

Chapter 2

Mappings and Algebras

Most of the statements and definitions in this chapter are formulated in the Hilbert space setting. The Hilbert space is always assumed to be finite-dimensional, so instead of operators one can consider matrices. The idea of block matrices provides quite a useful tool in matrix theory. Some basic facts on block matrices are given in Sect. 2.1. Matrices have two primary structures; one is of course their algebraic structure with addition, multiplication, adjoint, etc., and another is the order structure coming from the partial order of positive semidefiniteness, as explained in Sect. 2.2. Based on this order one can consider several notions of positivity for linear maps between matrix algebras, which are discussed in Sect. 2.6.

2.1 Block Matrices

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 the pairs are computed entrywise 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 $A : \mathcal{H} \rightarrow \mathcal{K}$ 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 four linear mappings:

$$A_{i1} : f_1 \mapsto g_i, \quad A_{i1} : \mathcal{H}_1 \rightarrow \mathcal{K}_i \quad (1 \leq i \leq 2)$$

and

$$A_{i2} : f_2 \mapsto g'_i, \quad A_{i2} : \mathcal{H}_2 \rightarrow \mathcal{K}_i \quad (1 \leq i \leq 2).$$

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

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} A_{11}f_1 + A_{12}f_2 \\ A_{21}f_1 + A_{22}f_2 \end{bmatrix}.$$

(The right-hand side is $A(f_1, f_2)$ written in the form of a column 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 , $1 \leq i, j \leq 2$. The linear operators $A_{ij} : \mathcal{H}_j \rightarrow \mathcal{K}_i$ have a matrix $[A_{ij}]$ with respect to these bases. Since

$$\{(e_t^1, 0) : 1 \leq t \leq m(1)\} \cup \{(0, e_u^2) : 1 \leq u \leq m(2)\}$$

is a basis in \mathcal{H} and similarly

$$\{(f_t^1, 0) : 1 \leq t \leq n(1)\} \cup \{(0, f_u^2) : 1 \leq u \leq n(2)\}$$

is a basis in \mathcal{K} , the operator A has an $(n(1) + n(2)) \times (m(1) + m(2))$ 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 a **block matrix**.

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

$$\begin{aligned} & \begin{bmatrix} [A_{11}] & [A_{12}] \\ [A_{21}] & [A_{22}] \end{bmatrix} \cdot \begin{bmatrix} [B_{11}] & [B_{12}] \\ [B_{21}] & [B_{22}] \end{bmatrix} \\ &= \begin{bmatrix} [A_{11}] \cdot [B_{11}] + [A_{12}] \cdot [B_{21}] & [A_{11}] \cdot [B_{12}] + [A_{12}] \cdot [B_{22}] \\ [A_{21}] \cdot [B_{11}] + [A_{22}] \cdot [B_{21}] & [A_{21}] \cdot [B_{12}] + [A_{22}] \cdot [B_{22}] \end{bmatrix}. \end{aligned}$$

In several cases we do not emphasize the entries of a block matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

However, if this matrix is self-adjoint we assume that $A = A^*$, $B^* = C$ and $D = D^*$. (These conditions include that A and D are square matrices, $A \in \mathbb{M}_n$ and $B \in \mathbb{M}_m$.)

The block matrix is used for the definition of **reducible matrices**. $A \in \mathbb{M}_n$ is reducible if there is a permutation matrix $P \in \mathbb{M}_n$ such that

$$P^t A P = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}.$$

A matrix $A \in \mathbb{M}_n$ is **irreducible** if it is not reducible.

For a 2×2 matrix, it is very easy to check the positivity:

$$\begin{bmatrix} a & b \\ \bar{b} & c \end{bmatrix} \geq 0 \quad \text{if and only if} \quad a \geq 0 \quad \text{and} \quad b\bar{b} \leq ac.$$

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 2.1 *Assume that A is invertible. The self-adjoint block matrix*

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \tag{2.1}$$

is 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 all vectors f_1 and f_2 . A computation gives that this condition is

$$\langle f_1, f_1 \rangle + \langle f_2, Cf_2 \rangle \geq -2\operatorname{Re} \langle Bf_2, f_1 \rangle.$$

If we replace f_1 by $e^{i\varphi} f_1$ 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 (2.1) 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 $C \geq B^* A^{-1} B$. □

Another important characterization of the positivity of (2.1) is the condition that $A, C \geq 0$ and $B = A^{1/2} W C^{1/2}$ with a contraction W . (Here the invertibility of A or C is not necessary.)

Theorem 2.1 has applications in different areas, see for example the Cramér–Rao inequality, Sect. 7.5.

Theorem 2.2 *For an invertible A , we have the so-called **Schur factorization***

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix} \cdot \begin{bmatrix} A & 0 \\ 0 & D - CA^{-1} B \end{bmatrix} \cdot \begin{bmatrix} I & A^{-1} B \\ 0 & I \end{bmatrix}. \quad (2.2)$$

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 (2.2) with $C = B^*$ is equivalent to the positivity of the middle factor of the right-hand side. This fact gives the second proof of Theorem 2.1.

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 an easy computation of the determinant and the inverse.

Theorem 2.3 *The determinant can be computed as follows.*

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det A \cdot \det (D - CA^{-1}B).$$

If

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

then $D - CA^{-1}B$ is called the **Schur complement** of A in M , and is denoted by M/A . Hence the determinant formula becomes $\det M = \det A \cdot \det (M/A)$.

Theorem 2.4 *Let*

$$M = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix}$$

be a positive invertible matrix. Then

$$M/C = A - BC^{-1}B^* = \sup \left\{ X \geq 0 : \begin{bmatrix} X & 0 \\ 0 & 0 \end{bmatrix} \leq \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \right\}.$$

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. □

Theorem 2.5 *For a block matrix*

$$0 \leq \begin{bmatrix} A & X \\ X^* & B \end{bmatrix} \in \mathbb{M}_n,$$

we have

$$\begin{bmatrix} A & X \\ X^* & B \end{bmatrix} = U \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} U^* + V \begin{bmatrix} 0 & 0 \\ 0 & B \end{bmatrix} V^*$$

for some unitaries $U, V \in \mathbb{M}_n$.

Proof: We can take

$$0 \leq \begin{bmatrix} C & Y \\ Y^* & D \end{bmatrix} \in \mathbb{M}_n$$

such that

$$\begin{bmatrix} A & X \\ X^* & B \end{bmatrix} = \begin{bmatrix} C & Y \\ Y^* & D \end{bmatrix} \begin{bmatrix} C & Y \\ Y^* & D \end{bmatrix} = \begin{bmatrix} C^2 + YY^* & CY + YD \\ Y^*C + DY^* & Y^*Y + D^2 \end{bmatrix}.$$

It follows that

$$\begin{bmatrix} A & X \\ X^* & B \end{bmatrix} = \begin{bmatrix} C & 0 \\ Y^* & 0 \end{bmatrix} \begin{bmatrix} C & Y \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & Y \\ 0 & D \end{bmatrix} \begin{bmatrix} 0 & 0 \\ Y^* & D \end{bmatrix} = T^*T + S^*S,$$

where

$$T = \begin{bmatrix} C & Y \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} 0 & 0 \\ Y^* & D \end{bmatrix}.$$

When $T = U|T|$ and $S = V|S|$ for unitaries $U, V \in \mathbb{M}_n$, then

$$T^*T = U(TT^*)U^* \quad \text{and} \quad S^*S = V(SS^*)V^*.$$

From the formulas

$$TT^* = \begin{bmatrix} C^2 + YY^* & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}, \quad SS^* = \begin{bmatrix} 0 & 0 \\ 0 & Y^*Y + D^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & B \end{bmatrix},$$

we have the result. □

Example 2.6 Similarly to the previous theorem we take a block matrix

$$0 \leq \begin{bmatrix} A & X \\ X^* & B \end{bmatrix} \in \mathbb{M}_n.$$

For a unitary

$$W := \frac{1}{\sqrt{2}} \begin{bmatrix} iI & -I \\ iI & I \end{bmatrix}$$

we notice that

$$W \begin{bmatrix} A & X \\ X^* & B \end{bmatrix} W^* = \begin{bmatrix} \frac{A+B}{2} + \operatorname{Im} X & \frac{A-B}{2} + i\operatorname{Re} X \\ \frac{A-B}{2} - i\operatorname{Re} X & \frac{A+B}{2} - \operatorname{Im} X \end{bmatrix}.$$

So Theorem 2.5 gives

$$\begin{bmatrix} A & X \\ X^* & B \end{bmatrix} = U \begin{bmatrix} \frac{A+B}{2} + \operatorname{Im} X & 0 \\ 0 & 0 \end{bmatrix} U^* + V \begin{bmatrix} 0 & 0 \\ 0 & \frac{A+B}{2} - \operatorname{Im} X \end{bmatrix} V^*$$

for some unitaries $U, V \in \mathbb{M}_n$. □

We have two remarks. If C is not invertible, then the supremum in Theorem 2.4 is $A - BC^\dagger B^*$, where C^\dagger is the Moore–Penrose generalized inverse. The supremum of that theorem can be formulated without the block matrix formalism. Assume that P is an ortho-projection (see Sect. 2.3). Then

$$[P]M := \sup\{N : 0 \leq N \leq M, \quad PN = N\}. \quad (2.3)$$

If

$$P = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad M = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix},$$

then $[P]M = M/C$. The formula (2.3) makes clear that if Q is another ortho-projection such that $P \leq Q$, then $[P]M \leq [P]QM$.

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

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} \cdot \begin{bmatrix} A^{-1} & 0 \\ 0 & (D - CA^{-1}B)^{-1} \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix}.$$

After multiplication on the right-hand side, we have the following:

$$\begin{aligned} \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} &= \begin{bmatrix} A^{-1} + A^{-1}BW^{-1}CA^{-1} & -A^{-1}BW^{-1} \\ -W^{-1}CA^{-1} & W^{-1} \end{bmatrix} \\ &= \begin{bmatrix} V^{-1} & -V^{-1}BD^{-1} \\ -D^{-1}CV^{-1} & D^{-1} + D^{-1}CV^{-1}BD^{-1} \end{bmatrix}, \end{aligned} \quad (2.4)$$

where $W = M/A := D - CA^{-1}B$ and $V = M/D := A - BD^{-1}C$.

Example 2.7 Let X_1, X_2, \dots, X_{m+k} be real random variables with (Gaussian) joint probability distribution

$$f_M(\mathbf{z}) := \sqrt{\frac{\det M}{(2\pi)^{m+k}}} \exp\left(-\frac{1}{2}\langle \mathbf{z}, M\mathbf{z} \rangle\right),$$

where $\mathbf{z} = (z_1, z_2, \dots, z_{m+k})$ and M is a positive definite real $(m+k) \times (m+k)$ matrix, see Example 1.43. We want to compute the distribution of the random variables X_1, X_2, \dots, X_m .

Let

$$M = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix}$$

be written in the form of a block matrix, where 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

$$f_M(\mathbf{x}_1, \mathbf{x}_2) = \sqrt{\frac{\det M}{(2\pi)^{m+k}}} \exp\left(-\frac{1}{2}\langle(\mathbf{x}_1, \mathbf{x}_2), M(\mathbf{x}_1, \mathbf{x}_2)\rangle\right)$$

on \mathbb{R}^m is the distribution

$$f_1(\mathbf{x}_1) = \sqrt{\frac{\det M}{(2\pi)^m \det D}} \exp\left(-\frac{1}{2}\langle\mathbf{x}_1, (A - BD^{-1}B^*)\mathbf{x}_1\rangle\right). \quad (2.5)$$

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

$$\begin{aligned} &\int \exp\left(-\frac{1}{2}\langle(\mathbf{x}_1, \mathbf{x}_2)M(\mathbf{x}_1, \mathbf{x}_2)^t\rangle\right) d\mathbf{x}_2 \\ &= \exp\left(-\frac{1}{2}(\langle A\mathbf{x}_1, \mathbf{x}_1\rangle - \langle DW\mathbf{x}_1, W\mathbf{x}_1\rangle)\right) \\ &\quad \times \int \exp\left(-\frac{1}{2}\langle D(\mathbf{x}_2 + W\mathbf{x}_1), (\mathbf{x}_2 + W\mathbf{x}_1)\rangle\right) d\mathbf{x}_2 \\ &= \exp\left(-\frac{1}{2}\langle(A - BD^{-1}B^*)\mathbf{x}_1, \mathbf{x}_1\rangle\right) \sqrt{\frac{(2\pi)^k}{\det D}} \end{aligned}$$

and obtain (2.5).

