

Lectures on Quantum Information Theory and Quantum Statistics

Draft, not for circulation



Dénes Petz

Alfréd Rényi Institute of Mathematics

POB 127, H-1364 Budapest, Hungary



9. Coarse-grainings

May 9, 2007

A quantum mechanical system is described by an algebra \mathcal{M} of operators and the dynamical variables (or observables) correspond to the self-adjoint elements. The evolution of the system \mathcal{M} can be described in the **Heisenberg picture** in which an observable $A \in \mathcal{M}$ moves into $\alpha(A)$, where α is a linear transformation. α is induced by a unitary in case of the time evolution of a closed system but it could be the irreversible evolution of an open system. The **Schrödinger picture** is dual, it gives the transformation of the states:

$$\langle \alpha(A), \rho \rangle = \langle A, \mathcal{E}(\rho) \rangle,$$

where \mathcal{E} is the state transformation and the duality means $\langle B, \omega \rangle = \text{Tr} \omega B$. State transformation is an essential concept in quantum information theory and its role is the performance of information transfer. However, state transformation may appear in a different context as well. Shortly speaking, the dual of a state transformation will be called **coarse-graining**, in particular in a statistical context.

Assume that a level of observation does not allow to know the expectation value of all observables, but only a part of them. In our algebraic approach we assume that this part is the self-adjoint subalgebra \mathcal{N} of the full algebra \mathcal{M} . The positive linear embedding $\alpha : \mathcal{N} \rightarrow \mathcal{M}$ is a coarse-graining. It provides partial information of the total quantum system \mathcal{M} . If the algebra \mathcal{N} is “small” compared with \mathcal{M} , then loss of information takes place and the problem of statistical inference is to reconstruct the real state of \mathcal{M} from partial information.

9.1 Basic examples

Assume that the Hilbert space describing our quantum system is \mathcal{H} . A completely positive identity preserving linear mapping from an algebra \mathcal{N} to $B(\mathcal{H})$ will be called **coarse-graining**. It will be mostly assumed that \mathcal{N} is a subalgebra of $B(\mathcal{K})$, where \mathcal{K} is another Hilbert space. In the algebraic approach followed here, this Hilbert space is not specified always. It is well-known that a completely unit-preserving mapping α satisfies the **Schwarz inequality**

$$\alpha(A^*A) \geq \alpha(A)^* \alpha(A). \quad (9.1)$$

Example 9.1 The simplest example of coarse-graining appears if a component of a composite system $\mathcal{H} \equiv \mathcal{H}_1 \otimes \mathcal{H}_2$ is neglected. Assume that we restrict ourselves to the

observables of the first subsystem. Then the embedding

$$\alpha : B(\mathcal{H}_1) \rightarrow B(\mathcal{H}), \quad \alpha : A \mapsto A \otimes I_2$$

is the relevant coarse-graining. This is not only positive but it is a multiplicative isometry, therefore we call it embedding.

The dual of α is the **partial trace** Tr_2 :

$$\langle \alpha(A), B \otimes C \rangle = \text{Tr}(A^*B) \otimes C = \text{Tr} A^*B \text{Tr} C = \langle A, \mathcal{E}(B \otimes C) \rangle,$$

where \mathcal{E} defined as $\mathcal{E}(B \otimes C) = B \text{Tr} C$ is the so-called partial trace over the second factor. \square

Example 9.2 Consider a composite system of n identical particles: $\mathcal{H} \otimes \mathcal{H} \otimes \dots \otimes \mathcal{H}$ and assume that we restrict ourselves to the symmetric observables. Each permutation of the particles induces a unitary U and set $\mathcal{N} := \{A : UAU^* = A \text{ for every permutation unitary } U\}$. The embedding of \mathcal{N} is again a coarse-graining. The algebra \mathcal{N} is not isomorphic to a full matrix algebra but it is a subalgebra of $B(\mathcal{H}^{n\otimes})$. \square

The coarse-graining of the previous two examples were embeddings which are multiplicative. A general coarse-graining is not fully multiplicative but it satisfies a restricted multiplicativity.

Lemma 9.1 *Let $\alpha : \mathcal{N} \rightarrow \mathcal{M}$ be a coarse-graining. Then*

$$\mathcal{N}_\alpha := \{A \in \mathcal{N} : \alpha(A^*A) = \alpha(A)^*\alpha(A) \text{ and } \alpha(AA^*) = \alpha(A)\alpha(A)^*\} \quad (9.2)$$

is a subalgebra of \mathcal{N} and

$$\alpha(AB) = \alpha(A)\alpha(B) \quad \text{and} \quad \alpha(BA) = \alpha(B)\alpha(A) \quad (9.3)$$

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

Proof. The proof is based only on the Schwarz inequality (9.1). 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 inequalities we have

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

The other identity is proven similarly. \square

We call the subalgebra \mathcal{N}_α the **multiplicative domain** of α .

Example 9.3 The algebra of a quantum system is typically non-commutative but the mathematical formalism supports commutative algebras as well. A measurement is usually modelled by a positive partition of unity $(F_i)_{i=1}^n$, where F_i is a positive operator in $\mathcal{M} = B(\mathcal{H})$ and $\sum_i F_i = I$. The mapping $\beta : \mathbb{C}^n \rightarrow \mathcal{M}$, $(z_1, z_2, \dots, z_n) \mapsto \sum_i z_i F_i$ is positive and unital mapping from the commutative algebra \mathbb{C}^n to the non-commutative algebra \mathcal{M} . Every positive unital mappings $\mathbb{C}^n \rightarrow \mathcal{M}$ corresponds in this way to a certain measurement. Measurements are coarse-grainings from commutative algebras.

The n -tuple (z_1, z_2, \dots, z_n) can be regarded as a diagonal matrix and \mathcal{N} is a subalgebra of a matrix algebra. \square

Assume that \mathcal{A} is a subalgebra of $B(\mathcal{H})$. If ρ is a density matrix in $B(\mathcal{H})$, then the **reduced density matrix** $\rho_0 \in \mathcal{A}$ is uniquely determined by the equations

$$\mathrm{Tr} \rho_0 A = \mathrm{Tr} \rho A \quad \text{for every } A \in \mathcal{A}.$$

The mapping $\mathcal{E} : \rho \mapsto \rho_0$ is a state transformation and it is a sort of generalization of the partial trace. Sometimes $\rho|_{\mathcal{A}}$ is written instead of ρ_0 . For example,

$$S(\rho|_{\mathcal{A}} \| \omega|_{\mathcal{A}}) := S(\rho_0 \| \omega_0)$$

if ρ and ω are density matrices.

9.2 Conditional expectations

Conditional expectation is a transformation of the observables. Let \mathcal{A} be a subalgebra of $B(\mathcal{H})$ such that $\mathcal{A}^* = \mathcal{A}$ and $I_{\mathcal{H}} \in \mathcal{A}$. More generally, in place of $B(\mathcal{H})$ we can consider a matrix algebra \mathcal{B} . A **conditional expectation** $E : \mathcal{B} \rightarrow \mathcal{A}$ is a unital positive mapping which has the property

$$E(AB) = AE(B) \quad \text{for every } A \in \mathcal{A} \quad \text{and} \quad B \in \mathcal{B}. \quad (9.4)$$

Choosing $B = I$, we obtain that E acts identically on \mathcal{A} . It follows from the positivity of E that $E(B^*) = E(B)^*$. Therefore,

$$E(BA) = E((A^*B^*)^*) = E(A^*B^*)^* = (A^*E(B^*))^* = E(B^*)^*A = E(B)A \quad (9.5)$$

for every $A \in \mathcal{A}$ and $B \in \mathcal{B}$.

Example 9.4 Let \mathcal{B} be a matrix algebra and $\rho \in \mathcal{B}$ an invertible density matrix. Assume that $\alpha : \mathcal{B} \rightarrow \mathcal{B}$ is a coarse-graining such that

$$\mathrm{Tr} \rho \alpha(B) = \mathrm{Tr} \rho B \quad \text{for every } B \in \mathcal{B}.$$

Let $\mathcal{A} := \{A \in \mathcal{B} : \alpha(A) = A\}$ be the set of fixed points of α . First we show that \mathcal{A} is an algebra. Assume that $A \in \mathcal{A}$. Then

$$\mathrm{Tr} \rho A^* A = \mathrm{Tr} \rho \alpha(A^* A) \geq \mathrm{Tr} \rho \alpha(A)^* \alpha(A) = \mathrm{Tr} \rho A^* A.$$

Since ρ is invertible, $\alpha(A^*A) = A^*A$ and $A^*A \in \mathcal{A}$ and similarly $AA^* \in \mathcal{A}$. This gives that $\mathcal{A} \subset \mathcal{N}_\alpha$. Therefore for $A_1, A_2 \in \mathcal{A}$ we have $\alpha(A_1A_2) = \alpha(A_1)\alpha(A_2) = A_1A_2$ and $A_1A_2 \in \mathcal{A}$. Hence we conclude that \mathcal{A} is a subalgebra of \mathcal{N}_α .

