The entropy of systems

J. L. Manríquez-Zepeda, J. Rueda-Paz, P. D. Filio-Aguilar, and L. López-García

Centro Universitario UAEM Valle de Chalco 3, Universidad Autónoma del Estado de México,
Edo. Méx. 56615, México.

The entropy is studied from the perspective of several formalisms. The origin of the concept in the classical thermodynamics is reviewed. It also is developed a step-by-step-clearly demonstration about it. After that, this formula is connected with the entropy of Shannon. Elementary concepts of the quantum mechanics are explained to demonstrate the von Neumann entropy formula. From the perspective of the information theory, the entropy of an optimal code in communications is the same that the expectation value of bits needed to transmit a message through a channel. In this way, systems and languages can be studied using the same concept of entropy. It is proposed the entropy concept as a measure of the variability of the states inside a system given a set of rules that operate inside it for a while. It is stressed the importance of the teaching of the entropy because this is a useful tool in the development of new technologies, for example the quantum communications.

Keywords: Von Neumann entropy; shannon entropy; information theory; history of entropy; systems and states; teaching entropy.

1. Introduction

The entropy is a very interesting and useful concept in sciences. In physic, entropy has been used to quantify the available energy that can be transformed into work [1]. Information theory defines the entropy as the optimal coding to compress or to transmit data [2]. Also, in Relativistic Quantum Information (QRI) the loss of entanglement in multipartite quantum systems is quantified by the entropy [3,4]. The concept reveals us the sharpness of the thinkers of nineteenth century that didn't give up before the big mystery of the piston that moving without apparent reason when they stop to supply heat on it. One of them, Boltzmann taken ideas of macroscopic world to the microscopic world and then make use of the calculus theory to infer the right distribution that were compatible with macroscopic measures [5,6]. Some time later, Shannon developed the entropy concept as a measure of the uncertainty of a source of information [7], becoming this concept in the cornerstone of the computer science.

However, in despite the importance of this concept, there exist misunderstandings. From the perspective of thermodynamics; entropy is related with the transformation of energy into work; from the perspective of statistical physic is a probability distribution of microstates given at thermodynamical equilibrium; in the educational strategy to teaching entropy, it consists in a measure of disorder; from the perspective of information theory is the mean of bits needed to transmit or storing a message [8,9]. Besides, entropy also has interpretations in cosmology, mathematical physics, economics, complexity theory, evolutionary biology, psychology, and sociology. It has been considered as if it were almighty over everything related with randomness, evolution and as an arrow of time. This approach is about thermal randomness, where, it must be considered the equilibrium of physical laws and probability distribution over systems [10]. In this way, the entropy concept is useful to understand important ideas in

quantum information theory and quantum computing. As entropy is related with all previous examples, this cannot be reduced to a formula without context. Teachers may explain to students all about of its history and development for helping them to build a complete understanding of the entropy in the new sciences [11].

In this work, the history about origins of the concept is recounted in Secs. 2 and 3. After that, it is derived an analytical expression of the entropy in Sec. 4. Then, an operative formula to compute entropy in computer sciences is explained in Sec. 5. In Sec. 7 is explained in a friendly way the elementary concepts of quantum mechanics in order to bring from the thermodynamic entropy formula in physics to the von Neumann entropy in quantum information theory. Later, in Sec. 8 is talked about the entanglement and how it can be measured by entropy of von Neumann [12]. Finally, at the end of this work a definition of entropy is proposed and some observations are made.

2. Origins of entropy in thermodynamics

In the historical framework of the development of the Industrial Revolution and the development of the steam engines, it was investigated how to optimize the output work by machines. Nicolas Leonard Sadi Carnot developed a thought experiment that involved a piston moved by the expansion and contraction of an ideal gas inside it [13].

The Carnot machine is a hermetically closed system with an ideal gas inside it. When piston at position C_1 (see Fig. 1), the ideal gas is in thermodynamic equilibrium. Then, the ideal gas is heated with a quantity Q_{in}, which cause this expands its volume as the piston upward up to the position C_2. After that, the ideal gas is stopped to heating. But, the piston continues get upward for a while up to reach the position C_3, then, it is said the system has reached the thermodynamic
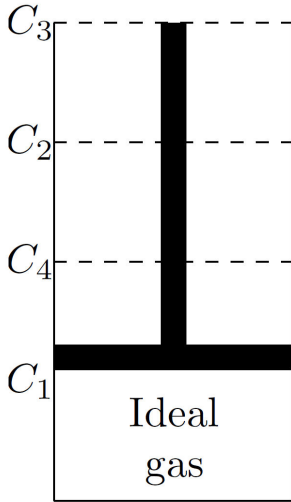
FIGURE 1. Hermetically sealed piston. Picture adapted from figure 5 in chapter 3 of the Ref. [13].

FIGURE 2. Phase diagram of a Carnot cycle. T_A and T_B are isothermal curves that cross through the states C_1, C_2 and C_4, C_3 of the ideal gas system, respectively. Picture adapted from figure 5 in chapter 3 of Ref. [13].

equilibrium. Later, it is extracted the quantity of heat Q_{out} of the ideal gas putting this one in contact with a cold source. As before, the extraction of heat is stopped at position C_4 (see Fig. 1), but again, the ideal gas continues contracting until the piston down warding to the position C_1 and it is said that the ideal gas has reached the thermodynamic equilibrium.

From Carnot's times, this experiment was observed as a system where pressure, volume and temperature change in such a way that each possible state fulfil that energy of the system (E) is constant. For better to study this one, phase diagrams like Fig. 2 were proposed:

Where the curve that joint C_1 to C_2 corresponds to the isothermal T_A and the curve that joint C_4 to C_3 corresponds to the isothermal T_B. Displacement of the piston is a net conversion of Q_{in} into work. The ideal gas system changes of (P_1, V_1, T_1) state to (P_2, V_2, T_2) state. On the other side, curves C_2 to C_3 and C_4 to C_1 are adiabatic curves. These curves show the behaviour of a system that does not exchange heat with the environment. This time, displacement of the piston must be paid by the internal energy of the ideal gas system.

The mystery in this thought experiment is that after stops heating the ideal gas, the volume must stop to increases in the context where the cause of expansion it supposed to be the heating of the ideal gas. But instead, counterintuitively the ideal gas even expands and push the piston a little more. During the isothermal curves, Q_{in} and Q_{out} are converted into work. But, unlike the isothermal process, while the adiabatic process occurs, something happens inside the ideal gas that causes the expansion of volume in C_2 to C_3 and the contraction in C_4 to C_1. Namely, Q_{in} changes the state of the ideal gas from (P_1, V_1, T_1) to (P_2, V_2, T_2), and later, still under the effects of Q_{in}, the system changes internally in such a way that after a time the ideal gas reaches the state (P_3, V_3, T_3). In the same way, Q_{out} yields the change of states from (P_3, V_3, T_3) to (P_4, V_4, T_4), and then, internally to (P_1, V_1, T_1). As the thermodynamic equilibrium means
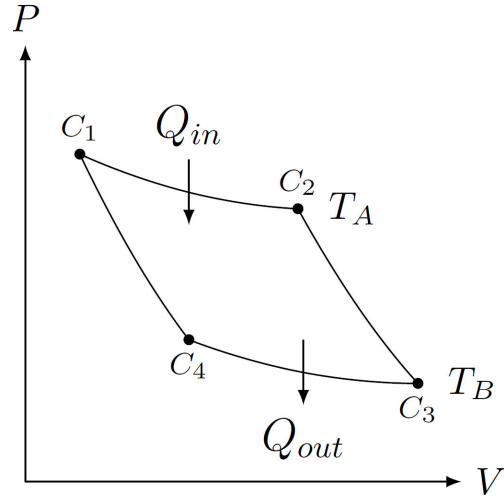
that all internal forces inside de ideal gas reached the equilibrium among them, these evidences suggested the existence of a variable inside the system that depends only on the internal configurations of the ideal gas [13]. This variable was called as "entropy" (S) since $\tau\rho o\pi\dot{\eta}$ means "turn" in Greek. In this work, it is proposed that the meaning that the nineteenth century scientists gave it to the entropy was to refer to an internal transformation, as if something were moving and rotating inside the system. Operationally, this entropy is computed by

$$\Delta S = \frac{\Delta Q}{T}, \tag{1}$$

because the change of the hypothetical internal configurations in the ideal gas may be associated to the supplied heat along an isothermal curve.

