes (\sigma_k\rho_B\sigma_k), \tag{10.49}$$

This formula shows that the teleportation protocol works for a density matrix ρ as well. The only modification is that when Bob receives the information that the outcome of Alice's measurement has been k, he must perform the transformation $D \mapsto \sigma_k D\sigma_k$ on his spin.

We can generalize the above protocol. Assume that $\mathcal{H}$ is n-dimensional and we have unitaries U_i $(1 \le i \le n^2)$ such that

$$\operatorname{Tr} U_i^* U_j = 0 \quad \text{if} \quad i \ne j\,. \tag{10.50}$$

These orthogonality relations guarantee that the operators U_i are linearly independent and they must span the linear space of all matrices.

Let $\Phi \in \mathcal{H} \otimes \mathcal{H}$ be a maximally entangled state vector and set

$$\Phi_i := (U_i \otimes I)\Phi \qquad (1 \leq i \leq n^2).$$

One can check that $(\Phi_i)_i$ is a basis in $\mathcal{H} \otimes \mathcal{H}$:

$$\begin{aligned} \langle (U_i \otimes I)\Phi, (U_j \otimes I)\Phi \rangle &= \langle \Phi, (U_i^* U_j \otimes I)\Phi \rangle = \operatorname{Tr} (U_i^* U_j \otimes I)|\Phi\rangle\langle\Phi| \\ &= \operatorname{Tr} \operatorname{Tr}_2 \Big(U_i^* U_j \otimes I)|\Phi\rangle\langle\Phi| \Big) = \operatorname{Tr} (U_i^* U_j) \operatorname{Tr}_2 |\Phi\rangle\langle\Phi| \\ &= \frac{1}{n} \operatorname{Tr} U_i^* U_j \,. \end{aligned}$$

Consider 3 quantum systems, similarly to the spin-$\frac{1}{2}$ case, assume that the system X and A are localized at Alice and B is at Bob. Each of these n-level systems are described an n dimensional Hilbert space $\mathcal{H}$. Let the initial state be

$$\rho_A \otimes |\Phi\rangle\langle\Phi|_{XB}.$$

The density ρ is to be teleported from Alice to Bob by the following protocol:

1. Alice measures the basis $(\Phi_i)_i$ on the quantum system $A + X$. The outcome of the measurement is an element of the set $\{1, 2, 3, \ldots, n^2\}$.

2. Alice communicates this outcome to Bob in a classical communication channel.

3. Bob applies the state transformation $D \mapsto U_k D U_k^*$ to his quantum system B if the message of Alice is "k" $(1 \leq k \leq n^2)$.

The measurement of Alice is described by the projections $E_i := |\Phi_i\rangle\langle\Phi_i| \otimes I_B$ $(1 \leq i \leq n^2)$. The state transformation $D \mapsto U_k D U_k^*$ corresponds to the transformation $A \mapsto U_k^* A U_k$, hence to show that the protocol works we need

$$\sum_k \operatorname{Tr} E_k(\rho \otimes |\Phi\rangle\langle\Phi|) E_k(I_{AX} \otimes U_k A U_k^*) = \operatorname{Tr} \rho A \tag{10.51}$$

for all $A \in B(\mathcal{H})$. Indeed, the left-hand-side is the expectation value of A after teleportation, while the right-hand-side is the expectation value in the state ρ.

Since this equation is linear both in ρ and in A, we may assume that $\rho = |\phi\rangle\langle\phi|$ and $A = |\psi\rangle\langle\psi|$. Then the right-hand-side is $|\langle\psi, \phi\rangle|^2$ and the left-hand-side is

$$\sum_k |\langle \phi \otimes \Phi, (U_k \otimes I)\Phi \otimes U_k^* \psi \rangle|^2 = \sum_k |\langle U_k^* \phi \otimes \Phi, \Phi \otimes U_k^* \psi \rangle|^2 \,.$$

