IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

# Differential Landauer's principle

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2013 EPL 101 50004

(http://iopscience.iop.org/0295-5075/101/5/50004)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 129.68.65.223
The article was downloaded on 08/09/2013 at 06:01

Please note that terms and conditions apply.

A LETTERS JOURNAL EXPLORING THE FRONTIERS OF PHYSICS

March 2013

EPL, **101** (2013) 50004
doi: 10.1209/0295-5075/101/50004

www.epljournal.org

# Differential Landauer's principle

Léo Granger and Holger Kantz

*Max-Planck-Institut für Physik komplexer Systeme - Nöthnitzerstr. 38, D-01187, Dresden, Germany, EU*

received 15 November 2012; accepted in final form 22 February 2013
published online 19 March 2013

PACS 05.70.Ln – Nonequilibrium and irreversible thermodynamics
PACS 05.40.-a – Fluctuation phenomena, random processes, noise, and Brownian motion
PACS 05.20.-y – Classical statistical mechanics

**Abstract** – Landauer's principle states that the erasure of information must be a dissipative process. In this paper, we carefully analyze the recording and erasure of information on a physical memory. On the one hand, we show that, in order to record some information, the memory has to be driven out of equilibrium. On the other hand, we derive a differential version of Landauer's principle: We link the rate at which entropy is produced at every time of the erasure process to the rate at which information is erased.

Copyright © EPLA, 2013

**Introduction. –** The fundamental role of information in thermodynamics is now well established. On the one hand, it is possible to convert information into useful work and use it as a "fuel" for some heat engine [1,2]. On the other hand, information itself must be recorded on some physical memory. Information processing is implemented via physical transformations operating on the memory and is thereby accompanied by thermodynamic costs [3–12]. The paradigmatic result concerning the costs of information processing is Landauer's erasure principle [3]. It is a statement about the thermodynamics of information erasure. In its original formulation, it states that the resetting to zero of a random bit necessarily leads to the dissipation of $k_{\mathrm{B}}T \log 2$ of energy, where $k_{\mathrm{B}}$ is Boltzmann's constant and $T$ is the temperature of the environment. Landauer's principle was celebrated as the resolution Maxwell's demon's paradox [13]. However, it is still a very controversial result, see [14] and references therein. In this paper, we carefully analyze the processes of recording and erasing information on a physical memory in a simple but generic scenario and we precisely derive the thermodynamic costs of these processes. While the problem usually addressed in the literature is the operation "reset to zero" performed on isolated one-bit memories [5,6,15–19], here we propose to analyze the recording and erasure of some information coming from an external source. We use the framework of stochastic thermodynamics [20–24] to model the memory and to derive the thermodynamics of these processes.

We consider the following scenario: A source randomly emits a symbol $\alpha_k$ out of $N$ possible symbols, $\alpha_1, \ldots, \alpha_N$. We wish to record which symbol appeared on some physical memory and then we wish to erase it. The memory should be able to be in at least $N$ states: One for each of the possible outcomes. However, it is convenient to have one more state serving as the "standard" state: It is the state of the memory when it is empty, *i.e.*, when nothing is recorded. Recording $\alpha_k$ simply means to drive the memory from the standard state to state number $k$. Erasing the content of the memory means to drive it back to the standard state by a process that is independent of the symbol stored.

In the following, we show that recording the information amounts to correlate the memory to the symbol emitted by the source. During the erasure process, these correlations gradually decrease to zero. We show that the rate of entropy produced is bounded from below by the rate at which the correlations decrease all along the erasure process. This is our second main result, eq. (10), which can be viewed as a differential version of Landauer's principle. Our first main result, eq. (5) is a statement about a different kind of costs for the correlations: Although the recording can in principle be performed reversibly, the memory can only store some information if it is out of equilibrium.

**Recording information on a physical system. –** In order to simplify the discussion, we use an overdamped particle subject to a random force of thermal origin as a memory. However, any system obeying the laws of stochastic thermodynamics would do. The particle is also subjected to a force $F(x, \lambda) = -\partial_x V(x, \lambda)$ derived from a conservative potential $V(x, \lambda)$. The potential is time dependent through the external control parameter

$\lambda$ that can be varied in order to control the state of the memory. The position of the particle evolves according to the Langevin equation:

$$\dot{x} = \mu F(x, \lambda) + \xi(t), \tag{1}$$

where $\mu$ is the mobility of the particle and $\xi(t)$ is a Gaussian white noise with $\langle \xi(t) \rangle = 0$ and $\langle \xi(t)\xi(t') \rangle = 2D\delta(t - t')$, $D$ being the diffusion constant. When the medium is in equilibrium at temperature $T$, the diffusion constant and the mobility are related by the Einstein relation $D = k_\mathrm{B} T\mu$. The probability density $\rho(x,t)$ to find the particle at position $x$ at time $t$ evolves according to the Fokker-Planck equation:

$$\partial_t \rho(x,t) = -\partial_x \left(\mu F(x,\lambda)\rho(x,t) - D\partial_x \rho(x,t)\right), \tag{2}$$

