CHAPTER 6

INFORMATION AND ENTROPY

Information and entropy have been discussed in Sections 1.3.2, 3.3.1, 4.7, etc. and we continue with a further elaboration (see in particular [**10, 23, 72, 146, 173, 174, 175, 240, 343, 388, 396, 400, 431, 446, 452, 481, 512, 634, 637, 639, 694, 740, 749, 755, 765, 766, 856, 914, 916, 915, 906, 976**]). As before we will again encounter relations to the quantum potential which serves as a persistent theme of development. There is an enormous literature on entropy and we try to select aspects which fit in with ideas of quantum diffusion and information theory.

1. THE DYNAMICS OF UNCERTAINTY

We begin with some topics from [**396**] to which we refer for certain tutorial aspects. Given events A_j $(1 \leq j \leq N)$ with probabilities μ_j of occurance in some game of chance with N possible outcomes one calls $log(\mu_j)$ an uncertainty function for A_j. We write the natural logarithm as log and recall that e.g. $log_2(b) = log(b)/ln(2)$ (the information theoretic base is taken as 2 in some contexts). The quantity (Shannon entropy)

$$\mathfrak{S}(\mu) = -\sum_1^N \mu_j log_2(\mu_j) \tag{1.1}$$

stands for the measure of the mean uncertainty of the possible outcomes of the game and at the same time quantifies the mean information which is accessible from an experiment (i.e. actually playing the game). Thus if one identifies the A_i as labels for discrete states of a system (1.1) can be interpreted as a measure of uncertainty of the state before this state is chosen and the Shannon entropy is a measure of the degree of ignorance concerning which possibility (event A_j) may hold true in the set of all $A_i^{'s}$ with a given a priori probability distribution (μ_i). Note also that $0 \leq \mathfrak{S}(\mu) \leq log_2(N)$ (since certainty means one entry with probability 1 and maximum uncertainty occurs when all events are equally probable with $\mu_j = 1/N$). There is some discussion of the Boltzman law $\mathfrak{S} = k_B log(W) = -k_B log(P)$ $(P = 1/W)$ and its relation to Shannon entropy, coarse graining, and differential entropy defined as

$$\mathfrak{S}(\rho) = -\int \rho(x) log(\rho(x)) dx \tag{1.2}$$

(cf. Sections 1.1.6 and 1.1.8). One recalls also the vonNeumann entropy

$$\mathfrak{S}(\hat{\rho}) = -k_B Tr(\hat{\rho} log(\hat{\rho})) \tag{1.3}$$

where $\hat{\rho}$ is the density operator for a quantum state ($\hat{\rho}log(\rho)$ is defined via functional calculus for selfadjoint operators (cf. [**916**]). For diagonal density operators with eigenvalues p_i this will coincide with the Shannon entropy $\sum p_i log(p_i)$. We go now directly to an extension of the discussion in Sections 1.1.6-1.1.8. It is known from [**862**] that among all one dimensional distributions $\rho(x)$ with a finite mean, subject to the condition that the standard deviation is fixed at σ, it is the Gaussian with half width σ which sets a maximum of the differential entropy. Thus for the Gaussian with $\rho(x) = (1/\sigma\sqrt{2\pi})exp[-(x-x_0)^2/2\sigma^2]$ one has

$$\mathfrak{S}(\rho) \leq \frac{1}{2}log(2\pi e\sigma^2) \Rightarrow \frac{1}{\sqrt{2\pi e}}exp[\mathfrak{S}(\rho)] \leq \sigma \tag{1.4}$$

A result of this is that the major role of the differential entropy is to be a measure of localization in the configuration space (note that even for relatively large mean deviations $\sigma < 1/\sqrt{2\pi e} \simeq .26$ the differential entropy $\mathfrak{S}(\rho)$ is negative. Consider now a one parameter family of probability densities $\rho_\alpha(x)$ on $\mathbf{R}$ whose first (mean) and second moments (variance) are finite. Write $\int x\rho_\alpha(x)dx = f(\alpha)$ with $\int x^2\rho_\alpha dx < \infty$. Under suitable hypotheses (implying that $\partial\rho_\alpha/\partial\alpha$ is bounded by a function $G(x)$ which together with $xG(x)$ is integrable on $\mathbf{R}$) one obtains

$$\int (x-\alpha)^2\rho_\alpha(x)dx \cdot \int \left(\frac{\partial log(\rho_\alpha)}{\partial\alpha}\right)^2 \rho_\alpha dx \geq \left(\frac{df(\alpha)}{d\alpha}\right)^2 \tag{1.5}$$

which results from

$$\frac{df}{d\alpha} = \int [(x-\alpha)\rho_\alpha^{1/2}]\left[\frac{\partial(log(\rho_\alpha))}{\partial\alpha}\rho_\alpha^{1/2}\right] dx \tag{1.6}$$

and the Schwartz inequality. Assume now that the mean value of ρ_α actually is α and fix at σ^2 the value of the variance $<(x-\alpha)^2>=<x^2>-\alpha^2$. Then (1.5) takes the familiar form

$$\mathfrak{F}_\alpha = \int \frac{1}{\rho_\alpha}\left(\frac{\partial\rho_\alpha}{\partial\alpha}\right)^2 dx \geq \frac{1}{\sigma^2} \tag{1.7}$$

where the left side is the Fisher information for ρ_α. This says that the Fisher information is a more sensitive indicator of the wave packet localization than the entropy power in (1.4). Consider now $\rho_\alpha = \rho(x-\alpha)$ so $\mathfrak{F}_\alpha = \mathfrak{F}$ is no longer dependent on α and one can transform this to the QM form (up to a factor of D^2 where $D = \hbar/2m$ which we acknowledge here via the symbol $\sim$)

$$\frac{1}{2}\mathfrak{F} = \frac{1}{2}\int \frac{1}{\rho}\left(\frac{\partial\rho}{\partial x}\right)^2 dx \sim \int \rho\cdot\frac{u^2}{2}dx \sim -<\tilde{Q}> \tag{1.8}$$

where $u = \nabla log(\rho)$ is the osmotic velocity field and the average $<\tilde{Q}>=\int\rho\cdot\tilde{Q}dx$ involves the quantum potential $\tilde{Q} = 2(\Delta\sqrt{\rho}/\sqrt{\rho})$ (cf. equations (6.1.13) - (6.1.16) where $Q = -(\hbar^2/2m)(\Delta\sqrt{\rho}/\sqrt{\rho})$, $\tilde{Q} = -(1/m)Q$, $D = \hbar/2m$, and $u = D\nabla log(\rho))$. Consequently $-<\tilde{Q}>\geq(1/2\sigma^2)$ for all relevant probability densities with any finite mean (with variance fixed at σ^2). We continue in this section with the notation $\tilde{Q} = 2(\Delta\sqrt{\rho}/\sqrt{\rho})$ and note that $D^2\tilde{Q}$ is in fact the correct $\tilde{Q} = -(1/m)Q$ (which occasionally arises here as well, in a diffusion context).

Next one defines the Kullback entropy $K(\theta, \theta')$ for a one parameter family of probability densities ρ_θ so that the distance between any two densities can be directly evaluated. Let $p_{\theta'}$ be the reference density and one writes

$$K(\theta, \theta') = K(\rho_\theta | \rho_{\theta'}) = \int \rho_\theta(x) log \frac{\rho_\theta(x)}{\rho_{\theta'}(x)} dx \tag{1.9}$$

(note this is positive and sometimes one refers to $\mathfrak{H}_c = -K$ as a conditional entropy). If one takes $\theta' = \theta + \Delta\theta$ with $\Delta\theta << 1$ then under a number of standard assumptions

$$K(\theta, \theta + \Delta\theta) \simeq \frac{1}{2} \mathfrak{F}_\theta \cdot (\Delta\theta)^2 \tag{1.10}$$

where $\mathfrak{F}_\theta$ denotes the Fisher information measure as in (1.7). More generally for a two parameter family $\theta \sim (\theta_1, \theta_2)$ of densities one has

$$K(\theta,\ \theta + \Delta\theta) \simeq \frac{1}{2} \sum \mathfrak{F}_{ij} \Delta\theta_i \Delta\theta_j;\ \mathfrak{F}_{ij} = \int \rho_\theta \frac{\partial log(\rho_\theta)}{\partial\theta_i} \frac{\partial log(\rho_\theta)}{\partial\theta_j} dx \tag{1.11}$$

For Gaussian densities at fixed σ with $\theta = \alpha$ one has then $K(\alpha, \alpha + \Delta\alpha) \simeq (\Delta\alpha)^2/2\sigma^2$. Various related formulas are derived and in particular one relates the Shannon entropy for a coarse grained density ρ_B to the differential entropy of the density ρ leading to a formula $\mathfrak{S}(\rho_B) - \mathfrak{S}(\rho'_B) \simeq \mathfrak{S}(\rho) - \mathfrak{S}(\rho')$. One considers also spatial Markov diffusion processes in $\mathbf{R}$ with a diffusion coefficient D which drive space-time inhomogeneous probability density densities $\rho(x,t)$. For example a free Brownian motion characterized by $v = -u = -D\nabla log(\rho(x,t))$ and diffusion current $j = v \cdot \rho$ obeys the continuity equation $\partial_t \rho = -\nabla j$ which is equivalent to the heat equation. As in Sections 1.1.6-1.1.8 and 6.1 we have the important relations

$$\tilde{Q} = 2D^2 \frac{\Delta\rho^{1/2}}{\rho^{1/2}} = \frac{1}{2} u^2 + D\nabla \cdot u;\ \partial_t v + (v \cdot \nabla) v = -\nabla \tilde{Q} \tag{1.12}$$

A straightforward generalization refers to a diffusive dynamics of a mass m in a conservative force field $F = -\nabla V$. The associated Smoluchowski diffusion with a forward drift $b(x) = F/m\beta$ is analyzed in terms of a Fokker-Planck (FP) equation $\partial_t \rho = D\Delta\rho - \nabla(b \cdot \rho)$ with initial data $\rho_0(x) = \rho(x, 0)$. For standard Brownian motion in an external force field one has $D = k_B T/m\beta$ where $\beta \sim$ friction, T is temperature and k_B is the Boltzman constant. With suitable hypotheses one has the following compatibility equations in the form of hydrodynamical conservation laws

$$\partial_t \rho + \nabla(v\rho) = 0;\ (\partial_t + v \cdot \nabla) v = \nabla(\Omega - \tilde{Q}) \tag{1.13}$$

where $\Omega(x)$ is the volume potential for the process, namely

$$\Omega = \frac{1}{2} \left(\frac{F}{m\beta} \right)^2 + D\nabla \cdot \left(\frac{F}{m\beta} \right) \tag{1.14}$$

Here$v = b - u = (F/m\beta) - D(\nabla\rho/\rho)$ defines the current velocity of Brownian partices in an external force field. With a solution ρ of the FP equation one associates a differential entropy $\mathfrak{S}(t) = -\int \rho log(\rho) dx$ which is typically not conserved.

With boundary conditions on ρ, $v\rho$, and $b\rho$ involving vanishing at boundaries or at infinity one obtains

$$\frac{d\mathfrak{S}}{dt} = \int \left[\rho(\nabla \cdot b) + D\frac{(\nabla\rho)^2}{\rho}\right] dx \tag{1.15}$$

One emphasizes that it is not obvious whether the differential entropy grows, decreases, or whatever. One can rewrite (1.15) in the forms

$$D\dot{\mathfrak{S}} = D < \nabla \cdot b > + < u^2 >= D < \nabla \cdot v >; \tag{1.16}$$

$$D\dot{\mathfrak{S}} =< v^2 > - < b \cdot v >= - < v \cdot u >$$

where $< \ >$ denotes the mean value relative to ρ. For $b = F/m\beta$ and $j = v\rho$ this leads to a characteristic "power release" expression

$$\frac{d\tilde{Q}}{dt} = \frac{1}{D}\int \frac{1}{m\beta} F \cdot jcx = \frac{1}{D} < b \cdot v > \tag{1.17}$$

Again $\dot{Q}$ can be positive (power removal) or negative (power absorption). In thermodynamic terms one deals here with the time rate at which the mechanical work per unit of mass is dissipated (removed from the reservoir) in the form of heat in the course of the Smoluchowski diffusion process - i.e. $k_B T\dot{\tilde{Q}} = \int F \cdot jdx$ where T is the temperature of the bath. For $b = 0$ (no external forces) one has $D\dot{\mathfrak{S}} = D^2 \int[(\nabla\rho)^2/\rho]dx = D^2\mathfrak{F} = -D^2 < \tilde{Q} >$ and one can also write

$$\frac{d\mathfrak{S}}{dt} = \left(\frac{d\mathfrak{S}}{dt}\right)_{in} - \frac{d\tilde{Q}}{dt} \tag{1.18}$$

from (1.15) and (1.16) (here $(\dot{\mathfrak{S}})_{in} = (1/D) < v^2 >$.

