Thermodynamics of Information

Juan M.R. Parrondo^a

^aDepartamento de Estructura de la Materia, Física Térmica y Electrónica and GISC, Universidad Complutense de Madrid, Ciudad Universitaria s/n, 28040, Madrid, Spain

ARTICLE INFO

Keywords: Feedback processes Information engines Information flows Landauer principle Maxwell demon Mutual information Non-equilibrium free energy Shannon entropy Szilárd engine

ABSTRACT

As early as 1867, two years after the introduction of the concept of entropy by Clausius, Maxwell showed that the limitations imposed by the second law of thermodynamics depend on the information that one possesses about the state of a physical system. A "very observant and neat-fingered being", later on named *Maxwell demon* by Kelvin, could arrange the molecules of a gas and induce a temperature or pressure gradient without performing work, in apparent contradiction to the second law. One century later, Landauer claimed that "information is physical", and showed that certain processes involving information, like overwriting a memory, need work to be completed and are unavoidably accompanied by heat dissipation. Thermodynamics of information analyzes this bidirectional influence between thermodynamics and information processing. The seminal ideas that Landauer and Bennett devised in the 1970's have been recently reformulated in a more precise and general way by realizing that informational states are out of equilibrium and applying new tools from non-equilibrium statistical mechanics.

1. Introduction

The connection between information and thermodynamics was first revealed by the celebrated Maxwell demon, a gedanken experiment that Maxwell devised shortly after his discovery of the distribution of velocities in an equilibrium gas. The velocity distribution revealed that temperature is proportional to the average kinetic energy of molecules, but their speeds cover a wide range of values. Maxwell considered two gases, a hot one A and a cold one B, confined respectively in two containers separated by a wall or diaphragm. Then he imagined an external agent or demon as "a finite being who knows the paths and velocities of all the molecules by simple inspection but who can do no work except open and close a hole in the diaphragm by means of a slide without mass". In a letter to his friend Tait, dated December 11th 1867, Maxwell describes for the first time the operation of the demon:

> Let him first observe the molecules in A and when he sees one coming the square of whose velocity is less than the mean square velocity of the molecules in B, let him open the hole and let it go into B. Next, let him watch for a molecule of B, the square of whose velocity is greater than the mean square velocity in A, and when it comes to the hole let him draw the slide and let it go into A, keeping the slide shut for all other molecules. Then the number of molecules in A and B are the same as at first, but the energy in A is increased and that in B diminished, that is, the hot system has got hotter and the cold colder and yet no work has been done, only the intelligence of a very observant and neat-fingered being has been employed.

🙎 parrondo@ucm.es (J.M.R. Parrondo)

http://seneca.fis.ucm.es/parr (J.M.R. Parrondo)
ORCID(s): 0000-0001-8525-3709 (J.M.R. Parrondo)

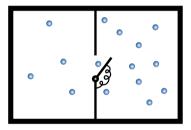


Figure 1: The autonomous pressure Maxwell demon.

Maxwell published the idea, four years after the letter to Tait, in a section of his book, *Theory of Heat* (1871). The section is entitled "Limitation of the second law of thermodynamics". Later on, Lord Kelvin coined the name Maxwell's demon for this "neat-fingered being", and it has been subject of constant controversy until nowadays (Leff and Rex, 1990). The original Maxwell demon opened at least two important lines of research.

The first one is based on a simpler version of the Maxwell demon, also known as pressure demon. In this variant, the demon opens the door only when a particle from the left half of the container is moving to the right half, that is, he allows particles to cross only from left to right. By doing so, the demon accumulates particles in the rightmost half of the container, increasing its density and inducing a pressure difference that can be subsequently used to obtain work. Notice that, in principle, a simple valve, like the one depicted in Fig. 1, could do the same job as the demon. However, this naive idea does not work because the valve is subjected to thermal fluctuations that randomly open the gate and allow particles to cross in the "wrong" direction. In fact, there is a fundamental impossibility of rectifying thermal fluctuations due to the reversibility of microscopic dynamics (Ehrich et al., 2019). The celebrated Smoluchowsky-Feynman ratchet (Feynman et al., 1966; Smoluchowski,

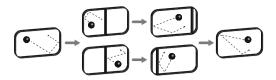


Figure 2: The Szilárd engine.

1912) is another example of autonomous Maxwell demon, which cannot rectify thermal fluctuations from a single bath. Feynman went further and assumed that the ratchet is immersed in a fluid at a temperature lower than the source of the fluctuations and proved that then the ratchet is indeed able to rectify and perform work. Nevertheless, the second law is not defeated since now the system is in contact with two thermal baths at different temperatures and the whole setup acts as a thermal machine. Feynman showed that the efficiency of such a motor cannot be higher than the Carnot efficiency, in agreement with the second law¹. This approach can be applied to generic classical systems in contact with thermostats and/or chemostats. This idea has yielded a fruitful line of research on Brownian motors (Reimann, 2002), systems that exhibit systematic motion or perform work using gradients of temperature and/or chemical potential, with applications in biology and nanotechnology (Bustamante et al., 2001; Feng et al., 2021). The role of information in Brownian motors is not completely clear, although we introduce in section 7 the concept of information flows in autonomous systems, which allows one to analyze some of these machines as an exchange between information and entropy (Allahverdyan et al., 2009; Horowitz and Esposito, 2014).

The second line of research directly concerns the use of information by the original Maxwell's demon (Leff and Rex, 1990; Parrondo et al., 2015; Sagawa, 2012a,b). Two questions immediately arise form the fact that the demon can beat the second law using information. First, can we derive a quantitative relationship between the entropy decrease and the information that the demon possesses? Or, in other words, how the acquisition of information modifies the second law? Second, can we restore the validity of the second law by finding an entropy cost associated to the acquisition and/or manipulation of information by the demon? Thermodynamics of information addresses these two questions. However, it is convenient to analyze them separately. We discuss the first question in section 4 and the second one in section 6.

In 1929, Szilárd introduced a simplified version of the Maxwell demon known as the Szilárd engine (Szilárd,

1929), which is more suitable for the resolution of these issues. It consists of a single-molecule gas at temperature T in a box of volume V. This means that the molecule thermalizes at temperature T in any collision with the wall, i.e., the outgoing velocity is a random variable distributed according to the corresponding equilibrium distribution. An external agent or demon performs the cyclic operations sketched in Fig. 2. First, he inserts a piston at the middle point of the container and measures in which of the two halves the particles lies. With this information, the demon can realize a reversible isothermal expansion by opposing a force to the pressure P exerted by the single-molecule gas. In the expansion, the demon extracts a work

$$W_{\text{extract}} = \int_{V/2}^{V} P dV' = \int_{V/2}^{V} \frac{kT}{V'} dV' = kT \ln 2$$
 (1)

where we have used the equation of ideal gases, PV' = NkT, with N = 1, and k is the Boltzmann constant. Finally, the piston is removed and the cycle is completed. In principle, the insertion and removal of the piston can be performed without any energy expenditure². Hence, the Szilárd engine is able to extract energy from a single thermal bath in a cycle, in apparent contradiction to the second law.

Some aspects of the original Szilárd cycle are not completely clear, specially the application of the ideal gas equation to the pressure on the piston, which is due to single collisions. Nevertheless, these issues can be resolved and, moreover, the Szilárd engine can be implemented in any system exhibiting a symmetry breaking, like an Ising model (Parrondo, 2001). In fact, there are experimental realizations of the Szilárd engine in a variety of systems: single-electron circuits (Koski et al., 2014; Pekola and Khaymovich, 2019), Brownian particles in optical traps (Roldán et al., 2014), or Brownian rotors in electrostatic potentials (Toyabe et al., 2010).

The Szilárd engine, as the original Maxwell demon, utilizes the information gathered in the measurement to decrease the entropy of the system. While in the Maxwell demon the information involved and the decrease of entropy are difficult to assess, in the Szilárd engine both have a precise value: information is acquired in a single measurement per cycle with two equiprobable outcomes, left (L) or right (R), and the entropy decreases by $k \ln 2$. This is why most of the studies on the connection between thermodynamics and information are based on the Szilárd engine. In fact, the Szilárd engine partially solves the first question mentioned above. Before the measurement, the particle can be in any of the two halves and, after the measurement, it occupies only one. The measurement is equivalent to a compression from a volume V to V/2, which reduces the entropy of the gas by $k \ln 2$. Such a compression would need a work $kT \ln 2$, counterbalancing the work (1) extracted in the expansion, but apparently the measurement is able to perform the compression without any energy expenditure.

