Article

Jarzyski’s Equality and Crooks’ Fluctuation Theorem for General Markov Chains with Application to Decision-Making Systems

Pedro Hack *, Sebastian Gottwald and Daniel A. Braun

Institute of Neural Information Processing, Ulm University, 89081 Ulm, Germany
* Correspondence: pedro.hack@uni-ulm.de

Abstract: We define common thermodynamic concepts purely within the framework of general Markov chains and derive Jarzynski’s equality and Crooks’ fluctuation theorem in this setup. In particular, we regard the discrete-time case, which leads to an asymmetry in the definition of work that appears in the usual formulation of Crooks’ fluctuation theorem. We show how this asymmetry can be avoided with an additional condition regarding the energy protocol. The general formulation in terms of Markov chains allows transferring the results to other application areas outside of physics. Here, we discuss how this framework can be applied in the context of decision-making. This involves the definition of the relevant quantities, the assumptions that need to be made for the different fluctuation theorems to hold, as well as the consideration of discrete trajectories instead of the continuous trajectories, which are relevant in physics.

Keywords: fluctuation theorems; Markov chains; information processing; decision-making

1. Introduction

Over the last 20 years, several advances in thermodynamics have led to the development of results relating equilibrium quantities to nonequilibrium trajectories. Those advances have crystallized in a new area of research, nonequilibrium thermodynamics, where these relations play a major role [1–3]. Among them, two of the most remarkable are Jarzynski’s equality [4–6] and Crooks’ fluctuation theorem [7,8], for which experimental evidence has been reported in several contexts: unfolding and refolding processes involving RNA [9,10], electronic transitions between electrodes manipulating a charge parameter [11], the rotation of a macroscopic object inside a fluid surrounded by magnets where the current of a wire attached to the macroscopic object is manipulated [12], and a trapped-ion system [13,14].

These two results have been derived under several assumptions in the context of nonequilibrium thermodynamics, including both deterministic [5,6] and stochastic dynamics [4,7,8,15]. Moreover, it has been argued that both results can be obtained as a consequence of Bayesian retrodiction in a physical context [16]. Here, we derive both of them using only the concepts from the theory of Markov chains. This allows us to both distinguish the mathematical from the physical assumptions underlying them and, thus, to make them available for application in other areas where the framework of thermodynamics may be useful. The distinction between mathematical and physical assumptions will be of particular importance for the definition of work, as we will see, since the usual definition based on physical considerations leads to an asymmetry of the definition in processes that run either forward or backward in time—see Figure 1 for a simple example. This is relevant, for instance, when analysing trajectories in terms of their work value, if we do not know whether they were recorded in the forward direction or whether they have been generated by playing them backwards. Ideally, we would like to be able to ascribe work values directly to trajectories without any additional information.
Figure 1. Relation between a forward process, its corresponding backward process, and the definition of work. We consider a trajectory $x = (x_0, x_1, x_2)$ and three energy functions $E_0, E_1, E_2$. The upper (bottom) line of arrows represents the forward (backward) process. A work step $W_i = E_i(x_{i-1}) - E_{i-1}(x_{i-1})$ is typically defined as the change in energy due to the external change of the energy function, whereas a heat step $Q_i = E_i(x_i) - E_i(x_{i-1})$ is defined as the change in energy due to internal state changes.

(a) Typical relation between the forward and backward processes in physics [8]. Work in the forward process would be $W_F = E_1(x_0) - E_0(x_0) + E_2(x_1) - E_1(x_1)$, whereas the backward work under the same definition would be $W_B = E_1(x_1) - E_2(x_1)$. Instead, backward work is usually defined as $W_B = E_1(x_1) - E_2(x_1) + E_0(x_0) - E_1(x_0)$ to fulfil the physical time reversal symmetry $W_F = -W_B$.

(b) Another typical protocol in physics [15,17]. In this case, the asymmetry is the other way round, where $E_2$ does not influence the forward process.

(c) Symmetric protocol where both forward and backward work follow the same definition with $W_F = E_1(x_1) - E_0(x_0)$ and $W_B = E_0(x_1) - E_1(x_1) = -W_F$. This is the protocol we propose in Section 4.

One of the application areas where the framework of thermodynamics has recently been investigated outside the realm of physics is the analysis of simple learning systems [1,18–26] and, in particular, the problem of decision-making under uncertainty and resource constraints [22,27–31]. The basic analogy follows from the idea that decision-making involves two opposing forces: (i) the tendency of the decision-maker toward better options (equivalently, to maximize a function called utility) and (ii) the restrictions on this tendency given by the limited information-processing capabilities of the decision-maker, which prevents him/her from always picking the best option and is usually modelled by a bound on the entropy of the probability distribution that describes the decision-maker’s behaviour. Thermodynamic systems are also explained in terms of two opposing forces, the first being the energy, which the system tries to minimize, and the second being the entropy, which prevents the minimization of the energy to its full extent. Thus, in both cases, we formally deal with optimization problems under information constraints, and thus, we can conceptualize both decision-making and thermodynamics in terms of information theory. In particular, we can consider the environment in which a decision is being made or a thermodynamic system is immersed as a source of information (in the form of either utility or energy), which, due to the noise (modelled by entropy), reaches the decision-
making or thermodynamic system with some error. This results in an imperfect response by the system.

The analogy between thermodynamics and decision-making is not restricted to the equilibrium case [22,27–31], but can be taken further from equilibrium to nonequilibrium systems. In particular, the aforementioned fluctuation theorems of Jarzynski and Crooks have been previously suggested to apply to decision-makers that adapt to changing environments [32]. In this previous work, hysteresis and adaptation were investigated in decision-makers, however, based on the physical convention of defining work differently for forward and backward processes. Here, we improve on the work there by replacing this convention with a different energy protocol that naturally entails a symmetric definition of work and by weakening the assumptions that are actually needed in order for the fluctuation theorems to hold in the context of general Markov chains. Given the fact that the literature on this topic belongs for the most part to thermodynamics, we adopt the thermodynamic notation here. In particular, we consider energy functions instead of utilities and take Markov chains as the starting point.

Our manuscript is organized as follows. In Section 2, we introduce the notions of work and other thermodynamic concepts that are inherent to Markov chains, that is, in contrast to the formalism in physics [7,8,15], we start with the assumption of a Markov chain and deduce all other concepts from that without the need to presuppose the existence of an external energy function. We discuss under what conditions these concepts are uniquely specified from the Markov chain. In Section 3, we use this framework to weaken the derivation of Jarzynski’s equality (Theorem 1) in the context of decision-making that was presented in [32]. In Section 4, we prove Crooks’ fluctuation theorem (Theorem 2 and Corollary 2) within the same setup. In particular, we use an additional assumption that is not mandatory in Crooks’ work [7,8,15], but here is needed given the inherent nature of our definition of work. In fact, we provide an example in which the new requirement is violated and, as a consequence, Crooks’ theorem is false everywhere (Proposition 3). In Section 5, we discuss how the concepts we have developed can be applied to decision-making systems. Moreover, we follow the terminology in [33]. Notice, for simplicity, we develop the results for discrete-time Markov chains with finite state spaces. However, the ideas can be translated, for example, to continuous state spaces by assuming for densities of Markov kernels the properties we use here for transition matrices. We include a few more details regarding this scenario in the discussion (Section 5), where we also briefly address the case of continuous-time Markov chains.

### 2. Thermodynamics for Markov Chains

#### 2.1. Definitions of Energy, Heat, and Work

In this section, we assume there is some stochastic process that can be modelled as a Markov chain and discuss the definition of the energy, partition function, free energy, heat, and dissipated work in such a context. We follow the terminology in [33].

We call a finite number of random variables $X = (X_n)_{n=0}^N$ over a finite state space $S$ a Markov chain if we have for any $0 < n \leq N$ that

$$\mathbb{P}(X_n = x_n | X_0 = x_0, \ldots, X_{n-1} = x_{n-1}) := \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1})$$

for all $(x_0, x_1, \ldots, x_n) \in S^{n+1}$. Notice that we can characterize $X$ by a probability distribution $p_0$, where $p_0(x) = \mathbb{P}(X_0 = x)$ and $N$ transition matrices $(M_n)_{n=1}^N$ given by

$$(M_n)_{xy} := \mathbb{P}(X_n = x | X_{n-1} = y)$$

for all $x, y \in S$ and $1 \leq n \leq N$. Notice that any transition matrix $M_n$ is a stochastic matrix, that is we have $\sum_{y \in S} (M_n)_{xy} = 1$ for all $y \in S$.

