5th Warsaw School of Statistical Physics

22-29 June 2013 Kazimierz Dolny, Poland

Editors: Bogdan Cichocki, Marek Napiórkowski and Jarosław Piasecki

WUW

5th Warsaw School of Statistical Physics

5th Warsaw School of Statistical Physics

Kazimierz Dolny, Poland

22 – 29 June 2013

Editors
B. Cichocki, M. Napiórkowski and J. Piasecki

Organized by
the Institute of Theoretical Physics, University of Warsaw
and Pro Physica Foundation

Cover design
Katarzyna Jarnuszkiewicz

Cover photograph
Bogdan Cichocki

Photographs of the lecturers and participants
Agnieszka Budek

Technical editor
Ewa Szymczak

ISBN 978-83-235-1561-6
ISBN 978-83-235-1739-9 PDF

Warsaw University Press
Sales Department: tel. +48 22 55 31 333
e-mail: dz.handlowy@uw.edu.pl
Internet Bookshop: http://www.wuw.pl/ksiegarnia

Foreword

Following the previous schools, which have taken place in Kazimierz every two years since 2005, the Institute of Theoretical Physics of the University of Warsaw organized the 5th Warsaw School of Statistical Physics (June 22th to 29th, 2013). The program of the school was essentially composed of six courses corresponding to various areas of research in the field of statistical physics. Six distinguished scientists presented pedagogical series of lectures bringing a clear explanation of basic theoretical ideas, and encouraging further research. The lectures were attended by PhD students, postgraduate researchers, and also by more experienced scientists interested in getting acquainted with a new field.

The present volume contains the texts of the courses. We are grateful to the invited speakers for their willingness to make their lecture notes ready for publication. We do hope the volume will be useful not only to the participants of the school but also to all those interested in the current development of ideas in statistical physics.

It is also a pleasure to acknowledge all those individuals and organizations (listed overleaf) who contributed to the success of the school.

Scientific Organizing Committee:
Bogdan Cichocki
Marek Napiórkowski
Jarosław Piasecki

For further information about the Schools see:
http://www.fuw.edu.pl/~wssph/

Sponsors

Faculty of Physics, University of Warsaw
Ministry of Science and Higher Education

Local Organizing Committee

Bogdan Cichocki - Chair
Paweł Jakubczyk - Secretary
Piotr Szymczak
Karol Wędołowski

Contents

1

Non-equilibrium Statistical Mechanics of the Stochastic Navier–Stokes Equations and Geostrophic Turbulence

Freddy Bouchet, Cesare Nardini, and Tomás Tangarife

Abstract Two-dimensional and geophysical turbulent flows have the property to self organize and create large scale coherent jets and vortices. This is for instance one of the major processes for the dynamics of Earth's atmosphere. Following Onsager initial insight, based on conjugated works by mathematicians and physicists, this fundamental physical process has found some explanations in the framework of statistical mechanics. An important step, initiated twenty years ago, has been the study of the equilibrium statistical mechanics for the 2D Euler and the related quasi-geostrophic models (the Miller-Robert-Sommeria theory).
Real geophysical and experimental flows are however dissipative and maintained by external forces. These lectures focus on recent theoretical development of the statistical mechanics of those non-equilibrium situations. Those progresses have been achieved using tools from field theory (path integrals and instantons), non-equilibrium statistical mechanics (large deviations, stochastic averaging). The aim of these lectures is to briefly introduce the theoretical aspects of this program in the simplest context: the 2D stochastic Euler or Navier-Stokes equations and the quasi-geostrophic equations.
We review path integral representations of stochastic processes, large deviations for transition probabilities, action minimization, instanton theory, for general mechanical systems forced by random forces. We will apply this framework in order to predict equilibrium and non-equilibrium phase transitions for the 2D Euler, Navier-Stokes, and quasi-geostrophic dynamics, and to predict the rates of rare transitions between two attractors in situations of first order phase transitions.
Kinetic theory of systems with long range interactions, both with and without stochastic external forces, are explained. Based on this kinetic theory, we predict non-equilibrium phase transitions, and discuss their recent experimental observations and numerical simulations.
Even if the model we have considered so far are too simple academic models, the

---

Freddy Bouchet
Laboratoire de physique, École Normale Supŕieure de Lyon et CNRS, 46 allée d'Italie, 69007 Lyon, France
email: Freddy.Bouchet@ens-lyon.fr

expected relevance of those approaches in the future for Earth atmosphere and climate dynamics is briefly discussed.

1 Introduction

1.1 Self-organization of two-dimensional and geophysical flows

Atmospheric and oceanic flows are three-dimensional (3D), but are strongly dominated by the Coriolis force mainly balanced by pressure gradients (geostrophic balance). The turbulence that develops in such flows is called geostrophic turbulence. Models describing geostrophic turbulence have the same type of additional invariants as those of the two-dimensional (2D) Euler equations. As a consequence, energy flows backward and the main phenomenon is the formation of large scale coherent structures (jets, cyclones and anticyclones). One such example is the formation of Jupiter's Great Red Spot, Fig. 1.

Fig. 1: Picture of Jupiter's Great Red Spot - a large scale vortex situated between bands of atmospheric jets. Photo courtesy of NASA: http://photojournal.jpl.nasa.gov/catalog /PIA00014.

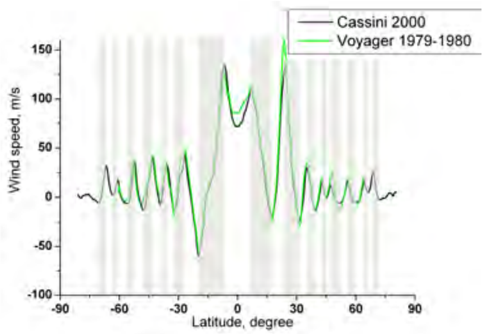


Fig. 2: Zonally averaged velocity profile in the upper troposphere of Jupiter. The flow is organised into alternating strong jets.

