Vol. 37 (2006) *ACTA PHYSICA POLONICA B* No 5

ENTROPY METHODS IN RANDOM MOTION*

Piotr Garbaczewski

Institute of Physics, University of Zielona Góra
Prof. Szafrana 4a, 65-516 Zielona Góra, Poland

(Received February 20, 2006)

Dedicated to Professor Peter Talkner on the occasion of his 60th birthday

We analyze a contrasting dynamical behavior of Gibbs–Shannon and conditional Kullback-Leibler entropies, induced by time-evolution of continuous probability distributions. The question of predominantly *purpose-dependent* entropy definition for non-equilibrium model systems is addressed. The conditional Kullback–Leibler entropy is often believed to properly capture physical features of an asymptotic approach towards equilibrium. We give arguments in favor of the usefulness of the standard Gibbs-type entropy and indicate that its dynamics gives an insight into physically relevant, but generally ignored in the literature, non-equilibrium phenomena. The role of physical units in the Gibbs–Shannon entropy definition is discussed.

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

1. Introduction

There are many notions of entropy. Except for the Clausius (thermodynamic) entropy, none of them may be considered unambiguously defined or to share the status of a physically universal quantity in the class of dynamical systems and phenomena, to the description of which a particular entropy notion has been possibly designed.

Let us reproduce the standard (albeit non-exhaustive) list of entropies. For classical dynamical systems one is tempted to use any of: Boltzmann, Gibbs, Shannon, Kullback–Leibler, Renyi, Tsallis, information/differential, topological, measure-theoretic and Kolmogorov–Sinai entropies. In the quantum case one encounters von Neumann, Wehrl and Leipnik entropies, plus more or less natural/obvious generalizations of, classical by provenance, Kullback–Leibler, Renyi and Tsallis entropies. The concrete entropy choice

* Presented at the XVIII Marian Smoluchowski Symposium on Statistical Physics, Zakopane, Poland, September 3–6, 2005.

is with no doubt the context (classical or quantum setting, specific model system, specific notion of state, microstate and macrostate) and purpose-dependent.

We shall follow associations born by non-equilibrium statistical physics phenomena, where in the time-dependent problems one encounters such issues like "trends" (convergence or divergence) towards stationary states plus Boltzmann-type theorems (temporal behavior of H-functionals), validity, limitations, possible violations, general rules of entropy evolution, meaning of the entropy "production"/dissipation and its temporal behavior.

The term *entropy methods* essentially refers to the mathematically rigorous discussion of the asymptotic (large time) behavior of solutions of various partial differential equations, in particular to these governing the dynamics of probability densities. One attempts to quantify the speed of con(div)ergence of measures that allow to differentiate among different solutions and their possibly different temporal properties.

To set the stage to the main theme of our considerations, let us invoke the simplest (naive) version of the Boltzmann H-theorem, valid in case of the rarified gas (mass m particles), without external forces, close to its thermal equilibrium, under an assumption of its space homogeneity, [1, 2].

If the probability density function $f(v)$ is a solution of the corresponding Boltzmann kinetic equation, then the Boltzmann H-function (which coincides with the negative of the Gibbs–Shannon entropy) $H(t) = \int f(v) \ln f(v) dv$ does not increase:

$$\frac{d}{dt} H(t) \leq 0 \,. \tag{1}$$

In particular, we know that there exists an invariant (asymptotic) density $f_*(v) \simeq \exp[-m(v-v_0)^2/2k_{\mathrm{B}}T]$ and $H(t)$ is a constant only if $f \doteq f_*(v)$.

Notice that in the one-dimensional case, the $L^1(R)$ density normalization coefficient reads $(m/2\pi\, k_{\mathrm{B}}T)^{1/2}$ and thence, *formally*, $H_* = \int f_* \ln f_* dv = -(1/2)\ln(2\pi\, e k_{\mathrm{B}}T/m)$ where e is the base of the natural logarithm. One must be aware of an apparent dimensional difficulty, [3], since an argument of the logarithm is not dimensionless.

Clearly, a consistent integration outcome for $H(t)$ should involve a dimensionless argument $k_{\mathrm{B}}T/m[v]^2$ instead of $k_{\mathrm{B}}T/m$, provided $[v]$ stands for *any* unit of velocity. Examples are $[v] = 1\,\mathrm{m/s}$ (here m stands for the SI length unit, and not for a mass parameter) or $10^{-5}\,\mathrm{m/s}$. To this end, it suffices to redefine H_* as follows, [3, 4]:

$$H_* \to H_*^{[v]} = \int f_* \ln([v]\, f_*) dv \,. \tag{2}$$

Multiplying f_* by $[v]$ we arrive at the dimensionless argument of the logarithm in the above.

We shall come back later to a deeper discussion of an impact of dimensional units on the general definition of the Gibbs–Shannon entropy

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

for a probability density function $\rho \in L^1(R^n)$.

The entropy methods basically refer to the large time asymptotic of the heat and Fokker–Planck equations, where in a mathematically oriented research all dimensional units, for the sake of clarity, are scaled away. Following [5], let us consider the heat equation in the re-scaled (no physical constants) form: $\partial_t u = \Delta u$ with $x \in R^n$, $t \in R^+$ and $u(., t = 0) = u_0(.) \geq 0$, $\int u_0(x) dx = 1$.

As $t \to \infty$, for any $u(x,t)$ we have $u(x,t) \simeq \rho(x,t) = (4\pi t)^{-n/2} \exp\left[-x^2/4t\right]$ in conformity with the standard wisdom [7] that a regular solution of the heat equation behaves asymptotically as a fundamental solution, once time goes to infinity.

There is a natural question to be addressed: what is the $t \to \infty$ rate of convergence of the so-called Kullback "distance"

$$\|u - \rho\|_{L^1}(t) \doteq \int |u(x,t) - \rho(x,t)|\, dx \tag{4}$$

between two densities. Since, for two density functions ρ and ρ' there holds the Csiszár–Kullback inequality, [6]:

$$\int \rho \ln \left(\frac{\rho}{\rho'}\right) dx \geq \frac{1}{2} \|\rho - \rho'\|^2_{L^1} \tag{5}$$

it is the Kullback–Leibler entropy

$$\mathcal{K}(\rho, \rho') \doteq \int \rho(x) \, \ln \frac{\rho(x)}{\rho'(x)}\, dx\,, \tag{6}$$

which actually stands for an upper bound upon a "distance measure" in the set of density functions.

