J. Phys. A: Math. Gen. 14 (1981) 1701–1718. Printed in Great Britain

H theorems in statistical fluid dynamics

George F Carnevale†‡, Uriel Frisch§ and Rick Salmon¶

‡National Center for Atmospheric Research¶, Boulder, Colorado 80307, USA
§Division of Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA
and
CNRS, Observatoire de Nice, Nice, France
¶Scripps Institution of Oceanography, La Jolla, California 92093, USA

Received 28 April 1980, in final form 6 November 1980

Abstract. It is demonstrated that the second-order Markovian closures frequently used in turbulence theory imply an H theorem for inviscid flow with an ultraviolet spectral cut-off. That is, from the inviscid closure equations, it follows that a certain functional of the energy spectrum (namely entropy) increases monotonically in time to a maximum value at absolute equilibrium. This is shown explicitly for isotropic homogeneous flow in dimensions $d \geq 2$, and then a generalised theorem which covers a wide class of systems of current interest is presented. It is shown that the H theorem for closure can be derived from a Gibbs-type H theorem for the exact non-dissipative dynamics.

1. Introduction

In the kinetic theory of classical and quantum many-particle systems, statistical evolution is described by a Boltzmann equation. This equation provides a collisional representation for nonlinear interactions, and through the Boltzmann H theorem it implies monotonic relaxation toward absolute canonical equilibrium. Here we show to what extent this same framework formally applies in statistical macroscopic fluid dynamics.

Naturally, since canonical equilibrium applies only to conservative systems, a strong analogy can be produced only by confining ourselves to non-dissipative idealisations of fluid systems. For example, such an idealised system is represented by the incompressible Navier–Stokes equation when an ultraviolet spectral cut-off is imposed and viscosity is set equal to zero. Also for such idealisations Liouville's theorem applies (Lee 1952); that is, the evolution of the system can be represented as incompressible flow in a phase space.

Unlike the dynamical equations for many-particle systems, the primitive field equations for fluids cannot generally be put in Hamiltonian form even in the non-

† Work supported in part by the National Science Foundation under Grant No DMR-77-10210. Based in part on a thesis submitted by GFC in partial fulfillment of the requirements for PhD, Harvard University, 1979.
¶ The National Center for Atmospheric Research is sponsored by the National Science Foundation.

dissipative idealisation (e.g. the inviscid Navier–Stokes equation is not derivable from a Hamiltonian in terms of Eulerian velocity fields (Millikan 1929)). Thus we are aiming for a generalisation of Boltzmann's H theorem for non-Hamiltonian systems which nonetheless obey Liouville's theorem. Such a generalisation is non-trivial for the following reason. The proof of Boltzmann's H theorem involves manipulations which invoke the collisional nature of the interaction integrals. This collisional structure is induced by the underlying Hamiltonian dynamics. The direct analogue of the Boltzmann equation is the second-order Markovian closure equation in statistical fluid dynamics. This equation can be obtained by techniques directly analogous to those used in the statistical quantum field theory derivation of the Boltzmann equation (Kadanoff and Baym 1962, Kraichnan 1959, 1971, Carnevale 1979, Carnevale and Martin 1981). An alternative derivation is the eddy-damped quasi-normal Markovian (EDQNM) procedure (Orszag 1977, Rose and Sulem 1978) which is generalised here in § 4. This equation does not, in general, provide a collisional interpretation of nonlinear interactions and so an H theorem cannot be implied by direct analogy with many-particle physics. A well known exception to this is the case of wave interactions in fluids where the wave amplitude equations derive from a Hamiltonian formalism. In that case there is a well defined collision cross section and the second-order Markovian closure equation is indeed a Boltzmann equation; hence, the H theorem follows directly (Hasselmann 1966, Webb 1978).

In § 2 we derive the H theorem for incompressible, isotropic, homogeneous, d-dimensional flow. We have also produced H theorems specific to more complicated systems of current interest (Carnevale 1979). These include MHD turbulence in both three dimensions (Pouquet *et al* 1976) and two dimensions (Pouquet 1978), a two-layer model of stratified flow (Salmon 1978a), and a model of flow over irregular topography (Holloway 1978). All these specific cases are covered by the generalised formalism of §§ 3, 4 and 5. Section 3 establishes an abstract notation and some general statistical mechanical results. It also provides us with a generalisation of Boltzmann's entropy as the information content of the second-order correlation matrix. In § 4 we first provide a derivation of the second-order Markovian closure equation in its most general form, and then deduce from this a generalised H theorem. Finally § 5 emphasises the fundamental role played by Liouville's theorem in the relaxation process.

Relaxation to equilibrium has been demonstrated in numerical simulation of inviscid two- and three-dimensional flow (Fox and Orszag 1973, Seyler *et al* 1975, Orszag 1977), and it is a well known aspect of second-order Markovian closure that the canonical equilibrium spectrum is a stationary solution (Orszag 1977). H theorems have been demonstrated previously for inviscid flow by Cook (1974) and Montgomery (1976); however, we believe that our presentation is the first comprehensive analytical study of this general aspect of second-order Markovian closure.

2. Isotropic homogeneous flow

The second-order Markovian closure equation for incompressible, isotropic, mirror-symmetric, homogeneous, d-dimensional flow is

$$\frac{\partial U_k}{\partial t}=\frac{4}{(d-1)^2}\int \mathrm{d}^d p\ \mathrm{d}^d q \delta^d(\boldsymbol{k}+\boldsymbol{p}+\boldsymbol{q})k^2\theta(t)_{kpq}\left(a^{(d)}_{kpq}U_pU_q-b^{(d)}_{kpq}U_qU_k\right)-2\nu k^2U_k \tag{2.1}$$

(cf Fournier and Frisch 1978, Rose and Sulem 1978)†. U_k is the modal energy defined by

$$U_k(t) = \int \frac{d^d r}{(2\pi)^d} \exp(-i\boldsymbol{k}\cdot\boldsymbol{r})\langle \boldsymbol{v}(\boldsymbol{x}+\boldsymbol{r}, t)\cdot\boldsymbol{v}(\boldsymbol{x}, t)\rangle \tag{2.2}$$

where the angular brackets imply an ensemble average. ν is the kinematic viscosity, and d is the spatial dimension. The $a^{(d)}_{kpq}$ and $b^{(d)}_{kpq}$ coefficients with the restriction $\boldsymbol{k}+\boldsymbol{p}+\boldsymbol{q}=0$ can be explicitly written as

$$a^{(d)}_{kpq} = \frac{1}{4k^2}\left(\frac{\sin\alpha}{k}\right)^2 \left[(p^2-q^2)^2+(d-2)k^2(p^2+q^2)\right] \tag{2.3}$$

$$b^{(d)}_{kpq} = \frac{1}{2k^2}\left(\frac{\sin\alpha}{k}\right)^2 \left[(p^2-q^2)(k^2-q^2)+(d-2)k^2p^2\right] \tag{2.4}$$

with α the angle between wavevectors $\boldsymbol{p}$ and $\boldsymbol{q}$.

The array $\theta(t)_{kpq}$ is called the triad relaxation time. Physically it may be thought of as the decay time for triple correlations, $\langle v_k v_p v_q\rangle$, due solely to the effects of third and higher-order cumulants. The prescription for obtaining the value of $\theta(t)_{kpq}$ depends on the particular procedure used in deriving (2.1) and additional modelling assumptions. For our purpose we need not choose a specific expression for $\theta(t)_{kpq}$, and we leave it arbitrary except for two restrictions on its form.

