Journal of Statistical Physics, Vol. 123, No. 2, April 2006 (© 2006)
DOI: 10.1007/s10955-006-9058-2

Differential Entropy and Dynamics of Uncertainty

Piotr Garbaczewski[1]

Received April 22, 2005; accepted September 6, 2005

We analyze the functioning of Gibbs-type entropy functionals in the time domain, with emphasis on Shannon and Kullback-Leibler entropies of time-dependent continuous probability distributions. The Shannon entropy validity is extended to probability distributions inferred from $L^2(R^n)$ quantum wave packets. In contrast to the von Neumann entropy which simply vanishes on pure states, the differential entropy quantifies the degree of probability (de)localization and its time development. The associated dynamics of the Fisher information functional quantifies nontrivial power transfer processes in the mean, both in dissipative and quantum mechanical cases.

KEY WORDS: Entropy functionals, entropy methods, Shannon entropy, Fisher information, dynamics of densities, entropy dynamics, Smoluchowski process, quantum evolution, information localization, uncertainty.

PACS NUMBERS: 05.45.+b, 02.50.-r, 03.65.Ta, 03.67.-a

1. INTRODUCTION

In a number of manifestations of the concept of entropy in physics and mathematics, information-theory based entropy methods were devised to investigate the large time behavior of solutions for various, mostly dissipative, partial differential equations. One obvious physical motivation, e.g. the Boltzmann H-theorem for a dilute gas, stands merely for a particular illustration of the general trend (convergence) towards equilibrium issue for solutions of a concrete (here, Boltzmann) equation.

There are many notions of entropy, like: Clausius thermodynamic, Boltzmann, Gibbs, Shannon, Kullback-Leibler, Renyi, Tsallis, von Neumann, Wehrl, Leipnik, information entropy, differential entropy, topological, measure-theoretic,

---

[1] Institute of Physics, University of Zielona Góra, 65-516 Zielona Góra, Poland.

Kolmogorov-Sinai entropies (we do not discriminate between classical and quantum regimes)). The possibility of having different entropies is not harmful, as long as one unambiguously sets them in correspondence with different physical situations and remembers that a concrete entropy choice is context and purpose-dependent, see Refs. (44, 61) for a deeper discussion of related intricacies.

To avoid any possible confusion, let us explain that in the present paper, we shall not directly invoke nor use the concept of a standard thermodynamical entropy for classical or quantum many-body systems, in equilibrium or out of equilibrium. We are interested in the general information-theory background for the entropy notion and use the word "entropy" in a wider, not necessarily thermodynamic sense. In particular, the standard thermodynamical entropy is regarded as a function on phase-space, while the differential entropy is a function of a probability distribution on whatever space.

We often refer to an explicit information-theoretic lore. Shannon, Kullback and von Neumann entropies are typical information theory tools which were designed to quantify the "information content". For quantum systems with vanishing von Neumann entropy (in pure states) one presumes a "complete information" about the state. For pure state, the differential entropy gives access to another "information level", associated with a probability distribution inferred from a $L^2(R^n)$ wave packet. It is perfectly suited to give account of the Schrödinger picture dynamics of quantum wave packets. In view of its role in the formulation of entropic indeterminacy relations, the term "information" is used in the paper in the technical sense, meaning the inverse of "uncertainty" and should not be taken literally.

In physical sciences, entropy is typically regarded as a measure of the degree of randomness and the tendency (trends) of physical systems to become less and less "organized". For "less organized" systems the notion of entropy becomes synonymous with the measure of uncertainty. We attribute a specific meaning to the term "organization". Namely, we are interested in quantifying the degree of the probability distribution "complexity" and "(de)localization" on the state space (needs to be specified) of the system.

The paper is structured as follows. We begin by recalling the standard lore of the Shannon information theory to attribute an unambiguous meaning to two principal notions, this of *information* and that of *uncertainty*. To this end various notions of "state" of a model system are invoked and suitable information measures are discussed.

Next we turn to the coarse-graining issue and set a connection between the Shannon entropy of a discrete probability measure and the differential entropy of a related (through a suitable limiting procedure) continuous probability density. We discuss various entropic inequalities for both differential and coarse-grained entropies of quantum mechanical densities.

In Section 3, the localization level of probability densities is analyzed by means of so-called entropy powers and of the Fisher information measure. We infer two chain inequalities, Eqs. (28) and (29), which imply that typically the differential entropy is a well behaved quantity, bounded both from below and above. The formalism is general enough to include quantum mechanical densities as merely the special case.

In Section 4 we set a conceptual framework for time-dependent problems. Since classical dynamical, stochastic and quantum systems (in their pure states) in general give rise to time-dependent probability densities and information entropies, we resolve the exemplary density dynamics in terms of Smoluchowski diffusion processes, albeit with no explicit random path (e.g. random variable) input.

The entropy and Fisher information evolution equations are established. Close links of the differential and conditional Kullback entropies are established for Smoluchowski diffusion processes, when asymptotic invariant densities enter the scene. We discuss a compliance of the induced continual power release in the course of the diffusion process with the mean energy conservation law, Eqs. (77) and (81).

In section 5 we analyze differential entropy dynamics and time evolution of the Fisher localization measure in quantum theory and next exemplify the general formalism for simple analytically solvable cases. The emergent continual power transfer effect has been analyzed in connection with the finite energy constraint for the mean energy of quantum motion, Eqs. (87) and (90).

Although uncertainty dynamics scenarios of Sections 4 and 5 are fundamentally different, nonetheless the respective methodologies appear to have an overlap, when restricted to steady states which support invariant densities for (reversible) stationary diffusion-type processes.

2. ENTROPY: INFORMATION AND UNCERTAINTY

Notions of entropy, information and uncertainty are intertwined and cannot be sharply differentiated. While entropy and uncertainty are—to some extent synonymous—"measures of ignorance" (lack of information, uncertainty), the complementary notion of information basically quantifies the *ability* of observers to make reliable predictions about the system(86,88,103): the more aware one is about chances of a concrete outcome, the lower is the uncertainty of this outcome. Normally, the growth of uncertainty is identified with an increase of entropy which in turn is interpreted as an information loss.

Following Ref. (99) let us recall that entropy—be it thermodynamical, von Neumann, Shannon, or any other conceivable candidate—has an exceptional status among physical quantities. As a *derived* quantity it does not show up in any fundamental equation of motion, and is surely *not* a constraint upon the quantum

(Schrödinger) dynamics per se. As an obvious remnant of the standard thermodynamical reasoning, one expects entropy to be a "state function" of the system (thermodynamical notions of equilibrium or near-equilibrium are implicit).

This *state* connotation is a source of ambiguities, since inequivalent notions of "state" are used in the description of physical systems, be them classical, thermodynamical and quantum.(61) Not to mention rather specialized meaning of "states" employed in the standard information theory.(17,86,103)

A primitive information-theory system is simply a *bit* whose two admissible states are binary digits 1 and 0. Its quantum equivalent is a *qubit* whose admissible states are vectors in a two-dimensional Hilbert space, hence an infinity of pure states of a two-level quantum system.

But, the information theory framework, when extended to more complicated systems, employs a plethora of notions of "state".(86,103) As very special cases we may mention a phase-space point as the determinative of the state of a classical dynamical system, or the macroscopic notion of a thermodynamical state in its classical and quantum versions.(61,99)

The notion of entropy when adopted to quantum systems appears to be purpose-dependent.(2,77,99) A symbolic mathematical representation of quantum states in terms of wave vectors and/or density operators is expected to provide an experimentally verifiable "information" about the system. To obtain a catalogue of the corresponding statistical predictions, an a priori choice of suitable observables (and thus measurement procedures) is necessary. Then, a casual interpretation of entropy as a measure of one's *uncertainty* about measurable properties of a system in a prescribed quantum state may acquire an unambiguous meaning.

For a given density operator $\hat{\rho}$, von Neumann entropy

$$\mathcal{S}(\hat{\rho}) = -k_B\, Tr(\hat{\rho} \ln \hat{\rho}) \tag{1}$$

is commonly accepted as a reliable measure of the departure from purity and often regarded as defining a notion of information content, to be experimentally extracted from of a quantum system in a given state. Since von Neumann entropy is insensitive to unitary transformations, the result exhibits an invariance under the change of the Hilbert space basis and the conservation in time for a closed system (when there is no information/energy exchange with the environment). Thus, Schrödinger dynamics has no impact on the von Neumann encoding of information, see e.g. also Refs. (55, 90) for a related discussion.

Pure states have vanishing von Neumann entropy ($\mathcal{S}(\hat{\rho}) = 0$ "for the pure states and only for them"(99)) and are normally considered as irrelevant from the quantum information theory perspective, since "one has complete information"(99) about such states. One may even say that a pure state is an unjustified over-idealization, since otherwise it would constitute e.g. a completely measured state of a system in an infinite Hilbert space.(1,78)

Note that as a side comment we find in Ref. (78) a minor excuse: "this idealization, often employed for position-momentum degrees of freedom, is usually an adequate approximation". This can be read as an answer to an objection of Ref. (33): "although continuous observables such as the position are familiar enough, they are really unphysical idealizations", c.f. also Ref. (40).

On the other hand, the position-momentum issue can be addressed in terms of the classic Shannon entropy which is known to be a natural measure of the amount of uncertainty related to measurements for pairs of observables, discrete and continuous on an equal footing, when a quantum system actually is in a *pure* state. The related entropic versions of Heisenberg-type indeterminacy (uncertainty) relations for finite and infinite quantum systems have received due attention in the literature.(7)–(51,78)

Obviously, there is no use of Shannon entropy if one is interested in verifying for mixed quantum states, how much actually a given state is mixed. On the other hand, von Neumann entropy appears to be useless in the analysis of $L^2(R)$ wave packets and their dynamical manifestations (time-dependent analysis) which are currently in the reach of experimental techniques.(100,104) If the state vector is regarded as an information (alternatively, predictions and uncertainty) resource, then questions like(23): "how much information in the state vector" and "information about what", may be considered meaningful.

Given the probability density $\rho(x)$ on R^n, we define the differential entropy,(27,88) as follows:

$$\mathcal{S}(\rho) = -\int \rho(x) \ln \rho(x) \, dx \, . \tag{2}$$

One may consider a subset $\Gamma \subset R^n$ to be a support of ρ instead of R; this is guaranteed by the convention that the integrand in Eq. (2) vanishes if ρ does. Note a minor but crucial notational difference between $\hat{\rho}$ and ρ.

We emphasize that in the quantum mechanical context, we shall invoke either position $\mathcal{S}(\rho)$ or momentum $\mathcal{S}(\tilde{\rho})$ information entropies, with no recourse to the "classical entropy" given in terms of classical phase-space distributions $f(q, p)$ or their Wigner/Husimi analogues.(47,99)

The notion of entropic uncertainty relations,(11,12,63,77) explicitly relies on the differential entropy input. Namely, an arithmetic sum of (presumed to be finite) momentum and position information entropies for any normalized $L^2(R^n)$ wave packet $\psi(x)$, is bounded from below:

$$\mathcal{S}(\rho) + \mathcal{S}(\tilde{\rho}) \geq n(1 + \ln \pi) \tag{3}$$

where n stands for the configuration space (respectively momentum space) dimension.(12) This feature is worth emphasizing, since neither $\mathcal{S}(\rho)$ nor $\mathcal{S}(\tilde{\rho})$ on their own are bounded from below or from above. Nonetheless, both take finite values in physically relevant situations and their sum is always positive.

Since a normalized wave function ψ represents a pure state of a quantum system whose dynamics is governed by the Schrödinger equation, only for stationary states the differential entropy $\mathcal{S}(\rho)$ is for sure a conserved quantity. In general, the Schrödinger picture evolution of $\psi(x,t)$ and so this of $|\psi(x,t)|^2 \doteq \rho(x,t)$ may give rise to a nontrivial dynamics of the information entropy associated with the wave packet $\psi(x,t)$.

Let us point out that most of the "entropic" research pertains to time-independent situations, like in case of stationary solutions of the Schrödinger equation. Notable exceptions are Refs. (67, 68, 84). On general non-quantum grounds an information (differential entropy) dynamics is addressed in Refs. (66, 88) and (30)–(50), see also Refs. (44, 57, 74, 80, 81, 83).

The differential entropy, by a number of reasons,(86,88) cannot quantify the "absolute amount of information carried by the state of the system" (Shannon's uncertainty), unless carefully interpreted. Up to measure preserving coordinate transformations the latter objection remains invalid.

This feature gave impetus to numerically assisted comparative studies of the Shannon information content of different pure states of a given quantum system and to the study of quantum system "complexity". Results are ranging from simple atoms to molecules, nuclei, aggregates of particles, many-body Bose and Fermi systems, and Bose-Einstein condensates, see e.g. Refs. (7)–(38). In these cases, Shannon's differential entropy appears to be a fully adequate measure for the complexity and localization degree of the involved stationary wave packets.

