Quantum effects

Prof. Dr. Michael M. Wolf

January 14, 2014

Contents

Incomplete time-line of quantum theory

1900-'25:	Heuristic quantum theory (Planck, Bohr, Einstein, ...)
'25-'27:	Mathematical language for quantum theory (Heisenberg, Schrödinger, ...)
'28-'32:	Mathematical foundation of quantum theory (Dirac, von Neumann, ...)
'35:	Enter entanglement (Einstein, Podolsky, Rosen, Schrödinger, ...)
'70s:	Quantum theory from the "c.p. perspective" (Davies, Lindblad, Kraus, Choi, ...)
'90s:	Enter Quantum Information Theory

1 Mathematical Preliminaries

1.1 Hilbert space – home of quantum theory

Throughout $\mathcal{H}$ will be a separable complex Hilbert space. This implies that:

- there is a scalar product $\langle\cdot,\cdot\rangle\colon \mathcal{H}\times\mathcal{H}\to\mathbb{C}$ and an induced norm $\|\psi\| := \langle\psi,\psi\rangle^{\frac{1}{2}}$ w.r.t. which $\mathcal{H}$ is complete.
  We will use the convention $\langle\bar{c}\psi,\varphi\rangle = c\langle\psi,\varphi\rangle = \langle\psi,c\varphi\rangle\ \forall c\in\mathbb{C}$.
- there is a countable basis $\{e_i\}$ s.t. $\langle e_i,e_j\rangle = \delta_{ij}$ and for any $\psi\in\mathcal{H}$ we have $\psi = \sum_i \langle e_i,\psi\rangle e_i$.
- by the Riesz representation theorem $\mathcal{H}$ is *self-dual*, i.e. for any continuous linear functional $f\colon\mathcal{H}\to\mathbb{C}$ there is a $\varphi\in\mathcal{H}$ s.t. $\forall\psi\in\mathcal{H}\colon f(\psi)=\langle\varphi,\psi\rangle$.

Important examples:

(i) $l_2(\mathbb{N}) := \left\{\psi\in\mathbb{C}^{\mathbb{N}} \,\middle|\, \sum_i |\psi_i|^2 < \infty\right\}$ with the scalar product $\langle\psi,\varphi\rangle := \sum_i \overline{\psi_i}\varphi_i$.

(ii) $\mathbb{C}^n$ with the analogous scalar product

(iii) $\mathcal{L}_2(\mathbb{R}) := \left\{f\colon\mathbb{R}\to\mathbb{C} \,\middle|\, \int_{\mathbb{R}} |f(x)|^2\,\mathrm{d}x < \infty\right\} / \sim$ where
$f\sim g \Leftrightarrow \int_{\mathbb{R}} |f(x)-g(x)|^2\,\mathrm{d}x = 0$ and $\langle\psi,\varphi\rangle := \int_{\mathbb{R}} \overline{\psi}(x)\varphi(x)\,\mathrm{d}x$

All separable Hilbert spaces are isomorphic to (i) or (ii). In particular $\mathcal{L}_2(\mathbb{R}) \simeq l_2(\mathbb{N})$. That is, there is a bijection between the Hilbert spaces that preserves all scalar products.

Dirac notation

- elements of $\mathcal{H}$: $|\psi\rangle$ called *ket*
- elements of the dual space: $\langle\varphi|$ called *bra*

  $\to$ scalarproduct (*bra-ket*): $\langle\varphi|\psi\rangle$
- rank-one operator (*ket-bra*) $|\psi\rangle\langle\varphi|\colon\mathcal{H}\to\mathcal{H}$ defined by $|\Phi\rangle\mapsto|\psi\rangle\langle\varphi|\Phi\rangle$
- If $k\in\mathbb{N}$, then $|k\rangle$ denotes the k'th element of an orthonormal basis called *computational basis*. Unless otherwise defined $\psi_k := \langle k|\psi\rangle$.

1.2 Bounded operators

With *operator* we mean a linear map between two vector spaces. An operator *on* V is one that maps V into itself.
$\mathcal{B}(\mathcal{H}) :=$ space of continuous linear mappings $A\colon \mathcal{H} \to \mathcal{H}$.
This becomes a Banach space w.r.t. the *operator norm* $\|A\| := \sup_{\psi \neq 0} \frac{\|A\psi\|}{\|\psi\|}$.

- From the definition we easily verify that $\|AB\| \leq \|A\|\|B\|$.
- If A is represented by a matrix, then $\|A\|$ is its largest singular value.

Lemma 1. *If* $A\colon \mathcal{H} \to \mathcal{H}$ *is linear, then the properties*

(i) A *is* bounded, *i.e.* $\|A\| < \infty$

(ii) A *is continuous at 0*

(iii) A *is continuous*

are all equivalent.

Definition 2. Let $A \in \mathcal{B}(\mathcal{H})$.

- The *adjoint* $A^* \in \mathcal{B}(\mathcal{H})$ is defined via $\langle \psi, A\varphi \rangle =: \langle A^*\psi, \varphi \rangle \ \forall \varphi, \psi \in \mathcal{H}$.
- If $A^* = A$, then A is called *Hermitian*.
- If $A^*A = AA^* = \mathbb{1}$, then A is called *unitary*.
- If $\langle \psi, A\psi \rangle \geq 0 \ \forall \psi \in \mathcal{H}$, then A is called *positive* and we write $A \geq 0$.

Remarks.

- If $A = A^*$ is represented by a matrix with elements $A_{kl} := \langle k|A|l \rangle$, then
$$A_{kl} = \overline{\langle l|A^*|k \rangle} = \overline{\langle l|A|k \rangle} = \overline{A_{lk}}.$$
- $A \geq 0 \Rightarrow A^* = A$.
- In physics literature A^* is written $A^\dagger$ (read: *"A dagger"*).
- Positivity induces a partial order within the set of Hermitian operators. One writes $A \geq B$ meaning $A - B \geq 0$.

Example (Pauli matrices)**.**

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are all unitary and Hermitian.

Lemma 3. *Let $A \in \mathcal{B}(\mathcal{H})$. Then the following are equivalent:*

(i) $A \geq 0$

(ii) $A^* = A \wedge \operatorname{spec}(A) \in [0, \infty)$

(iii) $\exists B \in \mathcal{B}(\mathcal{H})\colon A = B^* B$

In (iii) we can choose $B \geq 0$. Then this uniquely defines $B =: \sqrt{A}$.

Recall. If defined, the *trace* is given by $\operatorname{tr}[A] := \sum_i \langle \psi_i | A | \psi_i \rangle$ where the sum runs over all elements of an orthonormal basis of $\mathcal{H}$.

Its main properties are that $\operatorname{tr}[A]$ is basis independent and that $\operatorname{tr}[AB] = \operatorname{tr}[BA]$.

Definition 4. Let $\mathcal{F} \subseteq \mathcal{B}(\mathcal{H})$ be the subspace of finite rank operators.

- The set of *trace class* operators $S_1 \subseteq \mathcal{B}(\mathcal{H})$ is the completion of $\mathcal{F}$ w.r.t. the *trace norm* $\|A\|_1 := \operatorname{tr}|A|$ where $|A| := \sqrt{A^* A}$.
- The set of *Hilbert-Schmidt* operators $S_2 \subseteq \mathcal{B}(\mathcal{H})$ is the completion of $\mathcal{F}$ w.r.t. the *Hilbert-Schmidt norm* $\|A\|_2 := \sqrt{\operatorname{tr}(A^* A)}$.
- The set of *compact operators* $\mathcal{C} \subseteq \mathcal{B}(\mathcal{H})$ is the completion of $\mathcal{F}$ w.r.t. the operator norm.

Remarks.

- $S_i = \{A \in \mathcal{B}(\mathcal{H}) \mid \|A\|_i < \infty\}$ for $i = 1, 2$.
- $S_1 \subseteq S_2 \subseteq \mathcal{C} \subseteq \mathcal{B}(\mathcal{H})$ and $\|A\|_1 \geq \|A\|_2 \geq \|A\|$.
- $A \in \mathcal{C}$ iff there is a sequence $(\lambda_i \geq 0)_{i=1}^{\dim(\mathcal{H})}$ s.t. $\lambda_i \to 0$ if $\dim(\mathcal{H}) = \infty$ and two orthonormal sets $\{\Psi_i \in \mathcal{H}\}$, $\{\Phi_i \in \mathcal{H}\}$ so that $A = \sum_i \lambda_i |\Psi_i\rangle\langle\Phi|$ (*singular value decomposition*).
- Similarly, if $A = A^* \in \mathcal{C}$, then there is a spectral decomposition $A = \sum_i a_i |\Psi_i\rangle\langle\Psi_i|$, where $a_i \in \mathbb{R}$ and $\{\Psi_i\}$ is an orthonormal basis.
- S_1 is a two-sided ideal in $\mathcal{B}(\mathcal{H})$, i.e. $A \in S_1,\ B \in \mathcal{B}(\mathcal{H}) \Rightarrow AB, BA \in S_1$. In fact, $\mathcal{B}(\mathcal{H})$ is isomorphic to the dual of S_1: for every continuous linear functional $F\colon S_1 \to \mathbb{C}$ there is a $B \in \mathcal{B}(\mathcal{H})$ s.t. $F(A) = \operatorname{tr}[B^* A]$.
- S_2 becomes a Hilbert space with the Hilbert-Schmidt inner product $\langle B, A \rangle := \operatorname{tr}[B^* A]$.

2 Probabilistic structure of quantum theory

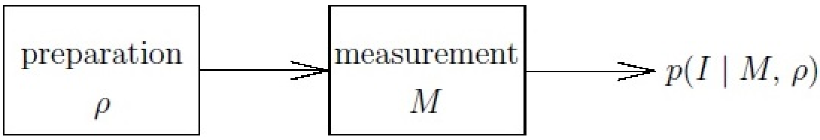
Quantum theory can be regarded as a general theoretical framework for physical theories. It consist out of a mathematical core which becomes a physical theory when adding a set of correspondence rules telling us which mathematical objects we have to use in different physical situations. In contrast to classical physical theories, these correspondence rules are not very intuitive as linear operators on Hilbert spaces are quite far from everyday life. It is truly remarkable that Heisenberg, Schrödinger, Dirac, Bohr, von Neumann together with all the other famous minds of this golden age come up with such a theory. In this chapter we will briefly review the mathematical formalism of quantum mechanics—abstracting from concrete physical realizations. All quantities will be dimensionless and $\hbar$ is set to one (although one should keep in mind that it is actually $10^{-34}Js$, in other words: really, really small).

It is often useful, and within quantum theory in fact necessary, to divide physical experiments into two parts: *preparation* and *measurement*. This innocent looking step already covers one of the basic differences between the quantum and the classical world, as in classical physics there is no need to talk about measurements in the first place. Note also that the division of a physical process into preparation and measurement is ambiguous, but, fortunately, quantum theoretical predictions do not depend on this choice. A genuine request is that a physical theory should predict the outcome of any measurement given all the information about the preparation, i.e., the initial conditions, of the system. Quantum mechanics teaches us that this is in general not possible and that all we can do is to predict the probabilities of outcomes in statistical experiments, i.e., long series of experiments where all relevant parameters in the procedure are kept unchanged. Thus, quantum mechanics does not predict individual events, unless the corresponding probability distribution happens to be tight. We will see later that there are good reasons to believe that this 'fuzziness' is not due to incompleteness of the theory and lacking knowledge about some *hidden variables* but rather part of natures character. In fact, *entanglement* will be the leading actor in that story. The fact that the appearance of probabilities is not only due to the ignorance of the observer, but at the very heart of the description, means that the measurement process can be regarded as a transition from possibilities to facts.

