

Article

Information Geometry in Classical and Quantum Systems

Eun-jin Kim 

School of Mathematics and Statistics, University of Sheffield, Sheffield, S3 7RH, UK

* Correspondence: e.kim@sheffield.ac.uk; Tel.: +44-114-222-3876

Abstract: A probabilistic description is essential for understanding the dynamics of stochastic systems far from equilibrium. To compare different Probability Density Functions (PDFs), it is extremely useful to quantify the difference among different PDFs by assigning an appropriate metric to probability such that the distance increases with the difference between the two PDFs. This metric structure then provides a key link between stochastic processes and geometry. For a non-equilibrium process, we define an infinitesimal distance at any time by comparing two PDFs at times infinitesimally apart and sum these distances in time. The total distance along the trajectory of the system quantifies the total number of different states that the system undergoes in time and is called the information length. By using this concept, we investigate classical and quantum systems and demonstrate the utility of the information length as a unique Lagrangian diagnostic to quantify the information change as a system continuously evolves in time and to map out attractor structure. We further elucidate quantum effects (uncertainty relation) and the dual role of the width of PDF in quantum systems.

Keywords: stochastic processes; Langevin equation; Fokker-Planck equation; information length; Fisher information; relaxation; chaos; attractor

1. Introduction

Stochastic processes are ubiquitous in nature and laboratories, and play a major role across traditional disciplinary boundaries. For a proper understanding and description of such processes, it is essential to utilize a probabilistic methodology such as a Probability Density Function (PDF). Furthermore, in order to compare different systems, it is invaluable to employ a measure which is independent of any particular realization of a system. This can very conveniently be achieved by using a geometric measure [1] in a statistical space by assigning a metric between PDFs such that the distance increases with the difference between the two PDFs. This metric structure then provides a key link between stochastic processes and information geometry. One important example is a dimensionless, statistical distance [2] based on the Fisher (or Fisher-Rao) metric [3], which represents the number of distinguishable states between two PDFs. For example, for a Gaussian distribution, statistically distinguishable states are determined by the standard deviation, which increases with the level of fluctuations; two PDFs which have the same standard deviation and differ in peak positions by less than one standard deviation are statistically indistinguishable. Previous work using this fluctuation-based metric include [4–14].

Compared with a metric defined for any given two PDFs, significantly much less work has been done in the case of a time-dependent PDF in non-equilibrium systems. A continuous change in PDFs in this case necessitates defining a distance at any time by comparing two PDFs at times infinitesimally apart and then integrating these distances over time (see Section 2). The cumulative change in information is mathematically quantified by the information length $\mathcal{L}(t)$, which is uniquely defined as a function of time for a given initial condition. This paper reviews the utility of $\mathcal{L}(t)$ in understanding information change and attractor structure in classical and quantum systems [15–23].

The remainder of the paper is organized as follows. §2 discusses information length and §3 investigates attractor structure. §4 and §5 present the analysis of classical music and quantum systems, respectively. Conclusions are found in §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 characteristic time scale τ over which the information changes can be computed by the correlation time of a time-dependent PDF, say $p(x, t)$, as follows.

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

As defined in Eq. (1), $\tau(t)$ has the dimension of time and serves as a dynamical time unit for information change. $\mathcal{L}(t)$ is the accumulated change in information between the initial and final times, 0 and t respectively, given by

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

Note that $\tau(t)$ in Eq. (1) can vary with time, so we need the integral for \mathcal{L} in Eq. (2). To make an analogy, we can consider an oscillator with a period $\tau = 2$ secs. Then, within the clock time 10 secs, there are 5 oscillations. When the period τ is changing with time, we need an integration of dt/τ over the time interval.

We now recall how $\tau(t)$ and $\mathcal{L}(t)$ in Eqs. (1)-(2) are related to the relative entropy (or Kullback-Leibler divergence) [17,18]. To this end, we consider $p_1 = p(x, t_1)$ and $p_2 = p(x, t_2)$ and the relative entropy $D(p_1, p_2) = \int dx p_2 \ln(p_2/p_1)$. We compute this to the second order in an infinitesimally small $|t_2 - t_1|$ by Taylor expansion. To this end, we calculate

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

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

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

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

By taking the limit where $t_2 \rightarrow t_1 = t$ ($p_2 \rightarrow p_1 = p$) and by using the total probability conservation $\int dx \partial_t p = 0$, Eqs. (3) and (5) above lead to

$$\lim_{t_2 \rightarrow t_1 = t} \frac{\partial}{\partial t_1} D(p_1, p_2) = \lim_{t_2 \rightarrow t_1 = t} \frac{\partial}{\partial t_2} D(p_1, p_2) = \int dx \partial_t p = 0, \quad (7)$$

while Eqs. (4) and (6) give

$$\lim_{t_2 \rightarrow t_1 = t} \frac{\partial^2}{\partial t_1^2} D(p_1, p_2) = \lim_{t_2 \rightarrow t_1 = t} \frac{\partial^2}{\partial t_2^2} D(p_1, p_2) = \int dx \frac{(\partial_t p)^2}{p} = \frac{1}{\tau(t)^2}. \quad (8)$$

Thus, by using Eqs. (7)-(8) and $D(p_1, p_1) = 0$, we obtain the Taylor expansion of $D(p_1, p_2)$ up to $O((dt)^2)$ ($dt = t_2 - t_1$) as

$$D(p_1, p_2) = \frac{1}{2} \left[\int dx \frac{(\partial_{t_1} p(x, t_1))^2}{p(x, t_1)} \right] (dt)^2 + O((dt)^3), \quad (9)$$

where $O((dt)^3)$ is higher order term in dt . We can then define the infinitesimal distance $dl(t_1)$ between t_1 and $t_1 + dt$ by

$$dl(t_1) = \sqrt{D(p_1, p_2)} = \frac{1}{\sqrt{2}} \sqrt{\int dx \frac{(\partial_t p(x, t_1))^2}{p(x, t_1)}} dt + O((dt)^{3/2}). \quad (10)$$

