Journal of Statistical Physics, Vol. 116, Nos. 1/4, August 2004 (© 2004)

New Kinds of Entropy

S. F. Edwards[1]

Received June 14, 2003; accepted January 20, 2004

Two new entropies are offered in this paper. Neither belongs to the traditional conservative statistical mechanics. The first recognizes that the discovery that there are states of granular or glassy systems which can be restricted by an appropriate experimental pathway implies that an entropy can be defined for them by deriving an appropriate Boltzmann equation in which the pathway to the steady state has an increasing entropy until the state characterized by a compactivity $X = \frac{\partial V}{\partial S}$ is reached. The second problem is to consider a state where energy flows in and flows out, leaving a steady state, e.g., a steady turbulent flow. Time dependent correlation functions are well defined and hence a probability of the history of the system is defined. There are many *ad hoc* studies of such systems. We ask whether such systems will possess an entropy defined by the usual $\int P \log P$ but now P is the probability of the entire history of the system and the integral is over all histories.

KEY WORDS: Granular entropy; Lagrangian entropy.

1. INTRODUCTION

The concept of entropy appears not only in physics, but in many other disciplines usually under the name of information: Shannon information, Fisher information etc. These are related to our familiar entropy in papers by Jaynes(1) and many others. In this paper I propose to study two different kinds but both of physics. The first concerns purely dissipative systems where energy is of minor importance, but the system can take up reproducible states which can be quantified. A prime example is that of a granular system.(2) By for example shaking a powder, any movement is rapidly terminated by friction, but the system can and does have a well defined volume. Ordinary statistical mechanics of a conservative system gives a

---

[1] Polymers and Colloids Group, Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, United Kingdom: e-mail: sfe11@phy.cam.ac.uk

0022-4715/04/0800-0029/0 © 2004 Plenum Publishing Corporation

unique state at zero temperature, but it is well known that the density of for example deformed spheres can vary from a minimum (minimum random closepacked) to a maximum (maximum random closepacked) across a range of values which can be reproduced under conditions discussed below. Such a system has no energy and (though it is rather meaningless) zero temperature. I will show that an appropriately prepared powder has an entropy which satisfies the essential Boltzmann structure that under the appropriate disturbance (to be developed in detail below), this entropy always increases $\frac{\partial S}{\partial t} \geqslant 0$ until a distribution is achieved wherein $\frac{\partial S}{\partial t} = 0$ and the system has a well defined "temperature" which I call the compactivity X,

$$X = \frac{\partial V}{\partial S}. \tag{1.1}$$

Note that this is very different from conventional statistical mechanics because for rigid perfectly rough grains, once established into a volume V and at the reversible curve, are indifferent to pressure changes. Applying external or body forces does not affect the configuration, but the configuration is essential to be able to work out the force distribution resulting from applied forces. The second new entropy stems from the fact time dependent correlation function are now commonplace in experimental physics and just as for example a density correlation function

$$\langle \rho(r)\, \rho(r') \rangle = \int \rho(r)\, \rho(r')\, e^{-\frac{H}{kT}}\, \mathrm{d}(all), \tag{1.2}$$

a time dependent correlation function

$$\langle \rho(r,t)\, \rho(r',t') \rangle = \int \rho(r,t)\, \rho(r',t')\, P([r])\, \mathrm{d}(all), \tag{1.3}$$

where $P([r])$ gives the distribution in time of all the coordinates (here generally labelled $[r]$) specifying the system. For example a harmonic oscillator $x(t)$ in a Brownian heat bath can have an equilibrium distribution

$$e^{-\frac{mx^2}{2kT}}, \tag{1.4}$$

but the distribution of the whole trajectory of the motion

$$\mu \dot{x} + \omega_0^2\, x = \eta \tag{1.5}$$

where η is noise

$$\langle \eta(t)\, \eta(t') \rangle = D\delta(t-t') \tag{1.6}$$

or

$$(i\mu\omega + \omega_0^2)\, x(\omega) = \eta(\omega) \tag{1.7}$$

$$\langle \eta(\omega)\, \eta(\omega') \rangle = D\delta(\omega+\omega') \tag{1.8}$$

from

$$P([\eta]) = e^{-\int \frac{d\omega}{D} |\eta(\omega)|^2} \tag{1.9}$$

hence

$$P([x]) = Je^{-\int \frac{d\omega}{D} (\mu^2\omega^2 |x|^2 + \omega_0^4 |x|^2)} \tag{1.10}$$