This computation gives a proof of Theorem 2.3 (for a real positive definite matrix) as well. If we know that $f_1(\mathbf{x}_1)$ is Gaussian, then its quadratic matrix can be obtained from formula (2.4). The covariance of X_1, X_2, \dots, X_{m+k} is M^{-1} . Therefore, the covariance of X_1, X_2, \dots, X_m is $(A - BD^{-1}B^*)^{-1}$. It follows that the quadratic matrix is the inverse: $A - BD^{-1}B^* \equiv M/D$. \square

Theorem 2.8 *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, \dots, X_n .*

Proof: A can be written as $C^* C$ for some

$$C = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1n} \\ C_{21} & C_{22} & \dots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} \end{bmatrix}.$$

Let B_i be the block matrix such that its i th row is the same as in C and all other elements are 0. Then $C = B_1 + B_2 + \dots + B_n$ and for $t \neq i$ we have $B_i^* B_t = 0$. Therefore,

$$A = (B_1 + B_2 + \dots + B_n)^* (B_1 + B_2 + \dots + B_n) = B_1^* B_1 + B_2^* B_2 + \dots + B_n^* B_n.$$

The (i, j) entry of $B_i^* B_t$ is $C_{ii}^* C_{tj}$; hence this matrix is of the required form. \square

Example 2.9 Let \mathcal{H} be an n -dimensional Hilbert space and $A \in B(\mathcal{H})$ be a positive operator with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. If $x, y \in \mathcal{H}$ are orthogonal vectors, then

$$|\langle x, Ay \rangle|^2 \leq \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n} \right)^2 \langle x, Ax \rangle \langle y, Ay \rangle,$$

which is called the **Wielandt inequality**. (It also appeared in Theorem 1.45.) The argument presented here includes a block matrix.

We can assume that x and y are unit vectors and we extend them to a basis. Let

$$M = \begin{bmatrix} \langle x, Ax \rangle & \langle x, Ay \rangle \\ \langle y, Ax \rangle & \langle y, Ay \rangle \end{bmatrix},$$

where A has a block matrix

$$\begin{bmatrix} M & B \\ B^* & C \end{bmatrix}. \quad (2.6)$$

We can see that $M \geq 0$ and its determinant is positive:

$$|\langle x, Ay \rangle|^2 \leq \langle x, Ax \rangle \langle y, Ay \rangle.$$

If $\lambda_n = 0$, then the proof is complete. Now we assume that $\lambda_n > 0$. Let α and β be the eigenvalues of M . Formula (1.27) tells us that

$$|\langle x, Ay \rangle|^2 \leq \left(\frac{\alpha - \beta}{\alpha + \beta} \right)^2 \langle x, Ax \rangle \langle y, Ay \rangle.$$

We need the inequality

$$\frac{\alpha - \beta}{\alpha + \beta} \leq \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}$$

when $\alpha \geq \beta$. This is true, since $\lambda_1 \geq \alpha \geq \beta \geq \lambda_n$. \square

As an application of the block matrix technique, we consider the following result, called the **UL-factorization** (or the Cholesky factorization).

Theorem 2.10 *Let X be an $n \times n$ invertible positive matrix. Then there is a unique upper triangular matrix T with positive diagonal such that $X = TT^*$.*

Proof: The proof is by mathematical induction on 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

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix}, \quad (2.7)$$

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 $C = 0$, then the positivity of (2.7) forces $B = 0$ so that we can apply the induction hypothesis to A . So we may assume that $C > 0$. If the number T_{22} is positive, then $T_{22} = \sqrt{C}$ is the unique solution and moreover

$$T_{12} = BC^{-1/2}, \quad T_{11}T_{11}^* = A - BC^{-1}B^*.$$

From the positivity of (2.7), we have $A - BC^{-1}B^* \geq 0$ by Theorem 2.1. The induction hypothesis gives that the latter can be written in the form $T_{11}T_{11}^*$ where T_{11} is upper triangular. Therefore T is upper triangular, too. \square

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 2.11 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 is equal to $(\text{Tr}_2 D) \otimes (\text{Tr}_1 D)$ (as noted in Example 1.56). We have

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

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

In quantum theory the non-separable positive operators are said to be **entangled**. The positive operator D is **maximally entangled** if it has minimal rank (meaning 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 effective procedure to decide if a positive operator in a tensor product space is separable or entangled.

2.2 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 2.12 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 includes 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 be m -dimensional, so the matrices are $m \times m$. Recall that the following conditions are equivalent:

- (1) $\|A - A_n\| \rightarrow 0$.
- (2) $A_n x \rightarrow Ax$ for every vector x .
- (3) $\langle x, A_n y \rangle \rightarrow \langle x, Ay \rangle$ for all vectors x and y .
- (4) $\langle x, A_n x \rangle \rightarrow \langle x, Ax \rangle$ for every vector x .
- (5) $\text{Tr}(A - A_n)^*(A - A_n) \rightarrow 0$.
- (6) $[A_n]_{ij} \rightarrow [A]_{ij}$ for every $1 \leq i, j \leq m$.

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

Theorem 2.13 *Let A_n be an increasing sequence of operators with an upper bound: $A_1 \leq A_2 \leq \dots \leq B$. Then there is an operator $A \leq B$ such that $A_n \rightarrow A$.*

Proof: Let $\phi_n(x, y) := \langle x, A_n y \rangle$ be a sequence of complex bilinear functionals. Then $\phi_n(x, x)$ is a bounded increasing real sequence and it is convergent. By 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

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

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

Example 2.14 Assume that $0 \leq A \leq I$ for an operator A . Define a sequence T_n of operators by recursion. Let $T_1 = 0$ and

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

T_n is a polynomial in A with real coefficients. Thus 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 $T_n \leq I$.

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

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

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

Theorem 2.13 tells us that T_n converges to an operator B . The limit of the recursion formula yields

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

Therefore $A = B^2$. This example is a constructive proof of Theorem 1.38. □

Theorem 2.15 Assume that $0 < A, B \in \mathbb{M}_n$ are invertible matrices and $A \leq B$. Then $B^{-1} \leq A^{-1}$.

Proof: The condition $A \leq B$ is equivalent to $B^{-1/2}AB^{-1/2} \leq I$ and the statement $B^{-1} \leq A^{-1}$ is equivalent to $I \leq B^{1/2}A^{-1}B^{1/2}$. If $X = B^{-1/2}AB^{-1/2}$, then we have to show that $X \leq I$ implies $X^{-1} \geq I$. The condition $X \leq I$ 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 $A \leq B$. It follows from (1.13) 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 (counting multiplicities).

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

Theorem 2.16 If $A \leq B$, then $\lambda_k(A) \leq \lambda_k(B)$ for all k .

This is a consequence of the minimax principle, Theorem 1.27.

Corollary 2.17 Let $A, B \in B(\mathcal{H})$ be self-adjoint operators.

- (1) If $A \leq B$, then $\text{Tr } A \leq \text{Tr } B$.
- (2) If $0 \leq A \leq B$, then $\det A \leq \det B$.

Theorem 2.18 (Schur's theorem) Let A and B be positive $n \times n$ matrices. Then

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

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 by Example 1.40. The general case reduces to this one. \square

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

Corollary 2.19 Assume that $0 \leq A \leq B$ and $0 \leq C \leq D$. Then $A \circ C \leq B \circ D$.

Proof: The equation

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

implies the statement. \square

Theorem 2.20 (Oppenheim's inequality) If $0 \leq A, B \in \mathbb{M}_n$, then

$$\det(A \circ B) \geq \left(\prod_{i=1}^n A_{ii} \right) \det B.$$

Proof: For $n = 1$ the statement is obvious. The argument will be by induction on n . We take the Schur complementation and the block matrix formalism

$$A = \begin{bmatrix} a & A_1 \\ A_2 & A_3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b & B_1 \\ B_2 & B_3 \end{bmatrix},$$

where $a, b \in [0, \infty)$. We may assume that $a, b > 0$. From the inductive assumption we have

$$\det(A_3 \circ (B/b)) \geq A_{2,2} A_{3,3} \dots A_{n,n} \det(B/b). \quad (2.8)$$

From Theorem 2.3 we have $\det(A \circ B) = ab \det(A \circ B/ab)$ and

$$\begin{aligned} A \circ B/ab &= A_3 \circ B_3 - (A_2 \circ B_2) a^{-1} b^{-1} (A_1 \circ B_1) \\ &= A_3 \circ (B/b) + (A/a) \circ (B_2 B_1 b^{-1}). \end{aligned}$$

The matrices A/a and B/b are positive, see Theorem 2.4. So the matrices

$$A_3 \circ (B/b) \quad \text{and} \quad (A/a) \circ (B_2 B_1 b^{-1})$$

are positive as well. Thus

$$\det(A \circ B) \geq ab \det(A_3 \circ (B/b)).$$

Finally the inequality (2.8) gives

$$\det(A \circ B) \geq \left(\prod_{i=1}^n A_{ii} \right) b \det(B/b).$$

Since $\det B = b \det(B/b)$, the proof is complete. \square

A linear mapping $\alpha : \mathbb{M}_n \rightarrow \mathbb{M}_n$ is called **completely positive** if it has the form

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

for some matrices V_i . The sum of completely positive mappings is completely positive. (More details concerning completely positive mappings can be found in Theorem 2.49.)

Example 2.21 Let $A \in \mathbb{M}_n$ be a positive matrix. The mapping $S_A : B \mapsto A \circ B$ sends positive matrices to positive matrices. Therefore it is a positive mapping.

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

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

and S_A is completely positive. \square

2.3 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 $x_1 \perp \mathcal{K}$, see Theorem 1.11. The linear mapping $P : x \mapsto x_0$ is called the (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 terminology **ortho-projection** is also used.

The partial ordering is very simple for projections, see Example 2.12. If P and Q are projections, then the relation $P \leq Q$ means that the range of P is included in the range of Q . An equivalent algebraic formulation is $PQ = P$. The largest projection in \mathbb{M}_n is the identity I and the smallest one is 0. Therefore $0 \leq P \leq I$ for any projection $P \in \mathbb{M}_n$.

Example 2.22 In \mathbb{M}_2 the non-trivial ortho-projections have rank 1 and they have the form

$$P = \frac{1}{2} \begin{bmatrix} 1 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & 1 - a_3 \end{bmatrix},$$

where $a_1, a_2, a_3 \in \mathbb{R}$ and $a_1^2 + a_2^2 + a_3^2 = 1$. In terms of the **Pauli matrices**

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (2.9)$$

we have

$$P = \frac{1}{2} \left(\sigma_0 + \sum_{i=1}^3 a_i \sigma_i \right).$$

An equivalent formulation is $P = |x\rangle\langle x|$, where $x \in \mathbb{C}^2$ is a unit vector. This can be extended to an arbitrary ortho-projection $Q \in \mathbb{M}_n(\mathbb{C})$:

$$Q = \sum_{i=1}^k |x_i\rangle\langle x_i|,$$

where the set $\{x_i : 1 \leq i \leq k\}$ is a family of orthogonal unit vectors in \mathbb{C}^n . (k is the rank of the image of Q , or $\text{Tr } Q$.) \square

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 2.23 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 also arrive at this statement by taking the adjoint.)

