PHYSICAL REVIEW E VOLUME 49, NUMBER 6 JUNE 1994

ARTICLES

Limitation on entropy increase imposed by Fisher information

B. Nikolov
Department of Physics, K. Preslawski University, 9700 Shumen, Bulgaria

B. Roy Frieden
Optical Sciences Center, University of Arizona, Tucson, Arizona 85721
(Received 14 December 1993)

Consider a system obeying conservation of flow, as in classical particle flow or in relativistic quantum mechanics. In such cases a probability density function $p(\mathbf{r}|t)$ may be used to describe the system, where $\mathbf{r}$ is particle position and t is time. Let $H(t)$ be the Shannon form of the Boltzmann entropy corresponding to $p(\mathbf{r}|t)$. It is found that $(dH/dt)_{\max}=\frac{1}{6}I(t)d/dt\langle r^2(t)\rangle$, where $I(t)$ is the Fisher information about the centroid of the system, and $\langle r^2(t)\rangle$ is the time-dependent mean-square particle position. A corollary is that, for classical particle flow obeying $\langle \mathbf{r}\rangle=0$, positional uncertainty $\sigma(t)$ must ever increase with time.

PACS number(s): 05.40.+j, 03.65.Bz, 41.75.Ht, 47.10.+g

I. INTRODUCTION

The trace of the Fisher information matrix

$$I=\int d\mathbf{r}\nabla p\cdot\nabla p/p \;, \tag{1}$$

is a scalar information quantity that has been shown [1,2] to derive a host of physical phenomena. [The integration limits in Eq. (1) and all following equations are understood to be infinite.] This includes such diverse phenomena as the complex Schrödinger wave equation, the Maxwell-Boltzmann distribution law, and Maxwell's equations. In Eq. (1), p is a probability density function $p(x)$ for a coordinate x whose unit demarks the physical scenario: e.g., a length in quantum mechanics, or a velocity in classical particle statistics.

Equation (1) has the following meaning. Suppose that a physical scenario consists of a particle or a system of particles (as in quantum mechanics or in gas dynamics, respectively). A particle is measured to be at position y,

$$y=\theta+x \;, \tag{2a}$$

where θ is (say) the mean particle position and x is a random excursion from θ. It is desired to know θ, and so an estimate of θ is formed from the data value y. How much information resides in the observation y about the sought-after mean θ?

Intuitively, if y is, on average, close to θ, then there is a large amount of such information in a reading y. Then, by Eq. (2a), x is small, so that high information implies small x, on average. The probability density function (PDF) for x, $p(x)$, should be narrow and peaked about the origin. Conversely, if $p(x)$ is broad there should be small information present.

This trend is obeyed by the one-dimensional version of Fisher information (1):

$$I=\int dx(dp/dx)^2/p \;,\quad p=p(x) \;. \tag{2b}$$

[Note: The division by p in the integrand does not cause problems as $p\rightarrow 0$. See Eqs. (8) and (9) below.] For example, if $p(x)$ is Gaussian with variance σ^2, then Eq. (2b) gives simply

$$I=1/\sigma^2 \;. \tag{2c}$$

Again, a narrow PDF law (indicated by small σ) gives large information I; and a broad PDF gives small I.

Now the width of a PDF measures the *degree of disorder* that exists in predicting a value x. A narrow, highly peaked PDF indicates small disorder, since most of the time (predictably) small x will occur. Conversely, a broad, flattened PDF indicates large disorder.

Hence, quantity I is a dual measure. It indicates both the degree to which the ideal value of a random variable may be measure-estimated (to coin a word), and simultaneously, the degree of disorder in the variable.

A related measure is H, the Shannon entropy (defined below) of the system. It is also a dual measure. It measures the number of distinguishable signals that may be transmitted by a communication channel [3,4] and the degree of disorder (thermodynamic disorder, in most cases) of a system of particles. It is well known that maximization of H gives rise to some classical distribution functions, notably the Boltzmann and Maxwell-Boltzmann ones [5,6].

Since I and H have analogous statistical properties, this suggests that extremization of I should perhaps likewise derive certain physical distribution functions. In fact, this is the case [1,2]. It is important, however, to stress the different origins of the two measures. Information I arises in a measure-estimation channel (described above), whereas H arises in a communication channel. The wealth of physical phenomena that follow from extremization of I perhaps suggests that physical phenomena are more expressions of the ability to measure-estimate than to transmit distinguishable signals. In other words, physical phenomena arise out of measure-estimation channels, and not communication channels.