When \mathcal{B} is endowed with the inner product $\langle B_1, B_2 \rangle := \text{Tr } \rho B_1^* B_2$, then α becomes a contraction and due to the **von Neumann ergodic theorem**

$$s_n(B) := \frac{1}{n}(B + \alpha(B) + \dots + \alpha^{n-1}(B)) \rightarrow E(B),$$

where $E : \mathcal{B} \rightarrow \mathcal{A}$ is an orthogonal projection. Since s_n 's are coarse-grainings, so is their limit E . In fact, E is a conditional expectation onto the fixed-point-algebra \mathcal{A} . It is easy to see that for $A \in \mathcal{A}$ and $B \in \mathcal{B}$, $s_n(AB) = A s_n(B)$. The limit $n \rightarrow \infty$ gives $E(AB) = AE(B)$.

A conditional expectation $E : \mathcal{B} \rightarrow \mathcal{A}$ was obtained as the limit of s_n under the condition that there is an invertible state left by α invariant. (We often refer to this result as the **Kovács-Szűcs theorem**.) \square

Heuristically, $E(B)$ is a kind of best approximation of B from \mathcal{A} . This is justified in the next example.

Example 9.5 Assume that τ is a linear functional on a matrix algebra \mathcal{B} such that

- (i) $\tau(B) \geq 0$ if $B \geq 0$,
- (ii) If $B \geq 0$ and $\tau(B) = 0$, then $B = 0$,
- (iii) $\tau(B_1 B_2) = \tau(B_2 B_1)$.

(These conditions say that τ is a positive, faithful and tracial functional.) \mathcal{B} becomes a Hilbert space when it is endowed with the inner product

$$\langle B_1, B_2 \rangle := \tau(B_1^* B_2).$$

Recall that $B \in \mathcal{B}$ is positive if (and only if) $\tau(BB_1) \geq 0$ for every positive B_1 .

Let \mathcal{A} be a unital subalgebra of \mathcal{B} . We claim that the orthogonal projection E with respect to the above defined inner product is a conditional expectation of \mathcal{B} onto \mathcal{A} .

Since $I \in \mathcal{A}$, we have $E(I) = I$ and E is unital. Let $B \in \mathcal{B}$ be positive. To show the positivity of E , we have to show that $E(B) \geq 0$. We have

$$\tau(A_0 E(B)) = \langle A_0^*, E(B) \rangle = \langle E(A_0^*), B \rangle = \langle A_0^*, B \rangle = \tau(A_0 B) \geq 0$$

for every positive $A_0 \in \mathcal{A}$. It follows that $E(B) \in \mathcal{A}$ is positive.

Condition (9.4) is equivalent to

$$\tau(A_1 E(AB)) = \tau(A_1 A E(B))$$

for every $A_1 \in \mathcal{A}$. This is true since

$$\tau(A_1 E(AB)) = \langle A_1^* A, E(AB) \rangle = \langle E(A_1^*), AB \rangle = \langle A_1^*, AB \rangle = \tau(A_1 AB)$$

and

$$\tau(A_1AE(B)) = \langle A^*A_1^*A, E(B) \rangle = \langle E(A^*A_1^*), B \rangle = \langle A^*A_1^*, B \rangle = \tau(A_1AB).$$

Both in the proof of the positivity and the module property (9.4) of E the tracial condition (iii) of τ was used. \square

A conditional expectation $E : \mathcal{B} \rightarrow \mathcal{A}$ is automatically **completely positive**. For $A_i \in \mathcal{A}$ and $B_i \in \mathcal{B}$ we have

$$\sum_{ij} A_i^* E(B_i^* B_j) A_j = E\left(\left(\sum_i B_i A_i\right)^* \left(\sum_j B_j A_j\right)\right) \geq 0 \quad (9.6)$$

due to the positivity and the module property of E .

The mapping \mathcal{E} in Example 3.6 is a conditional expectation from the $n \times n$ matrices to the algebra of diagonal $n \times n$ matrices.

Given a conditional expectation $E : B(\mathcal{H}) \rightarrow \mathcal{A}$ and a density matrix $\rho_0 \in \mathcal{A}$, the formula

$$\mathrm{Tr} \rho B = \mathrm{Tr} \rho_0 E(B) \quad (B \in B(\mathcal{H}))$$

determines a density ρ such that the reduced density is ρ_0 . The correspondence $\mathcal{E} : \rho_0 \mapsto \rho$ is a state transformation and called **state extension**. The state extension \mathcal{E} is the dual of the conditional expectation E . The converse is also true.

Theorem 9.1 *Let \mathcal{A} be a subalgebra of the matrix algebra \mathcal{B} . If $\mathcal{E} : \mathcal{A} \rightarrow \mathcal{B}$ is a positive trace-preserving mapping such that the reduced state of $\mathcal{E}(\rho_0)$ is ρ_0 for every density $\rho_0 \in \mathcal{A}$, then the dual of \mathcal{E} is a conditional expectation.*

Proof. The dual $E : \mathcal{B} \rightarrow \mathcal{A}$ is a positive unital mapping and $E(A) = A$ for every $A \in \mathcal{A}$. For a contraction B , $\|E(B)\|^2 = \|E(B)^* E(B)\| \leq \|E(B^* B)\| \leq \|E(I)\| = 1$. Therefore, we have $\|E\| = 1$.

Let P be a projection in \mathcal{A} and $B_1, B_2 \in \mathcal{B}$. We have

$$\begin{aligned} \|PB_1 + P^\perp B_2\|^2 &= \|(PB_1 + P^\perp B_2)^*(PB_1 + P^\perp B_2)\| \\ &= \|B_1^* P B_1 + B_2^* P^\perp B_2\| \\ &\leq \|B_1^* P B_1\| + \|B_2^* P^\perp B_2\| \\ &= \|PB_1\|^2 + \|P^\perp B_2\|^2. \end{aligned}$$

Using this, we estimate for an arbitrary $t \in \mathbb{R}$ as follows.

$$\begin{aligned} (t+1)^2 \|P^\perp E(PB)\|^2 &= \|P^\perp E(PB) + tP^\perp E(PB)\|^2 \\ &\leq \|PB + tP^\perp E(PB)\|^2 \\ &\leq \|PB\|^2 + t^2 \|P^\perp E(PB)\|^2 \end{aligned}$$

Since t can be arbitrary, $P^\perp E(PB) = 0$, that is, $PE(PB) = E(PB)$. We may write P^\perp in place of P :

$$(I - P)E((I - P)B) = E((I - P)B), \text{ equivalently, } PE(B) = PE(PB).$$

Therefore we conclude $PE(B) = E(PB)$. The linear span of projections is the full algebra \mathcal{A} and we have $AE(B) = E(AB)$ for every $A \in \mathcal{A}$. This completes the proof. \square

It is remarkable that in the proof of the previous theorem it was shown that if $E : \mathcal{B} \rightarrow \mathcal{A}$ is a positive mapping and $E(A) = A$ for every $A \in \mathcal{A}$, then E is a conditional expectation. This statement is called **Tomiyama theorem**. Actually, the positivity is equivalent to the condition $\|E\| = 1$ (when $E(I) = I$ is assumed). This has the consequence that in the definition of the conditional expectation, it is enough if (9.4) holds for $B = I$.

We say that the conditional expectation $E : B(\mathcal{H}) \rightarrow \mathcal{A}$ preserves the state ρ if

$$\operatorname{Tr} \rho B = \operatorname{Tr} \rho E(B) \quad \text{for every } B \in B(\mathcal{H}). \quad (9.7)$$

Takesaki's theorem tells about the existence of a conditional expectation.

Theorem 9.2 *Let $\mathcal{B} \simeq M_n(\mathbb{C})$ be a matrix algebra and \mathcal{A} be its subalgebra. Suppose that $\rho \in \mathcal{B}$ is an invertible density matrix. The following conditions are equivalent:*

(i) *A conditional expectation $E : \mathcal{B} \rightarrow \mathcal{A}$ preserving ρ exists.*

(ii) *For every $A \in \mathcal{A}$ and for the reduced density $\rho_0 \in \mathcal{A}$*

$$\rho^{1/2} A \rho^{-1/2} = \rho_0^{1/2} A \rho_0^{-1/2} \quad (9.8)$$

holds.

