Vol. 36 (2005) *ACTA PHYSICA POLONICA B* No 5

DYNAMICS OF UNCERTAINTY IN NONEQUILIBRIUM RANDOM MOTION*

Piotr Garbaczewski

Institute of Physics, University of Zielona Góra
65-516 Zielona Góra, Poland

(Received November 8, 2004)

Dedicated to Professor Andrzej Fuliński, with admiration

Shannon information entropy is a natural measure of probability (de)localization and thus (un)predictability in various procedures of data analysis for model systems. We pay particular attention to links between the Shannon entropy and the related Fisher information notion, which jointly account for the shape and extension of continuous probability distributions. Classical, dynamical and random systems in general give rise to time-dependent probability densities and associated information measures. The induced dynamics of Shannon and Fisher functionals reveals an interplay among various characteristics of the considered diffusion-type systems: information, uncertainty and localization while put against mean energy and its balance.

PACS numbers: 02.50.–r, 89.70.+c, 05.40.–a

1. Introduction

We shall investigate relationships between the dynamical features of the differential entropy (Shannon entropy of general time-dependent continuous probability densities), [1, 2], and so-called hydrodynamical conservation laws (mass/probability, momentum and energy balance in the mean) of the corresponding (ir)reversible diffusion-type process.

In part, our arguments derive from a standard trajectory interpretation in which random transport is modeled in terms of the Markovian process and its sample paths. The pertinent process obviously complies with the Fokker–Planck dynamics of an initially prescribed probability density, [3, 4].

* Presented at the XVII Marian Smoluchowski Symposium on Statistical Physics, Zakopane, Poland, September 4–9, 2004.

However, we would like to point out that a generic property of physically interesting cases is their conflict with rather stringent growth and Hölder continuity restrictions for drift and diffusion coefficient functions. Those bounds need to be respected for a mathematically consistent definition of the process and its transition density functions. In most of "typical" cases, the uniqueness and non-explosiveness of the process cannot be guaranteed, see however [5, 6], how to evade the explosiveness problem.

This formal defect of a theoretical framework is usually bypassed in a pragmatic computer-assisted research by neglecting the unwanted (even if annoying, interpreted as artifacts) contributions to the data. In view of low probability for troublesome events (explosive behavior), it is often taken for granted that a mathematical pedantry is here unnecessary and that the Langevin equation can be employed in the study of diffusion-type processes without any specific precautions. This attitude is omnipresent, when one attempts to solve explicitly the "obvious" Fokker–Planck or Smoluchowski diffusion equation, but does *not* inquire into an issue of transition probability density functions. Needless to say, with the latter step ignored, the random variable and random path notions are often maintained as legitimate elements of the analysis.

On the other hand, the previously mentioned restrictions on drift and diffusion coefficient functions may be relaxed in a controlled way to allow for a consistent theory. There is an obvious price to be paid, one should admit and learn to live with non-unique and possibly explosive stochastic processes, all of them being capable to drive accordingly a *unique* probability density. Examples of such milder (than usual) restrictions can be found in Refs. [5–7].

The previous obstacles motivate a principal peculiarity of our approach which is rooted in the fact that we extract relevant data exclusively from the (basically, spatial) probability density of the pertinent dynamical process and this density gradient, with no explicit mention of random or deterministic paths. Clearly, there are many distinct stochastic processes which can be associated with the once prescribed Fokker–Planck dynamics of a concrete probability density.

It is widely accepted in the literature to invoke relative Kullback–Leibler entropies as "distance measures" in the set of different probability densities. In particular, for comparison of different solutions of a given Fokker–Planck equation, [4, 10]. One often takes for granted that the Kullback entropy is a proper analog of Boltzmann's H-function in the diffusion process setting. The reason is that it never takes negative values, while the differential entropy does. Its time rate is negative, hence refers to a continually decreasing function in accordance with thermodynamical intuitions, which is not necessarily the case for Shannon entropies, [1].

There is one minor obstacle: a closer inspection shows that the Kullback entropy is mainly explored under standard severe restrictions upon drift and diffusion coefficients. See *e.g.* [4] for a verbal statement: "we assume that the drift coefficients have no singularities and that they do not allow the solutions to run away to infinity". Not surprisingly, in a statistical physics lore, a tacit assumption is that "all solutions of the Fokker–Planck equation finally agree if we wait long enough". Hence it is believed to be immaterial to discuss their behavior in other regimes, than close-to-equilibrium (asymptotic invariant density).

Another peculiarity of our approach is that we are not quite interested in "measuring a distance" between two different probability densities. We rather wish to make a comparison of the very *same* non-equilibrium density and its differential entropy at different stages of their time evolution. In particular, the difference of the respective entropy values at two time instants is a legitimate "distance measure" (information gain or loss) [1], the time rate of information entropy is also a well defined quantity. For those reasons, we deliberately avoid the use of the Kullback entropy and insist on investigating the role and potential utility of the Shannon-type information entropy *per se*.

Other motivations come from varied attempts to use information theory concepts as natural tools for quantifying signatures of disorder and its intrinsic dynamics (time rate of generation/propagation of disorder, information flow, entropy production rate). This involves an issue of non-equilibrium steady states and the time rate ("speed") of an asymptotic approach to equilibrium, when time-reversible stationary processes ultimately enter the game, [9]. Discussions [11–13] of a physical role of the probability density gradient in classical non-equilibrium thermodynamics of irreversible processes are worth mentioning, to place our discussion in a proper context.

An analysis of links [16, 17] between dynamical systems, weak noise and information entropy production is also useful to that end. An independent input comes from general studies of the dynamical origin of increasing entropy ("dynamical foundations of the evolution of entropy to maximal states"), entirely carried out with respect to time-dependent probability densities, [3, 17, 18] see also [4, 8].

