Review

Information Geometry, Fluctuations, Non-Equilibrium Thermodynamics, and Geodesics in Complex Systems

Eun-jin Kim

Center for Fluid and Complex Systems, Coventry University, Priory St, Coventry CV1 5FB, UK; ejk92122@gmail.com

Abstract: Information theory provides an interdisciplinary method to understand important phenomena in many research fields ranging from astrophysical and laboratory fluids/plasmas to biological systems. In particular, information geometric theory enables us to envision the evolution of non-equilibrium processes in terms of a (dimensionless) distance by quantifying how information unfolds over time as a probability density function (PDF) evolves in time. Here, we discuss some recent developments in information geometric theory focusing on time-dependent *dynamic* aspects of non-equilibrium processes (e.g., time-varying mean value, time-varying variance, or temperature, etc.) and their thermodynamic and physical/biological implications. We compare different distances between two given PDFs and highlight the importance of a path-dependent distance for a time-dependent PDF. We then discuss the role of the information rate $\Gamma = \frac{d\mathcal{L}}{dt}$ and relative entropy in non-equilibrium thermodynamic relations (entropy production rate, heat flux, dissipated work, non-equilibrium free energy, etc.), and various inequalities among them. Here, $\mathcal{L}$ is the information length representing the total number of statistically distinguishable states a PDF evolves through over time. We explore the implications of a geodesic solution in information geometry for self-organization and control.

1. Introduction

Information geometry refers to the application of the techniques of differential geometry to probability and statistics. Specifically, it uses differential geometry to define the metric tensor that endows the statistical space (consisting of probabilities) with the notion of distance [1–31]. While seemingly too abstract, it permits us to measure quantitative differences among different probabilities. It then makes it possible to link a stochastic process, complexity, and geometry, which is particularly useful in classifying a growing number of data from different research areas (e.g., from astrophysical and laboratory systems to biosystems). Furthermore, it can be used to obtain desired outcomes [6–10,15] or to understand statistical complexity [4].

For instance, the Wasserstein metric [6–10] was widely used in the optimal transport problem where the main interest is to minimize transport cost which is a quadratic function of the distance between two locations. It satisfies the Fokker-Planck equation for gradient flow which minimizes the entropy/energy functional [7]. For Gaussian distributions, the Wasserstein metric space consists of *physical distances*—Euclidean and positive symmetric matrices for the mean and variance, respectively (e.g., see [8]).

In comparision, the Fisher (Fisher-Rao) information [32] can be used to define a *dimensionless* distance in statistical manifolds [33,34]. For instance, the statistical distance *ds* represents the number of indistinguishable states as [5,33]

$$(ds)^2 \equiv \sum_j \frac{dp_j^2}{p_j} = \sum_j p_j (d \ln p_j)^2 = \sum_{j,\alpha,\beta} p_j \frac{\partial \ln p_j}{\partial \lambda^\alpha} \frac{\partial \ln p_j}{\partial \lambda^\beta} d\lambda^\alpha d\lambda^\beta = \sum_{\alpha,\beta} d\lambda^\alpha g_{\alpha\beta} d\lambda^\beta. \quad (1)$$

Here the Fisher information metric $g_{\alpha\beta} = \langle \frac{\partial \ln p_j}{\partial \lambda^\alpha} \frac{\partial \ln p_j}{\partial \lambda^\beta} \rangle = \sum_j p_j \frac{\partial \ln p_j}{\partial \lambda^\alpha} \frac{\partial \ln p_j}{\partial \lambda^\beta}$ provides natural (Riemannian) distinguishability metric on the space of probability distributions. λ^α's are the parameters of the probability p_j and the angular brackets represent the ensemble average over p_j. Note that Equation (1) is given for a discrete probability p_j. For a continuous Probability Density Function (PDF) $p(x)$ for a variable x, Equation (1) becomes $(ds)^2 = \int dx p(x)[d \ln(p(x)]^2 = \sum_{\alpha,\beta} d\lambda^\alpha g_{\alpha\beta} d\lambda^\beta$ where $g_{\alpha\beta} = \int dx \, p(x) \frac{\partial \ln(p(x))}{\partial \lambda^\alpha} \frac{\partial \ln(p(x))}{\partial \lambda^\beta}$.

For Gaussian processes, the Fisher metric is inversely proportional to the covariance matrices of fluctuations in the systems. Thus, in thermodynamic equilibrium, strong fluctuations lead to a strong correlation and a shorter distance between the neighboring states [34,35]. Alternatively, fluctuations determine the uncertainty in measurements, providing the resolution (the distance unit) that normalizes the distance between different thermodynamic states.

To appreciate the meaning of fluctuation-based metric, let us consider the (equilibrium) Maxwell-Boltzmann distribution $p(E_j) = \beta e^{-\beta E_j}$ for the energy state E_j

$$\frac{p(E_i)}{p(E_j)} = e^{-\beta(E_i - E_j)}. \quad (2)$$

Here $\beta = 1/k_B T$ is the inverse temperature; k_B is the Boltzmann constant; T is the temperature of the heat bath. In Equation (2), the thermal energy $k_B T = \langle E \rangle$ of the heat bath (the width/uncertainty of the probability) provides the resolution to differentiate different states $\Delta E = E_i - E_j$. The smaller is the resolution (temperature), the more distinguishable states (more accessible information in the system) there are. It agrees with the expectation that a PDF gradient (the Fisher-information) increases with information [32].

This concept has been generalized to non-equilibrium systems [36–43], including the utilization for controlling systems to minimize entropy production [38,40,42], the measurement of the statistical distance in experiments to validate theoretical predictions [41], etc. However, some of these works rely on the equilibrium distribution Equation (2) that is valid only in or near equilibrium while many important phenomena in nature and laboratories are often far from equilibrium with strong fluctuations, variability, heterogeneity, or stochasticity [44–52]. Far from equilibrium, there is no (infinite-capacity) heat bath that can maintain the system at a certain temperature, or constant fluctuation level. One of the important questions far from equilibrium is indeed to understand how fluctuation level $\beta(t)^{-1}$ changes with time. Furthermore, PDFs no longer follow the Maxwell-Boltzmann nor Gaussian distributions and can involve the contribution from (rare) events of large amplitude fluctuations [53–62]. Therefore, the full knowledge of time-varying PDFs and the application of information geometry to such PDFs have become of considerable interest.

Furthermore, while in equilibrium [63,64], information theoretical measures (e.g., Shannon information entropy) can be given thermodynamic meanings (e.g., heat), in non-equilibrium such interpretations are not always possible and equilibrium thermodynamic rules can break down locally (e.g., see [65,66] and references therein). Much progress on these issues has been made by different authors (e.g., [65–82]) through the development of information theory, stochastic thermodynamics, and non-equilibrium fluctuation theorems with the help of the Fisher information [32], relative entropy [83], mutual information [84,85], etc. Exemplary works include the Landauer's principle which links information loss to the ability to extract work [86,87]; the resolution of Maxwell's demon paradox [88]; black hole thermodynamics [89,90]; various thermodynamic inequal-

ity/uncertainty relations [65,68,91–97]; and linking different research areas (e.g., non-equilibrium processes to quantum mechanics [98–100], physics to biology [101]).

The paper aims to discuss some recent developments in the information geometric theory of non-equilibrium processes. Since this would undoubtedly span a broad range of topics, this paper will have to be selective and will focus on elucidating the *dynamic* aspect of non-equilibrium processes and thermodynamic and physical/biological implications. Throughout the paper, we highlight that time-varying measures (esp. variance) introduces extra complication in various relations, in particular, between the information geometric measure and entropy production rate. We make the efforts to make this paper self-contained (e.g., by including the derivations of some well-known results) wherever possible.

The remainder of this paper is organized as follows. Section 2 discusses different distances between two PDFs and the generalization for a time-dependent non-equilibrium PDF. Section 3 compares the distancs from Section 2. Section 4 discusses key thermodynamic relations that are useful for non-equilibrium processes. Section 5 establishes relations between information geometric quantities (in Section 2) and thermodynamics (in Section 4). In Section 6, we discuss the concept of a geodesic in information geometry and its implications for self-organization or designing optimal protocols for control. Conclusions are provided in Section 7.

2. Distances/Metrics

This section discusses the distance defined between two probabilities (Section 2.1) and along the evolution path of a time-dependent probability (Section 2.2). Examples and comparisons of these distances are provided in Section 3.1. For illustration, we use a PDF $p(x,t)$ of a stochastic variable x and differential entropy $S = -\int dx\, p(x,t) \ln(p(x,t))$ by using the unit $k_B = 1$.

2.1. Distance Between Two PDFs

We consider the distance between two PDFs $p_1 = p(x,t_1)$ and $p_2 = p(x,t_2)$ of a stochastic variable x at two times t_1 and t_2, respectively where $t_1 = t_2$ or $t_1 \neq t_2$ in general.

2.1.1. Wootters' Distance

The Wootters' distance [5,33] is defined in quantum mechanics by the shortest distance between the two p_1 and p_2 that have the wave functions ψ_1 and ψ_2 ($p_1 = |\psi_1|^2$ and $p_2 = |\psi_2|^2$), respectively. Specifically, for given p_1 and p_2, the distance $s(p_1, p_2)$ between p_1 and p_2 can be parameterized by infinitely many different paths between p_1 and p_2. Letting z be the affine parameter of a path, we have

$$s(p_1, p_2) = \int_1^2 ds = \int dz \frac{ds(z)}{dz} = \int dz \sqrt{\sum_{\alpha,\beta} \frac{d\lambda^\alpha}{dz} g_{\alpha\beta} \frac{d\lambda^\beta}{dz}}, \qquad (3)$$

where ds is given in Equation (1). Among all possible paths, the minimum of $s(p_1, p_2)$ is obtained for a particular path that optimizes the quantum distinguishability; the (Hilbert-space) angle between the two wave functions provides such minimum distance as

$$W[p_1, p_2] = \cos^{-1}\left[\int dx\, [p(x,t_1)]^{\frac{1}{2}} [p(x,t_2)]^{\frac{1}{2}}\right]. \qquad (4)$$

Equation (4) is for a pure state and has been generalized to mixed states (e.g., see [37,102] and references therein). Note that the Wootters' distance is related to the Hellinger distance [43].

2.1.2. Kullback-Leibler (K-L) Divergence/Relative Entropy

Kullback-Leibler (K-L) divergence between the two PDFs [83], also called relative entropy, is defined by

$$\mathcal{K}(p_1|p_2) \;=\; \int dx\, p(x,t_1) \ln\left(\frac{p(x,t_1)}{p(x,t_2)}\right). \tag{5}$$

Relative entropy quantifies the difference between a PDF p_1 and another PDF p_2. It takes the minimum zero value for identical two PDFs $p_1 = p_2$ and becomes large as p_1 and p_2 become more different. However, as it is defined in Equation (5), it is not symmetric between p_1 and p_2 and does not satisfy the triangle inequality. It is thus not a metric in a strict sense.

2.1.3. Jensen Divergence

The Jensen divergence (also called Jensen distance) is the symmetrized Kullback–Leibler divergence defined by

$$\mathcal{J}(p_1|p_2) \;=\; \frac{1}{2}\Big[\mathcal{K}(p_1|p_2) + \mathcal{K}(p_2|p_1))\Big]. \tag{6}$$

While the square root of the Jensen-Shannon divergence $\sqrt{\mathcal{J}(p_1|p_2)}$ is a metric [4,103], $\mathcal{J}(p_1|p_2)$ itself has also been used in examining statistical complexity (e.g., see [43,104,105]).

2.1.4. Euclidean Norm

In analysis of big data, the Euclidean norm [5,106] is used, which is defined by

$$|p_1 - p_2|^2 \;=\; \int dx\, \big[p(x,t_1) - p(x,t_2)\big]^2. \tag{7}$$

While Equation (7) has a direct analogy to the physical distance, it has a limitation in measuring statistical complexity due to the neglect of the stochastic nature [5]. For instance, the Wootters' distance in Equation (4) was shown to work better than Equation (7) in capturing complexity in the logistic map [5].

2.2. *Distance along the Path*

Equations (4)–(7) can be used to define the distance between the two given PDFs $p(x,t_1)$ and $p(x,t_2)$ at times t_1 and t_2 ($t_2 > t_1$). However, $p(x,t)$ at the intermediate time $t = (t_1, t_2)$ can take an infinite number of different values depending on the exact path that a system takes between $p(x,t_1)$ and $p(x,t_2)$. One example would be i) $p(x,t_1) = p(x,t_2) = p(x,t)$ for all $t = (t_1, t_2)$ and x, in comparison with ii) $p(x,t_1) = p(x,t_2)$ but $p(x,t) \neq p(x,t_1)$ and $p(x,t) \neq p(x,t_2)$. What is necessary is a *path-dependent* distance that depends on the exact evolution and the form of $p(x,t)$ for $t = (t_1, t_2)$.