If we consider ρ_t to be a solution of the heat equation with the initial data ρ_0 and take $\rho_\alpha(x) = (1/\sqrt{2\alpha\pi}) \exp[-x^2/2\alpha]$, then we may always find α and k such that $\rho_{\alpha+kt}$ has the same second moment as ρ_t. This implies an asymptotic $1/t$ decay of the initially prescribed Kullback–Leibler "distance", [5],

$$\mathcal{K}(\rho_t, \rho_{\alpha+kt}) \leq \mathcal{K}(\rho_0, \rho_\alpha) \left[\frac{\alpha}{(\alpha + kt)}\right]. \tag{7}$$

In view of the concavity of the function $f(w) = -w \ln w$, the Kullback–Leibler entropy is positive. This property if often contrasted with the fact the Gibbs–Shannon entropy $\mathcal{S}(\rho)$ may take negative values. Therefore, right at this point (anticipating further discussion) we introduce the *conditional* Kullback–Leibler entropy notion, which although non-positive by construction:

$$\mathcal{H}_{\rm c}(\rho, \rho') \doteq -\mathcal{K}(\rho, \rho') \tag{8}$$

is nonetheless one of the major tools in the study of an asymptotic convergence towards an invariant (equilibrium) density, [8, 9]. This entropy typically displays a prototype behavior (monotonic growth in time), expected to hold true if the entropy definition is to be compatible with the casual understanding of the second law of thermodynamics, [9].

Now, let us consider the drifted Fokker–Planck (Smoluchowski) equation $\partial_t f = \Delta f - \nabla (bf)$, where $f(.,t) = f_0 \geq 0$, $\int f_0(x) dx = 1$. We assume that the forward drift $b = b(x,t)$ has a gradient form. Let f_* be the stationary solution of the F–P equation, then an obvious question is: what is the $t \to \infty$ rate of convergence of $\|f - f_*\|_{L^1}(t) \doteq \int |f(x,t) - f_*(x)|\, dx$ towards the value 0?

The outcome, albeit not completely general, is that $f_t \doteq f(x,t)$, $t \geq 0$ decays *in relative entropy* to a Gaussian, the speed of such decay being exponential, [6]. This is typically encoded in the formula, [6,8,9] of the form

$$\mathcal{H}_{\rm c}(t) \simeq \exp(-\alpha t)\, \mathcal{H}_{\rm c}(0)\,, \tag{9}$$

where $\mathcal{H}_{\rm c}(t) \doteq \mathcal{H}_{\rm c}(f_t, f_*)$, with $\alpha > 0$ and $t > 0$. See also an explicit discussion of the Ornstein–Uhlenbeck process in [10].

In the course of the time evolution, the conditional entropy monotonically approaches its maximum at zero, [9]. This property is seldom shared by the Gibbs–Shannon entropy of the involved time-dependent probability density. The Gibbs entropy may grow, diminish, oscillate and show more complicated patterns of behavior, [9–11]. A physical relevance of such "strange" temporal properties, compare *e.g.* Eq. (1), is worth addressing and it is our main goal in the present paper.

2. Gibbs–Shannon and Kullback–Leibler entropies

A casual understanding of the entropy notion in physics is that entropy (tacitly one presumes to deal with its thermodynamic Clausius version) is a measure of the degree of randomness and the tendency (trend) of physical systems to become less and less organized. We attribute a very concrete meaning to the term *organization* — namely, we are interested in quantifying how good is the probability *localization* on the state space (whatever: configuration space, velocity or phase-space) of the system.

As a hint let us consider a probability measure $\mu = (\mu_1, \mu_2, \ldots, \mu_N)$ on a system of N points, *e.g.* $\sum_{j=1}^{N} \mu_j = 1$. The standard Shannon entropy reads $S(\mu) = -\sum_{j=1}^{N} \mu_j \log \mu_j \implies 0 \leq S(\mu) \leq \log N$ and its maximum corresponds to a uniform probability distribution $\mu_j = 1/N$ for all j.

If X is a discrete random variable taking values x_i with probabilities p_i, $i = 1, 2, \ldots, N$, the quantity $\mathcal{S}(X) = -\sum p_i \log p_i$ is called the Shannon entropy of a discrete random variable or the entropy of the probability distribution $(p_1, \ldots, p_N)$. If X takes infinitely many values $x_1, x_2, \ldots$ with probabilities $p_1, p_2, \ldots$, then the entropy $\mathcal{S}(X)$ is not necessarily finite.

As a side comment we recall that log has base 2 in which case the unit of entropy is called a bit (binary digit), while for ln with base e, the unit of entropy is called a nat (natural); we observe that $\log b \ \ln 2 = \ln b$.

For a continuous random variable X with values in $x \in R^n$ and the probability density $\rho(x)$ one usually defines the Shannon entropy of a continuous random variable (called the differential entropy of X) as:

$$\mathcal{S}(X) = -\int_{\Gamma} \rho(x) \log \rho(x) dx \,,$$

where $\Gamma \in R^n$ is the support set of X. One may also denote $\mathcal{S}(X) \doteq \mathcal{S}(\rho)$.

There is number of standard views about the discrete and continuous entropies. In the discrete case, the entropy quantifies randomness in an *absolute* way. In the continuous case there is no smooth limiting passage from the discrete to continuous entropy. Then, the entropy cannot work "as it is" as a measure of global randomness and one usually invokes a casual list of drawbacks: $\mathcal{S}(\rho)$ may be negative, may be unbounded both from below and above, is scaling (hence coordinate transformation) dependent.

Anyway, a difference of two Shannon entropies, necessarily evaluated with respect to the same coordinate system, $\mathcal{S}(\rho) - \mathcal{S}(\rho')$ is known to quantify an absolute change in the information/randomness content when passing from ρ to ρ' and is obviously scaling independent. The same observation extends to the time derivative of the Shannon entropy in case of time-dependent probability densities.

