Some Analogies Between Thermodynamics and Shannon Theory

Dragan Samardzija
Wireless Research Laboratory
Bell Laboratories, Alcatel-Lucent
Holmdel, NJ 07733, USA
Email: dragan@bell-labs.com

Abstract— In this paper we present some analogies between thermodynamics and certain Shannon theory results. We revisit the previously published results that relate notion of energy and information. We then introduce a thermodynamic system that could be used to store information. The ideal gas is considered. We present the corresponding thermodynamic analysis and establish equivalence with the additive white Gaussian noise (AWGN) channel capacity formula. Specifically, we show that the average energy needed for adiabatic compression of the ideal gas to $1/N$ of its initial volume is the same as the average energy needed to achieve the capacity $C = \log_2 N$ of the equivalent AWGN channel. In addition, the analysis is extended to show a link between the gas volume and minimum squared codeword distance. Furthermore, we show that the ideal gas which went through the adiabatic compression, and later settled according to the second law of thermodynamics, will reach an equilibrium state which is directly related to Shannon random coding and joint typicality decoding.

I. Introduction

In the early days of information theory a tentative link with thermodynamics and its probabilistic aspects was recognized. *Shannon entropy*, which was defined in the ground-breaking work [1], was named after the well known *thermodynamic entropy*. Thermodynamic entropy, as a measure of how energy is distributed among system components, was introduced by Clausius (circa 1850) while its probabilistic interpretation was attributed to Boltzmann (at the turn of the twentieth century). The functional resemblance between the entropies was recognized by Shannon and Von Neumann. An anecdote goes that Von Neumann advised Shannon to adopt this term because it would give him "... a great edge in debates because nobody really knows what entropy is anyway." [2]

Since the beginning, information theory took on a life of its own, but there has been a number of efforts to apply its results in different branches of physics. For example, certain results from information theory are applied in statistical and quantum mechanics [3], [4]. Furthermore, the notion of information became a central point in many studies in theoretical physics. In particular, there has been an interest in determining the limits on how much information can be stored in unit mass and what the limits on speed of information processing are (see [5] and references therein).

Attempts to understand energy limits of computation led to a number of results connecting thermodynamics and theory of computation [6]. In addition to computation problems, the studies have addressed the issue of how much energy has to be invested into a thermodynamic system to half the number of its available states (i.e., the system loses one bit of information needed to describe its state). Furthermore, it was shown that a knowledge of the system state leads to an ability to extract energy out of the system. For the given amount of knowledge, described by the number of information bits, there is a limit on the energy that can be extracted [7], [8]. The results will be reviewed later in the text. In this paper, we extend the basic ideas that are used to derive the above results, and focus on the analogies with Shannon theory and communication systems.

The paper is organized as follows. In Section II the basic properties of the ideal gas are presented. Furthermore, we revisit and summarize the previously reported results that relate notion of energy and information. In Section III we show that the average energy needed for adiabatic compression of the ideal gas to $1/N$ of its initial volume is the same as the average energy needed to communicate $C = \log_2 N$ bits per channel use over the equivalent additive white Gaussian noise (AWGN) channel (i.e., to achieve the channel capacity C). The analysis is extended to show a link between the gas volume and minimum squared codeword distance. Furthermore, we show that the ideal gas which went through the adiabatic compression, and later settled according to the second law of thermodynamics, will reach an equilibrium state which is directly related to Shannon random coding and joint typicality decoding. We conclude in Section IV.

II. Summary of Previously Published Results

In this section we first repeat certain communication theory results that are relevant to this study. We then present basic definitions and equations of thermodynamics. Furthermore, we revisit some results that connect notion of energy and information. Focusing on information and communication theory aspects, we present two theorems summarizing the results.