2.2.1. Information Rate

Calculating a path-dependent *distance* for a time-dependent PDF $p(x,t)$ requires the generalization of the distance in Section 2.1. To this end, we consider two (temporally) adjacent PDFs along the trajectory, say, $p(x,t)$ and $p(x,t+dt)$ and calculate the (infinitesimal) relative entropy between them in the limit $dt \to 0$ to the leading order in $O(dt)$:

$$\lim_{dt\to 0} \frac{1}{(dt)^2}\mathcal{K}[p(x,t+dt)|p(x,t)] = \lim_{dt\to 0} \frac{1}{(dt)^2}\mathcal{K}[p(x,t)|p(x,t+dt)]$$

$$= \lim_{dt\to 0} \frac{1}{(dt)^2}\mathcal{J}[p(x,t+dt)|p(x,t)] = \frac{1}{2}\int dx\, p(x,t)(\partial_t \ln p(x,t))^2 \equiv \frac{1}{2}\Gamma^2. \tag{8}$$

Here, we used $p(x, t + dt) = p(x, t) + (dt)\partial_t p + \frac{1}{2}(dt)^2(\partial_t p)^2 + O((dt)^3)$, $\ln(1 + r) = r - \frac{1}{2}r^2 + O(r^3)$ for $r \ll 1$, and $\int dx \partial_t p(x, t) = \int dx \partial_{tt} p(x, t) = 0$ because of the total probability conservation $\int dx p(x, t) = 1$. Due to the symmetry of $\mathcal{K}[p(x, t + dt)|p(x, t)]$ to leading order $O((dt)^2)$, $\mathcal{K}[p(x, t + dt)|p(x, t)] = \mathcal{J}[p(x, t + dt)|p(x, t)]$ to $O((dt)^2)$.

In Equation (8), the information rate Γ is defined by [15–29]

$$\Gamma^2(t) = \lim_{dt \to 0} \frac{2}{(dt)^2} J[p(x, t + dt)|p(x, t)] = \int dx p(x, t)(\partial_t \ln p(x, t))^2 = 4 \int dx \, (\partial_t q(t))^2. \tag{9}$$

Here, $q = \sqrt{p}$, and $\Gamma \geq 0$ by definition. We note that the last term in terms of q in Equation (9) can be used when $q = p = 0$. The dimensions of $\Gamma^2 \equiv \mathcal{E}$ and Γ are $(\text{time})^{-2}$ and $(\text{time})^{-1}$, respectively. They do not change their values under nonlinear, time-independent transformation of variables (see Appendix A). Thus, using the unit where the length is dimensionless, $\mathcal{E}$ and Γ can be viewed as the kinetic energy per mass and velocity, respectively. For this reason, Γ was called the velocity (e.g., in [15,17]).

Note that $\mathcal{E}$ can be viewed as the Fisher information [32] if time is interpreted as a parameter (e.g., [97]). However, time in classical mechanics is a passive quantity that cannot be changed by an external control. Γ is also called the entropy production rate in quantum mechanics [107]. However, as shown in Sections 4.1 and 4.4, the relation between Γ and thermodynamic entropy production rate is more complicated (see Equation (28)).

Γ in Equation (9) is the information rate representing how quickly new information is revealed as a PDF evolves in time. Here, $\Gamma^{-1} = \tau$ is the characteristic time scale of this information change in time. To show that Γ is related to fluctuation's smallest time scale [97], we assume that λ^α's are the estimators (parameters) of a $p(x, t)$ and use the Cramér-Rao bound on the Fisher information $g_{\alpha\beta} = \int dx p(x, t) \partial_{\lambda^\alpha}[\ln p(x, t)] \partial_{\lambda^\beta}[\ln p(x, t)] \geq C_{\alpha\beta}^{-1}$ where $C^{\alpha\beta} \equiv \langle \delta\lambda^\alpha \delta\lambda^\beta \rangle$ is the covariance matrix (e.g., see [32]); $\delta\lambda^\alpha = \lambda^\alpha - \langle \lambda^\alpha \rangle$ denotes fluctuation. Using $\frac{d \ln p}{dt} = \frac{\partial \ln p}{\partial \lambda^\alpha} \frac{d\lambda^\alpha}{dt}$ then leads to

$$\Gamma^2 = \sum_{\alpha\beta} \frac{d\lambda^\alpha}{dt} g_{\alpha\beta} \frac{d\lambda^\beta}{dt} \geq \sum_{\alpha\beta} \frac{d\lambda^\alpha}{dt} C_{\alpha\beta}^{-1} \frac{d\lambda^\beta}{dt}. \tag{10}$$

For the diagonal $g_{\alpha\beta} = g_\alpha \delta_{\alpha\beta}$, Equation (10) is simplified as

$$\sum_\alpha \frac{1}{\langle (\delta\lambda^\alpha)^2 \rangle} \left(\frac{d\lambda^\alpha}{dt} \right)^2 \leq \Gamma^2. \tag{11}$$

Equation (11) shows how the RMS fluctuation-normalized rate at which the parameter λ^α can change is bounded above by Γ. If there is only $\alpha = 1$ ($\lambda^\alpha = \lambda \delta_{\alpha,1}$), Equation (11) is further simplified:

$$\frac{1}{\sqrt{\langle (\delta\lambda)^2 \rangle}} \left| \frac{d\lambda}{dt} \right| \leq \Gamma, \tag{12}$$

clearly showing that λ normalized by its RMS fluctuations cannot change faster than the information rate.

Finally, it is worth highlighting that Equation (9) is general and can be used even when the parameters λ^α's and $g_{\alpha\beta}$ in Γ^2 in Equation (10) are unknown. Examples include the cases where PDFs are empirically inferred from experimental/observational data. Readers are referred to Refs. [21,23,28] for examples. It is only the special case where we have a complete set of parameters λ^α's of a PDF that we can express Γ using the Fisher information as in Equation (10). For instance, for a Gaussian $p(x, t)$ that is fully described by the mean value $\langle x \rangle$ and variance $\frac{1}{2\beta}$, $(\lambda^1, \lambda^2) = (\langle x \rangle, \beta)$.

2.2.2. Information Length

Since $\Gamma \propto \sqrt{\mathcal{J}[p(x, t + dt)|p(x, t)]}$ is a metric [103] as noted in Section 2.1, Γ is also a metric. Thus, we sum Γ along the trajectory to define a finite distance. Specifically, starting with an initial PDF $p(x, t = 0)$, we integrate $\Gamma(t)$ over time to obtain the dimensionless number as a function time as

$$\mathcal{L}(t) \quad = \quad \int_0^t dt_1 \Gamma(t_1). \tag{13}$$

$\mathcal{L}$ is the information length [15–31] that quantifies the total change in information along the trajectory of $p(x, t)$ or the total number of statistically distinguishable states it evolves through over time. [We note that different names (e.g., statistical length [108], or statistical distance [97]) were also used for $\mathcal{L}$.] It is important to note that unlike the Wootters' distance (the shortest distance among all possible paths between the two PDFs) in Equation (3) (e.g., [5]), $\mathcal{L}(t)$ in Equation (13) is fixed for a given time-evolving PDF $p(x, t)$.

By definition in Equation (13), $\mathcal{L}(t = 0) = 0$ and $\mathcal{L}(t)$ monotonically increases with time since $\Gamma \geq 0$ (e.g., see Figure A2 in [22]). $\mathcal{L}(t)$ takes a constant value only in a stationary state ($\Gamma = \partial_t p = 0$). One of its important consequences is that when $p(x, t)$ relaxes into a stationary PDF in the long time limit $t \to \infty$, $\Gamma(t) \to 0$ and $\mathcal{L}(t) \to \mathcal{L}_\infty$ as $t \to \infty$ where $\mathcal{L}_\infty$ is a constant depending on initial conditions and parameters. This property of $\mathcal{L}_\infty$ was used to understand attractor structure in a relaxation problem; specifically, Refs. [15,16,18,22,28] calculated $\mathcal{L}_\infty$ for different values of the mean position x_0 of an initial PDF and examined how $\mathcal{L}_\infty$ depends on x_0. Furthermore, Γ and $\mathcal{L}$ were shown to be useful for quantifying hysteresis in forward-backward processes [19], correlation and self-regulation among different players [23,25], and predicting the occurrence of sudden events [27] and phase transitions [23,25]. Some of these points are illustrated in Section 3.1.

3. Model and Comparison of Metrics

For discussing/comparing different metrics in Section 2 and statistical measures in Section 4, we use the following Langevin model [109]

$$\frac{dx}{dt} = f(x, t) + \xi = -\partial_x V(x, t) + \xi. \tag{14}$$

Here, $V(x, t)$ is, in general, a time-dependent potential which can include an internal potential and an external force; ξ is a short (delta)-correlated Gaussian noise with the following statistical property

$$\langle \xi(t) \xi(t') \rangle = 2D\delta(t - t'). \tag{15}$$

Here, the angular brackets represent the ensemble average over the stochastic noise ξ; $D \geq 0$ is the amplitude of ξ. It is important to note that far from equilibrium, the average (e.g., $\langle x(t) \rangle$) is a function of time, in general.

The exact PDFs can be obtained for the Ornstein-Uhlenbeck (O-U) process which has $V = \frac{\gamma}{2}(x - v(t))^2$ and $f = -\gamma(x - v(t))$ in Equation (14). Here, $v(t)$ is a deterministic function of time. Specifically, for the initial Gaussian PDF $p(x, 0)$

$$p(x, 0) = \sqrt{\frac{\beta_0}{\pi}} e^{-\beta_0 (x - x_0)^2}, \tag{16}$$

a time-dependent PDF remains Gaussian at all time:

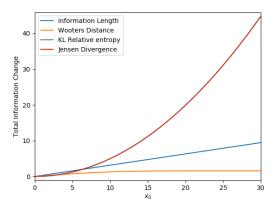$$p(x, t) \quad = \quad \int dx_1 p(x, t; x_1, 0) p(x_1, 0) = \sqrt{\frac{\beta}{\pi}} e^{-\beta(x - \langle x \rangle)^2}, \tag{17}$$

$$\frac{1}{2\beta(t)} \quad = \quad \frac{e^{-2\gamma t}}{2\beta_0} + \frac{D(1 - e^{-2\gamma t})}{\gamma}, \tag{18}$$

$$\langle x(t) \rangle \;\; = \;\; x_0 e^{-\gamma t} + \gamma \int_0^t dt_1 \, e^{-\gamma(t-t_1)} v(t_1). \tag{19}$$

In Equations (16)–(19), $x_0 = \langle x(t=0) \rangle$, $\beta_0 = \beta(t=0)$, and $\langle (\delta x)^2 \rangle = \frac{1}{2\beta} = \sigma^2$. Here, β, σ and σ^2 are the inverse temperature, standard deviation, and variance, respectively; β_0 and x_0 are the values of β and $\langle x \rangle$, respectively, at $t = 0$. Equation (18) shows that as $t \to \infty$, $\beta(t \to \infty) = \frac{\gamma}{2D}$. Note that we use both β and σ here to clarify the connections to the previous works [15,17,22,26–28].

3.1. Geometric Structure of Equilibrium/Attractors

To elucidate the main difference between the distances in Equations (4)–(6) and (13), we consider the relaxation problem by assuming $v(t) = 0$. In the following, we compare the distance between $p(x,0)$ and $p(x,t \to \infty)$ by using $p_1(x,t_1) = p(x,0)$ and $p_2(x,t_2) = p(x,t \to \infty)$ in Equations (4)–(6) and Equation (13). Analytical expressions for these distances are given in [22].
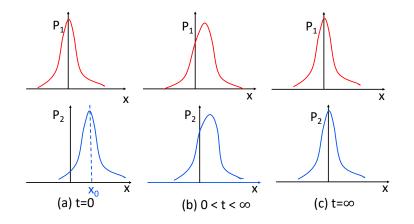
Each curve in Figure 1 shows how each distance depends on the initial mean position x_0. The four different curves are for $\mathcal{L}_\infty$ (in blue), Wootters' distance (in orange), K-L relative entropy (in green), and Jensen divergence (in red), respectively. The relative entropy and Jensen divergence exhibit similar behavior, the red and green color curves being superimposed on each other. Of note is a linear relation between $\mathcal{L}_\infty$ and x_0 in Figure 1. Such linear relation is not seen in other distances. This means that the information length is a unique measure that manifests a linear geometry around its equilibrium point in a linear Gaussian process [28,30]. Note that for a nonlinear force f, $\mathcal{L}_\infty$ has a power-law relation with x_0 for a sufficiently large x_0 [18,28]. These contrast with the behaviour in a chaotic system [16,28] where $\mathcal{L}_\infty$ depends sensitively on the initial condition and abruptly changes with x_0. Thus, the information length provides a useful tool to geometrically understand attractor structures in relaxation problems.