We sum $dl(t_1)$ for $t_1 = 0, dt, \dots, t - dt$ by using Eq. (10) and then take the limit of $dt \rightarrow 0$ as

$$\begin{aligned} l(t) &= \lim_{dt \rightarrow 0} [dl(0) + dl(dt) + dl(2dt) + dl(3dt) + \dots dl(t - dt)] \\ &= \lim_{dt \rightarrow 0} \left[\sqrt{D(p(x, 0), p(x, dt))} + \sqrt{D(p(x, dt), p(x, 2dt))} + \dots \sqrt{D(p(x, t - dt), p(x, t))} \right] \\ &\propto \int_0^t dt_1 \sqrt{\int dx \frac{(\partial_{t_1} p(x, t_1))^2}{p(x, t_1)}} = \mathcal{L}(t), \end{aligned} \quad (11)$$

where $\mathcal{L}(t)$ is the information length. Thus, \mathcal{L} is related to the sum of infinitesimal relative entropy. It is important to note that Eq. (11), that is, \mathcal{L} , depends not only on the initial and final PDFs $p(x, 0)$ and $p(x, t)$, but also on the particular path that a system takes. Thus, in general, $l(t)^2$ in Eq. (11) is not simply proportional to the relative entropy $D(p(x, 0), p(x, t))$, which depends only on $p(x, 0)$ and $p(x, t)$.

To recapitulate, Eq. (2) establishes a Lagrangian distance between the initial and final PDFs according to the total number of different states that a system passes through in time. For example, in equilibrium where $\frac{\partial p}{\partial t} = 0$ and hence $\tau(t_1) \rightarrow \infty$ for all time t_1 ; thus $dt_1/\tau(t_1) = 0$ in Eq. (2), and $\int_0^t dt_1/\tau(t_1) = 0$. This reflects that there is no flow of information in equilibrium. In the opposite limit of a small τ , information changes very quickly. Eq. (1) can be written in terms of Fisher metric [3] as shown in §III.

3. Attractor structure

Since $\mathcal{L}(t)$ represents the accumulated change in information (due to the change in PDF) at time t , $\mathcal{L}(t)$ approaches a constant value \mathcal{L}_∞ when a PDF settles into its final equilibrium PDF. The latter is a unique representation of the total number of statistically different states that a PDF evolves through in reaching a final PDF; the smaller \mathcal{L}_∞ , the smaller number of states that the initial PDF passes through to reach the final equilibrium. Therefore, \mathcal{L}_∞ provides us with a path-dependent, Lagrangian measure of the distance between a given initial and final PDF. This enables us to map out the attractor structure (the proximity of y_0 to an equilibrium) by choosing a narrow initial PDF at a different peak position y_0 and by measuring \mathcal{L}_∞ against y_0 . We are particularly interested in the behaviour of \mathcal{L}_∞ against y_0 depending on whether a system has a stable equilibrium point in the absence of noise or is chaotic.

3.1. Linear vs cubic forces

We first consider the case where a system has a stable equilibrium point in the absence of noise and investigate how \mathcal{L}_∞ is affected by different deterministic forces [17,18]. To this end, we consider

a stochastic variable x driven by a random noise ξ with the strength D , governed by the following Langevin equation for over-damped oscillators [24]:

$$\frac{dx}{dt} = F(x) + \xi. \quad (12)$$

Here, x is a random variable of interest, and ξ is a white noise with a short correlation time with the following property:

$$\langle \xi(t)\xi(t') \rangle = 2D\delta(t-t'). \quad (13)$$

Here, D is the strength of the forcing. We note that the dimension of D is $\text{length}^2/\text{time}$ since the dimensions of ξ and $\delta(t-t')$ are $\text{length}/\text{time}$ and $1/\text{time}$, respectively. $F(x)$ is a deterministic force, which can be interpreted as the gradient of the potential $U(x)$ as $F(x) = -\frac{\partial U(x)}{\partial x}$. We consider two types of F , which both have a stable equilibrium point $x = 0$; the first one is the linear force $F = -\gamma x$ ($U = \frac{\gamma}{2}x^2$) and the second is the cubic force $F = -\mu x^3$ ($U = \frac{\mu}{4}x^4$). Here γ and μ have dimensions of $1/\text{time}$ and $1/(\text{time} \times \text{length}^2)$, respectively. The linear system is the familiar Ornstein-Uhlenbeck (O-U) process, which has been widely used as a model for a noisy relaxation system in many areas of physical science and financial mathematics (e.g. [?]). The Fokker-Planck equation corresponding to Eqs. (12) and (13) is [24]

$$\frac{\partial}{\partial t} p(x, t) = \frac{\partial}{\partial x} \left[-F(x) + D \frac{\partial}{\partial x} \right] p(x, t). \quad (14)$$

As an initial PDF, we consider a Gaussian PDF with the inverse temperature β_0 as

$$p(x_0, 0) = \sqrt{\frac{\beta_0}{\pi}} \exp[-\beta_0(x_0 - y_0)^2]. \quad (15)$$

