J. Stat. Mech. (2010) P10009

The many faces of the second law

C Van den Broeck

University of Hasselt, B-3590 Diepenbeek, Belgium
E-mail: christian.vandenbroeck@uhasselt.be

Abstract. There exists no perpetuum mobile of the second kind. We review the implications of this observation on the second law, on the efficiency of thermal machines, on Onsager symmetry, on Brownian motors and Brownian refrigerators, and on the universality of efficiency of thermal machines at maximum power. We derive a microscopic expression for the stochastic entropy production, and obtain from it the detailed and integral fluctuation theorem. We close with the remarkable observation that the second law can be split in two: the total entropy production is the sum of two contributions each of which is growing independently in time.

Contents

1. Introduction: the perpetuum mobile of the second kind

The alchemists' search for the philosopher's stone is legendary. This magic object was supposed to make possible the transmutation of the elements. It would provide an elixir of life. But these 'early scientists' were also obsessed with the search for a perpetuum mobile. They proposed all kinds of devices, mechanical, magnetic, chemical and other [1]. Since radioactive decay and the structure of the atom have become understood, it has been known that the transmutation of elements is possible (but usually very difficult and costly). Also the elixir of life has been identified[1]. Concerning the perpetuum mobile, the search for such engines continues up to this very day. But (alas) so far to no avail. We now believe that two very basic principles of physics prevent the construction of perpetuum mobiles. One is the conservation of energy, making impossible the so-called perpetuum mobile of the first kind. This conservation principle derives from the translational invariance in time of the laws of physics. The other principle is more subtle and is still the subject of debate. It is, as we will explain in some detail below, related to the time-reversal symmetry of the physical laws. It touches on the foundations of statistical mechanics and thermodynamics, but it has also been related to the expansion of the universe, the collapse of the wavefunction in quantum mechanics, the role of chaos in making trajectories effectively unpredictable, and even to the physics of computation.

   To make a start, we will assume that a perpetuum mobile of the second kind is not possible, and investigate its consequences. More precisely, we suppose (stated somewhat vaguely) that it is impossible to extract work out of a single heat bath at equilibrium. By heat bath we understand a confined system that has been allowed to relax to an equilibrium state, and is only allowed to exchange energy with its surrounding via heat. A direct consequence is that the famous Escher waterfall [2] cannot occur in the physical

---

[1] According to an undisclosed source, it is called 'Belgian beer'.

world. In this purposefully misleading drawing it looks as if water is running around while 'falling down all the time'. Either it is indeed falling down and we extract energy out of the gravitational field, which is fine, but the center of gravity of the device then has to be going down. If this is not the case, energy is conserved, so there is no conflict with the concept of a perpetuum mobile of the first kind, but the flow can, according to our assumption, not be an equilibrium situation[2]. Otherwise, we have a perpetuum mobile of the second kind since we can extract work by using a watermill. From very early on [3], it has been the matter of passionate discussions whether this restriction could be circumvented when operating on a small scale. Indeed, even in a system at equilibrium at a finite temperature (and even at zero temperature in an 'empty system' if one takes into account vacuum fluctuations), there is a ongoing dynamical activity at small scale, and it is tempting to assume that one can extract energy out of these small scale fluctuations. A device that would perform such a task was first proposed by Maxwell, and is nowadays referred to as a Maxwell demon [3]. The debate raged on for a long time including a discussion about the thermodynamic cost of measurement and computation [4]. The consensus today is that there are no Maxwell demons, and the basic impossibility of a perpetuum mobile of the second kind prevails (see [5] for a further discussion). As a consequence, a ratchet device like the one conceived by Smoluchowski [6] and discussed further by Feynman [7] will indeed fail to rectify fluctuations. In conclusion, there are no macroscopic currents in a heat bath at equilibrium, and even at the microscopic scale, there are no net-currents (see footnote 2). In other words, even though time is present at the small scale, in the sense that the state of the system changes, there is no directionality of time: statistically speaking, any transition and its time-reverse are equally probable. There is no way to distinguish the past from the future. What we have enunciated here corresponds to the famous principle of detailed balance [8]. Together with a corollary result discussed below, the symmetry of Onsager coefficients, it earned Lars Onsager his Nobel prize. This principle has been the basis for several fundamental advances in statistical mechanics including the fluctuation dissipation theorem and Green–Kubo relations for transport coefficients. Much of what will be discussed below, and in particular the by now famous fluctuation theorem (see below), can be considered as a continuation and extension of this pioneering work.

2. Efficiency of thermal engines

The progress in design and construction of the steam engine by James Watt and others promoted the industrial revolution in the 19th century. There was an obvious interest in the efficiency of such machines. The impossibility of a perpetuum mobile of the second kind stipulates that the efficiency of a thermal engine operating with a single heat bath has to be zero, since no work can be extracted. Starting from this premise, the French engineer Sadi Carnot made a momentous observation. In his book 'Réflexions sur la puissance motrice du feu et sur les machines propres a développer cette puissance', he showed that the efficiency of a thermal machine, operating between two heat reservoirs

---

[2] The discussion given here is slightly misleading. Due to conservation of angular momentum, an overall circular flow can exist in a system in equilibrium. In the quantum mechanical context, one can also refer to superfluidity and superconductivity, corresponding to the existence of a non-dissipative flow. It is however customary to consider the frame of reference in which the overall flows, both momentum and angular momentum, are zero.

at temperatures T_h and T_c respectively, is at most given by

$$\eta = \frac{W}{Q_h} \leq \eta_C = 1 - \frac{T_c}{T_h}. \tag{1}$$

Here W and Q_h stand for the work and amount of heat coming from the hot reservoir, respectively. In view of the simplicity of the premise compared to the enormous impact of this result as the founding stone of thermodynamics, Carnot's derivation can be considered as one of the most impressive in the history of physics. It proceeds in two steps. First Carnot constructs a machine that achieves Carnot efficiency η_C. It consists of an ideal gas undergoing a cycle of isothermal expansion in contact with the hot reservoir, adiabatic expansion while disconnected from the reservoirs, isothermal compression in contact with the cold reservoir, and adiabatic compression while disconnected from the reservoirs. This cycle is performed such that the system is asymptotically close to equilibrium during the entire process. Since the equation of state is known for an ideal gas, the heat and work can be evaluated in each of the steps. By such an explicit calculation, Carnot finds that the efficiency is indeed equal to η_C. The fact that the cycle runs essentially at equilibrium provides a crucial by-product: the engine can also be operated in the other direction. It then functions as a heat pump delivering an amount of heat Q_h to the hot reservoir upon input of an amount of work W, again with their ratio given by the efficiency η_C, cf (1). The second step of the derivation is a proof *ex absurdo*: suppose there is another engine that can deliver work at a higher efficiency than Carnot efficiency, i.e., for a certain amount of heat Q_h it produces more work than specified by (1). Using the Carnot machine, one can however re-conduct to the hot reservoir all the heat Q_h extracted by the former engine, by using an amount of work specified by (1), i.e., by using less work than the first engine has produced. The net result is a remaining amount of work, obviously extracted from the cold reservoir alone, which is in contradiction with the premise.

It is revealing to rewrite (1) in another way, using conservation of energy, $W = Q_h - Q_c$, where Q_c is the heat delivered to the cold reservoir:

$$\Delta S_{tot} = \frac{Q_h}{T_h} - \frac{Q_c}{T_c} \geq 0. \tag{2}$$

Since Q_h/T_h and $-Q_c/T_c$ are the entropy changes in the hot and cold reservoirs, and the entropy of the gas is unchanged since it returns to the same state after a cycle, the upper bound of Carnot efficiency is nothing but a statement of the second law of thermodynamics, stipulating the increase of the (total) entropy [9][3]. Carnot efficiency is reached for zero overall entropy production, i.e., when all processes take place reversibly. As a corollary, we conclude that such a device can always be operated in both directions: reaching Carnot efficiency implies that it can function either as a thermal engine or as a refrigerator and a heat pump.

3. Onsager symmetry

The machine considered by Carnot operates in a cyclic manner. But the upper limit applies to any thermal engine, and in particular to a machine operating in a steady state

---

[3] Note that Clausius first derives from Carnot efficiency the fact that entropy is a state function, followed by a proof of the increase of total entropy.