The analogy between 2D turbulence and geophysical turbulence is further emphasized by the theoretical similarity between the 2D Euler equations, describing 2D flows, and the layered quasi-geostrophic or shallow water models, describing the largest scales of geostrophic turbulence: both are transport equations for a scalar quantity by a non-divergent flow, conserving an infinite number of invariants.

The formation of large scale coherent structures is a fascinating problem and an essential part of the dynamics of Earth's atmosphere and oceans. This is the main motivation for setting up a theory for the self-organization of 2D turbulence.

1.2 Statistical mechanics of the self-organization of two-dimensional and geophysical flows: Onsager's equilibrium route

Any turbulence problem involves a huge number of degrees of freedom coupled via complex nonlinear interactions. The aim of any theory of turbulence is to understand the statistical properties of the velocity field. It is thus extremely tempting to attack these problems from a statistical mechanics point of view.

Statistical mechanics is indeed a very powerful set of theoretical tools that allows us to reduce the complexity of a system down to a few thermodynamic parameters. As an example, the concept of phase transition allows us to describe drastic changes of the whole system when a few external parameters are changed. Statistical mechanics is the main theoretical approach we develop in these lectures. It succeeds in explaining many of the phenomena associated with two-dimensional turbulence [13].

This may seem surprising at first, as it is a common belief that statistical mechanics is not successful in handling turbulence problems. The reason for this belief is that most turbulence problems are intrinsically far from equilibrium. For instance, the forward energy cascade in three-dimensional turbulence involves a finite energy dissipation, no matter how small the viscosity (anomalous dissipation) (see for instance Onsager's insightful consideration of the non-conservation of energy by the three dimensional Euler equations [28]). As a result of this finite energy flux, three dimensional turbulent flows cannot be considered close to some equilibrium distribution.

By contrast, two-dimensional turbulence does not suffer from the anomalous dissipation of the energy, so equilibrium statistical mechanics, or close to equilibrium statistical mechanics makes sense when small fluxes are present. The first attempt to use equilibrium statistical mechanics ideas to explain the self-organization of two-dimensional turbulence dates from Onsager work in 1949 [51] (see [28] for a review of Onsager's contributions to turbulence theory). Onsager worked with the point-vortex model, a model that describes the dynamics of singular point vortices, first used by Lord Kelvin and which corresponds to a special class of solutions of the 2D Euler equations. The equilibrium statistical mechanics of the point-vortex model has a long and very interesting history, with wonderful pieces of mathematical achievements [1; 18; 21; 26; 27; 37; 39; 51].

The generalization of Onsager's ideas to the 2D Euler equations with a continuous vorticity field, taking into account all invariants, has been proposed in the beginning of the 1990s [45; 57; 58; 60], leading to the Miller–Robert–Sommeria theory (MRS theory). The MRS theory includes the previous Onsager theory and determines within which limits the theory will give relevant predictions and results.

The MRS theory deals with the microcanonical invariant measure. It predicts that most microscopic states (vorticity field) concentrate into a single macrostate (most vorticity fields basically have the same large scale velocity field). This explains why one should expect the flow to self-organize into this equilibrium macrostate. This equilibrium macrostate is characterized by the maximization of an entropy with some constraints related to dynamics invariants. The aim of section 3 is to sketch the derivation of this variational problem, which is the basis of the theory. Then application to the Great Red Spot of Jupiter will be briefly summarized.

These two points constitute a very brief overview of equilibrium statistical mechanics. Over the last fifteen years, the RSM equilibrium theory has been applied successfully to a large class of problems, for both the two-dimensional Euler and quasi-geostrophic equations. This includes many interesting applications, such as the predictions of phase transitions in different contexts, a model for the Great Red Spot and other Jovian vortices, and models of ocean vortices and jets. A detailed description of the statistical mechanics of 2D and geophysical flows (theory) and of these geophysical applications is presented in the review [13]. Older reviews or books [40; 42; 64], give a very interesting complementary viewpoint, stressing mainly the theory and laboratory experiments. The note by Y. Pomeau [54] gives also a very interesting comment on the reason why the two-dimensional Euler equations, by contrast with most other equilibrium approach for classical field theory, does not suffer from the Rayleigh-Jeans paradox (basically the fact that a classical field has an infinite heat capacity). This point is further discussed in [13]. Finally we stress that equilibrium statistical mechanics for two dimensional and geophysical flows is still a very active subject, with many contribution during the last few years [8; 24; 33–35; 48; 49; 55; 66; 70; 71; 73], many of them by bright young scientists.

As far as equilibrium statistical mechanics is concerned, the aim of these lecture is just to explain the basis of Miller–Robert–Sommeria theory, explain how to compute the entropy of macrostate and thus their probability through the use of large deviation theory. We discuss these points in section 3 at a level which is as elementary as possible.

1.3 Non-equilibrium statistical mechanics of the self-organization of two-dimensional and geophysical flows: statistical mechanics and dynamics

Most of natural turbulent flows are not freely evolving, they are rather constantly forced and dissipated. Then, in statistically stationary regimes, power input through external forces balance energy dissipation on average. In the limit of very small forces and dissipation, compared to conservative terms of the dynamics, it is expected to find a strong relation between these non-equilibrium flows and some of the states predicted by equilibrium statistical mechanics. In order to give a precise meaning to this general idea, and to deal with far from equilibrium situations, it is essential to develop also the non-equilibrium statistical mechanics of the 2D Euler, 2D Navier-Stokes and barotropic quasi-geostrophic equations. As we discuss below, this has been the subject of recent key advances in the applications of statistical mechanics to turbulent flows. This is actually the main subject of these lectures.

We present two non-equilibrium statistical mechanics approaches: the first deals with non-equilibrium first order phase transitions and the computation of transition rates using large deviations, and the second is a kinetic theory approach to the prediction of the large scale flows.

