

Article

# Investigating Information Geometry in Classical and Quantum Systems through Information Length

Eun-jin Kim 

School of Mathematics and Statistics, University of Sheffield, Sheffield S3 7RH, UK;  
e.kim@sheffield.ac.uk or e.kim@shef.ac.uk; Tel.: +44-114-222-3876

Received: 19 July 2018; Accepted: 1 August 2018; Published: 3 August 2018



**Abstract:** Stochastic processes are ubiquitous in nature and laboratories, and play a major role across traditional disciplinary boundaries. These stochastic processes are described by different variables and are thus very system-specific. In order to elucidate underlying principles governing different phenomena, it is extremely valuable to utilise a mathematical tool that is not specific to a particular system. We provide such a tool based on information geometry by quantifying the similarity and disparity between Probability Density Functions (PDFs) by a metric such that the distance between two PDFs increases with the disparity between them. Specifically, we invoke the information length  $\mathcal{L}(t)$  to quantify information change associated with a time-dependent PDF that depends on time.  $\mathcal{L}(t)$  is uniquely defined as a function of time for a given initial condition. We demonstrate the utility of  $\mathcal{L}(t)$  in understanding information change and attractor structure in classical and quantum systems.

**Keywords:** stochastic processes; Langevin equation; Fokker–Planck equation; information length; Fisher information; relaxation; chaos; attractor; probability density function

## 1. Introduction

Stochastic processes are ubiquitous in nature and laboratories, and play a major role across traditional disciplinary boundaries. Due to the randomness associated with stochasticity, the evolution of these systems is not deterministic but instead probabilistic. Furthermore, these stochastic processes are described by different variables and are thus very system-specific. This system-specificity makes it impossible to make comparison among different processes. In order to understand universality or underlying principles governing different phenomena, it is extremely valuable to utilise a mathematical tool that is not specific to a particular system. This is especially indispensable given the diversity of stochastic processes and the growing amount of data.

Information geometry provides a powerful methodology to achieve this goal. Specifically, the similarity and disparity between Probability Density Functions (PDFs) is quantified by a metric [1] such that the distance between two PDFs increases with the disparity between them. This was the very idea behind a statistical distance [2] based on the Fisher (or Fisher–Rao) metric [3] which represents the total number of statistically different states between two PDFs in Hilbert space for quantum systems. The analysis in [2] was extended to impure (mixed-state) quantum systems using a density operator by [4]. Other related work includes [5–12]. For Gaussian PDFs, a statistically different state is attained when the physical distance exceeds the resolution set by the uncertainty (PDF width).

This paper presents a method to define such distance for a PDF which changes continuously in time, as is often the case of non-equilibrium systems. Specifically, we invoke the information length  $\mathcal{L}(t)$  according to the total number of statistically different states that a system evolves through in time.  $\mathcal{L}(t)$  is uniquely defined as a function of time for a given initial condition. We demonstrate

the utility of  $\mathcal{L}(t)$  in understanding information change and attractor structure in classical and quantum systems [13–21].

This paper is structured as follows: Section 2 discusses information length and Section 3 investigates attractor structure. Sections 4 and 5 present the analysis of classical music and quantum systems, respectively. Conclusions are found in Section 6.

## 2. Information Length

Intuitively, we define the information length  $\mathcal{L}$  by computing how quickly information changes in time and then measuring the clock time based on that time scale. Specifically, the time-scale of information change  $\tau$  can be computed by the correlation time of a time-dependent PDF, say  $p(x, t)$ , as follows.

$$\frac{1}{\tau^2} = \int dx \frac{1}{p(x, t)} \left[ \frac{\partial p(x, t)}{\partial t} \right]^2. \quad (1)$$

From Equation (1), we can see that the dimension of  $\tau = \tau(t)$  is time and serves as a dynamical time unit for information change.  $\mathcal{L}(t)$  is the total information change between time 0 and  $t$ :

$$\mathcal{L}(t) = \int_0^t \frac{dt_1}{\tau(t_1)} = \int_0^t dt_1 \sqrt{\int dx \frac{1}{p(x, t_1)} \left[ \frac{\partial p(x, t_1)}{\partial t_1} \right]^2}. \quad (2)$$

In principle,  $\tau(t)$  in Equation (1) can depend on time, so we need the integral for  $\mathcal{L}$  in Equation (2). To make an analogy, we can consider an oscillator with a period  $\tau = 2$  s. Then, within the clock time 10 s, there are five oscillations. When the period  $\tau$  is changing with time, we need an integration of  $dt/\tau$  over the time interval.

We now recall how  $\tau(t)$  and  $\mathcal{L}(t)$  in Equations (1) and (2) are related to the relative entropy (Kullback–Leibler divergence) [15,16]. We consider two nearby PDFs  $p_1 = p(x, t_1)$  and  $p_2 = p(x, t_2)$  at time  $t = t_1$  and  $t_2$  and the limit of a very small  $\delta t = t_2 - t_1$  to do Taylor expansion of  $D[p_1, p_2] = \int dx p_2 \ln(p_2/p_1)$  by using

$$\frac{\partial}{\partial t_1} D[p_1, p_2] = - \int dx p_2 \frac{\partial_{t_1} p_1}{p_1}, \quad (3)$$

$$\frac{\partial^2}{\partial t_1^2} D[p_1, p_2] = \int dx p_2 \left\{ \frac{(\partial_{t_1} p_1)^2}{p_1^2} - \frac{\partial_{t_1}^2 p_1}{p_1} \right\}, \quad (4)$$

$$\frac{\partial}{\partial t_2} D[p_1, p_2] = \int dx \left\{ \partial_{t_2} p_2 + \partial_{t_2} p_2 [\ln p_2 - \ln p_1] \right\}, \quad (5)$$

$$\frac{\partial^2}{\partial t_2^2} D[p_1, p_2] = \int dx \left\{ \partial_{t_2}^2 p_2 + \frac{(\partial_{t_2} p_2)^2}{p_2} + \partial_{t_2}^2 p_2 [\ln p_2 - \ln p_1] \right\}. \quad (6)$$