¹Feynman also proved that the thermal ratchet can reach Carnot efficiency in the limit of zero power. However, this result has been subject to some criticism: an autonomous thermal machine can reach Carnot efficiency only in the case of tight coupling between heat flow and mechanical work, something that it is not possible to reach in the original setup by Smoluchowsky and Feynman (Parrondo and Español, 1996).

²This is not true for a quantum particle at low temperature because, due the Heisenberg uncertainty principle, the kinetic energy increases when it is confined in one of the halves (Zurek, 1990).

Thermodynamics of information refines and generalizes this argument (Parrondo et al., 2015). First, by incorporating concepts from information theory that quantify the information acquired in a measurement in a more precise and general way. Second, by realizing that the states resulting from measurements and other processes involving information are in general non-equilibrium states. Hence, thermodynamics of information can be considered as a branch of statistical mechanics that deals with a special class of non-equilibrium states bearing information.

The outline of this article is the following. In section 2, we review the concepts of information theory relevant to thermodynamics, like Shannon entropy and mutual information. We discuss in detail the relationship between Shannon entropy and thermodynamic entropy in section 3. In section 4, we show how the acquisition of information affects the second law. The resulting modified second law for feedback processes —those that follows a protocol depending on the outcome of a measurement— is the answer to the first question that we have posed above. In section 5, we derive the thermodynamic cost of some processes involving information in memories and other information devices, like the celebrated Landauer principle (Landauer, 1961). These costs help us to solve the second question, i.e., the restoration of the original second law in the Szilárd engine and any generic isothermal feedback processes. This is discussed in section 6, where we show that the feedback processes obey the original second law when one accounts for the cost of measurement and/or restoring the demon's memory, generalizing Bennett's analysis of the Szilárd engine (Bennett, 1982). In section 7 we introduce a formalism that allows one to study information flows (Allahverdyan and Saakian, 2008; Horowitz and Esposito, 2014) in autonomous Maxwell demons. Finally, in section 8 we present our main conclusions and discuss the physical nature of information.

2. Information

In his seminal papers published in 1948 (Shannon, 1948), Shannon introduced a measure of the uncertainty or ignorance that one has about a random object or discrete random variable X, which can be a generic message in a communication channel, a file in a computer, or the microstate of a physical system. Shannon uncertainty, also known as Shannon entropy³, is defined as

$$S(X) = S[p_X] = -\sum_{x} p_X(x) \log p_X(x).$$
 (2)

Here we have introduced two useful equivalent notations, namely, expressing the dependency of the uncertainty S on the random variable X or on its probability distribution $p_X(x)$. Uncertainty can be expressed in bits, if the logarithm

in Eq. (2) is in base 2, in nats, if it is the natural logarithm, or in units of thermodynamic entropy if we multiply nats by, for instance, the Boltzmann constant k. In the rest of the article, we will assume that the Shannon entropy is expressed in physical units of energy divided by temperature, without the need of writing explicitly the Boltzmann constant. Finally, for continuous random variables, the sum in (2) must be replaced by an integral (Cover and Thomas, 2006).

It is customary to interpret Shannon uncertainty as a measure of information, and this is indeed appropriate in certain situations. For example, when we consider the information content of a given instance of the random variable X. However, when we talk about information, specially in physics, or, more precisely, about obtaining information from a system, we have in mind an inquiry about the state X of the system, which consists in the measurement of a quantity Y. When we measure Y, we acquire information about X that results in a decrease of its uncertainty. This decrease is precisely the amount of information that Y provides about X:

$$I(X;Y) \equiv S(X) - S(X|Y) \tag{3}$$

which is called *mutual information* (Cover and Thomas, 2006; Shannon, 1948). Here S(X|Y) is the uncertainty of the posterior probability distribution $p_{X|Y}(x|y)$ averaged over the possible outcomes $p_Y(y)$, i.e.

$$S(X|Y) \equiv \sum_{y} p_{Y}(y) \left[-\sum_{x} p_{X|Y}(x|y) \log p_{X|Y}(x|y) \right]$$
$$= -\sum_{x,y} p_{XY}(x,y) \log p_{X|Y}(x|y). \tag{4}$$

Using the relationship between conditional and joint probabilities (Bayes' formula): $p_{XY}(x, y) = p_{X|Y}(x|y)p_Y(y)$, we can write the mutual information as

$$I(X;Y) = \sum_{x,y} p_{XY}(x,y) \log \frac{p_{XY}(x,y)}{p_X(x)p_Y(y)}.$$
 (5)

From this expression we can derive a number of interesting properties of the mutual information. First, mutual information is symmetric, i.e., Y provides the same information about X as X provides about Y. Second, using the properties of the logarithm, one can prove from Eq. (5) that I(X;Y) is always positive and vanishes if and only if X and Y are statistically independent (Cover and Thomas, 2006). Hence, mutual information is a measure of the correlation between X and Y. Third, if we measure without errors a quantity Y = f(X), then S(Y|X) = 0 and I(X;Y) = S(Y). In other words, the information provided by an error-free measurement is the uncertainty of the outcome.

Finally, Eq. (5) allows us to write the mutual information in three different ways:

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

$$= S(Y) - S(Y|X)$$

$$= S(X) + S(Y) - S(X,Y) \ge 0$$
 (6)

³In a famous conversation between Shannon and Von Neumann, the latter suggested Shannon to use the name entropy because (2) is similar to the definition of entropy in statistical mechanics and, more importantly, because "nobody really knows what entropy is". The conversation was reported by Shannon himself but later on he admitted that it probably never took place (Price, 1982).

where S(X,Y) is the Shannon entropy or uncertainty of the joint probability distribution $p_{XY}(x,y)$. The last equality indicates that correlations make the entropy sub-additive:

$$S(X,Y) = S(X) + S(Y) - I(X;Y).$$
 (7)

This expression will be useful when we consider the physical nature of Maxwell or Szilárd demons and interpret a measurement as the creation of correlations between the state of the demon and the state of the system.

3. Shannon and thermodynamic entropies

If Shannon uncertainty is identified as thermodynamic entropy, one immediately obtains from the definition of mutual information (3) that the entropy of a system X decreases by an amount I(X; M) when we measure a quantity M. For instance, in the case of the Szilárd engine with error-free measurements, X is the micro-state of the global system, particle plus bath, which obviously includes the position of the particle, and M = L, R (L for left, and R for right) is the outcome of the measurement, i.e., the half of the container where the particle lies after the insertion of the piston. Since M is a function of X and the two outcomes are equally probable, $I(X; M) = S(M) = k \ln 2$. Thus, the total entropy decreases by $k \ln 2$, and this decrease allows one to extract an energy $kT \ln 2$ from the thermal bath in the form of work by driving quasi-statically the system back to its initial state. This is a similar argument as the one that we sketched in section 1.

This explanation, however, requires a more careful consideration. The identification of Shannon and thermodynamic entropies is not correct, in general. The main reason is that the Shannon entropy $S[\rho(x,t)]$ of the probabilistic state of a classical system is invariant under Hamiltonian evolution. Here x denotes a micro-state of the system, i.e., the value of the positions and momenta of all its particles, and $\rho(x,t)$ is the probability density of observing a microstate x at time t. The same argument holds for quantum systems, replacing $\rho(x,t)$ by the density matrix $\rho(t)$ and the Shannon entropy by Von Neumann entropy, $\text{Tr}(\rho \ln \rho)$ (Schumacher and Westmoreland, 2010). Here, we will restrict our discussion to classical systems, although it is not difficult to extend it to quantum systems (Esposito et al., 2010).

To observe an increase of Shannon entropy, as dictated by the second law for irreversible processes, one has to consider information losses that degrade the probabilistic state $\rho(x,t)$, like coarse-graining or the inaccessibility of detailed information about certain degrees of freedom.