(iii) *For every $A \in \mathcal{A}$*

$$\rho^{1/2} A \rho^{-1/2} \in \mathcal{A} \quad (9.9)$$

holds.

Proof. Recall that the reduced density in \mathcal{A} is determined by the equation

$$\operatorname{Tr} \rho A = \operatorname{Tr} \rho_0 A \quad \text{for every } A \in \mathcal{A}.$$

Assume that $E : \mathcal{B} \rightarrow \mathcal{A}$ is a conditional expectation preserving ρ . We can consider \mathcal{B} as Hilbert spaces with the inner product

$$\langle B_1, B_2 \rangle = \operatorname{Tr} \rho B_1^* B_2.$$

Then the adjoint of the embedding $\mathcal{A} \rightarrow \mathcal{B}$ is the conditional expectation $E : \mathcal{B} \rightarrow \mathcal{A}$:

$$\langle A, B \rangle = \operatorname{Tr} \rho A^* B = \operatorname{Tr} E(A^* B) = \operatorname{Tr} \rho A^* E(B) = \langle A, E(B) \rangle$$

for $A \in \mathcal{A}$ and $B \in \mathcal{B}$.

Define a conjugate linear operator:

$$S : \mathcal{B} \rightarrow \mathcal{B}, \quad S(B) = B^* \quad (B \in \mathcal{B}).$$

We compute its adjoint S^* which is determined by the equations

$$\langle S(B_1), B_2 \rangle = \langle B_2, S^*(B_1) \rangle \quad (B_1, B_2 \in \mathcal{B}).$$

We show that $S^*(B_2) = \rho B_2^* \rho^{-1}$:

$$\langle S(B_1), B_2 \rangle = \text{Tr } \rho B_1 B_2 = \overline{\text{Tr } B_2^* B_1^* \rho} = \overline{\text{Tr } \rho B_1^* \rho B_2^* \rho^{-1}} = \overline{\langle B_1, \rho B_2^* \rho^{-1} \rangle}.$$

Due to the positivity of the conditional expectation, $ES = SE$. This implies that the positive operator $\Delta := S^*S$ leaves the subspace \mathcal{A} invariant. The action of Δ is

$$\Delta B = \rho B \rho^{-1} \quad (B \in \mathcal{B}).$$

For $A \in \mathcal{A}$, we have

$$\langle A, \Delta A \rangle = \text{Tr } \rho A A^* = \text{Tr } \rho_0 A A^* = \langle A, \Delta_0 A \rangle$$

if $\Delta_0 A = \rho_0 A \rho_0^{-1}$. Since the restriction of Δ to \mathcal{A} is Δ_0 , we have

$$\Delta_0^{1/2} A = \rho_0^{1/2} A \rho_0^{-1/2} = \Delta^{1/2} A = \rho^{1/2} A \rho^{-1/2} \quad (A \in \mathcal{A}).$$

This is exactly (9.8) and (i) \Rightarrow (ii) is proven.

The conditional expectation $F : \mathcal{B} \rightarrow \mathcal{A}$ preserving Tr exists, it is constructed in Example 9.5. The mapping

$$E_\rho(B) := \rho_0^{-1/2} F(\rho^{1/2} B \rho^{1/2}) \rho_0^{-1/2} \quad (9.10)$$

is completely positive and preserves the state ρ . Indeed,

$$\text{Tr } \rho E_\rho(B) = \text{Tr } \rho_0 E_\rho(B) = \text{Tr } F(\rho^{1/2} B \rho^{1/2}) = \text{Tr } (\rho^{1/2} B \rho^{1/2}) = \text{Tr } \rho B.$$

E_ρ is canonically determined and it is often called **generalized conditional expectation**.

We shall show that under Takesaki's condition (9.8), E_ρ is really a conditional expectation. Actually, we prove more. Assume that $A \in \mathcal{A}$ satisfies the condition $\rho^{1/2} A \rho^{-1/2} = \rho_0^{1/2} A \rho_0^{-1/2}$. Then

$$\begin{aligned} E_\rho(A) &= \rho_0^{-1/2} F(\rho^{1/2} A \rho^{-1/2}) \rho_0^{-1/2} = \rho_0^{-1/2} F(\rho_0^{1/2} A \rho_0^{-1/2}) \rho_0^{-1/2} \\ &= \rho_0^{-1/2} \rho_0^{1/2} A \rho_0^{-1/2} F(\rho) \rho_0^{-1/2} = A \rho_0^{-1/2} F(\rho) \rho_0^{-1/2} = A. \end{aligned}$$

It follows that $A \in \mathcal{A}$ is in the multiplicative domain of E_ρ . If this holds for every $A \in \mathcal{A}$, then E is really a conditional expectation. This completes the proof of (ii) \Rightarrow (i).

(ii) \Rightarrow (iii) is obvious and we show the converse. Endow \mathcal{B} with the Hilbert-Schmidt inner product. Condition (iii) tells us that the positive operator

$$\Delta B = \rho B \rho^{-1} \quad (B \in \mathcal{B})$$

leaves the subspace \mathcal{A} invariant. This is true also for Δ^{it} : For every $A \in \mathcal{A}$,

$$\rho^{it} A \rho^{-it} \in \mathcal{A}.$$

Now we apply Theorem 11.26 for the unitaries ρ^{it} . In the decomposition of \mathcal{A}

$$P_{(m,d)} \rho = \sum_{i=1}^{K(m,d)} \rho_i^L \otimes \rho_i^R \otimes E_{i,i}^{m,d},$$

since the permutation σ must be identity. We have

$$P_{(m,d)}\rho_0 = \sum_{i=1}^{K(m,d)} I_m \otimes \rho_i^R \otimes E_{i,i}^{m,d}.$$

Assume that $A \in \mathcal{A}$. Then $P_{(m,d)}A$ has the form

$$\sum_{i=1}^{K(m,d)} I_m \otimes A(m, d, i) \otimes E_{i,i}^{m,d}.$$

and

$$P_{(m,d)}\rho^{1/2}A\rho^{-1/2} = \sum_{i=1}^{K(m,d)} I_m \otimes (\rho_i^R)^{1/2}A(m, d, i)(\rho_i^R)^{-1/2} \otimes E_{i,i}^{m,d}$$

and this is the same as $P_{(m,d)}\rho_0^{1/2}A\rho_0^{-1/2}$. Hence (ii) is shown. \square

An important property of the relative entropy is related to conditional expectation.

Theorem 9.3 *Let $\mathcal{A} \subset B(\mathcal{H})$ be a subalgebra and $\rho, \omega \in B(\mathcal{H})$ be density matrices with reductions $\rho_0, \omega_0 \in \mathcal{A}$. Assume that there exists a conditional expectation $E : B(\mathcal{H}) \rightarrow \mathcal{A}$ onto \mathcal{A} which leaves ω invariant. Then*

$$S(\rho\|\omega) = S(\rho_0\|\omega_0) + S(\rho\|\rho \circ E). \quad (9.11)$$

Proof. Denoting the density of $\rho \circ E$ by ρ_1 , we need to show that

$$\text{Tr } \rho \log \rho - \text{Tr } \rho \log \omega = \text{Tr } \rho \log \rho_0 - \text{Tr } \rho \log \omega_0 + \text{Tr } \rho \log \rho - \text{Tr } \rho \log \rho_1.$$

The conditional expectation E leaves the states ω and $\rho \circ E$ invariant. According to (iii) in Theorem 9.8 this condition implies

$$\rho_1^{it}\omega^{-it} = \rho_0^{it}\omega_0^{-it} \quad (t \in \mathbb{R})$$

Differentiating this at $t = 0$, we have

$$\log \rho_1 - \log \omega = \log \rho_0 - \log \omega_0.$$

This relation gives the proof. \square

The state $\rho \circ E$ is nothing else than the extension of $\rho_0 \in \mathcal{A}$ with respect to the state ω . Therefore, the **conditional expectation property** (9.11) has the following heuristical interpretation. The relative entropy distance of ρ and ω comes from two sources. First, the distance of their restrictions to \mathcal{A} , second, the distance of ρ from the extension of ρ_0 (with respect to ω).

Note that the conditional expectation property is equivalently written as

$$S(\rho\|\omega) = S(\rho\|\rho \circ E) + S(\rho \circ E\|\omega). \quad (9.12)$$

Theorem 9.4 *Let \mathcal{B} be a matrix algebra and \mathcal{A} be its subalgebra. Suppose that $\omega, \rho \in \mathcal{B}$ are invertible density matrices. Then $E_\rho = E_\omega$ if and only if*

$$\rho_0^{-1/2} \rho^{1/2} = \omega_0^{-1/2} \omega^{1/2}$$

holds for the reduced densities $\omega_0, \rho_0 \in \mathcal{A}$.