If $p$ is a distribution, $M_n$ a transition matrix of $X$ for some fixed $n$, and we have $M_n p = p$, where $(M_n p)(x) := \sum_{y \in S} (M_n)_{xy} p(y)$, then we say $p$ is a stationary distribution of $M_n$. Note that this terminology applies to a single $M_n$, that is a stationary distribution $p$ of $M_n$ is stationary with respect to the (homogeneous) Markovian dynamics of that
fixed transition matrix, and generally not with respect to the (inhomogeneous) Markovian dynamics of $X$. If $p$ fulfills
\[
(M_n)_{xy}p(x) = (M_n)_{xy}p(y) \quad \forall x, y \in S,
\]
then we say $p$ satisfies detailed balance with respect to $M_n$. Note that such a $p$ is a stationary distribution of $M_n$. In case $M_n$ has a unique stationary distribution $p$ that satisfies detailed balance with respect to $M_n$, then we may simply say $M_n$ satisfies detailed balance. We say a transition matrix $M_n$ is irreducible if, for any pair $x, y \in S$, there exists an integer $m \geq 1$ such that $(M_n^m)_{xy} > 0$. Irreducible transition matrices have a useful property, which we present in Lemma 1 (see ([33], Corollary 1.17 and Proposition 1.19) for a proof).

**Lemma 1.** If a transition matrix is irreducible, then it has a unique stationary distribution. Furthermore, the stationary distribution has non-zero entries.

If $X$ has initial distribution $p_0$, irreducible transition matrices $(M_n)_{n=1}^N$ and $p_N$ is the unique stationary distribution of $M_N$, then we say the Markov chain $Y := (Y_n)_{n=0}^N$ with initial distribution $p_N$ and transition matrices $(M_{N-(n-1)})_{n=1}^N$ is the time reversal of $X$. Notice that there are different notions of time reversal. A discussion can be found in ([34], Section III).

Given a distribution $p$ on $S$ such that $p(x) > 0$ for all $x \in S$ and some $\beta > 0$, we can associate with $p$ an energy function $E$, that is a function $E : S \to \mathbb{R}$ such that
\[
p(x) = \frac{1}{Z} e^{-\beta E(x)} \quad \forall x \in S,
\]
where $Z := \sum_{x \in S} e^{-\beta E(x)}$ is called a partition function and $F := -\frac{1}{\beta} \log(Z)$ the corresponding free energy. Notice that, given a distribution $p$ with two energy functions $E$ and $E'$, there is a constant $c \in \mathbb{R}$ such that we have
\[
E(x) = E'(x) + c \quad \forall x \in S
\]
and, accordingly,
\[
F = F' + c
\]
where $F$ and $F'$ are the free energies for $p$ using $E$ and $E'$, respectively. Thus, each distribution where all entries are strictly positive has, up to a constant, a unique energy function associated with it.

For the remainder of this section, let $X := (X_n)_{n=0}^N$ be a Markov chain such that $p_0$ has non-zero entries and each transition matrix $M_n$ has a unique stationary distribution $p_n$ with non-zero entries. In this context, we call a family $E = (E_n)_{n=0}^N$ of functions $E_n : S \to \mathbb{R}$ a family of energies of $X$, if $E_n$ is an energy function of $p_n$ for all $0 \leq n \leq N$. We define the work of a realization $x = (x_0, x_1, ..., x_N) \in S^{N+1}$ of $X$, with respect to a family of energies $E$ of $X$, as
\[
W_{X,E}(x) := \sum_{n=0}^{N-1} E_{n+1}(x_n) - E_n(x_n).
\]
Given another family of energies $E' = (E'_n)_{n=0}^N$ of $X$, we have
\[
W_{X,E}(x) = W_{X,E'}(x) + (c_N - c_0),
\]
where $c_n := E_n - E'_n$ are constants (by (2)). Hence, without fixing a family of energies of $X$, the work defined in (4) is unique up to a constant. Whenever $X$ and $E$ are clear from the context, we may simply use $W$ instead of $W_{X,E}$ for brevity.
Similarly, the heat of a realization $x$ of $X$ with respect to a family of energies $E$ is given by

$$Q_{X,E}(x) := \sum_{n=1}^{N} E_n(x_n) - E_n(x_{n-1}).$$

(6)

Given another family of energies $E' := (E'_n)_{n=0}^{N}$ of $X$, we have

$$Q_{X,E}(x) = Q_{X,E'}(x)$$

(7)

by (2). We may, thus, use $Q_X$ instead of $Q_{X,E}$, or even $Q$ in case $X$ is clear.

Moreover, if $P_n$ is the free energy associated with $p_0$ for any $0 \leq n \leq N$, we call $\Delta F_{X,E} := F_N - F_0$ the free energy difference associated with $E$, satisfying

$$\Delta F_{X,E} = \Delta F_{X,E'} + (c_N - c_0)$$

(8)

by (3). While both $W_{X,E}$ and $\Delta F_{X,E}$ depend on the difference between the constants $c_N$ and $c_0$, the so-called dissipated work:

$$W_{X,E}^{d}(x) := W_{X,E}(x) - \Delta F_{X,E}$$

does not, that is, for any realization $x$ of $X$, we have

$$W_{X,E}^{d}(x) = W_{X,E}^{d}(x)$$

(9)

as a consequence of (5) and (8).

Notice, in the context of the Markov chain framework we have adopted here, the first law of thermodynamics is a direct consequence of the definitions of work and heat (see (12) below):

$$W_{X,E}(x) + Q_{X,E}(x) = E_N(x_N) - E_0(x_0).$$

The second law of thermodynamics can be obtained as well, which is a direct consequence of Jarzynski’s equality, as we will see in Section 5.

2.2. Main Result: Fluctuation Theorems for Markov Chains

Our main results are versions of Jarzynski’s equality [4–6] and Crooks’ fluctuation theorem [7,8,15] for Markov chains, the derivations of which can be found in Sections 3 and 4 below.

Consider a Markov chain $X = (X_n)_{n=0}^{N}$ on a finite state space whose initial distribution $p_0$ has non-zero entries and whose transition matrices $(M_n)_{n=1}^{N}$ are irreducible. Then, for any family of energies $E = (E_n)_{n=0}^{N}$ of $X$, we have

$$\left< e^{-\beta(W(X) - \Delta F)} \right> = 1,$$

where $\langle \cdot \rangle$ denotes the expectation operator and $W = W_{X,E}$, $\Delta F = \Delta F_{X,E}$. In physics, this result is known as Jarzynski’s equality, which has been shown in the past to hold under various conditions. In Section 3, we give a simple proof in the context of Markov chains as a direct consequence of the definitions of work and free energy (Theorem 1).

Moreover, if all transition matrices of $X$ satisfy detailed balance and $p_0$ is the stationary distribution of $M_1$, then for any possible work value, $w$, we have

$$\frac{P(W_{X,E} = w)}{P(W_{Y,E} = -w)} = e^{\beta(w - \Delta F)},$$

where $Y$ is the time reversal of $X$ with energies $\hat{E} = (\hat{E}_n)_{n=0}^{N}$. The analogous result in physics is known as Crooks’ fluctuation theorem. In Section 4, we prove a slightly more general version in the context of Markov chains without detailed balance (Theorem 2), which is then applied in Corollary 2 to obtain Crooks’ theorem.
3. Jarzynski’s Equality for Markov Chains

Jarzynski’s equation was originally derived for deterministic dynamics [5,6] (see also [35]) and later extended to stochastic dynamics [4] using a Master equation approach. Shortly after that, it was shown in the non-deterministic Markov chain context relying on assumptions about the time reversal of the dynamics [8]. In Theorem 1 below, we see that, in the context of Markov chains, Jarzynski’s equality is a straightforward consequence of the definitions of work and free energy. Importantly, it does not require any assumptions regarding time reversal, in contrast to the requirements in a previous decision-theoretic approach to fluctuation theorems in [32].

For the proof of Jarzynski’s equality for Markov chains, the basic observation is that we start with the expected value of a quantity closely related to the equilibrium distributions of our initial Markov chain $X$, namely $e^{-βW(X)}$. With this in mind, we define a new Markov chain $Y$ using the transition matrices of $X$ and the equilibrium distributions of the individual steps. In particular, we define it in a way such that we cancel the dependency of $e^{-βW(X)}$ on $X$ and end up with a constant, whose expected value over $Y$ is the constant itself. We include the details in the following theorem.

**Theorem 1** (Jarzynski’s equality for Markov chains). If $X = (X_n)_{n=0}^N$ is a Markov chain on a finite state space $S$ whose initial distribution $p_0$ has non-zero entries and whose transition matrices $(M_n)_{n=1}^N$ are irreducible, then we have, for any family of energies $E = (E_n)_{n=0}^N$ of $X$,

$$
\langle e^{-β(W(X)-ΔF)} \rangle = 1,
$$

(10)

where $W = W_{X,E}$ and $ΔF = ΔF_{X,E}$.