where J is the normalisation (= Jacobian $\eta \rightleftharpoons x$). The distribution (1.4) is for any one time in the equilibrium state. The distribution (1.10) gives the probability of finding an entire trajectory. Expressions like (1.10) appeared long ago in Dirac's book(3) in

$$\exp\left(-\frac{i}{\hbar}\int L\right) \tag{1.11}$$

and are familiar from the treatment of Brownian motion by a (real) Feynman path integral. Equation (1.10) is of course very simple, but what about a non-linear dissipative equation like Navier–Stokes for a stirred liquid. In a general notation, B is friction, C stirring

$$\frac{\partial A_k}{\partial t} + B_k A_k + C_{kjl} A_j A_l = \eta_k(t) \tag{1.12}$$

and η has some noise distribution. For a steady state there will be a probability distribution for $A_k = a_k$, $F(\ldots a_k \ldots)$ but this will be quite different for a driven, non conservative, system in comparison to the $\exp(-H/kT)$) of a Hamiltonian system. From F one can derive an entropy via Boltzmann's formula, and a rate of entropy production: however there must be a probability $P([A])$ which gives the whole history of the system. If one has a P, is there a significance to

$$S = -\int P \log P \prod_{\omega} \mathrm{d}A(\omega)? \tag{1.13}$$

Clearly S is not the conventional entropy, but it is an "information" and we can speculate that $P([A])$ can describe a "steady" ensemble of fluctuations, e.g., a turbulent flow in a pipe with time dependent correlation function $\langle u(r, t)\, u(r', t')\rangle$ which is a function of $t-t'$ alone. Can S be useful? And how? We emphasize that although steady but not equilibrium systems have been studied many years ago, notably by Onsager, they concern entropy production (e.g., Jaynes(1)), and as such strive to extend established thermodynamics. Here we study the disorder of a steady distribution, i.e., an "entropy" *not* an entropy production. The question is as to whether the quantity in Eq. (1.13) is useful in the evaluation of the distribution function P. It is indeed the case of a problem arising in the self-consistent expansion of the equation for (4.39). It has been found, in McComb's thesis(4) that modelling the steady state of a non-linear equation such as (1.12) by a Fokker–Planck structure contains an arbitrary degree of freedom which can be fixed by assuming that the entropy of the unresolved parameter of the steady distribution function is maximized. Explicitly, if the steady solution of (1.12) is $F(A)$, and we model by

$$\sum_k \frac{\partial}{\partial A_k}\left(D_k \frac{\partial}{\partial A_k} + \omega_k A_k\right) F(A) = 0 \tag{1.14}$$

we must find two quantities D_k and ω_k. One equation is easily found, but not the second. McComb proposed that

$$S = -\int F \log F = S(D, \omega) \tag{1.15}$$

is maximized by the best ω

$$\frac{\partial S}{\partial \omega_k} = 0. \tag{1.16}$$

Details of this procedure are in McComb's book "The Physics of Fluid Turbulence."(4) Here we ask: instead of $F(A_k)$ can we consider $P([A])$, i.e., $P(\ldots A_{k\,\omega} \ldots)$ the distribution of the whole history of the motion, and the P of (1.13) in analogous form to (1.16). Note that there is a considerable literature which uses such distributions but is focussed on entropy production, which is not what studied here, where Fokker–Planck (Hermite) expansion contains an *arbitrary* parameter and a method has the power to fix the parameter. The position adopted is that this parameter is chosen to minimize the information in the distribution. We emphasize that this is to get a "best solution;" it is not a new physical principle of entropy production.

2. THE ENTROPY OF A SYSTEM DOMINATED BY FRICTION

This section will be devoted to the entropy of powders, but there are now several papers which use computer simulation to explore related systems. Some years ago, Edwards and Oakeshott proposed$^{(5)}$ that a granular system which was "jammed," i.e., each grain is locked into position by its neighbours and which took up a configuration determined by its history but reproducible, would have a distribution controlled by its volume. Suppose the grains are specified by their contacts with their neighbours $\mathbf{C}^{\alpha\beta}$, then the volume would be a function of the $\mathbf{C}^{\alpha\beta}$'s. Call it W. Edwards and Oakeshott postulated that the microcanonical ensemble would be

$$e^{S}\int \delta(V-W)\,\Theta\,\mathscr{D}\mathbf{C}^{\alpha\beta}, \tag{2.17}$$

where Θ constrains the grains to touch and V is the volume. Here

$$S=\log\int \delta(V-W)\,\Theta\,\mathscr{D}\mathbf{C}^{\alpha\beta}. \tag{2.18}$$