4

rather than cyclical manner. A simple model for such a construction involves two applied forces. One is the thermal gradient ΔT inducing the powering heat flux. The other is a load force F which produces the work, by for example lifting a weight over a distance x. There are also two fluxes, $J_1 = \dot{Q}_h$, the heat flux leaving the hot reservoir, and $J_2 = \dot{x}$, the speed at which the force lifts the (appropriate) load. We will be concerned with the case of a small (but finite) T-gradient so that one can define a base temperature T. One identifies the thermodynamic forces $X_1 = \Delta T/T^2$ and $X_2 = F/T$,[4] which are driving the fluxes. Since the fluxes vanish in the absence of forces, one finds by a Taylor expansion of fluxes in terms of the forces to linear order:

$$J_1 = L_{11}X_1 + L_{12}X_2 \qquad J_2 = L_{21}X_1 + L_{22}X_2. \tag{3}$$

The coefficients L form a matrix of linear response, also called the Onsager matrix. The diagonal elements correspond to known material constants: L_{11} is essentially the heat conductivity and L_{22} the mobility. There is however no *a priori* reason why the off-diagonal elements would be zero[5]. In fact, they precisely correspond to the phenomena we are interested in: L_{21} describes motion in response to a heat flux, i.e., it is the thermal engine; L_{12} corresponds to a heat flux induced by a force, hence we are talking about a heat pump or a refrigerator.

Returning to the issue of efficiency, we recall that Carnot efficiency can only be approached when working in the vicinity of equilibrium. Hence both fluxes should become vanishingly small. If the determinant of the Onsager matrix is not zero this can only happen in the limit $X_1 \to 0, X_2 \to 0$, which is not the case we want to consider. There is, however, an alternative, namely to consider the limit of a zero determinant, $L_{11}L_{22} = L_{12}L_{21}$. This is equivalent to saying that both fluxes become proportional to each other: $J_1 = \epsilon J_2$, with ϵ the unit of energy linking both fluxes. In particular one has that $\epsilon = L_{11}/L_{21} = L_{12}/L_{22}$. In this case, the application of a stopping force to the flux J_2, namely $X_2^{\text{stop}} = -L_{21}X_1/L_{22}$, will result in a vanishing heat flux J_1. An engine with fluxes proportional to each other is called strongly coupled. Note that such a situation is by no means exceptional. It can be easily realized in nano- and bio-technology, and it is a natural constraint in chemical reactions, see also further examples given below.

Turning to the calculation of the efficiency for such a strongly coupled steady state engine, one finds:

$$\eta = \frac{W}{Q_h} = \frac{\dot{W}}{\dot{Q}_h} = \frac{-F\dot{x}}{\dot{Q}_h} = \frac{-TX_2J_2}{J_1} = -TX_2/\epsilon. \tag{4}$$

Note furthermore that $\eta_c = \Delta T/T = TX_1$. Upon applying a stopping force $X_2 = X_2^{\text{stop}} = -X_1L_{21}/L_{22}$, we find that reproducing Carnot efficiency $\eta = \eta_c$ implies the equality $L_{21}/(L_{22}\epsilon) = 1$, leading to the famous symmetry of the Onsager coefficients:

$$L_{12} = L_{21}. \tag{5}$$

We conclude that the impossibility of a perpetuum mobile of the second kind implies Carnot efficiency, which in turn implies Onsager symmetry (in strongly coupled systems).

---

[4] With these definitions one can show that the rate of total entropy production is given by $\dot{S}_{\text{tot}} = J_1X_1 + J_2X_2$, cf [10].

[5] As is well known, coefficients can be zero for symmetry reasons such as Curie's principle stating that scalar and vectorial properties in isotropic systems do not couple. In the present case, we consider a one-dimensional setting so that both the temperature gradient and the force are scalar variables.

Onsager in fact proved that this symmetry prevails even when the two fluxes are not proportional. The simplest way to derive this result today is by invoking the fluctuation theorem which will be presented below. Onsager symmetry in the context of thermal engines can be seen as an extension to the regime of linear response of the duality between the engine and refrigerator that we mentioned before for a reversible Carnot engine. A thermal engine and a mechanical refrigerator always appear in tandem also outside of the regime of reversible operation (but limited to linear response), and with equal intensity if the proper definitions of thermodynamic forces and fluxes are used.

There are many examples and applications of Onsager symmetry. We mention a surprising recent illustration, which is relevant in the realm of nanotechnology. While we are now able, not only to observe and measure at the very small scale, but also to design and assemble, the next crucial step is to add functionality. One problem is that fluctuations (and quantum effects) have to be taken into account at such a small scale. Copying the mode of operation from a macroscopic counterpart may not be the best way to go. A well documented example concerns the so-called Brownian motors [11]. Such motors have no macroscopic counterpart: they cease functioning in the macroscopic limit. Of specific interest to us is the thermal Brownian motor arising from the discussion of the Maxwell demon mentioned before. Feynman shows that this device, while it cannot rectify thermal fluctuations from a single heat bath, does function as a thermal engine when two reservoirs at different temperature are available. Feynman's mesoscopic analysis is however flawed: his construction cannot operate reversibly and hence does not reach Carnot efficiency [12] (there is a way around this difficulty which interestingly relies on the use of strongly coupled engines as discussed above, cf [13] for a further discussion). He neglected the thermal conduction that appears via the degree of freedom that couples both heat baths. This issue is also the basis of a debate on the foundations of macroscopic thermodynamics, as illustrated by the analysis of the so-called adiabatic piston [14]. In order to clarify the origin of these difficulties, an exact calculation from first principles calculation is welcome. This was achieved for a simplified version of the Feynman–Smoluchowski device [15]. Subsequently is was realized that due to Onsager symmetry, a thermal Brownian motor implies the existence of a Brownian refrigerator [16]. Making efficient small scale refrigerators appears to be a technological challenge. The Brownian refrigerator has the great advantage that it operates better as the device becomes smaller, justifying its name as the smallest refrigerator in the world [17].

4. Efficiency at maximum power

We stressed that reaching Carnot efficiency requires reversible, i.e., infinitely slow, operation. The power, work over time, for such a machine is therefore zero. Such a machine has limited technological interest. It is more natural to search for engines working at high power. This raises the question of the corresponding efficiency. After several earlier contributions [18], the issue was taken up by Curzon and Ahlborn in an influential paper [19]. They modified the Carnot cycle by assuming a difference of temperature between the reservoirs and the gas, while neglecting the dissipation associated to the adiabatic phases. They obtained the following remarkable result for the efficiency of this engine at maximum power (maximizing with respect to the gradients of temperature

between gas and reservoirs):

$$\eta_{CA} = 1 - \sqrt{\frac{T_c}{T_h}}. \tag{6}$$

The beauty of this result and the fact that it can be found in a number of other model systems raises the question of its universality. It turns out however that η_{CA} is neither a lower nor an upper bound of efficiency at maximum power. Furthermore, one needs to specify more precisely with respect to which variables the maximum is taken, and the latter may even not be unique. It is however revealing to study efficiency at maximum power when operating close to equilibrium. Considering a small temperature difference between hot and cold reservoir, it follows that $\eta_C = 1 - T_c/T_h$ is a small quantity and we can consider the expansion of η_{CA} in terms of η_C:

$$\eta_{CA} = 1 - \sqrt{1 - \eta_C} = \frac{\eta_C}{2} + \frac{\eta_C^2}{8} + \frac{\eta_C^3}{16} + \cdots. \tag{7}$$

In the region of linear response, i.e. to linear order in η_C, one can prove that there is a unique maximum of the power, and that the efficiency at this maximum is at most half of the Carnot efficiency [21]. The limit is achieved for strongly coupled systems, which was, as discussed above, also the prerequisite for achieving Carnot efficiency itself. The proof for the steady state engine described before, cf (3), goes as follows. In the regime of linear response, the power $X_2 J_2 = X_2(L_{21}X_1 + L_{22}X_2)$ has a maximum at half the stopping force $X_2 = X_2^{stop}/2 = -L_{21}X_1/(2L_{22})$. The corresponding efficiency, cf (4), is given by:

$$\eta = -TX_2/\epsilon = \eta_C \frac{L_{21}L_{12}}{2L_{11}L_{22}} = \frac{\eta_C}{2}. \tag{8}$$

The basic ingredient for the derivation is Onsager symmetry, which follows from Carnot efficiency which derives from the impossibility of a perpetuum mobile of the second kind. We conclude that the universal coefficient $1/2$ is a consequence of the impossibility of a perpetuum mobile of the second kind.