Because of thermal fluctuations, the position of the particle is random and we cannot control it. However, through a suitable choice of the time dependence of the control parameter $\lambda(t)$ (the protocol), we can control the distribution $\rho(x,t)$ describing the position of the particle. Hence, we will use such distributions to encode the different symbols $\alpha_k$.

Let $\{\phi_k\}$, $0 \leqslant k \leqslant N$ be $N+1$ distributions over the position of the particle: $\phi_0(x)$ is the standard state and for $1 \leqslant k \leqslant N$, $\phi_k(x)$ encodes the symbol $\alpha_k$. For the information to be unambiguously recorded, we need the states encoding different symbols to be perfectly distinguishable. This means that the corresponding distributions should not overlap: For each position $x$ there should be only one $k$ such that $\phi_k(x)$ is non-zero. In fact, the information is stored in the position of the particle. If, for a given position $x$, $\phi_k(x) > 0$ and $\phi_{k'}(x) > 0$, then observing the particle at $x$ we cannot decide whether $\alpha_k$ or $\alpha_{k'}$ is stored.

The recording process is the following. Assume that the symbol emitted by the source is $\alpha_k$. Initially, the memory is in the standard state $\rho(x,t_0) = \phi_0(x)$ and the control parameter takes the value $\lambda(t_0) = \lambda_0$. From time $t = t_0 < 0$ to $t = 0$, the control parameter is changed from its initial value $\lambda_0$ to some final value $\lambda_\mathrm{rec}$ according to some protocol $\lambda_k(t)$. The protocol depends on the symbol $\alpha_k$ that appeared and is such that the final state of the memory is $\rho(x,0) = \phi_k(x)$. However, once the information is recorded, we want to be able to manipulate the memory without knowing the information stored. Hence, the final value of the control parameter, $\lambda_\mathrm{rec}$ should be the same for all $k$ as illustrated in fig. 1. As a consequence, at the end of the recording process, the memory is out of equilibrium. In fact we want to allow it to be in one out of $N$ different states for a single value of $\lambda$. But for each value of $\lambda$, there is only one equilibrium state given by the Boltzmann distribution:

$$\rho_\mathrm{eq}(x,\lambda) = \exp\left(-\frac{V(x,\lambda) - F_\mathrm{eq}(\lambda)}{k_\mathrm{B} T}\right), \tag{3}$$

where $\exp(-F_\mathrm{eq}(\lambda)/k_\mathrm{B}T) = \int \exp(-V(x,\lambda)/k_\mathrm{B}T)\mathrm{d}x$ is the partition function and $F_\mathrm{eq}(\lambda)$ the equilibrium free energy as a function $\lambda$.

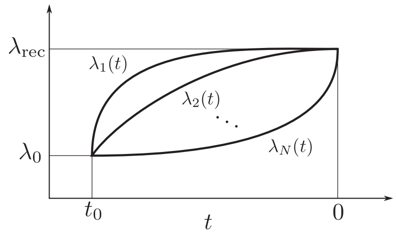


Fig. 1: Recording protocol: From time $t = t_0 < 0$ to $t = 0$, the control parameter $\lambda$ is driven from $\lambda(t_0) = \lambda_0$ to $\lambda(0) = \lambda_\mathrm{rec}$ in a way that depends on $k$, such that the final state of the memory is $\phi_k$, the state encoding $\alpha_k$, the symbol that was emitted by the source.

In stochastic thermodynamics, the distance of a non-equilibrium state $\rho$ to equilibrium $\rho_\mathrm{eq}$ is quantified by the Kullback-Leibler divergence, or relative entropy between $\rho$ and $\rho_\mathrm{eq}$ [22,25,26]:

$$D[\rho \| \rho_\mathrm{eq}] = \int \rho(x) \log \frac{\rho(x)}{\rho_\mathrm{eq}(x)} \mathrm{d}x. \tag{4}$$

In fact, this quantity is linked to the free energy of the non-equilibrium state $\rho$: $F[\rho] - F_\mathrm{eq} = k_\mathrm{B} T D[\rho \,\|\, \rho_\mathrm{eq}] \geqslant 0$ [20,22,26]. The relative entropy between two distributions is non-negative and it vanishes if and only if the two distributions are identical [27].