Figure 1. The distance against x_0 between $p(x,0)$ and $p(x,t \to \infty)$ for the O-U process. (Figure 1 in [22]).

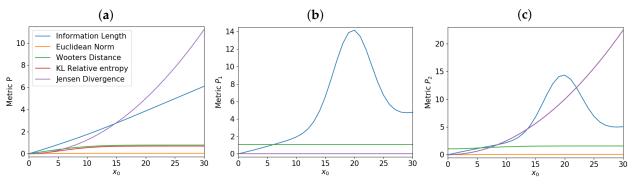### 3.2. Correlation between Two Interacting Components

We next show that the information length is also useful in elucidating the correlation between two interacting species such as two competing components relaxing to the same equilibrium in the long time limit. Specifically, the two interacting components with the time-dependent PDFs $P_1(x,t)$ and $P_2(x,t)$ are coupled through the Dichotomous noise [110,111] (see Appendix B) and relax into the same equilibrium Gaussian PDF $P_1(x,t \to \infty) = P_2(x,t \to \infty) = \frac{1}{2}P(x,t \to \infty)$ around $x = 0$ in the long time limit. Here, $P(x,t) = P_1(x,t) + P_2(x,t)$ is the total PDF. For the case considered below, $P(x,t)$ satisfies the O-U process (see Appendix B for details). We choose the initial conditions where $P_1(t=0) = P_1(t \to \infty)$ with zero initial mean value while $P_2(t=0)$ takes an initial mean value x_0. These are demonstrated in the cartoon figure, Figure 2a,c.

Figure 2. P_1 (top) and P_2 (bottom) at time $t = 0$ in panel (**a**), $t = (0, \infty)$ in panel (**b**), and $t \to \infty$ in panel (**c**). Note that $P_1(0 < t < \infty) \neq P_1(t = 0) (= P_1(t \to \infty))$.

Although $P_1(t = 0) = P_1(t \to \infty)$, at the intermediate time $t = (0, \infty)$, $P_1(x, t)$ evolves in time due to its coupling to P_2 and thus $P_1(x, t) \neq P_1(x, t = 0)$, as shown in Figure 2b. Consequently, $\mathcal{L}(t)$ calculated from P_1 monotonically increases to its asymptotic value $\mathcal{L}_\infty$ until it reaches the equilibrium (see Figure A2 in [22] for time-evolution of $\mathcal{L}$ from P_1 and P_2). On the other hand, P_2 with an initial mean value x_0 undergoes a different time evolution (unless $x_0 = 0$) until it reaches the equilibrium.

The distances in Equations (4)–(7) and (13) can be calculated from the total $P = P_1 + P_2$, P_1 and P_2 for different values of x_0. Results are shown in Figure 3a–c, respectively; (a) $P(x, 0)$ and $P(x, t \to \infty)$, (b) $P_1(x, 0)$ and $P_1(x, t \to \infty)$, and (c) $P_2(x, 0)$ and $P_2(x, t \to \infty)$, respectively. Specifically, for each value of x_0, we calculate the distances in Equations (4)–(7) and (13) by using $p_1(x, t_1) = P(x, 0)$ and $p_2(x, t_2) = P(x, t \to \infty)$ for Figure 3a; $p_1(x, t_1) = P_1(x, 0)$ and $p_2(x, t_2) = P_1(x, t \to \infty)$ for Figure 3b; $p_1(x, t_1) = P_2(x, 0)$ and $p_2(x, t_2) = P_2(x, t \to \infty)$ for Figure 3c. The same procedure above is then repeated for many other x_0's to show how each distance depends on x_0.

Figure 3. The distance between $P(x, 0)$ and $P(x, t \to \infty)$ against in x_0 in (**a**); $P_1(x, 0)$ and $P_2(x, t \to \infty)$ in (**b**); $P_2(x, 0)$ and $P_2(x, t \to \infty)$ in (**c**). (Figure 4 in [22]).

For the total P, a linear relation between $\mathcal{L}_\infty$ and x_0 is seen in Figure 3a (like in Figure 1). This linear relation is not seen in $\mathcal{L}_\infty$ calculated from either P_1 or P_2 in Figure3b or Figure 3c; a non-monotonic dependence of $\mathcal{L}_\infty$ in Figure 3b,c is due to large-fluctuations and strong-correlation between P_1 and P_2 during time-evolution for large x_0. What is quite remarkable is that in contrast to other distances, $\mathcal{L}_\infty$ calculated from P_1 and P_2 in Figure 3b,c exhibits a very similar dependence on x_0. It means that despite very different time-evolutions of P_1 and P_2 (see Figure 2), they undergo similar total change in information. These results suggest that strong coupling between two components can be inferred from their similar information length (see also [24,25]).

4. Thermodynamic Relations

To elucidate the utility of information geometric theory in understanding non-equilibrium thermodynamics, we review some of the important thermodynamic measures of irreversibility and dissipation [112] and relate them to information geometric measures Γ and $\mathcal{K}$ [29]. For illustration below, we use the model in Equations (14) and (15) unless stated otherwise. Corresponding to Equations (14) and (15) is the following Fokker-Planck equation [109]

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[f(x,t)p(x,t)\right] + D\frac{\partial^2 p(x,t)}{\partial x^2} = -\partial_x J(x,t), \tag{20}$$

where $J = fp - D\partial_x p$ is the probability current.

4.1. Entropy Production Rate and Flow

For non-equilibrium thermodynamics, we need to consider the entropy in the system S and the environment S_m, and the total entropy $S_T = S + S_m$. To clarify the difference among these, we go over some derivation by using $\partial_t p = -\partial_x J$ and $J = fp - D\partial_x p$ to obtain

$$\dot{S} = \frac{dS(x,t)}{dt} = -\int dx \partial_t p \ln p = \frac{dS_T(x,t)}{dt} - \frac{dS_m(x,t)}{dt}, \tag{21}$$

where,

$$\dot{S}_T = \frac{dS_T}{dt} = \int dx \left(\frac{1}{Dp}J^2\right), \qquad \dot{S}_m = \frac{dS_m}{dt} = \int dx \left(\frac{1}{D}Jf\right). \tag{22}$$

Here, we used integration by parts in t and x. $\dot{S}_T$ denotes the (total) entropy production rate, which is non-negative $\dot{S}_T \geq 0$ by definition, and serves as a measure of irreversibility [112]. The sign of $\dot{S}_m$ in Equation (22) represents the direction in which the entropy flows between the system and environment. Specifically, $\dot{S}_m > 0$ ($\dot{S}_m < 0$) when the entropy flows from the system (environment) to the environment (system). $\dot{S}_m$ is related to the heat flux $Q = D S_m$ from the system to the environment. The equality $\dot{S}_T = 0$ holds in an equilibrium reversible process. In this case, $\dot{S} = -\dot{S}_m = -\frac{Q}{D}$, which is the usual equilibrium thermodynamic relation. In comparison, when $\dot{S} = 0$, $\dot{S}_T = \dot{S}_m \geq 0$.

For the O-U process with $V = \frac{\gamma}{2}(x - v(t))^2$ and $f = -\gamma(x - v(t))$ in Equations (14), (17)–(19), (21) and (22) lead to (see [29] for details)

$$S(t) = \frac{1}{2}\left[1 + \ln\frac{\pi}{\beta}\right], \tag{23}$$

$$D\dot{S}_T = = \frac{(\partial_t\beta)^2}{8\beta^3} + (\partial_t\langle x\rangle)^2 = (\partial_t\sigma)^2 + (\partial_t\langle x\rangle)^2, \tag{24}$$

$$\dot{S} = -\frac{\partial_t\beta}{2\beta} = \frac{\partial_t\sigma}{\sigma}, \tag{25}$$

$$\dot{S}_m = \dot{S}_T - \dot{S}. \tag{26}$$

Here, we used $J = [f + 2D\beta(\delta x)]p = \left[-\frac{\partial_t\beta}{2\beta}(\delta x) + \partial_t\langle x\rangle\right]p$, $\partial_x p = -2\beta(\delta x)p$, $f = -\gamma(x - v(t))$, $\partial_t\langle x\rangle = \langle f\rangle$, $2D\beta - \gamma = -\frac{\partial_t\beta}{2\beta}$, and $\frac{\partial_t\beta}{\beta} = -2\frac{\partial_t\sigma}{\sigma}$.

In order to relate these thermodynamical quantities $\dot{S}_T$ and $\dot{S}$ above to the information rate Γ, we recall that for the O-U process [15,17,26–28],

$$\mathcal{E} = \Gamma^2 = 2\beta(\partial_t\langle x\rangle)^2 + \frac{(\partial_t\beta)^2}{2\beta^2} = \frac{1}{\sigma^2}\left[2(\partial_t\sigma)^2 + (\partial_t\langle x\rangle)^2\right]. \tag{27}$$

Equations (24) and (27) then give us

$$\Gamma^2 = \frac{D}{\sigma^2}\dot{S}_T + \dot{S}^2. \tag{28}$$

Interestingly, Equation (28) reveals that the entropy production rate needs be normalized by variance σ^2. This is because of the extensive nature of $\dot{S}_T$ unlike Γ or $\dot{S}$. That is, $\dot{S}_T$ changes its value when the variable x is rescaled by a scalar factor, say, α (> 0) as $y = \alpha x$. Furthermore, Equation (28) shows that the information rate Γ in general does not have a simple relation to the entropy production rate (c.f., [107]).

One interesting limit of Equation (28) is the case of constant $\beta(t)$ with $\dot{S} = 0$. In that case, Equation (24) becomes $D\dot{S}_T = (\partial_t \langle x \rangle)^2$ while Equations (13), (27) and (28) give us

$$\mathcal{L}(t) = \frac{\langle x(t) \rangle - \langle x(t=0) \rangle}{\sigma}, \tag{29}$$

$$\Gamma^2 = \frac{1}{\sigma^2}(\partial_t \langle x \rangle)^2 = \frac{D}{\sigma^2}\dot{S}_T. \tag{30}$$

Equation (29) simply states that $\mathcal{L}$ measures the total change in the mean value normalized by fluctuation level σ. Equation (30) manifests a linear relation between Γ^2 and $\dot{S}_T$ when $\partial_t \sigma = 0$, as invoked in the previous works (e.g., [107]). Furthermore, a linear relation between Γ^2 and $\dot{S}_T$ in Equation (30) implies that minimizing the entropy production $\int^t dt_1 \dot{S}_T$ along the trajectory corresponds to minimizing $\int_0^t dt_1 \Gamma^2(t_1)$, which, in turn, is equivalent to minimizing $\mathcal{L}(t)$ (see Section 5 for further discussions).

Finally, to demonstrate how entropy production rate and thermal bath temperature (D) are linked to the speed of fluctuations $c = \sigma\Gamma$ [97], we rewrite Equation (28) as

$$c = \sigma\Gamma = \left[D\dot{S}_T + \sigma^2\dot{S}^2 \right]^{\frac{1}{2}}. \tag{31}$$

For constant variance $\dot{\beta} = 0$, Equation (31) gives a simple relation $c = \sigma\Gamma = \sqrt{D\dot{S}_T}$.

4.2. Non-Equilibrium Thermodynamical Laws

To relate the statistical measures in Section 4.1 to thermodynamics, we let U (internal energy) be the average potential energy $U = \langle V \rangle$ and obtain (see also [66,113] and references therein)

$$\frac{dU}{dt} = \frac{d}{dt}\langle V \rangle \equiv \dot{W} - \dot{Q}, \tag{32}$$

where

$$\dot{W} = \int dx(\partial_t V)p = \langle \partial_t V \rangle, \tag{33}$$

$$\dot{Q} = -\int dx V(\partial_t p) = \int dx Jf = \langle f\dot{x} \rangle = D\dot{S}_m. \tag{34}$$

The power $\dot{W}$ represents the average rate at which the work is done to the system because of time-varying potential; the average work during the time interval $[t_0, t]$ is calculated by $W = \int_{t_0}^t dt' \dot{W}(t')$. On the other hand, $\dot{Q}$ represents the rate of dissipated heat.

Equation (32) establishes the non-equilibrium thermodynamic relation $\dot{U} = \dot{W} - \dot{Q}$. Physically, it simply means that the work done to the system $\dot{W}$ increases U while the dissipated heat to the environment $\dot{Q}$ decreases it. Equations (21), (32), and (34) permit us to define a non-equilibrium (information) free energy $\mathcal{F}(t) = U(t) - DS(t)$ [92] and its time-derivative

$$\dot{\mathcal{F}} = \dot{U} - D\dot{S} = \dot{W} - D\dot{S}_T, \tag{35}$$

where $\dot{\mathcal{F}} = \frac{d\mathcal{F}}{dt}$ and $\dot{U} = \frac{dU}{dt}$. Since $\dot{S}_T \geq 0$, Equation (35) leads to the following inequality

$$D\dot{S}_T \;=\; \dot{W} - \dot{\mathcal{F}} \equiv \dot{W}_D \geq 0, \tag{36}$$

where the non-negative dissipated power (lost to the environment) $\dot{W}_D$ is defined. Finally, the time-integral version of Equation (36) provides the bound on the average work performed on the system as $W - \Delta\mathcal{F} = W_D \geq 0$ (e.g., [68]).