1.3.1 Statistical mechanics of paths in phase space and non-equilibrium bistable turbulent flows

Many turbulent flows can evolve and self-organize towards two or more very different states. In some of these systems, the transitions between two of such states are rare and occur relatively rapidly. Examples include the Earth magnetic field reversals (over geological timescales) or in magnetic field reversal in MHD experiments (e.g. the Von Kármán Sodium (VKS) turbulent dynamo in Fig. 3) [3], Rayleigh-Bénard convection cells [17; 20; 50; 65], 2D turbulence [10; 41; 63] (see Fig. 4), 3D flows [56] and for ocean and atmospheric flows [62; 72]. The understanding of these transitions is an extremely difficult problem due to the large number of degrees of freedoms, large separation of timescales and the non-equilibrium nature of these flows.

However, for forced-dissipated turbulent systems it is unclear how to define the set of attractors for the dynamics. Although, in the limit of weak forcing and dissipation, one would expect that the set of attractors would converge to the ones of the deterministic equation. In the case of the 2D Euler equations, equilibrium statistical mechanics in the form of the equilibrium Miller-Robert–Sommeria theory allows for the prediction set of attractors for the dynamics. They are a subsets of the steady states of the 2D Euler equations, then equilibrium statistical mechanics gives a first partial answer to the question of attractors.

Moreover, simulations of the 2D Navier-Stokes equations in the weak force and dissipation limit showed that the dynamics actually concentrates precisely close to the set of the 2D Euler equations attractors [10]. Interestingly, the same simulation showed sporadic non-equilibrium phase transitions, where the system spontaneously switched between two apparently stable steady states resulting in a complete change in the macroscopic behavior (see figure 4). If the forces and dissipation are weak, then these transitions are actually extremely rare, occurring on a timescale much longer than the dynamical timescale.

In such situations, when the turbulent flow switches at random times from one type of attractor to another, a theoretical aim is to compute the transition rate. It is also often the case that most transition paths from one attractor to another concentrate to a single path, then a natural aim is to compute this most probable

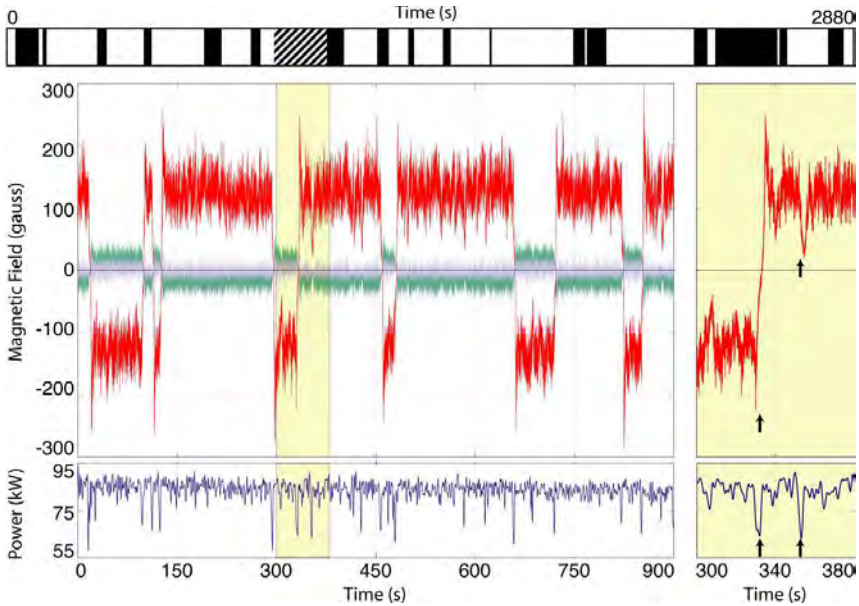


Fig. 3: Figure taken from [3] showing random transitions between meta-stable orientations of the magnetic field in an experimental turbulent dynamo. The main azimuthal component of the magnetic field is shown in red.

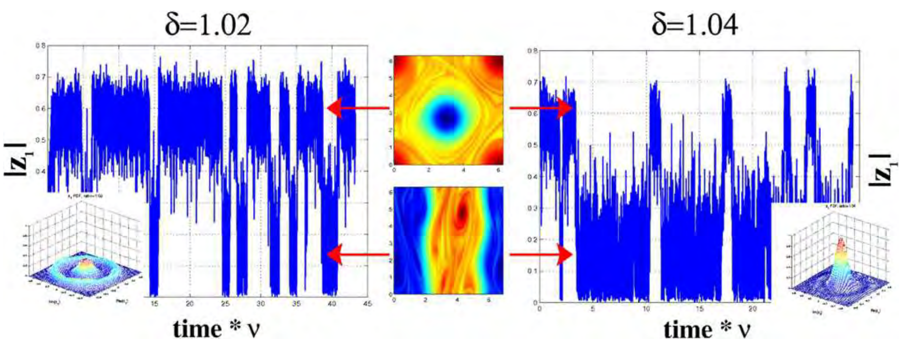


Fig. 4: Figure taken from [10] showing rare transitions (illustrated by the Fourier component of the largest y mode) between two large scale attractors of the periodic 2D Navier-Stokes equations. The system spends the majority of its time close to the vortex dipole and parallel flows configurations.

path. In order achieve those aims, we will use a path integral representation of the transition probabilities and study its semi-classical limit, in an asymptotic expansion where the small parameter is the one that determines both the force and dissipation amplitude. In this limit, if this semi-classical approach is relevant, one expects a large deviation result, similar to the one obtained through the Freidlin-Wentzell theory[30]. In order to illustrate in a pedagogical way the general approach, we will treat in these lectures the classical case of the Kramer model (computation of the transition rate for a particle in a double well potential). We will generalize the discussion to a set of Langevin dynamics that includes the two-dimensional Euler and Quasi-Geostrophic Langevin dynamics, and finally we will discuss partial results for the two-dimensional Navier-Stokes equations when detailed balance is not satisfied. Those different points are discussed in section 4.4.