3. Statistical point of view of the entropy

Soon, it was necessary a better understanding of the entropy concept after both microscopic and statistical new ideas in physics arrived in the twentieth century. Ludwig Boltzmann proposed the thought experiment in which there are N enclosed molecules of an ideal gas inside a rigid and hermetically sealed container where it is taken a snapshot in such a way that is captured all the dynamic of the system in just a scene [14]. Later, it is counted the molecules with energy ε_1, molecules with energy ε_2 and so on, in such a manner that it is said that at certain time t

There are n_1 molecules with energy ε_1
There are n_2 molecules with energy ε_2
$$\vdots$$
There are n_j molecules with energy ε_j
$$\vdots$$
There are n_k molecules with energy ε_k

The aim was to take the macroscopic concepts of world to the microscopic world. It is going to think that every set of molecules n_i with energy ε_i is enclosed in a rigid and hermetically sealed cell of volume dv. In this way, each set of n_i molecules is called a microstate. Then, for a certain constant energy E corresponds

$$W_{\{n_i\}} = \frac{N!}{\prod\limits_{i=1}^{k} n_i!}, \tag{2}$$

of microstates. There are k microstates for a given energy E that must fulfill the restrictions that

$$\sum_{i=1}^{k} n_i = N,$$
$$\sum_{i=1}^{k} \varepsilon_i n_i = E. \tag{3}$$

These conditions must be consistent with the values of pressure (P), volume (V) and temperature (T) of the macroscopic system.

The total number of microstates given a certain energy E is called a macrostate. Some examples are given for illustrative purposes.

Suppose $N = 3$ molecules, then $N = 6!$, whether E has a spectrum of $\{\varepsilon_1\}$ possible energies, then

$$W_1 = \frac{3!}{3!} = 1, \tag{4}$$

whether E were $\{\varepsilon_1, \varepsilon_2\}$, then energy distributes over the 3 molecules in an indistinguishable way as the ensemble $\{(n_1 = 2, \varepsilon_1), (n_2 = 1, \varepsilon_2)\}$ and the number of microstates is

$$W_2 = \frac{3!}{2!1!} = 3. \tag{5}$$

In the next case E has the spectrum $\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$, then energy distributes over the 3 molecules as the ensemble $\{(n_1 = 1, \varepsilon_1), (n_2 = 1, \varepsilon_2), (n_3 = 1, \varepsilon_3)\}$ and the number of microstates this time is

$$W_3 = \frac{3!}{1!1!1!} = 6. \tag{6}$$

Now, the question Boltzmann and contemporaries wanted to answer is what happens with energy E whether it can distribute in any possible way over the N molecules? namely, this distribution could be $\{E/N, E/N, \ldots, E/N\}$ or $\{E/N, 8E/N, E/N\}$ or anything. Due that, many macrostates are compatible with E, though not compatible with the restrictions Eq. (3). The total number of microstates compatible with energy E is

$$\Omega(E) = \sum_{\{\varepsilon\}^*} W_{\{n_i\}}, \tag{7}$$

where, Ω is a number that is function only of a given E, while $\{\varepsilon\}^*$ refers to all possible distributions of the energy E over the molecules of the ideal gas. It is necessary to stop here and see that the concepts here consider a system in thermodynamic equilibrium, in such a way that, there must be an overwhelmingly greater distribution compatible with the (P, V, T) state given the energy E. This idea is expressed in the following equation.

$$\Omega(E, P, V, T) = t(n), \tag{8}$$

where, it is defined $t(n)$ as a function of the distributions of the energy E, because $n = n(\{\varepsilon_i\})$. Boltzmann solves the problem using tools of calculus of university, he looks for critical points over $t(n)$ [14], by doing

$$\nabla t(n) = 0, \tag{9}$$

whereas, in the thermodynamic equilibrium the system does not change the number of particles and energy, thus

$$\nabla N = 0,$$
$$\nabla E = 0, \tag{10}$$

after that, in order to find the critical points according to the undetermined coefficients method of Joseph-Louis Lagrange, [15]

$$\nabla t(n) = \alpha \nabla N + \beta \nabla U. \tag{11}$$

At this point, there is a system of i first-degree differential equations, one for each set of n_i molecules with energy ε_i.

$$\left(\frac{\partial t(n)}{\partial n_i} - \alpha - \beta \varepsilon_i \right) dn_i = 0. \tag{12}$$

To solve Eq. (12) it is worked with the logarithm of $t(n)$ manipulating it till we achieve a relation that can help us.

$$\log(t(n)) = \log\left(\frac{N!}{\prod\limits_i n_i!} \right)$$
$$= \log(N!) - \log\left(\prod_i n_i! \right)$$
$$= \log(N!) - \sum_i \log(n_i!)$$
$$= N \log(N) - N - \sum_i \left(n_i \log(n_i) - n_i \right),$$
$$\log(t(n)) = N \log(N) - \sum_i n_i \log(n_i), \tag{13}$$

where, it has been used two times the Stirling approximation of $\log(N!)$ and some logarithms laws. Later, it is derived

with respect to some arbitrary n_j,

$$\frac{\partial \log(t(n))}{\partial n_j} = \frac{\partial}{\partial n_j}\left(N\log(N) - \sum_i n_i \log(n_i)\right),$$

$$\frac{d}{dt}(\log(t))\frac{\partial t(n)}{\partial n_j} = -\frac{\partial}{\partial n_j}\left(\sum_i n_i \log(n_i)\right),$$

$$\frac{1}{t(n)}\frac{\partial t(n)}{\partial n_j} = -\sum_i \frac{\partial}{\partial n_j}\left(n_i \log(n_i)\right)$$

$$= -\sum_i \left(\delta_{ij}\log(n_i) + n_i\frac{1}{n_i}\delta_{ij}\right),$$

$$\frac{\partial t(n)}{\partial n_j} = -t(n)\left(\log(n_j) - 1\right), \tag{14}$$

then, this result is replaced in Eq. (12) to obtain

$$-t(n_j)\left(\log(n_j) - 1\right) - \alpha - \beta\varepsilon_j = 0,$$

$$\log(n_j) = -\frac{\alpha}{t(n_j)} - \frac{\beta}{t(n_j)}\varepsilon_j + 1, \tag{15}$$