The intertwined dynamics of the differential (information) entropy and the probability localization properties (dynamics of uncertainty) appears to be an intrinsic physical feature of any formalism operating with general time-dependent (in the present paper, spatial) probability distributions.

2. Information entropy and its dynamics

Let us consider a classical dynamical system in R^n whose evolution is governed by equations of motion:

$$\dot{x} = f(x)\,, \tag{1}$$

where $\dot{x}$ stands for the time derivative and f is an R^n-valued function of $x \in R^n$, $x = \{x_1, x_2, \ldots, x_n\}$. The statistical ensemble of solutions of such dynamical equations can be described by a time-dependent probability density $\rho(x,t)$ whose dynamics is given by the generalized Liouville (in fact, continuity) equation

$$\partial_t \rho = -\nabla(f\rho)\,, \tag{2}$$

where $\nabla \doteq \{\partial/\partial x_1, \ldots, \partial/\partial x_n\}$.

With any continuous probability density $\rho \doteq \rho(x,t)$, where $x \in R^n$ and we allow for an explicit time-dependence, we can associate a probability density functional named Shannon entropy of a continuous probability distribution (convergence of an integral is presumed), [1]

$$\mathcal{S}(\rho) = -\int \rho \, \ln \rho \, dx\,. \tag{3}$$

In general, $\mathcal{S}(\rho) \doteq \mathcal{S}(t)$ depends on time. Let us take for granted that an interchange of time derivative with an indefinite integral is allowed (suitable precautions are necessary with respect to the convergence of integrals). Then, we readily get an identity, [14–16]:

$$\dot{\mathcal{S}} = \int \rho \, (\mathrm{div} f) dx \doteq \langle \nabla f \rangle\,. \tag{4}$$

Accordingly, the information entropy $\mathcal{S}(t)$ grows with time only if the dynamical system has positive mean flow divergence.

However, in general $\dot{\mathcal{S}}$ is not positive definite. For example, dissipative dynamical systems are characterized by the negative (mean) flow divergence. Fairly often, the divergence of the flow is constant, [14]. Then, an "amount of information" carried by a corresponding statistical ensemble (*e.g.* its density) increases, which is paralleled by the information entropy decay (decrease).

An example of a system with a point attractor (sink) at origin is a one-dimensional non-Hamiltonian system $\dot{x} = -x$. In this case $\mathrm{div} f = -1$ and $\dot{\mathcal{S}} = -1$. A discussion of dynamical systems with strange (multifractal) attractors, for which the Shannon information entropy decreases indefinitely (the pertinent steady states are no longer represented by probability density functions) can be found in [14, 16].

We note that for Hamiltonian systems, the phase–space flow is divergence-less, hence $\dot{\mathcal{S}} = 0$ which implies that "information is conserved" in Hamiltonian dynamics. Take for example a two-dimensional conservative system with $\dot{x} = p/m$ and $\dot{p} = (-\nabla V)$, where $H = p^2/2m + V(x)$. The classical equations of motion yield the standard Liouville equation (which is a special case of Eq. (4)):

$$\frac{\partial}{\partial t}\rho = -\frac{p}{m}\frac{\partial}{\partial x}\rho + (\nabla V)\frac{\partial}{\partial p}\rho \tag{5}$$

for the phase–space density $\rho(x, p)$. The corresponding divergence vanishes and the phase space volume is conserved. For non-Hamiltonian systems we may generically expect the phase–space volume contraction, expansion or both at different stages of time evolution, [14, 16].

In case of a general dissipative dynamical system (1), a controlled admixture of noise can stabilize dynamics and yield asymptotic invariant densities. For example, an additive modification of the right-hand side of Eq. (1) by white noise term $A(t)$ where $\langle A_i(s)\rangle = 0$ and $\langle A_i(s)A_j(s')\rangle = 2q\delta(s-s')\delta_{ij}$, $i = 1, 2, \ldots n$, implies the Fokker–Planck–Kramers equation

$$\partial_t \rho = -\nabla\,(f\,\rho) + q\Delta\rho\,, \tag{6}$$

where $\Delta \doteq \nabla^2 = \sum_i \partial^2/\partial x_i^2$. Accordingly, the differential entropy dynamics would take another form than this defined by Eq. (4)

$$\dot{\mathcal{S}} = \int \rho\,(\mathrm{div} f)dx + q\int \frac{1}{\rho}(\nabla\rho)^2\,dx\,. \tag{7}$$

Now, the dissipative term $\langle\nabla\, f\rangle < 0$ can be counterbalanced by a strictly positive stabilizing contribution $q\sum_i \int \frac{1}{\rho}(\partial\rho/\partial x_i)^2\,dx$. This allows to expect that, under suitable circumstances dissipative systems with noise may yield $\dot{\mathcal{S}} = 0$. In case of $\langle\nabla\, f\rangle \geq 0$, the information entropy would grow monotonically.

At this point, we depart from an explicit phase–space background for further discussion and consider exclusively *spatial* Markov diffusion processes with a diffusion coefficient D (constant or time-dependent, with standard dimensions of $k_{\mathrm{B}}T/m\beta$ where β is a friction coefficient, or $\hbar/2m$). We admit them to drive space–time inhomogeneous probability densities $\rho = \rho(\vec{x}, t)$ with $\vec{x} \in R^3$. The density gradient is introduced in conjunction with so-called osmotic velocity field $\vec{u} = D\vec{\nabla}\ln\rho$, *cf.* [18]. The probability density is to obey the continuity equation, with $\vec{v}$ set in correspondence with the previous vector-valued function $f \in R^n$

$$\partial_t\rho = -\vec{\nabla}\cdot(\vec{v}\,\rho)\,, \tag{8}$$

where a (postulated) decomposition: $\overrightarrow{v}(\overrightarrow{x},t) = \overrightarrow{v} \doteq \overrightarrow{b} - \overrightarrow{u}$ allows us to infer the related Fokker–Planck equation

$$\partial_t \rho = D\Delta\rho - \overrightarrow{\nabla} \cdot (\overrightarrow{b}\rho)\,, \tag{9}$$

with a forward drift function $\overrightarrow{b}(\overrightarrow{x},t)$.