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 conclude that $PQ = 0$, which is the orthogonality. \square

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 maximal projection, denoted by $P \wedge Q$, which is the orthogonal projection onto the intersection of the ranges of P and Q .

Theorem 2.24 Assume that P and Q are ortho-projections. Then

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

Proof: The operator $A := PQP$ is a positive contraction. Therefore the sequence A^n is monotone decreasing and Theorem 2.13 implies that A^n has the limit R . The operator R is self-adjoint. Since $(A^n)^2 \rightarrow R^2$ 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 $PQP \leq P$, $R \leq P$ follows. Taking the limit of $(PQP)^n Q$ $(PQP)^n = (PQP)^{2n+1}$, we have $RQR = R$. From this we have $R(I - Q)R = 0$ and $(I - Q)R = 0$. This gives $R \leq Q$.

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

Corollary 2.25 Assume that P and Q are ortho-projections and $0 \leq H \leq P, Q$. Then $H \leq P \wedge Q$.

Proof: Since $(I - P)H(I - P) = 0$ implies $H^{1/2}(I - P) = 0$, we have $H^{1/2}P = H^{1/2}$ so that $PHP = H$, and similarly $QHQ = H$. These imply $(PQP)^n H (PQP)^n = H$ and the limit $n \rightarrow \infty$ gives $RHR = H$, where $R = P \wedge Q$. Hence $H \leq R$. \square

Let P and Q be ortho-projections. If the ortho-projection R has the property $R \geq P, Q$, then the image of R includes the images of P and Q . The smallest such R projects to the linear subspace 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 . However, the so-called distributivity

$$A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C)$$

does not hold.

Example 2.26 We show that any operator $X \in \mathbb{M}_n(\mathbb{C})$ is a linear combination of ortho-projections. We write

$$X = \frac{1}{2}(X + X^*) + \frac{1}{2i}(iX - iX^*),$$

where $X + X^*$ and $iX - iX^*$ are self-adjoint operators. Therefore, it is enough to find linear combinations of ortho-projections for self-adjoint operators. This is essentially the spectral decomposition (1.11).

Assume that φ_0 is defined on projections of $\mathbb{M}_n(\mathbb{C})$ and it has the properties

$$\varphi_0(0) = 0, \quad \varphi_0(I) = 1, \quad \varphi_0(P + Q) = \varphi_0(P) + \varphi_0(Q) \quad \text{if } P \perp Q.$$

It is a famous theorem of **Gleason** that in the case $n > 2$ the mapping φ_0 has a linear extension $\varphi : \mathbb{M}_n(\mathbb{C}) \rightarrow \mathbb{C}$. The linearity implies φ is of the form

$$\varphi(X) = \text{Tr } \rho X \quad (X \in \mathbb{M}_n(\mathbb{C}))$$

for some matrix $\rho \in \mathbb{M}_n(\mathbb{C})$. However, from the properties of φ_0 we have $\rho \geq 0$ and $\text{Tr } \rho = 1$. Such a ρ is usually called a density matrix in the quantum applications. It is clear that if ρ has rank 1, then it is a projection. \square

In quantum information theory the traditional **variance** is

$$\text{Var}_\rho(A) = \text{Tr } \rho A^2 - (\text{Tr } \rho A)^2 \quad (2.10)$$

where ρ is a density matrix and $A \in \mathbb{M}_n(\mathbb{C})$ is a self-adjoint operator. This is a straightforward analogy of the variance in probability theory; a standard notation is $\langle A^2 \rangle - \langle A \rangle^2$ in both formalisms. We note that for two self-adjoint operators the corresponding notion is **covariance**:

$$\text{Cov}_\rho(A, B) = \text{Tr } \rho AB - (\text{Tr } \rho A)(\text{Tr } \rho B).$$

It is rather different from probability theory that the variance (2.10) can be strictly positive even in the case where ρ has rank 1. If ρ has rank 1, then it is an ortho-projection of rank 1, also known as a pure state.

It is easy to show that

$$\text{Var}_\rho(A + \lambda I) = \text{Var}_\rho(A) \quad \text{for } \lambda \in \mathbb{R}$$

and the concavity of the variance functional $\rho \mapsto \text{Var}_\rho(A)$:

$$\text{Var}_\rho(A) \geq \sum_i \lambda_i \text{Var}_{\rho_i}(A) \quad \text{if } \rho = \sum_i \lambda_i \rho_i.$$

(Here $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$.)

The formulation is easier if ρ is diagonal. We can change the basis of the n -dimensional space so that $\rho = \text{Diag}(p_1, p_2, \dots, p_n)$; then we have

$$\text{Var}_\rho(A) = \sum_{i,j} \frac{p_i + p_j}{2} |A_{ij}|^2 - \left(\sum_i p_i A_{ii} \right)^2. \quad (2.11)$$

In the projection example $P = \text{Diag}(1, 0, \dots, 0)$, formula (2.11) gives

$$\text{Var}_P(A) = \sum_{i \neq 1} |A_{1i}|^2$$

and this can be strictly positive.

Theorem 2.27 *Let ρ be a density matrix. Take all the decompositions such that*

$$\rho = \sum_i q_i Q_i, \quad (2.12)$$

where Q_i are pure states and (q_i) is a probability distribution. Then

$$\text{Var}_\rho(A) = \sup \left(\sum_i q_i \left(\text{Tr } Q_i A^2 - (\text{Tr } Q_i A)^2 \right) \right), \quad (2.13)$$

where the supremum is over all decompositions (2.12).

The proof will be an application of matrix theory. The first lemma contains a trivial computation on block matrices.

Lemma 2.28 *Assume that*

$$\rho = \begin{bmatrix} \rho^\wedge & 0 \\ 0 & 0 \end{bmatrix}, \quad \rho_i = \begin{bmatrix} \rho_i^\wedge & 0 \\ 0 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} A^\wedge & B \\ B^* & C \end{bmatrix}$$

and

$$\rho = \sum_i \lambda_i \rho_i, \quad \rho^\wedge = \sum_i \lambda_i \rho_i^\wedge.$$

Then

$$\begin{aligned} & \left(\text{Tr } \rho^\wedge (A^\wedge)^2 - (\text{Tr } \rho^\wedge A^\wedge)^2 \right) - \sum_i \lambda_i \left(\text{Tr } \rho_i^\wedge (A^\wedge)^2 - (\text{Tr } \rho_i^\wedge A^\wedge)^2 \right) \\ &= (\text{Tr } \rho A^2 - (\text{Tr } \rho A)^2) - \sum_i \lambda_i \left(\text{Tr } \rho_i A^2 - (\text{Tr } \rho_i A)^2 \right). \end{aligned}$$

This lemma shows that if $\rho \in \mathbb{M}_n(\mathbb{C})$ has a rank $k < n$, then the computation of a variance $\text{Var}_\rho(A)$ can be reduced to $k \times k$ matrices. The equality in (2.13) is rather obvious for a rank 2 density matrix and, by the previous lemma, the computations will be with 2×2 matrices.

Lemma 2.29 *For a rank 2 matrix ρ , equality holds in (2.13).*

Proof: By Lemma 2.28 we can make a computation with 2×2 matrices. We can assume that

$$\rho = \begin{bmatrix} p & 0 \\ 0 & 1-p \end{bmatrix}, \quad A = \begin{bmatrix} a_1 & b \\ \bar{b} & a_2 \end{bmatrix}.$$

Then

$$\text{Tr } \rho A^2 = p(a_1^2 + |b|^2) + (1-p)(a_2^2 + |b|^2).$$

We can assume that

$$\text{Tr } \rho A = p a_1 + (1 - p) a_2 = 0.$$

Let

$$Q_1 = \begin{bmatrix} p & c e^{-i\varphi} \\ c e^{i\varphi} & 1 - p \end{bmatrix},$$

where $c = \sqrt{p(1-p)}$. This is a projection and

$$\text{Tr } Q_1 A = a_1 p + a_2 (1 - p) + b c e^{-i\varphi} + \bar{b} c e^{i\varphi} = 2c \text{Re } b e^{-i\varphi}.$$

We choose φ such that $\text{Re } b e^{-i\varphi} = 0$. Then $\text{Tr } Q_1 A = 0$ and

$$\text{Tr } Q_1 A^2 = p(a_1^2 + |b|^2) + (1 - p)(a_2^2 + |b|^2) = \text{Tr } \rho A^2.$$

Let

$$Q_2 = \begin{bmatrix} p & -c e^{-i\varphi} \\ -c e^{i\varphi} & 1 - p \end{bmatrix}.$$

Then

$$\rho = \frac{1}{2} Q_1 + \frac{1}{2} Q_2$$

and we have

$$\frac{1}{2}(\text{Tr } Q_1 A^2 + \text{Tr } Q_2 A^2) = p(a_1^2 + |b|^2) + (1 - p)(a_2^2 + |b|^2) = \text{Tr } \rho A^2.$$

Therefore we have an equality. □

We denote by $r(\rho)$ the rank of an operator ρ . The idea of the proof is to reduce the rank and the block diagonal formalism will be used.

Lemma 2.30 *Let ρ be a density matrix and $A = A^*$ be in $\mathbb{M}_n(\mathbb{C})$. Assume the block matrix forms*

$$\rho = \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix}, \quad A = \begin{bmatrix} A_1 & A_2 \\ A_2^* & A_3 \end{bmatrix}$$

and $r(\rho_1), r(\rho_2) > 1$. We construct

$$\rho' := \begin{bmatrix} \rho_1 & X^* \\ X & \rho_2 \end{bmatrix}$$

such that

$$\text{Tr } \rho A = \text{Tr } \rho' A, \quad \rho' \geq 0, \quad r(\rho') < r(\rho).$$

Proof: The condition $\text{Tr } \rho A = \text{Tr } \rho' A$ is equivalent to $\text{Tr } X A_2 + \text{Tr } X^* A_2^* = 0$ and this holds if and only if $\text{Re } \text{Tr } X A_2 = 0$.

