Journal of Applied Mathematics and Stochastic Analysis, **15**:4 (2002), 341-361.

ANALYSIS OF STATISTICAL EQUILIBRIUM MODELS OF GEOSTROPHIC TURBULENCE

RICHARD S. ELLIS[1] and BRUCE TURKINGTON[2]
University of Massachusetts
Department of Mathematics and Statistics
Amherst, MA 01003 USA
E-mail: rsellismath.umass.edu, turk@math.umass.edu

KYLE HAVEN
New York University
Courant Institute of Mathematical Sciences
New York, NY 10012 USA
E-mail: haven@cims.nyu.edu

(Received January, 2002; Revised August, 2002)

Statistical equilibrium lattice models of coherent structures in geostrophic turbulence, formulated by discretizing the governing Hamiltonian continuum dynamics, are analyzed. The first set of results concern large deviation principles (LDP's) for a spatially coarse-grained process with respect to either the canonical and/or the microcanonical formulation of the model. These principles are derived from a basic LDP for the coarse-grained process with respect to product measure, which in turn depends on Cramér's Theorem. The rate functions for the LDP's give rise to variational principles that determine the equilibrium solutions of the Hamiltonian equations. The second set of results addresses the equivalence or nonequivalence of the microcanonical and canonical ensembles. In particular, necessary and sufficient conditions for a correspondence between microcanonical equilibria and canonical equilibria are established in terms of the concavity of the microcanonical entropy. A complete characterization of equivalence of ensembles is deduced by elementary methods of convex analysis.

[1]This research was supported by a grant from the Department of Energy (DE-FG02-99ER25376) and by a grant from the National Science Foundation (NSF-DMS-9700852).

[2]This research was supported by a grant from the Department of Energy (DE-FG02-99ER25376) and by a grant from the National Science Foundation (NSF-DMS-9971204).

The mathematical results proved in this paper complement the physical reasoning and numerical computations given in a companion paper, where it is argued that the statistical equilibrium model defined by a prior distribution on potential vorticity fluctuations and microcanonical conditions on total energy and circulation is natural from the perspective of geophysical applications.

Key words: Large Deviation Principle, Cramér's Theorem, Coarse Graining, Statistical Models of Turbulence, Hamiltonian Systems.

AMS subject classifications: 76F55, 60F10.

1. Introduction

The problem of developing a statistical equilibrium theory of coherent structures in two-dimensional turbulence or quasi-geostrophic turbulence has been the focus of numerous mathematics and physics papers during the past fifty years. In a landmark paper Onsager [23] studied a microcanonical ensemble of point vortices and argued that they cluster into a coherent large-scale vortex in the negative temperature regime corresponding to sufficiently large kinetic energies. The quantitative aspects of the point vortex model were elaborated by Joyce-Montgomery [14], and subsequently these predictions were compared with direct numerical simulations of the end state of freely decaying turbulence [21]. Another approach due to Kraichnan [15], called the energy-enstrophy model, was based on spectral truncation of the underlying fluid dynamics. This Gaussian model was extensively applied to geophysical fluid flows [4, 27, 28]; for instance, to the organization of barotropic, quasi-geostrophic flows over bottom topography [11]. The next major development was the Miller-Robert model [20, 26], which enforced all the invariants of ideal motion and included the previous models as special or limiting cases. Recently, this general model has been scrutinized from the theoretical standpoint [30] and from the perspective of practical applications [6, 18]. These investigations have clarified the relation between various formulations of the general model and the simpler models, and they have revealed a formulation that has firm theoretical justification and rich physical applications. This formulation of the statistical equilibrium theory of coherent structures is described in our companion paper [10], and a particularly striking application of this model is given in another paper [31], where it is implemented to predict the permanent jets and spots in the active weather layer of Jupiter.

The purpose of the current paper is to prove the theorems that legitimize the theoretical model developed in [10]. We therefore follow the line of development in that paper, in which the nonlinear partial differential equation governing the dynamics of the fluid — an infinite-dimensional Hamiltonian system — is used to motivate an equilibrium statistical lattice model, both in its microcanonical and canonical formulations. Our first set of main results pertains to the continuum limit of this model. These results are large deviation principles (LDP's) — that is, exponential-order refinements of the law of large numbers — for a spatially coarse-grained process which justify the definition of the equilibrium states for the model. We base our direct proofs of these LDP's on an elementary large deviation analysis of the coarse-grained process with respect to the product measure that underlies the lattice model.

Our second set of main results concerns the relationship between the microcanonical and canonical equilibrium states; in particular, the equivalence or nonequivalence

of the microcanonical and canonical ensembles with respect to energy and circulation. The most interesting result is that microcanonical equilibria are often not realized as canonical equilibria. Significant examples of this remarkable phenomenon are given in [10, 31], where its impact on hydrodynamic stability is also discussed.

The outline of the paper is as follows. In Section 2 we formulate the statistical equilibrium theory described in [10], and we define the coarse-grained process used in the analysis of the continuum limit. We state and prove the basic LDP for this process in Section 3, and then in Section 4 we present the LDP's for the microcanonical and canonical ensembles, relying on general theorems established in [9]. Then in Section 5 we turn to the issue of equivalence of those ensembles at the level of their equilibrium states, giving necessary and sufficient conditions for equivalence in terms of the concavity of the microcanonical entropy. In the concluding Section 6, we point to some of the physical implications of our results.

2. Statement of Problem

The continuum dynamics that underlies our statistical equilibrium lattice model can be viewed as a noncanonical Hamiltonian system with infinitely many degrees of freedom. Namely, it can be written in the form

$$\frac{\partial Q}{\partial t} = \mathcal{J}\frac{\delta H}{\delta Q}, \tag{2.1}$$

where $\delta H/\delta Q$ denotes the variational derivative, and $\mathcal{J}$ is an operator with the property that

$$\{F, G\} \doteq \int_{\Lambda} \frac{\delta F}{\delta Q}\mathcal{J}\frac{\delta G}{\delta Q}d\lambda$$

is a Poisson bracket [22]. Here $\Lambda \subset \mathbb{R}^2$ is a bounded spatial domain, and $d\lambda$ represents normalized Lebesgue measure on Λ, normalized so that $\lambda(\Lambda) = 1$. For the sake of definiteness, the flow domain Λ is chosen to be the channel

$$\Lambda = \{x = (x_1, x_2)\colon |\, x_1 \,| < \ell_1/2,\ |\, x_2 \,| < \ell_2/2\}; \tag{2.2}$$

thus $d\lambda = dx/\ell_1\ell_2$.

The scalar function $Q = Q(x, t)$ represents either the vorticity in two-dimensional flow or the potential vorticity in quasi-geostrophic flow. Q evolves according to

$$\frac{\partial Q}{\partial t} + [Q, \psi] = 0, \text{ where } [f, g] \doteq \frac{\partial f}{\partial x_1}\frac{\partial g}{\partial x_2} - \frac{\partial f}{\partial x_2}\frac{\partial g}{\partial x_1}. \tag{2.3}$$

In this formula $\psi = \psi(x, t)$ is the streamfunction defined by $Q = L\psi + b$, where L is a uniformly elliptic operator on Λ and $b = b(x)$ is a given continuous real-valued function on Λ; ψ satisfies appropriate boundary conditions on the boundary of Λ. Consequently, ψ can be recovered from Q via a Green's function $g(x, x')$:

$$\psi(x) = \int_{\Lambda} g(x, x')(Q(x') - b(x'))\lambda(dx').$$

Hence (2.3) is an equation in Q alone. This equation takes the Hamiltonian form (2.1) when $J = -[Q, \cdot]$ and

$$H(Q) = \tfrac{1}{2}\int_{\Lambda} \psi(x)(Q(x) - b(x))\lambda(dx) \tag{2.4}$$

$$= \tfrac{1}{2}\int_{\Lambda\times\Lambda} (Q(x) - b(x))g(x,x')(Q(x') - b(x'))\lambda(dx)\lambda(dx').$$

Thus H is a quadratic form in Q in which interactions between $Q(x)$ and $Q(x')$ are governed by the Green's function $g(x, x')$. This Hamiltonian formulation of two-dimensional or quasi-geostrophic flow is discussed in [12, 22, 27, 29].

The Hamiltonian dynamics (2.1) conserves the total energy H and the total circulation

$$C(Q) = \int_{\Lambda} (Q(x) - b(x))\lambda(dx). \tag{2.5}$$

In addition, there is an infinite family of Casimir invariants parametrized by continuous real-valued functions a and defined as

$$A(Q) \doteq \int_{\Lambda} a(Q(x))\lambda(dx). \tag{2.6}$$

For the particular choice of a channel domain Λ, the linear impulse

$$\int_{\Lambda} x_2(Q(x) - b(x))\lambda(dx)$$

is also conserved. For simplicity, we shall ignore it throughout this paper since it plays a role identical to C in the theory.