To make things simpler, we assume to have given a concrete functional expression for the time-independent forward drift $\overrightarrow{b}(\overrightarrow{x})$ (here, we do not bother about its detailed justification on phenomenological or model construction grounds) and fix initial/boundary data for the probability density ρ. We shall not demand the validity of standard mathematical restrictions (growth and Hölder continuity conditions), guaranteeing the existence of non-explosive solutions $\overrightarrow{X}(t)$ of the underlying stochastic differential equation, since that would exclude a vast number of physically interesting situations, when the corresponding partial differential (Fokker–Planck) equation nonetheless *has* well defined solutions of the initial/boundary value problem. Therefore, we prefer to investigate random diffusive motion in terms of probability densities, and *not* directly in terms of paths (sample trajectories) induced by random variable $\overrightarrow{X}(t)$.

With a solution $\rho(\overrightarrow{x},t)$ of the Fokker–Planck equation, we associate its differential (Shannon information) entropy $\mathcal{S}(t) = -\int \rho \ln \rho\, d^3x$ which typically is not time-independent, [14, 16]. The evolution (dynamics of information) and rate of change in time of the entropy $\mathcal{S}$ directly follow.

First, let us notice that in the particular case of $\overrightarrow{v} = -\overrightarrow{u}$ (*i.e.* $\overrightarrow{b} = 0$), where $\overrightarrow{u} = D\overrightarrow{\nabla} \ln \rho$, we infer the standard free Brownian motion outcome, [11]

$$\frac{d\mathcal{S}}{dt} = D \cdot \int \frac{(\overrightarrow{\nabla}\rho)^2}{\rho} d^3x > 0\,, \tag{10}$$

to be compared with the previously introduced stabilizing term in Eq. (7). Thus, information entropy definitely increases in the Brownian motion and its time rate may be interpreted as the rate of information decay (uncertainty increase) in the course of the diffusion process, in close parallel with the casual perception of the laws of thermodynamics.

While passing from the free Brownian motion to the forced one and more general diffusion-type processes, we shall demand the current velocity $\overrightarrow{v}(\overrightarrow{x},t)$ to be a gradient field $\overrightarrow{v} \doteq \overrightarrow{b} - \overrightarrow{u}$, where the forward drift $\overrightarrow{b}(\overrightarrow{x},t)$ of the process may be time-dependent.

Boundary restrictions upon ρ, $\overrightarrow{v}\rho$ and $\overrightarrow{b}\rho$ to vanish at spatial infinities (or at finite spatial volume boundaries) yield the information entropy balance equation

$$\frac{d\mathcal{S}}{dt} = \int \left[\rho\,(\overrightarrow{\nabla} \cdot \overrightarrow{b}) + D \cdot \frac{(\overrightarrow{\nabla}\rho)^2}{\rho}\right] d^3x\,, \tag{11}$$

to be compared with the previous, vanishing $\overrightarrow{b}$, case. We can rewrite this equation as follows

$$D\dot{\mathcal{S}} \doteq \langle \overrightarrow{u}^2 \rangle + D\langle \overrightarrow{\nabla} \cdot \overrightarrow{b} \rangle = D\langle \overrightarrow{\nabla} \cdot \overrightarrow{v} \rangle \,, \tag{12}$$

or equivalently

$$D\dot{\mathcal{S}} = \langle \overrightarrow{v}^2 \rangle - \langle \overrightarrow{b} \cdot \overrightarrow{v} \rangle = -\langle \overrightarrow{v} \cdot \overrightarrow{u} \rangle \,. \tag{13}$$

Note that we have employed an identity

$$\langle \overrightarrow{u}^2 \rangle = -D\langle \overrightarrow{\nabla} \cdot \overrightarrow{u} \rangle \,. \tag{14}$$

The osmotic velocity field, by its very definition, always has negative mean divergence.

The mean divergence of the current velocity field has no definite sign. Therefore, the monotonic increase of $\mathcal{S}(t)$ is guaranteed only if $\langle \overrightarrow{\nabla} \cdot \overrightarrow{v} \rangle > 0$, or equivalently $\langle \overrightarrow{\nabla} \cdot \overrightarrow{b} \rangle > \langle \overrightarrow{\nabla} \cdot \overrightarrow{u} \rangle$. Invariant probability densities are allowed when the information entropy remains constant in time: $d\mathcal{S}/dt = 0$, that is when $\langle \overrightarrow{\nabla} \cdot \overrightarrow{v} \rangle = 0$, *i.e.* $\langle \overrightarrow{\nabla} \cdot \overrightarrow{b} \rangle = \langle \overrightarrow{\nabla} \cdot \overrightarrow{u} \rangle$.

The simplest realization of the state of equilibrium is granted by $\overrightarrow{b} = \overrightarrow{u} = D\overrightarrow{\nabla} \ln \rho$, when the diffusion current identically vanishes: $\overrightarrow{v} = \overrightarrow{0}$. For familiar Smoluchowski diffusion processes whose drifts have the form $\overrightarrow{b} = -(1/m\beta)\overrightarrow{\nabla} V$, where V is time-independent, we immediately arrive at the classic equilibrium identity

$$-\frac{1}{k_{\mathrm{B}}T}\overrightarrow{\nabla} V = \overrightarrow{\nabla} \ln \,\rho \,, \tag{15}$$

with the implicit Einstein fluctuation–dissipation formula $D = k_{\mathrm{B}}T/m\beta$ (k_{B} is the Boltzmann constant).