Alternatively, although with reservations, one may pass to the familiar notion of the Kullback–Leibler entropy

$$\mathcal{K} = \int_{\Gamma} \rho \left(\ln \rho - \ln \rho' \right) dx \,,$$

non-negative and scaling-independent from the outset. However, one should keep in mind that it is the conditional Kullback–Leibler (K–L) entropy $\mathcal{H}_{\mathrm{c}} = -\mathcal{K}$, Eq. (8), which is often used in the literature in connection with the "entropy growth paradigm", [8]. The conditional K–L entropy takes only negative values and its upper bound actually equals zero, while $\mathcal{S}(\rho)$ may take positive or negative value depending on the particular choice of ρ.

Let us point out that a consistent exploitation of the conditional K–L entropy is restricted either to the large time-scale phenomena, see *e.g.* Eq. (7), or to the dynamical systems which have an invariant density, see Eq. (9). In the short time-scale regimes and for systems without invariant densities, the conditional Kullback–Leibler entropy is not an adequate tool.

Let us consider

$$\rho_{\alpha,\beta} = \beta\,\rho[\beta(x-\alpha)]\,, \tag{10}$$

where $\alpha \geq 0, \beta > 0$ are real parameters. The respective Shannon entropy reads:

$$\mathcal{S}(\rho_{\alpha,\beta}) = \mathcal{S}(\rho) - \ln\beta\,. \tag{11}$$

For general probability distributions $\rho(x)$ with a fixed variance σ we have $S(\rho) \leq 1/2\ln(2\pi e\sigma^2)$ and $S(\rho)$ becomes maximized if and only if ρ is a Gaussian. Therefore, we can write

$$(2\pi e)^{-1/2}\,\exp[\mathcal{S}(\rho_{\alpha,\beta})] \leq \frac{\sigma}{\beta}\,, \tag{12}$$

and give a meaning to the β-scaling transformation of $\rho(x-\alpha)$: the density is broadened if $\beta < 1$ and shrinks if $\beta > 1$.

Given a one parameter family of Gaussian densities $\rho_\alpha = \rho(x-\alpha)$, with the mean $\alpha \in R$ and the standard deviation fixed at σ. These densities share the very same value of Shannon entropy, independent of α:

$$\mathcal{S}_\sigma = \frac{1}{2}\ln\left(2\pi e\sigma^2\right)\,.$$

If we admit the standard deviation σ to be another free parameter, a two-parameter family $\rho_\alpha \to \rho_{\alpha,\sigma}(x)$ appears. Then

$$\mathcal{S}_{\sigma'} - \mathcal{S}_\sigma = \ln\left(\frac{\sigma'}{\sigma}\right)\,.$$

By denoting $\sigma \doteq \sigma(t) = \sqrt{2Dt}$ and $\sigma' \doteq \sigma(t')$ we make the non-stationary (heat kernel) density amenable to the "absolute comparison" formula at different time instants $t' > t > 0$: $(\sigma'/\sigma) = \sqrt{t'/t}$.

Indeed a fundamental solution of the heat equation $\partial_t\rho = D\Delta\rho$ reads

$$\rho(x,t) = \frac{1}{(4\pi Dt)^{1/2}}\exp\left(-\frac{x^2}{4Dt}\right)\,, \tag{13}$$

whose differential entropy equals $\mathcal{S}(t) = (1/2)\ln(4\pi eDt)$, or in the dimensionless form: $\mathcal{S}^{[x]}(t) = (1/2)\ln(4\pi eDt/[x]^2)$, where $[x]$ is any dimensional unit with the SI dimension of length.

Let ρ_υ denote a convolution of a probability density ρ with a Gaussian probability density having variance υ. The transition density (heat kernel) of the Wiener process generates such a convolution for any $\rho_0(x)$, with $\upsilon = \sigma^2 \doteq 2Dt$. Then, (de Bruijn) we have the entropy accumulation formula:

$$\frac{d\mathcal{S}}{dt} = D\,\mathcal{F} = D\int \frac{(\nabla\rho)^2}{\rho}dx > 0\,.$$

The monotonic growth of $\mathcal{S}(t)$ is paralleled by linear in time growth of the standard deviation $\sigma(t)$, hence quantifies the uncertainty (disorder) increase related to the "flattening" down of ρ.

Let us consider the Kullback entropy $\mathcal{K}(\theta,\theta')$ for a family of probability densities ρ_θ labeled by a parameter (one or more) θ, so that the "distance" between any two densities in this family can be directly evaluated. We take $\rho_{\theta'}$ as reference probability density. Then:

$$\mathcal{K}(\theta,\theta') \doteq \mathcal{K}(\rho_\theta|\rho_{\theta'}) = \int \rho_\theta(x)\,\ln\frac{\rho_\theta(x)}{\rho_{\theta'}(x)}\,dx\,. \tag{14}$$

It is particularly instructive to evaluate various K–L "distances" among members of a two-parameter family of $L^1(R)$-normalized Gaussian functions, labeled by independent parameters $\theta_1=\alpha$ and $\theta_2=\sigma$ (alternatively $\theta_2=\sigma^2$) such that $\theta \doteq (\theta_1,\theta_2)$. In the self-explanatory notation, for two different θ and θ' Gaussian densities there holds:

$$\mathcal{K}(\theta,\theta') = \ln\frac{\sigma'}{\sigma} + \frac{1}{2}\left(\frac{\sigma^2}{\sigma'^2} - 1\right) + \frac{1}{2\sigma'^2}(\alpha-\alpha')^2\,. \tag{15}$$

We may assume that θ' very little deviates from θ: $\theta' = \theta + \Delta\theta$. Then, we have

$$\mathcal{K}(\theta,\theta+\Delta\theta) \simeq \frac{1}{2}\sum_{i,j}\mathcal{F}_{ij}\;\Delta\theta_i\Delta\theta_j\,, \tag{16}$$

where $i,j, = 1,2$ and the Fisher information matrix $\mathcal{F}_{ij}$ has the form:

$$\mathcal{F}_{ij} = \int \rho_\theta \frac{\partial\ln\rho_\theta}{\partial\theta_i}\,\frac{\partial\ln\rho_\theta}{\partial\theta_j}\,dx\,. \tag{17}$$

In case of Gaussian densities, labeled by independent $\theta_1 = \alpha$, $\theta_2 = \sigma$, (or $\theta_2 = \sigma^2$) the Fisher matrix is diagonal.