Proof. Since

$$E_\rho(B) = F(\rho_0^{-1/2} \rho^{1/2} B \rho^{1/2} \rho_0^{-1/2}),$$

the condition obviously imply $E_\rho = E_\omega$. To see the converse, we refer to Theorem 9.7 and note that

$$\rho_0^{-1/2} \omega_0^{1/2} = \rho^{-1/2} \omega^{1/2}$$

is an equivalent form of the condition. □

A conditional expectation preserving a given state ρ does not always exist. This simple fact has a very far reaching consequences. Those concepts and arguments in classical probability which are based on conditioning have a very restricted chance to be extended to the quantum setting.

Example 9.6 Let ρ_{12} be a state of the composite system $\mathcal{H}_1 \otimes \mathcal{H}_2$. Then a conditional expectation $B(\mathcal{H}_1 \otimes \mathcal{H}_2) \rightarrow B(\mathcal{H}_1) \otimes \mathbb{C}I_2$ (preserving ρ_{12}) exists if and only if ρ_{12} is a product state.

When $\rho_{12} = \rho_1 \otimes \rho_2$, then

$$E(X \otimes Y) = X \operatorname{Tr} \rho_2 Y \quad (X \in B(\mathcal{H}_1), \quad Y \in B(\mathcal{H}_2)) \quad (9.13)$$

is a conditional expectation, it does not depend on ρ_1 .

From Takesaki's condition we obtain

$$\rho_1^{-1/2} \rho_{12}^{1/2} (X \otimes I) = (X \otimes I) \rho_1^{-1/2} \rho_{12}^{1/2}$$

for all $X \in B(\mathcal{H}_1)$. Therefore, $\rho_1^{-1/2} \rho_{12}^{1/2}$ is in the commutant and must have the form

$$\rho_1^{-1/2} \rho_{12}^{1/2} = I \otimes Y$$

for some $Y \in B(\mathcal{H}_2)$. This gives the factorization of ρ_{12} .

The example shows that the existence of a ρ_{12} -preserving conditional expectation is a very strong limitation for ρ_{12} . □

The previous example can be reformulated and it has an interesting interpretation, the **no cloning theorem**.

Example 9.7 Similarly to the previous example, consider $B(\mathcal{H}_1)$ as the subsystem of $B(\mathcal{H}_1 \otimes \mathcal{H}_2)$. The state transformation $\mathcal{E} : B(\mathcal{H}_1) \rightarrow B(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is a **state extension** if the reduced state of $\mathcal{E}(\rho)$ on the first subsystem is ρ itself, for every state ρ . Since the a state extension is the dual of a conditional expectation, the only possible **state extension** is

$$\mathcal{E}(\rho) = \rho \otimes \omega, \quad (9.14)$$

where ω is a fixed state of $B(\mathcal{H}_2)$.

Assume now that $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$. A state transformation $\mathcal{C} : B(\mathcal{H}) \rightarrow B(\mathcal{H} \otimes \mathcal{H})$ is called **cloning** if for any input state ρ both reduced density of the output $\mathcal{C}(\rho)$ is ρ itself. The transformation \mathcal{C} yields an output which is a pair of two subsystems, each of them in the state of the input. \mathcal{C} is interpreted as copying or cloning.

Since a cloning is a particular state extension, it must have the form (9.14) which contradicts to the definition of cloning (in the case when the dimension of \mathcal{H} is at least two). Therefore, cloning does not exist. One can arrive at the same conclusion under the weaker assumption that \mathcal{C} clones pure states only. \square

9.3 Commuting squares

Let \mathcal{A}_{123} be a matrix algebra with subalgebras $\mathcal{A}_{12}, \mathcal{A}_{23}, \mathcal{A}_2$ and assume that $\mathcal{A}_2 \subset \mathcal{A}_{12}, \mathcal{A}_{23}$. Then the diagram of embeddings $\mathcal{A}_2 \rightarrow \mathcal{A}_{12}, \mathcal{A}_2 \rightarrow \mathcal{A}_{23}, \mathcal{A}_{12} \rightarrow \mathcal{A}_{123}$ and $\mathcal{A}_{23} \rightarrow \mathcal{A}_{123}$ is obviously commutative, see Figure 9.1.

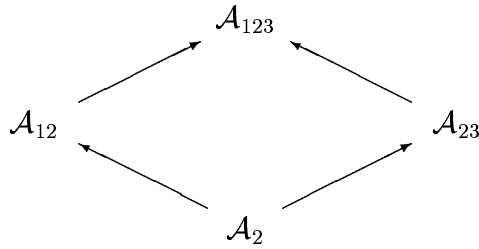


Figure 9.1: Subalgebras of \mathcal{A}_{123} : $\mathcal{A}_2 \subset \mathcal{A}_{12}, \mathcal{A}_{13}$.

Assume that a conditional expectation $E_{12}^{123} : \mathcal{A}_{123} \rightarrow \mathcal{A}_{12}$ exists. If the restriction of E_{12}^{123} to \mathcal{A}_{23} is a conditional expectation to \mathcal{A}_2 , then we call $(\mathcal{A}_{123}, \mathcal{A}_{12}, \mathcal{A}_{23}, \mathcal{A}_2, E_{12}^{123})$ a **commuting square**. The terminology comes from the commutativity of the diagram Figure 9.2 which consists of conditional expectations and embeddings.

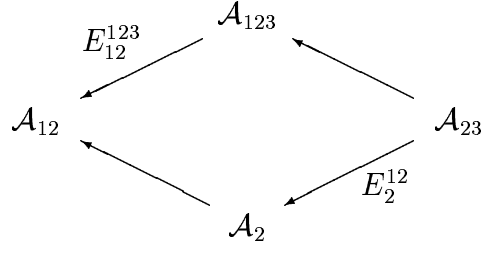
Theorem 9.5 *Let \mathcal{A}_{123} be a matrix algebra with subalgebras $\mathcal{A}_{12}, \mathcal{A}_{23}, \mathcal{A}_2$ and assume that $\mathcal{A}_2 \subset \mathcal{A}_{12}, \mathcal{A}_{23}$. Suppose that ω_{123} is a separating state on \mathcal{A}_{123} and the conditional expectations $E_{12}^{123} : \mathcal{A}_{123} \rightarrow \mathcal{A}_{12}$ and $E_2^{23} : \mathcal{A}_{23} \rightarrow \mathcal{A}_2$ preserving φ_{123} exist and $(\mathcal{A}_{123}, \mathcal{A}_{12}, \mathcal{A}_{23}, \mathcal{A}_2, E_{12}^{123})$ is a commuting square. Then the following conditions are equivalent:*

- (1) $E_{12}^{123}|_{\mathcal{A}_{23}} = E_2^{23}$. (2) $E_{23}^{123}|_{\mathcal{A}_{12}} = E_2^{12}$.
- (3) $E_{12}^{123} E_{23}^{123} = E_{23}^{123} E_{12}^{123}$ and $\mathcal{A}_{12} \cap \mathcal{A}_{23} = \mathcal{A}_2$.
- (4) $E_{12}^{123} E_{23}^{123} = E_2^{123}$. (5) $E_{23}^{123} E_{12}^{123} = E_2^{123}$.

The idea of the proof is to consider \mathcal{A}_{123} to be a Hilbert space endowed with the inner product

$$\langle A, B \rangle := \omega_{123}(A^* B) \quad (A, B \in \mathcal{A}_{123}).$$

Then E_{12}^{123} becomes an orthogonal projection onto the subspace \mathcal{A}_{12} of \mathcal{A}_{123} and all conditions can be reformulated in the Hilbert space. The details will be skipped.