The first result of this paper is the following inequality:

$$\sum_k P_k D[\phi_k \| \rho_\mathrm{eq}] \geqslant H, \tag{5}$$

where $\rho_\mathrm{eq}$ is the final equilibrium distribution, $P_k$ is the probability that the symbol $\alpha_k$ was emitted and $H = -\sum_k P_k \log P_k$ is the Shannon entropy of the source. The latter quantifies our *a priori* uncertainty about the symbol emitted by the source or the average information provided by the observation of the symbol emitted [27]. Inequality (5) above means, that the expected distance to equilibrium at the end of the recording process is greater than the average information provided by the emission of the symbol to be recorded. One can interpret $H$ as the amount of information stored in the memory. Inequality (5) can then be interpreted in the following way: In order to record some information, the system serving as a memory has to be driven out of equilibrium and the distance to equilibrium has to be greater than the information stored. Nevertheless, the recording process can in principle be performed reversibly in that there is no finite lower bound to the entropy that has to be produced.

Inequality (5) is very general and does not depend on the structure of the memory. The only assumptions are: i) The macroscopic states of the memory are given by distributions over the micro-states, ii) states of the memory encoding different symbols should be perfectly distinguishable, iii) there is a unique equilibrium distribution. Sometimes, due to technical restrictions, it might not be possible to prepare the memory in non-overlapping states and hence it would not be possible to fulfill condition ii). In this case the information cannot be completely unambiguously recorded, but nevertheless, as we will see later, one can still quantify the maximum amount of information, $I_{\max} < H$, that can be stored and inequality (5) will still be valid replacing $H$ by $I_{\max}$ in the right-hand side. We will address this point later, after having discussed the erasure process.

**Erasure: differential Landauer's principle.** – Erasure is a process bringing the memory back to its standard state without making use of the information stored. This is expressed by the fact that the erasure protocol does not depend on $k$. The initial value of the control parameter is $\lambda_{\mathrm{rec}}$ and the initial state of the memory is the state encoding the symbol that was emitted, *i.e.*, it is $\phi_k(x)$ with probability $P_k$. From time $t=0$ to time $t=t_1$ the control parameter is driven from $\lambda_{\mathrm{rec}}$ back to $\lambda_0$ in such a way that the final state is the standard state $\phi_0(x)$.

Assume that we stop the process at some intermediate time $t$. We would like to address the following issues: What is the information erased until that time? Or alternatively, what is the information still contained in the memory? And what is the minimum amount of entropy produced until that time?

In order to illustrate the information loss, consider the following situation: At time $t$, we see the particle at position $x$ and from this knowledge, we would like to infer which symbol was stored. The probability $P(k|x;t)$, that the symbol $\alpha_k$ was stored is given by the Bayes rule:

$$P(k|x;t) = \frac{\rho_k(x,t)P_k}{\rho_{\mathrm{m}}(x,t)}, \qquad (6)$$

where $\rho_k(x,t)$ is the distribution of the particle's position at time $t$ if $\alpha_k$ was stored and $\rho_{\mathrm{m}}(x,t) = \sum_k P_k\rho_k(x,t)$ is the marginal distribution of the position of the particle at time $t$. Its presence in the expression above ensures that the probability distribution $P(k|x;t)$ is normalized. Concretely, $\rho_k(x,t)$ is obtained by propagating $\phi_k(x)$ with the Fokker-Planck equation with the erasure protocol $\lambda(t)$. Initially, at $t=0$, $P(k|x;0)=0$ or 1 depending on whether $x$ belongs to the support of $\phi_k$ of not: The position of the particle contains the complete information about the symbol that was emitted. At intermediate time $t$, the $\rho_k(x,t)$ might overlap and we will have some uncertainty about the symbol that was stored upon seeing the particle at position $x$. This uncertainty is quantified by the Shannon entropy of the probability distribution $P(k|x;t)$:

$$h_{\mathrm{er}}(x,t) = -\sum_k P(k|x;t)\log P(k|x;t). \qquad (7)$$

The total information erased until time $t$ is the average uncertainty about the symbol that was stored upon knowing the position of the particle at time $t$:

$$H_{\mathrm{er}}(t) = \int \rho_{\mathrm{m}}(x,t)h_{\mathrm{er}}(x,t)\mathrm{d}x. \qquad (8)$$

At the beginning of the erasure process, no information is yet erased and $H_{\mathrm{er}}(0)=0$. At the end of the process, knowing the position of the particle does not reduce our uncertainty about the symbol that had been emitted and $H_{\mathrm{er}}(t_1)=H$. The information $I(t)$ still contained in the memory is the reduction in uncertainty about the symbol that was stored upon knowing the position of the particle at time $t$:

$$I(t) = H - H_{\mathrm{er}}(t). \qquad (9)$$

This is just the *mutual information* between the position of the particle and and the symbol originally recorded. It is a measure of how much information the position of the particle at time $t$ can still provide about the symbol originally stored [27]. During the erasure process, it decreases from $H$ to 0 as the information is gradually erased.