Since $\langle \eta_1 \otimes \Phi, \Phi \otimes \eta_2 \rangle = n^{-1} \langle \eta_1, \eta_2 \rangle$, the above expression equals to,

$$\sum_k |\langle U_k^* \phi \otimes \Phi, \Phi \otimes U_k^* \psi \rangle|^2 = \frac{1}{n^2} \sum_k |\langle U_k^* \phi, U_k^* \psi \rangle|^2 = |\langle \phi, \psi \rangle|^2 \,.$$

This proves (10.51) which can be written more abstractly:

$$\sum_k \operatorname{Tr}(\rho \otimes \omega)(E_k \otimes T_k(A)) = \operatorname{Tr} \rho A \tag{10.52}$$

for all $\rho, A \in B(\mathcal{H})$, where E_k is a von Neumann measurement on $\mathcal{H} \otimes \mathcal{H}$ and $T_k : B(\mathcal{H}) \to B(\mathcal{H})$ is a noiseless channel, $T_k(A) = U_k A U_k^*$ for some unitary U_k. In this style, the dense coding is the equation

$$\operatorname{Tr} \omega(T_k \otimes \mathrm{id}) E_\ell = \delta_{k,\ell}. \tag{10.53}$$

Recall that in the teleportation protocol we had a maximally entangled state $\omega = |\Phi\rangle\langle\Phi|$, a basis U_k of unitaries which determined the measurement

$$E_k = |\Phi_k\rangle\langle\Phi_k|, \qquad \Phi_i := (U_i \otimes I)\Phi$$

and the channels $T_k(A) = U_k A U_k^*$. These objects satisfy equation (10.53) as well, so the dense coding protocol works on n level systems.

Next we see how to find unitaries satisfying the orthogonality relation (10.50).

Example 10.25 Let $e_0, e_1, \ldots, e_{n-1}$ be a basis in the n-dimensional Hilbert space $\mathcal{H}$ and let X be the unitary operator permuting the basis vectors cyclically:

$$Xe_i = \begin{cases} e_{i+1} & \text{if } 0 \le i \le n-2, \\ e_0 & \text{if } i = n-1. \end{cases}$$

Let $q := e^{\mathrm{i}2\pi/n}$ and define another unitary by $Ye_i = q^i e_i$. It is easy to check that $YX = qXY$ or more generally the commutation relation

$$Z^k X^\ell = q^{k\ell} X^\ell Z^k \tag{10.54}$$

is satisfied. For $S_{j,k} = Z^j X^k$, we have

$$S_{j,k} = \sum_{m=0}^{n-1} q^{mj} |e_m\rangle\langle e_{m+k}| \quad \text{and} \quad S_{j,k} S_{u,v} = q^{ku} S_{j+u,k+v},$$

where the additions $m+k, j+u, k+v$ are understood modulo n. Since $\operatorname{Tr} S_{j,k} = 0$ when at least one of j and k is not zero, the unitaries

$$\{S_{j,k} \,:\, 0 \le j, k \le n-1\}$$

are pairwise orthogonal.

Note that $S_{j,k}$ and $S_{u,v}$ commute if $ku = jv \bmod n$. These unitaries satisfy a discrete form of the **Weyl commutation relation** and the case $n = 2$ simply reduces to the Pauli matrices: $X = \sigma_1$ and $Z = \sigma_3$. (This fact motivated our notation.) □



John von Neumann (1903–1957)

David Hilbert invited von Neumann in 1927 to Göttingen. That time Heisenberg lectured about the new field called quantum theory. Von Neumann transformed the physical things into a mathematical formalism. Therefore his book *Mathematische Grundlagen der Quantenmechanik* (published in 1932) contained the basic mathematical approach.

10.5 Notes and remarks

There are several books about the mathematical foundations of quantum mechanics. The book of von Neumann [39] has a historical value, it was published in 1932. Holevo's lecture note [27] is rather comprehensive and [13] treats unbounded linear operators of Hilbert spaces in details. My book [47] is close to the material of this chapter.

Theorem 10.6 was proved by M. A. Nielsen and J. Kempe, see Separable states are more disordered globally than locally, Phys. Rev. Lett., **86**, 5184-7 (2001).