The preceding considerations defined static systems, i.e., systems whose statistics remain fixed with time. We next consider systems whose statistics change with time. If these systems are isolated as well, then they obey an equation of continuity of flow (see below). We will find that the resulting flux of particles allow entropy H and information I to be interrelated. The rate of increase of entropy will be found to be limited by the available level of Fisher information I.

II. DISORDER IN STOCHASTIC SYSTEMS

Consider a stochastic system consisting of one or more particles. Particle position $\mathbf{r}=(x,y,z)$ is random, and specified by a conditional probability law $p(\mathbf{r}|t)$. This represents the probability of a particle at position $\mathbf{r}$, conditional upon (at) time t. The broader $p(\mathbf{r}|t)$ is (at a fixed t) as a function of $\mathbf{r}$, the more equally probable are all $\mathbf{r}$ values, and hence, the higher is the state of *disorder* for the system at that time. A measure of disorder that has this property is the Shannon form [3] of the Boltzmann entropy,

$$H(t)=-\int d\mathbf{r}\,p(\mathbf{r}|t)\ln p(\mathbf{r}|t)\ . \tag{3a}$$

This is sometimes called "communication" entropy, because of its use in the field of communication theory [4]. For a normal law $p(\mathbf{r}|t)$, as in the case of Brownian motion, where

$$p(\mathbf{r}|t)=\left[\frac{1}{\sqrt{2\pi}\sigma(t)}\right]^3 e^{-r^2/2\sigma^2(t)}\ , \tag{3b}$$

use of Eq. (3a) gives

$$H(t)=\tfrac{3}{2}(1+\ln 2\pi)+3\ln\sigma(t)\ . \tag{3c}$$

This shows that H varies monotonically as the "width" σ of $p(\mathbf{r}|t)$, which makes sense for a measure of disorder.

Differentiation of (3c) gives

$$dH/dt\equiv H_t=3\sigma_t/\sigma \tag{4}$$

or, identically,

$$H_t=\left[\frac{1}{2}\frac{d\sigma^2}{dt}\right]\left[\frac{3}{\sigma^2}\right]\ . \tag{5}$$

(We use the convention that letter subscripts denote derivatives.) The dependence of H_t upon the first factor is intuitive: Both H and σ measure the width of $p(\mathbf{r}|t)$, the extent of disorder, so that it is plausible that their time rates of change should also be proportional.

For general statistics, it will be found that the maximum value $H_t^{\max}$ of H_t follows Eq. (5), where more generally, $d\sigma^2/dt$ is replaced by $d\langle r^2\rangle/dt$ and factor $3/\sigma^2$ is replaced by I, the trace (1) of the Fisher information matrix,

$$I=\int d\mathbf{r}\frac{\nabla p\cdot\nabla p}{p}\ ,\quad p\equiv p(\mathbf{r}|t)\ . \tag{6}$$

[That $I=3/\sigma^2$ for the Gaussian case (3b) is easily verified by direct substitution into (6).] The proviso is that the phenomenon obey properties (a)–(d) listed in Sec. III. Quantity I is the information in a single measurement of particle position about the centroid of the pattern $p(\mathbf{r}|t)$.

It is interesting to consider why the derivative $H_t^{\max}$ might go as I. Mathematically, a similar relation is [7]

$$H_v=\tfrac{1}{2}I\ . \tag{7}$$

This holds for a probability law $p(x)$, where $x=y+z$, y obeys any PDF, y and z are independent, and z is Gaussian with variance v. Relation (7) shows that I can indeed vary as the derivative of H, and under fairly general conditions. However, a more fundamental physical reason for such a tie-in is as follows.

We saw above that H is a measure of disorder. In fact, both H *and* I are measures of disorder. That I measures disorder can be seen as follows. Work with a function $q(\mathbf{r}|t)$ obeying

$$p(\mathbf{r}|t)\equiv q^2(\mathbf{r}|t) \tag{8}$$

in (6), which directly gives

$$I=4\int d\mathbf{r}\,\nabla q\cdot\nabla q\ . \tag{9}$$

Hence I measures the gradient content in q (or p). This ties in with disorder. Consider a function q that is concentrated about the point $\mathbf{r}=\mathbf{0}$. This exhibits low disorder, since then the point $\mathbf{r}=\mathbf{0}$ is much more probable than other points. Also, by normalization property

$$\int d\mathbf{r}\,p(\mathbf{r}|t)=1\ , \tag{10}$$

the function must be very 'tall' about $\mathbf{r}=\mathbf{0}$. But then it has large gradient content in this vicinity. Hence, its I value given by (9) will be large. The result is that low disorder is accompanied by a large I. Conversely, and by the same reasoning, high disorder is associated with a small I. Thus, I is a monotonic measure of disorder. (The fact that it is an inverse measure is inconsequential: this could be remedied by simply working with its negative.)

