An Introduction to Large Deviations and Equilibrium Statistical Mechanics for Turbulent Flows

Corentin Herbert

Abstract Two-dimensional turbulent flows, and to some extent, geophysical flows, are systems with a large number of degrees of freedom, which, albeit fluctuating, exhibit some degree of organization: coherent structures emerge spontaneously at large scales. In this short course, we show how the principles of equilibrium statistical mechanics apply to this problem and predict the condensation of energy at large scales and allow for computing the resulting coherent structures. We focus on the structure of the theory using the language of large deviation theory.

1 Introduction

Various characterizations of turbulent flows can be encountered; the components they usually entail are a chaotic dynamics on a strange attractor [81], a large range of scales (i.e. a large number of degrees of freedom), and strong nonlinear effects due to the prevalence of inertia over molecular dissipation [29]. Such flows can be found in industrial problems, but also in nature, for instance in geophysical flows and astrophysical flows. The above mentioned properties typically mean that not much can be said about the system in a deterministic framework, and that one should try instead to predict statistical properties.

This is exactly the purpose of the field of statistical mechanics: given a dynamical system (or set of ordinary or partial differential equations) in a large phase space (the microscopic state), can we predict typical values for specific functions on phase space (the macroscopic observables) without knowing the exact trajectory in phase space? For a large class of systems, said to be *in equilibrium*, such typical values can be obtained by assuming that the microscopic variables are random and distributed according to probability measures built upon a few macroscopic quantities, the invariants of the dynamical system. A classical example is that of the ideal gas: the exact position and velocity of the molecules matters little to us, but knowing the

C. Herbert (✉)

National Center for Atmospheric Research, Boulder, CO 80307, USA

e-mail: cherbert@ucar.edu

relations between macroscopic quantities such as temperature, pressure, energy and entropy is fundamental.

The ideas of statistical mechanics have been applied successfully to a large number of models of physical phenomena. An example of achievement of this approach is the theory of phase transitions, in which systems such as the Ising model, a toy-model of ferromagnetism, have been instrumental. However, turbulent flows present a number of difficulties: (i) they are directly formulated as continuous fields (infinite number of degrees of freedom) and can have an infinity of conserved quantities, (ii) the interactions between constituents have a long range, (iii) in many practical applications, the system is driven out of equilibrium by external forces.

Although we shall not tackle issue (iii) at all in this chapter, we will try to show how (i) and (ii) are actually useful ingredients to make probabilistic predictions for the system. They are the cornerstones of a mean-field theory: interacting degrees of freedom can be treated as statistically independent random variables in the limit of a large number of degrees of freedom. A natural language to express these properties is that of *large deviations theory* [25, 46, 80]: the probability of the outcome of a given observable concentrates exponentially around a set of values when the size of the system goes to infinity. The focus of the chapter is on the presentation of the large deviation principles for carefully chosen observables for a discretized form of 2D turbulence. To show that the principles at work are very general, we shall underline the connection with simpler models such as variants of the Ising model of ferromagnetism. Although it is shown that the theory allows us to compute the equilibrium states of the system, we shall not dwell on the description of such equilibrium states; the reader is referred to the review articles [13, 53] on this topic. We shall also refrain from discussing the connections with earlier applications of statistical mechanics, like the point vortex approach of Onsager, reviewed in [28], or the Kraichnan approach to Galerkin truncated flows [44], only mentioned briefly in Sect. 3.6.1.

These notes are based on lectures given at the *Stochastic Equations for Complex Systems: Theory and Applications* summer school organized at the University of Wyoming in June 2014. They mostly serve a pedagogical purpose, and we shall not give proofs of the results with the required mathematical rigor. However, we have tried as much as possible to provide the original references for the interested readers. The presentation adopted here owes much to the references [10, 69, 89]. Note that the ideas discussed here are applicable to many other systems with long range interactions [14, 23] and in particular gravitational systems [17, 66], plasmas, cold atoms or toy models of statistical physics.

2 Models of Turbulent Flows

2.1 3D and 2D Hydrodynamics

We are mainly interested here in the behavior of incompressible fluid flows, which is governed by the Navier-Stokes equations:

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla P + \nu \Delta \mathbf{u}, \qquad (1)$$

$$\nabla \cdot \mathbf{u} = 0, \qquad (2)$$

where $\mathbf{u}$ is the velocity field, P the pressure and ν the viscosity. The equations can be recast into non-dimensional form by introducing a velocity scale U, a length scale L, the corresponding time scale or *eddy turnover time* $T = L/U$, and the Reynolds number $Re = UL/\nu$. In other words, the Reynolds number measures the ratio of the nonlinear term and the dissipative term, or equivalently, of inertia and viscosity [29, 45]. Since viscosity acts at small scales, it is also a measure of the range of scales characteristic of the flow: the smallest scale is the Kolmogorov scale $\ell_\eta = (\nu^3/\epsilon)^{1/4}$, where ϵ is the energy dissipation rate. Now, with $\epsilon = U^3/L$, we obtain $L/\ell_\eta = Re^{3/4}$. Hence, the effective number of degrees of freedom in 3D flows grows as $Re^{9/4}$: flows with Reynolds number on the order of 10^9 are not uncommon in nature (the atmosphere and the ocean for instance), leading to a very large typical number of degrees of freedom.

The Navier-Stokes equations can be recast in terms of the vorticity field $\boldsymbol{\omega} = \nabla \times \mathbf{u}$:

$$\partial_t \boldsymbol{\omega} + \mathbf{u} \cdot \nabla \boldsymbol{\omega} = \boldsymbol{\omega} \cdot \nabla \mathbf{u} + \nu \Delta \boldsymbol{\omega}. \qquad (3)$$

The first term on the right hand side corresponds to stretching of vorticity tubes. When we consider a flow on a two-dimensional surface rather than the three-dimensional space, this vorticity stretching term vanishes (the only non-vanishing component of vorticity ω is normal to the surface), yielding conservation of vorticity along streamlines in the inviscid ($\nu = 0$) case:

$$\partial_t \omega + \mathbf{u} \cdot \nabla \omega = 0. \qquad (4)$$

This difference between 2D and 3D flows have important consequences on their respective behavior. While 3D flows tend to transfer energy from the large scales to the small scales, where it is dissipated by viscosity, in a process referred to as a direct energy cascade [31, 41] (big vortices break up into smaller and smaller vortices), 2D flows, on the contrary, transfer energy from the small scales to the large scales, and this is called an inverse energy cascade [6, 44, 86, 87]. In this inverse cascade process, vortices merge to form larger and larger vortices [55]. Unless sufficient large scale dissipation (e.g. bottom friction) is present, the energy piles up at the largest available scales, forming a *condensate* which dominates the flow [6, 21, 85].

The physical problem we are interested in here is the inverse energy cascade and the emergence of large scale coherent structures.

2.2 Global Invariants

The equations of motion for 3D and 2D hydrodynamics have a Hamiltonian structure, although non-canonical: there exists a Poisson structure, but it is degenerate [64]. This degeneracy leads to the existence of invariants, described in this section.

Inviscid 3D flows have two global invariants, the energy and the helicity [84]:

$$E = \frac{1}{2} \int \mathbf{u}^2, \tag{5}$$

$$H = \int \mathbf{u} \cdot \boldsymbol{\omega}. \tag{6}$$

Helicity being sign indefinite, it does not in general constrain the nonlinear transfers sufficiently to hamper the direct energy cascade process [43] (see however, [5, 34] for particular cases). On the contrary, in 2D, vorticity conservation along streamlines leads to a family of invariants in addition to the energy (ψ being the stream function, defined by $\omega = -\Delta\psi$)

$$\mathcal{H} = \int_{\mathcal{D}} \omega(\mathbf{x})\psi(\mathbf{x})d\mathbf{x}, \tag{7}$$

the Casimir invariants:

$$I_g = \int_{\mathcal{D}} g(\omega(\mathbf{x}))d\mathbf{x}, \tag{8}$$

where g is an arbitrary function. As a particular case, all the moments (or L^p norms) of the vorticity field are conserved:

$$\Gamma_p = \frac{1}{|\mathcal{D}|} \int_{\mathcal{D}} \omega(\mathbf{x})^p d\mathbf{x}, \tag{9}$$

including the L^2 norm of the vorticity field, referred to as the *enstrophy*. It was anticipated early on [2, 42, 48] that the existence of a second, positive-definite, quadratic invariant, in addition to the energy, is sufficient to reverse the direction of the energy cascade. The basic idea is that enstrophy is stronger in the presence of small-scale activity: transferring energy towards the small-scales while keeping the total energy constant cannot be done if we also need to conserve enstrophy. This loose statement was made more precise by a number of analytic arguments [2, 30, 42, 48, 56, 63], and verified in experiments [67, 82] and high-resolution numerical

simulations [6]. Statistical mechanics provides one of these analytical arguments (see Sect. 3.6.1).

Conservation of the Casimir invariants can be formulated equivalently in terms of the moments of the vorticity field, as above, or in terms of the vorticity distribution. Indeed, the fraction of the domain area $|\mathcal{D}|$, occupied by the vorticity level σ, which can be written as

$$\gamma(\sigma) = \frac{1}{|\mathcal{D}|} \int_{\mathcal{D}} \delta(\omega(\mathbf{x}) - \sigma) d\mathbf{x}, \tag{10}$$

is conserved. We shall see that this form is particularly convenient in Sect. 3, but note that the two formulations are connected by the formula

$$\Gamma_p = \frac{1}{|\mathcal{D}|} \int_{\mathcal{D}} d\mathbf{x} \int_{\mathbb{R}} d\sigma \sigma^p \delta(\omega(\mathbf{x}) - \sigma) = \int_{\mathbb{R}} d\sigma \sigma^p \gamma(\sigma), \tag{11}$$

and, conversely, using an integral representation of the Dirac distribution,

$$\gamma(\sigma) = \frac{1}{2\pi} \sum_{p=0}^{+\infty} \frac{(-1)^p \Gamma_p}{p!} \delta^{(p)}(\sigma). \tag{12}$$

Finally, note that the vorticity distribution is normalized:

$$\int_{\mathbb{R}} \gamma(\sigma) d\sigma = 1. \tag{13}$$

2.3 Geophysical Flows

Although 2D flows are interesting in themselves, part of the motivation for studying them comes from their common features with geophysical flows. Indeed, in addition to the small aspect ratio of the atmosphere and the ocean, their dynamics is subjected to the effect of strong rotation and density stratification. These properties allow for an asymptotic regime which describes well the large-scale dynamics, the *quasi-geostrophic* regime [93]. This regime is very similar to 2D flows, because it reduces to a quantity, called *potential vorticity*, being advected by the flow, similarly to the vorticity (see (4)). In particular, the velocity field is purely horizontal. The only difference is that the fields also depend on the vertical, and whereas the vorticity is the laplacian of the stream function: $\omega = -\Delta\psi$ in 2D, here the potential vorticity is related to the stream function by a slightly more complicated linear differential operator. The existence of Casimir invariants similar to those of 2D flows leads again to an inverse cascade of energy and the formation of coherent structures at large scales [15, 72, 83]. Therefore, the considerations presented here may apply to such flows as well, and attempts to extend the theory in this context have flourished over the

past few years. For the sake of simplicity, we shall restrict ourselves here to the case of 2D flows; the interested reader may consult the literature on extensions to quasi-geostrophic flows in the barotropic case [12, 36, 61, 95], the baroclinic case [24, 33, 94, 96], shallow-water equations [18, 20], as well as the general references [13, 53, 54], for instance.