where, n_j is an extremal of $t(n)$ under the restrictions imposed by Eq. (3). This reasoning is what claims the second law of the thermodynamics about the thermodynamic equilibrium relating to a maximum of the entropy function at a given internal energy [16]. Remember that n_j is one of the distributions inside the domain of possible distributions given E. For that, n_j is the distribution of energies in the ensemble $\{(n_j, \varepsilon_j)\}$ such that, it does compatible with the (P, V, T) state given the constraints in Eqs. (3). Then, as $t(n_j)$ is the number of microstates in the ideal gas, it can be absorbed into the α and β constants, namely

$$\log(n_j) = -\alpha - \beta\varepsilon_j, \tag{16}$$

where, it can be obtained directly

$$n_i = A\,\mathsf{e}^{-\beta\varepsilon_i}, \tag{17}$$

where, it has been changed j by i, and $A = \mathsf{e}^{-\alpha}$. The distribution n_i is called as the Maxwell-Boltzmann distribution [17]. In this way,

$$\Omega(n_i) = \frac{N!}{\prod_i n_i!}, \tag{18}$$

where, Ω is a function of the ensemble $\{(n_i, \varepsilon_i)\}$. As the optimization only outputs one critical distribution n_i, it may be asserted that does not exist more than one ensemble compatible with the (P, V, T) state given the energy E, when the thermodynamic equilibrium has been reached.

4. Analytical expression of the entropy

At this moment yet, a complete understanding of entropy was unknown, but it was supposed that it was related with the number of microstates in thermodynamic equilibrium of the ideal gas. In symbols

$$S = f(\Omega), \tag{19}$$

now, for to infer an analytical expression, from perspective of thermodynamics, it may be added two thermodynamic systems with macroscopic entropies S_1 and S_2, which sums to

$$S = S_1 + S_2, \tag{20}$$

On the other hand, because of multiplicative nature of the combinatorial properties of Ω, the sum over the Ω_1 and Ω_2 macrostates is

$$\Omega = \Omega_1 \cdot \Omega_2, \tag{21}$$

therefore, applying f over Ω

$$f(\Omega) = f(\Omega_1) + f(\Omega_2), \tag{22}$$

where, it is derived with respect Ω_1 for getting

$$\frac{df(\Omega)}{d\Omega_1} = \frac{df(\Omega_1)}{d\Omega_1} + \overset{0}{\cancel{\frac{df(\Omega_2)}{d\Omega_1}}}, \tag{23}$$

then, it is used the chain rule over $f(\Omega)=f(\Omega_1\Omega_2)$.

$$\frac{df(\Omega)}{d\Omega}\frac{d(\Omega_1\Omega_2)}{d\Omega_1} = \frac{df(\Omega)}{d\Omega}\Omega_2 = \frac{df(\Omega_1)}{d\Omega_1}, \tag{24}$$

after that, it is multiplied side by side by Ω_1

$$\Omega\frac{df(\Omega)}{d\Omega} = \Omega_1\frac{df(\Omega_1)}{d\Omega_1}, \tag{25}$$

where, it was used $\Omega = \Omega_1\Omega_2$. Moreover, note that

$$\frac{d}{dx}\left(\log(x)\right) = \frac{1}{x},$$

$$d\left(\log(x)\right) = \frac{dx}{x}, \tag{26}$$

$$\frac{1}{d\left(\log(x)\right)} = \frac{x}{dx},$$

so that, as long as there is a variable over its differential, it can be replaced by the reciprocal of the differential of its logarithm. Using this one,

$$\frac{1}{d\left(\log(\Omega)\right)}df(\Omega) = \frac{1}{d(\log(\Omega_1))}df(\Omega_1). \tag{27}$$

Then, it is derived again, this time with respect to Ω_2.

$$\frac{d}{d\Omega_2}\left(\frac{df(\Omega)}{d(\log(\Omega))}\right) = \frac{d}{d\Omega_2}\left(\frac{df(\Omega_1)}{d(\log(\Omega_1))}\right) = 0, \tag{28}$$

i.e.,

$$\frac{d\,f(\Omega)}{d(\log(\Omega))} = k, \tag{29}$$

where k is a constant. This equation can be integrated immediately.

$$f(\Omega) = k \, \log(\Omega) + \Omega_0, \tag{30}$$

finally, using the symmetry of the problem, it is compared with Eq. (22) and concluded that $\Omega_0 = 0$.

$$\begin{aligned} f(\Omega) &= f(\Omega_1 \Omega_2) \\ &= k \, \log(\Omega_1 \Omega_2) + \Omega_0 \\ &= k \, \log(\Omega1) + k \, \log(\Omega_2) + \Omega_0 \\ &= f(\Omega_1) + f(\Omega_2) + \Omega_0, \end{aligned} \tag{31}$$

where it is observed that Ω_0 must be zero to be agree with Eq. (20). Then, the analytical expression for the entropy stands

$$S = k \, \log \, (\Omega). \tag{32}$$

What was achieved with this exploration was to know that entropy is proportional to the logarithm of the number of microstates of a systems when this have reached the thermodynamic equilibrium.

5. Entropy in the information theory

Information in the science of the information has the meaning of the astonishment of get a particular message from a set of possible messages. Namely, it is said that whether a highly unlikely event occurs, then it has been gained more information than when occurs an ordinary event []. An operative expression of this idea is

$$I = -\log_2\big(p(x)\big). \tag{33}$$

For example, let's suppose a tombola with 2 red balls inside it, then whether it is chosen a ball from inside the tombola, the information gained is

$$I = -\log_2\left(\frac{2 \text{ red balls}}{2 \text{ total balls}}\right) = -\log_2(1) = 0, \tag{34}$$

that is, it hasn't been gained new information from the experiment because the color of the next ball was known in advance. Even more, there isn't exist a way to use this system to contrast data, therefore it cannot be made a code.

Now suppose there is both a red and a blue ball inside the tombola. This time, when it is chosen a ball from inside the tombola,

$$I = -\log_2\left(\frac{1 \text{ red ball}}{2 \text{ total balls}}\right) = 1. \tag{35}$$

What this result means is that, this time it has been gained 1 bit of information. Information that can be useful to know the color of the next ball inside the tombola prior to drawn.

In this order of ideas, observe that, an expectation value of expression I leads to equation

$$E[I] = \sum_i p(i) I(n_i) = -\sum_i p(i) \log\left(\frac{n_i}{N}\right), \tag{36}$$

where, $p_i = n_i/N$ is the probability of an event i. *I.e.* it is asked about a mean value of information in a long sample. A clearly example is to guess the second letter of any random word in spanish language. Where the variability at first is

$$\begin{aligned} H_2 = -p_a \log_2 p_a - p_b \log_2 p_b - \ldots \\ - p_z \log_2 p_z = 4.04 \text{ bits}, \end{aligned} \tag{37}$$

but, if a hint is that the first letter of such word will be a zeta, then the second letter of a spanish word is constrained to be "a", "o" or "u"(since words with "ze" and "zi" at first are very unlikely).

$$\begin{aligned} H_1 = -p_a \log_2 p_a - p_o \log_2 p_o \\ - p_u \log_2 p_u = 0.91 \text{ bits}, \end{aligned} \tag{38}$$

afterwards,

$$I = -(H_2 - H_1) = 3.13 \text{ bits}, \tag{39}$$

where, it is said that it is the information gained after using the hint, because now, it is more likely to guess the word.