The quantum **de Finetti theorem** is about states of the infinite product

$$\mathcal{A} \otimes \mathcal{B} \otimes \mathcal{B} \otimes \dots$$

which remain invariant under the fine permutations of the factors $\mathcal{B}$. Such states are called **symmetric**. The theorem obtained from the paper M. Fannes, J. T. Lewis and A. Verbeure, Symmetric states of composite systems, Lett. Math. Phys. **15**(1988), 255–260. We have that symmetric states are in the closed convex hull of symmetric product states. (For other generalization of the de Finetti theorem, see [40].)

Example 10.22 is from the paper H. F. Hofmann and S. Takeuchi, Violation of local uncertainty relations as a signature of entanglement, Phys. Rev. A **68**(2003), 032103.

The unitaries S_{jk} in Example 10.25 give a discrete form of the **Weyl commutation relation**

$$U(h_1)U(h_2) = U(h_1 + h_2) \exp\left(\mathrm{i}\,\mathrm{Im}\,\langle h_1, h_2\rangle\right),$$

where the unitary family is labeled by the vectors of a Hilbert space [46]. The construction in Example 10.25 is due to Schwinger, Unitary operator basis, Proc. Nat. Acad. Sci. USA **46**(1960), 570–579.

The bases (e_i) and (f_j) has the property

$$|\langle e_i, f_j\rangle|^2 = \frac{1}{n}.$$

Such bases are called **complementary** or **unbiased**. They appeared in connection with the uncertainty relation, it is due to K. Kraus, see Chap. 16 of [40]. A family of mutually unbiased bases on an n-dimensional space has cardinality at most $n+1$. It is not known if the bound is reachable for any n. (It is easy to construct $n+1$ mutually unbiased bases if $n = 2^k$.) More details about complementarity are also in the paper D. Petz, Algebraic complementarity in quantum theory, J. Math. Phys. **51**, 015215 (2010).

10.6 Exercises

1. Show that the vectors $|x_1\rangle, |x_2,\rangle, \dots, |x_n\rangle$ form an orthonormal basis in an n-dimensional Hilbert space if and only if
$$\sum_i |x_i\rangle\langle x_i| = I.$$

2. Express the Pauli matrices in terms of the ket vectors $|0\rangle$ and $|1\rangle$.

3. Show that the Pauli matrices are unitarily equivalent.

4. Show that for a 2×2 matrix A the relation
$$\frac{1}{2}\sum_{i=0}^{3}\sigma_i A\sigma_i = (\mathrm{Tr}\, A)I$$
holds.

5. Let t be a real number and n be a unit vector in $\mathbb{R}^3$. Show that
$$\exp(\mathrm{i}tn\cdot\sigma) = \cos t\,(n\cdot\sigma) + \mathrm{i}\sin t\,(n\cdot\sigma).$$

6. Let v and w be complex numbers and let n be a unit vector in $\mathbb{R}^3$. Show that
$$\exp(v\sigma_0 + w(n\cdot\sigma)) = e^v\left((\cosh w)\,\sigma_0 + (\sinh w)\,n\cdot\sigma\right). \tag{10.55}$$

7. Let $|1\rangle, |2\rangle, |3\rangle$ be a basis in $\mathbb{C}^3$ and
$$\begin{aligned}
|v_1\rangle &= \frac{1}{\sqrt{3}}(|1\rangle + |2\rangle + |3\rangle),\\
|v_2\rangle &= \frac{1}{\sqrt{3}}(|1\rangle - |2\rangle - |3\rangle),\\
|v_3\rangle &= \frac{1}{\sqrt{3}}(|1\rangle - |2\rangle + |3\rangle),\\
|v_4\rangle &= \frac{1}{\sqrt{3}}(|1\rangle + |2\rangle - |3\rangle).
\end{aligned}$$

Compute $\sum_{i=1}^{4} |v_i\rangle\langle v_i|$.