When $L = -\Delta \doteq -\partial^2/\partial x_1^2 - \partial^2/\partial x_2^2$ and $b = 0$, (2.3) is the vorticity transport equation, which is an equivalent formulation of the Euler equations governing an incompressible, inviscid flow in two dimensions. When $L = -\Delta + r^{-2}$ and $b = \beta x_2 + \psi_2$, (2.3) is referred to as the quasi-geostrophic equations; the Coriolis parameter $\beta \geq 0$ and the Rossby deformation radius $r \in (0, \infty]$, and the effective bottom topography $\psi_2 = \psi_2(x_2)$ is a given real-valued function. With this choice of L and b, (2.3) governs the motion of a shallow, rotating layer of a homogeneous, incompressible, inviscid fluid in the limit of small Rossby number [27]. Throughout the rest of the paper, we refer to Q, as in this quasi-geostrophic case, as the potential vorticity.

2.1 Statistical Model

Since the governing dynamics (2.1) typically results in an intricate mixing of Q on a range of scales, we consider ensembles of solutions, instead of individual deterministic solutions. This statistical mechanics approach is standard for a Hamiltonian system with finitely many degrees of freedom in canonical form [2, 3, 24]. In order to carry over this methodology to the infinite dimensional Hamiltonian system (2.1), it is necessary to formulate an appropriate sequence of

lattice models obtained by discretizing the system. Next we define these probabilistic lattice models.

For each $n \in \mathbb{N}$, the flow domain Λ is uniformly partitioned into a_n microcells. We let Λ_n denote a uniform lattice of a_n sites indexed by s, where each $s \in \Lambda_n$ corresponds to a unique microcell $M(s)$ of Λ; each $M(s)$ satisfies $\lambda\{M(s)\} = 1/a_n$. The lattice discretizations of the Hamiltonian H and circulation C are given by

$$H_n(\{Q(s): s \in \Lambda_n\}) \doteq \frac{1}{2a_n^2} \sum_{s \in \Lambda_n} \sum_{s' \in \Lambda_n} (Q(s) - b_n(s)) g_n(s, s') (Q(s') - b_n(s')) \quad (2.7)$$

and

$$C_n(\{Q(s): s \in \Lambda_n\}) \doteq \frac{1}{a_n} \sum_{s \in \Lambda_n} (Q(s) - b_n(s)), \quad (2.8)$$

where $g_n(s, s')$ is the average over $M(s) \times M(s')$ of the Green's function $g(x, x')$, and $b_n(s)$ is the average over $M(s)$ of b:

$$g_n(s, s') \doteq a_n^2 \int_{M(s)} \int_{M(s')} g(x, x') \lambda(dx) \lambda(dx') \text{ and } b_n(s) \doteq a_n \int_{M(s)} b(x) \lambda(dx). \quad (2.9)$$

The Casimirs (2.6) can be discretized similarly as functions $A_n(\{Q(s): s \in \Lambda_n\})$.

We shall not attempt to prescribe a lattice dynamics under which the functions H_n, C_n and A_n are exact invariants. In fact, a lattice dynamics conserving the nonlinear enstrophies A_n is not known. The subtleties associated with the rigorous formulation of this statistical model are fully discussed elsewhere [10, 30]. Rather than pursue this direction any further here, we shall simply impose a joint probability distribution Π_n on the lattice potential vorticities $Q(s)$, $s \in \Lambda_n$; Π_n captures the invariance of the nonlinear enstrophies A and the phase volume dQ.

Let ρ be a probability distribution on $\mathbb{R}$ and let $\mathcal{Y}$ be the support of ρ. In order to simplify our presentation throughout the paper, we assume that $\mathcal{Y}$ is compact. The extension to the case of noncompact support can be carried out via appropriate modifications of the techniques that we use. We define the phase space Ω_n to be $\mathcal{Y}^{a_n}$, denote by $Q \in \Omega_n$ the vector of lattice potential vorticities $\{Q(s), s \in \Lambda_n\}$, and let Π_n be the product measure on Ω_n with one-dimensional marginals ρ. That is, for any Borel subset B of Ω_n, we define

$$\Pi_n\{B\} = \int_B \Pi_n(dQ) \doteq \int_B \prod_{s \in \Lambda_n} \rho(dQ(s)). \quad (2.10)$$

We refer to Π_n as the prior distribution of the lattice potential vorticities.

As is explained in [10], Π_n can be interpreted as a canonical ensemble with respect to lattice enstrophy A_n, in which the distribution ρ is determined by the function a. Alternatively, in the perspective adopted in the Miller-Robert model [19, 25], ρ is determined from the continuum initial data by

$$\rho(dy) = \int_\Lambda \delta_{Q(x,0)}(dy) \lambda(dx).$$

While this choice of prior distribution is conserved by the continuum equations, it is not preserved under passage to a discretized dynamics. In practice, therefore, it is better to view ρ

as a free parameter in the model. For discussions on the choice of the prior distribution in specific applications, see [10, §6.1] and [31, page 12347].

2.2 Definition of Ensembles

In terms of the prior distribution $\mathbf{\Pi}_n$, we can form two different statistical equilibrium models by using either the microcanonical ensemble or the canonical ensemble with respect to the invariants $\boldsymbol{H}_n$ and $\boldsymbol{C}_n$. Given $\boldsymbol{\kappa} > \mathbf{0}$, for any Borel subset $\boldsymbol{B}$ of $\mathbf{\Omega}_n$ the microcanonical ensemble is defined by

$$\boldsymbol{P}_{n,\kappa}^{E,\Gamma}\{\boldsymbol{B}\} \doteq \mathbf{\Pi}_n\{\boldsymbol{B} \mid \boldsymbol{H}_n \in [\boldsymbol{E} - \boldsymbol{\kappa}, \boldsymbol{E} + \boldsymbol{\kappa}], \boldsymbol{C}_n \in [\mathbf{\Gamma} - \boldsymbol{\kappa}, \mathbf{\Gamma} + \boldsymbol{\kappa}]\}. \quad (2.11)$$

This is the conditional probability distribution obtained by imposing the constraints that $\boldsymbol{H}_n$ and $\boldsymbol{C}_n$ lie within the (small) closed intervals around $\boldsymbol{E}$ and $\mathbf{\Gamma}$, respectively. Given real numbers $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, the canonical ensemble is defined by

$$\boldsymbol{P}_{n,\beta,\gamma}\{\boldsymbol{B}\} \doteq \tfrac{1}{\boldsymbol{Z}_n(\beta,\gamma)} \int_{\boldsymbol{B}} \exp(-\boldsymbol{a}_n\boldsymbol{\beta}\boldsymbol{H}_n(\boldsymbol{Q}) - \boldsymbol{a}_n\boldsymbol{\gamma}\boldsymbol{C}_n(\boldsymbol{Q}))\mathbf{\Pi}_n(d\boldsymbol{Q}), \quad (2.12)$$

where

$$\boldsymbol{Z}_n(\boldsymbol{\beta},\boldsymbol{\gamma}) \doteq \int_{\mathbf{\Omega}_n} \exp(-\boldsymbol{a}_n\boldsymbol{\beta}\boldsymbol{H}_n(\boldsymbol{Q}) - \boldsymbol{a}_n\boldsymbol{\gamma}\boldsymbol{C}_n(\boldsymbol{Q}))\mathbf{\Pi}_n(d\boldsymbol{Q}).$$

The real parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ correspond to the inverse temperature and the chemical potential.

2.3 Continuum Limit

The first set of main results of this paper concerns the continuum limits of the joint probability distributions (2.11) and (2.12). In this limit we obtain the continuum description of the potential vorticity that captures the large-scale behavior of the statistical equilibrium state. The key ingredient in this analysis is a coarse-graining of the potential vorticity field, defined by a local averaging of the random microstate $\boldsymbol{Q}$ over an intermediate spatial scale. This process is constructed as follows.