The first restriction is that θ_{kpq} remains unchanged under permutation of $\boldsymbol{k}$, $\boldsymbol{p}$ and $\boldsymbol{q}$‡. This symmetry condition ensures that equation (2.1) conserves the appropriate quadratic invariants of the inviscid Navier–Stokes equation. For $d=2$ there are two such invariants: the total energy, $E_{\mathrm{T}}=\frac{1}{2}\int d^2k U_k$, and the total enstrophy, $Z_{\mathrm{T}}=\frac{1}{2}\int d^2k\, k^2 U_k$. For $d>2$ with the assumption of isotropy there is only one non-trivial invariant, the total energy E_{T}. Furthermore, the symmetry of θ_{kpq} ensures that the absolute equilibrium spectrum predicted by statistical mechanics is a stationary solution of (2.1). Calculated from either a canonical or microcanonical ensemble (Kraichnan 1975, Basdevant and Sadourny 1975, Salmon *et al* 1976) the absolute equilibrium spectrum (assuming an ultraviolet cut-off) is

$$U_k = (a+bk^2)^{-1} \tag{2.5}$$

with a and b determined by E_{T} and Z_{T} for $d=2$, and with $b\equiv 0$ and a determined by E_{T} for $d>2$. The methods of equilibrium statistical mechanics which produce this result are discussed in § 3. Here we have stipulated an ultraviolet cut-off in order to avoid an ultraviolet catastrophe (i.e. infinite total energy for finite a and b). We shall have more to say about this cut-off later in this section.

The second restriction on θ_{kpq} is that it must be non-negative. For $d\geqslant 2$ this condition is sufficient to ensure that the energy spectrum predicted by (2.1) is non-negative for all wavenumbers at all times (Rose and Sulem 1978). For analytic extensions of (2.1) to $d<2$ even the non-negativity of θ_{kpq} fails to ensure realisable

† In Fournier and Frisch (1978), the coefficient C_d should be $C_d = 4\, S_{d-1}/(d-1)^2 S_d$. In Rose and Sulem (1978), the coefficient C_d should be $C_d = 8\, S_{d-1}/(d-1)^2\, S_d$.

‡ The direct interaction (Kraichnan 1959) derivation of (2.1) produces a $\theta(t)_{kpq}$ which involves an integral over the history of the spectrum and which is not symmetric in k, p, and q. However, near equilibrium one can invoke the fluctuation–dissipation theorem, and this leads to a symmetric expression for the direct interaction $\theta(t)_{kpq}$ (cf Kraichnan 1959).

results; thus we shall not consider (2.1) to be valid for $d<2$ and our discussion will only be for $d \geqslant 2$ (cf Fournier and Frisch 1978).

In the next section we shall give a full discussion of the general prescription for obtaining the entropy as a functional of the spectrum. For the moment we shall simply introduce the functional

$$S[U]=\frac{1}{2} \int \mathrm{d}^d k \ln U_k \tag{2.6}$$

as the functional which in this problem plays the role analogous to the Boltzmann entropy. This functional, first used by Edwards and McComb (1969), yields the same value for entropy in equilibrium as given by the canonical Gibbs prescription (cf § 3). We note that subject to the constraint of prescribed total energy (and enstrophy for $d=2$) this S is maximal for the equilibrium spectrum (2.5), as can be checked by using the method of Lagrange multipliers (cf § 3).

To calculate $\mathrm{d}S/\mathrm{d}t$ from (2.1) we need to assume that U_k^{-1} is finite. Physically this assumption implies that all modes have some energy (which may be arbitrarily small). Then using (2.1) with (2.3) and (2.4) we obtain

$$\begin{aligned} \frac{\mathrm{d}S}{\mathrm{d}t}=\int \mathrm{d}\, \tau_{kpq}^{(d)} \bigg[& \left(\frac{(p^2-q^2)^2}{U_k^2}+\frac{2(p^2-q^2)(q^2-k^2)}{U_k U_p}\right) \\ & +(d-2)\left(\frac{k^2(p^2+q^2)}{U_k^2}-\frac{2k^2p^2}{U_k U_p}\right)\bigg]-\int \mathrm{d}^d k\, \nu k^2 \end{aligned} \tag{2.7}$$

with

$$\mathrm{d}\tau_{kpq}^{(d)} \equiv \frac{1}{2(d-1)^2} \mathrm{d}^d k\, \mathrm{d}^d p\, \mathrm{d}^d q\, \delta^{(d)}(\boldsymbol{k}+\boldsymbol{p}+\boldsymbol{q}) \theta_{kpq}\left(\frac{\sin \alpha}{k}\right)^2 U_k U_p U_q .$$

Everywhere in our calculations we assume an ultraviolet cut-off (i.e. all integrals are cut off at a given maximum wavevector magnitude, $k_{\max}$). This insures that the integral $\int \mathrm{d}^d k \nu k^2$, is finite. This also avoids problems which would otherwise develop in considering the inviscid problem (i.e. $\nu=0$). Specifically, in the inviscid limit without ultraviolet cut-off and with an initial energy spectrum for which all moments are defined i.e.

$$\int k^n U_k\, \mathrm{d}^d k<\infty \qquad n \geqslant 0$$

singularities (i.e. diverging moments of the spectrum) may develop after a finite time (Fournier and Frisch 1978, Rose and Sulem 1978). This would spoil the convergence of the integral $\int \mathrm{d}\tau_{kpq}^{(d)} \ldots$ in (2.7). Assuming an ultraviolet cut-off ensures that the manipulations to follow remain valid at all times t.

We note that according to the law of sines and the symmetry of θ_{kpq}, the expression $\mathrm{d}\tau_{kpq}^{(d)}$ is fully symmetric under permutation of $\boldsymbol{k}$, $\boldsymbol{p}$ and $\boldsymbol{q}$. Thus by interchanging the integration variables in (2.7) and summing the resulting expressions we obtain

$$\begin{aligned} \frac{\mathrm{d}S}{\mathrm{d}t}=\int \mathrm{d}^d\, \tau_{kpq}^{(d)} \bigg[& \frac{1}{3}\left(\frac{p^2-q^2}{U_k}+\frac{q^2-k^2}{U_p}+\frac{k^2-p^2}{U_q}\right)^2 \\ & +(d-2)k^2p^2\left(\frac{1}{U_k}-\frac{1}{U_p}\right)^2\bigg]-\int \mathrm{d}^d k \nu k^2 . \end{aligned} \tag{2.8}$$

With $\nu=0$, $\mathrm{d}S/\mathrm{d}t$ is manifestly non-negative. The only analytic solution to $\mathrm{d}S/\mathrm{d}t=0$ is the equilibrium spectrum (2.5). The difference in the equilibrium spectrum for $d=2$ and $d>2$ is reproduced by the vanishing of the term proportional to $d-2$ in (2.8) for $d=2$.

Thus second-order Markovian closure with ultraviolet cut-off implies that for $\nu=0$ the functional S increases monotonically to its stationary value given by canonical equilibrium—the desired H theorem.

In numerical simulations of two-dimensional, viscous, unforced turbulence with a given ultraviolet cut-off $k_{\max}$ and given initial energy spectrum, two qualitatively different behaviours are observed depending on the strength of the viscosity (Deem and Zabusky 1971, Cook and Taylor 1972, Montgomery 1972, Taylor 1972, Fox and Orszag 1973). If the viscosity is 'sufficiently strong', then a damped turbulence state develops with inertial and dissipation ranges. If the viscosity is 'sufficiently small', then a quasi-equilibrium spectrum develops. Specifically, Fox and Orszag (1973) state that if the viscous time scale is large compared with the eddy turnover time for *all* retained scales (i.e. for all $k \leqslant k_{\max}$), then approach to equilibrium can be expected. We should be able to relate this, at least qualitatively, to the rate of change of the functional S. For a given spectrum the rate of change of total energy and enstrophy are proportional to the viscosity. If the viscosity is sufficiently small so that the integral $\int \mathrm{d}^d k \nu k^2$ in (2.7) is negligible initially, then an increasing entropy implies that the spectrum evolves toward an equilibrium form determined by the slowly varying values of total energy and enstrophy. Of course, the development of a quasi-equilibrium range in this case is artificial and not physically relevant because of the imposition of a $k_{\max}$ which loses the important physics of the dissipation range (Fox and Orszag 1973). On the other hand, the problem of flow over irregular topography (Bretherton and Haidvogel 1976, Herring 1977, Holloway 1978) perhaps represents a system for which this concept of entropy increase may be relevant even for dissipative dynamics. The equilibrium state for this system is characterised by strong correlations between flow and topography. Simulations show that flows initially uncorrelated with topography rapidly (i.e. on the order of an eddy turnover time) develop such correlations with topography, and this does not appear to be an artifact of spectral truncation (Holloway 1978). We might interpret this initial rapid development of correlations in terms of a rapidly increasing entropy functional.