Because of the entropy stands a degree of the variability of the microstates of a system, the information may be defined as

$$I = -(H_2 - H_1), \tag{40}$$

where, H_2 and H_1 correspond to the entropy of a system in two different states. Namely, it may be known the variability of the states of a system, but not the exact configuration of this [19].

From this perspective, entropy of the information or entropy of Shannon [20,21] has the explicit formula

$$H = -\sum_x p(x) \log_2\big(p(x)\big), \tag{41}$$

which, can be derived from Eq. (32) as it will be demonstrated.

In order to make a more comprehensive explanation, suppose a language, where the words can be encoded with just 3 bits and exhibits the probability distribution $p(x)$ of Table I.

TABLE I. Coding of a source of information. Data obtained from example 1.1.2 in Ref. [20].

word	codeword	$p(x)$
mi	000	1/2
ma	001	1/4
me	010	1/8
mo	011	1/16
mu	100	1/64
pi	101	1/64
pa	110	1/64
po	111	1/64

TABLE II. Optimal coding of a source of information.

word	codeword	$p(x)$
mi	0	1/2
ma	10	1/4
me	110	1/8
mo	1110	1/16
mu	111100	1/64
pi	111101	1/64
pa	111110	1/64
po	111111	1/64

But, it is required to transmit encoded messages with the most optimal dictionary through a channel of communication in order to saving costs and resources. It is possible to create the next code to fulfil the goal (see Table II).

Where, the shortest element 0 will be related to the most used word "*mi*" because this helps to compact the messages. Other words of the language are optimized in the same way to build this optimal code. The expectation value of bits required to transmit messages with this code is

$$E[\text{bits}] = \frac{1}{2}(1 \text{ bit}) + \frac{1}{4}(2 \text{ bits}) + \frac{1}{8}(3 \text{ bits})$$
$$+ \frac{1}{16}(4 \text{ bits}) + \frac{4}{64}(6 \text{ bits}) = 2 \text{ bits} \qquad (42)$$

Note that $1 = \log_2(2)$, $2 = \log_2(4)$ and so on with others, *i.e*

$$E[\text{bits}] = \frac{1}{2}\log_2(2) + \frac{1}{4}\log_2(4) + \frac{1}{8}\log_2(8)$$
$$+ \frac{1}{16}\log_2(16) + \frac{4}{64}\log_2(64) = 2 \text{ bits}, \qquad (43)$$

that is the same that

$$E[\text{bits}] = -\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{4}\log_2\left(\frac{1}{4}\right) - \frac{1}{8}\log_2\left(\frac{1}{8}\right)$$
$$- \frac{1}{16}\log_2\left(\frac{1}{16}\right) - \frac{4}{64}\log_2\left(\frac{1}{64}\right) = 2 \text{ bits}, \quad (44)$$

where the general form of this summation is the expectation value

$$E[\text{bits}] = -\sum_x p(x) \log\left(p(x)\right), \qquad (45)$$

where x is the length of bits in codewords. Surprisingly, it can be obtained a similar expression by expanding the Ω term of the entropy formula S using the macrostate of the distribution (n_i, ε_i) of molecules compatible with a system in its thermodynamic equilibrium state (see Eq. (18)),

$$S = k \, \log(\Omega) = k \, \log\left(\frac{N!}{\prod_i n_i!}\right)$$

$$= -k\sum_i \left(n_i \, \log(n_i) - n_i \, \log(N)\right)$$

$$= -k\sum_i n_i \, \log\left(\frac{n_i}{N}\right), \qquad (46)$$

where $k = 1/N$ fulfil a normalization constraint over the ensemble (n_i, ε_i), which leads to the Shannon entropy formula Eq. (41).

$$S = \sum_i p(x_i) \, \log\left(p(x_i)\right). \qquad (47)$$

Therefore, entropy in the information theory is the expectation value of bits needed to code in an optimal way a language.

$$H(n_i) = E[\log(p_i)]. \qquad (48)$$

This result is formally established in the noiseless coding theorem stated by Claude Shannon which is a cornerstone in the computer and information sciences [21,22].

6. What is entropy then?

To answer this question, it can be asked about the relation among entropy of Boltzmann S and entropy of Shannon H. Both seem to be the expectation value of the logarithm of a probability density. Instead, S is the expectation value of the ensemble (n_i, ε_i).

$$S(\Omega) = E[(n_i, \varepsilon_i)], \qquad (49)$$

when a thermodynamic system has reached the equilibrium of forces inside it. In some references associate this scenario to the Heat Death of systems, when all interactions are in equilibrium and thus the state of the system does not change anymore [23]. Entropy raises toward a maximum in this configuration because a high homogeneity among its microstates grows over time. It cannot be used this system to make a code because of this is a very plain source of information, and a code needs contrast among its states. By other hand, a language system has its own rules that tends to yield a different distribution of its elements, hence, it is a different entropy than in the thermodynamic systems and H does not

raise to maximum [24,25]. This entropy is a measure of the variability of elements in a language that is available to code data [26].

Other topic associated to the entropy is related to the measure of disorder. Let's see a friendly example to understand it: Think in your work office. When the cleaning man goes to tidy up your office, he puts all thing in the proper places. For instance, he saves the pens in the drawer of the desk and the notebook in the shelf. The cleaning man tells you that the office is ready. Then, when you go to work, you always leave the notebook and pens over the desk, next the computer. If someone asks the cleaning man, where is it the notebook? He would say that is over the shelf. But, the actual situation is that the notebook and pens are over the desk, next the computer. Whatever the reason that changes a system, new rules tend to set a new order. Then, entropy can be used to contrast the degree of variability of the states in any system, where the order inside the system is a concept that depends of the rules. From the perspective of the cleaning man, entropy is equal 0 stands the office is tidy, while entropy equal 1 means the office is messy. But, in contrast, for you, the entropy is equal 0 when you know where is every thing, even if it implies the office is messy.

As a reconciling concept it is proposed that the entropy is a measure of the variability of the states or the element of a system given the rules that operate over it.

7. The von Neumann entropy

The entropy for quantum systems is known as the von Neumann entropy and its expression is

$$H = -Tr\left(\hat{\rho}\,\log_2\left(\hat{\rho}\right)\right),\tag{50}$$

where ρ stands for the density matrix of a quantum system. In order to explain this formula, some concepts in quantum mechanics are introduced.

7.1. Quantum systems

Think in a marble machine $\mathcal{M}$ (see figure Fig. 3) with six boxes inside it. The boxes are connected by pipes. The arrows indicate the direction that marbles take whether let them pass through the pipes.
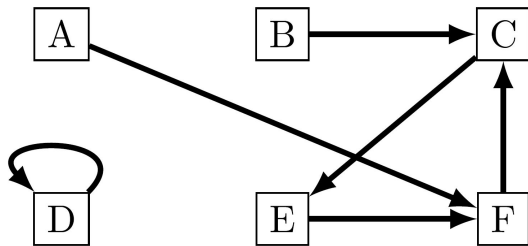
FIGURE 3. Scheme of machine $\mathcal{M}$. Arrows represent direction that marbles will take whether let them pass through the pipes. Picture adapted from example 3.1.1 in Ref. [27].
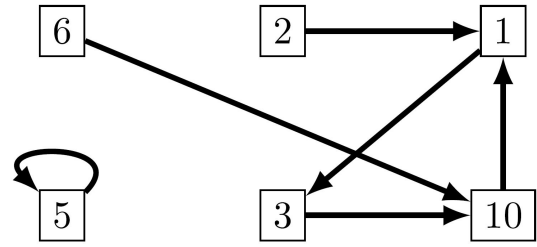
FIGURE 4. Scheme of machine $\mathcal{M}$ where it is filled boxes with 27 marbles. Picture adapted from example 3.1.1 in Ref. [27].
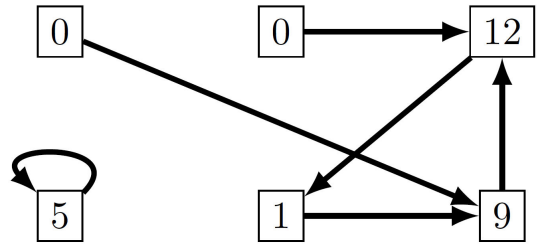
FIGURE 5. Scheme of machine $\mathcal{M}$ where it was allowed that the marbles pass through the pipes. Picture adapted from example 3.1.3 in Ref. [27].