Then, for the O-U process, we have an exact time-dependent PDF [17,18]:

$$p(x, t) = \sqrt{\frac{\beta(t)}{\pi}} \exp[-\beta(t)(x - \langle x \rangle)^2], \quad (16)$$

where

$$\langle x \rangle = y_0 e^{-\gamma t}, \quad (17)$$

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

Here $\langle x(t) \rangle$ is the mean position of the Gaussian profile, and y_0 is its initial value. Similarly, $\beta(t)$ is the inverse temperature, and β_0 is its initial value. As t tends to infinity, $\langle x \rangle \rightarrow 0$ and $\beta(t) \rightarrow \frac{\gamma}{2D} \equiv \beta_*$. To compare initial and final equilibrium states, it is convenient also to introduce $D_0 = \frac{\gamma}{2\beta_0}$. The variance at $t = 0$ and $t \rightarrow \infty$ is then given by $\langle (x_0 - y_0)^2 \rangle = \frac{1}{2\beta_0} = \frac{D_0}{\gamma}$ and $\langle x^2 \rangle = \frac{1}{2\beta_*} = \frac{D}{\gamma}$, respectively. We note from Eqs. (16)-(18) that when $D = D_0$, $\beta(t) = \beta_0 = \frac{\gamma}{2D}$ for all time. In this case, the Gaussian PDF simply moves from y_0 to 0 without changing its shape. If D is greater (lesser) than D_0 , it also broadens (narrows) as it moves.

For this Gaussian process, β 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 [18]

$$g_{ij} = \int dx \frac{1}{p(x, t)} \frac{\partial p}{\partial z^i} \frac{\partial p}{\partial z^j} = \begin{pmatrix} \frac{1}{2\beta^2} & 0 \\ 0 & 2\beta \end{pmatrix}, \quad (19)$$

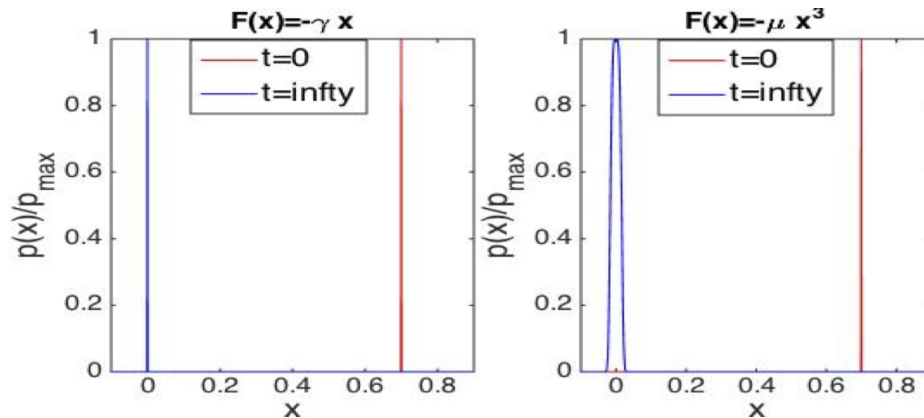


Figure 1. Initial (red) and final (blue) PDFs for the O-U process on the left and the cubic process on the right.

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

$$\frac{1}{\tau^2} = \frac{1}{2\beta^2} \left(\frac{d\beta}{dt} \right)^2 + 2\beta \left(\frac{d\langle x \rangle}{dt} \right)^2 = g_{ij} \frac{dz^i}{dt} \frac{dz^j}{dt}. \quad (20)$$

The derivation of the first relation in Eq. (20) is provided in Appendices A-B. Using Eqs. (2) and (20), we can calculate \mathcal{L} analytically for this O-U process (see also Appendices A-B).

In comparison, theoretical analysis can be done only in limiting cases such as small and large times for the cubic process [19,25]. In particular, the stationary PDF for large time can be shown to have the form

$$p(x) = \frac{2\beta_c^{\frac{1}{4}}}{\Gamma\left(\frac{1}{4}\right)} e^{-\beta_c x^4}, \quad (21)$$

where $\beta_c = \frac{\mu}{4D}$. Eq. (14) must be solved numerically to calculate $\mathcal{L}(t)$.

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 use a range of points around $x = y_0$ to make a narrow PDF (see Fig. 1) and measure \mathcal{L}_∞ . The question is how this \mathcal{L}_∞ depends on y_0 . We repeat the same procedure for the cubic process, as shown in Fig. 1, and examine how \mathcal{L}_∞ depends on y_0 . \mathcal{L}_∞ as a function of y_0 is shown for both linear (in red dotted line) and cubic (in blue solid line) processes in Fig. 2. In the linear case we can see a clear linear relation between y_0 and \mathcal{L}_∞ , meaning that the information length preserves the linearity of the system. This linear relationship holds for all D and D_0 . In contrast, for the cubic process, the relation is not linear, and the log-log plot on the right in Fig. 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 [18]. This demonstrates that nonlinear interaction tends to change geometric structure of a non-equilibrium process from linear to power-law scalings. In either cases here, \mathcal{L}_∞ 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 §3.2.

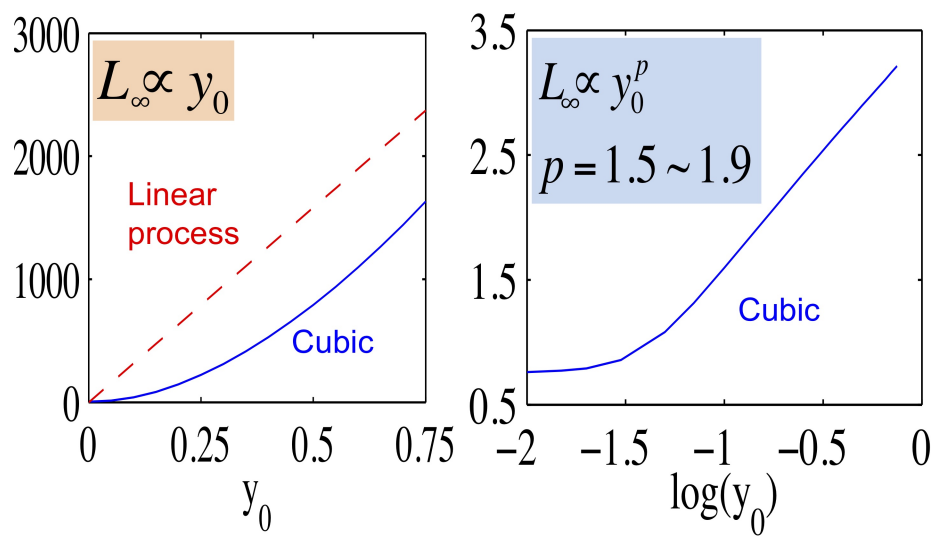


Figure 2. Left: \mathcal{L}_∞ against $\langle x(t=0) \rangle = y_0$ for the linear process in red dashed line and for the cubic process in blue solid line. Right: \mathcal{L}_∞ against $\langle x(t=0) \rangle = y_0$ for the cubic process on log-log scale [19].

