

Article

# Microscopically Reversible Pathways with Memory

Jose Ricardo Arias-Gonzalez 

Centro de Tecnologías Físicas, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain; ricardo.ariasgonzalez@upv.es or ricardo.ariasgonzalez@gmail.com

**Abstract:** Statistical mechanics is a physics theory that deals with ensembles of microstates of a system compatible with environmental constraints and that on average define a thermodynamic state. The evolution of a small system is normally subjected to changing constraints, as set by a protocol, and involves a stochastic dependence on previous events. Here, we generalize the dynamic trajectories described by a realization of a physical system without dissipation to include those in which the history of previous events is necessary to understand its future. This framework is then used to characterize the processes experienced by the stochastic system, as derived from ensemble averages over the available pathways. We find that the pathways that the system traces in the presence of a protocol entail different statistics from those in its absence and prove that both types of pathways are equivalent in the limit of independent events. Such equivalence implies that a thermodynamic system cannot evolve away from equilibrium in the absence of memory. These results are useful to interpret single-molecule experiments in biophysics and other fields in nanoscience, as well as an adequate platform to describe non-equilibrium processes.

**Keywords:** non-markovian; memory; pathway; stochastic; microscopic reversibility; statistical mechanics; Information Theory; phase space; single molecule; non-equilibrium



**Citation:** Arias-Gonzalez, J.R. Microscopically Reversible Pathways with Memory. *Mathematics* **2021**, *9*, 127. <https://doi.org/10.3390/math9020127>

Received: 16 May 2020  
Accepted: 6 January 2021  
Published: 8 January 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Reversibility refers to quasistatic processes that invert isentropically. Such processes involve a sufficiently slow dynamics to prevent heat flows; more in depth, they take place through a timeless succession of states along which there is no energy dissipation. Reversible processes are normally analyzed by equilibrium statistics: a so-called partition function is used to characterize the thermodynamic states described by a system by pondering all possible configurations of microstates compatible with each state [1]. For a system in which fate is not dependent on either the past or the present, like many macroscopic systems, both equilibrium and frictionless quasistatic processes can be examined through the same mathematical framework because such system is able to explore all the possible configurations in an indefinite time. In this paper, we introduce for the first time memory effects of full extent through an ab initio theory that reveals differential statistics.

Phase space is a long-standing concept to analyze the dynamics of physical systems, transcending the configuration space of position coordinates. From the classical to the quantum realm, the system—a point comprising a set of generalized position and momentum coordinates—traces a pathway (or trajectory) described by the time-ordered union of points in this space [2]. Phase-space trajectories determine the evolution of the thermodynamic microstates of similarly constructed systems under the same protocol through so-called processes, which involve ensemble averages over such trajectories [3]. The principle of microscopic reversibility asserts that the probability of any pathway of a system realization through phase space is equivalent to that of the time reversed pathway. While individual pathways are reversible, processes may not, thus entailing energy dissipation [4]. We herein apply the principle of microscopic reversibility to systems that recall not only the present but also the past by an unprecedented mathematical treatment, in which phase space trajectories are analyzed exactly.

Microscopic reversibility and phase space underlie the theory of dynamical systems in physics but they usually remain at the conceptual level. With the advent of nanoscience, the urge for understanding so-called *small systems* [5]—those for which energy exchanges are smaller or similar to the thermal level—is pushing these concepts to the direct applications. A small system typically measures from a few to hundreds of nanometers and contains from one to some thousands of atoms or molecules. Due to the scale nature, a small system is deemed to progress along only one of the possible trajectories by a certain protocol. Its energy balance in terms of work and heat as a function of the temperature has to be evaluated at the single phase-space pathway level. In addition, the protocol by which this system evolves has to be considered when ensemble-average thermodynamics are addressed, especially when memory effects are present.

Biological systems have become central players in this urge for understanding small systems. For many processes that take place in the cell, the study of each molecular trajectory individually is crucial for a complete comprehension of the role of fluctuations [6]. Biophysical processes have traditionally been analyzed by bulk (ensemble-average) strategies but the importance of tackling them at the single-molecule level and at the single-chemical reaction has raised much both scientific and technological interest in the last twenty years [7,8]. In this regard, replication, transcription, and translation in molecular biology, just to name a few, are processes in which thorough investigation requires single-molecule approaches [9]: nucleotides or amino acids are incorporated sequentially by a protein in which operation determines a certain copying direction and a mechanism, both of them responsible for chain stability and information fidelity. Sustained by physical interactions, these biological processes encompass naturally memory effects because nucleotide and amino acid polymers carry genetic meaning. Protein folding is another *in singulo* process that showcases the firm link between physical interactions and memory, and how this link brings thermodynamic consequences to the structural and functional fate of a polypeptide [10]. Certainly, both the synthesis protocol and the amino acid sequence stochastically guide protein folding dynamics through preferential phase-space trajectories [11]. Another biological example in which memory is a key ingredient is that of *learning*, in which organisms collect information to perform complex control tasks [12]. In general, small systems may present mutual and internal correlations due to physical interactions [13,14], which, steered by a protocol, make up their biography, a history of events that influence their future.

To analyze theoretically small systems with memory, thermodynamics uses stochastic variables and statistical mechanics [15,16], often considering evolutions where only the immediate present is necessary to inspect the future (known as memoryless or Markovian evolutions). Tsallis entropy generalization enabled non-extensive analyses of physical systems that keep memory on previous events, that is, of systems that recall not only the present but also the past events (non-Markovian evolutions) [17]; this generalization in turn compacts non-equilibrium dynamics elegantly. Open system formalisms, including spectral analysis, introduce memory effects through non-Markovian approaches to address the environment surrounding the system under study [18–22]. Complete memory effects have been considered in the study of spatial chains made up of physical subunits, including those with symbolic meaning, to access non-equilibrium dynamics [23] and information management in nanoscale systems [24]. They have also been taken into account to analyze abstract strings of symbols for the paradigm of communication [25]. However, full memory effects in temporal chains of events have never been treated from an exact perspective.

In the following, we develop a general framework for reversible pathways with memory to address the evolution of physical systems. Although not restricted to them, we will discuss this framework in the context of small systems, where strong interest resides in their stochastic behavior and in the wealth of current applications. We analyze microscopic pathways exactly, tracing explicitly the memory that a physical system retains along the pathway it describes, and we derive consistency constraints. Our theory discerns between protocol-driven and equilibrium pathways. We analyze their associated dynamics at the

level of the ensemble average, as taken over the pathways through which the system can evolve. We end up by illustrating our theory in the context of computing biomolecular systems, where memory effects are inherent to their thermodynamic modeling.

## 2. Analysis

### 2.1. Concepts and Terminology

The evolution of a small system is strongly affected by the thermal fluctuations. Under the action of external perturbations, the statistical characterization of any set of ensembles of the same system does not determine a unique thermodynamic description because different perturbations drive the system preferentially through different sets of pathways. Then, it is necessary to introduce the protocol,  $\lambda$ , which is a collection of control parameters that describes the mechanism and external constraints that act over the system [5]. Given an initial and a final microstate of the system at time instants  $t = t_0$  and  $t = t_f$ , respectively, every pathway (or trajectory) that connects them in phase space will be specified by a temporal sequence,  $\nu = \{x_{t_0}, \dots, x_t, \dots, x_{t_f}\}$ , of stochastic events  $X_t = x_t$  ( $x_t \in \mathcal{X}$ , being  $\mathcal{X}$  the domain of the random variables with cardinality  $|\mathcal{X}|$ , and  $t = t_0, \dots, t_f$ ) under protocol  $\lambda$ .

We will use the term *event* for  $x_t$  and *microstate* for the event at time  $t$  plus the ordered sequence of previous events that the system recalls. We will use the term *state* for the ensemble-average over the pathways that the system can follow until time  $t$  driven by protocol  $\lambda$ . This is the notion of macrostate in statistical mechanics [1], but since it is misleading in a nanoscale context, we prefer using *state*. In the limit in which the memory of a system at time  $t$  extends to its complete history,  $\{x_{t_0}, \dots, x_{t-1}\}$ , microstates and pathways are equivalent. In the limit of independent events, events and microstates are equivalent and the term pathway is not necessary since microstates do not depend on how they have been reached by the system. A state is solely determined by the protocol by which events have been driven. For the sake of conceptual clarity, we will also distinguish between *process* and pathway, the former being obtained under averages over ensembles of the latter. The principle of microscopic reversibility, expressed at the introduction, stands at the level of the pathway.

We will tackle sequences  $\nu$  as directional, stochastic chains with memory [26] in the time domain. Events experienced by the system at each time instant,  $x_t$ , involve a set of stochastic variables according to the degrees of freedom of the system:

$$x_t \equiv (q_1(t), \dots, q_D(t); p_1(t), \dots, p_D(t)), \quad (1)$$

or, abbreviated,  $x_t \equiv (q_h(t), p_h(t))$ ,  $h = 1, \dots, D$ , where  $D$  is the number of degrees of freedom and  $q_h$  and  $p_h$  are generalized space and momentum coordinates, respectively.

The microstate of the system at time  $t$  is not only determined by the value of  $X_t$  but also by how  $x_t$  has been reached. Therefore, the energy of each microstate is a function of the previous events, namely

$$E(x_t; x_{t-1}, \dots, x_{t_0}), \quad (2)$$

which account for the memory, i.e., the relative interactions of the present event,  $x_t$ , with its previous ones,  $\{x_{t-1}, \dots, x_{t_0}\}$ . As noted, we will represent the random variables corresponding to the memory of a microstate in a thermodynamic function by a semicolon followed by the variable values with decreasing subscript order.

In addition to the energy, central to the thermodynamic description of a physical system is the entropy, which characterizes the level of reversibility in the system's evolution by pondering the level of uncertainty in the stochastic events along phase-space pathways. In this regard, it is connected to the entropy defined in information theory, as reflected in the mathematical similarity between the Gibbs and the Shannon entropies (appearing later on) [27].

