Shannon Meets Carnot: Generalized Second Thermodynamic Law

Ori Shental[1] and Ido Kanter[2]
[1]Center for Magnetic Recording Research, University of California, San Diego
9500 Gilman Drive, La Jolla, CA 92093, USA
[2]Department of Physics, Bar-Ilan University, Ramat-Gan, 52900 Israel

The classical thermodynamic laws fail to capture the behavior of systems with energy Hamiltonian which is an explicit function of the temperature. Such Hamiltonian arises, for example, in modeling information processing systems, like communication channels, as thermal systems. Here we generalize the second thermodynamic law to encompass systems with temperature-dependent energy levels, $dQ = TdS + < d\mathcal{E}/dT > dT$, where $< \cdot >$ denotes averaging over the Boltzmann distribution and reveal a new definition to the basic notion of temperature. This generalization enables to express, for instance, the mutual information of the Gaussian channel as a consequence of the fundamental laws of nature - the laws of thermodynamics.

PACS numbers: 05.70.-a,89.70.-a,89.70.Cf,89.70Kn

The current scientific conception is that the theory of information is a creature of mathematics and has its own vitality independent of the physical laws of nature.

The laws of thermodynamics are fundamental laws of nature, and in particular, the second thermodynamic law linearly relates the change in the entropy, dS, to the amount of heat, dQ, absorbed by a system at equilibrium, $dQ = TdS$, thus defining the temperature, T, of the system [1]. Thermodynamics is primarily an intellectual achievement of the 19^{th} century. The first analysis of heat engines was given by the French engineer Sadi Carnot in his seminal 1824 publication, '*on the Motive Power of Fire and on Machines Fitted to Develop that Power*' [2, 3], laying the foundations to the second law of thermodynamics. This paper marks the start of thermodynamics as a modern science [4], which was subsequently evolved to the more general discipline of statistical mechanics [5].

Information theory is a statistical theory dealing with the limits and efficiency of informatics [6]. This theory is one of the major enablers of modern technologies in the information age, from compressed ZIP files, CDs, MP3s, DSL high-speed modems and mobile phones to the Voyager missions to deep space. Similar to the key role of Carnot's paper in the development of thermodynamics, the birth of information theory as an independent discipline is attributed to the landmark publication of Claude Shannon in 1948 '*A* Mathematical Theory of Communication' [7].

The generic problem in information processing is the transmission of information over a noisy channel. This central paradigm of information theory can be mathematically abstracted to having two random variables X and Y representing the desired information and its noisy replica, respectively. Noisy transmission can occur either via space from one geographical point to another, as happens in communications, or in time, for example, when writing or reading files from a hard disk in the computer.

The most important measures of information are entropy and mutual information. Information (Shannon) entropy, $H(\cdot)$, is a measure of the amount of uncertainty in a random variable, indicating how easily data can be compressed. Mutual information quantifies the amount of information in common between two random variables and it is used to upper bound the attainable rate of information transferred across a channel. To put differently, mutual information, $I(X;Y)$, measures the amount of information that can be obtained about one random variable (the channel input X) by observing another (the output Y). A basic property of the mutual information is that $I(X;Y) = H(X) - H(X|Y)$, hence knowing Y, we can save an average of $I(X;Y)$ bits in encoding X compared to not knowing Y.

The archetypal Gaussian channel is one of the most popular models in the field of informatics and it arises in numerous applications in information processing, modeling the relation between the latent input and the corrupted observed output [8]. Consider a real-valued channel with input and output random variables X and Y, respectively, of the form

$$Y = X + N, \tag{1}$$

where $N \sim \mathcal{N}(0, 1/\mathsf{snr})$ is a Gaussian noise independent of X, and $\mathsf{snr} \geq 0$ is the channel's signal-to-noise ratio (SNR). The input is taken from a probability distribution $P(X)$ with a bounded second moment. In this Letter, random variables are denoted by upper case letters and their values denoted by lower case letters.