In the limit  $t_2 \rightarrow t_1 = t$  ( $p_2 \rightarrow p_1 = p$ ), Equations (3)–(6) give us

$$\begin{aligned} \lim_{t_2 \rightarrow t_1} \frac{\partial}{\partial t_1} D[p_1, p_2] &= \lim_{t_2 \rightarrow t_1} \frac{\partial}{\partial t_2} D[p_1, p_2] = \int dx \partial_t p = 0, \\ \lim_{t_2 \rightarrow t_1} \frac{\partial^2}{\partial t_1^2} D[p_1, p_2] &= \lim_{t_2 \rightarrow t_1} \frac{\partial^2}{\partial t_2^2} D[p_1, p_2] = \int dx \frac{(\partial_t p)^2}{p} = \frac{1}{\tau^2}. \end{aligned} \quad (7)$$

Up to  $O((dt)^2)$  ( $dt = t_2 - t_1$ ), Equation (7) and  $D(p_1, p_1) = 0$  lead to

$$D[p_1, p_2] = \frac{1}{2} \left[ \int dx \frac{(\partial_t p(x, t))^2}{p(x, t)} \right] (dt)^2, \quad (8)$$

and thus the infinitesimal distance  $dl(t_1)$  between  $t_1$  and  $t_1 + dt$  as

$$dl(t_1) = \sqrt{D[p_1, p_2]} = \frac{1}{\sqrt{2}} \sqrt{\int dx \frac{(\partial_{t_1} p(x, t_1))^2}{p(x, t_1)}} dt. \tag{9}$$

By summing  $dl(t_i)$  for  $i = 0, 1, 2, \dots, n - 1$  (where  $n = t/dt$ ) in the limit  $dt \rightarrow 0$ , we have

$$\lim_{dt \rightarrow 0} \sum_{i=0}^{n-1} dl(idt) = \lim_{dt \rightarrow 0} \sum_{i=0}^{n-1} \sqrt{D[p(x, idt), p(x, (i+1))] dt} \propto \int_0^t dt_1 \sqrt{\int dx \frac{(\partial_{t_1} p(x, t_1))^2}{p(x, t_1)}} = \mathcal{L}(t), \tag{10}$$

where  $\mathcal{L}(t)$  is the information length. Thus,  $\mathcal{L}$  is related to the sum of infinitesimal relative entropy. It cannot be overemphasised that  $\mathcal{L}$  is a Lagrangian distance between PDFs at time 0 and  $t$  and sensitively depends on the particular path that a system passed through reaching the final state. In contrast, the relative entropy  $D[p(x, 0), p(x, t)]$  depends only on PDFs at time 0 and  $t$  and thus does not tell us about intermediate states between initial and final states.

### 3. Attractor Structure

Since  $\mathcal{L}(t)$  represents the accumulated change in information (due to the change in PDF) at time  $t$ ,  $\mathcal{L}(t)$  settles to a constant value  $\mathcal{L}_\infty$  when a PDF reaches its final equilibrium PDF. The smaller  $\mathcal{L}_\infty$ , the smaller number of states that the initial PDF passes through to reach the final equilibrium. Therefore,  $\mathcal{L}_\infty$  provides us with a unique representation of a path-dependent, Lagrangian measure of the distance between a given initial and final PDF. We will utilise this property to map out the attractor structure by considering a narrow initial PDF at a different peak position  $y_0$  and by measuring  $\mathcal{L}_\infty$  against  $y_0$ . We are particularly interested in how the behaviour of  $\mathcal{L}_\infty$  against  $y_0$  depends on whether a system has a stable equilibrium point or is chaotic.

#### 3.1. Linear vs. Cubic Forces

We first consider the case where a system has a stable equilibrium point when there is no stochastic noise and investigate how  $\mathcal{L}_\infty$  is affected by different deterministic forces [15,16]. We consider the following Langevin equation [22] for a variable  $x$ :

$$\frac{dx}{dt} = F(x) + \zeta. \tag{11}$$

Here,  $\zeta$  is a short (delta) correlated stochastic noise with the strength  $D$  as

$$\langle \zeta(t)\zeta(t') \rangle = 2D\delta(t - t'), \tag{12}$$

where the angular brackets denote the average over  $\zeta$  and  $\langle \zeta \rangle = 0$ . We consider two types of  $F$ , which both have a stable equilibrium point  $x = 0$ ; the first one is the linear force  $F = -\gamma x$  ( $\gamma > 0$  is the frictional constant) which is the familiar Ornstein–Uhlenbeck (O-U) process, a popular model for a noisy relaxation system (e.g., [23]). The second is the cubic force  $F = -\mu x^3$  where  $\mu$  represents the frictional constant. Note that, in these models, the dimensions of  $\gamma$  ( $s^{-1}$ ) and  $\mu$  ( $s^{-1}m^{-2}$ ) are different.