Figure 9.2: Commuting square, $E_{12}^{123}|_{\mathcal{A}_{23}} = E_2^{12}$

Example 9.8 Let $\mathcal{A}_{123} := B(\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3)$ be the model of a quantum system consisting of three subsystems let \mathcal{A}_{12} , \mathcal{A}_{23} and \mathcal{A}_2 be the subsystem corresponding to the subscript, formally $\mathcal{A}_{12} := B(\mathcal{H}_1 \otimes \mathcal{H}_2) \otimes \mathbb{C}I$, $\mathcal{A}_{23} := \mathbb{C}I \otimes B(\mathcal{H}_1 \otimes \mathcal{H}_2)$ and $\mathcal{A}_2 := \mathbb{C}I \otimes B(\mathcal{H}_2) \otimes \mathbb{C}I$. The conditional expectations preserving the tracial state τ (or Tr) from \mathcal{A}_{123} onto any subalgebra exist and unique. For example, the conditional expectation $E_{12}^{123} : \mathcal{A}_{123} \rightarrow \mathcal{A}_{12}$ is of the form

$$A \otimes B \otimes C \mapsto \tau(C)A \otimes B \otimes I.$$

Up to a scalar, this is the partial trace over \mathcal{H}_3 . Its restriction to \mathcal{A}_{23} is

$$I \otimes B \otimes C \mapsto \tau(C)I \otimes B \otimes I$$

which is really the conditional expectation of \mathcal{A}_{23} to \mathcal{A}_2 . □

Example 9.9 Assume that the algebra \mathcal{A} is generated by the elements $\{a_i : 1 \leq i \leq n\}$ which satisfy the **canonical anticommutation relations**

$$\begin{aligned} a_i a_j + a_j a_i &= 0 \\ a_i a_j^* + a_j^* a_i &= \delta_{i,j} I \end{aligned}$$

for $1 \leq i, j \leq n$. It is known that \mathcal{A} is isomorphic to a matrix algebra $M_{2^n}(\mathbb{C}) \simeq M_2(\mathbb{C})^{\perp} \otimes \cdots \otimes M_2(\mathbb{C})^{\perp}$, the isomorphism is called **Jordan-Wigner transformation**. The relations

$$\begin{aligned} e_{11}^{(i)} &:= a_i a_i^*, & e_{12}^{(i)} &:= V_{i-1} a_i, \\ e_{21}^{(i)} &:= V_{i-1} a_i^*, & e_{22}^{(i)} &:= a_i^* a_i, \end{aligned}$$

$$V_i := \prod_{j=1}^i (I - 2a_j^* a_j)$$

determine a family of mutually commuting 2×2 matrix units for $1 \leq i \leq n$. Since

$$a_i = \prod_{j=1}^{i-1} (e_{11}^{(j)} - e_{22}^{(j)}) e_{12}^{(i)},$$

the above matrix units generate \mathcal{A} and give an isomorphism between \mathcal{A} and $M_2(\mathbb{C}) \otimes \cdots \otimes M_2(\mathbb{C})$:

$$e_{i_1 j_1}^{(1)} e_{i_2 j_2}^{(2)} \cdots e_{i_n j_n}^{(n)} \longleftrightarrow E_{i_1 j_1} \otimes E_{i_2 j_2} \otimes \cdots \otimes E_{i_n j_n}. \quad (9.15)$$

(Here E_{ij} stand for the standard matrix units in $M_2(\mathbb{C})$.) It follows from this isomorphism that \mathcal{A} has a unique tracial state.

Let $I_{12}, I_{23} \subset \{1, 2, \dots, n\}$ and let $I_2 = I_{12} \cap I_{23}$. Moreover, let $\mathcal{A}_{123} := \mathcal{A}$ and $\mathcal{A}_{12}, \mathcal{A}_{23}, \mathcal{A}_2$ be generated by the elements a_i , where i belong to the set I_{12}, I_{23}, I_2 , respectively. It is known that

$$\mathcal{A}_{12} \cap \mathcal{A}_{23} = \mathcal{A}_2.$$

The trace preserving conditional expectation $E_{12}^{123} : \mathcal{A}_{123} \rightarrow \mathcal{A}_{12}$ exists and we have a commuting square $(\mathcal{A}_{123}, \mathcal{A}_{12}, \mathcal{A}_{23}, \mathcal{A}_2, E_{12}^{123})$. \square

9.4 Superadditivity

The content of the next theorem is the superadditivity of the relative entropy for a commuting square.

Theorem 9.6 *Assume that $(\mathcal{A}_{123}, \mathcal{A}_{12}, \mathcal{A}_{23}, \mathcal{A}_2, E_{12}^{123})$ is a commuting square, ω_{123} is a separating state on \mathcal{A}_{123} and E_{12}^{123} leaves this state invariant. Let ρ_{123} be an arbitrary state on \mathcal{A}_{123} and we denote by $\omega_{12}, \omega_{23}, \omega_2, \rho_{12}, \rho_{23}, \rho_2$ the restrictions of these states. Then*

$$S(\rho_{123} \|\omega_{123}) + S(\rho_2 \|\omega_2) \geq S(\rho_{12} \|\omega_{12}) + S(\rho_{23} \|\omega_{23}). \quad (9.16)$$

Proof. The conditional expectation property of the relative entropy tells us that

$$S(\rho_{123} \|\omega) = S(\rho_{12} \|\omega_{12}) + S(\rho_{123} \|\rho_{123} \circ E_{12}^{123}), \quad (9.17)$$

$$S(\rho_2 \|\omega_2) + S(\rho_{23} \|\rho_{23} \circ E_2^{23}) = S(\rho_{23}, \omega_{23}). \quad (9.18)$$

The monotonicity and the commuting square property give

$$S(\rho_{123} \|\rho_{123} \circ E_{12}^{123}) \geq S(\rho_{23} \|\rho_{123} \circ E_{12}^{123} | \mathcal{A}_{23}) = S(\rho_{23} \|\rho_2 \circ E_2^{23}). \quad (9.19)$$

By adding the equations (9.17), (9.18) and the inequality (9.19), we conclude the statement of the theorem.

It is worthwhile to note that the necessary and sufficient condition for the equality is

$$S(\rho_{123} \|\rho_{123} \circ E_{12}^{123}) = S(\rho_{23} \|\rho_{123} \circ E_{12}^{123} | \mathcal{A}_{23}). \quad (9.20)$$

This fact will be used later. (This condition means that the subalgebra \mathcal{A}_{23} is sufficient for the states ρ_{123} and $\rho_{123} \circ E_{12}^{123}$.) \square

The theorem implies the strong subadditivity of the von Neumann entropy for the tensor product structure as it is in (5.12), but a similar strong subadditivity holds for the CAR algebra as well.

9.5 Sufficiency

In order to motivate the concept of sufficiency, we first turn to the setting of classical statistics. Suppose we observe an N -dimensional random vector $X = (x_1, x_2, \dots, x_N)$,

characterized by the density function $f(x|\theta)$, where θ is a p -dimensional vector of parameters and p is usually much smaller than N . Assume that the densities $f(x|\theta)$ are known and the parameter θ completely determines the distribution of X . Therefore, θ is to be estimated. The N -dimensional observation X carries information about the p -dimensional parameter vector θ . One may ask the following question: Can we compress x into a low-dimensional statistic without any loss of information? Does there exist some function $t = Tx$, where the dimension of t is less than N , such that t carries all the useful information about θ ? If so, for the purpose of studying θ , we could discard the measurements x and retain only the low-dimensional statistic t . In this case, we call t a **sufficient statistic**. The following example is standard and simple.

Suppose a binary information source emits a sequence of -1 's and $+1$'s, we have the independent variables X_1, X_2, \dots, X_N such that $\text{Prob}(X_i = 1) = \theta$. The quantum mechanical example we may have in our mind is the measurement of a Pauli observable σ_1 on N identical copies of a qubit. The empirical mean

$$T(x_1, x_2, \dots, x_N) = \frac{1}{N} \sum_{i=1}^N x_i$$

can be used to estimate the parameter θ and it is a sufficient statistic. (Knowledge of the empirical mean is equivalent to the knowledge of the relative frequencies of ± 1 .)

Let \mathcal{B} be a matrix algebra. Assume that a family $\mathcal{S} := \{\rho_\theta : \theta \in \Theta\}$ of density matrices are given. $(\mathcal{M}, \mathcal{S})$ is called **statistical experiment**. The subalgebra $\mathcal{A} \subset \mathcal{B}$ is **sufficient** for $(\mathcal{B}, \mathcal{S})$ if there exists a coarse-graining $\alpha : \mathcal{B} \rightarrow \mathcal{A}$ such that

$$\text{Tr } \rho_\theta B = \text{Tr } \rho_\theta \alpha(B) \quad (\theta \in \Theta, \quad B \in \mathcal{B}). \quad (9.21)$$

If we denote by $\rho_{\theta,0}$ the reduced density of ρ_θ in \mathcal{A} , then we can formulate sufficiency in a slightly different way. $\mathcal{A} \subset \mathcal{B}$ is **sufficient** for $(\mathcal{B}, \mathcal{S})$ if there exists a completely positive trace-preserving mapping $\mathcal{E} : \mathcal{A} \rightarrow \mathcal{B}$ such that

$$\mathcal{E}(\rho_{\theta,0}) = \rho_\theta \quad (\theta \in \Theta). \quad (9.22)$$

Indeed, \mathcal{E} satisfies (9.22) if and only if its dual α satisfies (9.21).