The quasi-geostrophic regime breaks down at smaller scales, and we enter an intermediate regime, often referred to as stratified turbulence [50]. In this regime, even though we can still define a potential vorticity which is a Lagrangian invariant, it does not put as strong a constraint on the system as in the 2D case. Indeed, the fields (velocity, density) can be decomposed into a *balanced* part which contributes to potential vorticity, and *inertia-gravity waves*, which do not. As a result, the organization of the system in terms of inertial ranges and energy cascades is not so simple. High-resolution numerical simulations have indicated the existence of two inertial ranges with a constant and opposite flux of energy [70]. A possible interpretation is that the vortical modes are responsible for the inverse cascade of energy while the inertia-gravity waves have to do with the direct energy cascade. This interpretation is supported by a statistical mechanics argument [37], which is an adaptation of the Kraichnan argument (see Sect. 3.6.1) in the context of the restricted canonical ensemble [68].

Independently of the constraining effect of rotation and stratification (which can be seen as forces breaking isotropy), another direction of generalization which has been considered is that of 3D flows with symmetries, and especially axisymmetric flows [49, 62, 88]. This configuration is relevant for setups used in laboratory experiments, such as the von Karman experiment. It has been shown in particular that one could define a microcanonical measure using an approach analogous to that of Sect. 3.1, with, however, some considerable complications to treat the fluctuations of the poloidal field [88].

2.4 Discretized Form for 2D Euler Flows and Analogies with Toy Models of Magnetic Systems

Instead of the continuous vorticity field ω and the infinite dimensional phase space it belongs to, it may be more convenient to introduce finite dimensional models. Here we shall mostly consider a discretization on a square lattice with N sites equally spaced in the domain $\mathcal{D}$ (see Fig. 1), and the variables of interest are the values taken by vorticity at each site. In this form, the system can be related to some classical models of statistical physics.

2.4.1 Two-Vorticity Level System and Long-Range Ising Model

The Ising model is one of the most famous models in statistical physics. It can be seen as a toy model of ferromagnetism, but it has served as a testbed for a very large
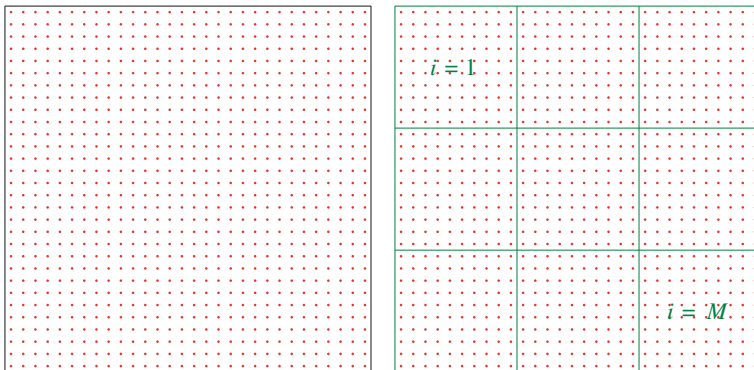
Fig. 1 Discretization and coarse-graining. *Left* We replace the 2D domain by a finite lattice and the continuous vorticity field by a N-dimensional vector whose components are the values of the vorticity at each site. *Right* We decompose the lattice in M cells, each containing $n = N/M$ sites. The coarse-grained vorticity is a M dimensional vector whose components are the average value of the vorticity in each cell

number of ideas going far beyond this particular problem [22]. It consists of a finite number N of spins $s_i \in \{-1, 1\}$ located on a lattice of arbitrary shape and dimension (although a square lattice is often considered) and interacting through a hamiltonian (per degree of freedom) of the form:

$$\mathcal{H}_I[\hat{s}] = -\frac{1}{N} \sum_{i,j=1}^{N} J_{ij} s_i s_j. \tag{14}$$

In this form, the hamiltonian is just any quadratic function. A standard choice of interaction is the nearest-neighbor model: $J_{ij} = J$ if the sites i and j are connected in the lattice, and $J_{ij} = 0$ if they are not. That way, aligned neighboring spins will contribute a term $-J$ to the hamiltonian, while anti-aligned neighboring spins will contribute J. If J is positive the system is called *ferromagnetic* and if it is negative the system is called *antiferromagnetic*. An observable of interest is the *magnetization* (per spin):

$$\mathcal{M}[\hat{s}] = \frac{1}{N} \sum_{i=1}^{N} s_i. \tag{15}$$

When one finds about the same proportion of positive and negative spins, the magnetization should vanish. Applying an external magnetic field, represented by a term of the form $-h \sum_i s_i$ in the hamiltonian, leads to alignment of spins, and therefore a non-vanishing magnetization. This is the standard behavior of so-called *paramagnetic* materials. By contrast, in *ferromagnetic* materials, spins may align spontaneously and yield unit magnetization (in absolute value) without imposing an external magnetic field (or, in experiments, the system retains its magnetization

when the applied magnetic field is switched off). The Ising model can be seen as a toy model of the paramagnetic-ferromagnetic transition.

In the above case, the system has short-range interactions, since only neighboring spins interact. Versions with long-range interactions can be built by allowing non-vanishing J_{ij} for distant sites i and j. For instance, one may assume that all the spins interact with all the other spins with the same coupling constant: $J_{ij} = 1/N$, where the $1/N$ ensures that the energy per degree of freedom is an intensive quantity. In this case, the Hamiltonian becomes a function of magnetization only:

$$\mathcal{H}_{IMF}[\hat{s}] = -\frac{1}{N^2} \sum_{i,j=1}^{N} s_i s_j = -\mathcal{M}[\hat{s}]^2. \tag{16}$$

This version of the Ising model is referred to as *mean-field*, because it is tantamount to saying that each spin feels the effect of a magnetic field created by all the other spins rather than the individual effect of each of his neighbors. Indeed, let us consider a given spin s_i; it provides a contribution $-1/Ns_i \sum_j J_{ij}s_j$, which is the same as a non-interacting spin under external magnetic field $1/N \sum_j J_{ij}s_j$ would. If we replace this magnetic field by the magnetization, we obtain the mean-field Hamiltonian. Note that the geometric shape (square, triangle, etc.) and the dimension of the lattice do not matter here since all the spins interact with the same intensity.

An advantage of the mean-field Ising model is that it has an exact solution in any dimension [3]. On the contrary, exact solutions for the standard, short-range Ising model are only know for dimension one [38] and two [65].

The discretized version of 2D flows described above is related to the Ising model in the following way: rather than allowing the vorticity to take any real value, we can restrict it to a two-level set $\{\sigma, -\sigma\}$. Then the system becomes analogous to the Ising model, with an interaction matrix given by the Green function of the Laplacian on the lattice. On a plane, this amounts to interactions proportional to the logarithm of the distance between sites: $J_{ij} \propto \ln|i - j|$ for $i \neq j$. This is a kind of long-range interaction. The difference with the Ising model is the presence of the vorticity distribution conservation constraint. This would amount to fixing the number of $+$ spins and the number of $-$ spins in the Ising model.

2.4.2 Energy-Enstrophy Ensemble and the Long-Range Spherical Model

Another variant of the Ising model consists in letting the spins s_i take any real value, while satisfying the global constraint $\sum_{i=1}^{N} s_i^2 = N$. Clearly, this constraint is satisfied in the standard Ising model with spins in $\{1, -1\}$. The name *spherical model* was coined for this variant because of the form of the global constraint, which means that the set of all spin values lies on the surface of a sphere in $\mathbb{R}^N$. It was introduced by Berlin and Kac [4] as an attempt to patch the divergence arising from assuming that the spins are distributed according to a normal distribution (the *Gaussian model*) while remaining exactly solvable in any dimension [3]. The observables of interest

(Hamiltonian, magnetization) are the same as for the Ising model. Versions with short-range [4] or long-range [39] interactions can again be considered by choosing different quadratic forms J_{ij}.

In their discretized version, 2D flows resemble a long-range spherical model if we only retain one Casimir invariant: the enstrophy. Indeed, enstrophy conservation implies $\sum_i \omega_i^2 = N\Gamma_2$. Again, the interaction matrix is given by the Green function of the Laplacian on the lattice. This connection is further investigated in Sect. 3.6. It has also been pointed out in a series of papers by Lim [51, 52].

3 Mean-Field Theory for 2D Flows

We provide here a heuristic presentation of the mean-field theory introduced independently by Miller [59, 60], Robert and Sommeria [75, 77], and further developed by many others. The presentation is inspired by the original work by Miller and the more recent references [10, 13, 69]. More rigorous mathematical proofs can be found in the original papers by Robert and coworkers [57, 58, 73–77] and Ellis, Turkington and coworkers [7, 8, 26, 27, 91].

3.1 Microcanonical Measure and Large Deviations for the Energy and Vorticity Distribution

The general idea is to consider the vorticity field ω, referred to as the *microstate*, as a random variable distributed according to the *microcanonical distribution*. In other words, we introduce a probability measure on the phase space $\Lambda = L^\infty(\mathcal{D})$, where $\mathcal{D}$ is a 2D domain (we shall mostly consider the case of a rectangular domain here). We are going to give a sketch of the construction of this measure as a limit of measures on finite-dimensional phase spaces corresponding to approximations of the continuous vorticity field. Then, we will be able to make predictions on the value of *macrostates*, i.e. observables $\mathcal{A} : \Lambda \longrightarrow \mathbb{R}$ (or more generally $\mathcal{A} : \Lambda \longrightarrow M$ where the space M is *macroscopic* in some sense, e.g. has a dimension much lower than Λ) on phase space, which, as we shall see, satisfy large deviation properties: they concentrate in probability around some specific values, the equilibrium states.

To keep things simple, we shall consider a finite number of vorticity levels $\mathfrak{S} = \{\sigma_1, \ldots, \sigma_K\}$. This amounts to saying that the vorticity distribution has the form $\gamma(\sigma) = \sum_{k=1}^K \gamma_k \delta(\sigma - \sigma_k)$. We consider the discretized system with N sites on the square lattice introduced in Sect. 2.4 (see Fig. 1), and define a *microstate* as being simply the value of the vorticity field at all the points of the lattice. Therefore the phase space is simply $\Lambda_N = \mathfrak{S}^N$.

Considering the conservation laws mentioned above, there are two observables of primary interest: the energy observable, i.e. the Hamiltonian, given by

$$\mathcal{H}_N : \hat{\omega} \in \Lambda_N \longmapsto \frac{1}{2N^2} \sum_{1 \leq i \neq j \leq N} G_{ij}^N \omega_i \omega_j, \tag{17}$$

with G_{ij}^N the Green function of the Laplacian on the lattice, and the vorticity distribution observables

$$\mathcal{G}_N^{(k)} : \hat{\omega} \in \Lambda_N \longmapsto \frac{1}{N} \sum_{i=1}^N \delta_{\omega_i, \sigma_k}. \tag{18}$$

Note that the set of accessible energies (i.e. the values taken by the observable $\mathcal{H}_N$) is finite and depends both on the vorticity levels σ_k and on the number of sites N. Ultimately, in the limit $N \rightarrow \infty$, we shall be interested in a continuum of energy levels. One approach to circumvent this difficulty is to consider in a first step energy shells with finite width ΔE, large enough so that each shell is attained by the energy observable for some microstates [69]. In the limit $N \rightarrow +\infty$, the results will not depend on the value of ΔE. To keep notations as simple as possible, we shall refrain from doing so here, but in all rigor one should understand $\mathcal{H}_N[\hat{\omega}] \in [E, E + \Delta E]$ whenever we write $\mathcal{H}_N[\hat{\omega}] = E$. In this framework, the set of microstates with vorticity distribution γ and energy E is

$$\Lambda_N(\gamma, E) = \{\hat{\omega} \in \Lambda_N \mid \mathcal{H}_N[\hat{\omega}] = E, \forall k \in [\![1, K]\!], \mathcal{G}_N^{(k)}[\hat{\omega}] = \gamma_k\}, \tag{19}$$

$$= \mathcal{H}_N^{-1}(\{E\}) \cap \bigcap_{k=1}^K \mathcal{G}_N^{(k)^{-1}}(\{\gamma_k\}). \tag{20}$$

This is a finite set whose cardinality we denote by $\Omega_N(\gamma, E) = \text{Card } \Lambda_N(\gamma, E)$.

