# I. Classical World

$X$ - discrete random variable

$$\Pr\{X = x_i\} = p_i$$

Entropy:

$$H(X) = -\sum_i p_i \log p_i$$

Conditional entropy:

$$H(Y|X) = \sum_i p_i H(Y|X = x_i)$$

Information (Shannon, 1948):

$$I(X;Y) = H(Y) - H(Y|X) = H(X) - H(X|Y)$$

Classical communication channel:

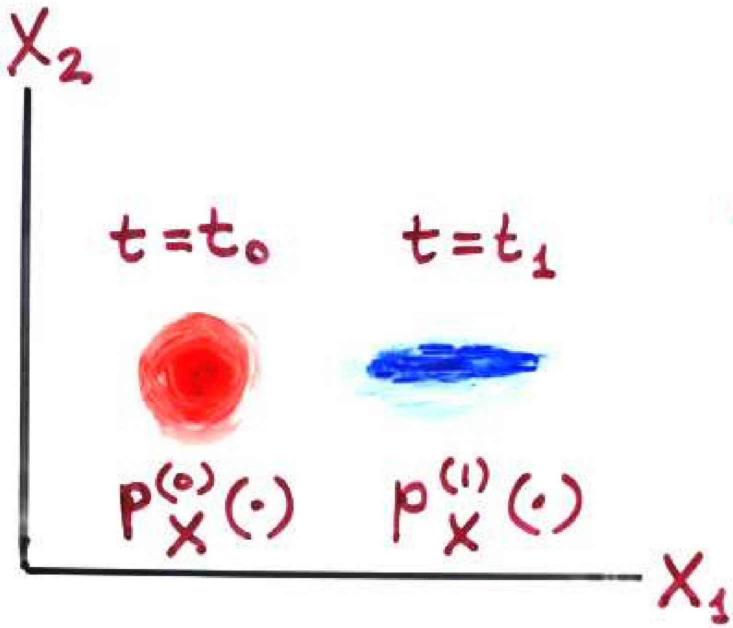
$$\underset{p_i,\ x_i}{X} \longrightarrow \underset{p_i,\ P_{Y|X}(\cdot|x_i)}{Y}$$

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

# Information and Irreversibility of Quantum Measurement

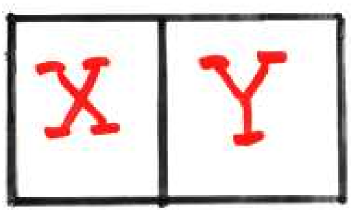
# Information and irreversibility (the Second Law of thermodynamics)

Let $X$ be a complete set of dynamic variables of a physical system




$$H_0(X) = H_1(X)$$

The "fine-grained" entropy is an integral of motion.




At time $t=0$:

$$p_{XY}^{(0)}(x_i, y_j) = p_X^{(0)}(x_i)\, p_Y^{(0)}(y_j)$$

and $H_0(X,Y) = H_0(X) + H_0(Y)$

At time $t=t_1$: $p_{XY}^{(1)}(x_i, y_j) \neq p_X^{(1)}(x_i)\, p_Y^{(1)}(y_j)$

Hence: $H_1(X) + H_1(Y) > H_1(X,Y) =$

$= H_0(X,Y) = H_0(X) + H_0(Y)$

Thus, $H(X)+H(Y)$ <u>increases</u> with time (the "coarse-grained" entropy)

Why?

$$[H_1(X)+H_1(Y)]-[H_0(X)+H_0(Y)]=$$

$$=H_1(X)+H_1(Y)-H_1(X,Y)=$$

$$=I_1(X,Y)$$

The "coarse-grained" entropy ignores information in one subsystem about the other.

# II. Quantum World

The concept of Shannon's information has to be modified: to change the form in order to preserve the meaning.

The meaning of Shannon's information: the number of information units (e.g. bits) that can be determined with infinitesimal probability of error, based on the results of a measurement.