Equivalent to the Langevin equation governed by Equations (11) and (12) is the Fokker–Planck equation [22]

$$\frac{\partial}{\partial t} p(x, t) = \frac{\partial}{\partial x} \left\{ -F(x) + D \frac{\partial}{\partial x} \right\} p(x, t). \tag{13}$$

As an initial PDF, we consider a Gaussian PDF

$$p(x_0, 0) = \sqrt{\frac{\beta_0}{\pi}} e^{-\beta_0(x_0 - y_0)^2}. \tag{14}$$

Then, for the O-U process, the PDF remains Gaussian for all time with the following form [15,16]:

$$p(x, t) = \sqrt{\frac{\beta(t)}{\pi}} e^{-\beta(t)(x-\langle x \rangle)^2}. \tag{15}$$

In Equations (14) and (15),  $\langle x \rangle = y_0 e^{-\gamma t}$  is the mean position and  $y_0$  is its initial value;  $\beta_0$  is the inverse temperature at  $t = 0$ , which is related to the variance at  $t = 0$  as  $\langle (x_0 - y_0)^2 \rangle = \frac{1}{2\beta_0} = \frac{D_0}{\gamma}$ . The fluctuations level (variance) changes with time, with time-dependent  $\beta(t)$  given by

$$\langle (x - \langle x \rangle)^2 \rangle = \frac{1}{2\beta(t)} = \frac{D(1 - e^{-2\gamma t})}{\gamma} + \frac{e^{-2\gamma t}}{2\beta_0}. \tag{16}$$

Note that, when  $D = D_0$ ,  $\beta(t) = \beta_0 = \frac{\gamma}{2D}$  for all  $t$ , PDF maintains the same width for all  $t$ .

For this Gaussian process,  $\beta$  and  $\langle x \rangle$  constitute a parameter space on which the distance is defined with the Fisher metric tensor [3]  $g_{ij}$  ( $i, j = 1, 2$ ) as [16]

$$g_{ij} = \int dx \frac{1}{p(x,t)} \frac{\partial p}{\partial z^i} \frac{\partial p}{\partial z^j} = \begin{pmatrix} \frac{1}{2\beta^2} & 0 \\ 0 & 2\beta \end{pmatrix}, \tag{17}$$

where  $i, j = 1, 2$ ,  $z^1 = \beta$ ,  $z^2 = \langle x \rangle$ . This enables us to recast  $\frac{1}{\tau^2}$  in Equation (1) in terms of  $g_{ij}$  as

$$\frac{1}{\tau^2} = \frac{1}{2\beta^2} \left( \frac{d\beta}{dt} \right)^2 + 2\beta \left( \frac{d\langle x \rangle}{dt} \right)^2 = g_{ij} \frac{dz^i}{dt} \frac{dz^j}{dt}. \tag{18}$$

The derivation of the first relation in Equation (18) is provided in Appendix A (see Equation (A2)). Using Equations (2) and (18), we can calculate  $\mathcal{L}$  analytically for this O-U process (see also Appendix A).

In comparison, theoretical analysis can be done only in limiting cases such as small and large times for the cubic process [17,24]. In particular, the stationary PDF for large time is readily obtained as

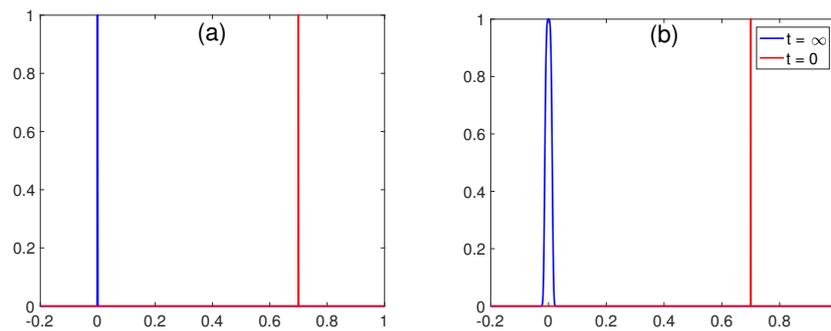
$$p(x) = \frac{2\beta_c^{\frac{1}{4}}}{\Gamma\left(\frac{1}{4}\right)} e^{-\beta_c x^4}, \tag{19}$$

where  $\beta_c = \frac{\mu}{4D}$ . For the exact calculation of  $\mathcal{L}(t)$ , Equation (13) is to be solved numerically.

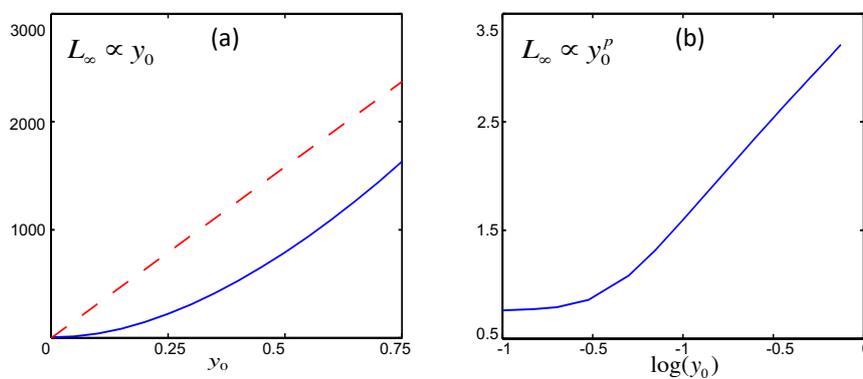
To summarise, due to the restoring forcing  $F$ , the equilibrium is given by a PDF around  $x = 0$ , Gaussian for linear force and quartic exponential for cubic force. If we were to pick any point in  $x$ , say  $y_0$ , we are curious about how close  $y_0$  is to the equilibrium and how  $F(x)$  affects it. To determine this, we make a narrow PDF around  $x = y_0$  (see Figure 1) at  $t = 0$  and measure  $\mathcal{L}_\infty$ . The question is how this  $\mathcal{L}_\infty$  depends on  $y_0$ . We repeat the same procedure for the cubic process, as shown in Figure 1, and examine how  $\mathcal{L}_\infty$  depends on  $y_0$ .

