Hilbert space methods for quantum mechanics

Dénes Petz

Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, POB 127, H-1364 Budapest, Hungary petz@renyi.hu

1 Hilbert spaces

The starting point of the quantum mechanical formalism is the **Hilbert space**. The Hilbert space is a mathematical concept, it is a space in the sense that it is a complex vector space which is endowed by an **inner** or **scalar product** $\langle \cdot , \cdot \rangle$. The linear space $\mathbb{C}^n$ of all n-tuples of complex numbers becomes a Hilbert space with the inner product

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

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

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

gives the inner product. The latter space is denoted by $L^2(\mathbb{R}^n)$ and it is infinite dimensional contrary to the n-dimensional space $\mathbb{C}^n$. Below we are mostly satisfied with finite dimensional spaces. The inner product of the vectors $|x\rangle$ and $|y\rangle$ will be often denoted as $\langle x|y\rangle$, this notation, sometimes called bra and ket, is popular in physics. On the other hand, $|x\rangle\langle y|$ is a linear operator which acts on the vector $|z\rangle$ as

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

Therefore,

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

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

1.1 Orthogonal expansions in a Hilbert space

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

(1) $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ $(x, y, z \in \mathcal{H})$,
(2) $\langle \lambda x, y\rangle = \overline{\lambda}\langle x, y\rangle$, $(\lambda \in \mathbb{C},\ x, y \in \mathcal{H})$,
(3) $\langle x, y\rangle = \overline{\langle y, x\rangle}$ $(x, y \in \mathcal{H})$,
(4) $\langle x, x\rangle \geq 0$ for every $x \in \mathcal{H}$ and $\langle x, x\rangle = 0$ only for $x = 0$.

These conditions imply the **Schwarz inequality**

$$\big|\langle x, y\rangle\big|^2 \leq \langle x, x\rangle\, \langle y, y\rangle. \tag{1}$$

The inner product determines a **norm**

$$\|x\| := \sqrt{\langle x, x\rangle} \tag{2}$$

which has the property

$$\|x + y\| \leq \|x\| + \|y\|\,.$$

$\|x\|$ is interpreted as the length of the vector x. A further requirement in the definition of a Hilbert space that every Cauchy sequence must be convergent, that is, the space is **complete**.

Exercise 1.1 *Show that*

$$\|x - y\|^2 + \|x + y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{3}$$

which is called parallelogram law.

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

Example 1.1 *Let $L^2[a, b]$ be the set of square integrable (complex-valued) functions on the interval $[a, b]$. This is a Hilbert space with the inner product*

$$\langle f, g\rangle := \int_a^b \overline{f(x)}\, g(x)\, dx$$

and with the norm

$$\|f\| := \sqrt{\int_a^b \|f(x)\|^2 \, dx} \, .$$

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

Example 1.2 *The infinite dimensional analogue of* $\mathbb{C}^n$ *is the space* $\ell^2(\mathbb{N})$*:*

$$\ell^2(\mathbb{N}); = \{x = (x_1, x_2, \ldots) : x_n \in \mathbb{C}, \sum_n |x_n|^2 < +\infty\}.$$

The inner product is

$$\langle x, x' \rangle := \sum_n \bar{x}_n x'_n \, .$$

The canonical basis in this spaces is the sequence δ_n $(n = 1, 2, \ldots)$*:*

$$\delta_n = (0, 0, \ldots, 1, 0, \ldots) \qquad \textit{(1 is at the nth place).}$$

□

Theorem 1.1 *Let* $x_1, x_2, \ldots$ *be a basis in a Hilbert space* $\mathcal{H}$*. Then for any vector* $x \in \mathcal{H}$ *the expansion*

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

holds.

Example 1.3 *In the space* $L^2[0, \pi]$ *the functions*

$$f_n(x) = \sqrt{\frac{2}{\pi}} \sin nx \tag{4}$$

form a basis. Any function $g \in L^2[0, \pi]$ *has an expansion* $g = \sum_n a_n f_n$*. The convergence is in the* L^2*-norm. (It is known from the theory of Fourier series that for a continuous* g *the expansion is convergent pointwise as well.)* □

Theorem 1.2 (Projection theorem) *Let* $\mathcal{M}$ *be a closed subspace of a Hilbert space* $\mathcal{H}$*. Any vector* $x \in \mathcal{H}$ *can be written in a unique way in the form* $x = x_0 + y$*, where* $x_0 \in \mathcal{M}$ *and* $y \perp \mathcal{M}$*.*

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

$$P(\lambda x + \mu y) = \lambda P x + \mu P y$$

Moreover, $P^2 = P$.

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

$$Ae_k = c_{1k}e_1 + c_{2k}e_2 + \ldots + c_{nk}e_n.$$

The numbers c_{ij} for an $n \times n$ matrix, it is called the **matrix** of the linear transformation A in the basis $e_1, e_2, \ldots, e_n$. When $B : \mathcal{H} \to \mathcal{H}$ is another linear transformation, the the matrix of the composition $A \circ B$ is the usual matrix product of the matrix of A and that of B. If a basis is fixed, then it induces a 1-1 correspondence betweem linear transformations and $n \times n$ matrices.

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

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

Exercise 1.2 *Show that* $\|AB\| \leq \|A\| \, \|B\|$.

Exercise 1.3 *Let f be a continuous function onthe interval $[a, b]$. Define a linear operator $M_f : L^2[a, b] \to L^2[a, b]$ as*

$$M_f g = fg.$$

(This is the multiplication by the function f.) Show that

$$\|M_f\| = \sup\{|f(x)| : x \in [a, b]\}.$$

1.2 The adjoint of a linear operator

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

$$\langle x, Ty\rangle_{\mathcal{K}} = \langle T^*x, y\rangle_{\mathcal{H}} \qquad (x \in \mathcal{H}, y \in \mathcal{K}). \tag{5}$$

$T \in B(\mathcal{H})$ is called self-adjoint if $T^* = T$. T is self-adjoint if and only if $\langle x, Tx\rangle$ is real for every vector $x \in \mathcal{H}$.

Exercise 1.4 *Show that any orthogonal projection is selfadjoint.*

Example 1.4 *Let $S : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ be the right-shift defined as $S\delta_n = \delta_{n+1}$ in the canonical basis. Then*

$$S^*(x_1, x_2, x_3, \ldots) = (x_2, x_3, x_4, \ldots)$$

In another way,

$$S^*\delta_1 = 0, \qquad S^*\delta_{n+1} = \delta_n \,.$$

S^* is called left-shift. □

Theorem 1.3 *The properties of the adjoint:*

(1) $(A+B)^* = A^* + B^*$, $(\lambda A)^* = \overline{\lambda} A^* \qquad (\lambda \in \mathbb{C})$,
(2) $(A^*)^* = A$, $(AB)^* = B^* A^*$,
(3) $(A^{-1})^* = (A^*)^{-1}$ *if* A *is invertible.*
(4) $\|A\| = \|A^*\|$

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

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

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

Example 1.6 *For any* $A \in B(\mathcal{H})$*, the operator* A^*A *is self-adjoint.* □

An invertible operator $U \in B(\mathcal{H})$ is called a **unitary** if $U^{-1} = U^*$.

