Entropy and Quantum Information

- In classical physics information is represented as a binary sequence.
- When we read out information that is carried by a classical system we reveal a certain bit value that even before the reading of information is performed.
- The particular sequence of bit values obtained can be considered to be physically defined by the properties of the classical system measured.
The information read is then measured by the Shannon measure of information. It can operationally be defined as the number of binary questions yes or no to determine the sequence of 0’s and 1’s.

In quantum physics information is represented by a sequence of qubits, each of which is defined in a two-dimensional Hilbert space. If we read out the information carried by the qubit, we have to project the state of the qubit onto the measurement basis $|0\rangle$, $|1\rangle$. 
The value obtained by the measurement has an element of irreducible randomness and therefore cannot be assumed to reveal the bit value or even a hidden property of the system existing before the measurement is performed.

Consider an urn filled with $N$ colored balls. There are $n_1, n_2, \ldots, n_m$ balls with various different colors: black, white, \ldots, red.

Now the urn is shaken, and we draw one after the other all balls from the urn.

To what extent can we predict the particular color sequence drawn?
If the various colors are present in equal proportions

We have no knowledge about the arrangement of the balls after shaking the urn

We are maximally uncertain about the color sequence drawn

The total number of different color sequences of $N$ balls made up of $m$ groups of black, white, . . . red balls indistinguishable within each group is

$$W = \frac{N!}{n_1! \cdot n_2! \cdots n_m!}$$
For $N$ sufficiently large:

\[ W = \frac{N!}{n_1! \cdot n_2! \cdots n_m!} \]

- Suppose now that one wishes to identify a specific color sequence of the drawn balls from the complete set of possible color sequences by asking questions “yes” or “no” can be given as an answer.
- There are $W = 2^{HN}$ possible different sequences, the minimal number of yes no questions is $HN$. 

\[
W = \frac{N!}{n_1! \cdot n_2! \cdots n_m!}
\]

\[
I = \sum_{i=1}^{n} p_i \log_2 p_i
\]

\[
p_1 = \frac{n_1}{N}, p_2 = \frac{n_2}{N}, \ldots, p_m = \frac{n_m}{N}
\]
- Equivalently, the Shannon information expressed in bits is the minimal number of yes-no questions necessary to determine which particular sequence of outcomes occurs, divided by $N$

- A particular color sequence is specified by writing down, in order, the yes’s and no’s encountering from the root to the specific leaf of the tree
If instead of balls with \( d \) colors we consider quantum systems whose individual properties are not defined before the measurements are performed.

Does the Shannon measure of information still define the information gain in the measurements appropriately?

Von Neumann entropy refers to the extension of classical entropy concepts to the field of quantum mechanics.
For any state (vector of unit length) we define the density matrix belonging to \( x \) be

\[
| x \rangle \otimes \langle x | = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \otimes \begin{pmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{pmatrix} = \begin{pmatrix} | x_1 \rangle^2 & x_1 x_2^* & \cdots & x_1 x_n^* \\ x_2 x_1^* & | x_2 \rangle^2 & \cdots & x_2 x_n^* \\ \vdots & \vdots & \ddots & \vdots \\ x_n x_1^* & x_n x_2^* & \cdots & | x_n \rangle^2 \end{pmatrix}
\]

\[ | x \rangle \otimes \langle x | = | x \rangle \langle x | \]

If \( p \) is a density matrix having \( n \) eigenvalues, then there is orthonormal set of eigenvectors of \( p \)

\[ p = \lambda_1 | x_1 \rangle \langle x_1 | + \lambda_2 | x_2 \rangle \langle x_2 | + \cdots + \lambda_n | x_n \rangle \langle x_n | \]
The spectral representation tells us that each state (density matrix) is a linear combination of one-dimensional projections which are defined as unit-length vectors.

We call the states corresponding to one-dimensional projections vector state or pure states.