It is not obvious at all that the differential (Shannon information) entropy needs to increase, when a given "attracting" state of equilibrium (invariant density) is being asymptotically approached. Entropy decay scenario seems to be equally likely in this situation.

A hint to this end: invoking the standard Smoluchowski diffusion, fix $\overrightarrow{b}(\overrightarrow{x})$, *i.e.* external force, and fine-tune an initial density $\rho_0(\overrightarrow{x})$ so that $\langle \overrightarrow{\nabla} \cdot \overrightarrow{b} \rangle < \langle \overrightarrow{\nabla} \cdot \overrightarrow{u} \rangle$ and, therefore, $\langle \overrightarrow{\nabla} \cdot \overrightarrow{v} \rangle < 0$. Realization: consider the one-dimensional example with $b(x) = -\gamma x$, $\gamma > 0$ and choose $\rho_0(x) = [1/(\sigma\sqrt{2\pi}] \exp[-x^2/2\sigma^2]$, implying $\langle \nabla \cdot u \rangle = -D/\sigma^2$. Finally adjust σ and/or γ to yield $D/\sigma^2 < \gamma$.

Let us also observe that, in view of $D\dot{\mathcal{S}} = -\langle \overrightarrow{v} \cdot \overrightarrow{u} \rangle$, by reintroducing the diffusion current $\rho\,\overrightarrow{v}$ and recalling that $\overrightarrow{u} = (D\overrightarrow{\nabla}\rho)/\rho$, we arrive at

$$D\frac{d\mathcal{S}}{dt} = -\int \left[\rho^{-1/2}(\rho\overrightarrow{v})\right] \cdot \left[\rho^{-1/2}(D\overrightarrow{\nabla}\rho)\right] d^3x \,. \tag{16}$$

By means of the Schwarz inequality we infer an upper bound on the magnitude of the information entropy time rate:

$$D\Big|\frac{d\mathcal{S}}{dt}\Big| \leq \left\langle \overrightarrow{v}^2 \right\rangle^{1/2} \left\langle \overrightarrow{u}^2 \right\rangle^{1/2} . \tag{17}$$

As a byproduct we realize that a necessary condition for $\frac{d\mathcal{S}}{dt} \neq 0$ is that *both* $\left\langle \overrightarrow{v}^2 \right\rangle$ and $\left\langle \overrightarrow{u}^2 \right\rangle$ are non-vanishing. A sufficient condition for $\frac{d\mathcal{S}}{dt} = 0$ is that *any* of $\left\langle \overrightarrow{v}^2 \right\rangle$, $\left\langle \overrightarrow{u}^2 \right\rangle$, or both vanish.

3. Information entropy balance in Smoluchowski diffusion process

Remembering that in the standard Brownian motion, essentially the same mathematical formalism applies to a single particle and to a statistical ensemble of identical noninteracting Brownian particles, we shall adopt to our purposes basic tenets of so-called thermodynamic formalism of isothermal diffusion processes, [5,7,19] (see also [20] and [21]), originally introduced in connection with nonequilibrium thermodynamics of single macromolecules immersed in an ambient fluid at a constant temperature, and promoted in [5] to the status of "stochastic macromolecular mechanics".

Let us discuss in more detail Eq. (13) for the differential entropy balance which is extremely persuasive in the special case of Smoluchowski diffusions. Indeed, then $\overrightarrow{b} \doteq \overrightarrow{F}/(m\beta)$ stands for an externally acting force, capable of performing a mechanical work which in turn may be converted into heat. We refer to the standard phase–space conceptual background, [8, 21].

According to [5, 7, 19] (we adjust their framework and notation to our purposes), close to equilibrium, one expects the information entropy to decrease in the course of the Smoluchowski diffusion process. The mean rate of the entropy loss per unit of mass, equals

$$\frac{d\mathcal{Q}}{dt} \doteq \frac{1}{D} \int \frac{1}{m\beta} \overrightarrow{F} \cdot \overrightarrow{j}\, d^3x = \frac{1}{D} \langle \overrightarrow{b} \cdot \overrightarrow{v} \rangle . \tag{18}$$

That can be rewritten otherwise: $k_{\mathrm{B}} T \dot{\mathcal{Q}} = \int \overrightarrow{F} \cdot \overrightarrow{j}\, d^3x$, where T is the temperature of the bath. In the formal thermodynamical lore, we deal here with the time rate at which the mechanical work is being dissipated into thermal environment in the form of (removed) heat. Let us point out that this interpretation is surely true under equilibrium conditions [5]. In general, far from equilibrium, the sign of $d\mathcal{Q}/dt$ remains indefinite and may refer to heat absorption (if negative) instead of heat removal.

The nonnegative term in Eq. (13) can be consistently interpreted, *cf.* [5], as the measure of the entropy gain per unit of time by the diffusion process.

Accordingly, we have

$$\frac{d\mathcal{S}}{dt} = \frac{d\mathcal{S}_{\rm gain}}{dt} - \frac{d\mathcal{Q}}{dt}\,, \tag{19}$$

where $d\mathcal{S}_{\rm gain}/dt \doteq (1/D)\left\langle \overrightarrow{v}^2 \right\rangle$. If the entropy gain is counterbalanced by heat removal, we may have $d\mathcal{S}/dt = 0$.

Let us mention that our "entropy gain" is named "entropy production" in Ref. [5]. In the earlier literature on the subject, [30], the entropy production name has been reserved to the accumulating entropy surplus which is being removed from the system under consideration to the environment. In our discussion, just to the contrary, the information entropy appears to be pumped into the system (*e.g.* the diffusion process) instead of being removed.