4.3. Relative Entropy as a Measure of Irreversibility

The relative entropy has proven to be useful in understanding irreversibilities and non-equilibrium thermodynamic inequality relations [91–94,114–116]. In particular, the dissipated work $W_D = W - \Delta\mathcal{F}$ (in Equation (36)) is related to the relative entropy between the PDFs in the forward and reverse processes

$$W_D = D\mathcal{K}[p_F(\gamma_F(t))|p_R(\gamma_R(t))]. \tag{37}$$

(e.g., see [91–94].) Here, $p_F(\gamma_F(t))$ and $p_R(\gamma_R(t))$ are the PDFs for the forward and reverse processes driven by the forward $\gamma_F(t)$ and reverse $\gamma_R(t)$ protocols, respectively. Using Equation (36) in Equation (37) immediately gives

$$\dot{S}_T = \frac{d}{dt}\mathcal{K}[p_F(\gamma_F(t))|p_R(\gamma_R(t))] \geq 0, \tag{38}$$

which is a proxy for irreversibility (see [115,116] for a slightly different expression of Equation (38)). It is useful to note that forward and reversal protocols are also used to establish various fluctuations theorems for different dissipative measures such as entropy production, dissipated work, etc. (see, e.g., [80] for a nice review and references therein).

However, we cannot consider forward and reversal protocols in the absence of a model control parameter that can be prescribed as a function of time. Even in this case, the relative entropy is useful in quantifying irreversibility through inequalities, and this is what we focus on in the remainder of Section 4.3.

To this end, let us consider a non-equilibrium state $p(x, t)$ which has an instantaneous non-equilibrium stationary state $p_s(x, t)$ and calculate the relative entropy between the two. Here, $p_s(x, t)$ is a steady solution of the Fokker-Planck equation $\partial_t p_s = 0$ in Equation (20) (e.g., see [29]). Specifically, one finds $p_s(x, t) = e^{-\frac{V(x,t) - \mathcal{F}_s(t)}{D}}$ by treating the parameters to be constant (being frozen to their instantaneous values at a given time). Here, V and $\mathcal{F}_s$ are the potential energy and the stationary free energy, respectively. For clarity, an example of $p_s(x, t)$ is given in Section 4.4.

The average of $\ln p_s(t)$ in the non-equilibrium state $p(x, t)$ can be expressed as follows:

$$\int dx\, p(x, t) \ln p_s(x, t) = -\frac{1}{D} \int dx\, p(x, t)(V(x, t) - \mathcal{F}_s(t)) = -\frac{1}{D}(U(t) - \mathcal{F}_s(t)). \tag{39}$$

Equations (35) and (39) then give us

$$\mathcal{F}(t) - \mathcal{F}_s(t) = D \int dx\, p(x, t) \ln\left[\frac{p(x, t)}{p_s(x, t)}\right] \equiv D\mathcal{K}[p(x, t)|p_s(x, t)] \geq 0. \tag{40}$$

Here, we used the fact the relative entropy is non-negative. Equation (40) explicitly shows that non-equilibrium free energy is bounded below by the stationary one $\mathcal{F} \geq \mathcal{F}_s$ (see also [1,92] and references therein for open Hamiltonian systems).

Equation (40) together with Equation (35) then lead to the following irreversible work W_{irr} [29,92]:

$$W_{irr} \equiv W - \Delta\mathcal{F}_s = D\Delta S_T + \Delta(\mathcal{F} - \mathcal{F}_s) = D\Delta S_T + D\Delta\mathcal{K}[p(x, t)|p_s(x, t)]. \tag{41}$$

Here, $\Delta \mathcal{K}[p(x,t)|p_s(x,t)] = \mathcal{K}[p(x,t)|p_s(x,t)] - \mathcal{K}[p(x,t_0)|p_s(x,t_0)]$, etc. The derivation of Equation (41) for open-driven Hamiltonian systems is provided in [92] (see their Equation (38)).

On the other hand, we directly calculate the time-derivative of $\mathcal{K}[p(x,t)|p_s(x,t)]$ in Equation (40) by using $p_s(x,t) = e^{-\frac{V(x,t) - \mathcal{F}_s(t)}{D}}$, $\dot{S}_T = \dot{S} + \dot{S}_m$, $\int dx p \partial_t V = \dot{W}$ and $\dot{Q} = -\int dx\, \partial_t p V = D\dot{S}_m$, and $W_{irr} = W - \Delta\mathcal{F}_s$:

$$\frac{d}{dt}\mathcal{K}[p(x,t)|p_s(x,t)] = -\dot{S} + \frac{1}{D}\frac{d}{dt}\left[\int dx V p - \mathcal{F}_s\right] = -\dot{S}_T + \frac{1}{D}\left[\dot{W} - \frac{d}{dt}\mathcal{F}_s\right]. \quad (42)$$

One can see easily that equating Equation (42) to $\frac{1}{D}[\dot{\mathcal{F}} - \dot{\mathcal{F}}_s]$ (from Equation (40)) simply recovers $\dot{W} - \frac{d}{dt}\mathcal{F} = D\dot{S}_T$ in Equation (35).

Finally, we obtain a differential form of Equation (41) by using $\dot{W}_{irr} = \dot{W} - \frac{d}{dt}\mathcal{F}_s$ in Equation (42) as follows

$$\dot{W}_{irr} = D\dot{S}_T + D\frac{d}{dt}\mathcal{K}[p(x,t)|p_s(x,t)]. \quad (43)$$

4.4. Example

We consider $v(t) = ut$ with a constant u so that $V = -\frac{\gamma}{2}(x - ut)^2$ in Equation (14). While the discussion below explicitly involves $v(t)$, the results are general and valid for the limiting case $v(t) = 0$. The case with $v(t) = 0$ is an example where the forward and reversal protocols do not exist while a non-equilibrium stationary state does.

For $f = -\gamma(x - ut)$, Equation (19) is simplified as follows

$$\langle x(t) \rangle = x_0 e^{-\gamma t} + ut - \frac{u}{\gamma}\left(1 - e^{-\gamma t}\right). \quad (44)$$

For the non-equilibrium stationary state with fixed γ and D, $\beta_s = \frac{\gamma}{2D}$ is also constant ($\frac{d}{dt}\mathcal{F}_s = 0$). Therefore, we have

$$p_s(x,t) = \sqrt{\frac{\beta_s}{\pi}}e^{-\beta_s(x-ut)^2}. \quad (45)$$

Then, we can find (see [29] for details)

$$\mathcal{K}[p(x,t)|p_s(x,t)] = \frac{1}{2}\left[-1 + \ln(\beta/\beta_s)\right] + \beta_s\left((\langle x \rangle - ut)^2 + \frac{1}{2\beta}\right), \quad (46)$$

$$D\dot{S}_T = \left[\left(-\gamma\langle x_0\rangle e^{-\gamma t} + u\left(1 - e^{-\gamma t}\right)\right)^2 + \frac{1}{2\beta}(2\beta D - \gamma)^2\right], \quad (47)$$

$$\dot{Q} = \left[-\gamma\langle x_0\rangle e^{-\gamma t} + u\left(1 - e^{-\gamma t}\right)\right]^2 - \frac{\gamma}{2\beta}(2\beta D - \gamma), \quad (48)$$

$$\dot{W} = -u\int dx\gamma(x - ut)p = -u\gamma\langle x - ut\rangle = u^2\left(1 - e^{-\gamma t}\right), \quad (49)$$

$$\frac{d}{dt}\mathcal{K}[p(x,t)|p_s(x,t)] = -\frac{1}{D}\left[(\partial_t\langle x\rangle)^2 + (\partial_t\sigma)^2 - u\partial_t\langle x\rangle\right] = -\dot{S}_T + \frac{1}{D}\dot{W}. \quad (50)$$

Here, we used Equations (23)–(26), $\langle x \rangle$ and β in Equations (44) and (18), respectively, and $\dot{W} = \langle -\gamma u(x - ut)\rangle = u\partial_t\langle x\rangle$.

It is worth looking at the two interesting limits of Equations (46)–(50). First, in the long time limit as $t \to \infty$, the following simpler relations are found:

$$\langle x \rangle \to u(t - \gamma^{-1}), 2\beta \to \frac{\gamma}{D} = 2\beta_s,$$

$$D\dot{S}_T = \dot{Q} = \dot{W} = (\sigma\Gamma)^2 \to u^2,$$

$$\dot{S} \to 0, \quad \frac{d}{dt}\mathcal{K}[p(x,t)|p_s(x,t)] \to 0. \tag{51}$$

Equation (51) illustrates how the external force $v(t) = ut \neq 0$ keeps the system out of equilibrium even in the long time limit, with non-zero entropy production and dissipation. When there is no external force $u = 0$, the system reaches equilibrium as $t \to \infty$, and all quantities in Equation (51) apart from β become zero.

The second is when the system is initially in equilibrium with $\beta(t = 0) = \beta(t \to \infty) = \frac{\gamma}{2D}$ and $\langle x_0 \rangle = 0$ and evolve in time as it is driven out of equilibrium by $u \neq 0$. As u does not affect variance, $\beta(t) = \beta_0 = \frac{\gamma}{2D}$ ($\partial_t \sigma = 0$) and $\dot{S} = 0$ for all time. In this case, we find

$$D\dot{S}_T = \dot{Q} = u^2(1 - e^{-\gamma t})^2 = (\sigma\Gamma)^2, \quad \dot{W} = u^2(1 - e^{-\gamma t}),$$

$$D\mathcal{K}[p(x,t)|p_s(x,t)] = \frac{u^2}{2\gamma}(1 - e^{-\gamma t})^2,$$

$$D\frac{d}{dt}\mathcal{K}[p(x,t)|p_s(x,t)] = u^2(1 - e^{-\gamma t})e^{-\gamma t}. \tag{52}$$

Equation (52) shows that $\dot{S}_T$, $\dot{Q}$, Γ^2, $\dot{W}$, and $\mathcal{K}$ start with zero values at $t = 0$ and monotonically increase to their asymptotic values as $t \to \infty$.

Finally, both cases considered above in Equations (51) and (52) have $\partial_t \sigma = 0$ and thus recover Equation (30):

$$\Gamma^2 = \frac{D\dot{S}_T}{\sigma^2} = \frac{\dot{Q}}{\sigma^2}. \tag{53}$$

5. Inequalities

Section 4 utilized the average (first moment) of a variable (e.g., $\langle V \rangle$) and the average of its first time derivative ($\langle \partial_t V \rangle = \dot{W}$) while the work $W = \int dt\, \dot{W}$ is defined by the time integral of $\dot{W} = \langle \partial_t V \rangle$ in Equation (33). This section aims to show that the rates at which average quantities vary with time are bounded by fluctuations and Γ. Since the average and time derivatives do not commute, we pay particular attention to when the average is taken.

To this end, let us first define the microscopic free energy $\mu = V + D\ln p$ (called the chemical potential energy in [113]). In terms of μ, we have $J = -p\partial_x \mu$ and $\langle \mu \rangle = U - DS = \mathcal{F}$. On the other hand,

$$\partial_t \mu = \partial_t V + D\frac{\dot{p}}{p}, \qquad \langle \partial_t \mu \rangle = \langle \partial_t V \rangle = \dot{W}. \tag{54}$$

$\langle \partial_t \mu \rangle = \dot{W}$ means that the average rate of change in the microscopic free is the power. From Equation (54), it follows

$$\langle (\partial_t \mu - \partial_t V)^2 \rangle = D^2 \int dx \frac{\dot{p}^2}{p} = D^2\Gamma^2. \tag{55}$$

Equation (55) establishes the relation between the microscopic free energy and Γ. Next, we calculate the time-derivative of $\mathcal{F}$

$$\frac{d}{dt}\mathcal{F} = \frac{d}{dt}\langle \mu \rangle = \langle \dot{\mu} \rangle + \int dx\,\mu\dot{p} = \dot{W} + \int dx\,\mu\dot{p}. \tag{56}$$

Using $\frac{d}{dt}\mathcal{F} = \dot{W} - D\dot{S}_T$ in Equation (56) gives $\dot{S}_T$ in terms of μ as

$$\dot{W} - \frac{d}{dt}\mathcal{F} = D\dot{S}_T = -\int dx\,\mu\dot{p}. \tag{57}$$

Equation (57) is to be used in Section 5.1 for linking $\dot{S}_T$ to Γ through an inequality.