Before stating the main theorem characterizing sufficient subalgebras, recall the concept of the **Connes' cocycle**. If ρ and ω are density matrices, then

$$[D\rho, D\omega]_t = \rho^{it} \omega^{-it}$$

is a one-parameter family of contractions in \mathcal{B} and it is called the Connes' cocycle of ρ and ω . When both densities are invertible, the Connes' cocycle consists of unitaries. The Connes' cocycle is the quantum analogue of the Radon-Nikodym derivative of measures.

The next result is called **sufficiency theorem**.

Theorem 9.7 *Let $\mathcal{A} \subset \mathcal{B}$ be matrix algebras and let $(\mathcal{B}, \{\rho_\theta : \theta \in \Theta\})$ be a statistical experiment. Assume that there are densities $\rho_n \in \mathcal{S} := \{\rho_\theta : \theta \in \Theta\}$ such that*

$$\omega := \sum_{n=1}^{\infty} \lambda_n \rho_n$$

is an invertible density for some constants $\lambda_n > 0$. Then the following conditions are equivalent.

- (i) \mathcal{A} is sufficient for $(\mathcal{B}, \mathcal{S})$.
- (ii) $S_\alpha(\rho_\theta||\omega) = S_\alpha(\rho_{\theta,0}||\omega_0)$ for all θ and for some $0 < |\alpha| < 1$.
- (iii) $[D\rho_\theta, D\omega]_t = [D\rho_{\theta,0}, D\omega_0]_t$ for every real t and for every θ .
- (v) The generalized conditional expectation $E_\omega : \mathcal{B} \rightarrow \mathcal{A}$ leaves all the states ρ_θ invariant.

Lemma 9.2 Let ρ_0 and ω_0 be the reduced densities of $\rho, \omega \in \mathcal{B}$ in \mathcal{A} . Assume that ω is invertible. Then $S_\alpha(\rho||\omega) = S_\alpha(\rho_0||\omega_0)$ implies

$$\rho_0^{it}\omega_0^{-it} = \rho^{it}\omega^{-it}$$

for every real number $t \in \mathbb{R}$.

Proof. The relative α -entropies can be expressed by the **relative modular operators** Δ acting on the Hilbert space \mathcal{B} and Δ_0 acting on \mathcal{A} :

$$\Delta A = \rho A \omega^{-1} \quad (A \in \mathcal{A}) \quad \text{and} \quad \Delta_0 B = \rho_0 B \omega_0^{-1} \quad (B \in \mathcal{B}).$$

The α -entropies are

$$S_\alpha(\rho||\omega) = \frac{1}{\alpha(1-\alpha)}(1 - \text{Tr} \omega^\alpha \rho^{1-\alpha}) \quad (9.23)$$

but for the sake of simplicity we can neglect the constants:

$$S_\alpha^0(\rho||\omega) = \text{Tr} \omega^\alpha \rho^{1-\alpha} = \langle \omega^{1/2}, \Delta^\beta \omega^{1/2} \rangle, \quad (9.24)$$

where $\beta = 1 - \alpha$. Assume that $0 < \alpha < 1$. (The case of negative α is treated similarly.) From the integral representation of Δ^β , we have

$$S_\alpha^0(\rho||\omega) = \frac{\sin \pi \beta}{\pi} \int_0^\infty \langle \omega^{1/2}, t^{\beta-1} \Delta(t + \Delta)^{-1} \omega^{1/2} \rangle dt$$

and

$$S_\alpha^0(\rho_0||\omega_0) = \frac{\sin \pi \beta}{\pi} \int_0^\infty \langle \omega_0^{1/2}, t^{\beta-1} \Delta_0(t + \Delta_0)^{-1} \omega_0^{1/2} \rangle dt.$$

Due to the monotonicity of quasi-entropies, there is an inequality between the two integrands. Therefore, the equality of the entropies is equivalent to the condition

$$\langle \omega^{1/2}, \Delta(t + \Delta)^{-1} \omega^{1/2} \rangle = \langle \omega_0^{1/2}, \Delta_0(t + \Delta_0)^{-1} \omega_0^{1/2} \rangle \quad (t > 0),$$

or

$$\langle \omega^{1/2}, (t + \Delta)^{-1} \omega^{1/2} \rangle = \langle \omega_0^{1/2}, (t + \Delta_0)^{-1} \omega_0^{1/2} \rangle \quad (t > 0). \quad (9.25)$$

These are equalities for numbers, we want to obtain equalities for operators.

The operator $V : \mathcal{A} \rightarrow \mathcal{B}$ defined as

$$V A \omega_0^{1/2} = A \omega^{1/2} \quad (A \in \mathcal{A})$$

is an isometry and

$$V^* \Delta V \leq \Delta_0.$$

The function $f_t(y) = (y + t)^{-1} - t^{-1}$ is operator monotone decreasing, so

$$(\Delta_0 + t)^{-1} - t^{-1}I \leq (V^*\Delta V + t)^{-1} - t^{-1}I.$$

Moreover f_t is operator convex and $f_t(0) = 0$, therefore

$$(V^*\Delta V + t)^{-1} - t^{-1}I \leq V^*(\Delta + t)^{-1}V - t^{-1}V^*V.$$

The two equations together give

$$K := (\Delta_0 + t)^{-1} - t^{-1}I \leq V^*(\Delta + t)^{-1}V - t^{-1}V^*V =: L$$

Recall that equation (9.25) is

$$\langle \omega_0^{1/2}, K\omega_0^{1/2} \rangle = \langle \omega_0^{1/2}, L\omega_0^{1/2} \rangle.$$

This implies that $\|(L - K)^{1/2}\omega_0^{1/2}\|^2 = 0$ and $K\omega_0^{1/2} = L\omega_0^{1/2}$, or

$$V^*(\Delta + t)^{-1}\omega^{1/2} = (\Delta_0 + t)^{-1}\omega_0^{1/2} \quad (9.26)$$

for all $t > 0$. Differentiating by t we have

$$V^*(\Delta + t)^{-2}\omega^{1/2} = (\Delta_0 + t)^{-2}\omega_0^{1/2}$$

and we infer

$$\begin{aligned} \|V^*(\Delta + t)^{-1}\omega^{1/2}\|^2 &= \langle (\Delta_0 + t)^{-2}\omega_0^{1/2}, \omega_0^{1/2} \rangle \\ &= \langle V^*(\Delta + t)^{-2}\omega^{1/2}, \omega_0^{1/2} \rangle \\ &= \|(\Delta + t)^{-1}\omega^{1/2}\|^2 \end{aligned}$$

When $\|V^*\xi\| = \|\xi\|$ holds for a contraction V , it follows that $VV^*\xi = \xi$. In the light of this remark we arrive at the condition

$$VV^*(\Delta + t)^{-1}\omega^{1/2} = (\Delta + t)^{-1}\omega^{1/2}$$

and from (9.26)

$$\begin{aligned} V(\Delta_0 + t)^{-1}\omega_0^{1/2} &= VV^*(\Delta + t)^{-1}\omega^{1/2} \\ &= (\Delta + t)^{-1}\omega^{1/2} \end{aligned}$$

The family of functions $g_t(x) = (t + x)^{-1}$ is very large and the Stone-Weierstrass approximation yields

$$Vf(\Delta_0)\omega_0^{1/2} = f(\Delta)\omega^{1/2} \quad (9.27)$$

for any continuous function f . In particular for $f(x) = x^{it}$ we have

$$\rho_0^{it}\omega_0^{-it}\omega^{1/2} = \rho^{it}\omega^{-it+1/2}. \quad (9.28)$$

This condition follows from the equality of α -entropies. \square

Now we can prove the theorem. (i) implies (ii) due to the monotonicity of the relative α -entropies. The key point is (ii) \Rightarrow (iii). This follows from Lemma 9.2.

(iii) \Rightarrow (iv) is rather straightforward:

$$\begin{aligned} \text{Tr } \rho_{\theta,0} E_{\omega}(B) &= \text{Tr } \rho_{\theta,0} \omega_0^{-1/2} F(\omega^{1/2} B \omega^{1/2}) \omega_0^{-1/2} = \text{Tr } \rho_{\theta,0} \omega_0^{-1/2} \omega^{1/2} B \omega^{1/2} \omega_0^{-1/2} \\ &= \text{Tr } \rho_{\theta} \omega^{-1/2} \omega^{1/2} B \omega^{1/2} \omega^{-1/2} = \text{Tr } \rho_{\theta} B, \end{aligned}$$

where $\rho_{\theta,0} \omega_0^{-1/2} = \rho_{\theta} \omega^{-1/2}$ was used.