8. Show the identity

$$(I_2 - \sigma_k) \otimes (I_2 + \sigma_k) + (I_2 + \sigma_k) \otimes (I_2 - \sigma_k) = I_4 - \sigma_k \otimes \sigma_k \tag{10.56}$$

for $k = 1, 2, 3$.

9. Let e and f be unit vectors in $\mathbb{C}^2$ and assume that they are eigenvectors of two different Pauli matrices. Show that

$$|\langle e, f\rangle|^2 = \frac{1}{2}.$$

10. Consider two complementary orthonormal bases in a Hilbert space. Let A and B operators such that $\operatorname{Tr} A = \operatorname{Tr} B = 0$, A is diagonal in the first basis, while B is diagonal in the second one. Show that $\operatorname{Tr} AB = 0$.

11. Show that the number of pairwise **complementary orthonormal bases** in an n-dimensional Hilbert space is at most $n + 1$. (Hint: Estimate the dimension of the subspace of traceless operators.)

12. Show that the quantum Fourier transform moves the standard basis to a complementary basis.

13. What is the 8×8 matrix of the controlled-swap gate? (This unitary is called **Fredkin gate**.)

14. Give the density matrix corresponding to the singlet state (10.8) and compute the reduced density matrices.

15. Let ρ be a density matrix. Show that ρ corresponds to a pure state if and only if $\rho^2 = \rho$.

16. Compute the dimension of the set of the extreme points and the dimension of the topological boundary of the $n \times n$ density matrices.

17. Compute the reduced density matrices of the state

$$\frac{1}{3}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

18. Let $0 < p < 1$. Show that the Kraus representation of the depolarizing channel on M_2 is

$$\mathcal{E}_{p,2}(A) = \frac{3p+1}{4}A + \frac{1-p}{4}\sigma_1 A\sigma_1 + \frac{1-p}{4}\sigma_2 A\sigma_2 + \frac{1-p}{4}\sigma_3 A\sigma_3.$$

19. Assume that $\mathcal{E} : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ is defined as

$$\mathcal{E}(A) = \frac{1}{n-1}(I \operatorname{Tr} A - A).$$

Show that $\mathcal{E}$ is positive but not completely positive.

20. Let p be a real number. Show that the mapping $\mathcal{E}_{p,2} : M_2 \to M_2$ is defined as

$$\mathcal{E}_{p,2}(A) = pA + (1-p)\frac{I}{2}\operatorname{Tr} A$$

is positive if and only if $-1 \leq p \leq 1$. Show that $\mathcal{E}_{p,2}$ is completely positive if and only if $-1/3 \leq p \leq 1$. (Hint: $\mathcal{E}_{p,2}$ is a Pauli channel.)

21. Let $\mathcal{E}_{p,2} : M_2 \to M_2$ be the depolarizing channel and denote by A^T the transpose of A. For which values of the parameter p will be $A \mapsto \mathcal{E}_{p,2}(A)^T$ a state transformation?

22. What is the spectrum of the linear mapping $\mathcal{E}_{p,2}$? May a positive mapping have negative eigenvalues?

23. Give the Kraus representation of the phase-damping channel.

24. Show that

$$\frac{1}{3}\begin{bmatrix} 5 & 0 & 0 & \sqrt{3} \\ 0 & 1 & \mathrm{i}\sqrt{3} & 0 \\ 0 & -\mathrm{i}\sqrt{3} & 3 & 0 \\ \sqrt{3} & 0 & 0 & 3 \end{bmatrix}$$

is the representing block-matrix of the Fuchs channel.

25. Show that the matrix

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\delta & 0 & 0 \\ 0 & 0 & \cos\gamma & 0 \\ \sin\gamma\sin\delta & 0 & 0 & \cos\gamma\cos\delta \end{bmatrix}$$

determines a state transformation of a qubit.

26. Compute the limit of $\mathcal{E}^n$ if $\mathcal{E}$ is the amplitude-damping channel.

27. Assume that $\mathcal{E} : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ acts as

$$\mathcal{E}(A)_{ij} = \delta_{ij} A_{ij},$$

that is, $\mathcal{E}$ kills all off-diagonal entries. Find the Kraus representation of $\mathcal{E}$.