No measurement ⟶ no information

Entropy of a quantum system:

$$H\{\rho\} = -\mathrm{Tr}\,\rho \ln \rho$$ (von Neumann, 1932)

# Quantum communication channel (with classical input)

$$X \longrightarrow \rho$$

$$p_i, x_i \qquad\qquad p_i, \rho_i$$

$$\rho = \sum_i p_i \rho_i$$

Entropy defect:

$$D = -\mathrm{Tr}\,\rho \ln \rho + \sum_i p_i \mathrm{Tr}\,\rho_i \ln \rho_i$$

Th.1 $\quad 0 \leq D \leq -\sum_i p_i \ln p_i = H(X)$ (1969)

= iff all $\rho_i$ are equal

= iff all $\rho_i$ are orthogonal

Information: (measurement!)

L – direct (von Neumann) measurement

K(L) – random variable resulting from the measurement

$I(K(L);X)$ – information from measurement $L$ about $X$.

"Accessible information" $\mathcal{I}$:

$$\mathcal{I} = \sup_{\{L\}} I(K(L);X)$$

**Th. 2** ("Entropy defect principle")

$$\mathcal{I} \leq D$$

(1969)

(= iff all $\rho_i$ commute) (Holevo 1973)

Why? Because of the irreversibility of quantum measurement.

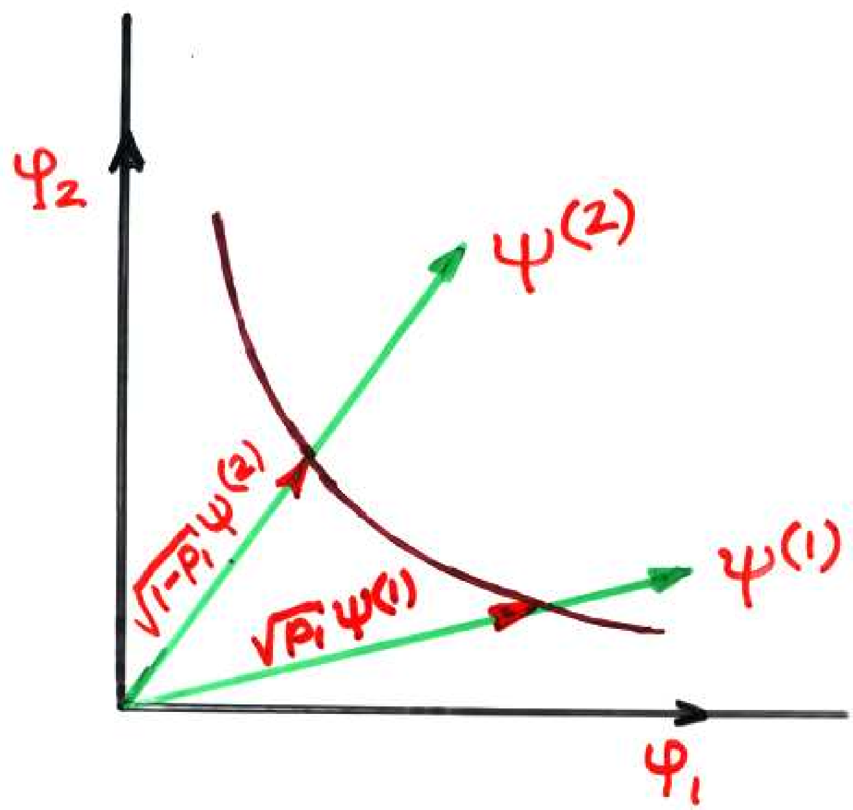
(In classics $\mathcal{I} = D$ always)

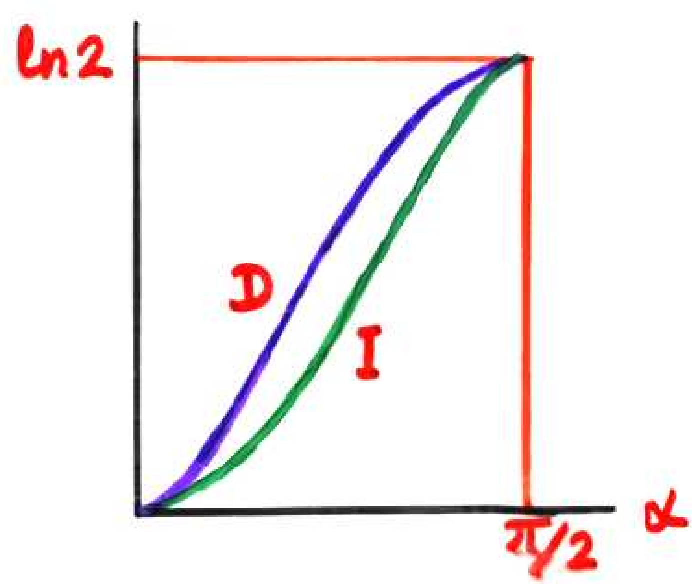
In general, $\mathcal{I}$ is hard to calculate.