**Proof.** Notice that, by Lemma 1, we have $p_n(x) > 0$ for all $x \in S$ and $0 \leq n \leq N$. We first define a new Markov chain $Y := (Y_n)_{n=0}^N$ with initial distribution $p_N$ and with transition matrices $(\hat{M}_n)_{n=1}^N$, where for all $x, y \in S$,

$$
(\hat{M}_{n+1})_{xy} := \frac{p_{N-n}(x)}{p_{N-n}(y)} (M_{N-n})_{yx}
$$

(11)

for $0 \leq n \leq N - 1$. Notice that $\hat{M}_n$ is a stochastic matrix for $1 \leq n \leq N$, as we have for all $y \in S$

$$
\sum_{x \in S} (\hat{M}_n)_{xy} = \sum_{x \in S} (M_{N+1-n})_{yx} \frac{p_{N+1-n}(x)}{p_{N+1-n}(y)}
= \frac{1}{p_{N+1-n}(y)} \sum_{x \in S} (M_{N+1-n})_{yx} p_{N+1-n}(x) = 1,
$$

where we applied (11) in the first equality and the fact that $p_n$ is a stationary distribution for $M_n$ for $1 \leq n \leq N$ by the assumption in the last equality. Thus, $Y$ is well defined. Note that, by definition we have

$$
P(Y_{n+1} = x_{n+1} | Y_n = x_n) = P(Y_{n+1} = x_{n+1} | Y_n = x_n, \ldots, Y_0 = x_0)
= (\hat{M}_{n+1})_{x_n, x_{n+1}}
$$

for $0 \leq n \leq N - 1$, $(x_0, \ldots, x_{n+1}) \in S^{n+2}$. We can now use $Y$ to show the result:
\[
\langle e^{-\beta W(X)} \rangle \overset{(i)}{=} \sum_{x_0,\ldots,x_N \in S} P(X_0 = x_0)(M_1)_{x_0} (M_2)_{x_1} \cdots (M_N)_{x_N} x_{N+1} \times \\
e^{-\beta E_1(x_0)} \cdots e^{-\beta E_N(x_{N-1})}
\]

\[
\overset{(ii)}{=} \sum_{x_0,\ldots,x_N \in S} P(X_0 = x_0)(\tilde{M}_N)_{x_N} e^{-\beta E_1(x_1)} \cdots (\tilde{M}_1)_{x_{N-1}} x_N \times \\
e^{-\beta E_N(x_N)} e^{-\beta E_1(x_0)} \cdots e^{-\beta E_N(x_{N-1})}
\]

\[
\overset{(iii)}{=} \frac{Z_N}{Z_0} \sum_{x_0,\ldots,x_N \in S} P(Y_0 = x_N)(\tilde{M}_1)_{x_{N-1}} x_N \cdots (\tilde{M}_1)_{x_0} x_1 
\]

\[
\overset{(iv)}{=} e^{-\beta \Delta F},
\]

where we use the Markov property in (i) and apply (11) in (ii). In (iii), we cancel the repeated terms coming from the definition of \((\tilde{M}_n)_{n=1}^N\) and from \(e^{-\beta W(x)}\), since we have

\[
\sum_{n=1}^N E_n(x_n) - E_n(x_{n-1}) = E_N(x_N) - E_1(x_0) + \sum_{n=1}^{N-1} E_n(x_n) - E_{n+1}(x_n)
\]

\[
= E_N(x_N) - E_0(x_0) - W(x) \quad (12)
\]

which leads to

\[
P(X_0 = x_0) e^{-\beta E_1(x_1)} \cdots e^{-\beta E_N(x_{N-1})} e^{-\beta E_1(x_0)} \cdots e^{-\beta E_N(x_{N-1})} \]

\[
= \frac{1}{Z_0} e^{-\beta (E_0(x_0) + E_N(x_N) - E_0(x_0) - W(x) + W(x))}
\]

\[
= \frac{Z_N}{Z_0} P(Y_0 = x_N).
\]

Lastly, we apply the definition of \(\Delta F\) and the normalization of \(Y\) in (iv). \(\Box\)

The main purpose of proving Theorem 1 is to show that Jarzynski’s equality can be obtained under milder conditions than the ones that were considered before in decision-making. It should be noted that, in contrast to the usual approaches such as [8], where \(Y\) is assumed to be the time reversal of \(X\), here, it is just a convenient mathematical object for the proof. A similar approach to Jarzynski’s equality in the context of continuous-time Markov chains can be found in [36]. Our approach to the discrete-time case in Theorem 1 is much simpler, since we do not need measure-theoretic concepts nor smoothness assumptions.

4. Crooks’ Fluctuation Theorem for Markov Chains

The original derivation of Crooks’ fluctuation theorem for Markovian dynamics [7,8,15] was carried out using a definition of work different from the one in (4). In this section, we derive the theorem for Markov chains using (4) and comment on the difference between these approaches in the discussion. As discussed in the Introduction, an additional hypothesis is needed for the result to hold in our setup. We derive Crooks’ fluctuation theorem using this additional assumption in Theorem 2 and Corollary 2 below, and then, in Proposition 3, we provide an example where this requirement is violated and the theorem is false everywhere.

Before proving the intermediate results and, finally, Crooks’ fluctuation theorem, let us briefly sketch the procedure we follow throughout this section. We start with Proposition 1, where we use the same Markov chain \(Y\) that we defined in the proof of Theorem 1 and
obtain, along similar lines, a more precise relation between $X$ and $Y$, in particular between the probability of some realization of $X$ and that of the same realization (with the events taking place in reversed order) of $Y$. As a matter of fact, we show that, for a given $X$, $Y$ is (roughly) the only Markov chain fulfilling such a relation. Then, in Proposition 2, we show how the driving signals of $X$ and $Y$ are related. In order to do so, we exploit the relation between the equilibrium distributions of $X$ and $Y$, which comes from the fact the both Markov chains share the same equilibrium distributions. By combining these two propositions, we reach a relation between the probability distributions of the driving signals of $X$ and $Y$ in Theorem 2. Lastly, in Corollary 2, we impose an extra condition on the equilibrium distributions of $X$ in order to obtain Crooks’ fluctuation theorem in its usual form. We proceed now to show the details involved in the argument we just presented. We start by proving Proposition 1.

**Proposition 1.** If $X = (X_n)_{n=0}^N$ is a Markov chain on a finite state space $S$ whose initial distribution $p_0$ has non-zero entries and whose transition matrices $(M_n)_{n=1}^N$ are irreducible, then there exists a unique Markov chain $Y = (Y_n)_{n=0}^N$ such that $Y_0 \sim p_N$, $(\hat{M}_{n+1})_{xx} = (M_{N-n})_{xx}$ for $0 \leq n \leq N - 1$, where $(\hat{M}_n)_{i,j}$ are the transition matrices of $Y$, and for $x = (x_0,x_1,\ldots,x_N) \in S_{N+1}$,

$$
\mathbb{P}(X_1 = x_1, \ldots, X_N = x_N|X_0 = x_0) = \mathbb{P}(Y_1 = x_{n-1}, \ldots, Y_N = x_0|Y_0 = x_N) e^{-\beta Q(x)}
$$

for any family of energies $E = (E_n)_{n=0}^N$ of $X$. Moreover, this unique $Y$ satisfies

$$
\mathbb{P}(X = x) = \mathbb{P}(Y = x^R) e^{\beta(W(x) - \Delta F)},
$$

where $W = W_{X,E}$, $\Delta F = \Delta F_{X,E}$, and $x^R := (x_N,\ldots,x_0)$ denotes the reversal of $x$. In particular, the probability of $X$ following $x$ is the same as the probability of $Y$ following $x^R$ if and only if $\Delta F = W(x)$.