The Gaussian channel (1) can be also viewed as a physical system [9, 10, 11], operating under the laws of thermodynamics. The microstates of the thermal system are equivalent to the hidden values of the input X. A comparison of the channel's a-posteriori probability distribution, given by Bayes' law

$$P(X=x|Y=y) = \frac{\sqrt{\mathsf{snr}}P(X=x)}{\sqrt{2\pi}p(Y=y)} \exp\Big(-\frac{\mathsf{snr}}{2}(y-x)^2\Big)$$
$$= \frac{\exp\big(-\mathsf{snr}(-xy + \frac{x^2}{2} - \frac{\log P(X=x)}{\mathsf{snr}})\big)}{\sum_{x\in\mathcal{X}} \exp\big(-\mathsf{snr}(-xy + \frac{x^2}{2} - \frac{\log P(X=x)}{\mathsf{snr}})\big)},$$

with the Boltzmann distribution law yields the following mapping of the inverse temperature and energy of the equivalent thermal system

$$\begin{aligned} \mathsf{snr} &\to \beta, \\ -xy + \frac{x^2}{2} - \frac{\log P(X=x)}{\beta} &\to \mathcal{E}(X=x|Y=y;\beta). \end{aligned} \quad (2)$$

In order to capture the temperature-dependent nature of the energy in systems like the communication channel, we generalize the formulation of the second law of thermodynamics $dS = dQ/T$. The differential of the partition function's logarithm, $\log \mathcal{Z}$, can be written as

$$d\log\mathcal{Z} = \frac{d\log\mathcal{Z}}{d\beta}d\beta.$$

Utilizing the identity $\log\mathcal{Z} = -\beta U + S$, one obtains

$$dS = d(\beta U) + \frac{d\log\mathcal{Z}}{d\beta}d\beta.$$

Since for T-dependent energy

$$\frac{d\log\mathcal{Z}}{d\beta} = -U - \beta\left\langle\frac{d\mathcal{E}(X)}{d\beta}\right\rangle,$$

we get

$$\begin{aligned} dS &= d(\beta U) - Ud\beta - \beta\left\langle\frac{d\mathcal{E}(X)}{d\beta}\right\rangle d\beta \\ &= \beta dU - \beta\left\langle\frac{d\mathcal{E}(X)}{d\beta}\right\rangle d\beta. \end{aligned}$$

Recalling that according to the first law $dU = dQ$ concludes the generalized second thermodynamic law

$$dS = \frac{dQ}{T} - \frac{1}{T}\left\langle\frac{d\mathcal{E}(X)}{dT}\right\rangle dT. \quad (3)$$

The generalized second law of thermodynamics (3) has a clear physical interpretation. For simplicity, let us assume that an examined system is characterized by a comb of discrete energy levels $\mathcal{E}1, \mathcal{E}2, \ldots$. The heat absorbed into the T-dependent system has the following dual effect: A first contribution of the heat, $dU - < d\mathcal{E}(X)/dT > dT$, increases the temperature of the system while the second contribution, $< d\mathcal{E}(X)/dT > dT$, goes for shifting the energy comb. However, the shift of the energy comb does *not* affect the entropy, since the occupation of each energy level remains the same, and the entropy is independent of the energy values which stand behind the labels $\mathcal{E}1, \mathcal{E}2, \ldots$. The change in the entropy can be done only by moving part of the occupation of one tooth of the energy comb to the neighboring teeth, and this can be achieved only by changing the temperature. Hence, the *effective heat* contributing to the entropy is $dQ - < d\mathcal{E}(X)/dT > dT$, and this is the physical explanation to the generalized second law (3). A schematic picture of communication-heat-engine is depicted in Figure 1, where the heat devoted to the change of the Hamiltonian is denoted by the term 'working channel'.

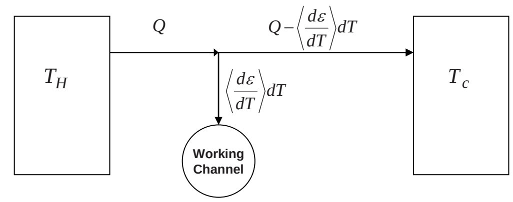

Figure 1: A schematic communication-heat-engine. H/C denote Hot/Cold temperatures.

Note that the first law remains unaffected, $dU = dQ$, since both ways of heat flow absorption into the system are eventually contributing to the average internal energy U. The generalized second law specifies the trajectory, the weight of each one of the two possible heat flows, at a given temperature.

The temperature, originally defined by the second law as $T = dQ/dS$, is redefined now as

$$\frac{1}{T} = \frac{dS/dQ}{1 - \frac{<d\mathcal{E}(X)/dT>}{dQ/dT}} = \frac{dS/dQ}{1 - \frac{<d\mathcal{E}(X)/dT>}{C_V(T)}}. \quad (4)$$

This redefinition has a more complex form and involves an implicit function of T, since the temperature appears on both sides of the equation. Note that the form of the denominator of (4) resembles an efficiency factor of the communication-heat-engine, and will be discussed at the concluding paragraphs.