We are going to introduce two probability measures on phase space: first, let us consider a *prior measure* $\mu^{(N)}$, which here is just the normalized counting measure: if $M \subset \Lambda_N$, $\mu^{(N)}(M) = \frac{\text{Card } M}{K^N}$. This amounts to saying that all the microstates are equiprobable: for any observable $\mathcal{A}_N : \Lambda_N \longrightarrow \mathbb{R}$, the probability of the outcome x is just the fraction of microstates for which $\mathcal{A}_N[\hat{\omega}] = x$. Now, we want to restrict that statement to all the microstates with a fixed energy and vorticity distribution, while assigning vanishing probability to all the other microstates. Hence, we introduce the (finite-N) microcanonical measure $\mu_{\gamma, E}^{(N)}$: if $M \subset \Lambda_N$, $\mu_{\gamma, E}^{(N)}(M) = \text{Card}(M \cap \Lambda_N(\gamma, E))/\Omega_N(\gamma, E)$. Hence, for an observable $\mathcal{A}_N$, the probability law of the random variable $\mathcal{A}_N[\hat{\omega}]$ is $\mu_{\gamma, E}^{(N)}(\mathcal{A}_N[\hat{\omega}] = x) = \mu_{\gamma, E}^{(N)}(\mathcal{A}_N^{-1}(\{x\}))$. Note that we have introduced indices γ and E to distinguish from probabilities computed with respect to the prior measure. Probabilities in the microcanonical ensemble are thus just conditional probabilities:

$$\mu_{\gamma,E}^{(N)}(\mathcal{A}_N[\hat{\omega}] = x) = \mu^{(N)}(\mathcal{A}_N[\hat{\omega}] = x|\mathcal{H}_N[\hat{\omega}] = E, \mathcal{G}_N^{(k)}[\hat{\omega}] = \gamma_k), \quad (21)$$

$$= \begin{cases} \dfrac{\mu^{(N)}(\mathcal{A}_N[\hat{\omega}]=x,\mathcal{H}_N[\hat{\omega}]=E,\mathcal{G}_N^{(k)}[\hat{\omega}]=\gamma_k)}{\mu^{(N)}(\mathcal{H}_N[\hat{\omega}]=E,\mathcal{G}_N^{(k)}[\hat{\omega}]=\gamma_k))} & \text{if } \mathcal{H}_N[\hat{\omega}]=E, \text{ and} \\ & \forall k \in [\![1,K]\!], \mathcal{G}_N^{(k)}[\hat{\omega}]=\gamma_k \\ 0 & \text{otherwise.} \end{cases} \quad (22)$$

As mentioned above, observables of particular interest are the hamiltonian $\mathcal{H}_N$ and the vorticity distribution observables $\mathcal{G}_N^{(k)}$. The joint probability to observe an energy E and a vorticity distribution γ, with respect to the prior measure, satisfies a large-deviation property, and the large deviation rate function is (up to an unimportant constant term) the opposite of the entropy $S(E, \gamma)$:

$$\mu^{(N)}(\mathcal{H}_N[\hat{\omega}] = E, \mathcal{G}_N^{(k)}[\hat{\omega}] = \gamma_k) = \frac{\Omega_N(\gamma, E)}{K^N} = e^{NS(E,\gamma)-N\ln K + o(N)}, \quad (23)$$

with

$$S(E, \gamma) = \lim_{N \to \infty} \frac{1}{N} \ln \Omega_N(\gamma, E). \quad (24)$$

3.2 Large Deviations for the Macrostates

We now introduce a new class of observables associated with the coarse-graining of the vorticity field. We decompose the lattice into M cells, each containing $n = N/M$ sites. For a microstate $\hat{\omega} \in \mathfrak{S}^N$, we shall denote the components as $\omega_{i\alpha}$ where $1 \leq i \leq M$ is the index of the cell and $1 \leq \alpha \leq n$ is the index of the site within the cell (see Fig. 1). The coarse-graining observable is given by

$$\mathfrak{C} : \hat{\omega} \in \mathfrak{S}^N \longmapsto \bar{\omega} \in \mathbb{R}^M, \text{ with } \bar{\omega}_i = \frac{1}{n} \sum_{\alpha=1}^{n} \omega_{i\alpha}, 1 \leq i \leq M. \quad (25)$$

More generally, we can define an observable which corresponds to the distribution of vorticity levels in each cell. It is just the empirical vector

$$\mathfrak{P} : \hat{\omega} \in \mathfrak{S}^N \longmapsto P = (p_{ik})_{\substack{1 \leq i \leq M \\ 1 \leq k \leq K}} \in \mathcal{M}_{M,K}(\mathbb{R}), \text{ with } p_{ik} = \frac{1}{n} \sum_{\alpha=1}^{n} \delta_{\omega_{i\alpha}, \sigma_k}. \quad (26)$$

Note that $\sum_{k=1}^{K} p_{ik} = 1$. Besides, the observable $\mathfrak{C}$ can be deduced from $\mathfrak{P}$ since $\bar{\omega}_i = \sum_{k=1}^{K} \sigma_k p_{ik}$ for $1 \leq i \leq M$. Let us refer to the elements of the image of $\mathfrak{P}$ as the *macrostates*. The set of microstates corresponding to a given macrostate

P is simply its pre-image $\mathfrak{P}^{-1}(P)$. The number of microstates realizing a given macrostate P will be denoted $W(P) = \text{Card } \mathfrak{P}^{-1}(P)$. It is easily computed that:

$$W(P) = \prod_{i=1}^{M} \binom{n}{np_{i1}} \binom{n - np_{i1}}{np_{i2}} \cdots \binom{n - np_{i1} - \cdots - np_{in-1}}{np_{in}}, \quad (27)$$

$$= \prod_{i=1}^{M} \frac{n!}{\prod_{k=1}^{K}(np_{ik})!}. \quad (28)$$

The vorticity distribution observables $\mathcal{G}_N^{(k)}$ take a constant value over an equivalence class $\mathfrak{P}^{-1}(P)$:

$$\mathcal{G}_N^{(k)}[\hat{\omega}] = \frac{1}{M} \sum_{i=1}^{M} p_{ik}, \quad (29)$$

so that if $\hat{\omega}_1, \hat{\omega}_2 \in \mathfrak{S}^N$ are such that $\mathfrak{P}[\hat{\omega}_1] = \mathfrak{P}[\hat{\omega}_2]$, then for $1 \leq k \leq K$, $\mathcal{G}_N^{(k)}[\hat{\omega}_1] = \mathcal{G}_N^{(k)}[\hat{\omega}_2]$. In other words, the equivalence kernel of the observable $\mathfrak{P}$ is finer than that of any of the observables $\mathcal{G}_N^{(k)}$. In practice, this means that we need not worry about enforcing the vorticity distribution constraint when counting the number of microstates realizing a given macrostate. For the energy observable, the situation is slightly more subtle: denoting $G_{i\alpha, j\beta}^{M,n}$ the Green function of the Laplacian on the lattice with the new indexing of the sites, the energy observable is given by:

$$\mathcal{H}_{N,M}[\hat{\omega}] = \frac{1}{2N^2} \sum_{\substack{1 \leq i,j \leq M \\ 1 \leq \alpha, \beta \leq n \\ (i,\alpha) \neq (j,\beta)}} G_{i\alpha, j\beta}^{M,n} \omega_{i\alpha} \omega_{j\beta}, \quad (30)$$

which is not necessarily constant over equivalence classes. However, it can be shown that the dominant contribution is the *mean-field* energy, i.e. the energy of the coarse-grained vorticity field:

$$\mathcal{H}_{N,M}[\hat{\omega}] = \frac{1}{2M^2} \sum_{1 \leq i \neq j \leq M} G_{ij}^{M} \bar{\omega}_i \bar{\omega}_j + o\left(\frac{1}{N}\right), \quad (31)$$

$$= \mathcal{H}_M[\mathfrak{C}[\hat{\omega}]] + o\left(\frac{1}{N}\right). \quad (32)$$

The above results are sometimes restated by saying that we have an *energy* (and here, also vorticity distribution) *representation function* [89] (see Fig. 2). It allows us to obtain the most probable states with respect to the microcanonical measure by obtaining a large deviation property with respect to the prior (unconstrained) measure.
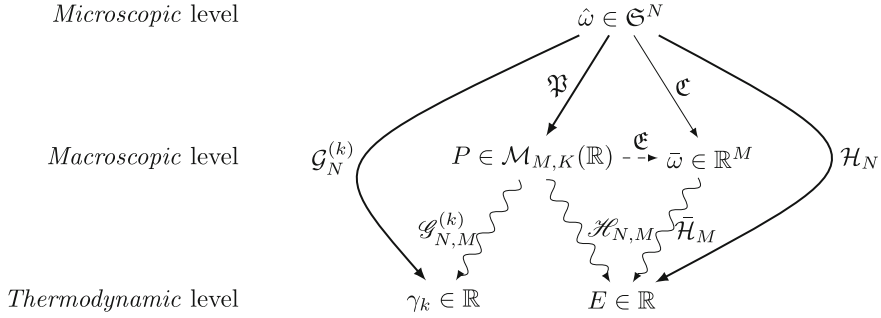
Fig. 2 The different levels of description of the system, and the observables/representation function relating them. Observables are represented with *straight arrows*, representation functions with *wiggly arrows* and contraction principles with *dashed arrows*. Observables for which a large deviation principle has been obtained directly are represented with *thick arrows*

Indeed, the unconstrained probability of observing a macrostate P is

$$\mu^{(N)}(\mathfrak{P}[\hat{\omega}] = P) = \mu^{(N)}(\mathfrak{P}^{-1}(P)), \tag{33}$$

$$= \frac{W(P)}{K^N}. \tag{34}$$

Using the Stirling approximation, it is easily shown that when $N \to \infty$, this probability satisfies a large deviation property:

$$\mu^{(N)}(\mathfrak{P}[\hat{\omega}] = P) = e^{N\mathscr{S}_{M,K}[P] + o(N)}, \tag{35}$$

where we have introduced the mean-field entropy

$$\mathscr{S}_{M,K}[P] = \lim_{N \to +\infty} \frac{1}{N} \ln \mu^{(N)}(\mathfrak{P}[\hat{\omega}] = P), \tag{36}$$

$$= -\frac{1}{M} \sum_{i=1}^{M} \sum_{k=1}^{K} p_{ik} \ln p_{ik}, \tag{37}$$

which again appears as a large deviation rate function (up to an additive constant and a minus sign), although this time it is a large deviation of an empirical vector (observable $\mathfrak{P}$) rather than a sample mean (energy observable $\mathcal{H}$). Hence, the above result should in all rigor be seen as a consequence of the Sanov theorem.