**Proof.** Consider the Markov chain $Y = (Y_n)_{n=0}^N$ defined in the proof of Theorem 1, which is well defined, since we have the same hypotheses. We can proceed in the same way as in Theorem 1 to obtain

$$
\mathbb{P}(X_1 = x_1, \ldots, X_N = x_N|X_0 = x_0) (i) \equiv (M_1)_{x_1,x_0} \cdots (M_N)_{x_N,x_{N-1}}
$$

$$
\equiv e^{-\beta E_1(x_1)} \cdots e^{-\beta E_N(x_N)}
$$

$$
\equiv e^{-\beta Q(x)} \mathbb{P}(Y_1 = x_{n-1}, \ldots, Y_N = x_0|Y_0 = x_N),
$$

where we applied the Markov property of $X$ in (i), (11) in (ii), and the definition of $Q$ plus the Markov property of $Y$ in (iii). This proves the existence of a Markov chain with the desired properties. Moreover, we have

$$
\mathbb{P}(X = x) (i) \equiv e^{-\beta E_0(x_0)} Z_0 e^{-\beta Q(x)} \mathbb{P}(Y_1 = x_{n-1}, \ldots, Y_N = x_0|Y_0 = x_N)
$$

$$
\equiv Z_0 e^{-\beta (Q(x) - \left(E_N(x_N) - E_0(x_0)\right))} \mathbb{P}(Y_0 = x_N, \ldots, Y_N = x_0)
$$

$$
\equiv e^{\beta (W(x) - \Delta F)} \mathbb{P}(Y = x^R),
$$

where we applied (4), the definition of conditional probability and the fact $X_0$ follows $p_0$ in (i), the definition of the conditional probability, the fact that $Y_0$ follows $p_N$ in (ii), and both the definitions of $Q$ and $\Delta F$ plus (12) in (iii).
Thus, in this case, dissipated work is an unambiguous measure of the discrepancy between the probability of observing a realization of $X$ and the probability of observing the same realization in the time reversal $Y$. Assume $Z = (Z_n)_{n=0}^N$ is a Markov chain with transition matrices $(M'_n)_{n=1}^N$ such that $Z_0 \sim p_N$, (13) holds, and $(M'_n)_{xX} = (M_n)_{xX}$ for all $x \in S$ and $0 \leq n \leq N - 1$. Consider some $n$ such that $1 < n < N$ and some $a, b \in S$ with $a \neq b$. By (13), we have
\[
P(X_1 = a, \ldots, X_{N-(n-2)} = a, X_{N-(n-1)} = b, \ldots, X_N = b | X_0 = a) e^{-\beta E_{N-(n-1)}(a)}
= \Pr(Z_1 = b, \ldots, Z_{n-1} = b, Z_n = a, \ldots, Z_N = a | Z_0 = b) e^{-\beta E_{N-(n-1)}(b)}.
\]
Applying the Markov property, the fact that $(M'_{n+1})_{xx} = (M_{n})_{xx}$ for $0 \leq n \leq N - 1$, and the definition of $E_{N-(n-1)}$, we obtain
\[
(M'_n)_{ab} = (M_{n-(n-1)})_{ba} \frac{p_{N-(n-1)}(a)}{p_{N-(n-1)}(b)} = (\hat{M}_n)_{ab}.
\]
Since the argument also works for $n = 1$ and $n = N$, the case $a = b$ holds by definition, and $Y_0, Z_0 \sim p_N$, we have $Z = Y$. □

We say $X$ is microscopically reversible [15] if (13) is satisfied for $Y$ being the time reversal of $X$, i.e., if the unique Markov chain $Y$ that exists by Proposition 1 has initial distribution $p_N$ and transition matrices $(M_{n})_{n=1}^N$. Notice, if this property holds, then (14) relates the probability of observing a realization $x$ of a Markov chain $X$ with that of observing the reversed realization $x^R$ in the time reversal $Y$ of $X$, that is when starting with the equilibrium distribution of the last environment and choosing according to the same conditional probabilities, but in reversed order. This is the case if and only if the transition matrices of $X$ satisfy detailed balance, as we show in the following lemma.

**Lemma 2.** If $X = (X_n)_{n=0}^N$ is a Markov chain on a finite state space $S$ with irreducible transition matrices $(M_n)_{n=1}^N$ and initial distribution with non-zero entries, then $M_n$ satisfies detailed balance for $1 \leq n \leq N$ if and only if $X$ is microscopically reversible.

**Proof.** If $X$ satisfies detailed balance, that is if each $p_n$ satisfies (1), then by the definition of $\hat{M}$ in (11), we have $\hat{M}_n = M_{N-(n-1)}$ for each $1 \leq n \leq N$. Hence, in this case, the Markov chain $Y$ constructed in Theorem 1 and Proposition 1 is the time reversal of $X$, and so, $X$ is microscopically reversible by definition.

It remains to show that, if $X$ is microscopically reversible, then its transition matrices satisfy detailed balance. Let $Y$ be the time reversal of $X$. Since, by assumption, $Y_0 \sim p_N$ and $\hat{M}_{N-(n-1)} = M_n$, where $(\hat{M}_n)_{n=1}^N$ are the transition matrices of $Y$, we can follow the proof of Proposition 1 to obtain for $1 \leq n \leq N$\[
(M_n)_{ab} = (\hat{M}_{n-(n-1)})_{ba} = (M_n)_{ba} \frac{p_n(a)}{p_n(b)}.
\]
Thus, for each $1 \leq n \leq N$, $p_n$ satisfies detailed balance with respect to $M_n$. □

In particular, this means that, if $X$ satisfies detailed balance, then the time reversal $Y$ of $X$ satisfies (14), that is
\[
P(X = x) = \Pr(Y = x^R) e^{\beta W(x) - \Delta F} = \Pr(Y = x^R) e^{\beta W^d(x)}
\]
for any family of energies $E = (E_n)_{n=0}^N$ of $X$, where $W^d := W^d_E$ is the dissipated work of $X$. Thus, in this case, dissipated work is an unambiguous measure of the discrepancy between the probability of observing a realization of $X$ and the probability of observing the same trajectory in reversed order in the time reversal of $X$. We have, hence, an unambiguous measure of hysteresis (see Section 5).

Before showing Theorem 2 and Crooks’ fluctuation theorem, we relate the work of $X$ with that of its time reversal.
Proposition 2. If \( X = (X_n)_{n=0}^N \) is a Markov chain on a finite state space \( S \) with the initial distribution with non-zero entries and irreducible transition matrices, then there exists a Markov chain \( Y = (Y_n)_{n=0}^N \) and a constant \( k \in \mathbb{R} \) such that

\[
W_{Y,E}(x^R) = -W_{X,E}(x) + E_1(x_0) - E_0(x_0) + k \quad \forall x \in S^{N+1},
\]

where \( E \) and \( \hat{E} \) are families of energies of \( X \) and \( Y \), respectively. Moreover, if the stationary distribution \( p_1 \) of \( M_1 \) coincides with the initial distribution \( p_0 \) of \( X \), then there exists a constant \( k \in \mathbb{R} \) such that

\[
W_{Y,E}(x^R) = -W_{X,E}(x) + k \quad \forall x \in S^{N+1}.
\]

Proof. Let \( Y = (Y_n)_{n=0}^N \) be the Markov chain defined in the proof of Theorem 1. Since work is well defined (up to a constant) for both \( X \) and \( Y \), by Lemma A2 (see Appendix A), we can use the relation between the energy functions of both chains in (A2) to show (16). We have

\[
W_{Y,E}(x^R) = \sum_{n=0}^{N-1} \hat{E}_{n+1}(x_n^R) - \hat{E}_n(x_n^R)
\]

where in (i) we defined \( \hat{E}_0 \) as a constant since \( \hat{E}_0 = E_N + k_0 \) and \( \hat{E}_1 = E_N + k_1 \) by Lemma A2 (see Appendix A). In (ii), we applied the definition of \( x^R \) and (A2), cancelled the repeated \( k_n \), defined as in Lemma A2, for all \( 1 < n < N \), and introduced \( k := k_n - k_1 + \hat{c} = (\hat{E}_N - E_1) - (\hat{E}_1 - E_N) + (\hat{E}_1 - \hat{E}_0) = (\hat{E}_N - E_1) - (\hat{E}_0 - E_N) \). In (iii), we rewrite the sum in terms of \( m := N - n \).

For the second statement, notice that, if \( p_0 \) is the stationary distribution of \( M_1 \), then there exists a constant \( c \) such that \( E_1 = E_0 + c \) by (2). Thus, we have \( k := (\hat{E}_N - E_0) - (\hat{E}_0 - E_N) = (E_1 - E_0) + (\hat{E}_N - E_1) - (\hat{E}_0 - E_N) = c + k' = E_1(x_0) - E_0(x_0) + k' \) for all \( x_0 \in S \), where \( k' \) is the constant in (16). \( \square \)

If detailed balance holds, then (16) and (17) relate the work along a realization \( x \) of \( X \) with the reversed realization \( x^R \) of its time reversal \( Y \). More precisely, we obtain the following corollary.

Corollary 1 (When work is odd under time reversal). If all transition matrices of \( X \) in Proposition 2 satisfy detailed balance and \( Y \) is the time reversal of \( X \), then the constants \( k \) in (16) and (17) can be taken to be zero.

Proof. For the constant in (16), we simply choose \( \hat{E}_N = E_1 \) and \( \hat{E}_0 = E_N \), and for the constant in (17), we choose \( E_N = E_0 \) and \( \hat{E}_0 = E_N \), which we can do since \( p_0 \) is the stationary distribution of \( M_1 \) and, by Lemma A2 (see Appendix A), also that of \( M_N \). \( \square \)

Remark 1. Note that choosing the energy functions in Corollary 1 is unnecessary whenever both \( X \) and \( Y \) are thermodynamic processes. Although energy is defined only up to a constant in thermodynamics, it would make no sense to pick the constants differently when dealing with a system where the same dynamics occur more than once. Thus, there, we have \( \hat{E}_n = E_{N-(n-1)} \) for \( 1 \leq n \leq N \), \( \hat{E}_0 = \hat{E}_1 \), and in case \( p_0 \) is a stationary distribution of \( M_1 \), \( E_1 = E_0 = \hat{E}_N \). In
particular, when taking $Y$ to be the time reversal of $X$ in thermodynamics, we always have $k = 0$ in (16), and in case $p_0$ is a stationary distribution of $M_1$, $k = 0$ in (17).

