INSTITUTE OF PHYSICS PUBLISHING NONLINEARITY

Nonlinearity **15** (2002) 239–255 PII: S0951-7715(02)14832-8

Nonequivalent statistical equilibrium ensembles and refined stability theorems for most probable flows

Richard S Ellis, Kyle Haven and Bruce Turkington

Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003, USA

Received 20 June 2000, in final form 4 December 2001
Published 17 January 2002
Online at stacks.iop.org/Non/15/239

Recommended by S Kida

Abstract

Statistical equilibrium models of coherent structures in two-dimensional and barotropic quasi-geostrophic turbulence are formulated using canonical and microcanonical ensembles, and the equivalence or nonequivalence of ensembles is investigated for these models. The main results show that models in which the energy and circulation invariants are treated microcanonically give richer families of equilibria than models in which they are treated canonically. For each model, a variational principle that characterizes its equilibrium states is derived by large deviation techniques. An analysis of the two different variational principles resulting from the canonical and microcanonical ensembles reveals that their equilibrium states coincide if and only if the microcanonical entropy function is concave. Numerical computations implemented for geostrophic turbulence over topography in a zonal channel demonstrate that nonequivalence of ensembles occurs over a wide range of the model parameters and that physically interesting equilibria are often omitted by the canonical model. The nonlinear stability of the steady mean flows corresponding to microcanonical equilibria is established by a new Lyapunov argument. These stability theorems refine the well-known Arnold stability theorems, which do not apply when the microcanonical and canonical ensembles are not equivalent.

Mathematics Subject Classification: 76E20, 76M35, 82B21

1. Introduction

In the half century since Onsager [31] explained the coalescence phenomenon of a collection of point vortices, statistical equilibrium theories have been remarkably successful in characterizing the long-lived, large-scale coherent structures in two-dimensional turbulence and barotropic quasi-geostrophic turbulence. Several different theories have been proposed: the Joyce–Montgomery theory, which applies to systems of point vortices [5, 15, 20, 21]; the Kraichnan energy–enstrophy theory, which uses only the quadratic invariants [6,17,23,35]; and the Miller–Robert theory, which enforces all the invariants of ideal fluid motion [27–29,33,34]. These theories share the common feature that they make definite predictions about the macroscopic behaviour of the fluid turbulence without requiring any detailed resolution of the underlying microscopic dynamics. As such, they provide an important methodology for constructing turbulence models self-consistently from fundamental principles. Nevertheless, each of these statistical equilibrium theories, being derived from a distinctive set of simplifying assumptions, has their own strengths and weaknesses with respect to their potential physical applications [24, 36].

In this paper, we investigate various formulations of the modern statistical equilibrium theory of coherent structures, and we conclude that one formulation has definite advantages over the others. Specifically, we contend that the natural formulation is the one based on a canonical ensemble with respect to the generalized enstrophy invariants (the higher moments of the potential vorticity) and a microcanonical ensemble with respect to the energy and circulation invariants. On the one hand, we argue that the canonical formulation of the enstrophy invariants leads to a theory that is better suited to realistic modelling of physical problems than the corresponding microcanonical formulation, and we point to several recent applications as evidence. On the other hand, we show that, with respect to the energy and circulation invariants, the theory based on the microcanonical ensemble is richer than the theory derived from the corresponding canonical ensemble, in the sense that many microcanonical equilibrium states are not realized as canonical equilibrium states. Our theoretical results suggest a unifying perspective on the equilibrium statistical theory of coherent structures in turbulence, in which a single probability distribution (the one-point statistics of the potential vorticity field) describes the unresolved small scales, and a few global invariants (the total energy and circulation) control the large scales of motion. To illustrate this approach, we implement the theory for a prototypical geophysical problem (a simple model of a band in the Jovian atmosphere), where the advantages gained by using our preferred formulation are manifestly clear.

This paper is organized as follows. In section 2 we define two different formulations of the statistical equilibrium theory for barotropic, quasi-geostrophic potential vorticity dynamics. Common to both of these models is a canonical ensemble with respect to the generalized enstrophy invariants, which furnishes the natural statistical description of the potential vorticity fluctuations. The two formulations differ in how they treat the energy and circulation invariants: one model uses a microcanonical ensemble, while the other uses a canonical ensemble. The exploration of this difference is the main theme of this paper. In section 3 we analyse the continuum limits of the probabilistic lattice models defined in section 2 and, using large deviation techniques, we derive the variational principles that characterize the most probable macrostates of the two models. We come to our main results in section 4, where we completely characterize the equivalence or nonequivalence of the two ensembles. In particular, we show that the usual correspondence between canonical and microcanonical equilibrium states breaks down unless the microcanonical entropy function is concave. In section 5 we observe that the nonlinear stability of the steady flows associated with equilibrium states follows from Arnold's Lyapunov argument if and only if those equilibrium states are realized by the canonical ensemble. In order to establish the nonlinear stability of the microcanonical equilibrium states that are omitted by the canonical ensemble, we refine Arnold's argument, invoking an idea from constrained optimization theory to construct the required Lyapunov functional. Finally, in section 6 we briefly present some computations that illustrate the general results obtained in sections 4 and 5 in the context of a particular physical problem. For this purpose we consider a simplified version of the model of the Jovian weather layer developed in [37], and we demonstrate that the nonequivalence-of-ensembles behaviour occurs throughout most of the physically relevant range of model parameters.

For the sake of brevity, we omit detailed proofs throughout this paper. The reader is referred to our other papers [3, 13, 14, 36] for complete discussions of the mathematics.

2. Formulation of the models

2.1. Underlying dynamics

Our statistical equilibrium models are based on a governing equation that includes as special cases the equations of motion for both purely two-dimensional turbulence and barotropic quasi-geostrophic turbulence. Namely, we consider the nonlinear advection equation

$$\frac{\partial Q}{\partial t} + \frac{\partial Q}{\partial x_1}\frac{\partial \psi}{\partial x_2} - \frac{\partial Q}{\partial x_2}\frac{\partial \psi}{\partial x_1} = 0, \tag{1}$$

in which $Q = Q(x_1, x_2, t)$ and $\psi = \psi(x_1, x_2, t)$ are real scalar fields related by the elliptic equation

$$Q = -\Delta\psi + r^{-2}\psi + b. \tag{2}$$

Here, Δ denotes the Laplacian on R^2, r is a given positive constant which may be infinity, and $b = b(x_1, x_2)$ is a specified continuous function. The nondivergent velocity field v is derived from the streamfunction ψ by $v = (\partial\psi/\partial x_2, -\partial\psi/\partial x_1)$. For the sake of definiteness, we consider a channel domain $\mathcal{X} = \{x = (x_1, x_2) : |x_1| < l_1/2, |x_2| < l_2/2\}$, and for boundary conditions we impose l_1-periodicity on ψ in x_1 and $\psi = 0$ at $x_2 = \pm l_2/2$.

When $r = \infty$ and $b = 0$, Q coincides with the vorticity $\omega = \partial v_2/\partial x_1 - \partial v_1/\partial x_2$, and (1), (2) reduce to the Euler equations for incompressible, inviscid flow. When a finite r and a nonvanishing b are included, Q is the potential vorticity, and (1), (2) become the barotropic, quasi-geostrophic equations, which govern a shallow rotating layer of homogeneous incompressible inviscid fluid in the limit of small Rossby number. The various terms in (2) are explained in the standard geophysical literature [32]: $-\psi$ represents the scaled free-surface perturbation; $b = \beta x_2 + h$, where β is the gradient of the Coriolis parameter $f = f(x_2)$ and h is the height of the bottom topography; $r = \sqrt{gH_0}/f_0$ is the Rossby deformation radius, determined by the gravitational acceleration g, the mean fluid depth H_0, and a mean value f_0. The same equations (1), (2) also govern the so-called $1\frac{1}{2}$-layer model, in which a shallow upper layer lies on a deep lower layer of denser fluid whose motion is unaffected by that in the upper layer. This model is often used to describe the observed weather layer of the Jovian atmosphere, which overlies a steady, zonal flow that creates an effective bottom topography [19, 26].