A second important property of thermodynamic entropy is its connection with the energetics of a process through Clausius equation, which relates the temperature T to the increase of the entropy, δS , and the energy, δE , of a system⁴:

$$\delta S = \frac{\delta E}{T}.\tag{8}$$

This equation is crucial to find the limitations that the second law imposes on energy extraction from thermal reservoirs. For instance, if a system is in contact with a thermal bath at temperature T, according to Eq. (8) the total entropy production in a process is

$$\Delta S_{\text{tot}} = \Delta S - \frac{Q}{T} \ge 0 \tag{9}$$

where ΔS is the change of entropy in the system and Q is the energy transferred from the bath to the system. If an external agent manipulates the system preforming a work W, the change of energy in the system is $\Delta E = Q + W$ and

$$T\Delta S_{\text{tot}} = T\Delta S - \Delta E + W = W - \Delta F \tag{10}$$

 ΔF being the increase of free energy, F=E-TS, in the system. The second law $\Delta S_{\rm tot} \geq 0$ for this isothermal process is equivalent to

$$W \ge \Delta F. \tag{11}$$

In particular, for a cycle $\Delta F = 0$ and the inequality, $W \ge 0$, tells us that it is impossible to extract energy from a thermal bath in a cyclic process.

For a system in contact with equilibrium thermal reservoirs, one can solve these two issues —namely, the invariance of Shannon entropy under Hamiltonian evolution and its connection with the energetics of a process— with the two following assumptions that specify the information available to an external observer (Esposito et al., 2010): first, the observer does not have access to the correlations between the system and the baths and, second, the only information about the state of the baths is their average energy. More precisely, the baths are described by thermal states with given average energies. In the case of a single bath with micro-states z, if the actual probabilistic state of the global system at time t is $\rho(x,z;t)$, this loss of information yields an effective state $\rho_{\rm obs}(x,z;t)$ given by

$$\rho(x, z; t) \to \rho_{\text{obs}}(x, z; t) = \rho(x, t) \rho_{\text{B eq}}(z; T(t)) \quad (12)$$

where $\rho(x, t)$ is the marginal probabilistic state of the system

$$\rho(x,t) = \int dz \, \rho(x,z;t) \tag{13}$$

and $\rho_{\mathrm{B,eq}}(z;T)$ is the thermal state of the bath

$$\rho_{\text{B,eq}}(z;T) = \frac{e^{-\beta H_{\text{B}}(z)}}{Z_{\text{B}}(\beta)}.$$
(14)

Here $H_{\rm B}(z)$ is the Hamiltonian of the bath, $\beta=1/(kT)$ the inverse temperature, and $Z_{\rm B}(\beta)$ the partition function. The temperature T(t) in (12) is set as the one that verifies

$$E_{\rm B}(t) \equiv \int dx dz \, H_{\rm B}(z) \, \rho(x,z;t)$$

⁴Clausius originaly introduced this equation as a definition of entropy. However, it is in fact a definition of temperature for systems at equilibrium, since energy and entropy are more fundamental magnitudes that can be obtained from dynamical properties, like the Hamiltonian and the phase space volume of regions of constant energy.

$$= \int dz H_{\rm B}(z) \rho_{\rm B,eq}(z;T(t)), \qquad (15)$$

i.e., both distributions have the same average bath energy $E_{\rm R}(t)$ at any time t.

The replacement of the actual state $\rho(x, z; t)$ by the observable state $\rho_{\text{obs}}(x, z; t)$ solves the two aforementioned issues: the invariance of Shannon entropy under Hamiltonian evolution and the relationship between entropy and energy given by the Clausius equation (8). First, applying (7), we find that the replacement (12) increases the Shannon entropy:

$$S[\rho_{\rm obs}(x,z;t)] - S[\rho(x,z;t)] = I(X;Z) + \Delta S_{\rm B} \ge 0 \ (16)$$

where $\Delta S_{\rm B} = S[\rho_{\rm B,eq}(z;T(t))] - S[\rho(z;t)]$. The inequality follows from the positiveness of the mutual information and the fact that the thermal sate (14) is the one that maximizes the Shannon entropy under condition (15).

Consider a process where we manipulate the system, whose Hamiltonian H(x;t) is now time-dependent. We also assume that initially system and bath are uncorrelated and the latter is at equilibrium, i.e., $\rho(x,z;0) = \rho(x,0)\rho_{\rm B,eq}(z;T(0))$. Then, $\rho(x,z;0) = \rho_{\rm obs}(x,z;0)$ and, since Shannon entropy is invariant under the Hamiltonian evolution, $S[\rho(x,z;t)] = S[\rho_{\rm obs}(x,z;0)]$. Hence, from Eq. (16), we obtain the entropy production in the process:

$$S[\rho_{\text{obs}}(x, z; t)] - S[\rho_{\text{obs}}(x, z; 0)] = I(X; Z) + \Delta S_{\text{B}} \ge 0.$$
(17)

This expression reflects the two information losses that we have introduced in (12): I(X;Z) is the information associated to the correlations between the system and the bath, whereas $\Delta S_{\rm B}$ is the increase of Shannon entropy that results from replacing the actual state of the bath by a thermal state (Esposito et al., 2010). Each of these terms is positive and we recover a second law for the Shannon entropy $S[\rho_{\rm obs}(x,z;t)]$. It is interesting to notice that both terms can be relevant in realistic situations, as shown in (Ptaszyński and Esposito, 2019).

Alternatively, the entropy production can also be calculated by using the explicit form of the thermal state (14):

$$S[\rho_{\rm obs}(x,z,t)] = S[\rho(x,t)] + k\beta(t)E_{\rm B}(t) + k\ln Z_{\rm B}(\beta(t)). \eqno(18)$$

Since $E_{\rm B} = -\partial \ln Z_{\rm B}/\partial \beta$, the time derivative of the Shannon entropy of the observable state can be written as

$$\dot{S}[\rho_{\text{obs}}(x, z, t)] = \dot{S}[\rho(x, t)] + \frac{\dot{E}_{B}(t)}{T(t)}.$$
 (19)

Integrating over the whole process with the initial condition discussed above, and taking into account that the bath is large enough so that the change of temperature is negligible⁵, we

finally get

$$S[\rho_{\text{obs}}(x, z; t)] - S[\rho_{\text{obs}}(x, z; 0)] = \Delta S - \frac{Q}{T}$$
 (20)

where $Q=E_{\rm B}(0)-E_{\rm B}(t)$ is the net energy transferred from the bath to the system. From Eq. (17), we conclude that the entropy production, as expressed in Eq. (20), is positive. Thus, we recover the same expression for the entropy production as in (9), i.e., the loss of information implied by (12) is compatible with Clausius equation (8). The standard thermodynamic argument to derive (11) can now be applied to the *non-equilibrium free energy*

$$\mathcal{F}(\rho, H) \equiv \langle H(x) \rangle - TS[\rho(x)] \tag{21}$$

yielding a bound to the work needed to complete an isothermal process:

$$W \ge \Delta \mathcal{F}. \tag{22}$$

This can be considered an extension of the second law (11) for isothermal processes involving non-equilibrium states, and reduces to the standard second law if the initial and final states are equilibrium, since the non-equilibrium free energy (21) is equal to the standard equilibrium one for thermal states like (14). Alternatively to our derivation based on information losses, (22) can be directly obtained from a master equation obeying local detailed balance (Esposito and Van den Broeck, 2011), which is the equation that describes systems in contact with reservoirs in the Markovian approximation.

The bound (22) is saturated for operationally reversible process, i.e., for processes where the system visits the same states as in the forward process when the protocol is reversed (Parrondo et al., 2015). This observation puts into question the utility of (22) for generic non-equilibrium states, since any irreversible relaxation to equilibrium prevents the bound to be reached. One strategy to overcome this issue consists in modifying instantaneously the Hamiltonian to convert the initial and final states into equilibrium states, although this protocol is a bit artificial and sometimes unrealizable ⁶ (Parrondo et al., 2015). However, the bound (22) can be tight as well for systems with slow and fast degrees of freedom. If there is a large separation of time scales and the fast degrees of freedom are equilibrated during the process, then the inequality (22) is met. As we will see in section 5, it turns out that this is the case of many relevant processes involving information.

4. Second law for feedback processes

In an isothermal feedback process, like the Szilárd engine, the external agent measures a magnitude M at a given time $t_{\rm m}$ and uses this information to complete an isothermal

⁵Notice that we have to consider a time-dependent temperature to derive the variation of Shannon entropy in the bath in Eq. (19). However, this variation can be neglected when we integrate this equation, if the heat capacity of the bath is large enough.