There are earlier references in the literature to the H theorem in the context of two-dimensional flow. Cook (1974) outlines the derivation of the H theorem by making reference to the work of Taylor and Thompson (1973). Newell and Aucoin (1971) suggested that an H theorem for the two-dimensional Rossby wave problem should follow by analogy to their derivation of an H theorem from the Boltzmann equation for 'classical phonons.'. This analogy is not strict; unlike the phonon problem the dynamics of Rossby wave interactions do not admit a Hamiltonian description and the interaction between Rossby waves cannot be represented by a collision cross section. However, for $d=2$ equation (2.1) is valid for anisotropic as well as isotropic spectra; in fact, it is identical in structure to the closure equation for Rossby waves on a β plane with the only difference occurring in the actual values of $\theta(t)_{kpq}$ (Holloway and Hendershott 1977). Hence, the inviscid Rossby wave problem also satisfies the H theorem with the same final equilibrium state (2.5), and only the rate of entropy increase is affected by the inclusion of wave propagation.

A different approach to the question of an H theorem for fluids is presented by Montgomery (1976). He investigates the behaviour of the probability distribution for velocity amplitudes as described by a BBGKY hierarchy. By making the same sort of approximations as are used to derive the closure equation (2.1), he obtains an equation for the marginal probability distribution of the Fourier amplitude of the velocity field. If y_i represents the real or imaginary part of the Fourier amplitude of velocity v_k and if $f_i(y_i)$ represents the marginal univariate probability distribution, then the direct analogue of the classical Boltzmann entropy is

$$S_{\mathrm{B}} = -\sum_i \int \mathrm{d}y_i\, f_i(y_i) \ln f_i(y_i). \tag{2.9}$$

Montgomery shows that this functional (2.9) satisfies an H theorem and infers that the H theorem in terms of the energy spectrum must hold. Our work is complementary to this in that we demonstrate an H theorem entirely in terms of the spectral equations. Note that when $f_i(y_i)$ is a Gaussian distribution with correlation given by U_k, (2.6) and (2.9) become identical up to an additive constant.

3. General statistical mechanics results

Before proceeding to the question of H theorems for more complicated systems, we need to establish a representation sufficiently general to apply to systems involving multiple species of fields (e.g. velocity, topography, magnetic, etc) and suitable for application of established statistical techniques. We assume a representation in which the dynamics are given in terms of a discretely indexed set of real independent variables, $\{y_i\}$. By independent we intend that there should be no holonomic constraints (i.e. diagnostic relations) on the variables y_i. For example, with periodic boundary conditions we would use the real and imaginary parts of the Cartesian components of the Fourier amplitudes $v_j(\boldsymbol{k})$, and eliminate redundant fields using the incompressibility relations, $\boldsymbol{k}\cdot\boldsymbol{v}(\boldsymbol{k})=0$, and the hermiticity constraint, $\boldsymbol{v}(\boldsymbol{k})=\boldsymbol{v}^*(-\boldsymbol{k})$. The discrete indexing is convenient for this section, but this condition can be relaxed afterwards.

The dynamical independent variables $\{y_i\}$ form a phase space. The motion of a point through this phase space represents the evolution of a particular system. The motion is governed by the primitive fluid field equations, $\dot{y}_i = \Omega_i(\{y\})$, where $\Omega_i(\{y\})$ is a quadratically nonlinear function of the field variables.

We introduce statistics by assuming an ensemble of systems with a distribution of initial conditions. $P(\{y\}, t)$ gives the distribution of phase space points as each system evolves independently. The quantities of interest to us are the statistical moments defined by the average

$$Y_{ij\ldots z} = \langle y_i y_j \ldots y_z \rangle \equiv \int (y_i y_j \ldots y_z) P(\{y\}, t) \prod_k \mathrm{d}y_k.$$

The distribution $P(\{y\}, t)$, evolves according to the continuity equation (cf Tolman 1938). Non-dissipative (e.g. inviscid) fluid dynamics satisfies the condition (Lee 1952, Leith 1971)

$$\sum_i \frac{\partial}{\partial y_i} \dot{y}_i = 0. \tag{3.1}$$

Hence the Liouville theorem

$$\frac{\partial P}{\partial t} + \sum_i \dot{y}_i \frac{\partial P}{\partial y_i} = 0 \tag{3.2}$$

holds; that is, the 'flow' of a distribution of points in phase space is incompressible. It follows from Liouville's theorem that if there are dynamical invariants of the primitive field equation, then any function of these invariants is a stationary probability distribution (cf Tolman 1938).

The absolute equilibrium distribution contains only the information given by prescribed values of the dynamical invariants. The invariants of concern in most fluid problems are quadratic, that is, they can be written in the form

$$I^{(l)} = I^{(l)}_{12} y_1 y_2. \tag{3.3}$$

Here and throughout we adopt the notational convention in which repeated numerical subscripts are to be taken as indices which are summed over unless otherwise stated. We call $I^{(l)}_{ij}$ the invariant kernel; the superscript (l) refers to the particular invariant considered (e.g. the total energy or helicity in three-dimensional flow). With no loss of generality we assume $I^{(l)}_{ij} = I^{(l)}_{ji}$. Invariants of other than quadratic order are of course possible (e.g. the spatial integral of any power of the vorticity in two-dimensional flow). However in most fluid problems higher-order invariants are not used in specifying the equilibrium ensemble for various reasons (e.g. they fail to survive as invariants in a spectral representation). We restrict ourselves to consideration of only quadratic invariants; for further discussion on this matter see Salmon *et al* (1976), Fyfe *et al* (1977), Kraichnan (1973). Fournier and Frisch (1978) and Thompson (1974).

If the initial ensemble is such that all systems have the same prescribed values of the dynamical invariants, then the appropriate equilibrium distribution is the microcanonical distribution (Tolman 1938). However, calculating moments with the microcanonical distribution is quite cumbersome. With invariants which involve only monomials of low order in the fields, it can be shown (Salmon *et al* 1976) that calculation of moments of low order compared with the total number of degrees of freedom is equivalent and far more convenient with the canonical equilibrium distribution specified by

$$\langle I^{(l)} \rangle = E^l. \tag{3.4}$$

This is suitable for our purposes since the closure equation is defined in terms of the average second-order moments. The canonical equilibrium distribution (cf Tolman 1938) is

$$P(\{y\}) = N \exp\left(-\sum_l \beta^{(l)} I^{(l)}\right) \tag{3.5}$$

where N is a normalisation constant. The coefficients $\beta^{(l)}$ are generalisations of the inverse temperature of particle statistical mechanics and are determined by the conditions (3.4). If we define the matrix I by

$$I_{ij} \equiv \sum_l \beta^{(l)} I^{(l)}_{ij} \tag{3.6}$$

then the canonical distribution with unit normalisation (cf Leslie 1973, Orszag 1977) has $N = |I|^{1/2}(\pi)^{-n/2}$ where $|I|$ is the determinant of the matrix I_{ij}. For this distribution to be valid, I must be positive definite. This condition is assured if the values of the $E^{(l)}$, which determine the $\beta^{(l)}$, are physically realisable. The equilibrium second-order correlation matrix, calculated from (3.5), is (cf Fisz 1963)

$$Y_{ij}^{\mathrm{eq}} = \tfrac{1}{2} I_{ij}^{-1}. \tag{3.7}$$

Note that Y_{ij}^{eq} is a positive definite matrix and its inverse is an invariant kernel.