28. Let $e_1, e_2, \ldots, e_n$ be a basis in the Hilbert space $\mathcal{H}$. Show that the vector

$$\frac{1}{\sqrt{n}} \sum_{i,j=1}^{n} U_{ij} e_i \otimes e_j$$

gives a maximally entangled state in $\mathcal{H} \otimes \mathcal{H}$ if and only if the matrix $(U_{ij})_{i,j=1}^n$ is a unitary.

29. Let Φ be a unit vector in $\mathcal{H} \otimes \mathcal{H}$ and let n be the dimension of $\mathcal{H}$. Show that Φ gives a maximally entangled state if and only if

$$\langle \eta_1 \otimes \Phi, \Phi \otimes \eta_2 \rangle = n^{-1} \langle \eta_1, \eta_2 \rangle$$

for every vector $\eta_1, \eta_2 \in \mathcal{H}$.

30. Let Φ be a maximally entangled state in $\mathcal{H} \otimes \mathcal{H}$ and W be a unitary on $\mathcal{H}$. Show that $(W \otimes I)\Phi$ gives a maximally entangled state.

31. Use the partial transposition and Theorem 10.9 to show that the density (10.39) is entangled.

32. Show that the matrix of an operator on $\mathbb{C}^2 \otimes \mathbb{C}^2$ is diagonal in the Bell basis if and only if it is a linear combination of the operators $\sigma_i \otimes \sigma_i$, $0 \leq i \leq 3$.

33. Use Theorem 10.6 to show that the density (10.39) is entangled.

34. Let $\mathcal{E} : M_n(\mathbb{C}) \to M_k(\mathbb{C})$ be a channel and assume that

$$\mathrm{id} \otimes \mathcal{E}(|\Phi\rangle\langle\Phi|)$$

is separable for a maximally entangled state $|\Phi\rangle\langle\Phi|$ in $M_n(\mathbb{C}) \otimes M_n(\mathbb{C})$, where id_n is the identity $M_n(\mathbb{C}) \to M_n(\mathbb{C})$. Show that $\mathcal{E}$ has the form

$$\mathcal{E}(\rho) = \sum_i \rho_i \mathrm{Tr}\, S_i \rho \tag{10.57}$$

where ρ_i is a family of states and S_i are positive matrices such that $\sum_i S_i = I$.

35. Show that

$$\frac{1}{12} \sum_{k=1}^{3} (I_4 - \sigma_k \otimes \sigma_k)$$

is a Werner state. Use identity (10.56) to show that it is separable.

36. Show that the range of $\mathcal{E}_{p,2} \otimes \mathcal{E}_{p,2}$ does not contain an entangled state if $\mathcal{E}_{p,2} : M_2(\mathbb{C}) \to M_2(\mathbb{C})$ is the depolarizing channel and $0 \leq p \leq 1/2$.

37. Assume that three qubits are in the pure state

$$\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$$

(called GHZ state, named after **Greeneberger-Horne-Zeilinger**). Show that all qubits are in a maximally mixed state and any two qubits are in a separable state.

38. Let $|\Psi\rangle\langle\Psi| \in B(\mathbb{C}^n \otimes \mathbb{C}^n)$ be a maximally entangled state ($n \geq 3$) and assume that for another state ρ we have $\| \, |\Psi\rangle\langle\Psi| - \rho \, \|_1 \leq 1/3$. Show that ρ is entangled.

39. Let $|\Psi\rangle\langle\Psi| \in B(\mathbb{C}^n \otimes \mathbb{C}^n)$ be a maximally entangled state and τ be the tracial state on $B(\mathbb{C}^n \otimes \mathbb{C}^n)$. When will be the **Werner state**

$$\rho_p := p|\Psi\rangle\langle\Psi| + (1-p)\tau \qquad (0 \leq p \leq 1) \tag{10.58}$$

entangled?