⁶Any state $\rho(x)$ can be converted into an equilibrium state by setting the Hamiltonian equal to $H(x) = -kT \ln \rho(x)$. Notice however that such Hamiltonian can be difficult or impossible to realize as, for instance, when $\ln \rho(x)$ is not a quadratic function of velocities.

process. The measurement is characterized by the conditional probability $p_{M|X}(m|x)$, where m is the measurement outcome and x is the microscopic state of the system. This includes error-free measurements, where $p_{M|X}(m|x)=1$ if m=m(x) and zero otherwise, but it can also describe measurements affected by noise or any other source of inaccuracy. Using the measurement outcome, the observer updates the probabilistic state of the system $\rho_X(x,t_{\rm m})$ according to Bayes rule

$$\rho_X(x, t_{\rm m}) \to \rho_{X|M}(x|m) = \frac{p_{M|X}(m|x)\rho_X(x, t_{\rm m})}{p_M(m)}.$$
 (23)

Here we are assuming the most common case, where x is a continuous random variable and m is discrete, and use the Latin letter p for probability distributions and the Greek one ρ for probability densities. Nevertheless, the argument can be easily generalized to other cases. For a given outcome m, the updated state $\rho_{X|M}(x|m)$ is in general a non-equilibrium state, even if the state before the measurement, $\rho_X(x,t_{\rm m})$, is equilibrium. We see here the first appearance of non-equilibrium states in processes involving information. The unconditional state of the system after the measurement is again

$$\sum_{m} p_{M}(m) \rho_{X|M}(x|m) = \rho_{X}(x, t_{m}), \tag{24}$$

i.e., the measurement does not alter the system. In section 6, we will establish this as a condition for an ideal classical measurement. Given an outcome m, the free energy after the measurement is

$$\mathcal{F}[\rho_{X|M}(x|m), H(x, t_{\rm m})]$$

$$= \int dx H(x, t_{\rm m}) \rho_{X|M}(x|m) - TS[\rho_{X|M}(x|m)]$$
(25)

where H(x,t) is the Hamiltonian of the system at time t. Averaging over the possible outcomes we get the non-equilibrium free energy after the measurement

$$\begin{aligned} \mathcal{F}_{\text{post}} &= \sum_{m} p_{M}(m) \mathcal{F}[\rho_{X|M}(x|m), H(x, t_{\text{m}})] \\ &= \int dx H(x, t_{\text{m}}) \rho(x, t_{\text{m}}) - TS(X|M). \end{aligned} \tag{26}$$

Using the definition of mutual information (7), we obtain that, as a consequence of the Bayes update, the non-equilibrium free energy of the system changes as

$$\Delta \mathcal{F}_{\text{meas}} \equiv \mathcal{F}_{\text{post}} - \mathcal{F}_{\text{pre}}$$

$$= T[(S(X) - S(X|M)]$$

$$= TI(X; M), \qquad (27)$$

where $\mathcal{F}_{pre} = \mathcal{F}[\rho(x,t_{\rm m}),H(x,t_{\rm m})]$ is the non-equilibrium free energy of the system immediately before the measurment.

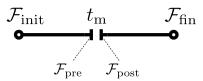


Figure 3: Scheme of a basic feedback process with a single measurement.

We can now apply Eq. (22) to the subprocesses that take place before and after the measurement. The total work in the feedback process is bound as

$$W_{\rm fb} \geq \mathcal{F}_{\rm fin} - \mathcal{F}_{\rm post} + \mathcal{F}_{\rm pre} - \mathcal{F}_{\rm init}$$

= $\Delta \mathcal{F} - \Delta \mathcal{F}_{\rm meas}$ (28)

where $\Delta \mathcal{F} = \mathcal{F}_{fin} - \mathcal{F}_{init}$ and the subscripts, init, pre, post, fin, indicate respectively the states at the different stages of the process: initial and final, and immediately before and after the measurement (see Fig. 3). Combining (27) and (28), we finally obtain

$$W_{\text{fb}} \ge \Delta \mathcal{F} - TI(X; M)$$
 (29)

which is the second law for feedback processes. This result was derived by Sagawa and Ueda for classical and quantum systems (Sagawa and Ueda, 2008), although the relevance of mutual information in thermodynamics was first pointed out by Lloyd and Touchette (Lloyd, 1989; Touchette and Lloyd, 2000). It indicates that the work necessary to complete a process decreases by an amount TI(X; M) if we use the information gathered in the measurement. In particular, for a cycle the work can be negative, i.e., we can extract a work $W_{\text{extract}} = -W_{\text{fb}}$ at least equal to TI(X; M). The Szilárd engine, where $I(X; M) = k \ln 2$ is an explicit example that saturates the bound.

Notice that, due to (24), the unconditional non-equilibrium free energy of the system does not change in the measurement. The mutual information I(X; M) in (27) and (29) appears because we average the conditional entropy over $p_M(m)$. The reason of averaging in this way is that, in a feedback process, the protocol depends on the outcome m, and so does the work. The bound (29) applies to the average work W_{fb} over all possible outcomes and protocols, each one obeying (22). If the information provided by the measurement is not used in the protocol from t_m to the final time of the process, then we recover the standard second law (22) with equilibrium free energies.

The bound (29) does not resolve the problem of reconciling the Szilárd engine with the original second law. However, it is an important result because it establishes a benchmark to the work that can be extracted in a feedback process. It also indicates that the mutual information I(X; M) is a thermodynamic resource and can be used to define efficiencies for information and hybrid engines (Saha et al., 2021; Schmitt et al., 2015) and to compare the performance of information motors and of chemical or thermal machines (Horowitz et al., 2013). It is also interesting to explore

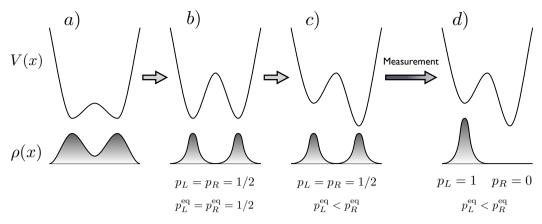


Figure 4: Informational states of a Brownian degree of freedom in a bistable potential. The rise of the barrier from a) to b) illustrates the creation of information (Roldán et al., 2014), whereas c) and d) show non-equilibrium probabilistic informational states, which are the result of the history and of a measurement, respectively.

which protocols saturate the bound (29). It is found that this happens for operationally reversible processes, those which visit the same states when the action of the external agent is reversed (Horowitz and Parrondo, 2011a). Notice however that the notion of reversibility of feedback processes requires a more careful discussion since it is not trivial to define the time reversal of a measurement (Horowitz and Parrondo, 2011b; Horowitz and Vaikuntanathan, 2010).

The second law (29) can be further refined by deriving fluctuation theorems. These theorems are exact equalities that hold for any process even far from equilibrium, and from which the second law can be obtained as a corollary (Horowitz and Vaikuntanathan, 2010; Ponmurugan, 2010; Sagawa and Ueda, 2009, 2010, 2012).

5. Informational states and Landauer's principle

Information devices, like memories, processors, or the biochemical machinery of DNA and RNA, are systems that can adopt different informational states that have a long lifetime and are interchangeable. These states are, for example, the bits in a hard drive, which can be 0 or 1, or the bases of DNA, which can be any of the four nucleotides: cytosine, guanine, adenine or thymine.

These two properties, long lifetimes and interchangeability, imply some kind of effective symmetry breaking in the system: its phase space is partitioned into several regions where the system remains for a very long time. Each region is an informational state and the system is no longer ergodic in the whole phase space. Here, we refer to an effective lack of ergodicity when there is a huge separation between the time scales of the dynamics within the regions and of the jumps between regions. This type of effective ergodicity breaking occurs due to phase transitions in finite systems or to the presence of high energy barriers with a negligible probability to be crossed. A paradigmatic example

is a Brownian degree of freedom in a double well potential, like the one depicted in figure 4.

In general, the phase space Γ of an information device is partitioned into a number of regions Γ_m , with m = $1, 2, \dots, \mathcal{M}$. We will assume that the dynamics within each region is very fast and that the system is in local equilibrium within every Γ_m . On the other hand, the system can be in any of the informational states with a certain probability p_m that depends on the history, measurements, etc. We show an explicit example in figure 4, based on a Brownian particle or degree of freedom in a double well potential. By raising the barrier far above kT, we create a memory from a) to b). If the potential is symmetric, each informational state m = L, R is populated with the same probability $p_L = p_R = 1/2$ and the system is in global equilibrium. However, if we change the energy of the wells, as in c), populations do not change due to the high barrier but the equilibrium probabilities $p_{R,L}^{eq}$ depart from 1/2. Hence, the system is no longer in global equilibrium. The informational state can also change due to a measurement, as in the transition from c) to d).