3.2. Chaotic attractor

Section 3.1 demonstrates that \mathcal{L}_∞ takes its minimum value at a stable equilibrium point [17,18]. We now show that in contrast, in a chaotic attractor, \mathcal{L}_∞ is minimum for an unstable equilibrium point [15]. To this end, we consider a chaotic attractor using a logistic map [15]. 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 [27]

$$x_{t+1} = 1 - ax_t^2, \quad (22)$$

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

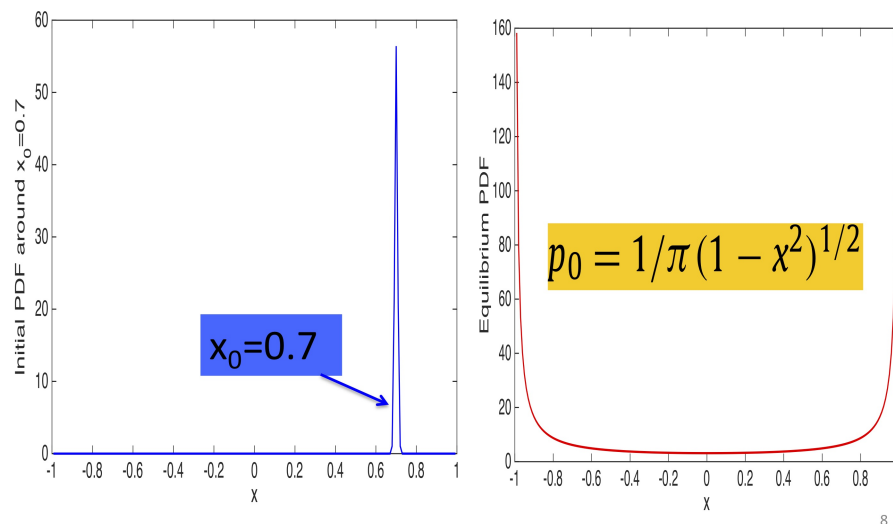


Figure 3. Left: An initial narrow PDF at the peak $x_0 = 0.7$. Right: The invariant density of a logistic map.

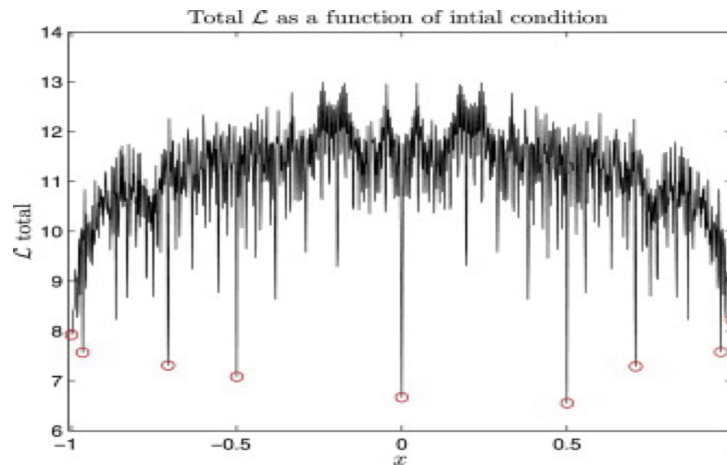


Figure 4. \mathcal{L}_∞ against the peak position $x = x_0$ of an initial PDF in the chaotic regime of a logistic map (from [15]).

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 Fig. 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}_∞ 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 Fig. 3. We show \mathcal{L}_∞ against x_0 in Fig. 4. A striking feature of Fig. 4 is an abrupt change in \mathcal{L}_∞ 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}_∞ 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}_∞ . This is quite similar to the sensitive dependence of the Lyapunov exponent on the initial condition [27]. That is, our \mathcal{L}_∞ provides a new methodology to test chaos. And, it is a good illustration of a chaotic equilibrium. Another interesting feature of Fig. 4 are several points with small values of \mathcal{L}_∞ , shown by red circles. In particular, $x_0 = 0.5$ has the smallest value of \mathcal{L}_∞ , 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}_∞ .

4. Music: can we see the music?

Our methodology can be applied to any system as long as time-dependent PDFs can be computed, e.g. from data. 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 [16]. Fig. 6 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 Fig. 6, which also manifests the most subtle change in information length for Vivaldi and Mozart.

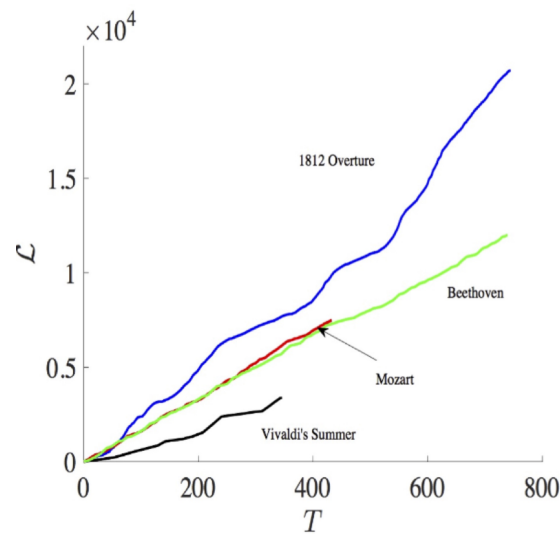


Figure 5. $\mathcal{L}(t)$ against time T for different composers (from [16]).