There exist unitaries U and W such that $U \rho_1 U^*$ and $W \rho_2 W^*$ are diagonal:

$$U \rho_1 U^* = \text{Diag}(0, \dots, 0, a_1, \dots, a_k), \quad W \rho_2 W^* = \text{Diag}(b_1, \dots, b_l, 0, \dots, 0)$$

where $a_i, b_j > 0$. Then ρ has the same rank, $k + l$, as the matrix

$$\begin{bmatrix} U & 0 \\ 0 & W \end{bmatrix} \rho \begin{bmatrix} U^* & 0 \\ 0 & W^* \end{bmatrix} = \begin{bmatrix} U \rho_1 U^* & 0 \\ 0 & W \rho_2 W^* \end{bmatrix}.$$

A possible modification of this matrix is $Y :=$

$$\begin{bmatrix} \text{Diag}(0, \dots, 0, a_1, \dots, a_{k-1}) & 0 & 0 & 0 \\ 0 & a_k & \sqrt{a_k b_1} & 0 \\ 0 & \sqrt{a_k b_1} & b_1 & 0 \\ 0 & 0 & 0 & \text{Diag}(b_2, \dots, b_l, 0, \dots, 0) \end{bmatrix} \\ = \begin{bmatrix} U \rho_1 U^* & M \\ M & W \rho_2 W^* \end{bmatrix}$$

and $r(Y) = k + l - 1$. So Y has a smaller rank than ρ . Next we take

$$\begin{bmatrix} U^* & 0 \\ 0 & W^* \end{bmatrix} Y \begin{bmatrix} U & 0 \\ 0 & W \end{bmatrix} = \begin{bmatrix} \rho_1 & U^* M W \\ W^* M U & \rho_2 \end{bmatrix}$$

which has the same rank as Y . If $X_1 := W^* M U$ is multiplied by $e^{i\alpha}$ ($\alpha > 0$), then the positivity condition and the rank remain. On the other hand, we can choose $\alpha > 0$ such that $\text{Re } \text{Tr } e^{i\alpha} X_1 A_2 = 0$. Then $X := e^{i\alpha} X_1$ is the matrix we wanted. \square

Lemma 2.31 *Let ρ be a density matrix of rank $m > 0$ and $A = A^*$ be in $\mathbb{M}_n(\mathbb{C})$. We claim the existence of a decomposition*

$$\rho = p\rho_- + (1 - p)\rho_+$$

such that $r(\rho_-) < m$, $r(\rho_+) < m$, and

$$\text{Tr } A \rho_+ = \text{Tr } A \rho_- = \text{Tr } \rho A.$$

Proof: By unitary transformation we can obtain the setup of the previous lemma:

$$\rho = \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix}, \quad A = \begin{bmatrix} A_1 & A_2 \\ A_2^* & A_3 \end{bmatrix}.$$

With ρ' as in the previous lemma we choose

$$\rho_+ = \rho' = \begin{bmatrix} \rho_1 & X^* \\ X & \rho_2 \end{bmatrix}, \quad \rho_- = \begin{bmatrix} \rho_1 & -X^* \\ -X & \rho_2 \end{bmatrix}.$$

Then

$$\rho = \frac{1}{2}\rho_- + \frac{1}{2}\rho_+$$

and the requirements $\text{Tr } A\rho_+ = \text{Tr } A\rho_- = \text{Tr } \rho A$ also hold. \square

Proof of Theorem 2.27: For rank 2 states, the theorem is true by Lemma 2.29. Any state with a rank larger than 2 can be decomposed into a mixture of lower rank states, according to Lemma 2.31, that have the same expectation value for A as the original ρ has. The lower rank states can then be decomposed into a mixture of states with an even lower rank, until we reach states of rank ≤ 2 . Thus, any state ρ can be decomposed into a mixture of pure states

$$\rho = \sum p_k Q_k$$

such that $\text{Tr } A Q_k = \text{Tr } A\rho$. Hence the statement of the theorem follows. \square

2.4 Subalgebras

A unital ***-subalgebra** of $\mathbb{M}_n(\mathbb{C})$ is a subspace \mathcal{A} that contains the identity I and is closed under matrix multiplication and adjoint. That is, if $A, B \in \mathcal{A}$, then so are AB and A^* . In what follows, to simplify the notation, we shall use the term subalgebra for all *-subalgebras.

Example 2.32 A simple subalgebra is

$$\mathcal{A} = \left\{ \begin{bmatrix} z & w \\ w & z \end{bmatrix} : z, w \in \mathbb{C} \right\} \subset \mathbb{M}_2(\mathbb{C}).$$

Since $A, B \in \mathcal{A}$ implies $AB = BA$, this is a commutative subalgebra. In terms of the Pauli matrices (2.9) we have

$$\mathcal{A} = \{z\sigma_0 + w\sigma_1 : z, w \in \mathbb{C}\}.$$

This example will be generalized. \square

Assume that P_1, P_2, \dots, P_n are projections of rank 1 in $\mathbb{M}_n(\mathbb{C})$ such that $P_i P_j = 0$ for $i \neq j$ and $\sum_i P_i = I$. Then

$$\mathcal{A} = \left\{ \sum_{i=1}^n \alpha_i P_i : \alpha_i \in \mathbb{C} \right\}$$

is a maximal commutative $*$ -subalgebra of $\mathbb{M}_n(\mathbb{C})$. The usual name is **MASA**, which is an acronym for Maximal Abelian Sub-Algebra.

Let \mathcal{A} be any subset of $\mathbb{M}_n(\mathbb{C})$. Then \mathcal{A}' , the **commutant** of \mathcal{A} , is given by

$$\mathcal{A}' = \{B \in \mathbb{M}_n(\mathbb{C}) : BA = AB \text{ for all } A \in \mathcal{A}\}.$$

It is easy to see that for any set $\mathcal{A} \subset \mathbb{M}_n(\mathbb{C})$, \mathcal{A}' is a subalgebra. If \mathcal{A} is a MASA, then $\mathcal{A}'' = \mathcal{A}$.

Theorem 2.33 *If $\mathcal{A} \subset \mathbb{M}_n(\mathbb{C})$ is a unital $*$ -subalgebra, then $\mathcal{A}'' = \mathcal{A}$.*

Proof: We first show that for any $*$ -subalgebra \mathcal{A} , $B \in \mathcal{A}''$ and any $v \in \mathbb{C}^n$, there exists an $A \in \mathcal{A}$ such that $Av = Bv$. Let \mathcal{K} be the subspace of \mathbb{C}^n given by

$$\mathcal{K} = \{Av : A \in \mathcal{A}\}.$$

Let P be the orthogonal projection onto \mathcal{K} in \mathbb{C}^n . Since, by construction, \mathcal{K} is invariant under the action of \mathcal{A} , $PAP = AP$ for all $A \in \mathcal{A}$. Taking the adjoint, $PA^*P = PA^*$ for all $A \in \mathcal{A}$. Since \mathcal{A} is a $*$ -algebra, this implies $PA = AP$ for all $A \in \mathcal{A}$. That is, $P \in \mathcal{A}'$. Thus, for any $B \in \mathcal{A}''$, $BP = PB$ and so \mathcal{K} is invariant under the action of \mathcal{A}'' . In particular, $Bv \in \mathcal{K}$ and hence, by the definition of \mathcal{K} , $Bv = Av$ for some $A \in \mathcal{A}$.

We apply the previous statement to the $*$ -subalgebra

$$\mathcal{M} = \{A \otimes I_n : A \in \mathcal{A}\} \subset \mathbb{M}_n(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C}) = \mathbb{M}_{n^2}(\mathbb{C}).$$

It is easy to see that

$$\mathcal{M}'' = \{B \otimes I_n : B \in \mathcal{A}''\} \subset \mathbb{M}_n(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C}).$$

Now let $\{v_1, \dots, v_n\}$ be any basis of \mathbb{C}^n and form the vector

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{C}^{n^2}.$$

Then

$$(A \otimes I_n)v = (B \otimes I_n)v$$

and $Av_j = Bv_j$ for every $1 \leq j \leq n$. Since $\{v_1, \dots, v_n\}$ is a basis of \mathbb{C}^n , this means $B = A \in \mathcal{A}$. Since B was an arbitrary element of \mathcal{A}'' , this shows that $\mathcal{A}'' \subset \mathcal{A}$. Since $\mathcal{A} \subset \mathcal{A}''$ is an automatic consequence of the definitions, this proves that $\mathcal{A}'' = \mathcal{A}$. \square

Next we study subalgebras $\mathcal{A} \subset \mathcal{B} \subset \mathbb{M}_n(\mathbb{C})$. A **conditional expectation** $\mathcal{E} : \mathcal{B} \rightarrow \mathcal{A}$ is a unital positive mapping which has the property

$$\mathcal{E}(AB) = A\mathcal{E}(B) \quad \text{for every } A \in \mathcal{A} \text{ and } B \in \mathcal{B}.$$

Choosing $B = I$, we obtain that \mathcal{E} acts identically on \mathcal{A} . It follows from the positivity of \mathcal{E} that $\mathcal{E}(C^*) = \mathcal{E}(C)^*$. Therefore, $\mathcal{E}(BA) = \mathcal{E}(B)A$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Another standard notation for a conditional expectation $\mathcal{B} \rightarrow \mathcal{A}$ is $\mathcal{E}_{\mathcal{A}}^{\mathcal{B}}$.

Theorem 2.34 *Assume that $\mathcal{A} \subset \mathcal{B} \subset \mathbb{M}_n(\mathbb{C})$. If $\alpha : \mathcal{A} \rightarrow \mathcal{B}$ is the embedding, then the dual $\mathcal{E} : \mathcal{B} \rightarrow \mathcal{A}$ of α with respect to the Hilbert–Schmidt inner product is a conditional expectation.*

Proof: From the definition

$$\text{Tr } \alpha(A)B = \text{Tr } A\mathcal{E}(B) \quad (A \in \mathcal{A}, B \in \mathcal{B})$$

of the dual, we see that $\mathcal{E} : \mathcal{B} \rightarrow \mathcal{A}$ is a positive unital mapping and $\mathcal{E}(A) = A$ for every $A \in \mathcal{A}$. For every $A, A_1 \in \mathcal{A}$ and $B \in \mathcal{B}$ we further have

$$\text{Tr } A\mathcal{E}(A_1 B) = \text{Tr } \alpha(A)A_1 B = \text{Tr } \alpha(AA_1)B = \text{Tr } AA_1\mathcal{E}(B),$$

which implies that $\mathcal{E}(A_1 B) = A_1\mathcal{E}(B)$. \square

Note that a conditional expectation $\mathcal{E} : \mathcal{B} \rightarrow \mathcal{A}$ has norm 1, that is, $\|\mathcal{E}(B)\| \leq \|B\|$ for every $B \in \mathcal{B}$. This follows from Corollary 2.45.

The subalgebras $\mathcal{A}_1, \mathcal{A}_2 \subset \mathbb{M}_n(\mathbb{C})$ cannot be orthogonal since I is in \mathcal{A}_1 and in \mathcal{A}_2 . They are called **complementary** or **quasi-orthogonal** if $A_i \in \mathcal{A}_i$ and $\text{Tr } A_i = 0$ for $i = 1, 2$ imply that $\text{Tr } A_1 A_2 = 0$.

Example 2.35 In $\mathbb{M}_2(\mathbb{C})$ the subalgebras

$$\mathcal{A}_i := \{a\sigma_0 + b\sigma_i : a, b \in \mathbb{C}\} \quad (1 \leq i \leq 3)$$

are commutative and quasi-orthogonal. This follows from the facts that $\text{Tr } \sigma_i = 0$ for $1 \leq i \leq 3$ and

$$\sigma_1\sigma_2 = i\sigma_3, \quad \sigma_2\sigma_3 = i\sigma_1, \quad \sigma_3\sigma_1 = i\sigma_2.$$

So $\mathbb{M}_2(\mathbb{C})$ has 3 quasi-orthogonal MASAs.

In $\mathbb{M}_4(\mathbb{C}) = \mathbb{M}_2(\mathbb{C}) \otimes \mathbb{M}_2(\mathbb{C})$ we can give five quasi-orthogonal MASAs. Each of them is the linear span of four operators in one of the following lines:

$$\begin{array}{cccc}
\sigma_0 \otimes \sigma_0, & \sigma_0 \otimes \sigma_1, & \sigma_1 \otimes \sigma_0, & \sigma_1 \otimes \sigma_1, \\
\sigma_0 \otimes \sigma_0, & \sigma_0 \otimes \sigma_2, & \sigma_2 \otimes \sigma_0, & \sigma_2 \otimes \sigma_2, \\
\sigma_0 \otimes \sigma_0, & \sigma_0 \otimes \sigma_3, & \sigma_3 \otimes \sigma_0, & \sigma_3 \otimes \sigma_3, \\
\sigma_0 \otimes \sigma_0, & \sigma_1 \otimes \sigma_2, & \sigma_2 \otimes \sigma_3, & \sigma_3 \otimes \sigma_1, \\
\sigma_0 \otimes \sigma_0, & \sigma_1 \otimes \sigma_3, & \sigma_2 \otimes \sigma_1, & \sigma_3 \otimes \sigma_2.
\end{array}$$

□

Theorem 2.36 Assume that $\{\mathcal{A}_i : 1 \leq i \leq k\}$ is a set of quasi-orthogonal MASAs in $\mathbb{M}_n(\mathbb{C})$. Then $k \leq n + 1$.