### 2.2. Protocol-Driven Pathways

The probability of a microstate, according to a canonical formalism, is:

$$p^{(\lambda)}(x_t|x_{t-1}, \dots, x_{t_0}) = \frac{e^{-\beta E(x_t;x_{t-1}, \dots, x_{t_0})}}{\sum_{x'_t} e^{-\beta E(x'_t;x_{t-1}, \dots, x_{t_0})}} = \frac{e^{-\beta E(x_t;x_{t-1}, \dots, x_{t_0})}}{Z^{(\lambda)}(;x_{t-1}, \dots, x_{t_0})}, \tag{3}$$

with  $\beta = 1/(kT)$  ( $T$  the absolute temperature and  $k$  the Boltzmann constant),  $\sum_{x_t} p^{(\lambda)}(x_t|x_{t-1}, \dots, x_{t_0}) = 1$  and the partition function given by

$$Z^{(\lambda)}(;x_{t-1}, \dots, x_{t_0}) \equiv \sum_{x'_t} e^{-\beta E(x'_t;x_{t-1}, \dots, x_{t_0})}. \tag{4}$$

Note that  $\sum_{x'_t} \equiv \sum_{q'_1(t), \dots, q'_D(t)} \sum_{p'_1(t), \dots, p'_D(t)}$ . We use the prime symbol in  $x_t$  to emphasize that the sum does not affect the previous events,  $x_{t_0}, \dots, x_{t-1}$ , since they are fixed at present time  $t$ . For completeness, consider that for  $t = t_0$ ,  $p^{(\lambda)}(x_{t_0}) = \exp(-\beta E(x_{t_0})) / \sum_{x'_{t_0}} \exp(-\beta E(x'_{t_0}))$ .  $p^{(\lambda)}(x_t|x_{t-1}, \dots, x_{t_0})$  is the probability that random variable  $X$  (which, as explained in the previous subsection, comprises a set of generalized position and momentum coordinates, Equation (1)) takes value  $x$  at time  $t$  provided that the values of this random variable for the previous instants,  $t - 1, \dots, t_0$ , are  $X_{t-1} = x_{t-1}, \dots, X_{t_0} = x_{t_0}$ , respectively.

The probability of a pathway is:

$$\begin{aligned} p_v^{(\lambda)} &\equiv \Pr^{(\lambda)}\{X_{t_0} = x_{t_0}, \dots, X_{t_f} = x_{t_f}\} = p^{(\lambda)}(x_{t_0}, \dots, x_{t_f}) \\ &= p^{(\lambda)}(x_{t_0})p^{(\lambda)}(x_{t_0+1}|x_{t_0}) \cdots p^{(\lambda)}(x_{t_f}|x_{t_f-1}, \dots, x_{t_0}) = \prod_{t=t_0}^{t_f} p^{(\lambda)}(x_t|x_{t-1}, \dots, x_{t_0}). \end{aligned} \tag{5}$$

$p^{(\lambda)}(x_{t_0}, \dots, x_{t_f})$  is the probability that a particular sequence of microstates take place, that is, that random variable  $X$  takes the specific sequence of values  $X_{t_0} = x_{t_0}, \dots, X_{t_f} = x_{t_f}$  from  $t_0$  to  $t_f$ . According to probability theory [28], it is obtained as a product of the conditional probabilities  $p^{(\lambda)}(x_t|x_{t-1}, \dots, x_{t_0})$ , for  $t = t_0$  to  $t = t_f$ . This probability can be expressed in a more compact form [26]:

$$p_v^{(\lambda)} = \frac{e^{-\beta E_v}}{Z_v^{(\lambda)}}, \tag{6}$$

such that  $\sum_{v=1}^N p_v^{(\lambda)} \equiv \sum_{x_{t_0}, \dots, x_{t_f}} p^{(\lambda)}(x_{t_0}, \dots, x_{t_f}) = 1$ . In this equation,

$$E_v \equiv E(x_{t_0}, \dots, x_{t_f}) = \sum_{t=t_0}^{t_f} E(x_t; x_{t-1}, \dots, x_{t_0}), \tag{7}$$

and  $Z_v^{(\lambda)}$  is the sequence-dependent partition function,

$$Z_v^{(\lambda)} \equiv \prod_{t=t_0}^{t_f} Z^{(\lambda)}(;x_{t-1}, \dots, x_{t_0}) \tag{8}$$

$$= \sum_{x'_{t_0}, \dots, x'_{t_f}} \exp\left(-\beta \sum_{t=t_0}^{t_f} E(x'_t; x_{t-1}, \dots, x_{t_0})\right) \equiv \sum_{v'(\lambda)=1}^N \exp(-\beta E_{v'}), \tag{9}$$

with  $E_{\nu' \nu}$  the two-sequence energy,

$$E_{\nu' \nu} \equiv \sum_{t=t_0}^{t_f} E(x'_t; x_{t-1}, \dots, x_{t_0}). \tag{10}$$

$N = |\mathcal{X}|^{\Delta t + 1}$ ,  $\Delta t = t_f - t_0$ , is the number of configurations along pathway  $\nu$ , which is the result of combining  $\Delta t + 1$  events and  $|\mathcal{X}| = |\mathcal{Q}|^D |\mathcal{P}|^D$  possibilities for each event ( $\mathcal{Q}$  and  $\mathcal{P}$  are the discrete domains of  $q_h$  and  $p_h$ ,  $h = 1, \dots, D$ , respectively, and  $|\mathcal{Q}|$  and  $|\mathcal{P}|$  indicate the number of elements in the range of  $\mathcal{Q}$  and  $\mathcal{P}$ , respectively) and subscript  $\nu'(\lambda)$  in the sigma symbol reminds that the sum over the multiple  $x'_t$  variables, which are correlated due to memory effects, has to be evaluated according to the constraints imposed by the protocol.

The probability of a pathway, Equation (6), is a function of its energy,  $E_\nu$  (see Equation (7)), which is a sum over the energies of the microstates that the system has passed through in its evolution between microstates  $x_{t_0}$  and  $x_{t_f}$ . The pathway energy is in essence the action,  $\mathcal{S}$ , in the discrete domain; more in depth, it is the abbreviated action,  $\mathcal{S}_0$ , minus the action itself, as defined in classical mechanics [2], divided by the time elapsed between the initial and final microstates, namely  $(\mathcal{S}_0(\nu) - \mathcal{S}(\nu)) / \Delta t$ . The probability of a pathway thus weights the dynamics of a system according to the exponential of the value that this functional difference takes on such a pathway.

The expected value of the instantaneous energy, which is a state function known as the internal energy, is:

$$\langle E_t \rangle_\lambda \equiv \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_t) E(x_t; x_{t-1}, \dots, x_{t_0}), \tag{11}$$

where we have used that

$$p^{(\lambda)}(x_{t_0}, \dots, x_t) = \sum_{x_{t+1}, \dots, x_{t_f}} p^{(\lambda)}(x_{t_0}, \dots, x_{t_f}), \tag{12}$$

straightforward from the expansion of the probability as a product of conditional probabilities, Equation (5), and the fact that  $\sum_{x_t} p^{(\lambda)}(x_t | x_{t-1}, \dots, x_{t_0}) = 1$ ,  $t = t_0, \dots, t_f$ . This equation for the probability of a truncated temporal sequence can be expressed as ( $t < t_f$ )

$$p^{(\lambda)}(x_{t_0}, \dots, x_t) = \frac{e^{-\beta E(x_{t_0}, \dots, x_t)}}{Z^{(\lambda)}(x_{t_0}, \dots, x_{t-1})}, \tag{13}$$

where, in analogy to Equations (7) and (8) for the whole temporal sequence, the energy and the partition function of a truncated sequence are, respectively:

$$E(x_{t_0}, \dots, x_t) = \sum_{i=t_0}^t E(x_i; x_{i-1}, \dots, x_{t_0}), \tag{14}$$

$$Z^{(\lambda)}(x_{t_0}, \dots, x_{t-1}) = \prod_{i=t_0}^t Z^{(\lambda)}(x_i; x_{i-1}, \dots, x_{t_0}). \tag{15}$$

The expected value of the pathway energy,

$$\langle E_\nu \rangle_\lambda \equiv \sum_{\nu=1}^N p_\nu^{(\lambda)} E_\nu = \sum_{x_{t_0}, \dots, x_{t_f}} p^{(\lambda)}(x_{t_0}, \dots, x_{t_f}) E(x_{t_0}, \dots, x_{t_f}), \tag{16}$$

is equal to the sum over the expected instant energies that conform the pathway, namely

$$\langle E_v \rangle_\lambda = \sum_{t=t_0}^{t_f} \langle E_t \rangle_\lambda. \tag{17}$$

The proof to this result can be found in Appendix A. Although the expected value of a pathway function, like the energy in Equation (16), does not depend on a specific pathway, we will keep the subscript “v” within the angle brackets to denote that it is an ensemble average over a pathway function.

We now introduce the entropy at  $t$  with reference to both the protocol and the pathway that the system is following as [29]:

$$S^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0}) \equiv -k \ln p^{(\lambda)}(x_t | x_{t-1}, \dots, x_{t_0}), \tag{18}$$

with expected value

$$\langle S_t^{(\lambda)} \rangle_\lambda = \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_t) S^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0}). \tag{19}$$

This entropy constitutes the state function used in thermodynamics [1,3], with the caveat that it observes memory effects. Such entropy is used in information theory under the name of conditional entropy [27].

We can also introduce the pathway entropy:

$$S_v^{(\lambda)} \equiv -k \ln p_v^{(\lambda)} = \sum_{t=t_0}^{t_f} S^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0}). \tag{20}$$

The last part of this equation is a direct consequence of Equations (5) and (18). Its expected value,

$$\langle S_v^{(\lambda)} \rangle_\lambda = \sum_{v=1}^N p_v^{(\lambda)} S_v^{(\lambda)} = \sum_{x_{t_0}, \dots, x_{t_f}} p^{(\lambda)}(x_{t_0}, \dots, x_{t_f}) S^{(\lambda)}(x_{t_0}, \dots, x_{t_f}), \tag{21}$$

fulfills an analogous expression to that of the energy, Equation (17), namely

$$\langle S_v^{(\lambda)} \rangle_\lambda = \sum_{t=t_0}^{t_f} \langle S_t^{(\lambda)} \rangle_\lambda; \tag{22}$$

see Appendix A for the proof.