We next consider the coefficient $1/8$ of the quadratic term in (7). It turns out that one can again prove its universality under the additional assumption of a left/right symmetry in the system [22]. This result is very surprising because all the traditional 'universal' results in nonequilibrium statistical mechanics are limited to the region of linear response. The above universality refers to the coefficient of the quadratic nonlinearity. To bridge the gap to the nonlinear regime, one needs to invoke the recent spectacular advances in our understanding of the second law, which we will regroup under the common name of the fluctuation theorem. Paradoxically, while the second law was originally intended for the description of macroscopically large systems, it is its application to small scale systems that is providing the essential breakthrough. The fluctuation theorem allows the above result to be derived, but more importantly it incorporates fully the impact of micro-reversibility, it clarifies the deeper meaning of entropy production, and it naturally extends thermodynamics into a new field 'stochastic thermodynamics' which deals with the thermodynamics of single trajectories. The new developments concerning the efficiency at maximum power in thermal engines have stimulated the study of several exactly solvable models [23]. They are all found to be consistent with the predicted universality.

5. Dissipated work

The fluctuation theorem is most easily derived from microscopic theory in one of the traditional scenarios of thermodynamics: the work performed on a system in contact with a heat bath. We first reproduce the result of macroscopic thermodynamics. A system is initially at equilibrium in contact with a heat bath at temperature T. It undergoes a transformation towards a new equilibrium state ending again in contact with the same heat bath. Note that the heat bath need not stay connected with the system during the process. It is sufficient that it is so at initial and final times. Let us call U and S the energy and entropy of the system, which are well defined for the system at equilibrium, in particular in the initial and final equilibrium states. One also introduces the corresponding free energy $F = U - TS$. When applied to small scale systems, we include the energetic terms coming from the interaction with the heat bath in the definition of these quantities. During the transformation, a certain amount of external work W is performed. This is an algebraic quantity with a positive sign if work is performed on the system. We further call Q_m the amount of heat dumped into the heat bath (or medium). The latter is supposed to be and remain at equilibrium at temperature T throughout the entire process. The change in its entropy is therefore $\Delta S_m = Q_\mathrm{m}/T$. We stress that while this result is in agreement with equilibrium statistical mechanics, in the sense that this relation is obtained if the heat bath is described by an equilibrium ensemble at all times, there is no microscopic foundation for this result. In particular the fact that the heat bath remains in its equilibrium distribution, while in contact with a system perturbed far from equilibrium, is strictly speaking not compatible with Hamiltonian dynamics except if the perturbation is infinitely small. The use of an idealized heat bath or an equivalent substitute can, as far as I know, not be avoided and forms the basis of any entropy calculation, see however [24]–[26] for some interesting new ideas. Conservation of energy implies that $\Delta U = W - Q_\mathrm{m}$. Furthermore since both the final and initial temperature of the system are equal to the temperature T of the heat bath, one has $\Delta F = \Delta U - T\Delta S$. One concludes that for the total entropy production $\Delta S_\mathrm{tot} = \Delta S + \Delta S_m$:

$$T\Delta S_\mathrm{tot} = W - \Delta F. \tag{9}$$

This result is very interesting for the following reasons. First it allows us to identify the total entropy production in terms of a mechanical quantity, the amount of work W. Note that ΔF is completely specified since it is the change in free energy of the system between its well-defined final and initial equilibrium states. Second this identification is valid independent of the size of the considered system. For a small system, we stress that the properties of the 'small' system, its energy, entropy and free energy, include the term related to the interaction of the system with the reservoir. Indeed, the crucial ingredient for the identification of the entropy production is the formula $\Delta S_m = Q/T$, which applies to the degrees of freedom specific to the reservoir, and not for the interaction energy. Alternatively, one can consider the limit of infinitely weak coupling between system and reservoir, so that the interaction energy can be neglected (but this also implies that the actual relaxation towards equilibrium via this very weak coupling will take a very long time). The case of a small system has a particular interest since the work will in general differ from one run to another. As a result the amount of energy evacuated to the reservoir and the concomitant entropy production given by (9) is also a fluctuating quantity. This

opens the door to the calculation of the probability distribution for the total entropy produced upon perturbing a small system in contact with a heat reservoir. Third, the above result is valid for any procedure in which the state of the system is changed from its initial to final equilibrium. One can expect that the amount of work will, even for a macroscopically large system, depend on the way this transformation is performed. For example, it is expected that the work will be larger for brutal transitions. In fact, returning to macroscopic thermodynamics (entailing the consideration of a macroscopically large system) we can invoke the second law of thermodynamics, stipulating the increase in total entropy $\Delta S_{\text{tot}} = W - \Delta F \geq 0$. This leads to the familiar statement that the amount of work performed on a system in contact with a heat bath, needed to switch between an initial and final equilibrium state, is at least equal to the difference in free energy. When the latter is negative, one can in principle extract work (i.e. W is then negative) from this decrease of free energy. The equality sign is reached for reversible transformations. One can wonder how this result relates to the entropy production for a small system. Both W and ΔS_{tot} are then random variables and there is no reason why the above inequality would persist. The macroscopic limit is recovered by putting a large number of such systems in parallel, which is tantamount to repeating the experiment many times. By the law of large numbers, we conclude that the second law applies to the averages of entropy production or work

$$\langle \Delta S_{\text{tot}} \rangle = \langle W \rangle - \Delta F \geq 0. \tag{10}$$

We now present a microscopic calculation, providing an explicit expression for W and hence for $\Delta S_{\text{tot}} = W - \Delta F$. This result is consistent with the above inequality, but provides a deeper formulation of the second law, and will lead to the replacement of its usual formulation as an inequality for the average entropy production by equalities known as the integral and detailed fluctuation theorem for the stochastic entropy production.

6. Fluctuation theorem

We set out to derive an exact microscopic expression for the work upon bringing a system from one canonical equilibrium state into another one [27]. The system is described by the Hamiltonian $H(q, p; \lambda)$, where q, p denote all microscopic degrees of freedom (position and momenta) and λ is a control parameter, that describes the energy exchange with an external mechanical device. For simplicity, we will assume that the Hamiltonian is an even function with respect to inversion of momenta. The system is initially assumed to be in canonical equilibrium at temperature T at the value λ_{A} of the control parameter. This can be achieved by putting it in contact with a heat bath at this temperature. For simplicity of the argument we will be assume that the heat bath is subsequently disconnected. At the end of the forward process, the system will be reconnected so that it returns to canonical equilibrium. The disconnection and reconnection can be replaced by an assumption of weak coupling. Note also that we can always include in the definition of the system a large part of the heat bath such that the boundaries of the system are during the time of the perturbation not affected by it and hence this boundary region is and remains at equilibrium throughout the process. The disconnection and reconnection should thus have no influence on the work performed nor the change in state of the system during the transition. During the experiment, the control parameter is changed following

a specific protocol $\lambda(t)$ from λ_A to λ_B. Since the reservoir is disconnected no energy is exchanged other than the work W performed by the external agent on the system. We will furthermore consider the time-reversed scenario, with the system initially at canonical equilibrium at the same temperature T, but at the value λ_B of the control parameter. The latter is now changed according to the exact time-reversed protocol. We will indicate with a superscript tilde all the variables that relate to this so-called reverse experiment.

We first set out to calculate the work $W(q, p; t)$ done along the forward process, for a specific microscopic trajectory, namely the one that passes through the phase point (q, p) at time t. Since the dynamics are deterministic, there is precisely one such trajectory. Let us call (q_0, p_0) and (q_1, p_1) the corresponding initial and final phase points uniquely determined by (q, p) (for the specified scenario $\lambda(t)$). Note also that there is a one-to-one correspondence with the time-reversed trajectory in the time-reversed protocol which, starting from $(q_1, -p_1)$, goes through $(q, -p)$ and, finally, into $(q_0, -p_0)$. For simplicity of notation, we will use the forward time to express times in both forward and backward scenarios. By conservation of total energy, one has that $W(q, p; t) = H(q_1, p_1; \lambda_B) - H(q_0, p_0; \lambda_A)$. Now, since the phase space density is conserved along any Hamiltonian trajectory, one has, in both the forward and backward processes that

$$\rho(q, p; t) = \rho(q_0, p_0; t_0) = \frac{\exp[-\beta H(q_0, p_0; \lambda_A)]}{Z_A} \tag{11}$$

$$\tilde{\rho}(q, -p; t) = \tilde{\rho}(q_1, -p_1; t_1) = \frac{\exp[-\beta H(q_1, -p_1; \lambda_B)]}{Z_B}, \tag{12}$$

where Z_A and Z_B are partition functions at the equilibrium states A and B, respectively, and $\beta = T^{-1}$ is the inverse temperature (for simplicity, we set Boltzmann's constant equal to 1). These expressions allow us to eliminate the Hamiltonian (which is supposed to be even in the momenta) at initial and final times in favor of the phase space density at any intermediate time point. We thus obtain the following basic expression for the work of the change of the total entropy, associated to the specific trajectory under consideration:

$$T \Delta S_{\text{tot}} = W(q, p, t) - \Delta F = T \ln \frac{\rho(q, p, t)}{\tilde{\rho}(q, -p, t)} = T \ln \frac{\text{Prob}(\text{path}(q, p, t))}{\tilde{\text{Prob}}(\tilde{\text{path}}(q, -p, t))}. \tag{13}$$

