PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Foundations of Non-Equilibrium Statistical Mechanics

J. P. Dougherty

Phil. Trans. R. Soc. Lond. A 1994 **346**, 259-305
doi: 10.1098/rsta.1994.0022

Foundations of non-equilibrium statistical mechanics†

By J. P. Dougherty

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, U.K.

Contents

Two separate 'schools' have offered what are claimed to be general foundations for the subject of non-equilibrium statistical mechanics, based respectively on subdynamics and information theory. These appear to be unrelated, to start from conflicting interpretations of probability, to generate different methodology and to be useful in non-overlapping applications. The present paper is a study of the relation between the work of the two schools. After reviewing them critically in a presentation that unifies the notation used, the connection between them

† This paper was produced from the author's disk by using the TEX typesetting system.

Phil. Trans. R. Soc. Lond. A (1994) **346**, 259–305
Printed in Great Britain

is developed. It is concluded, subject to some provisos, that they are equivalent. It is suggested that the subject could be unified on this basis, and that the methodologies be made more generally available.

1. Introduction

Statistical mechanics concerns the connection between microscopic and macroscopic dynamics. There is an immediate difficulty: the macroscopic properties of a specimen of matter (supposed in a condition of isolation) are characterized observationally by the tendency to evolve towards states of *equilibrium*. Microscopic dynamics, taken here to mean classical hamiltonian dynamics, reveals no such concept of equilibrium, and, as is well known, the governing equations are reversible (Loschmidt's paradox).

Accepting the undoubted existence of equilibrium, the subject is split into equilibrium and non-equilibrium statistical mechanics. Foundations for the equilibrium case are provided by the microcanonical and canonical distributions and although opinions vary as to how these are to be justified, they constitute a working basis for theoretical studies of the properties of matter *in equilibrium*.

By comparison, non-equilibrium statistical mechanics (NESM) is an undeveloped subject. One widespread opinion is that it cannot be given general foundations analogous to those of equilibrium statistical mechanics. In one sense that may be so, as there is virtually no limitation on the scope of the questions that may be asked about time-dependent phenomena. In the present paper we limit attention to a particular class of question. Given a macroscopic system, a definite level of reduced description is to be adopted, and we ask for evolution equations for the parameters or fields which specify that description. The 'motion' of the system as so described would then be capable of prediction. In practice, much of the work attempted in NESM falls into this class.

Within this narrower meaning to NESM, various writers have sought and proposed general foundations for the subject. Here we shall consider two broad and very contrasted approaches, which also, to a considerable extent, form personal 'schools', though it is not intended to suggest that the work of other writers may be disregarded.

The first of these is the 'Brussels school', associated with Professor I. Prigogine and numerous co-authors. This development starts from microscopic dynamics (Hamilton's or Schrödinger's equations) and it is regarded as important not to introduce any non-dynamical ideas, such as might be imported from information theory: everything should be derived from dynamics alone. Naturally approximations, perturbation expansions and the like have to be introduced if progress is to be made. The aim is thus a dynamical separation into the parameters to be retained in the desired description, and those to be discarded. This separation, if it exists, is to emerge from dynamics as an asymptotic property, valid for large time. There are several versions and extensions of the theory, but we shall concentrate here on the so-called method of *subdynamics* (Prigogine *et al.* 1969; Balescu 1975).

The approach of this school may be regarded as descended from that of Boltzmann, who, in setting up his equation in the kinetic theory of gases, sought

to take account of microscopic dynamics by detailed consideration of collisions. However, he found that he could discard the unwanted aspects of these only by means of his famous 'hypothesis of molecular chaos'. This line of development may be further traced through the Ehrenfest (1912) encyclopaedia article, the methodology of Chapman and Enskog (Chapman & Cowling 1970) and the well-known article of Bogoliubov (1962), who introduced a more general hypothesis than Boltzmann's. These developments use the concept of *probability*, but the underlying philosophy need not be emphasized, and most physicists would conceive it in terms of numerous repetitions of an experiment. The lack of precision is then traced to the practical impossibility of reproducing (in microscopic terms) the initial preparation of the system.

The second school, which we shall refer to as the 'Maxent school', is that associated with Professor E. T. Jaynes and his collaborators. It is important to realize that their approach amounts not merely to the development of a different methodology. It originates in an alternative interpretation for probability itself, namely the adoption of the bayesian viewpoint. In several recent articles, Jaynes has used the term 'predictive statistical mechanics', which well conveys the idea. The prediction of the future macroscopic evolution of a system cannot be done with certainty on the basis of initial macroscopic data (in physicists' words, there are fluctuations). This is a consequence of the continuing influence of the parameters or data which are supposed to be discarded in the selected description. The problem is thus subject to an intrinsic ambiguity, suggesting that the dynamical equations themselves are insufficient; the introduction of further hypotheses such as those of Boltzmann and Bogoliubov can be seen as attempts to escape the ambiguity. The work of this school proposes instead that a proper theory of statistical inference should replace these *ad hoc* hypotheses. Conventional statistical inference, of which the school is deeply critical, is not adopted; instead, the bayesian philosophy leads to the use of information theory, and hence to the algorithm of entropy maximization, to set up a theory of NESM. Details of the development of this work up to 1983 (including the criticism of statistical inference) can be conveniently read in a reprint collection of papers by Jaynes† (some of which are otherwise difficult of access) published in 1983; we may also refer to the text of Grandy (1988).

This approach may be regarded as descended from that of Gibbs; in the equilibrium case the 'method of most probable state' involves the same mathematics as that of entropy maximisation, and is very straightforward. The extension to the non-equilibrium case is not immediately obvious, and, for the class of problems considered in the present paper, is technically difficult; the attractive simplicity of the method in equilibrium statistical mechanics is not inherited. This is not surprising, as it is inevitable that the dynamical equations should have a detailed influence on the problem of time evolution. The early development of this approach (Jaynes, pp. 4–39) appeared in 1957, following that of information theory itself in the late 1940s. Jaynes records that he was in possession of a generalization to NESM in 1962; see Jaynes, p. 239, and his book review (1968), not included in the 1983 collection.

† Throughout the present article, reference to material contained in this book will be cited in the form 'Jaynes', followed by a page number.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

In parallel with the above development, there is also the independent work of Zubarev, which appeared in Soviet journals in the early 1960s, and is described in his book of 1974 (all of this work is available in English translation, but the book supersedes the journal articles for most purposes). This work starts from the point of view of Bogoliubov (and indeed under his influence), and makes no mention of bayesian inference, yet reaches results which (in the context of the class of problem considered here) can be transformed into those of the Maxent school (as we demonstrate below, § 17). Zubarev proceeds intuitively by adopting so-called 'quasi-equilibrium' distributions as the basis for an approximation scheme. In the subsequent development he cites an analogy with the Gell-Mann–Goldberger theory of scattering in quantum mechanics as justification for his methods.

At this point, it may well be asked how the work of the two 'schools' has been received by statistical physicists generally, i.e. by non-members. I think it fair to answer that their reception must have seemed disappointing. One reason may be the point, already referred to, that many physicists do not expect there to exist a general formalism along the lines considered. In other words, intuition and experience lead them to conclude that an unwise question is being asked. Anybody holding that view can of course be excused from devoting time to the matter, but he or she might nevertheless be interested to see it proved to be wrong!

In my opinion, the styles of publication adopted by the two main schools have not served their causes well. In the case of the Brussels school the number of publications has been very large, many of them cross-referenced internally within the school, in such a way as to overwhelm the outsider; the book by Balescu (1975) is a help, but this is itself somewhat eclectic within the school as a whole, and is rather daunting. On the other hand, at least until the appearance of a review article by Grandy (1980), which was later superseded by his book, (1988), the Maxent school suffered from paucity of publications, which tended to be in obscure places, and to have the appearance of preliminary announcements, so far as technical details are concerned, with heavy emphasis upon philosophical debate. And while it is plain that neither school sees merit in the other, this has to be deduced not from any passages of comment in depth, but rather from the almost complete absence of reference by either to the work of the other.

More cogently, it will be asked in this article whether the methods developed by the respective schools lead eventually to the same formalism, or at least finally to the same results. The formalisms have such a different appearance that one may well conclude that equivalence is unlikely, or far from easy to demonstrate. But here one should recall that practical results may turn out to be identical even if derived from apparently different distributions, an elementary example being the microcanonical and canonical distributions in equilibrium.

One might alternatively seek an application in which the two formalisms give demonstrably different results, so that one (at least) of them might be falsified. Unfortunately, the technical details of such calculations are so formidable that discrepancies might result merely from the approximations and expansions used. One instance where agreement has been recorded is the case of a dilute gas with interaction potential in the form of a normal distribution (Balescu 1975; Stirland 1983), for which the kinetic equation has been derived to lowest order in the interaction. Apart from this, comparative results for a common problem do not seem to be available.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The question of equivalence (in some sense to be made clear) seems to the present author to be of great importance. If the two formalisms are indeed equivalent, it may be possible to weld them together, or to present them as special cases or aspects of something more general. This would offer the opportunity to put NESM onto general foundations which would be more widely acceptable to the majority of interested physicists, and it might be possible to choose more freely from the very different methodologies (expansions, etc.) which they use.

By contrast, it would be of the greatest interest to philosophers of science if the two formalisms could be shown to be definitely inequivalent, for that might eventually provide evidence to support one or other of the available interpretations of probability (frequency theory, subjective theory, etc.; see Fine 1973; Gillies 1973) as they apply to physical theory. I have developed this further elsewhere, with more extensive discussion of irreversibility (Dougherty 1993).

Fortunately or unfortunately (depending on one's interests) it now appears to me that the formalisms are in fact equivalent. The remainder of this paper gives an outline of the arguments leading to that conclusion, though these cannot yet take the form of an unbroken, rigorous mathematical demonstration. In any case, the dynamical system must be presumed to possess a sufficient degree of stochasticity. (We use the word 'stochasticity' in a qualitative sense to suggest irregular orbits, sensitivity to initial values, and so on; 'instability' could also have been used. More precise concepts will be defined in § 5.) The subject of proving that particular dynamical systems, with a specified law of interaction, actually are ergodic, mixing, etc., is still in its infancy. We note here that the Maxent school would reject the relevance of stochasticity conditions, but we shall return to that point in its place (§ 22).

2. Notation

Here we briefly set out notation to be used, most of it being standard. We work with a classical hamiltonian many-body system consisting of N molecules and write

$$X = (\boldsymbol{q}_1, \ldots \boldsymbol{q}_N, \boldsymbol{p}_1, \ldots \boldsymbol{p}_N), \quad X \in \Gamma \tag{2.1}$$

for a phase point (microscopic state), in the phase space Γ (where $\Gamma \subset \mathbb{R}^n$ for some n), and let $\mathrm{d}\Omega = \mathrm{d}q_1 \ldots \mathrm{d}p_1 \ldots$ denote a volume element in Γ. Integration over Γ will involve the usual Lebesgue measure.

Microscopic evolution is governed by hamiltonian equations, and although these are of course immensely complicated and nonlinear, we appeal only to certain general properties that they are known or assumed to have. We shall generally assume here that the hamiltonian itself is time independent, so the equations are autonomous. (Attempts to dispense with that assumption have been made by some writers (e.g. Balescu & Misguich 1973; Jowett 1982; Coveney & George 1987).) All we need to state is that the *solution* of Hamilton's equations has the form

$$X(t) = \phi_t X(0), \tag{2.2}$$

where $\phi_t : \Gamma \to \Gamma$ is measure preserving and invertible with $\phi_{t+s} = \phi_t \phi_s$; ϕ_t is a *flow*.

As is well known, if ρ is a function $\Gamma \to \mathbb{R}$ (or $\Gamma \to \mathbb{C}$), the flow induces an

evolution of the form

$$\rho(X,t) = \rho\left(\phi_{-t}(X), 0\right). \tag{2.3}$$

This may be written

$$\rho(X,t) = U(t)\rho(X,0), \tag{2.4}$$

where $U(t)$ is a linear, real operator, with $U(t+s) = U(t)U(s)$ and $U(0) = I$. The measure-preserving property of ϕ_t gives ρ the character of a density on Γ which remains constant when convected with the flow. From this it follows that if the following conditions hold at one time, say $t = 0$

$$\text{positivity:} \qquad \rho(X) \geqslant 0, \tag{2.5}$$

$$\text{normalization:} \qquad \int_\Gamma \rho \, \mathrm{d}\Omega = 1, \tag{2.6}$$

they hold at all times. If so, ρ may be regarded as a probability density on Γ that evolves with time, and is the Liouville function. We can say this without prejudice to any choice of the philosophical interpretation of probability.

In practice, the enormous number of variables comprising X, as in (2.1), refer to numerous particles of a few chemical species, in which case the hamiltonian is symmetric to interchanges of like particles, as also is U. Although ρ need not be symmetric to such interchanges of like particles, if it is symmetric at $t = 0$ that symmetry is also preserved. As only the symmetric part has macroscopic significance, one usually restricts attention to symmetric solutions. (This applies even in classical mechanics; the symmetry conditions are of course built into the quantum case automatically.)

Equation (2.4) is itself the solution of a linear first-order partial differential equation, Liouville's equation, whose characteristics are the original orbits, (2.2). The coefficients of Liouville's equation follow from the original hamiltonian equations, but we shall write the equation formally

$$\dot{\rho} = L\rho, \tag{2.7}$$

where L is the *liouvillian*. As we are assuming that hamiltonian is time independent, so is the liouvillian. We then have a formal solution as in (2.4), with

$$U(t) = \exp(Lt). \tag{2.8}$$

If the conditions (2.5) and (2.6) are imposed, solutions do not form a vector space. It is often helpful to consider (2.7) in the more general setting in which we allow ρ to lie in a function space, E, which may also be complex. The choice of E is considered below in § 4. There is no difficulty about incorporating (or leaving out) the restriction of symmetry mentioned above. Then the liouvillian L is an *unbounded* operator, while U is a *unitary* operator. (Some writers define L so that an extra factor i appears in (2.7) and (2.8) to emphasize this last point.) This generalization enables the use of methods of functional analysis, such as spectral theory. The temporary suspension of (2.5) need not cause concern, since the Liouville equation preserves positivity. All will be well in applications since the set of non-negative solutions separates out from the generality of solutions, and the initial conditions will reimpose the positivity.

In the mathematical literature, the definition of U formed by equations (2.3) and (2.4) has the opposite sign convention, as obtained by replacing ϕ_{-t} by ϕ_t on

the right-hand side of (2.3); this follows the choice made in the pioneering paper of Koopman (1931). The two versions of U that result are merely adjoints of each other and have identical properties in the present context. We have preferred the physicist's choice that makes U the operator of Liouville's equation for the future time. The book by Lasota & Mackay (1985) uses both operators, referred to with the names 'Koopman' and 'Frobenius–Perrin', respectively.

Although a completed theory of this subject would only need to refer to continuous time, we shall often refer to results from ergodic theory in the simpler context of discrete time, so $t \in \mathbb{Z}$.

3. Observables

The introduction of a probability density (or measure) in the previous section enables us to compute means of random variables, or, in physicists' terminology, expectation values of observables. The choice of such observables needs careful consideration.

It will always be assumed that an observable is invariant with respect to permutations of like particles. This rules out observations that would purport to identify individual particles.

An elementary form of observable is one which is given by a real scalar function $a(X)$, so it could, in principle, be computed whenever the microscopic state is exactly known: it just depends on the positions and momenta of all the particles. We can then form the mean,

$$\langle a \rangle = \int_\Gamma a(X)\rho(X,t)\,\mathrm{d}\Omega\ . \tag{3.1}$$

We shall always assume that the prescription of $a(X)$ is not explicitly time dependent, so $\langle a \rangle$ can vary with time only if ρ does. The same applies to the more general observables to be introduced shortly. Removal of this limitation at some stage may be desirable, but is not attempted here.

Elementary examples of observables are the ones that appear in equilibrium, such as the total energy (i.e. the hamiltonian), momentum, magnetic moment. The normalization (2.6) is merely the statement $\langle 1 \rangle = 1$. Generally there is a set $\{a_k\}$ to which (3.1) applies, with values lying in a vector space A. Then, with $\rho \in E$, the notion of 'observation' corresponds to a linear map $E \to A$ that is expressed by a kernel $a_k(X)$, using (3.1).

In NESM, a somewhat more general form of observable is required to express the variables commonly used in theories of media. The point can be illustrated by the case of a gas composed of a single species of particles, so that X denotes the set $\{x_1, x_2, \ldots x_N\}$, x_i being the hamiltonian variables $(\boldsymbol{q}_i, \boldsymbol{p}_i)$ for the ith molecule. A possible choice for an observable is the one-particle distribution function,

$$f(x) = \int \rho\,(x, x_2, x_3 \ldots x_N)\ \mathrm{d}x_2 \mathrm{d}x_3 \ldots \mathrm{d}x_N, \tag{3.2}$$

where x is a new free variable similar in form to the x_i. In (3.2) the apparent asymmetry in the treatment of x_1 in contrast to $x_2, x_3 \ldots$ is not an objection owing to the symmetry of ρ; the expression could easily by symmetrized at the cost of a more clumsy form. Equation (3.2) can also be written in a way formally analogous to (3.1) provided the delta function $\delta(x - x_i)$ is permitted in the kernel

(the latter is then what plasma physicists call the Klimontovich function). There are similar two-particle and higher distribution functions which can in turn be re-expressed as correlation functions (Balescu 1975; Clemmow & Dougherty 1990). The outcome of this is that our observables still take the form of a linear map $M : E \to A$ though in this example A is now a function space, containing the functions $f(x)$, and the map is no longer represented by a continuous kernel.