$\mathcal{L}_\infty$  as a function of  $y_0$  is shown for both linear (in red dotted line) and cubic (in blue solid line) processes in Figure 2. In the linear case we can see a clear linear relation between  $y_0$  and  $\mathcal{L}_\infty$ , meaning that the information length preserves the linearity of the system. This linear relationship holds for all  $D$  and  $D_0$ . In particular, when  $D = D_0$ , we can show that  $\mathcal{L}_\infty = \frac{1}{\sqrt{D/\gamma}} y_0$  by taking the limit of  $t \rightarrow \infty$  ( $y \rightarrow 0$ ) in Equation (A10).

In contrast, for the cubic process, the relation is not linear, and the log-log plot on the right in Figure 2 shows a power-law dependence with the power-law index  $p$ . This power-law index  $p$  varies between 1.52 and 1.91 and depends on the width ( $\propto D_0^{1/2}$ ) of initial PDF and stochastic forcing amplitude  $D$ , as shown in [16]. This indicates that nonlinear force breaks the linear scaling of geometric structure and changes it to power-law scalings. In either cases here,  $\mathcal{L}_\infty$  has a smooth variation with  $y_0$  with its minimum value at  $y_0 = 0$  since the equilibrium point 0 is stable. This will be compared with the behaviour in chaotic systems in Section 3.2.



**Figure 1.** Initial (red) and final (blue) Probability Density Functions (PDFs) for the O-U process in (a) and the cubic process in (b).



**Figure 2.** (a):  $\mathcal{L}_\infty$  against  $\langle x(t = 0) \rangle = y_0$  for the linear process in red dashed line and for the cubic process in blue solid line; (b):  $\mathcal{L}_\infty$  against  $\langle x(t = 0) \rangle = y_0$  for the cubic process on log-log scale (data from [17]).

### 3.2. Chaotic Attractor

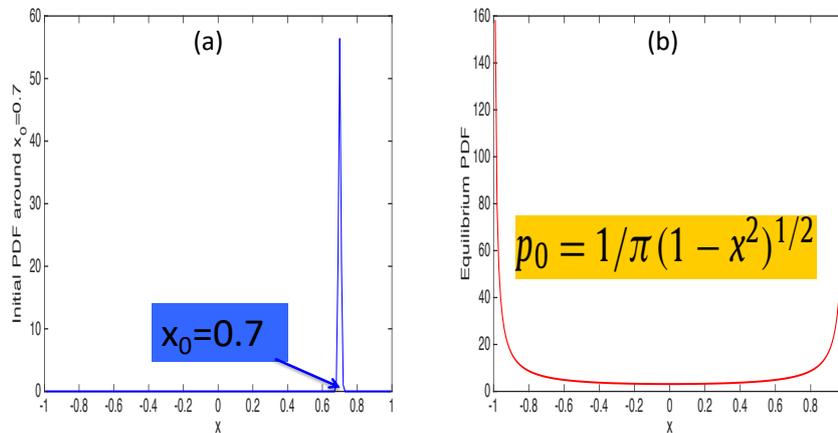
Section 3.1 demonstrates that the minimum value of  $\mathcal{L}_\infty$  occurs at a stable equilibrium point [15,16]. We now show that in contrast, in the case of a chaotic attractor, the minimum value of  $\mathcal{L}_\infty$  occurs at an unstable point [13]. To this end, we consider a chaotic attractor using a logistic map [13]. The latter is simply given by a rule as to how to update the value  $x$  at  $t + 1$  from its previous value at  $t$  as follows [25]

$$x_{t+1} = 1 - ax_t^2, \tag{20}$$

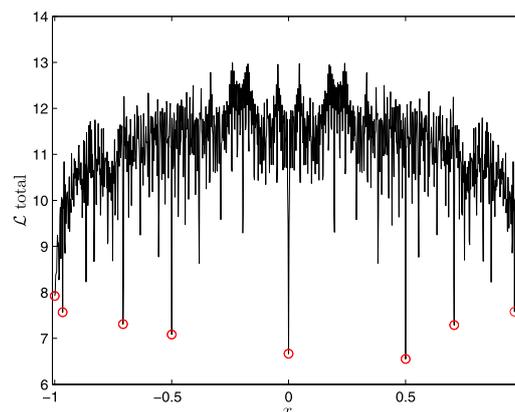
where  $x \in [-1, 1]$  and  $a$  is a parameter, which controls the stability of the system.

As we are interested in a chaotic attractor, we chose the value  $a = 2$  so that any initial value  $x_0$  evolves to a chaotic attractor given by an invariant density (shown in the right panel of Figure 3). A key question is then whether all values of  $x_0$  are similar as they all evolve to the same invariant density in the long time limit. To address how close a particular point  $x_0$  is to equilibrium, we (i) consider a narrow initial PDF around  $x_0$  at  $t = 0$ , (ii) evolve it until it reaches the equilibrium distribution, (iii) measure the  $\mathcal{L}_\infty$  between initial and final PDF, and (iv) repeat steps (i)–(iii) for many different values  $x_0$ . For example, for  $x_0 = 0.7$ , the initial PDF is shown on the left and final PDF on the right in Figure 3. We show  $\mathcal{L}_\infty$  against  $x_0$  in Figure 4. A striking feature of Figure 4 is an abrupt change in  $\mathcal{L}_\infty$  for a small change in  $x_0$ . This means that the distance between  $x_0$  and the final chaotic attractor depends sensitively on  $x_0$ . This sensitive dependence of  $\mathcal{L}_\infty$  on  $x(t = 0)$  means that a small change in the initial condition  $x_0$  causes a large difference in a path that a system evolves through and thus  $\mathcal{L}_\infty$ . This is a good illustration of a chaotic equilibrium and is quite similar to the sensitive dependence of the Lyapunov exponent on the initial condition [25]. That is, our  $\mathcal{L}_\infty$  provides a new methodology to

test chaos. Another interesting feature of Figure 4 are several points with small values of  $\mathcal{L}_\infty$ , shown by red circles. In particular,  $x_0 = 0.5$  has the smallest value of  $\mathcal{L}_\infty$ , indicating that the unstable point is closest to the chaotic attractor. That is, an unstable point is most similar to the chaotic attractor and thus minimises  $\mathcal{L}_\infty$ .