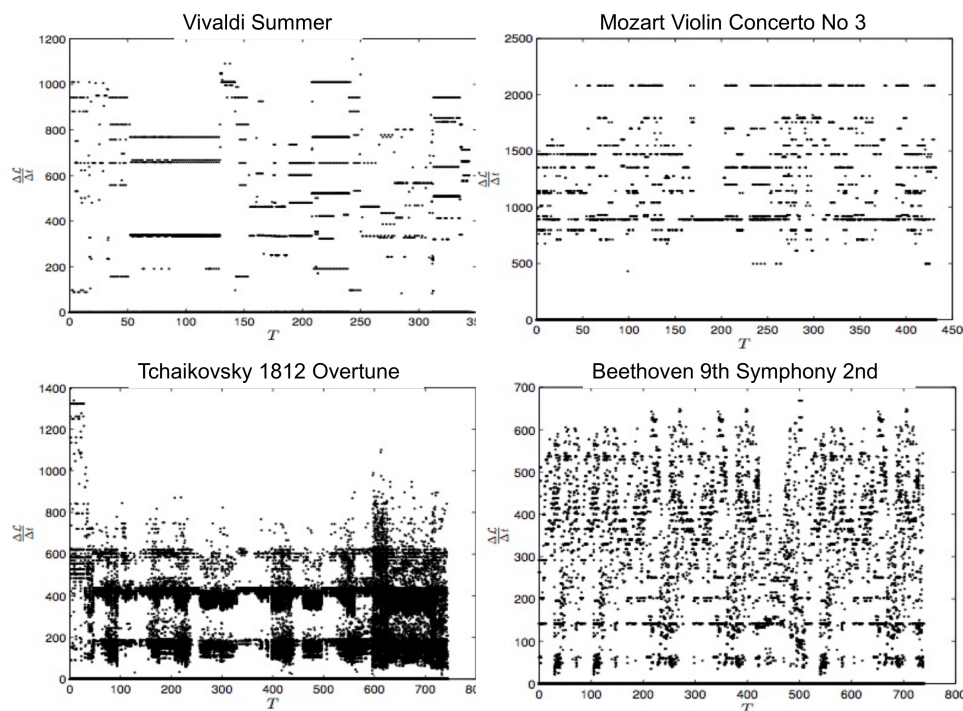


Figure 6. $\frac{1}{\tau} = \frac{d\mathcal{L}}{dt}$ for different composers shown in Fig. 5 (from [16]).

5. Quantum systems

Finally, we examine quantum effects on information length [23]. In Quantum Mechanics (QM), stochasticity arises due to the uncertainty relation $\Delta x \Delta p \geq \frac{\hbar}{2}$ even in the absence of an external noise. In particular, in the semi-classical limit, \hbar serves as a unit of information in the $p - x$ phase space, since each quantum state corresponds to a classical volume \hbar ; the total number of states is the classical volume of phase space divided by \hbar . A wider PDF corresponds to a QM wave function with larger variance, and occupies a larger x region in the phase space; it is thus expected to cause more change in information, opposite to what might be expected in classical systems (e.g. discussed in §III.A). To

elucidate quantum effects, for simplicity, we consider a particle of mass m under a constant force F and assume that it has an initial Gaussian wave function localised around $x' = 0$ [23]

$$\psi(x', 0) = \left[\frac{2\beta_0}{\pi} \right]^{\frac{1}{4}} e^{-\beta_0 x'^2 + ik_0 x'}, \quad (23)$$

where $k_0 = p(t=0)/\hbar$ is the wave number at $t=0$, $d_x = (2\beta_0)^{-1/2}$ is the width of the wave packet, and p is the momentum. A PDF $P(x, t) = |\psi(x, t)|^2$ is then found (e.g. see [23,28]):

$$P(x, t) = \sqrt{\frac{\beta}{\pi}} \exp \left[-\beta (x - \langle x \rangle)^2 \right], \quad (24)$$

where

$$\beta = \frac{2\beta_0 m^2}{m^2 + (2\hbar\beta_0 t)^2}, \quad \langle x \rangle = \frac{\hbar k_0 t}{m} + \frac{Ft^2}{2m}. \quad (25)$$

Here, v_0 is the constant velocity and the angular brackets denote the average. Eq. (24) obviously represents a Gaussian, with the mean $\langle x \rangle = \frac{\hbar k_0 t}{m} + \frac{Ft^2}{2m}$ and the variance $\Delta(t) = \langle (x - \langle x \rangle)^2 \rangle$

$$\Delta(t) = \frac{1}{2\beta} = \frac{1}{4\beta_0} + \frac{\beta_0 \hbar^2 t^2}{m^2} = \Delta(0) + \frac{\hbar^2 t^2}{4\Delta(0)m^2}. \quad (26)$$

The first term in Eq. (26) is due to the variance of the initial PDF, $\Delta(0) = \langle (x(0) - \langle x(0) \rangle)^2 \rangle = \frac{1}{4\beta_0}$. The second term represents the spreading of the wave packet/PDF in time due to quantum effects, which disappears in the classical limit $\hbar \rightarrow 0$ since $\beta = \beta(0) = 2\beta_0$. These quantum effects give rises to a super-diffusion $\propto t^2$, occurring faster than the Brownian motion $\Delta \propto t$, in the limit of a very narrow initial wave packet (as $\beta(0) \rightarrow \infty$).