The equilibrium statistical models that we study are based on the hypothesis that the microscopic dynamics (1), (2) is ergodic with respect to the ideal invariants. While there is no theoretical demonstration of this ergodic hypothesis, numerous observations and simulations of two-dimensional and geostrophic turbulence show that the self-induced straining of the potential vorticity effectively randomizes Q on a range of small scales. This behaviour is related to the direct cascade of enstrophy to small scales. At the same time, Q tends to organize into coherent structures on the large scales. This dual behaviour is associated with the inverse cascade of energy to large scales. The goal of the statistical equilibrium models is to characterize the typical steady mean flows that emerge and persist on the large scales without resolving the small scales of motion.

The invariants associated with (1), (2) that play the leading role in the theory are the total energy H and circulation C, defined by

$$H = \frac{1}{2}\int_{\mathcal{X}} \left[\left(\frac{\partial \psi}{\partial x_1}\right)^2 + \left(\frac{\partial \psi}{\partial x_2}\right)^2 + r^{-2}\psi^2 \right] \mathrm{d}x, \qquad C = \int_{\mathcal{X}} [Q - b]\,\mathrm{d}x. \quad (3)$$

In addition, the x_1-component of linear impulse, $M = \int_{\mathcal{X}} x_2[Q - b]\,\mathrm{d}x$, is also conserved in the channel $\mathcal{X}$; however, for the sake of brevity we will ignore this invariant in our subsequent discussion, since its role is completely analogous to that of C. These invariants depend on the large scales of motion, meaning that they are continuous with respect to spatial coarse-graining. By contrast, the generalized enstrophy invariants $A = \int_{\mathcal{X}} a(Q)\,\mathrm{d}x$, for an arbitrary nonlinear function a, are sensitive to the small-scale structure of Q. For this reason, we separate our formulation of the statistical equilibrium models below into two subsections: in the first, we define the probabilistic structure of the small scales, and in the second we impose the conditioning determined by the invariants for the large scales.

2.2. *Prior distribution on small-scale fluctuations*

The statistical equilibrium theory is derived by taking an appropriate continuum limit of a sequence of probabilistic lattice models, which we now introduce. The microstate Q is discretized on a lattice $\mathcal{L}_n$ having n sites, using a uniform partition of the intervals $-l_1/2 < x_1 < l_1/2$ and $-l_2/2 < x_2 < l_2/2$ into n_1 and n_2 parts, respectively, so that $n = n_1 n_2$. The domain $\mathcal{X}$ then consists of the disjoint union of n microcells $M(s)$ indexed by the sites $s \in \mathcal{L}_n$. The microstates in the lattice model are piecewise-constant fields Q relative to $\mathcal{L}_n$. The phase space for the lattice model is the product space $\Omega_n = \mathbb{R}^n$, whose points are the discretized fields $Q(s)$, $s \in \mathcal{L}_n$.

The small-scale fluctuations of the random microstate Q are described by the product measure

$$\Pi_n(\mathrm{d}Q) = \prod_{s \in \mathcal{L}_n} \rho(\mathrm{d}Q(s)) \qquad \text{on } \Omega_n, \qquad (4)$$

in which $\rho(\mathrm{d}y)$ is a given probability distribution on $\mathbb{R}$. Here and throughout this paper, y denotes a real variable running over the range of Q. With respect to the probability distribution Π_n, the microscopic fields Q consist of n independent identically distributed random variables over the n microcells in the lattice. We refer to the common distribution ρ as the *prior distribution*, signifying that it describes the statistical properties of the microstate Q before the conditioning due to the invariants (3) is imposed.

The product measure Π_n in (4) can be viewed as a canonical Gibbs measure with respect to a generalized enstrophy integral $A = \int_{\mathcal{X}} a(Q)\,\mathrm{d}x$. Indeed, when $\rho(\mathrm{d}y) = \mathrm{e}^{-a(y)}\,\mathrm{d}y$ for some continuous function a on R, we see that

$$\Pi_n(\mathrm{d}Q) = \mathrm{e}^{-nA_n(Q)} \prod_{s \in \mathcal{L}_n} \mathrm{d}Q(s) \qquad \text{with } A_n \doteq \frac{1}{n}\sum_{s \in \mathcal{L}_n} a(Q(s)).$$

The role of $\Pi_n(\mathrm{d}Q)$ in the lattice model is evident from this identity: it describes the statistical properties of the potential vorticity fluctuations on the lattice microscale consistent with the conservation of phase volume $\mathrm{d}Q = \prod \mathrm{d}Q(s)$ and a discretized generalized enstrophy A_n.

Statistical equilibrium theories of the long-time average behaviour of solutions to (1), (2) have tended to emphasize a microcanonical formulation with respect to the generalized enstrophy invariants [28, 29, 33, 34]. While such theories formulate well-defined and self-consistent lattice models, they have been criticized from a theoretical point of view in [3, 36] because they do not exhibit a lattice dynamics that conserves both the invariants A_n and the phase volume in Ω_n, and that converges to the underlying continuum dynamics on $\mathcal{X}$ as $n \to \infty$. In essence, these theories inhibit the flux of enstrophy from the lattice microscale to smaller, unresolved scales of motion excited by the exact ideal dynamics. The objection to a microcanonical perspective is even stronger from a practical point of view. Applications of the statistical equilibrium theory commonly involve dissipation and forcing at the small scales, and consequently the free evolution under ideal dynamics is rarely of any relevance. The generalized enstrophies, being sensitive to the higher-order moments of the small-scale potential vorticity fluctuations, are not continuous with respect to these nonideal perturbations. It is therefore undesirable to formulate a theory that depends upon specifying and conserving a family of generalized enstrophies.

When, instead, a canonical formulation with a specified prior distribution ρ is adopted, a much more flexible statistical equilibrium theory results. In practical applications, where the small-scale characteristics of the potential vorticity field Q typically involve many complex mechanisms, the choice of the prior distribution can be treated as a modelling issue. Often the properties of ρ can be inferred by comparing the predictions of the theory with observational data or numerical experiments. For instance, the long-time organization of coherent structures in freely decaying turbulence has been predicted by using a prior distribution that is compatible with the final state [4, 39]. Successful statistical equilibrium models of open ocean deep convection [8] and of the weather layer in the Jovian atmosphere [37] have been constructed by adapting the prior distribution to available data. The evolution of the large-scale structure of two-dimensional turbulence with weak driving and small dissipation has been modelled by means of an adiabatic approximation [9, 16, 24]. For all these reasons, we base our models on the canonical ensemble (4) parametrized by a probability distribution ρ, which is arbitrary apart from the technical requirement that it decays rapidly enough at infinity.

2.3. *Dynamical constraints on large-scale motions*

The statistical equilibrium lattice models that we consider are constructed from the product measure $\Pi_n(\mathrm{d}Q)$ and invariants H and C. With respect to these invariants, either the canonical or the microcanonical ensemble can be used. A main goal of this paper is to investigate the equivalence or nonequivalence of these two different ensembles, which we now define precisely.

The canonical model is defined by the Gibbs distribution

$$P_{n,\beta,\gamma}(\mathrm{d}Q) = Z_n(\beta,\gamma)^{-1}\exp(-n\beta H_n(Q) - n\gamma C_n(Q))\Pi_n(\mathrm{d}Q) \tag{5}$$

and is parametrized by $\beta, \gamma \in \mathbb{R}$, which play the roles of 'inverse temperature' and 'chemical potential', respectively. The partition function

$$Z_n(\beta,\gamma) = \int_{\Omega_n} \exp(-n\beta H_n(Q) - n\gamma C_n(Q))\Pi_n(\mathrm{d}Q)$$

normalizes the probability distribution $P_{n,\beta,\gamma}(\mathrm{d}Q)$ on Ω_n. On the other hand, the microcanonical model is defined by the conditional distribution

$$P_n^{E,\Gamma}(\mathrm{d}Q) = \Pi_n\left\{\mathrm{d}Q \mid H_n(Q) = E,\ C_n(Q) = \Gamma\right\}, \tag{6}$$

at given values E and Γ of energy and circulation, respectively. The functions H_n and C_n in (5) and (6) are lattice discretizations of the functionals H and C in (3); they are defined by

$$H_n(Q) = \frac{l_1^2 l_2^2}{2n^2} \sum_{s\in\mathcal{L}_n} \sum_{s'\in\mathcal{L}_n} [Q(s) - b(s)] g_n(s, s') [Q(s') - b(s')],$$
$$C_n(Q) = \frac{l_1 l_2}{n} \sum_{s\in\mathcal{L}_n} [Q(s) - b(s)],$$

where $g_n(s, s')$ denotes the average over $M(s) \times M(s')$ of the Green function $g(x, x')$ for the operator $-\Delta + r^{-2}$ on $\mathcal{X}$ with appropriate boundary conditions on $\partial\mathcal{X}$. These expressions approximate the energy and circulation of the piecewise-constant microstate $Q \in \Omega_n$.