One goes now to mean energy and the dynamics of Fisher information and considers $-\rho$ and s where $v = \nabla s$ as canonically conjugate fields; then one can use variational calculus to derive the continuity and FP equations together with the HJ type equations whose gradient gives the hydrodynamical conservation law

$$\partial_t s + (1/2)(\nabla s)^2 - (\Omega - \tilde{Q}) = 0 \tag{1.19}$$

Here the mean Lagrangian is

$$\mathfrak{L} = -\int \rho \left[\partial_t s + \frac{1}{2}(\nabla s)^2 - \left(\frac{u^2}{2} + \Omega\right)\right] dx \tag{1.20}$$

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

$$\mathfrak{H} = \int \rho \left[\frac{1}{2}(\nabla s)^2 - \left(\frac{u^2}{2} + \Omega\right)\right] dx = \frac{1}{2}(< v^2 > - < u^2 >) - < \Omega > \tag{1.21}$$

(note here $v = \nabla s$ satisfies $v = b - u$ with $u = D\nabla log(\rho)$ and we refer to Section 1.1 for clarification). One defines a thermodynamic force $F_{th} = v/D$ associated with the Smoluchowski diffusion with a corresponding potential $-\nabla\Psi = k_B T F_{th} = F - k_B T\nabla log(\rho)$ so in the absence of external forces $F_{th} = -\nabla log(\rho) = -(1/D)u$. The mean value of the thermodynamic force associates with the diffusion process an analogue of the Helmholz free energy $< \Psi >=< V > -T\mathfrak{S}_G$ where the dimensional version $\mathfrak{S}_G = k_B\mathfrak{S}$ of information entropy has been introduced (it is a configuration space analogue of the Gibbs entropy). Here the term $< V >$ plays the role of (mean) internal energy and assuming ρv vanishes at boundaries (or infinity) one obtains the time rate of change of Helmholz free energy at a constant temperature, namely

(1.22)
$$\frac{d}{dt} < \Psi >= -k_B T\dot{\tilde{Q}} - T\dot{\mathfrak{S}}_G \Rightarrow \frac{d}{dt} < \Psi >= -(k_B T)\left(\frac{d\mathfrak{S}}{dt}\right)_{in} = -(m\beta) < v^2 >$$

Now one can evaluate an expectation value of (1.19) which implies an identity $\mathfrak{H} = - < \partial_t s >$. Then using $\Psi = V + k_B T log(\rho)$ (with time independent V) one arrives at $\dot{\Psi} = (k_B T/\rho)\nabla(v\rho)$ and since $v\rho = 0$ at integration boundaries we get $< \dot{\Psi} >= 0$. Since $v = -(1/m\beta)\nabla\Psi$ define then $s(x,t) = (1/m\beta)\Psi(x,t)$ so that $< \partial_t s >= 0$ and hence $\mathfrak{H} = 0$ identically. This gives an interplay between the mean energy and the information entropy production rate in the form

$$\frac{D}{2}\left(\frac{d\mathfrak{S}}{dt}\right)_{in} = \frac{1}{2} < v^2 >= \int \rho\left(\frac{u^2}{2} + \Omega\right) dx \geq 0 \tag{1.23}$$

Next recalling (1.7)-(1.8) and setting $\mathfrak{F} = D^2\mathfrak{F}_\alpha$ one obtains

$$\mathfrak{F} =< v^2 > -2 < \Omega >\geq 0 \tag{1.24}$$

where $(1/2)\mathfrak{F} = - < \tilde{Q} >$ holds for probability densities with finite mean and variance. One also derives the following formulas (under suitable hypotheses)

$$\partial_t(\rho v^2) = -\nabla\cdot[(\rho v^3)] - 2\rho v\cdot\nabla(\tilde{Q} - \Omega); \tag{1.25}$$

$$\frac{d}{dt} < \Omega >=< v\cdot\nabla\Omega >;\ \frac{d}{dt}\mathfrak{F} = \frac{d}{dt}[< v^2 > -2 < \Omega >] = -2 < v\cdot\nabla\tilde{Q} >$$

Then since $\nabla\tilde{Q} = \nabla P/\rho$ where $P = D^2\rho\Delta log(\rho)$ (this is the real $\tilde{Q}$) the previous equation takes the form $\dot{\mathfrak{F}} = -\int \rho v\nabla\tilde{Q}dx = -\int v\nabla P dx$ which is an analogue of the familiar expression for the power release $(dE/dt) = F\cdot v$ with $F = -\nabla V$ in classical mechanics.

Next in [**396**] there is a discussion of differential entropy dynamics in quantum theory. Assume one has an arbitrary continuous function $\mathcal{V}(x,t)$ with dimensions of energy and consider the SE in the form $i\partial_t\psi = -D\Delta\psi + (\mathcal{V}/2mD)\psi$. Using $\psi = \rho^{1/2}exp(is)$ with $v = \nabla s$ one arrives at the standard equations $\partial_t\rho = -\nabla(v\rho)$ and $\partial_t s + (1/2)(\nabla s)^2 + (\Omega - \tilde{Q}) = 0$ where $\Omega = \mathcal{V}/m$ and $\tilde{Q}$ has the same form as in (1.12) (note a sign change of the $\Omega - \tilde{Q}$ term in comparison with (1.19)). These two equations generate a Markovian diffusion type process the probability density of which is propagated by a FP dynamics as before with drift $b = v - u$ (instead of $v = b - u$) where $u = D\nabla log(\rho)$ is an osmotic velocity field. Repeating the variational calculations one looks at (cf. (1.21))

$$\mathfrak{H} = \int \rho\left[\frac{1}{2}(\nabla s)^2 + \left(\frac{u^2}{2} + \Omega\right)\right] dx \tag{1.26}$$

Then

$$\mathfrak{H} = (1/2)[< v^2 > + < u^2 >] + < \Omega > = - < \partial_t s > \tag{1.27}$$

For time independent $\mathcal{V}$ one has $\mathfrak{H} = - < \partial_t s > = \mathcal{E} = const.$ and the FP equation propagates a probability density $|\psi|^2 = \rho$ whose differential entropy $\mathfrak{S}$ may nontrivially evolve in time. Maintaining the previous derivations involving $(\dot{\mathfrak{S}})_{in}$ one arrives at

$$(\dot{\mathfrak{S}})_{in} = \frac{2}{D}\left[\mathcal{E} - \left(\frac{1}{2}\mathfrak{F} + < \Omega >\right)\right] \geq 0 \tag{1.28}$$

One recalls $(1/2)\mathfrak{F} = - < \tilde{Q} > > 0$ so $\mathcal{E} - < \Omega > \geq (1/2)\mathfrak{F} > 0$. Hence the localization measure $\mathfrak{F}$ has a definite upper bound and the pertinent wave packet cannot be localized too sharply. Note also that $\mathfrak{F} = 2(\mathcal{E} - < \Omega >) - < v^2 >$ in general evolves in time (here $\mathcal{E}$ is a constant and $\dot{\Omega} = 0$). Using the hydrodynamical conservation laws one sees that the dynamics of Fisher information follows the rules

$$\frac{d\mathfrak{F}}{dt} = 2 < v\nabla\tilde{Q} >;\ \frac{1}{2}\dot{\mathfrak{F}} = -\frac{d}{dt}\left[\frac{1}{2} < v^2 > + < \omega >\right] \tag{1.29}$$

However $\dot{\mathfrak{F}} = \int v\nabla P dx$ where $P = D^2\rho\Delta log(\rho)$ and one interprets $\dot{\mathfrak{F}}$ as the measure of power transfer - keeping intact an overall mean energy $\mathfrak{H} = \mathcal{E}$. We refer to [**396**] for much more discussion and examples. We have concentrated on topics where the quantum potential appears in some form.

1.1. INFORMATION DYNAMICS. We go here to [**173, 174, 175**] and consider the idea of introducing some kind of dynamics in a reasoning process (Fisher information can apparently be linked to semantics - cf. [**907, 970**]). In [**173, 174**] one looks at the Fisher metric defined by

$$g_{\mu\nu} = \int_X d^4x p_\theta(x)\left(\frac{1}{p_\theta(x)}\frac{\partial p_\theta(x)}{\partial\theta^\mu}\right)\left(\frac{1}{p_\theta(x)}\right)\left(\frac{\partial p_\theta(x)}{\partial\theta^\nu}\right) \tag{1.30}$$

and constructs a Riemannian geometry via

$$\Gamma^\sigma_{\lambda\nu} = \frac{1}{2}g^{\nu\sigma}\left(\frac{\partial g_{\mu\nu}}{\partial\theta^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial\theta^\mu} - \frac{\partial g_{\mu\lambda}}{\partial\theta^\nu}\right); \tag{1.31}$$

$$R^\lambda_{\mu\nu\kappa} = \frac{\partial\Gamma^\lambda_{\mu\nu}}{\partial\theta^\kappa} - \frac{\partial\Gamma^\lambda_{\mu\kappa}}{\partial\theta^\nu} + \Gamma^\eta_{\mu\nu}\Gamma^\lambda_{\kappa\eta} - \Gamma^\eta_{\mu\kappa}\Gamma^\lambda_{\nu\eta}$$

Then the Ricci tensor is $R_{\mu\kappa} = R^\lambda_{\mu\lambda\kappa}$ and the curvature scalar is $R = g^{\mu\kappa}R_{\mu\kappa}$. The dynamics associated with this metric can then be described via functionals

$$J[g_{\mu\nu}] = -\frac{1}{16\pi}\int\sqrt{g(\theta)}R(\theta)d^4\theta \tag{1.32}$$

leading upon variation in $g_{\mu\nu}$ to equations

$$R^{\mu\nu}(\theta) - \frac{1}{2}g^{\mu\nu}(\theta)R(\theta) = 0 \tag{1.33}$$

Contracting with $g_{\mu\nu}$ gives then the Einstein equations $R^{\mu\nu}(\theta) = 0$ (since $R = 0$). J is also invariant under $\theta \to \theta + \epsilon(\theta)$ and variation here plus contraction leads to a contracted Bianchi identity. Constraints can be built in by adding terms $(1/2)\int\sqrt{g}T^{\mu\nu}g_{\mu\nu}d^4\theta$ to $J[g_{\mu\nu}]$. If one is fixed on a given probability distribution

$p(x)$ with variable θ^μ attached to give $p_\theta(x)$ then this could conceivably describe some gravitational metric based on quantum fluctuations for example. As examples a Euclidean metric is produced in 3-space via Gaussian $p(x)$ and complex Gaussians will give a Lorentz metric in 4-space. However it seems to be very restrictive to have a fixed $p(x)$ as the basis; it would be nice if one could vary the probability distribution in some more general manner and study the corresponding Fisher metrics (and this seems eminently doable with a Fisher metric over a space of probability distributions).

1.2. INFORMATION MEASURES FOR QM. We follow here [**749**] and derive the SE within an information theoretic framework somewhat different from the exact uncertainty principle of Hall and Reginatto (cf. Sections 1.1, 3.1, and 4.7). Begin with a SE for N particles in $d+1$ dimensions of the form $i\hbar\psi_t = [-(\hbar^2/2m)g_{ij}\partial_i\partial_j + V]\psi$ with $g_{ij} = \delta_{ij}/m_{[i]}$ where $i,j = 1,\cdots,dN$ and $[i]$ is the smallest integer $\geq i/d$. Use the Madelung transformation $\psi = \sqrt{\rho}exp(iS/\hbar)$ (cf. [**614**]) to get

(1.34)
$$\partial_t S + \frac{g_{ij}}{2}\partial_i S\partial_j S + V - \frac{\hbar^2}{8}g_{ij}\left(\frac{2\partial_i\partial_j\rho}{\rho} - \frac{\partial_i\rho\partial_j\rho}{\rho^2}\right) = 0;\ \partial_t\rho + g_{ij}\partial_i(\rho\partial_j S) = 0$$

These equations can be obtained from a variational principle, minimizing the action

(1.35)
$$\Phi = \int \rho\left[\partial_t S + \frac{g_{ij}}{2}\partial_i S\partial_j S + V\right]dx^{Nd}dt + \frac{\hbar^2}{8}I_F;$$
$$I_F = \int dx^{Nd}dt g_{ij}\rho(\partial_i log(\rho))(\partial_j log(\rho))$$

Here I_F resembles the Fisher information of [**369**] whose inverse sets a lower bound on the variance of the probabiliy distribution ρ via the Cramer-Rao inequaliity (see Section 1.1). (1.35) was used to derive the SE through a procedure analogous to the principle of maximum entropy in [**807, 806**] (cf. also Section 1.1). However the method of [**806**] does not explain a priori the form of information measure that should be used; i.e. why must the Fisher information be minimized rather than something else. The aim of [**749**] is to construct permissible information measures I. Thus the relevant action is

(1.36)
$$\mathfrak{A} = \int \rho\left[\partial_t S + \frac{g_{ij}}{2}\partial_i S\partial_j kS + V\right]dx^{Nd}dt + \lambda I$$

with λ a Lagrange multiplier. Varying this action will lead in general to a nonlinear SE

(1.37)
$$i\hbar\partial_t\psi = \left[-\frac{\hbar^2}{2}g_{ij}\partial_i\partial_j + V\right]\psi + F(\psi,\psi^\dagger)\psi$$

In order to have deformations of the linear theory that permit maximal preservation of the usual interpretation of the wave function one considers the following conditions:

(1) I should be real valued and positive definite for all $\rho = \psi^\dagger\psi$ and should be independent of V.

(2) I should be of the form $I = \int dx^{Nd} dt \rho H(\rho)$ where H is a function of $\rho(x,t)$ and its spatial derivatives. This will insure the weak superposition principle in the equations of motion.
(3) H should be invariant under scaling, i.e. $H(\lambda\rho) = H(\rho)$ which allows solutions of (1.37) to be renormalized, etc.
(4) H should be separable for the case of two independent subsystems for which the wave function factorizes, i.e. $H(\rho_1\rho_2) = H(\rho_1) + H(\rho_2)$.
(5) H should be Galilean invariant.
(6) The action should not contain derivatives beyond second order (Absence of higher order derivatives or AHD condition). This will insure that the multiplier λ, and hence Planck's constant, will be the only new parameter that is required in making the transition from classical to quantum mechanics.

The conditions 2-6 are already satisfied by the classical part of the action so it is quite minimalist to require them also of I. The homogeneity requirement 3 cannot be satisfied if H depends only on ρ; it must contain derivatives and the AHD and rotational invariance conditions imply then that $H = g_{ij}(U_1\partial_i U_2 \partial_j U_3 + V_1 \partial_i \partial_j V_2)$ where the U_i, V_i are functions of ρ. One can write then

$$H = g_{ij}\left(\frac{\partial_i\rho\partial_j\rho}{\rho^2}[U_1U_2'U_3'\rho^2 + V_1V_2''\rho^2] + \frac{\partial_i\partial_j\rho}{\rho}[V_1V_2'\rho]\right) \tag{1.38}$$

where the prime denotes a derivative with respect to ρ. Scaling conditions plus positivity and universality then lead to

$$I = \int dt dx^{Nd} \rho g_{ij} \frac{\partial_i\rho\partial_j\rho}{\rho^2} \tag{1.39}$$

Consequently the unique solution of the conditions 1-6 is the Fisher information measure and one arrives at the linear SE since the Lagrange multiplier must then have the dimension of $action^2$ thereby introducing the Planck constant. Note condition 4 was not used but it will be useful below. Further one notes that the AHD condition ensures that within the information theoretic approach the SE is the unique single parameter extension of the classical HJ equations. One also argues that a different choice of metric in the information term would in fact lead back to the original g_{ij} after a nonlinear gauge transformation; this suggests that a nonlinear SE is not automatically pathological. Further argument also shows that I should not depend on S. The main difference between this and the Hall-Reginatto method is to replace the exact uncertainly principle by condition 3.