The second main result of this paper is the following: The rate of entropy production is bounded from below by the rate of information erasure:

$$\dot{S}^{\mathrm{irr}} \geqslant k_{\mathrm{B}}\dot{H}_{\mathrm{er}} = -k_{\mathrm{B}}\dot{I}. \qquad (10)$$

We call this result the differential Landauer's principle. In fact, it is a precise and general statement of the thermodynamic costs of information erasure. Integrating the equation above between times $t$ and $t'$, with $0\leqslant t\leqslant t'\leqslant t_1$, we get a lower bound for the irreversible entropy production between time $t$ and $t'$:

$$\Delta S^{\mathrm{irr}}(t,t') \geqslant k_{\mathrm{B}}\left(I(t)-I(t')\right). \qquad (11)$$

Hence, the minimum amount of entropy produced is directly linked to the loss of correlation between the position of the particle and the symbol initially stored. Setting $t=0$ and $t'=t_1$ in eq. (11) above gives a general integral version of Landauer's principle:

$$\Delta S^{\mathrm{irr}} \geqslant k_{\mathrm{B}}H, \qquad (12)$$

linking the total amount $\Delta S^{\mathrm{irr}}$ of entropy produced to the amount $H$ of information erased.

In general, we might not be able to prepare the memory in non-overlapping states $\{\phi_k\}$. In this case, the information cannot be fully reliably stored. However, we can still quantify the maximum amount of information that we can store. It is given by eq. (9) for $t=0$: $I_{\max}=I(0)=H-H_{\mathrm{er}}(0)<H$, where $H_{\mathrm{er}}(0)>0$ due to the overlapping. In this case, inequality (5) still holds replacing $H$ by $I_{\max}$ in the right-hand side. Moreover, the lower bound to the total entropy produced during the erasure process, obtained by integrating inequality (10), is reduced: $\Delta S^{\mathrm{irr}}\geqslant k_{\mathrm{B}}I_{\max}$, implying that we do not have to pay for the information we were not able to record.

**Proof of eqs. (5) and (10).** – We now prove our two main results, inequalities (5) and (10). The information contained in the memory at time $t$ of the erasure process can be expressed as

$$I(t)=\sum_k P_k D[\rho_k(t)\|\rho_{\mathrm{m}}(t)]. \tag{13}$$

Using this expression, it is easy to show that the average distance to equilibrium satisfies

$$\sum_k P_k D[\rho_k(t)\|\rho_{\mathrm{eq}}(\lambda(t))]=I(t)+D[\rho_{\mathrm{m}}(t)\|\rho_{\mathrm{eq}}(\lambda(t))]. \tag{14}$$

At time $t=0$, this expression implies inequality (5), since $I(0)=H$, and the second term in the right-hand side is non-negative at all times. Actually, eq. (14) is a generalization of inequality (5). It relates the average distance to equilibrium to the information still contained in the memory at any time of the erasure process.

At time $t$ of the erasure process, if $\alpha_k$ was stored, the state of the memory is $\rho_k(x,t)$ introduced earlier. The thermodynamic entropy of the memory is given by the Shannon-Gibbs formula:

$$S_k(t)=-k_{\mathrm{B}}\int \rho_k(x,t)\log\rho_k(x,t)\mathrm{d}x. \tag{15}$$

Its variation satisfies the Clausius relation [22,28]:

$$\dot S_k(t)=\frac{\dot Q_k(t)}{T}+\dot S_k^{\mathrm{irr}}(t), \tag{16}$$

where $\dot Q_k(t)=\int V(x,\lambda(t))\partial_t\rho_k(x,t)\mathrm{d}x$ is the heat flux to the particle at time $t$ and $\dot S_k^{\mathrm{irr}}(t)\geqslant 0$ is the rate at which entropy is irreversibly produced.