$\Delta F = T \ln Z_B/Z_A$ is the free energy difference between the final and initial equilibrium states. Note also that the (phase space) probability densities $\rho(q, p, t)$ and $\tilde{\rho}(q, -p, t)$ for being in the microstates (q, p, t) and $(q, -p, t)$ at the forward time in the forward and reverse experiments are equal to the probability densities for observing the corresponding paths, which we represented by $\text{Prob}(\text{path}(q, p, t))$ and $\tilde{\text{Prob}}(\tilde{\text{path}}(q, -p, t))$. Equation (13) is the basic result which we are after: it gives the explicit expression for the trajectory dependent work. The trajectory is specified by the microscopic state at any intermediate time or, equivalently, by the entire trajectory itself (since there is a one-to-one correspondence between both). The probability distribution for the work is also in principle known since $\rho(q, p, t)$ is the probability density to observe the corresponding value. Let us first verify that the above expression is in agreement with the usual formulation of the second law. By performing the average of $W(q, p, t)$ with respect

to $\rho(q, p, t)$ over the entire phase space Γ, we find:

$$\langle \Delta S_{\text{tot}} \rangle = \frac{\langle W(q, p, t) \rangle - \Delta F}{T} = \int d\Gamma \, \rho(q, p, t) \ln \frac{\rho(q, p, t)}{\tilde{\rho}(q, -p, t)} = D(\rho(q, p, t) || \tilde{\rho}(q, -p, t))$$

(14)

which features the relative entropy of Kullback–Leibler distance $D(\rho(q, p, t) || \tilde{\rho}(q, -p, t))$ between the probability distributions $\rho(q, p, t)$ and $\tilde{\rho}(q, -p, t)$ [28, 29]. As is well known, it is a positive quantity. The expression (13) however allows us to derive a stronger result:

$$\langle e^{-\Delta S_{\text{tot}}} \rangle = \int d\Gamma \, \rho(q, p, t) \frac{\tilde{\rho}(q, -p, t)}{\rho(q, p, t)} = \int d\Gamma \, \tilde{\rho}(q, -p, t) = 1.$$

(15)

This is called the integral fluctuation theorem. This equality implies the 'traditional second law' $\langle \Delta S_{\text{tot}} \rangle \geq 0$ by application of Jensen's inequality. We now proceed to argue that whenever we have a integral fluctuation theorem, we also have a detailed fluctuation theorem. Indeed one can write:

$$\langle e^{-\Delta S_{\text{tot}}} \rangle = \int d\Delta \, S_{\text{tot}} P(\Delta S_{\text{tot}}) e^{-\Delta S_{\text{tot}}} = \int d\Delta \, S_{\text{tot}} \tilde{P}(-\Delta S_{\text{tot}}) = 1,$$

(16)

where we introduced the function $\tilde{P}$ defined by the relation:

$$\frac{P(\Delta S_{\text{tot}})}{\tilde{P}(-\Delta S_{\text{tot}})} = e^{\Delta S_{\text{tot}}}.$$

(17)

This relation is called the detailed fluctuation theorem. Note that $\tilde{P}$ is a probability distribution: it is positive and normalized. Note that the detailed fluctuation theorem in turn implies the integral fluctuation theorem (which in fact expresses the normalization of $\tilde{P}$), so they are equivalent statements. Without further specification, one cannot expect that $\tilde{P}$ is related in a simple way to the probability distribution P, other than via its 'definition' (17). However, the random variable under consideration, ΔS_{tot}, is equal to the logarithm of the ratio of two probabilities ρ and $\tilde{\rho}$, the latter being related via an involution ($\tilde{\tilde{\rho}} = \rho$), namely time-reversal. It is then easy to prove that $\tilde{P}$ is indeed the probability distribution when the entropy production is sampled with respect to the time-reversed experiment, as we anticipated by using the same superscript tilde notation, see [30] for more details.

Some additional comments are in order, identifying the integral or detailed fluctuation theorem as the basic and consistent formulation of the second law. First (14) states that the average entropy production is the amount of time-symmetry breaking, i.e. the level of difficulty in distinguishing snapshots of forward and backward experiments. In the case of a system of large size N, the time-symmetry breaking can become very prominent. The probability distribution in the forward experiment ρ will be dominated by typical configurations. When the system in far out of equilibrium, it will be exponentially unlikely (exponentially in N) to observe the same micro-configuration (with inverted momenta) in the backward experiment, leading to an extensive entropy (proportional to N). In this case, one will need to perform exponentially many experiments to identify the entropy production via the above formula, and other approaches (involving reduced description via elimination of fast variables) are more appropriate. As a somewhat paradoxical consequence, we note that entropy production as the breaking of time-symmetry can be

readily measured in small systems, where the arrow of time is less pronounced, but not in large systems where the arrow of time is 'exceedingly' dominant. Note also that the usual wording in macroscopic thermodynamics that entropy cannot decrease is strictly speaking incompatible with the above formula, since it would imply an infinitely large entropy production. The correct formulation is that decreases of total entropy are exponentially unlikely, as stipulated by the detailed fluctuation theorem.

Second we consider the limit of a reversible transformation, with the system being at equilibrium at all times. We can still follow single trajectories. Even at equilibrium there will be energy exchanges with the thermal bath and we can wonder about the corresponding total entropy change. Because of reversibility, the probability for a transition between two states with difference ΔS_{tot} in entropy is as large as the probability for the inverse transition $-\Delta S_{\text{tot}}$: $P(\Delta S_{\text{tot}}) = P(-\Delta S_{\text{tot}})$. In combination with the detailed fluctuation theorem this implies that the entropy production has to be essentially zero (with probability one) for all trajectories $\Delta S_{\text{tot}} \equiv 0$ or $P(\Delta S_{\text{tot}}) = \delta(\Delta S_{\text{tot}})$. It is revealing to show how this property follows from energy conservation. We split the total entropy change for a given trajectory into a medium contribution ΔS_m plus a system contribution ΔS: $\Delta S_{\text{tot}} = \Delta S_m + \Delta S$. Let Q and ΔH be the heat evacuated to the medium and energy change of the system. By energy conservation $\Delta H = -Q$. Obviously $\Delta S_m = Q/T$. We now define the entropy of the system while being in a specific state (q, p) following Seifert [43]: $S(q, p) = -\ln P(q, p)$. Since the system is at equilibrium one has $P(q, p) = P^{\text{eq}}(q, p) \sim e^{-\beta H(q,p)}$ and consequently $\Delta S = \beta \Delta H$. We conclude that total entropy being constant along the trajectories is tantamount to energy conservation. This should be contrasted with the heuristic discussion in macroscopic thermodynamics where one argues that there are fluctuations away from the state of maximum entropy.

Third, we note that since we know the work associated to every specific trajectory, we can also evaluate the average work of trajectories that behave in a specified manner, for example they pass through a specific subpart of phase space. This allows us to make the connection with the thermodynamics of computation and information processing and connect to the issues of Maxwell demons, etc [27, 31, 32].

Fourth, the expression for the stochastic trajectory dependent entropy given in (13) remains valid and is consistent with the one obtained in the context of a stochastic description. We briefly comment on the relationship between the two approaches, see also section 7 for a short derivation. First we mention an essential difference between the stochastic and Hamiltonian description. A Markovian stochastic description is irreversible from the start. So the relaxation of the system back to equilibrium, which we realize in the Hamiltonian context via the idealized heat bath, is built into the stochastic description. While relaxation is automatically taken care of in the stochastic approach, it is not obvious how the physics of the problem, and in particular fundamental physical properties, such as the underlying reversibility, is properly incorporated. After some early attempts [33] a procedure was proposed for the average entropy production for chemical models in [34] and for general Markov processes in [35]. The proper macroscopic entropy is recovered, but the bonus of the stochastic description is that the expression for the average entropy production includes the effect of the fluctuations. It could for example be shown that the fluctuations reinforce the minimum entropy production by adding a non-extensive negative contribution for steady states in the linear regime close to equilibrium [35]. We introduced at that time the term 'stochastic thermodynamics' to refer to this new

approach [36]. The real breakthrough came with the discovery of the work and fluctuation theorems, in which one studies the trajectory dependent quantities instead of the usual thermodynamic averages. The results obtained in this context deal with many different types of descriptions, ranging from Hamiltonian over thermostated system to stochastic processes, and involve a large number of significant contributions [37]. These separate results have created the impression that there are several fluctuation theorems, depending on which system, initial and final conditions, state or description are being used. We disagree with this opinion. The central and crucial observation is that the total stochastic entropy production is given by the logarithm of the ratio of two probabilities, one for a forward and another for the proper time-reversed experiment, as in (13), see also the next paragraph. It is also gratifying to realize that the stochastic analog can be derived directly from the above microscopic result by assuming that the stochastic description is for the physical system under consideration correct (i.e. that it is an exact consequence, in a limit which needs not to be specified, of the microscopic laws), cf [5, 38] for more details.