Example 1.7 *For any* $A = A^* \in B(\mathcal{H})$*, the operator*

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

is a unitary. □

Exercise 1.5 *Show that the product of any two unitary operators is a unitary.*

1.3 Tensor product of Hilbert spaces and operators

Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. Their **algebrac tensor product** cosists of the formal finte sums

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

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

$$\begin{aligned} &(x_1 + x_2) \otimes y = x_1 \otimes y + x_2 \otimes y, \quad (\lambda x) \otimes y = \lambda (x \otimes y)\,, \\ &x \otimes (y_1 + y_2) = x \otimes y_1 + x \otimes y_2, \quad x \otimes (\lambda y) = \lambda (x \otimes y)\,. \end{aligned} \tag{6}$$

The inner product is defined as

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

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

Example 1.8 *If $f \in \mathcal{H} := L^2(X,\mu)$ and $g \in \mathcal{K} := L^2(Y,\nu)$, then $f \otimes g$ can be interpreted as a function of two variables: $f(x)g(y)$.* □

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

If $e_1, e_2, \ldots$ and $f_1, f_2, \ldots$ are bases in $\mathcal{H}$ and $\mathcal{K}$, respectively, then $\{e_i \otimes e_j : i,j\}$ is a basis in the tensor product space. This shows that

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

Example 1.9 *In the Hilbert space $L^2(\mathbb{R}^2)$ we can get a basis if the space is considered as $L^2(\mathbb{R}) \otimes L^2(\mathbb{R})$. In the space $L^2(\mathbb{R})$ the Hermite functions*

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

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

$$\varphi_{nm}(x,y) := e^{-(x^2+y^2)/2}H_n(x)H_m(y) \qquad (n,m = 0,1,\ldots). \tag{7}$$

for a basis in $L^2(\mathbb{R}^2)$. □

Exercise 1.6 *Let $A \in B(\mathcal{H})$ and $B \in B(\mathcal{H})$ be operators on the finite dimensional spaces $\mathcal{H}$ and $\mathcal{K}$. Show that*

$$\det(A \otimes B) = (\det A)^m (\det B)^n,$$

where $n = \dim \mathcal{H}$ and $m = \dim \mathcal{K}$. (Hint: The determinant is the product of the eigenvalues.)

Exercise 1.7 *Show that $\|A \otimes B\| = \|A\| \cdot \|B\|$.*

Example 1.10 *Let $\{e_1, e_2, e_3\}$ be a basis in $\mathcal{H}$ and $\{f_1, f_2\}$ be a basis in $\mathcal{K}$. If $[A_{ij}]$ is the matrix of $A \in B(\mathcal{H}_1)$ and $[B_{kl}]$ is the matrix of $B \in B(\mathcal{H}_2)$, then*

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

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

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

□

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

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

$$v_1 \wedge v_2 \wedge \ldots \wedge v_k := \frac{1}{\sqrt{k!}} \sum_{\pi} (-1)^{\sigma(\pi)} v_{\pi(1)} \otimes v_{\pi(2)} \otimes \ldots \otimes v_{\pi(k)} \tag{8}$$

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

The computational rules for the antisymmetric tensors are similar to (6):

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

and

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

The subspace spanned by the vectors $v_1 \wedge v_2 \wedge \ldots \wedge v_k$ is called the kth antisymmetric tensor power of $\mathcal{H}$, in notation $\wedge^k \mathcal{H}$. So $\wedge^k \mathcal{H} \subset \otimes^k \mathcal{H}$. If $A \in B(\mathcal{H})$, then the transformation $\otimes^k A$ leaves the subspace $\wedge^k \mathcal{H}$ invariant. Its restriction is dented by $\wedge^k A$ which is equivalently defined as

$$\wedge^k A(v_1 \wedge v_2 \wedge \ldots \wedge v_k) = Av_1 \wedge Av_2 \wedge \ldots \wedge Av_k. \tag{9}$$

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

$$\{e_{i(1)} \wedge e_{i(2)} \wedge \ldots \wedge e_{i(k)} \ : \ 1 \leq i(1) < i(2) < \ldots < i(k)) \leq n\} \tag{10}$$

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

$$\binom{n}{k} \quad \text{ha} \quad k \leq n,$$

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

$$\wedge^n A = \lambda \times \text{identity} \tag{11}$$

Exercise 1.8 *Show that* $\lambda = \det A$ *in (11). Use this to prove that* $\det(AB) = \det A \times \det B$. *(Hint: Show that* $\wedge^k(AB) = (\wedge^k A)(\wedge^k B)$.*)*

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

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

where the summation is over all permutations π of the set $\{1, 2, \ldots, k\}$ again. The linear span of the symmetric tensors is the symmetric tensor power $\vee^k \mathcal{H}$. It has the basis

$$\{e_{i(1)} \vee e_{i(2)} \vee \ldots \vee e_{i(k)} \ : \ 1 \le i(1) \le i(2) \le \ldots \le i(k) \le n\}. \tag{12}$$

Exercise 1.9 *Give the dimension of* $\vee^k \mathcal{H}$ *if* $\dim(\mathcal{H}) = n$.

1.4 Positive operators

$T \in B(\mathcal{H})$ is called **positive** if $\langle x, Tx \rangle \ge 0$ for every vector $x \in \mathcal{H}$, in notation $T \ge 0$. A positive operator is self-adjoint.

Exercise 1.10 *Show that an orthogonal projection is positive.*

Theorem 1.4 *Let $T \in B(\mathcal{H})$ be a self-adjoint operator and $e_1, e_2, \ldots, e_n$ be a basis in the Hilbert space $\mathcal{H}$. T is positive if and only if for any $1 \le k \le n$ the determinant of the $k \times k$ matrix*

$$(\langle e_i, Te_j \rangle)_{ij=1}^k$$

is positive.

The spectrum, in particular the eigenvalues of a positive operator, lies in $\mathbb{R}^+$. Conversely, if all the eigenvalues are positive for a self-adjoint operator acting on a finite dimensional space, then it is positive. Positive matrices are also called positive semidefinite.

Let $A, B \in B(\mathcal{H})$ be self-adjoint operators. $A \le B$ if $B - A$ is positive.

Example 1.11 *Let $f : \mathbb{R}^+ \to \mathbb{R}$ be a smooth function. f is called* **matrix monotone** *if*

$$0 \le A \le B \quad \textit{implies that} \quad f(A) \le f(B) \,.$$

f is matrix monotone if and only for every positive operator A and X and for the real parameter $t \ge 0$,

$$\frac{\partial}{\partial t} \langle x, f(A + tX)x \rangle \ge 0$$

holds for every vector x which means that

$$\frac{\partial}{\partial t} f(A + tX) \ge 0.$$

We want to show that the squareroot function is matrix monotone. Let

$$F(t) := \sqrt{A + tX}\,.$$

It is enough to see that the eigenvalues of $F'(t)$ are positive. Differentiating the equality $F(t)F(t) = A + tX$, we get

$$F'(t)F(t) + F(t)F'(t) = X.$$

If $F'(t) = \sum_i \lambda_i E_i$ is the spectral decomposition, then

$$\sum_i \lambda_i (E_i F(t) + F(t)E_i) = X$$

and after multiplication by E_j from the left and from the right, we have for the trace

$$2\lambda_j \, Tr\, E_j F(t) E_j = Tr\, E_j X E_j.$$

Since both traces are positive, λ_j must be positive as well. □

Exercise 1.11 *Show that that the square function is not matrix monotone. (Hint: Choose A to be diagonal and*

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Use the argument of the previous example for 2×2 matrices.)