It may be filled $\mathcal{M}$ with 27 marbles as in Fig. 4. This configuration is called C_1.

Then, it is allowed the marbles go through pipes. $\mathcal{M}$ changes to the configuration called C_2 (see Fig. 5).

Namely $\mathcal{M}$ turns the configuration C_1 into C_2.

$$\begin{pmatrix} 6 \\ 2 \\ 1 \\ 5 \\ 3 \\ 10 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 \\ 0 \\ 12 \\ 5 \\ 1 \\ 9 \end{pmatrix}.\tag{51}$$

It is possible to describing the machine $\mathcal{M}$ by the matrix $M_{i,j}$.

$$M_{i,j} = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \end{array} \begin{array}{cccccc} A & B & C & D & E & F \\ \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

this is accomplished by putting a 1 in the entry (i,j) of $M_{i,j}$ provided a box i is connected by a pipe with a box j or with itself. For example, entry $(3,1)$ has a 1 because box B has a connection toward C.

After that, is possible to retrieve C_2 doing

$$M_{i,j}C_1 = C_2,\tag{52}$$

or, in matrix form,

$$
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0
\end{pmatrix}
\begin{pmatrix}
6 \\ 2 \\ 1 \\ 5 \\ 3 \\ 10
\end{pmatrix}
=
\begin{pmatrix}
0 \\ 0 \\ 12 \\ 5 \\ 1 \\ 9
\end{pmatrix}.
\tag{53}
$$

Above, it is showed how to describe a discrete system by mean of the matrix formalism. Quantum systems also can be described in the same fashion. Where, the behaviour of the quantum systems are modeled by matrices called operators, for instance, it is defined the operator $\hat{A}$. That operator acts over a column vector called state vector (for instance $|\psi_i\rangle$), which turns the state vector $|\psi_i\rangle$ into a new state vector $|\psi_j\rangle$ [27]. Thus, it can be written as

$$
\hat{A}|\psi_i\rangle = |\psi_j\rangle,
\tag{54}
$$

where $\hat{A}$ transforms the configuration $|\psi_i\rangle$ into the configuration $|\psi_j\rangle$ in the same way that $M_{i,j}$ transforms C_1 into C_2.

7.2. Eigenvalues and eigenstates

For a simple explanation of all the context in quantum mechanics the quantum property known as spin can be taken as an example, although it is worth mentioning that there exists many more. The spin is an angular momentum property of electrons with two possible configurations, spin $\uparrow$ and spin $\downarrow$. It can be viewed in a Stern-Gerlach experiment when a beam of electrons passing through a non-homogeneous electromagnetic field is deflected into two directions. (See scheme of Fig. 6).

The measurements of a quantum system can be used to build an operator that represents this property, by putting these along the diagonal of a matrix. So, in this manner, this operator is called observable. For instance, with the discrete outputs $\uparrow$ and $\downarrow$ in the Stern-Gerlach experiment, the observable operator is

$$
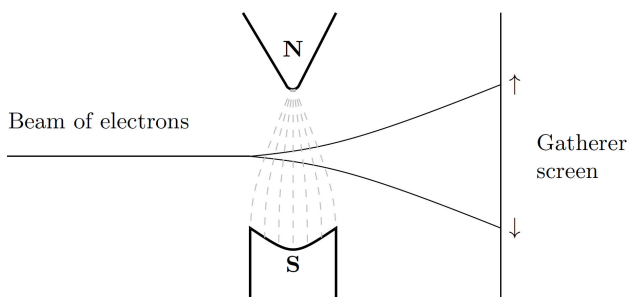\hat{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},
\tag{55}
$$

FIGURE 6. Scheme about the experiment of Stern-Gerlach. Picture adapted from similar in section 14 of Armando Martínez Tellez's blog: la-mecanica-cuantica.blogspot.com

where it is encoded $\uparrow$ and $\downarrow$ as 1 and -1 respectively. This measures are called the eigenvalues of the operator $\hat{Z}$. The eigenvalues $\{\lambda_i\}$ are related with the eigenstates $\{|\psi_i\rangle\}$ by

$$
\hat{A}|\psi_i\rangle = \lambda_i |\psi_i\rangle,
\tag{56}
$$

where, this formula is a symmetry of quantum systems that allows a remarkable simplification to solve most of the quantum mechanics problems [28].

The eigenstates are the possible configurations that can take a quantum system, to better understand this, remember both the C_1 and C_2 column vectors in the marble machine $\mathcal{M}$. An eigenstate is an abstract representation of the state of all the system, for this reason, an eigenstate and an eigenvalue are related, but they are not the same. *I.e.*, thinking in the marble machine $\mathcal{M}$, what are you measuring?.

In order to solve the eigenstates of an observable $\hat{Z}$, the characteristic equation must be solved. For $\hat{Z}$, the eigenvalues are $\{1, -1\}$ and its eigenstates are

$$
|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix},
\tag{57}
$$

where the Dirac notation is used in the left side of equations to short the notation, which is common in quantum mechanics.

7.3. The state vector

Quantum systems are fragile, with only watch them they are alter them. How can be measure something that it cannot be looked? To solve this big problem, in quantum mechanics the paradigm is that always it is measured a quantum system, is possible to get any measurement the system may take. Let's look how it works with a funny example: imagine a Christmas swapping of gifts, where people agreed the possible gifts will be scarves, gloves and cups. You received your gift inside a present box. What is the present into the box? From the perspective of the quantum theory, the gift is a probabilistic superposition of the three gifts (see Fig. 7).

Then, when you look inside the present box, it is said that the probabilistic wave function collapses and the gift rebuilds itself taking the form of actual gift. That is only one manner to model nondeterministic problems in quantum mechanics, nobody has told that physically your gift undergoes a matter transmutation.