Proof: The argument is rather simple. The traceless part of $\mathbb{M}_n(\mathbb{C})$ has dimension $n^2 - 1$ and the traceless part of a MASA has dimension $n - 1$. Therefore $k \leq (n^2 - 1)/(n - 1) = n + 1$. □

Determining the maximal number of quasi-orthogonal MASAs is a hard problem. For example, if $n = 2^m$, then $n + 1$ MASAs is possible, but for an arbitrary n there is no definite result.

The next theorem gives a characterization of complementarity.

Theorem 2.37 Let \mathcal{A}_1 and \mathcal{A}_2 be subalgebras of $\mathbb{M}_n(\mathbb{C})$ and denote Tr / n by τ . The following conditions are equivalent:

- (i) If $P \in \mathcal{A}_1$ and $Q \in \mathcal{A}_2$ are minimal projections, then $\tau(PQ) = \tau(P)\tau(Q)$.
- (ii) The subalgebras \mathcal{A}_1 and \mathcal{A}_2 are quasi-orthogonal in $\mathbb{M}_n(\mathbb{C})$.
- (iii) $\tau(A_1 A_2) = \tau(A_1)\tau(A_2)$ if $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$.
- (iv) If $\mathcal{E}_1 : \mathbb{M}_n(\mathbb{C}) \rightarrow \mathcal{A}_1$ is the trace-preserving conditional expectation, then \mathcal{E}_1 restricted to \mathcal{A}_2 is a linear functional (times I).

Proof: Note that $\tau((A_1 - \tau(A_1)I)(A_2 - \tau(A_2)I)) = 0$ and $\tau(A_1 A_2) = \tau(A_1)\tau(A_2)$ are equivalent. If they hold for minimal projections, they hold for arbitrary operators as well. Moreover, (iv) is equivalent to the property $\tau(A_1 \mathcal{E}_1(A_2)) = \tau(A_1(\tau(A_2)I))$ for every $A_1 \in \mathcal{A}_1$ and $A_2 \in \mathcal{A}_2$, and note that $\tau(A_1 \mathcal{E}_1(A_2)) = \tau(A_1 A_2)$. □

Example 2.38 A simple example of quasi-orthogonal subalgebras can be formulated with tensor products. If $\mathcal{A} = \mathbb{M}_n(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$, $\mathcal{A}_1 = \mathbb{M}_n(\mathbb{C}) \otimes \mathbb{C}I_n \subset \mathcal{A}$ and $\mathcal{A}_2 = \mathbb{C}I_n \otimes \mathbb{M}_n(\mathbb{C}) \subset \mathcal{A}$, then \mathcal{A}_1 and \mathcal{A}_2 are quasi-orthogonal subalgebras of \mathcal{A} . This comes from the property $\text{Tr}(A \otimes B) = \text{Tr} A \cdot \text{Tr} B$.

For $n = 2$ we give another example using the Pauli matrices. The 4-dimensional subalgebra $\mathcal{A}_1 = \mathbb{M}_2(\mathbb{C}) \otimes \mathbb{C}I_2$ is the linear span of the set

$$\{\sigma_0 \otimes \sigma_0, \sigma_1 \otimes \sigma_0, \sigma_2 \otimes \sigma_0, \sigma_3 \otimes \sigma_0\}.$$

Together with the identity, each of the following triplets linearly spans a subalgebra \mathcal{A}_j isomorphic to $M_2(\mathbb{C})$ ($2 \leq j \leq 4$):

$$\begin{aligned} &\{\sigma_3 \otimes \sigma_1, \sigma_3 \otimes \sigma_2, \sigma_0 \otimes \sigma_3\}, \\ &\{\sigma_2 \otimes \sigma_3, \sigma_2 \otimes \sigma_1, \sigma_0 \otimes \sigma_2\}, \\ &\{\sigma_1 \otimes \sigma_2, \sigma_1 \otimes \sigma_3, \sigma_0 \otimes \sigma_1\}. \end{aligned}$$

It is easy to check that the subalgebras $\mathcal{A}_1, \dots, \mathcal{A}_4$ are complementary.

The orthogonal complement of the four subalgebras is spanned by $\{\sigma_0 \otimes \sigma_3, \sigma_3 \otimes \sigma_0, \sigma_3 \otimes \sigma_3\}$. The linear span of this together with $\sigma_0 \otimes \sigma_0$ is a commutative subalgebra. \square

The previous example describes the general situation for $M_4(\mathbb{C})$. This will be the content of the next theorem. It is easy to calculate that the number of complementary subalgebras isomorphic to $M_2(\mathbb{C})$ is at most $(16 - 1)/3 = 5$. However, the next theorem says that 5 is not possible.

If $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, then the notation

$$x \cdot \sigma = x_1 \sigma_1 + x_2 \sigma_2 + x_3 \sigma_3$$

will be used and shall be called a Pauli triplet.

Theorem 2.39 *Assume that $\{\mathcal{A}_i : 0 \leq i \leq 3\}$ is a family of pairwise quasi-orthogonal subalgebras of $M_4(\mathbb{C})$ which are isomorphic to $M_2(\mathbb{C})$. For every $0 \leq i \leq 3$, there exists a Pauli triplet $A(i, j)$ ($j \neq i$) such that $\mathcal{A}'_i \cap \mathcal{A}_j$ is the linear span of I and $A(i, j)$. Moreover, the subspace linearly spanned by*

$$I \quad \text{and} \quad \left(\bigcup_{i=0}^3 \mathcal{A}_i \right)^\perp$$

is a maximal Abelian subalgebra.

Proof: Since the intersection $\mathcal{A}'_0 \cap \mathcal{A}_j$ is a 2-dimensional commutative subalgebra, we can find a self-adjoint unitary $A(0, j)$ such that $\mathcal{A}'_0 \cap \mathcal{A}_j$ is spanned by I and $A(0, j) = x(0, j) \cdot \sigma \otimes I$, where $x(0, j) \in \mathbb{R}^3$. Due to the quasi-orthogonality of $\mathcal{A}_1, \mathcal{A}_2$ and \mathcal{A}_3 , the unit vectors $x(0, j)$ are pairwise orthogonal (see (2.18)). The matrices $A(0, j)$ anti-commute:

$$\begin{aligned} A(0, i)A(0, j) &= i(x(0, i) \times x(0, j)) \cdot \sigma \otimes I \\ &= -i(x(0, j) \times x(0, i)) \cdot \sigma \otimes I = -A(0, j)A(0, i) \end{aligned}$$

for $i \neq j$. Moreover,

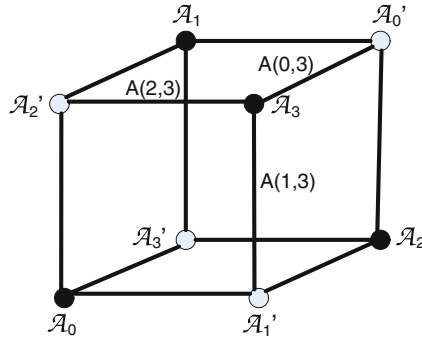
$$A(0, 1)A(0, 2) = i(x(0, 1) \times x(0, 2)) \cdot \sigma$$

and $x(0, 1) \times x(0, 2) = \pm x(0, 3)$ because $x(0, 1) \times x(0, 2)$ is orthogonal to both $x(0, 1)$ and $x(0, 2)$. If necessary, we can change the sign of $x(0, 3)$ so that $A(0, 1)A(0, 2) = iA(0, 3)$ holds.

Starting with the subalgebras $\mathcal{A}'_1, \mathcal{A}'_2, \mathcal{A}'_3$ we can similarly construct the other Pauli triplets. In this way, we arrive at the four Pauli triplets, the rows of the following table:

$$\begin{array}{cccc}
 \star & A(0, 1) & A(0, 2) & A(0, 3) \\
 A(1, 0) & \star & A(1, 2) & A(1, 3) \\
 A(2, 0) & A(2, 1) & \star & A(2, 3) \\
 A(3, 0) & A(3, 1) & A(3, 2) & \star
 \end{array} \tag{2.14}$$

When $\{\mathcal{A}_i : 1 \leq i \leq 3\}$ is a family of pairwise quasi-orthogonal subalgebras, then the commutants $\{\mathcal{A}'_i : 1 \leq i \leq 3\}$ are pairwise quasi-orthogonal as well. $\mathcal{A}'_j = \mathcal{A}_j$ and \mathcal{A}'_i have nontrivial intersection for $i \neq j$, actually the previously defined $A(i, j)$ is in the intersection. For a fixed j the three unitaries $A(i, j)$ ($i \neq j$) form a Pauli triplet up to a sign. (It follows that changing sign we can always reach the situation where the first three columns of table (2.14) form Pauli triplets. $A(0, 3)$ and $A(1, 3)$ anti-commute, but it may happen that $A(0, 3)A(1, 3) = -iA(2, 3)$.)



This picture shows a family $\{\mathcal{A}_i : 0 \leq i \leq 3\}$ of pairwise quasi-orthogonal subalgebras of $\mathbb{M}_4(\mathbb{C})$ which are isomorphic to $\mathbb{M}_2(\mathbb{C})$. The edges between two vertices represent the one-dimensional traceless intersection of the two subalgebras corresponding to two vertices. The three edges starting from a vertex represent a Pauli triplet.

Let $C_0 := \{\pm A(i, j)A(j, i) : i \neq j\} \cup \{\pm I\}$ and $C := C_0 \cup iC_0$. We want to show that C is a commutative group (with respect to the multiplication of unitaries).

Note that the products in C_0 have factors in symmetric position in (2.14) with respect to the main diagonal indicated by stars. Moreover, $A(i, j) \in \mathcal{A}(j)$ and $A(j, k) \in \mathcal{A}(j)'$, and these operators commute.

We have two cases for a product in C . Taking the product of $A(i, j)A(j, i)$ and $A(u, v)A(v, u)$, we have

$$(A(i, j)A(j, i))(A(i, j)A(j, i)) = I$$

in the simplest case, since $A(i, j)$ and $A(j, i)$ are commuting self-adjoint unitaries. The situation is slightly more complicated if the cardinality of the set $\{i, j, u, v\}$ is 3 or 4. First,

$$\begin{aligned} (A(1, 0)A(0, 1))(A(3, 0)A(0, 3)) &= A(0, 1)(A(1, 0)A(3, 0))A(0, 3) \\ &= \pm i(A(0, 1)A(2, 0))A(0, 3) \\ &= \pm iA(2, 0)(A(0, 1)A(0, 3)) \\ &= \pm A(2, 0)A(0, 2), \end{aligned}$$

and secondly,

$$\begin{aligned} (A(1, 0)A(0, 1))(A(3, 2)A(2, 3)) &= \pm iA(1, 0)A(0, 2)(A(0, 3)A(3, 2))A(2, 3) \\ &= \pm iA(1, 0)A(0, 2)A(3, 2)(A(0, 3)A(2, 3)) \\ &= \pm A(1, 0)(A(0, 2)A(3, 2))A(1, 3) \\ &= \pm iA(1, 0)(A(1, 2)A(1, 3)) \\ &= \pm A(1, 0)A(1, 0) = \pm I. \end{aligned} \tag{2.15}$$

So the product of any two operators in \mathcal{C} is again in \mathcal{C} .

Now we show that the subalgebra \mathcal{C} linearly spanned by the unitaries $\{A(i, j)A(j, i) : i \neq j\} \cup \{I\}$ is a maximal Abelian subalgebra. Since we know the commutativity of this algebra, we estimate the dimension. It follows from (2.15) and the self-adjointness of $A(i, j)A(j, i)$ that

$$A(i, j)A(j, i) = \pm A(k, \ell)A(\ell, k)$$

when i, j, k and ℓ are different. Therefore \mathcal{C} is linearly spanned by $A(0, 1)A(1, 0)$, $A(0, 2)A(2, 0)$, $A(0, 3)A(3, 0)$ and I . These are four different self-adjoint unitaries.