**Figure 3.** (a): an initial narrow PDF at the peak  $x_0 = 0.7$ ; (b): the invariant density of a logistic map.



**Figure 4.**  $\mathcal{L}_\infty$  against the peak position  $x = x_0$  of an initial PDF in the chaotic regime of a logistic map (Reprinted from *Physics Letters A*, 379, S.B. Nicholson & E. Kim, Investigation of the statistical distance to reach stationary distributions, 83-88, Copyright (2015), with permission from Elsevier).

#### 4. Music: Can We See the Music?

Our methodology is not system-specific and applicable to any stochastic processes. In particular, given any time-dependent PDFs that are computed from a theory, simulations or from data, we can compute  $\mathcal{L}(t)$  to understand information change. As an example, we apply our theory to music data and discuss information change associated with different pieces of classical music. In particular, we are interested in understanding differences among famous classical music in view of information change. To gain an insight, we used the MIDI file [26], computed time-dependent PDFs and the information length as a function of time [14].

Specifically, the midi file stores a music by the MIDI number according to 12 different music notes (C, C#, D, D#, E, F, F#, G, G#, A, A#, B) and 11 different octaves, with the typical time  $\Delta t$  between the two adjacent notes of order  $\Delta t \sim 10^{-3}$  s. In order to construct a PDF, we specify 129 statistically different states according to the MIDI number and one extra rest state (see Table 1 in [14]) and calculate an instantaneous PDF (see Figure S1 in [14]) from an orchestra music by measuring the frequency (the total number of times) that a particular state is played by all instruments at a given time. Thus,

the time-dependent PDFs are defined in discrete time steps with  $\Delta t \sim 10^{-3}$ , and the discrete version of  $\mathcal{L}$  (Equation (7) in [14]) is used in numerical computation. Figure 5 shows  $\mathcal{L}(t)$  against time for Vivaldi’s Summer, Mozart, Tchaikovsky’s 1812 Overture, and Beethoven’s Ninth Symphony 2nd movement. We observe the difference among different composers, in particular, more classical, more subtle in information change. We then look at the rate of information change against time for different music by calculating the gradient of  $\mathcal{L}$  ( $\frac{d\mathcal{L}}{dt} = 1/\tau$ ) in Figure 6, which also manifests the most subtle change in information length for Vivaldi and Mozart.

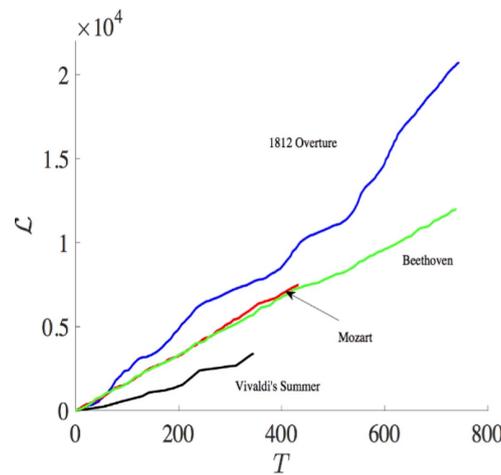


Figure 5.  $\mathcal{L}(t)$  against time  $T$  for different composers (from [14]).

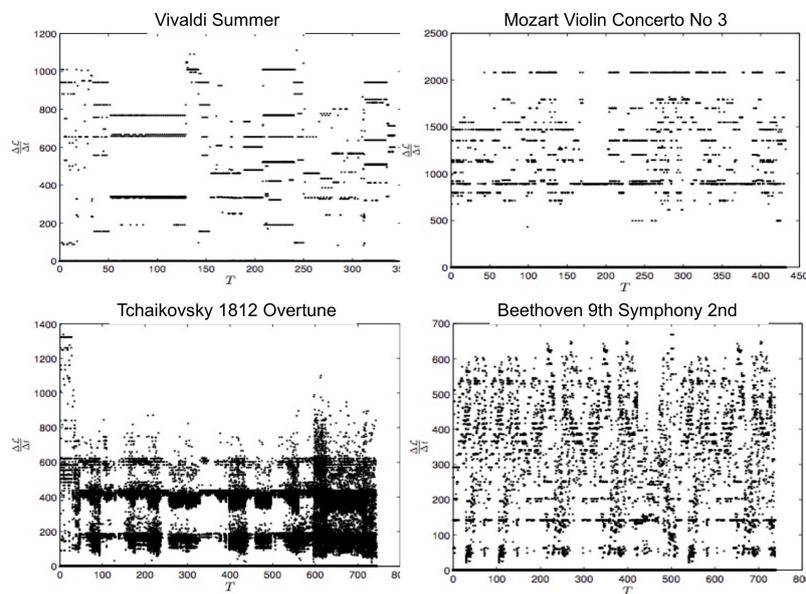


Figure 6.  $\frac{1}{\tau} = \frac{d\mathcal{L}}{dt}$  for different composers shown in Figure 5 (from [14]).