Finally, we introduce the Helmholtz free energy at  $t$  with reference to the protocol and the pathway as:

$$F^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0}) \equiv -kT \ln Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0}), \tag{23}$$

which actually does not depend on the present (instant  $t$ ), in contrast to the energy and the entropy, Equations (2) and (18), respectively. Its expected value (state function) is:

$$\langle F_t^{(\lambda)} \rangle_\lambda = \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_t) F^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0}). \tag{24}$$

The corresponding pathway Helmholtz free energy is:

$$F_v^{(\lambda)} \equiv -kT \ln Z_v^{(\lambda)} = \sum_{t=t_0}^{t_f} F^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0}). \tag{25}$$

The last part of this equation follows from Equations (8) and (23). Its expected value,

$$\langle F_v^{(\lambda)} \rangle_\lambda = \sum_{v=1}^N p_v^{(\lambda)} F_v^{(\lambda)} = \sum_{x_{t_0}, \dots, x_{t_f}} p^{(\lambda)}(x_{t_0}, \dots, x_{t_f}) F^{(\lambda)}(x_{t_0}, \dots, x_{t_f}), \tag{26}$$

fulfills

$$\langle F_v^{(\lambda)} \rangle_\lambda = \sum_{t=t_0}^{t_f} \langle F_t^{(\lambda)} \rangle_\lambda; \tag{27}$$

see Appendix A. The definitions of pathway thermodynamic potentials formally resemble those used in information theory for string of symbols [24] although their physical meaning is different.

In general, for a given thermodynamic potential “A”, it is possible to introduce its instantaneous and pathway versions under a protocol,

$$A^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0}), \tag{28}$$

$$A_v^{(\lambda)} \equiv \sum_{t=t_0}^{t_f} A^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0}), \tag{29}$$

respectively. State functions are constructed by taking expected values:

$$\langle A_t^{(\lambda)} \rangle_\lambda = \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_t) A^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0}), \tag{30}$$

$$\langle A_v^{(\lambda)} \rangle_\lambda = \sum_{v=1}^N p_v^{(\lambda)} A_v^{(\lambda)} = \sum_{x_{t_0}, \dots, x_{t_f}} p^{(\lambda)}(x_{t_0}, \dots, x_{t_f}) A^{(\lambda)}(x_{t_0}, \dots, x_{t_f}). \tag{31}$$

Instantaneous and pathway expected values fulfill:

$$\langle A_v^{(\lambda)} \rangle_\lambda = \sum_{t=t_0}^{t_f} \langle A_t^{(\lambda)} \rangle_\lambda. \tag{32}$$

From Equations (2), (18), and (23), and from Equations (7), (20), and (25), it is immediate to demonstrate the energy conservation for both instants and pathways, respectively:

$$F^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0}) = E(x_t; x_{t-1}, \dots, x_{t_0}) - TS^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0}), \tag{33}$$

$$F_v^{(\lambda)} = E_v - TS_v^{(\lambda)}. \tag{34}$$

In the same way, the energy is conserved for both instant and pathway state functions:

$$\langle F_t^{(\lambda)} \rangle_\lambda = \langle E_t \rangle_\lambda - T \langle S_t^{(\lambda)} \rangle_\lambda, \tag{35}$$

$$\langle F_v^{(\lambda)} \rangle_\lambda = \langle E_v \rangle_\lambda - T \langle S_v^{(\lambda)} \rangle_\lambda; \tag{36}$$

see Appendix A for the demonstrations of these results.

The internal energy for instants and pathways follows next laws:

$$\langle E_t \rangle_\lambda = - \left\langle \frac{\partial}{\partial \beta} \ln Z^{(\lambda)}(x_{t-1}, \dots, x_{t_0}) \right\rangle_\lambda, \quad \langle E_v \rangle_\lambda = - \left\langle \frac{\partial}{\partial \beta} \ln Z_v^{(\lambda)} \right\rangle_\lambda, \tag{37}$$

respectively. Corresponding expressions for the expected value of the entropies are

$$\langle S_t^{(\lambda)} \rangle_\lambda = - \left\langle \frac{\partial}{\partial T} F^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0}) \right\rangle_\lambda, \quad \langle S_v^{(\lambda)} \rangle_\lambda = - \left\langle \frac{\partial}{\partial T} F_v^{(\lambda)} \right\rangle_\lambda. \tag{38}$$

These four equations are demonstrated in Appendix A.

### 2.3. Equilibrium Pathways

The probability of a pathway described by the system under equilibrium conditions is:

$$p_v = \frac{\exp(-\beta E_v)}{Z}, \tag{39}$$

such that  $\sum_{v=1}^N p_v = 1$ . In this equation, we use the path energy, as given by Equation (7), and the partition function [1,3]

$$Z \equiv \sum_{x_{t_0}, \dots, x_{t_f}} \exp\left(-\beta \sum_{t=t_0}^{t_f} E(x_t; x_{t-1}, \dots, x_{t_0})\right) = \sum_{v=1}^N \exp(-\beta E_v), \tag{40}$$

where, in contrast to Equation (9), the sums are extended to both present and previous events. Certainly, there are no prime variables and the sums are nested in Equation (40), unlike in the protocol-driven Equation (9). Partition function  $Z$  is not built on the basis of a particular protocol; therefore, the resulting probability does not limit how the different pathways are accessed.

The equilibrium partition function  $Z$  measures the available pathways  $v$  that connect the microstates at  $t = t_0$  and  $t = t_f$ , constructed as temporal sequences of stochastic events  $X_t$  and statistically weighted by the exponential of their energies  $\sum_t E(x_t; x_{t-1}, \dots, x_{t_0})$ ; these energies, in turn, account for the memory, i.e., the relative interactions of every present microstate,  $x_t$ , with its previous ones,  $\{x_{t_0}, \dots, x_{t-1}\}$ . The sequence-dependent partition function,  $Z_v^{(\lambda)}$ , in contrast, measures the exponential energy-weighted pathways that connect these states considering that at each step,  $x_t$ , the previous events,  $\{x_{t-1}, \dots, x_{t_0}\}$ , are unchangeable and that the sequence of events is stochastically determined by protocol  $\lambda$ .

The probabilities expressed in Equation (39) represent pathways along which the external constraints to the system do not change with time. The dynamics so described are, therefore, timeless. As we will see, the expected values of the pathway thermodynamic potentials reduce to explicit functions of  $Z$ , as shown generally for equilibrium thermodynamics [1,3]. It is demonstrated (Appendix A) that, to recover Equation (39), having in mind that [28]

$$p_v \equiv \prod_{t=t_0}^{t_f} p(x_t | x_{t-1}, \dots, x_{t_0}), \tag{41}$$

the probability of a microstate under equilibrium pathways must be:

$$p(x_t | x_{t-1}, \dots, x_{t_0}) = \frac{e^{-\beta E(x_t; x_{t-1}, \dots, x_{t_0})} f(x_{t_0}, \dots, x_t)}{\sum_{x_t} e^{-\beta E(x_t; x_{t-1}, \dots, x_{t_0})} f(x_{t_0}, \dots, x_t)}, \tag{42}$$

where  $\sum_{x_t} p(x_t | x_{t-1}, \dots, x_{t_0}) = 1$  and

$$f(x_{t_0}, \dots, x_t) = \sum_{x_{t+1}, \dots, x_{t_f}} e^{-\beta E(x_{t+1}; x_t, \dots, x_{t_0})} \dots e^{-\beta E(x_{t_f}; x_{t_f-1}, \dots, x_{t_0})} \tag{43}$$

$$= \sum_{x_{t+1}, \dots, x_{t_f}} \exp\left(-\beta \sum_{i=t+1}^{t_f} E(x_i; x_{i-1}, \dots, x_{t_0})\right). \tag{44}$$

For completeness, it is important to note that

$$f(x_{t_0}, \dots, x_{t_f}) = 1, \tag{45}$$



and that

$$\sum_{x_{t_0}} e^{-\beta E(x_{t_0})} f(x_{t_0}) = Z. \tag{46}$$

It follows straightforward from Equation (44) that

$$f(x_{t_0}, \dots, x_t) = \sum_{x_{t+1}} e^{-\beta E(x_{t+1}; x_t, \dots, x_{t_0})} f(x_{t_0}, \dots, x_{t+1}). \tag{47}$$

We will call functions  $f(x_{t_0}, \dots, x_t)$ ,  $t = t_0, \dots, t_f$ , *partition factors* [30].

The probability of a microstate under equilibrium pathways, Equation (42), can be expressed as

$$p(x_t | x_{t-1}, \dots, x_{t_0}) = \frac{e^{-\beta E(x_t; x_{t-1}, \dots, x_{t_0})}}{Z(x_t; x_{t-1}, \dots, x_{t_0})}, \tag{48}$$

with

$$Z(x_t; x_{t-1}, \dots, x_{t_0}) \equiv \frac{f(x_{t_0}, \dots, x_{t-1})}{f(x_{t_0}, \dots, x_t)} \tag{49}$$

being the equilibrium instant partition function. Equation (49) reduces to Equation (4) for protocol-driven dynamics. In analogy to Equation (8), the equilibrium pathway partition function builds up as

$$Z = \prod_{t=t_0}^{t_f} Z(x_t; x_{t-1}, \dots, x_{t_0}). \tag{50}$$

The reduced probability for a truncated equilibrium temporal sequence is:

$$p(x_{t_0}, \dots, x_t) = \sum_{x_{t+1}, \dots, x_{t_f}} p(x_{t_0}, \dots, x_{t_f}) = \frac{e^{-\beta E(x_{t_0}, \dots, x_t)}}{Z(x_{t_0}, \dots, x_t)}, \tag{51}$$

where

$$Z(x_{t_0}, \dots, x_t) = \prod_{i=t_0}^t Z(x_i; x_{i-1}, \dots, x_{t_0}) = \frac{Z}{f(x_{t_0}, \dots, x_t)} \tag{52}$$

is the partition function for equilibrium truncated temporal sequences; see Appendix A.

For a thermodynamic potential “A”, the instantaneous and pathway versions under equilibrium conditions are represented by dropping superscript  $\lambda$  in Equations (28) and (29), namely

$$A(x_t; x_{t-1}, \dots, x_{t_0}), \tag{53}$$

$$A_v \equiv \sum_{t=t_0}^{t_f} A(x_t; x_{t-1}, \dots, x_{t_0}). \tag{54}$$

Their expected values (state functions) are constructed by using probability Equations (42) and (39), respectively:

$$\langle A_t \rangle = \sum_{x_{t_0}, \dots, x_t} p(x_{t_0}, \dots, x_t) A(x_t; x_{t-1}, \dots, x_{t_0}), \tag{55}$$

$$\langle A_v \rangle = \sum_{v=1}^N p_v A_v = \sum_{x_{t_0}, \dots, x_{t_f}} p(x_{t_0}, \dots, x_{t_f}) A(x_{t_0}, \dots, x_{t_f}), \tag{56}$$

fulfilling the next relation between instantaneous and pathway equilibrium state functions:

$$\langle A_v \rangle = \sum_{t=t_0}^{t_f} \langle A_t \rangle. \tag{57}$$

The proof to Equation (57) is formally equivalent to that for Equation (17).

