Investigation of the statistical distance to reach stationary distributions

Nicholson, S. B. & Kim, E. J.

Original citation & hyperlink:
Nicholson, SB & Kim, EJ 2015, 'Investigation of the statistical distance to reach stationary distributions' Physics Letters A, vol. 379, no. 3, pp. 83-88.
https://dx.doi.org/10.1016/j.physleta.2014.11.003

DOI 10.1016/j.physleta.2014.11.003
ISSN 0375-9601

Publisher: Elsevier

Investigation of the Statistical Distance to reach Stationary Distributions

S.B. Nicholson[1] and Eun-jin Kim[1]

[1]*Department of Applied Mathematics, University of Sheffield, Sheffield, S3 7RH, UK*

(Dated: August 17th 2013)

The thermodynamic length gives a Riemannian metric to a system's phase space. Here we extend the traditional thermodynamic length to the information length ($\mathcal{L}$) out of equilibrium and examine its properties. We utilise $\mathcal{L}$ as a useful methodology of analysing non-equilibrium systems without evoking conventional assumptions such as Gaussian statistics, detailed balance, priori-known constraints, or ergodicity and numerically examine how $\mathcal{L}$ evolves in time for the logistic map in the chaotic regime depending on initial conditions. To this end, we propose a discrete version of $\mathcal{L}$ which is mathematically well defined by taking a set theoretic approach. We identify the areas of phase space where the loss of information of the system takes place most rapidly. In particular, we present an interesting result that the unstable fixed points turn out to most efficiently drive the logistic map towards a stationary distribution through $\mathcal{L}$.

Corresponding author, Tel: 07794023712
E-mail address: smp11sbn@sheffield.ac.uk.

INTRODUCTION

A major goal in statistical mechanics is to understand how non-equilibrium systems evolve in time. The main reason this is a difficult problem is that much of the theory and machinery of traditional Boltzmann Gibbs statistics does not carry over to the non-equilibrium regime. Furthermore, non-equilibrium systems are not guaranteed to have well-defined time-independent constraints which can be utilized in the determination of the form of the probability density functions (PDFs). Another important issue which is addressed in this manuscript is the amount of phase space in the course of a system's (e.g. rapid) time-evolution, as they are not guaranteed to have explored all possible states in the phase space, invalidating any assumption of ergodicity. Thus, the presence (or, existence) of phase space with zero probabilities is a potential problem for any system starting from a set of non-equilibrium conditions.

A general measure that has proven to be very appealing theoretically is the thermodynamic length ($\mathcal{L}_{th}$). The thermodynamic length endows a phase space with a Riemannian metric, thus allowing one to measure the "distance" that a system travels between thermodynamic equilibrium states. These systems are governed by a set of control parameters λ^i which are the experimentally controllable variables of the system, the thermodynamic length is defined as,

$$\mathcal{L}_{th} = \int_0^\tau dt \sqrt{\frac{d\lambda^i}{dt} g_{ij} \frac{d\lambda^j}{dt}}. \tag{1}$$

The metric g_{ij} depends on the parameters of the system being analysed. Most previous studies used thermodynamic functions to define g_{ij} based on equilibrium states. For instance, Weinhold [19] used $g_{ij} = d^2U(V,S,N)/dx_i dx_j$ ($x_i = U, V, N$ for $i = 1, 2, 3$), where U is the internal energy which is a function of the extensive variables. In comparison, Rupeiner [14] used the second derivative of the entropy with respect to extensive variables (for other examples see [2, 5, 12]). Out of equilibrium the control parameters are often not known, making Weinhold and Rupeiners metrics inapplicable. Thus, we take the approach of Crooks [3] and use the probability distribution function $p(x,t)$ to define the Fisher information matrix [7] as follows:

$$g_{ij} = \sum_x p(x,t) \frac{\partial \log p(x,t)}{\partial \lambda^i} \frac{\partial \log p(x,t)}{\partial \lambda^j}. \tag{2}$$

Here $p(x,t)$ is the probability of finding the system in state x at time t, given that it evolved from an initial distribution $p(x,t_0)$ at an earlier time, $t > t_0$ and the conservation of total probability requires $p(x,t)$ follow, $\sum_x p(x,t) = 1$. The control parameters λ^i specify how the system evolves through the surface of accessible states specified by λ^i. In equilibrium thermodynamics these could be for example the temperature or pressure of the system [3].