FIGURE 7. Pictorial representation of the superposition of quantum states. $|\alpha|^2$, $|\beta|^2$ and $|\gamma|^2$ are the probabilities of get every gift.
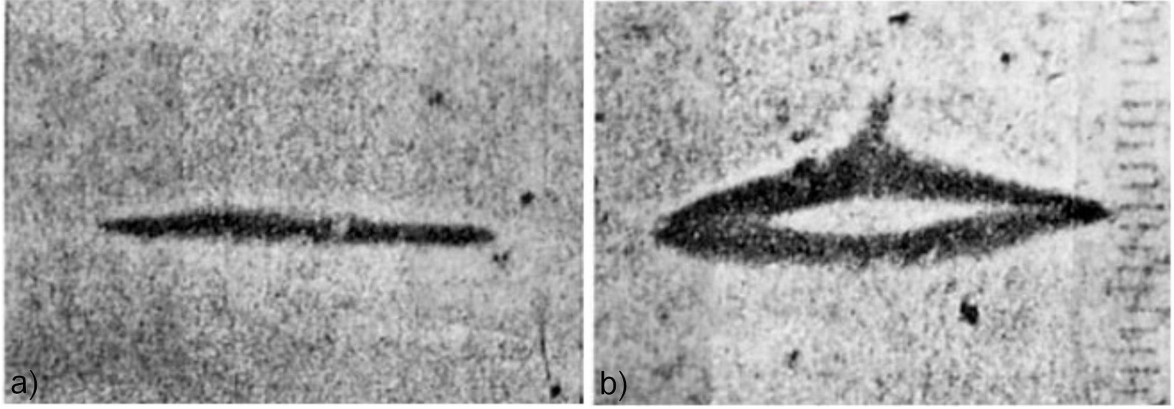
FIGURE 8. Projected spectrum over gatherer screen in the Stern-Gerlach experiment. a) picture shows no dispersion of beam when non electromagnetic field is present. b) picture shows the effect of activate the non-homogeneous electromagnetic field over the beam of electrons. It can be viewed two symmetrical populations on the vertical axe. Picture obtained from section 14 of Armando Martínez Tellez's blog: la-mecanica-cuantica.blogspot.com.

The superposition of possible states a system can take is called the state vector of a quantum system. For example, for one electron in the beam of the Stern-Gerlach experiment, the spin-electron quantum system may be

$|\psi_0\rangle = |0\rangle$ the spin-electron is $\uparrow$,

$|\psi_1\rangle = |1\rangle$ the spin-electron is $\downarrow$,

$|\psi_2\rangle = \dfrac{1}{\sqrt{2}}|0\rangle + \dfrac{1}{\sqrt{2}}|1\rangle$ either it can be one or other ,

$|\psi_3\rangle = \sqrt{\dfrac{1}{3}}|0\rangle + \sqrt{\dfrac{2}{3}}|1\rangle$ likely the spin is $\downarrow$,

$|\psi_4\rangle = \sqrt{\dfrac{1}{5}}|0\rangle + \sqrt{\dfrac{4}{3}}|1\rangle$ almost sure the spin is $\downarrow$,   (58)

where, all the above are examples of state vectors, where weighting factors left to eigenstates in Eq. (58) ensure the normalization of the probabilistic function.

7.4. The density matrix

We now are interested in finding out the probability distribution $p(x)$ of the spectrum of possible eigenvalues a quantum system can take. This distribution is found in the density matrix $\hat{\rho}$. That is a generalization of $p(x)$ in the classical world. Returning to the spin-electron quantum system example, it can be built one of the possible density matrices $\hat{\rho}_Z$ by taking one of its state vectors (for instance Eq. (58)-$|\psi_2\rangle$) and multiplying in this way

$$|\psi_Z\rangle\langle\psi_Z| = \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)\left(\frac{\langle 0| + \langle 1|}{\sqrt{2}}\right)$$
$$= \frac{|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|}{2}$$
$$= \hat{\rho}_Z,$$ (59)

where the state vector $|\psi_Z\rangle$ is multiplied by its conjugate transpose $\langle\psi_Z|$. In matrix form, $\hat{\rho}_Z$ is

$$\hat{\rho}_Z = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$ (60)

It can be viewed that, over the diagonal of $\hat{\rho}_Z$, it reads the values: $(1/2, 1/2)$. These are the values of probability $p(x)$ associated to get 1 and -1 in the Stern-Gerlach experiment (Look at the Fig. 8) [29]. Which means the population of electrons with spin $\uparrow$ and electrons with spin $\downarrow$ are 50-50% percent respectively. This last is the evidence that the probability distribution $p(x)$ of electrons inside the beam is $(p(\uparrow) = 1/2, p(\downarrow) = 1/2)$.

7.5. The von Neumann entropy formula

As it was demonstrated, the eigenvalues of an operator are the measurements of an experiment, in this way, it is natural to say that the probability distribution $p(x)$ is the diagonal of $\hat{\rho}_Z$. Nevertheless, it must be noted that $\hat{\rho}_Z$ is not the same that $\hat{Z}$, former is a density matrix and latter is an operator.

In order to get the von Neumann entropy. From the Shannon formula of the entropy (Eq. (41)) it expands the summation

$$H = -p_1\log_2(p_1) - p_2\log_2(p_2) - \cdots - p_n\log_2(p_n), \quad (61)$$

where each entry of $p(x)$ is an eigenvalue $\hat{\rho}_Z$. Hence

$$H = -\lambda_1\log_2(\lambda_1) - \lambda_2\log_2(\lambda_2) - \cdots - \lambda_n\log_2(\lambda_n), \quad (62)$$

where using the trace it can be simplified as

$$H = Tr \left[\begin{pmatrix} -\lambda_1 \log_2(\lambda_1) & & & \text{coherent} \\ & -\lambda_2 \log_2(\lambda_2) & & \text{terms} \\ \text{coherent} & & \ddots & \\ \text{terms} & & & -\lambda_n \log_2(\lambda_n) \end{pmatrix} \right]$$

$$= Tr \left[\begin{pmatrix} -\lambda_1 & & & \\ & -\lambda_2 & & \\ & & \ddots & \\ & & & -\lambda_n \end{pmatrix} \begin{pmatrix} \log_2(\lambda_1) & & & \\ & \log_2(\lambda_2) & & \\ & & \ddots & \\ & & & \log_2(\lambda_n) \end{pmatrix} \right]$$

$$= -Tr \left[\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \log_2 \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \right], \tag{63}$$

where it must be noted that the matrix is a density matrix, hence

$$H = -Tr \left(\hat{\rho} \, \log_2(\hat{\rho}) \right), \tag{64}$$

that is the von Neumann entropy formula [30].

8. Applications of the von Neumann entropy in quantum computing

The von Neumann entropy has important applications in quantum information theory, particularly in the development of both the Holevo-Schumacher-Westmoreland theorem (HSW) and the Loyd-Shor-Devetak (LSD) formula which are milestones in quantum communications [31]. Here it is exposed some simple cases where the von Neumann entropy is used to measure the variability of the density matrix.

For example, a state vector $|\psi\rangle_0 = |0\rangle$ with density matrix

$$\hat{\rho}_0 = |0\rangle \langle 0| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

has entropy

$$H = -(1) \log_2(1) - (0) \log_2(0) = 0, \tag{65}$$

where it is said that none information will be gained if measuring this system. Namely, since it was previously known the state of the quantum system, it is not surprising that entropy is zero. Other case is state vector $|\psi\rangle_Z = (1/\sqrt{2}) |0\rangle + (1/\sqrt{2}) |1\rangle$, where the eigenvalues of $\hat{\rho}_Z$ are $\{\lambda_1 = 1, \lambda_2 = 0\}$. In this scenario

$$H = -(1) \log_2(1) - (0) \log_2(0) = 0, \tag{66}$$

where again, it will not be gained new information at measuring the system. This is because, from the perspective of the quantum mechanics, the superposition of two states is a pure state of the system in their own right, therefore there must not be a surprise that a measurement of this system will be ↑ or ↓. That is to say, in quantum mechanics is natural to talk about undetermined states.