A. Additive White Gaussian Noise Channel

In this study we consider a complex AWGN channel. The output of the AWGN channel is

$$y = x + n \tag{1}$$

where x is the transmitted symbol, and n is the additive noise with complex Gaussian distribution $\mathcal{CN}(0, N_0)$, where

$N_0 = \mathrm{E}[nn^*]$ is the channel noise variance. The transmitted symbol x is assumed to be circularly symmetric random variable with complex Gaussian distribution $\mathcal{CN}(0, E_s)$, where $E_s = \mathrm{E}[xx^*]$ is the average energy of the transmitted symbol per channel use.

The maximum achievable communication data rate (i.e., capacity) of the above AWGN channel is

$$C = \log_2\left(1 + \frac{E_s}{N_o}\right) \qquad (2)$$

and

$$E_s = N_0(2^C - 1). \qquad (3)$$

By definition, C information bits are transmitted per channel use, therefore the average transmitted energy per one information bit is

$$E_b = \frac{E_s}{C} = N_0 \frac{(2^C - 1)}{C}. \qquad (4)$$

Furthermore, the minimum energy needed to infinitely slowly communicate one bit of information over the AWGN channel is

$$E_b^{min} = N_0 \lim_{C \to 0} \frac{2^C - 1}{C} = N_0 \ln 2. \qquad (5)$$

Note that it can be shown that a complex AWGN channel is equivalent to a real AWGN channel with two degrees of freedom [9]. Specifically, per each signal dimension the noise variance and the average energy of the transmitted symbol are $N_0/2$ and $E_s/2$, respectively.

B. Ideal Gas

In this study we use the ideal gas model. The ideal gas corresponds to a set of particles (i.e., molecules) that are confined within a container. The particles do not interact with each other, except during collisions, which are assumed to be perfectly elastic. Thus, the energy of each particle is exclusively translational kinetic energy. The internal energy U of the gas is the sum of energies of the particles.

A differential change of the internal gas energy is

$$dU = dQ - dW \qquad (6)$$

where dQ is the thermal energy, i.e., the heat exchanged with the surrounding environment, and dW is the differential mechanical work [10]. The above equation is a consequence of the conservation of energy, i.e., the first law of thermodynamics. The exchanged thermal energy is

$$dQ = T dS \qquad (7)$$

where T is the temperature of the gas, while dS is the differential change of its thermodynamic entropy. As mentioned earlier, the thermodynamic entropy is a fundamental quantity that describes how the internal energy U is distributed among the system components (i.e., in this case among the particles of the gas). Further, the differential mechanical work is

$$dW = p \, dV \qquad (8)$$

where p is the pressure that gas exerts on the walls of the container, while dV is the differential change of its volume.

The average energy of a particle, whose energy depends quadratically on its velocity, equals $kT/2$ per a degree of freedom, where k is Boltzmann's constant [11]. The change of the average internal energy U is

$$dU = \frac{LMk}{2} dT \qquad (9)$$

where L is the number of gas particles, and M is the number of degrees of freedom. Furthermore, the following well known relationship connects the gas pressure, volume and temperature as

$$p \, V = Lk \, T. \qquad (10)$$

Note that the above expression can be directly derived from the classical Newtonian mechanics under the ideal gas assumptions.

Without loss of generality, we will consider the ideal single-particle gas, $L = 1$, where the particle has two degrees of freedom, $M = 2$. For example, this assumption may correspond to a case of the particle being confined to a plane. Note that we will revisit these assumptions later in the text.

C. Volume Halving

In this subsection we assume that the walls of the container are thermally conductive. The container is submerged into a *thermal bath*, which is an infinitely large container at the constant temperature T. Let us compress the gas to one half of its initial volume. The compression is done infinitely slowly keeping the temperature of the gas constant. This process is known as isothermal compression. Thus, $dT = 0$ and the equation (6) becomes

$$0 = T dS - p dV = T dS - \frac{kT dV}{V}. \qquad (11)$$

After compressing the gas to $\widetilde{V} = V/2$, where V is its initial volume, from the above equation we get

$$\Delta \widetilde{W} = -\int_V^{V/2} \frac{kT dV}{V} = kT \ln 2 \qquad (12)$$

where $\Delta \widetilde{W}$ is the external mechanical work that is invested to perform the compression.