Since the PDF in Eq. (24) is Gaussian, we can use Eq. (20) to find (e.g. see [18])

$$\frac{1}{\tau^2} = 2t^2 \frac{1}{(T^2 + t^2)^2} + 2\beta_0 \frac{T^2}{T^2 + t^2} v_0^2 \left[1 + \frac{Ft}{\hbar k_0} \right]^2. \quad (27)$$

Here, we defined a characteristic time $T = \frac{m}{2\hbar\beta_0}$. It is interesting to rewrite T using $\Delta x(0)\Delta p(0) \sim \frac{\hbar}{2}$ so that $\frac{1}{4\beta_0} = \Delta(0) = (\Delta x(0))^2 \sim \frac{\hbar^2}{4} \frac{1}{(\Delta p(0))^2} \sim \frac{\Delta x(0)}{2(\Delta k(0))}$:

$$T \sim \frac{m\Delta x(0)}{\hbar(\Delta k(0))} \sim \frac{\Delta x(0)}{\Delta v_0}, \quad (28)$$

where $\Delta v_0 = \Delta p(0)/m$. Thus, T represents the characteristic time scale for the spreading of the initial Gaussian wave packet. Without the constant force F , the right-hand side in Eq. (27) decreases with time and τ increases, taking longer time for information change. However, with a constant force in Eq. (27), it is possible to have a constant value of τ (geodesic) for a sufficiently large t when the increase in the momentum $\propto Ft$ compensates the increase in the width of the PDF $\propto t$. Specifically, as $t \rightarrow \infty$, Eq. (27) is reduced to

$$\frac{1}{\tau} \rightarrow \frac{F}{\hbar\sqrt{2\beta_0}} \sim \frac{Fd_x}{\hbar}, \quad \mathcal{L} \rightarrow \frac{(Ft)d_x}{\hbar}. \quad (29)$$

In Eq. (29), Ft represents the momentum due to the constant force F while $d_x = (2\beta_0)^{-1/2}$ is the width of the initial wave packet. Thus, Ftd_x in Eq. (29) represents the phase volume covered by the motion due to the constant force F in the $p-x$ phase space, demonstrating information changes associated with the coverage of a phase volume \hbar due to the constant force F . 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 (see [29]):

$$P(p, t) = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha(p - (mv_0 + Ft))^2}, \quad \frac{1}{\tau^2} = 2\alpha F^2, \quad \mathcal{L} = \sqrt{2\alpha} Ft, \quad (30)$$

where $\alpha = \frac{1}{2\hbar^2\beta_0}$. Thus, a (non-zero) constant τ (geodesic) is induced by the force F . Furthermore, $\mathcal{L} = \sqrt{2\alpha} Ft \sim (Ft)\Delta(0)/\hbar$, similar to \mathcal{L} in Eq. (29). This is an elegant result showing the robustness of \mathcal{L} in either position or momentum representation for a geodesic solution, despite the complementary relation between position and momentum in quantum systems.

6. Conclusions

We investigated the information geometry of non-equilibrium processes involved 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 elapsed time t in units of τ . As a unique representation of the total number of statistically different states that a PDF evolves through in reaching a final PDF, \mathcal{L}_∞ was demonstrated to be a novel diagnostic for mapping out an attractor structure. In particular, \mathcal{L}_∞ preserves a linear geometry of a linear process while manifesting nonlinear geometry in a cubic (nonlinear) process; it takes its minimum value at the stable equilibrium point. In the case of a chaotic attractor, \mathcal{L}_∞ exhibits a sensitive dependence on initial conditions like a Lyapunov exponent. Thus, \mathcal{L}_∞ is a useful diagnostic for mapping out an attractor structure. To illustrate that \mathcal{L} can be applied to any data as long as time-dependent PDFs can be constructed from the data, we presented the analysis of different classical music (e.g. see [16]). Finally, the width of PDFs was shown to play a dual role in information length in quantum systems. It cannot be over-emphasized that \mathcal{L} is path-specific and is a dynamical measure of the metric, capturing the actual statistical change that occurs during time evolution. This path-specificity would be crucial when it is desirable to control certain quantities according to the state of the system (e.g. time-dependent PDFs) at any given time (e.g. see [18]). Due to the generality of our methodology, we envision a large scope for further applications to different phenomena.

Appendix A. \mathcal{L} for the O-U process

Given Eqs. (16)–(18), one can compute Eq. (1) by carrying out the analysis in Appendix B as follows:

$$\frac{1}{\tau^2} = \frac{1}{2\beta^2} \left(\frac{d\beta}{dt} \right)^2 + 2\beta \left(\frac{dy}{dt} \right)^2 = \frac{2\gamma^2}{T^2} (r^2 + qT). \quad (A1)$$

In (A1), $q = \beta_0\gamma y_0^2$, $r = 2\beta_0 D - \gamma$, and $T = 2\beta_0 D(e^{2\gamma t} - 1) + \gamma$, following the same notation as in [17–19]. Note that q is due to the difference between y_0 and $y(t \rightarrow \infty)$, while r is due to the difference between D_0 and D . Thus, the first term in (A1), involving r , represents the information change due to the change in PDF width when $D_0 \neq 0$, while the second term is due to the movement of the PDF (or the mean value of x). Recalling $D_0 = \frac{\gamma}{2\beta_0}$, we can recast r , q and T in Eq. (A1) as

$$q = \frac{\gamma^2 y_0^2}{2D_0}, \quad r = \gamma \left(\frac{D}{D_0} - 1 \right), \quad T = \gamma \left[\frac{D}{D_0} (e^{2\gamma t} - 1) + 1 \right]. \quad (A2)$$

From Eq. (A2), we can see that the dimension of q , r and T is the inverse of time. By using (A1) and (A2) in Eq. (2), we obtain

$$\mathcal{L} = \frac{1}{\sqrt{2}} \int_{T_i}^{T_f} \left\{ \frac{1}{T} \frac{1}{T+r} \sqrt{r^2 + qT} \right\} dT. \quad (\text{A3})$$