Now, in the microcanonical ensemble, the probability $\mu_{\gamma,E}^{(N)}(\mathfrak{P}[\hat{\omega}] = P)$ involves the joint (unconstrained) probability $\mu^{(N)}(\mathfrak{P}[\hat{\omega}] = P, \mathcal{H}_N[\hat{\omega}] = E, \mathcal{G}_N^{(k)}[\hat{\omega}] = \gamma_k)$. But due to the existence of the energy and vorticity distribution representation

functions, we have:

$$
\begin{aligned}
\mu^{(N)}(\mathfrak{P}[\hat{\omega}] = P, &\mathcal{H}_N[\hat{\omega}] = E, \mathcal{G}_N^{(k)}[\hat{\omega}] = \gamma_k) \\
&= \mu^{(N)}(\mathfrak{P}[\hat{\omega}] = P, \mathscr{H}_{N,M}[P] = E, \mathscr{G}_{N,M}^{(k)}[P] = \gamma_k),
\end{aligned} \tag{38}
$$

therefore,

$$
\mu_{\gamma,E}^{(N)}(\mathfrak{P}[\hat{\omega}] = P) = \begin{cases} \dfrac{\mu^{(N)}(\mathfrak{P}[\hat{\omega}]=P)}{\mu^{(N)}(\mathcal{H}_N[\hat{\omega}]=E,\mathcal{G}_N^{(k)}[\hat{\omega}]=\gamma_k))} & \begin{array}{l}\text{if } \mathscr{H}_{N,M}[P]=E, \text{ and} \\ \forall k \in [\![1,K]\!], \mathscr{G}_{N,M}^{(k)}[P]=\gamma_k \end{array} \\ 0 & \text{otherwise.} \end{cases} \tag{39}
$$

It follows that the probability of a given macrostate also satisfies a large deviation result with respect to the microcanonical measure:

$$
\mu_{\gamma,E}^{(N)}(\mathfrak{P}[\hat{\omega}] = P) = e^{-NI[P]+o(N)}, \tag{40}
$$

with the large deviation rate function

$$
I[P] = \begin{cases} S(E,\gamma) - \mathscr{S}_{M,K}[P] & \text{if } \mathscr{H}_{N,M}[P] = E, \text{ and } \forall k \in [\![1,K]\!], \mathscr{G}_{N,M}^{(k)}[P] = \gamma_k \\ +\infty & \text{otherwise.} \end{cases} \tag{41}
$$

Hence, the most probable macrostates with respect to the microcanonical measure are those which minimize the large deviation rate function, i.e. those which maximize the entropy $\mathscr{S}_{M,K}$ while satisfying the constraints on energy and vorticity distribution: they are solutions of a constrained variational problem. It is worthy of note that the Boltzmann-Gibbs entropy $\mathscr{S}_{M,K}$, defined in (37), evaluated at a solution P^* of the variational problem, agrees with the entropy $S(E, \gamma)$ defined from the Boltzman formula (24). This is not a coincidence, but a cornerstone of the mean-field approach. It can be understood in the language of large deviation theory as a *contraction principle* [89]. Roughly speaking, due to the existence of representation functions, the probability of observing an energy E and a vorticity distribution γ can be computed as the integral over all the macrostates (rather than the microstates) with these constraints: denoting

$$
\tilde{\Lambda}_{N,M}(\gamma, E) = \{P \in \mathcal{M}_{M,K}(\mathbb{R}) \mid \mathscr{H}_{N,M}[P] = E, \forall k \in [\![1, K]\!], \mathscr{G}_{N,M}^{(k)}[P] = \gamma_k\}, \tag{42}
$$

we have

$$\mu^{(N)}(\mathcal{H}_N[\hat{\omega}] = E, \mathcal{G}_N^{(k)}[\hat{\omega}] = \gamma_k) = \int_{\Lambda_N(\gamma, E)} \mu^{(N)}(d\hat{\omega}), \tag{43}$$

$$= \int_{\tilde{\Lambda}_{N,M}(\gamma, E)} \mu^{(N)}(\mathfrak{P}^{-1}(P)), \tag{44}$$

$$= \int_{\tilde{\Lambda}_{N,M}(\gamma, E)} e^{N\mathcal{S}_{M,K}[P]+o(N)}, \tag{45}$$

Using Laplace's approximation, the integral evaluates to

$$= \exp\left(N \max_{P \in \tilde{\Lambda}_{N,M}(\gamma, E)} \mathcal{S}_{M,K}[P] + o(N)\right). \tag{46}$$

As a conclusion, the most probables macrostates P^* with respect to the microcanonical measure satisfy $I[P^*] = 0$: they are solutions of the constrained variational problem:

$$S(E, \gamma) = \max_P \{\mathcal{S}_{M,K}[P] \mid \mathcal{H}_{N,M}[P] = E, \forall k \in [\![1, K]\!], \mathcal{G}_{N,M}^{(k)}[P] = \gamma_k\}. \tag{47}$$

3.3 Thermodynamic Limit and Mean-Field Equation

We are now interested in the macrostates obtained in the limit $M \rightarrow +\infty$. Letting also $K \rightarrow +\infty$, they are the probability distributions for fine-grained vorticity $\rho(\mathbf{r}, \sigma)$: $\rho(\mathbf{r}, \sigma)d\sigma$ is the probability that the vorticity at point $\mathbf{r}$ lies in the interval $[\sigma, \sigma + d\sigma]$. The local normalization condition $\int_{\mathbb{R}} \rho(\mathbf{r}, \sigma)d\sigma = 1$ must still be satisfied for each point $\mathbf{r} \in \mathcal{D}$. The coarse-grained vorticity field is now $\bar{\omega}(\mathbf{r}) = \int_{\mathbb{R}} \sigma \rho(\mathbf{r}, \sigma)d\sigma$. As explained above, the energy and vorticity distribution depend only on the macrostate ρ:

$$\mathcal{H}[\rho] = \frac{1}{2} \int_{\mathcal{D}^2} d\mathbf{r}d\mathbf{r}' \int_{\mathbb{R}^2} d\sigma d\sigma' G(\mathbf{r}, \mathbf{r}')\sigma\sigma'\rho(\mathbf{r}, \sigma)\rho(\mathbf{r}', \sigma'), \tag{48}$$

$$\mathcal{D}_\sigma[\rho] = \int_{\mathcal{D}} \rho(\sigma, \mathbf{r})d\mathbf{r}. \tag{49}$$

Similarly, the mean field entropy becomes

$$\mathcal{S}[\rho] = -\int_{\mathcal{D}} d\mathbf{r} \int_{\mathbb{R}} d\sigma \rho(\sigma, \mathbf{r}) \ln \rho(\sigma, \mathbf{r}). \tag{50}$$

The most probable macrostates are now those maximizing (50) while satisfying the energy and vorticity distribution constraints. They are solutions of the microcanonical variational problem:

$$S(E, \gamma) = \max_{\rho}\{\mathscr{S}[\rho] \mid \mathscr{H}[\rho] = E, \forall \sigma \in \mathbb{R}, \mathscr{D}_\sigma[\rho] = \gamma(\sigma)\}. \qquad (51)$$

The critical points of the variational problem are readily found: there exist Lagrange multipliers β and $\alpha(\sigma)$ such that the first variations vanish:

$$0 = \delta\mathscr{S} - \int_{\mathcal{D}} d\mathbf{r}\zeta(\mathbf{r}) \int_{\mathbb{R}} d\sigma \delta\rho(\sigma, \mathbf{r}) - \beta\delta\mathscr{H} - \int_{\mathbb{R}} d\sigma\alpha(\sigma) \int_{\mathcal{D}} d\mathbf{r}\delta\rho(\sigma, \mathbf{r}), \quad (52)$$

which leads to the Gibbs states

$$\rho^*(\sigma, \mathbf{r}) = \frac{e^{-\beta\sigma\overline{\psi}(\mathbf{r})-\alpha(\sigma)}}{\mathcal{Z}_\alpha(\beta\overline{\psi}(\mathbf{r}))}, \qquad (53)$$

where the coarse grained stream function $\overline{\psi}$ and the partition function $\mathcal{Z}_\alpha$ are given by

$$\overline{\psi} = -\Delta^{-1}\overline{\omega}, \qquad\qquad \mathcal{Z}_\alpha(u) = \int_{\mathbb{R}} e^{-\sigma u-\alpha(\sigma)}d\sigma. \qquad (54)$$

It follows that the coarse-grained vorticity field satisfies

$$\overline{\omega}(\mathbf{r}) = F_\alpha(\beta\overline{\psi}(\mathbf{r})), \qquad\qquad \text{with } F_\alpha(u) = -\frac{d\ln\mathcal{Z}_\alpha(u)}{du}. \qquad (55)$$

This is a (elliptic) partial differential equation, referred to as the *mean-field equation*, characterizing the most probable coarse-grained vorticity fields. Note that the equation is of the same form as the equation defining steady-states of the Euler equation: equilibrium states form a subclass of steady-states for which the function relating vorticity and stream function is fixed by the invariants of the system.

The equilibrium states of the system can thus be obtained by solving (55). In general, this is a difficult task. Analytical solutions have been obtained in the limit of a linear function F_α (the mean-field equation then reduces to a Helmholtz equation), using the method introduced by Chavanis and Sommeria [19], which consists in decomposing the vorticity field and stream function on a basis of Laplacian eigenfunctions. Numerical methods are also available: Turkington and Whitaker have proposed an algorithm to iteratively solve the variational problem described above [92], while Robert and Sommeria [78] have proposed relaxation equations where the dynamics maximize the entropy production rate, thereby reaching a maximum entropy state. We shall not describe in details these methods here, nor the solutions they yield. Note, however, that in general, they correspond to large scale coherent structures, like vortices or unidirectional (e.g. zonal) flows, depending on the

geometry of the domain: for instance dipole/monopole in a rectangular domain [19], dipole/unidirectional flow in a doubly periodic domain [11], Fofonoff flows on a beta-plane [61], and solid-body rotation/dipole/quadrupole/unidirectional flow on a sphere [32, 35, 36, 71].

3.4 Non-equivalence of Ensembles

3.4.1 Statistical Ensembles and Variational Problems

So far we have been using exclusively the microcanonical measure

$$\mu_{\gamma,E}^{(N)}(d\hat{\omega}) = \delta(\mathcal{H}_N[\hat{\omega}] - E) \prod_{k=1}^{K} \delta\left(\mathcal{G}_N^{(k)}[\hat{\omega}] - \gamma_k\right) \mu^{(N)}(d\hat{\omega}), \tag{56}$$

which assigns uniform probability to microstates with a given energy and vorticity distribution, and zero probability to other microstates. We could make different choices and consider the canonical measure

$$\mu_{\gamma,\beta}^{(N)}(d\hat{\omega}) = \frac{e^{-\beta\mathcal{H}_N[\hat{\omega}]}}{\mathcal{Z}} \prod_{k=1}^{K} \delta\left(\mathcal{G}_N^{(k)}[\hat{\omega}] - \gamma_k\right) \mu^{(N)}(d\hat{\omega}), \tag{57}$$

or the grand-canonical measure

$$\mu_{\alpha,\beta}^{(N)}(d\hat{\omega}) = \frac{e^{-\beta\mathcal{H}_N[\hat{\omega}] - \sum_{k=1}^{K} \alpha_k \mathcal{G}_N^{(k)}[\hat{\omega}]}}{\Xi} \mu^{(N)}(d\hat{\omega}), \tag{58}$$

and similarly in the thermodynamic limit $N \to +\infty$. If we replace the microcanonical measure in Sect. 3.2 by any of these two measures, we obtain *mutas mutandi* a large deviation principle for the macrostates. In the thermodynamic limit, the most probable macrostates (i.e. the equilibrium states) are therefore solutions of the following variational problems:

$$S(E, \gamma) = \max_{\rho}\{\mathcal{S}[\rho] \mid \mathcal{H}[\rho] = E, \forall\sigma \in \mathbb{R}, \mathcal{D}_\sigma[\rho] = \gamma(\sigma)\}, \tag{59}$$

$$F(\beta, \gamma) = \max_{\rho}\{\mathcal{S}[\rho] - \beta\mathcal{H}[\rho] \mid \forall\sigma \in \mathbb{R}, \mathcal{D}_\sigma[\rho] = \gamma(\sigma)\}, \tag{60}$$

$$J(\beta, \alpha) = \max_{\rho}\{\mathcal{S}[\rho] - \beta\mathcal{H}[\rho] - \int_{\mathbb{R}} d\sigma\alpha(\sigma)\mathcal{D}_\sigma[\rho]\}, \tag{61}$$

respectively for the microcanonical measure, the canonical measure and the grand-canonical measure. The maximized functions arise as large deviation rate functions, and the constraints stem from the definition of the ensembles as conditional probabilities and from the existence of representation functions. The entropy $S(E, \gamma)$, the

free energy $F(\beta, \gamma)$ and the *grand potential* $J(\beta, \alpha)$ are referred to generically as *thermodynamic potentials*.