The *preparation* of a quantum system is the set of actions which determines all probability distributions of any possible measurement. It has to be a procedure which, when applied to a statistical ensemble, leads to converging relative frequencies and thus allows us to talk about probabilities. Since many different preparations can have the same effect in the sense that all the resulting probability distributions coincide it is reasonable to introduce the concept of a *state*, which specifies the effect of a preparation regardless of how it has actually been performed. Note that, in contrast to classical mechanics, a quantum 'state' does not refer to the attributes of an individual system but rather describes a statistical ensemble—the effect of a preparation in a statistical experiment. Although it is quite common, one should neither assign states to single events nor interpret them as elements of reality.

To summarize:

- Quantum theory is a probabilistic (as opposed to a deterministic) theory
- It predicts probabilities of outcomes in statistical experiments. It does not predict outcomes of individual measurements (unless the probabilities happen to be 0 or 1).
- In contrast to other probabilistic theories, where randomness is solely due to subjective ignorance, quantum theory also contains randomness that is in a sense objective.
- A statistical experiment is in quantum theory divided into preparation and measurement.
- In contrast to classical physical theories quantum theory does not describe any physical property before a measurement.



$p(I|M,\rho)$ = probability of measuring an outcome in I if M characterizes the measurement and ρ describes the preparation.
We will now see how S and M are described mathematically and how $p(I|M,\rho)$ is computed from them.

2.1 Preparation

Every preparation is described by a *density operator* (or *state*):

Definition 5.

- A positive trace-class operator $\rho \in S_1(\mathcal{H}),\ \rho \geq 0$ with unit trace $\mathrm{tr}[\rho] = 1$ is called *density operator.*
- A density operator is called *pure* if $\mathrm{rank}(\rho) = 1$ (i.e. $\rho = |\psi\rangle\langle\psi|$) and *mixed* otherwise.

Note that for a given Hilbert space the set of density operators is convex, i.e. with ρ_1, ρ_2 also $\rho = \lambda\rho_1 + (1-\lambda)\rho_2$ is a valid density operator if $\lambda \in [0,1]$.
The interpretation is: if a preparation described by ρ_i is used with probability λ_i, then the new probabilistic preparation is described by $\rho = \sum_i \lambda_i \rho_i$.
The spectral decomposition $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ is a convex decomposition into pure states. If ρ is mixed, there are infinitely many such decompositions.

Example. $\mathcal{H} = \mathbb{C}^2$. In this case $\rho = \frac{1}{2}(\mathbb{1} + \vec{v}\cdot\vec{\sigma})$, $\vec{v} \in \mathbb{R}^3$, $\|\vec{v}\| \leq 1$ is the *Bloch sphere* parametrization of the set of all density operators ($\rightarrow$ see exercise). Physically this might model:

(1) an atom in a double-well potential. $\rho = |0\rangle\langle 0|$ and $\rho = |1\rangle\langle 1|$ would then correspond to the atom being left or right, respectively.

(2) a two-level atom with $\rho = |0\rangle\langle 0|$, $\rho = |1\rangle\langle 1|$ referring to the ground and exited state, respectively.

(3) the spin of an electron with $\rho = |0\rangle\langle 0| \hat{=}$ spin up, $\rho = |1\rangle\langle 1| \hat{=}$ spin down.

(4) polarization degrees of freedom of light. North-/south pole correspond to left-/right circular polarization while the east-/west pole correspond to horizontal/vertical polarization. The center $\rho = \frac{\mathbb{1}}{2}$ then describes unpolarized light.

2.2 Measurement

Every measurement is described by a *positive operator valued measure.* Let S be a set whose elements label all possible measurement outcomes, and $\mathbb{B}(S)$ the Borel sets over S w.r.t. some topology.
There are basically two practically relevant examples:

(i) S is a discrete set and $\mathbb{B}(S) = \{I \subseteq S\}$ is the power set.

(ii) $S = \mathbb{R}^n$ and $\mathbb{B}(S)$ are the Borel sets w.r.t. the usual topology.

If not specified otherwise, these are the Borel sets we use.

Definition 6. A *positive operator valued measure* (POVM) is a mapping $M\colon \mathbb{B}(S) \rightarrow \mathcal{B}(\mathcal{H})$ s.t.

(i) $\forall I \in \mathbb{B}(S)\colon M(I) \geq 0$.

(ii) If $\{I_k \in \mathbb{B}(S)\}$ is a countable partition of S with pairwise disjoint I_k, then $\sum_k M(I_k) = \mathbb{1}$.

Lemma 7. *Let $M\colon \mathbb{B}(S) \to \mathcal{B}(\mathcal{H})$ be a POVM. Then for $I, J \in \mathbb{B}(S)$:*

(1) $I \subseteq J \Rightarrow M(I) + M(J\backslash I) = M(J) \wedge M(I) \leq M(J)$.

(2) $M(I \cup J) \leq M(I) + M(J)$ with equality if $I \cap J = \emptyset$.

Proof. (1) Using (ii) in definition 6 twice, we get:

$$\mathbb{1} = \begin{cases} M(J) + M(S\backslash J) \\ M(I) + M(J\backslash I) + M(S\backslash J). \end{cases}$$

So

$$M(J) - M(I) = M(J\backslash I) \geq 0.$$

(2) Using (1) we get

$$M(J \cup I) = M(J) + M((J \cup I)\backslash J) \leq M(J) + M(I).$$

□

Definition 8.

- Let $P \in \mathcal{B}(\mathcal{H})$. P is called a *projector* (and the mapping $P\colon \mathcal{H} \to \mathcal{H}$ a *projection*) if $P^2 = P$. If P is in addition Hermitian, we call P a *Hermitian projector* (or $P\colon \mathcal{H} \to \mathcal{H}$ an *orthogonal projection*).
- We call a POVM $M\colon \mathbb{B}(S) \to \mathcal{B}(\mathcal{H})$ *sharp* if $\forall I \in \mathbb{B}(S)\colon M(I)^2 = M(I)$.

Example (Sharp measurements via spectral decomposition)**.** Consider $A = A^*$ compact. Then there is a spectral decomposition $A = \sum_i a_i P_i$ where $a_i \in \mathbb{R}$ and the P_i's are spectral projectors, i.e. $P_i \geq 0$, $P_i P_j = P_i \delta_{ij}$, $\sum_i P_i = \mathbb{1}$.
If $S = \operatorname{spec}(A)$ then $M\colon \mathbb{B}(S) \to \mathcal{B}(\mathcal{H})$, $M(I) := \sum_{i\colon a_i \in I} P_i$ is a sharp POVM.
Evidently, by this spectral decomposition trick we can not only obtain some but all sharp POVMs with a discrete number of outcomes.
If $S \subseteq \mathbb{R}$ is not countable, then similar turns out to be true if we consider *selfadjoint* (and possibly unbounded) operators instead of bounded Hermitian operators:

Definition 9. Let $A\colon D(A) \subseteq \mathcal{H} \to \mathcal{H}$ be an operator defined on a subspace $D(A)$ (called the *domain of* A) that is dense in $\mathcal{H}$. Define
$D(A^*) := \{\varphi \in \mathcal{H} \mid \exists \Phi \in \mathcal{H}\ \forall \psi \in D(A)\colon \langle \varphi, A\psi \rangle = \langle \Phi, \psi \rangle\}$ and
$A^*\colon D(A^*) \to \mathcal{H}$ via $A^*\colon \varphi \mapsto \Phi$.

- A is called *Hermitian* if $D(A) \subseteq D(A^*)$ and $A = A^*$ on $D(A)$.
- A is called *selfadjoint* if $D(A) = D(A^*)$ and $A = A^*$ on $D(A)$.

If $A \in \mathcal{B}(\mathcal{H})$, then $D(A) = D(A^*) = \mathcal{H}$ and Hermitian = selfadjoint.
In the literature *symmetric* is sometimes used synonymous to Hermitian.
In analogy to the above example, every sharp POVM with $S \subseteq \mathbb{R}$ can be constructed from the spectral decomposition of a selfadjoint operator.

Example. Let $\mathcal{H} = \mathbb{C}^2$ and assume

- ρ describes a two-level atom. Then the spectral projectors of σ_3 might correspond to measuring the atom in ground and excited state, respectively.

- ρ describes polarization degrees of light. Then the spectral projections of σ_1 and σ_3 might correspond to measuring horizontal/vertical or left/right-circular polarization.

2.3 Probabilities

Postulate 10. *The probability $p(I|M, \rho)$ of measuring an outcome in $I \in \mathbb{B}(S)$ if measurement and preparation are described by a POVM $M\colon \mathbb{B}(S) \to \mathcal{B}(\mathcal{H})$ and a density operator $\rho \in \mathcal{B}(\mathcal{H})$ respectively, is given by*

$$p(I|M, \rho) = \mathrm{tr}[M(I)\rho] \tag{2.1}$$

$=\colon p(I)$ *for short.*

- p is a probability distribution, since
  - (i) $M(I) \geq 0 \wedge \rho \geq 0 \Rightarrow p(I) \geq 0$.
  - (ii) $p(S) = \mathrm{tr}[M(S)\rho] = \mathrm{tr}[S] = 1$.
  - (iii) For a countable number of disjoint subsets $I_k \in \mathbb{B}(S)$:

$$p\left(\bigcup_k I_k\right) = \mathrm{tr}\left[M\left(\bigcup_k I_k\right)\rho\right] = \sum_k \mathrm{tr}[M(I_k)\rho] = \sum_k p(I_k).$$

- Physics tells us in principle how M and S have to be chosen in specific situations.

- In practise we typically know M and S for some simple cases together with some mathematical rules (yet to be formalized in this lecture) telling us how to reduce more general cases to these simple ones. The largest part of quantum theory (Schrödinger equation, composite systems, etc.) is about those rules and their consequences.

- Standard quantum theory courses often assume ρ to be pure and M to be sharp. We will soon see in which sense this is justified.

2.4 Observables and expectation values

Definition 11. Let $M \colon \mathbb{B}(S) \to \mathcal{B}(M)$ be a POVM over a discrete set S and $\{m_i \in \mathbb{R}\}$ be bounded measurement outcomes assigned to each $i \in S$ occuring with probability $p(i) := \mathrm{tr}[\rho\, M(i)]$. Then

- $\langle M \rangle := \sum_{i \in S} m_i\, p(i)$ is the *expectation value*,
- $\mathrm{var}(M) := \sum_{i \in S} m_i^2\, p(i) - \langle M \rangle^2$ is the *variance*.

Remarks.

- These can be measured in a statistical experiment.
- We can write
$$\langle M \rangle = \mathrm{tr}[\rho\, \hat{M}] =: \langle \hat{M} \rangle \tag{2.2}$$
if we define $\hat{M} := \sum_i m_i\, M(i)$.
- If the set of measurement outcomes is not bounded and/or not discrete, one can still define an analogous selfadjoint operator $\hat{M}$ so that $\langle \hat{M} \rangle$ returns the expectation value of M in ρ.
- If M is sharp, then $(M, \{m_i\}) \mapsto \hat{M}$ is one-to-one by the spectral theorem. In this case $\{m_i\} = \mathrm{spec}(\hat{M})$, the selfadjoint operator $\hat{M}$ is called *observable* and we have
$$\mathrm{var}(M) = \langle \hat{M}^2 \rangle - \langle \hat{M} \rangle^2 =: \mathrm{var}(\hat{M}). \tag{2.3}$$

Examples.

- If we assign the values ± 1 to ground and exited state of a two-level atom, or to spin up/down of an electron, or to left-/right circular polarization, then the corresponding observable is σ_3.
- Position and momentum of a particle in one dimension correspond to observables given by unbounded operators acting as $Q \colon \psi(x) \mapsto x\, \psi(x)$ and $P \colon \psi(x) \mapsto -i\, \frac{\partial}{\partial x} \psi(x)$ defined on a dense subspace of $\mathcal{L}_2(\mathbb{R})$. Note that
$$[Q, P] := QP - PQ = i\mathbb{1}, \tag{2.4}$$
which is called *canonical commutation relation* (CCR).

3 Uncertanty relations and joint measurability

Theorem 12. *For all observables A, B and density operators with support in $D(AB) \cap D(BA)$ we have*

$$\operatorname{var}(A)\operatorname{var}(B) \geq \frac{1}{4}|\langle[A,B]\rangle|^2 \tag{3.1}$$

where $[A,B] := AB - BA$ is the commutator*.*

Proof. Define $\tilde{A} := A - \langle A\rangle\mathbb{1},\ \tilde{B} := B - \langle B\rangle\mathbb{1}$. Then $\left\langle\tilde{A}\right\rangle = \left\langle\tilde{B}\right\rangle = 0$ and

$$\begin{aligned}
\operatorname{var}\left(\tilde{A}\right) &= \left\langle\tilde{A}^2\right\rangle - \left\langle\tilde{A}\right\rangle^2 \\
&= \left\langle A^2 + \langle A\rangle^2\mathbb{1} - 2\langle A\rangle^2\mathbb{1}\right\rangle \\
&= \operatorname{var}(A), \\
\operatorname{var}\left(\tilde{B}\right) &= \operatorname{var}(B).
\end{aligned}$$

Moreover $\left[\tilde{A},\tilde{B}\right] = [A,B]$.
Assume that $\rho = |\psi\rangle\langle\psi|$ is pure. Then using Cauchy-Schwarz and selfadjointness we obtain

$$\begin{aligned}
\operatorname{var}(A)\operatorname{var}(B) &= \left\langle\psi,\tilde{A}^2\psi\right\rangle\ \left\langle\psi,\tilde{B}^2\psi\right\rangle \\
&= \left\|\tilde{A}\psi\right\|^2\ \left\|\tilde{B}\psi\right\|^2 \\
&\geq \left|\left\langle\tilde{A}\psi,\tilde{B}\psi\right\rangle\right|^2 \\
&= \left|\left\langle\psi,\tilde{A}\tilde{B}\psi\right\rangle\right|^2 \\
&= \left|\operatorname{Re}\left\langle\psi,\tilde{A}\tilde{B}\psi\right\rangle\right|^2 + \left|\operatorname{Im}\left\langle\psi,\tilde{A}\tilde{B}\psi\right\rangle\right|^2 \\
&= \frac{1}{4}\left|\left\langle\psi,(\tilde{A}\tilde{B}+\tilde{B}\tilde{A})\psi\right\rangle\right|^2 + \frac{1}{4}\left|\left\langle\psi,\left[\tilde{A},\tilde{B}\right]\psi\right\rangle\right|^2 \\
&\geq \frac{1}{4}\left|\langle\psi,[A,B]\psi\rangle\right|^2
\end{aligned}$$

This proves the assertion for pure states. That this generalizes to the mixed state case will become obvious in the next section. □

- (3.1) implies that one cannot prepare a state in which the two observables A and B both have arbitrary small variance if $\langle [A,B] \rangle \neq 0$.
  In particular, for position and momentum we obtain

$$\mathrm{var}(Q)^{\frac{1}{2}}\,\mathrm{var}(P)^{\frac{1}{2}} \geq \frac{1}{2} \tag{3.2}$$

  (in units of $\hbar \approx 10^{-34} Js$).

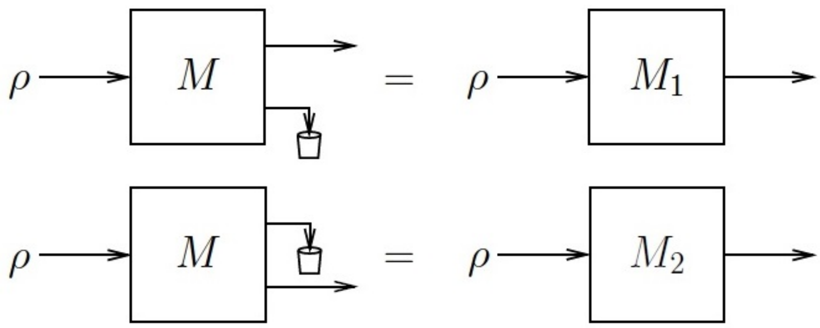
- (3.1) is called *uncertainty relation*. For observables corresponding to position and momentum, Heisenberg was the first to state such a relation in 1927 (albeit on mathematically and conceptionally shaky grounds).
  Theorem 12 goes back to Robertson, 1929.

- This uncertainty relation has nothing to do with 'disturbing one property by measuring the other'.

Definition 13. Two POVMs $M_i \colon \mathbb{B}(S_i) \to \mathcal{B}(\mathcal{H}),\ i = 1,2$ are called *jointly measurable*, if there is a POVM $M \colon \mathbb{B}(S_1 \times S_2) \to \mathcal{B}(\mathcal{H})$ such that

$$M(I \times S_2) = M_1(I)\ \forall I \in \mathbb{B}(S_1)$$
$$M(S_1 \times J) = M_2(J)\ \forall J \in \mathbb{B}(S_2).$$

This has an operational interpretation: there is a POVM that characterizes a measurement device with two output ports, such that if we ignore one of them, the statistics are indistinguishable from the ones obtained by measuring either M_1 or M_2, depending on which port we ignore. Since this holds irrespective of the state ρ, one can say that M measures M_1 and M_2 jointly.



In order to show what joint measurability has to do with commutativity, we need the following:

Lemma 14. *Let $A, B, C \in \mathcal{B}(\mathcal{H})$ be such that $0 \leq A \leq B$. Then*

$$BC = 0 \Rightarrow AC = 0.$$

Proof. Clearly $BC = 0 \Rightarrow C^*BC = 0$. However,

$$0 = C^*BC \geq C^*AC \geq 0,$$

so

$$0 = C^*AC = \left(\sqrt{A}C\right)^* \left(\sqrt{A}C\right),$$

which implies $\sqrt{A}C = 0$ and thus $AC = 0$. □

Theorem 15. *Two sharp POVMs $M_i\colon \mathbb{B}(S_i) \to \mathcal{B}(\mathcal{H}),\ i = 1, 2$ are jointly measurable iff*

$$[M_1(I_1), M_2(I_2)] = 0 \ \forall I_i \in \mathbb{B}(S_i).$$

Proof.

"⇐": Extending $M(I_1 \times I_2) := M_1(I_1)\, M_2(I_2)$ from rectangles to $\mathbb{B}(S_1 \times S_2)$ leads to the sought joint POVM.

"⇒":

$$\begin{aligned} M_1(I_1)\, M_2(I_2) &= M(I_1 \times S_2)\, M(S_1 \times I_2) \\ &= \left(M(I_1 \times I_2) + M\left(I_1 \times \bar{I}_2\right)\right) \left(M(I_1 \times I_2) + M\left(\bar{I}_1 \times I_2\right)\right) \\ &= M(I_1 \times I_2)^2 \end{aligned}$$

The three remaining terms, which are expected in the last line, vanish because by Lemma 14 for instance

$$M_1(I_1)\, M_1\left(\bar{I}_1\right) = 0 \Rightarrow M(I_1 \times I_2)\, M\left(\bar{I}_1 \times I_2\right) = 0$$

and similar for the other two.
Repeating the argument with interchanged elements finally yields

$$M_2(I_2)\, M_1(I_1) = M(I_1 \times I_2)^2 = M_1(I_1)\, M(I_2).$$

□

If the sharp POVMs are characterized by observables $\hat{M}_i \in \mathcal{B}(\mathcal{H})$, then the above condition is equivalent to

$$[\hat{M}_1, \hat{M}_2] = 0.$$

Remarks.

- For $\mathcal{H} \in \mathbb{C}^2$: different spin directions (or polarization types) cannot be jointly measured since $[\sigma_k, \sigma_l] = 2i\, \varepsilon_{klm}\, \sigma_m$.
- Position and momentum cannot be jointly measured since $[Q, P] = i\mathbb{1}$.
- Approximate joint measurements are still possible (and since $\hbar$, which we set equal to one, is very small on typical scales, these approximations may be quite accurate).

4 Composite systems

There are two elementary ways of how to combine two Hilbert spaces to form a new one: the direct sum and the tensor product.

Definition 16. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be complex Hilbert spaces. Their *direct sum* is the Hilbert space

$$\mathcal{H}_1 \oplus \mathcal{H}_2 := \{(\psi, \varphi) \in \mathcal{H}_1 \times \mathcal{H}_2\}$$

with inner product

$$\langle (\psi_1, \varphi_1), (\psi_2, \varphi_2) \rangle := \langle \psi_1, \psi_2 \rangle + \langle \varphi_1, \varphi_2 \rangle.$$

Instead of (ψ, φ) we also write $\psi \oplus \varphi$ for the elements of $\mathcal{H}_1 \oplus \mathcal{H}_2$.

Remarks.

- $\dim(\mathcal{H}_1 \oplus \mathcal{H}_2) = \dim \mathcal{H}_1 + \dim \mathcal{H}_2$
- If $\mathcal{H}_i = \mathbb{C}^{d_i}$, then $\mathcal{H}_1 \oplus \mathcal{H}_2 = \mathbb{C}^{d_1+d_2}$ and $\psi \oplus \varphi$ can be expressed as $(\psi_1, \ldots, \psi_{d_1}, \varphi_1, \ldots, \varphi_{d_2})$.
- The construction is associative: $\mathcal{H}_1 \oplus (\mathcal{H}_2 \oplus \mathcal{H}_3) = (\mathcal{H}_1 \oplus \mathcal{H}_2) \oplus \mathcal{H}_3 =: \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$.

Definition 17. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be complex Hilbert spaces, $\varphi_i \in \mathcal{H}_i$ and let $\varphi_1 \otimes \varphi_2 \colon \mathcal{H}_1 \times \mathcal{H}_2 \to \mathbb{C}$ denote the conjugate bilinear form defined by

$$\varphi_1 \otimes \varphi_2(\psi_1, \psi_2) := \langle \psi_1, \varphi_1 \rangle \langle \psi_2, \varphi_2 \rangle.$$

Let $\tilde{\mathcal{H}}$ be the set of all finite linear combinations of such conjugate bilinear forms. We define an inner product on $\tilde{\mathcal{H}}$ via

$$\langle \varphi_1 \otimes \varphi_2, \psi_1 \otimes \psi_2 \rangle := \langle \varphi_1, \psi_1 \rangle \langle \varphi_2, \psi_2 \rangle \tag{4.1}$$

and extend it by linearity to all of $\tilde{\mathcal{H}}$.
The *tensor product* $\mathcal{H}_1 \otimes \mathcal{H}_2$ is the Hilbert space obtained from the completion of $\tilde{\mathcal{H}}$ under this inner product.

Lemma 18. *If $\{\varphi_k\}$ and $\{\psi_l\}$ are orthonormal bases for $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then $\{\varphi_k \otimes \psi_l\}$ is an orthonormal basis of $\mathcal{H}_1 \otimes \mathcal{H}_2$.*

Remarks.

- $\dim(\mathcal{H}_1 \otimes \mathcal{H}_2) = \dim(\mathcal{H}_1) \dim(\mathcal{H}_2)$.

- If $\varphi \in \mathbb{C}^{d_1}$, $\psi \in \mathbb{C}^{d_2}$ then $\varphi \otimes \psi$ can be identified with the vector $(\varphi_1\psi, \ldots, \varphi_{d_1}\psi) \in \mathbb{C}^{d_1 d_2}$.
- The tensor product is associative and distributive over the direct sum, i.e.,

$$(\mathcal{H}_1 \oplus \mathcal{H}_2) \otimes \mathcal{H}_3 = (\mathcal{H}_1 \otimes \mathcal{H}_3) \oplus (\mathcal{H}_2 \otimes \mathcal{H}_3).$$

- Note that

$$(\alpha\varphi) \otimes \psi = \varphi \otimes (\alpha\psi) = \alpha(\varphi \otimes \psi) \text{ for } \alpha \in \mathbb{C}$$

  and

$$(\varphi_1 + \varphi_2) \otimes \psi = \varphi_1 \otimes \psi + \varphi_2 \otimes \psi.$$

Examples (from quantum physics)**.**

- The Hilbert space assigned to an electron is $\mathcal{L}_2(\mathbb{R}^3) \otimes \mathbb{C}^2$ where $\mathcal{L}_2(\mathbb{R}^3)$ and $\mathbb{C}^2$ correspond to the spatial and spin degrees of freedom, respectively.
- A neutron can decay into a proton, an electron and an electron anti-neutrino. The total Hilbert space assigned to this system is

$$\mathcal{H}_n \oplus \underbrace{\mathcal{H}_e \otimes \mathcal{H}_p \otimes \mathcal{H}_{\bar{\nu}_e}}_{\text{Hilbert spaces of the decay products}}.$$

- In general, if two systems at different locations are described by individual Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, then the composite system is described by $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Lemma 19. *Let $S_2(\mathcal{H}_1, \mathcal{H}_2)$ be the Hilbert space of all Hilbert-Schmidt class operators from $\mathcal{H}_1$ to $\mathcal{H}_2$ and let $\{e_k \otimes f_l\}$ be an orthonormal basis for $\mathcal{H}_1 \otimes \mathcal{H}_2$. Then by linearity the map*

$$|e_k\rangle \otimes |f_l\rangle \mapsto |f_l\rangle\langle e_k|$$

extends to a linear Hilbert space isomorphism between $\mathcal{H}_1 \otimes \mathcal{H}_2$ and $S_2(\mathcal{H}_1, \mathcal{H}_2)$.

Note that this means there is a linear bijection $\mathcal{H}_1 \otimes \mathcal{H}_2 \ni \Phi \mapsto M \in S_2(\mathcal{H}_1, \mathcal{H}_2)$ s.t. $\langle \Phi, \Phi \rangle^{\frac{1}{2}} = \|\Phi\| = \|M\| = \operatorname{tr}[M^* M]^{\frac{1}{2}}$.

Corollary 20 (Schmidt decomposition)**.** *For any $\Phi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ there exist orthonormal sets $\{e_i \in \mathcal{H}_1\}$ and $\{f_i \in \mathcal{H}_2\}$ and non-negative numbers $\lambda_i \geq 0$ s.t.*

$$\Phi = \sum_{i=1}^{D} \sqrt{\lambda_i}\, e_i \otimes f_i \tag{4.2}$$

where $D = \min\{\dim(\mathcal{H}_1), \dim(\mathcal{H}_2)\} \in \mathbb{N} \cup \{\infty\}$ and $\sum_{i=1}^{D} \lambda_i = \|\Phi\|^2$.
The λ_i's are uniquely determined by Φ.

Proof. We use the above isomorphism $\Phi \mapsto M \in S_2(\mathcal{H}_1, \mathcal{H}_2)$ and exploit that M is compact so that it admits a singular value decomposition $M = \sum_i \sqrt{\lambda_i}\, |f_i\rangle\langle e_i|$ where $\|M\|^2 = \mathrm{tr}[M^* M] = \sum_i \lambda_i$. Using the isomorphism (which is a linear map) backwards then leads to the sought result. □

Definition 21. Let $A_i \in \mathcal{B}(\mathcal{H}_i)$. Define

- $A_1 \oplus A_2 \colon \mathcal{H}_1 \oplus \mathcal{H}_2 \to \mathcal{H}_1 \oplus \mathcal{H}_2$ by $A_1 \oplus A_2(\varphi_1 \oplus \varphi_2) = (A_1\varphi_1) \oplus (A_2\varphi_2)$ and
- $A_1 \otimes A_2 \colon \mathcal{H}_1 \otimes \mathcal{H}_2 \to \mathcal{H}_1 \otimes \mathcal{H}_2$ by $A_1 \otimes A_2(\varphi_1 \otimes \varphi_2) = (A_1\varphi_1) \otimes (A_2\varphi_2)$

for all $\varphi_i \in \mathcal{H}_i$ and extended by linearity.

These constructions easily extend to unbounded operators, when we take domains into account.
If $A_i \in \mathcal{B}(\mathcal{H}_i)$ then $\|A_1 \oplus A_2\| = \max\{\|A_1\|, \|A_2\|\}$ and $\|A_1 \otimes A_2\| = \|A_1\| \cdot \|A_2\|$.
Consequently, $A_1 \oplus A_2$ and $A_1 \otimes A_2$ are bounded as well.
Moreover, if $A_i \geq 0$, then $A_1 \oplus A_2 \geq 0$ and $A_1 \otimes A_2 \geq 0$.

Example. If $\mathcal{H}_i = \mathbb{C}^d$, then $A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ is block diagonal and $A \otimes B$ corresponds to a matrix with entries $\begin{pmatrix} A_{11}B & \cdots & A_{1d}B \\ \vdots & \ddots & \vdots \\ A_{d1}B & \cdots & A_{dd}B \end{pmatrix} \in \mathbb{C}^{d^2 \times d^2}$.
If A and B are Hermitian matrices, then it is easy to verify that
$\mathrm{spec}(A \oplus B) = \mathrm{spec}(A) \cup \mathrm{spec}(B)$ and $\mathrm{spec}(A \otimes B) = \{\lambda\mu \mid \lambda \in \mathrm{spec}(A),\ \mu \in \mathrm{spec}(B)\}$.

Definition 22.

- The linear map $\mathrm{tr}_B \colon S_1(\mathcal{H}_A \otimes \mathcal{H}_B) \to S_1(\mathcal{H}_A),\ \rho \mapsto \rho_A$ defined via

$$\mathrm{tr}[\rho(X \otimes \mathbb{1})] =: \mathrm{tr}[\rho_A X] \ \forall X \in \mathcal{B}(\mathcal{H}_A) \tag{4.3}$$

  is called *partial trace*, and if ρ is a density operator, ρ_A is called the *reduced density operator* of ρ w.r.t. $\mathcal{H}_A$.

- A state ρ is called *product state* if there are ρ_i s.t. $\rho = \rho_1 \otimes \rho_2$.

The purpose of the partial trace in quantum theory is to obtain the description of a subsystem from the description of a larger composite system by discarding the other parts. If $\rho \in S_1(\mathcal{H}_A \otimes \mathcal{H}_B)$ is a density operator, then so is $\rho_A := \mathrm{tr}_B[\rho]$.

Corollary 23. *For $\rho \in S_1(\mathcal{H}_A \otimes \mathcal{H}_B)$ we have*

(i) $\mathrm{tr}[\rho_A] = \mathrm{tr}[\rho]$

(ii) $\rho \geq 0 \Rightarrow \rho_A \geq 0.$

Proof.

(i) $\mathrm{tr}[\rho_A] = \mathrm{tr}[\rho_A \mathbb{1}] = \mathrm{tr}[\rho(\mathbb{1} \otimes \mathbb{1})] = \mathrm{tr}[\rho]$.

(ii) $\langle\psi|\rho_A|\psi\rangle = \mathrm{tr}[\rho_A|\psi\rangle\langle\psi|] = \mathrm{tr}[\rho(\underbrace{|\psi\rangle\langle\psi| \otimes \mathbb{1}}_{\geq 0})] \geq 0$.

□

Measurements on subsystems

Let $M_A \colon \mathbb{B}(S) \to \mathcal{B}(\mathcal{H}_A)$ be a POVM that describes a measurement on a system that corresponds to a Hilbert space $\mathcal{H}_A$. Suppose this is a subsystem of a larger composite system to which $\mathcal{H}_A \otimes \mathcal{H}_B$ is assigned to. Then the measurement described as one on the subsystem corresponds to the POVM $M \colon \mathbb{B}(S) \to \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$

$$M(I) := M_A(I) \otimes \mathbb{1} \ \forall I \in \mathbb{B}(S).$$

Note that this is consistent with the definition of the reduced density operator ρ_A since

$$\mathrm{tr}[\rho_A M_A(I)] = \mathrm{tr}[\rho M_A(I) \otimes \mathbb{1}] = \mathrm{tr}[\rho M(I)].$$

Proposition 24. *Let $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a density operator. Then*

(i) if $\{e_k \in \mathcal{H}_A\}$ and $\{f_l \in \mathcal{H}_B\}$ are orthonormal bases, then

$$\langle e_i, \rho_A\, e_j\rangle = \sum_l \langle e_i \otimes f_l, \rho\, e_j \otimes f_l\rangle.$$

(ii) if $\rho = \rho_A \otimes \rho_B$, then ρ_A is the reduced density operator of ρ w.r.t. $\mathcal{H}_A$.

(iii) if $\rho = |\Phi\rangle\langle\Phi|$ and $\Phi = \sum_i \sqrt{\lambda_i}\, e_i \otimes f_i$ is a Schmidt decomposition, then $\rho_A = \sum_i \lambda_i\, |e_i\rangle\langle e_i|$ and $\rho_B = \sum_i \lambda_i\, |f_i\rangle\langle f_i|$ are the reduced density operators w.r.t. $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively.

Proof. $\to$ exercise. □

Note that in particular $\mathrm{spec}(\rho_A)\backslash\{0\} = \mathrm{spec}(\rho_B)\backslash\{0\}$ if $\rho = |\Phi\rangle\langle\Phi|$.

Theorem 25 (purification)**.** *If $\rho \in \mathcal{B}(\mathcal{H}_A)$ is a density operator, then there is a normalized $\Phi \in \mathcal{H}_A \otimes \mathcal{H}_B$ with $\mathcal{H}_B \simeq (\ker \rho)^\perp$ s.t. ρ is the reduced density operator of $|\Phi\rangle\langle\Phi|$ w.r.t. $\mathcal{H}_A$.*

Proof. From the spectral decomposition $\rho = \sum_i \lambda_i\, |e_i\rangle\langle e_i|$ we can construct $\Phi := \sum_{k\colon \lambda_k > 0} \sqrt{\lambda_k}\, e_k \otimes e_k$ which then has the claimed property. □

Remarks.

- $\Phi \in \mathcal{H}_A \otimes \mathcal{H}_B$ and the corresponding density operator $|\Phi\rangle\langle\Phi|$ are called *purification* of $\rho \in S_1(\mathcal{H}_A)$ if $\mathrm{tr}_B\, |\Phi\rangle\langle\Phi| = \rho$.

- If $\Phi_i \in \mathcal{H}_A \otimes \mathcal{H}_B^{(i)},\ i = 1, 2$ are purifications of ρ with $\mathcal{H}_B^{(1)} = (\ker \rho)^\perp$, then there is an isometry $V\colon \mathcal{H}_B^{(1)} \to \mathcal{H}_B^{(2)}$ so that

  $$|\Phi_2\rangle = (\mathbb{1} \otimes V)|\Phi_1\rangle.$$

  Conversely, every such isometry leads to a purification.

With this we can complete the proof of the uncertainty theorem (Theorem 12) and reduce the mixed state case to the pure state case:
Instead of ρ with observables A and B, consider a purification Φ of ρ and observables $\hat{A} := A \otimes \mathbb{1},\ \hat{B} := B \otimes \mathbb{1}$. Using $\mathrm{tr}[\rho\, XY] = \mathrm{tr}\left[|\Phi\rangle\langle\Phi|\, \hat{X}\hat{Y}\right]$ e.g. for $X, Y \in \{A, B, \mathbb{1}\}$ we can then express all the mixed state terms by the corresponding pure state terms for which the theorem has been proven already.

5 Correlations and entanglement

Lemma 26. *Let $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a density operator. Then $\rho = \rho_A \otimes \rho_B$ holds iff $\langle A \otimes B \rangle = \langle A \rangle \langle B \rangle := \langle A \otimes \mathbb{1} \rangle \langle \mathbb{1} \otimes B \rangle$ for all $A \in \mathcal{B}(\mathcal{H}_A), B \in \mathcal{B}(\mathcal{H}_B)$.*

Proof.

"$\Rightarrow$": Follows from the definition of the reduced density operators.

$$\begin{aligned}
\langle A \otimes B \rangle &= \operatorname{tr}[\rho_A \otimes \rho_B \, A \otimes B] \\
&= \operatorname{tr}[\rho_A \, A] \operatorname{tr}[\rho_B \, B] \\
&= \operatorname{tr}[\rho \, A \otimes \mathbb{1}] \operatorname{tr}[\rho \, \mathbb{1} \otimes B] \\
&= \langle A \rangle \langle B \rangle.
\end{aligned}$$

"$\Leftarrow$": For the converse note that $\{|i\rangle\langle j| \otimes |k\rangle\langle l|\}$ is an orthonormal basis for $S_2(\mathcal{H}_A \otimes \mathcal{H}_B)$. So if $\langle A \otimes B \rangle - \langle A \rangle \langle B \rangle = \operatorname{tr}[(\rho - \rho_A \otimes \rho_B)(A \otimes B)] = 0$ for all $A \otimes B = |i\rangle\langle j| \otimes |k\rangle\langle l|$, then $\rho = \rho_A \otimes \rho_B$.

□

If $\rho = \rho_A \otimes \rho_B$ is a product state, then the two subsystems are said to be *uncorrelated*.

Definition 27. A state $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is called *classically correlated* or *separable* if it can be approximated in trace norm by a convex combination of product states. That is, for any $\varepsilon > 0$ there are $\{\lambda_i \geq 0\}_{i=1}^N$ and product states $\rho_1^{(i)} \otimes \rho_2^{(i)} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ s.t.

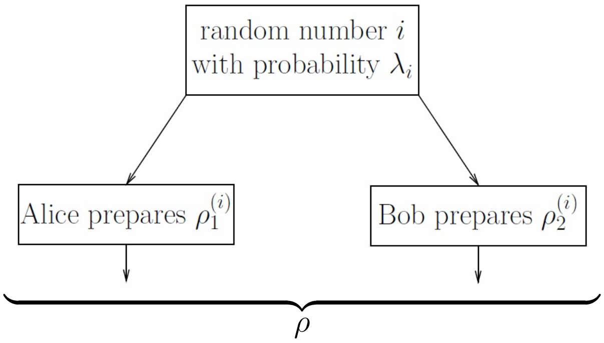
$$\left\| \rho - \sum_{i=1}^{N} \lambda_i \, \rho_1^{(i)} \otimes \rho_2^{(i)} \right\|_1 < \varepsilon.$$

If such an approximation is not possible, then the state is called *entangled*.

Remarks.

- The set of separable states is convex.
- Via the spectral decomposition we could w.l.o.g. restrict to pure product states in the definition.
- For finite dimensional Hilbert spaces any separable state admits an exact representation of this form (i.e., $\varepsilon = 0$) and (by Caratheodory's theorem) $N \leq (\dim(\mathcal{H}_A) \dim(\mathcal{H}_B))^2$ pure states are sufficient.

The interpretation of this definition is that $\rho = \sum_i \lambda_i \, \rho_1^{(i)} \otimes \rho_2^{(i)}$ describes the following preparation:



This allows correlations, which are, however, built up "classically" in the sense that they are based on the generation and distribution of a classical random number.

Corollary 28. *Let $\rho = |\Phi\rangle\langle\Phi| \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a pure density operator. Then the following are equivalent:*

(i) *Φ has Schmidt decomposition $\Phi = e \otimes f$.*

(ii) *ρ is a product state.*

(iii) *ρ is separable.*

Proof. (for finite dimension)

(iii)⇒(i) $|\Phi\rangle\langle\Phi| = \sum_i \lambda_i \, |\varphi_i\rangle\langle\varphi_i| \otimes |\psi_i\rangle\langle\psi_i|$ with $\|\varphi_i\| = \|\psi_i\| = 1$ and $\lambda_i > 0$.

Multiplying from both sides with $|\Phi\rangle\langle\Phi|$ yields:

$$|\Phi\rangle\langle\Phi| = \sum_i \lambda_i \, |\Phi\rangle\langle\Phi| \, \mu_i, \;\; \mu_i := |\langle\Phi, \varphi_i \otimes \psi_i\rangle|^2.$$

Since $\mu_i \in [0,1]$ and $\sum_i \lambda_i = 1$, we have to have $\lambda_i > 0 \Rightarrow \mu_i = 1$, which implies

$$|\Phi\rangle\langle\Phi| = |\varphi_i\rangle\langle\varphi_i| \otimes |\psi_i\rangle\langle\psi_i|.$$

(i)⇒(ii)⇒(iii) should be obvious.

□

6 Local hidden variable (LHV) theories and Bell inequalities

Aim. Description in terms of classical probability theory + physically reasonable assumptions.

Assumptions.

(i) One can assign definite values to quantities even before they are measured. These values may, however, be 'hidden' or unknown to the observer so that a probabilistic description may be necessary.

(ii) The properties of one subsystem should not immediately depend on what happens to a very distant other subsystem.

Example scenario. A source emits pairs of particles to distant observers (Alice and Bob) who perform ± 1 valued measurements each.

Assume that both observers choose one out of $m \in \mathbb{N}$ measurement devices denoted by A_x, B_x, with $x, y \in \{1, \dots, m\}$.
Let $p(a, b \mid x, y)$ be the probability that Alice obtains outcome $a \in \{\pm 1\}$ while Bob obtains outcome $b \in \{\pm 1\}$ in the same run of the experiment, if they have used devices A_x and B_y, respectively.
Consider the expectation value of their product, i.e., $\langle A_x B_y \rangle := \sum_{a,b \in \{\pm 1\}} ab\, p(a, b \mid x, y)$.

LHV Ansatz.

$$\langle A_x B_y \rangle = \int_\Omega A_x(\omega) B_y(\omega)\, dP(\omega) \tag{6.1}$$

with random variables $A_x, B_y \colon \Omega \to \{\pm 1\}$ and a probability measure P.
$\omega \in \Omega$ is the *hidden variable* and *locality* is expressed in the fact that A_x does not depend on y and B_y does not depend on x.

Bell inequalities.

Definition 29. Consider an $m \times m$ tuple $C \in \mathbb{R}^{m \times m}$ of empirically obtained expectation values $\langle A_x B_y \rangle =: C_{xy}$ for pairs of ± 1 valued measurements. Let $\mathcal{C} \subseteq \mathbb{R}^{m \times m}$ be the set of all such C's for which there exists an LHV description as in (6.1).
An inequality for C is called *Bell inequality* if it holds for all $C \in \mathcal{C}$.

Remarks.

- $\mathcal{C}$ is a closed convex polytope, so for a complete description of $\mathcal{C}$ a finite set of linear inequalities suffices.
- There are trivial Bell inequalities of the form $|\langle A_x B_y \rangle| \leq 1$.
- *Stochastic LHV theories* allow the random variables in (6.1) to have ranges in $[-1, 1]$ instead of $\{-1, +1\}$. It turns out, however, that this does not change $\mathcal{C}$ ($\rightarrow$ exercise).

Corollary 30. *Let $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be classically correlated and $M_x \colon \mathbb{B}(\{\pm 1\}) \rightarrow \mathcal{B}(\mathcal{H}_A)$ and similarly $M'_y \colon \mathbb{B}(\{\pm 1\}) \rightarrow \mathcal{B}(\mathcal{H}_B)$ be POVMs for $x, y \in \{1, \ldots, m\}$.*
Then $C \in \mathcal{C}$ holds for

$$C_{xy} := \sum_{a,b \in \{\pm 1\}} ab \operatorname{tr}\left[\rho\, M_x(a) \otimes M'_y(b)\right].$$

Proof. Define $A_x := \sum_{a \in \{\pm 1\}} a\, M_x(a) = M_x(1) - M_x(-1)$ and $B_y := \sum_{b \in \{\pm 1\}} b\, M'_y(b)$.
Then $-\mathbb{1} \leq A_x, B_y \leq \mathbb{1}$.
Suppose $\rho = \sum_\omega p_\omega\, \rho_A^{(\omega)} \otimes \rho_B^{(\omega)}$ is a convex combination of product states.
Then $C_{xy} = \operatorname{tr}[\rho\, A_x \otimes B_y] = \sum_\omega p_\omega \underbrace{\operatorname{tr}\left[\rho_A^{(\omega)} A_x\right]}_{=:A_y(\omega)} \underbrace{\operatorname{tr}\left[\rho_B^{(\omega)} B_y\right]}_{=:B_y(\omega)}$ is a stochastic LHV description with discrete probability space. $\square$

Consequently, unentangled quantum states can never violate any Bell inequality.
The simplest and most famous non-trivial Bell inequality is the Causer-Horne-Shimony-Holt (CHSH) inequality:

Theorem 31 (CHSH)**.** *Every LHV theory for the description of ± 1 valued measurements satisfies:*

$$\left|\langle A_1 B_1 \rangle + \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle - \langle A_2 B_2 \rangle\right| \leq 2. \tag{6.2}$$

Proof. L.h.s. $= \left|\int_\Omega A_1(\omega)\big(B_1(\omega) + B_2(\omega)\big) + A_2(\omega)\big(B_1(\omega) + B_2(\omega)\big)\, dP(\omega)\right|$.
For a fixed $\omega \in \Omega$ we can distinguish two cases:

1. $B_1(\omega) = B_2(\omega)$. Then $B_1(\omega) - B_2(\omega) = 0$ and $B_1(\omega) + B_2(\omega) \in \{\pm 2\}$.
2. $B_1(\omega) \neq B_2(\omega)$. Then $B_1(\omega) + B_2(\omega) = 0$ and $B_1(\omega) - B_2(\omega) \in \{\pm 2\}$.

In either case the integrand is in $\{-2, 2\}$ so that the average is in $[-2, 2]$. $\square$

The remarkable thing is, that this can be violated within quantum theory. For the quantum description, we assign POVMs M_x and M'_y to the measurement devices and introduce again $A_x := M_x(1) - M_x(-1), B_y := M'_y(1) - M'_y(-1)$.
With $\hat{\beta} := A_1 \otimes (B_1 + B_2) + A_2 \otimes (B_1 - B_2) \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ the l.h.s. of (6.2) is then given by $\left|\operatorname{tr}\left[\rho\hat{\beta}\right]\right|$.

Theorem 32 (CHSH violation & Cirelson bound)**.** *If $\hat{\beta}$ is as above, then for all density operators $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$:*

$$\left|\mathrm{tr}\left[\rho\hat{\beta}\right]\right| \leq 2\sqrt{2}.$$

That is, quantum theory violates the CHSH inequality at most by a factor $\sqrt{2}$.
Moreover, there exists a pure state on $\mathbb{C}^2 \otimes \mathbb{C}^2$ and observables $A_x, B_y \in \mathcal{B}(\mathbb{C}^2)$ with eigenvalues ± 1, s.t. equality holds in this equation.

Proof. Consider the map $A_1 \mapsto \mathrm{tr}\left[\rho\hat{\beta}\right]$. Since this is an affine functional the supremum and infimum over the convex set $-\mathbb{1} \leq A_1 \leq \mathbb{1}$ is attained for some extreme point for which $\mathrm{spec}(A_1) = \{\pm 1\}$ and thus $A_1^2 = \mathbb{1}$.
By the same reasoning we can assume that A_1, A_2, B_1, B_2 are all such that their square is $\mathbb{1}$. Using this property, direct computation leads to

$$\hat{\beta}^2 = 4\,\mathbb{1} \otimes \mathbb{1} + [A_2, A_1] \otimes [B_1, B_2].$$

We exploit this via positivity of the variance and obtain

$$\begin{aligned} \mathrm{tr}\left[\rho\hat{\beta}\right]^2 &\leq \mathrm{tr}\left[\rho\hat{\beta}^2\right] = 4 + \mathrm{tr}\left[\rho\,[A_2, A_1] \otimes [B_1, B_2]\right] &(6.3)\\ &\leq 4 + \big\|[A_2, A_1] \otimes [B_1, B_2]\big\| &(6.4)\\ &\leq 8, &(6.5) \end{aligned}$$

where the last inequality uses that $\|[A_1, A_2]\| \leq \|A_1A_2\| + \|A_2A_1\| \leq 2\|A_1\|\|A_2\| = 2$ and similarly for the B's.
(6.3)-(6.5) shows that $\mathrm{tr}\left[\rho\hat{\beta}\right] \leq 2\sqrt{2}$ as claimed.
In order to prove that equality can be achieved, assume that $\rho = |\psi\rangle\langle\psi|$ where $|\psi\rangle$ is an eigenvector of $\hat{\beta}$ with eigenvalue ν. Then equality holds in (6.3).
Now take $A_1 = B_1 = \sigma_1$ and $A_2 = B_2 = \sigma_2$ Pauli matrices, so that $\hat{\beta}^2 = 4\,\mathbb{1} + 4\,\sigma_3 \otimes \sigma_3$ has eigenvalues 0 and 8. Hence, ν can be chosen s.t. $\mathrm{tr}\left[\rho\hat{\beta}\right]^2 = \nu^2 = 8$. □

Remarks.

- In the early 80's, the violation of CHSH by a factor $\sqrt{2}$ has been verified experimentally. This was done using *down conversion* in a non-linear crystal, which produces entangled pairs of photons, whose polarization degrees of freedom violate CHSH.
- This is a remarkable step in the history of science, since a debate (initially mainly between Einstein and Bohr) that was originally considered metaphysical has eventually been decided by an experiment.

The argumentation can be generalized to more than two observables:

Consider $\langle A_x B_y \rangle =: C_{xy},\ x, y \in \{1, \dots, m\}$ as before, $\gamma \in \mathbb{R}^{m\times m}$ and define

$$\|\gamma\|_{\text{LHV}} := \sup_{a,b\in\{\pm 1\}^m} \left| \sum_{x,y} \gamma_{xy} a_x b_y \right|$$
$$\|\gamma\|_{\text{quantum}} := \sup_{\rho,\{A_x,B_y\}} \left| \sum_{x,y} \gamma_{xy} \operatorname{tr}[\rho\, A_x \otimes B_y] \right|$$

where $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is a density operator and $-\mathbb{1} \le A_x, B_y \le \mathbb{1}$.
The Cirelson bound for the CHSH inequality then reads:

$$\nu(\gamma) := \frac{\|\gamma\|_{\text{quantum}}}{\|\gamma\|_{\text{LHV}}} = \sqrt{2} \text{ for } \gamma = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Theorem 33 (General Cirelson bounds)**.**

(i) $\gamma \in \mathbb{R}^{2\times 2} \Rightarrow \nu(\gamma) \le \sqrt{2}$.

(ii) $\gamma \in \mathbb{R}^{m\times m}, m \in \mathbb{N} \Rightarrow \nu(\gamma) \le K_G < \frac{\pi}{2\ln(1+\sqrt{2})} \approx 1.782$.

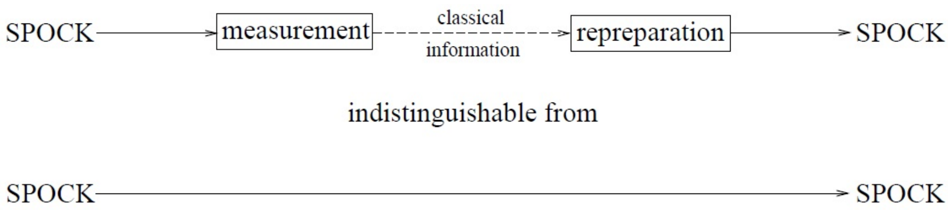
Remark.

- (i) means that the choice of coefficients in CHSH is optimal.
- (ii) is a non-trivial statement whose proof is based on a deep result of Grothendieck. K_G is called *Grothendieck's constant*, which is unknown but equal to the supremum of $\nu(\gamma)$ over all $\gamma \in \mathbb{R}^{m\times m}$ and all $m \in \mathbb{N}$.

7 Impossible machines

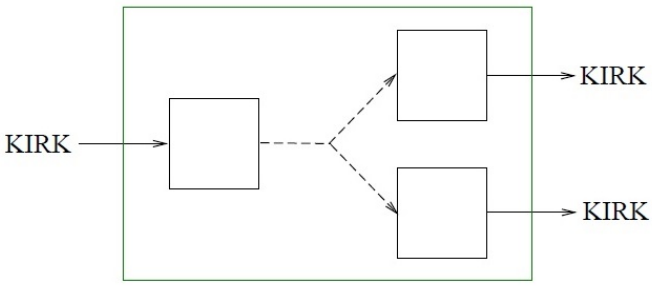
7.1 Classical teleportation

We call a *classical teleporter* a hypothetical machine that consists out of two devices, one for measuring and sending and one for recreation upon receiving the signal.
By assumption a classical teleporter is such that it measures (and possibly disintegrates) an unknown physical system (taken from a sufficiently large set of systems), sends classical information (a bit string) to the receiver, who then reprepares the system in such a way, that no experiment can distinguish the prepared system from the original one.



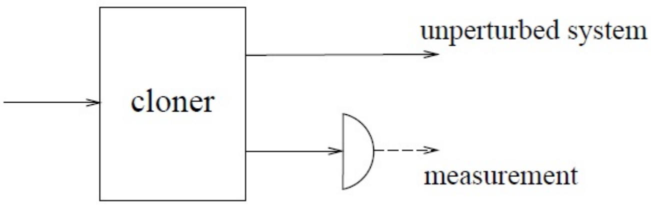
7.2 Cloning

We call a *cloner* a device that upon input of an unknown physical system outputs two copies, which are indistinguishable from the original. Since classical information can be copied, a classical teleporter would enable a cloner.



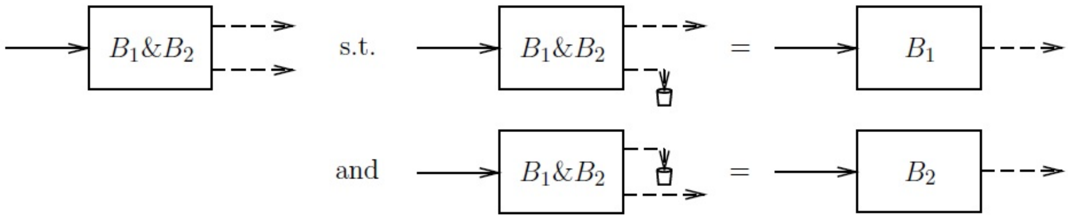
7.3 Measurement without disturbance

Is enabled by a cloning machine.



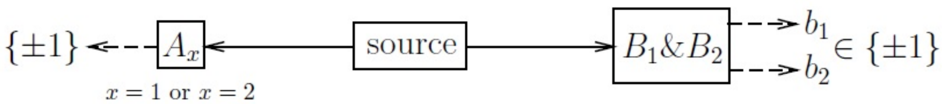
7.4 Joint measurement

We call two measurements with outcomes in S_1 and S_2 respectively *jointly measurable* if there exists a measurement device with outcomes in $S_1 \times S_2$ s.t. both marginal distributions are, on an arbitrary system, indistinguishable from those of the individual measurements. Clearly, a joint measurement device is enabled by measurement without disturbance.



7.5 Superluminal communication

Consider the setting of the CHSH inequality with the difference that Bob is supposed to have a joint measurement device for $B_1\&B_2$.



Define $\beta := \langle A_1B_1\rangle + \langle A_1B_2\rangle + \langle A_2B_1\rangle - \langle A_2B_2\rangle$ and $p_x(a_x, b_1, b_2) := p(a_x, b_1, b_2 \mid A_x)$. Since $B_1\&B_2$ is supposed to be a joint measurement device we have

$$\sum_{b_2} p_x(a_x, b_1, b_2) = p(a_x, b_1 \mid A_x, B_1) \tag{7.1}$$

$$\sum_{b_1} p_x(a_x, b_1, b_2) = p(a_x, b_2 \mid A_x, B_2) \tag{7.2}$$

both for $x \in \{1, 2\}$.

The protocol for *superluminal communication* is now as follows:
Alice encodes a bit she wants to send to Bob by either measuring A_1 or A_2. Bob "guesses" A_1 if $b_1 = b_2$ and A_2 otherwise. If Alice chooses A_1, then Bob is right with probability

$$\sum_{a_1,b_1,b_2} \left|\frac{b_1+b_2}{2}\right| |a_1| p_1(a_1,b_1,b_2).$$

Assume A_1 and A_2 are chosen equally often by Alice. Then the overall probability for Bob to guess correctly is

$$\begin{aligned}
&\frac{1}{2}\sum_{a_1,b_1,b_2} \left|\frac{b_1+b_2}{2}\right| |a_1|\, p_1(a_1,b_1,b_2) + \frac{1}{2}\sum_{a_2,b_1,b_2} \left|\frac{b_1-b_2}{2}\right| |a_2|\, p_2(a_2,b_1,b_2) \\
&\geq \frac{1}{4}\sum_{a_1,b_1,b_2} a_1(b_1+b_2)p_1(a_1,b_1,b_2) + \frac{1}{4}\sum_{a_2,b_1,b_2} a_2(b_1-b_2)p_2(a_2,b_1,b_2) \\
&= \frac{\beta}{4}.
\end{aligned}$$

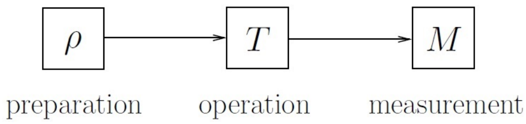
Where the last equation uses that Bob has a joint measurement device satisfying (7.1) and (7.2).
So if $\beta > 2$, i.e., if the CHSH inequality is violated, Bob's guess would be better than random. In that case he would have information about Alice's choice. Consequently, we have

CHSH violation and no superluminal communication
$\Rightarrow$ no joint measurement of the corresponding measurements
$\Rightarrow$ no measurement without disturbance
$\Rightarrow$ no cloning
$\Rightarrow$ no classical teleportation.

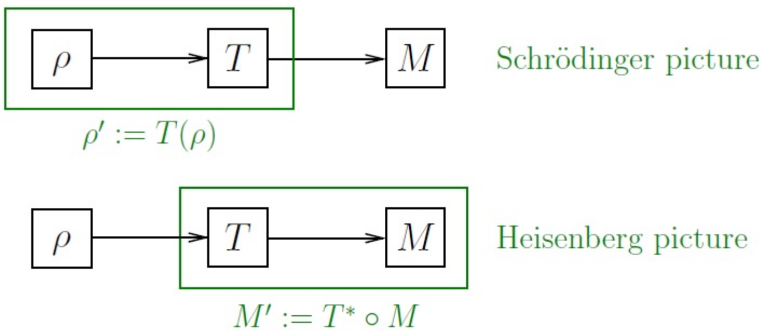
Note that if we regard CHSH violation as an empirical fact, then no quantum theory entered this argumentation. Any theory that allows for CHSH violation but not for superluminal communication is subordinate to this argumentation.