For $\boldsymbol{r} \in \mathbb{N}$ such that $\boldsymbol{r} < \boldsymbol{n}$, the flow domain $\mathbf{\Lambda}$ is uniformly partitioned into $\boldsymbol{m}_r$ macrocells denoted by $\boldsymbol{D}_{r,1}, \ldots, \boldsymbol{D}_{r,m_r}$. We assume that each macrocell satisfies $\boldsymbol{\lambda}\{\boldsymbol{D}_{r,k}\} = \mathbf{1}/\boldsymbol{m}_r$ and contains $\boldsymbol{a}_n/\boldsymbol{m}_r$ microcells $\boldsymbol{M}(\boldsymbol{s})$ and that

$$\boldsymbol{r} \xrightarrow{\lim} \infty \max_{\boldsymbol{k} \in \{\mathbf{1}, \ldots, \boldsymbol{m}_r\}} \operatorname{diam}(\boldsymbol{D}_{r,k}) = \mathbf{0}. \quad (2.13)$$

For instance, $\mathbf{\Lambda}$ can be dyadically refined into $\boldsymbol{a}_n = \mathbf{2}^{2n}$ microcells, and these can be collected into a partition of $\boldsymbol{m}_r = \mathbf{2}^{2r}$ macrocells $\boldsymbol{D}_{r,k}$ satisfying (2.13). With respect to the collection of macrocells $\boldsymbol{D}_{r,k}$, we introduce a doubly-indexed process $\boldsymbol{W}_{n,r} = \boldsymbol{W}_{n,r}(\boldsymbol{Q},\boldsymbol{x})$, defined for $\boldsymbol{Q} \in \mathbf{\Omega}_n$ and $\boldsymbol{x} \in \mathbf{\Lambda}$ by

$$\boldsymbol{W}_{n,r}(\boldsymbol{Q},\boldsymbol{x}) \doteq \sum_{k=1}^{m_r} \boldsymbol{S}_{n,r,k}\mathbf{1}_{\boldsymbol{D}_{r,k}}(\boldsymbol{x}), \quad (2.14)$$

where

$$\boldsymbol{S}_{n,r,k} = \boldsymbol{S}_{n,r,k}(\boldsymbol{Q}) \doteq \tfrac{1}{\boldsymbol{a}_n/\boldsymbol{m}_r} \sum_{\boldsymbol{s} \in \boldsymbol{D}_{r,k}} \boldsymbol{Q}(\boldsymbol{s}). \quad (2.15)$$

For $Q \in \Omega_n$ the piecewise-constant process $W_{n,r}(Q, \cdot)$ takes values in $L^2(\Lambda)$ and can be studied with respect to either the microcanonical ensemble (2.11) or the canonical ensemble (2.12). We refer to $W_{n,r}$ as the coarse-grained process.

With respect to each ensemble, our goal is to determine the set of L^2 functions on which $W_{n,r}$ concentrates in the continuum limit defined by first sending $n \to \infty$ and then $r \to \infty$. More precisely, we seek to find the smallest subsets of $L^2(\Lambda)$, $\mathcal{E}^{E,\Gamma}$ and $\mathcal{E}_{\beta,\gamma}$, such that

$$P_{n,\kappa}^{E,\Gamma}\{W_{n,r} \in \mathcal{E}^{E,\Gamma}\} \to 1 \text{ as } n \to \infty, r \to \infty, \text{ and } \kappa \to 0 \tag{2.16}$$

and

$$P_{n,\beta,\gamma}\{W_{n,r} \in \mathcal{E}_{\beta,\gamma}\} \to 1 \text{ as } n \to \infty \text{ and } r \to \infty. \tag{2.17}$$

The subsets $\mathcal{E}^{E,\Gamma}$ and $\mathcal{E}_{\beta,\gamma}$ consist of the most probable states for their respective ensembles and are referred to as the sets of equilibrium macrostates. In Section 4, each of these subsets is shown to arise by solving a variational principle. The variational principles are derived from the LDP's for $W_{n,r}$ with respect to the two ensembles. In order to prove these two LDP's, we first prove a basic LDP for $W_{n,r}$ with respect to the product measure Π_n defined in (2.10). This is done in the next section.

3. Basic LDP for the Coarse-Grained Process

Before stating the basic LDP for $W_{n,r}$ with respect to Π_n, we first discuss the form of the rate function I. Throughout the paper the term rate function is used to describe any lower semicontinuous function mapping a complete, separable metric space into $[0, \infty]$.

For $\alpha \in \mathbb{R}$ we define the cumulant generating function

$$c(\alpha) \doteq \log \int_{\mathbb{R}} e^{\alpha y} \rho(dy),$$

which is finite because ρ has compact support, and for $y \in \mathbb{R}$ we introduce the lower semi-continuous, conjugate convex function

$$i(y) \doteq \sup_{\alpha \in \mathbb{R}} \{\alpha y - c(\alpha)\}.$$

The normalized sums $S_{n,r,k}$ defined in (2.15) are sample means of the a_n/m_r i.i.d. random variables $Q(s)$, $s \in D_{r,k}$. Therefore, Cramér's theorem [5] implies that, for each $k = 1, \ldots, m_r$, the sequence $\{S_{n,r,k}, n \in \mathbb{N}\}$ satisfies an LDP with respect to Π_n with scaling constants a_n/m_r and rate function i. In other words, for any closed subset F of $\mathbb{R}$

$$\limsup_{n \to \infty} \tfrac{1}{a_n/m_r} \log \Pi_n\{S_{n,r,k} \in F\} \le -\inf_{y \in F} i(y),$$

and for any open subset G of $\mathbb{R}$

$$\liminf_{n \to \infty} \tfrac{1}{a_n/m_r} \log \Pi_n\{S_{n,r,k} \in G\} \ge -\inf_{y \in G} i(y).$$

The rate function for the LDP of $W_{n,r}$ with respect to Π_n is defined for $q \in L^2(\Lambda)$ to be

$$I(q) \doteq \int_\Lambda i(q(x))\lambda(dx). \tag{3.1}$$

Since i is nonnegative and convex, it follows that I is well-defined, nonnegative, and convex. We claim that I is lower semicontinuous with respect to the strong topology on $L^2(\Lambda)$. Assume that $q_n \to q$ in $L^2(\Lambda)$. Then there exists a subsequence q_{n_k} such that $\lim_{n_k\to\infty} I(q_{n_k}) = \liminf_{n\to\infty} I(q_n)$ and $q_{n_k} \to q$ a.s. Fatou's lemma and the lower semicontinuity of i on $\mathbb{R}$ yield the desired conclusion:

$$\liminf_{n\to\infty} I(q_n) = \liminf_{n_k\to\infty} \int_\Lambda i(q_{n_k}(x))\lambda(dx)$$

$$\geq \int_\Lambda \liminf_{n_k\to\infty} i(q_{n_k}(x))\lambda(dx) \geq \int_\Lambda i(q(x))\lambda(dx) = I(q).$$

We now state the two-parameter LDP for $W_{n,r}$ with respect to Π_n. For any subset A of $L^2(\Lambda)$, $I(A)$ denotes the infimum of I over A.

Theorem 3.1: *Consider $L^2(\Lambda)$ with the strong topology. The sequence $W_{n,r}$ satisfies the LDP on $L^2(\Lambda)$, in the double limit $n \to \infty$ and $r \to \infty$, with rate function I. That is, for any strongly closed subset F of $L^2(\Lambda)$*

$$\limsup_{r\to\infty} \limsup_{n\to\infty} \frac{1}{a_n} \log \Pi_n\{W_{n,r} \in F\} \leq -I(F),$$

and for any strongly open subset G of $L^2(\Lambda)$

$$\liminf_{r\to\infty} \liminf_{n\to\infty} \frac{1}{a_n} \log \Pi_n\{W_{n,r} \in G\} \geq -I(G).$$

We now prove the upper and lower large deviation bounds for $W_{n,r}$ in separate, but elementary steps. The following lemma is needed in both steps; it is an immediate consequence of Lemmas 2.5 — 2.8 in [17].

Lemma 3.2: *For each r the sequence $\{(S_{n,r,1}, \ldots, S_{n,r,m_r}), n \in \mathbb{N}\}$ satisfies the LDP on $\mathbb{R}^{m_r}$ with scaling constants a_n/m_r and the rate function*

$$(\nu_1, \ldots, \nu_m) \mapsto \sum_{k=1}^{m_r} i(\nu_k).$$

Proof of the Large Deviation Upper Bound: We define L_r^2 to be the set of $q \in L^2(\Lambda)$ of the form $q(x) = \sum_{k=1}^{m_r} \nu_k 1_{D_{r,k}}(x)$ for some $\nu_1, \ldots, \nu_{m_r} \in \mathbb{R}$. Let F be a strongly closed subset of $L^2(\Lambda)$. For $r \in \mathbb{N}$

$$F_r = \{(\nu_1, \ldots, \nu_{m_r}) \in \mathbb{R}^{m_r} : \sum_{k=1}^{m_r} \nu_k 1_{D_{r,k}}(x) \in F\}$$

is a closed subset of $\mathbb{R}^{m_r}$. By Lemma 3.2, it follows that

$$\limsup_{n\to\infty} \frac{1}{a_n} \log \Pi_n\{W_{n,r} \in F\}$$

$$= \tfrac{1}{m_n} \limsup_{n \to \infty} \tfrac{m_r}{a_n} \log \Pi_n \{(S_{n,r,1}, \ldots, S_{n,r,m_r}) \in F_r\}$$

$$\leq -\tfrac{1}{m_r} \inf \left\{ \sum_{k=1}^{m_r} i(\nu_k) \colon (\nu_1, \ldots, \nu_{m_r}) \in F_r \right\}$$

$$= -\inf \left\{ \int_\Lambda i(q(x)) \lambda(dx) \colon q \in F \cap L_r^2 \right\}$$

$$\leq -\inf \left\{ \int_\Lambda i(q(x)) \lambda(dx) \colon q \in F \right\}$$

$$= -I(F).$$

Sending $r \to \infty$ gives the desired large deviation upper bound. □

Proof of the Large Deviation Lower Bound: The following lemma is needed both here and later in the paper. We denote by $\| \cdot \|$ the norm on $L^2(\Lambda)$.