### 5. Quantum Systems

Finally, we examine quantum effects on information length [21]. In Quantum Mechanics (QM), the uncertainty relation  $\Delta x \Delta P \geq \frac{\hbar}{2}$  between position  $x$  and momentum  $P$  gives us an effect quite similar to a stochastic noise. We note here that we are using  $P$  to denote the momentum to distinguish it from a PDF ( $p(x, t)$ ). For instance, the trajectory of a particle in the  $x - P$  phase space is random and not smooth. Furthermore, the phase volume  $h$  plays the role of resolution in the phase space, one unit of information given by the phase volume  $h$ . Thus, the total number of states is given by the total

phase volume divided by  $h$ . This observation points out a potentially different role of the width of PDF in QM in comparison with the classical system since a wider PDF in QM occupies a larger region of  $x$  in the phase space, with the possibility of increasing the information.

To investigate this, for simplicity, we consider a particle of mass  $m$  under a constant force  $F$  and assume an initial Gaussian wave function around  $x' = 0$  [21]

$$\psi(x', 0) = \left(\frac{2\beta_0}{\pi}\right)^{\frac{1}{4}} e^{-\beta_0 x'^2 + ik_0 x'}, \tag{21}$$

where  $k_0 = P_0/\hbar$  is the wave number at  $t = 0$ ,  $D_x = (2\beta_0)^{-1/2}$  is the width of the initial wave function, and  $P_0$  is the initial momentum. A time-dependent PDF  $p(x, t)$  is then found as (e.g., see [21,27]):

$$p(x, t) = |\psi(x, t)|^2 = \sqrt{\frac{\beta(t)}{\pi}} e^{-\beta(t)(x - \langle x \rangle)^2}. \tag{22}$$

Here,

$$\beta(t) = \frac{2\beta_0 m^2}{m^2 + (2\hbar\beta_0 t)^2}, \quad \langle x \rangle = \frac{\hbar k_0 t}{m} + \frac{Ft^2}{2m}. \tag{23}$$

Equation (22) clearly shows that the PDF is Gaussian, with the mean  $\langle x \rangle = \frac{\hbar k_0 t}{m} + \frac{Ft^2}{2m}$  and the variance

$$\text{Var}(t) = \langle (x - \langle x \rangle)^2 \rangle = \frac{1}{4\beta} = \frac{1}{4\beta_0} + \frac{\beta_0 \hbar^2 t^2}{m^2} = \text{Var}(0) + \frac{\hbar^2 t^2}{4\text{Var}(0)m^2}. \tag{24}$$

In Equation (24),  $\text{Var}(0) = \langle (x(0) - \langle x(0) \rangle)^2 \rangle = \frac{1}{4\beta_0} = \frac{D_x^2}{2}$  is the initial variance. We note that the last term in Equation (24) increases quadratically with time  $t$  due to the quantum effect, the width of wave function becoming larger over time. Obviously, this effect vanishes as  $\hbar \rightarrow 0$ .

Since the PDF in Equation (22) is Gaussian, we can use Equation (18) to find (e.g., see [16])

$$\frac{1}{\tau^2} = 2t^2 \frac{1}{(T^2 + t^2)^2} + 2\beta_0 \frac{T^2}{T^2 + t^2} v_0^2 \left[1 + \frac{Ft}{\hbar k_0}\right]^2, \tag{25}$$

where  $T = \frac{m}{2\hbar\beta_0}$ , the time scale of the broadening of the initial wave function [21]. It is interesting to note that when there is no external constant force  $F$ , the two terms in Equation (25) decrease for large time  $t$ , making  $\tau$  large. The situation changes dramatically in the presence of  $F$  in Equation (25) as the second term approaches a constant value for large time. The region with the same value of  $\tau$  signifies that the rate of change in information is constant in time, and was argued to be an optimal path to minimise the irreversible dissipation (e.g., [16]). Physically, this geodesic arises when when the broadening of a PDF is compensated by momentum  $Ft$  which increases with time. Mathematically, the limit  $t \rightarrow \infty$  reduces Equation (25) and thus  $\mathcal{L}$  to

$$\frac{1}{\tau} \sim \frac{FD_x}{\hbar}, \quad \mathcal{L} \sim \frac{(Ft)D_x}{\hbar}. \tag{26}$$

Since  $Ft = P$  and  $D_x = (2\beta_0)^{-1/2}$  is the width of the wave function at  $t = 0$ ,  $FtD_x$  in Equation (26) represents the volume in the  $P - x$  phase space spanned by this wave function. This reflects the information changes associated with the coverage of a phase volume  $\hbar$ . Interestingly, similar results are also obtained in the momentum representation where  $\mathcal{L}$  is computed from the PDF  $p(P, t)$  in the momentum space:

$$p(P, t) = \sqrt{\frac{\lambda}{\pi}} e^{-\lambda(p - (mv_0 + Ft))^2}, \quad \frac{1}{\tau^2} = 2\lambda F^2, \quad \mathcal{L} = \sqrt{2\lambda} Ft, \tag{27}$$

where  $\lambda = \frac{1}{2\hbar^2\beta_0}$ . In Equation (27),  $\tau$  is obviously constant, and  $\mathcal{L}$  linearly increases with time  $t$ . We can see even a strong similarity between Equation (27) and Equation (26) as  $t \rightarrow \infty$  once using  $\mathcal{L} \propto \sqrt{2\lambda Ft} \sim (Ft)D_x/\hbar$ . In view of the complementary relation between position and momentum in quantum systems, the similar result for  $\mathcal{L}$  in momentum and position space highlights the robustness of the geodesic.