Finally, (iv) \Rightarrow (i) due to the definition of sufficiency. \square

The density ω appearing in the theorem is said to be **dominating** for the statistical experiment \mathcal{S} . Given a dominated statistical experiment \mathcal{S} , the subalgebra generated by the operators

$$\{\rho_{\theta}^{it} \omega^{-it} : t \in \mathbb{R}\}$$

is the smallest sufficient subalgebra. If there are $\theta_1, \theta_2 \in \Theta$ such that ρ_{θ_1} and ρ_{θ_2} do not commute, then there exists no sufficient commutative subalgebra for $\{\rho_{\theta} : \theta \in \Theta\}$.

The formulation of the sufficiency theorem was made a bit complicated, since the formulation is true for an arbitrary von Neumann algebra and for a family of normal (and not necessarily faithful) states. For two invertible states in finite dimension the relative entropy is always finite and we can have a simpler formulation of the **sufficiency theorem**.

Theorem 9.8 *Let $\omega, \rho \in \mathcal{B} = B(\mathcal{H})$ be invertible density matrices on a finite dimensional Hilbert space \mathcal{H} and let $\mathcal{A} \subset \mathcal{B}$ be a subalgebra. Denote by ρ_0 and ω_0 the reduced densities in \mathcal{A} . Then the following conditions are equivalent.*

- (i) \mathcal{A} is sufficient for $\{\rho, \omega\}$.
- (ii) $S_{\alpha}(\rho||\omega) = S_{\alpha}(\rho|\mathcal{A}||\omega|\mathcal{A})$ for some α such that $|\alpha| < 1$.
- (iii) $\rho^{it} \omega^{-it} = \rho_0^{it} \omega_0^{-it}$ for every real t .
- (iv) $\rho_0^{-1/2} \rho^{1/2} = \omega_0^{-1/2} \omega^{1/2}$.
- (v) The generalized conditional expectations $E_{\omega} : \mathcal{B} \rightarrow \mathcal{A}$ and $E_{\rho} : \mathcal{B} \rightarrow \mathcal{A}$ coincide.
- (vi) $\rho^{it} \omega^{-it} \in \mathcal{A}$ for all real t .

Proof. The equivalence (i) \iff (ii) \iff (iii) was proven in the previous theorem. (iii) \Rightarrow (iv) is obvious, (iv) \iff (v) is Theorem 9.4. (v) \Rightarrow (i) and (iii) \Rightarrow (vi) are trivial.

The real problem is to prove (vi) \Rightarrow (i). Let \mathcal{A}_0 be the algebra generated by $\{\rho^{it} \omega^{-it} : t \in \mathbb{R}\}$. Of course, $\mathcal{A}_0 \subset \mathcal{A}$ and we can write elements of \mathcal{A}_0 in the form

$$\bigoplus_{k=1}^K I_k^L \otimes A_k^R,$$

see Section 11.8. We have

$$\rho^{it} \mathcal{A}_0 \rho^{-it} \subset \mathcal{A}_0$$

for every $t \in \mathbb{R}$ and Theorem 11.27 tells us that

$$\rho = \bigoplus_{k=1}^K A_k^L \otimes A_k^R \quad \text{and} \quad \omega = \bigoplus_{k=1}^K B_k^L \otimes B_k^R.$$

Since $\rho\omega^{-1} \in \mathcal{A}_0$, $A_k^L(B_k^L)^{-1}$ is a constant multiple of the identity, we may assume that $A_k^L = B_k^L$. The reduced densities are

$$\rho_1 = \bigoplus_{k=1}^K (\text{Tr } A_k^L) I_k^L \otimes A_k^R \quad \text{and} \quad \omega_1 = \bigoplus_{k=1}^K (\text{Tr } A_k^L) I_k^L \otimes B_k^R.$$

We can conclude that

$$\rho = \rho_1 D \quad \text{and} \quad \omega = \omega_1 D,$$

where $D \in \mathcal{A}'_1$. This relation implies that $\rho^{it}\omega^{-it} = \rho_1^{it}\omega_1^{-it}$ for every $t \in \mathbb{R}$. So the subalgebra \mathcal{A}_0 is sufficient, and the larger subalgebra \mathcal{A} must be sufficient as well. \square

9.6 Markov states

Let ρ_{ABC} be a density matrix acting on the finite dimensional tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. The reduced density matrices will be denoted by ρ_{AB} , ρ_{BC} and ρ_B .

Theorem 9.9 *The following conditions are equivalent.*

(i) *The equality*

$$S(\rho_{ABC}) + S(\rho_B) = S(\rho_{AB}) + S(\rho_{BC})$$

holds in the strong subadditivity of the von Neumann entropy.

(ii) *If τ_A denotes the tracial state of $B(\mathcal{H}_A)$, then*

$$S(\rho_{ABC} \| \tau_A \otimes \rho_{BC}) = S(\rho_{AB} \| \tau_A \otimes \rho_B).$$

(iii) *There exist a subalgebra \mathcal{A} and a conditional expectation from $B(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ onto \mathcal{A} such that*

$$B(\mathcal{H}_A) \otimes \mathbb{C}I_B \otimes \mathbb{C}I_C \subset \mathcal{A} \subset B(\mathcal{H}_A) \otimes B(\mathcal{H}_B) \otimes \mathbb{C}I_C$$

and E leaves the state ρ_{ABC} invariant.

(iv) *The generalized conditional expectation*

$$E_\rho : B(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C) \rightarrow B(\mathcal{H}_A) \otimes B(\mathcal{H}_B) \otimes \mathbb{C}I_C$$

(with respect to ρ_{ABC}) leaves the operators in $B(\mathcal{H}_A) \otimes \mathbb{C}I_B \otimes \mathbb{C}I_C$ fixed.

(v) *There is a state transformation*

$$\mathcal{E} : B(\mathcal{H}_B) \rightarrow B(\mathcal{H}_B \otimes \mathcal{H}_C)$$

such that $(\text{id}_A \otimes \mathcal{E})(\rho_{AB}) = \rho_{ABC}$.

Proof. The equivalence (i) \iff (ii) is clear from the proof of the strong subadditivity in Section 5.4.

(iii) \implies (ii): We want to show that E leaves also the state $\tau_A \otimes \rho_{BC}$ invariant. First we establish that $E(I_A \otimes X_B \otimes X_C)$ commutes with every $X_A \otimes I_B \otimes I_C$, therefore $E(I_A \otimes X_B \otimes X_C) \in \mathbb{C}I_A \otimes B(\mathcal{H}_B) \otimes \mathbb{C}I_C$. We have

$$\begin{aligned} (\tau_A \otimes \rho_{BC})E(X_A \otimes X_B \otimes X_C) &= (\tau_A \otimes \rho_{BC})(X_A \otimes I_B \otimes I_C)E(I_A \otimes X_B \otimes X_C) \\ &= \tau_A(X_A) \text{Tr} \rho_{BC} E(I_A \otimes X_B \otimes X_C) \\ &= (\tau_A \otimes \rho_{BC})(X_A \otimes X_B \otimes X_C). \end{aligned}$$

It follows that

$$S(\rho_{ABC} \parallel \tau_A \otimes \rho_{BC}) = S(\rho_{ABC} \mid \mathcal{A} \parallel \tau_A \otimes \rho_{BC} \mid \mathcal{A}).$$

Since $S(\rho_{AB} \parallel \tau_A \otimes \rho_B)$ is between the left-hand-side and the right-hand-side, it must have the same value.

(iv) \implies (iii): E_ρ preserves ρ_{ABC} and so do its powers. Due to the Kovács-Szűcs theorem,

$$\frac{1}{n}(\text{id} + E_\rho + \dots + E_\rho^{n-1}) \rightarrow E$$

where E is a conditional expectation onto the fixed point algebra.

(ii) \implies (iv) and (v): Condition (v) in Theorem 9.8 tells us that the generalized conditional expectation E_ρ has the form

$$E_\rho(X) = (\tau_A \otimes \rho_B)^{-1/2} F((\tau_A \otimes \rho_{BC})^{1/2} X (\tau_A \otimes \rho_{BC})^{1/2}) (\tau_A \otimes \rho_B)^{-1/2}$$

where F is the conditional expectation preserving the normalized trace. Let \mathcal{A} be the fixed point algebra of E_ρ . One can compute that $E_\rho(X_A \otimes I_B \otimes I_C) = X_A \otimes I_B \otimes I_C$, therefore

$$B(\mathcal{H}_A) \otimes \mathbb{C}I_B \otimes \mathbb{C}I_C \subset \mathcal{A} \subset B(\mathcal{H}_A) \otimes B(\mathcal{H}_B) \otimes \mathbb{C}I_C. \quad (9.29)$$

and (iv) is obtained. Relation (9.29) gives that \mathcal{A} must be of the form $B(\mathcal{H}_A) \otimes \mathcal{A}_B \otimes \mathbb{C}I_C$ with a subalgebra \mathcal{A}_B of $B(\mathcal{H}_B)$. It follows that the dual of E_ρ has the form $\text{id}_A \otimes \mathcal{E}$ and we arrived at (v).