Let us now argue how the above exercise relates to information.

First, by halving the volume of the thermodynamic system, the number of available states where the gas particle could be is also halved. Therefore, we'll need one information bit less to describe its state compared to the initial full-volume state. Considering the thermodynamics of computation, the result in (12) is determined as the minimum energy needed for discarding one bit during the computation, i.e., information processing. This result is known as *Landauer's principle of erasure* and plays crucial role in determining how much energy is needed to perform a computation [6]–[8], [12].

Second, compressing the gas to one or the other half may be used to record, i.e., store one information bit. For example, if the gas is compressed to its left or right half it will be assigned logical '0' or '1', respectively.

We now relate the above analysis to the communication theory results. Specifically, for $N_0 = kT$, the above expression (12) becomes identical to the expression (5), i.e.,

$$\Delta \widetilde{W} = E_b^{min} \text{ for } N_0 = kT. \tag{13}$$

Focusing on information and communication aspects, we summarized the above analysis in the following.

Theorem 1: The same amount of energy ($kT \ln 2$ Joules) is needed to either store one information bit by the isothermal compression of the ideal gas or communicate one information bit over the equivalent AWGN channel (where $N_0 = kT$). In both cases, the compression and transmission are infinitely slow. $\square$

D. Energy Extraction

For the sake of simplicity let us assume that that a single-particle gas occupies the volume V, at the temperature T. As in the previous subsection, the gas container is in the thermal bath. The particle will be in either left or right half of the container with the equal probability. We assume that there is a mechanism to insert a piston either from left or right side of the container.

If there was a knowledge, i.e., information in which half of the container the particle is, the piston could be inserted from the opposite side. It is assumed that no energy is needed to perform the piston insertion. The procedure is depicted in Figure 1, for the particle being in the left half. One information bit will be sufficient to describe in which half of the container the particle is.
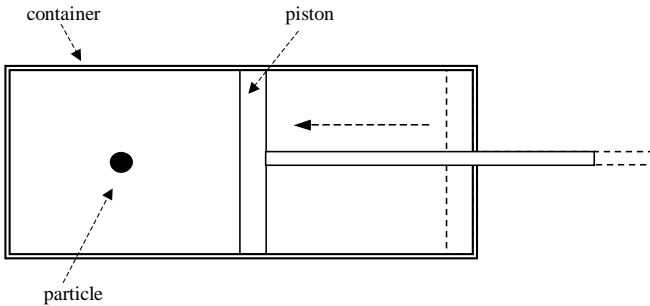

Fig. 1. The piston is inserted from the right while the particle is in the left half of the container.

After the insertion, the gas exerts pressure on the piston. Under the pressure, the piston is moving into the direction of the gas expansion. While expanding, the gas performs mechanical work. If the piston is moving infinitely slowly, upon reaching the initial volume V the gas has performed the total work that is equal to the one in (12). The expansion is isothermal.

The above system is known as *Szilard's engine* [8] and together with Landauer's principle is extensively used to analyze the famous thought experiment known as *Maxwell Demon* (see [13] for more details).

We summarized the above in the following.

Theorem 2: If there is a knowledge about the state of the ideal gas corresponding to one information bit, $kT \ln 2$ Joules of energy can be extracted from the gas via its isothermal expansion. The gas expansion and energy extraction are infinitely slow.$\square$

In other words, the energy value of one bit of information is $kT \ln 2$ Joules.

III. FURTHER ANALOGIES WITH SHANNON THEORY

Unlike the previous two subsections, here we assume that the gas is thermally isolated from its surroundings. Specifically, in the expression (6) the heat exchanged with the surrounding environment is $dQ = 0$. Thermodynamic processes with $dQ = 0$ are known to be *adiabatic*. As said earlier, we will study a single-particle gas, where the particle has two degrees of freedom.