1.3.2 Kinetic theory of zonal jets

One example of spontaneous emergence of large-scale coherent structures in geophysical flows is the formation of zonal (east-west) jets. The common pictures of Jupiter perfectly illustrate this fact: the surface flow is clearly organized into parallel, alternating zonal jets as shown in figure 2, with also the presence of giant and very stable vortices such as the Great Red Spot. Such large scale features are on one hand slowly dissipated, mainly due to a large-scale friction mechanism, and on the other hand maintained by the small-scale turbulence, through Reynolds' stresses. The main mechanism is thus a transfer of energy from the forcing scale (due to barotropic and baroclinic instabilities) to the turbulent scales and until the scale of the jets.

An important point in this phenomenology is the fact that the turbulent fluctuations are of very weak amplitude compared to the amplitude of the zonal jet, and that they evolve much faster. This means that the typical time scale of advection and shear of the fluctuations by the jet is much smaller that the typical time scale of formation or dissipation of the whole jet. This time scale separation is a very specific property of the geophysical large-scale structures.

In this turbulent context, the understanding of jet formation requires averaging out the effect of rapid turbulent degrees of freedom in order to describe the slow evolution of the jet structure. Such a task, an example of closure, is usually extremely hard to perform for turbulent flows. Using the time-scale separation mentioned earlier, we prove that it can be performed explicitly in this problem. This approach, called a kinetic theory by analogy with similar approaches in the statistical mechanics of systems with long range interactions, is presented in section 5.

1.4 A contemporary approach of statistical mechanics: large deviation theory

Onsager was the first to consider a statistical mechanics explanation of two-dimensional turbulent flows [51]. At the time he was scientifically active, Onsager made a large number of decisive contributions to statistical mechanics theory: solutions of the 2D Ising model, reciprocity relations, contributions to the statistical mechanics of electrolytes and turbulence, and so on. Since that time the theoretical approaches for treating statistical mechanics problems have been completely renewed. One of the main changes has been the use of the language of large deviation theory for more than 30 years. For instance, recent results in the understanding of equilibrium statistical mechanics problems, proving fluctuation theorems (Onsager's reciprocity relations generalized far from equilibrium), and in dealing with non-equilibrium statistical mechanics problems, are all related to large deviation theory.

Interestingly, as we discuss in these lectures, the route proposed by Onsager in his 1949 paper [51] in order to understand the self-organization of two-dimensional flows, led a few decades later to some of the first applications of large deviation theory to equilibrium statistical mechanics problems.

The theory of large deviations deals with the asymptotic behavior of the exponential decay of the probabilities of rare or extreme events. The associated limiting parameter is usually taken to be the number of observations, the number of particles, but can be other parameters, such as vanishing noise or the temperature of a chemical reaction, or large time. Large deviation theory can be considered a generalization of the central limit theorem, with the refinement of including information about the behavior of the tails of the probability density. The main result of large deviation theory is the large deviation principle, a result describing the leading asymptotic behavior of the tails or large deviations of the probability distribution in the limit $N \to \infty$. For instance, the large deviation principle for a random variable X_N is

$$\lim_{N\to\infty} -\frac{1}{N}\log[P(X_N = x)] = I(x), \tag{1}$$

where P is the probability density for the random variable X_N, and $I(x)$ is called the rate function. For instance, if $X_N = (1/N)\sum_{i=1}^{N} x_i$, where x_i are independent identically distributed random variables then $I(x)$ is given by Cramer's theorem.

Beside the applications described in the previous sections, the aim of these lectures is to explain and derive heuristically large deviation results for the equilibrium statistical mechanics of the two-dimensional Euler and quasi-geostrophic equations (equilibrium) and for the 2D Navier-Stokes or quasi-geostrophic equations with stochastic forces (non-equilibrium). The large deviation result for the equilibrium case (section 3) is derived through a generalization of Sanov theorem, and leads to a formula for the probability of macrostates for the microcanonical measures. The large deviation results for the non-equilibrium cases (section 4.4) are derived through semi-classical limits in path integrals (or equivalently the Freidlin-Wentzell framework) and lead to the evaluation of transition paths and transition probabilities for bistable turbulent flows, close to non-equilibrium phase transitions.

1.5 Organization of those lectures

In section 2, we state the equations of motion and their conservation laws. In section 3, we construct microcanonical invariant measures for the 2D Euler equations and discuss the entropy maximization problem in predicting the most probably steady states on the 2D Euler equation. In section 4.4, we discuss large deviations for non-equilibrium problems and illustrate this using a simple academic example, the problem of computation of transition rate for the Kramer problem, followed by the application to the 2D Navier-Stokes equations. Finally, in section 5 we discuss the kinetic theory of zonal jets for the barotropic quasi-geostrophic dynamics.

2 The 2D Euler, barotropic Quasi Geostrophic, and stochastic Navier–Stokes equations

2.1 Equations of motion

The aim of this section is to present the simplest model that describes two-dimensional and geophysical turbulent flows: the two-dimensional Navier-Stokes equation and the barotropic equation with stochastic forcing. In the limit when forces and dissipation go to zero, the two-dimensional Navier-Stokes equation reduces to the two-dimensional Euler equation. We describe the conservation laws for these equations and their influence on the dynamics. The review [13] gives a very brief introduction to geophysical fluid dynamics and the quasi-geostrophic model. A more complete introduction is found in textbooks of geophysical fluid dynamics [53; 68].