Since H and I are both measures of disorder, it is then not surprising that H or $H_t^{\max}$ could be expressed in terms of I. Another interesting tie-in between H and I is their predilection to derive distribution laws of physics when extremized (as discussed above). On this basis as well, it is plausible that I should enter into a limiting expression for H_t.

III. BOUNDARY CONDITIONS FOR AN ISOLATED SYSTEM

Consider a physical phenomenon involving discrete particle(s) that are in a general state of motion subject to the following conditions.

(a) The particles obey the equation of conservation of flow,

$$p_t(\mathbf{r}|t)+\nabla\cdot\mathbf{P}(\mathbf{r},t)=0\ , \tag{11}$$

where $\mathbf{P}$ is a measure of flow whose exact nature depends upon the application. Denote the components of $\mathbf{P}=(P_1,P_2,P_3)$. Numbered subscripts denote vector components, whereas letter subscripts denote derivatives.

(b) The system is isolated, so that there is no net flow across its boundaries,

$$\mathbf{P}(\mathbf{r},t)|_{\text{boundaries}}=\mathbf{0}\ . \tag{12}$$

Hence P obeys Dirichlet [8] boundary conditions. Furthermore, if the boundaries are at infinity then

$$\lim_{r\to\infty}\mathbf{P}(\mathbf{r},t)=\mathbf{0} \tag{13}$$

faster than $1/r^2$. It is noted that property (12) actually follows from the assumption of flow [Eq. (11)] and normalization condition (10). This is shown by integrating (11), $d\mathbf{r}$, to give

$$\int d\mathbf{r}\,\nabla\cdot\mathbf{P}=-\frac{\partial}{\partial t}\int d\mathbf{r}\,p\ .$$

The left-hand side is directly integrable to give $\mathbf{P}$ evaluated at the boundaries. Under condition (10) the right-hand side is zero. The combination gives (12).

(c) There is vanishing probability of a particle being on the boundary,

$$p(\mathbf{r}|t)|_{\text{boundaries}}=0\ . \tag{14}$$

Hence p also obeys Dirichlet boundary conditions. Furthermore, if the boundaries are at infinity,

$$\lim_{r\to\infty}p(\mathbf{r}|t)=0\ , \tag{15}$$

faster than $1/r^2$. The latter follows from the requirement (10) for normalization.

(d) Finally, condition (12) is so strong that

$$\mathbf{P}\ln p\,|_{\text{boundaries}}=\mathbf{0}\ . \tag{16}$$

That is, $\mathbf{P}\to\mathbf{0}$ faster than $\ln p\to-\infty$ at the boundaries.

On the basis of properties (a)–(d), the entropy $H(t)$ of the system has a time rate of change whose maximum value $H_t^{\max}(t)$ obeys

$$H_t^{\max}(t)=\tfrac{1}{6}I(t)\frac{d}{dt}\langle r^2(t)\rangle\ . \tag{17}$$

The proof follows.

IV. DERIVATION

The partial derivative $\frac{\partial}{\partial t}$ of Eq. (3a) gives

$$\begin{aligned}H_t&=-\frac{\partial}{\partial t}\int d\mathbf{r}\,p\ln p\\&=-\int d\mathbf{r}\,p_t\ln p-\int d\mathbf{r}\,p(1/p)p_t\end{aligned} \tag{18}$$

after differentiating under the integral sign. The second right-hand integral gives

$$\int d\mathbf{r}\,p(1/p)p_t=\frac{\partial}{\partial t}\int d\mathbf{r}\,p=0$$

by normalization requirement (10). Next, use the flow equation (11) in the first right-hand integral of Eq. (18). This gives

$$\begin{aligned}H_t&=\int d\mathbf{r}\,\nabla\cdot\mathbf{P}\ln p\\&\equiv\int\int\int dz\,dy\,dx\left[\frac{\partial}{\partial x}(P_1)+\frac{\partial}{\partial y}(P_2)\right.\\&\qquad\left.+\frac{\partial}{\partial z}(P_3)\right]\ln p\ .\end{aligned} \tag{19}$$