In thermodynamics, entropy and energy are measured in different units so one needs Boltzmann's constant

$$S=k\log\int \delta(E-H(p,q))\,\mathscr{D}p\,\mathscr{D}q. \tag{2.19}$$

Here it is not necessary, but it makes the formulae more recognisable if we write

$$S=\lambda\log\int \delta(V-W)\,\Theta\,\mathscr{D}\mathbf{C}^{\alpha\beta}. \tag{2.20}$$

This now can go to the canonical form

$$e^{-\frac{Y}{\lambda X}}=\int e^{-\frac{W}{\lambda X}}\Theta\,\mathscr{D}\mathbf{C}^{\alpha\beta}, \tag{2.21}$$

where

$$X = \frac{\partial V}{\partial S}, \tag{2.22}$$

and

$$Y = V - XS. \tag{2.23}$$

Studies using computer simulations have shown this formulation really works (see, e.g., Behringer$^{(6)}$ and references therein), but it has been repeatedly criticised as a mere hypothesis without a proof. I now discuss what a proof would mean, and then give a proof within the discussion given.

3. THE "SECOND LAW OF THERMODYNAMICS" DERIVED FOR POWDERS

Consider the powder to be grains of high friction and almost incompressible. I say "high" and "almost" for one has to have a way of an external agency to rearrange the grains. The first systematic exploration of this by the Chicago group$^{(7)}$ tapped a column of grains a fixed number of times, but in different runs used different forces. After destroying the initial low density state, it was found that the density lay on a reversible curve, so for a given N taps of force Γ, the density was $\rho(N, \Gamma)$. Here are some curves of recent work in Cambridge on the conductivity ($\sim$ density) of anthracite, vibrated with an amplitude A and frequency ω. The Γ and the A overcame the friction in certain places, but the movement is stifled by the friction. But different Γ or A produce more or less rearrangement, giving Fig. 1. The amplitude must not be so violent as the totally disrupt the powder, which is the case of $X \to \infty$. To analyze this situation recall the classic derivation of Boltzmann. He said: suppose a gas has a low enough density so that only two body collisions matter. This surely can be arranged. Suppose that at the collision no memory persists, the stosszahlansatz, surely this must also be fine, but many lives have been wasted worrying about it. Then, if $f(r, v)$ is the probability of finding a molecule with r, v

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial r} + \int K(v, v_1; v', v_1')(f(v)\, f(v_1) - f(v')\, f(v_1'))\, \mathrm{d}(all) = 0. \tag{3.24}$$