Finally, we check that the subalgebra \mathcal{C} is quasi-orthogonal to $\mathcal{A}(i)$. If the cardinality of the set $\{i, j, k, \ell\}$ is 4, then we have

$$\text{Tr } A(i, j)(A(i, j)A(j, i)) = \text{Tr } A(j, i) = 0$$

and

$$\text{Tr } A(k, \ell)A(i, j)A(j, i) = \pm \text{Tr } A(k, \ell)A(k, \ell)A(\ell, k) = \pm \text{Tr } A(\ell, k) = 0.$$

Moreover, because $\mathcal{A}(k)$ is quasi-orthogonal to $\mathcal{A}(i)$, we also have $A(i, k) \perp A(j, i)$, so

$$\text{Tr } A(i, \ell)(A(i, j)A(j, i)) = \pm i \text{Tr } A(i, k)A(j, i) = 0.$$

From this we can conclude that

$$A(k, \ell) \perp A(i, j)A(j, i)$$

for all $k \neq \ell$ and $i \neq j$. □

2.5 Kernel Functions

Let \mathcal{X} be a non-empty set. A function $\psi : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ is often called a **kernel**. A kernel $\psi : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ 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 $\{c_1, c_2, \dots, c_n\} \subset \mathbb{C}$ and $\{x_1, x_2, \dots, x_n\} \subset \mathcal{X}$.

Example 2.40 It follows from Schur's theorem that the product of positive definite kernels is a positive definite kernel as well.

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

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

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

The function $\psi : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ is called a **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 $\{c_1, c_2, \dots, c_n\} \subset \mathbb{C}$ and $\{x_1, x_2, \dots, x_n\} \subset \mathcal{X}$ when $\sum_{j=1}^n c_j = 0$.

The above properties of a kernel depend on the matrices

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

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

Lemma 2.41 Assume that the function $\psi : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ 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 the following:

Example 2.42 Assume that $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ is a C^1 -function with the property $f(0) = f'(0) = 0$. Let $\psi : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}$ be defined as

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

(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 $\varepsilon \rightarrow 0$, we have the positive definite kernel

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

Assume that $f(x) > 0$ for all $x > 0$. 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 a divided difference of the function $g(x) := x^2/f(x)$ on $(0, \infty)$. □

Theorem 2.43 (Schoenberg's theorem) Let \mathcal{X} be a non-empty set and let $\psi : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ 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 conditionally negative definite and so is

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

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

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

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 proves the case $t = 1$, and the argument is similar for general $t > 0$. \square

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

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

Therefore the results of this section have matrix consequences.

2.6 Positivity-Preserving Mappings

Let $\alpha : \mathbb{M}_n \rightarrow \mathbb{M}_k$ be a linear mapping. It is called **positive** (or positivity-preserving) if it sends positive (semidefinite) matrices to positive (semidefinite) matrices. α is **unital** if $\alpha(I_n) = I_k$.

The **dual** $\alpha^* : \mathbb{M}_k \rightarrow \mathbb{M}_n$ of α is defined by the equation

$$\text{Tr } \alpha(A)B = \text{Tr } A\alpha^*(B) \quad (A \in \mathbb{M}_n, B \in \mathbb{M}_k).$$

It is easy to see that α is positive if and only if α^* is positive and α is trace-preserving if and only if α^* is unital.

The inequality

$$\alpha(AA^*) \geq \alpha(A)\alpha(A)^*$$

is called the **Schwarz inequality**. If the Schwarz inequality holds for a linear mapping α , then α is positivity-preserving. If α is a positive mapping, then this inequality holds for normal matrices. This result is called the **Kadison inequality**.

Theorem 2.44 *Let $\alpha : \mathbb{M}_n(\mathbb{C}) \rightarrow \mathbb{M}_k(\mathbb{C})$ be a positive unital mapping.*

(1) *If $A \in \mathbb{M}_n$ is a normal operator, then*

$$\alpha(AA^*) \geq \alpha(A)\alpha(A)^*.$$

(2) *If $A \in \mathbb{M}_n$ is positive such that A and $\alpha(A)$ are invertible, then*

$$\alpha(A^{-1}) \geq \alpha(A)^{-1}.$$

Proof: A has a spectral decomposition $\sum_i \lambda_i P_i$, where the P_i 's are pairwise orthogonal projections. We have $A^*A = \sum_i |\lambda_i|^2 P_i$ and

$$\begin{bmatrix} I & \alpha(A) \\ \alpha(A)^* & \alpha(A^*A) \end{bmatrix} = \sum_i \begin{bmatrix} 1 & \lambda_i \\ \overline{\lambda_i} & |\lambda_i|^2 \end{bmatrix} \otimes \alpha(P_i).$$

Since $\alpha(P_i)$ is positive, the left-hand side is positive as well. Reference to Theorem 2.1 gives the first inequality.

To prove the second inequality, use the identity

$$\begin{bmatrix} \alpha(A) & I \\ I & \alpha(A^{-1}) \end{bmatrix} = \sum_i \begin{bmatrix} \lambda_i & 1 \\ 1 & \lambda_i^{-1} \end{bmatrix} \otimes \alpha(P_i)$$

to conclude that the left-hand side is a positive block matrix. The positivity implies our statement. \square

Corollary 2.45 *A positive unital mapping $\alpha : \mathbb{M}_n(\mathbb{C}) \rightarrow \mathbb{M}_k(\mathbb{C})$ has norm 1, i.e., $\|\alpha(A)\| \leq \|A\|$ for every $A \in \mathbb{M}_n(\mathbb{C})$.*

Proof: Let $A \in \mathbb{M}_n(\mathbb{C})$ be such that $\|A\| \leq 1$, and take the polar decomposition $A = U|A|$ with a unitary U . By Example 1.39 there is a unitary V such that $|A| = (V + V^*)/2$ and so $A = (UV + UV^*)/2$. Hence it suffices to show that $\|\alpha(U)\| \leq 1$ for every unitary U . This follows from the Kadison inequality in (1) of the previous theorem as

$$\|\alpha(U)\|^2 = \|\alpha(U)^* \alpha(U)\| \leq \|\alpha(U^*U)\| = \|\alpha(I)\| = 1.$$

\square

The linear mapping $\alpha : \mathbb{M}_n \rightarrow \mathbb{M}_k$ is called **2-positive** if

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \geq 0 \quad \text{implies} \quad \begin{bmatrix} \alpha(A) & \alpha(B) \\ \alpha(B^*) & \alpha(C) \end{bmatrix} \geq 0$$

when $A, B, C \in \mathbb{M}_n$.

Lemma 2.46 *Let $\alpha : \mathbb{M}_n(\mathbb{C}) \rightarrow \mathbb{M}_k(\mathbb{C})$ be a 2-positive mapping. If $A, \alpha(A) > 0$, then*

$$\alpha(B)^* \alpha(A)^{-1} \alpha(B) \leq \alpha(B^* A^{-1} B)$$

for every $B \in \mathbb{M}_n$. Hence, a 2-positive unital mapping satisfies the Schwarz inequality.

Proof: Since

$$\begin{bmatrix} A & B \\ B^* & B^*A^{-1}B \end{bmatrix} \geq 0,$$

the 2-positivity implies

$$\begin{bmatrix} \alpha(A) & \alpha(B) \\ \alpha(B^*) & \alpha(B^*A^{-1}B) \end{bmatrix} \geq 0.$$

So Theorem 2.1 implies the statement. \square

If $B = B^*$, then the 2-positivity condition is not necessary in the previous lemma, positivity is enough.

Lemma 2.47 *Let $\alpha : \mathbb{M}_n \rightarrow \mathbb{M}_k$ be a 2-positive unital mapping. Then*

$$\mathcal{N}_\alpha := \{A \in \mathbb{M}_n : \alpha(A^*A) = \alpha(A)^*\alpha(A) \text{ and } \alpha(AA^*) = \alpha(A)\alpha(A)^*\}$$

is a subalgebra of \mathbb{M}_n and

$$\alpha(AB) = \alpha(A)\alpha(B) \text{ and } \alpha(BA) = \alpha(B)\alpha(A)$$

holds for all $A \in \mathcal{N}_\alpha$ and $B \in \mathbb{M}_n$.

Proof: The proof is based only on the Schwarz inequality. Assume that $\alpha(AA^*) = \alpha(A)\alpha(A)^*$. Then

$$\begin{aligned} & t(\alpha(A)\alpha(B) + \alpha(B)^*\alpha(A)^*) \\ &= \alpha(tA^* + B)^*\alpha(tA^* + B) - t^2\alpha(A)\alpha(A)^* - \alpha(B)^*\alpha(B) \\ &\leq \alpha((tA^* + B)^*(tA^* + B)) - t^2\alpha(AA^*) - \alpha(B)^*\alpha(B) \\ &= t\alpha(AB + B^*A^*) + \alpha(B^*B) - \alpha(B)^*\alpha(B) \end{aligned}$$

for a real t . Divide the inequality by t and let $t \rightarrow \pm\infty$. Then

$$\alpha(A)\alpha(B) + \alpha(B)^*\alpha(A)^* = \alpha(AB + B^*A^*)$$

and similarly

$$\alpha(A)\alpha(B) - \alpha(B)^*\alpha(A)^* = \alpha(AB - B^*A^*).$$

Adding these two equalities we have

$$\alpha(AB) = \alpha(A)\alpha(B).$$

The other identity is proven similarly. \square

It follows from the previous lemma that if α is a 2-positive unital mapping and its inverse is 2-positive as well, then α is multiplicative. Indeed, the assumption implies $\alpha(A^*A) = \alpha(A)^*\alpha(A)$ for every A .

A linear mapping $\mathcal{E} : \mathbb{M}_n \rightarrow \mathbb{M}_k$ is called **completely positive** if

$$\mathcal{E} \otimes \text{id}_n : \mathbb{M}_n \otimes \mathbb{M}_n \rightarrow \mathbb{M}_k \otimes \mathbb{M}_n$$

is a positive mapping, where $\text{id}_n : \mathbb{M}_n \rightarrow \mathbb{M}_n$ is the identity mapping and $\mathcal{E} \otimes \text{id}_n$ is defined by

$$(\mathcal{E} \otimes \text{id}_n)([X_{ij}]_{i,j=1}^n) := [\mathcal{E}(X_{ij})]_{i,j=1}^n.$$

(Here, $B(\mathcal{H}) \otimes \mathbb{M}_n$ is identified with the $n \times n$ block matrices whose entries are operators in $B(\mathcal{H})$.) Note that if a linear mapping $\mathcal{E} : \mathbb{M}_n \rightarrow \mathbb{M}_k$ is completely positive in the above sense, then $\mathcal{E} \otimes \text{id}_m : \mathbb{M}_n \otimes \mathbb{M}_m \rightarrow \mathbb{M}_k \otimes \mathbb{M}_m$ is positive for every $m \in \mathbb{N}$.

Example 2.48 Consider the transpose mapping $\mathcal{E} : A \mapsto A^t$ on 2×2 matrices:

$$\begin{bmatrix} x & y \\ z & w \end{bmatrix} \mapsto \begin{bmatrix} x & z \\ y & w \end{bmatrix}.$$

\mathcal{E} is obviously positive. The matrix

$$\begin{bmatrix} 2 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$$

is positive. The extension of \mathcal{E} maps this to

$$\begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}.$$

This is not positive, so \mathcal{E} is not completely positive. □

Theorem 2.49 Let $\mathcal{E} : \mathbb{M}_n \rightarrow \mathbb{M}_k$ be a linear mapping. Then the following conditions are equivalent:

- (1) \mathcal{E} is completely positive.
- (2) The block matrix X defined by

$$X_{ij} = \mathcal{E}(E(ij)) \quad (1 \leq i, j \leq n) \tag{2.16}$$

is positive, where $E(ij)$ are the matrix units of \mathbb{M}_n .