If the device is at temperature T, the state of the system is then given by

$$\rho(x) = \sum_{m} p_m \frac{e^{-\beta H(x)}}{Z_m} \chi_m(x)$$
 (30)

where $x \in \Gamma$ is a micro-state, H(x) is the Hamiltonian, $\chi_m(x)$ is the index function of Γ_m , i.e., $\chi_m(x) = 1$ if $x \in \Gamma_m$ and zero otherwise, and Z_m is the restricted partition function

$$Z_m \equiv \int_{\Gamma_m} dx \, e^{-\beta H(x)}. \tag{31}$$

The probability distribution $\{p_m\}$, with $m=1,2,\ldots,\mathcal{M}$ is arbitrary and defines a probabilistic informational state. The main characteristic of the information device is that it can adopt any probabilistic informational state and that an external agent can drive the system, as in figure 4, from

one probabilistic informational state to another, writing or processing information in the device. A special probabilistic informational state is the one corresponding to global equilibrium $p_m = Z_m/Z$, where $Z = \sum_m Z_m$ is the global partition function.

The energetics of processes involving informational states can be analyzed using the non-equilibrium free energy (21), which for state (30) reads

$$\mathcal{F}(\rho, H) = \sum_{m} p_{m} F_{m} - TS(p_{m}) \tag{32}$$

where $F_m = -kT \ln Z_m$ is the restricted free energy and $S(p_m)$ is the Shannon entropy of the probabilistic informational state. As already mentioned, the non-equilibrium free energy equals the equilibrium one, $\mathcal{F}(\rho,H) = -kT \ln Z$ if the system is in global equilibrium. For a symmetric memory where $F_m = F_{\text{local}}$ is the same for all informational states m, we have

$$\mathcal{F}(\rho, H) = F_{local} - TS(p_m). \tag{33}$$

Now suppose that we manipulate a symmetric memory driving the system isothermally from a distribution p_m to p'_m . If the initial and final restricted free energies are equal, then the work needed to complete the process is bound by

$$W \ge T[S(p_m) - S(p'_m)]. \tag{34}$$

This equation tells us that we have to perform work to order a memory, $S(p'_m) < S(p_m)$, whereas we can extract work by disordering a symmetric memory $S(p'_m) > S(p_m)$.

Landauer's principle is a special case of Eq. (34). Landauer considered a simple process where a binary symmetric memory is initially in a random state $p_0 = p_1 = 1/2$, with $S(p_m) = k \ln 2$ and is driven to the informational state 0 with probability one, i.e., $p_0' = 1$ and $p_1' = 0$, with zero Shannon entropy. This process is called restore-to-zero, and is usually considered as the erasure of the initial random bit, although it is more appropriate to call it overwriting, since the initial unknown bit is replaced by a given one, 0 in our example. The work needed to complete Landauer's overwriting is

$$W \ge kT \ln 2 \tag{35}$$

which is the so-called Landauer's principle. The Landauer principle has been confirmed experimentally in different systems, like Brownian particles in optical traps (Berut et al., 2011) (see (Proesmans et al., 2020) for a discussion on experimental results on the energetics of overwriting at finite speed). Landauer's overwriting is a special case of a logically irreversible operation, since the initial informational state of the memory (input) cannot be recovered from the final one (output). It is a common misunderstanding to interpret Landauer's principle as if logical irreversibility implied thermodynamic irreversibility. This is not true. If the inequality (35) is met, the overwriting process is still logically irreversible but thermodynamically reversible because the decrease of entropy in the system is compensated by the increase of

entropy in the thermal bath due to heat dissipation. Hence, the total entropy production is zero. See (Sagawa, 2014) for a detailed discussion on the relation between thermodynamic and logical reversibility.

On the other hand, if a memory has a low Shannon entropy, one can extract work by disordering it. The thermodynamics of these information reservoirs that can act as thermodynamic resources has been analyzed in detail by Barato and Seifert (Barato and Seifert, 2014), and by Wolpert (Wolpert, 2019). Mandal and Jarzynski have devised a concrete realization of work extraction from an ordered memory using a kinetic model that reads the content of a tape (Mandal and Jarzynski, 2012).

6. Restoring the second law

We can now address the question of restoring the original second law in the Szilárd engine by considering the physical nature of the demon, who must possess a memory to register the measurement outcomes and is subjected to the thermodynamic limitations discussed in the previous section.

First, let us consider the energetics of a measurement. An ideal classical measurement consists of the interaction between a system and a measurement apparatus, which fulfills the following requirements: i) system and apparatus are initially uncorrelated; ii) the system is not affected by the interaction; and iii) system and apparatus do not interact before and after the measurement, i.e., the Hamiltonian of the global system is $H_{\text{sys}}(x) + H_{\text{app}}(y)$, where x and y are microstates of the system and the apparatus, respectively. Let Y and Y' be the random variables denoting the state of the apparatus before and after the measurement, respectively, and $p_{M|X}(m|x)$ the conditional probability characterizing the measurement. We will further assume that the outcome is a function of the apparatus micro-state M = m(Y') and that all the information about X is provided by the measurement outcome, i.e., $\rho_{X|Y}(x|y) = \rho_{X|M}(x|m(y))$.

If the initial state of the global system immediately before the measurement is

$$\rho_{XY}(x, y) = \rho_X(x)\rho_Y(y) \tag{36}$$

with non-equilibrium free energy

$$\mathcal{F}(X,Y) = \mathcal{F}(X) + \mathcal{F}(Y), \tag{37}$$

then the state after the measurement is

$$\rho_{XY'}(x,y) = \sum_{m} \rho_{X}(x) p_{M|X}(m|x) \rho_{Y'|M}(y|m)$$
 (38)

where $\rho_{Y'|M}(y|m) = \rho_{Y'}(y)/p_M(m)$ if m = m(y) and zero otherwise. The marginal probability density for the state of the system $\rho_X(x)$ does not change a s a consequence of the measurement. This is why we can still denote the random variable corresponding to the state of the system as X, whereas the apparatus changes from Y to Y' (see Fig. 5). With these assumptions, it is not hard to prove that I(X;Y') = I(X;M), since Y' provides information

about X through the outcome M. Since the energy of the system does not change due to the measurement, the non-equilibrium free energy of the global system (X, Y') immediately after the measurement can be written as

$$\mathcal{F}(X,Y') = \mathcal{F}(X) + \mathcal{F}(Y') + TI(X;Y')$$
$$= \mathcal{F}(X) + \mathcal{F}(Y') + TI(X;M)$$
(39)

where we have used Eq. (7) to express the total Shannon entropy S(X, Y') in terms of the mutual information I(X; Y'). Hence, the work needed to complete the measurement is

$$W_{\text{meas}} \ge \Delta \mathcal{F} = \Delta \mathcal{F}_Y + TI(X; M)$$
 (40)

where $\Delta \mathcal{F}_Y = \mathcal{F}(Y') - \mathcal{F}(Y)$. Since I(X; M) is positive, we see that measuring or, more generally, creating correlations between two systems, increases the free energy and, if not counterbalanced by $\Delta \mathcal{F}_Y$, needs work and heat dissipation to be completed.

As discussed in section 4, the demon can extract a work $W_{\text{extract}} = -W_{\text{fb}}$ with $W_{\text{fb}} \geq TI(X; M)$ if he uses the information acquired in the measurement in a cyclic process, where the system is driven back to its initial state X.

However, to complete the cycle, the apparatus must be also driven to its initial state *Y*. In doing so, the demon must perform a work

$$W_{\text{reset}} \ge \mathcal{F}(Y) - \mathcal{F}(Y') = -\Delta \mathcal{F}_Y.$$
 (41)

The total work in the process is then

$$W_{\text{tot}} = W_{\text{meas}} + W_{\text{fh}} + W_{\text{reset}} \ge 0. \tag{42}$$

Hence, the validity of the second law for feedback processes is restored when the entropy costs of measurement and/or resetting the demon's memory are taken into account. This is the generalization of Bennett's analysis of the Szilard engine. Bennett discussed in (Bennett, 1982) a realization of the Szilard engine where $W_{\rm meas}=0$ and the demon must overwrite the outcome of the measurement performing a work $kT \ln 2$, as dictated by Landauer's principle. This is achieved if $\Delta \mathcal{F}_Y = -TI(X; M)$.