5.1. General Inequality Relations

We now use $\int dx \dot{p} = 0$, $\int dx \dot{p}\langle A \rangle = \langle A \rangle \int dx \dot{p} = 0$ for any $A = A(x,t)$ and apply the Schwartz inequality $|\int^t dt_1 A(t_1)B(t_1)| \leq [\int^t dt_1 A^2(t_1)]^{\frac{1}{2}}[\int^t dt_1 B^2(t_1)]^{\frac{1}{2}}$ to Equations (21), (34) and (57) to obtain

$$|\dot{S}| = \left| \int dx \dot{p} \ln p \right| \leq \Gamma \left[\int dx p (\delta \ln p)^2 \right]^{1/2}, \tag{58}$$

$$|\dot{Q}| = \left| \int dx V \dot{p} \right| \leq \Gamma \left[\int dx p (\delta V)^2 \right]^{1/2}, \tag{59}$$

$$D\dot{S}_T = \left| \int dx \mu \dot{p} \right| \leq \Gamma \left[\int dx p (\delta \mu)^2 \right]^{1/2}. \tag{60}$$

Equation (60) (Equation (59)) establishes the inequality between entropy production rate (heat flux) and the product of the RMS fluctuations of the microscopic free energy (potential energy) and Γ. Since $\delta \mu = \delta V + D(\delta \ln p)$, we have

$$\langle (\delta \mu)^2 \rangle = \langle (\delta V)^2 \rangle + D^2 \langle (\delta \ln p)^2 \rangle + 2D \langle \delta V \delta \ln p \rangle. \tag{61}$$

These relations are to be used in Section 5.2 below.

5.2. Applications to the Non-Autonomous O-U Process

For a linear O-U process with $V = \frac{\gamma}{2}(x - v(t))^2$ and $f = -\gamma(x - v(t))$ in Equation (14), we use $\langle (\delta x)^2 \rangle = \frac{1}{2\beta}$, $\langle (\delta x)^4 \rangle = 3\langle (\delta x)^2 \rangle = \frac{3}{4\beta^2}$ and $\partial_t \langle x \rangle = -\gamma \langle x - v(t) \rangle$ to show

$$\delta V = -\partial_t \langle x \rangle \delta x + \frac{\gamma}{2}\left[(\delta x)^2 - \frac{1}{2\beta} \right]$$

$$\langle (\delta V)^2 \rangle = (\partial_t \langle x \rangle)^2 \sigma^2 + \gamma^2 \frac{\sigma^4}{2},$$

$$\delta \ln p = \frac{1}{2} - \beta (\delta x)^2,$$

$$\langle (\delta \ln p)^2 \rangle = \frac{1}{2},$$

$$\langle (\delta \ln p)(\delta V) \rangle = -\frac{\gamma}{4\beta}. \tag{62}$$

Using Equations (61)–(62) in Equations (58)–(60) together with $2D\beta - \gamma = -\frac{\partial_t \beta}{2\beta}$ and $\frac{\partial_t \beta}{\beta} = -2\frac{\partial_t \sigma}{\sigma}$ leads to

$$|\dot{S}| \leq \frac{1}{\sqrt{2}}\Gamma,$$

$$|\dot{Q}| \leq \Gamma \sigma \left[(\partial_t \langle x \rangle)^2 + \frac{\gamma^2 \sigma^2}{2} \right]^{\frac{1}{2}},$$

$$D\dot{S}_T \leq \Gamma \sigma \left[(\partial_t \langle x \rangle)^2 + \frac{1}{2}(\partial_t \sigma)^2 \right]^{\frac{1}{2}}. \tag{63}$$

Finally, it is useful to examine the extreme cases of Equation (63). First, when $\partial_t \sigma = 0$, Equation (63) holds as an equality as $D\dot{S}_T = (\partial_t \langle x \rangle)^2 = \sigma^2 \Gamma^2$ (see Equation (27)), recovering Equation (28) with $\partial_t \sigma = 0$. Second, when $\partial_t \langle x \rangle = 0$, Equation (63) again holds as an equality since $D\dot{S}_T = (\partial_t \sigma)^2$ and $\Gamma \sqrt{\frac{1}{2}\sigma^2 (\partial_t \sigma)^2} = (\partial_t \sigma)^2$.

6. Geodesics, Control and Hyperbolic Geometry

The section aims to discuss geodesics in information geometry and its implications for self-organization and control. To illustrate the key concepts, we utilize an analytically solvable, generalized O-U process given by

$$\frac{dx}{dt} = -\gamma(t)[x - v(t)] + \xi, \tag{64}$$

where $\gamma(t) > 0$ is a damping constant; $v(t)$ is a deterministic force which determines the time evolution of the mean value of x; ξ is a short (delta)-correlated noise with the time-dependent amplitude $D(t)$ in general, satisfying Equation (15).

For the initial condition in Equation (16), the mean value $\langle x \rangle \equiv y(t)$ and $\beta(t)$ are given by

$$y(t) = \langle x \rangle = x_0 e^{-\int_0^t dt_1 \gamma(t_1) dt_1} + \int_0^t dt_1 e^{-\int_0^t dt_1 [\gamma(t_1) - \gamma(t)] dt_1} \gamma(t_1) f(t_1), \tag{65}$$

$$\frac{1}{2\beta(t)} = \langle (x - \langle x \rangle)^2 \rangle = \frac{e^{-2\int_0^t dt_1 \gamma(t_1) dt_1}}{2\beta_0} + \int_0^t dt_1 e^{-2\int_0^t dt_1 [\gamma(t_1) - \gamma(t)] dt_1} 2D(t_1), \tag{66}$$

where $x_0 = \langle x(t = 0) \rangle$.

6.1. Geodesics–Shortest-Distance Path

A geodesics between the two spatial locations is a unique path with the shortest distance. A similar concept can be applied to information geometry to define a unique evolution path between the two given PDFs, say, $p(x, t_1)$ and $p(x, t_2)$ in the statistical space. The Wootters' distance in quantum mechanics in Equation (4) is such an example. For time-varying stochastic processes, there is an infinite number of different trajectories between the two PDFs at different times. The key question that we address in this section is how to find an exact time evolution of $p(x, t)$ when initial and final PDFs [15] are given. This is a much more difficult problem than finding a minimum distance between two PDFs (like the Wootter's distance). In the following, we sketch some main steps needed for finding such a unique evolution path (the so-called geodesics) between given initial and final PDFs by minimizing $\mathcal{L}$ (see [15] for detailed steps).

For the O-U process in Equation (64), a geodesic solution does not exist for constant γ, $v(t)$ and D. Thus, finding a geodesic solution boils down to determing suitable functions of $\gamma(t)$, $v(t)$ or $D(t)$ [15]. To be specific, let $p(x, t_0)$ and $p(x, t_F)$, respectively, be the PDFs at the time $t = t_0$ and t_F ($> t_0$) and find a geodesic solution by minimizing $\mathcal{L}(t) = \int_{t_0}^{t_F} dt' \Gamma(t')$. The latter is equivalent to minimizing $\int_{t_0}^{t_F} dt' \mathcal{E}(t')$ and to keeping Γ constant. (This geodesics is also called an optimal path (e.g., see [107]).) We rewrite $\mathcal{E}$ in Equation (27) for the O-U process in terms of $y = \langle x \rangle$

$$\mathcal{E} = \frac{1}{2\beta^2} \left(\frac{d\beta}{dt} \right)^2 + 2\beta \left(\frac{dy}{dt} \right)^2. \tag{67}$$

The Euler-Lagrange equation

$$0 = \frac{d\mathcal{E}}{d\beta} - \frac{d}{dt}\frac{d\mathcal{E}}{d\dot{\beta}}, \quad 0 = \frac{d\mathcal{E}}{dy} - \frac{d}{dt}\frac{d\mathcal{E}}{d\dot{y}} \tag{68}$$

($\dot{\beta} = \frac{d\beta}{dt}$ and $\dot{y} = \frac{dy}{dt}$) then gives us

$$\frac{d^2\beta}{dt^2} - \frac{1}{\beta}\left(\frac{d\beta}{dt}\right)^2 - 2\beta^2\left(\frac{dy}{dt}\right)^2 = 0, \tag{69}$$

$$\frac{d}{dt}\left[\beta\frac{dy}{dt}\right] = 0 \rightarrow \beta\frac{dy}{dt} = c, \tag{70}$$

where c is constant. An alternative method of obtaining Equations (69) and (70) is provided in Appendix C. The following equations are obtained from Equations (69) and (70) [15]

$$\left(\frac{d\beta}{dt}\right)^2 = -4c^2\beta + \alpha\beta^2, \tag{71}$$

$$\Gamma^2 = \frac{1}{2\beta^2}\dot{\beta}^2 + 2c^2\beta = \frac{\alpha}{2}, \tag{72}$$

where α is another (integration) constant. General solutions to Equations (70) and (71) for $c \neq 0$ were found in terms of hyperbolic functions as [15]

$$\beta(t) = \frac{4c^2}{\alpha}\cosh^2\left[\frac{1}{2}\sqrt{\alpha}(t-A)\right], \qquad y(t) = \frac{\sqrt{\alpha}}{2c}\tanh\left[\frac{1}{2}\sqrt{\alpha}(t-A)\right] - \frac{\sqrt{\alpha}}{2c} + B, \tag{73}$$
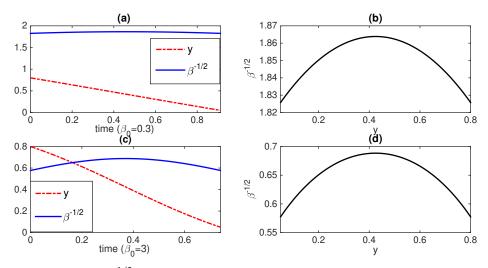
where A and B are constant.

Equation (73) can be rewritten using $\sigma = (2\beta)^{-\frac{1}{2}}$ and $z = \frac{y}{\sqrt{2}}$ as follows

$$(z - z_c)^2 + \sigma^2 = R^2, \ z_c = \frac{B}{\sqrt{2}} - sR, \ R = \frac{\Gamma}{2c}, \tag{74}$$

where s denotes the sign of c so that $s = 1$ when $c > 0$ while $s = -1$ when $c < 0$. Equation (74) is an equation of a circle for the variables z and σ with the radius R and the center z_c, defined in the upper-half plane where $\sigma \geq 0$. Thus, geodesic motions occur along the portions of a circle as long as $c \neq 0$ (as can be seen in Figure 4). A geodesic moves on a circle with a larger radius for a larger information rate Γ and speed and vice versus. This manifests the hyperbolic geometry in the upper half Poincaré model [13,117] where the half-plane represents z and $\sigma \neq 0$ (see also Appendix D). The constants c, α, A, and B determine the coordinate of the center and the radius of the circle R. These constants should be fixed by the fixed conditions at the initial $t = 0$ and final time t_F.
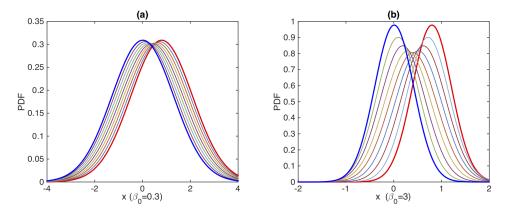
Having found the most general form of the geodesic solution for $y(t)$ and β, the next steps require finding the values of constant values c, α, A, B to satisfy the boundary conditions at $t = t_0$ and t_F, and then finding appropriate $\gamma(t)$, $D(t)$, and $v(t)$ that ensure the geodesic solutions. This means the O-U process should be controlled by $\gamma(t)$, $D(t)$ and $v(t)$ to ensure a geodesic solution.

Figure 4. y and $\beta^{-1/2}$ against time for $\beta_0 = 0.3$ and 3 in (**a**,**c**), respectively; the corresponding geodesic circular segments in the $(y, \beta^{-1/2})$ upper half-plane in (**b**,**d**), respectively. In both cases, $y_0 = \frac{5}{6}$ and $y_F = \frac{1}{30}$. (Figure 3 in [15]).

Figure 4 shows an example of a geodesic solution in the upper half-plane y and $\beta^{-1/2}$ when $\gamma(t) = 1$ is constant while $D(t)$ and $v(t)$ are time-dependent. The boundary conditions are chosen as $y(t_0) = y_0 = \frac{5}{6}$ and $y(t_F) = y_F = \frac{1}{30}$ in all panels (a)–(d). $\beta(t_0) = \beta_0 = \beta(t_F) = \beta_F = 0.3$ in panels (a) and (b) while $\beta_0 = \beta_F = 3$ in panels (c) and (d). Interestingly, circular-shape phase-portraits are seen in panels (b) and (d), reflecting hyperbolic geometry noted above (see also Appendix D) [13,117]. The speed at which the geodesic motion takes place in the phase portrait is determined by the constant value of $\Gamma = \sqrt{\frac{\alpha}{2}}$ (i.e., the larger α, the faster time evolution).