Second-order Markovian closures attempt to determine the evolution of the second-order correlations Y_{ij}. For simplicity we assume throughout that the average field vanishes identically. We wish to obtain a general prescription for the entropy as a functional of the Y_{ij}. Although it is possible to diagonalise the symmetric matrix Y_{ij}, it will be important not to do so when considering the general dynamical problem. This is because the diagonalisation transformation will not in general be independent of time. For example, in the barotropic flow over topography problem the correlation matrix involving the stream function ψ and topography h may initially have no off-diagonal elements $\langle \psi h \rangle$ but these will develop, and there is no way of finding a purely diagonal representation which will remain so. As another example, for isotropic, three-dimensional flow a diagonal representation is possible at all times, however, this is not so if the condition of isotropy is relaxed and a non-vanishing amount of helicity is admitted.

To obtain a general prescription for the entropy functional, we take an approach based on information theory. The functional

$$H[\mathrm{P}] \equiv \int P(\{y\}) \ln P(\{y\}) \prod_i \mathrm{d}y_i \tag{3.8}$$

is a measure of sharpness of the distribution P and can be used as the information content of the distribution (cf Everett 1973). If we are given the values of the moments Y_{ij} (with $Y_i \equiv 0$) and use these as constraints on the possible form of P, then the minimum value of $H[P]$ gives the information contained in just the knowledge of the moments Y_{ij}. By using the method of Lagrange multipliers (Salmon 1978b), we obtain the distribution, $\hat{P}$, which minimises $H[P]$ subject to the given constraints. This distribution, $\hat{P}$, is the multivariate Gaussian with second-order correlations, Y_{ij} (and vanishing average field, $Y_i \equiv 0$). By substituting $\hat{P}$ into the definition (3.8) we obtain

$$H[\hat{P}] = -\tfrac{1}{2} \ln |Y| - \tfrac{1}{2} n\ (\ln 2\pi + 1). \tag{3.9}$$

Y denotes the second-order correlation matrix and n is the total number of degrees of freedom. $H[\hat{P}]$ is the information given by a knowledge of the Y_{ij}. Heuristically we consider entropy to be the lack of information and thus write

$$S[Y] = -H[\hat{P}] = \tfrac{1}{2} \ln |Y| + \tfrac{1}{2} n\ (\ln 2\pi + 1). \tag{3.10}$$

When investigating the evolution of this entropy for a given system we may for convenience drop the constant term, $\tfrac{1}{2}n\ (\ln 2\pi + 1)$. Note that since the determinant $|Y|$ is invariant under similarity transformation, the functional $S[Y]$ is independent of choice of orthonormal basis for the field variables $\{y\}$ (cf Byron and Fuller 1969).

As a consistency check we demonstrate that the functional $S[Y]$ is maximal for Y_{ij} given by equilibrium statistical mechanics. To do this we perform the variational calculation

$$\frac{\delta}{\delta Y_{ij}} \Big(S[Y] - \sum_l \alpha^{(l)} I_{12}^{(l)} Y_{12} \Big) = 0. \tag{3.11}$$

The $\alpha^{(l)}$ are Lagrange multipliers to be adjusted to give the prescribed values $E^{(l)}$ of the invariants. Using the definitions of $|Y|$ and Y_{ij}^{-1} in terms of cofactors (Byron and Fuller

1969), and the symmetry of Y_{ij} we obtain

$$\frac{\delta}{\delta Y_{ij}} S[Y] = \tfrac{1}{2} Y_{ij}^{-1}. \tag{3.12}$$

Hence, $S[Y]$ is maximal for

$$Y_{ij}^{-1} = 2 \sum_{l} \alpha^{(l)} I_{ij}^{(l)}.$$

As the $\alpha^{(l)}$ are to be determined by the same conditions (3.4) that determine the $\beta^{(l)}$, which define the equilibrium distribution, we regain the equilibrium result (3.7). Thus the functional $S[Y]$ has the property of entropy of achieving its maximum value only in equilibrium.

Next we demonstrate that the functional $S[Y]$ obeys a generalised H theorem of the Gibbs type. The Gibbs type H theorem is an exact dynamical statement about the decrease of information contained in a 'smoothed' probability distribution (cf Tolman 1938). We assume, as before, that mean fields vanish identically.

Let $P(\{y\}, t)$ be the exact probability distribution determined from the initial conditions and Liouville's theorem. It can then be shown by using Liouville's theorem that the information $H[P]$ determined from the exact distribution does not change with time (cf Tolman 1938). That is,

$$\int P(t) \ln P(t) \prod_i \mathrm{d}y_i = \int P(0) \ln P(0) \prod_i \mathrm{d}y_i$$

where $P(t) \equiv P(\{y\}, t)$.

Let $\hat{P}(\{y\}, t)$ be the 'smoothed' distribution defined as a Gaussian distribution with correlations prescribed by the exact dynamics and zero mean fields; that is

$$\int y_i y_j \hat{P}(t) \prod_k \mathrm{d}y_k \equiv \int y_i y_j P(t) \prod_k \mathrm{d}y_k \equiv Y_{ij}(t).$$

Since $\hat{P}(t)$ is Gaussian its information content is simply the negative of our entropy functional given with the exact correlations $Y_{ij}(t)$. Furthermore, as discussed above the information content of $\hat{P}(t)$ must be less than or equal to the information content of $P(t)$, that is

$$\int \hat{P}(t) \ln \hat{P}(t) \prod_i \mathrm{d}y_i \leqslant \int P(t) \ln P(t) \prod_i \mathrm{d}y_i.$$

If we assume Gaussian initial conditions, then $P(0) = \hat{P}(0)$. Making this assumption we then have

$$\begin{aligned} S(t) - S(0) &\equiv \tfrac{1}{2} \ln |Y(t)| - \tfrac{1}{2} \ln |Y(0)| \\ &= -\int (\hat{P}(t) \ln \hat{P}(t) - \hat{P}(0) \ln \hat{P}(0)) \prod_k \mathrm{d}y_k \\ &= -\int (\hat{P}(t) \ln \hat{P}(t) - P(0) \ln P(0)) \prod_k \mathrm{d}y_k \\ &= -\int (\hat{P}(t) \ln \hat{P}(t) - P(t) \ln P(t)) \prod_k \mathrm{d}y_k \\ &\geqslant 0. \end{aligned} \tag{3.13}$$

That is, for both positive and negative time, $S(0)$ is the minimum value of $S(t)$ determined by the exact dynamics. We emphasise that this is an exact result of Liouville's theorem. The information content of the smoothed distribution $\hat{P}(t)$ cannot increase in time. This same result can also be demonstrated for the Boltzmann entropy (2.9) used by Montgomery (1976) (Carnevale 1979).

If $P(0)$ is the equilibrium distribution, then $S(t) = S(0)$ for all time. Otherwise, although the equality holds initially because $P(0)$ is defined Gaussian, we expect the inequality to develop as the interactions in phase space cause the moments of $P(t)$ to deviate from the initial Gaussian moments.