As we will see in the next section, putting Eq. (2) into Eq. (1) and summing over λ^i and λ^j gives us a distance in terms of probability distributions. It is important to note that Eq. (2) in general fulfils the requirements of a metric either whether the system is in equilibrium or not. Interestingly, in thermal equilibrium, using g_{ij} of Eq. (2) in Eq. (1) gives that $\mathcal{L}$ is proportional to the covariance of the forces conjugate to control parameters λ^i. That is, in equilibrium, thermodynamic length can be thought of as an integral in time over the fluctuations the system undergoes (see, e.g. [3]). Out of equilibrium, this is no longer true, and Eq. (2) is instead related to the integral of the covariance of fluctuations at different times [18].

A large body of theoretical work has already been developed for the thermodynamic length, starting with

Weinhold [19], Rupeiner [14] and Schlögl [16], continuing with [4] among others. There is however a distinct lack of numerical illustrations for the thermodynamic length. This is partly caused by the computational demand in time and difficulties associated with obtaining PDFs which are sufficiently accurate.

When our system evolves over a manifold of non-equilibrium states we will use the information length ($\mathcal{L}$) instead of the thermodynamic length. Using $\mathcal{L}$ the relaxation of an arbitrary configuration of the system will be numerically investigated as it relaxes to a stationary distribution. In this work stationary does not imply equilibrium, as equilibrium also requires the system satisfy detailed balance (as defined in section 3). In particular, we use a discrete map (the logistic map) as a typical example of a non-equilibrium system which also allows us to take computational advantages, as the simulation of maps is much less demanding and time-consuming than continuous systems. As noted above, for any non-equilibrium system having zero-valued probabilities $\mathcal{L}$ can be undefined. To overcome this we propose a discrete version of $\mathcal{L}$ which is mathematically well defined by taking a set theoretic approach.

The paper is organized as follows: Section 2 introduces the information length and provides its key properties in detail along with the definitions of our sets. These are followed in Section 3 where we show that $\mathcal{L}$ must increase for any PDFs other than the invariant (stationary) distribution. Section 4 will numerically examine the information length for the logistic map in the chaotic regime. In particular, using the logistic map we identify the areas of phase space where the conversion of the information of the system into work takes place most rapidly. We also show that the logistic map very often follows the path of minimum length. That is, the system follows the path of minimum information change. The importance of the minimum/optimal path has been noted in previous studies. For instance, J. Nulton et al utlilized this concept to link the thermodynamic length between equilibrium states to the "optimal" path in annealing processes [11]. In [5] it was suggested that experiments using biological motors would yield paths of minimum length. In Section 5, we show the curious result that the unstable fixed points rapidly drive the system to its stationary distribution. Conclusions are provided in Section 6.

INFORMATION LENGTH

To follow the path of a general non-equilibrium ensemble (e.g as it evolves towards equilibrium), we measure the Fisher-Rao information by using Eq. (2) in Eq. (1) and define the information length as follows,

$$\mathcal{L} = \int_0^\tau dt \sqrt{\sum_x \frac{1}{p(x,t)} \left[\frac{dp(x,t)}{dt}\right]^2}. \qquad (3)$$

Now distances are measured by the difference between consecutive PDFs. The difference in PDFs gives a measure of the statistical distance [11, 20]. The evolution of a system can then be envisioned as the trajectory in the probability space where the distance/metric at different times is provided by the statistical distance. As time is the only parameter, Eq. (3) is ideally suited for analysing experimental data, which we use exclusively in the remainder of the paper.

An alternative expression to Eq. (3) is often necessary to describe the evolution of non-equilibrium systems since $\mathcal{L}$ is undefined for PDFs with zero values (i.e. when $p(x,t) = 0$), as it is written in Eq. (3). This problem can be readily remedied by expressing $\mathcal{L}$ in terms of $q = \sqrt{p}$, as suggested by Wootters [20], which transforms Eq. (3) into the following form,

$$\mathcal{L} = 2\int_0^\tau dt \sqrt{\sum_x \left(\frac{dq(x,t)}{dt}\right)^2}, \qquad (4)$$

which no longer has this singularity problem. However if time is discrete, Eq. (4) is not equivalent to Eq. (3), making it necessary to look for a different form of $\mathcal{L}$. Thus, in this paper, we propose a set theoretical approach to overcome this problem, as presented shortly.