The canonical parameters β and γ are scaled by a factor n in (5). The scaling in this nonextensive continuum limit, which is different from the usual thermodynamic limit [29], ensures that the mean values $\langle H_n \rangle$ and $\langle C_n \rangle$ with respect to the canonical ensemble (5) tend to finite limits compatible with the fixed constraint values E and Γ for the microcanonical ensemble (6).

3. Maximum entropy principles

We now calculate the continuum limits of the canonical and microcanonical models. We obtain the variational principles governing the equilibrium macrostates of these two models by employing the theory of large deviations [7, 12] to analyse the asymptotic behaviour of a certain coarse-graining of the potential vorticity field Q.

3.1. Coarse-grained process

We introduce a macroscopic description of the potential vorticity field that complements the microscopic description inherent in $Q \in \Omega_n$, and we link these descriptions together through a *coarse-grained process* $\tilde{Q}_{n,m}$. To this end, we partition the domain $\mathcal{X}$ uniformly into $m = m_1 m_2$ macrocells $\mathcal{X}_{j_1,j_2}$, with $m_1 \ll n_1$, $m_2 \ll n_2$ and $j_1 = 1, \ldots, m_1$, $j_2 = 1, \ldots, m_2$. Each of the m macrocells $\mathcal{X}_{j_1,j_2}$ in this partition contains n/m sites of the lattice $\mathcal{L}_n$. Let $\tilde{Q}_{n,m}$ be the piecewise-constant, stochastic process defined by averaging the random microstate Q over each macrocell, namely

$$\tilde{Q}_{n,m}(x) = \frac{m}{n} \sum_{s\in\mathcal{X}_{j_1,j_2}} Q(s) \qquad \text{for all } x \in \mathcal{X}_{j_1,j_2}. \tag{7}$$

Also, let the space of macrostates q be the Hilbert space $L^2(\mathcal{X})$ with the usual norm $\|q\|^2 = \int_{\mathcal{X}} q^2 \, dx$. The coarse-grained process $\tilde{Q}_{n,m}$ takes values in $L^2(\mathcal{X})$. We are interested in the double limit in which $n \to \infty$ followed by $m \to \infty$, which we call the *continuum limit*. The spatial averaging inherent in $\tilde{Q}_{n,m}$ suggests that it exhibits a law-of-large-numbers behaviour in this continuum limit and that its asymptotics are characterized by a large deviation principle.

Our strategy is to deduce the limiting behaviour of $\tilde{Q}_{n,m}$ under either the canonical ensemble (5) or the microcanonical ensemble (6) from the corresponding behaviour under the product measure (4) determined by the prior distribution. Since the coarse-grained process $\tilde{Q}_{n,m}$ is a sample mean of the random variables $Q(s)$ over each macrocell, the basic large deviation principle for $\tilde{Q}_{n,m}$ with respect to $\Pi_n(dQ)$ is a straightforward extension of the classical Cramér's theorem for sample means of independent identically distributed random variables [7, 12]. Cramér's theorem involves the cumulant generating function

$$f(\eta) = \log \int_{\mathbb{R}} \exp(\eta y)\rho(\mathrm{d}y) \qquad (\eta \in \mathbb{R}) \tag{8}$$

associated with the prior distribution ρ, and the function i conjugate to f in the sense of the Legendre–Fenchel transform, namely

$$i(y) = \sup_{\eta} [\eta y - f(\eta)] \qquad (y \in \mathbb{R}). \tag{9}$$

Both f and i are convex real functions, and i achieves its unique minimum value of 0 at the mean value $\int y\rho(\mathrm{d}y)$. The reader is referred to [7, 12] for these definitions and properties.

In terms of these standard definitions, we define the functional

$$I(q) = \int_{\mathcal{X}} i(q(x))\,\mathrm{d}x \qquad (q \in L^2(\mathcal{X})). \tag{10}$$

In the terminology of large deviation theory, I is a rate function—it is a convex, lower semi-continuous functional mapping $L^2(\mathcal{X})$ into the extended interval $[0, +\infty]$. The basic large deviation principle for the coarse-grained process $\tilde{Q}_{n,m}$ with respect to $\Pi_n(\mathrm{d}Q)$ characterizes the asymptotic behaviour of $\tilde{Q}_{n,m}$ in terms of I. We can state a formal version of this large deviation principle as

$$\Pi_n\{\tilde{Q}_{n,m} \approx q\} \sim \mathrm{e}^{-nI(q)} \qquad \text{in the continuum limit } n \to \infty,\ m \to \infty \tag{11}$$

for any macrostate $q \in L^2(\mathcal{X})$. Here, the symbol $\approx$ means close in the strong topology of $L^2(\mathcal{X})$. Heuristically speaking, the asymptotic formula (11) provides an interpretation of the rate functional I as a negative entropy: $-I(q)$ quantifies the multiplicity of the microstates that correspond under the coarse-graining to a macrostate q. In probabilistic terms, it gives an exponential-order correction to the law-of-large-numbers behaviour of the coarse-grained process $\tilde{Q}_{n,m}$. In another paper [14], we state and prove the rigorous version of this basic theorem.

3.2. *Canonical model*

The continuum limit of the statistical equilibrium model governed by the canonical ensemble $P_{n,\beta,\gamma}(\mathrm{d}Q)$ defined in (5) can be characterized by the large deviation formula

$$P_{n,\beta,\gamma}\{\tilde{Q}_{n,m} \approx q\} \sim \mathrm{e}^{-nI_{\beta,\gamma}(q)} \qquad \text{in the continuum limit } n \to \infty,\ m \to \infty \tag{12}$$

for any $q \in L^2(\mathcal{X})$. In this formula,

$$I_{\beta,\gamma}(q) \doteq I(q) + \beta H(q) + \gamma C(q) - \Phi(\beta, \gamma), \tag{13}$$

where

$$\begin{aligned}\Phi(\beta, \gamma) &\doteq \min_{q \in L^2(\mathcal{X})} [I(q) + \beta H(q) + \gamma C(q)] \\ &= -\lim_{n\to\infty} \frac{1}{n} \log Z_n(\beta, \gamma).\end{aligned} \tag{14}$$

The precise statements and proofs of the large deviation limit (12) and the limit in (14) are given in our companion paper [14]. These results follow from Laplace asymptotics and a representation of the interaction functions H_n and C_n in terms of the coarse-grained process, namely

$$H_n(Q) = H(\tilde{Q}_{n,m}) + \mathrm{o}(1), \qquad C_n(Q) = C(\tilde{Q}_{n,m}) + \mathrm{o}(1). \tag{15}$$

Here the functionals H and C are evaluated on macrostates $q \in L^2(\mathcal{X})$, and the streamfunction $\psi = G(q-b)$ corresponding to q is determined by the Green operator G for $-\Delta+r^{-2}$, namely

$$Gz(x) = \int_{\mathcal{X}} g(x, x')z(x')\,\mathrm{d}x' \qquad (z \in L^2(\mathcal{X})). \tag{16}$$

The approximations (15) rely on the fact that the quadratic self-interaction term in H involves long-range interactions via $g(x, x')$, and the fact that C and the term in H arising from interaction with the topography are linear. A rigorous proof of such approximations is given [3, 14].

According to (12), the most probable macrostates q are those at which $I_{\beta,\gamma}(q)$ achieves its minimum value of 0. These macrostates constitute the *set of equilibrium states* associated with the canonical parameters $\beta, \gamma \in \mathbb{R}$, namely

$$\mathcal{E}_{\beta,\gamma} \doteq \{q \in L^2(\mathcal{X}) : I_{\beta,\gamma}(q) = 0\} = \arg\min[I + \beta H + \gamma C]. \tag{17}$$

Any macrostate q that does not lie in $\mathcal{E}_{\beta,\gamma}$ has an exponentially small probability of being observed as a coarse-grained state in the continuum limit. The macrostates in $\mathcal{E}_{\beta,\gamma}$ are therefore the overwhelmingly most probable among all possible macrostates of the turbulent system. Accordingly, the main predictions of the statistical equilibrium theory in its canonical form are derived by solving the unconstrained minimization problem whose objective functional is $I + \beta H + \gamma C$.

The first-order conditions for the variational problem determining equilibrium states $\bar{q} \in \mathcal{E}_{\beta,\gamma}$ are

$$0 = \delta(I + \beta H + \gamma C)(\bar{q}) = \int_{\mathcal{X}} [i'(\bar{q}) + \beta\bar{\psi} + \gamma]\,\delta q\,\mathrm{d}x, \tag{18}$$

where $\bar{\psi}$ is the streamfunction corresponding to $\bar{q}$, and δq denotes a variation in $L^2(\mathcal{X})$. From this calculation, we obtain the *mean-field equation*

$$\bar{q} = -\Delta\bar{\psi} + r^{-2}\bar{\psi} + b = f'(-\beta\bar{\psi} - \gamma), \tag{19}$$

using the fact that, since f and i are conjugate convex functions, f' and i' are inverse functions. The dependence f' of the mean potential vorticity $\bar{q}$ on the mean streamfunction $\bar{\psi}$ is determined via (8) by the statistical properties of the small-scale fluctuations encoded in the prior distribution ρ. The second-order conditions at an equilibrium state $\bar{q}$ are

$$\begin{aligned} 0 &\leqslant \delta^2(I + \beta H + \gamma C)(\bar{q}) \\ &= \int_{\mathcal{X}} \left\{ i''(\bar{q})(\delta q)^2 + \beta\left[\left(\frac{\partial\delta\psi}{\partial x_1}\right)^2 + \left(\frac{\partial\delta\psi}{\partial x_2}\right)^2 + r^{-2}(\delta\psi)^2\right]\right\}\mathrm{d}x. \end{aligned} \tag{20}$$

This condition is equivalent to the nonnegative-definiteness of the bounded, symmetric operator $i''(\bar{q}) + \beta G$ on $L^2(\mathcal{X})$. When this quadratic form is positive-definite at $\bar{q}$, the equilibrium set $\mathcal{E}_{\beta,\gamma} = \{\bar{q}\}$ necessarily consists of the single, nondegenerate solution.

3.3. Microcanonical model

The continuum limit of the statistical equilibrium model governed by the microcanonical ensemble $P_n^{E,\Gamma}(\mathrm{d}Q)$ defined in (6) is characterized by the large deviation formula

$$P_n^{E,\Gamma}\{\tilde{Q}_{n,m} \approx q\} \sim \mathrm{e}^{-nI^{E,\Gamma}(q)} \qquad \text{in the continuum limit } n \to \infty,\ m \to \infty \tag{21}$$

for any $q \in L^2(\mathcal{X})$. In this formula,

$$I^{E,\Gamma}(q) \doteq \begin{cases} I(q) + S(E, \Gamma) & \text{if } H(q) = E, \quad C(q) = \Gamma, \\ +\infty & \text{otherwise,} \end{cases} \tag{22}$$

where

$$S(E,\Gamma) \doteq -\min\{I(q) : H(q) = E,\ C(q) = \Gamma\}$$
$$= \lim_{n\to\infty} \frac{1}{n} \log \Pi_n\{H_n = E,\ C_n = \Gamma\}. \tag{23}$$

As in the analysis of the canonical model, the precise statements and proofs of the large deviation limit (21) and the limit in (23) are presented in [14]. Given the approximations (15) and the basic large deviation principle (11), these results follow from general large deviation principles for conditional distributions developed in [13]. (In the mathematically rigorous versions of these results, the microcanonical constraints are replaced by set containments such as $H_n \in [E-\epsilon, E+\epsilon]$, and the limit as $\epsilon \to 0$ is taken after the continuum limit.)

Among the macrostates lying on the microcanonical manifold $H = E$, $C = \Gamma$, the most probable macrostates q are those at which $I^{E,\Gamma}$ achieves its minimum value of 0. These macrostates compose the *set of equilibrium states* associated with given microcanonical parameters E, Γ, namely

$$\mathcal{E}^{E,\Gamma} \doteq \{q \in L^2(\mathcal{X}) : I^{E,\Gamma}(q) = 0\} = \arg\min\{I : H = E,\ C = \Gamma\}. \tag{24}$$

As in the canonical model, any macrostate that does not lie in the equilibrium set has an exponentially small probability of being observed.

The first-order conditions for a microcanonical equilibrium $\bar{q} \in \mathcal{E}^{E,\Gamma}$ are identical to (18), except that β and γ are the Lagrange multipliers associated with the energy and circulation constraints, respectively. Similarly, the mean-field equation (19) for the solution triple $(\bar{q}, \beta, \gamma)$ is unchanged. The second-order conditions, on the other hand, are fundamentally altered by shifting from the canonical to the microcanonical formulation. General principles of optimization imply that the nonnegativity condition (20) at a constrained minimizer $\bar{q}$ holds for all variations δq that are infinitesimally compatible with the constraints, but not necessarily for arbitrary variations δq [2, 30]. Thus, we find that the second-order conditions appropriate to a macrostate $\bar{q} \in \mathcal{E}^{E,\Gamma}$ are that (20) holds for all δq satisfying

$$\delta H(\bar{q}) = \int_{\mathcal{X}} \bar{\psi}\, \delta q \,\mathrm{d}x = 0 \qquad \text{and} \qquad \delta C(\bar{q}) = \int_{\mathcal{X}} \delta q \,\mathrm{d}x = 0. \tag{25}$$

In view of these side-conditions arising from the linearization of the constraints $H = E$ and $C = \Gamma$, the set of microcanonical equilibria is potentially larger than the corresponding set of canonical equilibria. In fact, this difference between the canonical and microcanonical equilibrium conditions at second-order underlies all of our subsequent development.

4. Equivalence and nonequivalence

We now turn our attention to the relation between the equilibrium sets $\mathcal{E}_{\beta,\gamma}$ for the canonical model and the equilibrium sets $\mathcal{E}^{E,\Gamma}$ for the microcanonical model. In most statistical equilibrium models, the canonical and microcanonical ensembles are equivalent, in the sense that there is a one-to-one correspondence between their equilibrium states. Surprisingly, however, for the theories of coherent structures in turbulence not all microcanonical equilibria are realized as canonical equilibria.

4.1. Thermodynamic functions

The properties of the thermodynamic functions in the microcanonical and canonical models determine the correspondence, or lack of correspondence, between equilibria for these two models. The fundamental thermodynamic function for the microcanonical model is the value function $S(E, \Gamma)$ in the constrained maximum entropy principle (23) whose solutions constitute the equilibrium set $\mathcal{E}^{E,\Gamma}$. Similarly, the fundamental thermodynamic function for the canonical model is the value function $\Phi(\beta, \gamma)$ in the free maximum entropy principle (14) whose solutions constitute the equilibrium set $\mathcal{E}_{\beta,\gamma}$. These two functions are conjugate functions in the sense of convex analysis, i.e.

$$\Phi(\beta, \gamma) = S^*(\beta, \gamma) \doteq \inf_{E,\Gamma} [\beta E + \gamma\Gamma - S(E, \Gamma)]. \tag{26}$$

The proof simply amounts to writing the free minimization in (14) in terms of the constrained minimization in (23):

$$\begin{aligned}\min_q [I + \beta H + \gamma\Gamma] &= \inf_{E,\Gamma} \min_q \{I + \beta H + \gamma\Gamma : H = E,\ C = \Gamma\} \\ &= \inf_{E,\Gamma} [\beta E + \gamma\Gamma - S(E, \Gamma)].\end{aligned}$$

We express the criteria for the equivalence of ensembles in terms of properties of the *microcanonical entropy* S. Let $\mathcal{A}$ denote the domain of S, the largest open subset of $\mathbb{R}^2$ consisting of admissible constraint pairs (E, Γ) for the microcanonical model; a constraint pair is admissible if $(E, \Gamma) = (H(q), C(q))$ for some $q \in L^2(\mathcal{X})$ with $I(q) < +\infty$. To simplify our presentation, let us assume that S is differentiable on its domain $\mathcal{A}$; experience with numerical solutions to (23) strongly suggests that this differentiability assumption is generally valid. Then, to every $(E, \Gamma) \in \mathcal{A}$ there corresponds a unique pair (β, γ) given by the usual thermodynamic formulae

$$\beta = \frac{\partial S}{\partial E}, \qquad \gamma = \frac{\partial S}{\partial \Gamma}, \tag{27}$$

and, under the usual equivalence of ensembles, we expect a correspondence between the microcanonical equilibrium set $\mathcal{E}^{E,\Gamma}$ and the canonical equilibrium set $\mathcal{E}_{\beta,\gamma}$. Such a correspondence holds, however, only when S is strictly concave on $\mathcal{A}$ [13]. In order to characterize the relation between these equilibrium sets in general, it is necessary to consider the concave hull of S, namely

$$S^{**}(E, \Gamma) = \Phi^*(E, \Gamma) \doteq \inf_{\beta,\gamma} [\beta E + \gamma\Gamma - \Phi(\beta, \gamma)]. \tag{28}$$

In general, $S^{**}(E, \Gamma) \geqslant S(E, \Gamma)$ for all $(E, \Gamma) \in \mathcal{A}$. The subset $\mathcal{C} \subseteq \mathcal{A}$ on which the concave hull S^{**} coincides with S plays a pivotal role in this characterization. The *concavity set* $\mathcal{C}$ consists of those points $(E, \Gamma) \in \mathcal{A}$ for which $S^{**}(E, \Gamma) = S(E, \Gamma)$, or, equivalently, for which S has a supporting plane, in the sense that

$$S(E', \Gamma') \leqslant S(E, \Gamma) + \beta(E' - E) + \gamma(\Gamma' - \Gamma) \qquad \text{for all } (E', \Gamma') \in \mathcal{A}. \tag{29}$$

4.2. Microcanonical and canonical equilibrium sets

In terms of the concavity set $\mathcal{C}$, the admissible set $\mathcal{A}$ decomposes into three disjoint sets, where (a) there is a one-to-one correspondence between microcanonical and canonical equilibria, (b) there is a many-to-one correspondence from microcanonical equilibria to canonical equilibria, and (c) there is no correspondence. We state this complete classification as follows:

(a) *Full equivalence.* If (E, Γ) belongs to $\mathcal{C}$ and there is a unique point of contact between S and its supporting plane at (E, Γ), then $\mathcal{E}^{E,\Gamma}$ coincides with $\mathcal{E}_{\beta,\gamma}$.

(b) *Partial equivalence.* If (E, Γ) belongs to $\mathcal{C}$ but there is more than one point of contact between S and its supporting plane at (E, Γ), then $\mathcal{E}^{E,\Gamma}$ is a strict subset of $\mathcal{E}_{\beta,\gamma}$. Moreover, $\mathcal{E}_{\beta,\gamma}$ contains all those $\mathcal{E}^{E',\Gamma'}$ for which (E', Γ') is also a point of contact.

(c) *Nonequivalence.* If (E, Γ) does not belong to $\mathcal{C}$, then $\mathcal{E}^{E,\Gamma}$ is disjoint from $\mathcal{E}_{\beta,\gamma}$. In fact, $\mathcal{E}^{E,\Gamma}$ is disjoint from all canonical equilibrium sets.

The proofs of these results are relatively elementary applications of convex analysis; they are given in [14]. A more abstract discussion of equivalence and nonequivalence of ensembles is presented in [13], where general results are established for a wide class of models and without the simplifying assumption that S is differentiable.

The most important outcome of this analysis is the nonequivalence behaviour, in which microcanonical equilibria for constraint values (E, Γ) lying outside the concavity set $\mathcal{C}$ are omitted by the canonical model. Remarkably, this behaviour is found to occur in many applications of the statistical equilibrium theory [8, 10, 15, 22, 37]. In section 6, we present a vivid illustration of the nonequivalence behaviour in a prototypical geophysical fluid dynamics problem. In light of these mathematical results and the observed fact that the concavity set is often a small subset of the admissible set, we see that the microcanonical formulation in the energy and circulation invariants produces a richer theory than the corresponding canonical formulation.

5. Nonlinear stability

In either the canonical or the microcanonical model, the equilibrium macrostates determine steady mean flows that are the most probable flows compatible with the given parameters of the model. This statistical property can be interpreted as a stability property: while the microstate evolves ergodically, the coarse-grained macrostate remains close to the steady mean flow with very high probability. Thus, the statistical equilibrium macrostates are stable with respect to perturbations on the microscopic scales. We now inquire whether these steady mean states are also stable in a strong sense with respect to macroscopic perturbations. Precisely, we investigate the evolution under ideal dynamics of any perturbed macroscopic state q that initially lies within a small, finite distance in $L^2(\mathcal{X})$ of an equilibrium macrostate $\bar{q}$.

Broadly speaking, we find that the canonical equilibrium states always satisfy the Arnold sufficient conditions for nonlinear stability [1, 18, 25], but that the microcanonical equilibrium states only satisfy those conditions when they are realized as canonical equilibrium states. To prove the nonlinear stability of the microcanonical equilibrium states in the nonequivalent case, we construct a more refined Lyapunov functional than that used to obtain the Arnold conditions.

5.1. Arnold stability theorems

We assume that, for given values of the canonical parameters β and γ, the equilibrium state $\bar{q} \in \mathcal{E}_{\beta,\gamma}$ is an isolated, nondegenerate minimizer of $I_{\beta,\gamma}$; otherwise, stability in a strict sense cannot be expected. In light of (20), a natural definition of nondegeneracy is that $\delta^2 I_{\beta,\gamma}(\bar{q})$ be strictly positive-definite at $\bar{q}$; i.e., for all variations $\delta q \in L^2(\mathcal{X})$,

$$\mu \int_{\mathcal{X}} (\delta q)^2 \, \mathrm{d}x \leqslant \delta^2 I_{\beta,\gamma}(\bar{q}) \tag{30}$$

for some positive constant μ independent of δq. The complementary upper bound

$$\delta^2 I_{\beta,\gamma}(\bar{q}) \leqslant \nu \int_{\mathcal{X}} (\delta q)^2 \, \mathrm{d}x \tag{31}$$

also holds, for a constant ν independent of δq. From the identity

$$\delta^2 I_{\beta,\gamma}(\bar{q}) = \delta^2(I + \beta H + \gamma\Gamma)(\bar{q}) = \int_{\mathcal{X}} [i''(\bar{q})(\delta q)^2 + \beta \, \delta q \, G \, \delta q] \, \mathrm{d}x,$$

it is obvious that the optimal constants μ and ν in (30) and (31) are the smallest and largest eigenvalues, respectively, of the operator $i''(\bar{q}) + \beta G$, where G is the Green operator (16). The nonlinear stability of $\bar{q}$ follows immediately by noticing that the rate function $I_{\beta,\gamma}$ for the large deviation principle (12) is a Lyapunov functional. Indeed, $I_{\beta,\gamma}$ is a conserved quantity for the dynamics (1), since it is a linear combination of H, C, and I, which coincides with a certain generalized enstrophy invariant. The familiar Lyapunov argument then yields

$$\frac{\mu}{2}\|q(\cdot,t)-\bar{q}\|^2 \leqslant I_{\beta,\gamma}(q(t)) = I_{\beta,\gamma}(q(0)) \leqslant 2\nu\|q(\cdot,0)-\bar{q}\|^2$$

for all $t > 0$. Thus, we conclude that any nondegenerate canonical equilibrium state $\bar{q} \in \mathcal{E}_{\beta,\gamma}$ determines a steady flow that is stable in the sense that, if $\|q(\cdot,0)-\bar{q}\|$ is sufficiently small, then for all time $t > 0$, $\|q(\cdot,t)-\bar{q}\| \leqslant c\|q(\cdot,0)-\bar{q}\|$ for some finite constant c.

In the familiar context of deterministic stability analysis, a specified steady flow may or may not satisfy the sufficient conditions for stability. In the statistical equilibrium theory, on the other hand, stability criteria amount to nondegeneracy conditions for equilibrium states. Apart from degeneracies, all positive temperature states ($\beta \geqslant 0$) are stable by the so-called first Arnold theorem, while all negative temperature states ($\beta < 0$) are stable by the so-called second Arnold theorem [1, 25]. In either case, the stability condition is that the bounded, symmetric operator $i''(\bar{q}) + \beta G$ be positive-definite. This condition can be translated into the familiar form used in hydrodynamic stability theory via the formula

$$\frac{\mathrm{d}\bar{q}}{\mathrm{d}\bar{\psi}} = -\frac{\beta}{i''(\bar{q})},$$

which follows from the mean-field equation (19) and the fact that f' and i' are inverse functions. The first Arnold theorem applies when $\mathrm{d}\bar{q}/\mathrm{d}\bar{\psi} < 0$, while the second Arnold theorem applies when $0 < \mathrm{d}\bar{q}/\mathrm{d}\bar{\psi} < \lambda_1$, where λ_1 is the smallest eigenvalue of $-\Delta + r^{-2}$.

5.2. Refined stability theorems

When a microcanonical equilibrium $\bar{q} \in \mathcal{E}^{E,\Gamma}$ does not lie in any canonical equilibrium set $\mathcal{E}_{\beta,\gamma}$, the Arnold stability conditions do not apply. Nevertheless, every nondegenerate equilibrium state for the microcanonical model determines a stable flow, as we now show by giving a more refined nonlinear stability analysis.

We assume that $\bar{q} \in \mathcal{E}^{E,\Gamma}$ is the isolated, nondegenerate minimizer of I at given microcanonical constraint values E and Γ. In the microcanonical model, the second-order conditions at $\bar{q}$ are subject to the side-conditions (25). Therefore, a natural definition of nondegeneracy for a microcanonical equilibrium state $\bar{q}$ is that (30) holds for all $\delta q \in L^2(\mathcal{X})$ satisfying the linearized constraints (25). The complementary upper bound (31) holds at $\bar{q}$ for arbitrary δq.

Our strategy for proving the stability of $\bar{q}$ is to construct a Lyapunov functional in the form

$$L^{E,\Gamma}_{\sigma,\tau}(q) \doteq I(q) + S(E,\Gamma) + \beta[H(q)-E] + \gamma[C(q)-\Gamma] + \frac{\sigma}{2}[H(q)-E]^2 + \frac{\tau}{2}[C(q)-\Gamma]^2, \tag{32}$$

where β and γ are the Lagrange multipliers for the energy and circulation constraints, respectively, and σ and τ are sufficiently large positive constants. The terms in (32) scaled by σ and τ penalize departures from the microcanonical constraints, and yet they do not change the values of the Lyapunov functional or its first variation at $\bar{q}$, which are

$$L^{E,\Gamma}_{\sigma,\tau}(\bar{q}) = 0, \qquad \delta L^{E,\Gamma}_{\sigma,\tau}(\bar{q}) = \delta(I + \beta H + \gamma C)(\bar{q}) = 0.$$

It is possible to choose finite constants σ and τ so that $L_{\sigma,\tau}^{E,\Gamma}$ has a nondegenerate, unconstrained minimum at $\bar{q}$ even in cases of nonequivalence or partial equivalence, when the microcanonical equilibrium $\bar{q}$ is not contained in the corresponding canonical equilibrium set. In such a case, (E, Γ) does not belong to the concavity set $\mathcal{C}$, and hence the tangent plane to S at (E, Γ) is not a supporting plane. Nevertheless, there is a supporting paraboloid to S at (E, Γ) determined by fixing σ and τ large enough so that

$$S(E', \Gamma') \leqslant S(E, \Gamma) + \beta(E' - E) + \gamma(\Gamma' - \Gamma) + \frac{\sigma}{2}(E' - E)^2 + \frac{\tau}{2}(\Gamma' - \Gamma)^2$$

for all $(E', \Gamma') \in \mathcal{A}$, with equality only when $(E', \Gamma') = (E, \Gamma)$. It follows immediately that $L_{\sigma,\tau}^{E,\Gamma}(q) > L_{\sigma,\tau}^{E,\Gamma}(\bar{q}) = 0$ for all $q \neq \bar{q}$. Since $L_{\sigma,\tau}^{E,\Gamma}(q)$ is a conserved quantity, the Lyapunov argument ensues.

Technically, it is also necessary to show that the second variation of the Lyapunov functional (32) is positive-definite at $\bar{q}$. The influence of the penalty terms on variations δq that are not tangential to the constraint manifold $H = E$, $C = \Gamma$ is transparent from the calculation

$$\delta^2 L_{\sigma,\tau}^{E,\Gamma}(\bar{q}) = \delta^2(I + \beta H + \gamma C)(\bar{q}) + \sigma \left\{ \int_{\mathcal{X}} \bar{\psi}\, \delta q \,\mathrm{d}x \right\}^2 + \tau \left\{ \int_{\mathcal{X}} \delta q \,\mathrm{d}x \right\}^2 . \quad (33)$$

We therefore omit the technical details involved in verifying the nondegeneracy condition; a full discussion will be given elsewhere. As in the canonical model, we obtain the nonlinear stability of the steady flow corresponding to a nondegenerate microcanonical equilibrium state $\bar{q} \in \mathcal{E}^{E,\Gamma}$: if $\|q(\cdot, 0) - \bar{q}\|$ is sufficiently small, then for all time $t > 0$, $\|q(\cdot, t) - \bar{q}\| \leqslant c\|q(\cdot, 0) - \bar{q}\|$ for some finite constant c.

The key idea behind this construction of a Lyapunov functional arises in the study of constrained optimization problems, where functionals having the form of $L_{\sigma,\tau}^{E,\Gamma}$ are known as 'augmented Lagrangians' [2, 30].

6. Numerical example

6.1. Barotropic flow over topography in a zonal channel

For the purposes of illustrating the general results obtained in sections 4 and 5, we now present a family of computed solutions to the microcanonical equilibrium problem for a particular choice of domain, topography, and prior distribution. We consider a $1\frac{1}{2}$-layer model of a single zone-belt domain in the Jovian weather layer [11, 19, 26], in which the model parameters are set to be consistent with the scales of the observed motions on Jupiter [11, 26, 37]. In dimensionless variables, the model is defined in a channel domain $\mathcal{X} = \{-2 < x_1 < 2,\ -0.5 < x_2 < 0.5\}$ with zonal effective topography $b = B_2 \sin(2\pi x_2)$. We take $r = 0.2$ and $B_2 = 0.3$, based on a length scale $L \approx 2 \times 10^6$ m and a velocity scale $U \approx 30\,\mathrm{m\,s^{-1}}$. In this regime we expect the topography-induced mean flow to be anticyclonic (negative vorticity) in the zone where b is positive, and cyclonic (positive vorticity) in the belt where b is negative. With this simple sinusoidal topography, the model captures the essential features of a single band in the mid-latitudes of Jupiter. A much more realistic determination of the effective bottom topography is given in [37].