The Gibbs-type H theorem does not give a quantitative evaluation of the rate of change of $S(t)$. For that we need to know how the second-order correlations change in time; this is provided by the second-order Markovian closure. As the closure equation involves only the instantaneous values of the second-order correlations, we should expect it to reflect that the information given by just the second-order correlations degrades with time. That is, we expect second-order Markovian closure to be consistent with $S(t)$ increasing with time.

The 'experimental' entropy for a system in equilibrium is defined through macroscopic parameters, and its value can be obtained through the Gibbs prescription

$$S_G = -\int P(\{y\}) \ln P(\{y\}) \prod_j \mathrm{d}y_j \tag{3.14}$$

with $P(\{y\})$ defined as the canonical distribution consistent with the macroscopic parameters (cf Jaynes 1965). We note that under our assumption that all the dynamical invariants are quadratic the canonical distribution is multivariate Gaussian; and hence S_G (3.14), the Boltzmann prescription S_B (2.9), and our functional $S[Y]$ (3.10) all give the same value for the entropy of the system in equilibrium.

4. Generalised formalism and H theorem

The primitive field equations for incompressible fluids involve a quadratic nonlinearity and may be written as

$$\dot{y}_1 = A_{123} y_2 y_3 + L_{12} y_2. \tag{4.1}$$

As in the previous section the y_i are the real independent dynamical variables. Both kernels A and L are real and with no loss of generality we take A_{123} to be symmetric in its last two indices. Again we denote instantaneous moments by $Y_{12\ldots z} \equiv \langle y_1 y_2 \ldots y_z \rangle$. To simplify, we assume that the average field, Y_1, vanishes identically; this implies that $A_{123} Y_{23}$ vanishes as well.

A hierarchy of moment evolution equations is obtained by using (4.1) to calculate the time derivative of successively higher-order monomials constructed from the $\{y\}$, and ensemble averaging. Due to the quadratic non-linearity, the evolution of the nth order moment depends on the $(n+1)$th order moment. Closures are attempts to reduce the infinite hierarchy to a closed set of equations (Monin and Yaglom 1971, Leslie 1973).

To emphasise possible Gaussian solutions, it proves convenient to recast the moment hierarchy in terms of cumulants, which for third and higher orders vanish for

Gaussian statistics. The second to fourth order cumulants are explicitly (assuming $Y_1 \equiv 0$)

$$Y^c_{12} = Y_{12}$$

$$Y^c_{123} = Y_{123}$$

$$Y^c_{1234} = Y_{1234} - Y_{12}Y_{34} - Y_{13}Y_{24} - Y_{14}Y_{23}.$$

The first two equations of the cumulant hierarchy are

$$\dot{Y}_{12} = A_{134}Y_{342} + A_{234}Y_{341} + L_{13}Y_{32} + L_{23}Y_{31} \tag{4.2}$$

$$\dot{Y}_{123} = 2(A_{145}Y_{24}Y_{35} + A_{245}Y_{34}Y_{15} + A_{345}Y_{14}Y_{25}) + A_{145}Y^c_{4523} + A_{245}Y^c_{4531} + A_{345}Y^c_{4512} + \mathrm{L}_{14}\mathrm{Y}_{423} + \mathrm{L}_{24}\mathrm{Y}_{431} + \mathrm{L}_{34}\mathrm{Y}_{412}. \tag{4.3}$$

As indicated in the introduction, second-order Markovian closure may be achieved in a variety of ways. The most straightforward in the present context is the EDQNM closure. This closure results from two distinct approximations or procedures—the eddy-damped quasi-normal procedure and Markovianisation.

The eddy-damped quasi-normal procedure parametrises the effect of fourth- and higher-order cumulants on third-order cumulants through eddy damping (Orszag 1970, Rose and Sulem 1978). Here we generalise this procedure by making the following replacement in equation (4.3)

$$A_{145}Y^c_{4523} + A_{245}Y^c_{4531} + A_{345}Y^c_{4512} \to -\eta_{11'22'33'}Y_{1'2'3'}. \tag{4.4}$$

We leave the 'eddy viscous' term $\eta_{11'22'33'}$ unspecified except to note that for consistency of the replacement (4.4), $\eta_{11'22'33'}$ should be symmetric under permutation of the indices in the pairs (1, 1′), (2, 2′) and (3, 3′) (i.e. $\eta_{11'22'33'} = \eta_{22'11'33'} =$ etc). With this parametrisation we can rewrite (4.2) as

$$\dot{Y}_{123} = 2C_{123} - \mu_{11'22'33'}Y_{1'2'3'}. \tag{4.5}$$

Here we have introduced the following shorthand definitions

$$C_{123} \equiv A_{145}Y_{42}Y_{53} + A_{245}Y_{43}Y_{51} + A_{345}Y_{41}Y_{52}$$

and

$$\mu_{11'22'33'} \equiv \lambda_{11'22'33'} + \eta_{11'22'33'}$$

with

$$\lambda_{11'22'33'} \equiv -L_{11'}\delta_{22'}\delta_{33'} - \delta_{11'}L_{22'}\delta_{33'} - \delta_{11'}\delta_{22'}L_{33'}.$$

Note that by definition C_{123} is symmetric under permutation of 1, 2 and 3.

The formal solution to (4.5) is

$$Y(t)_{123} = 2\int_0^\infty \mathrm{d}t'\, G(t,t')_{11'22'33'}C(t')_{1'2'3'} \tag{4.6}$$

with $G(t, t')_{11'22'33'}$ the retarded Green function. The procedure of Markovianisation may be viewed as an approximation in the following sense. The Green function will naturally be peaked for small values of the separation $(t-t')$. Thus it is reasonable to Taylor-expand $C(t')_{1'2'3'}$ in the integrand of (4.6) about time t. Markovianisation consists of dropping all terms in the expansion of order $(t-t')G(t,t')$ leaving

$$Y(t)_{123}=2\theta(t)_{11'22'33'}C(t)_{1'2'3'} \tag{4.7}$$

with θ defined by

$$\theta(t)_{11'22'33'}\equiv\int_0^\infty G(t,t')_{11'22'33'}\,\mathrm{d}t'. \tag{4.8}$$

$\theta_{11'22'33'}$ is the generalisation of the triad relaxation time (Kraichnan 1972); according to this construction, it is symmetric with respect to permutations of the pairs of indices $(1,1')$, $(2,2')$ and $(3,3')$. Using result (4.7) in equation (4.2) results in an evolution equation for second-order correlations involving only their instantaneous values. This generalised EDQNM equation is

$$\dot{Y}_{11'}-L_{12}Y_{21'}-L_{1'2}Y_{21}=2\theta_{1\bar{1}2\bar{2}3\bar{3}}A_{1'23}C_{\bar{1}\bar{2}\bar{3}}+2\theta_{1'\bar{1}2\bar{2}3\bar{3}}A_{123}C_{\bar{1}\bar{2}\bar{3}}. \tag{4.9}$$

A necessary condition for the validity of (4.9) is that it must conserve in the ensemble average all the quadratic invariants of the primitive equation, (4.1). By using (4.1) to compute the rate of change of a quadratic form $I^{(l)}=I^{(l)}_{12}y_1y_2$, and using the independence of the variables $\{y\}$, we find that an invariant kernel must satisfy the following conditions

$$I^{(l)}_{12}L_{23}+I^{(l)}_{32}L_{21}=0 \tag{4.10}$$

and

$$I^{(l)}_{14}A_{423}+I^{(l)}_{24}A_{431}+I^{(l)}_{34}A_{412}=0. \tag{4.11}$$

In the ordinary sense of non-dissipative (e.g. inviscid) dynamics, we understand that the linear kernel L_{ij} is restricted such that (4.10) is satisfied by the solutions of (4.11) (e.g. the kernel for the total energy satisfies (4.11) and (4.10) when viscosity vanishes). By using the pairwise symmetry of $\theta_{11'22'33'}$, it can be shown that according to (4.9) the temporal derivative of $\langle I^{(l)}\rangle\equiv I^{(l)}_{12}Y_{12}$ vanishes for any kernel $I^{(l)}_{ij}$ which satisfies (4.10) and (4.11). Thus, the generalised EDQNM equation conserves all the quadratic invariants of the primitive equation.