The relationship:

$$\overrightarrow{j} \doteq \rho\, D \overrightarrow{F}_{\rm th} \tag{20}$$

defines a thermodynamic force $\overrightarrow{F}_{\rm th}$ associated with the Smoluchowski diffusion

$$k_{\rm B}T\, \overrightarrow{F}_{\rm th} = \overrightarrow{F} - k_{\rm B}T\, \overrightarrow{\nabla} \ln\rho \doteq -\overrightarrow{\nabla}\Psi\,. \tag{21}$$

Notice that

$$\overrightarrow{v} = -\frac{1}{m\beta}\overrightarrow{\nabla}\Psi\,. \tag{22}$$

In the absence of external force (free Brownian motion), we obviously get $D\overrightarrow{F}_{\rm th} = -\overrightarrow{u}$, $\dot{\mathcal{Q}} = 0$ and $\dot{\mathcal{S}} = \dot{\mathcal{S}}_{\rm gain}$, hence delocalization coincides with the "diffusion of probability".

The mean value of the potential

$$\Psi = V + k_{\rm B}T \ln\rho \tag{23}$$

of the thermodynamic force defines the obvious diffusion process analogue of the Helmholtz free energy

$$\langle\Psi\rangle = \langle V\rangle - T\,\mathcal{S}_{\rm G}\,, \tag{24}$$

where the dimensional version of information entropy $\mathcal{S}_{\rm G} \doteq k_{\rm B}\mathcal{S}$ has been introduced (actually, it is a direct analog of the Gibbs entropy). The expectation value of the mechanical force potential $\langle V\rangle$ plays here the role of the mean internal energy.

By assuming that $\rho V \overrightarrow{v}$ vanishes at integration volume boundaries (or infinity), we easily get the time rate of Helmholtz free energy

$$\frac{d}{dt}\langle\Psi\rangle = -k_{\rm B}T\dot{\mathcal{Q}} - T\dot{\mathcal{S}}_{\rm G}\,, \tag{25}$$

where $k_{\rm B}T\dot{\mathcal{Q}} = \int \overrightarrow{F}\cdot\overrightarrow{j}\,d^3x$ and $T\dot{\mathcal{S}}_{\rm G} = \int (k_{\rm B}T\overrightarrow{F}_{\rm th} - \overrightarrow{F})\cdot\overrightarrow{j}\,d^3x$. In view of Eq. (19) we get

$$\frac{d}{dt}\langle\Psi\rangle = -(m\beta)\left\langle\overrightarrow{v}^2\right\rangle \,, \tag{26}$$

which is either negative or vanishes. Therefore, the Helmholtz free energy either remains constant in time or decreases as a function of time.

In the presence of external forces this property quantifies a possible asymptotic approach towards a minimum corresponding to an invariant density of the process. Indeed, a particular example of an equilibrium (invariant) density reads $\rho(x) = (1/Z)\exp(-V/k_{\rm B}T)$, where $Z = \int \exp(-V/k_{\rm B}T)\,dx$. Such ρ sets the pertinent minimum of $\langle\Psi\rangle$ at $\langle\Psi\rangle = \Psi = -k_{\rm B}T\ln Z$. This corresponds to $\Psi = V + k_{\rm B}T\ln\rho = \mathrm{const}$ and thus trivially implies $\overrightarrow{\nabla}\Psi = \overrightarrow{0} = \overrightarrow{v}$.

One should be aware that an invariant density as well may not exist: in case of free Brownian motion there is no invariant density.

4. Localization toolbox: Shannon entropy and Fisher information

For simplicity all of our further discussion will be restricted to one space dimension.

Let us consider the Gaussian probability density on the real line R as a reference density function

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-x_0)^2}{2\sigma^2}\right].$$

Among *all* one-dimensional distribution functions $\rho(x)$ with a finite mean, subject to the constraint that the standard deviation is fixed at σ, it is the Gauss function with half-width σ which sets a maximum of the differential entropy, [1]. For the record, let us add that if only the mean is given for probability density functions on R, then there is no maximum entropy distribution in their set.

The differential entropy of the Gauss density has a simple analytic form, independent of the mean value x_0 and maximizes an inequality

$$\mathcal{S}(\rho) \leq \frac{1}{2}\ln\left(2\pi e\sigma^2\right) \,. \tag{27}$$

This imposes a useful bound upon the so-called entropy power, [1]

$$\frac{1}{\sqrt{2\pi e}}\exp[\mathcal{S}(\rho)] \leq \sigma \,, \tag{28}$$

with an obvious bearing on the spatial localization of the density ρ, hence spatial (un)certainty of position measurements. We can say that almost surely, with probability 0.998, the probability is concentrated within the interval of the length 6σ which is centered about the mean value x_0 of the Gaussian density ρ.

The Shannon entropy of an arbitrary continuous probability density is unbounded form below and from above, but in the subset of all densities with a finite mean and a fixed variance σ^2, we actually have an upper bound set by Eq. (27). Note that not only for small, but also for relatively large mean deviation values $\sigma < 1/\sqrt{2\pi e} \simeq 0.26$ the differential entropy $\mathcal{S}(\rho)$ becomes negative.

Let us discuss to what extent, the Shannon entropy can be viewed as a measure of localization in the configuration space of the dynamical system.

Let us consider a one-parameter family of probability densities $\rho_\alpha(x)$ on R whose first (mean) and second moments (effectively, the variance) are finite. The parameter-dependence is here not completely arbitrary and we assume standard regularity properties that allow to differentiate various functions of ρ_α with respect to the parameter α under the sign of an (improper) integral.