1.5 The spectral theorem

The eigenvalues of a self-adjoint matrix are real and the eigenvectors corresponding to different eigenvalues are orthogonal. Therefore, the matrix (or the corresponding Hilbert space operator) can be written in the form

$$\sum_{i=1}^{k} \lambda_i E_i\,,$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the different eigenvalues and E_i is the orthogonal projection onto the subspace spanned by the igenvectors correspondg to the eigenvalue λ_i, $1 \le i \le k$. The spectral theorem extends this to arbitrary self-adjoint operator A. Then the spectrum is not necessary discrete and the finite sum is replaced by an integral.

Let $\mathcal{X}$ be a complete separable metric space and $\mathcal{H}$ be a Hilbert space. Assume that for each Borel set $B \subset \mathcal{X}$ a positive operator $E(B) \in B(\mathcal{H})$ is given such that

(1) $0 \le E(B) \le I$, $E(\emptyset) = 0$, $E(\mathbb{C}) = I$,

(2) If (B_i) is a sequence of pairwise disjoint Borel subset of $\mathcal{X}$ and $B = \cup_{i=1}^{\infty} B_i$, then

$$E(B)e = \sum_{i=1}^{\infty} E(B_i)e$$

for every vector $e \in \mathcal{H}$.

In this case E is called a **positive operator-valued measure**, shortly **POVM**. In the most important examples $\mathcal{X}$ is a finite set, the real line $\mathbb{R}$ or the unit circle $\mathbb{T}$.

We want to integrate a function $f : \mathcal{X} \to \mathbb{C}$ with respect an POVM on $\mathcal{X}$. When $\mathcal{X}$ is a finite set, then

$$\int_{\mathcal{X}} f(x)\, dE(x) = \sum_{x \in \mathcal{X}} f(x) E(\{x\})$$

is a finite sum. In the general case, the definition of the integral can be reduced to many integrals with respect to common measures. Given a vector $e \in \mathcal{H}$,

$$\mu_e(B) = \langle e, E(B)e \rangle$$

gives us a positive measure on the Borel sets of $\mathcal{X}$. We say that the integral $\int_{\mathcal{X}} f(x)\, dE(x) = T \in B(\mathcal{H})$, if

$$\langle e, Te \rangle = \int_{\mathcal{X}} f(x)\, d\mu_e(x)$$

holds for every $e \in \mathcal{X}$.

A POVM E is called **projection-valued measure** if $E(B)$ is a projection operator for every Borel set B, that is $E(B) = E(B)^2$.

Exercise 1.12 *Let E be a projection-valued measure and let B_1, B_2 be disjoint Borel set. Show that if a vector e is in the range of $E(B_1)$, then $E(B_2)e = 0$. (Therefore, $E(B_1)$ and $E(B_2)$ are orthogonal.)*

The next theorem is the **spectral theorem** for a bounded self-adjoint operator.

Theorem 1.5 *Let $A = A^* \in B(\mathcal{H})$. Then there exists a unique projection-valued measure on the real line such that*

$$A = \int \lambda\, dE(\lambda).$$

Moreover, if $B \subset \mathbb{R}$ and the spectrun of A are disjoint, then $E(B) = 0$ and

$$f(A) = \int f(\lambda)\, dE(\lambda)$$

for every continuos function defined on the spectrum of A.

The projection-valued measure in the theorem is called the **spectral measure** of the operator A. Similar result holds for unbounded self-adjoint operator A but in this case A and $f(A)$ are not everywhere defined operators. Similar theorem holds for unitary operators, then the spectral measure is on the unit circle.

2 Postulates of quantum mechanics

The first postulate of quantum mechanics tells that to each quantum mechanical system a Hilbert space $\mathcal{H}$ is associated. The (pure) physical states of the system correspond to unit vectors of the Hilbert space. This correspondance is not 1-1. When f_1 and f_2 are unit vectors, then the corresponding states identical if $f_1 = zf_2$ for a complex number z of modulus 1. Such z is often called **phase**.

2.1 n-level quantum systems

The **pure physical state** of the system determines a corresponding state vector up to a phase.

Example 1.12 *The 2 dimensional Hilbert space $\mathbb{C}^2$ is used to describe a 2-level quantum system called qubit. The canonical basis vectors $(1,0)$ and $(0,1)$ are usually denoted by $|\uparrow\rangle$ and $|\downarrow\rangle$, respectively. (An alternative notation is $|1\rangle$ for $(0,1)$ and $|0\rangle$ for $(1,0)$.) Since the polarization of a photon is an important example of a qubit, the state $|\uparrow\rangle$ may have the interpretation that the "polarization is vertical") and $|\downarrow\rangle$ means that the "polarization is horizontal".*

To specify a state of a qubit we need to give a real number x_1 and a complex number z such that $x_1^2 + |z|^2 = 1$. Then the state vector is

$$x_1 |\uparrow\rangle + z |\downarrow\rangle .$$

(Indeed, multiplying a unit vector $z_1 |\uparrow\rangle + z_2 |\downarrow\rangle$ by an appropriate phase, we can make the coefficient of $|\uparrow\rangle$ real and the corresponding state remains the same.)

Splitting z into real and imaginary parts as $z = x_2 + \mathrm{i}x_3$, we have the constraint $x_1^2 + x_2^2 + x_3^2 = 1$ for the parameters $(x_1, x_2, x_3) \in \mathbb{R}^3$.

Therefore, the space of all pure states of a qubit is conveniently visualized as the sphere in the three dimensional Euclidean space, it is called the **Bloch sphere**. □

Traditional quantum mechanics distinguishes between pure states and **mixed states**. Mixed states are described by **density matrices**. A density matrix or statistical operator is a positive operator of trace 1 on the Hilbert space. This means that the space has a basis consisting of eigenvectors of the

statistical operator and the sum of eigenvalues is 1. (In the finite dimensional case the first condition is automatically fulfilled.) The pure states represented by unit vectors of the Hilbert space are among the density matrices under an appropriate identification. If $x = |x\rangle$ is a unit vector, then $|x\rangle\langle x|$ is a density matrix. Geometrically $|x\rangle\langle x|$ is the orthogonal projection onto the linear subspace generated by x. Note that $|x\rangle\langle x| = |y\rangle\langle y|$ if the vectors x and y differ in a phase.

(A1) The physical states of a quantum mechanical system are described by statistical operators acting on the Hilbert space.

Example 1.13 *A state of the spin (of* $1/2$*) can be represented by the* 2×2 *matrix*

$$\frac{1}{2}\begin{bmatrix} 1+x_3 & x_1 - \mathrm{i}x_2 \\ x_1 + \mathrm{i}x_2 & 1-x_3 \end{bmatrix}. \tag{13}$$

This is a density matrix if and only if $x_1^2 + x_2^2 + x_3^2 \leq 1$. □

The second axiom is about observables.

(A2) The observables of a quantum mechanical system are described by self-adjoint operators acting on the Hilbert space.

A **self-adjoint operator** A on a Hilbert space $\mathcal{H}$ is a linear operator $\mathcal{H} \to \mathcal{H}$ which satisfies

$$\langle Ax,\, y\rangle = \langle x,\, Ay\rangle$$

for $x, y \in \mathcal{H}$. Self-adjoint operators on a finite dimensional Hilbert space $\mathbb{C}^n$ are $n \times n$ self-adjoint matrices. A self-adjoint matrix admits a **spectral decomposition** $A = \sum_i \lambda_i E_i$, where λ_i are the different eigenvalues of A and E_i is the orthogonal projection onto the subspace spanned by the eigenvectors corresponding to the eigenvalue λ_i. Multiplicity of λ_i is exactly the rank of E_i.