Fifth, one may wonder whether a formula like (13) or (14), even if they are exact and have a wide range of application, is useful and has, aside from its purely theoretical value, practical applications. Considering the formula (14) for the average entropy production, one can make the following observation. The integral is dominated by the typical realizations of the forward process, i.e. by the phase space points (q, p), or the paths passing through these points, for which $\rho(q, p; t)$ is large. The probability for these points or their corresponding trajectories has to be compared with the probability for the time-reversed trajectory in the backward process. When such a trajectory is atypical in the backward process, i.e., when it has a low probability, then there is significant entropy production. Let us turn the argument around. Entropy is an extensive quantity. Hence, in the case where one is not at equilibrium, the entropy is typically proportional to the number of degrees of freedom. To reproduce this result via (14) we conclude that the time-reversed versions of the typical forward trajectory have to be exponentially unlikely in the backward process. This is just another way of saying that entropy production is associated to the arrow of time. Playing the movie backward will look very different when the entropy is significant. The drawback is that, because of the extensivity of entropy and the corresponding exponential dependence of probabilities, the observation of the backward trajectories dominating the relative entropy in (14) is only feasible when we are either very close to equilibrium or when we are dealing with small systems (more precisely, with systems that have only a few degrees of freedom that are not at equilibrium). The situation is, not surprisingly, similar to that for the application of the work, see below [42]. This however is not the end of the story: there is one specific property that makes our formula quite useful, namely the chain rule of relative entropy. It states that the relative entropy of two probability distributions can only become smaller upon coarse-graining, i.e. upon considering reduced distributions. Hence

$$0 \leq D(P, \tilde{P}) \leq D(\rho \| \tilde{\rho}) = \langle \Delta S_{\text{tot}} \rangle, \tag{18}$$

for any coarse-grained distributions P and $\tilde{P}$ of ρ and $\tilde{\rho}$. In other words, any observed arrow of time, i.e., statistical distinction between forward and backward processes, gives rise to a lower bound for the entropy production. In the absence of any information on time-asymmetry, the lower bound is zero and one recovers the 'minimal' statement

incorporated in the second law, namely that the total average entropy cannot decrease. The above results also suggest that the choice of variables, which strongly display time-asymmetry, can capture most or all of the entropy production. For example, upon quenching an external potential for a Brownian particle in a heat bath, the only variable that is not at equilibrium at the instant of the quench is the position of this particle [27]. The full entropy production is reproduced by the time-asymmetry of this single variable. As one measures at later times after the quench, the signature of irreversibility moves into other variables, in particular the heat bath variables, and one needs to include their time-arrow to reproduce the total entropy production. These features can be nicely illustrated in more detail in exactly solvable harmonic models [39]. The fact that there are 'macroscopic expressions' for entropy production is also related to this observation. We are referring to situations in which the arrow of time is only present at the level of the observed macro-variables and all the other microscopic degrees of freedom are at equilibrium (this is sometimes referred to as the situation of local equilibrium).

Sixth, we focus on a specific version of the fluctuation theorem, namely the work theorem discovered by Jarzynski and refined by Crooks [40, 41]. To obtain this result, we just need to rewrite the detailed fluctuation theorem (15) in terms of the work rather than the stochastic entropy production:

$$\frac{P(W)}{\tilde{P}(-W)} = e^{\beta(W - \Delta F)}. \tag{19}$$

This result was first derived in the context of a stochastic description by Crooks [41]. It implies the famous Jarzynksi equality [40], the analog of the integral fluctuation theorem, derived originally in a Hamiltonian context very similar to ours:

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta F}. \tag{20}$$

Both the fluctuation theorem and its variant the above work theorem have been discussed, illustrated and verified in a large number of situations. We mention three of them. The first is a real experiment, namely the opening (forward experiment) and closing (backward experiment) of a hairpin RNA [44]. This can be achieved by using optical tweezers and the force versus displacement curves can be recorded. The surface under these curves represents the amount of work. By repeating the experiment many times, a histogram can be constructed. The histogram will depend on the protocol (speed of opening/closing). If the above relation (19) applies, the various histograms of $P(W)$ and $\tilde{P}(-W)$ should cross at the same point, namely $W = \Delta F$. The authors considered three different protocols (slow, medium and fast speed of operation) and it was found that the histograms indeed cross at approximately the same value of the work, a value which moreover is in good agreement with other estimates of the free energy difference between the open and closed hairpins.

As a second example we mention a computer experiment [45] in which an object of arrow-like shape is moved at constant speed between a given initial and final location. In the forward experiment the arrow is moving with its sharp angle pointing forward. In the backward experiment, it moves with its flat face pointing in the direction of motion. For a motion in an ideal gas the probability distributions for the work can be calculated analytically, and the Crooks' relation (19) is verified. The same relation is also verified when carrying out molecular dynamics simulations with the arrow moving through a gas

of hard discs. When the density is sufficiently dilute the ideal gas distributions for the work are recovered. Note however that the forward and backward distributions can differ very much in scale and shape. This is for instance the case when the motion of the arrow is supersonic (faster than the thermal speed of the gas particles). The forward work distribution still has a somewhat Gaussian shape, while the backward distribution has several extrema and is concentrated on much higher values of work. Yet the logarithm of the ratio of the two distributions is a linear function of the work, i.e., it is a straight line going furthermore nicely through the origin (corresponding to $\Delta F = 0$ in this case).

We finally turn to a thought experiment not unlike the one used by Einstein in his discovery of stimulated emission. We consider two black bodies at equilibrium at temperatures T_1 and T_2, respectively. At time $t = 0$ we open a window allowing photons of a specific frequency ω_0 to be exchanged. After a time t we close the window again. Let ΔN be the net number of photons that have moved from 1 to 2. After both reservoirs have returned to equilibrium (we assume they are large enough so that their final and initial temperatures are the same), their respective entropy changes are given by $\Delta S_1 = -\Delta N \hbar \omega_0 / T_1$ and $\Delta S_2 = \Delta N \hbar \omega_0 / T_1$, resulting in a total entropy change

$$\Delta S_{\text{tot}} = \Delta N \hbar \omega_0 \left(\frac{1}{T_2} - \frac{1}{T_1} \right). \tag{21}$$

We conclude that the probability distribution for ΔS_{tot} is essentially that for the photon count. At first one is tempted to assume that the photon arrival statistics is Poissonian, with the rate proportional to the density of photons, which is given by Planck's law. With this assumption, the total entropy production (21) does not, however, obey the fluctuation theorem. The point is that the arrival of photons is not independent. Photon bunching was first experimentally observed in the famous Hanbury Twiss experiment [46]. The intuitive explanation is that due to stimulated emission, photons have a certain tendency to bunch and arrive in pairs at a detector. The correct photon count distribution was first calculated from quantum field theory by Glauber [47], and is found to be a negative binomial (for large enough times). Upon inserting this distribution, one easily checks that the fluctuation theorem is indeed satisfied [48].