Boltzmann introduced

$$\frac{f f_1}{f' f_1'} = x > 0 \tag{3.25}$$

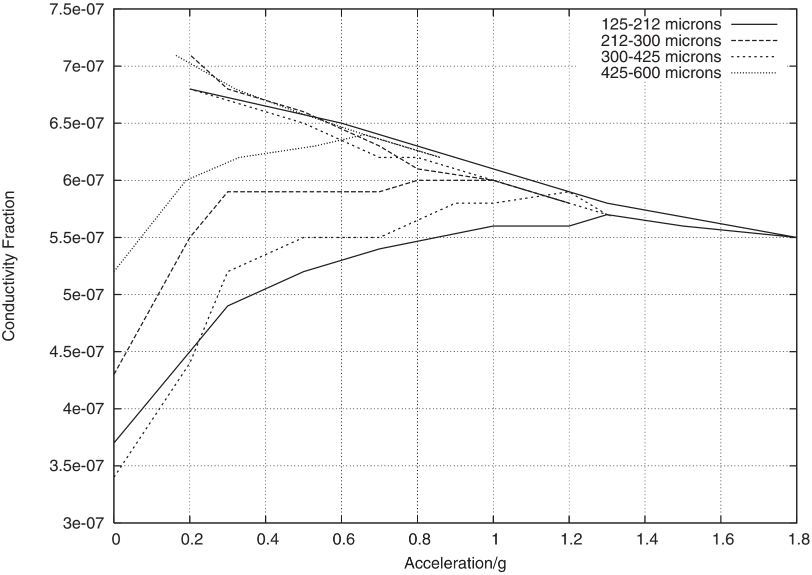

Fig. 1. Normalised conductivity fraction as a function of the dimensionless acceleration amplitude for grains of different sizes. The packings were prepared in a low packing fraction state by sequential deposition. The acceleration amplitude was first slowly increased and then decreased. The conductivity fraction was recorded after 10^5 taps at each value of the acceleration amplitude. The lower branches are irreversible and depend on the deposition history and particle size. The upper branches for grains of different sizes are reversible, i.e., upon subsequently raising the value of the acceleration amplitude again, the conductivity fraction retraces the values measured on the downward trajectory.

and

$$S = -k \int f \log f \tag{3.26}$$

and by exploring the symmetry of $\tilde{K}$ showed that

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \int \tilde{K}(x-1) \log x, \qquad \tilde{K} > 0 \tag{3.27}$$

and therefore, since

$$(x-1) \log x \geqslant 0, \tag{3.28}$$

$$\frac{\mathrm{d}S}{\mathrm{d}t} \geqslant 0. \tag{3.29}$$

$\tilde{K}$ contains the conservation of energy

$$v^2+v_1^2=v'^2+v_1'^2 \tag{3.30}$$

and hence $\frac{\mathrm{d}S}{\mathrm{d}t}=0$ when $f\sim e^{\epsilon v^2}$ or

$$f=e^{-\frac{mv^2}{2kT}} \tag{3.31}$$

in the usual notation and identification.

Clearly a granular system cannot operate 2-body contacts! In fact Newtonian force analysis shows that on average only four contacts are needed (and I believe it will always be exactly four but don't develop that here). it is easier to draw in 2D when three contacts obtain. A typical picture is Fig. 3. Concentrate attention on α, β_1, β_2, β_3. If all the others are fixed, these four can be moved around within certain limits. Suppose then that inflicting an external perturbation on the system finds a region, such as our α, β_1, β_2, β_3, wherein the threshold of friction can be overcome, but which leaves the surrounding matrix below threshold. Suppose one associated a volume W^α with the configuration of α, β_1, β_2, β_3. Then W^α is conserved when the particular tap or vibration is imposed, but the actual configuration of $W^\alpha(\alpha,\beta_1,\beta_2,\beta_3)$ is changed to $W^{\alpha'}$. This will happen throughout the system, but the overall volume is unchanged

$$\sum W^\alpha=W^{\alpha'}. \tag{3.32}$$

Thus we can construct a Boltzmann equation with the probability $f^\alpha(W^\alpha)$ of finding a configuration W^α, providing that we can associate a W^α with the particle α. If this is so, then

$$\frac{\partial f^\alpha}{\partial t}+\int K(f^\alpha f^{\beta_1}f^{\beta_2}f^{\beta_3}-f^{\alpha'}f^{\beta_1'}f^{\beta_2'}f^{\beta_3'})=0 \tag{3.33}$$

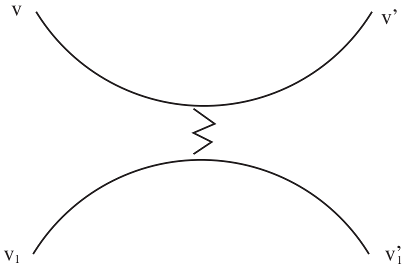

Figure 2.

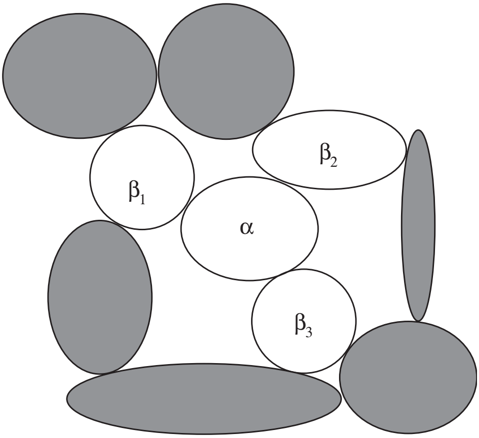

Fig. 3. First coordination shell of a reference grain in the "fixed" cage of second nearest neighbours.