To this end, we utilize a discrete version of Eq. (3),

$$\mathcal{L} = \sum_{t=1}^{\tau} \Delta t \sqrt{\sum_x \frac{1}{p(x,t)} \left(\frac{\Delta p(x,t)}{\Delta t}\right)^2} = \sum_{t=1}^{\tau} \Delta\mathcal{L}(t)\Delta t. \qquad (5)$$

Here, $\Delta p(x,t) = p(x,t') - p(x,t)$ where $t' = t+1$. Note that for discrete systems, t simply denotes the iteration number, taking the integer values as $t = 1, 2, \ldots \tau$ where τ is the total time of a given evolution. Consequently, for most of this work the time step $\triangle t$ is 1. That is, $p(x,t')$ ($t' = t+1$) and $p(x,t)$ are the two consecutive PDFs (i.e. the probability of being in state x at time t' or t, respectively) while $\Delta p(x,t) = p(x,t') - p(x,t)$ is the difference between these two consecutive PDFs. As we are dealing with numerical simulations, our state space $\mathbf{X}$ will be coarse grained into a finite number of disjoint sets x, which represent the new "states" of the system. For the logistic map, examined shortly, x is a one dimensional variable, $x \in [-1, 1]$. The probability of being in "state" x at time t is then, $p(x,t)$, where, $\sum_x p(x,t) = 1$. It is straightforward to generalize x to any higher dimensions, $\mathbf{x} = \{x_i, x_j, \ldots x_N\}$.

To guarantee that $\mathcal{L}$ is well defined for arbitrary discrete non-equilibrium systems, we need to account for states that have zero probability of being occupied along the system's evolution. That is, given a total state space $\mathbf{X}$, the probability of finding the system $p(x,t)$ in a particular states x can be zero ($p(x,t) = 0$ for some $x \in \mathbf{X}$). As a result we define the following two sets depending on the evolution of PDFs at two consecutive times t and t' as

$$\begin{aligned} Q_p &= \{x : p(x,t) \neq 0 | p(x,t') = 0\}, \\ Q_w &= \{x : p(x,t) \neq 0 | p(x,t') \neq 0\}. \end{aligned} \tag{6}$$

One possibility that is not included in the above equation is the case where $p(x,t) = 0$ and $p(x,t') \neq 0$, which can however be shown to have no contribution to $\mathcal{L}$. The subscript p in Q_p designates the unused probability of evolving over one time step, while w is the set that gives a measure of the available work in evolving over one time step, as shown later. To isolate the separate contributions to $\mathcal{L}$ from Q_p and Q_w, we define

$$L_{Q_p} = \sum_{x \in Q_p} \frac{p(x,t)}{(\Delta t)^2},$$

$$L_{Q_w} = \sum_{x \in Q_w} \frac{1}{p(x,t)} \left(\frac{\Delta p(x,t)}{\Delta t} \right)^2,$$

and express Eq. (5) as:

$$\mathcal{L} = \sum_{t=1}^{\tau} \Delta t \sqrt{L_{Q_p} + L_{Q_w}}. \tag{7}$$

A consequence of the system only occupying Q_p is that, the system evolves independently at two consecutive times, resulting in $\Delta \mathcal{L}(t) = 1$.

Next we will look how $\mathcal{L}$ is related to the measure of work of the system. The links between the thermodynamic length and the dissipation of a system are already well known, [15]. The Fisher information is also known to be related to the relative entropy; see [1] and the references within for the continuous case and [6] for the discrete case. For completeness, we show how the relative entropy $DS[p(x,t)|p(x,t')]$ from Procaccia and Levine's work [13] is related to $\mathcal{L}$. To this end we define the 'microscopic' relative entropy as,

$$\mathcal{DS}[p(x,t)|p(x,t')] = p(x,t) \log \left(\frac{p(x,t)}{p(x,t')} \right), \tag{8}$$

and express DS in terms of $\mathcal{DS}$,

$$\begin{aligned} DS[p(x,t)|p(x,t')] &= \sum_x p(x,t) \log \left(\frac{p(x,t)}{p(x,t')} \right), \\ &= \sum_x \mathcal{DS}[p(x,t)|p(x,t')]. \end{aligned} \tag{9}$$