Let us set $\alpha' = \alpha$ and consider $\sigma^2 = 2Dt$, $\Delta(\sigma^2) = 2D\Delta t$. Then $\mathcal{S}(\sigma'^2) - \mathcal{S}(\sigma^2) \simeq \Delta t/2t$, while $\mathcal{K}(\theta, \theta') \simeq (\Delta t)^2/4t^2$. Although, for finite increments Δt we have

$$\mathcal{S}(\sigma'^2) - \mathcal{S}(\sigma^2) \simeq \sqrt{\mathcal{K}(\theta, \theta')} \simeq \frac{\Delta t}{2t}\,,$$

the time derivative notion $\dot{\mathcal{S}}$ surely can be defined for the differential entropy, but is definitely meaningless in terms of the corresponding short time-scale Kullback "distance", *cf.* [10, 11].

We stress that no such obstacle arises in the standard cautious use of the conditional Kullback entropy $\mathcal{H}_c$, when an invariant density is in hands. Indeed, normally one of the involved densities is the stationary (reference) one $\rho_{\theta'}(x) \doteq \rho_*(x)$, while another is allowed to evolve in time $\rho_\theta(x) \doteq \rho(x,t)$, $t \in R^+$, thence $\mathcal{H}_c(t) \doteq -\mathcal{K}(\rho_t|\rho_*)$ and $d\mathcal{H}_c(t)/dt$ does make sense.

We recall that for the free Brownian motion there is no invariant density. As we have indicated before, Eq. (7), $\mathcal{H}_c(\rho_t, \rho_{t'})$, $t < t'$ still remains a useful tool, albeit in the asymptotic regime and for not too small values of $t' - t$.

3. Physical units in the entropy definition

Let us come back to an issue of physical units in the definition of a differential entropy. In fact, if x and p stand for one-dimensional phase space labels and $f(x,p)$ is a normalized phase-space density, $\int f(x,p)dxdp = 1$, then the related *dimensionless* differential entropy reads as follows, [4]:

$$\mathcal{S}_h = -\int (hf)\ln(hf)\frac{dxdp}{h} = -\int f\ln(hf)\,dxdp\,, \tag{18}$$

where $h = 2\pi\hbar$ is the tentatively accepted (there is no other mention of quantum theory) Planck constant. Let $\rho(x)$ and $\tilde{\rho}_h(p)$ be two independent, respectively, spatial and momentum space densities. We form the joint density

$$f(x,p) \doteq \rho(x)\tilde{\rho}_h(p) \tag{19}$$

and evaluate the differential entropy $\mathcal{S}_h$ for this density. Remembering that $\int \rho(x)dx = 1 = \int \tilde{\rho}_h(p)dp$, we have formally

$$\mathcal{S}_h = -\int \rho\ln\rho dx - \int \tilde{\rho}_h \ln\tilde{\rho}_h\,dp - \ln h = S^x + S^p - \ln h\,. \tag{20}$$

The formal use of the logarithm properties before executing integrations in $\int \tilde{\rho}_h \ln(h\tilde{\rho}_h)\,dp$, has left us with an issue of "literally taking the logarithm of a dimensional argument" *i.e.* that of $\ln h$.

We recall that $\mathcal{S}_h$ is a dimensionless quantity, while if x has dimensions of length, then the probability density has dimensions of inverse length and analogously in connection with momentum dimensions.

Let us denote $x \doteq r\delta x$ and $p \doteq \tilde{r}\delta p$ where labels r and $\tilde{r}$ are dimensionless, while δx and δp stand for respective position and momentum dimensional (hitherto — resolution) units. Then

$$-\int \rho \ln \rho dx - \ln(\delta x) \doteq -\int \rho \ln(\delta x \rho) dx \tag{21}$$

is a dimensionless quantity. Analogously

$$-\int \tilde{\rho}_h \ln \tilde{\rho}_h \, dp - \ln \delta p \doteq -\int \tilde{\rho}_h \ln(\delta p \tilde{\rho}_h) \, dp \tag{22}$$

is dimensionless. First left-hand side terms in two above equations we recognize as S^x and S^p, respectively.

Hence, formally we have arrived at a manifestly dimensionless decomposition

$$\begin{aligned} \mathcal{S}_h &= -\int \rho \ln(\delta x \rho) dx - \int \tilde{\rho}_h \ln(\delta p \tilde{\rho}_h) \, dp + \ln \frac{\delta x \delta p}{h} \\ &\doteq S^x_{\delta x} + S^p_{\delta p} + \ln \frac{\delta x \delta p}{h} \,, \end{aligned} \tag{23}$$

instead of the previous one, Eq. (20). The last identity Eq. (23) gives an unambiguous meaning to the preceding formal manipulations with dimensional quantities. Instead of the Planck constant h we can use any other unit with SI dimensions of action, say δh.

As a byproduct of our discussion, we have resolved the case of the spatially interpreted real axis, when x has dimensions of length, *cf.* also [4]: $S^x_{\delta x} = -\int \rho \ln(\delta x \rho) dx$ is the pertinent dimensionless differential entropy definition for spatial probability densities.

Example 1: Let us discuss an explicit example involving the Gauss density

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-x_0)^2}{2\sigma^2}\right] , \tag{24}$$

where σ is the standard deviation (its square stands for the variance). There holds $\mathcal{S}(\rho) = 1/2 \ln\left(2\pi e \sigma^2\right)$ which is a dimensionless outcome. If we pass to x with dimensions of length, then inevitably σ must have dimensions of length. It is instructive to check that in this dimensional case we have a correct dimensionless result:

$$S^x_{\delta x} = \frac{1}{2} \ln\left[2\pi e \left(\frac{\sigma}{\delta x}\right)^2\right] \tag{25}$$

to be compared with Eq. (21). Clearly, $S^x_{\delta x}$ vanishes if $\sigma/\delta x = (2\pi e)^{-1/2}$, hence at the dimensional value of the standard deviation $\sigma = (2\pi e)^{-1/2}\delta x$, compare *e.g.* [4].