Example 1.14 *In case of a quantum spin (of* $1/2$*) the matrices*

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

are used to describe the spin of direction x, y, z *(with respect to a coordinate system.) They are called* **Pauli matrices**. *Any* 2×2 *self-adjoint matrix is of the form*

$$A_{(x_0,x)} := x_0\sigma_0 + x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3$$

if σ_0 *stands for the unit matrix* I. *We also use the shorthand notation* $x_0\sigma_0 + x \cdot \sigma$.

The density matrix (13) can be written as

$$\tfrac{1}{2}(\sigma_0 + x \cdot \sigma), \tag{14}$$

where $\|x\| \leq 1$. Formula (14) makes an affine correspondence between 2×2 density matrices and the unit ball in the Euclidean 3-space. The extreme points of the ball correspond to pure state and any mixed state is the convex combination of pure states in infinitely many different ways. In higher dimension the situation is much more complicated. □

Any density matrix can be written in the form

$$\rho = \sum_i \lambda_i |x_i\rangle\langle x_i| \tag{15}$$

by means of unit vectors $|x_i\rangle$ and coefficients $\lambda_i \geq 0$, $\sum_i \lambda_i = 1$. Since ρ is self-adjoint such a decomposition is deduced from the spectral theorem and the vectors $|x_i\rangle$ may be chosen pairwise orthogonal eigenvectors and λ_i are the corresponding eigenvalues. Under this condition (15) is called **Schmidt decomposition**. It is unique if the spectrum of ρ is non-degenerate, that is, there is no multiple eigenvalue.

2.2 Measurements

Quantum mechanics is not deterministic. If we prepare two identical systems in the same state, and we measure the same observable on each, then the result of the **measurement** may not be the same. This indeterminism or stochastic feature is fundamental.

(A3) Let $\mathcal{X}$ be a finite set and for $x \in \mathcal{X}$ an operator $V_x \in B(\mathcal{H})$ be given such that $\sum_x V_x^* V_x = I$. Such an indexed family of operators is a model of a measurement with values in $\mathcal{X}$. If the measurement is performed in a state ρ, then the outcome $x \in \mathcal{X}$ appears with probability $\operatorname{Tr} V_x \rho V_x^*$ and after the measurement the state of the system is

$$\frac{V_x \rho V_x^*}{\operatorname{Tr} V_x \rho V_x^*}.$$

A particular case is the measurement of an observable described by a self-adjoint operator A with spectral decomposition $\sum_i \lambda_i E_i$. In this case $\mathcal{X} = \{\lambda_i\}$ is the set of eigenvalues and $V_i = E_i$. One compute easily that the expectation of the random outcome is $\operatorname{Tr} \rho A$. The functional $A \mapsto \operatorname{Tr} \rho A$ is linear and has two important properties: 1. If $A \geq 0$, then $\operatorname{Tr} \rho A \geq 0$, 2.$\operatorname{Tr} \rho I = 1$. These properties allow to see quantum states in a different way. If $\varphi : B(\mathcal{H}) \to \mathbb{C}$ is a linear functional such that

$$\varphi(A) \geq 0 \quad \text{if} \quad A \geq 0 \quad \text{and} \quad \varphi(I) = 1, \tag{16}$$

then there exists a density matrix ρ_φ such that

$$\varphi(A) = \operatorname{Tr} \rho_\varphi A. \tag{17}$$

The functional φ associates the expectation value to the observables A.

2.3 Composite systems

According to axiom (A1), a Hilbert space is associated to any quantum mechanical system. Assume that a **composite system** consists of the subsystems (1) and (2), they are described by the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. (Each subsystem could be a particle or a spin, for example.) Then we have

(A4) The composite system is described by the tensor product Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$.

When $\{e_j : j \in J\}$ is a basis in $\mathcal{H}_1$ and $\{f_i : i \in I\}$ is a basis in $\mathcal{H}_2$, then $\{e_j \otimes f_j : j \in J, i \in I\}$ is a basis of $\mathcal{H}_1 \otimes \mathcal{H}_2$. Therefore, the dimension of $\mathcal{H}_1 \otimes \mathcal{H}_2$ is $\dim \mathcal{H}_1 \times \dim \mathcal{H}_2$. If $A_i \in B(\mathcal{H}_i)$ $(i = 1, 2)$, then the action of the tensor product operator $A_1 \otimes A_2$ is determined by

$$(A_1 \otimes A_2)(\eta_1 \otimes \eta_2) = A_1\eta_1 \otimes A_2\eta_2$$

since the vectors $\eta_1 \otimes \eta_2$ span $\mathcal{H}_1 \otimes \mathcal{H}_2$.

When $A = A^*$ is an observable of the first system, then its expectation value in the vector state $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$, is

$$\langle \psi, (A \otimes I_2)\psi \rangle \,,$$

where I_2 is the identity operator on $\mathcal{H}_2$.

Example 1.15 *The Hilbert space of a composite system of two spins (of* $1/2$*) is* $\mathbb{C}^2 \otimes \mathbb{C}^2$*. In this space, the vectors*

$$e_1 := |\uparrow\rangle \otimes |\uparrow\rangle, \quad e_2 := |\uparrow\rangle \otimes |\downarrow\rangle, \quad e_3 := |\downarrow\rangle \otimes |\uparrow\rangle, \quad e_4 := |\downarrow\rangle \otimes |\downarrow\rangle$$

form a basis. The vector state

$$\phi = \frac{1}{\sqrt{2}}(|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle) \tag{18}$$

has a surprising property. Consider the observable

$$A := \sum_{i=1}^{4} i|e_i\rangle\langle e_i|,$$

which has eigenvalues $1, 2, 3, 4$ *and the corresponding eigenvectors are just the basis vectors. Measurement of this observable yields the values* $1, 2, 3, 4$ *with probabilities* $0, 1/2, 1/2$ *and* 0*, respectively. The 0 probability occurs when both spins are up or both are down. Therefore in the vector state* ϕ *the spins are anti-correlated.* □

We consider now the composite system $\mathcal{H}_1 \otimes \mathcal{H}_2$ in a state $\phi \in \mathcal{H}_1 \otimes \mathcal{H}_2$. Let $A \in B(\mathcal{H}_1)$ be an observable which is localized at the first subsystem. If we want to consider A as an observable of the total system, we have to define an extension to the space $\mathcal{H}_1 \otimes \mathcal{H}_2$. The tensor product operator $A \otimes I$ will do, I is the identity operator of $\mathcal{H}_2$.