where K is the kernel which depends on N, Γ or A, ω or whatever mechanism is being used. K contains (3.32). Hence the solution must be

$$f^{a} = e^{\frac{Y - W^{\alpha}}{X}} \tag{3.34}$$

where $X = \frac{\partial V}{\partial S}$ as expected by Edwards and Oakeshott.(5) This solution is the canonical ensemble when the microcanonical ensemble will be $\delta(V - W)$ and W will be $\sum W^{\alpha}$. Further to this however is that a good approximation to W^{α} can be derived in terms of $W^{\alpha} \cdots W^{\beta_3}$ where the W^{α} carry the label of the one grain via the tensor F^{α}_{ij}. In Fig. 4 we introduce $\mathbf{R}^{\alpha}$ which is the centroid of the contacts of α with β_1, β_2, β_3. Then if $\mathbf{R}^{\alpha\beta} = \mathbf{R}^{\alpha} - \mathbf{R}^{\beta}$

$$F^{\alpha}_{ij} = \sum_{\beta} (R^{\alpha}_i - R^{\beta}_i)(R^{\alpha}_j - R^{\beta}_j) \tag{3.35}$$

a good approximation to the area (volume) of the system is

$$W = 2 \sum_{\alpha}^{N} \sqrt{\operatorname{Det} F^{\alpha}_{ij}}. \tag{3.36}$$

Exact expressions can be obtained by Voronoi or Delaunay methods, but these do not have a single label. Another exact method has been given by Ball and Blumenfeld(9) which is appropriate to granular systems but is more complicated than the proposal (3.36). Experiments on packed grains by Brujic(11) show that (3.36) is a very good approximation to the volume of the system and we adopt it here to justify (3.33). This enables us to

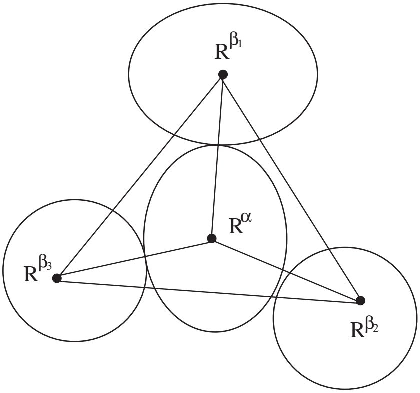

Fig. 4. Centroids of contact in the first coordination shell of a reference grain α.

consider the probability of the particular configuration of Fig. 5 to be $f^{\alpha} f^{\beta_1} f^{\beta_2} f^{\beta_3}$. Other configurations will have the appropriate number of f's. Notice the key point is that there is a definite volume $V(A, \omega)$ when A and ω are applied indefinitely. If a new volume is the result of A', ω', then when these parameters cease and one returns to A and ω, one returns to $V(A, \omega)$. This is the meaning of the reversible curve. Since there is a definite volume it means that the Mohr–Coulomb or similar criteria which tell one where rearrangements take place when A, ω are applied, lead to the distribution of regions $(a_1 \cdots a_i)$. We do not need to know where these are, or the detailed kernel K, to obtain the equivalent of an H-theorem, but of course if we wanted a fuller description of the physical processes we would need these. This is exactly the same as in the conventional Boltzmann equation. This can be considered for the whole material in the volume V. Applying A, ω or N, Γ one finds a series of domains whose threshold is breached and which rearrange. The steady state has at A, ω

$$V = \sum_{\text{domains}} W^{a_i}, \tag{3.37}$$

i.e.,

$$P = e^{\frac{Y - \sum_{a_i} W^{a_i}}{X}} \tag{3.38}$$

and X is $X(A, \omega)$ hence $S = S(A, \omega)$ and $V = V(A, \omega)$ as found in experiment. Having established that one can now effect many of the standard problems of statistical mechanics such as miscibility, stress propagation⁽⁸⁾ and so on.

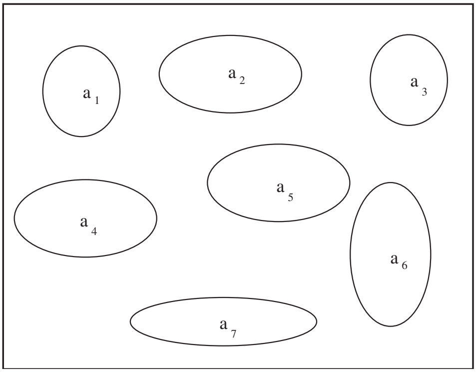

Fig. 5. Domains in the given volume V.

4. LAGRANGIAN ENTROPY

I use this phrase to contrast the idea of an entropy of an entire dynamical history which is steady in time, but contains all the time dependent correlation information, as in the example discussed above. Thus in contrast to the normal entropy which only discusses the values of the system at one definite time, the "Lagrangian Entropy" does not need to be defined in Hamiltonian system, and indeed seems most useful in systems where energy enters by some mechanism such as stirring, and leaves via friction. These can of course be tackled in the usual structure so that in (1.12) one can seek a $P(\ldots, A_k,\ldots; t)$ such that

$$\frac{\partial P}{\partial t}+\sum_k \frac{\partial}{\partial A_k}(B_k\ A_k+C_{kjl}\ A_jA_l)+\eta_k\frac{\partial P}{\partial A_k}=0 \tag{4.39}$$

when averaged over η gives (using same symbol P)

$$\frac{\partial P}{\partial t}-\sum_k \frac{\partial}{\partial A_k}D_k\frac{\partial P}{\partial A_k}+\sum_k \frac{\partial}{\partial A_k}(B_kA_k+C_{kjl}A_jA_l)\ P=0. \tag{4.40}$$

From this a transport equation can be deduced for

$$|A_k|^2=\phi_k \tag{4.41}$$

say

$$\frac{\partial \phi_k}{\partial t}+B_k\phi_k+\int K_{kjl}\phi_j\phi_l-\int K_{jkl}\phi_j\phi_k-\int K_{jlk}\phi_k\phi_l=D_k \tag{4.42}$$

where B_k gives loss due to friction, D_k the input via η and K transports the probability (or energy, or fluctuations according to what $|A_k|^2$ represents physically. The derivation of such an equation is given in McCombs book and recently renewed in detail by Edwards and Schwartz(10) and I will not repeat it here. In the steady state, provided the kernel K can be produced,

$$B_k\phi_k + \int K_{kjl}\phi_j\phi_l - \int K_{jkl}\phi_j\phi_k - \int K_{jlk}\phi_k\phi_l = D_k \tag{4.43}$$

summarizes the problem and ref. 10 gives a derivation of the kernel in terms of the effective lifetime of a mode k, ω_k. Assume this derivation is sound, we find that in KPZ for example B, D do not dominate the power law region of the solution, and

$$\int K_{kjl}\phi_j\phi_l - \int K_{jkl}\phi_j\phi_k - \int K_{jlk}\phi_k\phi_l = 0 \tag{4.44}$$

gives an excellent value for the index. The problem in such a derivation is that the model transport equation is

$$\frac{\partial P_0}{\partial t} - \int \frac{\partial}{\partial A_k}\left(D_k\,\frac{\partial}{\partial A_k} + \omega_k\,A_k\right)P_0 = 0 \tag{4.45}$$

where ϕ_k is $\frac{D_k}{\omega_k}$. Thus one expands (4.39) around (4.42) and finds a best self-consistent fit. The form (4.42) is the most general possible model from the point of view of making an expansion possible and one can fit

$$\int P = 1 \tag{4.46}$$

$$\int A_k\,A_j\,P = \delta(k+j)\,\phi_k \tag{4.47}$$

but one needs another condition to get ω_k and some mean life-time argument seems the obvious way to go. However Edwards and Schwartz(10) show that it is really rather difficult in situations such as fluid turbulence to define ω_k from (4.42) and it is better to work from the four-dimensional equation for $\langle A(r,t)\,A(r',t')\rangle$, i.e.,

$$\langle A_{k\omega}A_{j\sigma}\rangle = \delta(k+j)\,\delta(\omega+\sigma)\,\Phi_{k,\omega}. \tag{4.48}$$

Use capital letter K for k, ω. We want Φ_K and clearly will need a $\mathscr{D}_K$ to replace D_k and an Ω_{0K} to replace $i\omega + B\,k$. The equation of motion now becomes

$$\sum_K \frac{\partial}{\partial A_K}\left(\Omega_{0K} A_K + \sum MAA - f\right) P = 0. \tag{4.49}$$

A neat way to build up the analogue to (4.42) is to add a dimension s and consider our $d+1$ dimensional problem evolving in the s dimension, i.e.,

$$\frac{\partial A_K}{\partial s} = \left[\Omega_{0K} A_K + \sum MAA - f + g\right] \tag{4.50}$$

where $\langle g \rangle = 0$ and

$$\langle g_K(s)\, g_{-K}(s')\rangle = \mathscr{D}_K\, \delta(s - s'). \tag{4.51}$$