### 6. Conclusions

We investigated information geometry associated with stochastic processes in classical and quantum systems. Specifically, we introduced  $\tau(t)$  as a dynamical time scale quantifying information change and calculated  $\mathcal{L}(t)$  by measuring the total clock time  $t$  by  $\tau$ . As a unique Lagrangian measure of the information change,  $\mathcal{L}_\infty$  was demonstrated to be a novel diagnostic for mapping out an attractor structure. In particular,  $\mathcal{L}_\infty$  was shown to capture the effect of different deterministic forces through the scaling of  $\mathcal{L}_\infty$  again the peak position of a narrow initial PDF. For a stable equilibrium, the minimum value of  $\mathcal{L}_\infty$  occurs at the equilibrium point. In comparison, in the case of a chaotic attractor,  $\mathcal{L}_\infty$  exhibits a sensitive dependence on initial conditions like a Lyapunov exponent. We then showed the application of our method to characterize the information change associated with classical music (e.g., see [14]). Finally, we elucidated the effect of the width of a PDF on information length in quantum systems. Extension of this work to impure (mixed-state) quantum systems and investigation of Riemannian geometry on the space of density operators would be of particular interest for future work.

**Funding:** This research received no external funding.

**Acknowledgments:** This paper is a review and summary of work carried out over several years with numerous co-authors. Among these, I am particularly grateful to Rainer Hollerbach, Schuyler Nicholson and James Heseltine for their contributions and valuable discussions regarding different aspects of information length.

**Conflicts of Interest:** The author declares no conflict of interest.

### Appendix A. $\mathcal{L}$ for the O-U Process

To make this paper self-contained, we provide here the main steps for the derivation of  $\mathcal{L}$  for the O-U process [15,16]. We use  $y = \langle x \rangle = y_0 e^{-\gamma t}$  in  $p(x, t)$  in Equation (12) and differentiate it to find

$$\frac{\partial p}{\partial t} = \left[ \dot{\beta} \left( \frac{1}{2\beta} - (x - y)^2 \right) + 2\beta(x - y)y \right] p. \tag{A1}$$

Equations (A1) and (1) and using the properties of a Gaussian PDF [ $\langle (x - y)^2 \rangle = \frac{1}{2\beta}$ ,  $\langle (x - y)^4 \rangle = 3\langle (x - y)^2 \rangle^2$ ] lead to

$$\frac{1}{\tau^2} = \frac{1}{2\beta(t)^2} \left( \frac{d\beta}{dt} \right)^2 + 2\beta \left( \frac{dy}{dt} \right)^2. \tag{A2}$$

We express  $\beta$  in Equation (16) in terms of  $T = 2\beta_0 D(e^{2\gamma t} - 1) + \gamma$  as  $\beta = \frac{\gamma\beta_0 e^{2\gamma t}}{T}$ . Differentiating this and using  $r = 2\beta_0 D - \gamma$  then give

$$\frac{\dot{\beta}^2}{2\beta^2} = 2\gamma^2 r^2 \frac{1}{T^2}. \tag{A3}$$

Similarly, using  $\frac{dy}{dt} = -\gamma y_0 e^{-\gamma t}$ ,  $T = 2\beta_0 D(e^{2\gamma t} - 1) + \gamma$  and  $q = \beta_0 \gamma y_0^2$ , we obtain

$$2\beta y^2 = 2q\gamma^2 \frac{1}{T}. \tag{A4}$$

Using these results, Equations (A3) and (A4) in (A2) gives us

$$\frac{1}{\tau^2} = \frac{1}{2\beta^2} \left( \frac{d\beta}{dt} \right)^2 + 2\beta \left( \frac{dy}{dt} \right)^2 = \frac{2\gamma^2}{T^2} (r^2 + qT). \quad (\text{A5})$$

Again, in Equation (A5),  $q = \beta_0 \gamma y_0^2$ ,  $r = 2\beta_0 D - \gamma$ , and  $T = 2\beta_0 D (e^{2\gamma t} - 1) + \gamma$  [15–17]. It is worth noting that  $q$  and  $r$ , respectively, arise from the difference in mean position at  $t = 0$  and  $t \rightarrow \infty$  (i.e.,  $y_0 \neq y(t \rightarrow \infty)$ ) and in PDF width at  $t = 0$  and  $t \rightarrow \infty$  (i.e.,  $D_0 \neq D$ ). Thus, the first and second terms in Equation (A5) represent the information change due to the change in PDF width and the movement of the PDF, respectively. Using  $D_0 = \frac{\gamma}{2\beta_0}$ , we express  $r$ ,  $q$  and  $T$  in Equation (A5) as  $q = \frac{\gamma^2 y_0^2}{2D_0}$ ,  $r = \gamma \left( \frac{D}{D_0} - 1 \right)$ ,  $T = \gamma \left[ \frac{D}{D_0} (e^{2\gamma t} - 1) + 1 \right]$ . Equations (A5) and (2) then give us

$$\mathcal{L} = \frac{1}{\sqrt{2}} \int_{T_i}^{T_f} \left\{ \frac{1}{T} \frac{1}{T+r} \sqrt{r^2 + qT} \right\} dT, \quad (\text{A6})$$

where  $T_i = T(t = 0)$  and  $T_f = T(t)$ . To compute Equation (A6) for  $r \neq 0$ , we use  $Y = \sqrt{r^2 + qT}$  and integrate

$$\mathcal{L} = \frac{1}{\sqrt{2}} \left[ \ln \left( \frac{Y-r}{Y+r} \right) \right]_{Y_i}^{Y_f} + \frac{\sqrt{2}}{r} H, \quad H = \int_{Y_i}^{Y_f} \frac{qr - r^2}{Y^2 + qr - r^2} dY, \quad (\text{A7})$$

where  $Y_i = Y(t = 0)$  and  $Y_f = Y(t)$ . To calculate  $H$  in Equation (A7), we need to consider the two cases where  $q \geq r$  or  $q < r$ . First, when  $q \geq r$ , we use the change of the variable  $Y = \sqrt{qr - r^2} \tan \theta$  to find

$$H = \sqrt{qr - r^2} \left[ \tan^{-1} \left( \frac{Y}{\sqrt{qr - r^2}} \right) \right]_{Y_i}^{Y_f}. \quad (\text{A8})$$