Namely, let us denote $\int x\rho_\alpha(x)dx = f(\alpha)$ and $\int x^2\rho_\alpha dx < \infty$. We demand that as a function of $x \in R$, the modulus of the partial derivative $\partial\rho_\alpha/\partial\alpha$ is bounded by a function $G(x)$ which together with $xG(x)$ is integrable on R. This implies, the existence of $\partial f/\partial\alpha$ and an important inequality

$$\int (x-\alpha)^2\rho_\alpha dx \int \left(\frac{\partial \ln \rho_\alpha}{\partial\alpha}\right)^2 \rho_\alpha dx \geq \left(\frac{df(\alpha)}{d\alpha}\right)^2 , \tag{29}$$

directly resulting from

$$\frac{df}{d\alpha} = \int \left[(x-\alpha)\rho_\alpha^{1/2}\right] \left[\frac{\partial(\ln\rho_\alpha)}{\partial\alpha}\rho_\alpha^{1/2}\right] dx \tag{30}$$

via the standard Schwarz inequality, [22]. The equality appears if $\rho_\alpha(x)$ is the Gauss function with mean value α.

At this point let assume that the mean value of ρ_α actually equals α and we fix at σ^2 the value $\langle(x-\alpha)^2\rangle = \langle x^2\rangle - \alpha^2$ of the variance (in fact, standard deviation from the mean value) of the probability density ρ_α. The previous inequality now takes the familiar form

$$\mathcal{F}_\alpha \doteq \int \frac{1}{\rho_\alpha}\left(\frac{\partial\rho_\alpha}{\partial\alpha}\right)^2 dx \geq \frac{1}{\sigma^2} , \tag{31}$$

where an integral on the left-hand side is the so-called Fisher information of ρ_α, known to appear in various problems of statistical estimation theory, as well as an ingredient of a number of information — theoretic inequalities. In view of $\mathcal{F}_\alpha \geq 1/\sigma^2$, we realize that the Fisher information is more sensitive indicator of the probability density localization than the entropy power, Eq. (28).

Let us define $\rho_\alpha(x) \doteq \rho(x-\alpha)$. Then, the Fisher information can be readily transformed to the conspicuously quantum mechanical form (up to a factor D^2 with $D = \hbar/2m$)

$$\frac{1}{2}\mathcal{F}_\alpha = \frac{1}{2}\int \frac{1}{\rho}\left(\frac{\partial \rho}{\partial x}\right)^2 dx = \int \rho \frac{u^2}{2} dx = -\langle Q\rangle \,, \tag{32}$$

where $u \doteq \nabla \ln \rho$ (up to a factor D) represents an osmotic velocity field, [18, 24], and an average $\langle Q\rangle = \int \rho\, Q\, dx$ is carried out with respect to the function

$$Q = 2\frac{\Delta \rho^{1/2}}{\rho^{1/2}}\,. \tag{33}$$

As a consequence of Eq. (31), we have $-\langle Q\rangle \geq 1/2\sigma^2$ for all relevant probability densities with variance σ^2.

An important inequality, valid under an assumption $\rho_\alpha(x) = \rho(x-\alpha)$, has been proved in [31]

$$\frac{1}{\sigma^2} \leq (2\pi e)\, \exp[-2\mathcal{S}(\rho)] \leq \mathcal{F}_\alpha\,. \tag{34}$$

It tells us that the lower bound for the Fisher information is in fact given a sharper form by means of the (squared) inverse entropy power. Our two information measures appear to be correlated.

Under an additional decomposition/factorization ansatz (of the quantum mechanical $L^2(R^n)$ provenance) that $\rho(x) \doteq |\psi|^2(x)$, where a real or complex function $\psi = \sqrt{\rho}\exp(i\phi)$ is a normalized element of $L^2(R)$, another important inequality holds true, [31]

$$\mathcal{F}_\alpha = 4\int \left(\frac{\partial \sqrt{\rho}}{\partial x}\right)^2 dx \leq 16\pi^2\tilde{\sigma}^2\,, \tag{35}$$

provided the Fisher information takes finite values. Here, $\tilde{\sigma}^2$ is the variance of the "quantum mechanical momentum canonically conjugate to the position observable", up to (skipped) dimensional factors. In the above, we have exploited the Fourier transform $\tilde{\psi} \doteq (\mathcal{F}\psi)$ of ψ to arrive at $\tilde{\rho} \doteq |\tilde{\psi}|^2$ whose variance the above $\tilde{\sigma}^2$ actually is.

Let us point out that the Fisher information $\mathcal{F}(\rho)$ may blow up to infinity under a number of circumstances: when ρ approaches the Dirac delta behavior, if ρ vanishes over some interval in R or is discontinuous. We observe that $\mathcal{F} > 0$ because it may vanish only when ρ is constant everywhere on R, hence when ρ is *not* a probability density on R.

In view of two previous inequalities, we find out that not only the Fisher information, but also an entropy power may be bounded from below and above. Namely, we have

$$\frac{1}{\sigma^2} \leq \mathcal{F}_\alpha \leq 16\pi^2\tilde{\sigma}^2\,, \tag{36}$$

which implies $1/2\sigma^2 \leq -\langle Q\rangle \leq 8\pi^2\tilde{\sigma}^2$ and, furthermore,

$$\frac{1}{4\pi\tilde{\sigma}} \leq \frac{1}{\sqrt{2\pi e}}\,\exp[\mathcal{S}(\rho)] \leq \sigma\,. \tag{37}$$

Most important outcome of Eq. (37) is that the differential entropy $\mathcal{S}(\rho)$ typically may be expected to be a well behaved quantity: with finite lower and upper bounds. A standard statement in this regard is: Shannon entropy of a continuous probability density is neither bounded from below nor from above, [1, 2].