To compute Eq. (A3) for $r \neq 0$, we let $Y = \sqrt{r^2 + qT}$ and recast it as

$$\mathcal{L} = \frac{\sqrt{2}}{r} \int_{Y_i}^{Y_f} \left\{ \frac{r^2}{Y^2 - r^2} + \frac{qr - r^2}{Y^2 + qr - r^2} \right\} dY = \frac{1}{\sqrt{2}} \left[\ln \left(\frac{Y - r}{Y + r} \right) \right]_{Y_i}^{Y_f} + \frac{\sqrt{2}}{r} H, \quad (\text{A4})$$

where Y_i and Y_f are Y evaluated at T_i and T_f , and H is defined as

$$H = \int_{Y_i}^{Y_f} \frac{qr - r^2}{Y^2 + qr - r^2} dY. \quad (\text{A5})$$

Eq. (A5) is to be evaluated separately for two cases: $q \geq r$ and $q < r$. First, for $q \geq r$, we use $Y = \sqrt{qr - r^2} \tan \theta$ in (A5) to obtain

$$H = \sqrt{qr - r^2} \int \frac{\sec^2 \theta}{\tan^2 \theta + 1} d\theta = \sqrt{qr - r^2} \left[\tan^{-1} \left(\frac{Y}{\sqrt{qr - r^2}} \right) \right]_{Y_i}^{Y_f}. \quad (\text{A6})$$

Secondly, in the $q < r$ case, we let $Y = \sqrt{|qr - r^2|} \sec \theta = \sqrt{r^2 - qr} \sec \theta$ ($\cos \theta = \frac{\sqrt{r^2 - qr}}{Y}$) to obtain

$$H = -\sqrt{r^2 - qr} \int \frac{1}{\sin \theta} d\theta = -\frac{\sqrt{r^2 - qr}}{2} \left[\ln \left(\frac{Y - \sqrt{r^2 - qr}}{Y + \sqrt{r^2 - qr}} \right) \right]_{Y_i}^{Y_f}. \quad (\text{A7})$$

We note that Eq. (A7) is continuous across $q = r$. In Eq. (A4), the contribution from the difference in PDF width through $r \neq 0$ and that from the difference in mean value of x (e.g. PDF peaks) through $q \neq 0$ appear in both first and second terms.

In the case $r = 0$ where the initial and final PDFs have the same width, $\beta(t) = \beta_0$ for all time, and Eqs. (A1) and (2) give us

$$\mathcal{L} = \frac{1}{\sqrt{2}} \int_{T_i}^{T_f} \frac{\sqrt{q}}{T^{\frac{3}{2}}} dT = -\sqrt{2q} \left[\frac{1}{\sqrt{T}} \right]_{T_i}^{T_f}. \quad (\text{A8})$$

We use that for $r = 0$, $T = \gamma e^{2\gamma t}$, $T_i = \gamma$ at $t = 0$ and simplify Eq. (A8) as

$$\mathcal{L}(t) = \frac{\sqrt{\gamma} y_0}{\sqrt{D}} [1 - e^{-\gamma t}] = \frac{1}{\sqrt{D/\gamma}} [y_0 - y], \quad (\text{A9})$$

where $y = y_0 e^{-\gamma t} = \langle x \rangle$ is the mean position. Thus, \mathcal{L} in Eq. (A9) is the change in the mean position $y_0 - y$ between initial time 0 and time t measured in units of the resolution $\sqrt{\frac{D}{\gamma}}$. Interestingly, this resolution $\sqrt{\frac{D}{\gamma}}$ is the standard deviation, which is the square root of the variance $\langle (x - \langle x \rangle)^2 \rangle = \frac{D}{\gamma} = \frac{1}{2\beta} = \frac{1}{2\beta_0}$. In general when $q \neq 0$ and $r \neq 0$, \mathcal{L} results from the mixed contribution from the entropy change ($r \neq 0$) and the change in y ($q \neq 0$) measured in units of the resolution. In a more technical term, β and y in Eq. (A1) constitute a hyperbolic geometry upon a suitable change of variables (e.g. see [18]).

Appendix B. Derivation of Eq. (A1)

From $p(x, t)$ in Eq. (13), we obtain

$$\frac{\partial p}{\partial t} = \left[\dot{\beta} \left(\frac{1}{2\beta} - (x-y)^2 \right) + 2\beta(x-y)\dot{y} \right] p, \quad (\text{A10})$$

where we recall $y = \langle x \rangle = y_0 e^{-\gamma t}$. Using Eq. (A10) in Eq. (1) gives us

$$\begin{aligned} \frac{1}{[\tau(t)]^2} &= \int dx \left[\left(\frac{1}{2\beta} - (x-y)^2 \right) \dot{\beta} + 2\beta(x-y)\dot{y} \right]^2 p \\ &= \dot{\beta}^2 \left[\left(\frac{1}{2\beta} \right)^2 - \frac{1}{\beta} \langle (x-y)^2 \rangle + \langle (x-y)^4 \rangle \right] + 4\beta^2 \langle (x-y)^2 \rangle \dot{y}^2 \\ &= \frac{1}{2\beta(t)^2} \left(\frac{d\beta}{dt} \right)^2 + 2\beta \left(\frac{dy}{dt} \right)^2. \end{aligned} \quad (\text{A11})$$