In order to show that L_{Q_w} in Eq. (7) is related to $\mathcal{DS}$, we use $p(x,t') = p(x,t) + \Delta p(x,t)$ ($\Delta p(x,t) = p(x,t') - p(x,t)$) in $\mathcal{DS}$ as:

$$\begin{aligned} \mathcal{DS}[p(x,t)|p(x,t')] &= -p(x,t) \log \left[1 + \frac{1}{p(x,t)} \Delta p(x,t) \right] \\ &= -\Delta p(x,t), \end{aligned}$$

where $\log(1+x) \approx x$ was used above. Therefore, to leading order in Δp, the substitution of Eq. (10) into Eq. (7) gives,

$$\mathcal{L} = \sum_{t}^{\tau} \Delta t \sqrt{\sum_{x \in Q_p} \frac{p(x,t)}{\Delta t^2} + \sum_{x \in Q_w} \frac{1}{p(x,t)} \left(\frac{\mathcal{DS}[p(x,t)|p(x,t')]}{\Delta t} \right)^2}, \tag{10}$$

where $\Delta t = 1$ for our discrete system. Eq. (10) then demonstrates that $\mathcal{L}$ due to L_{Q_w} is directly related to the available work in the evolution of the system via the 'microscopic' relative entropy $\mathcal{DS}$ to leading order in Δp. Higher order terms in Δp are not, however, negligible for strongly non-equilibrium process involving rapid change in PDFs with large Δp, suggesting that the relation between $\mathcal{DS}$ and L_{Q_w} is more subtle, for instance, requiring the generalisation of Fisher information to higher moments [10]. The discrepancy between Eq. (7) and Eq. (10) will in fact be shown to occur in our numerical example to follow for rapid equilibration. It is worth noting at this point that $\mathcal{L}$ as well as the relative entropy (or entropy) is dependent on the resolution that one uses to coarse grain the system. As this is not a main issue of this paper, we simply use the same coarse graining for all simulations presented in this paper.

PROPERTIES OF $\mathcal{L}$

The main aim of this section is to examine the key properties of $\mathcal{L}$ and the information that can be inferred from $\mathcal{L}$ about a non-equilibrium system. In particular, we identify the condition for $\Delta\mathcal{L}(t) > 0$. We then examine its physical implications in terms of the loss of available work and the change in the macroscopic observables and fluctuations.

It is easy to see that the lower bound on $\Delta\mathcal{L}(t) = 0$ occurs when p is stationary (i.e. $p(x,t) = p(x,t')$). An interesting question is then if this stationary condition is a sufficient and necessary condition for $\Delta\mathcal{L}(t) = 0$ regardless of whether a system is in equilibrium or out of equilibrium. To answer this question, we utilise the current of probability which flows from state $y \to x$ in one time step, which is defined for a stationary system without summation as

$$J^s_{xy} = C_{xy}p_0(y) - C_{yx}p_0(x). \tag{11}$$

Here, C_{xy} is defined as the non-negative irreducible matrix of transition probabilities from states y to x,

$$C_{xy} = Pr\left(state\ at\ (t' > t)\ is\ x | state\ at\ t\ is\ y\right). \tag{12}$$

The distribution $p_0(x)$ is guaranteed to be a unique stationary distribution of C_{xy} due to C_{xy} being irreducible [8]. In matrix notation the invariance of p_0 is expressed as $Cp_0 = p_0$. $\sum_y J^s_{xy} = 0$ is guaranteed since J^s_{xy} follows Kirchoff's loop rule that the amount of current into a state is equal to the amount out of a state. We define stationary as the PDF being time independent, since we do not assume any knowledge of mean values. The system can be characterized as being reversible or not through $J^s_{xy} = 0$, or $J^s_{xy} \neq 0$ respectively. This is due to J^s_{xy} being a measure of microscopic irreversibility. In general, we can define a non-equilibrium current J_{xy} as,

$$J_{xy} = C_{xy}p_t(y) - C_{yx}p_t(x). \tag{13}$$

Summing Eq. (13) over y gives,

$$\sum_y J_{xy} = p_{t'}(x) - p_t(x) = \Delta p(x,t). \tag{14}$$

This allows us to link the operator C_{xy} to $\mathcal{L}$. Obviously, when $\sum_y J_{xy} = 0$, $p(x,t') = p(x,t)$, i.e. the distribution is stationary with $\Delta\mathcal{L}(t) = 0$. $\Delta\mathcal{L}(t) = 0$ is also guaranteed under the stricter condition of detailed balance which defines true equilibrium, when $J_{xy} = 0\ \forall\ x, y$ in Eqs. (11) or (13). Therefore, in view of the uniqueness of $p_0(x)$, we can infer that if $\sum_y J_{xy} = 0\ \forall\ x$ then $J_{xy} = J^s_{xy}$ and $\Delta\mathcal{L}(t) = 0$, meaning that the system is stationary. This shows that irreducibility is necessary for stationarity to uniquely imply $\Delta\mathcal{L} = 0$. If the system does not have an irreducible operator, then it is possible that $\Delta\mathcal{L}(t) = 0$ in general, as we will show in our numerical results.