Lemma 1.1 *Assume that $\mathcal{H}_1$ and $\mathcal{H}_2$ are finite dimensional Hilbert spaces. Let $\{e_j : j \in J\}$ be a basis in $\mathcal{H}_1$ and $\{f_i : i \in I\}$ be a basis in $\mathcal{H}_2$. Assume that*

$$\phi = \sum_{i,j} w_{ij}\, e_j \otimes f_i$$

is the expansion of a unit vector $\phi \in \mathcal{H}_1 \otimes \mathcal{H}_2$. Set W for the matrix which is determined by the entries w_{kl}. Then W^*W is a density matrix and

$$\langle \phi,\ (A \otimes I)\phi \rangle = Tr\, AW^*W\,.$$

Proof. Let E_{kl} be an operator on $\mathcal{H}_1$ which is determined by the relations $E_{kl}e_j = \delta_{lj}e_k$ $(k,l \in I)$. As a matrix, E_{kl} is called matrix unit, it is a matrix such that (k,l) entry is 1, all others are 0. Then

$$\begin{aligned}\langle \phi,\ (E_{kl} \otimes I)\phi \rangle &= \left\langle \sum_{i,j} w_{ij}\, e_j \otimes f_i,\ (E_{kl} \otimes I) \sum_{t,u} w_{tu}\, e_u \otimes f_t \right\rangle = \\ &= \sum_{i,j} \sum_{t,u} \overline{w}_{ij} w_{tu} \langle e_j,\ E_{kl} e_u \rangle \langle f_i, f_t \rangle = \\ &= \sum_{i,j} \sum_{t,u} \overline{w}_{ij} w_{tu}\, \delta_{lu}\delta_{jk}\delta_{it} = \sum_i \overline{w}_{ik} w_{il}\,.\end{aligned}$$

Then we arrived at the (k,l) entry of W^*W. Our computation may be summarized as

$$\langle \phi,\ (E_{kl} \otimes I)\phi \rangle = \mathrm{Tr}\, E_{kl}(W^*W) \qquad (k,l \in I).$$

Since any linear operator $A \in B(\mathcal{H}_1)$ is of the form $A = \sum a_{kl}E_{kl}$ $(a_{kl} \in \mathbb{C})$, taking linear combinations of the previous equations, we have

$$\langle \phi,\ (A \otimes I)\phi \rangle = \mathrm{Tr}\, A(W^*W)\,.$$

W^*W is obviously positive and

$$\mathrm{Tr}\, W^*W = \sum_{i,j} |w_{ij}|^2 = \|\phi\|^2 = 1\,.$$

Therefore it is a density matrix. □

This lemma shows a natural way from state vectors to density matrices. Given a density matrix ρ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ there are density matrices $\rho_i \in B(\mathcal{H}_i)$ such that

$$\mathrm{Tr}\,(A \otimes I)\rho = \mathrm{Tr}\, A\rho_1 \qquad (A \in B(\mathcal{H}_1)) \tag{19}$$

and

$$\mathrm{Tr}\,(I \otimes B)\rho = \mathrm{Tr}\, B\rho_2 \qquad (B \in B(\mathcal{H}_2)). \tag{20}$$

ρ_1 and ρ_2 are called **reduced density matrices**. (They are the quantum analogue of marginal distributions.)

The proof of Lemma 1.1 contains the reduced density of $|\phi\rangle\langle\phi|$ on the first system, it is W^*W. One computes similarly that the reduced density on the second subsystem, it is $(WW^*)^t$, where X^t denotes the transpose of the matrix X. Since W^*W and $(WW^*)^t$ have the same non-zero eigenvalues, the two subsystems are very strongly connected if the total system is in a pure state.

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces and let $\dim \mathcal{H}_1 = m$ and $\dim \mathcal{H}_2 = n$. It is well-known that the matrix of a linear operator on $\mathcal{H}_1 \otimes \mathcal{H}_2$ has a block-matrix form

$$U = (U_{ij})_{i,j=1}^m = \sum_{i,j=1}^m E_{ij} \otimes U_{ij},$$

relative to the lexicographically ordered product basis, where U_{ij} are $n \times n$ matrices. For example,

$$A \otimes I = (X_{ij})_{i,j=1}^m, \quad \text{where} \quad X_{ij} = A_{ij} I_n$$

and

$$I \otimes B = (X_{ij})_{i,j=1}^m, \quad \text{where} \quad X_{ij} = \delta_{ij} B.$$

Assume that

$$\rho = (\rho_{ij})_{i,j=1}^m$$

is a density matrix of the composite system, then

$$\mathrm{Tr}\,(A \otimes I)\rho = \sum_{i,j} A_{ij} \mathrm{Tr}\, I_n \rho_{ij} = \sum_{i,j} A_{ij} \mathrm{Tr}\, \rho_{ij}$$

and this gives that for the first reduced density matrix we have

$$(\rho_1)_{ij} = \mathrm{Tr}\, \rho_{ij}. \tag{21}$$

We can compute similarly the second reduced density ρ_2. Since

$$\mathrm{Tr}\,(I \otimes B)\rho = \sum_i \mathrm{Tr}\, B \rho_{ii}$$

we obtain

$$\rho_2 = \sum_{i=1}^m \rho_{ii}. \tag{22}$$

The reduced density matrices might be expressed by the **partial taces**. $\mathrm{Tr}_2 : B(\mathcal{H}_1) \otimes B(\mathcal{H}_2) \to B(\mathcal{H}_1)$ and $\mathrm{Tr}_1 : B(\mathcal{H}_1) \otimes B(\mathcal{H}_2) \to B(\mathcal{H}_2)$ are defined as

$$\mathrm{Tr}_2(A \otimes B) = A \mathrm{Tr}\, B, \qquad \mathrm{Tr}_1(A \otimes B) = \mathrm{Tr} AB\,. \tag{23}$$

We have

$$\rho_1 = \mathrm{Tr}_2 \rho \qquad \text{and} \qquad \rho_2 = \mathrm{Tr}_1 \rho\,. \tag{24}$$

Axiom (A4) tells about a composite quantum system consisting of two quantum components. In case of more quantum components, the formalism is similar, more tensor factors appear. It may happen that the quantum system under study has a classical and a quantum component, assume that the first component is classical. Then the description by tensor product Hilbert space is still possible. A basis $(|e_i\rangle)_i$ of $\mathcal{H}_1$ can be fixed and the possible density matrices of the joint system are of the form

$$\sum_i p_i |e_i\rangle\langle e_i| \otimes \rho_i^{(2)}\,, \tag{25}$$

where $(p_i)_i$ is a probability distribution and $\rho_i^{(2)}$ are densities on $\mathcal{H}_2$. Then the reduced state on the first component is the probability density $(p_i)_i$ (which may be regarded as a diagonal density matrix) and $\sum_i p_i\rho_i^{(2)}$ is the second reduced density.

Another postulate of quantum mechanics tells about the **time development** of a closed quantum system. If the system is not subject to any measurement in the time interval $I \subset \mathbb{R}$ and ρ_t denotes the statistical operator at time t, then

(A5) $\quad \rho_t = U(t,s)\rho_s U(t,s)^* \qquad (t,s \in I),$

where the **unitary propagator** $U(t,s)$ is a family of unitary operators such that

(i) $U(t,s)U(s,r) = U(t,r)$,
(ii) $(s,t) \mapsto U(s,t) \in B(\mathcal{H})$ is strongly continuous.

The first order approximation of the unitary $U(s,t)$ is the **Hamiltonian**:

$$U(t+\Delta t, t) = I - \frac{\mathrm{i}}{\hbar} H(t)\Delta t,$$

where $H(t)$ is the Hamiltonian at time t. If the Hamiltonian is time independent, then

$$U(s,t) = \exp\left(-\frac{\mathrm{i}}{\hbar}(s-t)H\right).$$

In the approach followed here the density matrices are transformed in time, this is the so-called **Schrödinger picture** of quantum mechanics. When discrete time development is considered, a single unitary U gives the transformation of the vector state in the form $\psi \mapsto U\psi$, or in the density matrix formalism $\rho \mapsto U\rho U^*$. When the unitary time development is viewed as a quantum algorithm in connection with quantum computation, the term **gate** is used instead of unitary. So the gates constitute an algorithm are simply unitary operators.