# Example.

Two pure states (1971, 1977, 1994)

$\psi^{(1)}$, $p_1$ and $\psi^{(2)}$, $p_2 = 1 - p_1$

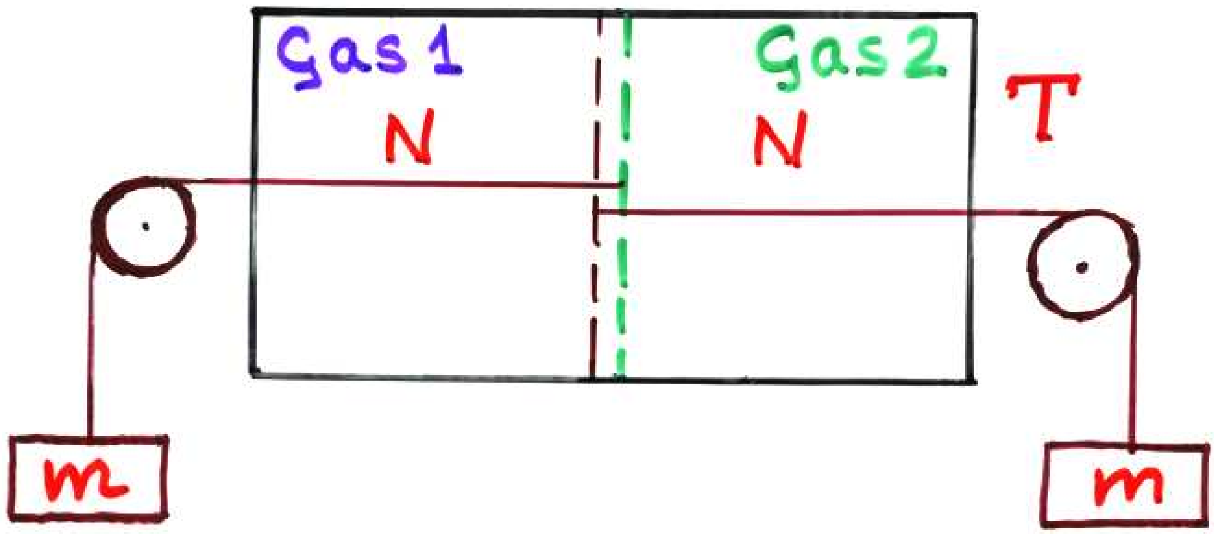



$p_1 = \frac{1}{2}$

# Thermodynamic meaning.

"Gedankenexperiment" –
– obtaining work by mixing gases




Classical case: the work is

$$R = 2NkT \ln 2 = 2NkTD$$

$\psi^{(1)}$ $\psi^{(2)}$

Quantum case: $\psi^{(1)}, \psi^{(2)}$ – two nonorthogonal states.

Is $R = 2NkTD$?

No! $R = 2NkTI$

Why? The semipermeable partitions perform a measurement over $\psi^{(1)}$ and $\psi^{(2)}$

# Indirect (POVM) measurements

Equivalent to von Neumann measurements in a tensor-product Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$

system ancilla

Can POVM measurements help?

Let $M$ be a POVM measurement

$$J = \sup_{\{M\}} I(K(M); X)$$

($K(M)$ – results of $M$)

Question: $J > I$ ?

Answer: in a finite-dim. Hilbert space $H_1$ in many cases $J > I$.