40. Compute the state of the spin X after the teleportation procedure between Alice and Bob.

41. Compute the joint state of the spins A and B after the teleportation procedure.

42. Show that relation (10.50) implies the condition

$$\sum_i U_i^* A U_i = n(\operatorname{Tr} A) I$$

for every $A \in B(\mathcal{H})$.

Index

adjoint operator, 12
annihilating
    polynomial, 18

bases
    complementary, 144
    mutually unbiased, 144
basis, 8
    Bell, 30, 123, 139
    product, 28, 131
bilinear form, 13
bipartite system, 131
Bloch
    ball, 116
    sphere, 114
block-matrix
    representation, 125
Boltzmann entropy, 40, 80
bra and ket, 10

Cayley transform, 16
channel
    amplitude-damping, 129
    depolarizing, 127
    entanglement breaking, 133
    Fuchs, 128
    phase-damping, 128
    transpose depolarizing, 130
    Werner-Holevo, 129
Choi, 125
commutation relation
    Weyl, 142, 143
complementary
    bases, 118, 145
    vector, 118
completely positive, 44, 124
composite system, 118
concave
    jointly, 95
conjecture
    BMV, 67
conjugate
    convex function, 92
contraction, 11
convex
    function, 90
    hal, 90
    set, 89
Cramer's rule, 26

decomposition
    polar, 38
    Schmidt, 21
    spectral, 21
dense coding, 139
density
    matrix, 115
determinant, 6
distinguished with certainty, 117
divided difference, 90

eigenvector, 19
entangled, 42
entangled state, 131
entanglement
    breaking channel, 133
    witness, 134
entropy
    Boltzmann, 40, 80
    quasi, 110
    Rényi, 69
    von Neumann, 64
environment, 123
estimator, 78

factorization
    Schur, 72
    UL-, 75
formula
    Lie-Trotter, 52
Fourier expansion, 9

gate, 122
  controlled-NOT, 122
  Fredkin, 145
  Hadamard, 122
Gaussian distribution, 39
geodesic, 81
geometric mean, 107
Gram-Schmidt procedure, 8

Hadamard
  gate, 122
  product, 44, 126
Hamiltonian, 121
Heinz mean, 108
Hilbert space, 5

inequality
  Cramér-Rao, 76
  Golden-Thompson-Lieb, 54
  Hadamard, 41, 95
  Jensen, 91
  Löwner-Heinz, 104
  Poincaré, 23
  Schwarz, 7
  Streater, 60
  transformer, 107
inner product, 6
  Hilbert–Schmidt, 8
inverse, 6, 26

Jensen inequality, 91
Jordan block, 18

kernel, 46
  positive definite, 46
Kraus representation, 125
Kronecker
  product, 31
  sum, 31

Lagrange, 19
Laplace transform, 56
Legendre transform, 92

matrix
  bias, 78
  covariance, 76
  Dirac, 35
  Fisher information, 76
  Pauli, 27, 52
  permutation, 13
  Toeplitz, 13
  tridiagonal, 19
  upper triangular, 10
matrix-unit, 5
maximally entangled, 42
mean
  geometric, 82, 107
  harmonic, 106
  Heinz, 108
  logarithmic, 108
measurement, 117
minimax principle, 23
mixed states, 115

Neumann series, 12
norm, 7
  Hilbert–Schmidt, 8
  operator, 11

operation elements, 124
operator
  conjugate linear, 13
  convex function, 96
  mean, 106
  positive, 37
  self-adjoint, 12
operator-sum representation, 124
orthogonal, 117
orthogonality, 8
orthonormal, 8

parallel sum, 105
parallelogram law, 14
partial trace, 31, 54, 121
Pauli matrices, 116
Pauli matrix, 27, 52
permanent, 34
phase, 114
polar decomposition, 38
polarization identity, 14
positive
  cone, 131
  matrix, 37
projection, 45
pure state, 114
purification, 132

quantum, 114
  de Finetti theorem, 137, 143
  Fourier transform, 122
quasi-entropy, 110