We shall take it quite generally that observation, or measurement, is specified by a linear map $M : E \to A$, and so write

$$M\rho = a. \tag{3.3}$$

For the present, the only formal statement we make about it is that it is *non-invertible*, so that a member of A does not imply a unique member of E. There is a sense in which A is very much 'smaller' than E, representing the loss of microscopic information. E is determined by *physics*, whereas M is at the choice of *physicists*, and results in the *reduced description* of the system expressed by members of A.

We comment later (§ 19) on the manner in which nonlinearity can reappear in macroscopic physics.

4. Use of densities

Our objective is a general theory, with associated methodology, showing the source of broad features of NESM, notably irreversibility. The required generality seems inevitably to be lost as soon as the details of particular hamiltonians or interactions are introduced. One recognizes, of course that eventually those details must be acted upon, and that therein lies the origin of the remarkable variety found in the properties of matter. But for the present we need to abstract from hamiltonian mechanics rather general, perhaps qualitative, properties as our starting point.

We remark that NESM does not apply to all systems, the exceptions being very simple systems. Examples are the simple pendulum and the two-body problem (as solved by Newton for the inverse square law). These are characterized by being *integrable*, with the result that both prediction and retrodiction are mathematically possible (as in eclipses), and there is no symmetry breaking with respect to time. So the properties we are looking for must fail in the case of simple systems. This points naturally to properties of non-integrable systems that are associated with ergodic theory, chaos, sensitivity to initial conditions, and so on. First, we make use of such measurable invariants of the motion as may exist (usually comprising only the energy). Let $\Sigma \subset \Gamma$ be a submanifold of Γ defined by specified values of the invariants. Then within Σ the motion is to be characterized in some way as chaotic. In other words it is *determinate* but not, in the long term, *predictable*.

The mathematical content of the previous assertion is that the initial value problem for such a dynamical system complies with the conditions for *existence* and *uniqueness*, but that solutions are *unstable* (although continuous) with respect to small changes in initial values. The equations for the evolution of the reduced description which are the object of NESM will in general be partial differential equations. For them, the standard Poincaré conditions (existence, uniqueness and continuity with respect to initial data) would define a well-posed problem. To be useful, these evolution equations need to be well-posed, at least for posi-

tive time and at least for some ranges of parameters (failure in the latter respect would be interpreted as 'turbulence'). Our question is how to achieve such reductions. Success supplies the physicist with the means to make predictions. These cannot be perfectly reliable, however. The reason for that is that the underlying microscopic dynamical system remains capable of pursuing motions not representable by the reduced description, as is obvious from the fact that the time reversal of any motion is microscopically possible. Any such non-predicted event is termed a *fluctuation*; for a large system, appreciable fluctuations are supposed to be *non-generic*, or *highly unlikely*. This is how *probability* arises. To carry out the reduction, we must search for mathematical structure that enables the exclusion of what is non-generic. The natural candidate for this is measure theory; the idea is that the unlikely events have very small measure, tending to zero in the limit of large systems. Adopting this means that dynamics is converted to a consideration of *densities*, and, as we have already anticipated in § 3, that is our approach in the present paper. The text of Lasota & Mackey (1985) provides an excellent source for the theory of densities as it applies to dynamical systems.

We mention in passing that other possibilities exist. Topology also supplies a concept (Baire category) for distinguishing generic from non-generic, which is not equivalent to measure, and one might follow the route of topological dynamics (Brown 1976). Since we always operate in subsets of $\mathbb{R}^n$, one can in fact draw freely upon both measure-theoretic and topological ideas. Algorithmic complexity has been invoked (Ford 1983; Zeldovich *et al.* 1990, p. 303), although it remains unclear how it would be applied. The possible role of information theory is discussed below. It may be that other notions still to be invented by mathematicians of the future will provide a better approach, but for the present we continue along the route already trodden by most writers, namely that of densities. Further arguments in favour of the measure-theoretic approach, and its connection with observation, may be found in Goldstein & Penrose (1981).

The mathematical structure underlying the treatment of the dynamical system by densities thus consists of $\{\Gamma, \mathcal{B}, \mu, \phi\}$, where $\mathcal{B}$ is the family of measurable subsets, μ is the measure, i.e. the standard Lebesgue measure in $\mathbb{R}^n$, and ϕ the measure-preserving map. When we apply the restriction to a submanifold Σ, as described above, there is a further technicality involved, as the measure has to be converted into a lower-dimensional one (introducing a jacobian factor) and the appropriate family of subsets of Σ should be defined; by an abuse of notation we will continue to designate these by μ and $\mathcal{B}$ respectively. A second measure, ν, on $\mathcal{B}$ specifies the weight (i.e. 'probability') to be ascribed to the presence of the system in any $B \in \mathcal{B}$. Provided ν is non-singular ($\nu(B) = 0$ whenever $\mu(A) = 0$) we can, by the Radon-Nikodym theorem, write

$$\mathrm{d}\nu = \rho(X)\,\mathrm{d}\mu, \tag{4.1}$$

where ρ is the density already introduced. One could allow for singular measures ν, so that a single point can have a non-zero probability, by letting ρ include delta-functions. In fact, we shall exclude that possibility, and in doing so we are making a fundamental assumption about the process of observation, which is inherent in the adoption of the measure theoretic approach (Goldstein & Penrose 1981): precise measurement of the values of the qs and ps to be attached to an event is impossible. For the same reason, we adopt the usual convention of ergodic

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

theory that statements concerning sets (e.g. $B \subset C$) are regarded as true if they are true with the exception of sets of measure zero.

The interpretation of ρ as a non-singular measure implies that the function space E should be of the Lebesgue class L_1,

$$\rho \in L_1(\Gamma) \quad \text{or} \quad L_1(\Sigma), \tag{4.2}$$

as the case may demand. In practice, owing to the extensive involvement with spectral theory, we usually work in Hilbert space, in which case L_1 is replaced with L_2.

5. Levels of stochasticity

Ergodic theory has led to the formulation of seven 'levels' of stochasticity that can be attained by an invertible measure-preserving map ϕ; there are similar considerations in the case of flows, ϕ_t. The levels are, in ascending order of stringency, (a) quasi-ergodic (or 'minimal'); (b) ergodic; (c) weak mixing; (d) mixing (or 'strong mixing'); (e) countable Lebesgue spectrum; (f) K-systems; (g) Bernoulli systems.

We review them briefly. Firstly, (a) differs from the others in being a topological property, namely that $\{\phi^n(x) : n \in \mathbb{Z}\}$ is *dense* in the manifold, so an orbit eventually reaches any open set. This expresses the idea that the orbits are not confined to any submanifold of lower dimension. It is of some historical interest as it can be traced back to Boltzmann's (incorrect) conjecture that (in the case of a flow) the orbit would be a space-filling curve. It will not concern us further.

Levels (b)–(g) work in the environment of measure theory and the associated spectral theory.

Level (b) articulates, in terms of measure theory, the same idea as (a). A measurable set $B \subset \Sigma$ is *invariant* under ϕ if $\phi(B) = B$. ϕ is *ergodic* if the only invariant sets are Σ itself and the null set. There are then no non-trivial subsets of Σ into which individual orbits can be confined.

Levels (a) and (b) do not imply rapid separation of a bundle of initially close orbits. That idea is expressed by a condition of the form,

$$\mu\left(A \cap \phi^n(B)\right) \to \mu(A)\mu(B) \quad \text{as} \quad n \to \infty, \quad \forall A, B \in \mathcal{B}. \tag{5.1}$$

A *mixing* system (level (d)) is defined by (5.1), the limit sign having its usual meaning. A *weak mixing* system (level (c)) is similarly defined, but the limit is in a weaker, Cesàro, sense. The somewhat technical distinction between the two will not be important here. The meaning of (5.1) can be understood by taking B to be a small set, representing the physicist's 'error bars' associated with the initial data at $n = 0$. For large n, the new position $\phi^n(B)$ is so widely and uniformly dispersed that the probability of finding the system in an observed set A is independent of the positions of either A or B, being merely proportional to their volumes. To put it in still another way, the information provided at $n = 0$ has become useless. The concept is a precise expression of the idea put forward intuitively by Gibbs concerning the spread of a drop of ink into a large volume of water, due to stirring.

It will be convenient to return to level (e) a little later. Meanwhile level (f), introduced by Kolmogorov, depends on the concept of a *sub-algebra*, say $\mathcal{A}$, of the algebra of measurable sets $\mathcal{B}$. This is simply a family of sets, belonging to

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

$\mathcal{B}$, that is itself closed under union and intersection. We can then define the union of two such subalgebras $\mathcal{A}^{(1)} \cup \mathcal{A}^{(2)}$ as the smallest subalgebra containing all members of both, while the intersection $\mathcal{A}^{(1)} \cap \mathcal{A}^{(2)}$ is the largest subalgebra common to them both. We then consider the action of ϕ on such a subalgebra, $\phi(\mathcal{A}) = \{\phi(A) : A \in \mathcal{A}\}$, which is itself a subalgebra. The characterization of a K-system is the requirement that a subalgebra $\mathcal{A}$ shall exist satisfying

$$\bigcup_{n=-\infty}^{\infty} \phi^n(\mathcal{A}) = \mathcal{B} \quad \text{and} \quad \bigcup_{n=-\infty}^{\infty} \phi^{-n}(\mathcal{A}) = \mathcal{N}, \tag{5.2}$$

$\mathcal{N}$ being the 'trivial' subalgebra containing only the sets of measure 1 or 0. (The definition also requires ϕ to be invertible, as we are assuming throughout; recall also that sets differing only in zero measure are identified.) Roughly speaking, ϕ churns up the subsets to such an extent that any measurable set can eventually be reached. A well-known example is the Baker transformation.

Level (g) provides a way to characterize a dynamical system as being 'completely random'. This is the case if a partition of Γ into disjoint subsets (which a physicist might call 'cells') exists such that the progression of the system through the cells reveals no correlation with respect to time. We omit detailed specification or discussion of this property, since it is too strong to play a useful part in NESM. Our list includes it only for completeness.

So far we have discussed the levels of stochasticity in terms of the properties of ϕ itself. If we are to describe observations by means of densities, we ask whether these properties are reflected in useful and easily recognised features of U or L. As ρ evolves according to

$$\left.\begin{array}{lll} \rho(t) = U^t\rho(0)\,, & t \in \mathbb{Z} & \text{(discrete time)} \\ \rho(t) = U(t)\rho(0), & \text{where } U(t) = \exp(Lt) & \text{(continuous time)} \end{array}\right\} \tag{5.3}$$

it is natural to investigate the *spectrum* of U. This implies that we adopt the Hilbert space model, $\rho \in L_2(X)$. As U is unitary, its spectrum lies on the unit circle $|\lambda| = 1$. In any event $\rho =$ constant is an eigenfunction with $\lambda = 1$.

From the rather large literature on this subject we select a few results for reference. They are discussed here in the discrete-time model.

ϕ is *ergodic* if and only if the eigenvalue $\lambda = 1$ is simple. (This does not rule out the possible existence of other eigenvalues on $|\lambda| = 1$: if they exist they are also simple, and form a multiplicative group; their eigenfunctions are not constant but have constant modulus.)

As ergodicity is the lowest level of stochasticity to be considered, the previous condition will always apply, and it is convenient to factor out the simple eigenvalue at $\lambda = 1$ by working with the restriction of U to $E^{\perp}$, which consists of functions that are orthogonal to the constants, i.e. functions having zero mean.

ϕ is *weak mixing* if and only if U has a continuous spectrum in $E^{\perp}$, i.e. it has no further eigenvalues.

The properties of being (strong) mixing and K-systems are also spectral invariants, but there do not appear to be transparent statements about the spectrum analogous to the preceding. Our property (e) is a spectral characterization that does not correspond to a simply expressed statement about ϕ itself. We assume that U has the above property equivalent to weak mixing. If now there exists a sequence of L_2 functions $\{f_j\}, j = 0, 1, 2, \ldots$, with $f_0 \equiv$ const., such that

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

$\{f_0\} \cup \{U^n f_j : (j \geqslant 1, n \in \mathbb{Z})\}$ forms an orthonormal base, then ϕ is said to have a *countable Lebesgue spectrum*. (The functions, other than f_0, forming the base are of course not eigenfunctions.)

It is known that each of the properties (a)–(g) implies its predecessor. The converse is in each case false, a fact that is in most cases not easy to show; this has been achieved by ingenious counter-examples. Details of these results are given in standard texts on ergodic theory (Cornfeld *et al.* 1981; Mañé 1989; Peterson 1983; Parry 1981; Walters 1982).

6. Decomposition of function space

We return to the idea, introduced in § 3, of the linear map $M : E \to A$ that represents measurement. The discarding of information by the map implies that it has a non-trivial kernel (null-space) which we write as

$$E_2 = \ker(M) = \{\rho : M\rho = \mathbf{0}\}. \tag{6.1}$$

If $a \in A$ is a physically possible observation, and the condition $M\rho = a$ is satisfied by some $\rho = \rho_1$, then it is also satisfied by $\rho = \rho_1 + \rho_2$ for any $\rho_2 \in E_2$. The question is how, in the context of this ambiguity, to make progress with calculations concerning the evolution of $\rho(t)$ and hence of $a(t)$. The following summarizes a general idea which underlies both the Brussels school and, subject to some modification, the work of Zubarev. It can also be conceived as generalizing the earlier methods mentioned in § 1. We put it forward initially in a form appropriate to the Brussels school.

A particular manifold $\mathcal{E}_1 \subset E$ is chosen, with the property that, for $a \in A$ there is a unique solution of $M\rho = a$ with $\rho \in \mathcal{E}_1$; we call this $\rho = \rho_1$. We can think of this as a preferred choice, or perhaps a first estimate, of the ρ to correspond to that value of a. Figure 1 shows this schematically: the vertical direction represents the 'unwanted' information in E_2, while the horizontal denotes the 'wanted' information specified by a; both directions actually denote infinitely many dimensions, and our showing them at right angles is purely for the convenience of the draughtsman. The use of a script letter reminds us that $\mathcal{E}_1$ is not necessarily a linear space. We now have a unique decomposition of any ρ:

$$\rho = \rho_1 + \rho_2 \quad \text{with} \quad \rho_1 \in \mathcal{E}_1, \quad \rho_2 \in E_2. \tag{6.2}$$

There is then a map $\mathcal{P} : E \to \mathcal{E}_1$ which discards ρ_2, so $\mathcal{P}\rho = \rho_1$. The result is a 'projection' onto $\mathcal{E}_1$, but if $\mathcal{E}_1$ is not a vector subspace $\mathcal{P}$ is nonlinear so is *not* a projection operator in the sense of the usual definition. It has the obvious property (sometimes called idempotent), that repeated applications have no further effect, so $\mathcal{P}^2 = \mathcal{P}$.

The above (so far purely formal) step has now to be combined with Liouville's equation. It will only be of value if it enables schemes for the solutions of the equation to the extent of providing the evolution $a(t)$; the essence of such schemes must be the elimination of ρ_2, or at least the removal or minimization of the effect of ρ_2 in some way to be made precise. Immediate elimination of ρ_2 would occur only if $\mathcal{P}$ commutes with L (or with U), for in that exceptional case ρ_1 stays in $\mathcal{E}_1$, while ρ_2 evolves independently and gives rise to no observed effect. It could therefore be made zero. The problem of NESM could then be justifiably be said

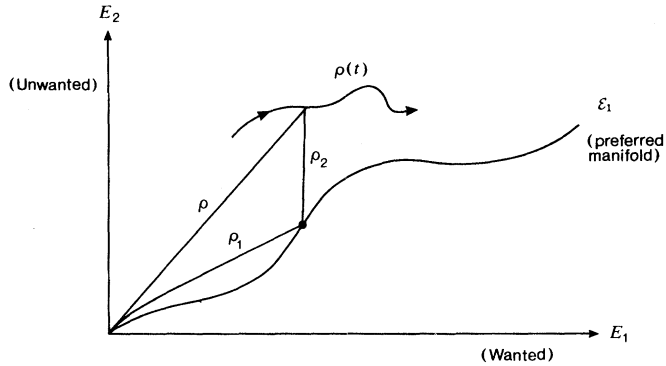

Figure 1. Decomposition of function space for solutions of the Liouville equation.

to have been solved: we would have identified an invariant set $\mathcal{E}_1 \subset E$ capable of answering all questions about the system at the level M.

In practice, no such $\mathcal{E}_1$ is obvious, so the only approach available appears to be that of making an initial choice of $\mathcal{E}_1$ and then performing formal manipulations in the hope of refining that choice. When looked at thus, the differences between the two schools (Brussels and Maxent) referred to in § 1 can be traced in part to the very different initial choices made in that respect, although there are other differences.