Using eq. (13), we get the following relation between the information $I(t)$ and the expected entropy $S(t)=\sum_k P_k S_k(t)$ of the memory:

$$S(t)=S_{\mathrm{m}}(t)-k_{\mathrm{B}}I(t), \tag{17}$$

where $S_{\mathrm{m}}(t)=-k_{\mathrm{B}}\int\rho_{\mathrm{m}}(x,t)\log\rho_{\mathrm{m}}(x,t)\mathrm{d}x$ is the entropy of the marginal distribution $\rho_{\mathrm{m}}(x,t)$. Since the Fokker-Planck equation (2) is linear and since the marginal distribution is a linear combination of solutions of this equation, it satisfies this equation as well. As a consequence, the variations of $S_{\mathrm{m}}(t)$ also satisfy some Clausius relation similar to eq. (16):

$$\dot S_{\mathrm{m}}(t)=\frac{\dot Q_{\mathrm{m}}(t)}{T}+S_{\mathrm{m}}^{\mathrm{irr}}(t), \tag{18}$$

where $\dot Q_{\mathrm{m}}(t)=\int V(x,\lambda(t))\partial_t\rho_{\mathrm{m}}(x,t)\mathrm{d}x$ and $S_{\mathrm{m}}^{\mathrm{irr}}(t)\geqslant 0$.

Noting that $\dot Q_{\mathrm{m}}(t)=\sum_k P_k\dot Q_k(t)$ and inserting eqs. (16) and (18) into eq. (17) yields for the average entropy production rate $\dot S^{\mathrm{irr}}(t)=\sum_k P_k\dot S_k^{\mathrm{irr}}(t)$:

$$\dot S^{\mathrm{irr}}(t)=\dot S_{\mathrm{m}}^{\mathrm{irr}}(t)-k_{\mathrm{B}}\dot I(t). \tag{19}$$

This proves eq. (10) since $\dot S_{\mathrm{m}}^{\mathrm{irr}}(t)\geqslant 0$.

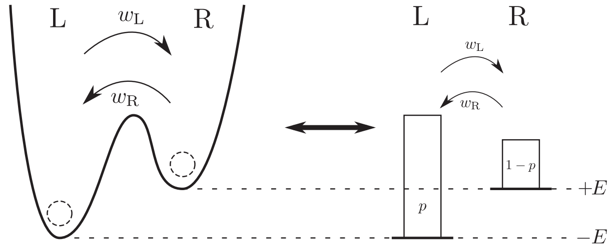


Fig. 2: A very simple model for a memory: A system that can make thermally activated transition between two states "L" and "R". This system can be controlled through the energy difference between the two states. Its macroscopic state is given by the probability $p$ to occupy the state "L". This system could represent a Brownian particle in a double-well potential, that can jump from one well to the other because of thermal fluctuations.

We used an over-damped Brownian particle as a memory in order to present our results clearly. However, one could use other types of systems obeying the laws of stochastic thermodynamics as a memory. In particular, systems with discrete phase space evolving according to master equations such as the systems considered in [23,29] yield the same results, where it is essential that they fulfill the Clausius relation, eq. (16). In fact in the following we will present the recording and erasure of the outcome of a binary random variable on a memory with two discrete micro-states.

**Example: a two states system as a memory.** – As an example, we consider the recording and erasure of the result of a binary random variable. Let $Y$ be a random variable with two possible outcomes. In analogy with the tossing of a biased coin, let "head" and "tail" be the two possible outcomes appearing, respectively, with probability $P$ and $1-P$. We want to record the outcome of $Y$ and then erase it. The amount of information we want to store is given by the Shannon entropy of $Y$, $H(P)=-P\log P-(1-P)\log(1-P)$.

The system serving as a memory is a two-states system in contact with an equilibrium heat bath at temperature $T$. This system could represent an over-damped Brownian particle in a double-well potential as depicted in fig. 2. In the following, we will again speak of a "particle" that can "jump" between two "wells" keeping in mind that the memory could be any two-states system in contact with a heat bath that can make transitions between the two states. Let "L" and "R" label the two wells and let $E_{\mathrm{L}}=-E$ and $E_{\mathrm{R}}=+E$ be their respective energies. We assume that we are able to control the energy difference $\Delta E=E_{\mathrm{R}}-E_{\mathrm{L}}=2E$ between the two wells by applying a force field. Hence, our control parameter is $E$.