The equilibrium instant and pathway energies are given by Equations (2) and (7), respectively, because they are independent of the protocol. However, state functions do depend on the presence or absence of a protocol because a different probability is used in the expressions of their expected values. Then, in analogy to Equations (11) and (16), these state functions read:

$$\langle E_t \rangle = \sum_{x_{t_0}, \dots, x_t} p(x_{t_0}, \dots, x_t) E(x_t; x_{t-1}, \dots, x_{t_0}), \tag{58}$$

$$\langle E_v \rangle = \sum_{v=1}^N p_v E_v = \sum_{x_{t_0}, \dots, x_{t_f}} p(x_{t_0}, \dots, x_{t_f}) E(x_{t_0}, \dots, x_{t_f}). \tag{59}$$

They keep the next relation:

$$\langle E_v \rangle = \sum_{t=t_0}^{t_f} \langle E_t \rangle, \tag{60}$$

which proof is formally equivalent to that for Equation (17).

The equilibrium instant and pathway entropies depend on the presence or absence of a protocol because they operate on a probability. In analogy to Equations (18) and (20), these thermodynamic potentials are built by using instant and pathway equilibrium probability Equations (39) and (42):

$$S(x_t; x_{t-1}, \dots, x_{t_0}) \equiv -k \ln p(x_t | x_{t-1}, \dots, x_{t_0}), \tag{61}$$

$$S_v \equiv -k \ln p_v = \sum_{t=t_0}^{t_f} S(x_t; x_{t-1}, \dots, x_{t_0}), \tag{62}$$

with expected values:

$$\langle S_t \rangle = \sum_{x_{t_0}, \dots, x_t} p(x_{t_0}, \dots, x_t) S(x_t; x_{t-1}, \dots, x_{t_0}), \tag{63}$$

$$\langle S_v \rangle = \sum_{v=1}^N p_v S_v = \sum_{x_{t_0}, \dots, x_{t_f}} p(x_{t_0}, \dots, x_{t_f}) S(x_{t_0}, \dots, x_{t_f}). \tag{64}$$

Finally, the equilibrium instant and pathway Helmholtz free energies, which also depend on the presence or absence of protocol because they operate on partition functions, are given in analogy to Equations (23) and (25) by:

$$F(x_t; x_{t-1}, \dots, x_{t_0}) \equiv -kT \ln Z(x_t; x_{t-1}, \dots, x_{t_0}), \tag{65}$$

$$F_v \equiv -kT \ln Z = \sum_{t=t_0}^{t_f} F(x_t; x_{t-1}, \dots, x_{t_0}), \tag{66}$$

respectively, with expected values:

$$\langle F_t \rangle = \sum_{x_{t_0}, \dots, x_t} p(x_{t_0}, \dots, x_t) F(x_t; x_{t-1}, \dots, x_{t_0}), \tag{67}$$

$$\langle F_v \rangle = \sum_{v=1}^N p_v F_v = \sum_{x_{t_0}, \dots, x_{t_f}} p(x_{t_0}, \dots, x_{t_f}) F(x_{t_0}, \dots, x_{t_f}). \tag{68}$$

The equilibrium entropy and free energy fulfill the next relations between pathway and instant expected values:

$$\langle S_v \rangle = \sum_{t=t_0}^{t_f} \langle S_t \rangle, \quad \langle F_v \rangle = \sum_{t=t_0}^{t_f} \langle F_t \rangle, \tag{69}$$

which proofs are formally equivalent to that for Equation (17).

Similar conservation energy equations to those found for protocol-driven dynamics are obtained for equilibrium. Their expressions (and proofs in Appendix A) follow by dropping superscript  $\lambda$  in Equations (33)–(36), namely.

$$F(x_t; x_{t-1}, \dots, x_{t_0}) = E(x_t; x_{t-1}, \dots, x_{t_0}) - TS(x_t; x_{t-1}, \dots, x_{t_0}), \tag{70}$$

$$F_v = E_v - TS_v, \tag{71}$$

$$\langle F_t \rangle = \langle E_t \rangle - T \langle S_t \rangle, \tag{72}$$

$$\langle F_v \rangle = \langle E_v \rangle - T \langle S_v \rangle. \tag{73}$$

In analogy to Equations (37) and (38), equilibrium processes follow the next relations for state functions:

$$\langle E_t \rangle = - \left\langle \frac{\partial}{\partial \beta} \ln Z(x_t; x_{t-1}, \dots, x_{t_0}) \right\rangle, \quad \langle E_v \rangle = - \frac{\partial}{\partial \beta} \ln Z, \tag{74}$$

respectively. Corresponding expressions for the entropies are

$$\langle S_t \rangle = - \left\langle \frac{\partial}{\partial T} F(x_t; x_{t-1}, \dots, x_{t_0}) \right\rangle, \quad \langle S_v \rangle = - \frac{\partial}{\partial T} F. \tag{75}$$

The right-hand side expressions in Equations (74) and (75) are known in the context of equilibrium thermodynamics in the absence of memory as  $U \equiv \langle E_v \rangle$ ,  $S \equiv \langle S_v \rangle = -k \sum_v p_v \ln p_v$ , and  $F \equiv \langle F_v \rangle = F_v = -kT \ln Z$  [1,3].

#### 2.4. Relations between Protocol-Driven and Equilibrium Treatments

The following relations for instantaneous partition functions hold:

$$\left\langle \frac{Z^{(\lambda)}(x_{t-1}, \dots, x_{t_0})}{Z(x_{t-1}, \dots, x_{t_0})} \right\rangle_\lambda = 1, \quad \left\langle \frac{Z(x_t; x_{t-1}, \dots, x_{t_0})}{Z^{(\lambda)}(x_{t-1}, \dots, x_{t_0})} \right\rangle = 1; \tag{76}$$

see Appendix A for their demonstrations. These expressions indicate that the equilibrium and protocol driven partition functions approach one another on average ratios.

Similar expressions can be found for partition functions along truncated sequences:

$$\left\langle \frac{Z^{(\lambda)}(x_{t_0}, \dots, x_{t-1})}{Z(x_{t_0}, \dots, x_t)} \right\rangle_\lambda = 1, \quad \left\langle \frac{Z(x_{t_0}, \dots, x_t)}{Z^{(\lambda)}(x_{t_0}, \dots, x_{t-1})} \right\rangle = 1, \tag{77}$$

which are trivially demonstrated from the definitions of  $p(x_{t_0}, \dots, x_t)$  and  $p^{(\lambda)}(x_{t_0}, \dots, x_t)$ . Particular cases of previous equations are those for the complete sequences, that is, when  $t = t_f$ :

$$\left\langle \frac{Z_v^{(\lambda)}}{Z} \right\rangle_\lambda = 1, \quad \left\langle \frac{Z}{Z_v^{(\lambda)}} \right\rangle = 1. \tag{78}$$

Straightforward consequences for the expected values of the probability ratios are next:

$$\left\langle \frac{p(x_t|x_{t-1}, \dots, x_{t_0})}{p^{(\lambda)}(x_t|x_{t-1}, \dots, x_{t_0})} \right\rangle_\lambda = 1, \quad \left\langle \frac{p^{(\lambda)}(x_t|x_{t-1}, \dots, x_{t_0})}{p(x_t|x_{t-1}, \dots, x_{t_0})} \right\rangle = 1, \tag{79}$$

$$\left\langle \frac{p(x_{t_0}, \dots, x_t)}{p^{(\lambda)}(x_{t_0}, \dots, x_t)} \right\rangle_\lambda = 1, \quad \left\langle \frac{p^{(\lambda)}(x_{t_0}, \dots, x_t)}{p(x_{t_0}, \dots, x_t)} \right\rangle = 1, \tag{80}$$

$$\left\langle \frac{p_\nu}{p_\nu^{(\lambda)}} \right\rangle_\lambda = 1, \quad \left\langle \frac{p_\nu^{(\lambda)}}{p_\nu} \right\rangle = 1. \tag{81}$$

The evolution of a system does not only depend on the protocol but also on the physical nature of the system itself. In this regard, partition functions and probability distributions for protocol-driven dynamics are equivalent to their associated equilibrium processes under ensemble averages, according to Equations (76)–(81). These relations indicate that, on average, the stochastic nature of any protocol-driven process is equivalent to that of the equilibrium process.

The Kullback-Leibler distance (or relative entropy) between two conditional probability distributions  $p$  and  $q$  are used as defined elsewhere [27]:

$$D(p(x_t|x_{t-1}, \dots, x_{t_0})||q(x_t|x_{t-1}, \dots, x_{t_0})) \equiv \sum_{x_{t_0}, \dots, x_t} p(x_{t_0}, \dots, x_t) \ln \frac{p(x_t|x_{t-1}, \dots, x_{t_0})}{q(x_t|x_{t-1}, \dots, x_{t_0})} = \left\langle \ln \frac{p(x_t|x_{t-1}, \dots, x_{t_0})}{q(x_t|x_{t-1}, \dots, x_{t_0})} \right\rangle_p. \tag{82}$$

They are positive for protocol-driven and equilibrium instantaneous probability distributions, as expected:

$$D(p^{(\lambda)}(x_t|x_{t-1}, \dots, x_{t_0})||p(x_t|x_{t-1}, \dots, x_{t_0})) = \left\langle \ln \frac{Z(x_t; x_{t-1}, \dots, x_{t_0})}{Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0})} \right\rangle_\lambda \geq 0, \tag{83}$$

$$D(p(x_t|x_{t-1}, \dots, x_{t_0})||p^{(\lambda)}(x_t|x_{t-1}, \dots, x_{t_0})) = \left\langle \ln \frac{Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0})}{Z(x_t; x_{t-1}, \dots, x_{t_0})} \right\rangle \geq 0, \tag{84}$$

where we have applied Jensen’s inequality [27] and Equation (76).