The non-constant term $E_1(x_0) - E_0(x_0)$ in (16), which remains even when $X$ satisfies detailed balance, follows from an asymmetry between $X$ and $Y$. In particular, $X$ goes from $E_0$ to $E_N$, whereas $Y$ goes from $E_N$ to $E_1$, because while $X_0 \sim p_0 \propto e^{-\beta E_0}$ and the stationary distribution of $M_N$ is $p_N \propto e^{-\beta E_N}$, which is also the initial distribution of $Y$, the stationary distribution of the final transition matrix $\tilde{M}_N$ of $Y$ is $p_1 \propto e^{-\beta E_1}$ (and not $p_0$). Furthermore, while $X$ may begin with a change in the energy function, since $p_1 \neq p_0$ is allowed, $Y$ does not, as $p_N$ is both the initial distribution of $Y$ and the stationary distribution of $\tilde{M}_1$. An example can be found in Figure 2. This asymmetry is erased if we assume that $p_0$ is the stationary distribution of $M_1$, in which case, for Markov chains $X$ that satisfy detailed balance, the work along any realization of $X$ has the opposite sign of the work along the reversed realization of $Y$. That is, thermodynamic work becomes odd under time reversal.

The following theorem contains Crooks’ fluctuation theorem as the special case when $X$ satisfies detailed balance (see Corollary 2 below).

![Figure 2](image-url)

**Figure 2.** Simple example showing how the asymmetry between $X$ and its time reversal $Y$ manifests in thermodynamics. We consider $x = (x_0, x_1, x_2)$ a trajectory, $(E_0, E_1, E_2)$ with $E_1(x_0) \neq E_0(x_0)$ the energy functions of $X$ (upper line of arrows), and by Remark 1, $(E_2, E_2, E_1)$ the energy functions of $Y$ (bottom line of arrows). We have $W_X(x) = E_2(x_1) - E_1(x_1) + E_1(x_0) - E_0(x_0)$ and $W_Y(x^R) = E_1(x_1) - E_2(x_1)$. As a result, $W_Y(x^R) = -W_X(x) + E_1(x_0) - E_0(x_0)$, in accordance with (16) and Corollary 1. Thus, thermodynamic work is not odd under time reversal in general.

**Theorem 2.** If $X = (X_n)_{n=0}^N$ is a Markov chain on a finite state space $S$ whose initial distribution $p_0$ has non-zero entries, whose transition matrices $(M_n)_{n=1}^N$ are irreducible, and where $p_0$ is the stationary distribution of $M_1$, that is $p_1 = p_0$, then there exists a Markov chain $Y = (Y_n)_{n=1}^N$ and a constant $k \in \mathbb{R}$ such that

$$\frac{\mathbb{P}(W_{X,E} = w)}{\mathbb{P}(W_{Y,\tilde{E}} = -w + k)} = e^{\beta(w - \Delta F)} \quad \forall w \in \text{supp}(\mathbb{P}(W_{X,E})), \tag{18}$$

where $E = (E_n)_{n=0}^N$ and $\tilde{E} = (\tilde{E}_n)_{n=0}^N$ are families of energies of $X$ and $Y$, respectively, and $\text{supp}(\mathbb{P}(W_{X,E}))$ denotes the support of the probability distribution of $W_{X,E}$, that is the values that can be taken by $W_{X,E}$ with non-zero probability.

**Proof.** Let $Y = (Y_n)_{n=0}^N$ be the Markov chain defined in the proof of Theorem 1. Note that, for any family of energies $E$ of $X$, there exists a constant $c$ such that $E_1 = E_0 + c$ by (2), since $p_0$ is the stationary distribution of $M_1$. Given some $w \in W_{X,E}(S^{N+1})$, we have

$$\mathbb{P}(W_{X,E} = w) = \sum_{x \in W_{X,E}(w)} \mathbb{P}( (X_0, \ldots, X_N) = x )$$

$$= \sum_{x \in W_{X,E}(w)} \mathbb{P}( (Y_0, \ldots, Y_N) = x^R )$$

$$= e^{\beta(w - \Delta F)} \sum_{x \in W_{X,E}(w)} \mathbb{P}( (Y_0, \ldots, Y_N) = x^R )$$

$$= e^{\beta(w - \Delta F)} \sum_{x \in S^{N+1}; W_{Y,\tilde{E}}(x^R) = -w + k} \mathbb{P}( (Y_0, \ldots, Y_N) = x^R )$$

$$= e^{\beta(w - \Delta F)} \mathbb{P}(W_{Y,\tilde{E}} = -w + k),$$
where we applied (14) in (i) and Proposition 2 plus the fact that $E_1 = E_0 + c$ in (ii). To obtain (18), it remains to show that $\mathbb{P}(W_{Y,E} = -w + k) > 0$ for all $w \in \text{supp}(\mathbb{P}(W_{X,E}))$. By definition, there exists some $x \in S^{N+1}$ such that $W_{X,E}(x) = w$ and $\mathbb{P}(X = x) > 0$. By Proposition 2, $W_{Y,E}(x^{R}) = -w + k$. Since $\mathbb{P}(X = x) > 0$ and the entries of the (unique) stationary distributions of $X$ are also non-zero by Lemma 1, we can use (11) plus the Markov property for both $X$ and $Y$ to show $\mathbb{P}(Y = x^{R}) > 0$, implying $\mathbb{P}(W_{Y,E} = -w + k) > 0$. 

As the special case when each transition matrix of $X$ in Theorem 2 satisfies detailed balance, we obtain Crooks’ fluctuation theorem modified by the additional assumption of $p_0 = p_1$.

**Corollary 2** (Crooks’ fluctuation theorem for Markov chains). If all transition matrices $(M_n)_{n=0}^{N-1}$ of the Markov chain $X = (X_n)_{n=0}^{N}$ in Theorem 2 satisfy detailed balance, then (18) holds with $k = 0$ and $Y$ being the time reversal of $X$, that is the Markov chain with initial distribution $p_N$ and transition matrices $(M_{N−(n−1)})_{n=1}^{N}$.

**Proof.** As can be seen from the proof of Theorem 2 and Corollary 1, in the case of detailed balance, we can choose $Y$ to be the time reversal of $X$. Moreover, the origin of the constant $k$ in Theorem 2 is Equation (17). By Corollary 1, this constant can be set to zero if the transition matrices of $X$ satisfy detailed balance.

Notice, in most of the literature on Crooks’ fluctuation theorem, one writes $\mathbb{P}^{F}(W)$ for the probability $\mathbb{P}(W_{X,E})$ of the work along the so-called forward process $X$ and $\mathbb{P}^{B}(W)$ for the probability $\mathbb{P}(W_{Y,E})$ of the work along the so-called backward process $Y$ (the time reversal of $X$), so that, by Corollary 2, under detailed balance, Equation (18) reads

$$\frac{\mathbb{P}^{F}(W = w)}{\mathbb{P}^{B}(W = -w)} = e^{\beta(w - \Delta f)}. \tag{19}$$

The condition that $p_0$ is the stationary distribution of $M_1$ is not necessary in Crooks’ original work [7,8,15,17]. Nonetheless, it is fundamental in our approach: Crooks’ fluctuation theorem can even be false for every work value if $p_0$ is not the stationary distribution of $M_1$, as we show in Proposition 3.

**Proposition 3.** If $p_0$ is not the stationary distribution of $M_1$, then there exist Markov chains where Crooks’ fluctuation theorem is false everywhere, despite the other assumptions in Theorem 2 and Corollary 2 being fulfilled.