7. The three faces of the second law

It is revealing to reproduce the formula for the total entropy production (13) for a system, whose dynamics is described by a Markovian stochastic process. This derivation also illustrates that the formula remains valid for any initial probability distribution, and is hence not limited to initial equilibrium distributions (such as the canonical one used above) or to steady state transformations. For simplicity, we assume that the system is described by a set of discrete non-degenerate states m with corresponding energy ϵ_m, with internal entropy of each state equal to zero. The time evolution for the probability distribution p_m in the state m is described by a Markovian master equation:

$$\frac{\mathrm{d}p_m}{\mathrm{d}t} = \Sigma_{m'} W_{m,m'} p_{m'}, \tag{22}$$

where $W_{m,m'}$ ($m \neq m'$) is the probability per unit time for a change in state from m' to m. Note that the diagonal element is negative and equal to minus the sum of the off-diagonal

elements:

$$\Sigma_m W_{m,m'} = 0. \tag{23}$$

This property guarantees conservation (normalization) of total probability. To proceed we need to introduce the basic elements relating the stochastic model to a physical system. We have associated to every state a specific energy. In view of energy conservation, jumps between states require energy exchange. Let us consider the case where this exchange takes place via ideal thermal reservoirs ν at temperatures T_ν. In principle, any reservoir can lead to a transition. Since the reservoirs do not interfere with each other (otherwise we would need to describe this interaction), one has $W_{m',m} = \sum_\nu W_{m',m}^\nu$. Furthermore, when all exchanges are switched off, except with a given reservoir ν, the system has to relax to the corresponding canonical distribution $p_m^{\mathrm{eq},\nu} \sim \exp(-\beta^\nu \epsilon_m)$ and detailed balance has to prevail. This implies:

$$\ln \frac{W_{m,m'}^\nu}{W_{m',m}^\nu} = \ln \frac{p_m^{\mathrm{eq},\nu}}{p_{m'}^{\mathrm{eq},\nu}} = \beta^\nu(\epsilon_m' - \epsilon_m) = \Delta S^\nu \tag{24}$$

where ΔS^ν is the entropy change in the reservoir ν upon absorbing the energy $\epsilon_m' - \epsilon_m$ from the system in its jump from m' to m. Note that these energy exchanges are modeled here as instantaneous jumps, in accordance with the Markovian description of the system. The simplest way to include an ideal frictionless device that exchanges work with the system is to operate with time-dependent transition rates $W_{m,m'}^\nu = W_{m,m'}^\nu(t)$, while preserving the validity of (24). This is for example achieved by the application of an external work producing/absorbing device that shifts the energy levels $\epsilon_m = \epsilon_m(t)$, implying a corresponding time-dependence of the transition rates. In contrast to the heat exchanges, which take place during the jumps, the work is exchanged continuously in time. It is easy to verify that these assumptions lead to proper expressions for work and heat, and are in agreement with overall energy conservation.

Turning to entropy considerations, we consider an experiment running from time $t = 0$ to a final time $t = T$. It is important to stress that an experiment is characterized not only by the way that the transition probabilities are changing in time, but also by the initial probability distribution $p_m(t = 0)$, corresponding to a specific preparation of the system. We also consider the time-reversed experiment, which runs with the time-reversed transition probabilities, and starts with the final probability distribution reached in the forward experiment at time $t = T$, namely $p_m(T)$. We stress that the initial distribution $p_m(t = 0)$ is arbitrary, but the backward experiment has to be run with the specific distribution $p_m(T)$ as initial distribution. A large part of the literature deals with the consideration of steady states, with $p_m(t = 0) = p_m(T) = p_m^{\mathrm{st}}$ and with time-independent rates. This has the 'advantage' that forward and backward experiments are identical. In this case one can drop the reference tilde for the time-reversed experiment so that the fluctuation theorem implies a symmetry property for the distribution $\tilde{P} = P$.

Let us now consider in the forward experiment the stochastic entropy change for a specific realization, in which the system starts at time $t = 0$ from state m_0, performs a number of jumps to $m_1, \ldots$ up to m_N at times $\tau_1, \ldots$ up to τ_N, staying there until time $t = T$. The sum over these jumps will be denoted by a sum over τ. We also need to specify for each of the jumps which reservoir ν is providing the energy, but omit this for simplicity of notation. The total entropy change consists of two contributions. We

need to define the stochastic entropy for the system when being in a particular state, say $m(t)$, at time t. The actual probability distribution in the experiment is $p_m(t)$. The stochastic entropy of the system is defined in accordance with information theory [28], as first proposed by Seifert [43], as $S = -\ln p_{m(t)}(t)$. The resulting entropy change of the system—for the above specified trajectory and the experiment—between final and initial state is thus $\Delta S = \ln p_{m_N}(T) - \ln p_{m_0}(0)$. This term is sometimes called the boundary term. Its omission in earlier derivations of the fluctuation theorem leads to a version that is only valid for asymptotic long times in systems with bounded energy. The entropy change of the surrounding medium is the sum of all the contributions from the successive jumps $\Delta S_m = \Sigma_\tau \Delta S^\nu(\tau)$, cf (24). The main point is now to show that the resulting total entropy production is exactly given by (13), in its path-probability version, namely:

$$\Delta S_{\text{tot}} = \Delta S + \Delta S_m = \ln \frac{\text{Prob(path)}}{\widetilde{\text{Prob}}(\widetilde{\text{path}})}. \tag{25}$$

Note that contrary to the deterministic situation, the paths in (25) are no longer determined by the state of the system at a given time but correspond to the complete history of the variables. To verify the above result, we need to calculate the ratio of path probabilities, for the forward versus backward path. This ratio is in fact easy to evaluate: a path is characterized by its initial state, and by the jumps, their nature (i.e. which heat bath provides the energy) and jump times. Let us first focus on sections of the path where no jump takes place. The probability for not having a jump during the stretches of each of these intervals is the same in the forward and backward experiments, since the system is for the considered paths in the same state, and the transition rates W are identical. Hence these probabilities drop out when calculating the ratio of the path probabilities. The logratio of the probabilities for the starting probabilities in the backward versus forward experiment is clearly $\ln(p_{m_N}(T)/p_{m_0}(0))$ (remember that we specified that the backward experiment starts with the final probability of the forward experiment). This quantity is precisely identical to the change in system entropy. Finally, the logratio of the probabilities for any jump say from m to m' forward and m' to m backward, via reservoir ν, $\ln W^\nu_{m,m'}/W^\nu_{m',m}$, is equal to the entropy change in the corresponding reservoir ΔS^ν, cf (24). The sum of these contributions is equal to the entropy production in all of the reservoirs (which we refer to as the surrounding medium): $\Delta S_m = \Sigma_\tau \Delta S^\nu(\tau)$. We conclude that the logratio of the path probability does reproduce the total stochastic entropy change for the specified trajectory.

The stochastic total entropy production is the logratio of two probabilities, which are related by an involution, namely time-reversal. Hence this stochastic total entropy obeys a detailed and integral fluctuation theorem, cf (17) and (15). The constituting subparts ΔS and ΔS_m, however, are not the logratios of probabilities and do not satisfy a fluctuation theorem. This is unfortunate especially for the case of the medium entropy change ΔS_m, since the latter is experimentally easily accessible. One can however consider the accumulated stochastic entropy production produced over long stretches of times, for example in nonequilibrium steady states. For a system with bounded entropy, one expects that $\Delta S_{\text{tot}} \sim \Delta S_m$ so that the medium stochastic entropy will satisfy a fluctuation theorem, but only so for asymptotically long times. For systems with unbounded energy, even the asymptotic fluctuation theorem has a limited range of validity, cf [49].

We recently discovered that there is, however, another way to split the total stochastic entropy into two contributions, with each part satisfying a separate fluctuation theorem at all times [30] see also [50]:

$$\Delta S_{\text{tot}} = \Delta S_a + \Delta S_{na}. \tag{26}$$

The contributions ΔS_a and ΔS_{na} have a simple and clear physical meaning: they correspond to contributions in the total entropy production associated to the nonequilibrium boundary conditions and those related to the driving. This splitting has been anticipated in special cases, with the adiabatic entropy production being a generalization of the concept of the housekeeping heat and the non-adiabatic entropy production related to system entropy and excess heat [51]. The crux of the derivation is the introduction of another symmetry operator for Markov processes, called the dual operator, which we will denote by a superscript $+$. Loosely speaking it neutralizes the steady state time-symmetry breaking coming from the boundary conditions. To explain this statement, we consider first the steady state regime of a Markov process with a time-independent transition matrix W. This matrix will in general not satisfy detailed balance with respect to the prevailing stationary probability distribution p^{st}, because it describes a nonequilibrium situation with 'conflicting' boundary constraints, for example the presence of various heat baths at different temperatures. The dual W^+ of the transition matrix W is defined by the following relation:

$$W_{m',m} p_m^{\text{st}} = W_{m,m'}^+ p_{m'}^{\text{st}}. \tag{27}$$

One easily verifies that it defines another Markov process with the same stationary probability distribution. More relevant is that this Markov process is statistically speaking the original Markov process but with the time-axis reversed. In other words if one applies the operations of duality and time-reversal in the steady state, they neutralize each other $\tilde{P}^+ = P$. Returning to the above case of a Markov process with time-dependent transition rates, one can perform the dual operation at all times on the transition matrix thus defining another Markov process. In general, however, this operation will not neutralize time-reversal because there can be a remaining symmetry breaking due to the fact that the system is actually not at the steady state. This will be the case when it did not start in the steady state, or when the transition rates are time-dependent. We refer to this other way of generating nonequilibrium states as the driving, and the corresponding entropy production is called the non-adiabatic contribution. The difference between the total and the non-adiabatic contributions will be called the adiabatic one.