Figure 5a,b are the corresponding PDF snapshots at different times (shown in different colors), demonstrating how the PDF evolves from the initial PDF in red to the final PDF in blue. In both cases, it is prominent that the PDF width ($\propto \beta^{-1/2}$) initially broadens and then becomes narrower.

Figure 5. Time evolution of PDFs against x: (**a**) $\beta_0 = 0.3$ corresponding to Figure 4a,b; (**b**) $\beta_0 = 3$ corresponding to Figure 4c,d. In both cases, $y_0 = \frac{5}{6}$ and $y_F = \frac{1}{30}$. The initial and final PDFs are shown by thick red and blue lines, respectively. (Figure 4a,b in [15]).

6.2. Comments on Self-Organization and Control

Self-organization (also called homeostasis) is the novel phenomena where order spontaneously emerges out of disorder and is maintained by different feedbacks in complex systems [45,52,53,118–123]. The extremum principles of thermodynamics such as the minimum entropy production (e.g., [119,121]) or maximum entropy entropy production (e.g., [122,123]) have been proposed by considering a steady state or an instant time in different problems.

However, far from equilibrium, self-organization can be a time-varying non-equilibrium process involving perpetual or large fluctuations (e.g., see [52–54]). In this case, the extreme of entropy production should be on accumulative entropy production over time rather than at one instant time nor in a steady state. That is, we should consider the time-integral of the entropy production $\dot{S}_T$, or equivalently, the time-integral of $\sqrt{\dot{S}_T}$. As seen from Equations (24) and (53), for a linear O-U process with a constant variance, there is an exact proportionality between $\sqrt{\dot{S}_T}$ and Γ. In this case, the extreme of $\mathcal{L}(t) = \int^t dt_1 \Gamma(t_1)$ would be the same as the extreme of $\int^t dt_1 \sqrt{\dot{S}_T}$. However, as noted previously, $\Gamma \propto \sqrt{\dot{S}_T}$ does not hold in general (e.g., see Equation (28)).

With these comments, we now look at the implications of a geodesic for self-organization, in particular, in biosystems. For the very existence and optimal functions of a living organiss, it is critical to minimize the dispersion of its physical states and to maintain its states within certain bounds upon changing conditions [124]. How fast its state changes in time can be quantified by the surprise rate $\partial_t [\ln(p(x,t))]$. Since $\int dx p(x,t) \partial_t \ln(p(x,t)) = 0$, we use its RMS value $\sqrt{\langle (\partial_t \ln p)^2 \rangle} = \Gamma$ (see Equation (9)) and realize that the total change over a finite time interval $[t_0, t_F]$ is nothing more than $\mathcal{L}(t) = \int_{t_0}^{t_F} dt_1 \Gamma(t_1)$. Thus, minimizing the accumulative/time-integral of the RMS surprise rate is equivalent to minimizing $\mathcal{L}$. Envisioning surprise rate as biological cost associated with changes (e.g., needed in updating

the future prediction based on the current state [124,125]), we can then interpret $\mathcal{L}$ as an accumulative biological cost. Thus, geodesic would be an optimal path that minimizes such an accumulative biological cost.

Ref [15] addressed how to utilize this idea to control populations (tumors). Specifically, the results in Section 6.1 were applied to a nonlinear stochastic growth model (obtained by a nonlinear change of variables of the O-U process), and the geodesic solution in Equation (73) was used to find the optimal protocols $v(t)$ and $D(t)$ in reducing a large-size tumor to a smaller one. Here, in this problem, $D(t)$ represents the heterogeneity of tumor cells (e.g., larger D for metastatic tumor) that can be controlled by gene reprogramming while $v(t)$ models the effect of a drug or radiation that reduces the mean tumor population/size.

7. Discussions and Conclusions

There has been a growing interest in information geometry from theoretical and practical considerations. This paper discussed some recent developments in information geometric theory, focusing on time-dependent *dynamic* aspects of non-equilibrium processes (e.g., time-varying mean value, time-varying variance, or temperature) and their thermodynamic and physical/biological implications.

In Sections 2 and 3, by utilizing a Langevin model of an over-damped stochastic process $x(t)$, we highlighted the importance of a *path-dependent* distance $\mathcal{L}$ in describing time-varying processes. In Sections 4 and 5, we elucidated the thermodynamic meanings of the relative entropy and the information rate Γ by relating them to the entropy production rate ($\dot{S}_T$), $\dot{S}$, heat flux ($Q = D\dot{S}_m$), dissipated work ($\dot{W}_D$), etc., and demonstrated the role of Γ in determining bounds (or speed limit) on thermodynamical quantities.

Specifically, in the O-U process, we showed the exact relation $\Gamma = \sqrt{\frac{D}{\sigma^2}\dot{S}_T + \dot{S}^2}$ (Equation (28)), which is simplified as $\sigma\Gamma = \sqrt{D\dot{S}_T}$ when $\partial_t\sigma = 0$ ($\sigma = \sqrt{\langle(\delta x)^2\rangle}$ is the standard deviation of x). Finally, Section 6 discussed geodesic and its implication for self-organization as well as the underlying hyperbolic geometry. It remains future works to explore the link between Γ and the entropy production rate in other (e.g., nonlinear) systems consisting of three or more interacting components or data from self-organizing systems (e.g., normal brain).

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A

In this Appendix, we show the invariance of Equation (9) when x changes as $y = F(x)$. Using the conservation of the probability, we then have

$$p(y,t) = p(x,t)\left|\frac{dx}{dy}\right| = p(x,t)\left|\frac{dF(x)}{dx}\right|^{-1}, \tag{A1}$$

Since $\frac{dF(x)}{dx}$ is independent of time t, it follows that $\partial_t p(y,t) = [\partial_t p(x,t)]\left|\frac{dF(x)}{dx}\right|^{-1}$. Using this and $dy = dx\left|\frac{dF(x)}{dx}\right|$, we have

$$\int dy \frac{1}{p(y,t)}\left[\frac{\partial p(y,t)}{\partial t}\right]^2 = \int dx \frac{1}{p(x,t)}\left[\frac{\partial p(x,t)}{\partial t}\right]^2. \tag{A2}$$

This shows that $p(x, t)$ and $p(y, t)$ give the same $\Gamma^2(t)$.

Appendix B. The Coupled O-U Process

The coupled O-U process for Figure 3 in Section 3.1 is governed by the Fokker-Planck equation [22]

$$\frac{\partial P_1}{\partial t} = \frac{\partial}{\partial x}[\gamma_1(x - \mu)P_1] + D\frac{\partial^2 P_1}{\partial x^2} - f_0 P_1 + g_0 P_2, \tag{A3}$$

$$\frac{\partial P_2}{\partial t} = \frac{\partial}{\partial x}[\gamma_2(x - \mu)P_2] + D\frac{\partial^2 P_2}{\partial x^2} + f_0 P_1 - g_0 P_2. \tag{A4}$$

Here, D is the strength of a short-correlated Gaussian noise given by Equation (15). These equations are the coupled O-U processes with the coupling constants f_0 and g_0 through the Dichotomous noise [110,111].

For simplicity, we use $\gamma_1 = \gamma_2 = \gamma$ and $f_0 = g_0 = \epsilon$ and the following initial conditions

$$P_1(x, 0) = \frac{1}{2}\sqrt{\frac{\beta_{10}}{\pi}}\exp[-\beta_{10}x^2], \tag{A5}$$

$$P_2(x, 0) = \frac{1}{2}\sqrt{\frac{\beta_{20}}{\pi}}\exp[-\beta_{20}(x - x_0)^2]. \tag{A6}$$

The solutions are given by

$$P_1(x, t) = \frac{1}{4}\left[\sqrt{\frac{\beta_1}{\pi}}(1 + e^{-2\epsilon t})e^{-\beta_1 x^2} + \sqrt{\frac{\beta_2}{\pi}}(1 - e^{-2\epsilon t})e^{-\beta_2(x - e^{-\gamma t}x_0)^2}\right], \tag{A7}$$

$$P_2(x, t) = \frac{1}{4}\left[\sqrt{\frac{\beta_1}{\pi}}(1 - e^{-2\epsilon t})e^{-\beta_1 x^2} + \sqrt{\frac{\beta_2}{\pi}}(1 + e^{-2\epsilon t})e^{-\beta_2(x - e^{-\gamma t}x_0)^2}\right], \tag{A8}$$

where

$$\frac{1}{2\beta_m} = \frac{e^{-2\gamma t}}{2\beta_{m0}} + \frac{D}{\gamma}\left(1 - e^{-2\gamma t}\right), \tag{A9}$$

for $m = 1, 2$. In the limit of $t \to \infty$, P_1 and P_2 in Equations (A7) and (A8) approach the same equilibrium distribution

$$P_m(x, t) = \frac{1}{2}\sqrt{\frac{\beta_m(t \to \infty)}{\pi}}e^{-\beta_m(t \to \infty)x^2}, \tag{A10}$$

where $\beta_m(t \to \infty) = \frac{\gamma}{2D}$. We note that the total PDF $P = P_1 + P_2$ satisfies the single O-U process where the initial PDF is given by the sum of Equations (A5) and (A6).

Figure 3 is shown for the fixed parameter values $\gamma = 0.1$, $D = 1$ and $\epsilon = 0.5$, $\beta_{20} = \beta_{10} = 0.05 = \beta(t \to \infty) = \frac{\gamma}{2D} = 0.05$. Different values of the initial mean position x_0 of P_2 are used to examine how metrics depend on x_0. As noted in Section 2.2, P_1 at $t = 0$ is chosen to be the same as the final equilibrium state which has the zero mean value and inverse temperature $\beta_{10} = 0.05$.

Appendix C. Curved Geometry: The Christoffel and Ricci-Curvature Tensors

A geodesic solution in Section 6.1 can also be found by solving the geodesic equation in general relativity (e.g., [31,107]). To this end, we let the two parameters be $\lambda^1 = \langle x \rangle = y$ and $\lambda^2 = \beta$ and express Equation (A11) in terms of the metric tensor g_{ij} as follows (see also Equation (10))

$$\mathcal{E} = \sum_{ij}\frac{d\lambda^i}{dt}g_{ij}\frac{d\lambda^j}{dt}, \tag{A11}$$

where

$$g_{ij} = \begin{pmatrix} \frac{1}{2\beta^2} & 0 \\ 0 & 2\beta \end{pmatrix}, \quad \lambda^i = \begin{pmatrix} \beta \\ y \end{pmatrix}. \tag{A12}$$

Note that while g_{ij} is diagonal, the 1-st diagonal component depends on β (the second parameter). That is, g_{ii} is not independent of j-th parameter for $j \neq i$ in general. From Equation (A12), we can find non-zero components of connection tensor $\Gamma_{ijk} = \frac{1}{2}\left[\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\right]$ ($\Gamma^i_{jk} = g_{im}\Gamma_{jkm}$)

$$\Gamma^1_{11} = -\frac{1}{\beta}, \Gamma^1_{22} = -2\beta^2, \Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{2\beta}. \tag{A13}$$

A geodesic equation $\frac{d^2\lambda^i}{dt^2} + \Gamma^i_{mk}\frac{d\lambda^m}{dt}\frac{d\lambda^k}{dt} = 0$ in terms of the Christoffel tensors becomes

$$\ddot{\beta} + \Gamma^1_{11}\dot{\beta}^2 + \Gamma^1_{22}\dot{y}^2 = 0, \tag{A14}$$
$$\ddot{y} + \Gamma^2_{12}\dot{\beta}\dot{y} + \Gamma^2_{21}\dot{\beta}\dot{y} = 0. \tag{A15}$$

Equations (A13)–(A15) give Equation (70). Note that if g_{ii} is independent of the λ^j ($j \neq i$) for all i and j, the Christoffel tensors have non-zero values only for Γ^i_{ii}, leading to a much simpler geodesic solution (e.g., see [31]).