1.3. PHASE TRANSITIONS. Referring here to [**512**] the introduction of a metric onto the space of parameters in models in statistical mechanics gives an alternative perspective on their phase structure. In fact the scalar curvature $\mathcal{R}$ plays a central role where for a flat geometry $\mathcal{R} = 0$ (noninteracting system) while $\mathcal{R}$ diverges at the critical point of an interacting one. Thus models are characterized by certain sets of parameters and given a probability distribution $p(x|\theta)$ and a sample x_i the object is to estimate the parameter θ. This can be done by maximizing the so-called likelihood function $L(\theta) = \prod_1^n p(x_i|\theta)$ or its logarithm. Thus one writes

(1.40)

$$log(L(\theta)) = \sum_1^n log(p(x_i|\theta));\ U(\theta) = \frac{d\,log(L(\theta))}{d\theta};\ Var[U(\theta)] = -\left[\frac{-d^2log(L(\theta))}{d\theta^2}\right]$$

The last term $Var[U(\theta))]$ is called the expected or Fisher information and we note that it is the same as (1.30) (see below) and in multidimensional form is expressed via

$$G_{ij}(\theta) = -E\left[\frac{\partial^2 log(p(x|\theta))}{\partial\theta_i\partial\theta_j}\right] = -\int p(x|\theta)\frac{\partial^2 log(p(x|\theta))}{\partial\theta_i\partial\theta_j}dx \tag{1.41}$$

In generic statistical-physics models one often has two parameters β (inverse temperature) and h (external field); in this case the Fisher-Rao metric is given by $G_{ij} = \partial_i\partial_j f$ where f is the reduced free energy per site and this leads to a scalar curvature

$$\mathcal{R} = -\frac{1}{2G^2}\begin{vmatrix} \partial_\beta^2 f & \partial_\beta\partial_h f & \partial_h^2 f \\ \partial_\beta^3 f & \partial_\beta^2\partial_h f & \partial_\beta\partial_h^2 f \\ \partial_\beta^2\partial_h f & \partial_\beta\partial_h^2 f & \partial_h^3 f \end{vmatrix} \tag{1.42}$$

where $G = det(G_{ij})$. In some sense $\mathcal{R}$ measures the complexity of the system since for $\mathcal{R} = 0$ the system is not interacting and (in all known systems) the curvature diverges at, and only at, a phase transition point. As an example under standard scaling assumptions one can anticipate the behavior of $\mathcal{R}$ near a second order critical point. Set $t = 1 - (\beta/\beta_c)$ and consider

$$f(\beta,h) = \lambda^{-1}f(t\lambda^{a_t}, h\lambda^{a_h}) = t^{1/a_t}\psi(ht^{-a_h/a_t});\ a_t = \frac{1}{\nu d};\ a_h = \frac{\beta\delta}{\nu d} \tag{1.43}$$

$a_t,\ a_h$ are the scaling dimensions for the energy and spin operators and d is the space dimension. For the scalar curvature there results

$$\mathcal{R} = -\frac{1}{2G^2}\begin{vmatrix} t^{(1/a_t)-2} & 0 & t^{(1/a_t)-2(a_h/a_t)} \\ t^{(1/a_t)-3} & 0 & t^{(1/a_t)-2(a_h/a_t)-1} \\ 0 & t^{(1/a_t)-2(a_h/a_t)-1} & t^{(1/a_t)-3(a_h/a_t)} \end{vmatrix}; \tag{1.44}$$

$$G \sim t^{(2/a_t)+2(a_h/a_t)-2} \Rightarrow \mathcal{R} \sim \xi^d \sim |\beta - \beta_c|^{\alpha-2}$$

where hyperscaling ($\nu d = 2 - \alpha$) is assumed and ξ is the correlation length. We refer to [**512**] for more details, examples, and references.

1.4. FISHER INFORMATION AND HAMILTON'S EQUATIONS. Going to [**755**] one shows that the mathematical form of the Fisher information I for a Gibb's canonical probability distribution incorporates important features of the intrinsic structure of classical mechanics and has a universal form in terms of forces and accelerations (i.e. one that is valid for all Hamiltonians of the form $T + V$). First one has shown that the Fisher information measure provides a powerful variational principle, that of extreme information, which yields most of the canonical Lagrangians of theoretical physics. In addition I provides an interesting characterization of the "arrow of time", alternative to the one associated with the Boltzman entropy (cf. [**776, 777**]). Following [**381, 384**] one considers a (θ, z) "scenario" in which we deal with a system specified by a physical parameter θ while z is a stochastic variable ($z \in \mathbf{R}^M$) and $f_\theta(z)$ is a probability density for z. One makes a measurement of z and has to infer θ, calling the resulting estimate $\tilde{\theta} = \tilde{\theta}(z)$. Estimation theory states that the best possible estimator $\tilde{\theta}(z)$, after a large number of samples, suffers a mean-square error e^2 from θ that obeys a relationship involving Fisher's I, namely $Ie^2 = 1$, where $I(\theta) = \int dz f_\theta(z)[\partial log(f_\theta(z))/\partial\theta]^2$ (only unbiased estimators with $<\tilde{\theta}>=\theta$ are in competition). The result here is that $Ie^2 \geq 1$ (Cramer-Rao bound). A case of great importance here concerns shift invariant distribution functions where the form does not change under θ displacements and one can write

$$I = \int dz f(z)\left(\frac{\partial log(f(z))}{\partial z}\right)^2 \tag{1.45}$$

If one is dealing with phase space where z is a M=2N dimensional vector with coordinates r and p then $I(z) = I(r) + I(p)$ (cf. [**755**]). Now assume that one wishes to describe a classical system of N identical particles of mass m with Hamiltonian

$$\mathfrak{H} = \mathfrak{T} + \mathfrak{V} = \sum_1^N \frac{p_i^2}{2m} + \sum_1^N V(r_i) \tag{1.46}$$

This is a simple situation but the analysis is not limited to such systems. Assume also that the system is in equilibrium at temperature T so that in the canonical ensemble the probability density is

$$\rho(r,p) = \frac{e^{-\beta\mathfrak{H}(r,p)}}{Z};\ Z = \int \frac{d^{3N}rd^{3N}p}{N!h^{3N}}e^{-\beta\mathfrak{H}(r,p)} \tag{1.47}$$

(here for h an elementary cell in phase space one writes $d\tau = d^{3N}rd^{3N}p/(N!h^{3N}$, $\beta = 1/kT$ with k the Boltzman constant, and Z is the partition function). Then from Hamilton's equations $\partial_p\mathfrak{H} = \dot{r}$ and $\partial_r\mathfrak{H} = -\dot{p}$ there results

$$-kT\frac{\partial log(\rho(r,p))}{\partial p} = \dot{r};\ -kT\frac{\partial log(\rho(r,p))}{\partial r} = -\dot{p} \tag{1.48}$$

One can now write the Fisher information measure in the form

$$I_\tau = \int \frac{d^{3N}rd^{3N}p}{N!h^{3N}}\rho(r,p)\mathfrak{A}(r,p);\ \mathfrak{A} = a\left(\frac{\partial log(\rho(r,p))}{\partial p}\right)^2 + b\left(\frac{\partial log(\rho(r,p))}{\partial r}\right)^2 \tag{1.49}$$

One needs two coefficients for dimensional balance (cf. [**755**]). One notes that

$$\frac{\partial log(\rho(r,p))}{\partial p} = -\beta\frac{\partial\mathfrak{H}}{\partial p};\ \frac{\partial log(\rho(r,p))}{\partial r} = -\beta\frac{\partial\mathfrak{H}}{\partial r} \tag{1.50}$$

leading to the Fisher information in the form

$$(kT)^2 I_\tau = a\left\langle\left(\frac{\partial\mathfrak{H}}{\partial p}\right)^2\right\rangle + b\left\langle\left(\frac{\partial\mathfrak{H}}{\partial r}\right)^2\right\rangle \Rightarrow I_\tau = \beta^2[a<\dot{r}^2> + b<\dot{p}^2>] \tag{1.51}$$

This gives the universal Fisher form for any Hamiltonian of the form (1.46) and we refer to [**607**] for connections to kinetic theory. Many other interesting results on Fisher can be found in [**381, 382, 755**].

1.5. UNCERTAINTY AND FLUCTUATIONS. We go first to [**38**] and recall the idea of a phase space distribution in the form (♣) $\mu(p,q) =< z|\rho|z >$ where ρ is the density matrix and $|z>$ denotes coherent states (cf. [**191, 757**] for coherent states). The chosen measure of uncertainty here is the Shannon information

$$I = -\int \frac{dpdq}{2\pi\hbar}\mu(p,q)log(\mu(p,q)) \tag{1.52}$$

The uncertainty principle manifests itself via the inequality (♠) $I \geq 1$ with equality if and only if ρ is a coherent state (cf. [**609, 987**]). In [**38**] one wants to generalize this to include the effects of thermal fluctuations in nonequilibrium systems and we sketch some of the ideas at least for equilibrium systems. There are in general three contributions to the uncertainty:

(1) The quantum mechanical uncertainty (quantum fluctuations) which is not dependent on the dynamics.
(2) The uncertainty due to spreading or reassembly of the wave packet. This is a dynamical effect and it may increase or decrease the uncertainty.
(3) The uncertainty due to the coupling to a thermal environment (diffusion and dissipation).

The time evolution I_t of I is studied for nonequilibrium systems and it is shown to generally settle down to monotone increase. I_t^{min} is a measure of the amount of quantum and thermal noise the system must suffer after a nonunitary evolution for time t (we do not deal with this here but refer to [**38**] for the nonequilibrium situation where the system decomposes into a distinguished system $\mathcal{S}$ plus the rest, referred to as the environment; the resulting time evolution of ρ is then nonunitary). In any event the lower bound I_t^{min} includes the effects of 1 and 3 but avoids 2.

One recalls the Shannon information (discussed earlier)

$$I(S) = -\sum_1^N p_i log(p_i);\ 0 \leq I(S) \leq log(N) \tag{1.53}$$

This is often referred to as entropy but here the word entropy is reserved for the vonNeumann entropy. In a similar manner, for continuous distributions (X a random variable with probability density $p(x)$ and $\int p(x)dx = 1$), the information of X is defined as

$$I(X) = -\int dxp(x)log(p(x)) \tag{1.54}$$

One emphasize that $p(x)$ here is a density (so it may be greater that 1 and $I(X)$ may be negative). However it retains its utility as a measure of uncertainty and e.g. for a Gaussian

$$p(x) = \frac{1}{[2\pi(\Delta x)^2]^{1/2}}exp\left(-\frac{(x-x_0)^2}{2(\Delta x)^2}\right);\ I(X) = log\left(2\pi e(\Delta x)^2\right)^{1/2} \tag{1.55}$$

Thus $I(X)$ is unbounded from below and goes to $-\infty$ as $\Delta x \to 0$ and $p(x)$ goes to a delta function. $I(X)$ is also unbounded from above but if the variance is fixed then $I(X)$ is maximized by the Gaussian distribution (1.55). Hence one has

$$I(X) \leq log\left(2\pi e(\Delta x)^2\right)^{1/2} \tag{1.56}$$

The generalization to more than one variable is straightforward, e.g.

$$I(X,Y) = -\int dxdyp(x,y)log(p(x,y)) \Rightarrow I(X,Y) \leq I(X) + I(Y) \tag{1.57}$$

where e.g. $I(X) = \int dyp(x,y)$. It is useful to introduce QM phase space distributions of the form

(1.58)

$$\mu(p,q) =< z|\rho|z >;\ < x|z >=< x|p,q >= \left(\frac{1}{2\pi\sigma_q^2}\right)^{1/4} exp\left(-\frac{(x-q)^2}{4\sigma_q^2} + ipx\right)$$

Here $< x|z >$ is a coherent state with $\sigma_q\sigma_p = (1/2)\hbar$ and there is a normalization $\int (dpdq/2\pi\hbar)\mu(p,q) = 1$. One can also show that

$$\mu(p,q) = 2\int dp'dq'exp\left(-\frac{(p-p')^2}{2\sigma_p^2} - \frac{(q-q')^2}{2\sigma_q^2}\right)W_\rho(p',q'); \tag{1.59}$$

$$W_\rho(p,q) = \frac{1}{2\pi\hbar}\int d\xi e^{-(i/\hbar)p\xi}\rho(q+(1/2)\xi, q-(1/2)\xi)$$

(Wigner function - cf. [**191, 192**]). One is interested in the extent to which $\mu(p,q)$ is peaked about some region in phase space and the Shannon information is a natural measure of the extent to which a probability distribution is peaked. Thus one takes as a measure of uncertainty the information

$$I(P,Q) = -\int \frac{dpdq}{2\pi\hbar}\mu(p,q)log(\mu(p,q)) \tag{1.60}$$

One expects there to be a lower bound for I and it should be achieved on a coherent state and this was in fact proved (cf. [**609, 987**]) in the form $I(P,Q) \geq 1$ with equality if and only if ρ is the density matrix of a coherent state $|z' >< z'|$. Further

$$log\left(\frac{e}{\hbar}\Delta_\mu q\Delta_\mu p\right) \geq I(Q) + I(P) \geq I(P,Q) \tag{1.61}$$

The variances here have the form

$$(\Delta_\mu q)^2 = (\Delta_\rho q)^2 + \sigma_q^2;\ (\Delta_\mu p)^2 = (\Delta_\rho p)^2 + \sigma_p^2 \tag{1.62}$$

where Δ_ρ denotes the QM variance and hence

$$\left((\Delta_\rho q)^2 + \sigma_q^2\right)\left((\Delta_\rho p)^2 + \sigma_p^2\right) \geq \hbar^2 \tag{1.63}$$

Minimizing (1.63) over σ_q (and recalling that $\sigma_q\sigma_p = (1/2)\hbar$) one obtains the standard uncertainty relation $\Delta x\Delta p \geq (\hbar/2)$. Now suppose one has a genuinely mixed state so that

$$\rho = \sum_n p_n|n >< n|;\ p_n < 1;\ \mu(p,q) = \sum p_n| < z|n > |^2 \tag{1.64}$$

The information of (1.64) will always satisfy $I(P,Q) \geq 1$ but this is a very low lower bound; indeed from the inequality

$$(1.65) \qquad -\left(\int dx f(x)g(x)\right) log\left(\int dy f(y)g(y)\right) \geq -\int dx g(x)g(x)log(x)$$

we have
(1.66)

$$I \geq -\int \frac{dpdq}{2\pi\hbar}\sum_n |<z|n>|^2 p_n log(p_n) = -\sum p_n log(p_n) = -Tr(\rho log(\rho)) \equiv S[\rho]$$

Thus I is bounded from below by the vonNeumann entropy $S[\rho]$ and this is a virtue of the chosen measure of uncertainty. One sees that I is a useful measure of both quantum and thermal fluctuations. It has a lower bound expressing the effect of quantum fluctuations which is connected to entropy and this in turn is a measure of thermal fluctuations.