Using the chain rule for relative entropies [27],

$$D(p(x_{t_0}, \dots, x_t)||q(x_{t_0}, \dots, x_t)) = \sum_{i=t_0}^t D(p(x_i|x_{i-1}, \dots, x_{t_0})||q(x_i|x_{i-1}, \dots, x_{t_0})), \tag{85}$$

it follows that

$$D(p^{(\lambda)}(x_{t_0}, \dots, x_t)||p(x_{t_0}, \dots, x_t)) = \left\langle \ln \frac{Z(x_{t_0}, \dots, x_t)}{Z^{(\lambda)}(x_{t_0}, \dots, x_{t-1})} \right\rangle_\lambda = \sum_{i=t_0}^t \left\langle \ln \frac{Z(x_i; x_{i-1}, \dots, x_{t_0})}{Z^{(\lambda)}(; x_{i-1}, \dots, x_{t_0})} \right\rangle_\lambda \geq 0, \tag{86}$$

$$D(p(x_{t_0}, \dots, x_t)||p^{(\lambda)}(x_{t_0}, \dots, x_t)) = \left\langle \ln \frac{Z^{(\lambda)}(x_{t_0}, \dots, x_{t-1})}{Z(x_{t_0}, \dots, x_t)} \right\rangle = \sum_{i=t_0}^t \left\langle \ln \frac{Z^{(\lambda)}(; x_{i-1}, \dots, x_{t_0})}{Z(x_i; x_{i-1}, \dots, x_{t_0})} \right\rangle \geq 0. \tag{87}$$

In particular,

$$D(p_\nu^{(\lambda)}||p_\nu) = \left\langle \ln \frac{Z}{Z_\nu^{(\lambda)}} \right\rangle_\lambda \geq 0, \tag{88}$$

$$D(p_\nu||p_\nu^{(\lambda)}) = \left\langle \ln \frac{Z_\nu^{(\lambda)}}{Z} \right\rangle \geq 0. \tag{89}$$

These formal results and those in Equation (78) appeared in previous works in the context of information chains [24,26]; we show them again for the sake of completeness in temporal chains.

2.5. Protocol-Driven Dynamics and Equilibrium Approach Each Other for Weak Memory Effects

We observed previously that sequence-dependent and equilibrium statistics are mutually convergent when a chain construction is sufficiently smoothly-dependent on its history [26]. We next formulate and demonstrate the Independence Limit theorem for instants, for which proof is in Appendix B:

**Theorem 1** (Instant Independence Limit). *Let  $v = \{x_{t_0}, \dots, x_t, \dots, x_{t_f}\}$  and  $v' = \{x'_{t_0}, \dots, x'_t, \dots, x'_{t_f}\}$  be two reversible pathways with memory along which a certain system may evolve. Let  $Z(x'_t; x'_{t-1}, \dots, x'_{t_0})$  and  $Z^{(\lambda)}(x_{t-1}, \dots, x_{t_0})$  be the equilibrium and protocol-driven instantaneous partition functions (Equations (49) and (4), respectively), relative to different histories until time  $t$ . If the normalized energy difference for microstate  $x'_t$  relative to distinct histories fulfills:*

$$|E(x'_t; x'_{t-1}, \dots, x'_{t_0}) - E(x'_t; x_{t-1}, \dots, x_{t_0})| / (kT) \rightarrow 0 \quad (kT > 0) \tag{90}$$

and the normalized energy differences for future events within the same pathway relative to distinct microstates  $x_t$  and  $x'_t$  at time  $t$  fulfill:

$$|E(x'_i; x'_{i-1}, \dots, x'_{t+1}, x'_t, x'_{t-1}, \dots, x'_{t_0}) - E(x'_i; x_{i-1}, \dots, x_{t+1}, x_t, x'_{t-1}, \dots, x'_{t_0})| / (kT) \rightarrow 0$$

for  $i = t + 1, \dots, t_f$ , (91)

then  $Z(x'_t; x'_{t-1}, \dots, x'_{t_0}) / Z^{(\lambda)}(x_{t-1}, \dots, x_{t_0}) \rightarrow 1$ .

This theorem provides the adequate link between protocol-driven and equilibrium pathways in the absence of friction. It states that when memory-induced variations in the energetic cost for advancing one temporal step are sufficiently low with respect to the thermal level ( $kT$ ), the evolution of the system by a defined mechanism approaches an equilibrium thermalization of independent events.

It is easy to prove (Appendix B) the two following corollaries for truncated and full partition functions:

**Corollary 1** (Truncated Pathway Independence Limit). *If, in addition to the condition expressed in Equation (91), the one in Equation (90) extends to all instants before  $t$ , namely*

$$|E(x'_i; x'_{i-1}, \dots, x'_{t_0}) - E(x'_i; x_{i-1}, \dots, x_{t_0})| / (kT) \rightarrow 0 \quad \text{for } i = t_0, \dots, t, \tag{92}$$

then  $Z(x'_{t_0}, \dots, x'_t) / Z^{(\lambda)}(x_{t_0}, \dots, x_{t-1}) \rightarrow 1$ .

**Corollary 2** (Full Pathway Independence Limit). *If, for all instants,  $t = t_0, \dots, t_f$ ,*

$$|E(x'_t; x'_{t-1}, \dots, x'_{t_0}) - E(x'_t; x_{t-1}, \dots, x_{t_0})| / (kT) \rightarrow 0, \tag{93}$$

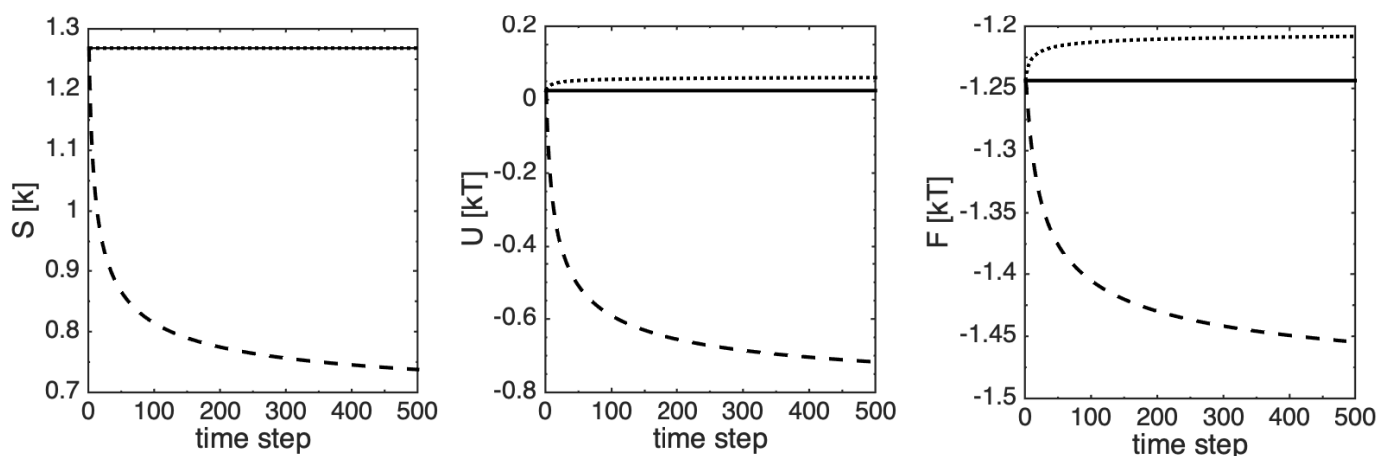
then  $Z / Z_v^{(\lambda)} \rightarrow 1$ .

Since the extension of Equation (90) to all instants comprises Equation (91), a sufficient condition to meet the *Independence Limit* for instants, truncated pathways and full pathways is  $|E(x'_t; x'_{t-1}, \dots, x'_{t_0}) - E(x'_t; x_{t-1}, \dots, x_{t_0})| / (kT) \rightarrow 0, \forall t = t_0, \dots, t_f$ , which is the condition of Equation (93) in the last corollary. Interestingly, this last corollary recovers the *Independence Limit Theorem* found elsewhere in the context of spatial chains with memory [26].

### 3. Applications

Computing systems, either natural, like the ones taking place in biological cells, or artificial, like tape-based technologies, generate stepwise evolutions through the intervention of Turing machines. They are of model significance to illustrate the above theory because they show how physical interactions within the tape and between the tape and the Turing-like processing machine influence both the spatial arrangements at the tape and the history of the whole system. In particular, biological replication, transcription, and translation comprise both a space- and time-dependent directional, stochastic chain with memory. More in depth, their dynamics consists of a single protein—a polymerase or ribosome—that builds a single template-directed polymer as a function of the time through incorporation of monomers one at a time.

Thermodynamic potentials for a growing biopolymer are simulated in Figure 1. The system exchanges energies near the thermal level at each time instant. Equilibrium pathways exploit memory effects greater than protocol-driven pathways because the system is assumed to explore indefinitely each monomer alternative at each time step, thus improving the chances for correct monomer selection under positive feedback. In particular, the entropy decreases, thus indicating a lower incidence of errors, and the internal and free energies become more negative, thus indicating a higher stability of the resulting polymer as time passes.



**Figure 1.** Entropy, internal energy, and Helmholtz free energy as a function of the time during the templated-directed synthesis of a biopolymer. The graphs show expected potentials per monomer considering independent events (solid lines), protocol-driven pathways (dotted lines) and equilibrium pathways (dashed lines). Each incorporated monomer either releases an energy of 0.5 kT if it fits correctly or absorbs an energy 0.5 kT otherwise. Monomers are selected from a pool of 4 different elements, only one fitting correctly at a time. The memory is modeled by a power law of the form  $1/(\Delta t)^{3/2}$ , where  $\Delta t$  is the number of elapsed time steps [24,26].

The protocol-driven results in Figure 1 just address the case in which the polymer only grows during its synthesis. Monomers may actually be removed during polymer synthesis due to proofreading and editing activities. In these situations, the system explores the available pathways in a longer time, thus approaching the equilibrium results (also shown in Figure 1) and thus reducing the incidence of errors. This applicative perspective can be modeled with the above theory by selecting a protocol that combines shrinking and growing dynamics.