The Brussels school make the choice that is clearly a natural one to try, namely that $\mathcal{E}_1$ should be a *linear* manifold, i.e. a vector subspace of E. We shall then denote it by the symbol E_1, and $\mathcal{P}$ becomes a projection which we denote by P. We can then introduce the complementary projection $Q : E \to E_2$ by

$$Q = I - P \tag{6.3}$$

and we can write

$$E = E_1 \oplus E_2 \tag{6.4}$$

It is elementary to show that $P^2 = P$, $Q^2 = Q$, $PQ = 0$.

(To avoid confusion of notation, it should be emphasized that our function space E does not, *a priori*, possess an inner product, so (6.4) does not imply that E_1 and E_2 are orthogonal. If we choose E to be L_2, i.e. Hilbert space, a concept of orthogonality, unrelated to the role of the functions as probability densities, is present from the start. In Hilbert space (although not in other Banach spaces) there is a convention that 'projection' means 'orthogonal projection' (as it is convenient for projections to be self-adjoint (see, for example, Lorch 1962, p. 72)), but here we wish to suspend that convention, adding explicitly the word 'orthogonal' if it applies. Note that papers of the Brussels school generally adopt the usual convention. A writer who is conscious of the need to translate the theory into quantum mechanics naturally adopts Hilbert space terminology.)

The 'refinement' of the choice of projection to reach the idea of subdynamics is achieved by a consideration of asymptotic analysis, with the aim of giving the best representation of the long-term evolution. This development occupies most of §§ 7–11.

The work of Zubarev uses 'quasi-canonical' distributions, meaning those which are similar in form to equilibrium distributions while providing the correct data

$a(t)$. These can also be characterized as those which maximize entropy subject to the value of $a(t)$, hence the connection with the Maxent school. These distributions form a manifold $\mathcal{E}_q$ which can be adopted as an initial choice for $\mathcal{E}_1$. That task involves nonlinear transformations and other modifications, and is much less transparent than for the Brussels case outlined above, see § 13 below (the originators of those results arrive at them via different routes). This development as a whole occupies §§ 12–17.

7. The master equation

We start with a formal calculation which provides useful background for the work of both 'schools'. The presentation is (like most of what follows) exploratory in style, as rigorous details are still incomplete. We also work non-perturbatively, as this seems essential if general principles are to be kept clear, while recognizing that perturbation expansions, etc., are unavoidable in applications, and that the form they take is itself an important theoretical topic.

In the present section we can temporarily dispense with the assumption that L is time-independent with only slight inconvenience, but the projections P and Q remain time-independent. We consider the inhomogeneous Liouville equation

$$\{\partial/\partial t - L(t)\}\rho = h\delta(t-t'), h \in E, \ \rho(t) = 0 \quad \text{for} \quad t < t' \tag{7.1}$$

with solution

$$\rho(t) = G(t,t')\,h,$$

where $G : E \to E$ is the *causal Green's operator*, and vanishes for $t < t'$. From this we can obtain the solution of the general initial value problem

$$\{\partial/\partial t - L\}\rho = h(t), \quad \rho(t_0) = \rho_0$$

in the form

$$\rho(t) = G(t,t_0)\,\rho_0 + \int_{t_0}^{t} G(t,t')\,h(t')\,\mathrm{d}t' \qquad (t \geqslant t_0). \tag{7.2}$$

Now use the decomposition $\rho_1 = P\rho \in E_1$, $\rho_2 = Q\rho \in E_2$ so that the Liouville equation $\dot{\rho} = L\rho$ becomes, in 'matrix' form

$$\begin{pmatrix} \dot{\rho}_1 \\ \dot{\rho}_2 \end{pmatrix} = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}, \tag{7.3}$$

where $L_{ij} : E_j \to E_i$ are obtained from L by equations like $L_{12} = PLQ$, etc. Of these, L_{11} and L_{22} map between like spaces, and we denote their causal Green's operators by $G_{11}(t,t')$, $G_{22}(t,t')$, while L_{12} and L_{21} map between unlike spaces and do not possess Green's operators. We now have

$$(\partial/\partial t - L_{11})\,\rho_1 = L_{12}\rho_2, \tag{7.4}$$

$$(\partial/\partial t - L_{22})\,\rho_2 = L_{21}\rho_1. \tag{7.5}$$

If $L_{12} = 0$, ρ_2 is not needed in computing ρ_1, corresponding to the trivial case where the 'unwanted' information is decoupled, so we now assume $L_{12} \neq 0$, and solve (7.5) for ρ_2, regarding ρ_1 as known and taking $t_0 = 0$ in the analogue of

(7.2):

$$\rho_2(t) = G_{22}(t,0)\rho_2(0) + \int_0^t G_{22}\,(t,t')\,L_{21}\,(t')\,\rho_1\,(t')\;\mathrm{d}t'.$$

Substituting this into (7.4):

$$\left\{\frac{\partial}{\partial t} - L_{11}(t)\right\}\rho_1(t) = D(t,0)\rho_2(0) + \int_0^t \psi\,(t,t')\,\rho_1\,(t')\;\mathrm{d}t'. \tag{7.6}$$

Here

$$D\,(t,t') = L_{12}(t)G_{22}\,(t,t') \qquad \psi\,(t,t') = L_{12}(t)G_{22}\,(t,t')\,L_{21}\,(t')\,. \tag{7.7}$$

This is the *master equation*. Note that ρ_2 has *not* been completely eliminated, but appears only in the first term, as $\rho_2(0)$. But the price paid for this is that ρ_1 appears in a more complicated way in (7.6) than it did in (7.4) since ρ_1 at earlier times than t is required, and we have an integro-differential equation. The phrases 'non-markovian' or 'system with memory' are used. The connection with the theory of stochastic processes originates from the right-hand side of (7.6) which appears to contain 'random' influences due to ρ_2, which is within the 'discarded' information. (The term 'master equation' is used only in a broad, generic, sense.)

Suppose now that we find, as the result of a carefully chosen P, that $D\,(t,t')$ decreases rapidly as $t-t'$ increases (although this would be difficult to demonstrate). The same would presumably be true of $\psi\,(t,t')$. Then the initial value term in (7.6) is a 'transient', which disappears after a short time, and in the integral only recent contributions ($t-t'$ small) are appreciable. If so, the system quickly 'forgets' $\rho_2(0)$ and (7.6) becomes approximately markovian.

This idea is referred to as the *hypothesis of rapid decay*. If valid, it achieves the desired separation of ρ_1 and ρ_2, although retaining influences from ρ_2. It provides the generalization of earlier ideas due to Boltzmann, and later due to Bogoliubov, referred to earlier. In Bogoliubov's work on the kinetic theory of gases, the one-particle distribution function is designated as the 'wanted' descriptor, and the higher distributions are conjectured to reach very rapidly an 'asymptotic' state which is then maintained evolving adiabatically as the one-particle function evolves. That idea, taken literally, would imply that we solve (7.5) by deleting the $\partial/\partial t$ term, so $\rho_2 = -L_{22}^{-1}L_{21}\rho_1$, whence (7.4) gives

$$\partial\rho_1/\partial t = (L_{11} - L_{12}L_{22}^{-1}L_{21})\,\rho_1. \tag{7.8}$$

This in fact oversimplifies Bogoliubov's method; moreover (in any context) it would not lead to irreversibility.

A less drastic implementation of the hypothesis of rapid decay is to treat (7.6) as follows. We omit the term $D(t,0)\rho_2(0)$ but retain the integral term; in that form the equation would treat the time $t=0$ preferentially, whereas our aim is to deduce a general evolution equation for ρ_1. To remove the special role of $t=0$ we replace the lower limit of integration with $t=-\infty$, with error similar to that already involved with $\rho_2(0)$. The result is (setting $t'=t-\tau$)

$$\left(\frac{\partial}{\partial t} - L_{11}\right)\rho_1(t) = \int_{\tau=0}^{\infty}\psi(t,t-\tau)\rho_1(t-\tau)\,\mathrm{d}\tau. \tag{7.9}$$

In this form, provided the integral converges (i.e. 'memory' of past events is

finally lost), (7.9) involves no approximation; a solution $\rho_1(t)$ of this, with its accompanying $\rho_2(t)$, would be proposed as the best way to predict $a(t)$ in the circumstances of the problem. The resulting set of 'trajectories' through E of form $\rho = \rho_1 + \rho_2$ comprise the 'refined' $\mathcal{E}_1$ discussed earlier. But it is obviously extremely awkward to manage in practice, and to supply the solution corresponding to a given $a(0)$.

8. Subdynamics

In describing this we shall revert to our usual situation in which L is supposed to be time-independent. In this case, the Green's operators that appeared in the previous section as functions of t, t' become functions of $t - t'$ only, and the integrals become convolutions. It then becomes very convenient to use the Laplace transform and we can proceed further than before. So we define Laplace transforms

$$\tilde{\rho}(p) = \int_0^\infty \mathrm{e}^{-pt}\rho(t)\,\mathrm{d}t$$

and the Liouville equation for $t \geqslant 0$ transforms to

$$(p - L)\ \tilde{\rho}(p) = \rho(0).$$

The formal solution, transformed back by means of the Bromwich contour, is

$$\rho(t) = \frac{1}{2\pi\mathrm{i}}\int_C (p - L)^{-1}\mathrm{e}^{pt}\,\mathrm{d}p\,\rho(0) = U(t)\rho(0),$$

where C is a vertical contour to the right of all singularities. Here $U(t)$ means the same as previously, (2.8), and takes the place of $G(t, t')$, so we have the formulae

$$\tilde{U}(p) = (p - L)^{-1},\ U(t) = \theta(t)\mathrm{e}^{Lt} = \frac{1}{2\pi\mathrm{i}}\int_C (p - L)^{-1}\mathrm{e}^{pt}\,\mathrm{d}p, \tag{8.1}$$

$\theta(t)$ being the Heaviside function. We ask whether the result could be obtained by direct evaluation of the integral. We note that the singularities of the integrand are, by definition, the spectrum of L. If we assume that our dynamical system is at least weak-mixing, and if we work in $E^\perp$ (discarding the constant solution) then we know from § 5 that the spectrum is purely continuous and lies on $\mathcal{R}(p) = 0$. So C may be put infinitesimally to the right of the imaginary axis. It cannot be closed in the left half-plane as it may not cross the spectral set. The method succeeds if the integral nevertheless converges, and as $t \to \infty$ the resulting ρ displays no oscillatory terms, in fact $\rho \to 0$. These points are possible because of the absence of a discrete spectrum and ρ being non-singular; a physicist would describe it as 'phase-mixing'. A rigorous treatment of the above steps requires the Hille–Yosida theorem in the theory of semigroups of linear operators, see Goldstein (1985), especially p. 17.

To develop subdynamics we proceed similarly, and take the Laplace transform of equations (7.3) to get

$$\begin{pmatrix} p - L_{11} & -L_{12} \\ -L_{21} & p - L_{22} \end{pmatrix}\begin{pmatrix} \tilde{\rho}_1 \\ \tilde{\rho}_2 \end{pmatrix} = \begin{pmatrix} \rho_1(0) \\ \rho_2(0) \end{pmatrix}. \tag{8.2}$$

(Here, where p appears, it should strictly be multiplying an identity operator

$P \to P$ or $Q \to Q$ as the case may be, but we hope that we may omit these and similar ones without causing confusion).

The elimination to solve for $\tilde{\rho}_1(p)$, $\tilde{\rho}_2(p)$ is elementary provided the ordering of the terms is preserved, giving in turn

$$\tilde{\rho}_2 = (p - L_{22})^{-1} \{\rho_2(0) + L_{21}\tilde{\rho}_1\}, \tag{8.3}$$

$$\{p - L_{11} - L_{12}(p - L_{22})^{-1} L_{21}\}\tilde{\rho}_1 = \rho_1(0) + L_{12}(p - L_{22})^{-1}\rho_2(0). \tag{8.4}$$

This is just the Laplace transform of the master equation (7.6), and we rewrite it using abbreviations,

$$\left.\begin{aligned} \tilde{\psi}(p) &= L_{12}(p - L_{22})^{-1} L_{21}, \quad \tilde{D}(p) = L_{12}(p - L_{22})^{-1}, \\ \tilde{C}(p) &= (p - L_{22})^{-1} L_{21}, \quad \tilde{U}_{11}(p) = \left\{p - L_{11} - \tilde{\psi}(p)\right\}^{-1}, \end{aligned}\right\} \tag{8.5}$$

to give

$$\tilde{\rho}_1 = \tilde{U}_{11}(\rho_1(0) + \tilde{D}\rho_2(0)). \tag{8.6}$$

Referring back to (8.3) one can compute $\tilde{\rho}_2$, and the eventual result is

$$\begin{pmatrix} \tilde{\rho}_1(p) \\ \tilde{\rho}_2(p) \end{pmatrix} = \begin{pmatrix} \tilde{U}_{11} & \tilde{U}_{11}\,\tilde{D} \\ \tilde{C}\tilde{U}_{11} & (p - L_{22})^{-1} + \tilde{C}\tilde{U}_{11}\tilde{D} \end{pmatrix} \begin{pmatrix} \rho_1(0) \\ \rho_2(0) \end{pmatrix}. \tag{8.7}$$

The notation used here loses sight of the symmetry between 1 and 2, but is convenient in the context of our aims; note that ψ and D have a similar role to the corresponding symbols in (7.7).

Attention is now directed to the operator $\tilde{U}_{11}(p) : E_1 \to E_1$ referred to as the *partial resolvent* associated with the original L and with P. When inverting the Laplace transforms, it is the singularities of this that will make contributions to the Bromwich integral. To find them one must consider the eigenvalue problem,

$$\left\{p - L_{11} - \tilde{\psi}(p)\right\}\rho_1 = 0, \quad \rho_1 \in E_1, \tag{8.8}$$

in which p enters nonlinearly, unlike the familiar case. The operator has no particular symmetry (such as hermiticity, etc.) so the eigenvalues might appear anywhere and there might be a continuous spectrum. There have been extensive investigations of this using the powerful mathematics of spectral deformation (Courbage 1982*a*, *b*, 1986; Obcemea & Brändas 1983), although these have not closed the matter. It seems likely, however, that there are circumstances where (8.8) gives rise to a discrete spectrum, in such a way that the left- and right-eigenvectors are complete in E_1, the eigenvalues being in $\mathcal{R}(p) < 0$, indicating decay, i.e. approach to equilibrium. The most slowly decaying modes would provide the asymptotic behaviour of the system for large t. By rather lengthy algebraic manipulations, not reproduced here, it is possible to show how, in that case, one can construct formally a new, non-orthogonal, projection $\Pi : E \to E$ which can replace P, and such that $\Pi(E)$ contains the above solutions in E_1 together with their accompanying ρ_2. The range of $I - \Pi$ then contains the remaining contributions to ρ_2. Apart from possible initial transients, this offers a solution to the programme we envisaged in § 6, as depicted schematically in figure 2, in which Π is deliberately drawn at a general angle to the original axes. We have thus reached, by a different route, the 'refinement', Π, of P. Balescu (1974) gives

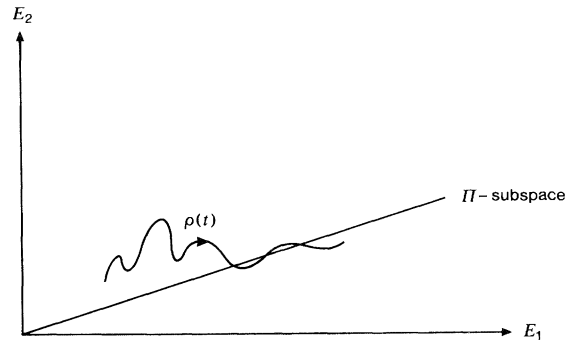

Figure 2. The idea of subdynamics.

explicit demonstrations that $\Pi^2 = \Pi$ and $L\Pi = \Pi L$. But these results must be equivalent to those obtained in the previous section, equation (7.9): the replacement of the lower limit of integration with $-\infty$ corresponds to the discarding of the more rapidly decaying modes. The two developments are of course related merely by Laplace transformation (after noting that in the present section we have restricted the work to the case of a time-independent liouvillian).

The previous paragraph does not do justice to the sophistication of the material summarized, but to the non-specialist this work can be conceived to be essentially an elaborate justification of the use of the hypothesis of rapid decay, and hence the use of (7.9). The decay occurs, at the selected level of description M, since the poles of the partial resolvent are indeed such as to imply decay, and the timescale appropriate for that level can be estimated from the smallest value of $-\mathcal{R}(p)$ that occurs.

Although subdynamics has generated a large literature, it remains the case that little is known about the mathematical conditions required for its validity (if indeed it is ever valid!). Early investigations in the quantum case were reported by Lanz *et al.* (1971), and references contained there. A recent paper of Coveney & Penrose (1992) offers another line of investigation.