From the previous discussion, we see that the Szilárd engine can be interpreted as a relatively simple exchange between work and the free energy TI(X; M) stored in the correlations between the system and the demon. For instance, if $\Delta \mathcal{F}_Y = 0$, the engine consists of creating correlations in the measurement, an operation that requires a work TI(X; M) and increases the free energy by the same amount, and then destroying these correlations in the feedback, where the same amount of work is now extracted.

The previous discussion can be presented in terms of entropy, instead of free energy, and extended to continuous time. Consider now that both the system and the demon are random processes, X(t) and Y(t), respectively. The time derivative of Eq. (7) reads

$$\dot{S}(X(t), Y(t)) = \dot{S}(X(t)) + \dot{S}(Y(t)) - \dot{I}(X(t); Y(t)) \tag{43}$$

and the total entropy production can be written as

$$\dot{S}_{\text{prod}} = \dot{S}(X(t)) + \dot{S}(Y(t)) - \dot{I}(X(t); Y(t)) + \dot{S}_{\text{env}}(t) \ge 0$$
 (44)

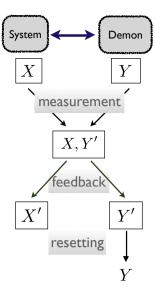


Figure 5: Sketch of the operations in a feedback cycle.

where $\dot{S}_{\rm env}(t)$ is the change per unit of time of the entropy of the environment at time t. For instance, if the system and/or the demon are in contact with several thermal baths at different temperatures T_i , then $\dot{S}_{\rm env}(t) = -\sum_i \dot{Q}_i(t)/T_i$, where $Q_i(t)$ is the energy transfer per unit of time from the i-th bath to the system.

The scheme of figure 5 can be seen as the different stages of a continuous evolution. In the measurement, the system does not change, hence $\dot{S}(X(t)) = 0$, and the integration of (44) yields

$$\Delta S_Y - \Delta I + \Delta S_{\text{env}}^{(\text{meas})} \ge 0. \tag{45}$$

Here, ΔI is the mutual information developed in the measurement. Notice however that the measurement in this description is any interaction between X and Y that creates correlations between the two systems, keeping X constant. During feedback, the demon does not change, hence $\dot{S}(Y(t)) = 0$. If the feedback process is cyclic, $\Delta S_X = 0$, and if all the correlations created during the measurement are destroyed, then

$$\Delta I + \Delta S_{\text{env}}^{\text{(fb)}} \ge 0. \tag{46}$$

Finally, during the demon resetting:

$$-\Delta S_Y + \Delta S_{\text{env}}^{(\text{reset})} \ge 0. \tag{47}$$

The sum of the three inequalities (45-47) yields the standard second law for the total entropy production

$$S_{\rm prod} = \Delta S_{\rm env}^{\rm (meas)} + \Delta S_{\rm env}^{\rm (fb)} + \Delta S_{\rm env}^{\rm (reset)} \ge 0. \eqno(48)$$

Therefore, Eq. (44) comprises all our discussion on the Szilárd engine but also extends to generic dynamics where two systems evolve in time creating and destroying correlations (Horowitz and Esposito, 2014).

7. Information flows

The previous discussion holds for an arbitrary evolution. However, for (44) to be meaningful, the evolution must be driven by an external agent (without feedback), which switches on and off the coupling between the system and the demon in the measurement and resets the demon to its initial state. Otherwise, the global system reaches a steady state where all terms in Eq. (44) vanish in average. Still, it is interesting to interpret as information engines autonomous motors that work in a non-equilibrium stationary state (NESS). For this purpose, it is useful to introduce the notion of *information flows* (Allahverdyan et al., 2009; Horowitz, 2015; Horowitz and Esposito, 2014):

$$\dot{I}_X(t) = \lim_{\tau \to 0^+} \frac{I(X(t+\tau); Y(t)) - I(X(t); Y(t))}{\tau} \quad (49)$$

and

$$\dot{I}_{Y}(t) = \lim_{\tau \to 0^{+}} \frac{I(X(t); Y(t+\tau)) - I(X(t); Y(t))}{\tau}.$$
 (50)

In the NESS, $\dot{I}_X(t) + \dot{I}_Y(t) = \dot{I}(X(t); Y(t)) = 0$, and $\dot{I}_X(t) = -\dot{I}_Y(t)$. This is why these quantities can be considered flows of a conserved quantity, although this interpretation is valid only in the stationary regime.

If $\dot{I}_X(t)$ is positive then the motion of X, with Y fixed, increases the correlation between the two systems. Accordingly to our discussion in the previous section, we can interpret it as X measuring Y. On the other hand, if $\dot{I}_X(t) < 0$, the motion of system X destroys correlations, like a demon using information to complete a feedback process and converting the free energy stored in correlations into work.

When each system is coupled to separate baths, one can derive a second law for the entropy production in the stationary regime due to the dynamics of each system, *X* or *Y*:

$$-\dot{I}_X + \dot{S}_{\text{env}}^X \ge 0$$

$$\dot{I}_X + \dot{S}_{\text{env}}^Y \ge 0$$
(51)

where $\dot{S}_{\mathrm{env}}^{X,Y}$ is the entropy increase in the reservoir connected to system X and Y, respectively. These partial second laws have been derived for multipartite continuous and discrete Markovian systems (Horowitz and Esposito, 2014; Horowitz, 2015) and they allow us to analyze the energetics of autonomous motors from the point of view of information flows. In Ref. (Horowitz and Esposito, 2014), Horowitz and Esposito apply it to two coupled quantum dots in contact with reservoirs with different temperatures and/or chemical potentials, a device that can operate as a pump or a refrigerator. Information flows and Eq. (51) do not only provide an interesting interpretation of one of the dots acting as a Maxwell demon on the other, but are also useful to derive limitations to the energetics that go beyond the application of the standard second law to the whole device. It is possible as well to evaluate separately the efficiency of the exchange between entropy and information in each dot (Horowitz and

Esposito, 2014). See also (Rosinberg and Horowitz, 2016) for an example of continuous classical degrees of freedom and (Ptaszyński and Esposito, 2019) for an extension to quantum systems.

Information flows are related to other magnitudes that quantify the exchange of information among different systems. In the context of data analysis, Schreiber introduced the so-called *transfer entropy* (Schreiber, 2000) to find causal relationships between two time series, whereas Liang and Kleeman (Liang and Kleeman, 2005) studied similar information transfers in dynamical systems. Although transfer entropies have been criticized as a tool to find causal relations (James et al., 2016), they have been proven useful to derive fluctuation theorems in bipartite systems (Hartich et al., 2014) and general causal networks (Ito and Sagawa, 2013). The relationship between transfer entropies and information flows has been studied in detail in (Allahverdyan et al., 2009; Cafaro et al., 2016; Horowitz, 2015; ?; Kiwata, 2022).

8. Conclusions: what is information?

Thermodynamics of information makes a precise assessment of the effect of information on the second law and clarifies the thermodynamic cost of information operations. The key concept is the mutual information between the state of a system and the outcome of a measurement since, as shown in (7), it is equal to the decrease of Shannon entropy due to the measurement, but also to the difference, due to correlations, between the entropy of a composite system and the sum of the entropies of its parts. To relate Shannon entropy with the energetics of a process, we have introduced in section 3 the loss of information that one would expect when a system is in contact with reservoirs at equilibrium. This analysis yields the definition of another relevant concept: the non-equilibrium free energy (21), which determines the minimal work necessary to complete an isothermal process connecting non-equilibrium states.

We have shown that the states relevant to information processing are out of equilibrium, either because they are the result of a Bayesian update, like (23), or because they describe memories with arbitrary probabilistic informational states, as in (30). In both cases, non-equilibrium free energy establishes limitations to the energetics of isothermal processes involving information. These limitations are extensions of the second law to informational non-equilibrium states. In this framework, information appears as a novel thermodynamic resource and those extensions of the second law are indispensable to assess the efficiency of its exploitation

We have also used non-equilibrium free energy to prove that the Szilárd engine and any isothermal feedback process are compatible with the standard second law of thermodynamics. For this purpose, we have considered the physical nature of the demon and, in doing so, we not only solve the fundamental question posed by Maxwell more than one century ago, but open as well the possibility of interpreting certain devices like motors, refrigerators, and other chemical or thermal machines, as transducers that convert information into work and vice-versa. For this interpretation to hold, thermal fluctuations and correlations must be essential ingredients in the functioning of these machines. This is why information thermodynamics is closely related to — or can even be considered as a branch of— a recent field of research, stochastic thermodynamics, which studies the energetics of processes at the micro- and nano-scale, incorporating fluctuations (Peliti and Pigolotti, 2021; Sekimoto, 2010).