Consider the first right-hand term in (19). The innermost integral is

$$\int dx\frac{\partial}{\partial x}(P_1)\ln p=P_1\ln p\,|_{\text{boundaries}}-\int dx(P_1/p)p_x$$

after integrating by parts,

$$=0-\int dx(P_1/p)p_x$$

by Dirichlet condition (16). Similar steps follows for the second and third right-hand terms in Eq. (19). The result is

$$H_t=-\int d\mathbf{r}\,\mathbf{P}\cdot\nabla p/p\ . \tag{20}$$

This is identical to

$$H_t=-\int d\mathbf{r}(\mathbf{P}/\sqrt{p}\,)\cdot(\sqrt{p}\,\nabla p/p)\ . \tag{21}$$

Hence H_t is expressed as the inner product over space $d\mathbf{r}$ of two vectors $(\mathbf{P}/\sqrt{p}\,)$ and $(\sqrt{p}\,\nabla p/p)$. By the Schwarz inequality, the magnitude of H_t is maximized when the two vectors are parallel, i.e., when each component in $\mathbf{r}$ space is proportional:

$$\mathbf{P}/\sqrt{p}=a\sqrt{p}\,\nabla p/p\ ,\quad a=a(t)\ . \tag{22}$$

Since the space of $\mathbf{r}$ does not include coordinate t, a is in general a function of t. Equation (22) simplifies to

$$\mathbf{P}=a\nabla p\ . \tag{23}$$

Substituting form (23) for $\mathbf{P}$ into Eq. (21) gives its maximum value,

$$H_t^{\max}=-a\int d\mathbf{r}\,\nabla p\cdot\nabla p/p=-aI \tag{24}$$

by definition (6). Thus the maximum rate of entropy increase for an isolated system is limited by the Fisher information level. Quantity $a(t)$ is next related to the width of $p(\mathbf{r}|t)$.

V. EVALUATION OF FACTOR $a(t)$

By Eq. (23),

$$\mathbf{P}\cdot\mathbf{r}=a\nabla p\cdot\mathbf{r}\ .$$

Integrating this $d\mathbf{r}$ gives

$$\int d\mathbf{r}\,\mathbf{P}\cdot\mathbf{r}=a\int d\mathbf{r}\,\nabla p\cdot\mathbf{r}$$
$$\equiv a\int\int\int dz\,dy\,dx(p_x x+p_y y+p_z z)\ . \qquad (25)$$

Consider the first right-hand term. Integrating by parts,

$$\int dx\,p_x x=xp|_{\text{boundaries}}-\int dx\,p(x,y,z|t)$$
$$=0-p(y,z|t)$$

by Dirichlet properties (14) and (15). Then

$$\int\int\int dz\,dy\,dx\,p_x x=-\int\int dy\,dz\,p(y,z|t)=-1\ . \qquad (26)$$

Similar identifies hold for the second and third right-hand terms of (25). The result is that

$$\int d\mathbf{r}\,\mathbf{P}\cdot\mathbf{r}=-3a$$

or

$$a=-\tfrac{1}{3}\int d\mathbf{r}\,\mathbf{P}\cdot\mathbf{r}\ . \qquad (27)$$

The latter integral is now shown to be related to the second moment of the law $p(\mathbf{r}|t)$,

$$\langle r^2\rangle=\int d\mathbf{r}\,r^2 p(\mathbf{r}|t)\ . \qquad (28)$$

Differentiate this as

$$\frac{\partial}{\partial t}\langle r^2\rangle=\int d\mathbf{r}\,r^2 p_t$$
$$=-\int d\mathbf{r}\,r^2\nabla\cdot\mathbf{P}$$

by flow Eq. (11),

$$\equiv-\int\int\int dz\,dy\,dx(x^2+y^2+z^2)\times\left[\frac{\partial P_1}{\partial x}+\frac{\partial P_2}{\partial y}+\frac{\partial P_3}{\partial z}\right]\ . \qquad (29)$$

Consider the contribution of term x^2 to the integral. First, regard its product by $\partial P_1/\partial x$. Integrating by parts,

$$\int dx\,x^2\frac{\partial P_1}{\partial x}=x^2P_1|_{\text{boundaries}}-2\int dx\,xP_1$$
$$=0-2\int dx\,xP_1 \qquad (30)$$

by boundary conditions (12) and (13). Next, regard the contribution of term $x^2\,\partial P_2/\partial y$ in (29). Switching the order of integration, we first integrate

$$\int dy\,x^2\frac{\partial P_2}{\partial y}=x^2P_2|_{\text{boundaries}}=0 \qquad (31)$$

by boundary conditions (12) and (13). The term $x^2\,\partial P_3/\partial z$ will contribute zero in the same way.

