

Fluctuation Theorems of Work and Entropy in Hamiltonian Systems

Sourabh Lahiri and Arun M Jayannavar

Fluctuation theorems are a group of exact relations that remain valid irrespective of how far the system has been driven away from equilibrium. Other than having practical applications, like determination of equilibrium free energy change from nonequilibrium processes, they help in our understanding of the second law and the emergence of irreversibility from time-reversible equations of motion at microscopic level. A vast number of such theorems have been proposed in literature, ranging from Hamiltonian to stochastic systems, from systems in steady state to those in transient regime, and for both open and closed quantum systems. In this article, we discuss about a few such relations, when the system evolves under Hamiltonian dynamics.

Introduction

Fluctuation theorems (FTs) are a group of exact relations that remain valid even when the system of interest is driven far away from equilibrium. For driven systems, fluctuations in heat, work, and entropy are not mere background noise, but satisfy strong constraints on the probability distributions of these fluctuating quantities. These relations are of fundamental importance in nonequilibrium statistical mechanics. Intensive research has been done in this direction in order to find such relations for thermodynamic quantities like work, heat or entropy changes.

They have resulted in conceptual understanding of how irreversibility emerges from reversible dynamics and of the second law of thermodynamics [1].

These theorems lead to the fact that the second law holds for aver-



Sourabh Lahiri is an Assistant Professor at Birla Institute of Technology, Mesra (Ranchi). He works in the field of nonequilibrium statistical mechanics.



Arun M Jayannavar is a senior scientist at Institute of Physics, Bhubaneswar and is interested in general condensed matter physics.

Keywords

Fluctuation theorem, Hamiltonian dynamics, feedback.



Equilibrium thermodynamics, that mainly involves equilibrium states and quasi-static reversible processes, is a very well-established subject. Many more results can be derived in the near-equilibrium regime where external driving force is weak. FTs, however, are valid for arbitrary strength of the force.

FTs rely on two main physical inputs. One is that in the absence of any driving, a system that is out of equilibrium will relax to its state of thermal equilibrium. The other is the principle of microscopic reversibility, described in (5).

Stochastic phase space trajectories imply that if the same experiment is performed repeatedly, the path traced out by the system in its phase space is in general, different in different experiments.

age quantities. However, there are atypical transient trajectories in phase space which violate the second law.

Two fundamental ingredients play a decisive role in the foundations of FT – the principle of micro reversibility, and the fact that thermal equilibrium is described by Gibbs canonical ensemble. Moreover, some of these relations have been found useful for practical applications like determining the change in equilibrium free energy in an irreversible process. Numerous FTs have been put forward in the last two decades [2]. Some of them are valid when the system is in a nonequilibrium steady state, while the others are valid in the transient regime. Fluctuation theorems have been proposed for Hamiltonian as well as stochastic dynamics, and quantum systems(both closed and open ones). Some of them have been tested experimentally [3].

To understand the significance of FTs, we first need to appreciate the fact that the phase space trajectories of small systems (at meso or nano scales) are stochastic trajectories because thermal fluctuations play a dominant role in their dynamics. In particular, if the evolution is Hamiltonian, then the stochasticity comes from the fact that the initial state of the system is sampled from some distribution that is not simply at delta-function. As a result, thermodynamic quantities like work, heat or entropy change are also stochastic and follow distributions instead of having a single value. The FTs provide stringent conditions on the symmetries of these distributions. Furthermore, the inequalities encountered in the second law can be readily obtained as corollaries from the FT relations. Thus, these theorems are stronger relations than the second law. We now move on to describe the FTs for work.

1. Fluctuation Theorems for Work

Initially, FTs were proven for the entropy production in steady state by Evans *et. al.* in [4], and was theoretically shown later in [5]. It was derived for Langevin dynamics in [6] and for general Markov processes in [7].

The work fluctuation theorems include the Jarzynski Equality (JE)



[8] and the Crooks work FT (CFT) [9]. These are also termed as nonequilibrium work theorems in literature [2]. They show that the equilibrium free energy difference between the final and initial values of the external parameter can be computed by measuring the nonequilibrium work done, instead of having to perform the experiment quasistatically. This has a lot of practical significance.