Proceeding in a more mathematical way, one can reach the same separation of the total entropy production by noting that with the combination of the dual operation superscript $+$ and the time-reversal superscript tilde, one can construct three symmetry operations, which are all involutions. Consider an experiment, characterized by Markovian stochastic dynamics, with specified initial probability distribution and a specified schedule for changing the transition rates. In this setting, we consider the entropy production associated to a realization or path. For simplicity of notation we will write $P = \text{Prob}(\text{path})$ for the probability density for this path. We also consider three other experiments, one with time-reversed transition rates starting with the final distribution of the forward experiment, one in which in addition the Markovian process is switched to the dual, and one in which the forward but dual dynamics is considered. The corresponding probabilities

for the path under consideration are denoted as $\tilde{P}$ and $\tilde{P}^+$ and P^+. The tilde operation also refers to a time-reversal of the path. We can thus introduce the following three entropy differences:

$$\Delta S_{\text{tot}} = \ln \frac{P}{\tilde{P}} \tag{28}$$

$$\Delta S_{na} = \ln \frac{P}{\tilde{P}^+} \tag{29}$$

$$\Delta S_a = \ln \frac{P}{P^+}. \tag{30}$$

Each of these quantities is the logratio of probabilities, being related to each other via an involution. Hence they obey, each separately, an integral and detailed fluctuation theorem:

$$\frac{P(\Delta S_{\text{tot}})}{\tilde{P}(-\Delta S_{\text{tot}})} = e^{\Delta S_{\text{tot}}} \tag{31}$$

$$\frac{P(\Delta S_{na})}{\tilde{P}^+(-\Delta S_{na})} = e^{\Delta S_{na}}, \qquad \frac{P(\Delta S_a)}{P^+(-\Delta S_a)} = e^{\Delta S_a}. \tag{32}$$

A short calculation furthermore allows us to verify that:

$$\Delta S_{\text{tot}} = \Delta S_{na} + \Delta S_a. \tag{33}$$

We thus reach the remarkable conclusion that we have split the second law of thermodynamics into two at its most fundamental level! We stress that we have done much more than identifying positive contributions to the total entropy production $\langle \Delta S_{\text{tot}} \rangle \geq 0$, $\langle \Delta S_{na} \rangle \geq 0$, and $\langle \Delta S_a \rangle \geq 0$: we are dealing here with the trajectory dependent stochastic quantities ΔS_{tot}, ΔS_a and ΔS_{na}. They each obey a detailed (and integral) fluctuation theorem.

8. Discussion

The second law was discovered and formulated about 150 years ago. While thermodynamics has played a crucial role in many important advances in physics, including the discovery of quantum mechanics, and while it is a field with an almost unlimited range of application, it is nowadays often viewed as an old if not old-fashioned field of research. Recent developments however show that this is not at all the case. One can formulate stochastic thermodynamics, consistent with a microscopic analysis. The second law inequality is thus replaced by an equality for a probability distribution of the stochastic entropy change. This equality derives from basic symmetries of the underlying physical laws. Carnot efficiency and Onsager symmetry are complemented with new laws such as the universal value of the coefficients in the expansion of efficiency at maximum power, and the symmetry for large deviation functions of currents. Last but not least, the second law appears to have new offspring under the form of two constitutive contributions, the adiabatic and non-adiabatic entropy productions, which each satisfy a

fluctuation theorem of their own. The importance of this result is underscored by the fact that it follows from the application of two symmetry operations: time-reversal of the driving and a duality operation of the transition matrix (which amounts to a kind of time-reversal of boundary conditions). It remains to be seen whether the implications of this duality are as far reaching as those of the second law itself, and whether there are alternative or supplementary divisions possible of the total entropy production into constitutive parts that satisfy fluctuation theorems.

Acknowledgment

References

[1] http://www.lhup.edu/~dsimanek/museum/people/people.htm
[2] Escher M C, http://en.wikipedia.org/wiki/Waterfall
[3] Maxwell J C, 1871 *The Theory of Heat* (London)
    Leff H S and Rex A F (ed), 2003 *Maxwell's Demon* (Bristol: Adam Hilger)
    Leff H S and Rex A F (ed), 2002 *Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing* (Boca Raton, FL: CRC Press)
[4] Szilard L, 1929 *Phys. Z.* **53** 840
    Brillouin L, 1951 *J. Appl. Phys.* **22** 334
    Landauer R, 1961 *IBM J. Res. Dev.* **5** 183
[5] Van den Broeck C, *Entropy and the arrow of time*, 2009 *The Complexity of Dynamical Systems: A Multidisciplinary Perspective* ed Johan Dubbeldam, Kirk Green and Daan Lenstra (New York: Wiley)
[6] von Smoluchowski M, 1912 *Phys. Z.* **13** 1069
[7] Feynman R P, Leighton R B and Sands M, 1963 *The Feynman Lectures in Physics* vol 1 (Reading, MA: Addison-Wesley)
[8] Onsager L, 1931 *Phys. Rev.* **37** 405
    Onsager L, 1931 *Phys. Rev.* **38** 2265
[9] Clausius R, 1850 *Poggendorff's Ann. Phys. Chem.* **79** 368, 500
[10] Prigogine I, 1947 *Etude Thermodynamique des Phenomènes Irreversibles* (Liège: Desoer)
    de Groot S R and Mazur P, 1984 *Non-Equilibrium Thermodynamics* (New York: Dover)
[11] Reimann P, 2002 *Phys. Rep.* **361** 57
    Astumian R D and Hanggi P, 2002 *Phys. Today* **55** 33
    Linke H (ed), *Ratchets: basics, experiments and applications*, 2002 *Appl. Phys.* A **75** 167
[12] Parrondo J M R and Espagnol P, 1996 *Am. J. Phys.* **64** 1125
    Sekimoto K, 1997 *J. Phys. Soc. Japan* **66** 1234
[13] Van den Broeck C, 2007 *Adv. Chem. Phys.* **135** 189
[14] Gruber Ch and Piasecki J, 1999 *Physica* A **268** 412
    Kestemont E, Van den Broeck C and Malek Mansour M, 2000 *Europhys. Lett.* **49** 143
[15] Van den Broeck C, Kawai R and Meurs P, 2004 *Phys. Rev. Lett.* **93** 090601
    Meurs P, Van den Broeck C and Garcia A, 2004 *Phys. Rev.* E **70** 1
    Van den Broeck C, Meurs P and Kawai R, 2005 *New J. Phys.* **7** 1
[16] Van den Broeck C and Kawai R, 2006 *Phys. Rev. Lett.* **96** 210601
    van den Broek M and Van den Broeck C, 2008 *Phys. Rev. Lett.* **100** 130601
[17] http://blogs.nature.com/news/blog/2006/05/mini-fridge-exploits-brownian.html
[18] Yvon J, *The Saclay reactor: two years experience on heat transfer by means of a compressed gas*, 1955 *Proc. Int. Conf. on Peaceful Uses of Atomic Energy (Geneva)* p 387
    Chambadal P, 1957 *Les Centrales Nuclaires* (Paris: Armand Colin)
    Novikov I I, 1958 *J. Nucl. Energy* II **7** 125
[19] Curzon F L and Ahlborn B, 1975 *Am. J. Phys.* **43** 22
[20] Band Y B, Kafri O and Salamon P, 1982 *J. Appl. Phys.* **53** 8
    De Vos A, 1985 *Am. J. Phys.* **53** 570

Berry R S, Kazakov V A, Sieniutycz S, Szwast Z and Tsvilin A M, 2000 *Thermodynamic Optimization of Finite-Time Processes* (Chichester: Wiley)

Salamon P, Nulton J D, Siragusa G, Andersen T R and Limon A, 2001 *Energy* **26** 307

Wu C and Sun F, 1999 *J. Non-Equilib. Thermodyn.* **24** 327

Hoffmann K H, Burzler J, Fischer A, Schaller M and Schubert S, 2003 *J. Non-Equilib. Thermodyn.* **28** 233

Durmayaz A, Sogut O S, Sahin B and Yavuz H, 2004 *Prog. Energy Comb. Sci.* **30** 175

[21] Van den Broeck C, 2005 *Phys. Rev. Lett.* **95** 190162

[22] Esposito M, Lindenberg K and Van den Broeck C, 2009 *Phys. Rev. Lett.* **102** 130602