Similarly the pairwise symmetry of $\theta_{11'22'33'}$ ensures that the canonical equilibrium correlation, Y^{eq}_{ij}, is a stationary solution of (4.9). This can be seen by substituting $Y^{\mathrm{eq}}_{ij}\equiv\frac{1}{2}I^{-1}_{ij}$ into (4.9) and noting that since I_{ij} is an invariant kernel it satisfies conditions (4.10) and (4.11).

In order to ensure that the functional $S=\frac{1}{2}\ln|Y|$ and its derivative

$$\frac{\mathrm{d}S}{\mathrm{d}t}=\tfrac{1}{2}Y^{-1}_{11'}\dot{Y}_{11'} \tag{4.12}$$

are well defined, we need to assume that the correlation matrix Y_{ij} is non-singular and hence invertible. Since Y_{ij} is by definition positive semi-definite, this additional assumption implies that it is positive definite.

Using (4.9) we obtain

$$\frac{\mathrm{d}S}{\mathrm{d}t}=2\theta_{1\bar{1}2\bar{2}3\bar{3}}A_{1'23}C_{\bar{1}\bar{2}\bar{3}}Y^{-1}_{11'}+L_{11}. \tag{4.13}$$

Then using the pairwise symmetry of $\theta_{11'22'33'}$ and applying some algebraic manipulation we obtain

$$\frac{\mathrm{d}S}{\mathrm{d}t}=\tfrac{2}{3}D_{123}Y_{1\bar{1}}Y_{2\bar{2}}Y_{3\bar{3}}\theta_{\bar{1}1'\bar{2}2'\bar{3}3'}D_{1'2'3'}+L_{11} \tag{4.14}$$

with D_{123} defined by

$$D_{123}\equiv Y_{14}^{-1}A_{423}+Y_{24}^{-1}A_{431}+Y_{34}^{-1}A_{412}. \tag{4.15}$$

The term L_{11}, the trace of the linear operator, vanishes if we assume that Liouville's theorem holds. That is $\delta\dot{y}_1/\delta y_1\equiv 0$ implies that $A_{113}=0$ and $L_{11}=0$. Thus we have

$$\frac{\mathrm{d}S}{\mathrm{d}t}=\tfrac{2}{3}D_n\chi_{n\bar{n}}\theta_{\bar{n}n'}D_{n'} \tag{4.16}$$

where, for convenience, we have adopted the shorthand notation $n=(1,2,3)$ for the indices and introduced the definition

$$\chi_{nn'}\equiv Y_{11'}Y_{22'}Y_{33'}.$$

Note that since $\chi_{nn'}$ is the Kronecker product of positive definite matrices, it is also positive definite (cf Marcus and Ming 1964).

Notice that for non-dissipative dynamics, D_{123} vanishes if and only if Y_{ij}^{-1} is an invariant kernel (cf equation (4.11)). Furthermore, since the generalised EDQNM equation conserves all quadratic invariants, it follows that D_{123} evolving according to the EDQNM equation can vanish if and only if Y_{ij} evolves to the canonical equilibrium matrix, $Y_{ij}^{\mathrm{eq}}=\frac{1}{2}I_{ij}^{-1}$, as defined in § 3.

Thus if we can demonstrate that the product $\chi_{n\bar{n}}\theta_{\bar{n}n'}$ appearing in (4.16) is positive definite, then S is stationary only in equilibrium and is otherwise monotonically increasing in time. Notice also that if the product $\chi\theta$ is positive semi-definite then the determinant, $|Y|$, cannot vanish since $\ln|Y|$ is non-decreasing in time. Thus the correlation matrix Y_{ij} with all positive eigenvalues cannot develop zero eigenvalues and hence must remain positive definite (i.e. realisable).

In all the particular problems mentioned in the introduction (i.e. MHD turbulence, two-layer flow and flow over irregular topography), for which H theorems have been demonstrated (Carnevale 1979), the fields are assumed homogeneous. Under this assumption the correlation matrix has block diagonal form. If we denote the interblock index by k (i.e. wavevector or other spectral index) and the intrablock index by α (e.g. α discriminates between different species of fields), then Y_{12} can be written as

$$Y_{12}=Y_{k_1k_2}^{\alpha_1\alpha_2}=Y_{\alpha_1\alpha_2}^{(k_1)}\delta_{k_1,k_2}\qquad\text{(no summation).} \tag{4.17}$$

Furthermore it is assumed that $\theta_{nn'}$ is positive definite, and that it does not mix intrablock elements. That is

$$\theta_{11'22'33'}\equiv\theta_{k_1k_2k_3}\delta_{k_1,k_1'}\delta_{k_2,k_2'}\delta_{k_3,k_3'}\delta_{\alpha_1,\alpha_1'}\delta_{\alpha_2,\alpha_2'}\delta_{\alpha_3,\alpha_3'},\qquad \theta_{k_1k_2k_3}>0. \tag{4.18}$$

These assumptions imply that $\chi_{nn'}\equiv Y_{11'}Y_{22'}Y_{33'}$ and $\theta_{nn'}$ are symmetric in n and n', and that χ and θ commute (i.e. $\omega_{n\bar{n}}\theta_{\bar{n}n'}=\theta_{n\bar{n}}\chi_{\bar{n}n'}$). Hence $\chi_{nn'}$ and $\theta_{nn'}$ can be simultaneously diagonalised by an orthonormal similarity transformation (cf Byron and Fuller 1969). This, together with the positive definiteness of $\chi_{nn'}$ and $\theta_{nn'}$, implies that the product

$\chi_{n\bar{n}}\theta_{\bar{n}n'}$ is indeed positive definite. Thus any homogeneous EDQNM equation which has a generalised interaction time $\theta_{11'22'33'}$ of the form (4.18) satisfies the H theorem.

It has been suggested in studies involving multiple species of fields (Salmon 1978a, Holloway 1978) that the triad relaxation time $\theta_{k_1k_2k_3}$ in (4.18) should also depend on the field species index. In that case (4.18) is replaced by $\theta_{nn'} = \theta^{(n)}\delta_{n,n'}$ with $\theta^{(n)} > 0$. If it is further assumed, as in Salmon's (1978a) equivalent-layers model of stratified flow, that the correlation matrix Y_{ij} is diagonal, then the Kronecker product, $\chi_{nn'} = Y_{11'}Y_{22'}Y_{33'}$ is also diagonal. The product of the positive definite diagonal matrices $\theta_{nn'}$ and $\chi_{nn'}$ is positive definite and hence the H theorem must hold. On the other hand, for the general case of block diagonal correlations, $Y_{11'} = Y^{(k_1)}_{\alpha_1\alpha_1'}\delta_{k_1,k_1'}$, the product is $\chi_{n\bar{n}}\theta_{\bar{n}n'} = \chi_{nn'}\theta^{(n')}$. To determine whether such a product is positive definite we form the symmetrised product $\chi_{n\bar{n}}\theta_{\bar{n}n'} + \chi_{n'\bar{n}}\theta_{\bar{n}n}$ and check whether all the principal subdeterminants are positive (Marcus and Ming 1964). The result is that for generally non-diagonal correlations $Y_{11'}$ the product is, in general, positive definite if and only if $\theta_{nn'}$ is of the more restrictive form (4.18). Thus we see that the simple diagonal model, $\theta_{nn'} = \theta^{(n)}\delta_{n,n'}$, for the relaxation time will not in general lead to an H theorem.