Most systems in nature never truly reach a stationary state, making the aforementioned discussion too idealized. We thus express $p(x,t)$ in general as the sum of its invariant distribution $p_0(x)$ and the fluctuations f as,

$$p(x,t) = p_0(x) + f(x,t). \tag{15}$$

The direct substitution of Eq. (15) into the discrete version of Eq. (5) over one time step yields,

$$\frac{\Delta\mathcal{L}}{\Delta t} = 2\sqrt{\sum_x \frac{(\triangle f)^2}{p(x,t)}}, \tag{16}$$

where $\Delta f = f(x,t') - f(x,t)$. The above expression immediately reveals that the driving force for the change in $\mathcal{L}$ is time-variation in the fluctuations from equilibrium. The same line of reasoning carries over to continuous systems as well.

The property of $\Delta\mathcal{L}(t) > 0$ for non-equilibrium and non-stationary systems makes $\Delta\mathcal{L}(t)$ a useful quantity to investigate non-equilibrium and non-stationary processes given some knowledge of the fluctuations in the system.

EQUILIBRATION AND SIMULATION

Having identified a physical meaning of $\mathcal{L}$ as a measure of available work, it is of interest to investigate a specific non-equilibrium system. To this end, we utilize the logistic map. The logistic map is used as it is a simple non-equilibrium system which exhibits much of the interesting properties of $\mathcal{L}$. But also since the logistic map can be considered one of the most difficult classes of systems to analyse using our sets, being a non-differentiable (in time) discrete system. The system being discrete in time means that for almost the entire evolution of the system we have zero-probability states, $p(x,t) = 0$ while $p(x,t') \neq 0$. This means that our set representation of the evolution is vital to avoid un-physical infinite lengths. We recall that the logistic map is governed by the following mapping

$$x_{t+1} = 1 - ax_t^2, \tag{17}$$

which describes the position of an orbit x_{t+1} at time $t+1$ as a function of its position x_t at the earlier time t. a is the control parameter, which is taken to be 2 for simulating a chaotic region. The stationary density for $a = 2$ is given by $p_0 = 1/\pi(1-x^2)^{1/2}$. In this chaotic region, the map has the two unstable fixed points $x = -1$ and $x = 1/2$, which turn out to play an interesting role in $\Delta\mathcal{L}(t)$ as shown later.

A key question of our interest is how an initial state far from equilibrium approaches $p_0(x)$ in probability space in terms of $\Delta\mathcal{L}(t)$. For instance, is there any unique property of $\Delta\mathcal{L}(t)$ that can be identified for all evolutions starting from different initial conditions? To answer this question, we perform numerical simulation of Eq. (17) starting from an initial PDF which is strongly localised at $x = x_0$, approximated by a delta function. For each simulation using different initial x_0, the domain, $[-1, 1]$ will be broken into M bins, with the width of each bin $\frac{2}{M}$. The number of bins used is a free parameter after all "There is no law of nature that defines the coarse grains" [17]. Here we have fixed the number so as to make each simulation comparable. $p(x, t)$ thus represents the probability of finding an orbit in bin x at time t.

Using the random initial distribution of $M = 9 \times 10^7$ points, centred at $x_0 = -0.553$, we first check the validity of approximating Eq. (5) with Eq. (10). Interestingly, Fig. (1) shows that $\mathcal{L}$ given in Eq. (5) plotted in the solid black line with solid dots agrees very well for most of the evolution with $\mathcal{L}$ given by Eq. (10) shown by the line with circles, respectively. It is seen from Fig. (1) that initially, the PDFs never overlap at the two consecutive times, occupying only set Q_p. For $12 < t \leq 15$, the PDFs overlap and change rather rapidly but still do not fill the whole state space. In this regime, approximation of the derivative seems to give errors, causing the difference in the results from Eq. (5) and Eq. (10). This is a clear manifestation of the difference between the local relative entropy and $\mathcal{L}$ in a strongly non-equilibrium evolution. For $15 < t \leq 20$, the PDFs fill out the entire domain $[-1, 1]$ but still are not in the stationary distribution. The less dramatic change of the PDF on each time step recovers a good agreement between Eq. (10) and Eq. (5). From $t = 21$ on, the system fluctuates around the stationary distribution and thus both equations are trivially near zero. We have checked that a similar agreement is also obtained for all other initial conditions that are considered in this paper.