A. Gas Compression and Information Storing

Let us compress the gas to $1/N$ of its initial volume V, $\widehat{V} = V/N$. The compression is done adiabatically, without any constraint on how fast it is performed (unlike the case of the isothermal compression that has to be infinitely slow).

Starting from the equations (6), (9), (10), for $L = 1$, $M = 2$ and $dQ = 0$ we arrive to

$$dU = -dW \Rightarrow k\,dT = -p\,dV \tag{14}$$

and

$$\frac{dT}{T} = -\frac{dV}{V}. \tag{15}$$

From the above

$$\ln \frac{\widehat{T}}{T} = -\ln \frac{\widehat{V}}{V} \tag{16}$$

where $\widehat{T}$ is the gas temperature after the compression and $\widehat{V} = V/N$. Thus,

$$\widehat{T} = T\,N. \tag{17}$$

The average external work that is invested to perform the compression is equal to the change of the average internal gas energy. In other words,

$$\Delta \widehat{W} = \Delta U = k(\widehat{T} - T) = kT(N - 1). \tag{18}$$

We now describe how the above process can be used to record, i.e, store information. The gas container of volume V is divided in N disjoint compartments. The compartments are of volume $\widehat{V} = V/N$. Each compartment will assigned to a unique string of information bits with the length of

$$C = \log_2 N \text{ bits.} \tag{19}$$

Storing of information is done by compressing and confining the ideal gas to one of the N compartments depending on C-bit-long binary word that is being stored. The confined gas may be viewed as data symbol. Figure 2 depicts a possible compartment and binary word arrangement for $C = 3$ and $N = 8$.
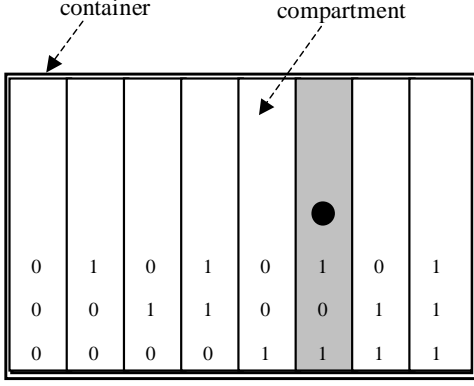
Fig. 2. The container is divided in $N = 8$ compartments, with a unique 3-bit binary word assigned to each of the compartments. In this particular example the gas particle is compressed and confined to the compartment corresponding to the binary word '101'.

Based on the expressions (18) and (19), the average energy invested to store the binary word is

$$\Delta \widehat{W} = kT(2^C - 1), \tag{20}$$

while the average energy per bit is

$$\Delta \widehat{W}_b = kT \frac{(2^C - 1)}{C}. \tag{21}$$

Note that for $N_0 = kT$, the results in (21) and (4) are identical, i.e.,

$$\Delta \widehat{W}_b = E_b \text{ for } N_0 = kT. \tag{22}$$

The following theorem summarizes the above analysis.

Theorem 3: The same amount of average energy is needed to either store C information bits by the adiabatic compression of the ideal gas or achieve the channel capacity C of the equivalent AWGN channel (where $N_0 = kT$). The ideal gas is compressed to $1/N$ of its original volume, where $N = 2^C$. □

Note the above theorem is more general than Theorem 1. It does not restrict the rate of information storing or transmission.