We are interested in the non-equilibrium dynamics associated to the two-dimensional stochastically forced barotropic equations (also called barotropic Quasi-Geostrophic equations):

$$\frac{\partial q}{\partial t} + \mathbf{v}\left[q-h\right]\cdot\nabla q = -\alpha\omega + \nu\Delta\omega + \sqrt{2\alpha}\eta, \tag{2}$$

$$\mathbf{v} = \mathbf{e}_z \times \nabla\psi, \quad q = \omega + h(y) = \Delta\psi + h, \tag{3}$$

where ω, $\mathbf{v}$ and ψ are respectively the vorticity, the non-divergent velocity, and the streamfunction. For simplicity, in these lectures we consider the dynamics on a periodic domain $\mathcal{D} = [0, 2\delta\pi) \times [0, 2\pi)$ with aspect ratio δ. Then ψ is periodic with the further condition $\int_{\mathcal{D}} \mathrm{d}\mathbf{r}\ \psi = 0$. q is the potential vorticity, and h is a given topography function with $\int_{\mathcal{D}} \mathrm{d}\mathbf{r}\ h = 0$. For $h = 0$, the barotropic equations reduces to the 2D Navier-Stokes equation.

The linear friction term $-\alpha\omega$ models large scale dissipation. We consider non-dimensional equations, where a typical energy is of order 1 (see [13]) such that ν is the inverse of the Reynolds number and α is the inverse of a Reynolds number based on the large scale friction. We assume that the Reynolds numbers satisfy $\nu \ll \alpha \ll 1$. In the limit of weak forces and dissipation $\lim_{\alpha\to 0}\lim_{\nu\to 0}$, the 2D Navier-Stokes equations converge to the two-dimensional Euler equations for finite time, but the type of forcing and dissipation determines to which set of attractors the dynamics evolve to over a very long time. The curl of the forcing $\eta(\mathbf{x},t)$ is a white in time Gaussian field defined by $\langle \eta(\mathbf{x},t)\eta(\mathbf{x}',t')\rangle = C(\mathbf{x}-\mathbf{x}')\delta(t-t')$, where C is the correlation function of a stochastically homogeneous noise.

The two-dimensional Euler equations ($h=0$), or the inertial barotropic equation ($h \neq 0$), are given by Eq. (2) with forces and dissipation set to zero ($\alpha = \nu = 0$).

2.2 Conservation laws for the inertial dynamics

The kinetic energy of the flow is given by

$$\mathcal{E}[q] = \frac{1}{2}\int_{\mathcal{D}} \mathrm{d}\mathbf{r}\, \mathbf{v}^2 = \frac{1}{2}\int_{\mathcal{D}} \mathrm{d}\mathbf{r}\, (\nabla\psi)^2 = -\frac{1}{2}\int_{\mathcal{D}} \mathrm{d}\mathbf{r}\, (q-h)\psi, \tag{4}$$

where the last equality is obtained with an integration by parts. The kinetic energy is conserved for the dynamics of the two-dimensional Euler and inertial barotropic equations i.e. $\mathrm{d}\mathcal{E}/\mathrm{d}t = 0$,. These equations also conserve an infinite number of functionals, named Casimirs. They are related to the degenerate structure of the infinite-dimensional Hamiltonian system and can be understood as invariants arising from Noether's theorem [61]. These functionals are of the form

$$\mathcal{C}_s[q] = \int_{\mathcal{D}} s(q)\,\mathrm{d}\mathbf{r}, \tag{5}$$

where s is any sufficiently regular function. We note that on a doubly-periodic domain the total circulation

$$\Gamma = \int_{\mathcal{D}} q\,\mathrm{d}\mathbf{r}, \tag{6}$$

is necessarily equal to zero: $\Gamma = 0$.

The infinite number of conserved quantities are responsible for the equations having an infinite (continuous) set of steady states (see section 2 in [13]). Any of the infinite number of steady states of the 2D Euler or inertial barotropic equations satisfy

$$\mathbf{v}\cdot\nabla q = 0.$$

For instance, if there is a functional relation between the potential vorticity and the streamfunction, i.e. $q = \Delta\psi = f(\psi)$, where f is any continuous function, then using 2 one easily check that $\mathbf{v}\cdot\nabla q = 0$. Physically, these states are important because some of them act as attractors for the dynamics.

There is also a strong empirical and numerical evidence that a complex evolution of the two-dimensional Euler equations leads most of the times to attractors that are steady states of the equations. The specific function f that is reached after a complex evolution can be predicted in certain situation using equilibrium statistical mechanical arguments presented in the next section (see [13] for more details).

2.3 The conservation of the vorticity distribution

The two-dimensional Euler and inertial barotropic equations conserve the distribution of potential vorticity, i.e. the total area of a specific potential vorticity level set is conserved. As we explain now the conservation of the potential vorticity distribution is equivalent to the conservation of all Casimirs.

We first prove that the potential vorticity distribution is conserved as a consequence of Casimir conservation laws. We consider the special class of Casimir (5):

$$C(\sigma) = \int_{\mathcal{D}} H(-q+\sigma)\,\mathrm{d}\mathbf{r}, \tag{7}$$

where $H(\cdot)$ is the Heaviside step function. The function $C(\sigma)$ returns the area occupied by all potential vorticity levels smaller or equal to σ. $C(\sigma)$ is an invariant for any σ and therefore any derivative of $C(\sigma)$ is also conserved. Therefore, the distribution of vorticity, defined as $D(\sigma) = C'(\sigma)$, where the prime denotes a derivation with respect to σ, is also conserved by the dynamics. The expression $D(\sigma)d\sigma$ is the area occupied by the vorticity levels in the range $\sigma \le q \le \sigma + d\sigma$.

Moreover, any Casimir can be written in the form

$$\mathcal{C}_f[q] = \int_{\mathcal{D}} \mathrm{d}\sigma\, f(\sigma)\, D(\sigma).$$

The conservation of all Casimirs, Eq. (5), is therefore equivalent to the conservation of $D(\sigma)$.

The conservation of the distribution of vorticity levels, as proven above, can also be understood from the equations of motion. We find that $Dq/Dt = 0$, showing that the values of the potential vorticity field are Lagrangian tracers. This means that the values of q are transported through the non-divergent velocity field, thus keeping the distribution unchanged.