8.1. The mixed state

In other hand, if a quantum system was a mix of two type of pure states (for instance Eq. (58) -$|\psi_0\rangle$ and Eq. (58) -$|\psi_2\rangle$), then this system must be described by the sum of the density matrix of each individual system. *I.e.*, if $\hat{\rho}_0 = |\psi_0\rangle \langle \psi_0|$ and $\hat{\rho}_2 = |\psi_2\rangle \langle \psi_2|$, a weighting sum of these density matrices is

$$\hat{\rho} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\hat{\rho} = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}, \quad \lambda_1 = \frac{1}{4} \left(\sqrt{2} + 2 \right),$$

$$\lambda_2 = \frac{1}{4} \left(2 - \sqrt{2} \right), \tag{67}$$

where, it has been calculated the eigenvalues of ρ to stress that this time, the eigenvalues do not correspond to the possible measurement of the system, simply because they are two mixed systems. But now,

$$H = -\frac{1}{4} \left(\sqrt{2} + 2 \right) \, \log_2 \left(\frac{1}{4} \left(\sqrt{2} + 2 \right) \right)$$

$$- \frac{1}{4} \left(2 - \sqrt{2} \right) \, \log_2 \left(\frac{1}{4} \left(2 - \sqrt{2} \right) \right)$$

$$= 0.600876 \text{ bits}, \tag{68}$$

where the weighting factors left to matrix densities is the probability of finding these states in the ensemble. In this case, the entropy is used to determine that the system is not a pure state.

8.2. The entanglement

Other example of von Neumann entropy is evaluating the degree of entanglement in multipartite quantum systems in non-inertial frames [32]. Entanglement is a quantum property that

does not have analogous in the classic world. The quantum world exhibits new physical rules that cannot be perceived in the classical due to a statistical vanish.

When the spins of two electrons interact, they flip its spin to fulfil with the angular momentum conservation law to sum zero [33]. In this way, two electrons e_1 and e_2 form a bipartite pure state called a singlet state ($|\psi_S\rangle$).

In general, a singlet state $|\psi_S\rangle$ can consider 4 possible configurations:

- $|0\rangle_{e_1} \otimes |0\rangle_{e_2}$, e_1 is *spin* $\uparrow$, e_2 is *spin* $\uparrow$

- $|0\rangle_{e_1} \otimes |1\rangle_{e_2}$, e_1 is *spin* $\uparrow$, e_2 is *spin* $\downarrow$

- $|1\rangle_{e_1} \otimes |0\rangle_{e_2}$, e_1 is *spin* $\downarrow$, e_2 is *spin* $\uparrow$

- $|1\rangle_{e_1} \otimes |1\rangle_{e_2}$, e_1 is *spin* $\downarrow$, e_2 is *spin* $\downarrow$

So, a state vector $|\psi_S\rangle$ should be

$$|\psi_S\rangle = c_{0,0} |0\rangle_{e_1} \otimes |0\rangle_{e_2} + c_{0,1} |0\rangle_{e_1} \otimes |1\rangle_{e_2}$$
$$+ c_{1,0} |1\rangle_{e_1} \otimes |0\rangle_{e_2} + c_{1,1} |1\rangle_{e_1} \otimes |1\rangle_{e_2}, \qquad (69)$$

where coefficients $c_{i,j}$ are weightings that ensure $|\psi_S\rangle$ there be normal. Nevertheless, because of the conservation laws only the antiparallel configurations are allowed and $|\psi_S\rangle$ actually is of the form

$$|\psi\rangle = c_{0,1} |0\rangle_{e_1} \otimes |1\rangle_{e_2} + c_{1,0} |1\rangle_{e_1} \otimes |0\rangle_{e_2}, \qquad (70)$$

Then, take attention in noting that, whether it is measured the spin state of the electron e_1 and gets $\uparrow$ ($|0\rangle$), it can be asserted that the spin state of the electron e_2 is $\downarrow$ ($|1\rangle$), without actually measuring this. It is proposed the follow definition of entanglement:

The entanglement is a phenomenon in quantum world where the physic conservation rules constraint several configurations in the state vector and increase overwhelmingly the probability to detect certain quantum states in a bipartite systems.

The entanglement is used in almost all quantum information task, as: quantum cryptography, data compression, quantum communications and teleportation [34], due that, there is so much interest on it. Moreover, as the relativistic effects of infinite accelerations affect the entanglement, there exists interest in investigating the entanglement of quantum states in non-inertial frames. The entropy of von Neumann can be used to measure the loss of entanglement in a multipartite quantum systems because the density matrix tends to mix as become more non-inertial [35]. For example, in a tetrapartite quantum system of a GHZ state in a non-inertial frame the loss of entanglement when one, two, three and all parties are non-inertial can be described as function of a parameter $r \in [0, infty]$ associated to the acceleration (see Fig. 9).

This plot shows the cases of the density matrix when one non-inertial observer (ρ_1), two non-inertial observers (ρ_2), three non-inertial observers (ρ_3) and all non-inertial observers (ρ_4). Entropy functions of the systems increase as
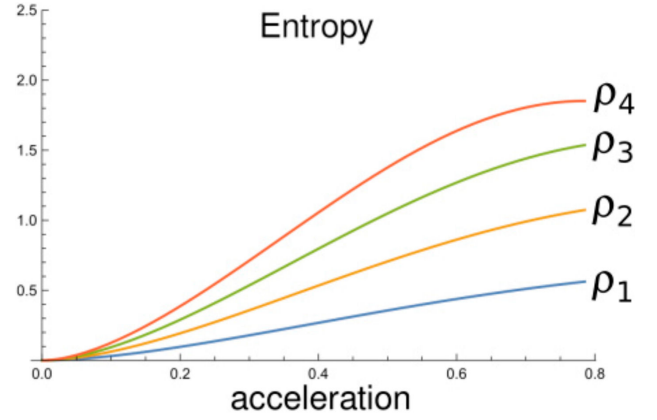

FIGURE 9. Comparison of entropy of von Neumann as one (ρ_1), two (ρ_2), three (ρ_3) and all (ρ_4) observers are non-inertial. The acceleration is parameterized by an angle $r \in [0, \pi/4]$ as $\cos r_i = (1 + e^{-2\pi\omega_i c/a_i})^{-1/2}$, where $\omega = \sqrt{k^2 + m^2}$ and a_i is the acceleration of the qubits. Picture adapted from figure 6 in Ref. [**?**].

acceleration increase, which translates into the loss of entanglement of the tetrapartite systems.

8.3. The channel capacity

Other relevant example it can be shown about issues of entropy is the channel capacity of a bipartite quantum system. Quantum channel capacity is the maximum rate of information it can be transmitted with no error using a quantum system [36]. One analytical expression for quantum channel capacity is

$$C_p = \log_a D + H(\rho_b) - H(\rho_{ab}), \qquad (71)$$

where H is the von Neumann entropy and D is the dimension of density operator ρ_a. This expression is useful for calculating the channel capacity of a bipartite system, but there also exist proposals to calculate quantum channel capacity in multipartite quantum systems greater than two [37].