A difference of two information entropies (evaluated with respect to the same coordinate system) $\mathcal{S}(\rho) - \mathcal{S}(\rho')$ undoubtedly quantifies an absolute change in the information content when passing from one state of a given system to another. Alternatively, to this end one may invoke the familiar notion of the relative Kullback entropy $-\int_\Gamma \rho\,(\ln\rho - \ln\rho')\,dx$,(86,88) provided ρ' is strictly positive. Cogent recommendations towards the use of the Shannon information measure, plainly against the Kullback option, can be found in Ref. (93). We shall come to this point later. For arguments just to the opposite see e.g. Ref. (87).

In the present paper, we predominantly use the differential entropy. In Section 4 we shall describe a number of limitations upon the use of the Kullback entropy. When both entropies can be safely used (we discuss selected models of diffusion type dynamics with asymptotic invariant densities), we establish direct links between the Shannon and Kullback entropy dynamics.

In the context of the induced (by time development of probability densities) "information dynamics" $\mathcal{S} \to \mathcal{S}(t)$,(88,92) it is the difference $\mathcal{S}(t) - \mathcal{S}(t')$ between the (presumed to be finite) information entropy values for the time-dependent state of the same physical system, considered at times $t' < t$, which properly captures the net uncertainty/information change in the respective time interval $[t', t]$. In particular, the rate in time of information entropy $\frac{d\mathcal{S}}{dt}$ is a well defined quantity characterizing the temporal changes (none, gain or loss) in the information content of a given $L^2(R^n)$ normalized wave packet $\psi(x, t)$ (strictly speaking, of the related probability density).

Information entropy and its intrinsic dynamics, like e.g. the information flow and information entropy production rate, quantify properties of general reversible and/or irreversible dynamical systems. Normally, the microscopic dynamics of such systems is expected to follow well defined trajectories (deterministic paths of a dynamical system or sample paths of a stochastic process) and those may be thought to induce a corresponding dynamics for statistical ensembles of trajectories.

It is seldom possible to have a sharp wisdom of the initial data $x_0 \in X$ for the trajectory dynamics taking place in a phase space X of the system. This imprecision extends to the terminal data ($x_0 \rightarrow x_t$ after time $t > 0$) as well. Therefore, even if one knows exact dynamical rules governing the behavior of individual trajectories in time, it is basically impossible to tell more about the system then: if its initial state can be found in a subset $A \subset X$ with a probability $prob(x_0 \in A)$, then after time t one can identify the terminal state of the system $x_t \in X$ in a subset $B \subset X$ with a probability $prob(x_t \in B)$. An evolution of *derived* probability densities eventually may come as a solution of appropriate partial differential transport equations.(16,30,60,66)

In the present paper we take a more general view and we bypass a concept of the underlying trajectory dynamics by emphasizing the role of transport equations themselves and their density solutions. Under such premises, we can safely address the *dynamics of uncertainty/information* generated by the Schrödinger picture quantum evolution of wave packets in closed (*no* system – reservoir/environment coupling) quantum mechanical systems.

3. DIFFERENTIAL ENTROPY

3.1. Prerequisites

The original definition of Shannon entropy conveys a dual meaning of both uncertainty and information measure. It is useful to interpret those features in a complementary (albeit colloquial) way: the less is the uncertainty of the system or its state, the larger (and more valuable) is the information we acquire as a result of the measurement (observation) upon the system, and in reverse.

We know that a result of an observation of any random phenomenon cannot be predicted a priori (i.e. before an observation), hence it is natural to quantify an uncertainty of this phenomenon. Let us consider $\mu = (\mu_1, \dots, \mu_N)$ as a probability measure on N distinct (discrete) events A_j, $1 \leq j \leq N$ pertaining to a model system. Assume that $\sum_{j=1}^{N} \mu_j = 1$ and $\mu_j = prob(A_j)$ stands for a probability for an event A_j to occur in the game of chance with N possible outcomes.

Let us call $-\log \mu_j$ an *uncertainty function* of the event A_j. Interestingly, we can coin here the name of the ("missing") *information function,* if we wish to interpret what can be learned via direct observation of the event A_j: the less probable is that event, the more valuable (larger) is the information we would retrieve through its registration.

Then, the expression

$$\mathcal{S}(\mu) = -\sum_{j=1}^{N} \mu_j \log \mu_j \tag{4}$$

stands for the measure of the *mean uncertainty* of the possible outcome of the game,(88) and at the same time quantifies the *mean information* which is accessible from an experiment (i.e. actually playing the game).

Thus, if we identify event values $A_1, \ldots, A_N$ with labels for particular discrete "states" of the system, we may interpret Eq. (4) as a measure of uncertainty of the "state" of the system, *before* this particular "state" it is chosen out of the set of all admissible ones. This well conforms with the standard meaning attributed to the Shannon entropy: it is a measure of the degree of ignorance concerning which possibility (event A_j) may hold true in the set $\{A_1, A_2, \ldots, A_N\}$ with a given a priori probability distribution $\{\mu_1, \ldots, \mu_N\}$.

Notice that:

$$0 \leq \mathcal{S}(\mu) \leq \log N \tag{5}$$

ranges from certainty (one entry whose probability equals 1 and thus no information is missing) to maximum uncertainty when a uniform distribution $\mu_j = 1/N$ for all $1 \leq j \leq N$ occurs. In the latter situation, all events (or measurement outcomes) are equiprobable and $\log N$ sets maximum for a measure of the "missing information".

By looking at all intermediate levels of randomness allowed by the inequalities Eq. (5) we realize that the lower is the Shannon entropy the less information about "states" of the system we are missing, i.e. we have more information about the system.

If the Shannon entropy increases, we actually loose an information available about the system. Consequently, the difference between two uncertainty measures can be interpreted as an information gain or loss.

3.2. Events, States, Microstates and Macrostates

The Boltzmann formula

$$\mathcal{S} = k_B \ln W \doteq -k_B \ln P \tag{6}$$

sets a link of entropy of the (thermodynamical) system with the probability $P = 1/W$ that an appropriate "statistical microstate" can occur. Here, W stands for a number of all possible (equiprobable) microstates that imply the prescribed macroscopic (e.g. thermodynamical) behavior corresponding to a *fixed* value of $\mathcal{S}$.

It is instructive to recall that if P is a probability of an event i.e. of a particular microstate, then $-\ln P$ (actually, with $\log_2$ instead of ln) may be interpreted(48) as "a measure of information produced when one message is chosen from the set, all choices being equally likely" ("message" to be identified with a "microstate"). Another interpretation of $-\ln P$ is that of a degree of uncertainty in the trial experiment.(103)

To get a better insight into the information-uncertainty intertwine, let us consider an ensemble of finite systems which are allowed to appear in any of $N > 0$ distinct elementary states. The meaning of "state" is left unspecified, although an "alphabet" letter may be invoked for convenience.

Let us pick up randomly a large sample composed of $G \gg 1$ single systems, each one in a certain (randomly assigned) state. We record frequencies $n_1/G \doteq p_1, \ldots, n_N/G \doteq p_N$ with which the elementary states of the type $1, \ldots, N$ do actually occur. This sample is a substitute for a "message" or a "statistical microstate" in the previous discussion.

Next, we identify the number of *all* possible samples of that fixed size G which would show up the very same statistics $p_1, \ldots, p_N$ of elementary states. We interpret those samples to display the same "macroscopic behavior".

It was the major discovery due to Boltzmann that the number W of relevant "microscopic states" can be approximately read out from each single sample and is directly related to the the introduced a priori probability measure $\mu_1, \ldots, \mu_N$, with an identification $p_i \doteq \mu_i$ for all $1 \leq i \leq N$, by the formula:

$$\ln W \simeq -G \sum_{i=1}^{N} p_i \ln p_i \doteq -G \cdot \mathcal{S}(\mu) \tag{7}$$

On the basis of this formula, we can consistently introduce $\mathcal{S}(\mu)$ as the *mean information* per each (i-th) elementary state of the N-state system, as encoded in a given sample whose size $G \gg 1$ is sufficiently large.(17,86)

3.3. Shannon Entropy Versus Differential Entropy

3.3.1. Coarse-Graining

For a given probability density function on R we can adopt the coarse-graining procedure,(66) giving account of an imprecision with which a spatial position x can be measured or estimated. Thus, we shall pass from density functions to approximating them piece-wise constant, histogram-type discontinuous functions.

We need to partition the configuration space R into a family of disjoint subsets (intervals) $\{B_k\}$ such that $\cup_k B_k \subseteq R$ and $B_i \cap B_j = \emptyset$ for $i \neq j$. We denote $\mu(B_k) \doteq \mu_k$ the length of the k-th interval, where μ stands for the Lebesgue measure on R.

A probability that a Gaussian random variable with the density ρ takes its value x in an interval B_k equals $prob(B_k) \doteq p_k = \int_{B_k} \rho(x)dx$. An average of the density ρ over B_k we denote $<\rho>_k = p_k/\mu_k$ where $\mu_k = \int_{B_k} dx$.

The probability density ρ coarse grained with respect to the partition $\{B_k\}$ reads:

$$\rho_B(x) \doteq \sum_k <\rho>_k 1_k(x) \tag{8}$$

where $1_k(x)$ is an indicator (characteristic) function of the set B_k, which is equal 1 for $x \in B_k$ and vanishes otherwise. Since $\int 1_k(x)dx = \mu_k$ it is clear that $\int \rho_B(x)dx = \sum_k <\rho>_k \mu_k = \sum_k p_k = 1$ where an interchange of the summation with integration is presumed to be allowed.

We choose a grating unit $\mu_k = r \ll 1$ for all k and notice that $<\rho>_k = p_k/r$ where $p_k \simeq \rho(x_k) \cdot r$ for certain $x_k \in B_k$. In view of the spatial localization properties of the Gauss function, we can safely assert that a finite interval L about a mean value x_0 may be used in the coarse graining procedure, instead of the full configuration space R. Effectively, we arrive at a *finite* partition on L with the resolution $L/G = r$ and then we can safely invoke the definition of $p_k \doteq P_k = r \cdot \rho(x_k)$.

For a coarse grained probability density we introduce a coarse grained Shannon entropy whose relationship to the original differential entropy is of major interest. We have:

$$\mathcal{S}(\rho_B) = -\sum_k p_k \ln p_k \simeq -\sum_k [r\rho(x_k)] \ln r - \sum_k [r\rho(x_k)] \ln[\rho(x_k)] \tag{9}$$

with a standard interpretation of the mean information per bin of length r. Here, if a partition (grating) unit r is small, one arrives at an approximate formula (we admit $|\ln r| \gg 1$):

$$\mathcal{S}(\rho_B) \simeq -\ln r - \int \rho(x) \ln[\rho(x)]dx = -\ln r + \mathcal{S}(\rho) \tag{10}$$

with the obvious proviso that $\mathcal{S}(\rho_B) \geq 0$ and $\mathcal{S}(\rho) \geq \ln r$.

In view of Eq. (10), as long as we keep in memory the strictly positive grating unit r, there is a well defined "regularization" procedure (add $-\ln r$ to $\mathcal{S}(\rho)$) which relates the coarse grained entropy with a given differential entropy. In a number of cases it is computationally simpler to evaluate the differential entropy, and then to extract—otherwise computationally intractable—coarse grained entropy.

Notice that one cannot allow a naive zero grating limit in Eq. (10), although r may be arbitrarily small. The smaller is the positive r, the better is an approximation of the differential entropy by the second term on the right-hand-side of Eq. (10).

We have inequalities $0 \leq \mathcal{S}(\rho_B) \leq G$ where $L = G \cdot r$. They extend to all approximately equal entries in Eq. (10). Since $-\ln r = -\ln L + \ln G$, we arrive at new inequalities:

$$\ln r \leq -\sum_k [r\rho(x_k)] \ln[\rho(x_k)] \leq \ln L \tag{11}$$

where $\sum_k [r\rho(x_k)] \ln[\rho(x_k)] \Rightarrow -\int \rho \ln \rho \, dx$ with $r \to 0$ and possibly $L \to \infty$. A conclusion is that the differential entropy is unbounded both form below and from the above. In particular, $\mathcal{S}(\rho)$ may take arbitrarily low negative values, in plain contrast to its coarse grained version $\mathcal{S}(\rho_B)$ which is always nonnegative.

Let us recall that the value of a convex function $x \ln x$ at the mean value of its argument $\langle x \rangle$, does not exceed the mean value $\langle x \ln x \rangle$ of the function itself. Then, in our notation which follows Eq. (12), we can directly employ an averaging over B_k:

$$\frac{1}{r}\int_{B_k} \rho \ln \rho dx \geq \left(\frac{1}{r}\int_{B_k} \rho dx\right)\left[\ln\left(\frac{1}{r}\int_{B_k} \rho dx\right)\right]. \tag{12}$$

Taking the minus sign, executing summations with respect to k (convergence of the series being presumed) and using Eqs. (15) and (16) we get:

$$\mathcal{S}(\rho) - \ln r \leq \mathcal{S}(\rho_B) \tag{13}$$

as a complement to Eq. (17), see e.g. also Refs. (11, 84). The difference of two coarse grained entropies, corresponding to the same partition but to different (coarse grained) densities, may be adequately approximated by the difference of the corresponding differential entropies:

$$\mathcal{S}(\rho_B) - \mathcal{S}(\rho'_B) \simeq \mathcal{S}(\rho) - \mathcal{S}(\rho'), \tag{14}$$

provided they take finite values(84,88):

3.3.2. Spatial Coarse Graining in Quantum Mechanics

The coarse grained entropy attributes the "mean information per bin of length r" to systems described by continuous probability densities and their differential entropies. Effectively one has a tool which allows to accompany the coarse grained density histogram (of p_k in the k-th bin on R) by the related histogram of uncertainties $-\ln p_k$, c.f. Section II.A where an uncertainty function has been introduced.

The archetypal example of position measurement in quantum mechanics presumes that position is measured in bins corresponding to the resolution of the measurement apparatus. This means that the continuous spectrum of the position observable is partitioned into a countable set of intervals (bins) whose maximum length we regard as a "resolution unit". For an interval $B_k \subset R$ we may denote p_k the probability of finding the outcome of a position measurement to have a value in B_k. We are free to set the bin size arbitrarily, especially if computer assisted procedures are employed.(73)

The formula Eq. (10) gives meaning to the intertwine of the differential and coarse grained entropies in the quantum mechanical context. When an analytic form of the entropy is in the reach, the coarse graining is straightforward. One should realize that most of the results known to date have been obtained numerically, hence with an implicit coarse-graining, although they were interpreted in terms of the differential entropy, see e.g. Refs. (47)–(72).

In connection with the entropic inequality Eq. (3) let us point out (77) that it is a generic property of normalized $L^2(R^n)$ wave functions that, by means of the Fourier transformation, they give rise to two interrelated densities (presently we refer to $L^2(R)$): $\rho = |\psi|^2$ and $\tilde{\rho} = |\mathcal{F}(\psi)|^2$ where

$$(\mathcal{F}\psi)(k) = \frac{1}{\sqrt{2\pi}} \int \psi(x) \exp(-ikx)\, dk \tag{15}$$

is the Fourier transform of $\psi(x)$. The inequality (3) for the corresponding (finite) differential entropies follows, here with $n = 1$.

By choosing arbitrarily small resolutions $r \ll 1$ and $\tilde{r} \ll 1$ we can introduce the respective coarse grained entropies, each fulfilling an inequality Eq. (13). Combining these inequalities with Eq. (3), we get the prototype entropic inequalities for coarse grained entropies:

$$\mathcal{S}(\rho_B) + \mathcal{S}(\tilde{\rho}_B) \geq 1 + \ln \pi - \ln(r \cdot \tilde{r}) \tag{16}$$

with the corresponding resolutions r and $\tilde{r}$.

By referring to Eq. (10) we realize that the knowledge of $\mathcal{S}(\rho_B)$, completely determines $\mathcal{S}(\tilde{\rho}_B)$ at the presumed resolution levels:

$$\mathcal{S}(\tilde{\rho}_B) \simeq 1 + \ln \pi - \ln(r \cdot \tilde{r}) - \mathcal{S}(\rho_B) \geq 0 \tag{17}$$

and in reverse. This in turn implies that in all computer generated position-momentum differential entropy inequalities, where the coarse graining is implicit, the knowledge of position entropy and of the resolution levels provide sufficient data to deduce the combined position-momentum outcomes, see also Refs. (90)–(72).

In standard units (with $\hbar$ reintroduced, see e.g. Ref. (54) for an explanation how to handle dimensional units in the entropy definition), the previous discussion pertains to quantum mechanical position–momentum entropic uncertainty relations. In the notation of Refs. (11, 14) we have:

$$S^x + S^p \geq 1 - \ln 2 - \ln\left(\frac{\delta x \cdot \delta p}{h}\right) \tag{18}$$

for measurement entropies with position and momentum "measuring device" resolutions δx and δp respectively, such that $\delta x \cdot \delta p \ll h$.

Let us point out that one should not confuse the above resolution units $r, \tilde{r}$ and $\delta x, \delta p$ with mean square deviation values which are present in the standard formulation of uncertainty relations. If, following conventions we define the squared standard deviation (i.e. variance) value for an observable A in a pure state ψ as $(\Delta A)^2 = (\psi, [A - \langle A \rangle]^2 \psi)$ with $\langle A \rangle = (\psi, A\psi)$, then for the position X and momentum P operators we have the following version of the entropic uncertainty relation (here expressed through so-called entropy powers, see e.g. Ref. (77)):

$$\Delta X \cdot \Delta P \geq \frac{1}{2\pi e} \exp[\mathcal{S}(\rho) + \mathcal{S}(\tilde{\rho})] \geq \frac{1}{2} \tag{19}$$

which is an alternative for Eq. (3): $n = 1$ and $\hbar \equiv 1$ being implicit.

4. LOCALIZATION: DIFFERENTIAL ENTROPY AND FISHER INFORMATION

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.(86) 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.

Let us consider the Gaussian probability density on the real line R as a reference density function: $\rho(x) = (1/\sigma\sqrt{2\pi}) \exp[-(x - x_0)^2/2\sigma^2]$. The differential entropy of the Gauss density has a simple analytic form, independent of the mean value x_0 and maximizes an inequality $\mathcal{S}(\rho) \leq \frac{1}{2} \ln(2\pi e \sigma^2)$. This imposes a useful bound upon the entropy power $\frac{1}{\sqrt{2\pi e}} \exp[\mathcal{S}(\rho)] \leq \sigma$ with an obvious bearing on the spatial localization of the density ρ, hence spatial (un)certainty of position measurements.

For arbitrary density functions the differential entropy Eq. (2) is unbounded form below and from above. Nonetheless we realize that in the subset of all densities with a finite mean and a fixed variance σ^2, we actually have an upper bound. However, in contrast to coarse grained entropies which are always nonnegative, even for relatively large mean deviation $\sigma < 1/\sqrt{2\pi e} \simeq 0.26$ the differential entropy $\mathcal{S}(\rho)$ is negative.

Therefore, quite apart from the previously discussed direct information theory links, c.f. Eqs. (10), (13) and (14), the major role of the differential entropy is to be a measure of localization in the "state space" (actually, configuration space) of the system.(46,79,95)

Let us consider a one-parameter family of probability densities $\rho_\alpha(x)$ on R where the mean value of ρ_α equals α and we fix at σ^2 the value $\langle (x - \alpha)^2 \rangle = \langle x^2 \rangle - \alpha^2$ of the variance (here, standard deviation) of the probability density ρ_α.

The following inequality holds true:

$$\mathcal{F}_\alpha \doteq \int \frac{1}{\rho_\alpha}\left(\frac{\partial \rho_\alpha}{\partial \alpha}\right)^2 dx \geq \frac{1}{\sigma^2} \tag{20}$$

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.$^{(28,32,45,46,89)}$ In view of $\mathcal{F}_\alpha \geq 1/\sigma^2$, we realize that the Fisher information is more sensitive indicator of the wave packet localization than the entropy power.

Let us define $\rho_\alpha(x) \doteq \rho(x-\alpha)$. Then, the Fisher information $\mathcal{F}_\alpha \doteq \mathcal{F}$ is no longer the mean value α-dependent and can be readily transformed to the conspicuously quantum mechanical form (up to a factor D^2 with $D = \hbar/2m$):

$$\frac{1}{2}\mathcal{F} = \frac{1}{2}\int \frac{1}{\rho}\left(\frac{\partial \rho}{\partial x}\right)^2 dx = \int \rho \cdot \frac{u^2}{2} dx = -\langle Q \rangle \tag{21}$$

where $u \doteq \nabla \ln \rho$ is named an osmotic velocity field$^{(41,75)}$ and an average $\langle Q \rangle = \int \rho \cdot Q dx$ is carried out with respect to the function

$$Q = 2\frac{\Delta \rho^{1/2}}{\rho^{1/2}}\,. \tag{22}$$

As a consequence of Eq. (20), we have $-\langle Q \rangle \geq 1/2\sigma^2$ for all relevant probability densities with *any* finite mean and variance fixed at σ^2.

When multiplied by D^2, the above expression for $Q(x)$ notoriously appears in the hydrodynamical formalism of quantum mechanics as the so-called de Broglie-Bohm quantum potential ($D = \hbar/2m$). It appears as well in the corresponding formalism for diffusion-type processes, including the standard Brownian motion (then, $D = k_B T/m\beta$, see e.g. Refs. (39, 41, 43).

An important inequality, valid under an assumption $\rho_\alpha(x) = \rho(x-\alpha)$, has been proved in Ref. (89), see also Refs. (20, 32):

$$\frac{1}{\sigma^2} \leq (2\pi e)\,\exp[-2\mathcal{S}(\rho)] \leq \mathcal{F} \tag{23}$$

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.

Let us point out that the Fisher information $\mathcal{F}(\rho)$ may blow up to infinity under a number of circumstances$^{(45)}$: 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.

The values of $\mathcal{F}(\rho_\alpha)$ and $\mathcal{S}(\rho_\alpha)$ are α-independent if we consider $\rho_\alpha(x) = \rho(x-\alpha)$. This reflects the translational invariance of the Fisher and Shannon information measures.(37)

Let us investigate an effect of the scaling transformation.(22,86,87) We denote

$$\rho_{\alpha,\beta} = \beta\, \rho[\beta(x-\alpha)] \tag{24}$$

where $\alpha > 0$, $\beta > 0$. The respective Shannon entropy reads:

$$\mathcal{S}(\rho_{\alpha,\beta}) = \mathcal{S}(\rho) - \ln\beta \tag{25}$$

which transforms the entropy power inequality to the form

$$(2\pi e)^{-1/2}\, \exp[\mathcal{S}(\rho_{\alpha,\beta})] \leq \sigma/\beta\,. \tag{26}$$

An obvious interpretation is that the β-scaling transformation of $\rho(x-\alpha)$ would broaden this density if $\beta < 1$ and would shrink when $\beta > 1$.

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(77,89):

$$\mathcal{F} = 4\int \left(\frac{\partial\sqrt{\rho}}{\partial x}\right)^2 dx \leq 16\pi^2\tilde{\sigma}^2\,, \tag{27}$$

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 $\tilde{\rho} \doteq |\tilde{\psi}|^2$ of Eq. (3) whose variance the above $\tilde{\sigma}^2$ actually is.

In view of two previous inequalities (23), (27) we find out that not only the Fisher information, but also an entropy power is bounded from below and above. Namely, we have:

$$\frac{1}{\sigma^2} \leq \mathcal{F} \leq 16\pi^2\tilde{\sigma}^2 \tag{28}$$

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

$$\frac{1}{4\pi\tilde{\sigma}} \leq \frac{1}{\sqrt{2\pi e}}\exp[\mathcal{S}(\rho)] \leq \sigma. \tag{29}$$

The differential entropy $\mathcal{S}(\rho)$ typically may be expected to be a well behaved quantity: with finite both lower and upper bounds.

We find rather interesting that the Heisenberg indeterminacy relationship Eq. (19), which is normally interpreted to set a lower bound on the experimentally accessible *phase-space* data (e.g. volume), according to Eq. (29) ultimately had appeared to give rise to lower and upper bounds upon the *configurational* (spatial) information measure and thence—the uncertainty (information) measure.

5. ENTROPY DYNAMICS FOR DETERMINISTIC AND DIFFUSIVE MOTIONS

5.1. Classical Dynamics

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

$$\dot{x} = f(x) \tag{30}$$

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

$$\partial_t \rho = -\nabla \cdot (f \rho) \tag{31}$$

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 probability density functional $\mathcal{S}(\rho)$, according to Eq. (2). 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$^{(30)}$:

$$\dot{\mathcal{S}} = \int \rho\, (div\, f) dx \doteq \langle \nabla \cdot f \rangle\, . \tag{32}$$

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 $div f = -1$ 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 Ref. (30).

We note that for Hamiltonian systems, the phase-space flow is divergenceless, 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. (32)):

$$\frac{\partial}{\partial t}\rho = -\frac{p}{m}\frac{\partial}{\partial x}\rho + (\nabla V)\frac{\partial}{\partial p}\rho \tag{33}$$

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.

In case of a general dissipative dynamical system (30), 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. (30) 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:

$$\partial_t \rho = -\nabla \cdot (f\rho) + q\Delta\rho \tag{34}$$

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. (32):

$$\dot{\mathcal{S}} = \int \rho\,(div\, f)dx + q\int \frac{1}{\rho}(\nabla\rho)^2\, dx. \tag{35}$$

Now, the dissipative term $\langle\nabla \cdot f\rangle < 0$ 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 $\langle\nabla \cdot f\rangle \geq 0$, the information entropy would grow monotonically.

5.2. Wiener Process

Let us consider the familiar heat kernel$^{(25)}$:

$$\rho(x, t) = \frac{1}{(4\pi Dt)^{1/2}}\exp\left(-\frac{x^2}{4Dt}\right) \tag{36}$$

with the diffusion coefficient D. It is a fundamental solution of the heat equation $\partial_t\rho = D\Delta\rho$ which is the Fokker-Planck equation for the Wiener process.

The differential entropy of the above time-dependent density reads:

$$\mathcal{S}(t) = (1/2)\ln(4\pi e Dt) \tag{37}$$

and its time evolution clearly displays the localization uncertainty growth. By means of the formula Eq. (35) we can quantify the differential entropy dynamics for all solutions of the heat equation.

Since the heat kernel determines the transition probability density for the Wiener process (free Brownian motion in R), by setting $x \to x - x'$ and $t \to t - t' \geq 0$, we can replace the previous $\rho(x, t)$ of Eq. (36) by $p(x - x', t - t')$. This transition density allows to deduce *any* given solution $\rho(x, t)$ of the heat equation from its past data, according to: $\rho(x, t) = \int p(x - x', t - t')\rho(x', t,)dx'$. In particular, we can consider the process starting at $t' = 0$ with any initial density $\rho_0(x)$.

Let ρ_υ denote a convolution of a probability density ρ with a Gaussian probability density having variance υ. The transition density of the Wiener process generates such a convolution for ρ_0, with $\upsilon = \sigma^2 \doteq 2Dt$. Then, de Bruijn identity,(46,89) $d\mathcal{S}(\rho_\upsilon)/d\upsilon = (1/2)\mathcal{F}(\rho_\upsilon)$, directly yields the information entropy time rate for $\mathcal{S}(\rho) = \mathcal{S}(t)$:

$$\frac{d\mathcal{S}}{dt} = D \cdot \mathcal{F} = D \cdot \int \frac{(\nabla\rho)^2}{\rho} dx > 0 . \tag{38}$$

The Fisher information $\mathcal{F}(\rho)$ is the $\alpha = 0$ version of the general definition given in Eqs. (20) and (21). The derivation of Eq. (38) amounts to differentiating an υ-dependent integrand under the sign of an improper integral.(28,45)

The monotonic growth of $\mathcal{S}(t)$ is paralleled by linear in time growth of $\sigma(t)$ and the decay of $\mathcal{F}$, hence quantifies the uncertainty (disorder) increase related to the "flattening" down of ρ, see also Refs. (37, 45). Let us indicate that the entropy time rate is positive, hence there is a definite "entropy accumulation" in the course of the Wiener process.

5.3. Kullback Entropy Versus Differential Entropy

We emphasize that in the present paper we have deliberately avoided the use of the relative Kullback-Leibler entropy.(58,66,88) This entropy notion is often invoked to tell "how far from each other" two probability densities are. In fact, for two density functions f and g one may invoke the Csiszár-Kullback inequality: $\int f \ln(f/g)dx \geq (1/2)\|f - g\|^2_{L^1}$, often interpreted to set an upper bound upon a "distance measure" in the set of density functions

The Kullback entropy is particularly useful if one investigates an approach of the system toward (or its deviation from) equilibrium, this being normally represented by a stationary density function.(60,82) In this context, it is employed to investigate a major issue of the dynamical origins of the increasing entropy, see Refs. (60, 64, 66). Consult also both standard motivations and apparent problems encountered in connection with the celebrated Boltzmann's H-theorem.(24,53,81)

However, a reliability of the Kullback entropy may be questioned in case of general parameter-dependent densities. In particular, this entropy fails to quantify properly certain features of a non-stationary dynamics of probability densities. Specifically if we wish to make a "comparison" of once given density function to itself, but at different stages (instants) of its time evolution.

Let us consider 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 are not differentiated by the information (differential) entropy and share its very same value $\mathcal{S}_\sigma = \frac{1}{2}\ln(2\pi e\sigma^2)$ independent of α.

If we admit σ to be another free parameter, a two-parameter family of Gaussian densities $\rho_\alpha \to \rho_{\alpha,\sigma}(x)$ appears. Such densities, corresponding to different values of σ and σ' do admit an "absolute comparison" in terms of the Shannon entropy, in accordance with Eq. (14):

$$\mathcal{S}_{\sigma'} - \mathcal{S}_\sigma = \ln\left(\frac{\sigma'}{\sigma}\right). \tag{39}$$

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 $t' > t > 0$: $(\sigma'/\sigma) = \sqrt{t'/t}$.

In the above we have "compared" differential entropies of quite akin, albeit different, probability densities. Among many inequivalent ways to evaluate the "divergence" between probability distributions, the relative (Kullback) entropy is typically used to quantify such divergence from the a priori prescribed reference density.(60,82)

We define the Kullback entropy $\mathcal{K}(\theta,\theta')$ for a certain θ-labelled family of probability densities ρ_θ, so that the "distance" between any two densities in this family can be directly evaluated. Let $\rho_{\theta'}$ stands for the prescribed (reference) probability density. We have(58,82,88):

$$\mathcal{K}(\theta,\theta') \doteq \mathcal{K}(\rho_\theta|\rho_{\theta'}) = \int \rho_\theta(x)\,\ln\frac{\rho_\theta(x)}{\rho_{\theta'}(x)}\,dx, \tag{40}$$

which, in view of the concavity of the function $f(w) = -w\ln w$, is positive.

Let us indicate that the negative of $\mathcal{K}$,

$$\mathcal{H}_c \doteq -\mathcal{K}, \tag{41}$$

named the conditional entropy,(88) is predominantly used in the literature(64,60,88) because of its affinity (regarded as a formal generalization) to the differential entropy. Then e.g. one investigates an approach of $-\mathcal{K}$ towards its maximum (usually achieved at the value zero) when a running density is bound to have a unique stationary asymptotic.(64)

Let us consider a two-parameter $\theta \doteq (\theta_1,\theta_2)$ family of densities. If we take $\theta_i' \doteq \theta_i + \Delta\theta_i$ with $\Delta\theta_i \ll 1$ for $i = 1,2$, the following approximate formula holds true under a number of standard assumptions(58):

$$\mathcal{K}(\theta, \theta + \Delta\theta) \simeq \frac{1}{2} \sum_{i,j} \mathcal{F}_{ij} \cdot \Delta\theta_i \Delta\theta_j \tag{42}$$

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

$$\mathcal{F}_{ij} = \int \rho_\theta \frac{\partial \ln \rho_\theta}{\partial \theta_i} \cdot \frac{\partial \ln \rho_\theta}{\partial \theta_j} \, dx. \tag{43}$$

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

It is useful to note (c.f. also Ref. (64)) that in self-explanatory notation, for two θ and θ' Gaussian densities there holds:

$$\mathcal{K}(\theta, \theta') = \ln \frac{\sigma'}{\sigma} + \frac{1}{2} \left(\frac{\sigma^2}{\sigma'^2} - 1 \right) + \frac{1}{2\sigma'^2} (\alpha - \alpha')^2. \tag{44}$$

The first entry in Eq. (44) coincides with the "absolute comparison formula" for Shannon entropies, Eq. (39). However for $|\theta' - \theta| \ll 1$, hence in the regime of interest for us, the second term dominates the first one.

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

$$\mathcal{S}(\sigma') - \mathcal{S}(\sigma) \simeq \sqrt{\mathcal{K}(\theta, \theta')} \simeq \frac{\Delta t}{2t}, \tag{45}$$

the time derivative notion $\dot{\mathcal{S}}$ can be defined exclusively for the differential entropy, and is meaningless in terms of the Kullback "distance".

Let us mention that no such obstacle arises in the standard cautious use of the relative Kullback entropy $\mathcal{H}_c$. Indeed, normally one of the involved densities stands for the stationary reference one $\rho_{\theta'}(x) \doteq \rho_*(x)$, while another evolves in time $\rho_\theta(x) \doteq \rho(x, t)$, $t \in R^+$, thence $\mathcal{H}_c(t) \doteq -\mathcal{K}(\rho_t | \rho_*)$, see e.g. Refs. (60, 64).

5.4. Entropy Dynamics in the Smoluchowski Process

We consider spatial Markov diffusion processes in R with a diffusion coefficient (constant or time-dependent) D and admit them to drive space-time inhomogeneous probability densities $\rho = \rho(x, t)$. In the previous section we have addressed the special case of the free Brownian motion characterized by the current velocity (field) $v \doteq v(x, t) = -u(x, t) = -D\nabla \ln \rho(x, t)$ and the diffusion current $j \doteq v \cdot \rho$ which obeys the continuity equation $\partial_t \rho = -\nabla j$, this in turn being equivalent to the heat equation.

It is instructive to notice that the gradient of a potential-type function $Q = Q(x, t)$, c.f. Eq. (22), entirely composed in terms of u:

$$Q = 2D^2 \frac{\Delta\rho^{1/2}}{\rho^{1/2}} = \frac{1}{2}u^2 + D\nabla \cdot u \tag{46}$$

almost trivially appears (i.e. merely as a consequence of the heat equation,(29,41)) in the hydrodynamical (momentum) conservation law appropriate for the free Brownian motion:

$$\partial_t v + (v \cdot \nabla)v = -\nabla Q\,. \tag{47}$$

A straightforward generalization refers to a diffusive dynamics of a mass m particle in the external field of force, here taken to be conservative: $F = F(x) = -\nabla V$. The associated Smoluchowski diffusion process with a forward drift $b(x) = \frac{F}{m\beta}$ is analyzed in terms of the Fokker-Planck equation for the spatial probability density $\rho(x, t)$(49,59,82,96):

$$\partial_t \rho = D\triangle\rho - \nabla(b \cdot \rho) \tag{48}$$

with the initial data $\rho_0(x) = \rho(x, 0)$.

Note that if things are specialized to the *standard* Brownian motion in an external force field, we know a priori (due to the Einstein fluctuation-dissipation relationship,(25)) that $D = \frac{k_B T}{m\beta}$, where β is interpreted as the friction (damping) parameter, T is the temperature of the bath, k_B being the Boltzmann constant.

We assume, modulo restrictions upon drift function,(35,57) to resolve the Smoluchowski dynamics in terms of (possibly non-unique) Markovian diffusion-type processes. Then, the following compatibility equations follow in the form of hydrodynamical conservation laws for the diffusion process(29,41):

$$\partial_t \rho + \nabla(v\rho) = 0 \tag{49}$$

$$(\partial_t + v \cdot \nabla)v = \nabla(\Omega - Q) \tag{50}$$

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:

$$\Omega = \frac{1}{2}\left(\frac{F}{m\beta}\right)^2 + D\nabla \cdot \left(\frac{F}{m\beta}\right), \tag{51}$$

where the functional form of Q is given by Eq. (46). Obviously the free Brownian law, Eq. (47), comes out as the special case.

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

$$v \doteq b - u = \frac{F}{m\beta} - D\frac{\nabla\rho}{\rho} \tag{52}$$

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.

With a solution $\rho(x,t)$ of the Fokker-Planck equation, we associate a differential (information) entropy $\mathcal{S}(t) = -\int \rho \, \ln \rho \, dx$ which is typically *not* a conserved quantity.$^{(30)-(47)}$ The rate of change in time of $\mathcal{S}(t)$ readily follows.

Boundary restrictions upon ρ, $v\rho$ and $b\rho$ to vanish at spatial infinities (or at finite spatial volume boundaries) yield the rate equation:

$$\frac{d\mathcal{S}}{dt} = \int \left[\rho \, (\nabla \cdot b) + D \cdot \frac{(\nabla \rho)^2}{\rho} \right] dx \tag{53}$$

to be compared with the previous, $b = 0$ case, Eq. (47).

Anticipating further discussion, let us stress that even in case of plainly irreversible diffusive dynamics, it is by no means obvious that the differential entropy should grow, decay (diminish) or show up a mixed behavior. It is often tacitly assumed that one should "typically" have $\dot{\mathcal{S}} > 0$ which is not true.$^{(22,81)}$

We can rewrite Eq. (53) in a number of equivalent forms, like e.g. (note that $\langle u^2 \rangle = -D\langle \nabla \cdot u \rangle$): $D\dot{\mathcal{S}} \doteq D\,\langle \nabla \cdot b \rangle + \left\langle u^2 \right\rangle = D\langle \nabla \cdot v \rangle$, but in conformity with observations of$^{(57,81)}$, we consider

$$D\dot{\mathcal{S}} = \langle v^2 \rangle - \langle b \cdot v \rangle = -\langle v \cdot u \rangle \tag{54}$$

as the major entropy balance equation. Here $\langle \cdot \rangle$ denotes the mean value with respect to ρ.

This balance equation is extremely persuasive, since $b = F/(m\beta)$ and $j = v\rho$ combine into a characteristic "power release" expression:

$$\frac{dQ}{dt} \doteq \frac{1}{D} \int \frac{1}{m\beta} F \cdot j \, dx = \frac{1}{D} \, \langle b \cdot v \rangle \, . \tag{55}$$

Like in case of not necessarily positive $\dot{\mathcal{S}}$, the "power release" expression $\dot{\mathcal{Q}}$ may be positive which represents the power removal to the environment, as well as negative which corresponds to the power absorption from the environment.

In the formal thermodynamical lore, in the above we deal here with the time rate at which the mechanical work per unit of mass may possibly be dissipated (removed to the reservoir) in the form of heat, in the course of the Smoluchowski diffusion process: $k_B T \dot{\mathcal{Q}} = \int F \cdot j \, dx$, with T being the temperature of the bath. When there is no external forces, we have $b = 0$, and then the differential entropy time rate formula for the free Brownian motion Eq. (38) reappears.

On the other hand, the positive terms in Eq. (54) and Eq. (38) represent the rate at which information entropy is put (pumped) into the diffusing system by the thermally active environment, thus causing a disorder/uncertainty growth. This particular "entropy production" rate may possibly be counterbalanced (to this end we need external forces) by the heat removal due to dissipation, according to:

$$\frac{d\mathcal{S}}{dt} = \left(\frac{d\mathcal{S}}{dt}\right)_{in} - \frac{d\mathcal{Q}}{dt} \tag{56}$$

where $\dot{\mathcal{Q}}$ is defined in Eq. (55) while $(\dot{\mathcal{S}})_{in} = (1/D)\langle v^2 \rangle$.

Remark: *In Refs. (57, 80, 81) a measure-theoretic and probabilistic justification was given to an interpretation of* $(1/D)\langle v^2 \rangle$ *as the entropy production rate of the (originally – stationary) diffusion process with the current velocity* v. *We would like to point out that traditionally,*$^{(74,83)}$ *a statistical mechanical notion of the entropy production refers to the excess entropy that is pumped out of the system. An alternative statement tells about the entropy production by the physical system into the thermostat. In the present discussion, an increase of the information entropy of the Smoluchowski process definitely occurs due to the thermal environment: the differential entropy is being generated (accumulated, produced) in the physical system by its environment.*

Of particular interest is the case of constant information entropy $\dot{\mathcal{S}} = 0$ which amounts to the existence of steady states. In the simplest case, when the diffusion current vanishes, we encounter the primitive realization of the state of equilibrium with an invariant density ρ. Then, $b = u = D\nabla \ln \rho$ and we readily arrive at the classic equilibrium identity for the Smoluchowski process:

$$-(1/k_B T)\nabla V = \nabla \ln \rho \tag{57}$$

which determines the functional form of the invariant density in case of a given conservative force field.$^{(57,82)}$ There is an ample discussion in Ref. (57) of how these properties match with time reversal of the stationary diffusion process and the vanishing of the entropy production (in our lore) rate $(\dot{\mathcal{S}})_{in}$.

Coming back to the general discussion, let us define the so-called thermodynamic force $F_{th} \doteq v/D$ associated with the Smoluchowski diffusion and introduce its corresponding time-dependent potential function $\Psi(x, t)$:

$$k_B T\, F_{th} = F - k_B T\, \nabla \ln \rho \doteq -\nabla \Psi. \tag{58}$$

Notice that $v = -(1/m\beta)\nabla\Psi$. In the absence of external forces (free Brownian motion), we obviously get $F_{th} = -\nabla \ln \rho = -(1/D)u$.

The mean value of the potential

$$\Psi = V + k_B T \ln \rho \tag{59}$$

of the thermodynamic force associates with the diffusion process an obvious analogue of the Helmholtz free energy:

$$\langle \Psi \rangle = \langle V \rangle - T\, \mathcal{S}_G \tag{60}$$

where the dimensional version $\mathcal{S}_G \doteq k_B \mathcal{S}$ of information entropy has been introduced (actually, it is a direct configuration-space analog of the Gibbs entropy). The expectation value of the mechanical force potential $\langle V \rangle$ plays here the role of (mean) internal energy.(74,81)

By assuming that $\rho V v$ vanishes at integration volume boundaries (or infinity), we easily get the time rate of Helmholtz free energy at a constant temperature T:

$$\frac{d}{dt} \langle \Psi \rangle = -k_B T \dot{\mathcal{Q}} - T \dot{\mathcal{S}}_G \, . \tag{61}$$

By employing Eq. (56), we readily arrive at

$$\frac{d}{dt} \langle \Psi \rangle = -(k_B T) \left(\frac{d\mathcal{S}}{dt} \right)_{in} = -(m\beta) \left\langle v^2 \right\rangle \tag{62}$$

which either identically vanishes (equilibrium) or remains negative.

Thus, Helmholtz free energy either remains constant in time or decreases as a function of time at the rate set by the information entropy "production" $\dot{\mathcal{S}}_{in}$. One may expect that actualy $\langle \Psi \rangle(t)$ drops down to a finite minimum as $t \to \infty$.

However, this feature is a little bit deceiving. One should be aware that a finite minimum as well may not exist, which is the case e.g. for the free Brownian motion. Also, multiple minima need to be excluded as well.

5.5. Kullback Entropy versus Shannon Entropy in the Smoluchowski Process

In the presence of external forces the property Eq. (62) may consistently quantify an asymptotic approach towards a minimum corresponding to an invariant (presumed to be unique) probability density of the process. Indeed, by invoking Eq. (57) we realize that

$$\rho_*(x) = \frac{1}{Z} \exp \left(-\frac{V(x)}{k_B T} \right) \tag{63}$$

where $Z = \int \exp(-V(x)/k_B T)\, dx$, sets the minimum of $\langle \Psi \rangle(t)$ at $\langle \Psi \rangle_* = \Psi_* = -k_B T \ln Z$.

Let us take the above $\rho_*(x)$ as a reference density with respect to which the divergence of $\rho(x,t)$ is evaluated in the course of the pertinent Smoluchowski process. This divergence is well quantified by the conditional Kullback entropy $\mathcal{H}_c(t)$. Let us notice that

$$\mathcal{H}_c(t) = -\int \rho \, \ln \left(\frac{\rho}{\rho_*} \right) dx = \mathcal{S}(t) - \ln Z - \frac{\langle V \rangle}{k_B T} \tag{64}$$

Consequently, in view of Eqs. (61) and (56), we get

$$\dot{\mathcal{H}}_c = \dot{\mathcal{S}} + \dot{\mathcal{Q}} = (\dot{\mathcal{S}})_{in} \geq 0 \tag{65}$$

so that $\frac{d}{dt}\langle\Psi\rangle = -(k_B T)\dot{\mathcal{H}}_c$. An approach of $\langle\Psi\rangle(t)$ towards the minimum proceeds in the very same rate as this of $\mathcal{H}_c(t)$ towards its maximum.

In contrast to $\dot{\mathcal{H}}_c$ which is non-negative, we have no growth guarantee for the differential entropy $\dot{\mathcal{S}}$ whose sign is unspecified. Nonetheless, the balance between the time rate of entropy production/removal and the power release into or out of the environment, is definitely correct.

We have $\dot{\mathcal{S}} \geq -\dot{\mathcal{Q}}$ and surely $\dot{\mathcal{Q}} < 0 \rightarrow \dot{\mathcal{S}} > 0$. If $\dot{\mathcal{Q}} > 0$, $\dot{\mathcal{S}}$ may take negative values down to the lower bound $-\dot{\mathcal{Q}}$. For the free Brownian motion, we have $\dot{\mathcal{Q}} = 0$ and thus $\dot{\mathcal{S}} \geq 0$. Let us notice that in view of Eq. (60) there holds $\langle\Psi\rangle = -k_B T\mathcal{H}_c(t) + \ln Z$.

5.6. One-Dimensional Ornstein-Uhlenbeck Process

It is quite illuminating to exemplify previous considerations by a detailed presentation of the standard one-dimensional Ornstein-Uhlenbeck process. We denote $b(x) = -\gamma x$ with $\gamma > 0$.

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 Eq. (48) preserves the Gaussian form of $\rho(x, t)$ while modifying the mean value $\alpha(t) = \alpha_0 \exp(-\gamma t)$ and variance according to

$$\sigma^2(t) = \sigma_0^2 \exp(-2\gamma t) + \frac{D}{\gamma}[1 - \exp(-2\gamma t)]\,. \tag{66}$$

Accordingly, since a unique invariant density has the form $\rho_* = \sqrt{\gamma/2\pi D}\exp(-\gamma x^2/2D)$ we obtain$^{(64)}$:

$$\mathcal{H}_c(t) = \exp(-2\gamma t)\mathcal{H}_c(\rho_0, \rho_*) = -\frac{\gamma\alpha_0^2}{2D}\exp(-2\gamma t) \tag{67}$$

while in view of our previous considerations, we have $\mathcal{S}(t) = (1/2)\ln[2\pi e\sigma^2(t)]$ and $\mathcal{F} = 1/\sigma^2(t)$. Therefore

$$\dot{\mathcal{S}} = \frac{2\gamma(D - \gamma\sigma_0^2)\exp(-2\gamma t)}{D - (D - \gamma\sigma_0^2)\exp(-2\gamma t)}\,. \tag{68}$$

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, though 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 $t \rightarrow \infty$ reads $(1/2)\ln[2\pi e(D/\gamma)$.

The differential entropy evolution is anti-correlated with this of the localization, since

$$\dot{\mathcal{F}} = -\frac{\gamma\dot{\mathcal{S}}}{[D - (D - \gamma\sigma_0^2)\exp(-2\gamma t)]^2}\,. \tag{69}$$

For all σ_0^2 the asymptotic value of $\mathcal{F}$ reads γ/D.

We have here a direct control of the behavior of the "power release" expression $\dot{\mathcal{Q}} = \dot{\mathcal{H}}_c - \dot{\mathcal{S}}$. Since

$$\dot{\mathcal{H}}_c = \left(\gamma^2\alpha_0^2/D\right)\exp\left(-2\gamma t\right) > 0, \tag{70}$$

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

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 $t_{change}(\alpha_0^2, \sigma_0^2)$.

5.7. Mean Energy and the Dynamics of Fisher Information

By considering $(-\rho)(x,t)$ and $s(x,t)$, such that $v = \nabla s$, as canonically conjugate fields, we can invoke the variational calculus. 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. (50)):

$$\partial_t s + \frac{1}{2}(\nabla s)^2 - (\Omega - Q) = 0, \tag{71}$$

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

$$\mathcal{L} = -\int \rho \left[\partial_t s + \frac{1}{2}(\nabla s)^2 - \left(\frac{u^2}{2} + \Omega\right)\right] dx. \tag{72}$$

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

$$\mathcal{H} \doteq \int \rho \cdot \left[\frac{1}{2}(\nabla s)^2 - \left(\frac{u^2}{2} + \Omega\right)\right] dx \tag{73}$$

i.e.

$$\mathcal{H} = (1/2)(< v^2 > - < u^2 >) - \langle\Omega\rangle\,.$$

We can evaluate an expectation value of Eq. (71) which implies an identity $\mathcal{H} = -\langle\partial_t s\rangle$. By invoking Eq. (59), with the time-independent V, we arrive at

$$\dot{\Psi} = \frac{k_B T}{\rho}\nabla(v\rho) \tag{74}$$

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

$$s(x,t) \doteq (1/m\beta)\Psi(x,t) \Longrightarrow \langle\partial_t s\rangle = 0 \tag{75}$$

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

We have thus arrived at the following interplay between the mean energy and the information entropy "production" rate:

$$\frac{D}{2}\left(\frac{dS}{dt}\right)_{in} = \frac{1}{2}\langle v^2\rangle = \int \rho \left(\frac{\vec{u}^{\,2}}{2} + \Omega\right) dx \geq 0, \tag{76}$$

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

By recalling the notion of the Fisher information Eq. (21) and setting $\mathcal{F} \doteq D^2\mathcal{F}_\alpha$, we can rewrite the above formula as follows:

$$\mathcal{F} = \langle v^2\rangle - 2\langle\Omega\rangle \geq 0 \tag{77}$$

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

We may evaluate directly the 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 $1/\sigma^2$, see e.g. Section 3.

Namely, by exploiting the hydrodynamical conservation laws Eq. (50) for the Smoluchowski process we get:

$$\partial_t(\rho v^2) = -\nabla \cdot [(\rho v^3)] - 2\rho v \cdot \nabla(Q - \Omega). \tag{78}$$

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

$$\frac{d}{dt}\left\langle v^2\right\rangle = 2\left\langle v \cdot \nabla(\Omega - Q)\right\rangle. \tag{79}$$

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

$$\frac{d}{dt}\langle\Omega\rangle = \langle v \cdot \nabla\Omega\rangle \tag{80}$$

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

$$\frac{d}{dt}\mathcal{F} = \frac{d}{dt}[\langle v^2\rangle - 2\langle\Omega\rangle] = -2\langle v \cdot \nabla Q\rangle. \tag{81}$$

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 to be compared with our previous discussion of the "heat dissipation" term Eq. (55).

For the Ornstein-Uhlenbeck process in one dimension we have $\langle v^2 \rangle(t) = (D/2)\dot{\mathcal{H}}_c = t(\gamma^2 \alpha_0^2/D)\exp(-2\gamma t)$, hence an asymptotic value 0, while $\langle u^2 \rangle(t) = (D/2)\mathcal{F}(t) \rightarrow \gamma/D$. Accordingly, we have $\langle \Omega \rangle(t) \rightarrow -\gamma/2D$.

6. DIFFERENTIAL ENTROPY DYNAMICS IN QUANTUM THEORY

6.1. Balance Equations

In the discussion of Smoluchowski diffusions, our major reference point was the conventional Fokker-Planck equation (48) for a probability density supporting a Markovian diffusion process. The (time-independent) drift function b was assumed to be known a priori (e.g. the conservative external forces were established on phenomenological or model construction grounds), while the initial and/or boundary data for the probability density of the process could be chosen (to a high degree) arbitrarily.

Under such "normal" circumstances, the hydrodynamical conservation laws (50) come out as a direct consequence of the Fokker-Planck equation. Also, the functional expression for Ω of the form (51) is basically known to arise if one attempts to replace an elliptic diffusion operator by a Hermitian (and possibly self-adjoint) one.(41,57,82)

We shall depart from the standard Brownian motion setting to more general Markovian diffusion-type processes which, while still respecting the Fokker-Planck equation, admit general time-dependent forward drifts. In fact, we invoke at this point a well defined stochastic counterpart of the Schrödinger picture quantum dynamics of wave packets,(20,21,35,39,41,57,75) where the notion of differential entropy and its dynamics finds a proper place. The dynamics of quantal probability densities is here resolved in terms of diffusion-type processes.

Let us assume to have chosen an arbitrary continuous (it is useful, if bounded from below) function $\mathcal{V} = \mathcal{V}(\vec{x}, t)$ with dimensions of energy. we consider the Schrödinger equation (set $D = \hbar/2m$) in the form

$$i\partial_t \psi = -D\Delta\psi + \frac{\mathcal{V}}{2mD}\psi. \tag{82}$$

The Madelung decomposition $\psi = \rho^{1/2}\exp(is)$ with the phase function $s = s(x, t)$ defining $v = \nabla s$ is known to imply two coupled equations: the standard continuity equation $\partial_t \rho = -\nabla(v\rho)$ and the Hamilton-Jacobi-type equation

$$\partial_t s + \frac{1}{2}(\nabla s)^2 + (\Omega - Q) = 0 \tag{83}$$

where $\Omega \doteq \mathcal{V}/m$ and the functional form of Q coincides with this introduced previously in Eq. (46). Notice a "minor" sign change in Eq. (83) in comparison with Eq. (71).

Those two equations form a coupled system, whose solutions describe a Markovian diffusion-type process: the probability density is propagated by a Fokker-Planck dynamics of the form Eq. (48) with the drift $b = v - u$ where $u = D\nabla \ln \rho$ is an osmotic velocity field.

We can mimic the calculus of variations steps of the previous section, so arriving at the Hamiltonian (actually, the mean energy of the quantum motion per unit of mass):

$$\mathcal{H} \doteq \int \rho \cdot \left[\frac{1}{2}(\nabla s)^2 + \left(\frac{u^2}{2} + \Omega\right)\right] dx, \tag{84}$$

to be compared with Eq. (73). There holds

$$\mathcal{H} = (1/2)\left[\left\langle v^2\right\rangle + \left\langle u^2\right\rangle\right] + \langle \Omega \rangle = -\left\langle \partial_t s\right\rangle . \tag{85}$$

Of particular interest (due to its relative simplicity) is the case of time-independent $\mathcal{V}$, when

$$\mathcal{H} = -\left\langle \partial_t s\right\rangle \doteq \mathcal{E} = const \tag{86}$$

is known to be a conserved finite quantity, which is not necessarily positive. Since generally $\mathcal{H} \neq 0$, we deal here with so-called finite energy diffusion-type processes.(21,35) The corresponding Fokker-Planck equation propagates a probability density $|\psi|^2 = \rho$, whose differential entropy $\mathcal{S}$ may quite nontrivially evolve in time.

Keeping intact the previous derivation procedures for $(\dot{\mathcal{S}})_{in}$ (while assuming the validity of mathematical restrictions upon the behavior of integrands), we encounter the information entropy balance equations in their general form disclosed in Eqs. (54)–(56). The related differential entropy "production" rate reads:

$$(\dot{\mathcal{S}})_{in} = \frac{2}{D}\left[\mathcal{E} - \left(\frac{1}{2}\mathcal{F} + \langle \Omega \rangle\right)\right] \geq 0\,, \,. \tag{87}$$

We recall that $\frac{1}{2}\mathcal{F} = -\langle Q \rangle > 0$ which implies $\mathcal{E} - \langle \Omega \rangle \geq \frac{1}{2}\mathcal{F} > 0$. Therefore, the localization measure $\mathcal{F}$ has a definite upper bound: the pertinent wave packet cannot be localized too sharply.

We notice that the localization (Fisher) measure

$$\mathcal{F} = 2(\mathcal{E} - \langle \Omega \rangle) - \langle v^2 \rangle \tag{88}$$

in general evolves in time. Here $\mathcal{E}$ is a constant and $\dot{\Omega} = 0$.

By invoking the hydrodynamical conservation laws, we find out that the dynamics of Fisher information follows an equation:

$$\frac{d\mathcal{F}}{dt} = +2\langle v\nabla Q\rangle \tag{89}$$

and that there holds

$$\frac{1}{2}\dot{\mathcal{F}} = -\frac{d}{dt}\left[\frac{1}{2}\langle v^2\rangle + \langle\Omega\rangle\right] \tag{90}$$

which is to be compared (notice the opposite sign of the right-hand expression) with the result we have obtained for Smoluchowski processes.

Obviously, now we have $\dot{\mathcal{F}} = +\int v\nabla P dx$, with the same functional form for P as before. We interpret $\dot{\mathcal{F}}$ as the measure of power transfer in the course of which the (de)localization "feeds" the diffusion current and *in reverse.* Here, we encounter a negative feedback between the localization and the proper energy of motion which keeps intact an overall mean energy $\mathcal{H} = \mathcal{E}$ of the quantum motion. See e.g. also Ref. (41).

In case of $v = 0$, we have $\mathcal{E} = \frac{1}{2}\mathcal{F} + \langle\Omega\rangle$ and no entropy "production" nor dynamics of uncertainty. There holds $\dot{\mathcal{S}} = 0$ and we deal with time-reversible stationary diffusion processes and their invariant probability densities $\rho(x)$.$^{(35,57)}$

Let us indicate that the phase function $s(x,t)$ shows up certain (remnant) features of the Helmholtz Ψ and $\langle\Psi\rangle$. This behavior is not unexpected, since e.g. the ground state densities (and other invariant densities of stationary states) are directly related to time-reversible stationary diffusion-type processes of Refs. (35, 57). We have $-\langle\partial_t s\rangle = \mathcal{E}$. In view of $v = \nabla s$ and assumed vanishing of $s\rho v$ at the integration volume boundaries, we get:

$$\frac{d}{dt}\langle s\rangle = \langle v^2\rangle - \mathcal{E}. \tag{91}$$

The previously mentioned case of *no* entropy "production" refers to $v = 0$ and thus $s = s_0 - \mathcal{E}\cdot t$.

We recall that the corresponding derivation of Eq. (62) has been carried out for $v = -(1/m\beta)\nabla\Psi$, with $\langle\dot{\Psi}\rangle = 0$). Hence, as close as possible link with the present discussion is obtained if we re-define s into $s_\Psi \doteq -s$. Then we have

$$\frac{d}{dt}\langle s_\Psi\rangle = \mathcal{E} - \langle v^2\rangle. \tag{92}$$

For stationary quantum states, when $v = 0$ identically, we get $\frac{d}{dt}\langle s_\Psi\rangle = \mathcal{E}$, in contrast to the standard Fokker-Planck case of $\frac{d}{dt}\langle\Psi\rangle = 0$.

Interestingly enough, we can write the generalized Hamilton-Jacobi equation, while specified to the $v = 0$ regime, with respect to s_Ψ. Indeed, there holds $\partial_t s_\Psi = \Omega - Q$, in close affinity with Eq. (71) in the same regime.

6.2. Differential Entropy Dynamics Exemplified

6.2.1. Free Evolution

Let us consider the probability density in one space dimension:

$$\rho(x,t) = \frac{\alpha}{[\pi(\alpha^4 + 4D^2t^2)]^{1/2}} \exp\left(-\frac{x^2\alpha^2}{\alpha^4 + 4D^2t^2}\right) \tag{93}$$

and the phase function

$$s(x,t) = \frac{2D^2x^2t}{\alpha^4 + 4D^2t^2} - D^2 \arctan\left(-\frac{2Dt}{\alpha^2}\right) \tag{94}$$

which determine a free wave packet solution of equations (82) and (83), i.e. obtained for the case of $\mathcal{V} \equiv 0$ with the initial data $\psi(x,0) = (\pi\alpha^2)^{-1/4}\exp(-x^2/2\alpha^2)$.

We have:

$$b(x,t) = v(x,t) + u(x,t) = \frac{2D(2Dt - \alpha^2)x}{\alpha^4 + 4D^2t^2} \tag{95}$$

and the the Fokker-Planck equation with the forward drift $b(x,t)$ is solved by the above ρ.

In the present case, the differential entropy reads:

$$\mathcal{S}(t) = \frac{1}{2}\ln\left[2\pi e\left\langle X^2\right\rangle(t)\right] \tag{96}$$

where $\left\langle X^2\right\rangle \doteq \int x^2\rho dx = (\alpha^4 + 4D^2t^2)/2\alpha^2$. Its time rate $D\dot{\mathcal{S}} = \langle v^2\rangle - \langle b \cdot v\rangle$ equals:

$$D\frac{d\mathcal{S}}{dt} = \frac{4D^3t}{\alpha^4 + 4D^2t^2} \geq 0 \tag{97}$$

for $t \geq 0$. Its large time asymptotic is D/t.

Furthermore, we have

$$D(\dot{\mathcal{S}})_{in} = \left\langle v^2\right\rangle = \frac{8D^4t^2}{\alpha^2(\alpha^4 + 4D^2t^2)} \tag{98}$$

with the obvious large time asymptotic value $2D^2/\alpha^2$: the differential entropy production remains untamed for all times.

Due to $\langle u^2\rangle = (2D^2\alpha^2)/(\alpha^4 + 4D^2t^2)$ there holds

$$\mathcal{E} = \frac{1}{2}(\langle v^2\rangle + \langle u^2\rangle) = \frac{D^2}{\alpha^2}. \tag{99}$$

Accordingly, the quantum mechanical analog of the entropy (rather than heat) "dissipation" term $-D \cdot \mathcal{Q}$ in the quantum case reads

$$-\langle b \cdot v\rangle = \frac{4D^3t(\alpha^2 - 2Dt)}{\alpha^2(\alpha^4 + 4D^2t^2)} \tag{100}$$

and while taking negative values for $t < \alpha^2/2D$, it turns out to be positive for larger times. Formally speaking, after a short entropy "dissipation" period we pass to the entropy "absorption" regime which in view of its D/t asymptotic, for large times is definitely dominated by $D(\dot{\mathcal{S}})_{in} \sim 2D^2/\alpha^2$.

Those differential entropy balance features do parallel a continual growth of the mean kinetic energy $(1/2)\langle v^2\rangle$ from an initial value 0 towards its asymptotic value $D^2/\alpha^2 = \mathcal{E}$. Note that the negative feedback is here displayed by the the behavior of $\langle u^2\rangle$ which drops down from the initial value $2D^2/\alpha^2$ towards 0. It is also instructive to notice that in the present case $\mathcal{F}(t) = D^2/\langle X^2\rangle(t)$. We can readily check that $\dot{\mathcal{F}} = d\langle u^2\rangle/dt = -d\langle v^2\rangle/dt$.

6.2.2. Steady State

We choose the probability density in the form:

$$\rho(x,t) = \left(\frac{\omega}{2\pi D}\right)^{1/2} \exp\left[-\frac{\omega}{2D}\,(x - q(t))^2\right] \tag{101}$$

where the classical harmonic dynamics with particle mass m and frequency ω is involved such that $q(t) = q_0\cos(\omega t) + (p_0/m\omega)\sin(\omega t)$ and $p(t) = p_0\cos(\omega t) - m\omega q_0\sin(\omega t)$.

One can easily verify that (48), and (83) hold true identically once we set $\mathcal{V} = \frac{1}{2}\omega^2x^2$ and consider:

$$s(x,t) = (1/2m)\,[xp(t) - (1/2)p(t)q(t) - mD\omega t]\,. \tag{102}$$

A forward drift takes the form:

$$b(x,t) = \frac{1}{m}p(t) - \omega\,(x - q(t)) \tag{103}$$

and the above ρ solves the corresponding Fokker-Planck equation.

The differential entropy is a constant equal $\mathcal{S} = (1/2)\ln(2\pi eD/\omega)$. Although trivially $d\mathcal{S}/dt = 0$, all previous arguments can be verified.

For example, we have $v = \nabla s = p(t)/2m$ and therefore an oscillating entropy "production" term $D(\dot{\mathcal{S}})_{in} = p^2(t)/4m^2$ which is balanced by an oscillating "dissipative" counter-term to yield $\dot{\mathcal{S}}$. Suitable expressions for $\langle s \rangle$ and $\langle \partial_t s \rangle$ easily follow.

Concerning the Fisher measure, we have obviously $\mathcal{F} = \omega/D$ which is a constant of motion.

6.2.3. Squeezed State

Let us consider$^{(67)}$ the squeezed wave function of the harmonic oscillator. We adopt the re-scaled units $\hbar = \omega = m = 1$, hence also $D = 1$. The solution of the Schröd inger equation $i\partial_t \psi = (-1/2)\Delta\psi + (x^2/2)\psi$ with the initial data $\psi(x, 0) = (\gamma^2\pi)^{-1/4} \exp(-x^2/2\gamma^2)$ and $\gamma \in (0, \infty)$, is defined in terms of the probability density:

$$\rho(x, t) = \frac{1}{(2\pi)^{1/2}\sigma(t)} \exp\left(-\frac{x^2}{2\sigma^2(t)}\right) \tag{104}$$

where

$$2\sigma^2(t) = \frac{1}{\gamma^2} \sin^2 t + \gamma^2 \cos^2 t \tag{105}$$

and the phase function

$$s(x, t) = \phi(t) + \frac{(1/\gamma^2 - \gamma^2) \sin 2t}{8\sigma^2(t)} x^2 \tag{106}$$

where $\phi(t) = \frac{t}{2} + \frac{\pi}{4} + \arctan(\gamma^2 \cot t)$.

Now, the differential entropy $\mathcal{S} = (1/2)\ln[2\pi e\sigma^2(t)]$ displays a periodic behavior in time, whose level of complexity depends on the particular value of the squeezing parameter γ. The previously mentioned negative feedback is here manifested through (counter)oscillations of the localization, this in conformity with the dynamics of $\sigma^2(t)$ and the corresponding oscillating dynamics of the Fisher measure $\mathcal{F} = 1/\sigma^2(t)$. .

See e.g. also Ref. (67) for a pictorial analysis and an instructive computer assisted discussion of the Schrödinger cat state (superposition of the harmonic oscillator coherent states with the same amplitude but with opposite phases), with the time evolution of the corresponding differential entropy.

6.2.4. Stationary States

In contrast to generic applications of the standard Fokker-Planck equation, where one takes for granted that there is a unique positive stationary probability density, the situation looks otherwise if we admit the Schrödinger equation as a primary dynamical rule for the evolution of (inferred) probability densities. For a chosen potential, all available stationary quantum states may serve the purpose, since then we have nonnegative (zeroes are now admitted) $\rho_*(x)$, and $v(x) = 0$ identically (we stay in one spatial dimension).

The standard harmonic oscillator may serve as an instructive example. One may e.g. consult Fig. 3 in Ref. (102) to check the behavior of both position and momentum differential entropies, and their sum, depending on the energy eigenvalue. All these stationary state values monotonically grow up with $n = 1, 2, \ldots, 60$,$^{(102)}$ and follow the pattern in the asymptotic regime $n \equiv 500$.$^{(67)}$

For convenience we shall refer to the Schrödinger eigenvalue problem with scaled away physical units. We consider (compare e.g. Eq. (82) with $D \to 1/2$)

$$\left[-\frac{1}{2}\Delta + \frac{x^2}{2}\right]\sqrt{\rho_*} = \left(n + \frac{1}{2}\right)\sqrt{\rho_*}\,. \tag{107}$$

In terms of a suitable Hamilton-Jacobi type equation we can address the same problem by seeking solutions of an equation

$$n + 1/2 = \Omega - Q \tag{108}$$

with respect to $\sqrt{\rho_*}$, provided we set $\Omega = x^2/2$, define $u = \nabla \ln \sqrt{\rho_*}$ and demand that $Q = u^2/2 + (1/2)\nabla \cdot u$.

For the harmonic oscillator problem, we can refer to standard textbooks. For each value of n we recover a corresponding unique stationary density: $\sqrt{\rho_*} \to \rho_n^{1/2}$ with $n = 0, 1, 2, \ldots$). We have:

$$\rho_n^{1/2}(x) = \frac{1}{(2^n n! \sqrt{\pi})^{1/2}} \exp\left(-\frac{x^2}{2}\right) H_n(x) \tag{109}$$

where $H_n(x)$ stands for the n-th Hermite polynomial: $H_0 = 1$, $H_1 = 2x$, $H_2 = 2(2x^2 - 1)$, $H_3 = 4x(2x^2 - 3)$, and so on.

We immediately infer e.g. $b_0 = -x \to Q = x^2/2 - 1/2$, next $b_1 = (1/x) - x \to Q = x^2/2 - 3/2$, and $b_2 = [4x/(2x^2 - 1)] - x \to Q = x^2/2 - 5/2$, plus $b_3 = [(1/x) + 4x/(2x^2 - 3)] \to Q = x^2 - 7/2$, that is to be continued for $n > 3$. Therefore Eq. (108) is here a trivial identity.

Obviously, except for the ground state which is strictly positive, all remaining stationary states are nonnegative.

An open problem, first generally addressed in Ref. (36), see also Ref. (13), is to implement a continuous dynamical process for which any of induced stationary densities may serve as an invariant asymptotic one. An obvious (Ornstein-Uhlenbeck) solution is known for the ground state density.

7. SELECTED USES OF ENTROPY FUNCTIONALS

One may raise an issue of what are the entropy functionals good for.

First, let us recall that the computer assisted research$^{(85)-(102)}$ on the complexity of atoms and nuclei (e.g. density distributions of nuclei, electrons in atoms and valence electrons in atomic clusters) infers densities from one- or many-body stationary states and quantifies complexity in terms of Shannon entropies in position and momentum spaces. The list of concrete Shannon entropy values for most common probability densities (pp. 486–7 in Ref. (27)) may not look inspiring. Nevertheless, the entropy values for quantum probability densities convey a useful information. See e.g. Refs. (26)–(38).

Quite aside from the omnipresence of entropy functionals in problems of mathematical statistics and statistical inference,$^{(28)}$ Shannon and Fisher functionals enter important inequalities that were used in alternative proofs of the celebrated central limit theorem.$^{(8,62)}$ The mathematically oriented research on entropy methods$^{(4)-(97)}$ addresses mostly an asymptotic (large time) behavior for solutions of various partial differential equations, and an issue of the speed of convergence towards equilibrium (invariant) density. The major tools are conditional entropies and convex Sobolev inequalities, see also Ref. (21). As a hint let us mention the validity of the Euclidean logarithmic Sobolev inequality for entropy functionals of the Ornstein-Uhlenbeck process: $\mathcal{S} \geq \frac{1}{2}[2 + \ln(2\pi) - \mathcal{F}]$.$^{(91)}$

In non-equilibrium statistical mechanics of gases one invokes entropy functionals to solve concrete physical problems: for example to address the second law of thermodynamics and the related Boltzmann H-theorem from a probabilistic point of view. At this point, we should mention a number of papers on the rigorous formulation of effects of noise on entropy evolution and attempts to justify the "entropy growth" paradigm for model systems.$^{(24,49,53,55,56,64,65)}$

The simplest (naive) version of the Boltzmann H-theorem in case of the rarified gas, without external forces, close to its thermal equilibrium and space homogeneous, tells: if the probability density function $f(v)$ is a solution of the Boltzmann kinetic equation, then the functional $H(t) = \int f(v) \ln f(v) dv$ does not increase: $\frac{d}{dt}H(t) \leq 0$. $H(t)$ is a constant only if $f \doteq f_*(v) \simeq \exp[-(v - v_0)^2/k_B T]$.

Remembering about the Csiszár-Kullback inequality $\int f \ln(f/g) dx \geq (1/2)\|f - g\|^2_{L^1}$, in which the Kullback-Leibler entropy appears as an upper bound upon a "distance measure" in the set of density functions, we may consider the behavior of solutions for heat and Fokker-Planck equations.

Let us consider $\partial_t u = \Delta u$ with $x \in R^n$, $t \in R^+$ and $u(., t = 0) = u_0(.) \geq 0$, $\int u_0(x) dx = 1$. As $t \to \infty$, we have $u(x, t) \simeq \rho(x, t) = (4\pi t)^{-n/2} \exp[-x^2/4t]$. The obvious question is: what is the $t \to \infty$ rate of convergence of $\|u - \rho\|_{L^1}(t) \doteq \int |u(x, t) - \rho(x, t)| dx$? Obviously one knows that a regular solution of the heat equation behaves asymptotically as a fundamental solution, once time goes to infinity.$^{(82)}$

This decay in time has been quantified in terms of the Kullback-Leibler entropy $\mathcal{K}(\rho_\theta|\rho_{\theta'})$. 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 $\mathcal{K}(\rho_t, \rho_{\alpha+kt}) \leq \mathcal{K}(\rho_0, \rho_\alpha)[\alpha/(\alpha + kt)]$.

If we consider $\partial_t f = \Delta f - \nabla \cdot (bf)$, where $f(., t = 0) = f_0 \geq 0$, $\int f_0(x)dx = 1$ and assume that the forward drift $b = b(x, t)$ has a gradient form, the natural question is: let f_* be the stationary solution of the F-P equation, 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 ρ_t decays in relative entropy to a Gaussian (Maxwellian), the speed of such decay is exponential.(4,64,65) See also our previous discussion of the Ornstein-Uhlenbeck process.

8. OUTLOOK

We emphasize the purpose dependence of all entropy functional definitions.

The conditional Kullback entropy is often considered as the only valid "entropy growth" justification, although it was *merely* designed to quantify "how distant two densities from each other are" with the large time asymptotic input implicit.

On the other hand, in the present paper we have analyzed the "distance" of a solution of an evolution equation at one time instant, to the very same solution at another time instant. The Shannon (differential, information) entropy is perfectly suited for such "shortest description length" (here, short time) analysis, and in particular for the study of rapid changes in time of the probability distribution.

The differential entropy balance equation contains terms with a direct physical interpretation of the power release or absorption. We have analyzed this issue for standard diffusion processes of non-equilibrium statistical physics and next generalized the formalism to the Schrödinger picture implemented dynamics of probability densities related to pure quantum states in $L^2(R)$, where an approach towards equilibrium is not expected to occur at all.

We have demonstrated that the differential entropy needs not to increase, even in case of plainly irreversible dynamics. The monotonic growth in time of the conditional Kullback entropy (when applicable), not necessarily should be related to the "dynamical origins of the increasing entropy".(66)

In case of Smoluchowski processes, the time rate of the conditional Kullback entropy was found to coincide with the corresponding differential (Shannon) entropy "production" rate. The differential entropy itself needs not to grow and may as well change its dynamical regime from growth to decay and in reverse, even with the entropy "production" involved.

The time evolution of the differential entropy and the Fisher information measure involves a nontrivial power transfer. In case of Smoluchowski processes this power release can be easily attributed to the entropy removal from the system or the entropy absorption (drainage) from the thermostat.

In the quantum mechanical regime, the inherent power transfer is related to metamorphoses of various forms of mean energy among themselves and needs not the notion of external to the system thermostat. There is a definite negative feedback between the degree of localization and the mean kinetic energy of the probability current, Eq. (90). Effectively, the time rate of the Fisher functional $\dot{\mathcal{F}}$ is a measure of the power transfer in the course of which the (de)localization "feeds" the probability current and in reverse.

ACKNOWLEDGEMENT

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

REFERENCES

1. C. Adami, Physics of information, arXiv:quant-ph/040505 (2004).
2. R. Alicki and M. Fannes, *Quantum Dynamical Systems*, Oxford University Press, Oxford (2001).
3. V. Ambegaokar and A. Clerk, Entropy and time, *Am. J. Phys.* **67**:1068–1073 (1999).
4. A. Arnold, et al., On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations, *Comm. Partial Diff. Equations* **26**:43–100 (2001).
5. B. C. Bag, Upper bound for the time derivative of entropy for nonequilibrium stochastic processes, *Phys. Rev.* **E 65**:046118 (2002).
6. C. C. Bag, et al., Noise properties of stochastic processes and entropy production, *Phys. Rev.* **E 64**:026110 (2001).
7. R. Balian, Random matrices and information theory, *Nuovo Cim.* **B 57**:183–103 (1968).
8. A. R. Barron, Entropy and the central limit theorem, *Annals Probab. Theory* **14**:336–342 (1986).
9. W. Beckner, Inequalities in Fourier analysis, *Ann. Math.* **102**:159–182 (1975).
10. K. Berndl, et al., On the global existence of Bohmian mechanics, *Commun. Math. Phys.* **173**:647–673 (1995).
11. Białynicki-I. Birula and J. Madajczyk, Entropic uncertainty relations for angular distributions, *Phys. Lett.* **A 108**:384–386 (1985).
12. Białynicki-I. Birula, and J. Mycielski, Uncertainty Relations for Information Entropy in Wave Mechanics, *Commun. Math. Phys.* **44**:129–132 (1975).
13. Ph. Blanchard and P. Garbaczewski, Non-negative Feynman-Kac kernels in Schrödinger's interpolation problem, *J. Math. Phys.* **38**:1–15 (1997).
14. R. Blankenbecler and M. H. Partovi, Uncertainty, entropy, and the statistical mechanics of microscopic systems, *Phys. Rev. Lett.* **54**:373–376 (1985).
15. A. V. Bobylev and G. Toscani, On the generalization of the Boltzmann H-theorem for a spatially homogeneous Maxwell gas, *J. Math. Phys.* **33**:2578–2586 (1992).
16. M. Bologna, et al., Trajectory versus probability density entropy, *Phys. Rev.* **E 64**:016223 (2001).
17. L. Brillouin, *Science and Information Theory*, Academic Press, NY (1962).

18. Ĉ. Brukner and A. Zeilinger, Conceptual inadequacy of the Shannon information in quantum measurements, *Phys. Rev.* **A 63**:022113 (2002).
19. V. Buyarov, et al., Computation of the entropy of polynomials orthogonal on an interval, *SIAM J. Sci. Comp.* to appear (2004), also *math.NA/0310238*.
20. E. A. Carlen, Superadditivity of Fisher's information and logarithmic Sobolev inequalities, *J. Funct. Anal.* **101**:194–211 (1991).
21. E. Carlen, Conservative diffusions, *Commun. Math. Phys.* **94**:293–315 (1984).
22. R. G. Catalan, J. Garay, and Lopez-R. Ruiz, Features of the extension of a statistical measure of complexity to continuous systems, *Phys. Rev.* **E 66**:011102 (2002).
23. C. M. Caves, and C. Fuchs, Quantum information: how much information in a state vector ?, *Ann. Israel Phys. Soc.* **12**:226–237 (1996).
24. C. Cercignani, *Theory and Application of the Boltzmann Equation*, Scottish Academic Press, Edinburgh (1975).
25. S. Chandrasekhar, Stochastic problems in physics and astronomy, *Rev. Mod. Phys.* **15**:1–89 (1943).
26. K. Ch. Chatzisavvas, Ch. C., Moustakidis and C. P. Panos, Information entropy, information distances and complexity of atoms, *J. Chem. Phys.* **123**:174111 (2005).
27. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, NY (1991).
28. H. Cramér, *Mathematical methods of statistics*, Princeton University Press, Princeton (1946).
29. R. Czopnik, and P. Garbaczewski, Frictionless Random Dynamics: Hydrodynamical Formalism, *Physica* **A 317**:449–471 (2003).
30. D. Daems and G. Nicolis, Entropy production and phase space volume contraction, *Phys. Rev.* **E 59**:4000–4006 (1999).
31. G. Deco, et al: Determining the information flow of dynamical systems from continuous probability distributions, *Phys. Rev. Lett.* **78**:2345–2348 (1997).
32. A. Dembo and T. Cover, Information theoretic inequalities, *IEEE Trans. Inf. Th.* **37**:1501–1518 (1991).
33. D. Deutsch, Uncertainty in quantum measurements, *Phys. Rev. Lett.* **50**:631–633 (1983).
34. J. Dunkel and S. A. Trigger, Time-dependent entropy of simple quantum systems, *Phys. Rev.* **A 71**:052102 (2005).
35. A. Eberle, *Uniqueness and Non-uniqueness of Semigroups Generated by Singular Diffusion Operators*, LNM vol. **1718**, Springer-Verlag, Berlin (2000).
36. R. Fortet, Résolution d'un systéme d'équations de M. Schrödingeer, *J. Math. Pures Appl.* **9**:83 (1040).
37. B. R. Frieden and B. H. Sofer, Lagrangians of physics and the game of Fisher-information transfer, *Phys. Rev.* **E 52**:2274–2286 (1995).
38. S. R. Gadre, et al., Some novel characteristics of atomic information entropies, *Phys. Rev.* **A 32**:2602–2606 (1985).
39. P. Garbaczewski and R. Olkiewicz, Feynman-Kac kernels in Markovian representations of the Schrödinger interpolating dynamics, *J. Math. Phys.* **37**:732–751 (1996).
40. P. Garbaczewski and W. Karwowski, Impenetrable barrriers and canonical quantization, *Am. J. Phys.* **72**:924–933 (2004).
41. P. Garbaczewski, Perturbations of noise: Origins of isothermal flows, *Phys. Rev.* **E 59**:1498–1511 (1999).
42. P. Garbaczewski, Signatures of randomness in quantum spectra, *Acta Phys. Pol.* **A 33**:1001–1024 (2002).
43. P. Garbaczewski, Stochastic models of exotic transport, *Physica* **A 285**:187–198 (2000).
44. S. Goldstein and J. L. Lebowitz, On the (Boltzmann) entropy of non-equilibrium systems, *Physica* **D 193**:53–66 (2004).

45. M. J. W. Hall, Exact uncertainty relations, *Phys. Rev.* **A 64**:052103 (2001).
46. M. J. W. Hall, Universal geometric approach to uncertainty, entropy and infromation, *Phys. Rev.* **A 59**: 2602–2615 (1999).
47. J. J. Halliwell, Quantum-mechanical histories and the uncertainty principle: Information-theoretic inequalities, *Phys. Rev.* **D 48**:2739–2752 (1993).
48. R. V. L. Hartley, Transmission of information, *Bell Syst. Techn. J.* **7**:535–563 (1928).
49. H. Hasegawa, Thermodynamic properties of non-equilibrium states subject to Fokker-Planck equations, *Progr. Theor. Phys.* **57**:1523–1537 (1977).
50. T. Hatano and S. Sasa, Steady-State Thermodynamics of Langevin Systems, *Phys. Rev. Lett.* **86**:3463–3466 (2001).
51. I. I. Hirschman, A note on entropy, *Am. J. Math.* **79**:152–156 (1957).
52. B. Hu, et al., Quantum chaos of a kicked particle in an infinite potential well, *Phys. Rev. Lett.* **82**:4224–4227 (1999).
53. K. Huang, *Statistical Mechanics*, Wiley, New York (1987).
54. R. S. Ingarden, A. Kossakowski, and M. Ohya, *Information Dynamics and Open Systems*, Kluwer, Dordrecht (1997).
55. E. T. Jaynes, Information theory and statistical mechanics.II., *Phys. Rev.* **108**:171–190 (1957).
56. E. T. Jaynes, Violations of Boltzmann's H Theorem in Real Gases, *Phys. Rev.* **A 4**:747–750 (1971).
57. D.-Q. Jiang, M. Qian, and M-P. Qian, *Mathematical theory of nonequilibrium steady states*, LNM vol. **1833**, Springer-Verlag, Berlin (2004).
58. S. Kullback, *Information Theory and Statistics*, Wiley, NY (1959).
59. J. Kurchan, Fluctuation theorem for stochastic dynamics, *J. Phys. A: Math. Gen.* **31**:3719–3729 (1998).
60. A. Lasota and M. C. Mackey, *Chaos, Fractals and Noise*, Springer-Verlag, Berlin (1994).
61. J. L. Lebowitz and Ch. Maes, Entropy - a Dialogue, pp. 269–273, in: *On Entropy*, Eds. A. Grevau, G. Keller, G. Warnecke, Princeton University Press, Princeton, (2003).
62. Y. V. Linnik, An information-theoretic proof of the central limit theorem, *Theory Probab. App.* **4**:288–299 (1959).
63. H. Maasen and J. B. M. Uffink, Generalized Entropic Uncertainty Relations, *Phys. Rev. Lett.* **60**:1103–1106 (1988).
64. M. C. Mackey and M. Tyran-Kamińska, Effects of noise on entropy evolution, arXiv.org preprint cond-mat/0501092 (2005).
65. M. C. Mackey and M. Tyran-Kamińska, Temporal behavior of the conditional and Gibbs entropies, arXiv.org preprint cond-mat/0509649 (2005).
66. M. C. Mackey, The dynamic origin of increasing entropy, *Rev. Mod. Phys.* **61**, 981–1015 (1989).
67. V. Majernik and T. Opatrný, Entropic uncertainty relations for a quantum oscillator, *J. Phys. A: Math. Gen.* **29**:2187–2197 (1996).
68. V. Majernik and L. Richterek, Entropic uncertainty relations for the infinite well, *J. Phys. A: Math. Gen.* **30**: (1997), L49-L54.
69. P. G. L. Mana, Consistency of the Shannon entropy in quantum experiments, *Phys. Rev.* **A 69**:062108 (2004).
70. S. E. Massen and Panos C. P., Universal property of the information entropy in atoms, nuclei and atomic clusters, *Phys. Lett.* **A 246**:530–532 (1998).
71. S. E. Massen, et al., Universal property of information entropy in fermionic and bosonic systems, *Phys. Lett.* **A 299**:131–135 (2002).
72. S. E. Massen, Application of information entropy to nuclei, *Phys. Rev.* **C 67**:014314 (2003).
73. M. McClendon and H. Rabitz, Numerical simulations in stochastic mechanics, *Phys. Rev.* **A 37**:3479–3492 (1988).

74. T. Munakata, A. Igarashi, and T. Shiotani, Entropy and entropy production in simple stochastic models, *Phys. Rev.* **E 57**:1403–1409 (1998).
75. E. Nelson, *Dynamical Theories of the Brownian Motion*, Princeton University Press, Princeton, 1967.
76. R. G. Newton, What is a state in quantum mechanics?, Am. J. Phys. **72**:348–350 (2004).
77. M. Ohya and D. Petz, *Quantum Entropy and Its use*, Springer-Verlag, Berlin, 1993.
78. M. H. Partovi, Entropic formulation of uncertainty for quantum measurements, *Phys. Rev. Lett.* **50**:1883–1885 (1983).
79. J. Pipek and I. Varga, Universal classification scheme for the spatial-localization properties of one-particle states in finite, d-dimensional systems, *Phys. Rev.* **A 46**:3148–3163 (1992).
80. H. Qian, M. Qian, and X. Tang, Thermodynamics of the general diffusion process: time-reversibility and entropy production, *J. Stat. Phys.* **107**:1129–1141 (2002).
81. H. Qian, Mesoscopic nonequilibrium thermodynamics of single macromolecules and dynamic entropy-energy compensation, *Phys. Rev.* **E 65**:016102 (2001).
82. H. Risken, *The Fokker-Planck Equation*, Springer-Verlag, Berlin, 1989.
83. D. Ruelle, Positivity of entropy production in nonequilibrium statistical mechanics, *J. Stat. Phys.* **85**:1–23 (1996).
84. J. Sánchez-Ruiz, Asymptotic formula for the quantum entropy of position in energy eigenstates, *Phys. Lett.* **A 226**:7–13 (1997).
85. M. S. Santhanam, Entropic uncertainty relations for the ground state of a coupled sysytem, *Phys. Rev.* **A 69**:042301 (2004).
86. C. E. Shannon, A mathematical theory of communication, *Bell Syst. Techn. J.* **27**:379–423, 623–656 (1948).
87. J. D. H. Smith, Some observations on the concepts of information-theoretic entropy and randomness, *Entropy*, **3**:1–11 (2001).
88. K. Sobczyk, Information Dynamics: Premises, Challenges and Results, *Mechanical Systems and Signal Processing* **15**:475–498 (2001).
89. A. J. Stam, Some inequalities satisfied by the quantities of information of Fisher and Shannon, *Inf. and Control* **2**:101–112 (1959).
90. A. Stotland, et al., The information entropy of quantum mechanical states, *Europhys. Lett.* **67**:700–706 (2004).
91. G. Toscani, Kinetic approach to the asymptotic behaviour of the solution to diffusion equation, *Rend. di Matematica* Serie VII **16**:329–346 (1996).
92. J. Trębicki and K. Sobczyk, Maximum entropy principle and non-stationary distributions of stochastic systems, *Probab. Eng. Mechanics* **11**:169–178 (1996).
93. M. Tribus and R. Rossi, On the Kullback information measure as a basis for information theory: Comments on a proposal by Hobson and Chang, *J. Stat. Phys.* **9**:331–338 (1973).
94. S. A. Trigger, Quantum nature of entropy increase for wave packets, *Bull. Lebedev. Phys. Inst.* **9**:44–51 (2004).
95. I. Varga and J. Pipek, Rényi entropies characterizing the shape and the extension of the phase-space representation of quantum wave functions in disordered systems, *Phys. rev.* **E 68**:026202 (2003).
96. J. M. G. Vilar and J. M. Rubi, Thermodynamics beyond local equilibrium, *Proc. Nat. Acad. Sci. (NY)* **98**:11081–11084 (2001).
97. J. Voigt, Stochastic opertors, Information and Entropy, *Commun. Math. Phys.* **81**:31–38 (1981).
98. J. Voigt, The H-Theorem for Boltzmann type equations, *J. Reine Angew. Math* **326**:198–213 (1981).
99. A. Wehrl, General properties of entropy, *Rev. Mod. Phys.* **50**:221–260 (1978).
100. S. A. Werner, and H. Rauch, *Neutron interferometry: Lessons in Experimental Quantum Physics*, Oxford University Press, Oxford, 2000.

101. R. J. Yañez, et al., Entropic integrals of hyperspherical harmonics and spatial entropy of D-dimensional central potentials, *J. Math. Phys.* **40**:5675–5686 (1999).
102. R. J. Yañez, Van W. Assche, J. S. Dehesa, Position and information entropies of the D-dimensional harmonic oscillator and hydrogen atom, *Phys. Rev.* A 50:3065–3079 (1994).
103. A. M. Yaglom and I. M. Yaglom, *Probability and Information*, D. Reidel, Dordrecht, 1983.
104. A. Zeilinger, et al., Single- and double-slit diffraction of neutrons, *Rev. Mod. Phys.* **60**:1067–1073 (1988).