Lemma 3.3: *Let q be any function in $L^2(\Lambda)$. For $r \in \mathbb{N}$ define*

$$q_r(x) \doteq \sum_{k=1}^{m_r} q_r(k) 1_{D_{r,k}}(x), \text{ where } q_r(k) \doteq m_r \int_{D_{r,k}} q(x) \lambda(dx).$$

Then as $r \to \infty$, $\| q - q_r \| \to 0$.

Proof: For any given $\epsilon > 0$ there exists a bounded Lipschitz function φ mapping Λ into $\mathbb{R}$ and satisfying $\| q - \varphi \| \leq \varepsilon$ [7]. Since the operator mapping $f \in L^2(\Lambda) \mapsto f_r$ is an orthogonal projection,

$$\| q_r - \varphi_r \| = \| (q - \varphi)_r \| \leq \| q - \varphi \| < \varepsilon. \tag{3.2}$$

Hence it suffices to estimate $\| \varphi - \varphi_r \|$, where φ is a Lipschitz function with constant $M < \infty$. A straightforward calculation coupled with (2.13) gives

$$\| \varphi - \varphi_r \| \leq M \cdot \left(\max_{k \in \{1, \ldots, m_r\}} \{\text{diam}(D_{r,k})\} \right) \to 0 \text{ as } r \to \infty. \tag{3.3}$$

This completes the proof. □

Given a strongly open subset G of $L^2(\Lambda)$, let q be any function in G and choose $\varepsilon > 0$ so that $B(q, \varepsilon)$, the ball in $L^2(\Lambda)$ with center q and radius ε, is a subset of G. Since $\| q - q_r \| \to 0$ as $r \to \infty$ (Lemma 3.3), we can also choose $N \in \mathbb{N}$ such that for all $r \in N$ $B(q_r, \varepsilon/2) \subset B(q, \varepsilon)$. Since

$$G_{r,\varepsilon} \doteq \left\{ (\nu_1, \ldots, \nu_{m_r}) \in \mathbb{R}^{m_r} \colon \sum_{k=1}^{m_r} \nu_k 1_{D_{r,k}}(x) \in B(q_r, \epsilon/2) \right\}$$

is an open subset of $\mathbb{R}^{m_r}$, it follows from Lemma 3.2 that for all $r \geq N$

$$\begin{aligned}
&\liminf_{n\to\infty} \tfrac{1}{a_n}\log\Pi_n\{W_{n,r} \in G\} \\
&\geq \liminf_{n\to\infty} \tfrac{1}{a_n}\log\Pi_n\{W_{n,r} \in B(q_r, \varepsilon/2)\} \\
&= \tfrac{1}{m_r}\liminf_{n\to\infty} \tfrac{m_r}{a_n}\log\Pi_n\{(S_{n,r,1}, \ldots, S_{n,r,m_r}) \in G_{r,\varepsilon}\} \\
&\geq -\tfrac{1}{m_r}\inf\left\{\sum_{k=1}^{m_r} i(\nu_k)\colon (\nu_1, \ldots, \nu_{m_r}) \in G_{r,\varepsilon}\right\} \\
&\geq -\tfrac{1}{m_r}\sum_{k=1}^{m_r} i(q_r(k)) = -\tfrac{1}{m_r}\sum_{k=1}^{m_r} i\left(m_r \int_{D_{r,k}} q(x)\lambda(dx)\right) \\
&\geq -\sum_{k=1}^{m_r}\int_{D_{r,k}} i(q(x))\lambda(dx) = -\int_{\Lambda} i(q(x))\lambda(dx) = -I(q).
\end{aligned}$$

The last inequality in this display follows from Jensen's inequality. We have proved that for arbitrary $q \in G$

$$\liminf_{r\to\infty}\liminf_{n\to\infty} \tfrac{1}{a_n}\log\Pi_n\{W_{n,r} \in G\} \geq -I(q).$$

Taking the supremum of $-I(q)$ over $q \in G$ yields the desired large deviation lower bound□

This completes the proof of Theorem 3.1.

4. LDP's with Respect to the Two Ensembles

In this section we present LDP's for $W_{n,r}$ with respect to the microcanonical and canonical ensembles defined in (2.11) and (2.12). The theorems follow from a general theory presented in [9]. The proofs of the LDP's depend in part on properties of various functionals given in the next subsection.

4.1 Properties of H_n, C_n, H, and C

We recall the definitions of the lattice energy H_n in (2.7), the circulation C_n in (2.8), and the functionals H and C in (2.4) and (2.5). The proofs of the LDP's with respect to the two ensembles rely on boundedness and continuity properties of $H(q)$ and $C(q)$ for $q \in L^2(\Lambda)$ and on the fact that uniformly over Ω_n, H_n and C_n are asymptotic to $H(W_{n,r})$ and $C(W_{n,r})$; the latter two quantities are defined by replacing $Q(x)$ in (2.4) and (2.5) by the $L^2(\Lambda)$-valued process $W_{n,r}(Q, x)$. The proof of this approximation for the circulation is immediate since it holds with equality. For H_n the proof depends on the fact that the vortex interactions are long-range, being determined by the Green's function $g(x, x')$. For this reason, H_n is well approximated by a function of the coarse-grained process $W_{n,r}$. In physical parlance, the turbulence model under consideration is an

asymptotically exact, local mean-field theory [20]. The next lemma proves the approximations as well as the boundedness and continuity properties of H and C needed for the LDP's.

Lemma 4.1: (a) *The sequence* $W_{n,r}$ *satisfies*

$$\lim_{r \to \infty} \lim_{n \to \infty} \sup_{Q \in \Omega_n} | H_n(Q) - H(W_{n,r}(Q)) | = 0 \tag{4.1}$$

and

$$\lim_{r \to \infty} \lim_{n \to \infty} \sup_{Q \in \Omega_n} | C_n(Q) - C(W_{n,r}(Q)) | = 0. \tag{4.2}$$

(b) H *and* C *are bounded on the union, over* $1 \le r < n$*, of the supports of the* Π_n*-distributions of* $W_{n,r}$.

(c) H *and* C *are continuous with respect to the weak topology on* $L^2(\Lambda)$ *and thus with respect to the strong topology on* $L^2(\Lambda)$.

Proof: (a) A short calculation shows that $C_n(Q) = C(W_{n,r}(Q))$ for all $Q \in \Omega_n$, thus proving (4.2). Carrying out the multiplication in the summand appearing in the definition of $H_n(Q)$, we write $H_n(Q) = \sum_{i=1}^4 H_n^{(i)}(Q)$, where $H_n^{(1)}$ has the summands $Q(s)g_n(s,s')Q(s')$, $H_n^{(2)}$ has the summands $Q(s)g_n(s,s')b_n(s')$, $H_n^{(3)}$ has the summands $b_n(s)g_n(s,s')Q(s')$, and $H_n^{(4)}$ has the summands $b_n(s)g_n(s,s')b_n(s')$. Similarly, carrying out the multiplication in the integrand appearing in the definition of

$$H(W_{n,r}(Q)) \doteq \frac{1}{2} \int_{\Lambda \times \Lambda} (W_{n,r}(Q,x) - b(x)) g(x,x') (W_{n,r}(Q,x') - b(x')) \lambda(dx) \lambda(x').$$

we write

$$H(W_{n,r}(Q)) = \sum_{i=1}^4 H^{(i)}(W_{nr}(Q)),$$

where $H^{(1)}(W_{n,r}(Q))$ has the integrand $W_{n,r}(Q,x)g(x,x')W_{n,r}(Q,x')$, $H^{(2)}(W_{n,r}(Q))$ has the integrand $W_{n,r}(Q,x)g(x,x')b(x')$, $H^{(3)}(W_{n,r}(Q))$ has the integrand $b(x)g(x,x')W_{n,r}(Q,x')$, and $H^{(4)}(W_{n,r}(Q))$ has the integrand $b(x)g(x,x')b(x')$.