(v) implies that $\mathcal{E}(\tau_A \otimes \rho_B) = \rho_{ABC}$ and (ii) must hold. \square

If a density matrix ρ_{ABC} acting on tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ satisfies the conditions of the previous theorem, then ρ_{ABC} will be called a **Markov state**.

If the notation

$$I(A : C \mid B)_\rho := S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_B) - S(\rho_{ABC})$$

is used, then the Markovianity of ρ is the condition $I(A : C \mid B)_\rho = 0$.

Example 9.10 Assume that $\mathcal{H}_B = \mathcal{H}_L \otimes \mathcal{H}_R$. Consider densities $\rho_{AL} \in B(\mathcal{H}_A \otimes \mathcal{H}_L)$ and $\rho_{RC} \in B(\mathcal{H}_R \otimes \mathcal{H}_C)$ and let $\rho_{ABC} = \rho_{AL} \otimes \rho_{RC}$. Then ρ_{ABC} is a Markov state, and we shall call it a product type. It is easy to check the strong additivity of the von Neumann entropy:

$$S(\rho_{ABC}) = S(\rho_{AL}) + S(\rho_{RC}), \quad S(\rho_B) = S(\rho_L) + S(\rho_R),$$

$$S(\rho_{AB}) = S(\rho_{AL}) + S(\rho_R), \quad S(\rho_{BC}) = S(\rho_L) + S(\rho_{RC}).$$

Let P and Q be orthogonal projections. Assume that ρ_{ABC} and ω_{ABC} are Markov states with support in P and Q , respectively. Computation of the von Neumann entropies yields that any convex combination $\lambda\rho_{ABC} + (1 - \lambda)\omega_{ABC}$ is a Markov state as well.

It is the content of the next theorem that every Markov state is the convex combination of orthogonal product type states. \square

Theorem 9.10 *Assume that ρ_{ABC} is a Markov state on the finite dimensional tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. Then \mathcal{H}_B has an orthogonal decomposition*

$$\mathcal{H}_B = \bigoplus_k \mathcal{H}_k^L \otimes \mathcal{H}_k^R$$

and for every k there are density matrices $\rho_{AL}^k \in B(\mathcal{H}_A \otimes \mathcal{H}_k^L)$ and $\rho_{RC}^k \in B(\mathcal{H}_k^R \otimes \mathcal{H}_C)$ such that ρ_{ABC} is a convex combination

$$\rho_{ABC} = \sum_k p_k \rho_{AL}^k \otimes \rho_{RC}^k.$$

Proof. Let $\omega := \tau_A \otimes \rho_{BC}$ and $\rho := \rho_{ABC}$. We know from Theorem 9.9 that Markovianity implies that the generalized conditional expectation

$$E_\omega : B(\mathcal{H}_A) \otimes B(\mathcal{H}_B) \otimes B(\mathcal{H}_C) \rightarrow B(\mathcal{H}_A) \otimes B(\mathcal{H}_B) \otimes \mathbb{C}I_C$$

with respect to the state ω is the same as E_ρ and the fixed point algebra \mathcal{A} has the property

$$B(\mathcal{H}_A) \otimes \mathbb{C}I_B \otimes \mathbb{C}I_C \subset \mathcal{A} \subset B(\mathcal{H}_A) \otimes B(\mathcal{H}_B) \otimes \mathbb{C}I_C.$$

It follows \mathcal{A} must be of the form $B(\mathcal{H}_A) \otimes \mathcal{A}_B \otimes \mathbb{C}I_C$ with a subalgebra \mathcal{A}_B of $B(\mathcal{H}_B)$. Elements of \mathcal{A}_B have the form

$$\bigoplus_{k=1}^K A_k^L \otimes I_k^R,$$

where $A_k^L \in B(\mathcal{H}_k^L)$, I_k^R is the identity on \mathcal{H}_k^R and

$$\mathcal{H}_B = \bigoplus_{k=1}^K \mathcal{H}_k^L \otimes \mathcal{H}_k^R.$$

In this way \mathcal{A} is isomorphic to

$$\bigoplus_{k=1}^K B(\mathcal{H}_A \otimes \mathcal{H}_k^L) \otimes \mathbb{C}I_k^R \otimes \mathbb{C}I_C.$$

Since $\rho^{it} \mathcal{A} \rho^{-it} \subset \mathcal{A}$ holds, Theorem 11.27 can be applied and gives the stated decomposition. \square

9.7 Notes

The conditional expectation in the matrix algebra (or von Neumann algebra) setting was introduced by Umegaki in [110] and Example 9.5 is due to him. Takesaki's theorem is from 1972. In the infinite dimensional von Neumann algebra case, the modular operators are unbounded and condition (9.9) is replaced by the equation

$$\Delta_{\mathcal{A}}^{it} A \Delta_{\mathcal{A}}^{-it} = \Delta_{\mathcal{B}}^{it} A \Delta_{\mathcal{B}}^{-it}$$

for real t 's. (Note that $\Delta_{\mathcal{A}}^{it}$ and $\Delta_{\mathcal{B}}^{it}$ are unitaries and the operators $\Delta_{\mathcal{A}}^z$ and $\Delta_{\mathcal{B}}^z$ are not everywhere defined for a complex $z \in \mathbb{C}$.) The generalized conditional expectation was introduced by Accardi and Cecchini, [86] is a suggested survey.

Example 9.4 is a particular case of the **Kovács-Szűcs theorem** which concerns a semigroup of coarse-grainings in von Neumann algebras. That was the first ergodic theorem in the von Neumann algebra setting.

In general von Neumann algebras, the **Connes' cocycle** is defined in terms of relative modular operators:

$$[D\psi, D\omega]_t = \Delta(\psi/\omega)^{it} \Delta(\omega/\omega)^{-it}$$

is a one-parameter family of contractions in the von Neumann algebra, although the relative modular operators $\Delta(\psi/\omega)$ are typically unbounded.

The **conditional expectation property** together with invariance, direct sum property, nilpotence and measurability can be used to axiomatize the relative entropy, see Chapter 2 of [79]. The conditional expectation property in full generality is Theorem 5.15 in the same monograph.

Sufficiency in the quantum setting was initiated by Petz in [85] for subalgebras and in [87] for coarse-grainings. Recent paper on the subject is [59] and [60] is survey with examples.

The concept of Markov state goes back to Accardi and Frigerio [1]. The original definition was formulated in terms of generalized or quasi-conditional expectation, see also Lemma 11.3 in [79].

Theorem 9.10 is due to Hayden et al [45], the presented proof follows [72]. The theorem holds in infinite dimensional Hilbert space if all the von Neumann entropies $S(\rho_{AB}), S(\rho_B), S(\rho_{BC})$ are finite, see [59].

9.8 Exercises

1. Let $\mathcal{A} \subset \mathcal{B}$ be matrix algebras, $\rho \in \mathcal{B}$ be a density matrix with reduced density $\rho_0 \in \mathcal{A}$. Endow \mathcal{A} with the inner product $\langle A_1, A_2 \rangle := \text{Tr } A_1^* \rho_0^{1/2} A_2 \rho_0^{1/2}$ and similarly let $\langle B_1, B_2 \rangle := \text{Tr } B_1^* \rho^{1/2} B_2 \rho^{1/2}$ be an inner product on \mathcal{B} . Show that the adjoint of the embedding $\mathcal{A} \rightarrow \mathcal{B}$ is the **generalized conditional expectation** E_ρ defined in (9.10).
2. Let \mathcal{B} be a matrix algebra and $\rho \in \mathcal{B}$ be an invertible density matrix. Assume that $\alpha : \mathcal{B} \rightarrow \mathcal{B}$ is a coarse-graining leaving the state ρ invariant and let \mathcal{A} be the set of fixed points of α . Show that for every $B \in \mathcal{B}$ the set

$$\overline{\text{conv}}\{\alpha^n(B) : n \in \mathbb{Z}^+\} \cap \mathcal{A}$$

is a singleton.

Example 9.11 Assume that ω_{ABC} is an invertible density matrix on the tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. Can we modify it to get a Markov state? The question is the quantum analogue of Example 5.1. Set

$$\Delta(\omega_{ABC}) := \inf\{S(\omega_{ABC} \parallel \rho_{ABC}) : \rho_{ABC} \text{ is a Markov state}\}. \quad (9.30)$$