5. Dynamics of uncertainty: mean energy versus localization

When multiplied by D^2, a potential-type function $Q = Q(x,t)$, *cf.* (33) notoriously appears in the hydrodynamical formalism of quantum mechanics as the so-called de Broglie–Bohm quantum potential ($D = \hbar/2m$), [24, 26]. It appears as well in the corresponding formalism for diffusion-type processes, including the standard Brownian motion (then, $D = k_{\mathrm{B}}T/m\beta$, see *e.g.* [23–25, 27, 29]. We have

$$Q = 2D^2\frac{\Delta\rho^{1/2}}{\rho^{1/2}} = \frac{1}{2}u^2 + D\nabla\cdot u\,, \tag{38}$$

and it is instructive to notice that the gradient of Q trivially appears (*i.e.* merely as a consequence of the heat equation, [23, 24, 26]) in the hydrodynamical (momentum) conservation law appropriate for the free Brownian motion

$$\partial_t v + (v\cdot\nabla)v = -\nabla Q\,. \tag{39}$$

We assume, modulo restrictions upon drift functions [6, 7], that the Smoluchowski dynamics can be resolved in terms of (possibly non-unique) Markovian diffusion-type processes. Then, the following compatibility equations follow in the form of local (hydrodynamical) conservation laws for the diffusion process, [24, 26]

$$\partial_t\rho + \nabla(v\rho) \;=\; 0\,, \tag{40}$$

$$(\partial_t + v\cdot\nabla)v \;=\; \nabla(\Omega - Q)\,, \tag{41}$$

where, not to confuse this notion with the previous force field potential V, we denote by $\Omega(x)$ the so-called volume potential for the process

$$\Omega = \frac{1}{2}\left(\frac{F}{m\beta}\right)^2 + D\nabla\left(\frac{F}{m\beta}\right) . \tag{42}$$

Obviously the free Brownian law, Eq. (39), comes out as the special case.

In the above (we use a short-hand notation $v \doteq v(x,t)$)

$$v \doteq b - u = \frac{F}{m\beta} - D\frac{\nabla\rho}{\rho} \tag{43}$$

defines the current velocity of Brownian particles in external force field. This formula allows us to transform the continuity equation into the Fokker–Planck equation and back.

By considering $(-\rho)(x,t)$ and $s(x,t)$, such that $v = \nabla s$, as canonically conjugate fields, we can invoke the variational calculus, [27, 28]. Namely, one may derive the continuity (and thus Fokker–Planck) equation together with the Hamilton–Jacobi type equation (whose gradient implies the hydrodynamical conservation law Eq. (41))

$$\partial_t s + \frac{1}{2}(\nabla s)^2 - (\Omega - Q) = 0 \, , \tag{44}$$

by means of the extremal (least, with fixed end-point variations) action principle involving the (mean) Lagrangian

$$\mathcal{L} = -\int \rho \left[\partial_t s + \frac{1}{2}(\nabla s)^2 - \left(\frac{u^2}{2} + \Omega\right)\right] dx \, . \tag{45}$$

The related Hamiltonian (which is the mean energy of the diffusion process per unit of mass) reads

$$\mathcal{H} \doteq \int \rho \left[\frac{1}{2}(\nabla s)^2 - \left(\frac{u^2}{2} + \Omega\right)\right] dx \, , \tag{46}$$

i.e.

$$\mathcal{H} = \tfrac{1}{2}\left(\langle v^2 \rangle - \langle u^2 \rangle\right) - \langle \Omega \rangle \, .$$

We can evaluate an expectation value of Eq. (44) which implies an identity $\mathcal{H} = -\langle \partial_t s \rangle$. By invoking the Smoluchowski diffusion and thus Eq. (24), with the time-independent V, we arrive at

$$\dot{\Psi} = \frac{k_{\rm B}T}{\rho}\nabla(v\rho) \, , \tag{47}$$

whose expectation value $\langle \dot{\Psi} \rangle$, in view of $v\rho = 0$ at the integration volume boundaries, identically vanishes. Since $v = -(1/m\beta)\nabla\Psi$, we define

$$s(x,t) \doteq \frac{1}{m\beta}\,\Psi(x,t) \Longrightarrow \langle \partial_t s \rangle = 0 \tag{48}$$

so that $\mathcal{H} \equiv 0$ identically.

We have thus arrived at the following interplay between the mean energy, localization and the information entropy gain

$$\frac{D}{2}\left(\frac{dS}{dt}\right)_{\text{gain}} = \int \rho \left(\frac{\overrightarrow{v}^2}{2}\right) dx = \int \rho \left(\frac{\overrightarrow{u}^2}{2} + \Omega\right) dx \geq 0\,, \tag{49}$$

generally valid for Smoluchowski processes with non-vanishing diffusion currents.

By recalling the notion of the Fisher information Eq. (32) and setting $\mathcal{F} \doteq D^2\mathcal{F}_\alpha$, we can rewrite the above formula as follows:

$$\mathcal{F} = \langle v^2 \rangle - 2\langle \Omega \rangle \geq 0\,, \tag{50}$$

where $\mathcal{F}/2 = -\langle Q \rangle > 0$ holds true for probability densities with finite mean and variance.

We may evaluate directly the localization/uncertainty dynamics of the Smoluchowski process, by recalling that the Fisher information $\mathcal{F}/2$ is the localization measure, which for probability densities with finite mean value and variance σ^2 is bounded from below by $1/\sigma^2$.