Example 2: In the Introduction we have discussed the simplest version of the Boltzmann H-theorem, where a suitable probability density function $f(v,t)$ determines temporal properties of the Boltzmann H-functional such that $H(t) = \int f(v) \ln f(v) dv$ does not increase, $\frac{d}{dt}H(t) \leq 0$. An invariant (asymptotic) density, in the one-dimensional case, has the form $f_*(v) = (m/2\pi k_{\mathrm{B}}T)^{1/2} \exp[-m(v - v_0)^2/2k_{\mathrm{B}}T]$. Therefore, $H_* = \int f_* \ln f_* dv = -(1/2)\ln(2\pi\, e k_{\mathrm{B}} T/m)$ and we are faced with a dimensional difficulty, [3]. A consistent integration outcome for $H(t)$ should involve $k_{\mathrm{B}}T/m[v]^2$ instead of $k_{\mathrm{B}}T/m$, provided $[v]$ stands for *any* unit of velocity. The redefinition $H_* \to H_*^{[v]} = \int f_* \ln([v]\, f_*) dv$, cures the dimensional obstacle.

We recall that under the scaling transformation Eq. (10) the respective Shannon entropy takes the form $\mathcal{S}(\rho_{\alpha,\beta}) = \mathcal{S}(\rho) - \ln\beta$. In case of Gaussian ρ, we get $\mathcal{S}(\rho_{\alpha,\beta}) = \ln[(\sigma/\beta)\sqrt{2\pi e}]$. Clearly, $\mathcal{S}(\rho_{\alpha,\beta})$ takes the value 0 at $\sigma = (2\pi e)^{-1/2}\beta$ in analogy with our previous dimensional considerations. If an argument of ρ is assumed to have dimensions, then the scaling transformation with the dimensional β may be interpreted as a method to restore the dimensionless differential entropy value.

4. Temporal behavior of entropies

4.1. Deterministic system

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

$$\dot{x} = f(x)\,, \tag{26}$$

where $\dot{x}$ stands for the time derivative and f is an R^n-valued function of $x \in R^n$, $x = \{x_1, x_2, \ldots, x_n\}$. A 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

$$\partial_t \rho = -\nabla\,(f\,\rho)\,, \tag{27}$$

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

With a continuous probability density $\rho \doteq \rho(x,t)$, where $x \in R^n$ and we allow for an explicit time-dependence, we associate a respective differential entropy functional $\mathcal{S}(\rho)$, where 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:

$$\dot{\mathcal{S}} = \int \rho\,(\mathrm{div}\, f) dx \doteq \langle \nabla\, f \rangle \,. \tag{28}$$

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. 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 $\dot{\mathcal{S}} = -1$. Further discussion of dynamical systems with strange (multifractal) attractors, for which the Shannon information (differential) entropy decreases indefinitely (the pertinent steady states are no longer represented by probability density functions) can be found in [12]. We note that for Hamiltonian systems, the phase-space flow has vanishing divergence, hence $\dot{\mathcal{S}} = 0$ which implies that "information is conserved" in Hamiltonian dynamics.

Let there be given an invertible dynamical system on R^2, with $f(x) \doteq Fx$, where F is a two-by-two real matrix and $x \in R^2$, [9]. A solution has the form $x(t) = \exp(tF)x(0)$, where the matrix operator $\exp(tF)$ is defined through the standard Taylor expansion formula. The solution of the Liouville equation with an initial probability density $f_0(x)$ is given by

$$f(x,t) = \exp[-(\mathrm{Tr}\, F)t]\, f_0(\exp(-tF)x) \,, \tag{29}$$

and hence

$$\mathcal{S}(f_t) = \mathcal{S}(f_0) + (\mathrm{Tr}\, F)t \Rightarrow \dot{\mathcal{S}}(f_t) = \mathrm{Tr}\, F \,. \tag{30}$$

Obviously $\mathrm{Tr}\, F = \lambda_1 + \lambda_2$, where $\lambda_i, i = 1, 2$ are the eigenvalues of F. We realize that $\mathcal{S}(f_t)$ grows indefinitely if $\mathrm{Tr}\, F > 0$ and diminishes indefinitely towards $-\infty$ if $\mathrm{Tr}\, F < 0$. There is no stationary density and the conditional entropy is not defined.

4.2. Random system

In case of a general dissipative dynamical system, 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. (26) by white noise term $A(t)$ where $\langle A_i(s) \rangle = 0$ and $\langle A_i(s) A_j(s') \rangle = \sqrt{2q}\delta(s - s')\delta_{ij}$, $i = 1, 2, \ldots\, n$, implies the Fokker–Planck–Kramers equation:

$$\partial_t \rho = -\nabla\, (f\, \rho) + q\Delta\rho \,, \tag{31}$$

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. (28):

$$\dot{\mathcal{S}} = \int \rho\,(\mathrm{div}\, f)dx + q \int \frac{1}{\rho}(\nabla\rho)^2\, dx\,. \tag{32}$$

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 $\dot{\mathcal{S}} = 0$. If $\langle \nabla\, f\rangle \geq 0$, then the differential (information) entropy would grow monotonically.

We shall discuss an example of a non-invertible system, provided by the standard one-dimensional Ornstein–Uhlenbeck process, [8, 10]. We choose the forward drift of the Fokker–Planck equation $\partial_t\rho = D\triangle\rho + \nabla[(\gamma x)\rho]$ with $\gamma > 0$ and $D > 0$ being the diffusion coefficient.

If an initial density is chosen in the Gaussian form, with the mean value α_0 and variance σ_0^2, the Fokker–Planck evolution preserves the Gaussian form of $\rho(x,t)$ while modifying the mean value $\alpha(t) = \alpha_0 \exp(-\gamma t)$ and variance

$$\sigma^2(t) = \sigma_0^2 \exp(-2\gamma t) + \frac{D}{\gamma}[1 - \exp(-2\gamma t)]\,. \tag{33}$$

Accordingly, since a unique invariant density has the form $\rho_* = \sqrt{\gamma/2\pi D}$ $\exp(-\gamma x^2/2D)$ we obtain:

$$\mathcal{H}_{\rm c}(t) = \exp(-2\gamma t)\mathcal{H}_{\rm c}(\rho_0, \rho_*) = -\frac{\gamma\alpha_0^2}{2D}\, \exp(-2\gamma t)\,, \tag{34}$$

i.e. a monotonic growth of the negative-valued conditional Kullback–Leibler entropy towards its maximum at zero:

$$\dot{\mathcal{H}}_c(t) = -2\gamma \exp(-2\gamma t)\, \mathcal{H}_{\rm c}(\rho_0, \rho_*) = \gamma^2 \frac{\alpha_0^2}{D}\, \exp(-2\gamma t) > 0\,. \tag{35}$$