   The existence of a large deviation principle for the macrostates does not depend on the particular choice of ensemble, but the most probable macrostates may depend on this choice. The task that we set out to investigate in this section is therefore how the different ensembles are related. The discussion closely follows the Refs. [26, 90].

3.4.2 Ensemble Equivalence at the Macrostate Level

First of all, it is clear from the structure of the variational problems and the Lagrange multiplier rule that they all have the same critical points. However, the critical points may be of different nature: a maximizer of one variational problem may be a saddle point of another variational problem for instance. Nevertheless, it is easily seen that a solution of a variational problem with a constraint relaxed (e.g. the canonical variational problem) is always a solution of the original constrained variational problem (e.g. the microcanonical problem). We can formalize this remark by introducing the sets of equilibrium states (i.e. solutions of the variational problems):

$$\mathcal{MC}(E, \gamma) = \{\rho \mid \mathcal{S}[\rho] = S(E, \gamma), \mathcal{H}[\rho] = E, \forall \sigma \in \mathbb{R}, \mathcal{D}_\sigma[\rho] = \gamma(\sigma)\},$$
$$(62)$$

$$\mathcal{C}(\beta, \gamma) = \{\rho \mid \mathcal{S}[\rho] - \beta\mathcal{H}[\rho] = F(\beta, \gamma), \forall \sigma \in \mathbb{R}, \mathcal{D}_\sigma[\rho] = \gamma(\sigma)\}, \quad (63)$$

$$\mathcal{GC}(\beta, \alpha) = \{\rho \mid \mathcal{S}[\rho] - \beta\mathcal{H}[\rho] - \int_\mathbb{R} d\sigma\alpha(\sigma)\mathcal{D}_\sigma[\rho] = J(\beta, \alpha)\}. \quad (64)$$

As per the above remark, we always have,

$$\forall \beta, \alpha, \forall \rho \in \mathcal{GC}(\beta, \alpha), \rho \in \mathcal{C}(\beta, \mathcal{D}_\sigma[\rho]) \text{ and } \rho \in \mathcal{MC}(\mathcal{H}[\rho], \mathcal{D}_\sigma[\rho]), \quad (65)$$
$$\forall \beta, \gamma, \forall \rho \in \mathcal{C}(\beta, \gamma), \rho \in \mathcal{MC}(\mathcal{H}[\rho], \gamma). \quad (66)$$

In particular,

$$\bigcup_{\beta, \alpha} \mathcal{GC}(\beta, \alpha) \subset \bigcup_{\beta, \gamma} \mathcal{C}(\beta, \gamma) \subset \bigcup_{E, \gamma} \mathcal{MC}(E, \gamma). \quad (67)$$

If the converse statements hold, i.e.

$$\forall E, \gamma, \forall \rho \in \mathcal{MC}(E, \gamma), \exists \beta \in \mathbb{R} : \rho \in \mathcal{C}(\beta, \gamma), \quad (68)$$

or

$$\forall \beta, \gamma, \forall \rho \in \mathcal{C}(\beta, \gamma), \exists \alpha : \rho \in \mathcal{GC}(\beta, \alpha), \quad (69)$$

we say, respectively, that the microcanonical and canonical ensembles are *equivalent at the macrostate level* or that the canonical and grand canonical ensembles are *equivalent at the macrostate level*. It is straightforward to see that it is a transitive relation, in the sense that if the microcanonical ensemble is equivalent to the canonical ensemble at the macrostate level, and if the canonical ensemble and the grand-canonical ensemble are equivalent at the macrostate level, then the microcanonical and the grand-canonical ensembles are equivalent a the macrostate level. Besides, if the grand-canonical ensemble is equivalent to the microcanonical ensemble at the macrostate level, then the canonical ensemble is equivalent to both the microcanonical and the grand-canonical ensembles at the macrostate level.

If the three ensembles are equivalent at the macrostate level, we have the equalities:

$$\bigcup_{\beta,\alpha} \mathscr{GC}(\beta,\alpha) = \bigcup_{\beta,\gamma} \mathscr{C}(\beta,\gamma) = \bigcup_{E,\gamma} \mathscr{MC}(E,\gamma). \tag{70}$$

3.4.3 Ensemble Equivalence at the Thermodynamic Level

Due to the definition of the thermodynamic potentials through the variational problems, connections exist between them as well. For the free energy for instance, we have

$$F(\beta,\gamma) = \max_{\rho,\mathcal{N}[\rho](\mathbf{x})=1} \{\mathscr{S}[\rho] - \beta\mathscr{H}[\rho] \mid \forall\sigma \in \mathbb{R}, \mathscr{D}_\sigma[\rho] = \gamma(\sigma)\},$$

$$= \max_{E\geq 0} \left(\max_{\rho,\mathcal{N}[\rho](\mathbf{x})=1,\mathscr{H}[\rho]=E} \{\mathscr{S}[\rho] - \beta E \mid \forall\sigma \in \mathbb{R}, \mathscr{D}_\sigma[\rho] = \gamma(\sigma)\} \right),$$

$$= \max_{E\geq 0} \left(S(E,\gamma) - \beta E \right).$$

This exactly means that the free energy is the *Legendre-Fenchel* transform of the entropy. The Legendre-Fenchel transform is a generalization of the Legendre transform to functions which need not be differentiable and convex [79]. Denoting the Legendre-Fenchel of an arbitrary function with a star (the variable with respect to which the transform is taken should be clear from the arguments of the function), we have the compact form:

$$F(\beta,\gamma) = S^\star(E,\gamma).$$

Similarly,

$$J(\beta,\alpha) = F^\star(\beta,\gamma).$$

We know that the Legendre transform is an involution [1]. This is not necessarily the case for the Legendre-Fenchel transform, because the Legendre-Fenchel transform of an arbitrary function is always a concave function, but it is true when the original function is concave. In general, when applying the Legendre-Fenchel

transform twice, we only obtain the concave hull of the original function [79]. Hence, the free energy F is always a concave function of β and the grand-potential J is always a concave function of its arguments, while $F^\star = S^{\star\star}$ is always a concave function of E, and is the smallest concave function satisfying $S(E, \gamma) \leq S^{\star\star}(E, \gamma)$. The equality holds if S is a concave function. Therefore, we say that the microcanonical and canonical ensemble are *equivalent at the thermodynamic level* if $S = F^\star = S^{\star\star}$, or equivalently, if S is a concave function of E. Similarly, the grand canonical and the canonical ensembles are equivalent at the thermodynamic level if $F = J^\star = F^{\star\star}$, i.e. if F is a concave function of γ.

Again, we have a transitivity property: equivalence of the grand canonical and canonical ensembles on the one hand, and of the canonical and microcanonical ensembles on the other hand implies equivalence of the grand canonical and microcanonical ensembles. Besides, if the grand canonical and the microcanonical ensembles are equivalent, then the canonical ensemble is equivalent to both the grand canonical and the microcanonical ensembles. In both these cases, the entropy S is a concave function of all its arguments.

3.4.4 Equivalence and Non-equivalence of Statistical Ensembles

The notions of ensemble equivalence at the macrostate level (Sect. 3.4.2) and at the thermodynamic level (Sect. 3.4.3) are connected. Indeed, the local concavity properties of the thermodynamic potential determine the possibility to invert the relation with the Lagrange multiplier, or in other words, the possibility that the macrostates can be obtained by solving a relaxed variational problem. Following [26], let us examine the three possibilities in the context of the microcanonical and canonical ensembles. Let us fix E, γ, then one of the three following assertions holds:

(i) **Total Ensemble Equivalence:** If $S = S^{\star\star}$ and S is not locally flat, then $\mathscr{MC}(E, \gamma) = \mathscr{C}(\beta, \gamma)$ for $\beta = \partial S/\partial E$.
(ii) **Marginal Ensemble Equivalence:** If $S = S^{\star\star}$ and S is locally flat, then $\mathscr{MC}(E, \gamma) \subsetneq \mathscr{C}(\beta, \gamma)$ for $\beta = \partial S/\partial E$.
(iii) **Ensemble Inequivalence:** If $S \neq S^{\star\star}$, then $\forall \beta \in \mathbb{R}, \mathscr{MC}(E, \gamma) \cap \mathscr{C}(\beta, \gamma) = \emptyset$.

3.5 Large Deviations for the Coarse-Grained Vorticity Field

In Sect. 3.2, we have considered how the probability of the outcome of a given observable (the distribution of fine-grained vorticity) behaves when the size of the system goes to infinity. We have found that it satisfies a large deviation property, which allows us to compute the most probable outcomes (see Sect. 3.3). From there, we are able to deduce what the most probable coarse-grained vorticity fields are. But can we apply the same methods directly to the coarse-graining observable to compute the most probable coarse-grained vorticity fields? In other words, can we obtain a large deviation principle directly for the observable $\mathfrak{C}$?

In general this is not straightforward, because we do not have a representation function for the vorticity distribution in terms of the coarse-grained vorticity field. Let us give an exemple in the simple case where we have only three levels of vorticity: $\mathfrak{S} = \{-1, 0, 1\}$. We have represented on Fig. 3 two microstates which lead to the same coarse-grained vorticity field, with different vorticity distributions. As a consequence, we cannot deduce a large deviation principle with respect to the microcanonical measure (or any of the other ensembles) from a large deviation principle with respect to the prior measure. In principle it remains possible to evaluate directly the probability of a coarse-grained vorticity field in the microcanonical ensemble, but this is a much more complicated combinatorial problem. However, in the special case of a two-level vorticity system, we do have a representation function for the vorticity distribution. We illustrate this in the following sections by making use of the analogy with the mean-field Ising model pointed out above.
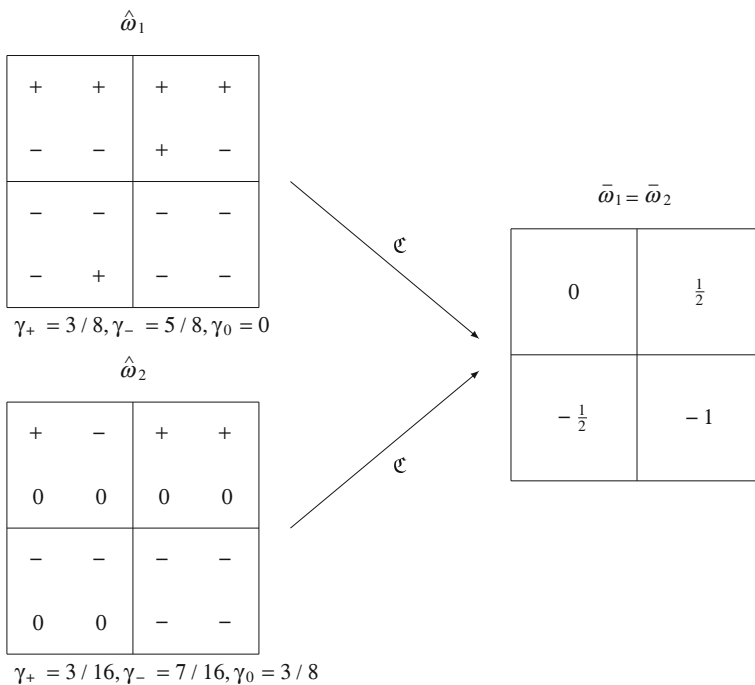
Fig. 3 Exemples of two microstates with different vorticity distribution, which are mapped to the same coarse-grained vorticity field by the operator $\mathfrak{C}$, in the three-level case: $\mathfrak{S} = \{-1, 0, 1\}$
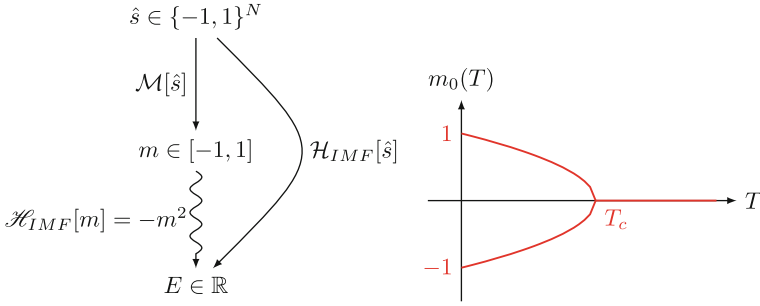
Fig. 4 Mean-field Ising model: observables (hamiltonian and magnetization) and representation function for the energy (*left*) and most probable magnetization as a function of temperature in the canonical ensemble (*right*)