Let $z \equiv (\mathbf{q}, \mathbf{p})$ denote the position and momenta of the system degrees of freedom. Initially, the system is prepared at equilibrium with a heat bath at temperature T , and at time $t = 0$ the bath is disconnected and an external parameter $\lambda(t)$ that acts on the system is switched on. Let $H_S(z, t)$ be the system Hamiltonian that depends explicitly on time, so that the initial state distribution of the system is given by,

$$\rho_{\lambda_0}(z_0) = \frac{e^{-\beta H_S(z_0)}}{Z_0}, \quad (1)$$

Z_0 being the initial partition function defined by

$$Z = \int dz_0 e^{-\beta H_S(z_0)}. \quad (2)$$

At the end of the process, at time $t = \tau$, the parameter is turned off. The probability of a trajectory $z(t)$ in phase space is given by $P[z(t)]$.

Correspondingly, a time-reversed process is defined which is described by the parameter $\lambda(\tau - t)$ acting on the system that is initially at equilibrium with the final value of the forward protocol. The reverse trajectory corresponding to the forward trajectory $z(t)$ is defined as $\tilde{z}(t) = z^*(\tau - t)$. Here, the asterisk implies that all the velocity variables switch signs. A typical forward and its corresponding reverse trajectory are depicted in *Figure 1*.

The initial state distribution of the system for the reverse process is then given by,

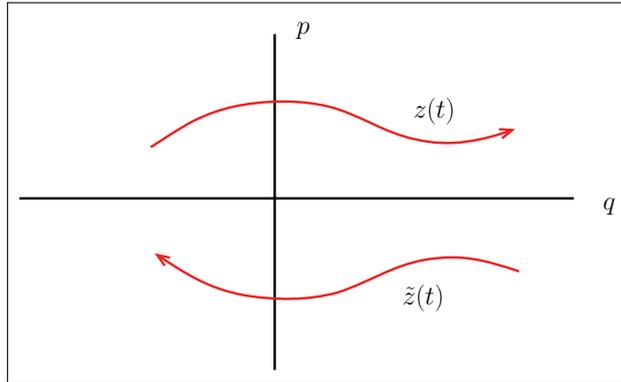
$$\tilde{\rho}_{\lambda_\tau}(z_\tau) = \frac{e^{-\beta H_S(z_\tau^*)}}{Z_\tau} \quad (3)$$

The probability of a reverse trajectory in the reverse process is given by $\tilde{P}[\tilde{z}(t)]$.

The external parameter is a *deterministic* function of time and drives the system out of equilibrium. For instance, if gas contained in a cylinder is compressed with a piston, then the volume $V(t)$ available to the gas would be the parameter $\lambda(t)$.



Figure 1. A typical forward and its corresponding reverse trajectory.



If we take the ratio between $P[z(t)]$ and $\tilde{P}[\tilde{z}(t)]$, then we get,

$$\frac{P[z(t)]}{\tilde{P}[\tilde{z}(t)]} = \frac{\rho_{\lambda_0}(z_0)}{\tilde{\rho}_{\lambda_\tau}(z_\tau)} = e^{\beta[H_S(z_\tau) - H_S(z_0) - \Delta F]}, \quad (4)$$

The one-to-one correspondence in Hamiltonian dynamics implies that once the initial state is chosen, the entire trajectory is determined.

where we have used the fact that there is a one-to-one correspondence between the entire trajectory and its initial point in a deterministic evolution. We have further assumed that the equilibrium Hamiltonian is symmetric under time reversal: $H_S(z_\tau^*) = H_S(z_\tau)$. ΔF is the free energy change of the system: $\Delta F = k_B T \ln \frac{Z_0}{Z_\tau}$. Since the total work W done on the system must be equal to the change in the Hamiltonian: $W = H_S(z_\tau) - H_S(z_0)$, we can write

$$\frac{P[z(t)]}{\tilde{P}[\tilde{z}(t)]} = e^{\beta(W - \Delta F)}. \quad (5)$$

This is the condition of microscopic reversibility. We remark that the same relation holds even for stochastic dynamics, although the derivation of the condition of microscopic reversibility is not so simple in that case.

Now, we will prove two FTs: one is called the Crooks FT, and the other is known as Jarzynski equality. Crooks FT directly follows



from (5):

$$\begin{aligned}
 P(W) &= \int \mathcal{D}[z(t)] \delta(W - W[z(t)]) P[z(t)] \\
 &= \int \mathcal{D}[z(t)] \delta(W - W[z(t)]) \tilde{P}[\tilde{z}(t)] e^{\beta(W - \Delta F)} \\
 &= e^{\beta(W - \Delta F)} \int \mathcal{D}[z(t)] \delta(W + W[\tilde{z}(t)]) \tilde{P}[\tilde{z}(t)] \\
 &= e^{\beta(W - \Delta F)} \tilde{P}(-W). \tag{6}
 \end{aligned}$$