Rényi entropy, 69
reduced
  density matrix, 120
relative entropy, 55, 60
Riemannian manifold, 80

Schmidt decomposition, 21, 132
Schrödinger, 21
Schrödinger picture, 122
Schur
  complement, 73, 106
  factorization, 72
self-adjoint operator, 115
separable
  positive matrix, 41
separable state, 131
singlet state, 132
singular value, 38
spectral decomposition, 21, 116
spectrum, 19
state
  entangled, 131
  Greeneberger-Horne-Zeilinger, 147
  maximally entangled, 132
  separable, 131
  singlet, 132
  symmetric, 143
  transformation, 123
  Werner, 138, 147
Streater inequality, 60
strong subadditivity, 54, 55

teleportation, 139
tensor product, 28
theorem
  Bernstein, 56
  Cramér-Rao, 77
  ergodic, 27
  Jordan canonical, 18
  Löwner, 103
  Lieb's concavity, 111
  quantum de Finetti, 137, 143
  Riesz-Fischer, 11
  Schoenberg, 47
  Schur, 44, 49
  Weyl's monotonicity, 43
time development, 121
trace, 6
transpose, 6

unitary, 13
unitary propagator, 121

variance, 135
vector
  cyclic, 19
von Neumann entropy, 64

weakly positive matrix, 41
Werner state, 138, 147
Weyl, 142
Wigner, 48

Bibliography

[1] T. Ando, Generalized Schur complements, Linear Algebra Appl. **27**(1979), 173–186.

[2] T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard products, Linear Algebra Appl. **26**(1979), 203–241.

[3] T. Ando, Totally positive matrices, Linear Algebra Appl. **90**(1987), 165–219.

[4] T. Ando, private communication, 2009.

[5] T. Ando, C-K. Li and R. Mathias, Geometric means, Linear Algebra Appl. **385**(2004), 305–334.

[6] T. Ando and D. Petz, Gaussian Markov triplets approached by block matrices, Acta Sci. (Szeged) **75**(2009), 265–281.

[7] R. Bellman, *Introduction to matrix analysis*, McGraw-Hill, 1960.

[8] J. Bendat and S. Sherman, Monotone and convex operator functions. Trans. Amer. Math. Soc. **79**(1955), 58–71.

[9] C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres and W. Wootters, Teleporting an unknown quantum state via dual classical and EPR channels, Physical Review Letters, **70**(1993), 1895–1899.

[10] D. Bessis, P. Moussa and M. Villani, Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics, J. Mathematical Phys. **16**(1975), 2318–2325.

[11] R. Bhatia, *Matrix Analysis*, Springer, 1997.

[12] G. Birkhoff, Tres observaciones sobre el algebra lineal, *Univ. Nac. Tucuman Rev. Ser. A* **5** (1946), 147–151.

[13] J. Blank, P. Exner and M. Havliček, *Hilbert space operators in quantum physics*, American Institute of Physics, 1994.

[14] M. D. Choi, Completely positive mappings on complex matrices, Linear Algebra Appl. **10**(1977), 285–290.

[15] J. B. Conway, *Functions of One Complex Variable I*, Second edition Springer-Verlag, New York-Berlin, 1978.

[16] I. Csiszár, Information type measure of difference of probability distributions and indirect observations, Studia Sci. Math. Hungar. **2**(1967), 299–318.

[17] W. F. Donoghue, Jr., *Monotone Matrix Functions and Analytic Continuation*, Springer-Verlag, Berlin-Heidelberg-New York, 1974.

[18] W. Feller, *An introduction to probability theory with its applications*, vol. II.

[19] Freud Róbert, *Lineáris algebra*, ELTE Eötvös Kiadó, 2001.

[20] F. Hansen and G. K. Pedersen, Jensen's inequality for operators and Löwner's theorem, *Math. Ann.* **258** (1982), 229–241.

[21] F. Hansen, Metric adjusted skew information, 2006.

[22] F. Hiai, Log-majorizations and norm inequalities for exponential operators, in *Linear operators* (Warsaw, 1994), 119–181, Banach Center Publ., **38**, Polish Acad. Sci., Warsaw, 1997.