9. Remarks and conclusions

The entropy concept born with the mystery inside a Carnot machine. When the methodology relate causes with effects fails, and an ideal gas continues expanding despite the supply of heat has been removed of system. Thinkers of the nineteenth century suspect about an internal transformation of ideal gas. The transformation was related with the thermodynamic equilibrium and the time needed to reach it. The formalism of the state functions leads to propose a new state function called entropy. Further on, the statistical formalism of thermodynamic helps to derive a distribution of molecules inside the ideal gas, where the state function results out to be the expectation value of the logarithm of the microstates associated to the distribution of the molecules. From the formalism of the information theory the entropy was used to characterize the variability of the elements in a language, in this

way, the difference among the entropy of Boltzmann S and the entropy of Shannon H is that the variability of elements in a thermodynamic systems tends to an homogeneous distribution when rules inside the systems reach its maximum, while the distribution of a language depends on the optimal code to transmit or store the data.

The core of the entropy concept rely on: to define a system and the rules that act inside it, to wait up until the interactions among elements of the system spread out over the system and at the end ask about the distribution of elements. High entropy means high variability of elements in a distribution, then it is said that is very uncertain an specific outcome in a drawn. But it is needed to contrast the elements of a system, in order to create a code. So, be careful with this, whether the system is nondeterministic, then the natural elements of the system are inherently a superposition of its possibilities. Order and arrow of time are applications of the concept of entropy because rules over the system impose a certain distribution of the elements further on the time. Many uses of the entropy were reviewed to stress the importance of the concept for instance, the entropy concept has been very useful to measure the entanglement and channel capacity on quantum states. Where both are relevant in the development of new quantum technologies.

1. Zemansky, Heat and Thermodynamics. (McGraw-Hill, 1957).

2. M. Hayashi, Quantum Information, an Introduction. (Springer, 1957).

3. F. Tamburini and I. Licata, Relativistic Quantum Information (Mdpi AG, 2020).

4. W. C. Qiang *et al.*, Entanglement property of the Werner state in accelerated frames, *Quantum Information Processing* **18** (2019) 1

5. L. García-Colín, Y sin embargo se mueven... Teoría cinética de la materia. (Fondo de Cultura Económica, 1955).

6. W. M. Saslow, A History of Thermodynamics: The Missing Manual, *Entropy* **22** (2019), `https://doi.org/10.3390/e22010077`.

7. E. T. Jaynes, Information Theory and Statistical Mechanics, *Phys. Rev.* **106** (1957) 620, `https://doi.org/10.1103/PhysRev.106.620`.

8. J. Haglund, F. Jeppsson, and H. Strömdahl, Different Senses of Entropy-Implications for Education, Entropy 12 (2010) 490, `https://doi.org/10.3390/e12030490`.

9. A. Ben-Naim, Entropy and Information Theory: Uses and Misuses, *Entropy* **21** (2019), 10.3390/e21121170

10. M. M. Kostic, The Second Law and Entropy Misconceptions Demystified, *Entropy* **22** (2020), `https://doi.org/10.3390/e22060648`.

11. R. Gaudenzi, Entropy? Exercices de Style, *Entropy* **21** (2019), `https://doi.org/10.3390/e21080742`.

12. C. H. Bennet, Notes on the history of reversible computation, *IBM Journal of Research and Development* **31** (1988) 16.

13. L. García-Colín, De la Máquina de Vapor al Cero Absoluto (Fondo de Cultura Económica, 2011), México.

14. L. García-Colín, Introducción a la Física Estadística (El colegio Nacional, 2011), México.

15. C. M. Bishop, Pattern Recognition and Machine Learning (Springer, 2006), Singapur.

16. E. T. Jaynes, Gibbs vs Boltzmann Entropies, *American Journal Of Physics* **33** (1964) 391.

17. L. Boltzmann, On the Relationship between the Second Fundamental Theorem of the Mechanical Theory of Heat and Probability Calculations Regarding the Conditions for Thermal Equilibrium (1877), Sharp, Kim, and Franz Matschinsky. 2015. Translation of Ludwig Boltzmann's Paper.

18. L. Brillouin, Science and Information Theory (Academic Press, 1960), EUA.

19. D. Welsh, Codes and Cryptography (Clarendon Press, Oxford, England, 1988).

20. T. Cover and J. Thomas, Elements of Information Theory (Wiley-Interscience, 2006), EUA.

21. C. E. Shannon, A Mathematical Theory of Communication, *The Bell System Technical Journal* **27** (1948) 379.

22. M. M. Wilde, Quantum Information Theory (Cambridge University Press, 2017), United Kingdom.

23. G. Kutrovátz, Heat Death in Ancient and Modern Thermodynamics, *Open Systems Information Dynamics* **08** (2001) 349, `https://doi.org/10.1023/a:1013901920999`.

24. Q. A. Wang, Probability distribution and entropy as a measure of uncertainty (2008), `https://hal.archives-ouvertes.fr/hal-00117452`, 11 pages. To be published in J. Physics A : Math. Theor. (2008).

25. E. Kharitonov *et al.*, Entropy Minimization In Emergent Languages (2019), `https://doi.org/10.48550/ARXIV.1905.13687`.

26. O. Gandrillon *et al.*, Entropy as a measure of variability and stemness in single-cell transcriptomics, *Current Opinion in Systems Biology* **27** (2021) 1, `https://doi.org/10.1016/j.coisb.2021.05.009`.

27. N. Yanofsky and M. Manucci, Quantum Computing For Computer Scientist (University Press, 2008), pp. 133-136, EUA.

28. R. Lemus and A. O. Hernández-Castillo, Symmetry projection, geometry and choice of the basis, *Rev. Mex. Fis. E* **61** (2015) 113.

29. T. A. Martínez, La Mecánica Cuántica: El experimento Stern-Gerlach, online blog (2009), URL `http://la-mecanica-cuantica.blogspot.com/2009/08/el-experimento-stern-gerlach.html`, Http://la-mecanica-cuantica.blogspot.com/2009/08/elexperimento-stern-gerlach.html.

30. M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, 2010), United Kingdom.

31. M. Hsieh and M. M. Wilde, Entanglement-Assisted on Information Theory Communication of Classical and Quantum Information, *IEEE Transactions* **56** (2010) 4682.

32. Q. Dong *et al.*, Tetrapartite entanglement measures of GHZ state with uniform acceleration, *Physica Scripta* **94** (2019) 1, `https://doi.org/10.1088/1402-4896/ab2111`.

33. J. J. Sakurai and J. Napolitano, Modern Quantum Mechanics (Addison-Wesley (Jim Smith), 201), pp. 199-245, EUA.

34. U. Pereg, C. Deppe, and H. Boche, Quantum broadcast channels with cooperating decoders: An information-theoretic perspective on quantum repeaters, *J. Math. Physics* **64** (2021), `https://doi.org/10.1063/5.0038083`.

35. A. J. Torres-Arenas *et al.*, Tetrapartite entanglement measures of W-class in noninertial frames, *Chinese Physics B* **28** (2019) 1.

36. C. H. Benett *et al.*, Entanglement-assisted classical capacity of noisy quantum channels, *Physical Review Letters* **83** (1999) 3081.

37. A. Sagheer and H. Hamdoun, Some Properties of Multi-Qubit States In Non-Inertial Frames, In IEEE Symposium on Swarm Intelligence (SIS) (2013) pp. 141-146.

38. D. McMahon, Quantum Computing Explained (Wiley- Interscience A John Wiley and Sons, 2008), EUA.