Based on the generalized second law, a thermodynamic expression for the mutual information of quasi-static communication channels can be derived. To make the derivation intelligible we start with the thermodynamic expression for the mutual information in the case of a temperature-independent Hamiltonian.

The entropy integral obtained from combining the three laws of thermodynamics for the thermal system being equivalent to the Gaussian channel, is

$$S(\beta) = S(X|Y=y;\beta) = -\int_\beta^\infty \gamma\frac{dU(Y=y;\gamma)}{d\gamma}d\gamma,$$

where the internal energy, $U(Y=y;\beta) = \mathbb{E}_{X|Y}\{\mathcal{E}(X|Y=y)\}$, is the energy averaged over all possible values of X, given y. In this contribution, the symbol $\mathbb{E}_X\{\cdot\}$ denotes expectation of the random object within the brackets with respect to the subscript random variable. The posterior information (Shannon) entropy, $H(X|Y;\beta)$, (in nats) of the channel can be expressed via the thermodynamic entropy conditioned on $Y=y$, $S(X|Y=y;\beta)$, as

$$\begin{aligned} H(X|Y;\beta) &= \mathbb{E}_Y\{S(X|Y=y;\beta)\} \\ &= -\mathbb{E}_Y\left\{\int_\beta^\infty \gamma\frac{dU(Y;\gamma)}{d\gamma}d\gamma\right\}. \end{aligned}$$

The input's entropy can also be reformulated in a similar manner, since $H(X) = H(X|Y;\beta=0)$. Hence,

$$H(X) \;=\; -\mathbb{E}_Y\left\{\int_0^\infty \gamma \frac{dU(Y;\gamma)}{d\gamma} d\gamma\right\}.$$

Now, the input-output mutual information can be described via thermodynamic quantities, namely the energy, $\mathcal{E}$, and inverse temperature, β, as

$$\begin{aligned} I(X;Y) &= I(\beta) \triangleq H(X) - H(X|Y;\beta) \\ &= -\mathbb{E}_Y\left\{\int_0^\beta \gamma \frac{dU(Y;\gamma)}{d\gamma} d\gamma\right\} \qquad (5) \\ &= -\big[\gamma \mathbb{E}_Y\{U(Y;\gamma)\}\big]_0^\beta + \mathbb{E}_Y\left\{\int_0^\beta U(Y;\gamma) d\gamma\right\}. \end{aligned}$$

For the generalized second thermodynamic law one has to insert the effective heat contribution to the entropy and the mutual information is given by

$$\begin{aligned} I(X;Y) &= -\big[\gamma \mathbb{E}_Y\{U(Y;\gamma)\}\big]_0^\beta + \qquad (6) \\ &\mathbb{E}_Y\left\{\int_0^\beta \Big(U(Y;\gamma) + \gamma \mathbb{E}_{X|Y}\Big\{\frac{d\mathcal{E}(X|Y;\gamma)}{d\gamma}\Big\}\Big) d\gamma\right\}. \end{aligned}$$

Note that this thermodynamic expression for the mutual information holds for any channel which can be described by a thermal system exhibiting quasi-static heat transfer.

For the Gaussian channel with a standard Gaussian input, $\mathcal{N}(0,1)$, we get $\log(P(X))/\beta = -x^2/(2\beta)$ and the energy (2) of the Gaussian channel system becomes an explicit function of β, given by $\mathcal{E}(X=x|Y=y;\beta) = -xy + x^2(1+\beta)/2\beta$. The derivative of this function with respect to β yields $d\mathcal{E}(X=x|Y=y;\beta)/d\beta = -x^2/(2\beta^2)$. The a-posteriori probability density function is $p(X=x|Y=y;\beta) = \mathcal{N}(\beta y/(1+\beta), 1/(1+\beta))$. Hence, the internal energy is

$$U(Y=y;\beta) = \mathbb{E}_{X|Y}\{\mathcal{E}(X|Y=y;\beta)\} = -\frac{y^2\beta}{2(1+\beta)} + \frac{1}{2\beta}.$$

The derivative of the energy averaged over all possible inputs is

$$\mathbb{E}_{X|Y}\left\{\frac{\partial \mathcal{E}(X|Y=y;\beta)}{\partial \beta}\right\} = -\frac{1}{2\beta^2}\left(\frac{1}{1+\beta} + \frac{\beta^2 y^2}{(1+\beta)^2}\right),$$

and the marginal probability density function of the output is given by $p(Y=y) = \mathcal{N}(0, (1+\beta)/\beta)$. Incorporating all these results into the mutual information expression (6) one can easily show that

$$I(X;Y) = \frac{1}{2}\log(1+\beta),$$

and the celebrated formula for the Shannon capacity [7] is derived from the perspective of thermodynamics.