This is the Crooks FT for work. Here, $\mathcal{D}[z(t)]$ is a functional integral, which in discrete time would correspond to $dz_0 dz_1 \cdots dz_n$, the subscripts representing time. In the third line, we have used the fact that $W[z(t)] = -W[\tilde{z}(t)]$, and the delta-function allows us to take $e^{\beta W - \Delta F}$ out of the integral. Now, if we multiply both sides of the above relation by $e^{-\beta(W - \Delta F)}$ and integrate over W , then using the normalization condition for $\tilde{P}(-W)$, we obtain,

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta F}. \tag{7}$$

Thus, from (6) and (7), we find that the change in the equilibrium free energy of the system can be computed from nonequilibrium work measurements, which removes the necessity of using a quasi-static process for obtaining ΔF .

We will now show that JE leads to an analogue of the maximum Work theorem, an alternative statement for the second law. First, we prove the relation $\langle e^r \rangle \geq e^{\langle r \rangle}$ for some variable r , using the identity $e^r \geq 1 + r$ for all values of r , the equality being satisfied only at $r = 0$. The proof goes as follows [10]:

$$\begin{aligned}
 \langle e^r \rangle &= \langle e^{r + \langle r \rangle - \langle r \rangle} \rangle = e^{\langle r \rangle} \langle e^{r - \langle r \rangle} \rangle \\
 &\geq e^{\langle r \rangle} \langle 1 + r - \langle r \rangle \rangle = e^{\langle r \rangle}. \tag{8}
 \end{aligned}$$

This is the so-called Jensen's inequality for exponential functions. Using this in (7), we find that,

The probability density function $P_y(g(x) = y)$ is defined as $P_y(y) = \int_{-\infty}^{\infty} P_x(x) \delta(y - g(x)) dx$. The extension, when x is replaced by the full trajectory, is easy to see.

In an isothermal reversible process, the mean work done on the system equals the difference in Helmholtz free energies at initial and final times: $\langle W \rangle = \Delta F$

The fluctuation theorems for work and entropy lead to the second law inequalities as corollaries.



$$e^{-\beta\Delta F} \geq e^{-\beta\langle W \rangle} \Rightarrow \langle W \rangle \geq \Delta F. \quad (9)$$

Multiplying both sides by -1 , we have $-\langle W \rangle \leq -\Delta F$, which is an analogue of the statement of the maximum work theorem: *For all thermodynamic processes between some initial and final states, the delivery of work is a maximum for a reversible process.* This maximum work extracted is given by the free energy difference $-\Delta F$.

2. Experimental Tests of the Work Fluctuation Theorems

The CFT as well as JE have been tested experimentally on various systems such as RNA hairpin, dragged colloidal particle, torsion pendulum, electric circuits, etc., [11–14]. We describe here one of them, namely the experiments performed on RNA hairpin [3], in order to provide a visualization of the applications of these theorems. Note, however, that the experiments deal mostly with stochastic dynamics, since the system is not isolated from the reservoir. However, as mentioned before, the FTs have been derived for stochastic dynamics as well. In this sense, these theorems are very general.

The experimental setup is as follows.

The two ends of an RNA hairpin are connected to two beads, one of whose position has been fixed using a strong laser trap. The other bead is moved by an actuator, according to a protocol $\lambda(t)$, which for simplicity was chosen to be a linear function of time. Initially, the RNA hairpin is at equilibrium in the folded state. The experiment was repeated a large number of times, and ΔF was computed using (7). The computed value agreed very well with the independent estimate using a quasistatic process.

Verification of CFT was carried out by generating an ensemble of realizations of both the forward and the reverse process [14]. In the present context, the reverse process consists of reducing the distance between the two ends of the RNA molecule, so that it is

The forward process consists of unfolding an RNA molecule, and the reverse process is its refolding. The force vs. extension curves shows hysteresis, from which W can be calculated.



a refolding process, as opposed to the unfolding or forward process. Of course, the initial distribution of the reverse process has to be the thermal one, corresponding to the stretched RNA hairpin with the end-to-end distance being the same as the one reached at the end of the forward process. Distributions of work were constructed in the forward and the reverse processes, and $P(W)$ and $\tilde{P}(-W)$ were plotted together. As per (6), when $P(W) = \tilde{P}(-W)$, then $W = \Delta F$. Thus, the intersection of the two distributions should provide the value of the change in free energy. Such distributions were constructed for different speeds of the driving protocol, keeping the initial and final end-to-end distance of the RNA hairpin the same. It was observed that for all pairs of distributions $P(W)$ and $\tilde{P}(-W)$, the intersection points were placed on the same vertical line, which is consistent with the fact that the free energy difference must be the same in all the cases, since the initial and final values of the protocols were same.