(Example: Holevo, 1973)

But: $J \leq D$ (Holevo, 1973)

(= iff all $\rho_i$ commute)

$d^2$-dim space is sufficient for an optimal measurement of an ensemble of $d$ signal states

(Davis, 1979)

Moreover:

**Th. 3.** If the signal Hilbert space $H_1$ is infinite-dimensional, then

$$J = I$$

(1981, 1994)

However, quantum channels are superadditive (Peres, Wootters, 1991)

Finally,

**Th. 4.** For measurements in the tensor-product space of long sequences of signal states:

$$\lim_{n\to\infty} \frac{J_n}{n} = D$$

(Schumacher, Westmoreland 1997)
(Holevo, 1998)

Two quantum systems, A and B.
How to characterize information that one of the systems carries about the other?

Correlation entropy:
(R. Stratonovich, 1955)

$$C(A,B) = H(A) + H(B) - H(A,B) \geq 0$$

Triangle inequalities:

$$H(A,B) + H(A) - H(B) \geq 0$$

$$H(A,B) + H(B) - H(A) \geq 0$$

Correlation entropy has <u>no</u> meaning of information.

# Information in one quantum system about another

(1999, 2002)

Consider a quantum system that consists of two subsystems $A$ and $B$. Let $\alpha$ and $\beta$ be random variables corresponding to the results of measurements of complete sets of variables in ($A$ + ancilla $A$) and ($B$ + ancilla $B$).

**Lemma.** For any $\alpha$ and $\beta$,

$$H(A,B) \le H(\alpha,\beta)$$
$$H(A) \le H(\alpha)$$
$$H(B) \le H(\beta)$$

**Proof.** Follows from Klein's lemma.

**Def. 1** The conditional entropy of subsystem $A$ conditioned by subsystem $B$ is:

$$H(A|B) = \inf_{\{\beta\}} H(A|\beta)$$

As in the classical case, conditioning can only decrease the entropy of a system.

**Th. 5.** $$H(A|B) \leq H(A)$$

**Proof.** $$\inf_{\{\beta\}} H(\alpha|\beta) \leq H(\alpha)$$

$$H(A|B) = \inf_{\{\beta\}} \sum_{b} \Pr\{\beta = b\} \inf_{\{\alpha\}} H(\alpha|\beta = b) \leq$$

$$\leq \inf_{\{\alpha\}} \inf_{\{\beta\}} H(\alpha|\beta) \leq \inf_{\{\alpha\}} H(\alpha) = H(A)$$

In classical information theory

$$H(\alpha,\beta) = H(\beta) + H(\alpha|\beta)$$

**Th. 6.** $H(A,B) \leq H(B) + H(A|B)$

Mutual information:

**Def. 2.** Information in subsystem $B$ about subsystem $A$ (and vice versa) is

$$I(A;B) = \sup_{\{\alpha\},\{\beta\}} I(\alpha;\beta)$$

Classically, $I(\alpha;\beta) = H(\alpha) - H(\alpha|\beta)$.

**Th. 7.** $I(A;B) \leq H(A) - H(A|B)$

**Proof.** It follows from the entropy defect bound that for any $\beta$

$$I(A;\beta) = \sup_{\{\alpha\}} I(\alpha;\beta) \le H(A) - H(A|\beta),$$

where equality holds iff all $\rho(A|\beta=b)$ commute and $\alpha$ corresponds to the measurement in the basis where all $\rho(A|\beta=b)$ are diagonal.

Then, $$I(A;B) = \sup_{\{\alpha\},\{\beta\}} I(\alpha;\beta) \le$$

$$\le \sup_{\{\beta\}} \left[H(A) - H(A|\beta)\right] = H(A) - \inf_{\{\beta\}} H(A|\beta)$$

$$= H(A) - H(A|B)$$

Example: superdense coding.

$A, B$ - members of an EPR pair

$C$ - classical random variable that takes on 4 values corresponding to the Bell states of $A$ and $B$.

$$I(A;C) = I(B;C) = 0$$

Moreover, $I(\alpha,\beta;C) = 1$ bit

But $I([A,B];C) = 2$ !

This is a paradox, since in classics

$$I(\alpha,\beta;C) = I(\alpha;C) + I(\beta;C|\alpha) \leq$$
$$\leq I(\alpha;C) + H(\beta|\alpha) \qquad (*)$$

In our case $I(A;C) = I(\alpha;C) = 0$, and $H(A|B) = H(\alpha|\beta) = 1$. Thus, inequality opposite to $(*)$ holds here:

$$I(A,B;C) > I(A;C) + H(B|A)$$