### 4. Discussion

We have extended thermodynamics to systems that evolve along specific sequences of events driven by changing constraints in the frictionless limit. To do that, we have introduced pathway- and protocol-dependent functions, including thermodynamic poten-

tials, that characterize the microstates of the system in the presence of memory. Similarly, we have analyzed the equilibrium evolution of systems that keep memory of previous events by introducing pathway-dependent functions. Under ensemble averages, protocol-driven and equilibrium thermodynamic functions become state functions because they lose their dependence over the pathways. Our theory discriminates between microscopic reversibility and equilibrium, which, as presented in the theorem, converge to the same concept when memory effects become negligible. Given that randomness is characterized by probability, we find that, on average terms, the stochastic nature of protocol-driven and equilibrium processes are equivalent because the expected value of the ratio between any protocol-driven probability and the equilibrium probability is always unity, with independence of which the protocol is. Finally, we have applied our framework in the context of template-directed biopolymer synthesis, a molecular process that objectifies the connection between thermodynamic and information entropies. This nanoscale illustration—in which a protein replicates, transcribes, or translates an information carrier—shows that information requires energy and that information is another manifestation of entropy (like heat to energy).

When memory is present, the number of different microstates that a system can attain increases with the number of events that it can recall because each microstate involves a configurational history of events. When memory extends to all previous events at each time step, there is a bijection between available microstates until a particular time  $t$  and pathways to reach them; the system actually resets whenever it explores new pathways in these conditions. If memory effects can be cut off down to a finite number of previous events, as, for example, in Markov (memoryless) dynamics or in the case of independent events, microstates can be recurrently visited within a particular pathway, that is, without starting over. The lower the number of nearest temporal neighbors to be considered in the memory, the lower the revisiting period. This revisiting period can be assumed as the so-called Poincaré recurrence time, which increases with the number of past events that stochastically influence the present. In the limit in which the origin of the system is at  $t = -\infty$  and the memory extends to all previous events, the Poincaré recurrence time tends to infinity because the system cannot restart.

The evolution of a system is a consequence of the existence of a protocol, which represents constraints that change with time. If the protocol is sufficiently smooth in the time dependence, the system has time to visit many microstates before the constraints has substantially changed, thus evolving near or at equilibrium. When the environment evolves very rapidly (involving that the protocol has a sharp time dependence), the system cannot follow its changes due to memory effects. For long times, the fact that the system cannot relax to its original state implies that it presents hysteresis and evolves away from equilibrium.

The existence of a protocol, therefore, biases the pathways through which a system progresses between two microstates. If the memory of the system is very long, the system does not lose correlations with previous events at short periods of time. In these conditions, it is not possible to assume ensemble averages and time averages as interchangeable, which make evolutions no longer ergodic.

Considering that fluctuation theorems assume both microscopic reversibility and Markovianity [5,6], and that Markovianity is a special case of non-Markovianity, we speculate that our theory may describe non-equilibrium thermodynamics within a unified framework.

**Funding:** Work supported by Ministerio de Ciencia e Innovación, grant number PID2019-107391RB-I00.

**Acknowledgments:** The author wishes to thank the Universtat Politècnica de València for general support through the *Attraction Talent* program.

**Conflicts of Interest:** The author declares no conflict of interest. The founding sponsor had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

### Appendix A. Proofs

We next demonstrate the expressions of Sections 2.2–2.4.

**Proof of Equation (17).** From Equations (16), (7), and (12), it follows that

$$\begin{aligned} \langle E_V \rangle_\lambda &= \sum_{x_{t_0}, \dots, x_{t_f}} p^{(\lambda)}(x_{t_0}, \dots, x_{t_f}) \sum_{t=t_0}^{t_f} E(x_t; x_{t-1}, \dots, x_{t_0}) \\ &= \sum_{t=t_0}^{t_f} \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_t) E(x_t; x_{t-1}, \dots, x_{t_0}) = \sum_{t=t_0}^{t_f} \langle E_t \rangle_\lambda, \end{aligned}$$

which proves Equation (17). □

The proof to Equation (22) is similar to this for Equation (17) and can be found in Reference [27] in the context of Information Theory for the therein called conditional entropy. The proofs to Equation (27) and to Equation (32) for a general thermodynamic potential “A” (following Equations (28) and (29)) parallel those for Equations (17) and (22).

**Proof of Equation (35).** From Equations (19), (18), and (3), it follows that

$$\begin{aligned} \langle S_t^{(\lambda)} \rangle_\lambda &= \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_t) S^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0}) \\ &= -k \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_t) \ln p^{(\lambda)}(x_t | x_{t-1}, \dots, x_{t_0}) \\ &= -k \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_t) \left( -\beta E(x_t; x_{t-1}, \dots, x_{t_0}) - \ln Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0}) \right) \\ &= \frac{1}{T} \left( \langle E_t \rangle_\lambda - \langle F_t^{(\lambda)} \rangle_\lambda \right), \end{aligned}$$

where we have used Equations (11), (23), and (24). □

**Proof of Equation (36).** This expression appears by taking sums over subscript *t* on Equation (35) and using Equations (17), (22), and (27). This result was demonstrated in a previous work through a different strategy [24]. □

**Proof of Equation (37).** From Equations (11) and (37), left, satisfies

$$\begin{aligned} \langle E_t \rangle_\lambda &= \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_t) E(x_t; x_{t-1}, \dots, x_{t_0}) \\ &= \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_{t-1}) p^{(\lambda)}(x_t | x_{t-1}, \dots, x_{t_0}) E(x_t; x_{t-1}, \dots, x_{t_0}) \\ &= \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_{t-1}) \frac{\exp(-\beta E(x_t; x_{t-1}, \dots, x_{t_0}))}{Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0})} E(x_t; x_{t-1}, \dots, x_{t_0}) \\ &= - \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_{t-1}) \frac{1}{Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0})} \frac{\partial}{\partial \beta} \exp(-\beta E(x_t; x_{t-1}, \dots, x_{t_0})) \\ &= - \sum_{x_{t_0}, \dots, x_{t-1}} p^{(\lambda)}(x_{t_0}, \dots, x_{t-1}) \frac{\partial}{\partial \beta} \sum_{x_t} p^{(\lambda)}(x_t | x_{t-1}, \dots, x_{t_0}) \\ &\quad - \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_t) \frac{\partial}{\partial \beta} \ln Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0}) = - \left\langle \frac{\partial}{\partial \beta} \ln Z_t^{(\lambda)} \right\rangle_\lambda, \end{aligned}$$



where we have used Equations (3), (5), and (12). To abbreviate, we have inserted  $\frac{\partial}{\partial\beta} \ln Z_t^{(\lambda)} \equiv \frac{\partial}{\partial\beta} \ln Z^{(\lambda)}(;x_{t-1}, \dots, x_{t_0})$ . Equation (37), right, is obtained by taking sums over subscript  $t$  on Equation (37), left. More in depth, since

$$\begin{aligned} & - \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_t) \frac{\partial}{\partial\beta} \ln Z^{(\lambda)}(;x_{t-1}, \dots, x_{t_0}) \\ &= - \sum_{x_{t_0}, \dots, x_{t_f}} p^{(\lambda)}(x_{t_0}, \dots, x_{t_f}) \frac{\partial}{\partial\beta} \ln Z^{(\lambda)}(;x_{t-1}, \dots, x_{t_0}) = - \sum_{\nu=1}^N p_\nu^{(\lambda)} \frac{\partial}{\partial\beta} \ln Z^{(\lambda)}(;x_{t-1}, \dots, x_{t_0}) \end{aligned}$$

so that

$$\begin{aligned} & - \sum_{t=t_0}^{t_f} \left\langle \frac{\partial}{\partial\beta} \ln Z_t^{(\lambda)} \right\rangle_\lambda = - \left\langle \frac{\partial}{\partial\beta} \sum_{t=t_0}^{t_f} \ln Z_t^{(\lambda)} \right\rangle_\lambda = - \left\langle \frac{\partial}{\partial\beta} \ln \prod_{t=t_0}^{t_f} Z_t^{(\lambda)} \right\rangle_\lambda \\ &= - \sum_{\nu=1}^N p_\nu^{(\lambda)} \frac{\partial}{\partial\beta} \ln Z_\nu^{(\lambda)} = - \left\langle \frac{\partial}{\partial\beta} \ln Z_\nu^{(\lambda)} \right\rangle_\lambda, \end{aligned}$$

then, considering  $\sum_{t=t_0}^{t_f} \langle E_t \rangle_\lambda = \langle E_\nu \rangle_\lambda$  from above, it follows that

$$\sum_{t=t_0}^{t_f} \left[ \langle E_t \rangle_\lambda = - \left\langle \frac{\partial}{\partial\beta} \ln Z_t^{(\lambda)} \right\rangle_\lambda \right] \rightarrow \langle E_\nu \rangle_\lambda = - \left\langle \frac{\partial}{\partial\beta} \ln Z_\nu^{(\lambda)} \right\rangle_\lambda \tag{A1}$$

which proves Equation (37), right. This equation was also demonstrated in a previous work by an alternative strategy [24]. □

**Proof of Equation (38).** From Equations (35), (23), and (24), and Equation (37), left, in this order, Equation (38), left, satisfies

$$\left\langle S_t^{(\lambda)} \right\rangle_\lambda = - \frac{1}{T} \left\langle F_t^{(\lambda)} - E_t^{(\lambda)} \right\rangle_\lambda = - \frac{1}{T} \left\langle -kT \ln Z_t^{(\lambda)} + \frac{\partial}{\partial\beta} \ln Z_t^{(\lambda)} \right\rangle_\lambda = - \left\langle \frac{\partial}{\partial T} F_t^{(\lambda)} \right\rangle_\lambda.$$

To abbreviate, we have inserted  $\frac{\partial}{\partial T} F_t^{(\lambda)} \equiv \frac{\partial}{\partial T} F^{(\lambda)}(x_t; x_{t-1}, \dots, x_{t_0})$ . Equation (38), right, is obtained by taking sums over subscript  $t$  on Equation (38), left. Following the above demonstration for  $\sum_{t=t_0}^{t_f} \langle E_t \rangle_\lambda = \langle E_\nu \rangle_\lambda$ , it is easy to see that  $\sum_{t=t_0}^{t_f} \langle S_t^{(\lambda)} \rangle_\lambda = \langle S_\nu^{(\lambda)} \rangle_\lambda$ . Likewise, considering the above demonstration for  $-\sum_{t=t_0}^{t_f} \left\langle \frac{\partial}{\partial\beta} \ln Z_t^{(\lambda)} \right\rangle_\lambda = - \left\langle \frac{\partial}{\partial\beta} \ln Z_\nu^{(\lambda)} \right\rangle_\lambda$ , it follows that

$$- \sum_{t=t_0}^{t_f} \left\langle \frac{\partial}{\partial T} F_t^{(\lambda)} \right\rangle_\lambda = - \left\langle \frac{\partial}{\partial T} \sum_{t=t_0}^{t_f} F_t^{(\lambda)} \right\rangle_\lambda = - \sum_{\nu=1}^N p_\nu^{(\lambda)} \frac{\partial}{\partial T} F_\nu^{(\lambda)} = - \left\langle \frac{\partial}{\partial T} F_\nu^{(\lambda)} \right\rangle_\lambda.$$