From now on, we restrict ourselves to a K-level vorticity distribution. We make this choice for pedagogical reasons, but a generalization of the discussion of next section to a continuous vorticity distribution is straightforward. The K-level vorticity distribution is defined as

$$D(\sigma) = \sum_{k=1}^{K} A_k \delta(\sigma - \sigma_k), \tag{8}$$

where A_k denotes the area occupied by the vorticity value σ_k. The areas A_k are not arbitrary, their sum is the total area $\sum_{k=1}^{K} A_k = |\mathcal{D}|$. Moreover, the constraint (6), imposes the constraint $\sum_{k=1}^{K} A_k \sigma_k = 0$.

3 Equilibrium statistical mechanics and the mean field variational problem as a large deviation result

3.1 Large deviation theory in 2D turbulence, the equilibrium mean field variational problem

The first large deviation results in two-dimensional turbulence have been obtained in the context of the theory for the 2D Euler equations. Michel and Robert [44] have studied the large deviation of Young measures and have suggested that the entropy of the Miller–Robert–Sommeria theory is the analogue of a large deviation rate function. By considering a prior distribution for the vorticity invariants, in a framework where the invariants are considered in a canonical ensemble rather than in a microcanonical one, Boucher and collaborators [5] have given a derivation of a large deviation result based on finite dimensional approximations of the vorticity field. The beginning of the nineties has also been a time of intense study of the statistical mechanics of the point vortex model [4; 18; 27; 28; 38; 39], a special class of solution of the two-dimensional Euler equations. Among those study, large deviations results for the equilibrium measures where also obtained.

The aim of this section is to present a heuristic construction of microcanonical invariant measures for the 2D Euler equations. This construction primarily follows the initial ideas of the previous works [5; 44], but is much simplified. Moreover, for pedagogical reasons, the reading of this heuristic presentation does not imply any knowledge of large deviation theory and avoids any technical discussion. These measures are constructed using finite dimensional approximation of the vorticity field, with N^2 number of degrees of freedom. N^2 is then the large deviation parameter and the entropy appears as the analogue of the large deviation rate function.

In order to state the main result, let us define $p(\mathbf{r},\sigma)$ as the local probability to observe vorticity values equal to σ at point $\mathbf{r}$: $p(\mathbf{r},\sigma) = \langle \delta(\omega(\mathbf{r}) - \sigma) \rangle$, where δ is the Dirac delta function (we consider averaging $\langle \cdot \rangle$ over the microcanonical measure, see section 3.2). We also define $\overline{\omega}(\mathbf{r}) = \int \mathrm{d}\sigma\, \sigma p(\mathbf{r},\sigma)$ the local vorticity average. Then the large deviation rate function for $p(\mathbf{r},\sigma)$ is $S(E_0) - S[p, E_0]$ where

$$S[p, E_0] = \mathcal{S}[p] \equiv \int_{\mathcal{D}} \sum_k p_k \log p_k \, \mathrm{d}\mathbf{r} \tag{9}$$

if the constraints $\mathcal{N}[p] = 1$, $\forall k,\ A\,[p_k] = A_k$ and $\mathcal{E}[\overline{\omega}] = E_0$ are satisfied, and $S[p, E_0] = -\infty$ otherwise, and where

$$S(E_0) = \sup_{\{p \mid \mathcal{N}[p]=1\}} \left\{ \mathcal{S}[p] \mid \mathcal{E}[\bar{\omega}] = E_0, \forall k\ A[p_k] = A_k \right\}, \tag{10}$$

with E_0, A_k and $\mathcal{N}$, the energy, the vorticity distribution, and the probability normalization defined in section 3.3 respectively.

The interpretation of this result is that the most probable value for the local probability is the maximizer of the variational problem (10), and that the probability to observe a departure from this most probable state is exponentially large, with parameter N^2 and rate function (9). Furthermore, the classical mean field equation for the streamfunction ψ can be derived from (9), as discussed in reference [6].

In next sections, we define precisely the microcanonical measure for the 2D Euler equations (section 3.2) and prove that the entropy $S[p, E_0]$ is a large deviation rate function for p (section 3.3). This justifies the mean field variational problem (10).

3.2 Microcanonical measure

In order to properly construct a microcanonical measure, we discretize the vorticity field on a uniform grid with N^2 grid points, define a measure on the corresponding finite-dimensional space and take the limit $N \to \infty$. A uniform grid has to be chosen in order to comply with a formal Liouville theorem for the 2D Euler equations [14; 59].

We denote the lattice points by $\mathbf{r}_{ij} = \left(\frac{i}{N}, \frac{j}{N}\right)$, with $0 \leq i, j \leq N-1$ and denote $\omega_{ij} \equiv \omega(\mathbf{r}_{ij})$ the vorticity value at point $\mathbf{r}_{ij}$. The total number of points is N^2.

As discussed in the previous section, we assume $D(\sigma) = \sum_{k=1}^{K} A_k \delta(\sigma - \sigma_k)$. For this finite-$N$ approximation, our set of microstates (configuration space) is then

$$X_N = \left\{ \omega^N = (\omega_{ij})_{0 \leq i,j \leq N-1} \mid \forall i, j\ \omega_{ij} \in \{\sigma_1, \ldots, \sigma_K\} \right.$$
$$\left. \text{and } \forall k\ \ \# \{\omega_{ij} \mid \omega_{ij} = \sigma_k\} = N^2 A_k \right\}.$$

Here, $\#(A)$ is the cardinal of set A. We note that X_N depends on $D(\sigma)$ through A_k and σ_k (see (8)).