The limit (4.1) follows if we prove for $i = 1,2,3,4$ that

$$\lim_{r \to \infty} \lim_{n \to \infty} \sup_{Q \in \Omega_n} | H_n^{(i)}(Q) - H^{(i)}(W_{n,r}(Q)) | = 0. \tag{4.3}$$

The details of the proof are given for $i = 1$; the cases $i = 2,3,4$ are handled similarly. The verification of (4.3) for $i = 1$ relies on the fact that since ρ has compact support $\mathcal{Y}$, there exists $C < \infty$ such that $| Q(s) | \le C$ for any $n \in \mathbb{N}$, $Q \in \Omega_n$, and $s \in \Lambda_n$.

For $1 \le k, k' \le m_r$ and x and x' in Λ we define

$$g_r(x,x') \doteq \sum_{k,k'=1}^{m_r} g_r(k,k') 1_{D_{r,k}}(x) 1_{D_{r,k'}}(x'), \tag{4.4}$$

where

$$g_r(k,k') \doteq m_r^2 \int\limits_{D_{r,k}} \int\limits_{D_{r,k'}} g(x,x')\lambda(dx)\lambda(dx'). \tag{4.5}$$

Substituting the definition (2.14) — (2.15) of $W_{n,r}(Q,x)$, we obtain

$$H^{(1)}(W_{n,r}(Q)) = -\frac{1}{2}\sum_{k,k'=1}^{m_r}\frac{m_r^2}{a_n^2}\sum_{\substack{s\in D_{r,k}\\ s'\in D_{r,k'}}} Q(s)Q(s')\int\limits_{D_{r,k}}\int\limits_{D_{r,k'}} g(x,x')\lambda(dx)\lambda(dx')$$

$$= -\frac{1}{2}\sum_{k,k'=1}^{m_r}\sum_{\substack{s\in D_{r,k}\\ s'\in D_{r,k'}}} \frac{Q(s)Q(s')g_r(k,k')}{a_n^2}.$$

Thus for any $1 \le r < n$,

$$\sup_{Q\in\Omega_n} | H_n^{(1)}(Q) - H^{(1)}(W_{n,r}(Q)) |$$

$$= \sup_{Q\in\Omega_n} \frac{1}{2}\left|\sum_{k,k'=1}^{m_r}\sum_{\substack{s\in D_{r,k}\\ s'\in D_{r,k'}}} Q(s)Q(s')\left[\frac{g_n(s,s')-g_r(k,k')}{a_n^2}\right]\right|$$

$$\le \frac{C^2}{2}\sum_{k,k'=1}^{m_r}\sum_{\substack{s\in D_{r,k}\\ s'\in D_{r,k'}}} \frac{|g_n(s,s')-g_r(k,k')|}{a_n^2}$$

$$= \frac{C^2}{2}\sum_{k,k'=1}^{m_r}\sum_{\substack{s\in D_{r,k}\\ s'\in D_{r,k'}}} \left|\int\limits_{M(s)}\int\limits_{M(s')}[g(x,x')-g_r(x,x')]\lambda(dx)\lambda(dx')\right|$$

$$\le \frac{C^2}{2}\int\limits_{\Lambda}\int\limits_{\Lambda} | g(x,x')-g_r(x,x') | \lambda(dx)\lambda(dx')$$

$$= \frac{C^2}{2} \| g-g_r \|_{L^1(\Lambda\times\Lambda)} \le \frac{C^2}{2} \| g-g_r \|_{L^2(\Lambda\times\Lambda)}.$$

The Green's function $g(x, x')$ is square-integrable on $\Lambda \times \Lambda$. Hence for the set $\Lambda \times \Lambda$ with the partition $D_{r,k} \times D_{r,k'}$, the function g_r defined in (4.4) – (4.5) is the analogue of q_r defined in Lemma 3.3. By an analogous proof, one shows that $\| g - g_r \|_{L^2(\Lambda\times\Lambda)} \to 0$. This limit and the preceding display yield the desired limit (4.3) in the case $i = 1$.

(b) For any $1 \le r < n$, $Q \in \Omega_n$, and $s \in \Lambda_n$, we have $| Q(s) | \le C$ and thus $\sup_{x\in\Lambda} | W_{n,r}(Q, x) | \le C$. Hence the range of the support of the Π_n-distribution of $W_{n,r}$ is a subset of $\{q \in L^2(\Lambda): \| q \|_\infty \le C\}$. Since b is bounded on Λ and $g(x, x') \in L^2(\Lambda \times \Lambda)$, it follows that H and C are bounded on the union of the supports of the Π_n-distributions of $W_{n,r}$.

(c) The weak continuity of $C(q) = \int_\Lambda (q - b) d\lambda$ is obvious. If $q^n \to q$ weakly in $L^2(\Lambda)$, then by a property of the Green's function

$$\int_\Lambda g(\,\cdot\,, x')(q^n(x') - b(x'))\lambda(dx') \to \int_\Lambda g(\,\cdot\,, x')(q(x') - b(x'))\lambda(dx') \text{ strongly in } L^2(\Lambda).$$

It follows that $H(q^n) \to H(q)$. The proof of the lemma is complete. □

4.2 Microcanonical Model

The LDP for $W_{n,r}$ with respect to the microcanonical ensemble is stated in Theorem 4.2. We say that a constraint pair (E, Γ) is admissible if $(E, \Gamma) = (H(q), C(q))$ for some $q \in L^2(\Lambda)$ with $I(q) < \infty$, and we let $\mathcal{A}$ denote the largest open subset of $\mathbb{R}^2$ consisting of admissible constraint pairs. We call this domain $\mathcal{A}$ the admissible set for the microcanonical model. For a fixed constraint pair $(E, \Gamma) \in \mathcal{A}$, the rate function for the LDP is defined for $q \in L^2(\Lambda)$ to be

$$I^{E,\Gamma}(q) \doteq \begin{cases} I(q) + S(E, \Gamma) & \text{if } H(q) = E, C(q) = \Gamma, \\ \infty & \text{otherwise,} \end{cases} \tag{4.6}$$

where $S(E, \Gamma)$ is defined by

$$S(E, \Gamma) \doteq -\inf_{\widetilde{q}} \{I(\widetilde{q}\,): H(\widetilde{q}\,) = E, C(\widetilde{q}\,) = \Gamma\}. \tag{4.7}$$

$S(E, \Gamma)$ is called the microcanonical entropy.

Theorem 4.2: *Let $(E, \Gamma) \in \mathcal{A}$ and consider $L^2(\Lambda)$ with the strong topology. With respect to $P^{E,\Gamma}_{n,\kappa}$, $W_{n,r}$ satisfies the LDP on $L^2(\Lambda)$, in the triple limit $n \to \infty$, $r \to \infty$ and $\kappa \to 0$, with rate function $I^{E,\Gamma}$ defined in (4.6). That is, for any strongly closed subset F of $L^2(\Lambda)$*

$$\lim_{\kappa \to 0} \limsup_{r \to \infty} \limsup_{n \to \infty} \tfrac{1}{a_n} \log P^{E,\Gamma}_{n,\kappa}\{W_{n,r} \in F\} \le -I^{E,\Gamma}(F),$$

and for any strongly open subset G of $L^2(\Lambda)$

$$\lim_{\kappa \to 0} \liminf_{r \to \infty} \liminf_{n \to \infty} \tfrac{1}{a_n} \log P^{E,\Gamma}_{n,\kappa}\{W_{n,r} \in G\} \ge -I^{E,\Gamma}(G).$$

Proof: This LDP is a three-parameter analogue, involving the triple limit $n \to \infty$, $r \to \infty$, and $\kappa \to 0$, of the two-parameter LDP given in Theorem 3.2 of [9]. The proof of that theorem is easily modified to the present three-parameter setting using the basic LDP in Theorem 3.1 and the properties of H_n, C_n, H, and C given in Lemma 4.1. □

For any admissible choice of E and Γ we define the set $\mathcal{E}^{E,\Gamma}$ of microcanonical equilibrium macrostates to be the set of $q \in L^2(\Lambda)$ such that $I^{E,\Gamma}(q) = 0$. It follows from the definition of $I^{E,\Gamma}$ that elements of $\mathcal{E}^{E,\Gamma}$ are exactly those functions that solve the following constrained minimization problem:

$$\begin{gathered}\text{minimize } I(\tilde{q}\,) \text{ over all } \tilde{q} \in L^2(\Lambda) \text{ subject to the constraints} \\ H(\tilde{q}\,) = E \text{ and } C(\tilde{q}\,) = \Gamma.\end{gathered} \tag{4.8}$$

The optimum value in this variational principle determines the microcanonical entropy S defined in (4.7).