Then,

$$\sum_{t=t_0}^{t_f} \left[ \left\langle S_t^{(\lambda)} \right\rangle_\lambda = - \left\langle \frac{\partial}{\partial T} F_t^{(\lambda)} \right\rangle_\lambda \right] \rightarrow \langle S_\nu^{(\lambda)} \rangle_\lambda = - \left\langle \frac{\partial}{\partial T} F_\nu^{(\lambda)} \right\rangle_\lambda, \tag{A2}$$

which proves Equation (38), right. This equation was also demonstrated in a previous work by an alternative strategy [24]. □

**Proof of Equation (42).** We want to demonstrate that  $p_\nu \equiv p(x_{t_0}, \dots, x_{t_f}) = \prod_{t=t_0}^{t_f} p(x_t | x_{t-1}, \dots, x_{t_0})$ , where  $p_\nu$  is given by Equation (39) and  $p(x_t | x_{t-1}, \dots, x_{t_0})$  by Equation (42).

Certainly, Equation (42) can be expressed as:

$$p(x_t|x_{t-1}, \dots, x_{t_0}) = e^{-\beta E(x_t;x_{t-1}, \dots, x_{t_0})} \frac{f(x_{t_0}, \dots, x_t)}{f(x_{t_0}, \dots, x_{t-1})}$$

by using property Equation (47). Then,

$$p_v = e^{-\beta E_v} \prod_{i=t_0}^{t_f} \frac{f(x_{t_0}, \dots, x_t)}{f(x_{t_0}, \dots, x_{t-1})} = e^{-\beta E_v} \frac{f_{t_0} f_{t_0+1} f_{t_0+2} \dots f_{t_f-2} f_{t_f-1} f_{t_f}}{Z f_{t_0} f_{t_0+1} f_{t_0+2} \dots f_{t_f-2} f_{t_f-1}} = \frac{e^{-\beta E_v}}{Z},$$

where we have used Equations (45) and (46). We have labeled partition factors  $f_t \equiv f(x_{t_0}, \dots, x_t), t = t_0, \dots, t_f$ , to abbreviate. □

**Proof of Equations (51) and (52).** Equation (51) is obtained by taking sums in the pathway probability product decomposition of Equation (41) from the most recent event variable,  $x_{t_f}$ , to the furthest event variable,  $x_{t+1}$ .

With regard to the second part of Equation (51),

$$p(x_{t_0}, \dots, x_t) = \prod_{i=t_0}^t \frac{e^{-\beta E(x_i;x_{i-1}, \dots, x_{t_0})}}{Z(x_i;x_{i-1}, \dots, x_{t_0})} = \frac{e^{-\beta E(x_{t_0}, \dots, x_t)}}{\prod_{i=t_0}^t Z(x_i;x_{i-1}, \dots, x_{t_0})}.$$

For  $Z(x_{t_0}, \dots, x_t) \equiv \prod_{i=t_0}^t Z(x_i;x_{i-1}, \dots, x_{t_0})$ , we find

$$Z(x_{t_0}, \dots, x_t) = \prod_{i=t_0}^t \frac{\sum_{x_i} e^{-\beta E(x_i;x_{i-1}, \dots, x_{t_0})} f(x_{t_0}, \dots, x_i)}{f(x_{t_0}, \dots, x_i)} = \frac{Z}{f(x_{t_0}, \dots, x_t)},$$

where we have used Equation (46). □

**Proof of Equation (76).** The left part expands as:

$$\begin{aligned} \left\langle \frac{Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0})}{Z(x_t; x_{t-1}, \dots, x_{t_0})} \right\rangle_\lambda &= \sum_{x_{t_0}, \dots, x_t} p^{(\lambda)}(x_{t_0}, \dots, x_t) \frac{Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0})}{Z(x_t; x_{t-1}, \dots, x_{t_0})} \\ &= \sum_{x_{t_0}, \dots, x_{t-1}} p^{(\lambda)}(x_{t_0}, \dots, x_{t-1}) \sum_{x_t} p^{(\lambda)}(x_t|x_{t-1}, \dots, x_{t_0}) \frac{Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0})}{Z(x_t; x_{t-1}, \dots, x_{t_0})} \\ &= \sum_{x_{t_0}, \dots, x_{t-1}} p^{(\lambda)}(x_{t_0}, \dots, x_{t-1}) \sum_{x_t} p(x_t|x_{t-1}, \dots, x_{t_0}) = 1. \end{aligned}$$

The proof to the right part is formally analogous. □

**Appendix B. Proof of the Independence Limit Theorem for Instants, Truncated, and Full Pathways**

We begin with the proof for the theorem for instants. We will use the following inequality:

$$|e^z - 1| \leq e^{|z|} |z|, \tag{A3}$$

which was proven elsewhere [31]. For mild memory effects, we express the condition of Equation (90) as

$$\beta |E(x'_t; x'_{t-1}, \dots, x'_{t_0}) - E(x'_t; x_{t-1}, \dots, x_{t_0})| \leq \epsilon, \quad (\beta \equiv 1/(kT) > 0). \tag{A4}$$

Firstly, we demonstrate that

$$\frac{Z^{(\lambda)}(; x'_{t-1}, \dots, x'_{t_0})}{Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0})} \xrightarrow{\epsilon \rightarrow 0} 1. \tag{A5}$$

**Proof of Equation (A5).** From Equation (4),

$$\begin{aligned}
 & \left| Z^{(\lambda)}(; x'_{t-1}, \dots, x'_{t_0}) - Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0}) \right| = \left| \sum_{x'_t} \left[ e^{-\beta E(x'_t; x'_{t-1}, \dots, x'_{t_0})} - e^{-\beta E(x'_t; x_{t-1}, \dots, x_{t_0})} \right] \right| \\
 &= \left| \sum_{x'_t} e^{-\beta E(x'_t; x_{t-1}, \dots, x_{t_0})} \left[ e^{-\beta (E(x'_t; x'_{t-1}, \dots, x'_{t_0}) - E(x'_t; x_{t-1}, \dots, x_{t_0}))} - 1 \right] \right| \\
 &\leq \sum_{x'_t} e^{-\beta E(x'_t; x_{t-1}, \dots, x_{t_0})} \left| e^{-\beta (E(x'_t; x'_{t-1}, \dots, x'_{t_0}) - E(x'_t; x_{t-1}, \dots, x_{t_0}))} - 1 \right| \\
 &\leq \sum_{x'_t} e^{-\beta E(x'_t; x_{t-1}, \dots, x_{t_0})} \beta |E(x'_t; x'_{t-1}, \dots, x'_{t_0}) - E(x'_t; x_{t-1}, \dots, x_{t_0})| e^{\beta |E(x'_t; x'_{t-1}, \dots, x'_{t_0}) - E(x'_t; x_{t-1}, \dots, x_{t_0})|} \\
 &\leq \sum_{x'_t} e^{-\beta E(x'_t; x_{t-1}, \dots, x_{t_0})} \varepsilon e^\varepsilon = Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0}) \varepsilon e^\varepsilon,
 \end{aligned}$$

where we have used Equation (A3) followed by Equation (A4). Then,

$$\left| \frac{Z^{(\lambda)}(; x'_{t-1}, \dots, x'_{t_0})}{Z^{(\lambda)}(; x_{t-1}, \dots, x_{t_0})} - 1 \right| \leq \varepsilon e^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} 0,$$

which demonstrates Equation (A5). □

Secondly, we demonstrate that

$$\frac{Z(x'_t; x'_{t-1}, \dots, x'_{t_0})}{Z^{(\lambda)}(; x'_{t-1}, \dots, x'_{t_0})} \xrightarrow{\varepsilon \rightarrow 0} 1. \tag{A6}$$

**Proof of Equation (A6).** Using Equation (47), Equation (49) can be expressed as:

$$Z(x'_t; x'_{t-1}, \dots, x'_{t_0}) = \sum_{x_t} e^{-\beta E(x_t; x'_{t-1}, \dots, x'_{t_0})} \frac{f(x'_{t_0}, \dots, x'_{t-1}, x_t)}{f(x'_{t_0}, \dots, x'_t)}.$$

Then,

$$\begin{aligned}
 & \left| Z(x'_t; x'_{t-1}, \dots, x'_{t_0}) - Z^{(\lambda)}(; x'_{t-1}, \dots, x'_{t_0}) \right| = \left| \sum_{x_t} e^{-\beta E(x_t; x'_{t-1}, \dots, x'_{t_0})} \left[ \frac{f(x'_{t_0}, \dots, x'_{t-1}, x_t)}{f(x'_{t_0}, \dots, x'_t)} - 1 \right] \right| \\
 &\leq \sum_{x_t} e^{-\beta E(x_t; x'_{t-1}, \dots, x'_{t_0})} \left| \frac{f(x'_{t_0}, \dots, x'_{t-1}, x_t)}{f(x'_{t_0}, \dots, x'_t)} - 1 \right|. \tag{A7}
 \end{aligned}$$

To continue, we need to prove that

$$\frac{f(x'_{t_0}, \dots, x'_{t-1}, x_t)}{f(x'_{t_0}, \dots, x'_t)} \xrightarrow{\varepsilon \rightarrow 0} 1. \tag{A8}$$

For mild memory effects, we express the condition of Equation (91) as

$$\beta |\Delta E(x'_i; x'_{i-1}, \dots, x'_{t+1}, x'_t, x_t, x'_{t-1}, \dots, x'_{t_0})| \leq \varepsilon, \tag{A9}$$

for  $i = t + 1, \dots, t_f$ , where

$$\begin{aligned}
 & \Delta E(x'_i; x'_{i-1}, \dots, x'_{t+1}, x'_t, x_t, x'_{t-1}, \dots, x'_{t_0}) \\
 &\equiv E(x'_i; x'_{i-1}, \dots, x'_{t+1}, x'_t, x'_{t-1}, \dots, x'_{t_0}) - E(x'_i; x'_{i-1}, \dots, x'_{t+1}, x_t, x'_{t-1}, \dots, x'_{t_0}). \tag{A10}
 \end{aligned}$$