[23] F. Hiai and H. Kosaki, *Means of Hilbert space operators*, Lecture Notes in Mathematics, 1820. Springer-Verlag, Berlin, 2003.

[24] F. Hiai and D. Petz, Riemannian geometry on positive definite matrices related to means, Lin. Alg. Appl. **430**(2009), 3105–3130.

[25] T. Hida, Canonical representations of Gaussian processes and their applications. Mem. Coll. Sci. Univ. Kyoto. Ser. A. Math. **33**1960/1961, 109–155.

[26] T. Hida and M. Hitsuda, *Gaussian processes*. Translations of Mathematical Monographs, **120**. American Mathematical Society, Providence, RI, 1993.

[27] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory*, North-Holland, Amsterdam, 1982.

[28] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, 1985.

[29] Kérchy László, *Bevezetés a véges dimenziós vektorterek elméletébe*, Polygon, 1997.

[30] Kérchy László, *Hilbert terek operátorai*, Polygon, 2003.

[31] F. Kubo and T. Ando, Means of positive linear operators, Math. Ann. **246**(1980), 205–224.

[32] P.D. Lax, *Functional Analysis*, John Wiley & Sons, 2002.

[33] P.D. Lax, *Linear algebra and its applications*, John Wiley & Sons, 2007.

[34] C.-K. Li and R. Mathias, The Lidskii-Mirsky-Wielandt theorem – additive and multiplicative versions, Numer. Math. **81** (1999), 377–413.

[35] E. H. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture, Advances in Math. **11**(1973), 267–288.

[36] E. H. Lieb and R. Seiringer, Equivalent forms of the Bessis-Moussa-Villani conjecture, J. Statist. Phys. **115**(2004), 185–190.

[37] K. Löwner, Über monotone Matrixfunctionen, Math. Z. **38**(1934), 177–216.

[38] S. Neshveyev and E. Størmer, *Dynamical entropy in operator algebras*, Springer-Verlag, Berlin, 2006.

[39] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin, 1932. English translation: Mathematical foundations of quantum mechanics, Dover, New York, 1954. Hungarian translation: A kvantummechanika matematikai alapjai, Akadémiai Kiadó, 1980.

[40] M. Ohya and D. Petz, *Quantum Entropy and Its Use*, Springer-Verlag, Heidelberg, 1993. Second edition 2004.

[41] D. Petz, A variational expression for the relative entropy, Commun. Math. Phys., **114**(1988), 345–348.

[42] D. Petz, *An invitation to the algebra of the canonical commutation relation*, Leuven University Press, Leuven, 1990.

[43] D. Petz, Quasi-entropies for states of a von Neumann algebra, Publ. RIMS. Kyoto Univ. **21**(1985), 781–800.

[44] D. Petz, Monotone metrics on matrix spaces. Linear Algebra Appl. **244**(1996), 81–96.

[45] D. Petz, Quasi-entropies for finite quantum systems, Rep. Math. Phys., **23**(1986), 57-65.

[46] D. Petz, *Algebra of the canonical commutation relation*, Leuven University Press, 1990.

[47] D. Petz, *Quantum information theory and quantum statistics*, Springer, 2008.

[48] D. Petz and H. Hasegawa, On the Riemannian metric of α-entropies of density matrices, Lett. Math. Phys. **38**(1996), 221–225.

[49] D. Petz and R. Temesi, Means of positive numbers and matrices, SIAM Journal on Matrix Analysis and Applications, **27**(2006), 712-720.

[50] Rózsa Pál, *Lineáris algebra és alkalmazásai*, Tankönyvkiadó, 1991.

[51] E. Schrödinger, Probability relations between separated systems, Proc. Cambridge Philos. Soc. **31**(1936), 446–452.

[52] F. Zhang, *The Schur complement and its applications*, Springer, 2005.

[53] W. K. Wooters and W. H. Zurek, A single quantum cannot be cloned, Nature, **299**(1982), 802–803.