- (3) There are operators $V_t : \mathbb{C}^n \rightarrow \mathbb{C}^k$ ($1 \leq t \leq k^2$) such that

$$\mathcal{E}(A) = \sum_t V_t A V_t^*. \quad (2.17)$$

(4) For finite families $A_i \in \mathbb{M}_n(\mathbb{C})$ and $B_i \in \mathbb{M}_k(\mathbb{C})$ ($1 \leq i \leq n$), the inequality

$$\sum_{i,j} B_i^* \mathcal{E}(A_i^* A_j) B_j \geq 0$$

holds.

Proof: (1) implies (2): The matrix

$$\sum_{i,j} E(ij) \otimes E(ij) = \frac{1}{n} \left(\sum_{i,j} E(ij) \otimes E(ij) \right)^2$$

is positive. Therefore,

$$(\text{id}_n \otimes \mathcal{E}) \left(\sum_{i,j} E(ij) \otimes E(ij) \right) = \sum_{i,j} E(ij) \otimes \mathcal{E}(E(ij)) = X$$

is positive as well.

(2) implies (3): Assume that the block matrix X is positive. There are orthogonal projections P_i ($1 \leq i \leq n$) on \mathbb{C}^{nk} such that 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 we define $V_t : \mathbb{C}^n \rightarrow \mathbb{C}^k$ by

$$V_t |i\rangle = P_i |f_t\rangle.$$

($|i\rangle$ 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 hence

$$\mathcal{E}(E(ij)) = P_i X P_j = \sum_t V_t E(ij) V_t^*.$$

Since this holds for all matrix units $E(ij)$, we obtain

$$\mathcal{E}(A) = \sum_t V_t A V_t^*.$$

(3) implies (4): Assume that \mathcal{E} is of the form (2.17). Then

$$\begin{aligned} \sum_{i,j} B_i^* \mathcal{E}(A_i^* A_j) B_j &= \sum_t \sum_{i,j} B_i^* V_t (A_i^* A_j) V_t^* B_j \\ &= \sum_t \left(\sum_i A_i V_t^* B_i \right)^* \left(\sum_j A_j V_t^* B_j \right) \geq 0 \end{aligned}$$

follows.

(4) implies (1): We consider

$$\mathcal{E} \otimes \text{id}_n : \mathbb{M}_n \otimes \mathbb{M}_n \rightarrow \mathbb{M}_k \otimes \mathbb{M}_n.$$

Since any positive operator in $\mathbb{M}_n \otimes \mathbb{M}_n$ is the sum of operators in the form $\sum_{i,j} A_i^* A_j \otimes E(ij)$ (Theorem 2.8), it is enough to show that

$$Y := \mathcal{E} \otimes \text{id}_n \left(\sum_{i,j} A_i^* A_j \otimes E(ij) \right) = \sum_{i,j} \mathcal{E}(A_i^* A_j) \otimes E(ij)$$

is positive. On the other hand, $Y = [Y_{ij}]_{i,j=1}^n \in \mathbb{M}_k \otimes \mathbb{M}_n$ is positive if and only if

$$\sum_{i,j} B_i^* Y_{ij} B_j = \sum_{i,j} B_i^* \mathcal{E}(A_i^* A_j) B_j \geq 0.$$

The positivity of this operator is assumed in (4). Hence (1) follows. \square

The representation (2.17) is called the **Kraus representation**. The block matrix X defined by (2.16) is called the **representing block matrix** (or the **Choi matrix**).

Example 2.50 We take $\mathcal{A} \subset \mathcal{B} \subset \mathbb{M}_n(\mathbb{C})$ and a conditional expectation $\mathcal{E} : \mathcal{B} \rightarrow \mathcal{A}$. Using condition (4) of the previous theorem we can argue that \mathcal{E} is completely positive. For $A_i \in \mathcal{A}$ and $B_i \in \mathcal{B}$ we have

$$\sum_{i,j} A_i^* \mathcal{E}(B_i^* B_j) A_j = \mathcal{E}\left(\left(\sum_i B_i A_i\right)^* \left(\sum_j B_j A_j\right)\right) \geq 0$$

and this is enough. \square

The next example is slightly different.

Example 2.51 Let \mathcal{H} and \mathcal{K} be Hilbert spaces and (f_i) be a basis in \mathcal{K} . For each i define the linear operator $V_i : \mathcal{H} \rightarrow \mathcal{H} \otimes \mathcal{K}$ by $V_i e = e \otimes f_i$ ($e \in \mathcal{H}$). These operators are isometries with pairwise orthogonal ranges and the adjoints act as $V_i^*(e \otimes f) = \langle f_i, f \rangle e$.

The **partial trace** $\text{Tr}_2 : B(\mathcal{H} \otimes \mathcal{K}) \rightarrow B(\mathcal{H})$ introduced in Sect. 1.7 can be written as

$$\text{Tr}_2(A) = \sum_i V_i^* A V_i \quad (A \in B(\mathcal{H} \otimes \mathcal{K})).$$

The reason for the terminology is the formula $\text{Tr}_2(X \otimes Y) = X \text{Tr } Y$. The above expression implies that Tr_2 is completely positive. It is actually a conditional expectation up to a constant factor. \square

Example 2.52 The trace $\text{Tr} : \mathbb{M}_k(\mathbb{C}) \rightarrow \mathbb{C}$ is completely positive if $\text{Tr} \otimes \text{id}_n : \mathbb{M}_k(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C}) \rightarrow \mathbb{M}_n(\mathbb{C})$ is a positive mapping. However, this is a partial trace which is known to be positive (even completely positive).

It follows that any positive linear functional $\psi : \mathbb{M}_k(\mathbb{C}) \rightarrow \mathbb{C}$ is completely positive. Since $\psi(A) = \text{Tr } DA$ for some positive D , ψ is the composition of the completely positive mappings $A \mapsto D^{1/2} A D^{1/2}$ and Tr . \square

Example 2.53 Let $\mathcal{E} : \mathbb{M}_n \rightarrow \mathbb{M}_k$ be a positive linear mapping such that $\mathcal{E}(A)$ and $\mathcal{E}(B)$ commute for any $A, B \in \mathbb{M}_n$. We want to show that \mathcal{E} is completely positive.

Any two self-adjoint matrices in the range of \mathcal{E} commute, so we can change the basis so that all of them become diagonal. It follows that \mathcal{E} has the form

$$\mathcal{E}(A) = \sum_i \psi_i(A) E_{ii},$$

where E_{ii} are the diagonal matrix units and ψ_i are positive linear functionals. Since the sum of completely positive mappings is completely positive, it is enough to show that $A \mapsto \psi(A)F$ is completely positive for a positive functional ψ and for a positive matrix F . The complete positivity of this mapping means that for an $m \times m$ block matrix X with entries $X_{ij} \in \mathbb{M}_n$, if $X \geq 0$ then the block matrix $[\psi(X_{ij})F]_{i,j=1}^m$ should be positive. This is true, since the matrix $[\psi(X_{ij})]_{i,j=1}^m$ is positive (due to the complete positivity of ψ). \square

Example 2.54 A linear mapping $\mathcal{E} : \mathbb{M}_2 \rightarrow \mathbb{M}_2$ is defined by the formula

$$\mathcal{E} : \begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix} \mapsto \begin{bmatrix} 1+\gamma z & \alpha x - i\beta y \\ \alpha x + i\beta y & 1-\gamma z \end{bmatrix}$$

where α, β, γ are real parameters.

The condition for positivity is

$$-1 \leq \alpha, \beta, \gamma \leq 1.$$

It is not difficult to compute the representing block matrix as follows:

$$X = \frac{1}{2} \begin{bmatrix} 1+\gamma & 0 & 0 & \alpha+\beta \\ 0 & 1-\gamma & \alpha-\beta & 0 \\ 0 & \alpha-\beta & 1-\gamma & 0 \\ \alpha+\beta & 0 & 0 & 1+\gamma \end{bmatrix}.$$

This matrix is positive if and only if

$$|1 \pm \gamma| \geq |\alpha \pm \beta|.$$

In quantum information theory this mapping \mathcal{E} is called the **Pauli channel**. □

Example 2.55 Fix a positive definite matrix $A \in \mathbb{M}_n$ and set

$$T_A(K) = \int_0^\infty (t+A)^{-1} K (t+A)^{-1} dt \quad (K \in \mathbb{M}_n).$$

This mapping $T_A : \mathbb{M}_n \rightarrow \mathbb{M}_n$ is obviously positivity-preserving and approximation of the integral by a finite sum also shows the complete positivity.

If $A = \text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, then we see from integration that the entries of $T_A(K)$ are

$$T_A(K)_{ij} = \frac{\log \lambda_i - \log \lambda_j}{\lambda_i - \lambda_j} K_{ij}.$$

Another integration gives that the mapping

$$\alpha : L \mapsto \int_0^1 A^t L A^{1-t} dt$$

acts as

$$(\alpha(L))_{ij} = \frac{\lambda_i - \lambda_j}{\log \lambda_i - \log \lambda_j} L_{ij}.$$

This shows that

$$T_A^{-1}(L) = \int_0^1 A^t L A^{1-t} dt.$$

To show that T_A^{-1} is not positive, we take $n = 2$ and consider

$$T_A^{-1} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & \frac{\lambda_1 - \lambda_2}{\log \lambda_1 - \log \lambda_2} \\ \frac{\lambda_1 - \lambda_2}{\log \lambda_1 - \log \lambda_2} & \lambda_2 \end{bmatrix}.$$

The positivity of this matrix is equivalent to the inequality

$$\sqrt{\lambda_1 \lambda_2} \geq \frac{\lambda_1 - \lambda_2}{\log \lambda_1 - \log \lambda_2}$$

between the geometric and logarithmic means. The opposite inequality holds, see Example 5.22, and therefore T_A^{-1} is not positive. \square

The next result tells us that the **Kraus representation** of a completely positive mapping is unique up to a unitary matrix.

Theorem 2.56 *Let $\mathcal{E} : \mathbb{M}_n(\mathbb{C}) \rightarrow \mathbb{M}_m(\mathbb{C})$ be a linear mapping which is represented as*

$$\mathcal{E}(A) = \sum_{t=1}^k V_t A V_t^* \quad \text{and} \quad \mathcal{E}(A) = \sum_{t=1}^k W_t A W_t^*$$

with operators $V_t, W_t : \mathbb{C}^n \rightarrow \mathbb{C}^m$. Then there exists a $k \times k$ unitary matrix $[c_{tu}]$ such that

$$W_t = \sum_u c_{tu} V_u \quad (1 \leq t \leq k).$$

Proof: Without loss of generality we may assume that $m \geq n$. Indeed, we can embed $\mathbb{M}_m = B(\mathbb{C}^m)$ into a bigger $\mathbb{M}_{m'} = B(\mathbb{C}^{m'})$ and consider \mathcal{E} as a mapping $\mathbb{M}_n \rightarrow \mathbb{M}_{m'}$. Let x_i be a basis in \mathbb{C}^m and y_j be a basis in \mathbb{C}^n . Consider the vectors

$$v_t := \sum_{j=1}^n x_j \otimes V_t y_j \quad \text{and} \quad w_t := \sum_{j=1}^n x_j \otimes W_t y_j.$$

We have

$$|v_t\rangle\langle v_t| = \sum_{j,j'} |x_j\rangle\langle x_{j'}| \otimes V_t |y_j\rangle\langle y_{j'}| V_t^*$$

and

$$|w_t\rangle\langle w_t| = \sum_{j,j'} |x_j\rangle\langle x_{j'}| \otimes W_t |y_j\rangle\langle y_{j'}| W_t^*.$$

Our hypothesis implies that

$$\sum_t |v_t\rangle\langle v_t| = \sum_t |w_t\rangle\langle w_t|.$$

Lemma 1.24 tells us that there is a unitary matrix $[c_{tu}]$ such that

$$w_t = \sum_u c_{tu} v_u.$$

This implies that

$$W_t y_j = \sum_u c_{tu} V_u y_j \quad (1 \leq j \leq n).$$

Hence we conclude the statement of the theorem. \square

2.7 Notes and Remarks

Theorem 2.5 is from the paper J.-C. Bourin and E.-Y. Lee, Unitary orbits of Hermitian operators with convex or concave functions, *Bull. London Math. Soc.* **44**(2012), 1085–1102.