**Proof.** We consider a state space with three components $S = \{A, B, C\}$, a Markov chain with two steps $X = (X_0, X_1)$ and $a, b, c > 0$, where $b < c$ and $a + b + c = 1$. We take $E_0$ as the energy function associated with $X_0$, where $E_0(A) = \log \frac{1}{3}$, $E_0(B) = \log \frac{1}{2}$ and $E_0(C) = \log \frac{1}{3}$, and $E_1$ as the energy function associated with $X_1$, where $E_1(A) = E_0(A)$, $E_1(B) = E_0(C)$ and $E_1(C) = E_0(B)$. Taking $\beta = 1$, we obtain both free energies $F_1$ and $F_0$ equal to one. Notice that we have $p_0 = (a, b, c)$, with non-zero entries, and $p_1 = (a, c, b)$. Notice, also, that $W_X(S^2) = \{w_A, w_B, w_C\}$, where $w_C := E_1(C) - E_0(C) > 0$, $w_B := E_1(B) - E_0(B) < 0$, and $w_A := E_1(A) - E_0(A) = 0$. We fix $\tilde{E}_1 = \tilde{E}_0 = E_1$. We can easily see that (19) is not defined for $w = w_B, w_C$ and that it is false, although defined, for $w = w_A$. We have $W_Y(x) = 0 \ \forall x \in S^2$, since $p_1$ is both the starting distribution of the time reversal of $X$ and the stationary distribution of its only transition matrix. Thus, we have $\mathbb{P}^{F}(W = w_C) = \mathbb{P}(X_0 = C) = c > 0$ and $\mathbb{P}^{B}(W = -w_C) = 0$, which means (19) is not defined at $W = w_C$. We obtain, analogously, that it is not defined for $W = w_B$. For $W = w_A$, we have $\mathbb{P}^{B}(W = -w_A) = 1$ and

$$\mathbb{P}^{F}(W = w_A) = \mathbb{P}(X_0 = A) = a < 1 = e^{\beta w_A}.$$
which means (19) is defined and false there. Although the argument is independent of the transition matrix for $X$, fix

$$M_1 := \begin{pmatrix} a & a & a \\ c & c & c \\ b & b & b \end{pmatrix}$$

for completeness, since it has non-zero entries and fulfills detailed balance with respect to $p_1$. □

5. Discussion: Application to Decision-Making

The bridge connecting thermodynamics and decision-making is an optimization principle directly inspired by the maximum entropy principle [27,30,37,38]. In particular, given a finite set of possible choices $S$, the optimal behaviour is given by the distribution $p \in \mathbb{P}_S$ that optimally trades utility and uncertainty according to the following optimization principle:

$$p = \arg \max_{q \in \mathbb{P}_S} \left\{ H(q) \mid \mathbb{E}_q[U] \geq U_0 \right\}, \quad (20)$$

where $\mathbb{P}_S$ is the set of probability distributions over $S$, $H$ is the Shannon entropy, $\mathbb{E}_q[\cdot]$ denotes the expected value over $q \in \mathbb{P}_S$, and $U : S \to \mathbb{R}$ is a utility function, that is a function that assigns larger values to options in $S$ that are preferred by the decision-maker. Note that the main difference between (20) and the maximum entropy principle is the substitution of an energy function $E : S \to \mathbb{R}$ by a utility $U$, which behaves as a negative energy function $U = -E$ (in the sense that it is the force that opposes uncertainty). As a result of their similarity, both principles yield the same result, namely the Boltzmann distribution:

$$p(x) = \frac{1}{Z} e^{\beta U(x)} = \frac{1}{Z} e^{-\beta E(x)} \quad \forall x \in S,$$

where $\beta$ is a trade-off parameter between uncertainty and utility/energy and $Z$ is a normalization constant.

The analogy between thermodynamics and decision-making can be taken further by considering not only their optimal distribution, but how they transition between different distributions in the path towards the optimal one. Here, the notion of uncertainty is useful again, although this time, it is relative to the optimal distribution $p$. More specifically, we can think of the transitions the decision-maker undergoes as being driven by the reduction of uncertainty with respect to the optimal distribution $p$, which can be modelled by the dual of $p$-majorization [39]. In this approach, given $q, q' \in \mathbb{P}_S$, the decision-maker transitions from $q$ to $q'$, which we denote by $q \preceq_p q'$, if $q'$ is closer to $p$ than $q$ (see [39] for a rigorous definition of the dual of $\preceq_p$ and, hence, of closer). For us, the important fact about $\preceq_p$ is that, as it turns out ([40], Theorem 2), the transitions that are allowed by $\preceq_p$ are precisely the ones that result from applying a transition matrix that has $p$ as a stationary distribution. More precisely,

$$q \preceq_p q' \iff q' = M_p q,$$

where $M_p$ is a matrix whose rows are normalized and fulfills $M_p p = p$, that is a stochastic matrix for which $p \in \mathbb{P}_S$ is a stationary distribution [33]. Importantly, the transition matrix assumption is common in the study of both thermodynamic and decision-making systems [8,15].

In our current study, the situation we have in mind is that of a decision-maker that has to make a sequence of decisions under varying environmental conditions. We model this as a stochastic process that behaves like a Markov chain and introduce, starting from a Markov chain, the thermodynamic tools we use to describe it: energy, partition function, free energy, work, heat, and dissipated work. The behaviour of the decision-maker then corresponds to a decision vector $(x_0, x_1, \ldots, x_n)$ collected over $n$ potentially different environments. In this most general decision-making scenario, where the environment is changing over time, the optimal distribution $p$ is changing as well. Thus, we can regard the decision-making
process as a sequence of transition matrices $M_p$, where each $M_p$ corresponds to a particular environment with optimal response $p$. We could imagine, for example, a gradient descent learner that would converge to $p$ for any given environment, presuming we allow for sufficient gradient update steps. Otherwise, the gradient learner, or any other optimization-based decision-making agent (e.g., following a Metropolis–Hastings optimization scheme), would lag behind the environmental changes and the environment would outpace the learner. In this general decision-making scenario, we can then study the relation between the optimal behaviour and the non-optimal one by fluctuation theorems [1–3] like Jarzynski’s equality [4–6] and Crooks’ fluctuation theorem [7,8]. While we focus on these two fluctuation theorems in our current study, similar arguments may be suitable to transfer other fluctuation theorems that have been considered in the thermodynamic literature (see, for example, [1]) to a decision-making scenario.

Jarzynski’s equality in decision-making:

Although the strong requirements in Lemma 2 were used in [8] to derive Jarzynski’s equality through (15) and were assumed in the only approach we know for it in decision-making [32], weaker assumptions that do not involve the time reversal of $X$ are sufficient (cf. Theorem 1). The same properties have been used to derive Jarzynski’s equality following a different method in [4].

Example application: Jarzynski’s theorem:

In a decision scenario, the energy becomes a loss function that a decision-maker is trying to minimize. If this loss function changes over time, we can conceptually distinguish changes in loss that are induced externally by changes in the environment (e.g., given data), from changes in loss due to internal adaptation when a learning system changes its parameter settings. The externally induced changes in loss correspond to the physical concept of work and drive the adaptation process. Hence, we can consider the decision-theoretic equivalent of physical work as a driving signal: the (negative) surprise experienced by the decision-maker, given that it adds the (negative) surprise that he/she experiences at each step (which can be quantified by the difference in energy/utility evaluated at the decision-maker’s state when the environment changes) [32,41]. With this in mind, we can use Jarzynski’s equality to obtain a bound on the decision-maker’s expected surprise. In particular, by applying Jensens’ inequality on Jarzynski’s equality, one obtains

$$ \Delta F \leq \mathbb{E}[W(X)]. \tag{22} $$

(Note that this is a version of the second law of thermodynamics [2].) Hence, (22) provides a bound on the expected surprise. While a similar bound has been previously pointed out for decision-making systems [32], here, we re-derive it under a novel energy protocol and with weaker assumptions regarding time reversibility.

Crooks’ theorem in decision-making:

Even though the assumption that $p_0$ must be the stationary distribution of $M_1$ seems to restrict the applicability of Crooks’ fluctuation theorem in our decision-theoretic setup when compared to the usual thermodynamics one (see Corollary 2 and Proposition 3), it is actually not an issue from an experimental point of view when the Markov chains correspond to thermodynamic processes. This is the case because of the way one is able to sample from a Boltzmann distribution given a thermodynamic system. One of the assumptions in Corollary 2 is that $X_0$ should follow such a distribution for $E_0$. For this to be fulfilled, one needs to wait until the system relaxes to such a state. Because of that, one can think of any trajectory as having an additional point that was also sampled from the Boltzmann distribution for $E_0$. Thus, the assumption that $p_0$ is the equilibrium distribution of $M_1$ is always fulfilled, and the experimental range of validity of Crooks’ fluctuation theorem in our setup remains equal to the one in nonequilibrium thermodynamics [8,15].
In particular, the new constraint is fulfilled in previous experimental setups supporting the theorem (see, for example, [9] or [11]).

Example application: Crooks’ theorem:

Hysteresis is a well-known effect that takes place in some physical systems and refers to the difference in the system’s behaviour when interacting with a series of environments compared to its response when facing the same conditions in reversed order [2]. The same idea also applies to decision-making systems. In fact, hysteresis has been reported in both simulations of decision-making systems [32], as well as in biological decision-makers recorded experimentally [41,42]. Given that it refers to the difference between decisions when the order in which the environments are presented is reversed, (15) and (19) constitute quantitative measures of hysteresis. In particular, Reference [41] used this measure successfully to quantify hysteresis in human sensorimotor adaptation, where human learners had to solve a simple motor coordination task in a dynamic environment with changing visuomotor mappings. While a simple Markov model proved adequate to model sensorimotor adaptation, it should be noted that more complex learning scenarios involving long-term memory and abstraction would not be captured by such a simple model.