Using the eigenvector basis, it is
\[ \lambda_1 + \lambda_2 + \cdots + \lambda_n = Tr(p) = 1 \]

Since \( p \) is positive \( \lambda_i = \langle x_i | px_i \rangle \geq 0 \).

Density matrix looks very much like a probability distribution of orthogonal vector states
\[ p = p_1 | x_1 \rangle \langle x_1 | + p_2 | x_2 \rangle \langle x_2 | + \cdots + p_n | x_n \rangle \langle x_n | \]
Given the density matrix $p$, von Neumann defined the entropy as

$$S(p) = -Tr(p \ln p)$$

- It is a proper extension of the Gibbs entropy (and the Shannon entropy) to the quantum case
- We note that the entropy $S(p)$ times the Boltzmann constant equals the thermodynamical or physical entropy

If the system is finite (finite dimensional matrix representation) the entropy describes the departure of our system from a pure state
- In other words, it measures the degree of mixture of our state describing a given finite system
- $S(p)$ is only zero for pure states
- $S(p)$ is maximal and equal to $\ln N$ for a maximally mixed state, $N$ being the dimension of the Hilbert space
- $S(p)$ is invariant under changes in the basis of $p$, that is $S(p)=S(UpU^*)$, with $U$ a unitary transformation

- $S(p)$ is concave, that is, given a collection of positive numbers $\lambda_i$ and density operators $p_i$ we have $S\left(\sum_{i=1}^{k}\lambda_i p_i\right) = \sum_{i=1}^{k}\lambda_i S(p_i)$
- $S(p)$ is additive
  - Given two density matrices $p_A$, $p_B$ describing different systems $A$ and $B$, then $S(p_A \otimes p_B) = S(p_A) + S(p_B)$
- Quantum information, and changes in quantum information, can be quantitatively measured by van Neumann entropy

- A Fourier transform of $f$ is a function of frequency $\nu$
  - Let be $\Delta \nu$ the frequency range
  - It can be proved that $\Delta t \Delta \nu \geq \frac{1}{4} \pi$
  - If $\Delta t$ is small, $f$ corresponds to a small interval of the whole function
    - Lower frequencies are not represented
  - For higher frequency range, a big interval of the function
    - Good frequency representation, bad time resolution, $\Delta t$ is big
One cannot know what spectral components exist at what instances of times

What one can know are the time intervals in which certain band of frequencies exist, which is a resolution problem

The problem has to do with the width of the window function that is used

- Narrow window $\Rightarrow$ good time resolution, poor frequency resolution
- Wide window $\Rightarrow$ good frequency resolution, poor time resolution

Conjugate pairs

- Another unexpected property of the nature:
- Physical variables come in „conjugate“ pairs
  - Position and momentum
  - Energy and time
- Both of which cannot be simultaneously measured with accuracy
Uncertainty principle

- If $p$ and $q$ are conjugate pairs
  - Position and momentum
  - Energy and time
- Both cannot be simultaneously measured with arbitrarily high accuracy

---

Uncertainty principle

- A Fourier transform of $f$ is a function of frequency $\nu$
- Let be $\Delta \nu$ the frequency range
  \[ \Delta t \Delta \nu \geq \frac{1}{4\pi} \]
- Planck:
  - Energy is bounded in discrete packets called quanta
  - Every energy quantum is associated with an oscillatory phenomenon, having a certain frequency
Regard observables as a self-adjoint operator

\[ A = \theta_1 E(\theta_1) + \theta_2 E(\theta_2) + \cdots + \theta_m E(\theta_m) \quad A = A^* \]

where \( E(\theta_i) \) are projections to orthogonal subspaces

The probability of observing \( \theta_i \) when the system is in state \( x_i \) is \( \langle x | E(\theta_i) x \rangle \)

We can compute the expected value of an observable \( A \) in state \( x \in H_n \)