At any time the particle might jump from one well to the other due to thermal fluctuations of the heat bath. The probability per unit time, that the particle jumps from the right to the left well (from the left to the right well) is given by Kramer's rate $w_\mathrm{R} = \tau^{-1}\exp(E/k_\mathrm{B}T)$ ($w_\mathrm{L} = \tau^{-1}\exp(-E/k_\mathrm{B}T)$), where $\tau$ is a time linked to the height of the potential barrier between the two wells. The macroscopic state of the memory is fully described by the probability $p$ that the particle is in the left well. The latter evolves in time according to the following master equation:

$$\dot{p} = -w_\mathrm{L} p + w_\mathrm{R}(1-p)\,. \tag{20}$$

The rates satisfy detailed balance $w_\mathrm{R}/w_\mathrm{L} = \exp(-\Delta E/k_\mathrm{B}T)$ and the equilibrium is given by the Boltzmann factor

$$p_\mathrm{eq}(E) = \exp\left(-\frac{E_\mathrm{L}-F}{k_\mathrm{B}T}\right), \tag{21}$$

where the equilibrium free energy $F(E)$ is linked to the partition function $Z(E) = \exp(-F(E)/k_\mathrm{B}T) = 2\cosh(E/k_\mathrm{B}T)$. The thermodynamic entropy of the memory is given by $S(p) = -k_\mathrm{B}(p\log p + (1-p)\log(1-p))$ and the heat received per unit time is given by $\dot{Q} = \dot{p}\Delta E = 2\dot{p}E$. The instantaneous entropy production rate is given by the Clausius relation $\dot{S}^\mathrm{irr} = \dot{S} - \dot{Q}/T \geqslant 0$.

Now, we need to decide which states we use to encode the different possible outcomes of $Y$. The idea is to say that the left well encodes one of the outcomes, say "head" and the right well the other. In other words, we want the state $p_\mathrm{h} = 1$ to encode "head" and the state $p_\mathrm{t} = 0$ to encode "tail". However, this supposes that we are able to create an infinite energy difference between the two wells. In the following, we assume that there is a maximum value $E_\mathrm{max} \geqslant 0$, such that $|E| \leqslant E_\mathrm{max}$. Then the best we can do is drive the memory to $p_\mathrm{max} = p_\mathrm{eq}(E_\mathrm{max})$ or $1 - p_\mathrm{max} = p_\mathrm{eq}(-E_\mathrm{max})$. Hence, we will use $p_\mathrm{h} = p_\mathrm{max}$ to encode "head" and $p_\mathrm{t} = 1 - p_\mathrm{max}$ to encode "tail". Introducing $p_\mathrm{m} = Pp_\mathrm{h} + (1-P)p_\mathrm{t}$, the marginal probability for the particle to occupy the left well at the end of the recording process, we can compute the maximum amount of information that can be stored in this memory using eq. (13):

$$I_\mathrm{max} = PD(p_\mathrm{h}\|p_\mathrm{m}) + (1-P)D(p_\mathrm{t}\|p_\mathrm{m}), \tag{22}$$

where $D(p\|q) = p\log\frac{p}{q} + (1-p)\log\frac{1-p}{1-q}$ is the Kullback-Leibler divergence between two distributions over a binary random variable. Figure 3(a) shows $I_\mathrm{max}/H$ as a function of $E_\mathrm{max}/k_\mathrm{B}T$. For $E_\mathrm{max}$ of the order of $k_\mathrm{B}T$, the maximum information content is clearly smaller than $H$. For $E_\mathrm{max} \gg k_\mathrm{B}T$, $I_\mathrm{max} \simeq H$.

Now that we have the states encoding the different possible outcomes, we have to specify the standard state. Traditionally in the literature, the standard state is: "the particle is in the left (or right) well" [3,5,19]. However, we can use any distribution over the two wells as the standard state. A simple choice is the equidistribution: $p_0 = 1/2$.

The recording process is schematically sketched in fig. 4. If "head" appears, the parameter $E$ is driven from 0 to $E_\mathrm{max}$ and the memory is let to relax towards equilibrium $p_\mathrm{h} = p_\mathrm{max}$ and if "tail" appears, we drive $E$ to $-E_\mathrm{max}$ and let the memory relax towards $p_\mathrm{t} = 1 - p_\mathrm{max}$. At the end of the recording process, we instantaneously drive $E$ back to zero, so that it has the same value in the two cases. As can be seen in fig. 3(b), at the end of the recording process, the Kullback-Leibler divergence to the equilibrium state is greater than the maximum information stored.

The erasure process is very simple: We keep $E = 0$ and simply let the memory relax towards equilibrium (see fig. 5). If "head" was recorded, the initial state of the memory is $p_\mathrm{h}(0) = p_\mathrm{max}$ and if "tail" was recorded, the

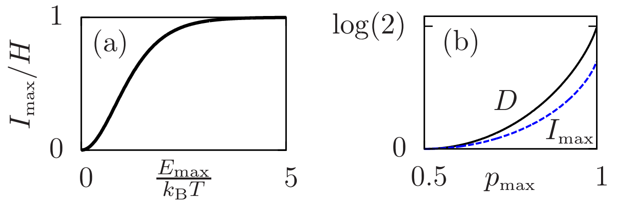


Fig. 3: (Colour on-line) (a) $I_\mathrm{max}/H$ as a function of $E_\mathrm{max}/k_\mathrm{B}T$ for $P = 0.8$. The result is very similar for other values of $P$. (b) Kullback-Leibler divergence $D = PD(p_\mathrm{h} \,\|\, p_\mathrm{eq}) + (1-P)D(p_\mathrm{t} \,\|\, p_\mathrm{eq})$ to equilibrium at the end of the recording process and maximum information $I_\mathrm{max}$ stored in the memory as a function of $p_\mathrm{max}$.