Finally, to appreciate the curved geometry associated with this geodesic solution, we proceed to calculate the Riemann curvature tensor $R^i_{kmn} = \partial_m\Gamma^i_{nk} + \Gamma^i_{mp}\Gamma^p_{nk} - \partial_n\Gamma^i_{mk} - \Gamma^i_{np}\Gamma^p_{mk}$ and the Ricci tensor $R_{ij} = R^k_{ikj}$ from Equation (A13) and find the following non-zero components [15]

$$R^1_{212} = -R^1_{221} = -\beta, \quad R^2_{112} = -R^2_{121} = \frac{1}{4\beta^2}. \tag{A16}$$

Non-zero curvature tensors represent that the metric space is curved with a finite curvature. Specifically, we find the Ricci tensor $R_{ij} = R^k_{ikj}$ and curvature R:

$$R_{11} = -\frac{1}{4\beta^2}, \quad R_{22} = -\beta, \quad R_{12} = R_{21} = 0, \tag{A17}$$

$$R = g_{ij}R_{ij} = -1 \tag{A18}$$

The negative curvature is typical of hyperbolic geometry. Finally, using $R = -1$, we calculate the Einstein field equation

$$G_{ij} = R_{ij} - \frac{1}{2}R\,g_{ij} = \begin{pmatrix} -\frac{1}{4\beta^2} & 0 \\ 0 & -\beta \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \frac{1}{2\beta^2} & 0 \\ 0 & 2\beta \end{pmatrix} = 0. \tag{A19}$$

Since $G_{ij} = 8\pi T_{ij}$ where T_{ij} is the stress-energy tensor, we see that $T_{ij} = 0$ for this problem.

Appendix D. Hyperbolic Geometry

The Hyperbolic geometry in the upper-half plane [13,117] becomes more obvious when Equation (A12) is expressed in terms of the two parameters $\langle x(t)\rangle$ and $\sigma(t)$ where x and y axes represent $\langle x(t)\rangle$ and $\sigma(t)$ with the metric tensor

$$g^H_{ij} = \begin{pmatrix} \frac{2}{\sigma^2} & 0 \\ 0 & \frac{1}{\sigma^2} \end{pmatrix}. \tag{A20}$$

References

1. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; Wiley: New York, NY, USA, 1991.
2. Parr, T.; Da Costa, L.; Friston, K.J. Markov blankets, information geometry and stochastic thermodynamics. *Philos. Trans. R. Soc. A* **2019**, *378*, 20190159. [CrossRef] [PubMed]
3. Oizumi, M.; Tsuchiya, N.; Amari, S. Unified framework for information integration based on information geometry. *Proc. Nat. Am. Soc.* **2016**, *113*, 14817. [CrossRef] [PubMed]
4. Kowalski, A.M.; Martin, M.T.; Plastino, A.; Rosso, O.A.; Casas, M. Distances in Probability Space and the Statistical Complexity Setup. *Entropy* **2011**, *13*, 1055–1075. [CrossRef]
5. Martin, M.T.; Plastino, A.; Rosso, O.A. Statistical complexity and disequilibrium. *Phys. Lett. A* **2003**, *311*, 126–132. [CrossRef]
6. Gibbs, A.L.; Su, F.E. On choosing and bounding probability metrics. *Int. Stat. Rev.* **2002**, *70*, 419–435. [CrossRef]
7. Jordan, R.; Kinderlehrer, D.; Otto, F. The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.* **1998**, *29*, 1–17. [CrossRef]
8. Takatsu, A. Wasserstein geometry of Gaussian measures *Osaka J. Math.* **2011**, *48* 1005–1026.
9. Lott, J. Some geometric calculations on Wasserstein space. *Commun. Math. Phys.* **2008**, *277*, 423–437. [CrossRef]
10. Gangbo, W.; McCann, R.J. The geometry of optimal transportation. *Acta Math.* **1996**, *177*, 113–161. [CrossRef]
11. Zamir, R. A proof of the Fisher information inequality via a data processing argument. *IEEE Trans. Inf. Theory* **1998**, *44*, 1246–1250. [CrossRef]
12. Otto, F.; Villani, C. Generalization of an Inequality by Talagrand and Links with the Logarithmic Sobolev Inequality. *J. Funct. Anal.* **2000**, *173*, 361–400. [CrossRef]
13. Costa, S.; Santos, S.; Strapasson, J. Fisher information distance. *Discrete Appl. Math.* **2015**, *197*, 59–69. [CrossRef]
14. Ferradans, S.; Xia, G.-S.; Peyré, G.; Aujol, J.-F. Static and dynamic texture mixing using optimal transport. *Lecture Notes Comp. Sci.* **2013**, *7893*, 137–148.
15. Kim, E.; Lee, U.; Heseltine, J.; Hollerbach, R. Geometric structure and geodesic in a solvable model of nonequilibrium process. *Phys. Rev. E* **2016**, *93*, 062127. [CrossRef] [PubMed]
16. Nicholson, S.B.; Kim, E. Investigation of the statistical distance to reach stationary distributions. *Phys. Lett. A* **2015**, *379*, 83–88. [CrossRef]
17. Heseltine, J.; Kim, E. Novel mapping in non-equilibrium stochastic processes. *J. Phys. A* **2016**, *49*, 175002. [CrossRef]
18. Kim, E.; Hollerbach, R. Signature of nonlinear damping in geometric structure of a nonequilibrium process. *Phys. Rev. E* **2017**, *95*, 022137. [CrossRef] [PubMed]
19. Kim, E.; Hollerbach, R. Geometric structure and information change in phase transitions. *Phys. Rev. E* **2017**, *95*, 062107. [CrossRef]
20. Kim, E.; Jacquet, Q.; Hollerbach, R. Information geometry in a reduced model of self-organised shear flows without the uniform coloured noise approximation. *J. Stat. Mech.* **2019**, *2019*, 023204. [CrossRef]
21. Anderson, J.; Kim, E.; Hnat, B.; Rafiq, T. Elucidating plasma dynamics in Hasegawa-Wakatani turbulence by information geometry. *Phys. Plasmas* **2020**, *27*, 022307. [CrossRef]
22. Heseltine, J.; Kim, E. Comparing information metrics for a coupled Ornstein-Uhlenbeck process. *Entropy* **2019**, *21*, 775. [CrossRef] [PubMed]
23. Kim, E.; Heseltine, J.; Liu, H. Information length as a useful index to understand variability in the global circulation. *Mathematics* **2020**, *8*, 299. [CrossRef]
24. Kim, E.; Hollerbach, R. Time-dependent probability density functions and information geometry of the low-to-high confinement transition in fusion plasma. *Phys. Rev. Res.* **2020**, *2*, 023077. [CrossRef]
25. Hollerbach, R; Kim, E.; Schmitz, L. Time-dependent probability density functions and information diagnostics in forward and backward processes in a stochastic prey-predator model of fusion plasmas. *Phys. Plasmas* **2020**, *27*, 102301. [CrossRef]
26. Guel-Cortez, A.J.; Kim, E. Information Length Analysis of Linear Autonomous Stochastic Processes. *Entropy* **2020**, *22*, 1265. [CrossRef] [PubMed]
27. Guel-Cortez, A.J.; Kim, E. Information geometric theory in the prediction of abrupt changes in system dynamics. *Entropy* **2021**, *23*, 694. [CrossRef]
28. Kim, E. Investigating Information Geometry in Classical and Quantum Systems through Information Length. *Entropy* **2018**, *20*, 574. [CrossRef]
29. Kim, E. Information geometry and non-equilibrium thermodynamic relations in the over-damped stochastic processes. *J. Stat. Mech. Theory Exp.* **2021**, *2021*, 093406. [CrossRef]
30. Parr, T.; Da Costa, L.; Heins, C.; Ramstead, M.J.D.; Friston, K.J. Memory and Markov Blankets. *Entropy* **2021**, *23*, 1105. [CrossRef]
31. Da Costa, L. ; Thomas, P.; Biswa, S.; Karl, F.J. Neural Dynamics under Active Inference: Plausibility and Efficiency of Information Processing. *Entropy* **2021**, *23*, 454. [CrossRef]
32. Frieden, B.R. *Science from Fisher Information*; Cambridge University Press: Cambridge, UK, 2004.
33. Wootters, W. Statistical distance and Hilbert-space. *Phys. Rev. D* **1981**, *23*, 357–362. [CrossRef]
34. Ruppeiner, G. Thermodynamics: A Riemannian geometric model. *Phys. Rev. A.* **1079**, *20*, 1608. [CrossRef]
35. Salamon, P.; Nulton, J.D.; Berry, R.S. Length in statistical thermodynamics. *J. Chem. Phys.* **1985**, *82*, 2433–2436. [CrossRef]
36. Nulton, J.; Salamon, P.; Andresen, B.; Anmin, Q. Quasistatic processes as step equilibrations. *J Chem. Phys.* **1985**, *83*, 334. [CrossRef]

37. Braunstein, S.L.; Caves, C.M. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.* **1994**, *72*, 3439. [CrossRef]
38. Diósi, L.; Kulacsy, K.; Lukács, B.; Rácz, A. Thermodynamic length, time, speed, and optimum path to minimize entropy production. *J. Chem. Phys.* **1996**, *105*, 11220. [CrossRef]
39. Crooks, G E. Measuring thermodynamic length. *Phys. Rev. Lett.* **2007**, *99*, 100602. [CrossRef]
40. Salamon, P.; Nulton, J.D.; Siragusa, G.; Limon, A.; Bedeaus, D.; Kjelstrup, D. A Simple Example of Control to Minimize Entropy Production. *J. Non-Equilib. Thermodyn.* **2002**, *27*, 45–55. [CrossRef]
41. Feng, E.H.; Crooks, G.E. Far-from-equilibrium measurements of thermodynamic length. *Phys. Rev. E.* **2009**, *79*, 012104. [CrossRef]
42. Sivak, D.A.; Crooks, G.E. Thermodynamic Metrics and Optimal Paths. *Phys. Rev. Lett.* **2012**, *8*, 190602. [CrossRef] [PubMed]
43. Matey, A.; Lamberti, P.W.; Martin, M.T.; Plastron, A. Wortters' distance resisted: A new distinguishability criterium. *Eur. Rhys. J. D* **2005**, *32*, 413–419.
44. d'Onofrio, A. Fractal growth of tumors and other cellular populations: linking the mechanistic to the phenomenological modeling and vice versa. *Chaos Solitons Fractals* **2009**, *41*, 875. [CrossRef]
45. Newton, A.P.L.; Kim, E.; Liu, H.-L. On the self-organizing process of large scale shear flows. *Phys. Plasmas* **2013**, *20*, 092306. [CrossRef]
46. Kim, E.; Liu, H.-L.; Anderson, J. Probability distribution function for self-organization of shear flows. *Phys. Plasmas* **2009**, *16*, 0552304. [CrossRef]
47. Kim, E.; Diamond, P.H. Zonal flows and transient dynamics of the L-H transition. *Phys. Rev. Lett.* **2003**, *90*, 185006. [CrossRef] [PubMed]
48. Feinberg, A.P.; Irizarry, R.A. Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1757. [CrossRef] [PubMed]
49. Wang, N.X.; Zhang, X.M.; Han, X.B. The effects of environmental disturbances on tumor growth. *Braz. J. Phys.* **2012**, *42*, 253. [CrossRef]
50. Lee, J.; Farquhar, K.S.; Yun, J.; Frankenberger, C.; Bevilacqua, E.; Yeung, E.; Kim, E.; Balázsi, G.; Rosner, M.R. Network of mutually repressive metastasis regulators can promote cell heterogeneity and metastatic transitions. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E364. [CrossRef] [PubMed]
51. Lee, U.; Skinner, J.J.; Reinitz, J.; Rosner, M.R.; Kim, E. Noise-driven phenotypic heterogeneity with finite correlation time. *PLoS ONE* **2015**, *10*, e0132397.
52. Haken, H. *Information and Self-Organization: A Macroscopic Approach to Complex Systems*, 3rd ed., Springer: Berlin/Heidelberg, Germany, 2006; pp. 63–64.
53. Kim, E. Intermittency and self-organisation in turbulence and statistical mechanics. *Entropy* **2019**, *21*, 574. [CrossRef]
54. Aschwanden, M.J.; Crosby, N.B.; Dimitropoulou, M.; Georgoulis, M.K.; Hergarten, S.; McAteer, J.; Milovanov, A.V.; Mineshige, S.; Morales, L.; Nishizuka, N.; et al. 25 Years of Self-Organized Criticality: Solar and Astrophysics. *Space Sci. Rev.* **2016**, *198*, 47–166. [CrossRef]
55. Zweben, S.J.; Boedo, J.A.; Grulke, O.; Hidalgo, C.; LaBombard, B.; Maqueda, R.J.; Scarin, P.; Terry, J.L. Edge turbulence measurements in toroidal fusion devices. *Plasma Phys. Contr. Fusion* **2007**, *49*, S1–S23. [CrossRef]
56. Politzer, P.A. Observation of avalanche-like phenomena in a magnetically confined plasma. *Phys. Rev. Lett.* **2000**, *84*, 1192–1195. [CrossRef] [PubMed]
57. Beyer, P.; Benkadda, S.; Garbet, X.; Diamond, P.H. Nondiffusive transport in tokamaks: Three-dimensional structure of bursts and the role of zonal flows. *Phys. Rev. Lett.* **2000**, *85*, 4892–4895. [CrossRef]
58. Drake, J.F.; Guzdar, P.N.; Hassam, A.B. Streamer formation in plasma with a temperature gradient. *Phys. Rev. Lett.* **1988**, *61*, 2205–2208. [CrossRef]
59. Antar, G.Y.; Krasheninnikov, S.I.; Devynck, P.; Doerner, R.P.; Hollmann, E.M.; Boedo, J.A.; Luckhardt, S.C.; Conn, R.W. Experimental evidence of intermittent convection in the edge of magnetic confinement devices. *Phys. Rev. Lett.* **2001**, *87*, 065001. [CrossRef]
60. Carreras, B.A.; Hidalgo, E.; Sanchez, E.; Pedrosa, M.A.; Balbin, R.; Garcia-Cortes, I. van Milligen, B.; Newman, D.E.; Lynch, V.E. Fluctuation-induced flux at the plasma edge in toroidal devices. *Phys. Plasmas* **1996**, *3*, 2664–2672. [CrossRef]
61. De Vries, P.C.; Johnson, M.F.; Alper, B.; Buratti, P.; Hender, T.C.; Koslowski, H.R.; Riccardo, V. JET-EFDA Contributors, Survey of disruption causes at JET. *Nucl. Fusion* **2011**, *51*, 053018. [CrossRef]
62. Kates-Harbeck, J.; Svyatkovskiy, A.; Tang, W. Predicting disruptive instabilities in controlled fusion plasmas through deep learning. *Nature* **2019**, *568*, 527. [CrossRef]
63. Landau, L.; Lifshitz, E.M. Statistical Physics: Part 1. In *Course of Theoretical Physics*; Elsevier Ltd.: New York, NY, USA, 1980; Volume 5.
64. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 623. [CrossRef]
65. Jarzynski, C.R. Equalities and Inequalities: Irreversibility and the Second Law of Thermodynamics at the Nanoscale. *Annu. Rev. Condens. Matter Phys.* **2011**, *2*, 329–351. [CrossRef]
66. Sekimoto, K. *Stochastic Energetic*; Lecture Notes in Physics 799; Springer: Heidelberg, Germany, 2010. [CrossRef]
67. Jarzynski, C.R. Comparison of far-from-equilibrium work relations. *Physique* **2007**, *8*, 495–506. [CrossRef]
68. Jarzynski, C.R. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.* **1997**, *78*, 2690–2693. [CrossRef]