Likewise the term y^2 contributes

$$-2\int dy\,yP_2$$

to the integral dy [compare with (30)], etc. for the term in z^2. The result is that

$$\frac{\partial}{\partial t}\langle r^2\rangle=2\int d\mathbf{r}\,\mathbf{P}\cdot\mathbf{r}\ . \qquad (32)$$

We may now combine results. By Eqs. (27) and (32),

$$a=a(t)=-\tfrac{1}{6}\frac{\partial}{\partial t}\langle r^2(t)\rangle=-\tfrac{1}{6}\frac{d}{dt}\langle r^2(t)\rangle\ . \qquad (33)$$

The partial derivative becomes the ordinary derivative in this single argument t case. Then Eq. (24) becomes

$$\frac{dH^{\max}(t)}{dt}=\tfrac{1}{6}I(t)\frac{d}{dt}\langle r^2(t)\rangle\ . \qquad (34)$$

This shows that the rate of increase of entropy is limited jointly by two effects: (i) the rate of expansion of the width of distribution $p(\mathbf{r}|t)$; and (ii) the Fisher information in $p(\mathbf{r}|t)$. The latter is, in particular, the information in a single data position reading about the centroid, or mean, position in the pattern $p(\mathbf{r}|t)$. This information varies with time, as the pattern $p(\mathbf{r}|t)$ changes its shape.

Regarding shape, an interesting corollary of Eq. (34) follows when H, the Shannon entropy (3a), is also the Boltzmann entropy for the system. Then H obeys the second law of thermodynamics, so that the left-hand side of (34) is positive (or zero). Also, by definition (6), so is I. Then (34) gives

$$\frac{d}{dt}\langle r^2(t)\rangle\geq 0\ . \qquad (35)$$

This states a general theorem about the motion of (now) thermodynamic particles in an isolated system obeying continuity of flow. The particle mean-square spread must increase, or remain fixed, at each instant of time. It can never decrease toward the origin.

VI. APPLICATION TO CLASSICAL PARTICLE FLOW

Consider an isolated system of one or more material particles moving under the influence of mutual attraction or repulsion, due to gravitational, electrical, or other properties. The particles may also collide with each other and/or with the walls of a confining container. If the particles are confined to a container, then the container walls are physical boundaries for the system. In this scenario the flow vector is [9]

$$\mathbf{P}(\mathbf{r},t)=p(\mathbf{r}|t)\mathbf{v}(\mathbf{r},t)\ , \qquad (36)$$

where $\mathbf{v}$ is particle velocity. Results (34) and (35) will hold for such a system, provide conditions (11)–(16) hold. We show next that these conditions hold.

Since the system is isolated, there is no net flow of particles in or out. Then conditions (11) and (12) hold by definition. Moreover, normalization (10) holds because it is consistent with (12) (as previously shown).

Condition (12) is now

$$p(\mathbf{r}|t)\mathbf{v}(\mathbf{r},t)|_{\text{boundaries}}=\mathbf{0}\ . \qquad (37)$$

We have to decide which of the two constituents p or $\mathbf{v}$ is zero at the boundaries. If $\mathbf{v}=0$ but $p\neq 0$ at the boundaries, this implies that once a particle locates at a boundary it cannot move away from it. And since $p\neq 0$ there, ultimately every particle will locate there. The result would be a degenerate collapse of the system into a state where all particles remain piled up at the boundaries. We do not consider this kind of specialized system. Hence, the alternative solution to (37) is taken generally to be $\mathbf{v}\neq 0$, but

$$p(\mathbf{r}|t)|_{\text{boundaries}}=0 \; . \tag{38}$$

Since normalization condition (10) holds,

$$\lim_{r\to\infty} p(\mathbf{r}|t)=0 \tag{39}$$

faster than $1/r^2$. Hence, condition (13) holds.

Conditions (14) and (15) are now the same as conditions (38) and (39), respectively, which already have been shown to hold.

Finally, condition (16) holds because

$$\mathbf{P}\ln p=\mathbf{v}p\ \ln p$$

and of course

$$\lim_{p\to 0} p\ \ln p=0 \; . \tag{40}$$

The zero limit for p is taken because of (38).

Hence, results (34) and (35) hold for this system of particles.

VII. QUANTUM-MECHANICAL PARTICLE

A particle is moving under the influence of a general scalar potential function $V(\mathbf{r})$. Its speed can be comparable to that of light. It is confined to a space that has finite or infinite bounds. Such a particle is known to obey an equation of continuity of flow Eq. (11). Also, both $p(\mathbf{r}|t)$ and $\mathbf{P}(\mathbf{r},t)$ are quadratic in a wave function $\psi(\mathbf{r}|t)$. The exact nature of these relations depends upon the type of particle present. For the Dirac spin-$\frac{1}{2}$ particle [10]

$$p(\mathbf{r}|t)=\psi^*(\mathbf{r}|t)\psi(\mathbf{r}|t) \tag{41a}$$

and

$$\mathbf{P}(\mathbf{r},t)=-c\psi^*(\mathbf{r}|t)[\alpha]\psi(\mathbf{r}|t) \; . \tag{41b}$$

Wave function $\psi(\mathbf{r}|t)$ is a 4-vector. Also, ψ^* denotes the Hermitian conjugate of ψ, and $[\alpha]$ is the usual 3-vector of 4×4 matrices α_1, α_2, and α_3, where the subscripts denote x, y, and z components, respectively. We next show that boundary conditions (12)–(16) hold for such a system.

Since the particle cannot exist beyond the boundary, the potential $V(\mathbf{r})$ must be infinite there. Also, the wave function ψ must continuously approach zero as $\mathbf{r}$ approaches the boundary [11],

$$\lim_{r\to \text{boundary}} \psi(\mathbf{r}|t)=0 \; . \tag{42}$$

Then since both p and $\mathbf{P}$ increase quadratically as ψ in Eqs. (41a) and (41b), necessarily

$$p(\mathbf{r}|t)|_{\text{boundary}}=0 \tag{43a}$$

and

$$\mathbf{P}(\mathbf{r},t)|_{\text{boundary}}=0 \; . \tag{43b}$$

Hence, boundary conditions (12) and (14) are obeyed. Also, because of Eq. (43b) normalization condition (10) follows, as previously shown. This, in turn, implies property (15). Properties (13) and (16) remain to be shown.

We turn to property (13), which requires a boundary at infinity. Because property (15) holds, by Eq. (41a) ψ must fall off with r as $1/r$ or faster, and hence by Eq. (41b) $\mathbf{P}$ falls off with r as $1/r^2$ or faster. Hence (13) is obeyed.

Finally, we consider property (16). By Eq. (41b), the x component P_1 of $\mathbf{P}$ obeys

$$P_1=-c(\psi_1\psi_2\psi_3\psi_4)^*\begin{bmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{bmatrix}\begin{bmatrix}\psi_1\\\psi_2\\\psi_3\\\psi_4\end{bmatrix}, \tag{44}$$

where we used the known [12] 4×4 matrix for α_1. After the indicated matrix products in (44), the result is

$$P_1=-c(\psi_1^*\psi_4+\psi_2^*\psi_3+\psi_3^*\psi_2+\psi_4^*\psi_1) \; .$$

Then, also using Eq. (41a),

$$P_1\ln p=-c(\psi_1^*\psi_4+\psi_2^*\psi_3+\psi_3^*\psi_2+\psi_4^*\psi_1)\times\ln(|\psi_1|^2+|\psi_2|^2+|\psi_3|^2+|\psi_4|^2) \; . \tag{45}$$

According to condition (16), we are interested in the limiting form of this expression as $r\to$ boundaries, where by Eq. (43a) all components $\psi_i=0$. We may first set $\psi_2=\psi_3=0$ in (45), giving

$$P_1\ln p|_{\text{boundary}}=-2c\ \mathrm{Re}(\psi_1^*\psi_4)\ln(|\psi_1|^2+|\psi_4|^2) \; . \tag{46}$$

The right-hand side is of the form 0 ln 0, and so has to be evaluated by a limiting process.

Denote a given boundary point by $\mathbf{R}$. Expand each ψ_1 and ψ_4 in Taylor series about point $\mathbf{R}$, dropping all quadratic and higher terms since limit $\mathbf{r}\to\mathbf{R}$ will be taken. The result is

$$\psi_i(\mathbf{r}|t)=\psi_i(\mathbf{R}|t)+d\mathbf{r}\cdot\nabla\psi_i(\mathbf{R}|t) \; , \quad d\mathbf{r}=\mathbf{R}-\mathbf{r} \; , \quad i=1,4 \; , \tag{47}$$

where

$$\lim d\mathbf{r}\to 0 \tag{48}$$

now defines the boundary. The first right-hand term in (47) is zero, since all $\psi_i=0$ on the boundary. Then, substituting Eq. (47) into (46) gives

$$P_1\ln p|_{\text{boundary}}=-2c\ \mathrm{Re}[(d\mathbf{r}\cdot\nabla\psi_1^*)(d\mathbf{r}\cdot\nabla\psi_4)]\times\ln[|d\mathbf{r}\cdot\nabla\psi_1|^2+|d\mathbf{r}\cdot\nabla\psi_4|^2]|_{d\mathbf{r}=0} \; . \tag{49}$$

Taking the limit $dy=dz=0$ gives

$$P_1 \ln p|_{\text{boundary}} = -2c(dx)^2 \text{Re}[(\psi_{1x}^*)(\psi_{4x})] \ln[(dx)^2|\psi_{1x}|^2 + (dx)^2|\psi_{4x}|^2]|_{dx=0} . \tag{50}$$

This is of the form (Au) ln (Bu), $u=(dx)^2$. In the limit $u \to 0$ it gives 0, as at Eq. (40). Retracing steps (44)–(50) for either of the other components P_2, P_3 gives the same result.

In summary, the equation of flow (11) and the boundary value conditions (12)–(16) hold, so that result (34) follows. [Note: However, Eq. (35) does not hold here, because in quantum mechanics H of Eq. (1) does not represent Boltzmann entropy [13], a requisite for derivation of (35).] Result (34) states that, in quantum mechanics, the maximum rate of change of disorder $H_t^{\max}$ is limited by the Fisher measure I of disorder.

VIII. SUMMARY

A system that obeys conservation of flow (11) and Dirichlet boundary conditions (12)–(16) has a maximum possible rate of entropy increase obeying Eq. (34). This shows that the maximum rate is jointly proportional to the rate of increase of the second moment, or spread, of the probability law $p(\mathbf{r}|t)$, and to the Fisher information I. The latter is the information in an observed position $\mathbf{r}$ about the centroid position of the system. This result follows both for classical particle flow and for a particle obeying Dirac's formulation of relativistic quantum mechanics.

A corollary of Eq. (34) is that, for classical particle flow where the mean particle position is zero, the uncertainty in locating a particle can never decrease with time. This follows from additional use of the second law of thermodynamics, and represents a case where a thermodynamic law limits the ability to locate or measure.

We regard result (34) as an addendum to the second law. That is, entropy shall increase, but (now) not by too much.

There are many other phenomena obeying Eq. (11) of flow, such as, e.g., charge-current flow in electromagnetism. To the extent that boundary conditions (12)–(16) are obeyed, result (34) will follow for these phenomena as well.

[1] B. R. Frieden, Phys. Rev. A **41**, 4265 (1990).
[2] B. R. Frieden, Physica A **180**, 262 (1993).
[3] C. E. Shannon, Bell System Tech. J. **27**, 379 (1948); **27**, 623 (1948).
[4] F. M. Reza, *An Introduction to Information Theory* (McGraw-Hill, New York, 1961).
[5] J. W. Gibbs, *The Scientific Papers* (Dover, New York, 1961), Vol. I.
[6] E. T. Jaynes, Am. J. Phys. **33**, 391 (1965).
[7] N. M. Blachman, IEEE Trans. Inf. Theory **IT-11**, 267 (1965).
[8] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Part I.
[9] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* [Ref. (8)], p. 172.
[10] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics,* [Ref. (8)], p. 261.
[11] L. I. Schiff, *Quantum Mechanics,* 2nd ed. (McGraw-Hill, New York, 1955), pp. 29 and 30.
[12] L. I. Schiff, *Quantum Mechanics,* 2nd ed. (McGraw-Hill, New York, 1955), p. 326.
[13] R. H. Fowler, *Statistical Mechanics,* 2nd ed. (Cambridge University Press, London, 1966), pp. 43, 72, and 73.