3.5.1 Mean-Field Ising Model

Remember the mean-field Ising model described in Sect. 2.4.1. We have mentioned above that there is a representation function for the energy in terms of the magnetization (Fig. 4). Therefore it is sufficient to obtain a large deviation principle for the magnetization with respect to the unconstrained measure. If N_+ (resp. N_-) is the number of $+$ (resp. $-$) spins, the magnetization is given by $\mathcal{M}[\hat{s}] = (N_+ - N_-)/N$, and we have $N_+ + N_- = N$. In other words, $N_\pm/N = (1 \pm \mathcal{M}[\hat{s}])/2$. Hence, the unconstrained probability to observe a given magnetization is

$$\mu^{(N)}(\mathcal{M}[\hat{s}] = m) = \frac{N!}{2^N N_+! N_-!} \tag{71}$$

$$= e^{N\mathscr{S}[m] + o(N)}, \tag{72}$$

where the mean-field entropy is given by (up to an unimportant constant $\ln 2$)

$$\mathscr{S}[m] = -\frac{1+m}{2} \ln\left(\frac{1+m}{2}\right) - \frac{1-m}{2} \ln\left(\frac{1-m}{2}\right), \tag{73}$$

which proves that the magnetization observable satisfies a large deviation principle. It is customary to work in the canonical ensemble (see Sect. 3.4), and the most probable states are therefore solutions of the variational problem:

$$F(\beta) = \min_{m \in [-1,1]} \left(\beta \mathscr{H}_{IMF}[m] - \mathscr{S}[m]\right), \tag{74}$$

where F is the free energy (per spin). Using $\mathscr{H}_{IMF}[m] = -m^2$ and (73), it is easily shown that for $\beta = 1/(kT)$ smaller than a critical value $\beta_c = 1/(kT_c)$ (high temperature T), there is a unique solution $m = 0$, while for β larger than the critical value

(low temperature), there are two non-zero solutions $\pm m_0(T)$. The most probable magnetization as a function of the temperature is represented on Fig. 4.

3.5.2 Two-Level System

We have noted above that when the vorticity level set is made of two opposite values, $\mathfrak{S} = \{\sigma_0, -\sigma_0\}$ (with $\sigma_0 > 0$), the system becomes analogous to the mean-field Ising model studied above. The only difference is the vorticity distribution conservation constraint (and the interaction coefficients). This amounts to keeping fixed the number of $+$ and $-$ spins in the Ising model, or equivalently, to fixing the magnetization. But the magnetization here is nothing but the circulation Γ_1. Therefore, conservation of the Casimir invariants in the discretized two-level model reduces to conservation of the circulation.

Another way to see this is to show explicitly that there exists a representation function for the vorticity distribution in this case. The coarse-graining operator $\mathfrak{C}$ takes value in a discrete subset of $\mathbb{R}^M$: denoting $\bar{\mathfrak{S}}_n = \{\left(\frac{2k}{n} - 1\right)\sigma_0, 0 \le k \le n\}$, the image of the operator is $\bar{\mathfrak{S}}_n^M$. Here, k corresponds to the number of sites with value σ_0 in each coarse-graining cell. The relation between k and $\bar{\omega}_i$ can be inverted: $k = n(1 + \bar{\omega}_i/\sigma_0)/2$, and we obtain

$$\gamma_+ = \frac{1}{N} \sum_{i=1}^{M} \frac{n}{2}\left(1 + \frac{\bar{\omega}_i}{\sigma_0}\right), \tag{75}$$

$$= \frac{1}{2} + \frac{\Gamma_1}{2\sigma_0}, \tag{76}$$

and similarly,

$$\gamma_- = \frac{1}{2} - \frac{\Gamma_1}{2\sigma_0}. \tag{77}$$

Note that, as expected, $\gamma_+ + \gamma_- = 1$ (Eq. (13)) and $(\gamma_+ - \gamma_-)\sigma_0 = \Gamma_1$ (Eq. (11)). Now, it is an easy task to evaluate the unconstrained probability of a given coarse-grained vorticity field:

$$\mu^{(N)}(\mathfrak{C}[\hat{\omega}] = \bar{\omega}) = \frac{\text{Card } \mathfrak{C}^{-1}[\bar{\omega}]}{2^N} = 2^{-N} \prod_{i=1}^{M} \frac{n!}{(n/2(1 + \bar{\omega}_i/\sigma_0))!(n/2(1 - \bar{\omega}_i/\sigma_0))!}, \tag{78}$$

$$= e^{N\bar{\mathscr{S}}_{M,2}[\bar{\omega}] + o(N)}, \tag{79}$$

with the entropy of the coarse-graining observable (up to an unimportant constant $\ln 2$)

$$\bar{\mathscr{S}}_{M,2}[\bar{\omega}] = -\frac{1}{2M} \sum_{i=1}^{M} \left[\left(1 + \frac{\bar{\omega}_i}{\sigma_0}\right) \ln\left(1 + \frac{\bar{\omega}_i}{\sigma_0}\right) + \left(1 - \frac{\bar{\omega}_i}{\sigma_0}\right) \ln\left(1 - \frac{\bar{\omega}_i}{\sigma_0}\right) \right]. \tag{80}$$

The contraction principle ensures that, as can be checked explicitly,

$$\bar{\mathscr{S}}_{M,2}[\bar{\omega}] = \max_{P \in \mathcal{M}_{M,2}(\mathbb{R})} \{\mathscr{S}_{M,2}[P] \mid \mathfrak{E}[P] = \bar{\omega}\}. \tag{81}$$

By the same token as in Sect. 3.2, it follows that the most probable coarse-grained vorticity fields are solutions of the constrained variational problem:

$$S(E, \gamma) = \max_{\bar{\omega} \in \tilde{\mathfrak{S}}_n^M} \{\bar{\mathscr{S}}_{M,2}[\bar{\omega}] \mid \bar{\mathcal{H}}_M[\bar{\omega}] = E, \bar{\mathcal{G}}_M^{(+)}[\bar{\omega}] = \gamma_+, \bar{\mathcal{G}}_M^{(-)}[\bar{\omega}] = \gamma_-\}, \tag{82}$$

or equivalently,

$$S(E, \Gamma_1) = \max_{\bar{\omega} \in \tilde{\mathfrak{S}}_n^M} \{\bar{\mathscr{S}}_{M,2}[\bar{\omega}] \mid \bar{\mathcal{H}}_M[\bar{\omega}] = E, \bar{\mathcal{M}}_M[\bar{\omega}] = \Gamma_1\}. \tag{83}$$

Straightforward computations show that the critical points of the variational problem are solutions of the equation:

$$\bar{\omega} = \sigma_0 \tanh\left(\frac{(\beta\bar{\psi} + \alpha_1)\sigma_0}{2}\right). \tag{84}$$

3.5.3 Fragile Constraints and Constrained Casimir Variational Problem

It has been observed by several authors that none of the Casimir invariants Γ_p (moments of the vorticity field) except the first (circulation) can be obtained from the coarse-grained vorticity field $\bar{\omega}$. For this reason they are often referred to as *fragile invariants*, in the sense that they do not survive coarse-graining. This is exactly the same as saying that there is no representation function for the Casimir invariants (or, equivalently, for the vorticity distribution $\gamma(\sigma)$) in terms of the coarse-grained vorticity field $\bar{\omega}$, except in the particular case mentioned above. However, with an arbitrary vorticity distribution, a large deviation principle can still be obtained by contraction, as illustrated above in the two-layer case. This provides a variational problem for the most probable coarse-grained vorticity field, even though it still relies on an auxiliary maximization on the distribution P for the vorticity levels.

Because of their fragile nature, and because an infinite number of invariants is difficult to handle in practice, it was suggested [16, 27] to treat these invariants in a

canonical ensemble, and to consider the Lagrange parameter $\alpha(\sigma)$ as a prior vorticity distribution chosen based on physical intuition of the problem at hand. This provides a subset of solutions of the microcanonical variational problem, but not necessarily the full set (see Sect. 3.4). However, this variational problem, expressed in terms of the distribution P for the vorticity levels, is equivalent to minimizing, with respect to the coarse-grained vorticity field $\bar{\omega}$, the so-called *Casimir functionals* $\int_{\mathcal{D}} s(\omega)$ with fixed energy, where s is a convex function, choosing for s the Legendre-Fenchel transform of $\ln \mathcal{Z}_\alpha$ [9].

3.6 The Energy-Enstrophy Measure

3.6.1 Gibbs Measure for Galerkin Truncated Flows

In this section we investigate the statistical mechanics of the 2D Euler equations resulting from simplifying the conservation constraints: we retain only the energy and the enstrophy invariants. This was actually one of the starting points for statistical mechanics of turbulent flows: Lee in 3D [47] and Kraichnan in 2D [42] considered Fourier series of the dynamical fields truncated at a given order N. In 3D, the only invariants are the energy and the helicity, while in 2D, Kraichnan considered the energy:

$$\mathcal{H}_N[\hat{\omega}] = \frac{1}{2} \sum_{i=1}^{N} \frac{|\hat{\omega}_i|^2}{\lambda_i}, \tag{85}$$

and the enstrophy

$$\mathcal{G}_N[\hat{\omega}] = \frac{1}{2} \sum_{i=1}^{N} |\hat{\omega}_i|^2, \tag{86}$$

where the truncated vorticity field is given by

$$\hat{\omega}(x) = \sum_{i=1}^{N} \hat{\omega}_i \phi_i(x), \quad \Delta \phi_i = -\lambda_i \phi_i, 0 \le \lambda_i \le \lambda_{i+1}. \tag{87}$$

It is assumed that the truncated vorticity field $\hat{\omega}$ is a random variable distributed according to the canonical (Gibbs) measure:

$$\mu_{\beta,\alpha}^{(N)}(d\hat{\omega}) = \frac{1}{\mathcal{Z}} e^{-\beta \mathcal{H}_N[\hat{\omega}] - \alpha \mathcal{G}_N[\hat{\omega}]} \prod_{i=1}^{N} d\hat{\omega}_i. \tag{88}$$

This is a Gaussian probability density, well-defined if $\beta + \alpha \lambda_i > 0$ for all i. This condition leads to three possible regimes: (i) $\beta < 0, \alpha > 0$, (ii) $\beta > 0, \alpha > 0$, (iii) $\beta > 0, \alpha < 0$. In each case, Kraichnan considered the energy spectrum $\mathcal{E}_i[\hat{\omega}] = \frac{|\hat{\omega}_i|^2}{2\lambda_i}$ and computed its average value with respect to the Gibbs measure [42, 44]: $\langle \mathcal{E}_i \rangle = \frac{1}{2(\beta + \alpha \lambda_i)}$. In the negative temperature ($\beta < 0$) regime, the spectrum peaks at the gravest mode ϕ_1; there is even an *infrared divergence* when $\beta \to -\alpha \lambda_1$. This is classically interpreted as an indication that not only nonlinear interactions in 2D flows tend to transfer energy towards the large scales (the *inverse cascade*), but there is a tendency for energy to accumulate in the gravest mode to form a *condensate* [6, 21, 85]. Note that the average value of each vorticity mode vanishes by symmetry: $\langle \hat{\omega}_i \rangle = 0$, because a given vorticity field and its opposite have the same probability in the canonical ensemble. Of course, in reality, the system will spontaneously break the symmetry and choose a vorticity field, which can be computed in the limit of large N using large deviations results for the macrostates as we did above. In the energy-enstrophy ensemble, averaging over the set of equilibrium states indeed yields a vanishing mean value, thereby showing that statistical mechanics is more about most probable states than average values.

3.6.2 Large Deviations in the Microcanonical Ensemble

Using the same notations as in the previous paragraph, one may assume that the truncated vorticity field is distributed according to the microcanonical measure

$$\mu_{E,\Gamma_2}^{(N)}(d\hat{\omega}) = \frac{1}{\Omega_N(E, \Gamma_2)} \delta(\mathcal{H}_N[\hat{\omega}] - E)\delta(\mathcal{G}_N[\hat{\omega}] - \Gamma_2)\prod_{i=1}^{N} d\hat{\omega}_i, \quad (89)$$

instead of the Gibbs measure $\mu_{\beta,\alpha}^{(N)}(d\hat{\omega})$, with the structure function given by

$$\Omega_N(E, \Gamma_2) = \int \delta(\mathcal{H}_N[\hat{\omega}] - E)\delta(\mathcal{G}_N[\hat{\omega}] - \Gamma_2)\prod_{i=1}^{N} d\hat{\omega}_i. \quad (90)$$

Rather exceptionally, since both constraints involve quadratic functions, the structure function can be computed explicitly, using integral representations of the Dirac distributions [10], and thus also the entropy:

$$\Omega_N(E, \Gamma_2) = e^{NS(E,\Gamma_2)+o(N)}, \quad (91)$$

$$S(E, \Gamma_2) = \frac{1}{2}\ln(\Gamma_2 - 2\lambda_1 E). \quad (92)$$

Note that Bouchet and Corvellec [10] have also checked with explicit computations that this entropy defined as the joint large deviation rate function for

the energy and enstrophy observables (i.e. the Boltzmann formula $S(E, \Gamma_2) = \lim(1/N) \ln \Omega_N(E, \Gamma_2)$, given in (92)) coincides with the entropy defined through the variational problem for the macrostates, as expected from the contraction principle (Eq. (47)). A similar computation of the structure function was carried out by Kastner and Schnetz [40] for the mean-field spherical model defined in Sect. 2.4.2.

From the joint large deviation principle for the energy and enstrophy, we can deduce a large deviation principle for the energy spectrum observable [10]:

$$\mu_{E,\Gamma_2}^{(N)}(\mathcal{E}_i[\hat{\omega}] = E_i) = e^{N \mathscr{S}_{E,\Gamma_2}^{(i)}[E_i] + o(N)}, \tag{93}$$

with

$$\mathscr{S}_{E,\Gamma_2}^{(1)}[E_1] = \begin{cases} \frac{1}{2} \ln(\Gamma_2 - 2\lambda_2 E + 2(\lambda_2 - \lambda_1)E_1) & \text{if } 0 < E_1 < E \\ -\infty & \text{otherwise} \end{cases}, \tag{94}$$

and for $i > 1$,

$$\mathscr{S}_{E,\Gamma_2}^{(i)}[E_i] = \begin{cases} \frac{1}{2} \ln(\Gamma_2 - 2\lambda_1 E - 2(\lambda_i - \lambda_1)E_i) & \text{if } 0 \leq E_i \leq E \\ -\infty & \text{otherwise} \end{cases}. \tag{95}$$

The large deviation rate functions are monotonous: $\mathscr{S}_{E,\Gamma_2}^{(1)}[E_1]$ is an increasing function of E_1, and $\mathscr{S}_{E,\Gamma_2}^{(i)}[E_i]$ are decreasing functions of E_i. Therefore, the most probable energy spectrum in the limit $N \to +\infty$ has all its energy in the gravest mode. This can be seen as the microcanonical counterpart of the Kraichnan argument presented in Sect. 3.6.1. It provides further theoretical evidence for the spectral condensation in 2D turbulence.

The above discussion on the vanishing of the average truncated vorticity field also applies in the microcanonical ensemble. The mean-field theory allows to compute the most probable macrostates: we find a linear mean-field equation for the coarse-grained vorticity field: $\bar{\omega} = \beta/(2\alpha)\bar{\psi}$, which is easily solved and yields $\bar{\omega} = \sqrt{2\lambda_1 E}\phi_1$, in agreement with the above prediction.

4 Conclusion

In this chapter, we have given a brief introduction to the methods of equilibrium statistical mechanics applied to models of turbulent flows, focusing on the case of two-dimensional flows. The main purpose of the course was to show, in the context of a lattice discretization of the system, how some well-chosen observables, such as the distribution of fined-grained vorticity levels, concentrate in probability around a set of equilibrium values. Such properties are conveniently expressed using the theory of large deviations. In fact, we have closely followed the principles of equilibrium

statistical mechanics formulated in the language of large deviations, as exposed for instance in [89]. A major ingredient in deriving the large deviation results is the long-range character of the interactions, because it leads to the existence of a representation function for the energy. This is a major simplification, as it allows us to compute the probability of a macrostate with respect to the uniform measure and then deduce the probability with respect to the microcanonical measure. We have emphasized this point by considering another observable, the coarse-grained vorticity field (for which there is no representation function for the vorticity distribution) and by making the analogy with a simpler system, the mean-field Ising model.

The large deviation principle leads to a variational problem characterizing the most probable macrostates. This allows to compute coarse-grained vorticity fields which should correspond in practice to the final state of the system, if ergodicity holds. This provides a statistical explanation of the spontaneous emergence of coherent structures in two-dimensional flows. The equilibrium states obtained may depend on the choice of probability measure in phase space: we have discussed the relations between the standard ensembles of statistical mechanics and given a connection with the concavity properties of the entropy.

In the simpler context of the energy-enstrophy measure, we have explained that the energy spectrum observable also satisfies a large deviation principle, which shows that the most probable state has all its energy condensed in the gravest mode. This is physically consistent with the familiar ideas of inverse cascade of energy and energy condensation for 2D flows.

References

1. Arnold VI (1989) Mathematical methods of classical mechanics, 2nd edn. Springer, New York
2. Batchelor G (1969) Computation of the energy spectrum in homogeneous two-dimensional turbulence. Phys Fluids 12(Suppl. II):233–239. doi:10.1063/1.1692443
3. Baxter RJ (1982) Exactly solved models in statistical mechanics. Academic Press, London
4. Berlin TH, Kac M (1952) The spherical model of a ferromagnet. Phys Rev 86:821. doi:10.1103/PhysRev.86.821
5. Biferale L, Musacchio S, Toschi F (2012) Inverse energy cascade in three-dimensional isotropic turbulence. Phys Rev Lett 108(16):164501. doi:10.1103/PhysRevLett.108.164501
6. Boffetta G, Ecke RE (2012) Two-dimensional turbulence. Annu Rev Fluid Mech 44:427. doi:10.1146/annurev-fluid-120710-101240
7. Boucher C, Ellis RS, Turkington B (1999) Spatializing random measures: doubly indexed processes and the large deviation principle. Ann Probab 27:297–324
8. Boucher C, Ellis RS, Turkington B (2000) Derivation of maximum entropy principles in two-dimensional turbulence via large deviations. J Stat Phys 98(5–6):1235–1278
9. Bouchet F (2008) Simpler variational problems for statistical equilibria of the 2D Euler equation and other systems with long range interactions. Physica D 237:1976–1981. doi:10.1016/j.physd.2008.02.029
10. Bouchet F, Corvellec M (2010) Invariant measures of the 2D Euler and Vlasov equations. J Stat Mech P08021. doi:10.1088/1742-5468/2010/08/P08021
11. Bouchet F, Simonnet E (2009) Random changes of flow topology in two-dimensional and geophysical turbulence. Phys Rev Lett 102:094504. doi:10.1103/PhysRevLett.102.094504

12. Bouchet F, Sommeria J (2002) Emergence of intense jets and Jupiter's Great Red Spot as maximum-entropy structures. J Fluid Mech 464:165–207. doi:10.1017/S0022112002008789
13. Bouchet F, Venaille A (2012) Statistical mechanics of two-dimensional and geophysical flows. Phys Rep 515:227–295. doi:10.1016/j.physrep.2012.02.001
14. Campa A, Dauxois T, Ruffo S (2009) Statistical mechanics and dynamics of solvable models with long-range interactions. Phys Rep 480:57–159. doi:10.1016/j.physrep.2009.07.001
15. Charney JG (1971) Geostrophic turbulence. J Atmos Sci 28:1087–1094. doi:10.1175/1520-0469(1971)028<1087:GT>2.0.CO;2
16. Chavanis PH (2003) Generalized thermodynamics and Fokker-Planck equations: applications to stellar dynamics and two-dimensional turbulence. Phys Rev E 68:036108
17. Chavanis PH (2006) Phase transitions in self-gravitating systems. Int J Mod Phys B 20:3113. doi:10.1142/S0217979206035400
18. Chavanis PH, Dubrulle B (2006) Statistical mechanics of the shallow-water system with an a priori potential vorticity distribution. C R Phys 7:422–432. doi:10.1016/j.crhy.2006.01.007
19. Chavanis PH, Sommeria J (1996) Classification of self-organized vortices in two-dimensional turbulence: the case of a bounded domain. J Fluid Mech 314:267–297. doi:10.1017/S0022112096000316
20. Chavanis PH, Sommeria J (2002) Statistical mechanics of the shallow water system. Phys Rev E 65:026302. doi:10.1103/PhysRevE.65.026302
21. Chertkov M, Connaughton C, Kolokolov I, Lebedev V (2007) Dynamics of energy condensation in two-dimensional turbulence. Phys Rev Lett 99(8):084501. doi:10.1103/PhysRevLett.99.084501
22. Cipra BA (1987) An introduction to the Ising model. Am Math Mon 94(10):937–959
23. Dauxois T, Ruffo S, Arimondo E, Wilkens M (eds) (2002) Dynamics and thermodynamics of systems with long range interactions. Lecture notes in physics, vol 602. Springer, New York. doi:10.1007/3-540-45835-2
24. DiBattista M, Majda AJ (2001) Equilibrium statistical predictions for baroclinic vortices: the role of angular momentum. Theor Comput Fluid Dyn 14(5):293–322
25. Ellis RS (1985) Entropy, large deviations, and statistical mechanics. Springer, New York
26. Ellis RS, Haven K, Turkington B (2000) Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles. J Stat Phys 101:999–1064. doi:10.1023/A:1026446225804
27. Ellis RS, Haven K, Turkington B (2002) Nonequivalent statistical equilibrium ensembles and refined stability theorems for most probable flows. Nonlinearity 15:239. doi:10.1088/0951-7715/15/2/302
28. Eyink G, Sreenivasan K (2006) Onsager and the theory of hydrodynamic turbulence. Rev Mod Phys 78:87–135. doi:10.1103/RevModPhys.78.87
29. Falkovich G (2011) Fluid mechanics. Cambridge University Press, Cambridge
30. Fjortoft R (1953) On the changes in the spectral distribution of kinetic energy for twodimensional, nondivergent flow. Tellus 5:225–230
31. Frisch U (1995) Turbulence, the legacy of A.N. Kolmogorov. Cambridge University Press, Cambridge
32. Herbert C (2013) Additional invariants and statistical equilibria for the 2D Euler equations on a spherical domain. J Stat Phys 152:1084–1114. doi:10.1007/s10955-013-0809-6
33. Herbert C (2014) Nonlinear energy transfers and phase diagrams for geostrophically balanced rotating-stratified flows. Phys Rev E 89:033008. doi:10.1103/PhysRevE.89.033008
34. Herbert C (2014) Restricted partition functions and inverse energy cascades in parity symmetry breaking flows. Phys Rev E 89:013010. doi:10.1103/PhysRevE.89.013010
35. Herbert C, Dubrulle B, Chavanis PH, Paillard D (2012) Phase transitions and marginal ensemble equivalence for freely evolving flows on a rotating sphere. Phys Rev E 85:056304. doi:10.1103/PhysRevE.85.056304
36. Herbert C, Dubrulle B, Chavanis PH, Paillard D (2012) Statistical mechanics of quasi-geostrophic flows on a rotating sphere. J Stat Mech P05023. doi:10.1088/1742-5468/2012/05/P05023