Consider now the situation of thermal equilibrium. Let the density matrix be thermal, $\rho = Z^{-1}exp(-\beta H)$ where $Z = Tr(e^{-\beta H})$ is the partition function and $\beta = 1/kT$. Then

$$(1.67) \qquad <z|\rho|z> = \frac{1}{Z}\sum e^{-\beta E_n}|<z|n>|^2$$

where $|n>$ are energy eigenstates with eigenvalue E_n. For simplicity look at a harmonic oscillator for which

$$(1.68) \quad H = \frac{1}{2}\left(\frac{p^2}{M} + M\omega^2 q^2\right);\ |<z|n>|^2 = \frac{|z|^{2n}}{n!}e^{-|z|^2};\ E_n = \hbar\omega(n + (1/2))$$

Here $z = (1/2)[(q/\sigma_q) + i(p/\sigma_p)]$ where $\sigma_q\sigma_q = (1/2)\hbar$ and $\sigma_q = (\hbar/2M\omega)^{1/2}$ (cf. [**191, 559, 757**] for coherent states). There results

$$(1.69) \qquad \mu(q,p) = <z|\rho|z> = \left(1 - e^{-\beta\hbar\omega}\right) exp\left(-(1 - e^{-\beta\hbar\omega})|z|^2\right)$$

The information (1.60) is then (•) $I = 1 - log(1 - e^{-\beta\hbar\omega})$ which is exactly what one expects; as $T \to 0$ one has $\beta \to \infty$ and the uncertainty reduces to the Lieb-Wehrl result $I(P,Q) \geq 1$ expressing purely quantum fluctuations. For nonzero temperature however the uncertainty is larger tending to the value $-log(\beta\hbar\omega)$ as $T \to \infty$ which expresses purely thermal fluctuations. It is interesting to compare (•) with the entropy $S = -Tr(\rho log(\rho))$. Here the partition function is $Z = [2Sinh((1/2)\beta\hbar\omega)]^{-1}$ and the entropy is then $S = -\beta(\partial_\beta(log(Z)) + log(Z)$ or

$$(1.70) \qquad S = -log[2Sinh((1/2)\beta\hbar\omega)] + (1/2)\beta\hbar\omega Coth[(1/2)\beta\hbar\omega]$$

For large T one has then $S \simeq -log(\beta\hbar\omega)$ coinciding with I but $S \to 0$ as $T \to 0$ while I goes to a nontrivial lower bound. Hence one sees that I is a useful measure of uncertainty in both the quantum and thermal regimes. We refer also to [**4**] where an information theoretic uncertainty relation including the effects of thermal fluctuations at thermal equilibrium has been derived using thermofield dynamics (cf. [**950**]); their information theoretic measure is however different than that in [**38**]. On goes next to non-equilibrium systems and proves for linear systems that, for each t, I has a lower bound I_t^{min} over all possible initial states. It coincides with the Lieb-Wehrl bound in the absence of an environment and is related to the vonNeumann entropy in the long time limit. We refer to [**38**] for details.

2. A TOUCH OF CHAOS

For quantum chaos we refer to [**33, 96, 97, 185, 218, 282, 359, 435, 484, 491, 517, 518, 584, 659, 747, 787, 936**] and begin here with [**747**]. Chaos is quantitatively measured by the Lyapunov spectrum of characteristic exponents which represent the principal rates of orbit divergence in phase space, or alternatively by the Kolmogorov-Sinai (KS) invariant, which quantifies the rate of information production by the dynamical system. Chaos is conspicuously absent in finite quantum systems but the chaotic nature of a given classical Hamiltonian produces certain characteristic features in the dynamical behavior of its quantized version; these features are referred to as quantum chaos (cf. [**218, 435**]). They include short term instabilities and diffusive behavior versus dynamical localization and other effects. One is concerned here with an approach to the information dynamics of the quantum-classical transition based on the HJ formalism with the KS invariant playing a central role. The extension to the quantum domain is accomplished via the orbits introduced by Madelung and Bohm (cf. [**129, 614**]); these are natural extensions of the classical phase space flow to QM and provide the required bridge across the transition. One striking result is that the quantum KS invariant for a given Madelung-Bohm (MB) orbit is equal to the mean decay rate of the probability density along the orbit. Further one shows that the quantum KS invariant averaged over the ensemble of MB orbits equals the mean growth rate of configuration space information and a general and rigorous argument is given for the conjecture that the standard quantum-classical correspondence (or the classical limit) breaks down for classically chaotic Hamiltonians.

We give only a sketch of results here. Thus consider a classical system of N degrees of freedom described by canonical variables (q_i, p_i) with $1 \leq i \leq N$ and denote the Hamiltonian as $H(\mathbf{q}, \mathbf{p}, t)$ with Hamilton principal function $S(\mathbf{q}, t, \mathbf{p}_0)$ where $\mathbf{p}_0$ being the initial momenta. In matrix form Hamilton's equations are $\dot{\xi} = \mathcal{J}\nabla_\xi H(\xi, t)$ where ξ stands for the 2N dimensional phase space vector $(\mathbf{q}, \mathbf{p})$. Here $\mathcal{J}$ is a real antisymmetric matrix of order 2N with a $2 \otimes N$ block form $(0_N, I_N, -I_N, 0_N)$ which is a listing of blocks in the order $(11, 12, 21, 22)$. The tangent dynamics of the system is described by the $2N \times 2N$ nonsingular matrix $\mathcal{T}_{\mu\nu}(t, \xi_0) = \partial\xi_\mu(t, \xi_0)/\partial\xi_{0\nu}$ (the sensitivity matrix) where $\xi(t, \xi_0)$ is the trajectory starting from ξ_0 at time t_0. One can in fact write ($\tilde{S} \equiv S^T$ - matrix transpose)

$$\mathcal{T} = (S^{-1}_{\mathbf{p}_0\mathbf{q}}, -S_{\mathbf{p}_0\mathbf{q}}S_{\mathbf{p}_0\mathbf{p}_0}, S_{\mathbf{q}\mathbf{q}}S^{-1}_{\mathbf{p}_0\mathbf{q}}, \tilde{S}_{\mathbf{p}_0\mathbf{q}} - S_{\mathbf{q}\mathbf{q}}S^{-1}_{\mathbf{p}_0\mathbf{q}}S_{\mathbf{p}_0\mathbf{p}_0}) \tag{2.1}$$

where $(S_{\mathbf{p}_0\mathbf{q}})_{ij} = \partial^2 S/\partial_{q_j}\partial_{p_0 i}$. It is shown that one can write $\mathcal{T}$ in an upper triangular block form $\Gamma = \Omega(\Theta)\mathcal{T}$ where $\Omega(\Theta) = (cos(\Theta), -Sin(\Theta), Sin(\Theta), Cos(\Theta))$ and $-\sigma = Tan(\Theta) = -S_{\mathbf{q}\mathbf{q}}$. Here Θ is a real symmetric matrix of order N while Ω is orthogonal and symplectic (symplectic phase matrix). The upper triangular form $(\Gamma_{11}, \Gamma_{12}, 0_N, \Gamma_{22})$ of Γ satisfies $\Gamma_{11}^{-1} = \tilde{\Gamma}_{22}$ and the upper half of the Lyapunov spectrum is obtained from the singular values of Γ_{11} (see [**747**]). In particular the

Kolmogorov-Sinai (KS) entropy is given via

$$k = lim_{t\to\infty} log[det(\Gamma_{11})]/t \tag{2.2}$$

For illustration consider the standard form $H = \mathbf{p}^2/2+V(\mathbf{q},t)$ with N-dimensional vectors $\mathbf{q}$, $\mathbf{p}$. Then

$$k =< Tr(\sigma) >_{p.v.};\ < f >= lim_{t\to\infty}\frac{1}{t}\int_0^t f(t')dt' \tag{2.3}$$

where p.v. stipulates a principal value evaluation (σ will have simple pole behavior near singularities and the principal value contribution vanishes). Since $Tr(\sigma) = \nabla_q^2 S$ along the orbit (2.3) simply states that the KS invariant equals the time average of the Laplacian of the action along the orbit. Now the MB formalism associates a phase space flow with a quantum system via

$$\psi = exp[iS(\mathbf{x},t)/\hbar + R(\mathbf{x},t)];\ \dot{\mathbf{q}}(t,\mathbf{q}_0,\mathbf{p}_0) = \mathbf{p} = \nabla S[\mathbf{q},t] \tag{2.4}$$

It can be verified that the expectation value of any observable in the state ψ is given by its average over the ensemble of orbits thus defined (e.g. Ehrenfest's equations arise in this manner). The correspondence thus allows us to define the quantum KS invariant for a given orbit as

$$\mathbf{k} =< \nabla^2 S >_{p.v.} \tag{2.5}$$

(the averaging process is with respect to the time along the MB orbit to which S is restricted). Now intuitively one would expect that orbits neighboring a hypothetical chaotic orbit in the ensemble diverge from it on the average thus causing the orbit density along the chaotic orbit to decrease with a mean rate related to $\mathbf{k}$. This is fully realized here as one sees by considering the equation of motion for $R(\mathbf{x},t)$ as inherited from the SE, namely $\partial_t R+\nabla R\cdot\nabla S = -(1/2)\nabla^2 S$. The characteristic curves for this equation are the MB orbits so that it takes the following form along these orbits;

$$\frac{dR}{dt} = -\frac{1}{2}\nabla^2 S \Rightarrow \mathbf{k} = -2\left\langle\frac{dR}{dt}\right\rangle = -\left\langle\frac{d\,log(|\psi|^2)}{dt}\right\rangle_{p.v.} \tag{2.6}$$

This says that the quantum KS invariant for a given orbit is the mean decay rate of the probability density along the orbit. Comparing this to a classical system where $k \neq 0$ while $\mathbf{k} = 0$ for the quantum version one sees that the classical limit cannot hold for chaotic Hamiltonians and since chaotic classical Hamiltonians are certainly more common than regular ones the idea of classical limit is not a reliable test for quantum systems. Finally let $\bar{\mathbf{k}}$ be the MB ensemble average, which is the same as the QM expectation value, leading to

$$\bar{\mathbf{k}} = lim_{t\to\infty}\frac{1}{t}\int dq|\psi|^2 log(|\psi|^2) \tag{2.7}$$

which is an information entropy measure. The discussion here is very incomplete but should motivate further investigation and we refer to [**747**] for more detail (cf. also [**976, 977**] involving chaos, fractals, and entropy.

2.1. CHAOS AND THE QUANTUM POTENTIAL. The paper [**745**] offers an interesting perspective on the quantum potential. Thus consider a system of n particles with the SE

$$i\hbar\partial_t\psi = \left[\sum_1^n \left(\frac{-\hbar^2}{2m_i}\right)\nabla_i^2 + V\right]\psi;\ \nabla_i = \left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial y_i}, \frac{\partial}{\partial z_i}\right) \tag{2.8}$$

(here $\mathbf{x}_i = (x_i, y_i, z_i)$). Set $\psi = Rexp[i(S/\hbar)]$ and there results as usual

$$\partial_t S + \sum_1^n (\nabla_i S)^2 (2m_i)^{-1} + Q + V = 0;\ \partial_t R^2 + \sum_1^n \nabla_i \cdot \left(\frac{R^2\nabla_i S}{m_i}\right) = 0 \tag{2.9}$$

where $Q = -\sum_1^n (\hbar^2/2m_i R)\nabla_i^2 R$. Now just as the causal form of the HJ equation contains the additional term Q so the causal form of Newton's second law contains Q as follows

$$\dot{P}_i = -\nabla_i V - \nabla_i Q;\ P = \sum_1^n P_i;\ \dot{P} = \frac{dP}{dt} = -\sum_1^n \nabla_i V - \sum_1^n \nabla_i Q \tag{2.10}$$

The author cites a number of curious and conflicting statements in the literature concerning the effect of the quantum potential on Bohmian trajectories, for clarification of which he observes that for an isolated system one has

$$-\sum \nabla_i V = 0;\ \dot{P} = 0;\ -\sum \nabla_i Q = -\sum F_i = 0 \tag{2.11}$$

Thus the sum of all the quantum forces is zero so $F_i = \sum_{j\neq i}^n (-F_j)$. Thus the net quantum force on a given particle is the result of all the other particles exerting force on this particle via the intermediary of the quantum potential. This then is his explanation for the guidance role of the wave function.

Next it is noted that removing Q from the HJ equation is equivalent to adding the term

$$\left(\frac{\hbar^2}{2m}\right) exp(iS/\hbar)\nabla^2 R = \left(\frac{\hbar^2}{2m}\right)|\psi|^{-1}\psi\nabla^2|\psi| = -Q\psi \tag{2.12}$$

to the SE so that the effective Hamiltonian becomes

$$H_{eff} = -\left(\frac{\hbar^2}{2m}\right)\nabla^2 + V + \left(\frac{\hbar^2}{2m}\right)|\psi|^{-1}\nabla^2|\psi| \tag{2.13}$$

Since $H_{eff} = H_{eff}(\psi)$ depends on ψ the superposition principle no longer applies. When $\phi \neq \psi$ we have

$$\int(\phi^* H_{eff}\psi - \psi H_{eff}\phi^*)d\tau = \left(\frac{\hbar^2}{2m}\right)\int \phi^*\psi[|\psi|^{-1}\nabla^2|\psi| - |\phi|^{-1}\nabla^2|\phi|] \neq 0 \tag{2.14}$$

so H_{eff} is not Hermitian. Hence the time development operator $exp[(i/\hbar)H_{eff}t]$ is not unitary and the time dependent SE is a nonunitary flow. Then since $i\hbar\partial_t(\psi^*\psi) = \psi^* H_{eff}\psi - \psi H_{eff}\psi^*$ one has

$$\partial_t \int |\psi - \phi|^2 d\tau = \left(\frac{\hbar^2}{2m}\right)\int i[\psi^*\phi - \phi^*\psi][|\psi|^{-1}\nabla^2|\psi| - |\phi|^{-1}|\phi|]d\tau \neq 0 \tag{2.15}$$

Consider then the case where two initial conditions for the time dependent SE differ only infinitesimally. As time progresses the two corresponding wave functions can become quite different, indicating the possibity of deterministic chaos, and this is a consequence of H_{eff} being a functional of the state upon which it is acting. If the term $(\hbar^2/2m)|\psi|^{-1}\nabla^2|\psi|$ is removed from (2.13) one is left with a Hermitian Hamiltonian and the normalization of $(\psi - \phi)$ is time independent, so there can be no deterministic chaos. Thus in particular Q acts as a constraining force preventing deterministic chaos (cf. also [**623**]).