3. Fluctuation Theorems for Change in Total Entropy

We now discuss the FTs for the entropy change of an isolated system. These theorems come in two forms – the detailed FT which is valid only for transitions between steady states, and an integral FT that is valid for all systems and leads to a statement similar to the second law of thermodynamics [15].

Let us consider a system that is prepared with an initial distribution $p(z_0, 0)$ which need not be a thermal distribution. It is then evolved under a protocol $\lambda(t)$ up to time $t = \tau$, and its final distribution is given by $p(z_\tau, \tau)$. The system’s entropy at any point in time is defined as,

$$s(t) = -\ln p(z_t, t), \tag{10}$$

so that over an ensemble, the average entropy is given by the Gibbs form:

$$\langle s(t) \rangle = -\int dz_t p(z_t, t) \ln p(z_t, t). \tag{11}$$

Any reasonable definition of a fluctuating thermodynamic quantity should, *on averaging*, agree with the standard definition used in macroscopic thermodynamics.



The total change in entropy in the process is then given by the difference between the system's entropy at the final and the initial times as,

$$\Delta s(t) = \ln \frac{p(z_0, 0)}{p(z_\tau, \tau)}. \quad (12)$$

Now, the probability of $\Delta s(t)$ is given by,

$$\begin{aligned} P(\Delta s) &= \int dz_0 dz_\tau \delta \left[\Delta s - \ln \frac{p(z_0, 0)}{p(z_\tau, \tau)} \right] P[z_0, z_\tau] \\ &= \int dz_0 dz_\tau \delta \left[\Delta s - \ln \frac{p(z_0, 0)}{p(z_\tau, \tau)} \right] P(z_\tau | z_0) p(z_0, 0) \\ &= \int dz_0 dz_\tau \delta \left[\Delta s - \ln \frac{p(z_0, 0)}{p(z_\tau, \tau)} \right] P(z_\tau | z_0) p(z_\tau, \tau) \frac{p(z_0, 0)}{p(z_\tau, \tau)} \\ &= e^{\Delta s} \int dz_0 dz_\tau \delta \left[\Delta s - \ln \frac{p(z_0, 0)}{p(z_\tau, \tau)} \right] P(\tilde{z}_0 | \tilde{z}_\tau) \tilde{p}(\tilde{z}_\tau, \tau) \\ &= e^{\Delta s} \int dz_0 dz_\tau \delta \left[\Delta s + \ln \frac{\tilde{p}(\tilde{z}_\tau, \tau)}{\tilde{p}(\tilde{z}_0, 0)} \right] P(\tilde{z}_0 | \tilde{z}_\tau) \tilde{p}(\tilde{z}_\tau, \tau) \\ &= \tilde{P}(-\Delta s) e^{\Delta s}, \end{aligned} \quad (13)$$

where we have used the fact that under deterministic evolution, $P(z_\tau | z_0) = P(\tilde{z}_0 | \tilde{z}_\tau)$, and we have assumed that the initial and the final distributions are symmetric under time-reversal: $p(z_0, 0) = \tilde{p}(\tilde{z}_0, 0)$, and $p(z_\tau, \tau) = \tilde{p}(\tilde{z}_\tau, \tau)$. This is the detailed fluctuation theorem for entropy change, and the corresponding integral FT is obtained by simply multiplying both sides by $e^{-\Delta s}$ and integrating over Δs :

$$\langle e^{-\Delta s} \rangle = 1. \quad (14)$$

The RHS is unity due to normalization of $\tilde{P}(-\Delta s)$. Once again, just as in the case of JE, we can apply the Jensen's inequality to the above theorem, which would give the statement of second law: $\langle \Delta s \rangle \geq 0$, i.e., *the average entropy of an isolated system*



never decreases with time. The FT for change in total entropy have been experimentally studied in [16, 17].