Using the above expression we define the energy shell $\Gamma_N(E_0, \Delta E)$ as

$$\Gamma_N(E_0, \Delta E) = \left\{ \omega^N \in X_N \mid E_0 \leq \mathcal{E}_N \left[\omega^N \right] \leq E_0 + \Delta E \right\},$$

where

$$\mathcal{E}_N = \frac{1}{2N^2} \sum_{i,j=0}^{N-1} \mathbf{v}_{ij}^2 = -\frac{1}{2N^2} \sum_{i,j=0}^{N-1} \omega_{ij} \psi_{ij},$$

is the finite-N approximation of the system energy, with $\mathbf{v}_{ij} = \mathbf{v}(\mathbf{r}_{ij})$ and $\psi_{ij} = \psi(\mathbf{r}_{ij})$ being the discretized velocity field and streamfunction field, respectively. ΔE is the width of the energy shell. Such a finite width is necessary for our discrete approximation, as the cardinal of X_N is finite. Then the set of accessible energies on X_N is also finite. Let $\Delta_N E$ be the typical difference between two successive achievable energies. We then assume that $\Delta_N E \ll \Delta E \ll E_0$. The limit measure defined below is expected to be independent of ΔE in the limit $N \to \infty$.

The fundamental assumption of statistical mechanics states that each microstate in the configuration space is equiprobable. By virtue of this assumption, the probability to observe any microstate is $\Omega_N^{-1}(E_0, \Delta E)$, where $\Omega_N(E_0, \Delta E)$ is the number of accessible microstates, i.e. the cardinal of the set $\Gamma_N(E_0, \Delta E)$. The finite-N specific Boltzmann entropy is defined as

$$S_N(E_0, \Delta E) = \frac{1}{N^2} \log \ \Omega_N(E_0, \Delta E). \tag{11}$$

The microcanonical measure is then defined through the expectation values of any observable A. For any observable $A[\omega]$ (for instance a smooth functional of the vorticity field), we define its finite-dimensional approximation by $A_N\left[\omega^N\right]$. The expectation value of A_N for the microcanonical measure reads

$$\left\langle \mu_N(E_0, \Delta E),\ A_N\left[\omega^N\right]\right\rangle_N \equiv \left\langle A_N\left[\omega^N\right]\right\rangle_N \equiv \frac{1}{\Omega_N(E_0, \Delta E)} \sum_{\omega^N \in \Gamma_N(E_0, \Delta E)} A_N\left[\omega^N\right].$$

The microcanonical measure μ for the 2D Euler equation is defined as a limit of the finite-N measure:

$$\langle \mu(E_0),\ A[\omega]\rangle \equiv \lim_{N\to\infty} \left\langle \mu_N(E_0, \Delta E),\ A_N\left[\omega^N\right]\right\rangle_N.$$

The specific Boltzmann entropy is then defined as

$$S(E_0) = \lim_{N\to\infty} S_N(E_0, \Delta E). \tag{12}$$

3.3 The mean field variational problem as a large deviation result

Computing the Boltzmann entropy by direct evaluation of Eq. (12) is usually an intractable problem. However, we shall proceed in a different way and show that this alternative computation yields the same entropy in the limit $N \to \infty$. We give heuristic arguments in order to prove that the computation of the Boltzmann entropy Eq. (12) is equivalent to the maximization of the constrained variational problem (10) (called a mean field variational problem). This variational problem is the foundation of the RSM approach to the equilibrium statistical mechanics for the 2D Euler equations. The essential message is that the entropy computed from the mean field variational problem (10) and from Boltzmann's entropy definition (12) are the equal in the limit $N \to \infty$. The ability to compute the Boltzmann entropy through this type of variational problems is one of the cornerstones of statistical mechanics.

Our heuristic derivation is based on the same type of combinatorics arguments as the ones used by Boltzmann for the interpretation of its H function in the theory of relaxation to equilibrium of a dilute gas. This derivation doesn't use the technicalities of large deviation theory. The aim is to actually obtain the large deviation interpretation of the entropy and to provide a heuristic understanding using basic mathematics only. The modern mathematical proof of the relationship between the Boltzmann entropy and the mean field variational problem involves Sanov theorem.

Macrostates are set of microscopic configurations sharing similar macroscopic behaviors. Our aim is to properly identify macrostates that fully describe the main features of the largest scales of 2D turbulent flow, and then to compute their probability or entropy.

Let us first define macrostates through local coarse-graining. We divide the $N \times N$ lattice into $(N/n) \times (N/n)$ non-overlapping boxes each containing n^2 grid points (n is an even number, and N is a multiple of n). These boxes are centered on sites $(i,j) = (In, Jn)$, where integers I and J verify $0 \leq I, J \leq N/n - 1$. The indices (I,J) label the boxes.

For any microstate $\omega^N \in \Gamma_N$, let f_{IJ}^k be the frequency to find the value σ_k in the box (I,J)

$$F_{IJ}^k(\omega^N) = \frac{1}{n^2} \sum_{i=I-n/2+1}^{I+n/2} \sum_{j=J-n/2+1}^{J+n/2} \delta_{\mathrm{d}}(\omega_{ij} - \sigma_k),$$

where $\delta_{\mathrm{d}}(x)$ is equal to one whenever $x = 0$, and zero otherwise. We note that for all (I,J), $\sum_{k=1}^{K} F_{IJ}^k(\omega^N) = 1$.

A macrostate $p_N = \left\{p_{IJ}^k\right\}_{0 \leq I,J \leq N/n-1; 1 \leq k \leq K}$, is the set of all microstates of $\omega^N \in X_N$ such that $F_{IJ}^k(\omega^N) = p_{IJ}^k$ for all I, J, and k (by abuse of notation, and for simplicity, $p_N = \left\{p_{IJ}^k\right\}_{0 \leq I,J \leq N/n-1; 1 \leq k \leq K}$ refers to both the set of values and to the set of microstates having the corresponding frequencies). The entropy of the macrostate is defined as the logarithm of the number of microstates in the macrostate

$$S_N[p_N] = \frac{1}{N^2} \log\left(\# \left\{\omega^N \in X_N \,\middle|\, \text{ for all } I, J, \text{ and } k,\ \ F_{IJ}^k(\omega^N) = p_{IJ}^k\right\}\right). \quad (13)$$