REMARK 6.2.1. There are many different aspects of quantum chaos and the perspective of [**?**] just mentioned does not deal with everything covered in the references already cited (cf. also [**633, 983, 1000, 1001**] for additional referencers). We are not expert enough to attempt any kind of in depth coverage but extract here briefly from a few papers. First from [**1000**] one notes that the dBB theory of quantum motion provides motion in deterministic orbits under the influence of the quantum potential. This quantum potential can be very intricate because it generates wave interferences and further numerical work has shown the presence of chaos and complex behavior of quantum trajectories in various systems (cf. [**746**]). In [**1000**] one indicates that movement of the zeros of the wave function (called vortices) implies chaos in the dynamics of quantum trajectories. These vortices result from wave function interferences and have no classical explanation. In systems without magnetic fields the bulk vorticity $\nabla \times \mathbf{v}$ in the probability fluid is determined by points where the phase S is singular (which can occur when the wave function vanishes). Due to singlevaluedness of the wave function the circulation $\Gamma = \int_C \dot{\mathbf{r}} d\mathbf{r} = (2\pi n/m)$ around a closed contour C encircling a vortex is quantized with n an integer and the velocity must diverge as one approaches a vortex. This leads to a universal mechanism producing chaotic behavior of quantum trajectories (cf. also [**746, 1001**]).

Next in [**983**] one speaks of the edge of quantum chaos (the border between chaotic and non-chaotic regions) where the Lyapunov exponent goes to zero; it is then replaced by a generalized Lyapunov coefficient describing power-law rather than exponential divergence of classical trajectories. In [**983**] one characterizes quantum chaos by comparing the evolution of an initially chosen state under the chaotic dynamics with the same state evolved under a perturbed dynamics (cf. [**761**]). When the initial state is in a regular region of a mixed system (one with regular and chaotic regions) the overlap remains close to one; however when the initial state is in a chaotic zone the overlap decay is exponential. It is shown that at the edge of quantum chaos there is a region of polynomial overlap decay. Here the overlap is defined as $O(t) = | < \psi_u(t)|\psi_p(t) > |$ where $\psi_u$is the state evolved under the unperturbed system operator and ψ_p is the state evolved under the perturbed operator.

In various papers (e.g. [**185, 484, 485, 517, 518, 584**]) one characterizes quantum chaos via the quantum action. This is defined via

$$\tilde{S}[x] = \int dt \frac{\tilde{m}}{2}\dot{x}^2 - \tilde{V}(x) \tag{2.16}$$

for a given classical action

$$S[x] = \int dt \frac{m}{2}\dot{x}^2 - V(x) \tag{2.17}$$

so that the QM transition amplitude is

$$G(x_f, t_f; x_i, t_i) = \tilde{Z} exp\left[\frac{i}{\hbar}\,\tilde{\Sigma}\Big| + x_i, t_i{}^{x_f,t_f}\right]; \tag{2.18}$$

$$\tilde{\Sigma}\Big|_{x_i,t_i}^{x_f,t_f} = \tilde{S}[\tilde{x}_{cl}]\Big|_{x_i,t_i}^{x_f,t_f} = \int_{t_i}^{t_f} dt \frac{\tilde{m}}{2}\dot{\tilde{x}}_{cl}^2 - \tilde{V})\tilde{x}_{cl})\Big| + x_i{}^{x_f}$$

where $\tilde{x}_{cl}$ is the classical path corresponding to the action $\tilde{S}$. One requires here 2-point boundary conditions $\tilde{x}_{cl}(t = t_i) = x_i$ and $\tilde{x}_{cl}(t = t_f) = x_f$ and $\tilde{Z}$ stands for a dimensionful normalisation factor. The parameters of the quantum action (i.e. mass and potential) are independent of the boundary points but depend on the transition time $T = t_f - t_i$. A general existence proof is lacking but such quantum actions exist in many interesting cases. Then quantum chaos is defined as follows. Given a classical system with action S the corresponding quantum system displays quantum chaos if the corresponding quantum action $\tilde{S}$ in the asymptotic regime $T \to \infty$ generates a chaotic phase space.

3. GENERALIZED THERMOSTATISTICS

We refer to [**38, 262, 382, 384, 694, 696, 755, 778, 779**] for discussion of various entropies based on deformed exponential functions (generalizations of the Boltzman-Gibbs formalism for equilibrium statistical physics), the entropies of Beck-Cohen, Kaniadakis, Renyi, Tsallis, etc., maximum entropy ideas, escort density operators, and a host of other matters in generalized theormstatistics. We sketch here first a few ideas following the third paper in [**694**]. Thus a model of thermostatistics is described by a density of states $\rho(E)$ and a probability distribution $p(E)$ and for a system in thermal equilibrium at temperature T one has

$$p(E) = \frac{1}{Z(T)}e^{-E/T};\ Z(T) = \int dE\rho(E)e^{-E/T} \tag{3.1}$$

(Boltzman's constant is set equal to one here). Thermal averages are defined via $< f >= \int dE\rho(E)p(E)f(E)$ (this is a simplified treatment with T not made explicit - i.e. $p(E) \sim p(E,T)$). A microscopic model of thermostatistics is specified via an energy functional $H(\gamma)$ over phase space Γ which is the set of all possible microstates. Using $\rho(E)dE = d\gamma$ one can write

$$< f >= \int_\Gamma d\gamma p(\gamma)f(\gamma);\ p(\gamma) = \frac{e^{-H(\gamma)/T)}}{Z(T)};\ Z(T) = \int_\Gamma d\gamma e^{-H(\gamma)/T)} \tag{3.2}$$

In the quantum case the integration is replaced by a trace to obtain

$$< f >= \frac{1}{Z(T)}Tr\, exp(-H/T)f;\ Z(T) = Tr\, exp(-H/T) \tag{3.3}$$

In relevant examples of thermostatistics the density of states $\rho(E)$ increases as a power law $\rho(E) \sim E^{\alpha N}$ with N the number of particles and $\alpha > 0$. There is an energy - entropy balance where the increase of density of states $\rho(E)$ compensates for the exponential decrease of probability density $p(E)$ with a maximum of $\rho(E)p(E)$ reached at some macroscopic energy far above the ground state energy. One can write $\rho(E)p(E) = (1/Z)exp(log(\rho(E)) - E/T)$ with the argument of the exponential maximal when E satisfies

$$\frac{1}{\rho(E)}\rho'(E) = \frac{1}{T} \tag{3.4}$$

where $\rho'(E)$ is the derivative $d\rho/dE$. If $\rho(E) \sim E^{\alpha N}$ then $E \sim \alpha NT$ follows which is the equipartition theorem. The form of the theory here indicaters that the actual form of the probability distribution is not very essential; alternative expressions for $p(E)$ are acceptable provided they satisfy the equipartition theorem and reproduce thermodynamics. One begins here by generalizing the equipartition result (3.4) and postulates the existence of an increasing positive function $\phi(x)$ defined for $x \geq 0$ such that ($\bullet$) $(1/T) = -[p'(E)/\phi(p(E))]$ holds for all E and T. Then the equation for the maximum of $\rho(E)p(E)$ becomes

$$0 = \frac{d}{dE}[\rho(E)p(E)] = \rho'(E)p(E) - \frac{1}{T}\rho(E)\phi(p(E)) \equiv \frac{\rho'(E)}{\rho(E)} = \frac{1}{T}\frac{\phi(p(E))}{p(E)} \tag{3.5}$$

The Boltzman-Gibbs case is recovered when $\phi(x) = x$. Now ($\bullet$) fixes the form of the probability distribution $p(E)$; to see this introduce a function $log_\phi(x)$ via

$$log_\phi(x) = \int_1^x \frac{1}{\phi(y)}dy \tag{3.6}$$

The inverse is $exp_\phi(x)$ and from the identity $1 = exp'_\phi(log_\phi(x))log'_\phi(x)$ there results ($\blacklozenge$) $\phi(x) = exp'_\phi(log_\phi(x))$. Hence ($\bullet$) can be written as

$$p'(E) = -\frac{1}{T}exp'_\phi[log_\phi(p(E))] \Rightarrow p(E) = exp_\phi(G_\phi(T) - (E/T)) \tag{3.7}$$

The function $G_\phi(T)$ is the integration constant and it must be chosen so that $1 = \int dE\rho(E)p(E)$ is satisfied. The formula (3.7) resembles the Boltzman-Gibbs distribution but the normalization constant appears inside the function $exp_\phi(x)$; for $\phi(x) = x$ one has then $G_\phi(T) = -log(Z(T))$.

In general it is difficult to determine $G_\phi(T)$ but an expression for its temperature derivative can be obtained via escort probabilities (cf. [**146, 943**]). The general definition is

$$P(E) = \frac{1}{Z(T)}\phi(p(E));\ Z(T) = \int dE\rho(E)\phi(p(E)) \tag{3.8}$$

Then expectation values for $P(E)$ are denoted by

$$< f >_* = \int dE\rho(E)P(E)f(E) \tag{3.9}$$

Note $P(E) = p(E)$ in the Boltzman-Gibbs case $\phi(x) = x$. Now calculate using (♦) and (3.8) to get

$$(3.10)\qquad \frac{d}{dT}p(E) = exp'_{\phi}(G_{\phi}(T) - (E/T))\left(\frac{d}{dT}G_{\phi}(T) + \frac{E}{T^2}\right) =$$
$$= Z(T)P(E)\left(\frac{d}{dT}G_{\phi}(T) + \frac{E}{T^2}\right)$$

from which follows (recall $\int dE\rho(E)p(E) = 1$)

$$(3.11)\qquad 0 = \int dE\rho(E)\frac{d}{dT}p(E) = Z(T)\frac{d}{dT}G_{\phi}(T) + \frac{1}{T^2}Z(T) < E >_{*} \Rightarrow$$
$$\Rightarrow \frac{d}{dT}G_{\phi}(T) = -\frac{1}{T^2} < E >_{*}$$

Note also that combining (3.10) and (3.11) one obtains

$$(3.12)\qquad \frac{d}{dT}p(E) = \frac{1}{T^2}Z(T)P(E)(E- < E >_{*})$$

One wants now to show that generalized thermodynamics is compatible with thermodynamics begins by establishing thermal stability. Internal energy $U(T)$ is defined via $U(T) =< E >$ with $p(E)$ given by (3.7), so using (3.12) one obtains (3.13)

$$\frac{d}{dT}U(T) = \int dE\rho(E)E\frac{d}{dT}p(E) = \int dE\rho(E)\frac{E}{T^2}Z(T)P(E)(E- < E >_{*}) =$$
$$= \frac{1}{T^2}Z(T)(< E^2 >_{*} - < E >_{*}^{2}) \geq 0$$

Hence average energy is an increasing function of T but thermal stability requires more so define ϕ entropy (relative to $\rho(E)dE$ via

$$(3.14)\quad S_{\phi}(p) = \int dE\rho(E)[(1 - p(E))F_{\phi}(0) - F_{\phi}(p(E))];\ F_{\phi}(x) = \int_1^x dylog_{\phi}(y)$$

One postulates that thermodynamic entropy $S(T)$ equals the value of the above entropy $S_{\phi}(p)$ with p given by (3.7). Then

$$(3.15)\qquad \frac{d}{dT}S(T) = \int dE\rho(E)(-log_{\phi}(p(E)) - F_{\phi}(0))\frac{d}{dT}p(E) =$$
$$= \int dE\rho(E)\left(-G_{\phi}(T) + \frac{E}{T} - F_{\phi}(0)\right)\frac{d}{dT}p(E) = \frac{1}{T}\frac{d}{dT}U(T)$$

(recall that $p(E)$ is normalized to 1). This shows that temperature T satisfies the thermodynamic relation $(1/T) = dS/dU$ and since E is an increasing function of T one concludes that S is a concave function of U; this is called thermal stability. One can also introduce the Helmholz free energy $F(T)$ via the well known $F(T) = U(T) - TS(T)$ so from (3.15) it follows that

$$(3.16)\qquad \frac{d}{d\beta}\beta F(T) = U(T)\ \ (\beta = 1/T)$$

Going back to (3.11) which is similar to (3.16) with $F(T)$ replaced by $TG_{\phi}(T)$ and with $U(T) =< E >$ replaced by $< E >_{*}$ the comparison shows that $TG_{\phi}(T)$

is the free energy associated with the escort probability distribution $P(E)$ up to a constant independent of T.