When  $q < r$ , we let  $Y = \sqrt{r^2 - qr} \sec \theta$  and find

$$H = -\frac{\sqrt{r^2 - qr}}{2} \left[ \ln \left( \frac{Y - \sqrt{r^2 - qr}}{Y + \sqrt{r^2 - qr}} \right) \right]_{Y_i}^{Y_f}. \quad (\text{A9})$$

When  $D = D_0$  ( $r = 0$ ),  $\beta(t) = \beta_0$  for all  $t$ . Thus, Equation (2) can easily be calculated directly from Equation (A5) with the result

$$\mathcal{L} = \frac{1}{\sqrt{2}} \int_{T_i}^{T_f} \frac{\sqrt{q}}{T^{\frac{3}{2}}} dT = -\sqrt{2q} \left[ \frac{1}{\sqrt{T}} \right]_{T_i}^{T_f} = \frac{1}{\sqrt{D/\gamma}} [y_0 - y], \quad (\text{A10})$$

where again  $y = \langle x \rangle = y_0 e^{-\gamma t}$ .

## References

- Gibbs, A.L.; Su, F.E. On choosing and bounding probability metrics. *Int. Stat. Rev.* **2002**, *70*, 419–435. [CrossRef]
- Wootters, W.K. Statistical distance and Hilbert space. *Phys. Rev. D* **1981**, *23*, 357. [CrossRef]
- Frieden, B.R. *Science from Fisher Information*; Cambridge University Press: Cambridge, UK, 2000.
- Braunstein, S.L.; Caves, C.M. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.* **1994**, *72*, 3439. [CrossRef] [PubMed]
- Feng, E.H.; Crooks, G.E. Far-from-equilibrium measurements of thermodynamic length. *Phys. Rev. E* **2009**, *79*, 012104. [CrossRef] [PubMed]
- Ruppeiner, G. Thermodynamics: A Riemannian geometric model. *Phys. Rev. A* **1979**, *20*, 1608–1613. [CrossRef]
- Schlögl, F. Thermodynamic metric and stochastic measures. *Z. Phys. B Condens. Matter* **1985**, *59*, 449–454. [CrossRef]

8. Nulton, J.; Salamon, P.; Andresen, B.; Anmin, Q. Quasistatic processes as step equilibrations. *J. Chem. Phys.* **1985**, *83*, 334–338. [[CrossRef](#)]
9. Sivak, D.A.; Crooks, G.E. Thermodynamic metrics and optimal paths. *Phys. Rev. Lett.* **2012**, *8*, 190602. [[CrossRef](#)] [[PubMed](#)]
10. Plastino, A.R.; Casas, M.; Plastino, A. Fisher's information, Kullback's measure, and H-theorems. *Phys. Lett. A* **1998**, *246*, 498–504. [[CrossRef](#)]
11. Poletini, M.; Esposito, M. Nonconvexity of the relative entropy for Markov dynamics: A Fisher information approach. *Phys. Rev. E* **2013**, *88*, 012112. [[CrossRef](#)] [[PubMed](#)]
12. Naudts, J. Quantum statistical manifolds. *Entropy* **2018**, *20*, 472. [[CrossRef](#)]
13. Nicholson, S.B.; Kim, E. Investigation of the statistical distance to reach stationary distributions. *Phys. Lett. A* **2015**, *379*, 83–88. [[CrossRef](#)]
14. Nicholson, S.B.; Kim, E. Structures in sound: Analysis of classical music using the information length. *Entropy* **2016**, *18*, 258. [[CrossRef](#)]
15. Heseltine, J.; Kim, E. Novel mapping in non-equilibrium stochastic processes. *J. Phys. A Math. Theor.* **2016**, *49*, 175002. [[CrossRef](#)]
16. 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](#)]
17. Kim, E.; Hollerbach, R. Signature of nonlinear damping in geometric structure of a nonequilibrium process. *Phys. Rev. E* **2017**, *95*, 022137. [[CrossRef](#)] [[PubMed](#)]
18. Hollerbach, R.; Kim, E. Information geometry of non-equilibrium processes in a bistable system with a cubic damping. *Entropy* **2017**, *19*, 268. [[CrossRef](#)]
19. Kim, E.; Tenkès, L.M.; Hollerbach, R.; Radulescu, O. Far-from-equilibrium time evolution between two gamma distributions. *Entropy* **2017**, *19*, 511. [[CrossRef](#)]
20. Tenkès, L.M.; Hollerbach, R.; Kim, E. Time-dependent probability density functions and information geometry in stochastic logistic and Gompertz models. *J. Stat. Mech. Theory Exp.* **2017**, *2017*, 123201. [[CrossRef](#)]
21. Kim, E.; Lewis, P. Information length in quantum system. *J. Stat. Mech. Theory Exp.* **2018**, *2018*, 043106. [[CrossRef](#)]
22. Risken, H. *The Fokker–Planck Equation: Methods of Solutions and Applications*; Springer: Berlin, Germany, 2013.
23. Klebaner, F. *Introduction to Stochastic Calculus with Applications*; Imperial College Press: London, UK, 2012.
24. Kim, E.; Hollerbach, R. Time-dependent probability density function in cubic stochastic processes. *Phys. Rev. E* **2016**, *94*, 052118. [[CrossRef](#)] [[PubMed](#)]
25. Ott, E. *Chaos in Dynamical Systems*; Cambridge University Press: Cambridge, UK, 2002.
26. Kern Scores. Available online: <http://kernscores.stanford.edu/> (accessed on 12 July 2016).
27. Andrews, M. Quantum mechanics with uniform forces. *Am. J. Phys.* **2018**, *78*, 1361–1364. [[CrossRef](#)]



© 2018 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).