2.4 State transformations

Assume that $\mathcal{H}$ is the Hilbert space of our quantum system which initially has a statistical operator ρ (acting on $\mathcal{H}$). When the quantum systems is not closed, it is coupled to another system, called **environment**. The environment has a Hilbert space $\mathcal{H}_e$ and statistical operator ρ_e. Before interaction the total system has density $\rho_e \otimes \rho$. The dynamical change caused by the interaction is implemented by a unitary and $U(\rho_e \otimes \rho)U^*$ is the new statistical operator and the reduced density $\tilde{\rho}$ is the new statistical operator of the quantum system we are interested in. The affine change $\rho \mapsto \tilde{\rho}$ is typical for quantum mechanics and called **state transformation**. In this way the map $\rho \mapsto \tilde{\rho}$ is defined on density matrices but it can be extended by linearity to all matrices. In this way we obtain a trace preserving and positivity preserving linear transformation.

The above defined state transformation can be described in several other forms, reference to the environment could be omitted completely. Assume that ρ is an $n \times n$ matrix and ρ_e is of the form $(z_k\overline{z_l})_{kl}$ where $(z_1, z_2, \dots, z_m)$ is a unit vector in the m dimensional space $\mathcal{H}_e$. (ρ_e is pure state.) All operators acting on $\mathcal{H}_e \otimes \mathcal{H}$ are written in a block matrix form, they are $m \times m$ matrices with $n \times n$ matrix entries. In particular, $U = (U_{ij})_{i,j=1}^m$ and $U_{ij} \in M_n$. If U is a unitary, then U^*U is the identity and this implies that

$$\sum_i U_{ik}^* U_{il} = \delta_{kl} I_n \tag{26}$$

Formula (22) for the reduced density matrix gives

$$\begin{aligned}
\tilde{\rho} &= \sum_i (U(\rho_e \otimes \rho)U^*)_{ii} \\
&= \sum_{i,k,l} U_{ik}(\rho_e \otimes \rho)_{kl}(U^*)_{li} \\
&= \sum_{i,k,l} U_{ik}(z_k\overline{z}_l\rho)(U_{il})^* \\
&= \sum_i \Big(\sum_k z_k U_{ik}\Big)\rho\Big(\sum_l z_l U_{il}\Big)^* \\
&= \sum_i A_i \rho A_i^*
\end{aligned}$$

where the operators $A_i := \sum_k z_k U_{ik}$ satisfy

$$\sum_p A_p^* A_p = I \tag{27}$$

due to (26) and $\sum_k |z_k|^2 = 1$.

Theorem 1.6 *Any state transformation $\rho \mapsto \mathcal{E}(\rho)$ can be written in the form*

$$\mathcal{E}(\rho) = \sum_p A_p \rho A_p^*,$$

where the operator coefficients satisfy (27). *Conversely, all linear mappings of this form are state transformations.*

The first part of the theorem was obtained above. To prove the converse part, we need to solve the equations

$$A_i := \sum_k z_k U_{ik} \qquad (i = 1, 2, \ldots, m).$$

Choose simply $z_1 = 1$ and $z_2 = z_3 = \ldots = z_m = 0$ and the equations reduce to $U_{p1} = A_p$. This means that the first column is given from the block matrix U and we need to determine the other columns such a way that U should be a unitary. Thanks to the condition (27) this is possible. Condition (27) tells us that the first column of our block matrix determines an isometry which extends to a unitary. □

The coefficients A_p in the **operator-sum representation** are called the **operation elements** of the state transformation. The terms quantum (state) operation and channeling transformation are also often used instead of state transformation.

The state transformations form a convex subset of the set of all positive trace preserving linear transformations. (It is not known what the extreme points of this set are.)

$\mathcal{E}$ is called **completely positive** if $\mathcal{E} \otimes id_n$ is positivity preserving for the identical mapping $id_n : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ on any matrix algebra.

Theorem 1.7 *Let $\mathcal{E} : M_n(\mathbb{C}) \to M_k(\mathbb{C})$ be a linear mapping. Then $\mathcal{E}$ is completely positive if and only if it admits a representation*

$$\mathcal{E}(A) = \sum_u V_u A V_u^* \tag{28}$$

by means of some linear operators $V_u : \mathbb{C}^n \to \mathbb{C}^k$.

This result was first proven by Kraus. It follows that stochastic mappings are completely positive and the operator-sum representation is also called **Kraus representation**. Note that this representation is not unique.

Let $\mathcal{E} : M_n(\mathbb{C}) \to M_k(\mathbb{C})$ be a linear mapping. $\mathcal{E}$ is determined by the block-matrix $(X_{ij})_{1 \le i,j \le k}$, where

$$X_{ij} = \mathcal{E}(E_{ij}) \tag{29}$$

(Here E_{ij} denote the matrix units.) This is the **block-matrix representation** of $\mathcal{E}$.

Theorem 1.8 *Let $\mathcal{E} : M_n(\mathbb{C}) \to M_k(\mathbb{C})$ be a linear mapping. Then $\mathcal{E}$ is completely positive if and only if the representing block-matrix $(X_{ij})_{1\le i,j\le k} \in M_k(\mathbb{C}) \otimes M_n(\mathbb{C})$ is positive.*

Example 1.16 *Consider the transpose mapping $A \mapsto A^t$ on 2×2 matrices:*

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

The representing block-matrix is

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

This is not positive, so the transpose mapping is not completely positive. □

Example 1.17 *Consider a positive trace-preserving transformation $\mathcal{E} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ such that its range consists of commuting operators. We show that $\mathcal{E}$ is automatically a state transformation.*

Since a commutative subalgebra of $M_m(\mathbb{C})$ is the linear span of some pairwise orthogonal projections P_k, one can see that $\mathcal{E}$ has the form

$$\mathcal{E}(A) = \sum_k P_k \, Tr \, F_k A, \tag{30}$$

where F_k is a positive operator in $M_n(\mathbb{C})$, it induces the coefficient of P_k as a linear functional on $M_n(\mathbb{C})$.

We want to show the positivity of the representing block-matrix:

$$\sum_{ij} E_{ij} \otimes \Big(\sum_k P_k \, Tr\,(F_k E_{ij}) \Big) = \sum_k \Big(\sum_{ij} E_{ij} \otimes P_k \Big) \circ \Big(\sum_{ij} E_{ij} \, Tr\,(F_k E_{ij}) \otimes I \Big),$$

where $\circ$ denotes the Hadamard (or entry-wise product) of $nm \times nm$ matrices. Recall that according to Schur's theorem the **Hadamard product** *of positive matrices is positive. The first factor is*

$$[P_k, P_k, \dots, P_k]^*[P_k, P_k, \dots, P_k]$$

and the second factor is $F_k \otimes I$, both are positive.

Consider the particular case of (30) where each P_k is of rank one and $\sum_{k=1}^r F_k = I$. Such a family of F_k's describe a measurement which associates the r-tuple $(Tr\,\rho F_1, Tr\,\rho F_2, \dots, Tr\,\rho F_r)$ to the density matrix ρ. Therefore a measurement can be formulated as a state transformation with diagonal outputs. □

The Kraus representation and the block-matrix representation are convenient ways to describe a state transformation in any finite dimension. In the 2×2 case we have the possibility to expand the mappings in the basis $\sigma_0, \sigma_1, \sigma_2, \sigma_3$.