Here, $\dot{\beta} = \frac{d\beta}{dt}$ and $\dot{y} = \frac{dy}{dt}$; we used $\langle (x-y)^2 \rangle = \frac{1}{2\beta}$ and $\langle (x-y)^4 \rangle = 3 \left(\frac{1}{2\beta} \right)^2$. To obtain the last equation in (A1), it is useful to express β in Eq. (18) in the following form:

$$\beta = \frac{1}{\frac{1}{\beta_1} + \frac{1}{\beta_0} e^{-2\gamma t}} = \frac{1}{\frac{2D(1-e^{-2\gamma t})}{\gamma} + \frac{1}{\beta_0} e^{-2\gamma t}} = \frac{\gamma \beta_0 e^{2\gamma t}}{T}. \quad (\text{A12})$$

where we used $T = 2\beta_0 D(e^{2\gamma t} - 1) + \gamma$. By differentiating Eq. (A12), we then obtain

$$\dot{\beta} = -2\gamma \beta^2 e^{-2\gamma t} \left[\frac{2D}{\gamma} - \frac{1}{\beta_0} \right] = -2\beta^2 e^{-2\gamma t} [2\beta_0 D - \gamma] \frac{1}{\beta_0}. \quad (\text{A13})$$

Eq. (A13) and $r = 2\beta_0 D - \gamma$ then give us

$$\frac{\dot{\beta}^2}{2\beta^2} = 2\gamma^2 r^2 \frac{1}{T^2}. \quad (\text{A14})$$

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

$$2\beta \dot{y}^2 = 2q \gamma^2 \frac{1}{T}. \quad (\text{A15})$$

Finally, using (A14)-(A15) in (A11) gives us (A1).

References

1. Gibbs, A.L.; Su, F.E. On choosing and bounding probability metrics. *Int. Stat. Rev.* **2002**, *70*, 419–435.
2. Wootters, W.K. Statistical distance and Hilbert space. *Phys. Rev. D* **1981**, *23*, 357.
3. Frieden, B.R. *Science from Fisher Information* (Cambridge Univ. Press, 2000).
4. Ruppeiner, G. Thermodynamics: A Riemannian geometric model. *Phys. Rev. A* **1979**, *20*, 1608.
5. Schlögl, F. Thermodynamic metric and stochastic measures. *Z. Phys. B Cond. Matt.* **1985**, *59*, 449–454.
6. Feng, E.H.; Crooks, G.E. Far-from-equilibrium measurements of thermodynamic length. *Phys. Rev. E* **2009**, *79*, 012104.
7. Braunstein, S.L.; Caves, C.M. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.* **1994**, *72*, 3439.
8. Strobel, H.; Muessel, W.; Linnemann, D.; Zibold, T.; Hume, D.B.; Pezzé, L.; Smerzi, A.; Oberthaler, M.K. Fisher information and entanglement of non-Gaussian spin states. *Science* **2014**, *345*, 424–427.
9. Nulton, J.; Salamon, P.; Andresen, B.; Anmin, Q. Quasistatic processes as step equilibrations. *J. Chem. Phys.* **1985**, *83*, 334–338.

10. Crooks, G.E. Measuring thermodynamic length. *Phys. Rev. Lett.* **2007**, *99*, 100602.
11. Sivak, D.A.; Crooks, G.E. Thermodynamic metrics and optimal paths. *Phys. Rev. Lett.* **2012**, *8*, 190602.
12. Salamon, P.; Nulton, J.D.; Siragusa, G.; Limon, A.; Bedeaus, D.; Kjelstrup, S. A simple example of control to minimize entropy production. *J. Non-Equilib. Thermodyn.* **2002**, *27*, 45–55.
13. Plastino, A.R.; Casas, M.; Plastino, A. Fisher's information, Kullback's measure, and H-theorems. *Phys. Lett. A* **1998**, *246*, 498.
14. Polettini, M.; Esposito, M. Nonconvexity of the relative entropy for Markov dynamics: A Fisher information approach. *Phys. Rev. E* **2013**, *88*, 012112.
15. Nicholson, S.B.; Kim, E. Investigation of the statistical distance to reach stationary distributions. *Phys. Lett. A* **2015**, *379*, 83–88.
16. Nicholson, S.B.; Kim, E. Structures in sound: Analysis of classical music using the information length. *Entropy* **2016**, *18*, 258.
17. Heseltine, J.; Kim, E. Novel mapping in non-equilibrium stochastic processes. *J. Phys. A* **2016**, *49*, 175002.
18. 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.
19. Kim, E.; Hollerbach, R. Signature of nonlinear damping in geometric structure of a nonequilibrium process. *Phys. Rev. E* **2017**, *95*, 022137.
20. Hollerbach, R.; Kim, E. Information geometry of non-equilibrium processes in a bistable system with a cubic damping. *Entropy* **2017**, *19*, 268.
21. Kim, E.; Tenkès, L.-M.; Hollerbach, R.; Radulescu, O. Far-from-equilibrium time evolution between two gamma distributions. *Entropy* **2017**, *19*, 511.
22. Tenkès, L.-M.; Hollerbach, R.; Kim, E. Time-dependent probability density functions and information geometry in stochastic logistic and Gompertz models. *J. Stat. Mech. Theo. & Exp.* **2017**, 123201.
23. Kim, E.; Lewis, P. Information length in quantum system. *J. Stat. Mech. Theo. & Exp.* **2018**, 043106.
24. Risken, H. *The Fokker-Planck equation: methods of solutions and applications* (Berlin: Springer) (2013).
25. Kim, E.; Hollerbach, R. Time-dependent probability density function in cubic stochastic processes. *Phys. Rev. E* **2016**, *94*, 052118.
26. Kern Scores. Available online: <http://kernscores.stanford.edu/>.
27. Ott, E. *Chaos in Dynamical Systems*; Cambridge University Press: Cambridge, 2002.
28. Andrews, M. Quantum mechanics with uniform forces, *Am. J. Phys.* **2018**, *78*, 1361.
29. R.W. Robinett and L.C. Bassett, Foundation of Phys. Lett. **17**, 607 (2004).