B. Volume and Codeword Distance

In this subsection we first consider a case when the average energy in (20) is invested to compress the gas, but the initial temperature is different than T, i.e.,

$$T' = T + \Delta T. \tag{23}$$

After the compression, instead of reaching $\widehat{V} = V/N$, the gas occupies different volume

$$\widehat{V}' = \widehat{V} + \Delta V. \tag{24}$$

The temperature after the compression becomes

$$\widehat{T}' = \frac{\widehat{U}'}{k} = T(N-1) + T' = TN + \Delta T \tag{25}$$

where U' is the average internal energy of the gas after the compression. From the equation (15), the volume $\widehat{V}'$ is

$$\widehat{V}' = \frac{V(T + \Delta T)}{TN + \Delta T}. \tag{26}$$

The rate of the compressed volume change due to the difference in the initial temperature is

$$\mu_V = \frac{\Delta V}{\widehat{V}} = \frac{\Delta T(N-1)}{TN + \Delta T}. \tag{27}$$

We now pose an equivalent communication problem. The average energy of the transmitted symbol is given in (3). The noise variance is different than N_0, i.e.,

$$N_0' = N_0 + \Delta N_0. \tag{28}$$

Consequently, the maximum achievable data rate is

$$C' = \log_2 \left(1 + \frac{E_s}{N_0'}\right). \tag{29}$$

In order to achieve the capacity C in (2), the transmitter sends one of 2^{nC} codewords, where $n \to \infty$. The codeword belongs to a set of 2^{nC} typical sequences that are uniformly distributed in the signal space [9], [14]. Due to the transmit power constrain, the volume of the signal space is proportional to $(nE_s)^n$. A fraction of the signal space volume that is allocated per codeword is

$$S = \alpha \frac{(n E_s)^n}{2^{nC}} \text{ for } n \to \infty, \tag{30}$$

where α is a proportionality constant. From the above, the squared distance between the neighboring codewords is

$$s = \beta S^{1/n} = \gamma \frac{n E_s}{2^C} \tag{31}$$

where β is a proportionality constant and $\gamma = \beta \alpha^{1/n}$. s is the minimum squared distance between the codewords, and it is proportional to the squared radius of the sphere of exclusion around each codeword (where different codeword spheres must not intersect). Similarly, for a channel with the capacity C' in (29), the above squared distance is

$$s' = s + \Delta s = \gamma \frac{n E_s}{2^{C'}}. \tag{32}$$

The rate of the minimum squared codeword distance change, due to the difference in the noise variance is

$$\mu_s = \frac{\Delta s}{s} = \frac{2^C - 2^{C'}}{2^{C'}} = \frac{\Delta N_0(2^C - 1)}{N_0 2^C + \Delta N_0}. \tag{33}$$

Note that for $N_0 = kT$, $\Delta N_0 = k\Delta T$, and $N = 2^C$ the results in (27) and (33) are identical, i.e.,

$$\mu_V = \mu_s \text{ for } N_0 = kT \text{ and } \Delta N_0 = k\Delta T. \tag{34}$$

The following theorem summarizes the above results.

Theorem 4: Using the same amount of average energy, the ideal gas can be adiabatically compressed to the volume $\widehat{V}$ that depends on the initial temperature T, i.e., $\widehat{V} = \widehat{V}(T)$. Furthermore, using the same amount of average transmit energy, the minimum squared distance s between the capacity achieving

codewords depends on the noise variance, i.e., $s = s(N_0)$. For the equivalent AWGN channel (where $N_0 = kT$ and $\Delta N_0 = k\Delta T$), the rates of the volume and squared codeword distance change are identical, i.e.,

$$\frac{\widehat{V}(T + \Delta T) - \widehat{V}(T)}{\widehat{V}(T)} = \frac{s(N_0 + \Delta N_0) - s(N_0)}{s(N_0)}. \square \quad (35)$$

C. Second Law and Random Coding

Let us perform the following information storing procedure. As described earlier, it is done by compressing and confining the ideal gas to one of the N compartments, storing $C = \log_2 N$ bits of information. Using n gas containers (where $n \to \infty$), the total of nC bits of information is stored. After the compression, the containers are brought into a thermal contact, while still being thermally isolated from the rest of the environment. Upon the contact, in each of the containers the gas particle is released from the compartment it was confined to as the result of information storing. The particle is free to move within its container and exchange thermal energy with the particles in other containers. The process will lead to a thermal equilibrium of the system. In the following we examine the properties of that equilibrium.