Any trace preserving mapping $\mathcal{E} : M_2(\mathbb{C}) \to M_2(\mathbb{C})$ has a matrix

$$T = \begin{bmatrix} 1 & 0 \\ t & T_3 \end{bmatrix}$$

with respect to this basis, where $T_3 \in M_3$ and

$$\mathcal{E}(w_0\sigma_0 + w \cdot \sigma) = w_0\sigma_0 + (t + T_3 w) \cdot \sigma. \tag{31}$$

The following examples of state transformations are given in term of the T-representation:

Example 1.18 (Pauli channels) *$t = 0$ and $T_3 = Diag\,(\alpha, \beta, \gamma)$. Density matrices are sent to density matrices if and only if*

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

for the real parameters α, β, γ.

It is not difficult to compute the representing block-matrix, we have

$$X = \begin{bmatrix} \frac{1+\gamma}{2} & 0 & 0 & \frac{\alpha+\beta}{2} \\ 0 & \frac{1-\gamma}{2} & \frac{\alpha-\beta}{2} & 0 \\ 0 & \frac{\alpha-\beta}{2} & \frac{1-\gamma}{2} & 0 \\ \frac{\alpha+\beta}{2} & 0 & 0 & \frac{1+\gamma}{2} \end{bmatrix}. \tag{32}$$

X is unitarily equivalent to the matrix

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

This matrix is obviously positive if and only if

$$|1 \pm \gamma| \geq |\alpha \pm \beta|. \tag{33}$$

This positivity condition holds when $\alpha = \beta = \gamma = p > 0$. Hence the next example gives a channeling transformation. □

3 Some applications

In the traditional approach to quantum mechanics, a physical system is described in a Hilbert space: Observables correspond to self-adjoint operators and statistical operators are associated with the states. Von Neumann associated an entropy quantity to a statistical operator in 1927 [15] and the discussion was extended in his book [16].

3.1 Von Neumann entropy

Von Neumann's argument was a gedanken experiment on the ground of phenomenological thermodynamics which is not repeated here, only his conclusion. Assume that the density ρ is the mixture of orthogonal densities ρ_1 and ρ_2, $\rho = p\rho_1 + (1-p)\rho_2$. Then

$$pS(\rho_1) + (1-p)S(\rho_2) = S(\rho) + \kappa p \log p + \kappa(1-p)\log(1-p)\,, \tag{34}$$

where S is a certain thermodynamical entropy quantity, relative to the fixed temperature and molecule density. (Remember that the ortogonality of states has a particular meaning in quantum mechanics.) From the two-component mixture, we can easily move to an arbitrary density matrix $\rho = \sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i|$ and we have

$$S(\rho) = \sum_i \lambda_i S(|\varphi_i\rangle\langle\varphi_i|) - \kappa \sum_i \lambda_i \log \lambda_i. \tag{35}$$

This formula reduces the determination of the (thermodynamical) entropy of a mixed state to that of pure states. The so-called **Schatten decomposition** $\sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i|$ of a statistical operator is not unique although $\langle\varphi_i, \varphi_j\rangle = 0$ is assumed for $i \neq j$. When λ_i is an eigenvalue with multiplicity, then the corresponding eigenvectors can be chosen in many ways. If we expect the entropy $S(\rho)$ to be independent of the Schatten decomposition, then we are led to the conclusion that $S(|\varphi\rangle\langle\varphi|)$ must be independent of the state vector $|\varphi\rangle$. This argument assumes that there are no super-selection sectors, that is, any vector of the Hilbert space can be a state vector. (Von Neumann's argument was somewhat different, see the original paper [15] or [23].) If the entropy of pure states is defined to be 0 as a kind of normalization, then we have the **von Neumann entropy** formula:

$$S(\rho) = -\kappa \sum_i \lambda_i \log \lambda_i = \kappa \mathrm{Tr}\, \eta(\rho) \tag{36}$$

if λ_i are the eigenvalues of ρ and $\eta(t) = -t\log t$. For the sake of simplicity the multiplicative constant κ will mostly be omitted.

It is worthwhile to note that if $S(\rho)$ is interpreted as the uncertainty carried by the statistical operator ρ, then (34) seems to be natural,

$$S(p\rho_1 + (1-p)\rho_2) = pS(\rho_1) + (1-p)S(\rho_2) + H(p, 1-p)\,, \tag{37}$$

holds for an orthogonal mixture and Shannon's classical information measure is involved. The **mixing property** (37) essentially determines the von Neumann entropy and tells us that the relation of orthogonal quantum states is classical. A detailed axiomatic characterization of the von Neumann entropy is Theorem 2.1 in [19].

Theorem 1.9 *Let ρ_1 and ρ_2 be density matrices and $0 < p < 1$. The following inequalities hold:*

$$p\,S(\rho_1) + (1-p)S(\rho_2) \le S(p\rho_1 + (1-p)\rho_2) \\ \le p\,S(\rho_1) + (1-p)S(\rho_2) + H(p, 1-p).$$

Proof. The first inequality is an immediate consequence of the concavity of the function $\eta(t) = -t \log t$. In order to obtain the second inequality we benefit from the formula

$$\operatorname{Tr} A\big(\log(A+B) - \log A\big) \\ = \int_0^\infty \operatorname{Tr} A(A+t)^{-1}B(A+B+t)^{-1}\,dt \ge 0 \qquad (A, B \ge 0)$$

and infer

$$\operatorname{Tr} p\rho_1\, \log(p\,\rho_1 + (1-p)\rho_2) \ge \operatorname{Tr} p\rho_1\, \log p\rho_1$$

and

$$\operatorname{Tr} (1-p)\rho_2\, \log(p\rho_1 + (1-p)\rho_2) \ge \operatorname{Tr} (1-p)\rho_2\, \log(1-p)\rho_2\,.$$

Adding the latter two inequalities we obtain the second inequality of the theorem. □

The von Neumann entropy is the trace of a continuous function of the density matrix, hence it is an obviously continuous functional on the states. However, a more precise estimate for the continuity will be required in approximations. Such an estimate is due to **Fannes**.

Theorem 1.10 *Let ρ_1 and ρ_2 be densities on a d-dimensional Hilbert space. If $\|\rho_1 - \rho_2\|_1 < \frac{1}{3}$, then the inequality*

$$|S(\rho_1) - S(\rho_2)| \le \|\rho_1 - \rho_2\|_1 \log d + \eta(\|\rho_1 - \rho_2\|_1)$$

holds. ($\|X\|_1 := Tr(X^*X)^{1/2}$).

The proof is found in [7] or [19]. Note that on an infinite dimensional Hilbert space the von Neumann entropy is not continuous (but it is such restricted to a set $\{\rho : S(\rho) \le c\}$).

Most properties of the von Neumann entropy will be deduced from the behavior of the relative entropy, see [19].

3.2 Fidelity

How close are two quantum states? There are many possible answers to this question. Restricting ourselves to pure states, we have to consider two unit vectors. $|\varphi\rangle$ and $|\psi\rangle$. Quantum mechanics has used the concept of transition probability $|\langle\varphi \mid \varphi\rangle|^2$ for a long time. This quantity is phase invariant, it lies between 0 and 1. It equals to 1 if and only if the two states coincide that is, $|\varphi\rangle$ equals to $|\psi\rangle$ up to a phase.