The most obvious generalization now involves $\phi(x) = x^q$ with $q > 0$ and this essentially produces the Tsallis entropy where one has

$$(3.17)\quad log_q(x) = \int_1^x dyy^{-q} = \frac{1}{1-q}(x^{1-q} - 1);\ exp_q(x) = [1 + (1-q)x]_+^{1/(1-q)}$$

(cf. [**?**]). The probability distribution (3.17) becomes
(3.18)

$$p(E) = [1 + (1-q)(G_q(T) - (E/T))]_+^{1/(1-q)} = \frac{1}{z_q(T)}[1 - (1-q)\beta_q^*(T)E]_+^{1/(1-q)}$$

$$z_q(T) = (1 + (1-q)G_q(T))^{1/(1-q)};\ \beta_q^*(T) = z_q(T)^{1-q}/T$$

A nice feature of the Tsallis theory is that the correspondence between $p(E)$ and the escort $P(E)$ leads to a dual structure $q \leftrightarrow 1/q$; indeed

$$(3.19)\qquad P(E) = \frac{1}{Z_q(T)}p(E)^q \Rightarrow p(E) = \frac{1}{Z_{1/q}(T)}P(E)^{1/q}$$

Moreover there is also a $q - 2 \leftrightarrow q$ duality; given $log_\phi(x)$ a new deformed $log_\psi(x)$ is obtained via

$$(3.20)\quad log_\psi(x) = (x-1)F_\phi(0) - xF_\phi(1/x);\ \frac{1}{\psi(x)} = F_\phi(0) - F_\phi(1/x) + \frac{1}{x}log_\phi(1/x)$$

and for $\phi = x^q$ one has $\psi = (2-q)x^{2-q}$. One notes also that the definition (3.14) of entropy $S_\phi(p)$ can be written as

$$(3.21)\qquad S_\phi(p) = \int dE\rho(E)p(E)log_\psi(1/p(E))$$

and with $\psi(x) = x^q$ we get the Tsallis entropy(cf. also [**1025, 1027**])

$$(3.22)\qquad S_q(p) = \int dE\rho(E)\frac{1}{1-q}(p(E)^q - p(E))$$

3.1. NONEXTENSIVE STATISTICAL THERMODYNAMICS. We go here to [**944**] for an lovely introduction and extract liberally. The Boltzman-Gibbs entropy is given via

$$(3.23)\qquad S_{BG} = -k\sum_1^W p_i log(p_i);\ \sum_1^W p_i = 1$$

Here p_i is the probability for the system to be in the i^{th} microstate and k is the Boltzman constant k_B (taken now to be 1). If every microstate has the same probability $p_i = 1/W$ then $S_{BG} = klog(W)$. The entropy (3.23) can be shown to be nonnegative, concave, extensive, and stable (or experimentally robust). By extensive one means that if A and B are two independent systems (i.e. $p_{ij}^{A+B} = p_i^A p_j^B$) then

$$(3.24)\qquad S_{BG}(A+B) = S_{BG}(A) + S_{BG}(B)$$

One can still not derive this form of entropy (3.23) from first principles. There is also good reason to conclude that physical entropies different from (3.23) would be more appropriate for anomalous systems. In this spirit the Tsallis entropy was proposed in [**945**] and the property thereby generalized is extensivity. One discusses motivations etc. in [**841**] and in particular observes that the function

$$y = \frac{x^{1-q} - 1}{1-q} = log_q(x) \tag{3.25}$$

satisfies

$$log_q(x_A x_B) = log_q(x_A) + log_q(X_B) + (1-q)(log_q(X_A))(log_q(X_B)) \tag{3.26}$$

Now rewrite (3.23) in the form ($k = 1$)

$$S_{BG} = -\sum_1^W p_i log(p_i) = \sum_1^W p_i log(1/p_i) = \left\langle log\frac{1}{p_i}\right\rangle \tag{3.27}$$

The quantity $log(1/p_i)$ is called surprise or unexpectedness and one thinks of a q-surprise $log_q(1/p_i)$ in defining

$$S_q = \left\langle log_q\frac{1}{p_i}\right\rangle = \sum_1^W p_i log_q(1/p_i) = \frac{1 - \sum_1^W p_i^q}{q-1} \tag{3.28}$$

In the limit $q \to 1$ one gets $S_1 = S_{BG}$ and assuming equiprobability $p_i = 1/W$ one gets

$$S_q = \frac{W^{1-q} - 1}{1-q} = log_q(W) \tag{3.29}$$

Consequently S_q is a genuine generalization of the BG entropy and the pseudo-additivity of the q-logarithm implies (restoring momentarily k)

$$\frac{S_q(A+B)}{k} = \frac{S_q(A)}{k} + \frac{S_q(B)}{k} + (1-q)\frac{S_q(A)}{k}\frac{S_q(B)}{k} \tag{3.30}$$

if A and B are two independent systems (i.e. $p_{ij}^{A+B} = p_i^A p_j^B$). Thus $q = 1,\ q < 1$, and $q > 1$ respectively correspond to the extensive, superextensive, and subextensive cases and the q-generalization of statistical mechanics is referred to as nonextensive statistical mechanics. (3.30) is true for independent A and B but if A and B are correlated in some way one can ask if extensivity would hold for some q. For example a system whose elements are correlated at at scales might correspond to $W(N) \sim N^\rho\ \rho > 0$ with entropy

$$S_q(N) = log_q W(N) \sim \frac{N^{\rho(1-q)} - 1}{1-q} \tag{3.31}$$

and extensivity is obtained if and only if $q = 1 - (1/\rho) < 1$ or $S_q(N) \propto N$. Shannon and Khinchin gave early similar sets of axioms for the form of the entropy functional, both leading to (3.23). These were generalized in [**5, 781, 780, 843**] leading to the entropy

$$S(p_1, \cdots, p_W) = k\frac{1 - \sum_1^W p_i^q}{q-1} \tag{3.32}$$

and it was shown that S_q is the only possible entropy extending the Boltzman-Gibbs entropy maintaining all the basic properties except extensivity for $q \neq 1$.

Some other properties are also discussed, e.g. bias, concavity, and stability. First note

$$S_{BG} = -\left[\frac{d}{dx}\sum_1^W p_i^x\right]_{x=1} \tag{3.33}$$

(x here is referred to as a bias). Similarly

$$S_q = -\left[D_q \sum_1^W p_i^x\right]_{x=1};\ D_q h(x) = \frac{h(qx)-h(x)}{qx-1} \tag{3.34}$$

(Jackson derivative) and this may open the door to quantum groups (see e.g. [**192**]). As for concavity consider for $p_i'' = \mu p_i + (1-\mu)p_i'$ $(0 < \mu < 1)$ concavity defined via

$$S(\{p_i''\}) \geq \mu S(\{p_i\}) + (1-\mu)S(\{p_i'\}) \tag{3.35}$$

It can be shown that S_q is concave for every $\{p_i\}$ and $q > 0$. This implies theormdynamic stability in the framework of statistical mechanics (i.e. stability of the system with regard to energetic perturbations). This means that the entropy functional is defined such that the stationary state (thermodynamic equilibrium) makes it extreme.

There are also other generalizations of the BG entropy and we mention the Renyi entropy

$$S_q^R = \frac{log\sum_1^W p_i^q}{1-q} = \frac{log[1+(1-q)S_q]}{1-q} \tag{3.36}$$

and an entropy due to Landsberg, Vedral, Rajagopal, Abe defined via

$$S_q^N = S_q^{LVRA} = \frac{1-\frac{1}{\sum_1^W p_i^q}}{1-q} = \frac{S_q}{1+(1-q)S_q} \tag{3.37}$$

These are however not concave nor experimentally robust and seem unsuited for thermodynamical purposes; on the other hand Renyi entropy seems useful for geometrically characterizing multifractals.

Various connections of S_q to thermodynamics are indicated in [**944**] and we mention here first the Legendre structure. Thus for all values of q

$$\frac{1}{T} = \frac{\partial S_q}{\partial U_q};\ T = \frac{1}{k\beta};\ U_q = -\frac{\partial}{\partial\beta}log_q Z_q; \tag{3.38}$$

$$log_q Z_q = \frac{Z_q^{1-q}-1}{1-q} = \frac{\bar{Z}^{1-q}-1}{1-q} - \beta U_q;\ F_q = U_q - TS_q = -\frac{1}{\beta}log_q Z_q$$

Here $U_q \sim$ internal energy and $F_q \sim$ free energy and the specific heat is

$$C_q = T\frac{\partial S_q}{\partial T} = \frac{\partial U_q}{\partial T} = -T\frac{\partial^2 F_q}{\partial T^2} \tag{3.39}$$

Finally a list of other properties follows supporting the thesis that S_q is a correct road for generalizing the BG theory (see [**944**] for details and references); we mention a few here via

(1) Boltzmann H-theorem (macroscopic time irreversibility) $q(dS_q/dt) \geq 0$ $(\forall q)$
(2) Ehrenfest theorem: For an observable $\hat{O}$ and a Hamiltonian $\hat{H}$ one has $d < \hat{O} >_q /dt = (i/\hbar) < [\hat{H}, \hat{O} >_q$ $(\forall q)$
(3) Pesin theorem (connection between sensitivity to initial conditions and the entropy production per unit time). Define the q-generalized Kolmogorov-Sinai entropy as

$$K_q = lim_{t\to\infty} lim_{W\to\infty} lim_{N\to\infty} \frac{< S_q > (t)}{t} \tag{3.40}$$

where N is the number of initial conditions, W is the number of windows in the partition (fine graining), and t is discrete time (cf. also [**593**]). The q-generalized Lyapunov coefficient λ_q can be defined via sensitivity to initial conditions

$$\xi = lim_{\Delta x(0)\to 0} \frac{\Delta x(t)}{\Delta x(0)} = e_q^{\lambda_q t} \tag{3.41}$$

(focusing on a 1-D system, basically $x(t+1) = g(x(t))$ with g nonlinear). It was proved in [**73**] that for unimodal maps $K_q = \lambda_q$ if $\lambda_q > 0$ and $K_q = 0$ otherwise. More explicitly $K_1 = \lambda_1$ if $\lambda_1 \geq 0$ (and $K_1 = 0$ if $\lambda_1 < 0$). But if $\lambda_1 = 0$ then there is a special value of q such that $K_q = \lambda_q$ if $\lambda_q \geq 0$ (and $K_q = 0$ if $\lambda_q < 0$).

We refer also to [**30, 31, 785, 636**] for other results and approaches to thermodynamics, temperature, fluctuations, etc. in generalized thermostatistics and to [**595**] for relativistic nonextensive thermodynamics.

4. FISHER PHYSICS

The book [**381**] purports (with notable success) to unify several subdisciplines of physics via Fisher information and this theme appears also in many papers, e.g. [**217, 262, 239, 382, 383, 384, 385, 660, 755, 776, 777, 778, 779, 780, 781, 949**]. We sketch some of this here and note in passing an interesting classical-quantum trajectory in [**386**] which differs from a Bohmian trajectory (cf. also [**93, 269, 376, 579, 629, 812, 999**]). First let us sketch some summary items from [**381**] and then provide some details. Thus in Chapter 12 of [**381**] Frieden lists (among other things) the following items:

(1) Writing $p = q^2$ in the standard formulas one can express the Fisher information as $I = 4\int dx(dq/dx)^2$ with q a real probability amplitude for fluctuations in measurement. Under suitable conditions (see below) the information I obeys an I-theorem $dI/dt \leq 0$. In the same spirit by which a positive increment in therodynamic time corresponds to an increase in Boltzman entropy there is a positive increment in Fisher time defined by a decrease in information I (the two times do not always agree). Let θ be

the measured phenomenon and define the Fisher temperature T_θ via

$$\frac{1}{T_\theta} = -k_\theta \frac{\partial I}{\partial \theta} \;\; (k_\theta = const.) \tag{4.1}$$

When θ is taken to be the sysem energy E then the Fisher temperature has analogous properties to the ordinary Boltzman temperature, in particular there is a perfect gas law $\bar{p}V = k_E T_E I$ where $\bar{p}$ is the pressure. The I theorem can be extended to a multiparameter, multicomponent scenario with

$$I = 4\int dx \sum_n \nabla q_n \cdot \nabla q_n \tag{4.2}$$

(2) Any measurement of physical parameters initiates a transformation of Fisher information $J \to I$ connecting the phenomenon with the "intrinsic data". The phenomenological or "bound" information is denoted by J and the acquired information is I; J is ultimately identified by an invariance principle that characterizes the measured phenomenon. In any exchange of information one must $\delta J = \delta I$ (conservation law) and for $K = I - J$ one arrives at a variational principle (extreme physical information or EPI) $K = I - J = extremum$. Since $J \geq I$ always the EPI zero principle involves $I - \kappa J = 0$ $(0 \leq \kappa \leq 1)$. These equations follow (independently of the axiomatic approach taken and of the I-theorem) if there is e.g. a unitary transformation connecting the measurement space with a physically meaningful conjugate space. In this manner one arrives at the Lagrangian approach to physics, often using the Fourier transform to connect I and J. This seems a little mystical at first but many convincing examples are given involving the SE, wave equations, KG equation, Dirac equation, Maxwell equations, Einstein equations, WDW equation, etc.