A little problem here f is *not* dependent on s and hence is a quenched variable. Proceeding as before we reach

$$\Omega_{0K}\Phi_K - 2\sum \frac{M\,M}{\Omega + \Omega + \Omega}\Phi\Phi - \sum \frac{|M|^2}{\Omega + \Omega + \Omega}\Phi\Phi = \frac{\mathscr{D}_0}{\Omega_{-K}}. \tag{4.52}$$

The strange fact is that the expansion (4.52), when the full series is employed is valid for <u>a</u>ny Ω_k. So we need another criterion for the <u>b</u>est Ω_k. We emphasize here that if we calculate S form P it is not related at all to entropy production, it is simply a measure of the information in P and will be $S(\phi, \Omega)$. It seems an interesting hypothesis to try $\frac{\partial S}{\partial \Omega} = 0$ as defining the missing equation for Ω. If we do a crude version with $\Omega = i\omega + \omega_k$ this reduces to $\frac{\partial S}{\partial \omega_k}$ where S is now obtained from the steady distribution. This has been successfully applied by Edwards and McComb to evaluate the front-factor in Kolmogoroff turbulence and is described in McComb's book on turbulence$^{(4)}$ cited in Section 1. An obvious proposal is to use the power expansion for P which is

$$P = P_0\left(1 - \sum \frac{M A_J A_L A_{-K}}{(\sum \Omega)\, A_K} + \cdots\right) \tag{4.53}$$

to evaluate S and apply

$$\frac{\partial S}{\partial \Omega} = 0 \tag{4.54}$$

or if we do a crude version where $\Omega = i\omega + \omega_k$

$$\frac{\partial S}{\partial \omega_k} = 0 \tag{4.55}$$

This has been successfully applied by Edwards and McComb$^{(12)}$ and is developed in McComb's book on turbulence.$^{(4)}$ But what about histories?

We now reach the difficult point. Why should S be maximized? Is it possible to produce a Boltzmann like argument that $\frac{\partial S}{\partial s} \geqslant 0$ reaching a maximum at an expression analogous to $e^{-\frac{H}{kT}}$. The only way I can see of doing this is to consider "s" as a representation of (irrelevant) boundary conditions, meaning that the ensemble of systems which have different initial conditions or different but irrelevant boundary conditions will "fill" $d+1$ dimensional space as fully as possible. The key point here is that conventional statistical mechanics has concentrated on Hamiltonian systems. The real world of friction and random inputs does not have theorems. it can only be described by explicitly derived transport equations. We have shown that a much more general approach should be possible, an approach which cannot be a naive extension by Hamiltonian statistical mechanics.

ACKNOWLEDGMENTS

This work has benefitted from my collaborations with Moshe Schwartz, David McComb and Dmitri Grinev. It is dedicated to Elliot Lieb, my first research student, who is as rigorous and accurate a worker as I am vague; but all sorts are needed in a world.

REFERENCES

1. E. T. Jaynes, *Phys. Rev.* **106**:620 (1957); *ibid.* **108**:171 (1957); see also R. D. Rosenkrantz and E. T. Jaynes, *Papers on Probability, Statistics, and Statistical Physics* (Kluwer, 1983).
2. A. Mehta, ed., *Granular Matter: An Interdisciplinary Approach* (Springer-Verlag, New York, 1993).
3. P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1930).
4. W. D. McComb, *The Physics of Fluid Turbulence* (Oxford University Press, 1990).
5. S. F. Edwards and R. B. S. Oakeshott, *Physica A* **157**:1080 (1989).
6. R. P. Behringer, *Nature* **415**:594 (2002).
7. E. R. Nowak, J. B. Knight, M. L. Povinelli, H. M. Jaeger, and S. R. Nagel, *Powder Technology* **94**:79 (1997).
8. S. F. Edwards and D. V. Grinev, *Phys. Rev. Lett.* **82**:5397 (1999).
9. R. C. Ball and R. Blumenfeld, *Phys. Rev. Lett.* **88**:115505-1 (2002).
10. S. F. Edwards and M. Schwartz, *Physica A* **303**:357 (2002).
11. J. Brujic, Ph.D. thesis (University of Cambridge, 2004).
12. S. F. Edwards and W. D. McComb, *J. Phys. A* **2**:157 (1969).