[23] Schmiedl T and Seifert U, 2007 *Phys. Rev. Lett.* **98** 108301

de Cisneros B J and Hernandez A C, 2007 *Phys. Rev. Lett.* **98** 130602

Schmiedl T and Seifert U, 2008 *Europhys. Lett.* **81** 20003

Then H and Engel A, 2008 *Phys. Rev.* E **77** 041105

Izumida Y and Okuda K, 2008 *Europhys. Lett.* **83** 60003

Tu Z C, 2008 *J. Phys. A: Math. Theor.* **41** 312003

Gomez-Marin A, Schmiedl T and Seifert U, 2008 *J. Chem. Phys.* **129** 024114

Schmiedl T and Seifert U, 2008 *Europhys. Lett.* **83** 30005

Izumida Y and Okuda K, 2009 *Phys. Rev.* E **80** 021121

Izumida Y and Okuda K, 2009 *Prog. Theor. Phys. Suppl.* **178** 163

Rutten B, Esposito M and Cleuren B, 2009 *Phys. Rev.* B **80** 235122

Esposito M, Kawai R, Lindenberg K and Van den Broeck C, 2010 *Europhys. Lett.* **89** 20003

Esposito M, Kawai R, Lindenberg K and Van den Broeck C, 2010 *Phys. Rev.* E **81** 041106

Zhou Y and Segal D, 2010 *Phys. Rev.* E **82** 011120

Gaveau B, Moreau M and Schulman L S, 2010 *Phys. Rev. Lett.* **105** 060601

[24] Popescu S, Short A J and Winter A, 2006 *Nat. Phys.* **2** 758

Goldstein S, Lebowitz J L, Tumulka R and Zanghi N, 2006 *Phys. Rev. Lett.* **96** 050403

Porporato A, Rigby J R and Daly E, 2007 *Phys. Rev. Lett.* **98** 094101

[25] Esposito M, Lindenberg K and Van den Broeck C, 2010 *New J. Phys.* **12** 013013

[26] Reimann P, 2008 *Phys. Rev. Lett.* **101** 190403

[27] Kawai R, Parrondo J M R and Van den Broeck C, 2007 *Phys. Rev. Lett.* **98** 080602

Parrondo J M R, Van den Broeck C and Kawai R, 2009 *New J. Phys.* **11** 073008

[28] Cover T M and Thomas J A, 2006 *Elements of Information Theory* 2nd edn (Hoboken, NJ: Wiley)

[29] Mackey M C, 1989 *Rev. Mod. Phys.* **61** 981

Qian H, 2001 *Phys. Rev.* E **63** 042103

Jiang D Q, Qian M and Zhang F-X, 2003 *J. Math. Phys.* **44** 4176

Maes C and Netocyny K, 2003 *J. Stat. Phys.* **110** 269

Gaspard P, 2004 *J. Stat. Phys.* **117** 599

Costa M, Goldberg A and Peng C-K, 2005 *Phys. Rev. Lett.* **95** 198102

[30] Esposito M and Van den Broeck C, 2010 *Phys. Rev. Lett.* **104** 090601

Esposito M and Van den Broeck C, 2010 *Phys. Rev.* E **82** 011143

Van den Broeck C and Esposito M, 2010 *Phys. Rev.* E **82** 011144

[31] Landauer R, 1961 *Adv. Chem. Phys.* **5** 183

Bennett C H, *Logical reversibility of computation*, 1973 *IBM J. Res. Dev.* **17** 525

Hopfield J J, 1974 *Proc. Nat. Acad. Sci.* **71** 4135

[32] Piechocinska B, 2000 *Phys. Rev.* A **61** 062314

Sagawa T and Ueda M, 2009 *Phys. Rev. Lett.* **102** 250602

Sagawa T and Ueda M, 2010 *Phys. Rev. Lett.* **104** 020602

[33] Klein M J and Meijer P H E, 1954 *Phys. Rev.* **96** 250

[34] Schnakenberg J, 1976 *Rev. Mod. Phys.* **48** 571

[35] Luo J L, Van den Broeck C and Nicolis G, 1984 *Z. Phys.* B **56** 165

[36] Van den Broeck C, *Stochastic thermodynamics*, 1986 *Selforganization by Nonlinear Irreversible Processes* ed W Ebeling and H Ulbricht (Berlin: Springer) pp 57–61

[37] Bochkov G N and Kuzovlev Y E, 1981 *Physica* A **106** 443

Bochkov G N and Kuzovlev Y E, 1981 *Physica* A **106** 480

Evans D, Cohen E G D and Morris G P, 1993 *Phys. Rev. Lett.* **71** 2401

Evans D J and Searles D J, 1994 *Phys. Rev.* E **50** 1645

Gallavotti G and Cohen E G D, 1995 *Phys. Rev. Lett.* **74** 2694

Gallavotti G and Cohen E G D, 1995 *J. Stat. Phys.* **80** 931

Kurchan J, 1998 *J. Phys. A: Math. Gen.* **31** 3719

Sekimoto K, 1998 *Prog. Theor. Phys. Suppl.* **130** 17

Lebowitz L and Spohn H, 1999 *J. Stat. Phys.* **95** 333

Maes C, 1999 *J. Stat. Phys.* **95** 367

Liphardt J, Dumont S, Smith S B, Tinoco I Jr and Bustamante C, 2002 *Science* **296** 1832

Evans D J and Searles D J, 2002 *Adv. Phys.* **51** 1529

Ritort F, 2003 *Semin. Poincare* **2** 195

Imperato A and Peliti L, 2006 *Phys. Rev.* E **74** 026106

Tietz C, Schuler S, Speck T, Seifert U and Wrachtrup J, 2006 *Phys. Rev. Lett.* **97** 050602

Derrida B, Gaspard P and Van den Broeck C (ed), *Work, dissipation, and fluctuations in nonequilibrium physics*, 2007 *C.R. Physique* **8** (special issue) 483

Speck T, Blickle V, Bechinger C and Seifert U, 2007 *Europhys. Lett.* **79** 30002

Maes C and Netocny K, 2007 *J. Math. Phys.* **48** 053306

Andrieux D and Gaspard P, 2007 *J. Stat. Phys.* **127** 107

Andrieux D, Gaspard P, Ciliberto S, Garnier N, Joubaud S and Petrosyan A, 2007 *Phys. Rev. Lett.* **98** 150601

Seifert U, 2008 *Eur. Phys. J.* B **64** 423

Jarzynski C, 2008 *Eur. Phys. J.* B **64** 331

Esposito M, Harbola U and Mukamel S, 2009 *Rev. Mod. Phys.* **81** 1665

[38] Gomez-Marin A, Parrondo J M R and Van den Broeck C, 2008 *Europhys. Lett.* **82** 50002

[39] Gomez-Marin A, Parrondo J M R and Van den Broeck C, 2008 *Phys. Rev.* E **78** 011107

[40] Jarzynski C, 1997 *Phys. Rev. Lett.* **78** 2690

Jarzynski C, 1997 *Phys. Rev.* E **56** 5018

[41] Crooks G E, 1998 *J. Stat. Phys.* **90** 1481

Crooks G E, 1999 *Phys. Rev.* E **60** 2721

Crooks G E, 2000 *Phys. Rev.* E **61** 2361

[42] Jarzynski C, 2006 *Phys. Rev.* E **73** 046105

[43] Seifert U, 2005 *Phys. Rev. Lett.* **95** 040602

[44] Collin D, Ritort F, Jarzynski C, Smith S B, Tinoco I Jr and Bustamante C, 2005 *Nature* **437** 231

[45] Cleuren B, Van den Broeck C and Kawai R, 2006 *Phys. Rev. Lett.* **96** 050601

[46] Hanbury Brown R and Twiss R Q, 1956 *Nature* **178** 1046

[47] Glauber R J, 1962 *Phys. Rev.* **10** 84

[48] Cleuren B and Van den Broeck C, 2007 *Europhys. Lett.* **79** 30001

[49] Farago J, 2002 *J. Stat. Phys.* **107** 781

Farago J, 2004 *Physica* A **331** 69

van Zon R and Cohen E G D, 2003 *Phys. Rev. Lett.* **91** 110601

Visco P, 2006 *J. Stat. Mech.* P06006

Baiesi M, Jacobs T, Maes C and Skantzos N S, 2006 *Phys. Rev.* E **74** 021111

[50] Esposito M, Harbola U and Mukamel S, 2007 *Phys. Rev.* E **76** 031132

[51] Oono Y and Paniconi M, 1998 *Prog. Theor. Phys. Suppl.* **130** 29

Hatano T and Sasa S I, 2001 *Phys. Rev. Lett.* **86** 3463

Speck T and Seifert U, 2006 *Europhys. Lett.* **74** 391

Chernyak V Y, Chertkov M and Jarzynski C, 2006 *J. Stat. Mech.* P08001

Harris R J and Schutz G M, 2007 *J. Stat. Mech.* P07020