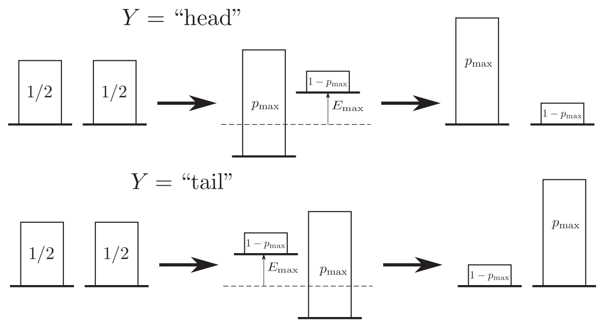


Fig. 4: Schematic picture of the recording process in the two possible scenarios. The control parameter $E$ is driven to $E_\mathrm{max}$ or $-E_\mathrm{max}$ depending on the outcome of $Y$. Once equilibrium is reached, $E$ is quickly driven back to 0.

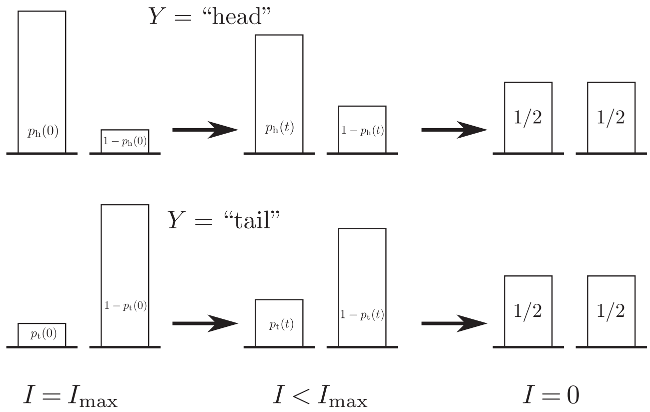


Fig. 5: Schematic picture of the erasure process. The memory, initially out of equilibrium is simply let to equilibrate. The information is erased as the overlap between $p_\mathrm{h}(t)$ and $p_\mathrm{t}(t)$ increases.

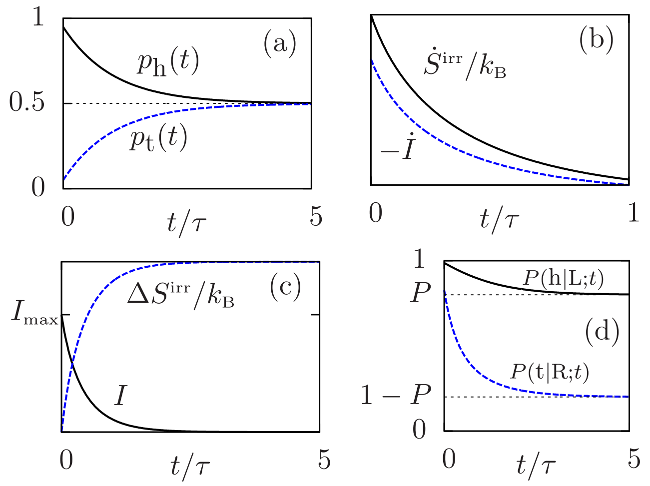


Fig. 6: (Colour on-line) Time evolution of various quantities during the erasure process for $p_{max} = 0.95$ and $P = 0.8$: (a) Evolution of the state of the memory during the erasure process. (b) Inequality (10). (c) Information content and total entropy produced during the erasure process. (d) Conditional probability that "head" ("tail") was recorded observing the particle in the left (right) well at time $t$ of the erasure process.