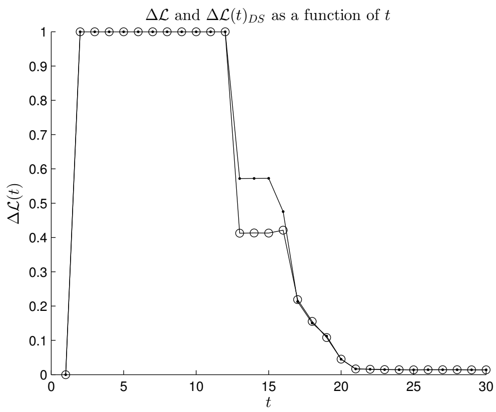

FIG. 1: Plot of the discrete version of $\mathcal{L}$ equation (5) against time in black which shows a good agreement with equation (10) plotted in black with circles. Both use $M = 1 \times 10^6$ initial points who all start as a delta function around, $x_o = 0.3826834$.

The evolution of $\mathcal{L}$ starting at $x_0 = -0.553$ is shown in Fig. (2) where we see that the system for almost its entire evolution follows the minimum path, i.e. a straight line. For $0 < t \leq 12$, PDFs do not overlap on each time step (as mentioned above), and thus have a slope of $\Delta\mathcal{L}(t) = 1$. When $12 < t \leq 16$, $\Delta\mathcal{L}(t)$ also follows a slope of $\Delta\mathcal{L}(t) = 0.41453$. For $16 < t \leq 20$, there is a non-linear transition towards the stationary distribution. Finally, for $t > 20$, the system has approximately come to the stationary distribution, resulting in almost no increase in $\mathcal{L}$.

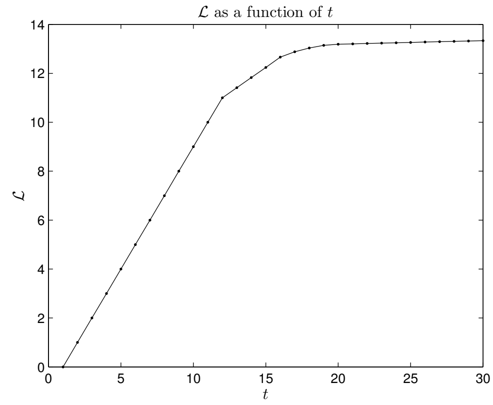

FIG. 2: The evolution of $\mathcal{L}$ starting from $x_0 = -0.553$ using $M = 9 \times 10^7$ orbits.

The finite discretization of the domain for numerical simulation artificially takes areas of measure zero, such as the sink at $x = -1$ and increases their influence to areas of non-zero measure. That is, orbits for short time periods may land very near a sink. On the next time step, due to their proximity near the sink their small movement again lands them in the same bin, this creates the appearance of a fixed point. This results in the decrease of the slope at $t = 12$ in Fig. (2). Here the PDFs overlap once they have landed in the bin which has the $x = -1$ fixed point. Figure (3) shows two consecutive PDFs near $x = -1$. Since only part of the orbits are able to leak out of the Bin containing $x = -1$ the PDFs overlap on subsequent time steps between -1 and approximately -0.75, thus also occupying Q_w. This results in the slope of $\Delta\mathcal{L}(t) < 1$. Yet since the rate in which orbits leave the bin that includes $x = -1$ is constant, the reduced slope is also constant. The constant slope is equivalent to the system taking the path of minimum available work through Eq. (10).

To understand how the initial position x_0 and the unstable fixed points are related to $\mathcal{L}$, we plot in Fig. (4) the total change in $\mathcal{L}$ starting from different initial delta functions uniformly spread over the domain. The total change in $\mathcal{L}$ between $t = 0$ and the final time when the evolution reaches its invariant density varies with the initial position x_0. Interestingly, the total change

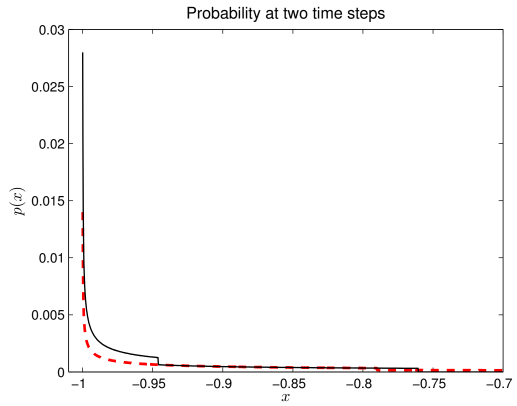

FIG. 3: $P(x, 13)$ plotted in black and $p(x, 14)$ is plotted in red with the dashed line.