The formulation of the statistical equilibrium theory for this model is complete once we specify a prior distribution ρ. As is stressed in section 2, the choice of the prior distribution is a key element in the theory because it encodes the statistics of the unresolved small-scale turbulence. Here we adopt the family of prior distributions deduced in [37] from observations of Jupiter. Namely, we select the family of gamma distributions

$$\rho_\epsilon(\mathrm{d}y) = \frac{1}{|\epsilon|\Gamma(\epsilon^{-2})}\exp(\epsilon^{-2}[\log(1+\epsilon y) - (1+\epsilon y)])\,\mathrm{d}y \qquad (-\infty < y \leqslant |\epsilon|^{-1}),$$

having their mean, variance, and skewness normalized to be

$$\int y\rho_\epsilon(\mathrm{d}y) = 0, \qquad \int y^2\rho_\epsilon(\mathrm{d}y) = 1, \qquad \int y^3\rho_\epsilon(\mathrm{d}y) = 2\epsilon.$$

The parameter ϵ is taken to be negative, meaning that the skewness 2ϵ of $\rho_\epsilon(\mathrm{d}y)$ is anticyclonic. Besides having desirable physical properties, these prior distributions have the virtue that their cumulant generating functions $f_\epsilon(\eta)$, defined in (8), and the associated functionals I_ϵ, defined in (10), can be calculated. In particular,

$$I_\epsilon(q) = \int_{\mathcal{X}} [\epsilon^{-1}q - \epsilon^{-2}\log(1+\epsilon q)]\mathrm{d}x,$$

and the mean-field equation (19) corresponding to the prior distribution ρ_ϵ is

$$\bar{q} = -\Delta\bar{\psi} + r^{-2}\bar{\psi} + B_2\sin 2\pi x_2 = \frac{-\beta\bar{\psi} - \gamma}{1 - \epsilon(-\beta\bar{\psi} - \gamma)}. \tag{34}$$

It is evident that ϵ determines the magnitude of the principal nonlinear term in this equation. As $\epsilon \to 0$, ρ_ϵ approaches the standard normal distribution, and the statistical equilibrium theory reduces to the well-known energy–enstrophy theory [6, 23, 35]. In the computations discussed below, we fix $\epsilon = -0.04$.

To compute the solutions to the variational problem for the microcanonical model analysed in section 3, we implement the globally convergent iterative algorithm developed in [10, 38]. For given constraint values E and Γ in the admissible range, this algorithm finds the equilibrium macrostate $\bar{q} = \bar{q}(x; E, \Gamma) \in \mathcal{E}^{E,\Gamma}$ that minimizes $I_\epsilon(q)$ subject to $H(q) = E$, $C(q) = \Gamma$.

6.2. Computed results

In figure 1, we display the velocity field of the computed solution $\bar{q}$ for $E = 3.74$ and $\Gamma = 12.0$. This solution consists of a zonal shear flow with strong eastward and westward jets together with an embedded vortex. The total circulation over the domain is cyclonic, while the coherent vortex is an anticyclone. Qualitatively, this solution resembles the observed coherent structures in the mid-latitude zone-belt domains on Jupiter.

We use the prototypical solution computed in figure 1 as a point of reference for examining the behaviour of the microcanonical entropy $S(E, \Gamma) = -I_\epsilon(\bar{q})$ as a function of E and Γ. In figure 2 we plot $\beta(E, 12.0)$ over the range $2.0 < E < 4.5$; in figure 3 we plot $\gamma(3.74, \Gamma)$ over

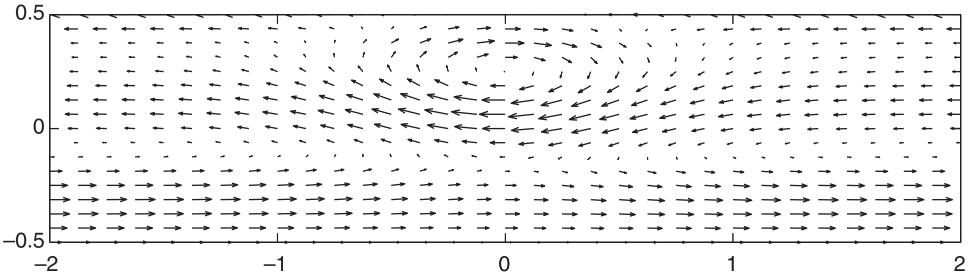

Figure 1. Velocity field of the microcanonical equilibrium state for $E = 3.74$ and $\Gamma = 12.0$.

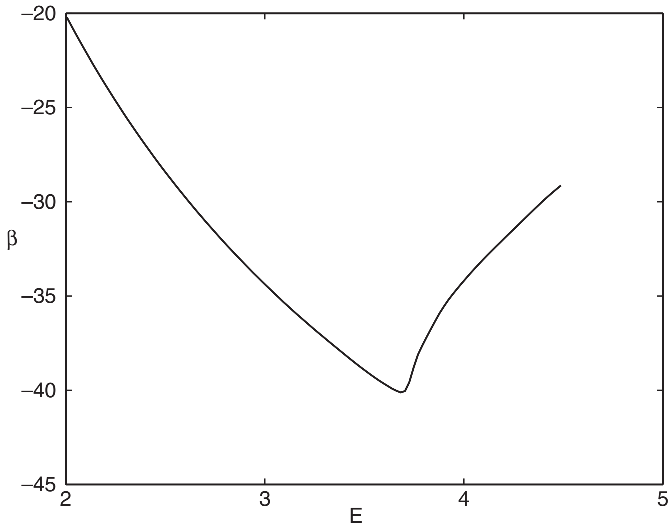

Figure 2. Inverse temperature β plotted versus energy E with fixed circulation $\Gamma = 12.0$.

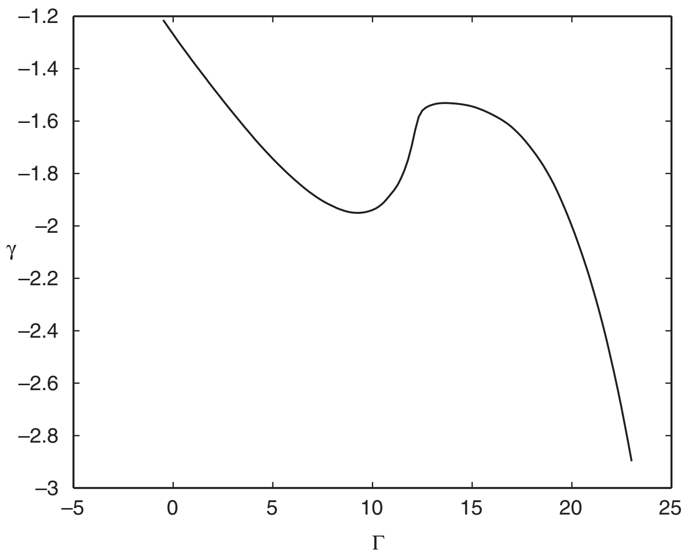

Figure 3. Chemical potential γ plotted versus circulation Γ with fixed energy $E = 3.74$.

the range $0 < \Gamma < 23$. In both of these plots we discover that the curves are not monotonically decreasing. In the light of (27) we conclude that both sections of the entropy function are not concave. This demonstrates that the equivalence between the microcanonical and canonical ensembles breaks down over a significant parameter range around the physically realistic flow with $E = 3.74$ and $\Gamma = 12.0$. Moreover, a further analysis of the tangent planes to S over this range of pairs (E, Γ) shows that the concavity set $\mathcal{C}$ is a relatively small subset of the admissible set $\mathcal{A}$.

In the light of this result, and similar results obtained in other implementations, we conclude that the statistical equilibrium theory based on a canonical ensemble in energy and circulation is inadequate for modelling the typical behaviour exhibited by the Jovian atmosphere. It is especially remarkable that the canonical formulation omits the physical regime of greatest interest in the Jovian context, namely the emergence of coherent anticyclonic vortices within a zonal shear having multiple flow reversals. For instance, the computed solutions with $\Gamma = 12.0$ transition from zonal shear flows to flows containing a coherent vortex near $E = 3.74$; for smaller E the computed solution is purely zonal, while for larger E it contains a vortex that intensifies as E increases. This transition nearly coincides with the strongest nonconcavity of the entropy. Similarly, and equally surprising, the nonequivalence with respect to Γ also occurs in this regime. By contrast, the microcanonical formulation produces realistic coherent structures throughout this regime.