Remark: With the definition (11), it can be shown that the entropy $\langle s(t) \rangle$ is a constant, independent of time: $\langle \Delta s \rangle = 0$, so that the inequality in the second law statement becomes inconsequential. However, the change in entropy along individual realizations can be non-zero, which is why (14) is a non-trivial result. Further, the same relation is true even for a stochastic system, where the system and the heat reservoir together forms an isolated system. Defining the system entropy and bath entropy changes separately and adding them up gives the total change in entropy of the full isolated system. In this case, the second law statement becomes $\langle \Delta s_{sys} \rangle + \langle \Delta s_{bath} \rangle \geq 0$, where the average total entropy will increase with time, unless the process is an equilibrium one. Note that this way of defining total entropy is different from simply taking the negative logarithm of the state distribution of the combined system (consisting of the system and the heat bath).

4. Work Fluctuation Theorems in Quantum Systems

Work fluctuation theorems have been extended to quantum systems [18, 19]. They have been proven for both open and closed quantum systems [20]. The fact that work can only be defined (see below) by performing two-point measurements on the Hamiltonian of the system was clarified in [21]. Here, we provide the derivation for a system which is initially at equilibrium with a heat bath, but at time $t = 0$ it has been disconnected from the bath and allowed to evolve unitarily under the action of an external time-dependent protocol $\lambda(t)$, just as in the case of a classical isolated system.

Let us consider an isolated quantum system undergoing unitary evolution under a time-dependent Hamiltonian $\hat{H}(t)$. The initial state of the system has been sampled from canonical distribution, and thereafter, the system and, the heat bath have been discon-

Since we know only the classical probabilities of the system being in different energy states, it is an example of a *mixed state* (as opposed to a pure state), and is formulated in terms of density operators.



In a unitary evolution, just like in classical Hamiltonian evolution, the final state is determined, given an initial state. This unique mapping also allows us to write the relation $P(m|n) = \tilde{P}(n|m)$.

nected. The initial density operator is thus given by,

$$\hat{\rho}(\lambda_0) = \frac{e^{-\beta\hat{H}(0)}}{Z_0}, \tag{15}$$

where $Z = \text{Tr } e^{-\beta\hat{H}}$. At time $t = 0$, a projective measurement is performed on the Hamiltonian. Let the state of the system collapse to the eigenstate $|n\rangle$ with an eigenvalue E_n . Then, the system is allowed to evolve unitarily till time $t = \tau$ when another projective measurement is performed. Let the final eigenstate be $|m\rangle$ and the corresponding eigenvalue be E_m . The work done on the system is simply equal to the change in the energy of the system

$$W = E_m - E_n. \tag{16}$$

The density operator at the beginning of the reverse process is given by,

$$\hat{\rho}(\lambda_\tau) = \frac{e^{-\beta\hat{H}(\tau)}}{Z_\tau}.$$

The probability of obtaining the eigenvalue E_n at time $t = 0$ is given by $p_0^{eq}(n) = \text{Tr} [\hat{\rho}(\lambda_0)|n\rangle\langle n|] = e^{-\beta E_n}/Z$. Similarly, the probability of obtaining eigenvalue E_m at the beginning of the reverse process is given by $p_\tau^{eq}(m) = \text{Tr} [\hat{\rho}(\lambda_\tau)|m\rangle\langle m|] = e^{-\beta E_m}/Z_\tau$. The probability of work is then given by,

$$\begin{aligned} P(W) &= \sum_{mn} \delta[W - (E_m - E_n)] P(m|n) p_0^{eq}(n) \\ &= \sum_{mn} \delta[W - (E_m - E_n)] |\langle m|U_\lambda(\tau, 0)|n\rangle|^2 \frac{e^{-\beta E_n}}{Z_0} \\ &= \sum_{mn} \delta[W - (E_m - E_n)] |\langle n|U_{\tilde{\lambda}}(0, \tau)|m\rangle|^2 \frac{e^{-\beta(E_m - W)}}{Z_\tau} \frac{Z_\tau}{Z_0} \\ &= e^{\beta(W - \Delta F)} \sum_{mn} \delta[W - (E_m - E_n)] |\langle n|U_{\tilde{\lambda}}(0, \tau)|m\rangle|^2 \frac{e^{-\beta E_m}}{Z_\tau} \\ &= e^{\beta(W - \Delta F)} \sum_{mn} \delta[-W - (E_n - E_m)] \tilde{P}(n|m) p_\tau^{eq}(m) \\ &= e^{\beta(W - \Delta F)} \tilde{P}(-W). \end{aligned} \tag{17}$$

While going from the first to the second line, we have used the fact that unitary evolution is reversible in time: $|\langle m|U_\lambda(\tau, 0)|n\rangle|^2 =$



$\langle n|U_{\tilde{\lambda}}(0, \tau)|m\rangle|^2$, if the external protocol is time-reversed. This is the Crooks fluctuation theorem for a quantum system. The Jarzynski equality follows from the CFT readily by cross multiplication followed by integration over W .