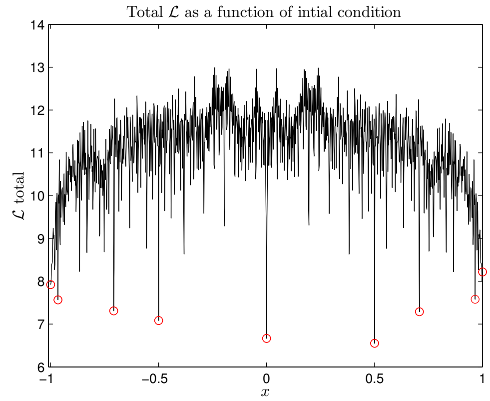

FIG. 4: The evolution of $\mathcal{L}$ as a function of time for many initial conditions spread over the domain. Most initial conditions travel a distance of between 13 and 16 before reaching $p_0(x)$. The points $x_0 = [-1, -0.96, -0.708, -0.5, 0, 0.5, 0.708, 0.96, 1]$ whose initial conditions are marked with circles, start at or quickly occupy the bin of a fixed point and thus reach $p_0(x)$ in a far shorter distance.

in $\mathcal{L}$ takes the minimum value for the initial ensembles starting from or quickly entering the two unstable fixed points $x = -1, 1/2$. Some of these initial conditions $x_0 = [-1, -0.96, -0.708, -0.5, 0, 0.5, 0.708, 0.96, 1]$ are marked with the circles in Fig. (4). All of these initial conditions reach fixed points in 5 iterations or less. Since $\mathcal{L}$ represents the statistical distance between the initial PDFs and the final, invariant density, this means that the unstable fixed points are what is most efficiently converting available work into wasted work such as heat. Phrased another way, the fixed points reduce the information of the PDF, bringing each PDF nearer to the invariant density, which is the distribution with the highest disorder [9]. If one wished to then prolong the distance to a stationary distribution or conversely find the shortest path to said distribution, one simply finds the path that rarely (or quickly) comes into the vicinity of an unstable fixed point.

There are more complicated paths the system can take to the stationary distribution than those presented above. For instance, starting at $x_0 = 0.7071$ gives us Fig. (5), where we can see four main phases involved in its evolution. The first is again only occupying Q_p for $0 < t \leq 4$. For $4 < t < 7$, the orbits all fall into the bin that includes $x = -1$. This does not contradict the results in Section 2 where we showed that for a system governed by an irreducible operator $\Delta\mathcal{L}(t) > 0$ out of equilibrium. For if one were to build an operator from the paths the system has taken it would indeed be reducible. For $7 < t \leq 16$ the orbits have escaped the attractor and the slope is again less than one. Finally the system quickly reaches the stationary distribution for $t > 16$.

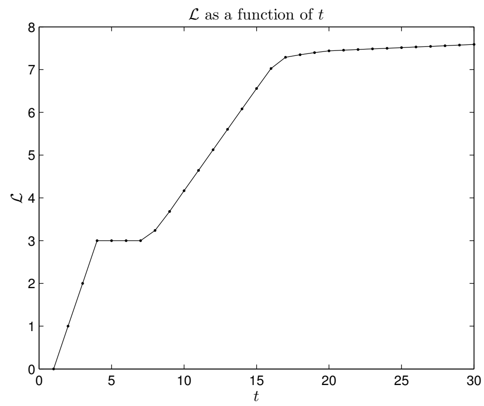

FIG. 5: The evolution of $\mathcal{L}$ starting from $xo = 0.7071$. The evolution is divided up into four main phases. $0 < t \leq 4$, all $x \in Q_p$, $4 < t \leq 7$ all orbits are in the bin that holds the $x = -1$ fixed point, though the operator that would be made from the orbits is reducible. $7 < t \leq 16$, $\Delta\mathcal{L}(t) < 1$ as the PDFs overlap and the information changes. $t > 16$ the system settles into $p_0(x)$.

CONCLUSION

In this paper we have investigated both theoretically and numerically the information length using our set theoretic approach. We have shown that $\frac{d\mathcal{L}}{dt} > 0$ is guaranteed for systems out of equilibrium as long as the system is evolved under an irreducible operator.