The **Wielandt inequality** has an extension to matrices. Let A be an $n \times n$ positive matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Let X and Y be $n \times p$ and $n \times q$ matrices such that $X^*Y = 0$. The generalized inequality is

$$X^*AY(Y^*AY)^-Y^*AX \leq \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n} \right)^2 X^*AX,$$

where a generalized inverse $(Y^*AY)^-$ is included: $BB^-B = B$. See Song-Gui Wang and Wai-Cheung Ip, A matrix version of the Wielandt inequality and its applications to statistics, *Linear Algebra Appl.* **296**(1999), 171–181.

The lattice of ortho-projections has applications in quantum theory. The cited **Gleason theorem** was obtained by A. M. Gleason in 1957, see also R. Cooke, M. Keane and W. Moran, An elementary proof of Gleason's theorem, *Math. Proc. Cambridge Philos. Soc.* **98**(1985), 117–128.

Theorem 2.27 is from the paper D. Petz and G. Tóth, Matrix variances with projections, *Acta Sci. Math. (Szeged)*, **78**(2012), 683–688. An extension of this result is in the paper Z. Léka and D. Petz, Some decompositions of matrix variances, to be published.

Theorem 2.33 is the double commutant theorem of **von Neumann** from 1929; the original proof was for operators on an infinite-dimensional Hilbert space. (There is a relevant difference between finite and infinite dimensions; in a finite-dimensional space all subspaces are closed.) The conditional expectation in Theorem 2.34 was first introduced in the paper H. Umegaki, Conditional expectation in an operator algebra,

Tôhoku Math. J. **6**(1954), 177–181, and it is related to the so-called **Tomiyama theorem**.

The maximum number of complementary MASAs in $\mathbb{M}_n(\mathbb{C})$ is a popular subject. If n is a prime power, then $n + 1$ MASAs can be constructed, but $n = 6$ is an unknown problematic case. (The expected number of complementary MASAs is 3 here.) It is interesting that n MASAs cannot exist in $\mathbb{M}_n(\mathbb{C})$ for any $n > 1$, see the paper [83] of M. Weiner.

Theorem 2.39 is from the paper H. Ohno, D. Petz and A. Szántó, Quasi-orthogonal subalgebras of 4×4 matrices, Linear Algebra Appl. **425**(2007), 109–118. It was conjectured that in the case $n = 2^k$ the algebra $\mathbb{M}_n(\mathbb{C})$ cannot have $N_k := (4^k - 1)/3$ complementary subalgebras isomorphic to \mathbb{M}_2 , but it was proved that there are $N_k - 1$ copies. 2 is not a typical prime number in this situation. If $p > 2$ is a prime number, then in the case $n = p^k$ the algebra $\mathbb{M}_n(\mathbb{C})$ has $N_k := (p^{2k} - 1)/(p^2 - 1)$ complementary subalgebras isomorphic to \mathbb{M}_p , see the paper H. Ohno, Quasi-orthogonal subalgebras of matrix algebras, Linear Algebra Appl. **429**(2008), 2146–2158.

Positive and conditionally negative definite kernel functions 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, vol. 100. Springer, New York, 1984. (It is noteworthy that conditionally negative definite is called there ‘negative definite’.)

2.8 Exercises

1. Show that

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \geq 0$$

if and only if $B = A^{1/2} Z C^{1/2}$ for a matrix Z with $\|Z\| \leq 1$.

2. Let $X, U, V \in \mathbb{M}_n$ and assume that U and V are unitaries. Prove that

$$\begin{bmatrix} I & U & X \\ U^* & I & V \\ X^* & V^* & I \end{bmatrix} \geq 0$$

if and only if $X = UV$.

3. Show that for $A, B \in \mathbb{M}_n$ the formula

$$\begin{bmatrix} I & A \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \begin{bmatrix} I & A \\ 0 & I \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}$$

holds. Conclude that AB and BA have the same eigenvectors.

4. Assume that $0 < A \in \mathbb{M}_n$. Show that $A + A^{-1} \geq 2I$.

5. Assume that

$$A = \begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix} \geq 0.$$

Show that $\det A \leq \det A_1 \times \det A_2$.

6. Assume that the eigenvalues of the self-adjoint matrix

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix}$$

are $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ and the eigenvalues of A are $\beta_1 \leq \beta_2 \leq \dots \leq \beta_m$. Show that

$$\lambda_i \leq \beta_i \leq \lambda_{i+n-m}.$$

7. Show that a matrix $A \in \mathbb{M}_n$ is irreducible if and only if for every $1 \leq i, j \leq n$ there is a power k such that $(A^k)_{ij} \neq 0$.

8. Let $A, B, C, D \in \mathbb{M}_n$ and $AC = CA$. Show that

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(AD - CB).$$

9. Let $A, B, C \in \mathbb{M}_n$ and

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \geq 0.$$

Show that $B^* \circ B \leq A \circ C$.

10. Let $A, B \in \mathbb{M}_n$. Show that $A \circ B$ is a submatrix of $A \otimes B$.

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

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

13. Let $A_1, A_2, \dots, A_k \in \mathbb{M}_n^{sa}$ and $A_1 + A_2 + \dots + A_k = I$. Show that the following statements are equivalent:

- (1) All operators A_i are projections.
- (2) For all $i \neq j$ the product $A_i A_j = 0$ holds.
- (3) $\text{rank}(A_1) + \text{rank}(A_2) + \dots + \text{rank}(A_k) = n$.

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

15. The matrix $M \in \mathbb{M}_n(\mathbb{C})$ is defined as

$$M_{ij} = \min\{i, j\}.$$

Show that M is positive.

16. Let $A \in \mathbb{M}_n$ and define the mapping $S_A : \mathbb{M}_n \rightarrow \mathbb{M}_n$ by $S_A : B \mapsto A \circ B$. Show that the following statements are equivalent.

- (1) A is positive.
- (2) $S_A : \mathbb{M}_n \rightarrow \mathbb{M}_n$ is positive.
- (3) $S_A : \mathbb{M}_n \rightarrow \mathbb{M}_n$ is completely positive.

17. Let A, B, C be operators on a Hilbert space \mathcal{H} and $A, C \geq 0$. Show that

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \geq 0$$

if and only if $|\langle Bx, y \rangle| \leq \langle Ay, y \rangle \cdot \langle Cx, x \rangle$ for every $x, y \in \mathcal{H}$.

18. Let $P \in \mathbb{M}_n$ be idempotent, i.e. $P^2 = P$. Show that P is an ortho-projection if and only if $\|P\| \leq 1$.
19. Let $P \in \mathbb{M}_n$ be an ortho-projection and $0 < A \in \mathbb{M}_n$. Prove the following formulas:

$$[P](A^2) \leq ([P]A)^2, \quad ([P]A)^{1/2} \leq [P](A^{1/2}), \quad [P](A^{-1}) \leq ([P]A)^\dagger.$$

20. 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 $\mathbb{R} \times \mathbb{R}$.

21. Show that the equality

$$A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C)$$

is not true for ortho-projections.

22. Assume that the kernel $\psi : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ 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.

23. Assume that the kernel $\psi : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ 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.

24. Show that the kernel $\psi(x, y) = (\sin(x - y))^2$ is negative semidefinite on $\mathbb{R} \times \mathbb{R}$.
25. Show that the linear mapping $\mathcal{E}_{p,n} : \mathbb{M}_n \rightarrow \mathbb{M}_n$ defined as

$$\mathcal{E}_{p,n}(A) = pA + (1-p)\frac{I}{n}\text{Tr } A$$

is completely positive if and only if

$$-\frac{1}{n^2-1} \leq p \leq 1.$$

26. Show that the linear mapping $\mathcal{E} : \mathbb{M}_n \rightarrow \mathbb{M}_n$ defined as

$$\mathcal{E}(D) = \frac{1}{n-1}(\text{Tr}(D)I - D^t)$$

is a completely positive unital mapping. (Here D^t denotes the transpose of D .) Show that \mathcal{E} has a negative eigenvalue. (This mapping is called the **Holevo-Werner channel**.)

27. Define $\mathcal{E} : \mathbb{M}_n \rightarrow \mathbb{M}_n$ by

$$\mathcal{E}(A) = \frac{1}{n-1}(I \text{Tr } A - A).$$

Show that \mathcal{E} is positive but not completely positive.

28. Let p be a real number. Show that the mapping $\mathcal{E}_{p,2} : \mathbb{M}_2 \rightarrow \mathbb{M}_2$ defined as

$$\mathcal{E}_{p,2}(A) = pA + (1-p)\frac{I}{2}\text{Tr } A$$

is positive if and only if $-1 \leq p \leq 1$. Show that $\mathcal{E}_{p,2}$ is completely positive if and only if $-1/3 \leq p \leq 1$.

29. Show that $\|(f_1, f_2)\|^2 = \|f_1\|^2 + \|f_2\|^2$.
 30. Give the analogue of Theorem 2.1 when C is assumed to be invertible.
 31. 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.

32. Let $\dim \mathcal{H} = 2$ and $0 \leq A, B \in B(\mathcal{H})$. Show that there is an orthogonal basis such that

$$A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \quad B = \begin{bmatrix} c & d \\ d & e \end{bmatrix}$$

with positive numbers $a, b, c, d, e \geq 0$.

33. 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 $-A \leq B \leq A$.

34. Determine the inverses of the matrices

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & b \\ -d & c & -b & a \end{bmatrix}.$$

35. Give the analogue of the factorization (2.2) when D is assumed to be invertible.

36. 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

$$Q = A - BD^{-1}B^* - CE^{-1}C^*, \quad P = Q^{-1}BD^{-1}, \quad R = Q^{-1}CE^{-1}.$$

37. Find the determinant and the inverse of the block matrix

$$\begin{bmatrix} A & 0 \\ a & 1 \end{bmatrix}.$$

38. Let $A \in \mathbb{M}_n$ be an invertible matrix and $d \in \mathbb{C}$. Show that

$$\det \begin{bmatrix} A & b \\ c & d \end{bmatrix} = (d - cA^{-1}b)\det A$$

where $c = [c_1, \dots, c_n]$ and $b = [b_1, \dots, b_n]^t$.

39. Prove the concavity of the variance functional $\rho \mapsto \text{Var}_\rho(A)$ defined in (2.10). The concavity is

$$\text{Var}_\rho(A) \geq \sum_i \lambda_i \text{Var}_{\rho_i}(A) \quad \text{if} \quad \rho = \sum_i \lambda_i \rho_i$$

when $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$.

40. For $x, y \in \mathbb{R}^3$ and

$$x \cdot \sigma := \sum_{i=1}^3 x_i \sigma_i, \quad y \cdot \sigma := \sum_{i=1}^3 y_i \sigma_i$$

show that

$$(x \cdot \sigma)(y \cdot \sigma) = \langle x, y \rangle \sigma_0 + \mathbf{i}(x \times y) \cdot \sigma, \quad (2.18)$$

where $x \times y$ is the vectorial product in \mathbb{R}^3 .

<http://www.springer.com/978-3-319-04149-0>

Introduction to Matrix Analysis and Applications

Hiai, F.; Petz, D.

2014, VIII, 332 p. 3 illus., Softcover

ISBN: 978-3-319-04149-0