initial state of the memory is $p_t(0) = 1 - p_{max}$. The initial information contained in the memory is $I(0) = I_{max}$. As time goes on, $p_h(t)$ and $p_t(t)$ converge towards equilibrium, which is also the standard state $p_{eq} = p_0 = 1/2$. During this process, the information decreases and reaches 0 in the limit $t \gg \tau$, where $p_h(t) = p_t(t) = p_0$, see fig. 6. The rate of entropy production $\dot{S}^{irr}/k_B$ (in units of $k_B$) is indeed greater than the rate of information erasure $-\dot{I}$ at all times. Figure 6(d) shows the time evolution of the conditional probabilities $P(h|L;t) = P p_h(t)/p_m(t)$ that "head" was stored observing the particle in the left well and $P(t|R;t) = (1-P)(1-p_t(t))/(1-p_m(t))$ that "tail" was stored observing the particle in the right well at time $t$. At the beginning of the erasure process, they are close to 1. During the erasure process, they decrease, respectively, towards $P$ and $1-P$, the *a priori* probabilities that "head" and "tail" were stored, respectively. As $P(h|L;t)$ and $P(t|R;t)$ decrease, it becomes more and more difficult to know whether "head" or "tail" was stored upon knowing in which well the particle is: The position of the particle gradually loses the information about the symbol that was initially stored.

**Conclusion.** – We have presented the thermodynamics of recording and erasing information on a physical memory within the framework of stochastic thermodynamics. Recording some information means to correlate the memory to the external source and the information contained in the memory is quantified by the mutual information between the symbol recorded and the micro-state of the memory. On the one hand, in order to store some information, the memory must be out of equilibrium and its average distance to equilibrium (measured in terms of Kullback-Leibler divergence) must be greater than the information stored. On the other hand, when the information is erased, entropy is produced at a rate greater than the information erasure rate (up to a factor $k_B$). This result is a differential generalization of Landauer's principle, precisely stating the thermodynamic costs of erasing information without making reference to manipulations on one-bit memories.

REFERENCES


[1] Szilard L., *Z. Phys.*, **53** (1929) 840.
[2] Sagawa T. and Ueda M., *Phys. Rev. E*, **85** (2012) 021104.
[3] Landauer R., *IBM J. Res. Dev.*, **5** (1961) 183.
[4] Leff H. S. and Rex A. F., *Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing* (Institute of Physics, Bristol) 2002.
[5] Bérut A., Arakelyan A., Petrosyan A., Ciliberto S., Dillenschneider R. and Lutz E., *Nature*, **483** (2012) 187.
[6] Orlov A. O., Lent C. S., Thorpe C. C., Boechler G. P. and Snider G. L., *Jpn. J. Appl. Phys.*, **51** (2012) 06FE10.
[7] Granger L. and Kantz H., *Phys. Rev. E*, **84** (2011) 061110.
[8] Sagawa T. and Ueda M., *Phys. Rev. Lett.*, **102** (2009) 250602.
[9] Jacobs K., *Phys. Rev. E*, **86** (2012) 040106.
[10] Sagawa T. and Ueda M., *Phys. Rev. Lett.*, **109** (2012) 180602.
[11] Mandal D. and Jarzynski C., *Proc. Natl. Acad. Sci. U.S.A.*, **109** (2012) 11641.
[12] Horowitz J. M., Sagawa T. and Parrondo J. M. R., arXiv:1210.6448 (2012).
[13] Bennett C. H., *Int. J. Theor. Phys.*, **21** (1982) 905.
[14] Norton J. D., *Stud. Hist. Philos. Sci. Part B: Stud. Hist. Philos. Mod. Phys.*, **42** (2011) 184.
[15] Shizume K., *Phys. Rev. E*, **52** (1995) 3495.
[16] Fahn P. N., *Found. Phys.*, **26** (1996) 71.
[17] Piechocinska B., *Phys. Rev. A*, **61** (2000) 062314.
[18] Plenio M. B. and Vitelli V., *Contemp. Phys.*, **42** (2001) 25.
[19] Dillenschneider R. and Lutz E., *Phys. Rev. Lett.*, **102** (2009) 210601.
[20] Gaveau B. and Schulman L., *Phys. Lett. A*, **229** (1997) 347.
[21] Seifert U., *Phys. Rev. Lett.*, **95** (2005) 040602.
[22] Esposito M. and Van den Broeck C., *EPL*, **95** (2011) 40004.
[23] Esposito M., *Phys. Rev. E*, **85** (2012) 041125.
[24] Seifert U., *Rep. Prog. Phys.*, **75** (2012) 126001.
[25] Vaikuntanathan S. and Jarzynski C., *EPL*, **87** (2009) 60005.
[26] Deffner S. and Lutz E., arXiv:1201.3888 (2012) .
[27] Cover T. M. and Thomas J. A., *Elements of Information Theory* (John Wiley & Sons) 2006.
[28] Van den Broeck C. and Esposito M., *Phys. Rev. E*, **82** (2010) 011144.
[29] Esposito M. and Van den Broeck C., *Phys. Rev. E*, **82** (2010) 011143.