A positive definite relaxation time matrix $\theta_{nn'}$ assures us that there are no non-trivial solutions D_n to $\theta_{nn'}D_{n'} = 0$, and hence entropy is stationary only in equilibrium. However, since $D_n (= Y^{-1}_{14}A_{423} + Y^{-1}_{24}A_{431} + Y^{-1}_{34}A_{412})$ is not arbitrary, the condition of positive definiteness for $\theta_{nn'}$ can be relaxed in particular cases with the same result. For homogeneous systems the non-linear kernel A_{123} vanishes unless the wavevectors $\boldsymbol{k}_1$, $\boldsymbol{k}_2$ and $\boldsymbol{k}_3$ form a triad (i.e. their lengths form the sides of a triangle). This implies that with homogeneous correlations, $Y_{11'} = Y^{(k_1)}_{\alpha_1\alpha_1'}\delta_{k_1,k_1'}$, D_n also vanishes except for triads $(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$. Thus we need only restrict $\theta_{k_1k_2k_3}$ of (4.18) to be positive on triads to insure that $\mathrm{d}S/\mathrm{d}t$ vanishes only in equilibrium. Although it is usually physically reasonable to assume that the triad relaxation time $\theta_{k_1k_2k_3}$ is non-vanishing for all triads, certain limiting cases where we would wish to relax this condition come to mind. For example, in the low-amplitude limit of Rossby waves on a β plane (cf Holloway and Hendershott 1977, Longuet-Higgins and Gill 1967) it can be shown that $\theta_{k_1k_2k_3}$ vanishes unless the triad wavevectors $(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ have Rossby frequencies $\boldsymbol{\omega}_k = -\beta k_x/k^2$ which satisfy a resonance condition, $\omega_{k_1} + \omega_{k_2} + \omega_{k_3} = 0$. However, because of the structure of the nonlinear kernel A_{123} for this problem (cf Holloway and Hendershott 1977) there is still no non-trivial solution D_n to $\mathrm{d}S/\mathrm{d}t = 0$ (i.e. entropy is stationary only in equilibrium).

As a final comment on this section we note that we used the assumption that Liouville's theorem holds (i.e. $\delta\dot{y}_1/\delta y_1 \equiv 0$) simply to assert that the trace of the linear operator L_{11} vanished. Alternatively we could have just assumed L_{11} vanishes but not invoked the full Liouville's theorem, which also requires $A_{113} \equiv 0$. However if the canonical equilibrium solution exists then its inverse, which is an invariant kernel, must satisfy condition (4.11), that is

$$Y^{-1}_{14}A_{423} + Y^{-1}_{24}A_{431} + Y^{-1}_{34}A_{412} = 0.$$

This implies that

$$\mathrm{A}_{113} = -\tfrac{1}{2} Y^{-1}_{13}A_{124}Y_{24}. \tag{4.19}$$

Under the assumption that the average field vanishes identically, which implies $A_{124}Y_{24} \equiv 0$, (4.19) reduces to $A_{113} = 0$. Thus the existence of Y^{eq} (with zero average fields) implies that Liouville's theorem holds.

5. Relationship between exact dynamics and the EDQNM H theorem

We have shown in § 3 that the exact dynamics implies a Gibbs-type H theorem (i.e. $S(t)-S(0)\geqslant 0$ for any t for Gaussian initial conditions). In § 4 we demonstrated that the EDQNM closure for homogeneous fields with the usual triad relaxation time θ_{kpq} implies a Boltzmann-type H theorem (i.e. $\mathrm{d}S/\mathrm{d}t\geqslant 0$ for all positive t). Here we wish to demonstrate that the EDQNM H theorem can be derived directly from the Gibbs-type H theorem.

We use the notation of § 4, and begin with the primitive equation (4.1). Under the assumptions that the average field vanishes identically and that the initial distribution is Gaussian, we immediately obtain from (4.1) the result

$$\dot{Y}_{11'}(t)-L_{12}Y_{21'}(t)-L_{1'2}Y_{21}(t)=2tA_{1'23}C_{123}(0)+2tA_{123}C_{1'23}(0)+0(t^2) \tag{5.1}$$

with

$$C_{123}(0)\equiv A_{145}Y_{24}(0)Y_{35}(0)+A_{245}Y_{34}(0)Y_{15}(0)+A_{345}Y_{14}(0)Y_{25}(0).$$

Using this result we obtain

$$\frac{\mathrm{d}S}{\mathrm{d}t}\equiv Y^{-1}_{1\bar{1}'}(t)\dot{Y}_{11'}(t)=L_{11}+2tA_{123}C_{1'23}(0)Y^{-1}_{1\bar{1}'}(0)+0(t^2). \tag{5.2}$$

Under the assumption that Liouville's theorem holds (i.e. $\delta\dot{y}_1/\delta y_1\equiv 0$), the term L_{11} in (5.2) vanishes identically. Furthermore according to the analysis of § 3, Liouville's theorem implies that $S(t)-S(0)\geqslant 0$; therefore, $\mathrm{d}S/\mathrm{d}t|_{t=0}=0$ and $\mathrm{d}^2S/\mathrm{d}t^2|_{t=0}\geqslant 0$. Thus Liouville's theorem implies

$$2A_{123}C_{1'23}(0)Y^{-1}_{1\bar{1}'}(0)\geqslant 0 \tag{5.3}$$

for arbitrary initial correlations $Y_{ij}(0)$.

Notice that (5.3) is precisely the result we need for an H theorem for second-order Markovian closure with relaxation time $\theta_{nn'}=\delta_{nn'}$. That is $\mathrm{d}^2S/\mathrm{d}t^2|_{t=0}$ for the exact dynamics is $\mathrm{d}S/\mathrm{d}t$ for the EDQNM closure with this simple relaxation time. This result can be extended to cover the more complicated case in which we assume homogeneous fields (i.e. $Y_{11'}=Y^{(k)}_{\alpha_1\alpha_1'}\delta_{k_1,k_1'}$) with relaxation time

$$\theta_{11'22'33'}=\theta_{k_1k_2k_3}\delta_{n,n'} \tag{5.4}$$

by simply replacing A_{123} by $\sqrt{\theta_{k_1k_2k_3}}A_{123}$ and $Y_{11'}$ by $Y^{(k)}_{\alpha_1\alpha_1'}\delta_{k_1,k_1'}$ in the above equations (assuming that the modified dynamics also satisfies Liouville's theorem). Thus the H theorem for second-order Markovian closure of § 4 for homogeneous fields and relaxation time (5.4) is a direct consequence of Liouville's theorem. This is also true in the case of a species dependent triad relaxation time $\theta^{(n)}$ with purely diagonal correlations $Y_{11'}=Y^{(1)}\delta_{1,1'}$ as can be checked by replacing A_{123} by $\sqrt{\theta^{(n)}}A_{123}$ in the above equations.

6. Conclusion

The primitive field equations for fluids represent conservative dynamical systems when spectral truncation is imposed and when viscosity and other sources of linear dissipation are neglected. Furthermore, the non-dissipative equations satisfy Liouville's theorem, and canonical statistical mechanics prescribes an equilibrium state determined solely by the dynamically conserved quantities. Although there is no proof that these systems are mixing or ergodic we may reasonably expect for 'most' initial conditions relaxation toward equilibrium, and this is verified to a certain extent by numerical simulation (Patterson 1973, Fox and Orszag 1973).