9. Comments on subdynamics

The first question that will naturally be asked is that of the origin of irreversibility, given that the original hamiltonian equations were reversible. The mathematical origin lies in the use of the Laplace transform for an initial-value problem, together with the treatment of $\tilde{U}_{11}(p)$ which is defined initially in $\mathcal{R}(p) > 0$ and is continued analytically to $\mathcal{R}(p) < 0$ to search for its poles. The resulting deformations in contours lead to the asymmetric handling of time, and can also be mimicked by the insertion of convergence factors of the form $\exp(\pm\epsilon)$ or $\exp(\pm i\epsilon)$ as is common in other fields like scattering theory, and as we do here in § 13. One can see this working in a simple context in the derivation of the retarded potential by the use of transform methods, leading to the choice of the 'causal' solution. A more subtle example which makes a closer comparison with subdynamics is that of Landau damping in plasma physics. In that problem, an observer who can only detect the electric field of the wave inevitably sees it decay, although if

he could observe the details of the electron distribution, he would see a continuous spectrum of motions that do not decay. See Crawford & Hislop (1989) for a development of this in the context of spectral deformation theory.

The eventual decay of that part of ρ_1, with its accompanying ρ_2, that denotes departure from equilibrium, is reflected in a similar approach to equilibrium on the part of $a(t)$, suggesting the the equations of motion for $a(t)$ derived by this means will comply with the third condition for a well-posed problem, as we wanted in § 4.

The implementation of the theory in the form of practical calculations is found, not surprisingly, to be very laborious. Here, the text of Balescu (1975), which contains very extensive developments, should be consulted. His approach is, however, entirely perturbation-theoretic, with the unperturbed case being that of zero interaction potential, thus following the Mayer expansion of equilibrium statistical mechanics. The projection P is that which reduces the description to the one-particle distribution. The calculations are carried out with widespread use of diagram techniques, thereby making another agreeable historical connection with Mayer, who originated those techniques.

10. Entropy in NESM

So far, we have not mentioned *entropy* in our explanation of subdynamics, and indeed any practical calculations could proceed without consideration of entropy. But it is natural to ask whether some variable or functional can be identified as 'entropy'. One would expect it to coincide with the ordinary meaning of entropy in equilibrium statistical mechanics and thermodynamics, to be additive in the case of disjoint systems, and of course to comply with the *law of increasing entropy*, or second law of thermodynamics. This is a subject liable to the utmost confusion, as writers use the word entropy to mean different things (information entropy, Gibbs entropy, Boltzmann entropy, Kolmogorov entropy, ...). The first two conditions (equilibrium value and additivity) point strongly to the use of the Gibbs entropy

$$S_G[\rho] = -\int_\Gamma \rho \ln \rho \, \mathrm{d}\Omega \tag{10.1}$$

(we set Boltzmann's constant equal to 1 for convenience). This fails to give the third condition, for as is well known, S as given by (10.1) is invariant. In fact, any functional of the form

$$S[\rho] = \int_\Gamma h(\rho) \, \mathrm{d}\Omega \tag{10.2}$$

is invariant, as an immediate consequence of the measure-preserving property of the flow, h being any measurable function.

The decomposition of phase space offers an escape from this difficulty, as we could define S so as to involve ρ_1 and ρ_2 in different ways, so it ceases to be a functional expressible in the form (10.2), and we may be able to comply with all three of the above conditions. The physicist should be satisfied if $S(t)$ is given as a functional of $a(t)$. Such an entropy could perhaps be called an 'observed entropy', though I hesitate to contribute to a yet further confusion of nomenclature! If we accept this approach, observed entropy for a non-equilibrium state cannot be computed from a knowledge of the hamiltonian together with the distribution

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

function ρ. Its computation requires a choice of the level of observation, M, and consideration of the projection P or more generally $\mathcal{P}$. But within an actual application of these methods, it should be computable (in principle) to any desired degree of accuracy for any non-equilibrium state.

Unfortunately, the words 'objective' and 'subjective' have been used to contrast an expression such as (10.2) and an 'observed entropy' as described here. This has led to a somewhat acrimonious but confusing debate, which has been helpfully clarified by Denbigh & Denbigh (1985). My own view is that it is inevitable that a concept along the lines of the previous paragraph is required for non-equilibrium entropy. At one extreme, that of a complete microscopic description, entropy disappears (along with irreversibility), while at the other extreme, that of equilibrium states, where the invariants of the motion comprise the only data, S_G is appropriate but constant. Between these extremes, i.e. in the domain of NESM, all theory is constructed in a way that inherently involves the level of description. As expressions like (10.2) do not work, we need not be surprised that the construction of entropy similarly involves the level of description. This is reinforced by recalling that the introduction of irreversibility itself required considerable technical subtlety and involved the level of description in a detailed way.

11. Further developments of subdynamics

The idea promoted by the Brussels school for the definition of non-equilibrium entropy is to introduce an invertible operator $\Lambda : E \to E$ satisfying

$$\Pi = \Lambda P \Lambda^{-1} \tag{11.1}$$

exhibiting the similarity relation between P and Π. Then a new representative $\rho_{\mathrm{p}}(t) \in E$ (the 'physical representation') is defined by

$$\rho_{\mathrm{p}} = \Lambda^{-1}\rho. \tag{11.2}$$

Finally, it is claimed that

$$\int_{\Gamma} |\rho_{\mathrm{p}}|^2 \, \mathrm{d}\Omega \tag{11.3}$$

is a decreasing function, and serves as a Lyapunov functional for the evolution of the system; its negative could therefore be used to define an entropy. For more details, Prigogine *et al.* (1977) should be consulted. Although this may indeed qualify as an 'observed entropy' as described above, it has not been widely accepted; it should be noted that this procedure is not included in Balescu (1975). More generally, (11.3) could be replaced by any concave functional, so such an entropy is not unique.

In further developments, members of the Brussels school, with others (Misra 1978; Misra *et al.* 1979) have sought to detach the consideration of the Lyapunov functional from the choice of the level of observation. The more abstract question is asked: given a liouvillian L is it possible to find a Λ such that if ρ_{p} is defined as in (11.2), then the resulting evolution equation for ρ_{p} has (11.3) as a Lyapunov functional. Still more generally, is it possible to find such a transformation with the result that ρ_{p} evolves in the manner of a Markov process, i.e. forms a Markov semi-group. Necessary and sufficient conditions for this to happen are derived,

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

which turn out to be similar to what we designated as level (e) in § 5 above. These investigations are of interest as they ask about properties of L alone under which NESM can succeed. However, the claim that such a theory is superior, for not discussing the discarding of information, seems to me to be misplaced: the level of description is an essential attribute to any calculation of NESM. There is also the point that the transformation to the Markov semi-group cannot be what is required in every case as it does not encompass the behaviour of some materials whose evolution depends on previous history, while the master equation can do so.

These considerations have also led to the introduction of what has been called an 'internal time operator'. We offer more comments on this in § 20 below.

12. The quasi-canonical distribution

We now turn to the investigation of the Zubarev and Maxent approaches to NESM. In this section we put forward a definition which will be used extensively in the sequel. Its purpose is to provide a choice of ρ to serve as the 'opening' $\rho_1 \in \mathcal{E}_1$ introduced in § 6. We have to satisfy the normalization

$$\int_\Gamma \rho \, \mathrm{d}\Omega = 1 \tag{12.1}$$

and, given the map $M : E \to A$ and a vector-valued $a \in A$, to satisfy

$$M\rho = a. \tag{12.2}$$

Suppose now that a functional of the form (10.2) is specified by

$$S[\rho] = \int_\Gamma h(\rho) \, \mathrm{d}\Omega, \tag{12.3}$$

where now h is twice differentiable, and the rule for selecting the preferred ρ is that it should maximize (12.3) subject to (12.1) and (12.2). We will take it (cf. the discussion in § 3) that (12.2) is in fact of the form of an integral over Γ with a suitable kernel (also written M); then a straightforward application of Lagrange's undetermined multipliers gives

$$h'(\rho) = c_0 + \lambda.M, \tag{12.4}$$

where $c_0 \in \mathbb{R}$ is the multiplier for (12.1), and λ is the multiplier for (12.2), so technically λ is in the space dual to A. The inner product in (12.4) denotes a sum if A is finite dimensional, otherwise an integral over the variables involved in the function space A, leaving $X \in \Gamma$ as the free variable of (12.4). Provided $h'(\rho)$ is monotonic, so $h(\rho)$ itself is *concave* (it will need to be bounded above, $h'(\rho) < 0$), (12.4) will possess a unique solution $\rho = \rho_1$ for each c_0, λ. We must also satisfy (12.1) and (12.2), which are linear. When this has been done, we shall have a unique set $\{c_0, \lambda, \rho_1\}$ for each specified macroscopic state a. This means we shall be in the position envisaged in § 6, having constructed the manifold $\mathcal{E}_1$ containing the values of ρ_1, and with it the idempotent map $\mathcal{P}$.

An elementary case of this procedure results from the choice $h(\rho) = -\frac{1}{2}\rho^2$, so (12.4) gives $\rho = -c_0 - \lambda.M$, whereupon the whole construction becomes linear. We are then back at the start of subdynamics. Similarly, given any $F(X) > 0$, one could use $h(\rho) = -\frac{1}{2}F\rho^2$. Actually, our procedure can be generalized still

further, as there is no reason why $S[\rho]$ need be restricted to being algebraic in ρ: it could also include derivatives. Provided S is quadratic in ρ, the procedure will be linear. We conjecture that a suitable choice might provide a link with the theory of § 11, but that will not be pursued here.

We now focus attention on the important case where $S = S_G$, so $h(\rho) = -\rho \ln \rho$. The resulting ρ will be called the *quasi-canonical* distribution ρ_q, and $\mathcal{E}_1$ is the *Maxent manifold* $\mathcal{E}_q$. With this $h(\rho)$ we have $h'(\rho) = -\ln \rho - 1$ so (12.4) becomes

$$\ln \rho_q = -\lambda_n - \lambda.M, \tag{12.5}$$

where $\lambda_n = 1 + c_0$ now takes care of the normalization, and our quasi-canonical distribution has the form

$$\rho_q = \exp(-\lambda_n - \lambda.M) \tag{12.6}$$

subject to (12.1) and (12.2).

The normalization can be handled, and λ_n eliminated, by the same means as in the equilibrium case, defining a *partition functional*,

$$Z[\lambda] = \int_\Gamma \exp(-\lambda.M)\, \mathrm{d}\Omega, \tag{12.7}$$

whereby

$$\rho_q = Z^{-1} \exp(-\lambda.M) \tag{12.8}$$

now subject only to (12.2). There are also formal expressions for a in terms of functional derivatives of $\ln Z$ with respect to λ, which are analogous to those of equilibrium statistical mechanics. The value of S that results from (12.6) is

$$S[\rho_q] = \int_\Gamma (\lambda_n + \lambda.M) \exp(-\lambda_n - \lambda.M)\, \mathrm{d}\Omega, \tag{12.9}$$

for which there are also alternative expressions.

We now have a map $a \mapsto \rho_q \in \mathcal{E}_q$, although defined in an awkward, implicit, way.

Before proceeding with our main development, we digress briefly to consider ρ_q itself. In general $\mathcal{P}_q$ does not commute with L, so time evolution cannot be expressed simply in terms of quasi-canonical distributions. Given the observational value of a at $t = 0$, one could start with the corresponding ρ_q and let $\rho(t)$ evolve thereafter by Liouville's equation. This is somewhat uncritically adopted by Robin (1990), although I believe that subsequent perturbation expansions lead eventually to acceptable results in his chosen context. More generally, one wants to avoid giving a privileged role to $t = 0$.

Another possibility was put forward by Lewis (1967), and I suspect by several writers (e.g. Richardson 1960; Levine 1979; Duering *et al.* 1985), it being superficially attractive. We remark that, given $\rho(t)$, Liouville's equation gives $\dot{\rho}$, from which

$$\dot{a} = M\dot{\rho} = ML\rho. \tag{12.10}$$

By asserting that $\rho(t)$ should at each instant be the ρ_q corresponding to $a(t)$ there results a universal algorithm for NESM since we have $\dot{a}$ in terms of a. However, this is simply incorrect: the resulting macroscopic equations are reversible, and the underlying $\rho_q(t)$ does not satisfy Liouville's equation. Nevertheless the results from this procedure are used in practice: they can be identified as the

adiabatic evolution through states of local equilibrium. In the case of continuum mechanics for example it gives an 'ideal' fluid with zero thermal conductivity and viscosity; in the kinetic theory of gases it compels ρ to factorize into a product of one-particle functions (see Mayer 1960) and leads to the Vlasov equation of plasma physics. Moreover, not surprisingly, the entropy $S[\rho_{\mathrm{q}}]$ for such an evolution remains constant (Dougherty 1972). This can easily be seen as follows. Let $\dot{\rho}_{\mathrm{q}}$ be the rate of change that results from the above construction, while $\dot{\rho}_L$ is the rate that would be given by Liouville's equation. Then

$$\int M\,(\dot{\rho}_{\mathrm{q}} - \dot{\rho}_L)\,\mathrm{d}\Omega = 0$$

by construction (we have reinstated the integral sign previously implied by our notation) and

$$\int (\dot{\rho}_{\mathrm{q}} - \dot{\rho}_L)\,\mathrm{d}\Omega = 0$$

by the normalization. But, recalling the variational argument leading to (12.4),

$$\begin{aligned}\dot{S}\,[\rho_{\mathrm{q}}] - \dot{S}\,[\rho_L] &= \int h'\,(\rho_{\mathrm{q}})\,(\dot{\rho}_{\mathrm{q}} - \dot{\rho}_L)\,\mathrm{d}\Omega \\ &= \int (c_0 + \lambda.\,M)\,(\dot{\rho}_{\mathrm{q}} - \dot{\rho}_L)\,\mathrm{d}\Omega = 0.\end{aligned}$$

As $\dot{S}\,[\rho_L] = 0$ (as noted in § 9),

$$\dot{S}\,[\rho_{\mathrm{q}}] = 0 \tag{12.11}$$

also. The above idea can be conceived in terms of discrete time-steps, say Δt; during a step ρ obeys Liouville's equation (preserving S but migrating away from $\mathcal{E}_{\mathrm{q}}$) and at the end of the step an application of $\mathcal{P}_{\mathrm{q}}$ returns it to $\mathcal{E}_{\mathrm{q}}$, with an increase ΔS in S. The reason why it fails to give the law of increasing entropy is that in the limit $\Delta t \to 0$, ΔS is $O\,(\Delta t^2)$. To meet this, Lewis proposed that Δt remain finite. But that is hardly acceptable as a piece of general theory, as it requires separate investigations, guided by physical intuition, to select the appropriate Δt (the relaxation time); for example, Boltzmann did this, apparently subconsciously, in regarding collisions as instantaneous.

Suppose now that at equation (12.3) we had made a different choice for the function $h(\rho)$, provided that h is bounded above and concave. The reader will easily check that everything we have developed in this section in respect of S_G would have a natural analogue; it would be necessary to replace the exponential function in (12.6) by the inverse of the function h'. What, then, is the particular merit in taking S to be S_G? To the physicist this choice has the obvious attraction that ρ_{q} is in some sense as close as possible to thermodynamic equilibrium subject to (12.2); similarly one might claim that it is 'locally' in equilibrium, as in the Chapman–Enskog method. It may be reasonable to expect that (by molecular processes such as collisions) the non-equilibrium aspects of ρ have relaxed rapidly to bring ρ close to ρ_{q}, as in Bogoliubov's discussion. Alternatively, one can invoke information theory, as we will describe later. However, for the present no particular justification for the choice $\rho_1 = \rho_{\mathrm{q}}$ is needed if it is to be used in procedures like that of § 7. The basis for choice would merely be that it performed

better, for example gave better approximations or convergence, or was easier to handle (though if so there might be helpful explanations of that fact).

13. Time evolution in Maxent

Here we derive our version of Zubarev's results using the notation and methodology that we have developed so far; subsequently we will outline the connection with Zubarev's presentation.

The previous discussion suggests that we should write

$$\rho = \rho_{\mathrm{q}} + \rho_2, \qquad \rho_{\mathrm{q}} \in \mathcal{E}_{\mathrm{q}}, \quad \rho_2 \in E_2 \tag{13.1}$$

and proceed with a calculation parallel to that of § 7 to derive another version of the master equation. That might be possible in principle, but when it is attempted we lose the linear mathematics of § 7, as $\mathcal{E}_{\mathrm{q}}$ is a curved manifold. It is at this stage that we need to amend the procedure outlined in § 6. We make a nonlinear transformation to a new representation:

$$\eta(X,t) = -\ln \rho(X,t). \tag{13.2}$$

It is well known that if ρ satisfies Liouville's equation, so does any $g(\rho)$, where the real function g is monotonic and differentiable. So, in particular,

$$\partial\eta/\partial t = L\eta. \tag{13.3}$$

Thus any spectral characterization that may be available in respect of ρ applies equally to η. By (12.5) we can write, in obvious notation,

$$\eta_{\mathrm{q}} = -\ln \rho_{\mathrm{q}} = \lambda_n + \lambda.M, \tag{13.4}$$

which shows that the class of 'preferred' distributions in E, namely $\mathcal{E}_{\mathrm{q}}$, is conveniently mapped into a vector subspace in the new representation, with $\{\lambda_n, \lambda\}$ as a parametrization. This may appear to be awkward, but it is analogous to the use of temperature in equilibrium statistical mechanics.