5. Extended Fluctuation Theorems in Presence of Feedback

Now, let us consider a system subjected to a feedback-dependent protocol. This means that the functional form of the protocol with time will depend on outcomes of measurements made on some property of the system. Such systems have been studied in recent literature (see [2] and the references therein). A simple approach to derive these theorems has been discussed in [22]. To keep our discussions simple, we will always assume that the measurements are made on the particle state in phase space for a classical system and projective measurements are made on the Hamiltonian for a quantum system. Let us first study the classical system undergoing Hamiltonian evolution. At time $t = 0$, a given protocol $\lambda_0(t)$ is switched on. The system evolves under this protocol till time $t = t_m$ when a measurement is made on the state of the particle, and let the outcome be m which can be different from the true state z_m due to measurement inaccuracies. This outcome is obtained with probability $p(m|z_m)$. Depending on the outcome, the protocol is changed to $\lambda_m(t)$ which acts on the system till the final time $t = \tau$. The probability of a forward trajectory is given by,

$$\begin{aligned} P[z(t), m] &= p_0^{eq}(z_0)P_{\lambda_0}[z_0 \rightarrow z_m]p(m|z_m)P_{\lambda_m}[z_m \rightarrow z_\tau] \\ &= P_m[z(t)]p(m|z_m). \end{aligned} \tag{18}$$

The reverse process is generated by choosing one of the forward processes (corresponding to one of the values of m) and blindly time-reversing this protocol without applying any feedback. The probability of a reverse trajectory is,

$$\tilde{P}[\tilde{z}(t), m] = \tilde{P}_m[\tilde{z}(t)]p(m). \tag{19}$$

In feedback driven protocols, some information is extracted about the state of the driven system at an intermediate time, and thereafter, the external drive is modified so as to enhance the efficiency of the process.

Modified FTs in presence of feedback also modify the corresponding second law inequalities.

The sequence is as follows: System starts from equilibrium \rightarrow Deterministic evolution from $t = 0$ to $t = t_m \rightarrow$ Measurement at $t = t_m \rightarrow$ Deterministic evolution from $t = t_m$ to $t = \tau$.



The mutual information quantifies the information gained about the actual state z_m through the measurement made at $t = t_m$.

Dividing (18) by (19), we get,

$$\frac{P[z(t), m]}{\tilde{P}[\tilde{z}(t), m]} = e^{\beta(W[z(t), m] - \Delta F(m)) + I}. \quad (20)$$

Here, we have used the condition of microscopic reversibility for any given set of forward and reverse protocols:

$$\frac{P_m[z(t)]}{\tilde{P}_m[\tilde{z}(t)]} = e^{\beta(W[z(t), m] - \Delta F(m))}, \quad (21)$$

and have defined the mutual information I between m and z_m as,

$$I = \ln \frac{p(m|z_m)}{p(m)}. \quad (22)$$

Thus, with (20), we obtain,

$$\langle e^{-\beta(W - \Delta F) - I} \rangle = 1. \quad (23)$$

Application of Jensen's inequality then leads to,

$$\langle W - \Delta F \rangle \geq -k_B T \langle I \rangle, \quad (24)$$

Instead of running the movie of a genuine forward process, one can perform and record a movie of the corresponding reverse process, and then run the movie backwards. Question is: what chance do we stand of distinguishing these two scenarios?

where,

$$\langle I \rangle = \int dm \int dz_m p(m, z_m) \ln \frac{p(m|z_m)}{p(m)}. \quad (25)$$

Since $\langle I \rangle$ is a Kullback–Leibler divergence, it is a non-negative quantity [23], which means that the dissipated work $W_d \equiv W - \Delta F$ can become negative on average. In other words, on an average, more work can be extracted from the system than the change in free energy. This is the modified second law in presence of information gain and feedback.



6. Guessing the Direction of Time's Arrow

When we observe a movie of some process, we can easily make out whether the movie is being run forward or backward. For instance, if the movie shows that broken shards of a flower vase arrange themselves into the original flower vase, then we can say with confidence that the movie is being run backward. In another example, if we see that a dye, well mixed with water is getting concentrated to form the original solid dye, then we can be sure that the movie is being run backward. However, such a directionality in the 'time's arrow' gets blurred when we go to mesoscopic or smaller systems, where thermal fluctuations can give rise to unexpected events. In such cases, $W > \Delta F$ does not necessarily mean that the process is running forward in time, and *vice versa*. In what follows, we explain the quantification of the ability to guess time's arrow from a hypothetical guessing game, following the treatment in [1].