As is explained in section 5, the most probable flows corresponding to constraint pairs (E, Γ) in the nonequivalence set are nonlinearly stable, even though their stability cannot be deduced from known sufficient conditions. In particular, the computed flows corresponding to constraint values near the reference values $E = 3.74$ and $\Gamma = 12.0$ have negative inverse temperature β and fail the Arnold second stability condition

$$0 < \frac{d\bar{q}}{d\bar{\psi}} < \lambda_1 = \pi^2 + r^{-2}. \tag{35}$$

Generally, this sufficient condition is too crude to be applicable to the jets and spots of the Jovian atmosphere. Nonetheless, our refined stability results guarantee the nonlinear stability of all microcanonical equilibrium states corresponding to admissible pairs (E, Γ), provided that only a technical nondegeneracy condition is fulfilled. (The linear rate of convergence exhibited by the iterative algorithm suggests that the required nondegeneracy is indeed satisfied.) Consequently, it is not necessary to verify a restrictive condition such as (35) to ensure the stability of most probable flows, and it is incorrect to assume that a steady flow that strongly violates the Arnold conditions is unstable. In essence, the known sufficient conditions are derived by utilizing conservation of a linear combination of two independent invariants, while the conservation of each of these quantities separately constrains the evolution of perturbations and leads to wider stability conditions. In this way, our refined stability results remove much of the confusion stemming from the fact that the permanent coherent structures in the Jovian atmosphere appear to lie beyond the reach of hydrodynamic stability conditions [11].

Acknowledgments

The authors thank Michael Kiessling and Andrew Majda for helpful conversations that stimulated aspects of this work. This research was supported by grants from the Department of Energy (DE-FG02-99ER25376) and the National Science Foundation (DMS-9700852 to RSE and DMS-9971204 to BT).

References

[1] Arnold V 1982 *Mathematical Methods in Classical Mechanics* (New York: Springer)
[2] Bertsekas D P 1982 *Constrained Optimization and Lagrange Multiplier Methods* (New York: Academic)
[3] Boucher C, Ellis R S and Turkington B 2000 Derivation of maximum entropy principles in two-dimensional turbulence via large deviations *J. Stat. Phys.* **98** 1235–78
[4] Brands H, Stulemeyer J and Pasmanter R A 1997 A mean field prediction of the asymptotic state of decaying 2D turbulence *Phys. Fluids* A **9** 2815–7
[5] Caglioti E, Lions P L, Marchioro C and Pulvirenti M 1992 A special class of stationary flows for two-dimensional Euler equations: a statistical mechanics description *Commun. Math. Phys.* **143** 501–25
[6] Carnavale G and Frederiksen J 1987 Nonlinear stability and statistical mechanics of flow over topography *J. Fluid Mech.* **175** 157–81

[7] Dembo A and Zeitouni O 1998 *Large Deviations Techniques and their Applications* 2nd edn (New York: Springer)
[8] DiBattista M and Majda A 1999 An equilibrium statistical model for the spreading phase of open-ocean convection *Proc. Natl Acad. Sci. USA* **96** 6009–13
[9] DiBattista M, Majda A and Grote M 2001 Meta-stability of equilibrium statistical structures for prototype geophysical flows with damping and driving *Physica* D **151** 271–304
[10] DiBattista M, Majda A and Turkington B 1998 Prototype geophysical vortex structures via large-scale statistical theory *Geophys. Astrophys. Fluid Dyn.* **89** 235–83
[11] Dowling T E 1995 Dynamics of Jovian atmospheres *Ann. Rev. Fluid Mech.* **27** 293–334
[12] Ellis R S 1985 *Entropy, Large Deviations, and Statistical Mechanics* (New York: Springer)
[13] Ellis R S, Haven K and Turkington B 2000 Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles *J. Stat. Phys.* **101** 999–1064
[14] Ellis R S, Haven K and Turkington B 2001 Analysis of statistical equilibrium models of geostrophic turbulence *Technical Report* University of Massachusetts, Amherst
[15] Eyink G L and Spohn H 1993 Negative-temperature states and large-scale, long-lived vortices in two-dimensional turbulence *J. Stat. Phys.* **70** 833–86
[16] Grote M J and Majda A 1997 Crude closure dynamics through large scale statistical theories *Phys. Fluids* **9** 3431–42
[17] Holloway G 1986 Eddies, waves, circulation and mixing: statistical geofluid mechanics *Ann. Rev. Fluid Mech.* **18** 91–147
[18] Holm D, Marsden J, Ratiu T and Weinstein A 1985 Nonlinear stability of fluid and plasma equilibria *Rev. Mod. Phys.* **123** 1–116
[19] Ingersoll A P and Cuong P G 1981 Numerical model of long-lived Jovian vortices *J. Atmos. Sci.* **38** 2067–76
[20] Joyce G and Montgomery D C 1973 Negative temperature states for the two-dimensional guiding center plasma *J. Plasma Phys.* **10** 107–21
[21] Kiessling M 1993 Statistical mechanics of classical particles with logarithmic interactions *Commun. Pure Appl. Math.* **46** 27–56
[22] Kiessling M and Lebowitz J 1997 The micro-canonical point vortex ensemble: beyond equivalence *Lett. Math. Phys.* **42** 43–56
[23] Kraichnan R 1975 Statistical dynamics of two-dimensional flow *J. Fluid Mech.* **67** 155–75
[24] Majda A and Holen M 1997 Dissipation, topography, and statistical theories of large-scale coherent structures *Commun. Pure Appl. Math.* **50** 1183–234
[25] Marchioro C and Pulvirenti M 1994 *Mathematical Theory of Incompressible Nonviscous Fluids (Appl. Math. Sci.* vol 96*)* (New York: Springer)
[26] Marcus P S 1993 Jupiter's Great Red Spot and other vortices *Ann. Rev. Astrophys.* **31** 523–73
[27] Michel J and Robert R 1994 Large deviations for Young measures and statistical mechanics of infinite dimensional dynamical systems with conservation law *Commun. Math. Phys.* **159** 195–215
[28] Miller J 1990 Statistical mechanics of Euler equations in two dimensions *Phys. Rev. Lett.* **65** 2137–40
[29] Miller J, Weichman P and Cross M C 1992 Statistical mechanics, Euler's equations, and Jupiter's red spot *Phys. Rev.* A **45** 2328–59
[30] Minoux M 1986 *Mathematical Programming: Theory and Algorithms* (Chichester: Wiley)
[31] Onsager L 1949 Statistical hydrodynamics *Suppl. Nuovo Cim.* **6** 279–87
[32] Pedlosky J 1987 *Geophysical Fluid Dynamics* 2nd edn (New York: Springer)
[33] Robert R 1991 A maximum-entropy principle for two-dimensional perfect fluid dynamics *J. Stat. Phys.* **65** 531–53
[34] Robert R and Sommeria J 1991 Statistical equilibrium states for two-dimensional flows *J. Fluid Mech.* **229** 291–310
[35] Salmon R, Holloway G and Henderschott M C 1976 The equilibrium statistical mechanics of simple quasi-geostrophic models *J. Fluid Mech.* **75** 691–703
[36] Turkington B 1999 Statistical equilibrium measures and coherent states in two-dimensional turbulence *Commun. Pure Appl. Math.* **52** 781–809
[37] Turkington B, Majda A, Haven K and DiBattista M 2001 Statistical equilibrium predictions of jets and spots on Jupiter *Proc. Natl Acad. Sci. USA* **98** 12346–50
[38] Turkington B and Whitaker N 1996 Statistical equilibrium computations of coherent structures in turbulent shear layers *SIAM J. Sci. Comput.* **17** 1414–33
[39] Turkington B and Whitaker N 2001 Direct numerical simulations and statistical equilibrium theories of two-dimensional turbulence *Technical Report* University of Massachusetts, Amherst