Namely, by exploiting the hydrodynamical conservation laws Eq. (41) for the Smoluchowski process we get

$$\partial_t(\rho v^2) = -\nabla\left[(\rho v^3)\right] - 2\rho v\,\nabla(Q - \Omega)\,. \tag{51}$$

We assume to have secured conditions allowing to take a derivative under an indefinite integral, and assume that of ρv^3 vanishes at the integration volume boundaries. This implies the following expression for the time derivative of $\left\langle v^2 \right\rangle$

$$\frac{d}{dt}\left\langle v^2 \right\rangle = 2\left\langle v\,\nabla(\Omega - Q)\right\rangle\,. \tag{52}$$

Proceeding in the same vein, in view of $\dot{\Omega} = 0$, we find that

$$\frac{d}{dt}\langle \Omega \rangle = \langle v\,\nabla\Omega \rangle \tag{53}$$

and so the equation of motion for $\mathcal{F}$ follows

$$\frac{d}{dt}\mathcal{F} = \frac{d}{dt}\left[\langle v^2 \rangle - 2\langle \Omega \rangle\right] = -2\langle v\,\nabla Q \rangle\,. \tag{54}$$

Since we have $\nabla Q = \nabla P/\rho$ where $P = D^2 \rho\, \Delta \ln \rho$, the previous equation takes the form $\dot{\mathcal{F}} = -\int \rho v \nabla Q dx = -\int v \nabla P dx$, which is an analog of the familiar expression for the power release ($dE/dt = F \cdot v$, with $F = -\nabla V$) in classical mechanics.

This should be compared with our previous discussion of the "heat dissipation" term. Indeed, $\dot{\mathcal{F}} = \int j\,(-2\nabla Q) dx$, while the expression for the heat dissipation rate had the form $k_{\mathrm{B}} T \dot{\mathcal{Q}} = \int j\,(-\nabla V) dx$.

Let us notice that $\dot{\mathcal{F}} > 0$ would tell us that the localization improves, clearly at the expense of the energy supply (power injection) from the environment. $\dot{\mathcal{F}} < 0$ indicates a localization decay and corresponds to the energy absorption (power release) by the environment.

We may typically expect the decrease of the localization measure $\mathcal{F}$ and the continual energy/heat absorption by the Smoluchowski diffusion process. This effect can be attributed to the active role of the thermal environment which generally leads to a delocalization of the initially localized probability density, unless the invariant measures enter the game. The power release complies with the identity $\mathcal{H} \equiv 0$ since "obviously" the diffusion process proceeds in an open system. The latter property should be contrasted with the behavior of so-called finite energy diffusions, [18, 24, 32].

The paper has been supported by the Polish Ministry of Scientific Research and Information Technology under the grant No. PBZ-MIN-008/P03/2003.

REFERENCES

[1] C.E. Shannon, *Bell Syst. Techn. J.* **27**, 379 (1948); *Bell Syst. Techn. J.* **27**, 623 (1948).

[2] K. Sobczyk, *Mechanical Systems and Signal Processing* **15**, 475 (2001).

[3] A. Lasota, M.C. Mackey, *Chaos, Fractals and Noise*, Springer-Verlag, Berlin 1994.

[4] H. Risken, *The Fokker–Planck Equation*, Springer-Verlag, Berlin 1989.

[5] D-Q. Jiang, M. Qian, M-P. Qian, *Mathematical Theory of Nonequilibrium Steady Ststes*, LNM **1833**, Springer-Verlag, Berlin 2004.

[6] A. Eberle, *Uniqueness and Non-uniqueness of Semigroups Generated by Singular Diffusion Operators*, LNM **1718**, Springer-Verlag, Berlin 2000.

[7] H. Qian, M. Qian, X. Tang, *J. Stat. Phys.* **107**, 1129 (2002).

[8] K. Huang, *Statistical Mechanics*, Wiley, New York 1963.

[9] *Dynamics of Dissipation*, LNP **597**, Eds. P. Garbaczewski and R. Olkiewicz, Springer-Verlag, Berlin 2002.

[10] S. Kullback, *Information Theory and Statistics*, Wiley, NY 1959.

[11] P. Gaspard, *Chaos, Scattering and Statistical Mechanics*, Cambridge University Press, Cambridge 1998.

[12] P. Gaspard, G. Nicolis, J.R. Dorfman, *Physica A* **323**, 294 (2003).

[13] L. Rondoni, E.G.D. Cohen, *Physica D* **168–169**, 341 (2002).

[14] L. Andrey, *Phys. Lett.* **A111**, 45 (1985).

[15] A.R. Plastino, A. Daffertshofer, *Phys. Rev. Lett.* **93**, 138701 (2004).

[16] D. Daems, G. Nicolis, *Phys. Rev.* **E59**, 4000 (1999).

[17] M.C. Mackey, *Rev. Mod. Phys.* **61**, 981 (1989).

[18] E. Nelson, *Dynamical Theories of the Brownian Motion*, Princeton University Press, Princeton 1967.

[19] H. Qian, *Phys. Rev.* **E65**, 016102 (2001).

[20] T. Hatano, S. Sasa, *Phys. Rev. Lett.* **86**, 3463 (2001).

[21] J.M.G. Vilar, J.M. Rubi, *Proc. Nat. Acad. Sci.* **98**, 11081 (2001).

[22] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, Princeton 1946.

[23] B.T. Geilikman, *Zh. Eksp. Teor. Fiz.* **17**, 830 (1947).

[24] P. Garbaczewski, *Phys. Rev.* **E59**, 1498 (1999).

[25] P. Garbaczewski, *Physica A* **285**, 187 (2000).

[26] R. Czopnik, P. Garbaczewski, *Physica A* **317**, 449 (2003).

[27] G.A. Skorobogatov, *Rus. J. Phys. Chem.* **61**, 509 (1987).

[28] M.J.W. Hall, M. Reginatto, *J. Phys. A: Math. Gen.* **35**, 3289 (2002).

[29] P. Garbaczewski, R. Olkiewicz, *J. Math. Phys.* **37**, 732 (1996).

[30] D. Ruelle, *J. Stat. Phys.* **85**, 1 (1996).

[31] A.J. Stam, *Inf. and Control* **2**, 101 (1959).

[32] P. Garbaczewski, *Differential Entropy and the Dynamics of Uncertainty*, `quant-ph/0408192`.