To proceed with the general solution of Liouville's equation we write

$$\eta(X,t) = \eta_{\mathrm{q}} + \zeta = \lambda_n + \lambda.M + \zeta, \tag{13.5}$$

where $\zeta(X,t)$ is a new phase function. This split $\eta = \eta_{\mathrm{q}} + \zeta$ is analogous to the split $\rho = \rho_1 + \rho_2$ used deriving the master equation in § 7. The aim in the remainder of this section is to follow that derivation in trying to eliminate $\zeta(0)$ instead of $\rho_2(0)$, but it is no longer the case that ζ corresponds in some way to the kernel, E_2.

It may be well to clarify the nature of the terms occurring in (13.5). Here, $\lambda_n(t)$ is a number, depending only on t; $\lambda(t)$ is a vector in A', the dual of the space of observable functions, also depending on t; $M(X)$ is a kernel in $A \otimes E$ which is independent of t, the dependence on A being suppressed in our notation but is what is involved in the 'dot product', while ζ is a function in E that depends on time. We have the Green's function $U(t) = \exp(Lt)$ and note that $U(t)c = c$ if c is a number (i.e. independent of X). We now define a *time-dependent kernel* analogous to the Heisenberg representation of quantum mechanics (although we use no quantum physics),

$$M(X,t) = U(-t)M(X), \tag{13.6}$$

satisfying

$$\frac{\partial}{\partial t}M(X,t) = -LM(X,t) \quad \text{with} \quad M(X,0) = M(X). \tag{13.7}$$

Suppose temporarily that we have to solve Liouville's equation for ζ in an interval (t_0, t) in circumstances where λ_n and λ have *prescribed* values throughout the interval. This postpones to a later stage the question of how those values were acquired. To solve (13.3) requires that we solve

$$\frac{\partial \zeta}{\partial t} - L\zeta = -\frac{\partial}{\partial t}\left(\lambda_n + \lambda.M\right) + L\left(\lambda_n + \lambda.M\right), \tag{13.8}$$

of which the right-hand side is a known inhomogeneous term. But there is no need to continue along those lines as the answer can be written down immediately from

$$\lambda_n(t) + \lambda(t).M + \zeta(t) = \exp\left\{L\left(t - t_0\right)\right\}\left\{\lambda_n\left(t_0\right) + \lambda\left(t_0\right).M + \zeta\left(t_0\right)\right\}$$

and it is readily checked that it can be rewritten as

$$\zeta(t) = \mathrm{e}^{L(t-t_0)}\zeta\left(t_0\right) - \mathrm{e}^{Lt}\left[\lambda_n\left(t'\right) + \lambda\left(t'\right).M\left(X,t'\right)\right]_{t_0}^{t} \tag{13.9}$$

(noting (13.6)).

We now ask whether it is possible to develop the theory along the lines of § 7, whereby t_0 can be pushed arbitrarily into the past, and the dependence on $\zeta\left(t_0\right)$ removed; as we saw there that is not in general possible but may become possible if L satisfies a condition such as having a continuous spectrum (in the space orthogonal to constants of the motion). If this is indeed so, the results can be obtained by a device that ensures causal behaviour; in § 7 this consisted of using the Laplace transform. Here we can do it by inserting a small extra term in the left-hand side of (13.8), so that it becomes

$$\partial\zeta/\partial t - (L - \epsilon)\zeta = \ldots, \tag{13.10}$$

where $\epsilon > 0$ is small and gives a slight decay to the homogeneous term; the Green's function for the left-hand side becomes $\exp\{(L-\epsilon)t\}$. No such alteration is made to the right-hand side of (60), but (as is again easily checked) we can rewrite it in a new form

$$-\mathrm{e}^{Lt}\frac{\partial}{\partial t}\left\{\lambda_n(t) + \lambda(t).M(X,t)\right\}. \tag{13.11}$$

If (13.11) is temporarily called $F(X,t)$, the solution to the modified equation (13.8), for a given ϵ, and over a given interval (t_0, t) can be written down using the Green's function, and is

$$\zeta_\epsilon(t) = \mathrm{e}^{(L-\epsilon)(t-t_0)}\zeta\left(t_0\right) + \int_{t_0}^{t} \mathrm{e}^{(L-\epsilon)\left(t-t'\right)}F\left(X,t'\right)\,\mathrm{d}t'.$$

By further easy manipulations this becomes

$$\zeta_\epsilon(t) = \mathrm{e}^{(L-\epsilon)(t-t_0)}\zeta\left(t_0\right) - \mathrm{e}^{(L-\epsilon)t}\int_{t_0}^{t} \mathrm{e}^{\epsilon t'}\frac{\partial}{\partial t'}\left\{\lambda_n\left(t'\right) + \lambda\left(t'\right).M\left(X,t'\right)\right\}\,\mathrm{d}t'.$$

So far this is without approximation or assumption of any kind. (If now $\epsilon \to 0$,

(13.9) is recovered.) Subject to what has already been discussed, we now let $t_0 \to -\infty$, keeping $\epsilon > 0$, with the result that the term in $\zeta(t_0)$ disappears, and we reach the causal solution:

$$\zeta_\epsilon(t) = -\mathrm{e}^{(L-\epsilon)t} \int_{-\infty}^{t} \mathrm{e}^{\epsilon t'} \frac{\partial}{\partial t'} \{\lambda_n(t') + \lambda(t').M(X,t')\}\, \mathrm{d}t'.$$

Writing $t' = t + \tau$ and shifting the time argument of the Heisenberg kernel so as to absorb the factor in front of the integral

$$\zeta_\epsilon(t) = -\int_{-\infty}^{0} \mathrm{e}^{\epsilon\tau} \frac{\partial}{\partial \tau} \{\lambda_n(t+\tau) + \lambda(t+\tau).M(X,\tau)\}\, \mathrm{d}\tau. \tag{13.12}$$

This is the desired result. It has the form of a Laplace transform with respect to negative time, $-\tau$, with ϵ as the transform variable. The limit $\epsilon \to 0$ is a delicate one as the integrand does not $\to 0$ for $\tau \to -\infty$ so the integral diverges, although it can be redefined by the theory of summability, and the limit treated as a Tauberian one (Widder 1946, p. 195).

If we carry out the obvious integration by parts in (13.12), which is merely a standard step in Laplace transform theory (transform of a derivative), there results

$$\zeta_\epsilon(t) = -\{\lambda_n(t) + \lambda(t).M\} + \epsilon \int_{-\infty}^{0} \mathrm{e}^{\epsilon\tau} \{\lambda_n(t+\tau) + \lambda(t+\tau).M(X,\tau)\}\, \mathrm{d}\tau.$$

The first part of this, which is independent of ϵ, is merely the negative of η_q, the 'assumed' contribution to η in (13.5), so when our result is substituted into (13.5), it cancels, leaving an integral over past time for η itself:

$$\eta(X,t) = \epsilon \int_{-\infty}^{0} \mathrm{e}^{\epsilon\tau} \{\lambda_n(t+\tau) + \lambda(t+\tau).M(X,\tau)\}\, \mathrm{d}\tau. \tag{13.13}$$

This remarkable expression needs further discussion, most of which we postpone. For the present we must emphasize that although (13.13) appears to be a formula for a solution to the Liouville equation, it is not an explicit solution on account of the presence of the operator $M(X,\tau)$, which is merely an abbreviation for the process of solution. What (13.13) actually does is to organize the *selection* of solutions of Liouville's equation appropriate for the level of description M. Each such choice is prescribed by a *history* $\{\lambda_n(t), \lambda(t)\}$ and implies (in principle) a resulting history $\{a(t)\}$. The theory of this section has *not* shown how to extract from this any explicit procedure for predicting $a(t)$ given $a(0)$. On reviewing the preceding (and somewhat clumsy-looking) derivation, it will be confirmed that it actually follows a similar line to the derivation, in § 7, of the master equation. In each case, unwanted details of ρ are eliminated by supposing that their influence decays when the initial instant is pushed into the remote past. In each case the formidable details of solving Liouville's equation are concealed by formal operators (ψ in § 7, M in the present section). The presentations differ because of the nonlinear transformation from ρ to η, but one can only conclude that the derivations either both succeed or both fail, and that where thay succeed they

reach the same result, i.e. the same selection of solutions, although differently expressed.

14. Zubarev's derivation

Equation (13.13) above, was given by Zubarev (1961). The resulting Liouville function $\rho = \exp(-\eta)$ is called by him the *non-equilibrium statistical operator*, the word 'operator' being used since that is the nature of ρ in Zubarev's quantum-mechanical context. We now outline his approach, making reference to his book (Zubarev 1974). He uses the name 'quasi-equilibrium' for what we have called 'quasi-canonical' distributions. His § 21 contains the basic material; besides treating the underlying physics as quantum mechanical he selects the desired description (M in our notation) to be specifically that of fluid dynamics, although commenting that other choices could be adopted. The method for deriving (13.13) is one of tentative construction and verification, and appears rather abstruse. The motivation begins with a consideration of the idea of the hierarchy of relaxation times, following Bogoliubov. From this emerges the concept of 'local integrals of the motion', which are objects whose time evolution occurs on a scale far slower than molecular evolution. As the distributions of equilibrium statistical mechanics arise in the form of functions of invariants of the motion, it is natural to attempt to construct distributions for NESM from these local integrals; this line of argument leads eventually to (13.13). A further technical point is that he omits the Lagrange multiplier for normalization, λ_n, but subsumes the process of normalization in a quotient, Q, which corresponds to the partition function of equilibrium theory. The basic result, equivalent to our (13.13), is reached at his equation (21.10a). It has to be verified by substitution that the resulting ρ satisfies Liouville's equation in the limit $\epsilon \to 0$.

The introduction of the factor ϵ attracts considerable discussion. The action of the terms involving ϵ in (13.13) is referred to as 'taking the invariant part' of the integrand, meaning the smoothing out of oscillating terms (denoted with a wavy line over symbols). This is traced to treatments of the scattering theory of Schrödinger's equation, in particular the work of Gell-Mann & Goldberger (1953). The prescription is that the thermodynamic limit ($V \to \infty$, where V is the volume) should be taken before letting $\epsilon \to 0$. This has the effect of selecting retarded solutions (or outgoing waves) or alternatively of removing the effect of reflections from boundaries. We recall that finite quantum systems cannot achieve a level of stochasticity higher than that of ergodic, but they can do so in the thermodynamic limit (although unfortunately that introduces other technical difficulties). It is only in that limit that a continuous spectrum is possible. We discuss some of these matters further in § 23. It should also be noted that, in taking the limit $\epsilon \to 0$ in (13.13), the parameters λ_n and λ are functions of ϵ.

Another possible attitude to the replacement of L by $L - \epsilon$ is referred to by Zubarev and may be enlarged upon here. Such a replacement in the full Liouville's equation would mean that the map is no longer measure-preserving, but its use in respect of the term in ζ, as in (13.10), still leaves the full equation measure-preserving since the necessary modification in λ_n ensures that automatically. When one uses the 'ensemble' concept of a large number of copies of the physical system, the term $\epsilon\zeta$ implies a mechanism for the destruction of some of the copies, while the adjustments due to λ_n imply a restoration by new copies in

different states. Physically this could be interpreted in terms of random external influences such as reservoirs, in other words an appeal to 'open systems'. This form of modification is curiously reminiscent of the so-called 'BGK' term used as a simple model for the Boltzmann equation (see, for instance, the third edition (1970) of Chapman & Cowling's book—it does not appear in earlier editions). In his Appendix II, Zubarev discusses the connection between his approach and that of McLennan, who developed methods involving open systems (see McLennan (1963) for a summary).

Elsewhere in the book, Zubarev presents a number of applications of his method to transport processes, correlations, relaxation processes and so on.

15. Entropy maximization

To set the scene, we outline the general arguments leading to the adoption of maximum entropy prescriptions for assigning probabilities in science. This summary does not pretend to be a substitute for the expositions of Jaynes and others. We make reference to the reprint collection Jaynes (1983).

Where complete description or observation is impracticable, one must be content with reduced descriptions, and when these are formalized as means, correlations and the like, the concept of probability seems naturally and inevitably to be imposed on the scientist. Once a probability distribution has been supplied, the extraction of further useful statements is a mathematical activity requiring expertise, but not one involving conceptual difficulty. The assignment of probabilities in the first place is a more contentious matter (Fine 1973; Gillies 1973).

The Maxent approach starts from the assertion that when a probability distribution $\{p_i\}$ is given, $(1 \leqslant i \leqslant k)$, it conveys *information* to the recipient. According to Shannon's theory, this can be measured quantitatively (in bits) by

$$\sum_{i=1}^{k} p_i \log_2 p_i.$$

The remaining uncertainty, or lack of information, can thus be measured, in suitable units, and with an arbitrary additive constant, by the *information entropy*:

$$S_I = -\sum_i p_i \ln p_i \tag{15.1}$$

(we revert to natural logarithms only for convenience).

The underlying philosophy is that if $\{p_i\}$ is unknown, and one is asked to make some prediction, then one should select $\{p_i\}$ to *maximize* S_I subject to $\sum p_i = 1$ and subject to such further (but partial) information as may be available. The basis for this, stated simply, is that to do otherwise would be to claim a higher level of information than had been provided. The above outlines the earlier version of the idea, explained at length in Jaynes's Brandeis lectures of 1963 (collected papers p. 39ff; see also pp. 233–237). The same conclusions, that S_I should be maximized can alternatively be reached, on the basis of very mild axioms, as a general principle of inference (Shore & Johnson 1980), and this approach is now the preferred one.

To accept these arguments means adopting an interpretation of 'probability' that departs from the more commonly accepted one of probabilities resulting

from numerous repetitions of an experiment. Neither mathematics nor theoretical physics has the resources to adjudicate between these possible approaches to this question, and our attitude in this article, as part of our comparative study, is to investigate the consequences of the hypothesis without appearing to endorse it.

If no other information is available, one has $p_i = k^{-1}$ for each i, a uniform distribution indicating complete ignorance; this can also be identified with what is called the 'prior' probability distribution. If now further information, e.g. by observations, becomes available, it is taken to mean the supply of the *expectation values* of random variables f_j, $j = 1, 2, \ldots, \ell$ where $\ell < k$. It is well known that the set $\{p_i\}$ that maximizes S_I subject to the additional constraints,

$$\sum_i f_j(i) p_i = \langle f_j \rangle = a_j \quad \text{(say)}, \tag{15.2}$$

has the form

$$\log p_i = -\lambda_n - \sum_j \lambda_j f_j(i), \tag{15.3}$$

where $\{\lambda_j\}$ are Lagrange multipliers. For a more careful investigation of this, and the expressions in terms of a partition function, we refer to Jaynes, pp. 45–52.

It is worth noting that the above formalism has been applied in other areas, an important example being image-processing. Here, an imperfect image can be 'sharpened up' by Maxent procedures. Work along these lines has progressed rapidly since about 1980, in parallel with advances in computer science and numerical analysis, which are required to handle the massive sets of equations that have to be solved. This has been of great benefit to astronomy and medical science (to mention only two of many fields). These applications are so impressive as to leave no doubt that this approach to probability must be accepted and welcomed in appropriate circumstances. That does not, however, excuse us from considering critically its application to statistical mechanics.

There is an initial difficulty arising from the fact that, in classical mechanics, the probability distribution $\rho(X)$ is a function of continuous variables, rather than a label i as above. On making the obvious adaptation of the theory one notices that the analogue of the uniform prior is not unique; a change from X to a new representation of phase space could imply that the new prior is not uniform, contradicting what would have been deduced had the new variables been in use from the start. This is indeed a worry in other contexts, as illustrated by Bertrand's paradox (Jaynes, p. 134ff). In statistical mechanics, it seems to be universally accepted that one is to start in cartesian coordinates, with the corresponding momenta. Then, one restricts changes of representation to canonical transformations (possibly involving the time, as in moving frames). As is well known, canonical transformations are measure-preserving (like time evolution). One then has a consistent scheme, but it could be objected that other possibilities have not been eliminated. One further point that can be offered here is that in the quantum mechanics of a spatially finite system, the energy eigenstates supply a natural discrete set on which probabilities $\{p_i\}$ are to be assigned—a feature often used in textbooks to argue that quantum statistical mechanics (in equilibrium) is more readily understood than classical. We then invoke the correspondence principle: when the classical limit is considered, asymptotic results

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

about the density of the eigenstate spectrum lead inevitably to the adoption of the uniform measure in classical phase, as above.

If thermodynamic equilibrium is defined to be a time-invariant macroscopic state which is completely specified by the values of invariants of the motion, then the above procedure of Maxent is identical with the familiar derivation of standard results. Usually the energy is the only such constant, as the effects of boundaries prevent the appearance of others, e.g. angular momentum, which the physicist might look for. If others do exist, the family of equilibrium solutions is different from what it is when the energy is the only one, although the latter is a subset of the former. It is important to emphasize that the Maxent formalism cannot answer the question 'what invariants are there?', though it can supply ρ to the physicist who has observed values of some or all of those invariants that exist. In my view, the question of what they are is a legitimate part of the subject, which is in principle to be answered by ergodic theory even though one admits that very little progress has been made (see Jaynes, p. 232, on this point).