According to the second law of thermodynamics, in the equilibrium the energy is distributed among the system components in the most disorganized way. In other words, the thermodynamic entropy of the system is maximized with respect to the probability density function (PDF) of the gas particle energy, i.e.,

$$p(E) = \arg\max_{q(E)} \left(-k \int_0^\infty q(E) \, \log_2 q(E) \, dE \right) \quad (36)$$

under the constraints

$$\int_0^\infty E \, p(E) \, dE = \widehat{U} \quad (37)$$

where $\widehat{U} = kTN = KT2^C$, is the average internal energy of the gas after the compression. Further,

$$\int_0^\infty p(E) \, dE = 1. \quad (38)$$

The solution of the above constrained optimization is

$$p(E) = A \exp\left(-\frac{E}{kTN} \right), \quad E \geq 0, \quad (39)$$

where A the normalization constant (to satisfy the constraint in (38)). The PDF $p(E)$ is known as *Boltzmann distribution* and it is distribution of energies of the ideal gas particles (which are identical but distinguishable) [11]. Boltzmann distribution corresponds to the energy of circularly symmetric complex random variable with Gaussian distribution $\mathcal{CN}(0, kTN)$. In other words,

$$E = |X + iY|^2 = X^2 + Y^2 \quad (40)$$

where X and Y are independent and identically distributed each with Gaussian distribution $\mathcal{N}(0, kTN/2)$.

Based on the above analysis, we conclude the following. Upon reaching the equilibrium, in each container, the gas particle will have energy that is drawn from Boltzmann distribution in (39). As we assumed, each particle has two degrees of freedom. Each degree of freedom has a quantity (e.g., a velocity) that is drawn from Gaussian distribution $\mathcal{N}(0, kTN/2)$. Therefore we can abstract that each particle has a complex quantity $X + iY$ assigned to it. Thus, upon reaching the thermal equilibrium, the particles in n gas containers that are used to store total of nC bits of information, will be associated with a sequence of complex quantities (for $n \to \infty$).

Immediately after the compression, each particle is confined to the volume $\widehat{V} = V/N = V/2^C$. The particle can be anywhere within the volume $\widehat{V}$. Consequently, the number of equally likely initial states is proportional to $\widehat{V}$ [11]. This means that for each string of nC information bits, there are K complex sequences, where

$$K \propto \left(\frac{V}{2^C} \right)^n, \quad \text{for } n \to \infty. \quad (41)$$

In other words, each string of nC information bits corresponds to a unique set of initial states and thus to a unique set of complex sequences with the cardinality K. There are 2^{nC} disjoint sets, one per each string of nC information bits.

Based on the equivalent AWGN channel that was introduced in the previous subsections (where $E_s = kT(N - 1)$ and $N_0 = kT$) the average energy of each complex quantity is

$$\bar{E} = \widehat{U} = KTN = E_s + N_0. \quad (42)$$

Based on the above, upon reaching the equilibrium, the state of the system corresponds to a sequence of received samples at the output of the equivalent AWGN channel. Each sample corresponds to a state of one gas container. The average energy per each sample is given in (42) while the samples are drawn from the complex Gaussian distribution. To determine which string of nC bits was stored, i.e., sent over the equivalent channel, the receiver has to determine which set the received sample sequence belongs to (out of 2^{nC} disjoint sets). This is equivalent to the decoding procedure based on joint typicality for the optimal channel coding that is described in [9]. The following theorem summarizes the above analysis.