The game consists of watching a movie and trying to guess whether it is being run forward or backward. Given a trajectory $z(t)$ in phase space, the likelihood that it is the forward process is given by $P[F|z(t)]$, and that of it being the reverse process running backward is $P[R|z(t)]$. The total likelihood of the movie being run either forward or backward is unity:

$$P[F|z(t)] + P[R|z(t)] = 1. \quad (26)$$

Bayes' theorem tells us,

$$P[F|z(t)] = \frac{P[z(t)|F]P_F}{P[z(t)]}. \quad (27)$$

Similarly,

$$P[R|z(t)] = \frac{P[z(t)|R]P_R}{P[z(t)]}. \quad (28)$$



Here, $P[z(t)|F]$ is the probability of the trajectory $z(t)$ in the forward process, which is simply $P[z(t)]$ in our earlier notation. Similarly, $P[z(t)|R]$ is the probability that the trajectory $z(t)$ is observed, given that the movie of the reverse process is being run backward, so that $P[z(t)|R] = \tilde{P}[\tilde{z}(t)]$. P_F is the prior probability of a forward process and P_R is that for the reverse process. If the forward and reverse processes are shown based on the outcome of a tossed fair coin, then $P_F = P_R = 1/2$. Dividing (27) by (28), we obtain,

$$\frac{P[F|z(t)]}{P[R|z(t)]} = e^{\beta(W-\Delta F)}, \quad (29)$$

where we have used the relation for microscopic reversibility (5). Now, from (26) and (29), we finally arrive at the following form of the likelihood:

$$P[F|z(t)] = \frac{1}{1 + e^{-\beta(W-\Delta F)}}. \quad (30)$$

Thus, if $W \gg \Delta F$, then $P[F|z(t)] \approx 1$, and we can say with a high degree of confidence that the movie is running forward in time. In the other extreme, if $W \ll \Delta F$, then $P[F|z(t)] \approx 0$, and we can say with a high degree of confidence that the actual process was the reverse one, whose movie is being run backward.

7. Fluctuation Theorems for Stochastic Processes

Till now, we have discussed about FTs when the evolution of the system in phase space is governed by Hamiltonian dynamics. It is important to note that they have also been derived the context of stochastic dynamics, when the evolution of the system is governed by the Langevin equation of motion, or when the probability distribution obeys the master equation. Thus, the theorems are quite robust and can be shown to be valid for a large variety of systems. In the stochastic evolution, however, one needs to incorporate the notion of dissipated heat, which is the energy that the



system dissipates into the environment with which it interacts. In that case, the total entropy change is given by the summation of the change in system's entropy and the change in the entropy of its environment. Then the total entropy change, in general, does not vanish on average, rather it is always non-negative, becoming zero only in the limit of quasistatic driving. For further details, see [2] and the references therein.

Work fluctuation theorems have been derived, for open quantum particles too, by considering unitary evolution of the combined system consisting of the system of interest and the surrounding heat bath [24].

8. Conclusion

In this short article, we have discussed the fluctuation theorems for work and entropy change for isolated systems undergoing Hamiltonian evolution in classical case or unitary evolution in the quantum case. We have seen that the concept of microscopic reversibility directly gives rise to the Crooks FT which in turn leads to the Jarzynski equality. Use of Jensen's inequality on JE or on the integral FT for entropy change leads to statements similar to those of the second law. These theorems get modified if the external protocol is feedback-driven, and so does the second law inequality. In this situation, the extracted work can become more than the free energy difference between the final and the initial states. We have also discussed the likelihood with which one can guess correctly the direction of time's arrow by looking at the movie of a process.

Acknowledgment

Arun M Jayannavar thanks DST, India for financial support (through J. C. Bose National Fellowship).

In Hamiltonian dynamics, the fluctuations of thermodynamic quantities arose entirely from the randomness of the initial state. In a stochastic process, the system and the bath remain connected throughout, so that collisions with bath molecules provide an additional source of randomness.