The differential entropy:

$$\mathcal{S}(t) = \frac{1}{2} \ln\,[2\pi e \sigma^2(t)] \tag{36}$$

shows another temporal behavior

$$\dot{\mathcal{S}} = \frac{2\gamma(D - \gamma\sigma_0^2) \exp(-2\gamma t)}{D - (D - \gamma\sigma_0^2) \exp(-2\gamma t)}\,. \tag{37}$$

We observe that if $\sigma_0^2 > D/\gamma$, then $\dot{\mathcal{S}} < 0$, while $\sigma_0^2 < D/\gamma$ implies $\dot{\mathcal{S}} > 0$.

In both cases the behavior of the differential entropy is monotonic, although its growth or decay do critically rely on the choice of σ_0^2. Irrespective of σ_0^2 the asymptotic value of $\mathcal{S}(t)$ as $t \to \infty$ reads $(1/2)\ln[2\pi e(D/\gamma)]$. It is useful to note, that in the special case of $\sigma_0^2 = D/\gamma$ the differential entropy is a constant of motion, while the conditional K–L entropy nonetheless does grow, asymptotically approaching the value zero according to Eq. (35).

Summarizing, we can say that the conditional Kullback–Leibler entropy of the Ornstein–Uhlenbeck process grows monotonically in time, while the temporal behavior of the Gibbs–Shannon (differential) entropy depends on statistical properties (half-width σ_0) of the initial ensemble density. This pattern of temporal behavior appears to be generic to a large class of dynamical systems, [9].

To find out whether there is anything deeper in the above apparent differences in the temporal behavior of the Gibbs–Shannon and Kullback–Leibler entropies associated with the same time-dependent probability density, except for the *a priori* presumed existence of the reference invariant density, let us consider the one-dimensional Fokker–Planck equation for any Smoluchowski process. We assume

$$\partial_t \rho = D \triangle \rho - \nabla(b\rho)\,, \tag{38}$$

with a forward drift $b = b(x,t)$ of the gradient form $b = -\nabla\Phi$ and attribute to a diffusion coefficient D dimensions of $\hbar/2m$ or $k_{\rm B}T/m\beta$.

Furthermore, we introduce the velocity fields: $u(x,t) = D\nabla \ln \rho(x,t)$ and $v(x,t) = b(x,t) - u(x,t)$. The current velocity $v(x,t)$, in view of $\partial_t \rho = -\nabla(v\rho)$ which is an equivalent form of Eq. (38), contributes to the diffusion current $j = v\rho$.

For the differential entropy $\mathcal{S}(t) = -\int \rho(x,t)\, \ln \rho(x,t)\, dx$, while imposing boundary restrictions that $\rho, v\rho, b\rho$ vanish at spatial infinities or finite interval borders, we readily get the entropy balance equation of the form Eq. (32), with the minor modification *i.e.* the replacement of q by D. We are, however, interested in its equivalent form (easily derivable under previously listed boundary restrictions), [10, 11]:

$$D\dot{\mathcal{S}} = \left\langle v^2 \right\rangle - \langle b\, v\rangle\ . \tag{39}$$

Remembering that we deal with the Smoluchowski process, we set (adjusting dimensional constants): $b = (D/k_{\rm B}T)\, F$. Exploiting $j \doteq v\rho$ and demanding $F = -\nabla V$ we infer:

$$\dot{\mathcal{S}} = \frac{1}{D}\, \left\langle v^2 \right\rangle - \dot{\mathcal{Q}}\,, \tag{40}$$

where the first (positive) term on the right-hand side stands for the differential entropy accumulation rate (entropy gain by the system).

The second term contains the $\dot{\mathcal{Q}}$ entry:

$$\dot{\mathcal{Q}} \doteq \frac{1}{k_{\mathrm{B}}T} \int F\, j\, dx = \frac{1}{D} \langle b\, v \rangle \ , \tag{41}$$

which, if positive ($\dot{\mathcal{Q}} > 0$ is not a must, [10]), allows to interpret $-\dot{\mathcal{Q}}$ as the entropy dissipation rate, *i.e.* an entropy transfer to the environment in the form of the surplus heat. Note that $k_{\mathrm{B}}T\dot{\mathcal{Q}} = \int F\, j\, dx$ has a conspicuous form of the fairly standard power release expression *i.e.* the time rate at which the mechanical work per unit of mass is returned back to the thermal reservoir (or absorbed if $\dot{\mathcal{Q}} < 0$) in the form of heat.

Under current premises, there exists a stationary solution of the Fokker–Planck equation

$$\rho_*(x) = \frac{1}{Z} \exp\left(-\frac{V(x)}{k_{\mathrm{B}}T}\right) , \tag{42}$$

where $Z = \int \exp(-V(x)/k_{\mathrm{B}}T)\, dx$.

Let us take $\rho_*(x)$ as a reference density with respect to which the divergence of $\rho(x,t)$ is quantified in terms of the conditional K–L entropy. Then

$$\mathcal{H}_{\mathrm{c}}(t) = -\int \rho\, \ln\left(\frac{\rho}{\rho_*}\right)\, dx = \mathcal{S}(t) - \ln Z - \frac{\langle V \rangle}{k_{\mathrm{B}}T}\,, \tag{43}$$

and straightforwardly, because of

$$\frac{d}{dt}\langle V \rangle = -k_{\mathrm{B}}T\dot{\mathcal{Q}} \tag{44}$$

we arrive at

$$\dot{\mathcal{H}}_{\mathrm{c}} = \dot{\mathcal{S}} + \dot{\mathcal{Q}} \geq 0\,. \tag{45}$$

At this point, we can come back to a continued discussion of the Ornstein–Uhlenbeck process. Namely, we have here a direct control of the behavior of the "power release" expression $\dot{\mathcal{Q}} = \dot{\mathcal{H}}_{\mathrm{c}} - \dot{\mathcal{S}}$. Since

$$\dot{\mathcal{H}}_{\mathrm{c}} = \frac{\gamma^2\alpha_0^2}{D}\, \exp(-2\gamma t) > 0\,, \tag{46}$$

in case of $\dot{\mathcal{S}} < 0$ we encounter a continual power supply $\dot{\mathcal{Q}} > 0$ by the thermal environment (alternatively, power absorption by the system).