It should also be noted that prescribing the value of the energy can be interpreted in two ways. In the first, the value is regarded as exact, and a part of the description of the system. In that case, probability zero is ascribed to those parts of Γ with incorrect values of the energy, and a non-vanishing measure is attached to Σ, the energy hypersurface (see § 4). After including the necessary jacobian factor, Maxent then leads to a uniform distribution on Σ, i.e. the microcanonical distribution. The second interpretation assumes that the given value of energy is an expectation value, in which case Maxent leads to the canonical distribution. In practice, for large systems, the two routes lead to equivalent results in nearly all predictions (an obvious exception arises if one asks about the variance of the energy!). We mention this to alert the reader to the idea that the Maxent philosophy is less clear-cut than it is sometimes made to appear. Of course, similar remarks apply where there are further invariants besides the energy.

16. Jaynes's derivation of NESM

We turn now to NESM. This may be defined as an activity in which some (at least) of the data supplied consists of values of observables which are *not* invariants. Moreover, the data may refer to various times or ranges of times; it may consist of different observables at different times. In short, it might be anything! Similarly, the predictions asked for could consist of the values of any observables, not necessarily the same ones as occur in the data, and at any times.

In this very general context, Jaynes has indicated how the Maxent principles and procedures should be formulated and applied. The material appears in several places in the collected papers (Jaynes 1983), but the most extended exposition occurs in Section D of a Conference paper of 1979 (see p. 287ff). This work also appears in Grandy (1988). The following is an attempt to present parts of it in the style and notation already used here, together with my own comments.

Data are available at times t_1, t_2, ..., and maximum entropy is supposed to be used to establish a Liouville function $\rho(t)$ which is genuinely time-dependent, and satisfies Liouville's equation. It will be sufficient to obtain ρ at a particular time as (in principle) Liouville's equation then enables it to be computed at other times. Without loss of generality, the particular time is designated as $t = 0$. Accepting the discussions of § 15, the problem is to maximize $S[\rho]$ at $t = 0$

subject to constraints imposed at other times t_1, t_2, ... (together with the usual normalization). The Heisenberg operator defined in (13.6) is a convenient tool. Suppose firstly that there is a single observable, given by a kernel $m_1(X)$, whose value at time t_1 is given as a_1, so, in our previous notation,

$$m_1(X)\rho(X,t_1) = a_1. \tag{16.1}$$

When this is converted to the corresponding implication at $t = 0$ we have

$$m_1(X,t_1)\rho(0) = a_1, \tag{16.2}$$

where $m_1(X,t_1)$ is the Heisenberg operator. More generally, there will be a family of such conditions for observables m_1, m_2, ... at times t_1, t_2, ...; the resulting formal procedure for maximizing $S[\rho]$ at $t = 0$ introduces Lagrange multipliers λ_1, λ_2, ..., together with λ_n for the normalization, in a way that is readily understood though likely to be very awkward to implement. Following the usual method (as in the steps from (12.1) to (12.6) above) it results in

$$\ln \rho(X,0) = -\lambda_n - \sum_k \lambda_k m_k\,(X,t_k)\,. \tag{16.3}$$

A further extension is also straightforward: the set of 'information-gathering' times $\{t_k\}$ might become a continuum, with the result that summation over k would be replaced with integration, and λ_k would become a function, $\lambda(t')$, say.

Suppose now that the information-gathering times are finite, i.e. do not extend to $t = \pm\infty$. Then the above formulation is undoubtedly the correct implementation of the philosophical principles contained in Maxent; indeed it could be claimed to include the most general questions that could be asked. If so, the class of questions considered in the present article, namely the evolution $a(t)$ of a specific macroscopic description, starting from $a(0)$, could be expected to emerge as an application. Unfortunately, it does not do so without further amendment.

The procedure leading to (16.2) is translationally invariant with respect to time, notwithstanding the particular use made of $t = 0$; it gives a $\rho(t)$ which would maximize $S[\rho]$ at any time subject to the given data (noting that it satisfies Liouville's equation, so S is constant). Some of that data could therefore be in the future, i.e. some of $\{t_k\}$ may be positive. In using the word 'prediction' it would be natural to arrange that $t_k \leqslant 0$, so that evolution in $t > 0$ emerges free of any further constraints. But the method given here could equally apply to times *earlier* than all the 'information-gathering' times, so giving 'retrodiction' for $\rho(t)$ at times before the data. As the underlying equations are time reversible (subject to reversal of velocities and magnetic fields), any solution $\rho(t)$ obtained from the procedure could be replaced with its time inverse, and the resulting observations m_1, m_2, ... at the corresponding sequence of times in reverse, to give another solution obtainable from the procedure. Thus the method contains nothing to select, from the class of all solutions of Liouville's equation, those which show macroscopic variables tending towards equilibrium. It cannot, therefore, lead to macroscopic evolution equations for $a(t)$ which (like the heat conduction equation) are time-asymmetric. On further consideration, this is not surprising, as Maxent does not say anything about time asymmetry and the procedure outlined above contains no ingredient that would create asymmetry.

Applying the method to the evolution of $a(t)$ when given $a(0)$ leads to the following. At $t = 0$ the resulting ρ is $\rho_{\mathrm{q}}(0)$, and it then evolves by Liouville's

equation in $t \geqslant 0$, so supplying $a(t)$ (and similarly, in $t < 0$ as well). This leads to a situation that we rejected in § 12. It puts ρ on $\mathcal{E}_q$ at the privileged time $t = 0$, from which it migrates at other times, and the resulting law of evolution for $a(t)$ fails to be translationally invariant with respect to t, (or, in subdynamics terminology, this projection operator fails to commute with the liouvillian). The ideas we described in § 12 (Lewis's principle) represent a crude attempt to correct this: by allowing evolution for finite time-steps and remaximizing S at the end of each step, plausible results can be obtained and the law of increasing entropy ensured (Lewis 1967). It can only be concluded that, for all its apparent generality, the time-dependent Maxent as described so far simply does not answer our question. It does not introduce irreversibility and it does not give macroscopic evolution laws that are translationally invariant with respect to time.

Jaynes (p. 289) notes the unsatisfactory point that ρ has the form ρ_q at $t = 0$, calling this 'somewhat of a mystery'. It would mean that irreversible processes are not yet in progress at $t = 0$, whereas intuition would suggest they should be already in progress. This point is offered to motivate the use of entropy maximization with respect to data at more than one time. (As a further motivation (p. 294), systems that exhibit 'memory', or hysteresis, are mentioned, and while it is true that an eventual theory would have to encompass these, they are not discussed in depth in the present paper, and are not needed for motivation here.) The following is my account of how the theory should be developed.

The idea is that a macroscopic evolution $a(t)$ shall be accompanied by a solution of Liouville's equation $\rho(t)$, in such a way that, besides satisfying $M\rho = a$ at all t, it shall also have the property that, at each t, $\rho(t)$ shall be that which maximizes $S[\rho]$ subject to all the *earlier* values of $a(t')$, $t' \leqslant t$. So the time-dependent Maxent principle is applied not merely to the initial data like $a(0)$ or to items of experimental data at miscellaneous times, as we envisaged before; it is to be applied to $a(t')$ itself as it is computed from the theory, in a recursive fashion.

The condition to be satisfied by $\rho(t)$ in respect of the observable at the earlier time t' is analogous to (16.2):

$$M\,(X, t' - t)\,\rho(t) = a\,(t')\,, \qquad t' \leqslant t. \tag{16.4}$$

It is unclear how far back in time t' should extend, so we provisionally set a lower limit $t' \geqslant t_0$. We now proceed as usual with the new constraint (16.4); this introduces a Lagrange multiplier $\overline{\lambda}\,(t')$ which, besides being (as previously) vector-valued to handle the chosen set of observables, is now also a function of the parameter t'. The variational procedure leads to an analogue of (16.3),

$$\ln \rho(X, t) = -\lambda_n - \int_{t_0}^{t} \overline{\lambda}\,(t')\,.M\,(X, t' - t)\,\mathrm{d}t', \tag{16.5}$$

where λ_n deals with the normalization (the reason for the notation with overbar will appear below).

This is the essential result. However, unless we allow the limit $t_0 \to -\infty$, it would appear to be no advance on the ideas described earlier in this section, as the 'information-gathering' times are of finite extent. If on the other hand we set $t_0 = -\infty$ there is a new situation. A preferred direction of time has been imposed, and irreversibility may be described. The symmetry breaking can be traced to the physicist's notion of prediction, i.e. to his role as the observer and

to his mental processes; in other words to what is also called the 'physiological arrow of time'. The limit $t_0 \to -\infty$ also leaves (16.5) translationally invariant with respect to shifts in time. But it raises a possible difficulty with respect to convergence.

Although Jaynes, in the 1979 article, does not propose the limit $t_0 \to -\infty$ in general terms, he discusses as an application the question of near-equilibrium properties, and here (p. 302) he includes that limit and shows that the problem leads to a Wiener–Hopf integral equation for $\lambda(t')$.

Adopting that limit in the context of general theory, we record the final result,

$$\ln \rho(X,t) = -\lambda_n - \int_{-\infty}^{t} \overline{\lambda}(t') . M(X, t'-t)\, \mathrm{d}t', \tag{16.6}$$

or, with $\eta = -\ln \rho$ and $t' = t + \tau$,

$$\eta(X,t) = \lambda_n + \int_{-\infty}^{0} \overline{\lambda}(t+\tau) . M(X,\tau)\, \mathrm{d}\tau. \tag{16.7}$$

This has to be used in conjunction with the constraints.

17. Further consideration of Maxent results

On comparing equations (13.13) and (16.7) it will be obvious that we have reached the same basic result by two very different routes. The equations are not, however, identical, so we must firstly comment on the differences. One of them concerns the handling of the normalization, which is plainly a mere technicality and need not be pursued here. The other difference arises from the treatment of the lower limit of integration at $\tau \to -\infty$; in our own derivation of (13.13) the convergence of the integral was not assumed, hence the device of inserting the convergence factor with the instruction to let $\epsilon \to 0$ at the end. Zubarev's own treatment also uses this device. As with subdynamics, it would be very helpful to establish analytic conditions under which this procedure is correct either in the sense of convergence or summability. Such conditions would certainly involve the level of stochasticity of the dynamical system, as discussed in § 5 above, requiring a continuous spectrum or perhaps some stronger characterization; however, I expect that the required condition would also involve, jointly, properties of the kernel M, i.e. the choice of observable. When such conditions hold, we may let $\epsilon \to 0$ in (13.13), and (after dealing with the normalization), it coincides with (16.7) if we identify $\epsilon\lambda(\epsilon)$ of (13.13) with $\overline{\lambda}$ of (16.7).

Next we comment on the derivation of (16.7) presented in § 16. There is an important distinction to be made between two activities to which Maxent has been applied. The first of these, which may be called *data processing* has as its input a collection of data in the form of numbers; it may be a large collection but is insufficient to prescribe the object completely. The desired output is further numbers, which would have been calculable exactly if complete data were known, but are to be estimated on the basis of incomplete data. This activity is not in general accompanied by time evolution subject to known differential equations; for example the applications to imperfect images in astronomy, or fingerprints, have no time evolution. (In contrast, an application to weather forecasting would include dynamical equations.) The second activity, which may be called *theoretical statistical mechanics* has as its input a differential equation, namely Liouville's

equation (which in turn was obtained from the law of molecular interaction) along with boundary conditions, and has as its output a new set of differential (or other types of) equations. The latter form the starting-point for branches of macroscopic physics, such as continuum mechanics, fluids, plasma physics, solid state physics, and so on. They enable the macroscopic physicist to work on those fields without ever having to refer back to the molecular dynamics; he will not need to do any molecular data processing, with the hazard of instability with respect to initial data.

Maxent as applied to data processing operates with great generality: provided the input data are not actually contradictory it can supply an answer to any question, with the obvious proviso that the answer is only an estimate and may turn out not to be true. Maxent as applied to *theoretical statistical mechanics* is more subtle, once one ventures beyond the case of thermal equilibrium. In NESM, as we saw in the previous section, to get the results the macroscopic physicist is looking for, which are contained in (16.7), one must be able to lower the limit of integration to $-\infty$. Although we have not formulated the condition of stochasticity required for this, it is plain that such a condition is indeed required, for otherwise statistical mechanics would apply to integrable systems, which is false. One has to conclude that Maxent alone cannot supply the foundations of NESM which form the subject of this paper.

We now have, in §§ 13 and 16, two derivations of (16.7). They both start from the existence of a liouvillian, L, and the adoption of a macroscopic description specified by M, and both are contingent upon the convergence of the integral. But the derivation of § 13 makes no use of Maxent! The integrals over past time, which entered (16.7) as a result of imposing the maximization of S with respect to earlier possible evolutions, appear in (13.13) as a result of the more mundane operation of eliminating ζ. So, in any instance where there is a valid equation of the form (16.7), it can be reached without appeal to the Maxent principle.

It would appear that, when the limit $t_0 \to -\infty$ is applied, the Maxent conditions become logically superfluous. I offer the following interpretation of this outcome. In the 'theoretical statistical mechanics' mode of thought one is only trying to explain 'histories' $a(t)$ which are macroscopic evolutions that occur naturally. The fact that they occur at all, and reproducibly, signifies that the 'unwanted' microscopic information can be eliminated. That process of elimination has been mathematically developed here in two distinct but similar presentations, namely in § 7 and § 13; success is contingent upon the elimination of the initial value of the unwanted data in the limit that the initial time is pushed into the remote past, i.e. our limit $\tau \to -\infty$ in the integrals that we have discussed. It is remarkable that the same integrals appear from the algorithm of retrospective Maxent (§ 16). In my view, this has encouraged a questionable application of the Maxent philosophy to the present problem. Since the result is obtained uniquely by the procedure of § 13, there is no scope for maximizing the entropy. The illusion that it is being maximized can be likened to the situation in the case of a variational problem with discrete parameters, such as Jaynes's dice problem (Jaynes, p. 41ff). If it should happen (unnoticed) that the number of constraints equals the number of unknown variables, one could proceed to introduce undetermined multipliers and carry out what looks like the normal manipulations to solve for the multipliers and the variables themselves. Although this would not lead to error, it would create a mistaken impression that maximization was being achieved.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

I suggest that a correct application of the Maxent formulation occurs in a case where data about $a(t)$ has indeed been supplied in a finite time interval, and it is desired to make the best prediction of subsequent evolution. The supplied data need not form a natural evolution: it could represent a fluctuation. The procedure leading to (16.3) would than give the best estimate of that subsequent evolution. One could do this even in circumstances where the integral over past time fails to converge, so there is no reduced evolution equation. Such applications should be regarded as 'data processing'. But when an evolution equation exists, it very quickly takes over and determines the future macroscopic motion almost uniquely, as implied by the dynamics, with no scope for maximizing entropy. I regard this as almost tautological: if we restrict ourselves to the macroscopically predictable, than the system contains whatever features are required to produce unique answers without invoking entropy maximization.

Once (16.7) has been reached, by whatever route, the redundancy arising from incomplete information has been resolved. But it remains a formidable problem to extract from (16.7) the required equations for evolution of $a(t)$. It is unclear how to eliminate events for $t < 0$ from the standard problem of predicting $a(t)$ in $t > 0$ from $a(0)$. The essence of this is, however, similar to our discussions of the Master equation in § 7. A physically useful description will be such as to have a property of 'rapid decay', whereby the integral is appreciable only for small τ. If this is not so one will have a system exhibiting hysteresis at the selected level of description.

Of course, these problems can only be tackled in the case of simple systems and with the aid of approximations, expansions and so on. Some idea of what is available can be seen in chapter IV of Zubarev's book, and there are applications in the papers of Jaynes (1983 collection) and the text of Grandy (1988). The methodology can be aided by a generalization of the partition function, but this will not be described here.

18. Equivalence of the two schools

In the previous section it became clear that the approaches of Zubarev and of the Jaynes group were in effect the same, although our presentation could be criticized for lack of rigour. As already hinted there the Brussels and Zubarev approaches can also be linked, as follows. In both developments, the aim is the (seemingly unsound) one of solving initial value problems for the Liouville equation in circumstances where the initial data are incomplete. This is only possible if the equation itself contains features whereby the influence of the missing data is quickly lost or decayed in some sense. If that is the case, a special class of solutions is identified by pushing the initial instant back to $t \to -\infty$ and ignoring the effects of the missing data. In the Brussels case this calculation is carried out in terms of ρ itself, and appears in § 7 above; the development of subdynamics in § 8 provides a useful way of organizing the subsequent calculations. In the Zubarev case, as developed by the present writer, it appears in § 13 and it differs from § 7 as a result of the change of variable from ρ to $\eta = -\ln \rho$. Although this makes considerable technical differences, the mathematical task is the same. We take this to imply that, in those circumstances where a result is possible, it will be the same in either representation. A rigorous investigation of this statement must

await the appearance of a satisfactory account of the conditions under which these procedures are possible.

Both the derivations presented here manipulate the unkown variable by means of a decomposition $\rho = \rho_1 + \rho_2$ or $\eta = \eta_q + \zeta$. The reader should avoid the supposition that these are directly related!