Suggested Reading

- [1] C Jarzynski, Equalities and Inequalities : Irreversibility and the Second Law of Thermodynamics at the Nanoscale, *Annu. Rev. Condens. Matter Phys.*, **2**, 329, 2010.
- [2] U Seifert, Stochastic Thermodynamics, Fluctuation Theorems and Molecular Machines, *Rep. Prog. Phys.*, **75**, 126001, 2012.
- [3] C Bustamante, J Liphardt and F Ritort, The Nonequilibrium Thermodynamics of Small Systems, *Physics Today*, **58**, 43, 2005.
- [4] D J Evans, E G D Cohen and G P Morriss, Probability of Second Law Violations in Shearing Steady States, *Phys. Rev. Lett.*, **71**, 2401, 1993.
- [5] G Gallavotti and E G D Cohen, Dynamical Ensembles in Nonequilibrium Statistical Mechanics, *Phys. Rev. Lett.*, **74**, 2694, 1995.
- [6] J Kurchan, Fluctuation Theorem for Stochastic Dynamics, *J. Phys. A: Math. Gen.*, **31**, 3719, 1998.
- [7] J L Lebowitz and H Spohn, A Gallavotti-Cohen Type Symmetry in the Large Deviation Functional for Stochastic Dynamics, *J. Stat. Phys.*, **95**, 333, 1999.
- [8] C Jarzynski, Nonequilibrium Equality for Free Energy Differences, *Phys. Rev. Lett.*, **78**, 2690, 1997.
- [9] G E Crooks, Entropy Production Fluctuation Theorem and the Nonequilibrium Work Relation for Free Energy Differences, *Phys. Rev. E*, **60**, 2721, 1999.
- [10] D Chandler, *Introduction to Modern Statistical Mechanics*, Oxford University Press, USA, 1987.
- [11] G M Wang, E M Sevick, E Mittag, D J Searles and D J Evans, Experimental Demonstration of Violations of the Second Law of Thermodynamics for Small Systems and Short Time Scales, *Phys. Rev. Lett.*, **89**, 050601, 2002.
- [12] J Liphardt, S Dumont, S B Smith, I Tinoco Jr and C Bustamante, Equilibrium Information from Nonequilibrium Measurements in an Experimental Test of Jarzynski's Equality, *Science*, **296**, 1832, 2002.
- [13] R van Zon, S Ciliberto and E G D Cohen, Power and Heat Fluctuation Theorems for Electric Circuits, *Phys. Rev. Lett.*, **92**, 130601, 2004.
- [14] D Collin, F Ritort, C Jarzynski, S B Smith, I Tinoco Jr and C Bustamante, Verification of the Crooks Fluctuation Theorem and Recovery of RNA Folding Free Energies, *Nature*, **437**, 231, 2005.
- [15] U Seifert, Stochastic Thermodynamics: Principles and Perspectives, *Eur. Phys. J. B*, **64**, 423, 2008.
- [16] G M Wang, J C Reid, D M Carberry, D R M Williams, E M Sevick and D J Evans, Experimental Study of the Fluctuation Theorem in a Nonequilibrium Steady State, *Phys. Rev. E*, **71**, 046142, 2005.
- [17] S Joubaud, N B Garnier and S Ciliberto, Fluctuations of the Total Entropy Production in Stochastic Systems, *Europhys. Lett.*, **82**, 30007, 2008.
- [18] J Kurchan, arxiv/cond-mat: 0007360.
- [19] Y Morikuni and H Tasaki, Quantum Jarzynski-Sagawa-Ueda Relations, *J. Stat. Phys.*, **143**, 1, 2010.
- [20] M Campisi, P Hänggi and P Talkner, Colloquium: Quantum Fluctuation Relations: Foundations and Applications, *Rev. Mod. Phys.*, **83**, 771, 2011.



- [21] P Talkner, E Lutz and P Hänggi, Fluctuation Theorems: Work is Not an Observable, *Phys. Rev. E*, 75, 050102, 2007.
- [22] S Lahiri, S Rana and A M Jayannavar, Fluctuation Theorems in the Presence of Information Gain and Feedback, *J. Phys. A: Math. Theor.*, 45, 065002, 2012.
- [23] T M Cover and J A Thomas, *Elements of Information Theory*, Wiley-Interscience, 1991.
- [24] M Campisi, P Talkner and P Hänggi, Fluctuation Theorem for Arbitrary Open Quantum Systems, *Phys. Rev. Lett.*, 102, 210401.

Address for Correspondence

A M Jayannavar
Institute of Physics
Sachivalaya Marg
Bhubaneswar 751 005, India.
Email: jayan@iopb.res.in

Sourabh Lahiri
Department of Physics
Birla Institute of Technology
Mesra
Ranchi, Jharkhand 835215,
India.
Email:
sourabhlahiri@gmail.com