In case of $\dot{\mathcal{S}} > 0$ the situation is more complicated. For example, if $\alpha_0 = 0$, we can easily check that $\dot{\mathcal{Q}} < 0$, *i.e.* we have the power drainage from the environment for all $t \in R^+$. More generally, the sign of $\dot{\mathcal{Q}}$ is negative for $\alpha_0^2 < 2(D - \gamma\sigma_0^2)/\gamma$. If the latter inequality is reversed, the sign of $\dot{\mathcal{Q}}$ is not uniquely specified and suffers a change at a suitable time instant $t_{\mathrm{change}}(\alpha_0^2, \sigma_0^2)$.

Interestingly enough, in the special case of $\sigma_0^2 = D/\gamma$ *i.e.* $\dot{\mathcal{S}} = 0$, we encounter

$$\dot{\mathcal{H}}_c = \dot{\mathcal{Q}} \geq 0\,, \tag{47}$$

i.e. a direct connection between the entropy increase and heat removal (to the thermostat) *time rates*, which counterbalance each other.

4.3. Phase-space dynamics

One may argue that the reported above, rather unexpected, insight into the nontrivial power transfer processes is an artifact of the one-dimensional spatial (Smoluchowski) projection of the phase-space motion. Let us, therefore, indicate arguments to the contrary.

For Hamiltonian systems the phase-space flow is divergence-less. Indeed, let us consider a two-dimensional conservative system $\dot{x} = p/m$ and $\dot{p} = -\nabla V$ where $H = p^2/2m + V(x)$. Obviously, $\mathrm{div} f = 0$ which implies $\dot{\mathcal{S}} = 0$. In particular this extends to the standard harmonic oscillator with $V(x) = (m\omega^2/2)x^2$.

For the harmonic oscillator with friction, $\dot{x} = v$, $\dot{x} = -(\gamma/m)v - (\omega^2/m)x$, we can adopt the observations of Subsection 3.1 with the two-by-two matrix F, whose first row contains only zeroes, while $(F)_{21} = -\omega^2/m$, $(F)_{22} = -\gamma/m$. Consequently $\mathrm{Tr}\, F = -\gamma/m$.

A solution of the corresponding Liouville-type equation was discussed in Subsection 4.1. The Gibbs–Shannon entropy evolves in time according to Eq. (30): $\mathcal{S}(t) = \mathcal{S}(0) - (\gamma t)/m$ and $\mathcal{S} \to -\infty$ as $t \to \infty$. Since $\gamma > 0$, we have $\dot{\mathcal{S}} = -\gamma/m < 0$. There is no stationary density and hence no $\mathcal{H}_c(t)$.

An admixture of noise in the velocity/momentum rate equation in the damped harmonic oscillator case allows for the existence of a stationary density. Let us consider, [8, 9], an example of the noisy damped harmonic oscillator: $\dot{x} = p/m$, $\dot{p} = -(\gamma/m)p - (\omega^2/m)p + \xi(t)$ where the white noise term ξ is normalized as follows $\langle \xi(t) \rangle = 0$, $\langle \xi(t)\xi(t') \rangle = \sigma\delta(t-t')$. The corresponding Fokker–Planck–Kramers equation for the probability density $f(x,v)$, with $v = p/m$ is

$$\frac{\partial f}{\partial t} = -\frac{\partial (vf)}{\partial x} + \frac{1}{m}\frac{\partial [(\gamma v + \omega^2 x)f]}{\partial v} + \frac{\sigma^2}{2m^2}\frac{\partial^2 f}{\partial v^2}\,, \tag{48}$$

and has a unique stationary solution:

$$f_*(x,v) = \frac{\gamma\omega\sqrt{m}}{\pi\sigma^2} \exp\left[-\frac{\gamma}{\sigma^2}\left(\omega^2 x^2 + m v^2\right)\right]\,. \tag{49}$$

A detailed, in part computer-assisted, analysis of the temporal behavior of Gibbs–Shannon and conditional K–L entropies evaluated for density solutions of the above Kramers equation, with the initial data

$$f_0(x,v) = \frac{1}{2\pi\sigma_x^2\sigma_v^2} \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{v^2}{2\sigma_v^2}\right) \tag{50}$$

has been made in Ref. [9]. We shall summarize the outcomes of this investigation.

In three basic regimes: overdamped $\gamma^2 > 4\omega^2$, critical $\gamma^2 = 4\omega^2$ and underdamped $\gamma^2 < 4\omega^2$ cases, the conditional Kullback–Leibler entropy quantifies an approach of $f(x,v,t)$ towards $f_*(x,v)$ in terms of the monotonic growth pattern (this statement includes also the case of $\dot{\mathcal{H}}_{\mathrm{c}}(t) = 0$).

The situation is entirely different, if we consider the Gibbs–Shannon entropy of $f(x,v,t)$. Let us denote $\sigma_* = \sigma^2/2\gamma\omega^2$ and $\alpha_x = \sigma_x^2 - \sigma_*$, $\alpha_v = \sigma_v^2 - \omega^2\sigma_*$. The behavior of $\mathcal{S}(t)$ sensitively depends on the mutual relations (signs, vanishing or non-vanishing of any or both *etc.*) between α_x and α_v and all details can be found in Ref. [9].

In the overdamped and critical cases, five independent temporal behaviors are admitted. First three are of the monotonic type, since $\dot{\mathcal{S}}$ is vanishing, positive or negative. The fourth one admits a change of sign of $\dot{\mathcal{S}}$ at certain $t_0 > 0$ from positive to negative plus the same scenario in reverse. The fifth temporal scenario shows a passage through $\dot{\mathcal{S}}$-positive, negative and again positive stages of evolution plus the reverse (negative, positive, negative) option.