We point out an important consequence of the large deviation upper bound in Theorem 4.2. Let A be a Borel subset of $L^2(\Lambda)$ whose closure $\overline{A}$ satisfies $\overline{A} \cap \mathcal{E}^{E,\Gamma} = \emptyset$. Then $I^{E,\Gamma}(\overline{A}\,) > 0$, and for all sufficiently small $\kappa > 0$ and all sufficiently large r and n

$$P_{n,\kappa}^{E,\Gamma}\{W_{n,r} \in A\} \leq \exp(-a_n I^{E,\Gamma}(\overline{A}\,)/2). \tag{4.9}$$

Since the right-hand side of (4.9) converges to 0 as $n \to \infty$, macrostates Q not lying in $\mathcal{E}^{E,\Gamma}$ have an exponentially small probability of being observed as a coarse-grained state in the continuum limit. The macrostates in $\mathcal{E}^{E,\Gamma}$ are therefore the overwhelmingly most probable among all possible macrostates of the turbulent system.

4.3 Canonical Model

The LDP for $W_{n,r}$ with respect to the canonical ensemble is stated in Theorem 4.3. For any real values of β and γ, the rate function for the LDP is defined for $q \in L^2(\Lambda)$ to be

$$I_{\beta,\gamma}(q) \doteq I(q) + \beta H(q) + \gamma C(q) - \Phi(\beta,\gamma), \tag{4.10}$$

where $\Phi(\beta,\gamma)$ is defined by

$$\Phi(\beta,\gamma) = \inf_{\tilde{q}} \{I(\tilde{q}\,) + \beta H(\tilde{q}\,) + \gamma C(\tilde{q}\,)\}. \tag{4.11}$$

$\Phi(\beta,\gamma)$ is called the canonical free energy.

Theorem 4.3: *Let $(\beta,\gamma) \in \mathbb{R}^2$ and consider $L^2(\Lambda)$ with the strong topology. With respect to $P_{n,\beta,\gamma}$, $W_{n,r}$ satisfies the LDP on $L^2(\Lambda)$, in the double limit $n \to \infty$ and $r \to \infty$, with rate function $I_{\beta,\gamma}$ defined in* (4.10). *That is, for any strongly closed subset F of $L^2(\Lambda)$*

$$\limsup_{r \to \infty} \limsup_{n \to \infty} \tfrac{1}{a_n} \log P_{n,\beta,\gamma}\{W_{n,r} \in F\} \leq -I_{\beta,\gamma}(F),$$

and for any strongly open subset G of $L^2(\Lambda)$

$$\liminf_{r \to \infty} \liminf_{n \to \infty} \tfrac{1}{a_n} \log P_{n,\beta,\gamma}\{W_{n,r} \in G\} \geq -I_{\beta,\gamma}(G),$$

Proof: This LDP is a two-parameter analogue, involving the double limit $n \to \infty$ and $r \to \infty$, of the one-parameter LDP given in Theorem 2.4 of [9]. The proof of that theorem relies on an application of Laplace's principle. The proof is easily modified to the present

two-parameter setting using the basic LDP in Theorem 3.1 and the properties of H_n, C_n, H, and C given in Lemma 4.1. As is shown in Theorem 1.3.4 of [8], the boundedness properties of H and C given in Lemma 4.1(b) allow Laplace's principle to be applied in this setting. □

For a particular choice of β and γ we define the set $\mathcal{E}_{\beta,\gamma}$ of canonical equilibrium macrostates to be the set of $q \in L^2(\Lambda)$ such that $I_{\beta,\gamma}(q) = 0$. That is, the elements of $\mathcal{E}_{\beta,\gamma}$ are exactly those functions that solve the following unconstrained minimization problem:

$$\text{minimize } I(\tilde{q}) + \beta H(\tilde{q}) + \gamma C(\tilde{q}) \text{ over all } \tilde{q} \in L^2(\Lambda). \tag{4.12}$$

The optimum value in this variational principle determines the canonical free energy $\Phi(\beta, \gamma)$ defined in (4.11). Comparing this minimization problem with the corresponding constrained minimization problem (4.8) arising in the microcanonical ensemble, we see that the parameters β and γ in the former are Lagrange multipliers dual to the two constraints in the latter.

As in the microcanonical case (see (4.9) and the associated discussion), the large deviation upper bound in Theorem 4.3 implies that the macrostates in $\mathcal{E}_{\beta,\gamma}$ are overwhelmingly the most probable among all possible macrostates of the turbulent system.

5. Equivalence of Ensembles

In this section we classify the possible relationships between the sets of equilibrium macrostates $\mathcal{E}^{E,\Gamma}$ and $\mathcal{E}_{\beta,\gamma}$. We begin by establishing that each of these equilibrium sets is nonempty. Then we turn to our second set of main results, which concern the equivalence or lack of equivalence between the microcanonical and canonical ensembles. In many statistical mechanical models it is common for the microcanonical and canonical ensembles to be equivalent, in the sense that there is a one-to-one correspondence between their equilibrium macrostates. However, in models of coherent structures in turbulence there can be microcanonical equilibria that are not realized as canonical equilibria. As is shown in our companion paper [10], and in a real physical application in [31], nonequivalence occurs often in the turbulence models and produces some of the most interesting coherent mean flows.

5.1 Existence of Equilibrium States

We continue to assume that the support $\mathcal{Y}$ of ρ is compact. It follows from this assumption that the $I(q)$ grows faster than quadratically as $\| q \| \to \infty$; the straightforward proof is left to the reader. Consequently, for any $M < \infty$ the sets $K_M \doteq \{q \in L^2(\Lambda): I(q) \le M\}$ are precompact with respect to the weak topology of $L^2(\Lambda)$; this follows from the fact that each K_M is a subset of a closed ball in $L^2(\Lambda)$. Moreover, $I(q)$ is lower semicontinous with respect to the weak topology on $L^2(\Lambda)$. This property is an immediate consequence of the relationship

$$I(q) = \sup_{r \in \mathbb{N}} I(q_r), \tag{5.1}$$

where q_r is the piecewise constant approximating function defined in Lemma 3.3. Indeed, since each function $q \mapsto I(q_r)$ is weakly lower semicontinuous, the weak lower semicontinuity of $I(q)$ follows. Thus the sets K_M are closed with respect to the weak topology. This property, in combination with the weak precompactness of these sets, implies that the sets K_M are compact with respect to the weak topology. Concerning the proof of (5.1), Jensen's inequality guarantees that $I(q_r) \le I(q)$, while Lemma 3.3 and the strong lower

semicontinuity of I proved right after its definition in (3.1) yield $\liminf_{r\to\infty} I(q_r) \geq I(q)$; (5.1) is thus proved.

The existence of microcanonical and canonical equilibrium states now follows immediately from the direct methods of the calculus of variations. In the microcanonical case, for any admissible $(E,\Gamma) \in \mathcal{A}$ a minimizing sequence exists that converges weakly to a minimizer $q \in \mathcal{E}^{E,\Gamma}$, by virtue of the properties of the objective functional I just derived and the continuity, with respect to the weak topology, of the constraint functionals H and C (Lemma 4.1(c)). In the canonical case, since I grows superquadratically as $\| q \| \to \infty$ while H grows quadratically and C linearly, it follows that for any $(\beta,\gamma) \in \mathbb{R}^2$ a minimizing sequence exists that converges weakly to a minimizer $q \in \mathcal{E}_{\beta,\gamma}$. The details of this routine analysis are omitted.

Typically, the solutions of the variational problems (4.8) and (4.12) for the microcanonical and canonical ensembles, respectively, are expected to be unique. Indeed, numerical solutions of these problems, such as those carried out in [10], demonstrate that apart from degeneracies and bifurcations the equilibrium sets $\mathcal{E}^{E,\Gamma}$ and $\mathcal{E}_{\beta,\gamma}$ are singleton sets. In the next subsection, without any special assumptions concerning uniqueness of equilibrium solutions for either model, we give complete and general results about the correspondence between these sets. In addition, the results given in the next subsection require no continuity or smoothness assumptions of the equilibrium solutions with respect to the model parameters (E,Γ) or (β,γ).

5.2 Dual Thermodynamic Functions

In the analysis to follow, we show how the properties of the thermodynamic functions for the microcanonical and canonical ensembles determine the correspondence, or lack of correspondence, between equilibria for these two ensembles. For the microcanonical ensemble the thermodynamic function is the microcanonical entropy $S(E,\Gamma)$ defined in the constrained variational principle (4.7), the solutions of which constitute the equilibrium set $\mathcal{E}^{E,\Gamma}$. For the canonical ensemble the thermodynamic function is the free energy $\Phi(\beta,\gamma)$ defined in the unconstrained variational principle (4.11), the solutions of which constitute the equilibrium set $\mathcal{E}_{\beta,\gamma}$.