For the Gaussian channel with Bernoulli-1/2 input, *i.e.* $P(X=1) = P(X=-1) = 1/2$, the $X^2/2 = 1/2$ and $\log P(X=x)/\beta$ terms of the Gaussian channel's energy (2) are independent of X and can be canceled out by the same terms coming from the partition function, leaving us with the expression, independent of β, $\mathcal{E}(X=x|Y=y) = -xy$. The internal energy is $U(Y=y;\beta) = \mathbb{E}_{X|Y}\{\mathcal{E}(X|Y=y)\} = -y\tanh(\beta y)$ and the marginal probability density function of the output is then given by $p(Y=y) = (\sqrt{\beta}/(2\sqrt{2\pi}))[exp(-\beta(y-1)^2/2) + exp(-\beta(y+1)^2/2)]$. Incorporating these definitions into eq. (5) for the case of energy function which is independent of the temperature one can verify that

$$I(\beta) = \beta - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2}\right)\log\cosh(\beta - \sqrt{\beta}y) dy,$$

which is identical to the known Shannon-theoretic result (see, *e.g.*, [12, eq. (18)] and [13, p. 274]).

We turn now to a discussion on the physical implication of the generalized second thermodynamic law and the redefinition of the temperature.

At equilibrium, a system with a T-dependent Hamiltonian is occupying the macroscopic state which maximizes the entropy. We argue that such generalized systems obey the following traditional physical picture: (a) Two systems which are in thermal contact and at equilibrium have the same (redefined) temperature. (b) For two systems which are in thermal contact and at generalized temperature T_1 and T_2, heat always flows from high temperature to low temperature. In addition to this traditional picture, (c) systems fulfilling $< d\mathcal{E}/dT > \geq 0$ achieve an equilibrium state, whereas systems obeying $< d\mathcal{E}/dT > < 0$ are unstable to thermal fluctuations, hence cannot describe a physical system at equilibrium.

For generalized systems, the redefined inverse temperature is $1/T = dS/[dQ - < d\mathcal{E}/dT > dT]$. Hence, β has the same meaning as for the classical second law but with an effective heat contributing to the entropy. When two systems are in thermal contact the total change in the entropy is $dS = dS_1 + dS_2$. A necessary condition for a maximum of the total entropy, dS, is that the slope of the entropy with an effective heat is the same for both systems. Otherwise, heat will flow from a system with the lower slope to the one with the higher slope and the total entropy will increase. Hence it is clear that at a thermal equilibrium the redefined temperatures of the two systems have to obey $T_1 = T_2$ as for the conventional thermodynamic picture.

We present now a self-consistent argument to prove that a generalized system is unstable in the event of $< d\mathcal{E}/dT > < 0$. Assume a generalized system is in thermal equilibrium with a reservoir, and time-dependent fluctuations (of heat transfers) between the measured system and the reservoir. Fluctuations that increase (decrease) the internal energy of the generalized system move the energy comb to 'right' ('left'), such that a higher (lower) energy is assigned to each one of the teeth of the comb. For any macroscopic state there is a degeneracy, *i.e.*, many microscopic states obeying the same macroscopic parameters. We assume that the degeneracy of the system increases in a move to the right tooth, corresponding to a relatively higher energy level.

For the simplicity of the discussion below, assume that the generalized system is always occupying one tooth of the energy comb only. In an event where the system absorbs some heat from the reservoir by fluctuations, there are two 'options' for the system how to 'invest' the absorbing heat. The first option is to move the energy comb to the 'right', and the second one is to keep the current position of the energy comb, but to move the system to a higher (right position) tooth. The first option does not change the entropy, while the second one increases the entropy. Hence, the system adopts the second option. Similarly when the system omits heat to the reservoir the system prefers to move the energy comb to the 'left', instead of moving to a lower tooth. Such processes will repeat again and again, until the energy comb saturates the leftmost position, and the system moves to the highest-energy tooth, independent of the temperature of the reservoir. When the 'left' move of the energy comb is unbounded, the process will not terminate. Note that indeed for the Gaussian channel (see eq. 2) $< d\mathcal{E}/dT >=< x^2/2 > > 0$, and the system achieves an equilibrium state.