8 Quantum channels and time evolution

Recall that preparation and measurement devices are, within quantum theory, described by density operators and POVMs respectively. Assume we insert some operation between preparation and measurement. (This "operation" may just be free time evolution of the system, if we vary the time between preparation and measurement.)



There are two complementary perspectives for describing this situation within our framework: either we regard T as part of a new preparation (*Schrödinger picture*) or as part of a new measurement (*Heisenberg picture*).



In the Schrödinger picture $T\colon \rho \mapsto \rho'$ has to map density operators onto density operators. In order to remain consistent with the probabilistic interpretation of convex combinations, T has to be a linear map.

Definition 34. Let $T\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ be a linear map.

- T is called *positive* if $A \geq 0 \Rightarrow T(A) \geq 0$.
- T is called *unital* if $T(\mathbb{1}) = \mathbb{1}$.
- T is called *trace-preserving* if $T\colon S_1(\mathcal{H}) \to S_1(\mathcal{H}')$ and $\operatorname{tr}[T(A)] = \operatorname{tr}[A]\ \forall A \in S_1(\mathcal{H})$.
- T is called *completely positive* if $T \otimes \mathrm{id}_n$ is positive for all $n \in \mathbb{N}$, where id_n is the identity map on $\mathbb{C}^{n\times n}$.

Definition 35. $T\colon S_1(\mathcal{H}) \to S_1(\mathcal{H}')$ is called a *quantum channel* in the Schrödinger picture if it is

(i) linear,

(ii) trace-preserving, and

(iii) completely positive.

Each quantum channel $T\colon S_1(\mathcal{H}) \to S_1(\mathcal{H}')$ defined in the Schrödinger picture induces a quantum channel $T^*\colon \mathcal{B}(\mathcal{H}') \to \mathcal{B}(\mathcal{H})$ in the Heisenberg picture via

$$\mathrm{tr}[B\,T(A)] \overset{!}{=} \mathrm{tr}[T^*(B)\,A] \;\forall B \in \mathcal{B}(\mathcal{H}')\;\forall A \in S_1(\mathcal{H}).$$

In order to see that this is well defined note that for fixed $B \in \mathcal{B}(\mathcal{H})$ $A \mapsto \mathrm{tr}[B\,T(A)]$ is a linear functional on $S_1(\mathcal{H})$. Hence, there is a $\tilde{\mathcal{B}} \in \mathcal{B}(\mathcal{H})$ so that
$\mathrm{tr}[B\,T(A)] = \mathrm{tr}[\tilde{\mathcal{B}}\,A]\;\forall A \in S_1(\mathcal{H})$. This defines the map $T^*\colon B \mapsto \tilde{B}$, which is sometimes called *dual* to T. T^* is completely positive iff T is. Moreover:

Corollary 36. *Let $T\colon S_1(\mathcal{H}) \to S_1(\mathcal{H}')$ be a linear map. Then*

$$T^*(\mathbb{1}) = \mathbb{1} \Leftrightarrow T \text{ is trace-preserving.}$$

Proof.

"$\Rightarrow$": $\mathrm{tr}[T(A)] = \mathrm{tr}[T^*(\mathbb{1})A] = \mathrm{tr}[A]$.

"$\Leftarrow$": $0 = \mathrm{tr}[(\mathbb{1} - T^*(\mathbb{1}))A]\;\forall A \in S_1(\mathcal{H})$ e.g. $A = |\Phi\rangle\langle\Psi|$.
So $(\mathbb{1} - T^*(\mathbb{1}))|\Phi\rangle = 0\;\forall \Phi \in \mathcal{H}$ and thus $T^*(\mathbb{1}) = \mathbb{1}$.

□

In order to see why complete positivity is a necessary condition and that mere positivity is not enough consider the transposition:

Definition 37. We write $\Theta\colon \mathbb{C}^{d\times d} \to \mathbb{C}^{d\times d}$ for the *transposition* w.r.t. a given basis. That is, $\Theta\colon A \mapsto \Theta(A) := A^T$.
The map $\Theta \otimes \mathrm{id}$ on $\mathcal{B}\left(\mathbb{C}^d \otimes \mathcal{H}\right)$ is called *partial transposition*.

Note that $(\Theta \otimes \mathrm{id})(A \otimes B) = A^T \otimes B$ extends to all of $\mathcal{B}(\mathbb{C}^d \otimes \mathcal{H})$ by linearity.

Lemma 38. *For the* flip operator $\mathbb{F}\colon \mathbb{C}^d \otimes \mathbb{C}^d \to \mathbb{C}^d \otimes \mathbb{C}^d$ *with* $\mathbb{F} := \sum_{i,j=1}^{d} |i\rangle\langle j| \otimes |j\rangle\langle i|$ *w.r.t. a fixed product orthonormal basis we have:*

(i) $\mathbb{F}\,|\varphi\rangle \otimes |\psi\rangle = |\psi\rangle \otimes |\varphi\rangle\;\forall \varphi, \psi \in \mathbb{C}^d$.

(ii) $\mathrm{tr}[\mathbb{F}\,A \otimes B] = \mathrm{tr}[AB]\;\forall A, B \in \mathbb{C}^{d\times d}$.

(iii) $(\mathrm{id}\otimes\Theta)(\mathbb{F}) = d\,|\Omega\rangle\langle\Omega|$ *with* $|\Omega\rangle := \frac{1}{\sqrt{d}}\sum_{i=1}^{d}|i\rangle\otimes|i\rangle$.

(iv) $\mathrm{spec}(\mathbb{F}) = \{-1,+1\}$.

Proof.

(i)-(iii) follow directly from inserting the definition of $\mathbb{F}$. For instance:

$$(\mathrm{id}\otimes\Theta)(\mathbb{F}) = \sum_{i,j=1}^{d}|i\rangle\langle j|\otimes(|j\rangle\langle i|)^T = \sum_{i,j=1}^{d}|i\rangle\langle j|\otimes|i\rangle\langle j|.$$

(iv) Since $\mathbb{F}^2 = \mathbb{1}\otimes\mathbb{1}$ we have $\mathrm{spec}(\mathbb{F})\subseteq\{-1,1\}$. Moreover, $|\psi\rangle\otimes|\psi\rangle$ and $|\psi\rangle\otimes|\varphi\rangle - |\varphi\rangle\otimes|\psi\rangle$ are eigenvectors of $\mathbb{F}$ w.r.t. eigenvalues ±1.

□

Corollary 39. Θ *is a unital and trace-preserving linear map on* $\mathcal{B}\left(\mathbb{C}^d\right)$ *that is positive but not completely positive.*

Proof. The only non-obvious property should be the last one. This follows from the previous Lemma since $(\mathrm{id}_d\otimes\Theta)(|\Omega\rangle\langle\Omega|) = \frac{1}{d}\mathbb{F} \ngeq 0$ although $|\Omega\rangle\langle\Omega| \geq 0$. □

Consequently, Θ cannot describe a physically reasonable time evolution.
Examples and building blocks for quantum channels:

1. Extensions: $T\colon S_1(\mathcal{H}) \to S_1(\mathcal{H}\otimes\tilde{\mathcal{H}})$, $T(\rho) := \rho\otimes\tilde{\rho}$ for a fixed density operator $\tilde{\rho}$ is completely positive since $\mathcal{B}(\mathbb{C}^n\otimes\mathcal{H}) \ni A \geq 0 \Rightarrow (\mathrm{id}_n\otimes T)(A) = A\otimes\tilde{\rho}\geq 0$ and trace-preserving since $\mathrm{tr}[T(B)] = \mathrm{tr}[B\otimes\tilde{\rho}] = \mathrm{tr}[B]\,\mathrm{tr}[\tilde{\rho}] = \mathrm{tr}[B]$.

2. Partial trace: $\mathrm{tr}_B\colon S_1(\mathcal{H}_A\otimes\mathcal{H}_B)\to S_1(\mathcal{H}_A)$, $X\mapsto \mathrm{tr}_B[X]$.
   Tensorized with id_n this remains a partial trace (albeit from a larger initial space), so it is completely positive and trace preserving.

3. Unitary evolution: Let $U\in\mathcal{B}(\mathcal{H})$. Then $T(A) := UAU^*$ is completely positive since $(\mathrm{id}_n\otimes T)(P) = (\mathbb{1}\otimes U)P(\mathbb{1}\otimes U)^*$.
   Moreover, $\mathrm{tr}[T(A)] = \mathrm{tr}[U^*UA] = \mathrm{tr}[A]$ if $U^*U = \mathbb{1}$.

Remarks.

- If U is a unitary, then T as well as T^{-1} are quantum channels, i.e., the evolution is (in principle) physically reversible. The converse also holds, i.e., if T and T^{-1} are quantum channels, then there is a unitary s.t. $T(A) = UAU^*$.
- The time evolution of a *closed system* (i.e., one that does not interact with its environment) is described by unitary evolution. In this case the time-dependence of U is typically given via the *Schrödinger equation* or its counterpart in the Heisenberg picture the *Heisenberg equation.*

If the input and output spaces match, then compositions as well as convex combinations of quantum channels are quantum channels again.
The three above examples turn out to be the main building blocks:

Theorem 40 (Stinespring dilation in the Schrödinger picture)**.** *Let* $T\colon S_1(\mathcal{H}) \to S_1(\mathcal{H})$ *be a quantum channel in the Schrödinger picture. Then there is a Hilbert space* $\mathcal{H}_E$*, a pure state* $\rho_E \in S_1(\mathcal{H}_E)$ *and a unitary* $U \in \mathcal{B}(\mathcal{H} \otimes \mathcal{H}_E)$ *s.t.*

$$T(\rho) = \mathrm{tr}_E \left[U(\rho \otimes \rho_E)U^*\right].$$

If $\dim(\mathcal{H}) = d < \infty$*, then one can choose* $\dim(\mathcal{H}_E) = d^2$*.*

Corollary 41 (Kraus representation)**.** *If* $T\colon S_1(\mathcal{H}) \to S_1(\mathcal{H})$ *is a quantum channel, then there exists a countable set* $\{K_x \in \mathcal{B}(\mathcal{H})\}$ *with* $\sum_x K_x^* K_x = \mathbb{1}$ *such that*

$$T(\rho) = \sum_x K_x \rho K_x^* \;\forall \rho \in S_1(\mathcal{H}).$$

Proof. Let $\rho_E = |o\rangle\langle o|$ be part of an orthonormal basis $\{|x\rangle \in \mathcal{H}_E\}$ and let $|j,x\rangle := |j\rangle \otimes |x\rangle$ be a corresponding product orthonormal basis for $\mathcal{H} \otimes \mathcal{H}_E$. Then

$$\langle i|T(|m\rangle\langle n|)|j\rangle = \sum_x \langle i,x|U|m,o\rangle\langle n,o|U^*|j,x\rangle.$$

So we can choose $K_x \in \mathcal{B}(\mathcal{H})$ such that $\langle i|K_x|m\rangle := \langle i,x|U|m,o\rangle$. □

The operators K_x are called *Kraus operators*.
Simple consequences of the quantum channel formalism:

Theorem 42 (No-cloning theorem)**.** *Let* $\mathcal{H}$ *be any Hilbert space with* $\dim(\mathcal{H}) > 1$*. Then there is no quantum channel* $T\colon S_1(\mathcal{H}) \to S_1(\mathcal{H}\otimes\mathcal{H})$ *such that* $\forall \rho \in S_1(\mathcal{H})\colon T(\rho) = \rho\otimes\rho$*.*