The basis for the equivalence of ensembles is the conjugacy between S and Φ; that is, Φ is the Legendre-Fenchel transform of S [13, 32]. This basic property is easily verified as follows:

$$\begin{aligned}
\Phi(\beta,\gamma) &= \min_{q}\{I(q) + \beta H(q) + \gamma C(q))\} \qquad (5.2)\\
&= \inf_{E,\Gamma} \min_{q}\{I(q) + \beta H(q) + \gamma C(q) : H(q) = E, C(q) = \Gamma\}\\
&= \inf_{E,\Gamma}\{\beta E + \gamma\Gamma - S(E,\Gamma)\}\\
&= S^*(\beta,\gamma).
\end{aligned}$$

Consequently, Φ is a concave function of (β,γ), which runs over $\mathbb{R}^2$. By contrast, S itself is not necessarily concave. The concave hull of S is furnished by the conjugate function of Φ, namely, $\Phi^* = S^{**}$, which satisfies the inequality

$$S(E,\Gamma) \leq \inf_{\beta,\gamma}\{\beta E + \gamma\Gamma - \Phi(\beta,\gamma)\} = S^{**}(E,\Gamma). \qquad (5.3)$$

The relationship between microcanonical equilibria and canonical equilibria depends crucially on the concavity properties of the microcanonical entropy S. To this end, we introduce the subset $\mathcal{C} \subseteq \mathcal{A}$ on which the concave hull S^{**} coincides with S; that is, $(E,\Gamma) \in \mathcal{C}$ if and only if $S^{**}(E,\Gamma) = S(E,\Gamma)$. Equivalently, $\mathcal{C}$ consists of those points $(E,\Gamma) \in \mathcal{A}$ for which there exists some $(\beta,\gamma) \in \mathbb{R}^2$ such that

$$S(E',\Gamma') \leq S(E,\Gamma) + \beta(E'-E) + \gamma(\Gamma'-\Gamma) \tag{5.4}$$

for all $(E',\Gamma') \in \mathbb{R}^2$. This condition means that S has a supporting plane, with normal determined by (β,γ), at the point (E,Γ). Such points (E,Γ) are precisely those points of $\mathcal{A}$ at which S has a nonempty superdifferential; this set consists of all (β,γ) for which (5.4) holds [13, 32]. As we will see in the next subsection, the concavity set $\mathcal{C}$ plays a pivotal role in the criteria for equivalence of ensembles.

5.3 Correspondence between Equilibrium Sets

The following theorem ensures that for constraint pairs in $\mathcal{C}$ the microcanonical equilibria are contained in a corresponding canonical equilibrium set, while for constraint pairs in $\mathcal{A}\backslash\mathcal{C}$ the microcanonical equilibria are not contained in any canonical equilibrium set.

Theorem 5.1:

(a) *If* $(E,\Gamma) \in \mathcal{A}$ *belongs to* $\mathcal{C}$*, then* $\mathcal{E}^{E,\Gamma} \subseteq \mathcal{E}_{\beta,\gamma}$ *for some* (β,γ).

(b) *If* $(E,\Gamma) \in \mathcal{A}$ *does not belong to* $\mathcal{C}$*, then* $\mathcal{E}^{E,\Gamma} \cap \mathcal{E}_{\beta,\gamma} = \emptyset$ *for all* (β,γ).

Proof: (a) If (E,Γ) belongs to $\mathcal{C}$, then there exists $(\beta,\gamma) \in \mathbb{R}^2$ such that (5.4) holds for all $(E',\Gamma') \in \mathbb{R}^2$. Thus for all (E',Γ')

$$\beta E + \gamma\Gamma - S(E,\Gamma) \leq \beta E' + \gamma\Gamma' - S(E',\Gamma').$$

By (5.2) and (4.11) this inequality implies that

$$\begin{aligned}\beta E + \gamma\Gamma - S(E,\Gamma) &\leq \inf_{(E',\Gamma')}\{\beta E' + \gamma\Gamma' - S(E',\Gamma')\} \\ &= \Phi(\beta,\gamma) = \min_{q}\{I(q) + \beta H(q) + \gamma C(q)\}.\end{aligned}$$

To show the claimed set containment, take any $\overline{q} \in \mathcal{E}^{E,\Gamma}$ and note that $H(\overline{q}) = E$, $C(\overline{q}) = \Gamma$, and $I(\overline{q}) = -S(E,\Gamma)$. Substituting these expressions into the preceding display yields

$$I(\overline{q}) + \beta H(\overline{q}) + \gamma C(\overline{q}) \leq \min_{q}\{I(q) + \beta H(q) + \gamma C(q)\}.$$

Since $\mathcal{E}_{\beta,\gamma}$ consists of the minimizers of $I + \beta H + \gamma C$, it follows that $\overline{q} \in \mathcal{E}_{\beta,\gamma}$. This completes the proof of (a).

(b) By (5.2) the hypothesis implies that for all $(\beta,\gamma) \in \mathbb{R}^2$

$$S(E,\Gamma) < \beta E + \gamma\Gamma - \Phi(\beta,\gamma).$$

Then any $\overline{q} \in \mathcal{E}^{E,\Gamma}$ satisfies

$$\begin{aligned}I(\overline{q}) + \beta H(\overline{q}) + \gamma C(\overline{q}) &= -S(E,\Gamma) + \beta E + \gamma\Gamma \\ &> \Phi(\beta,\gamma) = \min_{q}\{I(q) + \beta H(q) + \gamma C(q)\}.\end{aligned}$$

Thus $\overline{q}$ does not minimize $I + \beta H + \gamma C$ and hence does not belong to $\mathcal{E}_{\beta,\gamma}$. Since (β, γ) is arbitrary, this completes the proof of (b). □

According to this theorem, whenever $\mathcal{C} \neq \mathcal{A}$, there are microcanonical equilibria that are not realized by any canonical equilibria. On the other hand, all canonical equilibria are contained in some microcanonical equilibrium set, and $\mathcal{C}$ is exhausted by the constraint pairs realized by all canonical equilibria. These further results are given in the next theorem.

Theorem 5.2:

(a) *The concavity set $\mathcal{C}$ consists of all constraint pairs realized by the canonical equilibria; that is,*

$$\mathcal{C} = \bigcup\{H(\mathcal{E}_{\beta,\gamma}), C(\mathcal{E}_{\beta,\gamma})) : (\beta, \gamma) \in \mathbb{R}^2\}. \tag{5.5}$$

(b) *Each canonical equilibrium set $\mathcal{E}_{\beta,\gamma}$ consists of all microcanonical equilibria whose constraint pairs are realized by $\mathcal{E}_{\beta,\gamma}$; that is, for any (β, γ)*

$$\mathcal{E}_{\beta,\gamma} = \bigcup\{\mathcal{E}^{E,\Gamma} : (E, \Gamma) \in (H(\mathcal{E}_{\beta,\gamma}), C(\mathcal{E}_{\beta,\gamma}))\}. \tag{5.6}$$

Proof: (a) The containment of $\mathcal{C}$ in the union is immediate from Theorem 5.1(a). To show the opposite containment we argue by contradiction, supposing that for some (β, γ) and some $\overline{q} \in \mathcal{E}_{\beta,\gamma}$, $(E, \Gamma) = (H(\overline{q}), C(\overline{q})) \in \mathcal{A} \backslash \mathcal{C}$. Then by (5.3) we find that

$$S(E, \Gamma) < \beta E + \gamma \Gamma - \Phi(\beta, \gamma) = -I(\overline{q}) \leq S(E, \Gamma).$$

We thus obtain the desired contradiction. This completes the proof of part (a).

(b) The containment of $\mathcal{E}_{\beta,\gamma}$ in the union is straightforward. Let $\overline{q} \in \mathcal{E}_{\beta,\gamma}$ and set $E = H(\overline{q})$ and $\Gamma = C(\overline{q})$. Then $I(\overline{q}) + \beta E + \gamma \Gamma \leq I(q) + \beta H(q) + \gamma C(q)$ for all q. For those q that satisfy the constraints $H(q) = E$, $C(q) = \Gamma$, we therefore find that $I(\overline{q}) \leq I(q)$. Hence $\overline{q} \in \mathcal{E}^{E,\Gamma}$.

The opposite containment is also straightforward. If $E = H(\tilde{q})$ and $\Gamma = C(\tilde{q})$ for some $\tilde{q} \in \mathcal{E}_{\beta,\gamma}$, then for any $\overline{q} \in \mathcal{E}^{E,\Gamma}$ we have $I(\overline{q}) \leq I(\tilde{q})$. Since $\tilde{q} \in \mathcal{E}_{\beta,\gamma}$, we obtain

$$\min_q\{I(q) + \beta H(q) + \gamma C(q)\} = I(\tilde{q}) + \beta E + \gamma \Gamma$$

$$\geq I(\overline{q}) + \beta H(\overline{q}) + \gamma C(\overline{q}).$$

Hence $\overline{q} \in \mathcal{E}_{\beta,\gamma}$. This completes the proof of part (b). □

Theorems 5.1 and 5.2 allow us to classify the microcanonical constraint parameters (E, Γ) according to whether or not equivalence of ensembles holds for those parameters. In fact, the admissible set $\mathcal{A}$ can be decomposed into three disjoint sets, where

(1) there is a one-to-one correspondence between microcanonical and canonical equilibria,

(2) there is a many-to-one correspondence from microcanonical equilibria to canonical equilibria, and

(3) there is no correspondence.