The expected value

\[ E_x(A) = \theta_1 \langle x | E(\theta_1) x \rangle + \cdots + \theta_m \langle x | E(\theta_m) x \rangle \]
\[ E_x(A) = \langle x | \theta_1 E(\theta_1) \rangle + \cdots + \theta_m \langle x | E(\theta_m) \rangle \]
\[ E_x(A) = \langle x | Ax \rangle \]
The self-adjoint operator is defined by the spectral representation

\[ A = \theta_1 E(\theta_1) + \theta_2 E(\theta_2) + \cdots + \theta_m E(\theta_m) \quad A = A^* \]

...orthonormal basis of \( H_n \) that consists of eigenvectors of \( A \), eigenvalues

Now if we measure a particular observable \( \hat{A} \) of a quantum state \( |\psi\rangle \) and always obtain the same answer, then \( |\psi\rangle \) must be eigenstate of \( \hat{A} \) (It is sharp)

\( |\psi\rangle \) is eigenvector of \( \hat{A} \), \( a \) is real

\[ \hat{A} \cdot |\psi\rangle = a \cdot |\psi\rangle \quad a \in \mathbb{R} \]
Now let us consider measuring a different observable represented by the operator $\hat{B}$, $|\psi\rangle$ for which is not an eigenstate of $\hat{B}$.

If we measure the observable $\hat{B}$ of the quantum system in state $|\psi\rangle$ we would typically obtain different answers each time we made the measurements.

The notation $\langle \hat{A} \rangle = \langle \hat{A} \rangle_\psi = \langle \psi | A \psi \rangle$ refers to the average value that you would obtain if one puts repeatedly the system in the same state $|\psi\rangle$ and one would repeatedly measure the observable.

$$\mu = E_\psi(A) = \langle \hat{A} \rangle = \langle \hat{A} \rangle_\psi = \langle \psi | A \psi \rangle$$
The root mean square deviation of the observables is given by (standard deviation)

$$\Delta \hat{A} = \sqrt{\langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2} \quad \Delta \hat{B} = \sqrt{\langle \hat{B}^2 \rangle - \langle \hat{B} \rangle^2}$$

$$\text{Var}_\psi(A) = E_\psi((A - \mu)^2) = \| (A - \mu) \psi \|^2$$

The quantities $\Delta \hat{A}$ and $\Delta \hat{B}$ quantify the uncertainty with which the values of the observables are known.

As $|\psi\rangle$ is an eigenstate of the observable $\hat{A}$, then because the same answer is obtained each time the observable $\hat{A}$ is measured $\Delta \hat{A} = 0$.

However as $|\psi\rangle$ is not the eigenstate of $\hat{B}$, $\Delta \hat{B} \neq 0$.
The question arise as to what happens if we try to measure both observables. The answer depends on the order in which we make the measurements. If we measured observable $\hat{A}$ first, then the act of measurement would not perturb the state since is already eigenstate of $\hat{A}$. If we measured the observable $\hat{B}$ first, then as $|\psi\rangle$ is not eigenstate of $\hat{B}$, the act of measuring $|\psi\rangle$ will perturb the state of the system. We look at the difference between measurements performed in each order. To do it we construct the “commutator” operator $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$.
It can be shown that if two observables are measured simultaneously, the uncertainty in their joint values must always obey the inequality (Heisenberg Uncertainty):

\[
\Delta \hat{A} \Delta \hat{B} \geq \frac{1}{2} \left| \left[ \hat{A}, \hat{B} \right] \right|
\]

\[
\Delta \hat{A} \Delta \hat{B} \geq \frac{1}{2} \left| \left[ \hat{A}, \hat{B} \right] \right| \left( \hat{A} \hat{B} \right)
\]

\[
\text{Var}_\psi(A) \text{Var}_\psi(B) \geq \frac{1}{4} \left| \left( \hat{A}, \hat{B} \right) \right|^2
\]