The following, although not contributing to the logical structure of the present paper, may be noted as something of a curiosity. We revert to a point discussed previously in § 12. There we considered the possibility of 'entropy' functionals, defined as in (12.3), allowing the function $h(\rho)$ to take a different form from that of Gibbs (in which $h(\rho) = -\rho \ln \rho$). Here $h(\rho)$ is required to be concave and bounded above, so h' is monotonic; see the paragraph following (12.4) and the final paragraph of § 12. Consider now how the work of § 13 would be modified if such a modified 'entropy' had been used. We would need to redefine $\eta(X, t)$ to be $h'(\rho(X, t))$, but the remainder of the development up to (13.13) would be formally the same. The stochastic and spectral properties of the equation are the same (as we noted at (13.3)). But in implementing (13.13) one has to make use of the condition $M\rho = a$ at each t, and this makes the technical details very different for differing choices of $h(\rho)$; the meanings of the $\lambda(t)$ and $\lambda . M$ are likewise different. Similarly, we can consider how the work of § 16 would be modified if a new choice of entropy were to be adopted for use in Maxent, purely as a mathematical exercise and without concern for philosophical considerations. Again, the work would be formally the same, arriving at an analogue of (16.7), with the technical details (imposing the constraints) different. As before, it would remain the case that the intended maximization of entropy would be illusory. If now (as was inticipated in § 12) one chooses $h(\rho) = -\frac{1}{2}\rho^2$, so that $\eta = -\rho$, there is after all no change of variable (apart from the trivial change of sign), with the result that the analogue of the work of § 13 is now simply a new presentation of the Brussels approach.

A further important difference between the two schools has to be considered. We explained in § 4 that invariants of the motion should preferably be 'factored out', so that the motion in Γ is already confined to a submanifold Σ, within which no further invariants are available, although for clarity we used a notation that did not emphasize the point. We can think of Σ as the energy surface for simplicity, as usually the energy is the only invariant. So the support of ρ is restricted to Σ, corresponding to the microcanonical distribution in the equilibrium case. As I understand it, this is always the situation in the work of the Brussels school. In contrast, the Zubarev and Maxent publications permit the support of ρ to include the whole of Γ, which is the analogue of the canonical distribution. This represents independent motion in each Σ with a measure attached. It is well known (see Khinchin 1949) that in thermal equilibrium, and asymptotically (i.e. for large systems), the canonical and microcanonical distributions are equivalent for macroscopic purposes. The essence of this equivalence is that for those layers in Γ that have energy higher than the actual value of the energy, the contribution made is rapidly attenuated by the exponential factor in the Gibbs distribution, while for the layers with a lower energy the contribution is rapidly reduced owing to the smaller volume occupied. It seems reasonable to assume that the same would happen in NESM, but I offer no formal investigation here.

The relationship between the work of the two schools is not a direct one, in the sense that both obtain explicitly the same solutions of Liouville's equation.

No such solutions have been presented here. The most one can say is that the 'staging posts' for solution, namely (7.9) and (13.13) have been related. We have argued that in the (unkown) circumstances in which results can be obtained by continuing from those equations to reach explicit answers, those answers would be the same; that conclusion is itself further qualified by the point made in the preceding paragraph.

The reader may feel disappointed that our claims of equivalence are not supported by more instances of specific formulae obtained from the rival approaches. But, as explained in § 1, actual calculations are highly intractable, and I have not succeeded in producing such results.

19. Comparison of methodology

Setting aside the conceptual differences between them, we can ask about the relative advantages of the formalisms developed by the two schools, as regards the further development of the subject.

The overwhelming advantage of the Brussels school is that by working with ρ itself the underlying mathematics is linear. It is for this reason that the Brussels school is able to advance the work well beyond the formal derivation of the Master equation of § 7. The development of subdynamics (see § 8), and in particular the discussion of equation (8.8) forms an important extension to the subject as it enables one to consider quantitatively the long-time asymptotic behaviour of the system in its approach to equilibrium. As Balescu (1975) has shown, the use of perturbation expansions, although laborious, allows specific examples to be worked out. There is thus a body of results which seem to be inaccessible from the Maxent formalism. The perturbation expansions used start from non-interacting particles as the unperturbed situation. While that would seem to be the obvious suggestion, it is in fact widely suspected that this cannot provide explanations of phenomena like long-time tails in kinetic theory. There is also the point that (at least in the absence of boundary conditions) the system of non-interacting particles is not ergodic; if it is to become stochastic even at the lowest level (ergodic) the perturbation is in some sense singular, i.e. changes the system qualitatively. Finally, the use of this perturbation method means that one is in effect repeating *ab initio* the calculations of the equilibrium case (virial expansion).

The work just mentioned has led to impressive achievements by the Brussels school. On the other hand, their attempts to carry general theory further forward, as outlined in § 11, seem to have provided little further insight, and to have attracted little notice.

As a final comment on the Brussels school we add a caution, for although calculations start from a linear formalism, the nature of the subject is liable to reintroduce nonlinearity. In the case of gases, the use of expansions of the Mayer type, with r-particle distributions and correlations, is a case in point. Here, an ingenious embedding of E into a larger space, containing additional objects called 'correlation patterns' is used (see Balescu 1975).

The advantage of the Maxent formulation is that it works with expressions that are as close as possible in form to the canonical distribution of equilibrium. As a result, it works best in circumstances where the physical state of the system is near to equilibrium. This suggests an alternative notion of 'perturbation', in which the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

unperturbed case is that of equilibrium. This leads naturally to the consideration of transport coefficients, linear response theory and irreversible thermodynamics. Zubarev's (1974) and Grandy's (1988) texts should be consulted. It seems very likely that (as with the canonical distribution) there are areas of the subject where the formalism reached via Maxent supplies more tractable calculations. Such advantages may well be explicable in terms arising from the origin of the method. The passages in Jaynes concerning his concept of 'calibre' (p. 421), are indications in that direction.

It will be noticed that the fields of application that are natural to the two approaches, chosen merely on the basis of methodology, are very nearly disjoint. This explains the point, mentioned earlier, that lack of common examples makes it difficult to test them against each other. It makes the further point that a broadly based approach to NESM would seek to make both methods available; that is an object of the present paper.

20. The law of increasing entropy

We now resume the discussion started in § 10. We noted there that the problem of macroscopic evolution, treated in the style of the Brussels school, need not involve entropy. We have subsequently shown that the results of the Maxent approach are not mathematically dependent on entropy, notwithstanding the impression that the evolution may be regarded as 'supervised' by a maximum entropy condition (just as in elementary dynamics motion is supervised by Hamilton's principle). These considerations alone are not sufficient to determine what choice (if any) to make for a non-equilibrium entropy, and, as already noted, any entropy defined as in (10.2), i.e. as a functional specified in terms of $h(\rho)$, is constant in time.

One could leave the matter there, regarding it as a fruitless diversion. But such is the prestige (or notoriety) of the *the law of increasing entropy*, that a paper under the title of the present one ought to offer some explanation of what it means. Let us accept that, for *equilibrium* states, it should be the Gibbs entropy (10.1). This ensures that it has the conventional thermodynamic interpretation for virtual transitions between neighbouring equilibrium states, i.e. $\mathrm{d}S = \delta Q/T$, so maintaining the connection with the historical engineering origin of the concept. One then also finds that entropy increases in a change (not necessarily infinitesimal) between equilibrium states, e.g. where a partition is removed, so that the initial state ceases to be equilibrium; the intervening non-equilibrium states are not considered. Turning now to the question of attaching a time-dependent entropy to non-equilibrium states, it is highly desirable that this should retain the property of being additive for disjoint systems. That also enables entropy to be given a local characterization in appropriate circumstances, by which we mean that neighbourhoods in physical space can be regarded as subsystems possessing individual contributions to the total entropy, so one can introduce an entropy density. (The 'appropriate circumstances' concern the short-range nature of the interaction potential, and do not always apply, e.g. in Vlasov plasmas. If the concept of entropy density, with the associated concept of entropy flux, exists, it is often claimed that there is a local version of the law of increasing entropy, the Clausius–Duhem inequality (see, for example, Meixner 1969).)

The requirement of additivity implies the Gibbs entropy, but if the $\rho(t)$ used

there is the 'preferred' ρ selected by any of the theories discussed above, and therefore satisfying Liouville's theorem, S_G is invariant. One possibility is to replace $\rho(t)$ by the quasi-equilibrium density $\rho_{\mathrm{q}}(t)$ appropriate to $a(t)$, so the 'observed entropy' of § 10 is

$$S_{\mathrm{obs}}(t) = -\int \rho_{\mathrm{q}} \ln \rho_{\mathrm{q}} \, \mathrm{d}\Omega. \tag{20.1}$$

This can in principle be computed for any non-equilibrium evolution, whether or not ρ_{q} appears in the methodology for making the calculations. Consider such a calculation: initially $a(0)$ is not an equilibrium state, by hypothesis, while $a(\infty)$ is that equilibrium compatible with the conserved quantities, whose $\rho_{\mathrm{q}}(\infty)$ is such as to maximize (20.1). Hence $S_{\mathrm{obs}}(0) < S_{\mathrm{obs}}(\infty)$ by construction. Underlying this, however, is the preferred $\rho(t)$ (one must not call it the 'actual' ρ!), satisfying Liouville's equation and such that $M(\rho - \rho_{\mathrm{q}}) = 0$. For this, S_G is constant, with

$$S_G < S_{\mathrm{obs}}(0) < S_{\mathrm{obs}}(\infty). \tag{20.2}$$

This is straightforward provided one bears in mind that $\rho(t)$ does *not* evolve to an equilibrium distribution, but continues to contain contributions in E_2, see (6.1).

The above is an adaptation of an argument given in a well-known article of 1965 by Jaynes (see p. 77 of the collected papers). There, the observational level was, specifically, the one-particle distribution $f(\boldsymbol{x}, \boldsymbol{v}, t)$ in a gas, so S_{obs} is just a multiple of the Boltzmann entropy

$$S_B = -\int f \log f \mathrm{d}^3x \mathrm{d}^3v. \tag{20.3}$$

In that context, one can go further than (20.2). Accepting Boltzmann's hypothesis of molecular chaos, and the neglect of triple and higher collisions, Boltzmann proved that $\dot{S}_{\mathrm{B}} \geqslant 0$, with equality only at equilibrium, i.e. $S_{\mathrm{B}}(t)$ increases *monotonically.*

Other contexts, for example heat conduction, suggest that a *law of increasing entropy* should hold in the strong form of *monotonic* increase, but the argument leading to (20.2) only demonstrates that the final value is maximal. On this point, Khinchin (1949, p. 149) remarks that 'such a proof has not yet been given', and it would seem that that remains the case to this day. Besides, there are apparent exceptions, namely the so-called 'spin-wave echoes' in solids, and the similar 'plasma wave echoes' (see Clemmow & Dougherty 1990). Temporary decrease in entropy has also been considered in molecular biology in theories of embryonic cell differentiation (Schiffmann 1991). These examples suggest that there is no need for an aura of mystery, with invocation of Maxwell demons, to be associated with the point. What is needed is a dynamic characterization of the properties of liouvillians and observables that ensures that entropy increases monotonically, as it nearly always does.

(It may be noted in passing that Maxwell's demons, along with coarse-graining and other obsolete sources of confusion that have plagued the subject, are wisely omitted from the presentations of both the 'schools'.)

Finally, we recall from § 11 that among the additional developments emanating from the Brussels school is the notion of an internal 'time' operator, somewhat analogous to a Lyapunov function (see Misra & Prigogine (1982). This, if it exists, provides a quantity that increases monotonically, so indicating the sense of an

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

irreversible process at any instant, i.e. expresses the Second Law. But the 'internal time' is not related numerically to entropy in a simple way.

21. Equilibrium statistical mechanics revisited

As NESM is so complicated a subject, previous attempts to justify the foundations of *equilibrium* statistical mechanics have generally dealt with that question alone, i.e. not as part of the wider problem of NESM.

If we were to do that in the spirit of the present paper, we would (see § 4) insist on working with densities rather than individual orbits. We would then assume that the energy is the only integral (noting that that this could easily be generalized) and we would consider the restriction of the motion and the densities to an energy surface Σ. We would *assume* it to be ergodic (we discuss this further in § 22). Returning to the spectral properties listed in § 5, we note that ergodicity implies that $\lambda = 1$ is a *simple* eigenvalue, the eigenfunction being $\rho = \text{const}$. Thus the only time-independent solution available is $\rho = \text{const.}$ in Σ, and we have the microcanonical distribution. As we have noted before, this can be replaced with the canonical distribution if the number of particles is large.

The Maxent physicist would also work with ρ, which he would call 'probability'. He would know that energy is constant (whether or not the motion is ergodic) and if the value of the energy is the only data, he would use bayesian inference, hence the maximization of the Gibbs entropy, to reach the canonical distribution.

However, neither of these lines of argument has been widely adopted, and most physicists would give the answer that was developed in the 1930s (or dismiss the question as pedantic!). That answer is presented in complete detail in Khinchin (1949), and elsewhere. There is for example an elegant summary in Reed & Simon (1972, pp. 54 and 237). In this view, observations refer to an actual orbit $X(t)$, and the fact that observables other than the energy (such as pressure, magnetization, etc.) are found to have constant values is explained by saying a time-averaging is taking place. If the ergodic property is assumed, $X(t)$ eventually visits any open set in Σ, so that an average over a very long time results in a sampling of all of Σ. The difficult part, Birkhoff's ergodic theorem, establishes that the resulting average equals the microcanonical average.

Jaynes has, on several occasions, criticized this account of equilibrium (see pp. 104–106 of the collected papers for a good discussion). The time taken for $X(t)$ to make a reasonable sampling of Σ would be enormous even for quite a small system, typically more than the age of the universe. As it certainly far exceeds the time taken for such systems to reach equilibrium from a typical non-equilibrium state, it is obvious that such an explanation is unacceptable, and that it could not form a part of the subject in a treatment that also encompasses NESM.

We agree fully with this criticism. But Birkhoff's ergodic theorem is only one item in the whole subject of ergodic theory as set out in the books in our reference list; other relevant results include those we have summarized in § 5. Unfortunately, Jaynes uses the point about time averages as a reason for rejecting all things 'ergodic', and here we disagree.

The point of view of the present paper is that a common explanation is required for equilibrium and NESM alike. It follows that the explanation of the equilibrium states is their role as the eventual asymptotic state of non-equilibrium processes.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

As Liouville's equation by itself does not provide equilibrium, $\rho(t)$ does not 'forget' its past. It is for this reason that the notion of observables, and the decomposition of function space (§6) is required, enabling the 'unknowable' details to be relegated to E_2 at large t. The way this works is different for different choices of observable, hence the richness and difficulty of the subject. Whatever the choice, the *deviations* from equilibrium must eventually decay—that is the acknowledged experimental situation. How can they be guaranteed to do so? Broadly, the answer can only be that the system should possess at least the characterization listed in §5 as weak mixing, ensuring that the spectrum in $E^{\perp}$ should be continuous. Incidentally, it shows Gibbs's foresight at its most impressive! Once this is understood, and it is desired to study the properties of equilibrium states in their own right, one may as well redefine ρ to be one of the standard distributions, discarding the 'noise' from E_2.

If specifically required (e.g. for pedagogical purposes) to offer an explanation of the equilibrium distributions without reference to NESM, we would propose the one given at the start of this section. Our attitude to time averaging is the following. For NESM to be possible, a system much be at least weak mixing. By the hierarchy property described in §5 it must *a fortiori* be ergodic. Hence the conditions of Birkhoff's theorem are satisfied, so it is the case that the long-time average of any observable is the same as its equilibrium value. But the status of that result is only a mathematical curiosity, and it is not used as the basis of any explanation. Moreover it is not surprising, as the time for averaging far exceeds any initial period for equilibration.

22. Availability of stochastic properties

A question of concern is that of whether, and if so to what extent, real physical systems possess the properties described in §5, for if not the theories described here are empty, having no implementation. There is of course the well-known result of Sinai (1970); even that applies to a rather artificial case, and the proof of ergodicity is lengthy and technically difficult. A few other special results have been obtained more recently, for example by Knauf (1987), but there remains a large gap, which has been a source of criticism. One could simply reply that it will no doubt be closed by the efforts of future mathematicians, but that may be over-optimistic in view of the great variety of laws of force and types of molecules.

That variety accounts for the diversity of matter as observed macroscopically, but we must recall that it has common qualitative features, specifically the approach to equilibrium. Some writers like to emphasize that the Second Law is one of the most general statements in science. So, in attempting to provide foundations for NESM, it is desirable that they should lead to a formalism from which these common features can be explained, even though its use may well be laborious in practice. This is the reason for seeking helpful abstract properties of large dynamical systems, and it is hard to see how the approach to equilibrium could be otherwise explained in general terms.

I find this convincing evidence that the spectrum of macroscopic liouvillians must, in practice, be continuous or, if not, failure in that respect must occur only on unrepresentative energy surfaces, or alternatively failure would take the form of spectra which are not continuous but contain enormous numbers of discrete eigenvalues (I avoid misusing the technical term 'dense'). In the last case, macro-

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

scopic behaviour would effectively be irreversible over a very long timescale, and if that timescale becomes cosmological, it is no longer possible to disentangle NESM from cosmology itself. This happens roughly when the separation of the relevant frequency eigenvalues is of the order of the reciprocal of the age of the universe.

The past 15 years has seen an explosion in research on low-dimensional non-linear dynamics, prompted by the availability of microcomputers. This research has revealed that the subject contains a wealth of complicated structure and phenomena that were not anticipated. What are the implications of this for NESM? As most of the results are specialized, the answer is 'none'. However, we offer some comments on the so-called KAM theorem.