In order to simplify the precise statement of this classification, let us assume that the microcanonical entropy $S(E, \Gamma)$ is differentiable on its domain $\mathcal{A}$. Then, for each microcanonical parameter $(E, \Gamma) \in \mathcal{A}$ there is a corresponding canonical parameter (β, γ) determined locally by the familiar thermodynamic relations

$$\beta = \frac{\partial S}{\partial E}, \quad \gamma = \frac{\partial S}{\partial \Gamma}.$$

Under this assumption, we have the following classification, which is a consequence of Theorems 5.1 and 5.2.

(1) *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 $E_{\beta,\gamma}$.

(2) *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.

(3) *Nonequivalence.* If (E, Γ) does not belonging 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.

For a complete discussion of results of this kind, we refer the reader to our paper [9], where we state and prove the corresponding results in a much more general setting and without the simplifying assumption that S is differentiable.

6. Concluding Discussion

The theorems given in this paper support the theory developed in our companion paper [10], where we argue in favor of a statistical equilibrium model for geostrophic turbulence that is defined by a prior distribution on potential vorticity and microcanonical constraints on energy and circulation. The first set of main results in the present paper furnishes an especially simple methodology for deriving the equilibrium equations and associated LDP's for that model. The second set of main results shows that the microcanonical ensemble has richer families of equilibrium solutions than the corresponding canonical ensemble, which omits those microcanonical equilibrium macrostates corresponding to parameters (E, Γ) not lying in the concavity set of the microcanonical entropy. The computations included in [10] for coherent flows in a channel with zonal topography demonstrate that these omitted states constitute a substantial portion of the parameter range of the physical model. As this discussion shows, the equivalence-of-ensemble issue is a fundamental one in the context of these local mean-field theories of coherent structure whenever the phenomenon of self-organization into coherent states is modeled.

In recent work [31], the statistical theory presented in [10] and analyzed in the present paper is applied to the active weather layer of the atmosphere of Jupiter. When appropriately fit to a $1 + 1/2$-layer quasi-geostrophic model, the theory correctly produces large-scale features that agree qualitatively and quantitatively with features observed by the Voyager and Galileo missions. For instance, for the channel domain in the southern hemisphere containing the Great Red Spot and White Ovals [31, Figure 2], the theory produces both the alternating east-west zonal shear flow and two anticyclonic vortices embedded in it. In particular, the size and position of the vortices closely match the size and position of the GRS and White Ovals, while the zonally-averaged velocity profile is extremely close to the profile deduced by Limaye [16]. Similarly, for a northern hemisphere channel that contains no permanent vortices, the theory correctly predicts a zonal shear flow with no embedded vortices. This agreement between theory and observation demonstrates the practical utility of the model analyzed in the preceding sections of the present paper.

In Section 5.1 of [10], the nonlinear stability of the equilibrium states for the model is shown by applying a Lyapunov argument. The required Lyapunov functional is constructed from the rate function I for the coarse-grained process $W_{n,r}$ and the constraint functionals H and C. In the canonical case, this construction is identical with a well-known method due to Arnold [1], now called the energy-Casimir method [12]. On the other hand, in the

microcanonical case when nonequivalence prevails, this method breaks down and hence it is necessary to give a more refined stability analysis. The required refinement, which makes use of the concept of the augmented Lagrangian from constrained optimization theory, is presented in [10]. For example, the Jovian flows realized in [31] fall in this regime. Thus the variational principles derived from the large deviation analysis in Section 4 and examined in Section 5 also have important implications for hydrodynamic stability criteria.

References

[1] Arnold, V.I., *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York 1978.

[2] Balescu, R., *Equilibrium and Nonequilibrium Statistical Mechanics*, John Wiley & Sons, New York 1975.

[3] Balian, R., *From Microphysics to Macrophysics*, vol. I, (trans. by D. ter Haar and J.G. Gregg), Springer-Verlag, Berlin 1991.

[4] Carnevale, G.F. and Frederiksen, J.S, Nonlinear stability and statistical mechanics of flow over topography, *J. Fluid Mech.* **175** (1987), 157-181.

[5] Dembo, A. and Zeitouni, O., *Large Deviations Techniques and Applications*, 2nd ed., Springer-Verlag, New York 1998.

[6] DiBattista, M.T., Majda, A.J., and Grote, M.J., Meta-stability of equilibrium statistical structures for prototype geophysical flows with damping and driving, *Phys. D* **151** (2001), 271-304.

[7] Dudley, R.M., *Real Analysis and Probability*, Chapman & Hall, New York 1989.

[8] Ellis, R.S. and Dupuis, P., *A Weak Convergence Approach to the Theory of Large Deviations*, John Wiley & Sons, New York 1997.

[9] Ellis, R.S., Haven, K., and Turkington, B., Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles, *J. Stat. Phys.* **101** (2000), 999-1064.

[10] Ellis, R.S., Haven, K., and Turkington, B., Nonequivalent statistical equilibrium ensembles and refined stability theorems for most probable flows, *Nonlinearity* **15** (2002), 239-255.

[11] Holloway, G., Eddies, waves, circulation, and mixing: statistical geofluid mechanics, *Ann. Rev. Fluid Mech.* **18** (1986), 91-147.

[12] Holm, D., Marsden, J., Ratiu, T., and Weinstein, A., Nonlinear stability of fluid and plasma equilibria, *Rev. Mod. Phys.* **123** (1985), 1-116.

[13] Ioffe, A.D., and Tihomirov, V.M., *Theory of External Problems*, Elsevier North-Holland, New York 1979.

[14] Joyce, G. and Montgomery, D., Negative temperature states for the two-dimensional guiding center plasma, *J. Plasma Phys.* **10** (1973), 107-121.

[15] Kraichnan, R., Statistical dynamics of two-dimensional flow, *J. Fluid Mech.* **67** (1975), 155-175.

[16] Limaye, S.S., Jupiter: new estimates of the mean zonal flow at the cloud level, *Icarus* **65** (1986), 335-352.

[17] Lynch, J. and Sethuraman, J., Large deviations for processes with independent increments, *Ann. Prob.* **15** (1987), 610-627.

[18] Majda, A.J. and Holen, M., Dissipation, topography and statistical theories of large scale coherent structure, *Commun. Pure Appl. Math.* **50** (1997), 1183-1234.

[19] Miller, J., Statistical mechanics of Euler equations in two dimensions, *Phys. Rev. Lett.* **65** (1990), 2137-2140.

[20] Miller, J., Weichman, P., and Cross, M.C., Statistical mechanics, Euler's equations, and Jupiter's red spot, *Phys. Rev. A* **45** (1992), 2328-2359.

[21] Montgomery, D., Matthaeus, W.H., Stribling, W.T., Martinez, D., and Oughton, S., Relaxation in two dimensions and the sinh-Poisson equation, *Phys. Fluids A* **4** (1992), 3-6.
[22] Morrison, P.J., Hamiltonian description of the ideal fluid, *Rev. Mod. Phys.* **70** (1998), 467-521.
[23] Onsager, L., Statistical hydrodynamics, *Nuovo Cimento* **6** (1949), 279-287.
[24] Parisi, G., *Statistical Field Theory*, Addison-Wesley, California 1988.
[25] Robert, R., A maximum-entropy principle for two-dimensional perfect fluid dynamics, *J. Stat. Phys.* **65** (1991), 531-553.
[26] Robert, R. and Sommeria, J., Statistical equilibrium states for two-dimensional flows, *J. Fluid Mech.* **229** (1991), 291-310.
[27] Salmon, R., *Lectures on Geophysical Fluid Dynamics*, Oxford Univ. Press, New York 1998.
[28] Salmon, R., Holloway, G., and Hendershott, M.C., The equilibrium statistical mechanics of simple quasi-geostrophic models, *J. Fluid Mech.* **75** (1976), 691-701.
[29] Shepherd, T.G., Symmetries, conservation laws, and Hamiltonian structures in geophysical fluid dynamics, *Adv. Geophys.* **32** (1990), 287-338.
[30] Turkington, B., Statistical equilibrium measures and coherent states in two-dimensional turbulence, *Commun. Pure Appl Math.* **52** (1999), 781-809.
[31] Turkington, B., Majda, A., Haven, K., and DiBattista, M., Statistical equilibrium predictions of jets and spots on Jupiter, *Proc. Nat. Acad. Sci. USA* **98** (2001), 12346-12350.
[32] Zeidler, E., *Nonlinear Functional Analysis and its Applications III: Variational Methods and Optimization*, Springer-Verlag, New York 1985.