37. Herbert C, Pouquet A, Marino R (2014) Restricted equilibrium and the energy cascade in rotating and stratified flows. J Fluid Mech 758:374–406. doi:10.1017/jfm.2014.540
38. Ising E (1925) Beitrag zur theorie des ferromagnetismus. Z Phys 31:253–258
39. Joyce GS (1966) Spherical model with long-range ferromagnetic interactions. Phys Rev 146:349
40. Kastner M, Schnetz O (2006) On the mean-field spherical model. J Stat Phys 122:1195–1214
41. Kolmogorov AN (1941) The local structure of turbulence in incompressible viscous fluid for very large Reynolds' numbers. Dokl Akad Nauk SSSR 30:301. doi:10.1098/rspa.1991.0075
42. Kraichnan RH (1967) Inertial ranges in two-dimensional turbulence. Phys Fluids 10:1417–1423. doi:10.1063/1.1762301
43. Kraichnan RH (1973) Helical turbulence and absolute equilibrium. J Fluid Mech 59:745–752
44. Kraichnan RH, Montgomery DC (1980) Two-dimensional turbulence. Rep Prog Phys 43:547. doi:10.1088/0034-4885/43/5/001
45. Landau L, Lifchitz E (1971) Physique Théorique, Tome VI: Mécanique des fluides. Mir, Moscou
46. Lanford OE (1973) Entropy and equilibrium states in classical statistical mechanics. In: Lenard A (ed) Statistical mechanics and mathematical problems. Lecture notes in physics, vol 20. Springer, Berlin, pp 1–113
47. Lee TD (1952) On some statistical properties of hydrodynamical and magneto-hydrodynamical fields. Q Appl Math 10:69–74
48. Leith C (1968) Diffusion approximation for two-dimensional turbulence. Phys Fluids 11:671–673. doi:10.1063/1.1691968
49. Leprovost N, Dubrulle B, Chavanis PH (2006) Dynamics and thermodynamics of axisymmetric flows: theory. Phys Rev E 73:046308. doi:10.1103/PhysRevE.73.046308
50. Lilly DK (1983) Stratified turbulence and the mesoscale variability of the atmosphere. J Atmos Sci 40:749–761
51. Lim CC (2001) A long range spherical model and exact solutions of an energy enstrophy theory for two-dimensional turbulence. Phys Fluids 13:1961
52. Lim CC (2012) Phase transition to super-rotating atmospheres in a simple planetary model for a nonrotating massive planet: exact solution. Phys Rev E 88(6):066304. doi:10.1103/PhysRevE.86.066304
53. Lucarini V, Blender R, Herbert C, Pascale S, Ragone F, Wouters J (2014) Mathematical and physical ideas for climate science. Rev Geophys 52:809–859. doi:10.1002/2013RG000446
54. Majda AJ, Wang X (2006) Nonlinear dynamics and statistical theories for basic geophysical flows. Cambridge University Press, Cambridge
55. McWilliams JC (1984) The emergence of isolated coherent vortices in turbulent flow. J Fluid Mech 146:21–43. doi:10.1017/S0022112084001750
56. Merilees PE, Warn H (1975) On energy and enstrophy exchanges in two-dimensional non-divergent flow. J Fluid Mech 69(04):625–630
57. Michel J, 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. doi:10.1007/BF02100491
58. Michel J, Robert R (1994) Statistical mechanical theory of the great red spot of Jupiter. J Stat Phys 77:645–666. doi:10.1007/BF02179454
59. Miller J (1990) Statistical mechanics of Euler equations in two dimensions. Phys Rev Lett 65:2137–2140. doi:10.1103/PhysRevLett.65.2137
60. Miller J, Weichman PB, Cross MC (1992) Statistical mechanics, Euler's equation, and Jupiter's Red Spot. Phys Rev A 45:2328–2359. doi:10.1103/PhysRevA.45.2328
61. Naso A, Chavanis PH, Dubrulle B (2011) Statistical mechanics of Fofonoff flows in an oceanic basin. Eur Phys J B 80:493–517. doi:10.1140/epjb/e2011-10440-8
62. Naso A, Monchaux R, Chavanis PH, Dubrulle B (2010) Statistical mechanics of Beltrami flows in axisymmetric geometry: theory reexamined. Phys Rev E 81:066318. doi:10.1103/PhysRevE.81.066318
63. Nazarenko SV (2010) Wave turbulence. Lecture notes in physics, vol 825. Springer

64. Olver PJ (2000) Applications of Lie groups to differential equations. Graduate texts in mathematics. Springer
65. Onsager L (1944) Crystal statistics. I. A two-dimensional model with an order-disorder transition. Phys Rev 65:117
66. Padmanabhan T (1990) Statistical mechanics of gravitating systems. Phys Rep 188:285–362. doi:10.1016/0370-1573(90)90051-3
67. Paret J, Tabeling P (1997) Experimental observation of the two-dimensional inverse energy cascade. Phys Rev Lett 79:4162–4165
68. Penrose O, Lebowitz JL (1979) In: Montroll EW, Lebowitz JL (eds) Towards a rigorous molecular theory of metastability. Fluctuation phenomena, Chap. 5. North-Holland, Amsterdam, p 293
69. Potters M, Vaillant T, Bouchet F (2013) Sampling microcanonical measures of the 2D Euler equations through Creutz's algorithm: a phase transition from disorder to order when energy is increased. J Stat Mech P02017. doi:10.1088/1742-5468/2013/02/P02017
70. Pouquet A, Marino R (2013) Geophysical turbulence and the duality of the energy flow across scales. Phys Rev Lett 111:234,501. doi:10.1103/PhysRevLett.111.234501
71. Qi W, Marston JB (2014) Hyperviscosity and statistical equilibria of Euler turbulence on the torus and the sphere. J Stat Mech P07020. doi:10.1088/1742-5468/2014/07/P07020
72. Rhines PB (1979) Geostrophic turbulence. Annu Rev Fluid Mech 11:401–441
73. Robert R (1989) Concentration et entropie pour les mesures d'Young. C R Acad Sci Paris, Sér I 309:757
74. Robert R (1990) Etats d'équilibre statistique pour l'écoulement bidimensionnel d'un fluide parfait. C R Acad Sci Paris, Série I 311:575
75. Robert R (1991) A maximum-entropy principle for two-dimensional perfect fluid dynamics. J Stat Phys 65:531–553. doi:10.1007/BF01053743
76. Robert R (2000) On the statistical mechanics of 2D Euler equation. Commun Math Phys 212:245–256. doi:10.1007/s002200000210
77. Robert R, Sommeria J (1991) Statistical equilibrium states for two-dimensional flows. J Fluid Mech 229:291–310. doi:10.1017/S0022112091003038
78. Robert R, Sommeria J (1992) Relaxation towards a statistical equilibrium state in two-dimensional perfect fluid dynamics. Phys Rev Lett 69:2776–2779. doi:10.1103/PhysRevLett.69.2776
79. Rockafellar R (1970) Convex analysis. Princeton University Press, Princeton
80. Ruelle D (1969) Statistical mechanics: rigorous results. Benjamin, Amsterdam
81. Ruelle D (1989) Chaotic evolution and strange attractors. Lezioni Lincee, Accademia Nazionale dei Lincei
82. Rutgers MA (1998) Forced 2D turbulence: experimental evidence of simultaneous inverse energy and forward enstrophy cascades. Phys Rev Lett 81:2244–2247
83. Salmon R (1998) Lectures on geophysical fluid dynamics. Oxford University Press, New York
84. Serre D (1984) Les invariants du premier ordre de l'équation d'Euler en dimension trois. Physica D 13:105–136. doi:10.1016/0167-2789(84)90273-2
85. Smith LM, Yakhot V (1993) Bose condensation and small-scale structure generation in a random force driven 2D turbulence. Phys Rev Lett 71:352–355
86. Sommeria J (2001) Course 8: two-dimensional turbulence. In: Lesieur M, Yaglom A, David F (eds) New trends in turbulence, Les Houches theoretical physics summer school, Chap. 8. Springer
87. Tabeling P (2002) Two-dimensional turbulence: a physicist approach. Phys Rep 362:1–62
88. Thalabard S, Dubrulle B, Bouchet F (2014) Statistical mechanics of the 3D axisymmetric Euler equations in a Taylor-Couette geometry. J Stat Mech P01005
89. Touchette H (2009) The large deviation approach to statistical mechanics. Phys Rep 478:1–69. doi:10.1016/j.physrep.2009.05.002
90. Touchette H, Ellis RS, Turkington B (2004) An introduction to the thermodynamic and macrostate levels of nonequivalent ensembles. Physica A 340:138–146. doi:10.1016/j.physa.2004.03.088

91. Turkington B (1999) Statistical equilibrium measures and coherent states in two-dimensional turbulence. Comm Pure Appl Math 52:781–809. doi:10.1002/(SICI)1097-0312(199907)52:7<781:AID-CPA1>3.0.CO;2-C
92. Turkington B, Whitaker N (1996) Statistical equilibrium computations of coherent structures in turbulent shear layers. SIAM J Sci Comput 17:1414. doi:10.1137/S1064827593251708
93. Vallis GK (2006) Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation. Cambridge University Press, Cambridge
94. Venaille A (2012) Bottom-trapped currents as statistical equilibrium states above topographic anomalies. J Fluid Mech 699:500. doi:10.1017/jfm.2012.146
95. Venaille A, Bouchet F (2011) Oceanic rings and jets as statistical equilibrium states. J Phys Oceanogr 41:1860. doi:10.1175/2011JPO4583.1
96. Venaille A, Vallis GK, Griffies SM (2012) The catalytic role of beta effect in barotropization processes. J Fluid Mech 709:490–515. doi:10.1017/jfm.2012.344