The two sets that contribute to $\mathcal{L}$ are Q_p which is the amount of probability not being used in the systems evolution. Due to conservation of probability, when one PDF does not intersect with the PDF at the next time step only set Q_p is occupied and $\Delta\mathcal{L}(t) = 1$, meaning the system has no correlation with itself in time. When the system's PDFs start intersecting at the two subsequent times we have non-zero Q_w and the rate of change in $\mathcal{L}$ decreases in time. This is because the available work attributed at each state (measured with $\mathcal{DS}$) is reduced through the conversion of available probability in Q_p. The logistic map was used to corroborate our results. An interesting result of this simulation is that the system almost always follows the minimum path. The only time it appears to deviate from this is when the system is transitioning from a non-stationary distribution that fills the entire phase space to the invariant distribution. We also showed the special role of unstable fixed points as the most efficient areas of state space to convert a non-equilibrium distribution into the invariant density for the logistic map. This curious result may warrant further investigation as to the scope of its generality in other systems. Future work will also include a more detailed investigation between the total change in $\mathcal{L}$ and the structure of attractors (e.g. various unstable orbits).

We emphasize that our set methodology capitalizes on the attributes of the information length lending strong generality to the systems that can be studied while illustrating the relationship between the distance a system travels in state space and the available from that evolution. This is an improvement on other methods that rely on assumptions such as that of detailed balance, distinct PDFs (such as Gaussian, etc) or ergodicity in the system. S.B. Nicholson would like to thank Stephen Chaffin for his many useful discussions.

[1] A. Plastino A.R. Plastino, M. Casas. Fisher's information, kullback's measure, and h-theorems. Physics Letters A, 246, 1998.

[2] G. P. Beretta. Modeling non-equilibrium dynamics of a discrete probability distribution: General rate equation for maximal entropy generation in a maximum-entropy landscape with time-dependent constraints. Entropy, 10:160–180, 2008.

[3] G. E. Crooks. Measuring thermodynamic length. Physical Review Letters, 99:10060–2–10060–4, 2007.

[4] L. Diósi, K. Kulacsy, B. Lukács, and A. Rácz. Thermodynamic length, time, speed, and optimum path to minimize entropy. Journal of Chemical Physics, 105:11220–11225, 1996.

[5] E. H. Feng and G. E. Crooks. Far-from-equilibrium measurements of thermodynamic length. Physical review E, 79:012104–1–012104–4, 2009.

[6] B. R. Frieden. Lagrangians of physics and the game of fisher-information transfer. Physical Review E, 52:2274–2286, 1995.

[7] B. R. Frieden. Science from Fisher Information, volume 2. Cambridge University Press, 2004.

[8] R. A. Horn and C. R. Johnson. Matrix Analysis. Cambridge University Press, 2009.

[9] E.T. Jaynes. Information theory and statistical mechanics. Physical Review, 106:620–630, 1957.

[10] E. Lutwak, D. Yang, and G. Zhang. Cramer-rao and moment-entropy inequalities for renyi entropy and generalized fisher information. IEEE Transactions on Information Theory, 51:473–478, 2005.

[11] J. Nulton, P. Salamon, B. Andresen, and Q. Anmin. Quasistatic processes as step equilibrations. J. Chem. Phys, 83:334–338, 1985.

[12] M. Polettini and M. Esposito. Nonconvexity of the relative entropy for markov dynamics: A fisher information approach. Physical Review E., 88, 2013.

[13] I. Procaccia and R. D. Levine. Potential work: A statistical-mechanical approach for systems in disequilibrium. The Journal of Chemical Physics, 65:3357–3364, 1976.

[14] G. Rupeiner. Thermodynamics: A riemannian geometric model. Physical Review A, 20:1608–1613, 1979.

[15] P. Salamon and R. S. Berry. Thermodynamic length and dissipated availability. Physical Review Letters, 51:1127–1130, 1983.

[16] F. Schlögl. Thermodynamic metric and stochastic measures. Z. Phys. B, 59:449–454, 1985.

[17] L.S. Schulman and B. Gaveau. Course grains: The emergence of space and order. Foundations of Physics, 31:713–731, 2001.

[18] D. A. Sivak and G. E. Crooks. Thermodynamic metrics and optimal paths. Physical Review Letters, 108:190602–1–190602–5, 2012.

[19] F. Weinhold. Metric geometry of equilibrium thermodynamics. Journal of Chemical Physics, 63:2479–2483, 1975.

[20] W.K. Wootters. Statistical distance in hilbert space. Physical Review D, 23:357–362, 1981.