Detailed balance:

Detailed balance is not required neither for Jarzynski’s equality (Theorem 1), nor for the more general form of Crooks’ fluctuation theorem we presented in Theorem 2. It is, however, required in order to choose $Y$ to be the time reversal of $X$, which leads to Crooks’ fluctuation theorem in Corollary 2.

While the definition of detailed balance (1) we adopted here is standard in the Markov chain literature, there is some ambiguity regarding its use in thermodynamics, where it has, at least, two more meanings. It is used both for the weaker condition that the Boltzmann distribution $p_n$ is a stationary distribution of $M_n$ for $1 \leq n \leq N$ [4] and as a synonym of microscopic reversibility [8,15]. Although we have shown microscopic reversibility and detailed balance are indeed equivalent under some conditions (see Lemma 2), we followed its definition in [15], which is not the only one in the literature (see [35,43,44]).

Notice what is called a stationary distribution in the literature on Markov chains is referred to as a nonequilibrium steady state in thermodynamics [45]. In order for it to be an equilibrium state, it needs to fulfil detailed balance (1) with respect to the transition matrix in question. Notice, also, detailed balance is not fulfilled in several applications of nonequilibrium thermodynamics throughout physics [46] and biology [47].

Continuous-time Markov chains:

Notice, in the case that we have a continuous-time Markov chain, work becomes an integral, where the integrand for $X$ at $x$ and the one for $Y$ at $x^R$ differ, aside from the sign, in a single point. Thus, work is odd under time reversal and the assumption that $p_0 = p_1$ can be dropped in both Theorem 2 and Corollary 2. However, the technical tools required to show Crooks’ fluctuation theorem or Jarzynski’s equality are technically more involved in the continuous-time case, as one can see in [36].

Continuous state space:

In case the state space is continuous, the results can be derived in a similar fashion. What we ought to notice is that, in this scenario, the role of the transition matrices is played by the densities of the Markov kernels (see for example [48]). These densities allow us to write conditions such as detailed balance analogously to how we do in the discrete case. In the case of Jarzynski’s equality for a Markov chain on a continuous state space $X$, one can see that the result follows like the one with a discrete state space. To convince ourselves, the only thing to take into account is the substitution of the sum in the expected value by the integral and that of the probability distribution by the density of the Markov kernel. Then, following the proof of Theorem 1, we can define a stochastic process $Y$
whose density Markov kernels are defined through the stationary distributions and density Markov kernels of $X$, in analogy to how we defined them in Theorem 1. The rest follow exactly in the same fashion. Crooks’ fluctuation theorem requires a longer explanation, but, essentially, follows from the same considerations.

6. Conclusions

In this paper, we investigated the potential of thermodynamic fluctuation theorems to serve probabilistic laws of decision-making. In particular, we derived two thermodynamic fluctuation theorems, Jarzynski’s equality and Crooks’ theorem, in the context of general Markov chains $X$. We started by defining several thermodynamic concepts for Markov chains and discussing how these definitions do not correspond in general to the ones used in thermodynamics. Right after that, we derived Jarzynski’s equality in Theorem 1 without any assumption involving the time reversal of $X$. Thus, we improved on the previous attempt to derive it in the context of decision-making [32], which was based on the physical conventions pioneered in [8]. Regarding Crooks’ fluctuation theorem, we showed in Theorem 2, Corollary 2, and Proposition 3 that, in our decision-theoretic setup, it requires the additional assumption that the initial distribution of $X$ must be the stationary distribution of its first transition matrix. This results from the fact that our notion of work is inherent to Markov chains, which contrasts with the definition used in previous derivations [8,15,17], where the work along the forward and backward paths was calculated in a different way for physical reasons. Instead, we calculated the work along the two paths in the same way, in order for the quantities involved in the final result to have an interpretation that is relevant for decision-making.

Author Contributions: Conceptualization, P.H.; methodology, P.H.; validation, D.A.B. and S.G.; formal analysis, P.H.; resources, D.A.B. and S.G.; writing—original draft preparation, P.H.; writing—review and editing, D.A.B. and S.G.; visualization, P.H. and S.G.; supervision, D.A.B. and S.G.; project administration, D.A.B. and S.G.; funding acquisition, D.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Research Council, Grant Number ERC-StG-2015-ERC, Project ID: 678082, “BRISC: Bounded Rationality in Sensorimotor Coordination”.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of the data; in the writing of the manuscript; nor in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

| Abbreviation | Description |
|--------------|-------------|
| MDPI         | Multidisciplinary Digital Publishing Institute |
| DOAJ         | Directory of open access journals |
| TLA          | Three letter acronym |
| LD           | Linear dichroism |

Appendix A

Here, we show two results relating a Markov chain $X$ to the Markov chain $Y$ defined by $X$ as in Theorem 1. First, in Lemma A1, we relate the irreducibility of transition matrices for $X$ with that of $Y$.

Lemma A1. If $X = (X_n)_{n=0}^N$ is a Markov chain on a finite state space $S$ whose transition matrices are irreducible, then the transition matrices of the Markov chain $Y = (Y_n)_{n=0}^N$ defined in Theorem 1 are irreducible.
Proof. We denote by $\{(M_n)_{n=1}^{N}\}$ the transition matrices of $X$ and by $\{(\tilde{M}_n)_{n=1}^{N}\}$ the ones of $Y$. Consider $x, y \in S$ and $\tilde{M}_n$ for some $1 \leq n \leq N$. Since $M_{N+1-n}$ is irreducible, there exist $m \in \mathbb{N}$ and $x_0, x_1, \ldots, x_m$ where $x_0 = y$ and $x_m = x$ such that

$$P(X_m = x_m, \ldots, X_1 = x_1 | X_0 = x_0) = (M_{N+1-n})_{x_m,x_{m-1}} \cdots (M_{N+1-n})_{x_1,x_0} > 0.$$  \hspace{2cm} (A1)

Hence, we have

$$P(Y_m = x_m, \ldots, Y_1 = x_0 | Y_0 = x_m) \overset{(i)}{=} (\tilde{M}_n)_{x_m,x_{m-1}} \cdots (\tilde{M}_n)_{x_1,x_0} = (M_{N+1-n})_{x_m,x_{m-1}} \cdots (M_{N+1-n})_{x_1,x_0}> 0, \hspace{2cm} (A2)$$

where we applied the Markov property of $Y$ in (i), (11) in (ii), and (A1) plus the fact that, as $M_{N+1-n}$ is irreducible, $p_{N+1-n}(x) > 0 \forall x \in S$ by Lemma 1 in (iii).

The connection between the irreducibility of the transition matrices of $X$ and $Y$ in Lemma A1 results in a relation between their energy functions, which we prove in Lemma A2.

Lemma A2. If $X = (X_n)_{n=0}^{N}$ is a Markov chain on a finite state space $S$ whose initial distribution $p_0$ has non-zero entries and whose transition matrices $(M_n)_{n=1}^{N}$ are irreducible, $Y = (Y_n)_{n=0}^{N}$ is the Markov chain defined in Theorem 1, $E = (E_n)_{n=0}^{N}$ is a family of energy functions of $X$, and $\tilde{E} = (\tilde{E}_n)_{n=0}^{N}$ is a family of energy functions of $Y$, then there exist constants $(k_n)_{n=0}^{N}$ such that

$$\tilde{E}_n = E_n + k_n$$

for $0 \leq n \leq N - 1$. \hspace{2cm} (A3)

Proof. Since all transition matrices of $X$ are irreducible and the same holds for $Y$ by Lemma A1, we can apply Lemma 1 to $Y$ and obtain that each of its transition matrices has a unique stationary distribution composed of non-zero entries. Thus, energy is well defined, up to a constant, for $Y$. Since $Y_0$ follows $p_N$ by definition, we have automatically there exists a constant $k_0$ such that $\tilde{E}_0 = E_N + k_0$. To obtain (A2), notice that we have $\tilde{M}_{n+1} p_{N-n} = p_{N-n}$ for $0 \leq n \leq N - 1$ by the definition of $(\tilde{M}_n)_{n=0}^N$:

$$\overline{(\tilde{M}_{n+1} p_{N-n})}(y) = \sum_{x \in S} ((\tilde{M}_{n+1})_{yx} p_{N-n}(x)) = \sum_{x \in S} (M_{N-n})_{xy} p_{N-n}(y) = p_{N-n}(y),$$

where we applied (11) in the second equality and the fact that $M_{N-n}$ is a stochastic matrix in the third. Thus, for $0 \leq n \leq N - 1$, $p_{N-n}$ is the unique stationary distribution of $\tilde{M}_{n+1}$, and there exists a constant $k_{n+1}$ such that $\tilde{E}_{n+1} = E_{N-n} + k_{n+1}$.