Non-integrable systems often contain special simple solutions. In the case of gravitating particles, the three-body problem is non-integrable, but there are known simple solutions. For example the particles (which need not be of equal mass) can move on the vertices of an equilateral triangle rotating in its plane. One might be tempted to guess that such individual solutions would occupy zero measure in phase space. The KAM theorem shows that this is not so, as phase space generally contains ergodic and non-ergodic regions both having positive measure. (If this were not so, the solar system would be unstable!) NESM as presented here can survive this point if the non-ergodic regions are of far lower measure than the ergodic regions; it would reveal the possibility of special initial conditions (which have small but non-zero measure) leading to behaviour in disagreement with macroscopic physics. The question is therefore the proportion of measure occupied by the non-ergodic regions. It would be helpful to know whether that proportion diminishes as the number of particles increases, as it could then be safely ignored at $N = 10^{23}$. There have been some studies of infinite hamiltonian systems (Pöschel 1988), but the question is still open.

As we concluded in § 18 that the theories of the two schools (or three, if Zubarev is to be recognized as separate) are ultimately compatible, the question raised in this section has the same implications whichever approach is used. We note here the attention given to the matter. In their texts, neither Balescu nor Zubarev is much concerned about the point, though on p. 529 Balescu declares that his book is about those systems for which his methodology works. The case is different for Jaynes. On the one hand (p. 291) he tells us that maximum entropy 'is the only principle needed to construct ensembles which predict any experimentally reproducible effect, reversible or irreversible'. In view of the discussion of our § 17, we cannot agree with that, as the use of maximum entropy is unnecessary. On the other hand, the reference to 'experimentally reproducible effect' is perceptive, and is enlarged on a little later (p. 297) where we are told that 'it is not the business of statistical mechanics to predict everything that can be observed in nature, only what can be observed reproducibly'. We agree with that, but not with the apparent attitude that the 'reproducible' cases are to be identified by trial and error, or experience. We believe that it is the business of theoretical statistical mechanics to identify what underlies the reproducible case, at least in general terms. Our answer is the requirement that the integrals over past time that occur in any of these approaches (in (7.9), Brussels school; (13.13), Zubarev; (16.7) Jaynes) should converge. That is turn imposes a stochastic condition on the liouvillian, such as we have discussed in the present section.

It seems to me almost certain that conditions of some kind must also be placed

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

upon the level of observation, M, for reproducibility to be obtained. One possibility is to assume that the system has a countable Lebesgue spectrum, (property (e) of § 5). Then the set $\{U^n f_j\}$ might be partitioned into two subsets corresponding to what is observed and what is discarded, respectively. Complete separation would occur if the partition depends only upon j, with a weaker condition being sufficient to ensure separation in the sense required for NESM. But it is readily seen that weaker conditions will suffice in some cases. In those parts of E that are orthogonal to $M(X,t)$ for all t (if any), the behaviour of L is of no consequence, so the system could even possess a point spectrum (i.e. fail to be mixing). This does not seem to have been developed, though Spohn (1975) has made some investigations.

23. Note on quantum mechanics

Our exposition of NESM so far has been based entirely on classical, rather than quantum, mechanics. It is preferable to restrict attention to one or other in presenting the above material, as alternating between the two increases the scope for confusion. Using classical mechanics permits the connection with the standard ideas of ergodic theory, and moreover leads to an explanation of irreversibility. But the real world is to be described by quantum mechanics, and the writings of the rival 'schools' often adopt the quantum mechanical setting, so we now outline the translation of the formalism into quantum-mechanical language, and draw attention to the modification required.

Starting from elementary quantum mechanics one has a wave function $\psi \in \mathcal{H}$ in the appropriate Hilbert space, taking account of the usual symmetry or antisymmetry. Statistical mechanics is reached by introducing the so-called density matrix $\rho \in E$, where $E = \mathcal{H} \otimes \mathcal{H}$. The equation of motion, derived from Schrödinger's equation, is the familiar quantum Liouville equation

$$\mathrm{i}\hbar \frac{\mathrm{d}\rho}{\mathrm{d}t} = [H, \rho] = H\rho - \rho H, \tag{23.1}$$

where $\hbar$ is Planck's constant. This equation can be written $\dot{\rho} = L\rho$ as before, but now $L : E \to E$ has the form of a quantum mechanical 'superoperator', while $H : \mathcal{H} \to \mathcal{H}$ is the ordinary quantum mechanical hamiltonian operator. The expectation value of a quantum observable Q is

$$\langle Q \rangle = \mathrm{tr}\,(Q\rho) \tag{23.2}$$

and the normalization is simply $tr(\rho) = 1$. As ρ, in its original form, is a function of two 'copies' of the coordinates that originally appeared in ψ, the 'trace' in (23.2) is an integration over these variables, so it is of the same mathematical form as in the classical case. We can then extend this notion from a single observable to a collection of them labelled by new parameters, so reaching the familiar $M\rho = a$ where $M : E \to A$ is a linear map, just as at (3.3) above.

To work out these ideas in practice, considerable effort has to be devoted to the quantum mechanical representations suitable for the task in hand. One possibility is to transform to an occupation-number representation, or even to the Fock space formalism if the number of particles is not prescribed, so leading to the grand canonical distribution in the equilibrium theory. Another route is to transform ρ to the Wigner function picture, giving a closer correspondence in the classical

theory, and enabling the construction of the r-particle distributions and the Mayer expansion. These developments appear in texts, and with unusual completeness and clarity in Balescu (1975), but we shall not require the details here as we can work at a more abstract level.

One technicality which we need to note is that if ρ satisfies (23.1) than so does $f(\rho)$ where f is at least twice-differentiable. Although less obvious than in the classical version, this is readily shown, and enables us to work with $\eta = -\log \rho$ and deal with the Gibbs entropy.

Equipped with the information summarized in this section, it can be seen that the material in §§ 6–20 can be read *without formal change* as an exposition of quantum statistical mechanics; only a reinterpretation of the formalism is required, as it is based on nothing more than the equations $\dot{\rho} = L\rho$ and $M\rho = a$.

There must be an eventual difference, of course, and it can be traced to the nature of Liouville's equation. In classical mechanics, the equation is of *first* order in the qs and ps, whereas in quantum mechanics it is of *second* order. As is well known, any quantum system confined to a finite part of physical space has a discrete spectrum. Thus a finite system is not mixing, and one must use an argument similar to that mentioned in the previous section. For a system composed of, say, 10^{23} molecules, the spectrum contains such an astronomically large number of eigenvalues that the separation between them is minute. The reciprocal of that separation (expressed in terms of frequencies) gives an estimate of the time that would elapse before the failure of mixing would be detectable. That argument is in turn modified at very low temperature, by quantum degeneracy, and here one does indeed find effects – superconductivity and superfluidity – where irreversibility is defied.

The subtle nature of the limit $\hbar \to 0$, whereby the spectrum changes from discrete to continuous has recently been re-examined by Berry (1987) and by others, and is the subject of intense current research. The corresponding transition régime for macroscopic properties in NESM is likely to be equally complex.

24. Conclusion

The aim of this paper has been to show how the apparently disparate approaches to NESM can be related, and that they can be seen as part of a common underlying structure. For systems that are mixing in the limit $\hbar \to 0$, that structure enables the construction of evolution equations for observables at a selected observational level. The macroscopic origins of irreversibility can be traced, and although it must eventually be related to cosmology, it is possible to see how irreversible behaviour appears mathematically for large but finite isolated systems, and how the necessary calculations are to be based (in principle) without resort to any additional assumptions.

I am grateful to have been able to discuss this subject with both Professor I. Prigogine and Professor E. T. Jaynes. The fact that my viewpoint differs considerably from both of theirs does not detract from the enormous stimulation that has resulted from their writings. While this typescript was in the course of revision, I had hoped to discuss this work with Professor D. N. Zubarev at the Conference STATPHYS18 in Berlin, 2–8 August 1992, as he had also been an important influence; it was with sadness that I learned that he had died in a traffic accident on 29 July. I am also grateful in various ways to Professor Michael Redhead, Professor Mario Rasetti, Professor Radu Balescu, Dr John Skilling, Professor Oliver Penrose and Dr Peter Coveney.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

References

Balescu, R. 1975 *Equilibrium and non-equilibrium statistical mechanics.* Wiley.

Balescu, R. & Misguich, J. H. 1973 Fontenay-aux-Roses Report EUR-CEA-FC-691.

Berry, M. V. 1987 The Bakerian Lecture, 1987. Quantum chaology. *Proc. R. Soc. Lond.* A**413**, 183–198.

Bogoliubov, N. N. 1962 Problems of a dynamical theory in statistical physics. In *Studies in statistical mechanics*, vol. 1 (ed. J. de Boer & G. E. Uhlenbeck), pp. 5–118. North-Holland. (Russian original of 1946.)

Brown, J. R. 1976 *Ergodic theory and topological dynamics.* Academic Press.

Chapman, S. & Cowling, T. G. 1970 *Mathematical theory of non-uniform gases*, 3rd edn. Cambridge University Press.

Clemmow, P. C. & Dougherty, J. P. 1990 *Electrodynamics of particles and plasmas.* Addison-Wesley.

Cornfeld, I. P., Fomin, S. V. & Sinai, Y. G. 1981 *Ergodic theory.* Springer-Verlag.

Courbage, M. 1982 *a* Mathematical problems of irreversible statistical mechanics for quantum systems. I. Analytical continuation of the collision and destruction operators by spectral deformation method. *J. math. Phys.* **23**, 646–651.

Courbage, M. 1982 *b* Mathematical problems of irreversible statistical mechanics for quantum systems. II. On the singularities of $(\psi(z) - z)^{-1}$ and the pseudo-Markovian equation. Application to Lee model. *J. math. Phys.* **23**, 652–658.

Courbage, M. 1986 Remarks on a recent article by C. Obcemea and E. Brändas on the theory of subdynamics. *Annls Phys.* **166**, 250-252.

Coveney, P. V. & George, C. 1987 On the time-dependent formulation of analytical continuation in non-equilibrium statistical mechanics. *Physica* A**141**, 403–426.

Coveney, P. V. & Penrose, O. 1992 On the validity of the Brussels formalism in statistical mechanics. *J. Phys.* A**25**, 4947–4966.

Crawford, J. D. & Hislop, P. 1989 Application of the method of spectral deformation to the Vlasov–Poisson system. *Ann. Phys. N.Y.* **189**, 265–317.

Denbigh, K. G. & Denbigh, J. S. 1985 *Entropy in relation to incomplete knowledge.* Cambridge University Press.

Dougherty, J. P. 1972 Lewis's principle in non-equilibrium statistical mechanics. DAMTP Preprint.

Dougherty, J. P. 1993 Explaining statistical mechanics. *Stud. History Phil. Sci.* **24**. (In the press.)

Duering, E., Otero, D., Plastino, A. & Proto, A. N. 1985 Information theory and Riemann spaces: an invariant formulation. *Phys. Rev.* A**32**, 2455–2461.

Ehrenfest, P. & Ehrenfest, T. 1912 *The conceptual foundations of the statistical approach to mechanics* (trans. C. Moravec). Cornell University Press, 1959.

Fine, T. L. 1973 *Theories of probability.* Academic Press.

Ford, J. 1983 How random is a coin toss? *Physics Today* (April issue), pp. 40–47.

Gell-Mann, M. & Goldberger, M. L. 1953 The formal theory of scattering. *Phys. Rev.* **91**, 398–408.

Gillies, D. A. 1973 *An objective theory of probability.* Methuen.

Goldstein, J. A. 1985 *Semigroups of linear operators and applications.* Oxford.

Goldstein, S. & Penrose, O. 1981 A non-equilibrium entropy for dynamical systems. *J. statist. Phys.* **24**, 325–343.

Grandy, W. T. 1980 Principle of maximum entropy and irreversible processes. *Phys. Rep.* **62**, 175–266.

Grandy, W. T. 1988 *Foundations of statistical mechanics*, vol. 2. D. Reidel.

Jaynes, E. T. 1983 *Papers on probability, statistics and statistical mechanics* (ed. D. Rosenkrantz). D. Reidel.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Jaynes, E. T. 1968 Review of 'Principles of statistical mechanics – the information theory approach'. *IEE Trans.* **IT14**, 611–612.

Jowett, J. M. 1982 Statistical mechanics of systems with strongly time-dependent hamiltonians. Ph.D. thesis, Cambridge University.

Khinchin, A. I. 1949 *Mathematical foundations of statistical mechanics.* Dover.

Knauf, A. 1987 Ergodic and topological properties of Coulombic periodic potentials. *Communs math. Phys.* **110**, 89–112.

Koopman, B. O. 1931 Hamiltonian systems and transformations in Hilbert space. *Proc. natn. Acad. Sci. USA* **17**, 315–318.

Lanz, L., Lugiato, L. A. & Ramella, G. 1971 On the existence of independent subdynamics in quantum statistics. *Physica* **54**, 94–136.

Lasota, A. & Mackey, M. C. 1985 *Probabilistic properties of deterministic systems.* Cambridge University Press.

Levine, R. D. 1979 Maximal entropy procedures for molecular and nuclear collisions. In *The maximum entropy formalism* (ed. R. D. Levine & M. Tribus), pp. 247–272. MIT Press.

Lewis, R. M. 1967 A unifying principle in statistical mechanics. *J. math. Phys.* **8**, 1448–1459.

Lorch, E. 1962 *Spectral theory.* Oxford University Press.

Mañé, R. 1989 *Ergodic theory and differentiable dynamics.* Springer-Verlag.

Mayer, J. E. 1960 Ensembles of maximum entropy. *J. chem. Phys.* **33**, 1484–1487.

McLennan, J. A. 1963 The formal statistical theory of transport processes. *Adv. chem. Phys.* **5**, 261–317.

Meixner, J. 1969 Processes in simple thermodynamic materials. *Arch. ration. Mech. Analysis* **33**, 33–53.

Misra, B. 1978 Nonequilibrium entropy, Lyapounov variables and ergodic properties of classical systems. *Proc. natn. Acad. Sci. USA* **75**, 1627–1631.

Misra, B., Prigogine, I. & Courbage, M. 1979 From deterministic dynamics to probabilistic descriptions. *Physica* **A98**, 1–26.

Misra, B. & Prigogine, I. 1982 Time, probability and dynamics. In *Long-time prediction in dynamics* (ed. C. W. Horton Jr, L. E. Reichl & V. G. Szebehely), pp. 21–43. Wiley.

Obcemea, C. & Brändas, E. 1983 Analysis of Prigogine's theory of subdynamics. *Annls Phys.* **151**, 383–430.

Parry, W. 1981 *Topics in ergodic theory.* Cambridge University Press.

Petersen, K. 1983 *Ergodic theory.* Cambridge University Press.

Pöschel, J. 1988 A general infinite dimensional KAM theorem. University of Bonn preprint.

Prigogine, I., George, C. & Henin, F. 1969 Dynamical and statistical descriptions of N-body systems. *Physica* **45**, 418–434.

Prigogine, I., George, C., Henin, F. & Rosenfeld, L. 1973 A unified formulation of dynamics and thermodynamics. *Chemica Scripta* **4**, 5–32.

Prigogine, I., Mayné, F., George, C. & de Haan, M. 1977 Microscopic theory of irreversible processes. *Proc. natn. Acad. Sci. USA* **74**, 4152–4156.

Reed, M. & Simon, B. 1972 *Methods of modern mathematical physics*, vol. 1. Academic Press.

Richardson, J. M. 1960 The hydrodynamical equations of a one component system derived from non-equilibrium statistical mechanics. *J. math Analysis Appl.* **1**, 12–60.

Robin, W. A. 1990 Non-equilibrium thermodynamics. *J. Phys.* **A23**, 2065–2085.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Schiffmann, Y. 1991 An hypothesis: phosphorylation fields as the source of positional information and cell differentiation – (cAMP,ATP) as the universal morphogenetic Turing couple. *Prog. Biophys. molec. Bio.* **56**, 79–105.

Shore, J. E. & Johnson, R. W. 1980 Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inform. Theory* **IT26**, 26–37; see also errata, **IT29**, 942–943.

Sinai, Y. G. 1970 Dynamical systems with elastic reflections. *Russian math. Surv.* **25**(2), 137–189.

Spohn, H. 1975 Spectral properties of Liouville operators and their physical interpretation. *Physica* A**80**, 323–338.

Stirland, S. J. 1983 Theories of non-equilibrium statistical mechanics. Ph.D. thesis, Cambridge University.

Walters, P. 1982 *An introduction to ergodic theory.* Springer-Verlag.

Widder, D. V. 1946 *The Laplace transform*, 2nd edn. Princeton University Press.

Zeldovich, Y. B., Ruzmaikin, A. A. & Sokoloff, D. D. 1990 *The almighty chance.* World Scientific.

Zubarev, D. N. 1974 *Non-equilibrium statistical thermodynamics.* Plenum.

Zubarev, D. N. 1961 The statistical operator for non-equilibrium systems. *Dokl. Akad. Nauk. SSSR* **140**, 92–95. (English transl., 1962, *Sov. Phys. Dokl.* **6**, 776–778.)

Received 23 January 1992; accepted 7 July 1993

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES