Mathematical Aspects of Quantum Information Theory:

Lecture 1

Dario Trevisan

Università di Pisa
dario.trevisan@unipi.it

Plan

1. Introduction to the course
2. Postulates of Quantum Mechanics
   - Classical physics and probability
   - (Elementary) quantum systems and their states
   - Measurements and observables

From classical to quantum information theory

Information theory studies the laws of storage and communication of information:

- Traditionally initiated as a field in the 1940s by C. Shannon,
- a scientific field at the intersection of probability theory, statistics, computer science, statistical mechanics, information engineering, electrical engineering…

Quantum information theory:

- studies limitations and new possibilities by the quantum mechanical aspects of nature,
- independent research area since the 1990s,
- based on the postulates of quantum mechanics $\rightarrow$ further mathematical tools, in particular functional analysis (operator theory).

Aim of this course

- An introduction to the main mathematical aspects of quantum information theory.
- Main result: quantify how information deteriorates when transmitted through a quantum noisy communication channel, via a quantum coding theorem, extending the classical Shannon fundamental limit.
- No prior knowledge in classical information theory, nor in quantum mechanics, is required.
- Target audience: mathematicians with a background in probability, analysis or mathematical physics.

Structure of the course

We plan in total 6 lectures (in the morning, 11:00-13:00) and 6 problem sessions, to discuss examples and exercises (in the afternoon, 14:30-15:30).

1. 23/01 (Mon): postulates of Quantum mechanics
2. 27/01 (Fri): Quantum channels
3. 03/02 (Fri): Inequalities
4. 08/02 (Thu): Distances between quantum states
5. 13/02 (Mon): Quantum entropy
6. 17/02 (Fri): A coding theorem

Teaching material

The exposition selects from monographs from various authors (Nielsen-Chuang, Holevo, Wilde, Alicki-Fannes, Meyer, . . .)

At the webpage
`http://people.dm.unipi.it/trevisan/teaching.html`
you can download

1. Lecture notes (updated just before each lecture)
2. Slides (also annotated after the lecture)
3. Recordings (if possible)

I am also available for discussions online:

- email: `dario.trevisan@unipi.it`,
- Skype: dario-trevisan

If you plan to give the final exam, ask me about it!

Plan

Classical physics

A system is described via three mathematical objects:

1. A set Ω (the phase space): $\omega \in \Omega$ represent a possible *state* of the system.
2. Observables, i.e., $X : \Omega \to \mathcal{X}$, with possible outcomes $\mathcal{X}$, representing quantities obtained via physical measurements.
3. Transformations $T_t : \Omega \to \Omega$ representing the time evolution of the system.

Elementary probability

Analogues of these three objects can be found:

1. The (finite) set Ω (sample space), with $\omega \in \Omega$ describing the possible outcome of a random experiment.

   However, states are probability distributions $\rho : \Omega \to [0, 1]$, such that $\sum_{\omega \in \Omega} \rho(\omega) = 1$.

2. Random variables X on Ω, with values in $\mathcal{X}$. Events $V \subseteq \Omega$, model logical statements (i.e., either true or false) are naturally associated with indicator random variables 1_V with $\mathcal{X} = \{0, 1\}$.
3. Stochastic processes describe time evolutions of Ω, e.g. example Markov chains.

Quantum mechanics and its postulates

- *Quantum mechanics* is a physical theory, supported by a vast experimental evidence ($\sim$ 100 years old), describing accurately phenomena at very small scales (atoms, molecules, light),
- probabilistic features: it only predicts the odds that some event will occur.
- Its axioms describe three mathematical objects (states, observables, time evolution) following the above scheme, but with a twist (non-commutativity!)
- We first describe the elementary setting, i.e., finite-dimensional systems (roughly corresponding to Ω finite).
- We next introduce the C^*-algebra formalism to cover infinite-dimensional cases.

Plan

An elementary quantum system is described by a finite-dimensional complex Hilbert space $(H, \langle\cdot|\cdot\rangle)$.

1. The scalar product is linear in the right variable and anti-linear in the left variable.

2. We use Dirac's ket notation $|\psi\rangle \in H$,

3. bra vectors $\langle\varphi| \in H^*$ denote

$$\langle\varphi| : H \to \mathbb{C}, \qquad |\psi\rangle \mapsto \langle\varphi|\psi\rangle.$$

4. Riesz map $|\psi\rangle \mapsto \langle\psi|$ is anti-linear

5. Families of vectors $(|\psi_i\rangle)_{i\in I}$ will be often written as $(|i\rangle)_{i\in I}$.

Single qubit system

- When $H = \mathbb{C}^d$, the standard basis is written $|i\rangle_{i=1}^{d}$
- It is actually more common to count from 0, i.e., $(|k\rangle)_{k=0}^{d-1}$, and call it the computational basis.
- For $d = 2$, any vector is represented as
- The case $d = 2$ is called single qubit system, for general d one uses the term qudit system.

State vectors

1. Naively, H corresponds to Ω, but a less redundant description would be in terms of the complex projective space over H.

2. It is more convenient to keep the linear structure of H and define as state vector any $|\psi\rangle \in H$ with unit norm, i.e.,

$$\langle\psi|\psi\rangle = \|\psi\|^2 = 1.$$

3. Physically, $|\psi\rangle$ will be indistinguishable from a multiple $e^{i\theta}|\psi\rangle$ with $\theta \in \mathbb{R}$ (called a *phase*).

4. State vectors may be also called wave functions or slight improperly pure states.

Amplitudes and quantum superposition

- Even if H is finite dimensional, the set of state vectors is infinite.
- For an orthonormal basis $(|i\rangle)_{i \in I} \subseteq H$ any state vector can be represented as a quantum superposition

$$|\psi\rangle = \sum_{i \in I} \alpha_i |i\rangle$$

where $\alpha_i = \langle i|\psi\rangle \in \mathbb{C}$ are amplitudes satisfying

$$\sum_{i \in I} |\alpha_i|^2 = 1.$$

- The squared moduli $|\alpha_i|^2 = |\langle i|\psi\rangle|^2$ can be interpreted as probabilities,
- but $|\psi\rangle$ is *not* a classical probability distribution over the $|i\rangle$'s with density $|\alpha_i|^2$.
- Changing a single phase in an amplitude may give a different state vector!

Density operators

- Quantum analogues of probability distributions are density operators.
- Pure state associated to a state vector $|\psi\rangle \in H$: orthogonal projection

$$P_{|\psi\rangle}$$

- We then take convex combinations:

$$\rho = \sum_{i \in I} p_i |\psi_i\rangle \langle\psi_i|,$$

with $|\psi_i\rangle \in H$ state vectors, and $p_i \in [0, 1]$ a probability distribution, i.e., $\sum_{i \in I} p_i = 1$.

Abstract characterization

Density operators $\rho \in \mathcal{S}(H)$ are

1. self-adjoint (or Hermitian) $\rho^* = \rho$,
2. positive $\rho \geq 0$,
3. with unit trace $\mathrm{tr}[\rho] = 1$.

By spectral theorem (in finite dimensions):

Density matrix (with respect to a basis)

1. Fix an orthonormal basis $(|i\rangle)_{i \in I} \subseteq$ (with $|I| = d = \dim(H)$)
2. Any operator $A : H \to H$ can be represented as a matrix
$$A_{ij} = \langle i|Aj\rangle \, .$$
3. A density operator corresponds to a density matrix $(\rho_{ij})_{i,j \in I}$ such that
4. The diagonal $(\rho_{ii})_{i \in I}$ induces a classical probability distribution over I.
5. We identify classical probability distributions with diagonal matrices (however it depends on the basis!).

Exercises

Exercise: (Hilbert-Schmidt scalar product) Let H be an elementary quantum system and $A, B \in \mathcal{L}(H)$. Prove that

$$(A, B) \mapsto \mathrm{tr}[A^* B]$$

defines a scalar product on $\mathcal{L}(H)$ (called Hilbert-Schmidt scalar product). By choosing an orthonormal basis $(|i\rangle)_{i \in I}$, write explicitly its expression in terms of the matrices representing A and B.

Exercise: Given a density operator $\rho \in \mathcal{S}(H)$ on an elementary quantum system H, define its *purity* as $\mathrm{tr}[\rho^2]$.

1. Prove that the purity always belongs to the interval $[\dim(H)^{-1}, 1] \subseteq (0, 1]$.
2. Prove that ρ is a pure state if and only if its purity equals 1.

Plan

Observables

Aim: define the analogue of functions over a classical phase space, or of random variables on a sample space, taking values in a (finite) set $\mathcal{X}$. In brief:

- Observables $A \in \mathcal{O}(H)$ on a quantum system H are defined as self-adjoint operators $A : H \mapsto H$.
- The spectrum, i.e. the set of eigenvalues $\sigma(A) \subseteq \mathbb{R}$ plays the role of the "set of possible values' of the observable A when measured through a (hypothetical) device

These observables correspond to real-valued random variables. What about a general $\mathcal{X}$?

Indicator observables

Let us follow a path akin to elementary probability theory: we model logical propositions ("events") about an elementary quantum system H as subspaces $V < H$:

- the 0-dimensional $V = \{0\}$ represents a false statement
- the whole $V = H$ represents a true statement
- One-dimensional subspaces spanned by a state vector $|\psi\rangle$ can be interpreted as the proposition

"the quantum system is in the state associated to $|\psi\rangle$".

To each V, we associate its indicator observable $\mathbb{1}_V : H \mapsto H$, the orthogonal projection operator on V, which is

- self-adjoint $\mathbb{1}_V = \mathbb{1}_V^*$,
- $\mathbb{1}_V^2 = \mathbb{1}_V$,
- hence $\sigma(\mathbb{1}_V) = \{0, 1\}$.

Measuring $\mathbb{1}_V$

The observable $\mathbb{1}_V$ is interpreted as a physical device that, when applied to the system, yields outcomes 1 if V holds or 0 if V does not hold, with some probability according to the state of the system $\rho \in \mathcal{S}(H)$.

We postulate that

1. by measuring $\mathbb{1}_V$, the probability of observing that V holds (i.e., outcome is 1) is given by (Born's rule):

$$\mathbb{P}_\rho(V) = \mathbb{P}(\mathbb{1}_V = 1) := \mathrm{tr}[\mathbb{1}_V \rho] \in [0, 1]$$

   If $\rho = |\psi\rangle\langle\psi|$ is a pure state, this equals

$$\mathrm{tr}[\mathbb{1}_V \rho] = \mathrm{tr}\mathbb{1}_V |\psi\rangle\langle\psi| = \langle\psi|\mathbb{1}_V \psi\rangle = \|\mathbb{1}_V \psi\|^2 .$$

2. having measured $\mathbb{1}_V$ and observed that V holds, the state of the system is updated from ρ to the density operator (collapse of the state):

$$\rho_V = \frac{\mathbb{1}_V \rho \mathbb{1}_V}{\mathbb{P}_\rho(V)}$$

Interpretation of Born's rule and collapse of the state

Compare

$$\rho_V = \frac{\mathbb{1}_V \rho \mathbb{1}_V}{\mathbb{P}_\rho(V)} \quad \text{with the classical rule:} \quad \mathbb{P}(\cdot|V) = \frac{\mathbb{P}(\cdot \text{ and } V)}{\mathbb{P}(V)}.$$

- The interpretation of the measurement postulate can be various (akin to Bayesian vs frequentist in probability and statistics)
- What do states and probabilities in quantum mechanics represent? are they states of knowledge (subjective) or objective?
- We may (safely?) interpret that they describe relative frequencies in the ideal infinite limit of a repeated sequence of independent experiments, in a prepared situation (measurements yield classical i.i.d. sequences).

Example: projection on a state

- Let V be generated by a state vector $|\varphi\rangle \in H$, so that the indicator is $\mathbb{1}_V = |\varphi\rangle\langle\varphi| \in \mathcal{O}(H)$. Notice that it coincides with the associated density operator.
- $\mathbb{1}_V = |\varphi\rangle\langle\varphi| \in \mathcal{O}(H)$ yields therefore outcome 1 if "the quantum system is in the state associated to $|\psi\rangle$", with probability

$$\mathbb{P}_\rho(V) = \mathrm{tr}[\rho\,|\varphi\rangle\langle\varphi|] = \langle\varphi, \rho\varphi|.\rangle$$

In the classical case this would give probability either 0 or 1!

- After measuring $\mathbb{1}_V$ and observing outcome 1, the state collapses to the pure state associated to $|\varphi\rangle$.

The case that measuring $\mathbb{1}_V$ yields outcome 0

What about measuring $\mathbb{1}_V$ yields outcome 0, i.e. V does not hold?

1. Write $\mathbb{1}_V = \mathbb{1}_H - \mathbb{1}_{V^\perp}$, where $V^\perp < H$ is the orthogonal subspace (interpret $V^\perp$ as the negation of V).

2. Thus, it happens with probability

$$\mathbb{P}_\rho(V^\perp) = \mathbb{P}_\rho(\mathbb{1}_V = 0) = \operatorname{tr}[\mathbb{1}_{V^\perp}\rho] = 1 - \operatorname{tr}[\mathbb{1}_V\rho] = 1 - \mathbb{P}_\rho(V),$$

3. The density operator updates in this case to

$$\rho_{V^\perp} = \frac{\mathbb{1}_{V^\perp}\rho\mathbb{1}_{V^\perp}}{\mathbb{P}_\rho(V^\perp)}.$$

Measuring but not observing

Can we describe the state of the system after $\mathbb{1}_V$ has been measured but the outcome has been observed?

- We postulate it to be the convex combination

$$\rho_V \mathbb{P}_\rho(V) + \rho_{V^\perp} \mathbb{P}_\rho(V^\perp) = \mathbb{1}_V \rho I_V + \mathbb{1}_{V^\perp} \rho I_{V^\perp}. \tag{1}$$

- Compare with the law of total probability:

$$\mathbb{P}(\cdot) = \mathbb{P}(\cdot | V)\mathbb{P}(V) + \mathbb{P}(\cdot | V^c)\mathbb{P}(V^c),$$

- But in the quantum case the state is not ρ (in general).

- Intepretation: when $\mathbb{1}_V$ is measured, it interacts with the system, and perturbs its state (one cane give better description via so-called *de-coherence* phenomenon)

Joint measurements and compatibility

- Given $V, W < H$ with indicator observables $\mathbb{1}_V, \mathbb{1}_W \in \mathcal{O}(H)$, they are called compatible if they commute:

$$\mathbb{1}_V \mathbb{1}_W = \mathbb{1}_W \mathbb{1}_V \quad \text{or} \quad [\mathbb{1}_V, \mathbb{1}_W] = 0.$$

- In such a case, measuring first $\mathbb{1}_V$ and then $\mathbb{1}_W$ yields joint outcomes in $\{0,1\}^2$ with the same probability distribution as measuring in the opposite order:

$$\mathbb{P}_\rho(\text{first } \mathbb{1}_V = 1, \text{ then } \mathbb{1}_W = 1) = \mathbb{P}_\rho(V)\mathbb{P}_{\rho_V}(W) = \mathrm{tr}[\rho \mathbb{1}_V] \cdot \frac{\mathrm{tr}[\mathbb{1}_V \rho \mathbb{1}_V \mathbb{1}_W]}{\mathrm{tr}[\rho \mathbb{1}_V]}.$$

- Moreover, the state also updates to a well-defined state, e.g. after observing that both V and W hold:

$$\rho_{V,W} = \frac{\mathbb{1}_W \rho_V \mathbb{1}_W}{\mathbb{P}_{\rho_V}(\mathbb{1}_W = 1)} = \frac{\mathbb{1}_W \mathbb{1}_V \rho \mathbb{1}_V \mathbb{1}_W}{\mathbb{P}_\rho(V, W)},$$

- Notice that if V and W are orthogonal, $\mathbb{1}_V \mathbb{1}_W = 0$, they are compatible.

In the incompatible case, probabilities and the updated states may depend on the order in which the measurements are performed.

- Consider one-dimensional V, W, i.e.,

$$\mathbb{1}_V = |\varphi_0\rangle\langle\varphi_0|, \quad \mathbb{1}_W = |\varphi_1\rangle\langle\varphi_1|,$$

  and assume that $\rho = |\psi\rangle\langle\psi|$.

- Then,

$$\mathbb{P}_{|\psi\rangle}(\text{first } V\text{, then } W) = \mathrm{tr}[\rho \mathbb{1}_V I_W \mathbb{1}_V] = |\langle\varphi_1|\varphi_0\rangle\langle\varphi_0|\psi\rangle|^2 .$$

- Measuring first $\mathbb{1}_W$ and then $\mathbb{1}_V$ instead gives as observed outcomes that both W and V hold with probability

$$\mathbb{P}_{|\psi\rangle}(\text{first } W\text{, then } V) = |\langle\varphi_1|\varphi_0\rangle\langle\varphi_0|\psi\rangle|^2 ,$$

  which is different e.g. if $\langle\varphi_0|\varphi_1\rangle \neq 0$, $\langle\varphi_0|\psi\rangle \neq 0$ but $\langle\varphi_1|\psi\rangle = 0$.

Measurements with outcomes in a (finite) set $\mathcal{X}$

- Recall that random variable X with values in $\mathcal{X}$ can be identified with its system of alternatives $(\{X = x\})_{x \in \mathcal{X}}$.
- By considering the associated indicator variables, this amounts to require that

$$1_{\{X=x\}} 1_{\{X=y\}} = 0 \quad \text{for every } x \neq y \in \mathcal{X}, \text{ and} \quad \sum_{x \in \mathcal{X}} 1_{\{X=x\}} = 1_{\Omega}.$$

- By analogy we define (elementary) measurement X on a quantum system H as a collection of closed subspaces $X = (V_x)_{x \in \mathcal{X}}$ – or indicator observables $X = (\mathbb{1}_{V_x})_{x \in \mathcal{X}}$ – such that

$$\mathbb{1}_{V_x} \mathbb{1}_{V_y} = 0 \quad \text{for every } x \neq y \in \mathcal{X}, \text{ and} \quad \sum_{x \in \mathcal{X}} \mathbb{1}_{V_x} = \mathbb{1}_H.$$

- Such a family of operators is an elementary instance of a projection-valued measure (PVM).

From

$$\mathbb{1}_{V_x}\mathbb{1}_{V_y} = 0 \quad \text{for every } x \neq y \in \mathcal{X}, \text{ and } \quad \sum_{x \in \mathcal{X}} \mathbb{1}_{V_x} = \mathbb{1}_H,$$

we see that

- the V_x's are *compatible*, hence they can be measured in any order yielding the outcomes with well-defined probabilities.

- We refer to such operation as *measuring* X.

- if the system is in the state ρ and X is measured, the probability that V_x holds is

- The "distribution of X" is $(\mathbb{P}_\rho(X = x))_{x \in \mathcal{X}}$.

- If X is measured but the outcome is not observed, the density operator ρ updates to

Compatible measurements

- We say that two measurements $X = (V_x)_{x \in \mathcal{X}}$, $Y = (W_y)_{y \in \mathcal{Y}}$ are compatible if $\mathbb{1}_{V_x} \mathbb{1}_{W_x} = \mathbb{1}_{W_x} \mathbb{1}_{V_x}$ for every $x \in \mathcal{X}$, $y \in \mathcal{Y}$.
- In such a case, measuring X and Y yields observed values x, y with a probability $\mathbb{P}_\rho(X = x, Y = y)$ which does not depend on the order of the measurements, and also a well-defined updated state $\rho_{|X=x, Y=y}$.

Observables as real-valued measurements

We identify quantum observables as measurements with values in $\mathcal{X} \subseteq \mathbb{R}$:

- Given $X = (V_x)_{x \in \mathcal{X}}$ with $\mathcal{X} \subseteq \mathbb{R}$, we define

$$A_X = \sum_{x \in \mathcal{X}} x \mathbb{1}_{V_x} \in \mathcal{O}(H),$$

  so that $\sigma(A_X) = \mathcal{X}$.
- Viceversa, given $A \in \mathcal{O}(H)$, use the spectral theorem to represent

$$A = \sum_{\lambda \in \sigma(A)} \lambda \mathbb{1}_{\{A=\lambda\}},$$

  where $\{A = \lambda\}$ denotes the eigenspace associated to $\lambda \in \sigma(A)$.

The distribution of A is the collection of probabilities, for $\lambda \in \sigma(A)$,

$$\mathbb{P}_\rho(A = \lambda) = \mathrm{tr}[\rho \mathbb{1}_{A=\lambda}],$$

and can be used to define/compute e.g. mean, variance etc. of A:

$$(A)_\rho = \sum_{\lambda \in \sigma(A)} \lambda \mathbb{P}_\rho(A = \lambda) = \mathrm{tr}[\rho A].$$

Functional calculus and compatible observables

- If $A \in \mathcal{O}(H)$, and $f : \sigma(A) \to \mathbb{R}$, define

$$f(A) = \sum_{\lambda \in \sigma(A)} f(\lambda) \mathbb{1}_{\{A=\lambda\}} \in \mathcal{O}(H)$$

- Then

$$(f(A))_\rho = \sum_{\lambda \in \sigma(A)} f(\lambda) \mathbb{P}_\rho(A = \lambda) = \mathrm{tr}[\rho f(A)].$$

- Two observables $A, B \in \mathcal{O}(H)$ are compatible (in the sense that the associated measurements are compatible) if and only if they commute, $[A, B] = 0$.

Exercises

Exercise: Let H be an elementary quantum system and $A, B \in \mathcal{L}(H)$. Discuss the validity of the following statements.

1. If $A, B \in \mathcal{O}(H)$, then $\mathrm{tr}[AB] \in \mathbb{R}$.
2. If $\mathrm{tr}[AB] \in \mathbb{R}$ for every $B \in \mathcal{O}(H)$, then necessarily $A \in \mathcal{O}(H)$.
3. If $A, B \in \mathcal{O}_{\geq}(H)$, then $\mathrm{tr}[AB] \geq 0$.
4. If $A \in \mathcal{O}(H)$ and $\mathrm{tr}[AB] \geq 0$ for every $B \in \mathcal{O}_{\geq}(H)$, then necessarily $A \geq 0$.

Exercises

Exercise (A quantum Jensen inequality) On an elementary quantum system H, consider an observable $A \in \mathcal{O}(H)$. Let $f : \sigma(A) \to \mathbb{R}$ be convex, i.e.

$$f\left(\sum_{x\in\sigma(A)} x p_x\right) \leq \sum_{x\in\sigma(A)} f(x) p_x$$

for every probability distribution $(p_x)_{x\in\sigma(A)}$.

For every density operator $\rho \in \mathcal{S}(H)$, prove the following inequality:

$$f((A)_\rho) \leq (f(A))_\rho.$$