69. Evans, D.J.; Cohen, G.D.; Morriss, G.P. Probability of Second Law Violations in Shearing Steady States. *Phys. Rev. Lett.* **1993**, *71*, 2401. [CrossRef]
70. Evans, D.J.; Searles, D.J. The Fluctuation Theorem. *Adv. Phys.* **2012**, *51*, 1529–1585. [CrossRef]
71. Gallavotti, G.; Cohen, E.G.D. Dynamical Ensembles in Nonequilibrium Statistical Mechanics. *Phys. Rev. Lett.* **1995**, *74*, 2694. [CrossRef]
72. Kurchan, J. Fluctuation theorem for stochastic dynamics. *J. Phys. A Math. Gen.* **1998**, *31*, 3719. [CrossRef]
73. Searles, D.J.; Evans, D.J. Ensemble dependence of the transient fluctuation theorem. *J. Chem. Phys.* **2000**, *13*, 3503. [CrossRef]
74. Seifert, U. Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Phys. Rev. Lett.* **2005**, *95*, 040602. [CrossRef] [PubMed]
75. Abreu, D.; Seifert, U. Extracting work from a single heat bath through feedback. *EuroPhys. Lett.* **2011**, *94* 10001. [CrossRef]
76. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **2012**, *75*, 26001. [CrossRef]
77. Spinney, R.E.; Ford, I.J. Fluctuation relations: A pedagogical overview. *arXiv* **2012**, arXiv:1201.6381S.
78. Haas, K.R.; Yang, H.; Chu, J.-W. Trajectory Entropy of Continuous Stochastic Processes at Equilibrium. *J. Phys. Chem. Lett.* **2014**, *5*, 999. [CrossRef]
79. Van den Broeck, C. Stochastic thermodynamics: A brief introduction. *Phys. Complex Colloids* **2013**, *184*, 155–193.
80. Murashita, Y. Absolute Irreversibility in Information Thermodynamics. *arXiv* **2015**, arXiv:1506.04470.
81. Tomé, T. Entropy Production in Nonequilibrium Systems Described by a Fokker-Planck Equation. *Braz. J. Phys.* **2016**, *36*, 1285–1289. [CrossRef]
82. Salazar, D.S.P. Work distribution in thermal processes. *Phys. Rev. E* **2020**, *101*, 030101. [CrossRef] [PubMed]
83. Kullback, S. Letter to the Editor: The Kullback-Leibler distance. *Am. Stat.* **1951**, *41*, 340–341.
84. Sagawa, T. *Thermodynamics of Information Processing in Small Systems*; Springer: Berlin/Heidelberg, Germany, 2012.
85. Nielsen, M.A.; Chuang, I.L. *Quantum Computation and Quantum Information*; Cambridge University Press: Cambridge, UK, 2000.
86. Landauer, R. Irreversibility and heat generation in the computing process. *IBM J. Res. Dev.* **1961**, *5*, 183–191. [CrossRef]
87. Bérut, A.; Arakelyan, A.; Petrosyan, A.; Ciliberto, S.; Dillenschneider, R.; Lutz, E. Experimental verification of Landauer's principle linking information and thermodynamics. *Nature* **2012**, *483*, 187. [CrossRef]
88. Leff, H.S.; Rex, A.F. *Maxwell's Demon: Entropy, Information, Computing*; Princeton University Press: Princeton, NJ, USA, 1990.
89. Bekenstein, J.D. How does the entropy/information bound work? *Found. Phys.* **2005**, *35*, 1805. [CrossRef]
90. Capozziello, S.; Luongo, O. Information entropy and dark energy evolution. *Int. J. Mod. Phys. D* **2018**, *27*, 1850029. [CrossRef]
91. Kawai, R.; Parrondo, J.M.R.; Van den Broeck, C. Dissipation: The phase-space perspective. *Phys. Rev. Lett.* **2007**, *98*, 080602. [CrossRef] [PubMed]
92. Esposito, M.; Van den Broeck, C. Second law and Landauer principle far from equilibrium. *Europhys. Lett.* **2011**, *95*, 40004. [CrossRef]
93. Horowitz, J.; Jarzynski, C.R. An illustrative example of the relationship between dissipation and relative entropy. *Phys. Rev. E* **2009**, *79*, 021106. [CrossRef] [PubMed]
94. Parrondo, J.M.R.; van den Broeck, C.; Kawai, R. Entropy production and the arrow of time. *New J. Phys.* **2009**, *11*, 073008. [CrossRef]
95. Deffner, S.; Lutz, E. Information free energy for nonequilibrium states. *arXiv* **2012**, arXiv:1201.3888.
96. Horowitz, J.M.; Sandberg, H. Second-law-like inequalities with information and their interpretations. *New J. Phys.* **2014**, *16*, 125007. [CrossRef]
97. Nicholson, S.B.; García-Pintos, L.P.; del Campo, A.; Green, J.R. Time-information uncertainty relations in thermodynamics. *Nat. Phys.* **2020**, *16*, 1211–1215. [CrossRef]
98. Flego, S.P.; Frieden, B.R.; Plastino, A.; Plastino, A.R.; Soffer, B.H. Nonequilibrium thermodynamics and Fisher information: Sound wave propagation in a dilute gas. *Phys. Rev. E* **2003**, *68*, 016105. [CrossRef]
99. Carollo, A.; Spagnolo, B.; Dubkov, A.A.; Valenti, D. On quantumness in multi-parameter quantum estimation. *J. Stat. Mech. Theory E* **2019**, *2019*, 094010. [CrossRef]
100. Carollo, A.; Valenti, D.; Spagnolo, B. Geometry of quantum phase transitions. *Phys. Rep.* **2020**, *838*, 1–72 . [CrossRef]
101. Davies, P. Does new physics lurk inside living matter?. *Phys. Today* **2020**, *73*, 34. [CrossRef]
102. Sjöqvist, E. Geometry along evolution of mixed quantum states. *Phys. Rev. Res.* **2020**, *2*, 013344. [CrossRef]
103. Briët, J.; Harremoës, P. Properties of classical and quantum Jensen-Shannon divergence. *Phys. Rev. A* **2009**, *79*, 052311. [CrossRef]
104. Casas, M.; Lambertim, P.; Lamberti, P.; Plastino, A.; Plastino, A.R. Jensen-Shannon divergence, Fisher information, and Wootters' hypothesis. *arXiv* **2004**, arXiv:quant-ph/0407147.
105. Sánchez-Moreno, P.; Zarzo, A.; Dehesa, J.S. Jensen divergence based on Fisher's information. *J. Phys. A Math. Theor.* **2012**, *45*, 125305. [CrossRef]
106. López-Ruiz, L.; Mancini, H.; Calbet, X. A statistical measure of complexity. *Phys. Lett. A* **1995**, *209*, 321–326. [CrossRef]
107. Cafaro, C.; Alsing, P.M. Information geometry aspects of minimum entropy production paths from quantum mechanical evolutions. *Phys. Rev. E* **2020**, *101*, 022110. [CrossRef]
108. Ashida, K.; Oka, K. Stochastic thermodynamic limit on E. coli adaptation by information geometric approach. *Biochem. Biophys. Res. Commun.* **2019**, *508*, 690–694. [CrossRef]
109. Risken, H. *The Fokker-Planck Equation: Methods of Solution and Applications*; Springer: Berlin, Germany, 1996.

110. Van Den Brock, C. On the relation between white shot noise, Gaussian white noise, and the dichotomic Markov process. *J. Stat. Phys.* **1983**, *31*, 467–483. [CrossRef]

111. Bena, I. Dichotomous Markov Noise: Exact results for out-of-equilibrium systems (a brief overview). *Int. J. Mod. Phys. B* **2006**, *20*, 2825–2888. [CrossRef]

112. Onsager, L.; Machlup, S. Fluctuations and Irreversible Processes. *Phys. Rev.* **2951**, *91*, 1505–1512. [CrossRef]

113. Parrondo, J.M.R.; de Cisneros, B.J.; Brito R. Thermodynamics of Isothermal Brownian Motors. In *Stochastic Processes in Physics, Chemistry, and Biology*; Freund, J.A., Pöschel, T., Eds.; Lecture Notes in Physics; Springer: Berlin/Heidelberg, Germany, 2000; Volume 557. [CrossRef]

114. Gaveau, B.; Granger, L.; Moreau, M.; Schulman, L. Dissipation, interaction, and relative entropy. *Phys. Rev. E* **2014**, *89*, 032107. [CrossRef] [PubMed]

115. Ignacio A.; Martínez, G.B.; Jordan M.H.; Juan M.R.P. Inferring broken detailed balance in the absence of observable currents. *Nat. Commun.* **2019**, *10*, 3542.

116. Roldán, É.; Barral, J.; Martin, P.; Parrondo, J.M.R.; Jülicher, F. Quantifying entropy production in active fluctuations of the hair-cell bundle from time irreversibility and uncertainty relations. *New J. Phys.* **2021**, *23*, 083013. [CrossRef]

117. Chevallier, E.; Kalunga, E.; Angulo, J. Kernel Density Estimation on Spaces of Gaussian Distributions and Symmetric Positive Definite Matrices. 2015. Aavailable online: hal.archives-ouvertes.fr/hal-01245712 (accessed on 29 September 2021).

118. Nicolis, G.; Prigogine, I. *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order through Fluctuations*; John Wiley and Son: New York, NY, USA, 1977.

119. Prigogine, I. Time, structure, and fluctuations. *Science* **1978**, *201*, 777–785. [CrossRef]

120. Jaynes, E.T. The Minimum Entropy Production Principle. *Ann. Rev. Phys. Chem.* **1980**, *31*, 579–601. [CrossRef]

121. Mehdi, N. On the Evidence of Thermodynamic Self-Organization during Fatigue: A Review. *Entropy* **2020**, *22*, 372.

122. Dewar, R.C. Information theoretic explanation of maximum entropy production, the fluctuation theorem and self-organized criticality in non-equilibrium stationary states. *J. Phys. A. Math. Gen.* **2003**, *36*, 631–641. [CrossRef]

123. Sekhar, J.A. Self-Organization, Entropy Generation Rate, and Boundary Defects: A Control Volume Approach. *Entropy* **2021**, *23*, 1092. [CrossRef] [PubMed]

124. Philipp, S.; Thomas, F.; Ray, D.; Friston, K.J. Exploration, novelty, surprise, and free energy minimization. *Front. Psychol.* **2013**, *4*, 1–5.

125. Friston, K.J. The free-energy principle: A unified brain theory? *Nat. Rev. Neurosci.* **2012**, *11*, 127–138. [CrossRef] [PubMed]