From Equation (44),

$$\begin{aligned} & \left| f(x'_{t_0}, \dots, x'_{t-1}, x_t) - f(x'_{t_0}, \dots, x'_t) \right| \\ &= \left| \sum_{x'_{t+1}, \dots, x'_{t_f}} \exp\left(-\beta \sum_{i=t+1}^{t_f} E(x'_i; x'_{i-1}, \dots, x'_{t_0})\right) \left[ \exp\left(\beta \sum_{i=t+1}^{t_f} \Delta E(x'_i; x'_{i-1}, \dots, x'_t, x'_t, x'_{t-1}, \dots, x'_{t_0})\right) - 1 \right] \right| \\ &\leq \sum_{x'_{t+1}, \dots, x'_{t_f}} \exp\left(-\beta \sum_{i=t+1}^{t_f} E(x'_i; x'_{i-1}, \dots, x'_{t_0})\right) \left| \exp\left(\beta \sum_{i=t+1}^{t_f} \Delta E(x'_i; x'_{i-1}, \dots, x'_t, x'_t, x'_{t-1}, \dots, x'_{t_0})\right) - 1 \right| \\ &\leq \sum_{x'_{t+1}, \dots, x'_{t_f}} \exp\left(-\beta \sum_{i=t+1}^{t_f} E(x'_i; x'_{i-1}, \dots, x'_{t_0})\right) (t_f - t) \varepsilon e^{(t_f-t)\varepsilon} = f(x'_{t_0}, \dots, x'_t) (t_f - t) \varepsilon e^{(t_f-t)\varepsilon}, \end{aligned}$$

where we have used Equation (A3) followed by Equation (A9). Then,

$$\left| \frac{f(x'_{t_0}, \dots, x'_{t-1}, x_t)}{f(x'_{t_0}, \dots, x'_t)} - 1 \right| \leq (t_f - t) \varepsilon e^{(t_f-t)\varepsilon} \xrightarrow{\varepsilon \rightarrow 0} 0,$$

which demonstrates Equation (A8). As a consequence, from Equation (A7),

$$\left| Z(x'_t; x'_{t-1}, \dots, x'_{t_0}) - Z^{(\lambda)}(x'_{t-1}, \dots, x'_{t_0}) \right| \leq Z^{(\lambda)}(x'_{t-1}, \dots, x'_{t_0}) (t_f - t) \varepsilon e^{(t_f-t)\varepsilon}.$$

Then,

$$\left| \frac{Z(x'_t; x'_{t-1}, \dots, x'_{t_0})}{Z^{(\lambda)}(x'_{t-1}, \dots, x'_{t_0})} - 1 \right| \leq (t_f - t) \varepsilon e^{(t_f-t)\varepsilon} \xrightarrow{\varepsilon \rightarrow 0} 0,$$

which demonstrates Equation (A6). □

Finally, multiplying Equation (A5) by Equation (A6), we find:

$$\frac{Z(x'_t; x'_{t-1}, \dots, x'_{t_0})}{Z^{(\lambda)}(x_{t-1}, \dots, x_{t_0})} \xrightarrow{\varepsilon \rightarrow 0} 1,$$

which completes the proof of the theorem for instants.

The proof for the truncated sequence version is straightforward from the above results using Equations (15) and (52):

$$\frac{Z(x'_{t_0}, \dots, x'_t)}{Z^{(\lambda)}(x_{t_0}, \dots, x_{t-1})} = \prod_{i=t_0}^t \frac{Z(x'_i; x'_{i-1}, \dots, x'_{t_0})}{Z^{(\lambda)}(x_{i-1}, \dots, x_{t_0})} \xrightarrow{\varepsilon \rightarrow 0} 1.$$

The proof for the full pathways follows from the Equations (8) and (50) and from the fact that  $Z/Z_v^{(\lambda)} = Z(x'_{t_0}, \dots, x'_{t_f})/Z^{(\lambda)}(x_{t_0}, \dots, x_{t_f-1})$  (also see Reference [26,31]).

### References

1. Chandler, D. *Introduction to Modern Statistical Mechanics*; Oxford University Press: New York, NY, USA, 1987.
2. Goldstein, H.; Poole, C.; Safko, J. *Classical Mechanics*, 3rd ed.; Addison Wesley: San Francisco, CA, USA, 2002.
3. Pathria, R.K.; Beale, P.D. *Statistical Mechanics*, 3rd ed.; Academic Press: Boston, MA, USA, 2011.
4. Crooks, G.E. On thermodynamic and microscopic reversibility. *J. Stat. Mech.* **2011**. [CrossRef]
5. Bustamante, C.; Liphardt, J.; Ritort, F. The Nonequilibrium Thermodynamics of Small Systems. *Phys. Today* **2005**, *58*, 43–48. [CrossRef]
6. Ritort, F. Nonequilibrium fluctuations in small systems: From physics to biology. *Adv. Chem. Phys.* **2008**, *137*, 31–123.
7. Bustamante, C. In singulo biochemistry: When Less Is More. *Annu. Rev. Biochem.* **2008**, *77*, 44–50. [CrossRef]
8. Arias-Gonzalez, J.R. Single-molecule portrait of DNA and RNA double helices. *Integr. Biol.* **2014**, *6*, 904–925. [CrossRef]

9. Bustamante, C.; Cheng, W.; Mejia, Y.X. Revisiting the Central Dogma One Molecule at a Time. *Cell* **2011**, *144*, 480–497. [[CrossRef](#)]
10. Moret, M.A.; Bisch, P.M.; Mundim, K.C.; Pascutti, P. G. New Stochastic Strategy to Analyze Helix Folding. *Biophys. J.* **2002**, *82*, 1123–1132. [[CrossRef](#)]
11. Alexander, L.M.; Goldman, D.H.; Wee, L.M.; Bustamante, C. Non-equilibrium dynamics of a nascent polypeptide during translation suppress its misfolding. *Nat. Commun.* **2019**, *10*, 2709. [[CrossRef](#)]
12. Goldt, S.; Seifert, U. Stochastic Thermodynamics of Learning. *Phys. Rev. Lett.* **2017**, *118*, 010601. [[CrossRef](#)]
13. Jarzynski, C. Stochastic and Macroscopic Thermodynamics of Strongly Coupled Systems. *Phys. Rev. X* **2017**, *7*, 011008. [[CrossRef](#)]
14. Strasberg, P.; Esposito, M. Stochastic thermodynamics in the strong coupling regime: An unambiguous approach based on coarse graining. *Phys. Rev. E* **2017**, *95*, 062101. [[CrossRef](#)] [[PubMed](#)]
15. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **2012**, *75*, 126001. [[CrossRef](#)] [[PubMed](#)]
16. Seifert, U. From Stochastic Thermodynamics to Thermodynamic Inference. *Annu. Rev. Condens. Matter Phys.* **2019**, *10*, 171–192. [[CrossRef](#)]
17. Tsallis, C. Possible Generalization of Boltzmann-Gibbs Statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [[CrossRef](#)]
18. Breuer, H.-P.; Laine, E.-M.; Piilo, J.; Vacchini, B. Colloquium: Non-Markovian dynamics in open quantum systems. *Rev. Mod. Phys.* **2016**, *88*, 021002. [[CrossRef](#)]
19. Rivas, A.; Huelga, S.F.; Plenio, M.B. Quantum non-Markovianity: Characterization, quantification and detection. *Rep. Prog. Phys.* **2014**, *77*, 094001. [[CrossRef](#)]
20. Taranto, P.; Pollock, F.A.; Milz, S.; Tomamichel, M.; Modi, K. Quantum Markov Order. *Phys. Rev. Lett.* **2019**, *122*, 140401. [[CrossRef](#)]
21. Thomas, G.; Siddharth, N.; Banerjee, S.; Ghosh, S. Thermodynamics of non-Markovian reservoirs and heat engines. *Phys. Rev. E* **2018**, *97*, 062108. [[CrossRef](#)]
22. Yu, S.; Wang, Y.-T.; Ke, Z.-J.; Liu, W.; Meng, Y.; Li, Z.-P.; Zhang, W.-H.; Chen, G.; Tang, J.-S.; Li, C.-F.; et al. Experimental Investigation of Spectra of Dynamical Maps and their Relation to non-Markovianity. *Phys. Rev. Lett.* **2018**, *120*, 060406. [[CrossRef](#)]
23. Ventéjou, B.; Sekimoto, K. Progressive quenching: Globally coupled model. *Phys. Rev. E* **2018**, *97*, 062150. [[CrossRef](#)]
24. Arias-Gonzalez, J.R. Thermodynamic framework for information in nanoscale systems with memory. *J. Chem. Phys.* **2017**, *147*, 205101. [[CrossRef](#)]
25. Arias-Gonzalez, J.R. Writing, Proofreading and Editing in Information Theory. *Entropy* **2018**, *20*, 368. [[CrossRef](#)] [[PubMed](#)]
26. Arias-Gonzalez, J.R. Information management in DNA replication modeled by directional, stochastic chains with memory. *J. Chem. Phys.* **2016**, *145*, 185103. [[CrossRef](#)] [[PubMed](#)]
27. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 1st ed.; Wiley-Interscience: Hoboken, NJ, USA, 1991.
28. Fisz, M. *Probability Theory and Mathematical Statistics*, 3rd ed.; Krieger Publishing Company: Malabar, FL, USA, 1980.
29. Seifert, U. Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Phys. Rev. Lett.* **2005**, *95*, 040602. [[CrossRef](#)]
30. Arias-Gonzalez, J.R. Entropy involved in fidelity of DNA replication. *PLoS ONE* **2012**, *7*, e42272. [[CrossRef](#)] [[PubMed](#)]
31. Arias-Gonzalez, J.R.; Aleja, D. Comment on “Information management in DNA replication modeled by directional, stochastic chains with memory” [*J. Chem. Phys.* *145*, 185103 (2016)]. *J. Chem. Phys.* **2020**, *152*, 047101. [[CrossRef](#)]