References
1. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. Rep. Prog. Phys. 2012, 75, 126001. [CrossRef] [PubMed]
2. Jarzynski, C. Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale. Annu. Rev. Condens. Matter Phys. 2011, 2, 329–351. [CrossRef]
3. Jarzynski, C. Nonequilibrium work relations: Foundations and applications. Eur. Phys. J. B 2008, 64, 331–340. [CrossRef]
4. Jarzynski, C. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. Phys. Rev. E 1997, 56, 5018. [CrossRef]
5. Jarzynski, C. Nonequilibrium work theorem for a system strongly coupled to a thermal environment. J. Stat. Mech. Theory Exp. 2004, 2004, P09005. [CrossRef]
6. Jarzynski, C. Nonequilibrium equality for free energy differences. Phys. Rev. Lett. 1997, 78, 2690. [CrossRef]
7. Crooks, G.E. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. Phys. Rev. E 1999, 60, 2721. [CrossRef]
8. Crooks, G.E. Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems. J. Stat. Phys. 1998, 90, 1481–1487. [CrossRef]
9. Collin, D.; Ritort, F.; Jarzynski, C.; Smith, S.B.; Tinoco, I.; Bustamante, C. Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. Nature 2005, 437, 231–234. [CrossRef]
10. Lipphardt, J.; Dumont, S.; Smith, S.B.; Tinoco, I.; Bustamante, C. Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski’s equality. Science 2002, 296, 1832–1835. [CrossRef]
11. Saira, O.P.; Yoon, Y.; Tanttu, T.; Möttönen, M.; Averin, D.; Pekola, J.P. Test of the Jarzynski and Crooks fluctuation relations in an electronic system. Phys. Rev. Lett. 2012, 109, 180601. [CrossRef] [PubMed]
12. Douarche, F.; Ciliberto, S.; Petroyan, A.; Rabbiosi, I. An experimental test of the Jarzynski equality in a mechanical system. Europhys. Lett. 2005, 70, 593. [CrossRef]
13. An, S.; Zhang, J.N.; Um, M.; Lv, D.; Lu, Y.; Zhang, J.; Yin, Z.Q.; Quan, H.; Kim, K. Experimental test of the quantum Jarzynski equality with a trapped-ion system. Nat. Phys. 2015, 11, 193–199. [CrossRef]
14. Smith, A.; Lu, Y.; An, S.; Zhang, X.; Zhang, J.N.; Gong, Z.; Quan, H.; Jarzynski, C.; Kim, K. Verification of the quantum nonequilibrium work relation in the presence of decoherence. New J. Phys. 2018, 20, 013008. [CrossRef]
15. Crooks, G.E. Path-ensemble averages in systems driven far from equilibrium. Phys. Rev. E 2000, 61, 2361. [CrossRef]
16. Buscemi, F.; Scarani, V. Fluctuation theorems from Bayesian retrodiction. Phys. Rev. E 2021, 103, 052111. [CrossRef] [PubMed]
17. Crooks, G.E. Excursions in Statistical Dynamics; University of California, Berkeley: Berkeley, CA, USA, 1999.
18. Goldt, S.; Seifert, U. Stochastic thermodynamics of learning. Phys. Rev. Lett. 2017, 118, 010601. [CrossRef]
19. Perunov, N.; Marsland, R.A.; England, J.L. Statistical physics of adaptation. Phys. Rev. X 2016, 6, 021036. [CrossRef]
20. England, J.L. Dissipative adaptation in driven self-assembly. Nat. Nanotechnol. 2015, 10, 919–923. [CrossRef]
21. Still, S.; Sivak, D.A.; Bell, A.J.; Crooks, G.E. Thermodynamics of prediction. Phys. Rev. Lett. 2012, 109, 120604. [CrossRef]
22. Ortega, P.A.; Braun, D.A. Thermodynamics as a theory of decision-making with information-processing costs. Proc. R. Soc. A Math. Phys. Eng. Sci. 2013, 469, 20120683. [CrossRef]
23. Parr, T.; Da Costa, L.; Friston, K. Markov blankets, information geometry and stochastic thermodynamics. Philos. Trans. R. Soc. 2020, 378, 20190159. [CrossRef] [PubMed]
24. Da Costa, L.; Friston, K.; Heins, C.; Pavliotis, G.A. Bayesian mechanics for stationary processes. Proc. R. Soc. A 2021, 477, 20210518. [CrossRef] [PubMed]
25. Gottwald, S.; Braun, D.A. The two kinds of free energy and the Bayesian revolution. PLoS Comput. Biol. 2020, 16, 1–32. [CrossRef] [PubMed]
26. Boyd, A.B.; Crutchfield, J.P.; Gu, M. Thermodynamic machine learning through maximum work production. New J. Phys. 2022, 24, 083040. [CrossRef]
27. Wolpert, D.H. Information theory—the bridge connecting bounded rational game theory and statistical physics. In Complex Engineered Systems; Springer: Berlin/Heidelberg, Germany, 2006; pp. 262–290.
28. Tishby, N.; Polani, D. Information theory of decisions and actions. In Perception-Action Cycle; Springer: Berlin/Heidelberg, Germany, 2011; pp. 601–636.
29. Ortega, P.A.; Braun, D.A. Information, utility and bounded rationality. In Proceedings of the International Conference on Artificial General Intelligence, San Francisco, CA, USA, 15–18 October 2011; Springer: Berlin/Heidelberg, Germany, 2011; pp. 269–274.
30. Genewein, T.; Leibfried, F.; Grau-Moya, J.; Braun, D.A. Bounded rationality, abstraction, and hierarchical decision-making: An information-theoretic optimality principle. Front. Robot. AI 2015, 2, 27. [CrossRef]
31. Wolpert, D.H. The free energy requirements of biological organisms; implications for evolution. Entropy 2016, 18, 138. [CrossRef]
32. Grau-Moya, J.; Krüger, M.; Braun, D.A. Non-equilibrium relations for bounded rational decision-making in changing environments. Entropy 2018, 20, 1. [CrossRef]
33. Levin, D.A.; Peres, Y. Markov Chains and Mixing Times; American Mathematical Soc.: Providence, USA, 2017; Volume 107.
34. Yang, Y.J.; Qian, H. Unified formalism for entropy production and fluctuation relations. Phys. Rev. E 2020, 101, 022129. [CrossRef]
35. Cohen, E.; Mauzerall, D. A note on the Jarzynski equality. J. Stat. Mech. Theory Exp. 2004, 2004, P07006. [CrossRef]
36. Ge, H.; Qian, M. Generalized Jarzynski’s equality in inhomogeneous Markov chains. J. Math. Phys. 2007, 48, 053302. [CrossRef]
37. Jaynes, E.T. Information theory and statistical mechanics. Phys. Rev. 1957, 106, 620. [CrossRef]
38. Jaynes, E.T. Probability Theory: The Logic of Science; Cambridge University Press: Cambridge, UK, 2003.
39. Joe, H. Majorization and divergence. J. Math. Anal. Appl. 1990, 148, 287–305. [CrossRef]
40. Gottwald, S.; Braun, D.A. Bounded rational decision-making from elementary computations that reduce uncertainty. Entropy 2019, 21, 375. [CrossRef] [PubMed]
41. Hack, P.; Lindieg-Leon, C.; Gottwald, S.; Braun, D.A. Thermodynamic fluctuation theorems govern human sensorimotor learning. arXiv 2022, arXiv:2209.00941.
42. Turnham, E.J.; Braun, D.A.; Wolpert, D.M. Facilitation of learning induced by both random and gradual visuomotor task variation. J. Neurophysiol. 2012, 107, 1111–1122. [CrossRef] [PubMed]
43. Crooks, G.E. On thermodynamic and microscopic reversibility. J. Stat. Mech. Theory Exp. 2011, 2011, P07008. [CrossRef]
44. Tolman, R.C. The principle of microscopic reversibility. Proc. Natl. Acad. Sci. USA 1925, 11, 436. [CrossRef]
45. Zhang, X.J.; Qian, H.; Qian, M. Stochastic theory of nonequilibrium steady states and its applications. Part I. Phys. Rep. 2012, 510, 1–86. [CrossRef]
46. Tang, Y.; Yuan, R.; Chen, J.; Ao, P. Work relations connecting nonequilibrium steady states without detailed balance. Phys. Rev. E 2015, 91, 042108. [CrossRef]
47. Battle, C.; Broedersz, C.P.; Fakhri, N.; Geyer, V.F.; Howard, J.; Schmidt, C.F.; MacKintosh, F.C. Broken detailed balance at mesoscopic scales in active biological systems. Science 2016, 352, 604–607. [CrossRef] [PubMed]
48. Chib, S.; Greenberg, E. Understanding the Metropolis-Hastings algorithm. Am. Stat. 1995, 49, 327–335.