The inequality follows from Cauchy-Schwarz inequality:

\[
\|Ax\|^2 \|Bx\|^2 \geq |\langle Bx | Ax \rangle|^2 \geq \left| \text{Im} \{\langle Bx | Ax \rangle \} \right|^2
\]

\[
= \frac{1}{4} |2 \text{Im} \{\langle Bx | Ax \rangle \}|^2
\]

\[
= \frac{1}{4} |\langle Bx | Ax \rangle - \langle Bx | Ax \rangle|^2
\]

\[
= \frac{1}{4} |\langle Bx | Ax \rangle - \langle Ax | Bx \rangle|^2
\]

\[
= \frac{1}{4} |\langle ABx | x \rangle - \langle BAx | x \rangle|^2
\]

\[
= \frac{1}{4} |(AB - BA)x | x \rangle|^2
\]
Gives one form of the Robertson-Schrödinger relation:

\[
\frac{1}{4} |\langle [A,B|x] \rangle|^2 \leq \|Ax\|^2 \|Bx\|^2.
\]

\[
\Delta \hat{A} \Delta \hat{B} \geq \frac{1}{2} |\langle [\hat{A},\hat{B}] \rangle|
\]

\[
\Delta \hat{A} \Delta \hat{B} \geq \frac{1}{2} \langle x | [\hat{A},\hat{B}] x \rangle
\]

\[
Var_\psi(A)Var_\psi(B) \geq \frac{1}{4} \langle x | [\hat{A},\hat{B}] x \rangle^2
\]
Polarization

- Reduce a message to a sequence of bits and then create a stream of photons placed in a certain quantum state corresponding to these bits.
- The photon property we are interested in is called polarization.
It is possible to create photons with its electric fields oscillating in any desired plane.

- Polarized photons whose electric fields oscillate in a plane either 0° or 90° to some line “rectilinear”
- Polarized photons whose electric fields oscillate in a plane either 45° or 135° “diagonal”

Binary 0 represented by 0° and 45°
Binary 1 represented by 90° and 135°

**Measuring the polarization**

- It is necessary to measure the polarization
- We perform the measurements with calcium carbonate crystal
- It has the property of bifrigece
  - Electrons are not bound with equal strength
- Photon passing through the crystal will feel a different electromagnetic force depending on the orientation of its electric field
- Polarized photons whose electric fields oscillate in a plane either 0° or 90° to some line “rectilinear”
- If the polarization axis is aligned so that vertical 90° polarized photons pass through it
- A photon with horizontal polarization 0° will also pass through the crystal but it will emerge shifted
Polarized photons whose electric fields oscillate in a plane either 45° or 135° “diagonal”
- Uncertainty principle says that the polarizer provides no information about the original polarization
- The calcite crystal has to be aligned in diagonal
- It is impossible to measure both rectilinear and diagonal polarization exactly
- Any attempt to measure rectilinear polarization perturbs the diagonal polarization and vice versa
  - Why? The commutator describing both measurements does not vanish

We can design a secret protocol for exchanging a secret key
- It can be guaranteed, that nobody interferes with the message
- An eavesdropping can be detected
Alice makes to encode bits as polarized photons (first row)

Then each bit she chose to encode it either rectilinear (+) or in the diagonal polarization (x), the choice is made randomly

Alice then sends the photons she created to Bob (third row)

Bob receipts the photons (first row)

He chooses a polarizer orientation with which he measures the direction of polarization (randomly)

He reconstructs the bits (third row)
- Detection of eavesdropping
- Alice and Bob compare a subset of bits which were generated with same polarization
- The must be equal!

If they are not equal, it means that some one other has measured them, there was an eavesdropping!
Once Alice and Bob decided that the channel is secure, Alice tells Bob what polarization she used for each of her bits.

- Bob compares his polarization and read the bits, tells Alice about his polarization.
- These bits are now only known to Bob and Alice!
IBM’s Watson Research Center 1989