Let us examine now two systems at $T_1 < T_2$, $< d\mathcal{E}_1/dT_1 > > 0$ and $< d\mathcal{E}_2/dT_2 > > 0$. Assume self-consistently that the first (second) system absorbs (emits) heat. The total change in the entropy can be expressed as

$$\begin{aligned} 0 \ &< dS = dS_1 + dS_2 = \frac{dQ}{T_1}\big(1 - \frac{T_1}{T_2}\big) \\ &- < \frac{d\mathcal{E}_1}{dT_1} > \frac{dT_1}{T_1}\Big[1 - \Big(\frac{< \frac{d\mathcal{E}_2}{dT_2} >}{< \frac{d\mathcal{E}_1}{dT_1} >}\Big)\Big(\frac{T_1}{T_2}\Big)\Big(\frac{dT_2}{dT_1}\Big)\Big], \quad (7) \end{aligned}$$

where it is clear that $dT_1 > 0$ and $dT_2 < 0$. (If $dT_1 < 0$ then the first system absorbs more heat than dQ, *i.e.*, $dQ- < d\mathcal{E}_1/dT_1 > dT_1$.) Now it can be easily verified that the last term (second line) of the r.h.s of (7) is negative. Hence, a necessary condition for the total entropy, dS, to be positive is that the first term of the r.h.s. of (7) is positive, $T_2 > T_1$, implying that heat flows from high to low temperatures.

T-dependent (independent) Hamiltonians arise also for discrete (continuous) input symbols. An example for such T-dependent Hamiltonians is the Gaussian channels with non-equiprobable binary inputs, $P(\pm 1) \neq 1/2$, whereas an example for a T-independent Hamiltonian is the Gaussian channel with a uniform input distribution. It is clear that the characterization of communication channels with T-dependent/independent Hamiltonians does not follow the type of the input/output symbols. What is the origin, from the point of view of physics, for the T-dependent Hamiltonians? In order to answer this question one has to examine a possible conflict between the nature of the heat and the constraint imposed by the communication channel.

For a communication channel governed by a T-dependent Hamiltonian, *e.g.* eq. (2), the distribution of a microscopic state $exp(-\beta H) \propto P(X)$ which is independent of β. This is a nonphysical postulate assigned to the hidden values of the input X, a degree of freedom in the physical system. The distribution of a degree of freedom (*e.g.*, spin) governed by a Hamiltonian system with/without an inversion symmetry can be nonuniform, but it always depends on the temperature. There is a natural distribution inspired by the heat and a distribution imposed by the communication channel. In case of a conflict between these two distributions the constraint imposed by the channel wins, and the system has to devote some heat to implement it. A heat 'piston' is required to bend the natural heat-distribution to that of the channel, and this is the origin for the *working channel* by the communication-heat-engine (Figure 1).

The non-prefect communication-heat-engine has also a meaningful fingerprint in the additional (third term) of the generalized mutual information, eq. (6), which is negative since $< d\mathcal{E}/d\beta > < 0$, $(< d\mathcal{E}/dT > > 0)$. It indicates that the waste of heat playing against the laws of nature decreases the mutual information in comparison to the same model (communication channel) with the lack of the T-dependent term in the Hamiltonian.

[1] F. Reif, *Fundamentals of Statistical and Thermal Physics* (Mcgraw-Hill, 1965).
[2] S. Carnot, *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance* (Bachelier, Paris, 1824).
[3] E. Mendoza, ed., *Reflections on the Motive Power of Fire: And Other Papers on the Second Law of Thermodynamics* (Dover Publications, New York, 2005).
[4] E. Mendoza, Physics Today p. 32 (1961).
[5] L. E. Reichl, *A Modern Course in Statistical Physics* (University of Texas Press, 1980).
[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (John Wiley and Sons, 1991).
[7] C. E. Shannon, Bell System Technical Journal **27**, 379 (1948).
[8] R. G. Gallager, *Principles of Digital Communication* (Cambridge University Press, Cambridge, UK, 2008).
[9] N. Sourlas, Nature **339**, 693 (1989).
[10] P. Ruján, Phys. Rev. Lett. **70**, 2968 (1993).
[11] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing* (Oxford University Press, Oxford, UK, 2001).
[12] D. Guo, S. Shamai (Shitz), and S. Verdú, IEEE Trans. Inform. Theory **51**, 1261 (2005).
[13] R. E. Blahut, *Principles and Practice of Information Theory* (Addison-Wesley, Reading, MA, 1987).