There is much more summary material in [**381**] which we omit here. A certain amount of metaphysical thinking seems necessary and Frieden remarks that John Wheeler (cf. [**988**]) anticipated a lot of this in his remarks that "All things physical are information-theoretic in origin and this is a participatory universe....Observer participancy gives rise to information and information gives rise to physics." Going now to [**381**] recall $I = \int dx[(p')^2/p] = 4\int dx(q')^2$ for $p = q^2$ and one derives the inequality $e^2 I \geq 1$ as follows. Look at estimators $\hat{\theta}$ satisfying

$$< \hat{\theta}(y) - \theta >= \int dy[\hat{\theta}(y) - \theta]p(y|\theta) = 0 \tag{4.3}$$

where $p(y|\theta)$ describes fluctuations in data values y. Hence

$$\int dy(\hat{\theta} - \theta)\frac{\partial p}{\partial \theta} - \int dyp = 0 \tag{4.4}$$

Use now $\partial_\theta p = p(\partial log(p)/\partial\theta)$ and normalization to get $\int dy(\hat{\theta}-\theta)(\partial log(p)/\partial\theta)p = 1$ which becomes

(4.5)

$$\int dy\left[\frac{\partial log(p)}{\partial\theta}\sqrt{p}\right][(\hat{\theta}-\theta)\sqrt{p}] = 1 \Rightarrow \left[\int dy\left(\frac{\partial log(p)}{\partial\theta}\right)^2 p\right]\left[\int dy(\hat{\theta}-\theta)^2 p\right] \geq 1$$

For $e^2 = \int dy(\hat{\theta}-\theta)^2 p$ this gives immediately $e^2 I \geq 1$. One notes that if $p(y|\theta) = p(y-\theta)$ then I is simply $I = \int dx(\partial log(p(x))/\partial x)^2 p(x)$ where $x \sim y-\theta$. We recall also the Shannon entropy as $H = -\int dxp(x)log(p(x))$ and the Kullback-Leibler entropy is defined as

$$G = -\int dxp(x)log\frac{p(x)}{r(x)} \tag{4.6}$$

where $r(x)$ is a reference probability distribution function (PDF). Consider now a discrete form of Fisher information

(4.7)

$$I = (\Delta x)^{-1}\sum_n \frac{[p(x_{n+1}) - p(x_n)]^2}{p(x_n)} = (\Delta x)^{-1}\sum_n p(x_n)\left[\frac{p(x_n+\Delta x)}{p(x_n)} - 1\right]^2$$

Here $p(x_n+\Delta x)/p(x_n)$is close to 1 for Δx small and one writes $[p(x_n+\Delta x)/p(x_n)]-1 = \nu$. Then $log(1+\nu) \sim \nu - (\nu^2/2)$ or $\nu^2 = 2[\nu - log(1+\nu)]$. Hence I becomes

$$I = -2(\Delta x)^{-1}\sum_n p(x_n)log\frac{p(x_n+\Delta x)}{p(x_n)} + \tag{4.8}$$

$$2(\Delta x)^{-1}\sum_n p(x_n+\Delta x) - 2(\Delta x)^{-1}\sum_n p(x_n)$$

But each of the last two terms is $(\Delta x)^{-1}$ by normalization so they cancel leaving

$$I = -\frac{2}{\Delta x}\sum p(x_n)log\frac{p(x_n+\Delta x)}{p(x_n)} \to -\frac{2}{\Delta x}G[p(x), p(x+\Delta x)] \tag{4.9}$$

One notes (cf. [**381**]) that I results as a cross information between $p(x)$ and $p(x+\Delta x)$ for many different types of information measure, e.g. Renyi and Wooters information and in this sense serves as a kind of "mother" information. Next the I-theorem says that $dI/dt \leq 0$ and this can be seen as follows. Start with (4.9) in the form

$$I(t) = -2lim_{\Delta x\to 0}(\Delta x)^{-2}\int dx\,plog\frac{p_{\Delta_x}}{p};\ p_{\Delta_x} = p(x+\Delta x|t);\ p = p(x|t) \tag{4.10}$$

Under certain conditions (cf. [**381**]) p obeys a FK equation

$$\frac{\partial p}{\partial t} = -\frac{d}{dx}[D_1(x,t)p] + \frac{d^2}{dx^2}[D_2(x,t)p] \tag{4.11}$$

where D_1 is a drift function and D_2 a diffusion function. Then it is shown (cf. [**776, 811**]) that two PDF such as p and p_{Δ_x} that obey the FP equation have a cross entropy satisfying an H-theorem

$$G(t) = -\int dx\,plog\frac{p}{p_{\Delta_x}};\ \frac{dG(t)}{dt} \geq 0 \tag{4.12}$$

Hence I obeys an I theorem $dI/dt \leq 0$. We refer to [**381**] for more on temperature, pressure, and gas laws.

For multivariable situations one writes $I = 4\int dx \sum \nabla q_n \cdot \nabla q_n$ with $p_n = q_n^2$. An interesting notation here is

$$(4.13)\quad \psi_n = \frac{1}{\sqrt{N}}(q_{2n-1} + iq_{2n})\ \ (n = 1, \cdots, N/2);\ \sum_1^{N/2} \psi_n^* \psi_n = \frac{1}{N}\sum q_n^2 = p(x)$$

In such situations one finds for $I_n = 4\int dx \nabla q_n \cdot \nabla q_n$ and $I = \sum I_n$ (cf. [**381**])

$$(4.14)\qquad I_n = -\frac{2}{(\Delta x)^2} G_n[p_n(x|t), p_n(x + \Delta x|t)];\ \frac{\partial I_n}{\partial t} \leq 0; I(t) \to\ min.$$

Now one looks at minimization problems for I where $\delta I[\mathbf{q}(\mathbf{x}|t)] = 0$ and for anything meaningful to happen the physics has to be introduced via constraints and covariance (we refer to [**381**] for a more thorough discussion of these matters). Thus one is considering $K = I - J$ and the physics is introduced via J. One can write e.g. $I = \int dx \sum i_n(x)$ and $J = \int dx \sum j_n(x)$ where $i_n = 4\nabla q_n \cdot \nabla q_n$. In general now the functional form of J follows from a statement about invariance for the system. Examples of invariance are (i) unitary transformations such as that between the space and momentum space in QM (ii) gauge invariance as in EM or gravitational theory (iii) a continuity equation for the flow, usually involving sources. The answer $\mathbf{q}$ for EPI is completely dependent on the particular $J(\mathbf{q})$ for that problem and that in turn depends completely on the invariance principle that is used. If the invariance principle is not sufficiently strong in defining the system then one can expect the EPI output $\mathbf{q}$ to be only approximately correct. One has $I \leq J$ generally but $I = J$ for an optimally strong invariance principle. Note $\kappa = I/J$ measures the efficiency of the EPI in transferring Fisher information from the phenomenon (specfied by J) to the output (specified by I). Thus $\kappa < 1$ indicates that the answer $\mathbf{q}$ is only approximate. When the invariance principle is the statement of a unitary transformation between the measurement space and a conjugate coordinate space then the solution to the requirement $I - \kappa J = 0$ will simply be the reexpression of I in the conjugate space; when this holds then one can show that in fact $I = J$ (i.e. $\kappa = 1$). In this situation the out put $\mathbf{q}$ will be "correct", i.e. not explicitly incorrect due to ignored quantum effects for example. There are in fact nonquantum and nonunitary theories for which $\kappa = 1$ (or in fact any real number) and the nature of κ is not yet fully understood.

Let us call attention also to the information demon of Frieden and Soffer (cf. [**381, 384**]). For real Fisher coordinates x the EPI process amounts to carrying through a zero sum game betweem an observer (who wants to acquire maximal information) and an information demon (who wants to minimize the information transfer) with a limited resource of intrinsic information. The demon represents nature (and always wins or breaks even of course) and $K = I - J \leq 0$. Further since $\Delta I = K$ one has $\Delta I \leq 0$ while $\Delta t \geq 0$; hence the I-theorem follows.

We run through the EPI procedure here for the KG equation which illustrates many points. Define $x_1 = ix,\ x_2 = iy,\ x_3 = iz,\ x_4 = ct$ with $r = (x, y, z)$ and

$\mathbf{x} = (x_1, x_2, x_3, x_4)$ and use the ψ_n notation of (4.13). From $I = 4\int dx \sum \nabla q_n \cdot \nabla q_n$ we get

$$I = 4Nc\sum_1^{N/2}\int\int drdt\left[-(\nabla\psi_n)^*\cdot\nabla\psi_n + \left(\frac{1}{c^2}\right)^2(\partial_t\psi_n)^*(\partial_t\psi_n)\right] \tag{4.15}$$

The invariance principle here involves a unitary Fourier transformation from x to μ in the form

(4.16)

$$(ir, ct) \to (i\mu/\hbar, E/c\hbar);\ \psi_n(r,t) = \frac{1}{(2\pi\hbar)^2}\int\int d\mu dE\phi_n(\mu, E)e^{-i(\mu\cdot r - Et)/\hbar}$$

One recalls

$$\int\int drdt\psi_m^*\psi_n = \int\int d\mu dE\phi_m^*\phi_n \tag{4.17}$$

Differentiating in (4.16) one has $(\nabla\psi_n, \partial_t\psi_n) \to (-i\mu\phi_n/\hbar, iE\phi_n/\hbar)$ and via $\nabla\psi_n \sim -i\mu\phi_n/\hbar$ one gets

$$\int\int drdt(\nabla\psi_n)^*\cdot\nabla\psi_n = \frac{1}{\hbar^2}\int\int d\mu dE|\phi_n(\mu,E)|^2\mu^2; \tag{4.18}$$

$$\int\int drdt(\partial_t\psi_n)^*\partial_t\psi_n = \frac{1}{\hbar^2}\int\int d\mu dE|\phi_n(\mu,E)|^2E^2$$

Putting this in (4.15) gives

$$I = \left(\frac{4Nc}{\hbar^2}\right)\sum_1^{N/2}\int\int d\mu dE|\phi_n(\mu,E)|^2\left(-\mu^2 + \frac{E^2}{c^2}\right) = J \tag{4.19}$$

This is the invariance principle for the given scenario. The same value of I can be expressed in the new space (μ, E) where it is called J and J is then the bound (physical) information. Now one has from (4.17)

$$c\int\int drdt|\psi_n|^2 = c\int\int d\mu dE|\phi_n|^2\ \ (n = 1, \cdots, N/2) \tag{4.20}$$

Summing over n and using $p = \sum_1^{N/2}\psi_n^*\psi = (1/N)\sum q_n^2$ with normalization gives

$$1 = \int d\mu de P(\mu, E);\ P(\mu,E) = c\sum_1^{N/2}|\phi_n(\mu,E)|^2 \tag{4.21}$$

so P is a PDF in the (μ, E) space. One obtains then

(4.22)

$$I = J = \frac{4N}{\hbar^2}\int\int d\mu dEP(\mu,E)\left(-\mu^2 + \frac{E^2}{c^2}\right);\ J = \left(\frac{4N}{\hbar^2}\right)\left\langle -\mu^2 + \frac{E^2}{c^2}\right\rangle$$

One must have J a universal constant here so $-\mu^2 + (E^2/c^2) = const. = A^2(m,c)$ where A is some function of the rest mass m and c (which are the only other parameters ($\hbar$ must also be a constant). By dimensional analysis $A = mc$ so $E^2 = c^2\mu^2 + m^2c^4$ which links mass, momentum, and energy. This defines coordinates μ and E as momentum and energy values. One has then $I = 4N(mc/\hbar)^2 = J$ and the intrinsic information I in the 4-position of a particle is proportional to the square of its intrinsic energy mc^2. Since J is a universal constant (see comments below), c is fixed, and given that $\hbar$ has been fixed, one concludes that the rest mass m is a universal constant. Since I measures the capacity of the observed phenomenon to provide information about (in this case) 4-length it follows that I should translate into a figure for the ultimate fluctuation (resolution) length that is intrinsic to QM. Here the information is $I = (4N/L^2)$ with $L = \hbar/mc$ the reduced Compton wavelength. If all N estimates have the same accuracy some argument then leads to $e_{min} = L$ and e_{min} corresponds to a minimal resolution length (i.e. ability to know). Finally putting things together one gets

$$J = \frac{4Nm^2c^3}{\hbar^2}\int\int d\mu dE\sum_1^{N/2}\phi_n^*\phi_n = \frac{4Nm^2c^3}{\hbar^2}\int\int drdt\sum_1^{N/2}\psi_n^*\psi \tag{4.23}$$

$$K = I - J = \tag{4.24}$$

$$= 4Nc\sum_1^{N/2}\int\int drdt\left[-(\nabla\psi_n)^*\cdot\nabla\psi_n + \left(\frac{1}{c^2}\right)\partial_t\psi_n^*\partial_t\psi_n - \frac{m^2c^2}{\hbar^2}\psi_n^*\psi_n\right]$$

There is much more material in [**381**] to enhance and refine the above ideas. There are certain subtle features as well. In 4-dimensions the Fourier transform is unitary and covariance is achieved in all variables (treating t separately as in $q_n(x|t)$ is not a covariant formalism). EPI treats all phenomena as being statistical in origin and every Euler-Lagrange (EL) equation determines a kind of QM for the particular phenomenon (think here of the q_n as fields). This includes classical electromagnetism for example where the vector potential A is considered as a kind of probability "amplitude" for photons. In 4-D the Lorentz transformation satisfies the requirement that Fisher information I is invariant under a change of reference frame and this property is transmitted to J and K. Thus invariance of accuracy (or of error estimation) under a change of reference frame leads to the Lorentz transformation and to the requirement of covariance. Historically the classical Lagrangian has often been a contrivance for getting the correct answers and a main idea in [**381**] is to present a systematic approach to deriving Lagrangians. The Lagrangian represents the physical information $k(\mathbf{x}) = \sum k_n(\mathbf{x})$, $k_n(\mathbf{x}) = i_n(\mathbf{x}) - j_n(\mathbf{x})$, and $\int k(\mathbf{x})$ is the total physical information K for the system. The solution to the variational problem for the Lagrangian can represent then (for real coordinates) the payoff in a mathematical information game (e.g. the KG equation is a payoff expression). We exhibit now a derivation of the SE from [**381**] to show robustness of the EPI scheme. The position of a particle of mass m is measured as a value $y = x + \theta$ where x is a random excursion whose probability amplitude law $q(x)$ is sought. Since the time t is being ignored here one is in effect looking for a stationary solution to the problem. Note the issue of covariance does not arise here and the time dependent SE is not treated since in particular it is not covariant; it can however be obtained from the KG equation as a nonrelativistic limit. Assume that the particle is moving in a conservative field of scalar potential $V(x)$ with total energy W conserved. One defines complex wave functions as before and can write

$$I = 4N\sum_1^{N/2}\int dx\left|\frac{d\psi_n(x)}{dx}\right|^2 \tag{4.25}$$

A Fourier transform space is defined via $\psi_n(x) = (1/\sqrt{2\pi\hbar}\int d\mu\phi_n(\mu)exp(-i\mu x/\hbar)$ where $\mu \sim$ momentum. The unitary nature of this transformation guarantees the validity of the EPI variational procedure. One uses the Parseval theorem to get

$$I = \frac{4N}{\hbar^2}\int d\mu\mu^2\sum_n|\phi_n(\mu)|^2 = J \tag{4.26}$$

This corresponds to (4.19) and is the invariance principle for the given measurement problem. The x-coordinate expressions analogous to (4.20) and (4.21) show that the sum in (4.26) is actually an expectation $J = (4N/\hbar^2) < \mu^2 >$. Now use the specifically nonrelativistic approximation that the kinetic energy E_{kin} of the particle is $\mu^2/2m$ and then

$$J = \frac{8Nm}{\mu^2} < E_{kin} >= \frac{8Nm}{\hbar^2} < [W - V(x)] >= \tag{4.27}$$
$$= \frac{8Nm}{\hbar^2}\int dx[W - V(x)]\sum|\psi_n(x)|^2$$

where the last expression is the PDF $p(x)$. This J is the bound information functional $J[q] = J(\psi)$ and $\kappa = 1$ here. This leads to a variational problem

$$K = N\sum_1^{N/2}\int dx\left[4\left|\frac{d\psi_n(x)}{dx}\right|^2 - \frac{8m}{\hbar^2}[W - V(x)]|\psi_n(x)|^2\right] = extremum \tag{4.28}$$

The Euler-Lagrange equations are then ($\psi^*_{nx} = \partial\psi^*_n/\partial x$)

$$\frac{d}{dx}\left(\frac{\partial\mathfrak{L}}{\partial\psi^*_{nx}}\right) = \frac{\partial\mathfrak{L}}{\partial\psi^*_n};\ \psi''_n(x) + \frac{2m}{\hbar^2}[W - V(x)]\psi_n(x) = 0 \tag{4.29}$$

which is the stationary SE. Since the form of equation (4.29) is the same for each index value n the scenario admits $N = 2$ degrees of freedom $q_n(x)$ or one complex degree of freedom $\psi(x)$; hence the SE defines a single complex wave function. Since this derivation works with a real coordinate x the information transfer game is being played here and the payoff is the Schrödinger wave function.