Proof. This is a consequence of linearity. Let $\{|\psi_i\rangle\langle\psi_i|\}_{i=1}^n$ be a set of orthogonal pure states and $\{\lambda_i \neq 0\}$ probabilities. If there would be such a map, then

$$\begin{array}{rcl} \sum_i \lambda_i\, T(|\psi_i\rangle\langle\psi_i|) & = & \sum_i \lambda_i |\psi_i\rangle\langle\psi_i| \otimes |\psi_i\rangle\langle\psi_i| \text{ has rank } n, \text{ while} \\ \| & & \\ T\left(\sum_i \lambda_i |\psi_i\rangle\langle\psi_i|\right) & = & \sum_{i,j} \lambda_i\lambda_j |\psi_i\rangle\langle\psi_i| \otimes |\psi_j\rangle\langle\psi_j| \text{ has rank } n^2. \end{array}$$

□

Theorem 43 (No-signalling theorem)**.** *Let* $\rho \in S_1(\mathcal{H}_A \otimes \mathcal{H}_B)$ *be a bipartite state,* $T\colon S_1(\mathcal{H}_A) \to S_1(\mathcal{H}_A')$ *a quantum channel on Alice' side and* $M\colon \mathbb{B}(S) \to \mathcal{B}(\mathcal{H}_B)$ *a POVM on Bob's side. Then the statistics of the outcomes of Bob's measurement do not depend on Alice' channel.*

Proof.

$$\begin{aligned}\operatorname{tr}[(T\otimes \mathrm{id})(\rho)(\mathbb{1}\otimes M(I))] &= \operatorname{tr}[\rho\,(T^*\otimes \mathrm{id})(\mathbb{1}\otimes M(I))]\\ &= \operatorname{tr}[\rho\, T^*(\mathbb{1})\otimes M(I)]\\ &= \operatorname{tr}[\rho\,\mathbb{1}\otimes M(I)]\\ &= \operatorname{tr}[\rho_B\, M(I)]\end{aligned}$$

□

In this sense quantum theory is a local theory.

Lemma 44. *If $H \in S_1(\mathcal{H})$ is Hermitian, then*

$$\|H\|_1 = \sup_{-\mathbb{1}\le X\le \mathbb{1}} \operatorname{tr}[XH].$$

Proof. Using the spectral decomposition we can find positive operators $H_+, H_- \in S_1(\mathcal{H})$ s.t. $H = H_+ - H_-$ and $H_+H_- = 0$. Then $\sup_{-\mathbb{1}\le X\le \mathbb{1}} \operatorname{tr}[XH_+] - \operatorname{tr}[XH_-] \le \operatorname{tr}[H_+ + H_-] = \|H\|_1$. Equality is attained for $X = H_+^0 - H_-^0$ where $H_\pm^0$ are the projectors onto the ranges of $H_\pm$. □

Theorem 45 (Quantum channels are trace-norm contractions)**.** *If $T\colon S_1(\mathcal{H}) \to S_1(\mathcal{H}')$ is a quantum channel, then*

$$\forall \rho_1, \rho_2 \in S_1(\mathcal{H})\colon \|T(\rho_1) - T(\rho_2)\|_1 \le \|\rho_1 - \rho_2\|_1.$$

Proof. Note that T^* as any positive map is order preserving on the set of Hermitian operators since $A \ge B \Leftrightarrow A - B \ge 0 \Rightarrow T^*(A-B) \ge 0 \Leftrightarrow T^*(A) \ge T^*(B)$.
So if $-\mathbb{1} \le X \le \mathbb{1}$, then $-\mathbb{1} = -T^*(\mathbb{1}) \le T^*(X) \le T^*(\mathbb{1}) = \mathbb{1}$. This implies that

$$\begin{aligned}\|T(\rho_1-\rho_2)\|_1 &= \sup_{-\mathbb{1}\le X\le \mathbb{1}} \operatorname{tr}[T^*(X)(\rho_1-\rho_2)]\\ &\le \sup_{-\mathbb{1}\le Y\le \mathbb{1}} \operatorname{tr}[Y(\rho_1-\rho_2)]\\ &= \|\rho_1-\rho_2\|_1.\end{aligned}$$

□

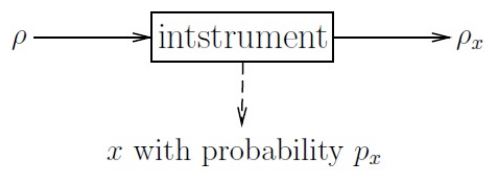
9 Instruments

In many cases a measurement involves a repreparation of the system. For instance, if we measure the spin of an electron in some direction, then the electron still exists after the measurement, but its state may have changed. This situation is described by an *instrument*, which we will for simplicity only discuss for a countable set S of measurement outcomes:

Definition 46. Let S be countable. An *instrument* in the Schrödinger picture is given by a set $\{T_x \colon S_1(\mathcal{H}) \to S_1(\mathcal{H}')\}$ of completely positive linear maps that satisfy that $\sum\limits_{x\in S} T_x$ is trace-preserving.

Interpretation. x is a measurement outcome. Upon input of ρ it is attained with probability $p_x := \operatorname{tr}[T_x(\rho)]$.
Note that positivity of T_x ensures that $p_x \geq 0$ and $\sum\limits_x p_x = \sum\limits_x \operatorname{tr}[T_x(\rho)] = 1$.

The state after having measured x is $\rho_x := \frac{T_x(\rho)}{p_x}$.



Remark. Instruments can be considered both generalizations and special cases of quantum channels.
In order to arrive at the latter point of view, let $\{|x\rangle\}_{x\in S}$ be an orthonormal basis and consider the quantum channel

$$T(\rho) := \sum_{x\in S} T_x(\rho) \otimes |x\rangle\langle x|.$$

Theorem 47 (No information without disturbance)**.**
Consider an instrument that does not disturb the input on average in the sense that for all density operators ρ we have $\sum\limits_x p_x\rho_x = \rho$.
Then p_x is independent of ρ, i.e., $\operatorname{tr}[T_x(\rho)] = \operatorname{tr}[T_x(\rho')]$ for any density operator ρ'.

Proof. (for a finite dimensional $\mathcal{H}$) We exploit that every linear map on $S_1(\mathcal{H})$ is uniquely characterized by its Choi-Jamiolkowski operator.
First note that $\sum_x p_x \rho_x = \rho$ is equivalent to $\sum_x T_x = \mathrm{id}$.
Hence the Choi-Jamiolkowski operators have to coincide: if $\tau_x := (\mathrm{id} \otimes T_x)(|\Omega\rangle\langle\Omega|)$, then $\sum_x \tau_x = |\Omega\rangle\langle\Omega|$.
Since $\tau_x \geq 0$ we have to have $\tau_x = q_x |\Omega\rangle\langle\Omega|$ for some $q_x \in [0, 1]$ with $\sum_x q_x = 1$.
Therefore $T_x = q_x \,\mathrm{id}$ so that $p_x = \mathrm{tr}[T_x(\rho)] = q_x$ is independent of ρ. □

10 Teleportation

We saw already that classical teleportation is impossible, if we accept no-signalling and the violation of the CHSH inequality. In the following we will see that a small but crucial modification namely *entanglement assisted teleportation* is possible.

Lemma 48. *For $d \in \mathbb{N}$ consider the set of d^2 unitaries*

$$U_j := \sum_{r=0}^{d-1} \eta^{r j_2} |r + j_1\rangle\langle r|, \ \eta := e^{\frac{2\pi i}{d}}, \ j = (j_1, j_2) \in \mathbb{Z}_d \times \mathbb{Z}_d$$

where addition inside the ket is modulo d and $\{|x\rangle \in \mathbb{C}^d\}_{x \in \mathbb{Z}_d}$ is an orthonormal basis. Then

(i) $\{U_j\}$ is a basis of operators in $\mathbb{C}^{d\times d}$ that is orthogonal w.r.t. the Hilbert-Schmidt inner product. More specifically, $\mathrm{tr}[U_i^* U_j] = d\, \delta_{i,j}$.

(ii) $U_i U_j = \eta^{i_2 j_1} U_{i+j}$ (with addition modulo d). Thus $U_j^{-1} = \eta^{j_1 j_2} U_{-j}$.

(iii) For $d = 2$ the set reduces to the set of Pauli matrices with identity, i.e., $(\mathbb{1}, \sigma_x, \sigma_y, \sigma_z) = (U_{(0,0)}, U_{(1,0)}, iU_{(1,1)}, U_{(0,1)})$.

Corollary 49. *Let $|\Omega\rangle := \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k\rangle \otimes |k\rangle$ be maximally entangled. Then with $j \in \mathbb{Z}_d \times \mathbb{Z}_d$ the set $|\Omega_j\rangle := (U_j \otimes \mathbb{1})|\Omega\rangle$ is an orthonormal basis of maximal entangled states in $\mathbb{C}^d \otimes \mathbb{C}^d$.*

Proof. The Ω_j's are orthonormal due to $\langle \Omega_i | \Omega_j \rangle = \mathrm{tr}[U_i^* U_j \otimes \mathbb{1} |\Omega\rangle\langle\Omega|] = \frac{1}{d}\,\mathrm{tr}[U_i^* U_j] = \delta_{i,j}$. Since there are d^2 Ω_j's, they form an orthonormal basis. □

Consider the following scenario described on $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2} \otimes \mathcal{H}_B$ with $\mathcal{H}_{A_1} \simeq \mathcal{H}_{A_2} \simeq \mathcal{H}_B$: Assume $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$ and $\mathcal{H}_B$ describe Alice's and Bob's part respectively and that the initial state is pure and described by $|\Psi\rangle_{A_1} \otimes |\Omega\rangle_{A_2 B}$, where $|\Omega\rangle := \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k\rangle \otimes |k\rangle$ is maximally entangled.
The task is now to "teleport" the state $|\Psi\rangle$ from Alice to Bob by

(i) performing a measurement on Alice's side

(ii) sending the measurement outcome to Bob

(iii) performing an operation on Bob's side.

None of these steps is allowed to depend on Ψ.

(i) The measurement:
Take $S = \mathbb{Z}_d \times \mathbb{Z}_d$ and $M(j) = |\Omega_j\rangle\langle\Omega_j|$ where $|\Omega_j\rangle := (U_j \otimes \mathbb{1})|\Omega\rangle$.
We want to determine the state ρ_j on Bob's side conditioned on Alice having obtained outcome j. Suppose this happens with probability p_j.
Then with $\underline{\overline{\Psi}} := |\Psi\rangle\langle\Psi|,\ \omega := |\Omega\rangle\langle\Omega|$

$$\begin{aligned}\operatorname{tr}[\rho_j B] &= \frac{1}{p_j} \operatorname{tr}\left[\left(\underline{\overline{\Psi}}_{A_1} \otimes \omega_{A_2 B}\right)(M_j \otimes B)\right] \\ &= \frac{1}{p_j} \operatorname{tr}\left[\left(\left(U_j^* \underline{\overline{\Psi}} U_j\right)_{A_1} \otimes \omega_{A_2 B}\right)(\omega_{A_1 A_2} \otimes B)\right] \\ &= \frac{1}{d^2 p_j} \operatorname{tr}\left[U_j^* \underline{\overline{\Psi}} U_j B\right] \ \forall B \in \mathbb{C}^{d\times d}.\end{aligned}$$

Applying this to $B = \mathbb{1}$, we see that $p_j = \frac{1}{d^2}$ independent of Φ.
Hence after the measurement Bob's state is $\rho_j = U_j^* \underline{\overline{\Psi}} U_j$ if the outcome was j.

(ii)-(iii) If Alice informs Bob about the measurement outcome, he can apply the channel $T_j(\rho) := U_j \rho U_j^*$ so that finally he possesses the state $\underline{\overline{\Psi}}$.