In this article, we have restricted ourselves mostly to classical systems. The reason is that, in an isothermal process, a quantum system rapidly adopts a state which is diagonal in the eigenbasis of the Hamiltonian. In these scenarios, coherences are lost and the system can be studied with the same tools as those developed for classical systems. In fact, most of the concepts that we have introduced in the previous sections can be straightforwardly extended to quantum systems. Nevertheless, in the last years, there has been an increasing interest in phenomena that combine in a nontrivial way quantum and thermal effects (Anders and Esposito, 2017). Zurek (Zurek, 1990) and Lloyd (Lloyd, 1997) already found distinctive aspects of the quantum Szilárd engine and Kim et al explored the effect of quantum statistics in multiparticle Szilárd cycles with bosons and fermions (Kim et al., 2011). The mutual information between quantum systems can be split into a classical and a purely quantum part due to entanglement, known as quantum discord, and whose role in quantum Maxwell demons have been analyzed in (Park et al., 2013; Zurek, 2003). Finally, some setups involving the coupling between two quantum states can be interpreted as Maxwell demons that use information. This is the case of the experiment reported in (Cottet et al., 2017) where the coupling between a transmon qubit and the radiation in a microwave cavity can be used to extract work.

From a more fundamental point of view, the interplay between information and matter is nowadays considered a crucial issue for our understanding of the physical world. In 1992, Wheeler summarized in the celebrated lemma "it from bit" (Wheeler, 1995) the idea that information is the basic concept upon which every physical theory should be built. Since then, a number of physicists and philosophers have attempted to follow this research program. Thermodynamics of information, on the other hand, pursues more modest goals, but it does provide some insight into how information is implemented in physical systems. In this article, we have seen two basic aspects of the physical nature of information. The first one is the characteristics of the informational states introduced in section 5, namely, long lifetimes and interchangeability. The second one is the interpretation of information in feedback engines as an exchange between work and the free energy TI(X; Y) stored in the correlations between two systems, X and Y. Each of these two aspects has yielded interesting lines of research. The first one explores the energetics of processes in systems where there is a huge separation of time scales (Roldán et al., 2014), whereas

the second one has inspired the notion of information flows in autonomous systems (Horowitz and Esposito, 2014). We believe that any reflection about information as a fundamental concept in physics should account for and/or utilize these two aspects.

Acknowledgements

This article summarizes work done over more than ten years and it would not have been possible without the interaction with many colleagues, friends, and students from different summer schools. A special mention deserve Jordan Horowitz, with whom I have done a great part of my contributions to the field, and Takahiro Sagawa, who has been one of the most active and influencing researchers in the recent application of stochastic thermodynamics to information processing. Together with Jordan and Takahiro, we wrote a review article (Parrondo et al., 2015), which shaped most of my ideas on the thermodynamics of information. The preparation of this review has been financially supported by the Spanish Government (Grant FLUID. Ref: PID2020-113455GB-I00) and the Foundational Questions Institute (FQXi) under the program "Information as Fuel" (Grant NANOQIT. Ref FQXi-IAF19-01). I appreciate the hospitality of Natalia Ares at University of Oxford and discussions with the whole NANOQIT team on the content and organization of this article. Finally, I am indebted to Jordan Horowitz, Takahiro Sagawa, and Massimiliano Esposito for their careful reading of the article and their helpful suggestions.

References

Allahverdyan, A.E., Janzing, D., Mahler, G., 2009. Thermodynamic efficiency of information and heat flow. Journal of Statistical Mechanics: Theory and Experiment 2009, P09011. doi:10.1088/1742-5468/2009/09/P09011. publisher: IOP Publishing.

Allahverdyan, A.E., Saakian, D.B., 2008. Thermodynamics of adiabatic feedback control. Europhys. Lett. 81, 30003.

Anders, J., Esposito, M., 2017. Focus on quantum thermodynamics. New Journal of Physics 19, 010201. doi:10.1088/1367-2630/19/1/010201.

Barato, A., Seifert, U., 2014. Stochastic thermodynamics with information reservoirs. Phys. Rev. E 90, 042150.

Bennett, C., 1982. The thermodynamics of computation—a review. Int. J. Theor. Phys. 21, 905–940. Reprinted in (Leff and Rex, 1990).

Berut, A., Arakelyan, A., Petrosyan, A., Ciliberto, S., Dillenschneider, R., Lutz, E., 2011. Experimental verification of landauer's principle linking information and thermodynamics. Nature 483, 187–189.

Bustamante, C., Keller, D., Oster, G., 2001. The Physics of Molecular Motors. Acc. Chem. Res. 34, 412–420.

Cafaro, C., Ali, S.A., Giffin, A., 2016. Thermodynamic aspects of information transfer in complex dynamical systems. Physical Review E 93, 022114. doi:10.1103/PhysRevE.93.022114. publisher: American Physical Society.

Cottet, N., Jezouin, S., Bretheau, L., Campagne-Ibarcq, P., Ficheux, Q., Anders, J., Auffèves, A., Azouit, R., Rouchon, P., Huard, B., 2017. Observing a quantum Maxwell demon at work. Proceedings of the National Academy of Sciences 114, 7561–7564. doi:10.1073/pnas.1704827114. publisher: Proceedings of the National Academy of Sciences.

Cover, T.M., Thomas, J.A., 2006. Elements of Information Theory. Second ed., Wiley-Interscience, New York, NY.

- Ehrich, J., Esposito, M., Barra, F., Parrondo, J.M., 2019. Micro-reversibility and thermalization with collisional baths. Physica A: Statistical Mechanics and its Applications, 122108doi:https://doi.org/10.1016/j.physa. 2019.122108.
- Esposito, M., Van den Broeck, C., 2011. Second law and landauer principle far from equilibrium. Europhys. Lett. 95, 40004.
- Esposito, M., Lindenberg, K., den Broeck, C.V., 2010. Entropy production as correlation between system and reservoir. New Journal of Physics 12, 013013. doi:10.1088/1367-2630/12/1/013013.
- Feng, Y., Ovalle, M., Seale, J.S.W., Lee, C.K., Kim, D.J., Astumian, R.D., Stoddart, J.F., 2021. Molecular Pumps and Motors. J. Am. Chem. Soc. 143, 5569–5591.
- Feynman, R.P., Leighton, R.B., Sands, M., 1966. The Feynman Lectures on Physics. Vol I. Addison-Wesley. Chap. 46.
- Hartich, D., Barato, A.C., Seifert, U., 2014. Stochastic thermodynamics of bipartite systems: transfer entropy inequalities and a maxwell's demon interpretation. J. Stat. Mech.: Theor. Exp., P02016.
- Horowitz, J.M., 2015. Multipartite information flow for multiple Maxwell demons. Journal of Statistical Mechanics: Theory and Experiment 2015, P03006. doi:10.1088/1742-5468/2015/03/P03006. publisher: IOP Publishing.
- Horowitz, J.M., Esposito, M., 2014. Thermodynamics with continuous information flow. Phys. Rev. X 4, 031015.
- Horowitz, J.M., Parrondo, J.M.R., 2011a. Designing optimal discrete-feedback thermodynamic engines. New J. Phys. 13, 123019.
- Horowitz, J.M., Parrondo, J.M.R., 2011b. Thermodynamic reversibility in feedback processes. Europhys. Lett. 95, 10005.
- Horowitz, J.M., Sagawa, T., Parrondo, J.M.R., 2013. Imitating chemical motors with optimal information motors. Physical Review Letters 111, 010602.
- Horowitz, J.M., Sandberg, H., 2014. Second-law-like inequalities with information and their interpretations. New Journal of Physics 16, 125007. doi:10.1088/1367-2630/16/12/125007. publisher: IOP Publishing.
- Horowitz, J.M., Vaikuntanathan, S., 2010. Nonequilibrium detailed fluctuation theorem for discrete feedback. Phys. Rev. E 82, 061120.
- Ito, S., Sagawa, T., 2013. Information thermodynamics on causal networks. Physical Review Letters 111, 180603.
- James, R.G., Barnett, N., Crutchfield, J.P., 2016. Information Flows? A Critique of Transfer Entropies. Physical Review Letters 116, 238701. doi:10.1103/PhysRevLett.116.238701. publisher: American Physical Society.
- Kim, S.W., Sagawa, T., De Liberato, S., Ueda, M., 2011. Quantum Szilard engine. Physical Review Letters 106, 070401.
- Kiwata, H., 2022. Relationship between Schreiber's transfer entropy and Liang-Kleeman information flow from the perspective of stochastic thermodynamics. Physical Review E 105, 044130. doi:10.1103/PhysRevE. 105.044130. publisher: American Physical Society.
- Koski, J.V., Maisi, V.F., Pekola, J.P., Averin, D.V., 2014. Experimental realization of a szilard engine with a single electron. Proceedings of the National Academy of Sciences 111, 13786–13789. doi:10.1073/pnas. 1406966111.
- Landauer, R., 1961. Irreversibility and heat generation in the computing process. IBM Journal of Research and Development 5, 183–191. doi:10. 1147/rd.53.0183.
- Leff, H.S., Rex, A.F. (Eds.), 1990. Maxwell's Demon: Entropy, Information, Computing. Princeton University Press, New Jersey.
- Liang, X.S., Kleeman, R., 2005. Information Transfer between Dynamical System Components. Physical Review Letters 95, 244101. doi:10.1103/ PhysRevLett.95.244101. publisher: American Physical Society.
- Lloyd, S., 1989. Use of mutual information to decrease entropy: Implications for the second law of thermodynamics. Phys. Rev. A 39, 5378– 5386.
- Lloyd, S., 1997. Quantum-mechanical Maxwell's demon. Phys. Rev. A 56, 3374–3382.
- Mandal, D., Jarzynski, C., 2012. Work and information processing in a solvable model of maxwell's demon. Proc. Nat. Ac. Sci. .
- Ptaszyński, K., Esposito, M., 2019. Thermodynamics of quantum information flows. Phys. Rev. Lett. 122, 150603. doi:10.1103/PhysRevLett.122.