We call the square root of the transition probability **fidelity**: $F(|\varphi\rangle, |\psi\rangle) := |\langle\varphi \mid \psi\rangle|$. Shannon used a nonnegative distortion measure, and we may regard $1 - F(|\varphi\rangle, |\psi\rangle)$ as a distortion function on quantum states.

Under a quantum operation pure states could be transformed into mixed states, hence we need extension of the fidelity:

$$F(|\varphi\rangle\langle\varphi|, \rho) = \sqrt{\langle\varphi \mid \rho \mid \varphi\rangle}\,, \tag{38}$$

or in full generality

$$F(\rho_1, \rho_2) = \mathrm{Tr}\,\sqrt{\rho_1^{1/2}\rho_2\,\rho_1^{1/2}} \tag{39}$$

for positive matrices ρ_1 and ρ_2. This quantity was studied by Uhlmann in a different context [25] and he proved a variational formula:

Theorem 1.11

$$F(\rho_1, \rho_2) = \inf\left\{\sqrt{Tr(\rho_1 G)\ Tr\,(\rho_2 G^{-1})} : 0 \le G \ \textit{is invertible}\right\}$$

From Theorem 1.11 the symmetry of $F(\rho_1, \rho_2)$ is obvious and we can easily deduce the **monotonicity of the fidelity** under state transformation:

$$\begin{aligned} F(\mathcal{E}(\rho_1), \mathcal{E}(\rho_2))^2 &\ge \mathrm{Tr}\,\mathcal{E}(\rho_1)G\,\mathrm{Tr}\,\mathcal{E}(\rho_2)G^{-1} - \varepsilon \\ &\ge \mathrm{Tr}\,\rho_1\mathcal{E}^*(G)\,\mathrm{Tr}\,\rho_2\mathcal{E}^*(G^{-1}) - \varepsilon\,, \end{aligned}$$

where $\mathcal{E}^*$ is the adjoint of $\mathcal{E}$ with respect to the Hilbert-Schmidt inner product, $\varepsilon > 0$ is arbitrary and G is chosen to be appropriate. It is well-known that $\mathcal{E}^*$ is unital and positive, hence $\mathcal{E}^*(G)^{-1} \ge \mathcal{E}^*(G^{-1})$.

$$\begin{aligned} \mathrm{Tr}\,\rho_1\mathcal{E}^*(G)\,\mathrm{Tr}\,\rho_2\mathcal{E}^*(G^{-1}) &\ge \mathrm{Tr}\,\rho_1\mathcal{E}^*(G)\,\mathrm{Tr}\,\rho_2\mathcal{E}^*(G)^{-1} \\ &\ge F(\rho_1, \rho_2)^2\,. \end{aligned}$$

In this way the monotonicity is concluded:

Theorem 1.12 *For a state transformation $\mathcal{E}$ the inequality*

$$F(\mathcal{E}(\rho_1), \mathcal{E}(\rho_2)) \ge F(\rho_1, \rho_2)$$

holds.

Another remarkable operational formula is

$$\begin{aligned} F(\rho_1, \rho_2) = \max\{|\langle\psi_1|\psi_2\rangle| : \mathcal{E}(|\psi_1\rangle\langle\psi_1|) = \rho_1, \\ \mathcal{E}(|\psi_2\rangle\langle\psi_2|) = \rho_2 \text{ for some state transformation } \mathcal{E}\}. \end{aligned} \tag{40}$$

This variational expression reduces the understanding of the fidelity of arbitrary states to the case of pure states. The monotonicity property is implied by this formula easily.

Convergence in fidelity is equivalent with convergence in trace norm: $F(\rho_n, \rho_n') \to 1$ if and only if $\mathrm{Tr}\,|\rho_n - \rho_n'| \to 0$. This property of the fidelity is a consequence of the inequalities

$$1 - F(\rho_1, \rho_2) \le \frac{1}{2}\mathrm{Tr}\,|\rho_1 - \rho_2| \le \sqrt{1 - F(\rho_1, \rho_2)}\,. \tag{41}$$

References

1. P.M. Alberti and A. Uhlmann, *Stochasticity and partial order. Doubly stochastic maps and unitary mixing*, VEB Deutscher Verlag Wiss., Berlin, 1981.
2. R. Alicki and M. Fannes, Continuity of the quantum conditional information, J. Phys A: Math. Gen. **34**(2004), l55–L57.
3. J. Blank, P. Exner and M. Havliček, *Hilbert space operators in quantum physics*, American Institute of Physics, 1994.
4. R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1996.
5. O. Bratteli and D.W. Robinson, *Operator Algebras and Quantum Statistical Mechanics II*, Springer-Verlag, New York-Heidelberg-Berlin, 1981
6. J. L. Dodd and M. A. Nielsen, A simple operational interpretation of fidelity, arXiv e-print quant-ph/0111053
7. M. Fannes, A continuity property of the entropy density for spin lattice systems, Commun. Math. Phys. **31**(1973), 291–294.
8. F. Hansen and G.K. Pedersen, Jensen's inequality for operator and Löwner's theorem, Math. Anal. **258**(1982), 229–241
9. C.W. Helstrom *Quantum detection and estimation theory.* Academic Press, New York, 1976.
10. F. Hiai and D. Petz, The proper formula for relative entropy and its asymptotics in quantum probability, Commun. Math. Phys. **143**(1991), 99–114.
11. A.S. Holevo, *Probabilistic and statistical aspects of quantum theory*, North-Holland, Amsterdam, 1982.
12. A.S. Holevo, *Statistical structure of quantum theory*, Springer, 2001.
13. G. Lindblad, Completely positive maps and entropy inequalities, Commun. Math. Phys. **40**(1975), 147–151.
14. A.W. Marshall and I. Olkin, *Inequalities: Theory of majorization and its applications*, Academic Press, New York, 1979.
15. J. von Neumann, Thermodynamik quantummechanischer Gesamheiten, Gött. Nach. **1**(1927), 273-291.
16. J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin, 1932.
17. M. A. Nielsen and J. Kempe, Separable states are more disordered globally than locally, Phys. Rev. Lett., **86**, 5184-7 (2001).
18. M. A. Nielsen and D. Petz, A simple proof of the strong subadditivity, Quantum Inf. Comp., **5**(2005), 507–513,
19. M. Ohya and D. Petz, *Quantum Entropy and Its Use*, Springer, 1993
20. M. Ohya, D. Petz, N. Watanabe, On capacities of quantum channels, Prob. Math. Stat. **17**(1997), 179–196.
21. D. Petz, Quasi-entropies for finite quantum systems, Rep. Math. Phys. **21**(1986), 57–65.
22. D. Petz, *Algebra of the canonical commutation relation*, Leuven University Press, 1990.
23. D. Petz, Entropy, von Neumann and the von Neumann entropy, in *John von Neumann and the Foundations of Quantum Physics*, eds. M. Rédei and M. Stöltzner, Kluwer, 2001.
24. E. Schrödinger, Probability relations between separated systems, Proc. Cambridge Philos. Soc. **31**(1936), 446–452.
25. A. Uhlmann, The "transition probability" in the state space of a *-algebra, Rep. Mathematical Phys. **9**(1976), 273–279.

26. A. Uhlmann, Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory, Commun. Math. Phys. **54**(1977), 21–32.
27. V. Vedral, The role of relative entropy in quantum information theory, Rev. Modern Phys. **74**(2002), 197–234.