On the other hand, the non-dissipative primitive field equations are time-reversal invariant. Thus we can always construct some initial state or conjure some type of classical demon to demonstrate evolution away from equilibrium. Consequently an exact statistical theory must be capable of representing evolution either toward or away from equilibrium depending on the particular circumstances and, therefore, must be incompatible with a monotonicity theorem like Boltzmann's H theorem. This must also be true for approximate treatments that maintain the time-reversal symmetry of the primitive dynamics such as the DIA (Kraichnan 1959), which can be shown to be an exact description of certain model dynamics. On the strength of Liouville's theorem we proved (cf § 3) that the exact dynamics implies a Gibbs-type H theorem (i.e. for Gaussian initial conditions $S(t) \geq S(0)$), which does not imply any particular time arrow. As explained in § 5, the Markovian closure procedure, for which $t = 0$ is not a distinguished time, changes this into an irreversible H theorem of the Boltzmann type. It is the Markovian character of closures such as the TFM (Kraichnan 1971) and the EDQNM (Orszag 1977, Rose and Sulem 1978) that results in monotonic relaxation toward equilibrium.

Unlike the case of the Boltzmann equation where the H theorem is of direct physical relevance, our H theorems for closure are generally not immediately relevant to the essentially irreversible dynamics of viscous untruncated flow (Orszag 1977, Rose and Sulem 1978). Consider a given finite wavenumber band (k_1, k_2). The H theorem implies that the nonlinear interactions *within* this band will tend to bring about absolute equilibrium. Even when viscosity is negligible (e.g. in the inertial range), interactions with neighbouring wavenumber bands will generally prevent absolute equilibrium from being established. Nevertheless, if the characteristic time for evolution to absolute equlibrium is of the same order as the time for exchanging excitation between neighbouring wavenumber bands (e.g. a local turnover time), then some features of the exact non-equilibrium dynamics may be controlled by absolute equilibrium. In evidence of this the direction in which energy, enstrophy, helicity, magnetic helicity, etc, cascade in various inertial ranges in hydrodynamic or magnetohydrodynamic flows can be predicted by absolute equilibrium calculations (Kraichnan 1967, 1973, Frisch *et al* 1975). Statistical mechanics methods have also recently found applications in explaining the changes in kinetic energy spectra in numerical weather prediction and climatic models resulting from resolution changes (Frederiksen and Sawford 1980). Although such models (e.g. Manabe *et al* 1970, Lambert and Meriless 1978) contain both forcing and dissipation the spectral change with increasing resolution is very similar to that found for inviscid models.

Acknowledgments

One of us (GFC) wishes to acknowledge the inspiration provided by Paul C Martin. We wish to thank Greg Holloway, Harvey A Rose, Jack Herring, David Montgomery, Phil Thompson, and Jorgen Frederiksen for very helpful discussions. Some of the financial support was provided by the National Science Foundation, Harvard University, and the National Center for Atmospheric Research. UF thanks the National Center for Atmospheric Research for its hospitality.

References

Basdevant C and Sadourny R 1975 *J. Fluid Mech.* **69** 673–88

Bretherton F P and Haidvogel D B 1976 *J. Fluid Mech.* **78** 129–54

Byron F W and Fuller R W 1969 *Mathematics of Classical and Quantum Systems* (Reading, Mass. Addison-Wesley)

Carnevale G F 1979 *PhD thesis* Harvard University

Carnevale G F and Martin 1981 in preparation

Cook I 1974 *J. Plasma Phys.* **12** 501–7

Cook I and Taylor J B 1972 *Phys. Rev. Lett.* **28** 82–4

Deem G S and Zabusky N J 1971 *Phys. Rev. Lett.* **27** 396–9

Edwards S F and McComb W D 1969 *J. Phys. A: Gen. Phys.* **2** 157–71

Everett H 1973 *The Theory of the Universal Wave Function* in *Many-Worlds Interpretation of Quantum Mechanics*; ed S DeWitt and N Grahm (Princeton NJ: Princeton University Press)

Fisz M 1963 *Probability Theory and Mathematical Statistics* (New York: Wiley)

Fournier J D and Frisch U 1978 *Phys. Rev.* A **17** 747–62

Fox D G and Orszag S A 1973 *Phys. Fluids* **16** 169–71

Frederiksen J S and Sawford B L 1980a *J. Atmos. Sci.* **37** 717–32

Frisch U, Pouquet A, Leorat J and Mazure A 1975 *J. Fluid Mech.* **68** 769–78

Fyfe D, Joyce G and Montgomery D 1977 *J. Plasma Phys* **17** 317–98

Hasselmann K 1966 *Rev. Geophys.* **4** 1–32

Herring J R 1977 *J. Atmos. Sci.* **34** 1731–50

Holloway G 1978 *J. Phys. Ocean* **8** 414–27

Holloway G and Hendershott M C 1977 *J. Fluid Mech.* **82** 747–65

Jaynes E T 1965 *Am. J. Phys.* **33** 391–8

Kadanoff L P and Baym G 1962 *Quantum Statistical Mechanics* (Reading, Mass: Benjamin)

Kraichnan R H 1959 *J. Fluid Mech.* **5** 497–543

—— 1965 *Phys. Fluids* **8** 575–98

—— 1967 *Phys. Fluids* **10** 1417–23

—— 1971 *J. Fluid Mech.* **47** 525–35

—— 1972 *J. Fluid Mech.* **56** 287–304

—— 1973 *J. Fluid Mech.* **59** 745–52

—— 1975 *J. Fluid Mech.* **67** 155–75

Lambert S J and Merilees 1978 *Atmosphere-Ocean* **16** 197–211

Lee T D 1952 *Quart. J. Appl. Math.* **10** 69–74

Leith C E 1971 *J. Atmos. Sci.* **28** 145–61

Leslie D C 1973 *Developments in the Theory of Turbulence* (Oxford: Clarendon)

Longuet-Higgins M S and Gill A E 1967 *Proc. R. Soc.* A **299** 120–44

Manabe S, Smagorinsky J, Holloway J L and Stone H M 1970 *Mon. Wea. Rev.* **98** 175–212

Marcus M and Ming H 1964 *A Survey of Matrix Theory and Matrix Inequalities* (Boston: Allyn and Bacon) p 70

Millikan C B 1929 *Phil. Mag.* **7** 641–62

Monin A and Yaglom A M 1971 *Statistical Fluid Mechanics: Mechanics of Turbulence* ed J L Lumley (Cambridge: MIT Press)

Montgomery D 1972 *Phys. Lett.* **39A** 7–8

—— 1976 *Phys. Fluids* **19** 802–10

Newell A C and Aucoin P 1971 *J. Fluid Mech.* **49** 593–609

Orszag S A 1970 *J. Fluid Mech.* **41** 363–86

—— 1977 *Statistical Theory of Turbulence* in *Fluid Dynamics* 1973 *Les Houches Summer School of Theoretical Physics* ed R Balian and J L Peube (New York: Gordon and Breach)

Patterson G S 1973 private communication

Pouquet A 1978 *J. Fluid Mech.* **88** 1–16

Pouquet A, Frisch U and Leorat J 1976 *J. Fluid Mech.* **77** 321–54
Rose H A and Sulem P L 1978 *J. Physique* **39** 441–84
Salmon R 1978a *Geophys. Astrophys. Fluid Dynamics* **10** 25–52
—— 1978b *Lectures* Oregon State University, unpublished
Salmon R, Holloway G and Hendershott M C 1976 *J. Fluid Mech.* **75** 691–703
Seyler C E, Salu Y, Montgomery D and Knorr D 1975 *Phys. Fluids* **18** 803–12
Taylor J B 1972 *Phys Lett* **40A** 1–2
Taylor J B and Thompson W B 1973 *Phys. Fluids* **16** 111–7
Thompson P D 1974 *J. Atmos. Sci.* **31** 1453–7
Tolman R C 1938 *The Principles of Statistical Mechanics* (Oxford: Oxford University Press)
Webb D J 1978 *Deep-Sea Research* **25** 279–298