REMARK 6.4.1. There are generalizations of EPI to nonextensive information measures in [**217, 262, 239**] (cf. also [**755, 756, 776, 778**]).

4.1. LEGENDRE THERMODYNAMICS. We go to the last paper in [**382**] which provides a discussion of Fisher thermodynamics and the Legendre transformation. It is shown that the Legendre transform structure of classical thermodynamics can be replicated without change if one replaces the entropy S by the Fisher information I. This produces a thermodynamics capable of treating equilibrium and nonequilibrium situations in a traditional manner. We recall the Shannon information measure $S = -\sum P(i)log[P(i)]$; it is known that if one chooses the Boltzmann constant as the informational unit and identifies Shannon's entropy with the thermodynamic entropy then the whole of statistical mechanics can be elegantly reformulated without any reference to the idea of ensemble. The success of thermodynamics and statistical physics depends crucially on the Legendre structure and one shows now that such relationships all hold if one replaces S by the Fisher information measure. We recall that for $\int g(x,\theta)dx = 1$ one writes $I = \int dx g(x,\theta)[\partial_\theta g/g]^2$ and for shift invariant g one has $I = \int dx[(g')^2/g]$. There are two approaches to using Fisher information, EPI and minimum Fisher information (MFI), and both lead to the same results here. We write (shifting to a probability function f)

$$\int dx f(x,\theta) = 1;\ I[f] = \int dx F_{Fisher}(f);\ F_{Fisher}(f) = f(x)[f'/f]^2 \tag{4.30}$$

Assume that for M functions $A_i(x)$ the mean values $< A_i >$ are known where

$$< A_i >= \int dx A_i(x) f(x) \tag{4.31}$$

This represents information at some appropriate (fixed) time t. The analysis will use MFI (or EPI) to find the probability distribution $f_I = f_{MFI}$ that extremizes I subject to prior conditions $< A_i >$ and the result will be given via solutions of a stationary Schrödinger like equation. The Fisher based extremization problem has the form $(F(f) = F_{Fisher}(f))$

$$\delta_f\left[I(f) - \alpha < 1 > - \sum_1^M \lambda_i < A_i >\right] = 0 \equiv \tag{4.32}$$
$$\delta_f\left[\int dx\left(F(f) - \alpha f - \sum_1^M \lambda_i A_i f\right)\right] = 0$$

Variation leads to $((\alpha, \lambda_1, \cdots, \lambda_M)$ are Lagrange multipliers)

$$\int dx\delta f\left[(f)^{-2}\left(\frac{\partial f}{\partial x}\right)^2 + \frac{\partial}{\partial x}\left(\frac{2}{f}\frac{\partial f}{\partial x}\right) + \alpha + \sum_1^M \lambda_i A_i\right] = 0 \tag{4.33}$$

and on account of the arbitrariness of δf this yields

$$(f)^{-2}(f')^2 + \frac{\partial}{\partial x}(2/f)f') + \alpha + \sum_1^M \lambda_i A_i = 0 \tag{4.34}$$

The normalization condition on f makes α a function of the λ_i and we assume $f_I(x,\lambda)$ to be a solution of (4.34) where $\lambda \sim (\lambda_i)$. The extreme Fisher information is then

$$I = \int dx f_I^{-1}(\partial_x f_I)^2 \tag{4.35}$$

Now to find a general solution of (4.34) define $G(x) = \alpha + \sum_1^M \lambda_i A_i(x)$ and write (4.34) in the form

$$\left[\frac{\partial log(f_I)}{\partial x}\right]^2 + 2\frac{\partial^2 log(f_I)}{\partial x^2} + G(x) = 0 \tag{4.36}$$

Make the identification $f_I = (\psi)^2$ now the introduce a new variable $v(x) = \partial log(\psi(x))/\partial x$. Then (4.36) becomes

$$v'(x) = -\left[\frac{G(x)}{4} + v^2(x)\right] \tag{4.37}$$

which is a Riccati equation. This leads to

$$u(x) = exp\left[\int^x dx[v(x)]\right] = exp\left[\int^x dx\frac{dlog(\psi)}{dx}\right] = \psi; \tag{4.38}$$

$$-\frac{1}{2}\psi''(x) - \frac{1}{8}\sum_1^M \lambda_i A_i(x)\psi(x) = \frac{\alpha}{8}\psi(x)$$

where the Lagrange multiplier $\alpha/8$ plays the role of an energy eigenvalue and the sum of the $\lambda_i A_i(x)$ is an effective potential function $U(x) = (1/8)\sum_1^M \lambda_i A_i(x)$. We note (in keeping with the Lagrangian spirit of EPI) that the Fisher information measure corresponds to the expectation value of the kinetic energy of the SE. Note also that (4.38) has multiple solutions and it is reasonable to suppose that the solution leading to the lowest I is the equilibrium one. Now standard thermodynamics uses derivatives of the entropy S with respect to λ_i and $< A_i >$ and we start from (4.35) and write after an integration by parts

$$\frac{\partial I}{\partial \lambda_i} = \int dx \frac{\partial f_I}{\partial \lambda_i}\left[-f_I^{-2}(f_I')^2 - \frac{\partial}{\partial x}\left(\frac{2}{f_I}f_I'\right)\right] \tag{4.39}$$

Comparing this to (4.34) one arrives at

$$\frac{\partial I}{\partial \lambda_i} = \int dx \frac{\partial f_I}{\partial \lambda_i}\left[\alpha + \sum_1^M \lambda_j A_j\right] \tag{4.40}$$

which on account of normalization yields

$$\frac{\partial I}{\partial \lambda_i} = \sum_1^M \lambda_j \frac{\partial}{\partial \lambda_i}\int dx f_I A_j(x) \equiv \frac{\partial I}{\partial \lambda_i} = \sum_1^N \lambda_j \frac{\partial}{\partial \lambda_i} < A_j > \tag{4.41}$$

This is a generalized Fisher-Euler theorem whose thermodynamic counterpart is the derivative of the entropy with respect to the mean values. One computes easily

$$\sum_i \frac{\partial I}{\partial \lambda_i}\frac{\partial \lambda_i}{\partial < A_j >} = \sum_i\sum_k \lambda_k \frac{\partial < A_k >}{\partial \lambda_i}\frac{\partial \lambda_i}{\partial < A_j >} \Rightarrow \frac{\partial I}{\partial < A_j >} = \lambda_j \tag{4.42}$$

as expected. The Lagrange multipliers and mean values are seen to be conjugate variables and one can also say that $f_I = f_I(\lambda_1, \cdots, \lambda_M)$.

Now as the density f_I formally depends on $M+1$ Lagrange multipliers, normalization $\int dx f_I(x) = 1$ makes α a function of the λ_i and we write $\alpha = \alpha(\lambda_1, \cdots, \lambda_M)$. One can assume that the input information refers to the λ_i and not to the $< A_i >$. Introduce then a generalized thermodynamic potential (Legendre transform of I) as

$$\lambda_J(\lambda_1, \cdots, \lambda_M) = I(< A_1 >, \cdots, < A_M >) - \sum_1^M \lambda_i < A_i > \tag{4.43}$$

Then

$$\frac{\partial\lambda_J}{\partial\lambda_i} = \sum_1^M \frac{\partial I}{\partial <A_j>}\frac{\partial <A_j>}{\partial\lambda_i} - \sum_1^M \lambda_j \frac{\partial <A_j>}{\partial\lambda_i} - <A_i) = - <A_i> \quad (4.44)$$

where (4.42) has been used. Thus the Legendre structure can be summed up in

$$\lambda_J = I - \sum_1^M \lambda_i <A_i>;\ \frac{\partial\lambda_J}{\partial\lambda_i} = - <A_i>;\ \frac{\partial I}{\partial <A_i>} = \lambda_i; \quad (4.45)$$

$$\frac{\partial\lambda_i}{\partial <A_j>} = \frac{\partial\lambda_j}{\partial <A_i>} = \frac{\partial^2 I}{\partial <A_i>\partial <A_j>};$$

$$\frac{\partial <A_j>}{\partial\lambda_i} = \frac{\partial <A_i>}{\partial\lambda_j} = -\frac{\partial^2\lambda_J}{\partial\lambda_i\partial\lambda_j}$$

As a consequence one can recast (4.41) in the form

$$\frac{\partial I}{\partial\lambda_i} = \sum_1^M \lambda_j \frac{\partial}{\partial\lambda_j} <A_i> \quad (4.46)$$

Thus the Legendre transform structure of thermodynamics is entirely translated into the Fisher context.

4.2. FIRST AND SECOND LAWS. We go here to [**779**] where one shows the coimplication of the first and second laws of thermodynamics. Thus macroscopically in classical phenomenological thermodynamics the first and second laws can be regarded as independent statements. In statistical mechanics an underlying microscopic substratum is added that is able to explain thermodynamics itself. Of this substratum a microscopic probability distribution (PD) that controls the population of microstates is a basic ingredient. Changes that affect exclusively microstate population give rise to heat and how these changes are related to energy changes provides the essential content of the first law (cf. [**809**]). In [**779**] one shows that the PD establishes a link between the first and second laws according to the following scheme.

- Given: An entropic form (or an information measure) S, a mean energy U and a temperature T, and for any system described by a microscopic PD p_i a heat transfer process via $p_i \to p_i + dp_i$ then
- If the PD p_i maximizes S this entails $dU = TdS$ and alternatively
- If $dU = TdS$ then this predetermines a unique PD that maximizes S.

For the second law one wants to maximize entropy S with M appropriate constraints A_k which take values $A_k(i)$ at the microstate i; the constrains have the form

$$<A_k> = \sum_i p_i A_k(i) \ \ (k = 1, \cdots M) \quad (4.47)$$

The Boltzman constant is k_B and assume that $k = 1$ in (4.47) corresponds to the energy E with $A_1(i) = \epsilon_i$ so that the above expression specializes to

$$U = <A_1> = \sum p_i\epsilon_i \quad (4.48)$$

One should now maximize the "Lagrangian" Φ given by

$$\Phi = \frac{S}{k_B} - \alpha \sum_i p_i - \beta \sum_i p_i \epsilon_i - \sum_2^M \lambda_k \sum_i p_i A_k(p_i) \tag{4.49}$$

in order to obtain the actual distribution p_i from the equation $\delta_{p_i}\Phi = 0$. Since here one is interested just in the "heat" part the last term on the right of (4.49) will not be considered. It is argued that if p_i changes to $p_i + dp_i$ because of $\delta_{p_i}\Phi = 0$ one will have

$$0 = \frac{dS}{k_B} - \beta dU \tag{4.50}$$

(note $\sum_i \delta p_i = 0$ via normalization). Since $\beta = 1/k_B T$ we get $dU = TdS$ so MaxEnt implies the first law.

The central goal here is to go the other way so assume one has a rather general information measure of the form

$$S = k \sum_i p_i f(p_i) \tag{4.51}$$

where $k \sim k_B$. The sum runs over a set of quantum numbers denoted by i (characterizing levels of energy ϵ_i) that specify an appropriate basis in Hilbert space, $\mathcal{P} = \{p_i\}$ is an (as yet unknown) probability distribution with $\sum p_i = constant$, and f is an arbitrary smooth function of the p_i. Assume further that mean values of quantities A that take the value A_i with probability p_i are evaluated via

$$< A >= \sum_i A_i g(p_i) \tag{4.52}$$

In particular the mean energy U is given by $U = \sum_i \epsilon_i g(p_i)$. Assume now that the set $\mathcal{P}$ changes in the fashion

$$p_i \to p_i + dp_i; \ \sum dp_i = 0 \tag{4.53}$$

(the last via $\sum p_i = constant$. This in turn generates corresponding changes dS and dU and one is thinking here of level population changes, i.e. heat. To insure the first law one assumes ($\bullet$) $dU - TdS = 0$ and as a consequence of ($\bullet$) a little algebra gives (up to first order in the dp_i the condition

$$\epsilon_i g'(p_i) - kT[f(p_i) + p_i f'(p_i)] = 0 \tag{4.54}$$

This equation is now examined for several situations. First look at Shannon entropy with

$$f(p_i) = -log(p_i); \ g(p_i) = p_i \tag{4.55}$$

In this situation (4.54) becomes

$$-\epsilon_i = kT[log(p_i) + 1] \Rightarrow p_i = \frac{1}{e} exp(-\epsilon_i/kT) \tag{4.56}$$

After normalization this is the canonical Boltzmann distribution and this is the only distribution that guarantees obedience to the first law for Shannon's information measure. A posteriori this distribution maximizes entropy as well with U as a constraint which establishes a link with the second law. Several other measures are considered, in particular the Tsallis measure, and we refer to [**779**] for details. In summary if one assumes entropy is maximum one immediately derives the first law and if you assume the first law and an information measure this predetermines a probability distribution that maximizes entropy.

REMARK 6.4.2. There is currently a great interest in acoustic wave phenomena, sound and vortices, acoustic spacetime, acoustic black holes, etc. A prime source of material involves superfluid physics à la Volovik [**968, 969**] and Bose-Einstein condensates (see e.g. [**39, 40, 86, 81, 101, 115, 369, 370, 912, 913, 963**]). We had originally written out material from [**912, 913**] in preparation for sketching some material from [**968**]. However we realized that there is simply too much to include in this book; at least 2-3 more chapters would be needed to even get off the ground.