The underdamped case shows even more intriguing patterns of temporal behavior. Namely, in addition to the monotonic negative or positive signs of $\dot{\mathcal{S}}$ we have also a conspicuous damped oscillation of $\mathcal{S}(t)$, where $\dot{\mathcal{S}}$ changes sign indefinitely, but an amplitude of oscillations performed by $\mathcal{S}(t)$ continually diminishes.

All these *diverse* temporal patterns are special for the Gibbs–Shannon entropy. They are in turn accompanied by a *unique* pattern of the strictly monotonic growth (or none) $\dot{\mathcal{H}}_{\mathrm{c}}(t) \geq 0$ which is displayed by the conditional Kullback–Leibler entropy, [9].

In close analogy with our considerations pertaining to the nontrivial power transfers between an open dynamical system and its thermal environment, *cf.* Subsection 4.2, let us notice that the invariant density Eq. (49) has the form analogous to this of ρ_*, Eq. (42). Indeed, we have:

$$f_*(x,v) = \frac{1}{Z} \exp\left[-\frac{2\gamma}{\sigma^2} E_{\mathrm{cl}}(p,x)\right] \tag{51}$$

with $1/Z = (\gamma\omega\sqrt{m})/(\pi\sigma^2)$ and $E_{\mathrm{cl}}(p,x) = p^2/2m + V(x)$ with $V(x) = \omega^2 x^2/2$ is an energy of a classical harmonic oscillator at the $(x, p = mv)$ phase-space point.

Accordingly, we have

$$\mathcal{H}_c(t) = -\int f \, \ln\left(\frac{f}{f_*}\right) dxdv = \mathcal{S}(t) - \ln Z - \frac{2\gamma}{\sigma^2}\langle\, E_{\rm cl}\rangle\,, \tag{52}$$

where $\mathcal{S}(t) = -\int f \ln f dxdv$. Therefore, it is an intrinsic property of our dynamical system that $\dot{\mathcal{H}} = \dot{\mathcal{S}} + \dot{\mathcal{Q}} \geq 0$, where we define

$$\frac{d}{dt}\langle E_{\rm cl}\rangle \doteq -\frac{\sigma^2}{2\gamma}\dot{\mathcal{Q}} \tag{53}$$

and clearly, $\dot{\mathcal{Q}}$ is the direct analogue of the previously introduced power/heat transfer rate in the mean, *cf.* Eqs. (41) and (44).

Remark: Let us add that in the study of *quantum* open systems weakly coupled to thermal reservoirs, the heat bath is known to drive an open system to its equilibrium state at the same temperature, [13,14]. In terms of quantum mean values, evaluated by means of reduced density operators, the analogues of the first and second laws of thermodynamics were derived. They stay in close affinity with our "thermodynamical" formulas (40), (45) and (53), here obtained in the purely classical (*e.g.* non-quantum) framework. Another viewpoint on the uses of entropy methods in quantum theory, in connection with closed quantum systems, can be found in Refs. [11, 15].

5. Conclusions

Standard notions of thermodynamical entropy are basically used under equilibrium or near-equilibrium conditions. The primary built-in concept is an equilibrium (steady) state and the behavior of entropy in the time domain is seldom addressed.

If one attempts to analyze a dynamics of an approach towards the prescribed steady state, it is necessary to pass to the time domain where the non-equilibrium and often rapid dynamical processes take place. Various notions of entropy may be designed to quantify such non-equilibrium phenomena.

Our analysis of simple diffusion-type models indicates that the very notion of entropy, except perhaps for the standard Clausius thermodynamical entropy, is non-universal and purpose-dependent. In particular, the conditional Kullback–Leibler entropy is regarded (in reference to the "purpose") to be the only valid entropy growth justification in terms of model systems, [8, 9], (that in conformity with the standard interpretation of the second law of thermodynamics for closed systems).

However, a deeper insight into the underlying physical phenomena (power/heat transfer processes in the mean) is available only through the differential (Gibbs–Shannon) entropy, whose temporal behavior is generically inconsistent with the "entropy growth" pattern. Moreover, the Gibbs–Shannon entropy balance equation contains the conditional Kullback–Leibler entropy time rate as an explicit non-negative "entropy production" or rather "entropy accumulation" term, see *e.g.* Subsections 4.2 and 4.3. The entropy dissipation may proceed through the previously mentioned mean power transfer mechanism, however, the involved "heat transfer" expression $\dot{\mathcal{Q}}$ is not necessarily positive-definite.

The conditional Kullback–Leibler entropy is an appropriate tool in case of "slow" processes, and in the asymptotic (large) time regime. The Gibbs–Shannon (differential, information) entropy is perfectly suited for the "shortest description length analysis", in particular for the study of rapid changes in time of the probability distribution involved.

The paper has been supported by the Polish Ministry of Science and Information Society Technologies under the (solicited) grant No PBZ-MIN-008/P03/2003.

REFERENCES

[1] K. Huang, *Statistical Mechanics*, Wiley, New York 1987.

[2] C. Cercignani, *Theory and Application of the Boltzmann Equation*, Scottish Academic Press, Edinburgh 1975.

[3] L. Brillouin, *Science and Information Theory*, Academic Press, New York 1962.

[4] R.S. Ingarden, A. Kossakowski, M. Ohya, *Information Dynamics and Open Systems*, Kluwer, Dordrecht 1997.

[5] G. Toscani, *Rend. di Matematica*, Serie VII **16**, 329 (1996).

[6] A. Arnold *et al.*, *Comm. Partial Diff. Equations* **26**, 43 (2001).

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

[8] M.C. Mackey, M. Tyran-Kamińska, `cond-mat/0501092`.

[9] M.C. Mackey, M. Tyran-Kamińska, `cond-mat/0509649`.

[10] P. Garbaczewski, *Phys. Lett.* **A341**, 33 (2005).

[11] P. Garbaczewski, *Entropy* **7[4]**, 253 (2005).

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

[13] R. Alicki, *J. Phys. A: Math. Gen.* **12**, L103 (1979).

[14] R. Alicki, Invitation to Quantum Dynamical Semigroups, in *Dynamics of Dissipation*, Eds. P. Garbaczewski and R. Olkiewicz, LNP vol. 597, pp. 239–264, Springer, Berlin 2002.

[15] P. Garbaczewski, *Phys. Rev.* **A72**, 056101 (2005).