- 150603.
- Park, J.J., Kim, K.H., Sagawa, T., Kim, S.W., 2013. Heat engine driven by purely quantum information. Physical Review Letters 111, 230402.
- Parrondo, J.M.R., 2001. The Szilard engine revisited: Entropy, macroscopic randomness, and symmetry breaking phase transitions. Chaos 11, 725– 733.
- Parrondo, J.M.R., Español, P., 1996. Criticism of Feynman's analysis of the ratchet as an engine. American Journal of Physics 64, 1125–1130. doi:10.1119/1.18393.
- Parrondo, J.M.R., Horowitz, J.M., Sagawa, T., 2015. Thermodynamics of information. Nature Physics 11, 131–139. doi:10.1038/nphys3230.
- Pekola, J., Khaymovich, I., 2019. Thermodynamics in Single-Electron Circuits and Superconducting Qubits. Annual Review of Condensed Matter Physics 10, 193–212. doi:10.1146/annurev-conmatphys-033117-054120. _eprint: https://doi.org/10.1146/annurev-conmatphys-033117-054120.
- Peliti, L., Pigolotti, S., 2021. Stochastic Thermodynamics. Princeton University Press.
- Ponmurugan, M., 2010. Generalized detailed fluctuation theorem under nonequilibrium feedback control. Phys. Rev. E 82, 031129. doi:10.1103/ PhysRevE.82.031129.
- Price, R., 1982. Oral-History:Claude E. Shannon. Engineering and Technology History Wiki (ETHW) URL: https://ethw.org/Oral-History:Claude_E._Shannon.
- Proesmans, K., Ehrich, J., Bechhoefer, J., 2020. Finite-time landauer principle. Phys. Rev. Lett. 125, 100602. doi:10.1103/PhysRevLett.125. 100602.
- Ptaszyński, K., Esposito, M., 2019. Entropy Production in Open Systems:
 The Predominant Role of Intraenvironment Correlations. Physical Review Letters 123, 200603.
- Reimann, P., 2002. Brownian motors: noisy transport far from equilibrium. Phys. Rep. 361, 57–265.
- Roldán, E., Martínez, I.A., Parrondo, J.M.R., Petrov, D., 2014. Universal features in the energetics of symmetry breaking. Nature Physics.
- Rosinberg, M.L., Horowitz, J.M., 2016. Continuous information flow fluctuations. EPL (Europhysics Letters) 116, 10007. doi:10.1209/0295-5075/116/10007. publisher: IOP Publishing.
- Sagawa, T., 2012a. Thermodynamics of Information Processing in Small Systems*). Progress of Theoretical Physics 127, 1–56. doi:10.1143/PTP. 127.1.
- Sagawa, T., 2012b. Thermodynamics of Information Processing in Small Systems. Springer Theses, Springer.
- Sagawa, T., 2014. Thermodynamic and logical reversibilities revisited. Journal of Statistical Mechanics: Theory and Experiment 2014, P03025. doi:10.1088/1742-5468/2014/03/p03025.
- Sagawa, T., Ueda, M., 2008. Second law of thermodynamics with discrete quantum feedback control. Physical Review Letters 100, 080403.
- Sagawa, T., Ueda, M., 2009. Minimal energy cost for thermodynamic information processing: Measurement and information erasure. Physical Review Letters 102, 250602.
- Sagawa, T., Ueda, M., 2010. Generalized jarzynski equality under nonequilibrium feedback control. Physical Review Letters 104, 090602.
- Sagawa, T., Ueda, M., 2012. Fluctuation theorem with information exchange: Role of correlations in stochastic thermodynamics. Physical Review Letters 109, 180602.
- Saha, T.K., Lucero, J.N.E., Ehrich, J., Sivak, D.A., Bechhoefer, J., 2021.
 Maximizing power and velocity of an information engine. Proceedings of the National Academy of Sciences 118, e2023356118. doi:10.1073/pnas.2023356118.
- Schmitt, R.K., Parrondo, J.M.R., Linke, H., Johansson, J., 2015. Molecular motor efficiency is maximized in the presence of both power-stroke and rectification through feedback. New Journal of Physics 17. doi:10.1088/ 1367-2630/17/6/065011.
- Schreiber, T., 2000. Measuring Information Transfer. Physical Review Letters 85, 461–464. doi:10.1103/PhysRevLett.85.461. publisher: American Physical Society.
- Schumacher, B., Westmoreland, M., 2010. Quantum Processes Systems, and Information. Cambridge University Press, USA.

- Sekimoto, K., 2010. Stochastic Energetics. volume 799 of Lect. Notes Phys. Springer, Berlin Heidelberg.
- Shannon, C.E., 1948. A mathematical theory of communication. Bell System Technical Journal 27, 379–423. doi:https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.
- Smoluchowski, M.v., 1912. Experimentell nachweisbare, der Üblichen Thermodynamik widersprechende Molekularphenomene. Phys. Zeitshur. 13, 1069.
- Szilárd, L., 1929. über die entropieverminderung in einem thermodynamischen system bei eingriffen intelligenter wesen. Zeitschrift für Physik 53, 840–856. doi:10.1007/BF01341281. On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. Translated into English and reprinted in (Leff and Rex, 1990).
- Touchette, H., Lloyd, S., 2000. Information-theoretic limits of control. Physical Review Letters 84, 1156–1159.
- Toyabe, S., Sagawa, T., Ueda, M., Muneyuki, E., Sano, M., 2010. Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality. Nature Physics 6, 988–992.
- Wheeler, J.A., 1995. Toward "it from bit", in: Anandan, J.S., Safko, J.L. (Eds.), Quantum Coherence and Reality, World Scientific. p. 281. doi:10.1142/9789814533294.
- Wolpert, D.H., 2019. The stochastic thermodynamics of computation. Journal of Physics A: Mathematical and Theoretical 52, 193001. doi:10. 1088/1751-8121/ab0850. publisher: IOP Publishing.
- Zurek, W., 1990. Maxwell's demon, Szilard's engine and quantum measurements, in: Leff, H.S., Rex, A.F. (Eds.), Maxwell's Demon: Entropy, Information, Computing. Princeton University Press, New Jersey.
- Zurek, W., 2003. Quantum discord and Maxwell's demons. Phys. Rev. A 67, 012320.