Theorem 5: nC information bits are stored in n containers with the ideal gas. In each container, C bits are stored by adiabatic compression to one of 2^C compartments. Upon the compression, the containers are brought into a thermal contact, and particles are freed from the compartments. Based on the second law of thermodynamics the system will reach the equilibrium. The state of the system will correspond to the output of the equivalent AWGN channel with optimally encoded transmission. To retrieve the stored information a decoding based on joint typicality could be used. $\square$

D. Comments on Degrees of Freedom and Number of Particles

As said earlier, we have assumed that the gas particle has two degrees of freedom (e.g., the particle is confined to a

plane). Similarly, the capacity in (2) corresponds to a channel with two degrees of freedom, e.g., a complex AWGN channel [9].

The analysis that is presented in the previous subsections can be directly extended to a more general case where both the gas particle and the AWGN channel have M degrees of freedom (i.e., M-dimensional systems)[1]. Specifically, in the case of the M-dimensional AWGN channel, per each signal dimension the noise variance and the average energy of the transmitted symbol have to be N_0/M and E_s/M, respectively. Its capacity is

$$C_M = \frac{M}{2} \log_2 \left(1 + \frac{E_s}{N_0}\right). \tag{43}$$

Furthermore, we have considered a single-particle ideal gas. If we consider an L-particle ideal gas, instead of $N_0 = kT$ and $\Delta N_0 = k\Delta T$, $N_0 = LMkT/2$ and $\Delta N_0 = LMk\Delta T/2$ should be used, respectively.

Using the above assumptions, we can repeat the analysis given in the previous subsections. It can be shown that it will lead to the equivalent analogies that are presented in Theorems 3, 4 and 5, but for M degrees of freedom and L gas particles.

IV. CONCLUSIONS

In this paper we presented some analogies between thermodynamics and Shannon theory. We showed that the average energy needed for adiabatic compression of the ideal gas to $1/N$ of its initial volume is the same as the average energy needed to achieve the capacity $C = \log_2 N$ of the equivalent AWGN channel. In addition, we presented the link between the gas volume and minimum squared codeword distance. Furthermore, we related the thermal equilibrium of the corresponding system to the equivalent AWGN channel and optimal encoding and decoding that are based on Shannon theory.

Extending the above analogies to study limits on energy and time needed to perform encoding and decoding is of interest for our future work.

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pt. I, pp. 379–423, 1948; pt. II, pp. 623–656, 1948.
[2] R. P. Feynman, *Feynman lectures on computation*, Perseus Publishing, 1996.
[3] E. T. Jaynes, "Information theory and statistical mechanics," *Phys. Rev.*, vol. 106, pp. 620–630, 1957.
[4] E. T. Jaynes, "Information theory and statistical mechanics II," *Phys. Rev.*, vol. 108, pp. 171–190, 1957.
[5] S. Lloyd, "Ultimate limits on computation," *Nature*, vol. 406, pp. 1047–1054, 2000.
[6] C. H. Bennett, "Thermodynamics of computation - a review," *Intl. J. Theoretical Physics*, vol. 21, pp. 905–940, 1982.
[7] T. W. Lynch, "The energy content of knowledge," *IEEE Workshop on Physics and Computation Physcomp'94*, pp. 78–82, 1994.
[8] W. H. Zurek, "Algorithmic randomness, physical entropy, measurements and the demon of choice," *eprint arXiv:quant-ph/9807007*, 1998.
[9] T. M. Cover and J. A. Thomas, *Elements of information theory*, Wiley-Interscience, 1991.
[10] E. Fermi, *Thermodynamics*, Dover Publications, 1956.
[11] C. Kittel and H. Kroemer, *Thermal physics*, Freeman, 2nd edition, 1980.
[12] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM J. of Res. and Dev.*, vol. 5, pp. 183–191, 1961.
[13] H.S. Leff and A. F. Rex, *Maxwell demon: entropy, information, computing*, Adam Hilger, 1990.
[14] S. Verdú and T. Han, "The role of the asymptotic equipartition property in noiseless source coding," *IEEE Trans. on Information Theory*, vol. 43, pp. 847–857, 1997.

---

[1]For example, for three degrees of freedom, $M = 3$, the gas particle freely moves through the three-dimensional space within the container.