Following an argument by Boltzmann, it is a classical exercise in statistical mechanics, using combinatorics and the Stirling formula, to prove that in the limit $N \gg n \gg 1$, without taking into account of the area constraints A_k, the entropy of the macrostate would converge to

$$S_N[p_N] \overset{N \gg n \gg 1}{\sim} \mathcal{S}_N[p_N] = -\frac{n^2}{N^2} \sum_{I,J=0}^{N/n-1} \sum_{k=1}^{K} p_{IJ}^k \log\ p_{IJ}^k$$

if $\forall I, J,\ \mathcal{N}[p_{IJ}] = 1$, and $S_N[p_N] \sim -\infty$ otherwise, where $\mathcal{N}[p_{IJ}] \equiv \sum_k p_{IJ}^k$. The area constraints are easily expressed as constraints over p_N: $A_N\left[p_N^k\right] \equiv \frac{n^2}{N^2} \sum_{I,J=0}^{N/n-1} p_{IJ}^k = A_k$ and $\forall I, J,\ \mathcal{N}[p_{IJ}] = 1$. An easy generalization of the above formula gives

$$S_N[p_N] \overset{N \gg n \gg 1}{\sim} \mathcal{S}_N[p_N]$$

if $\forall k,\ A_N\left[p_N^k\right] = A_k$, and $S_N[p_N] \sim -\infty$ otherwise. In the theory of large deviation, this result could have been obtained using Sanov's theorem. We now consider a new macrostate (p_N, E_0) which is the set of microstates ω^N with energy $\mathcal{E}_N\left[\omega^N\right]$ verifying $E_0 \leq \mathcal{E}_N\left[\omega^N\right] \leq E_0 + \Delta E$ (the intersection of $\Gamma_N(E_0, \Delta E)$ and p_N). For a given macrostate p_N, not all microstates have the same energy. The constraint on the energy thus can not be recast as a simple constraint on the macrostate p_N. Then one has to treat the energy constraint in a more subtle way. The energy is

$$\mathcal{E}_N\left[\omega^N\right] = -\frac{1}{2N^2}\sum_{i,j=0}^{N-1}\omega_{ij}^N\psi_{ij}^N.$$

The streamfunction ψ_{ij}^N is related to ω^N through

$$\psi_{ij} = \frac{1}{N^2}\sum_{i',j'=0}^{N-1}G_{ij,i'j'}\omega_{i'j'}^N,$$

where $G_{ij,i'j'}$ is the Laplacian Green function in the domain $\mathcal{D}$. In the limit $N \gg n \gg 1$, the variations of $G_{ij,i'j'}$ for (i',j') running over the small box (I,J) are vanishingly small. Then $G_{ij,i'j'}$ can be well approximated by their average value over the boxes $G_{IJ,I'J'}$. Then

$$\psi_{ij} \simeq \psi_{IJ} \equiv \frac{1}{N^2}\sum_{I',J'=0}^{N/n-1}G_{IJ,I'J'}\sum_{i'=I-n/2+1}^{I+n/2}\;\sum_{j'=J-n/2+1}^{J+n/2}\omega_{i'j'}^N = \frac{n^2}{N^2}\sum_{I',J'=0}^{N-1}G_{IJ,I'J'}\overline{\omega_{IJ}^N},$$

where the coarse-grained vorticity is defined as

$$\overline{\omega_{IJ}^N} = \frac{1}{n^2}\sum_{i'=I-n/2+1}^{I+n/2}\;\sum_{j'=J-n/2+1}^{J+n/2}\omega_{i'j'}^N.$$

We note that, over the macrostate p_N, the coarse-grained vorticity depends on p_N only:

$$\overline{\omega_{IJ}^N} = \sum_{k=1}^{K}p_{IJ}^k\sigma_k \;\text{ for }\; \omega^N \in p_N.$$

Using similar arguments, it is easy to conclude that in the limit $N \gg n \gg 1$ the energy of any microstate of the macrostate p_N is well approximated by the energy of the coarse-grained vorticity

$$\mathcal{E}_N\left[\omega^N\right] \overset{N\gg n\gg 1}{\sim} \mathcal{E}_N\left[\overline{\omega_{IJ}^N}\right] = -\frac{n^2}{2N^2}\sum_{I,J=0}^{N/n-1}\overline{\omega_{IJ}^N}\psi_{IJ}^N.$$

Then the Boltzmann entropy of the macrostate is

$$S_N[p_N,E_0] \overset{N\gg n\gg 1}{\sim} \mathcal{S}_N[p_N] \tag{14}$$

if $\forall k,\ \mathcal{N}[p_N^k] = 1$, $A_N\left[p_N^k\right] = A_k$ and $\mathcal{E}_N\left[\overline{\omega_{IJ}^N}\right] = E_0$, and $S_N[p_N,E_0] \sim -\infty$ otherwise.

Consider $P_{N,E_0}(p_N)$ to be the probability density to observe the macrostate p_N in the finite-N microcanonical ensemble with energy E_0. By definition of the microcanonical ensemble of the entropy $S_N(E_0)$ (see Eq. (11) and the preceding paragraph), we have

$$\log P_{N,E_0}(p_N) \overset{N\to\infty}{\sim} N^2 \left[S_N[p_N, E_0] - S_N(E_0)\right]. \tag{15}$$

From the general definition of a large deviation result given by Eq. (1), we clearly see that formula (14) is a large deviation result for the macrostate p_N in the microcanonical ensemble. The large deviation parameter is N^2 and the large deviation rate function is $-S_N[p_N, E_0] + S_N(E_0)$.

We now consider the continuous limit. The macrostates p_N^k are now seen as the finite-N approximation of p_k, the local probability to observe $\omega(\mathbf{r}) = \sigma_k$: $p_k(\mathbf{r}) = \langle \delta(\omega(\mathbf{r}) - \sigma_k)\rangle$. The macrostate is then characterized by $p = \{p_1, \ldots, p_K\}$. Taking the limit $N \gg n \gg 1$ allows us to define the entropy of the macrostate (p, E_0) as

$$S[p, E_0] = \mathcal{S}[p] \equiv \sum_k \int_{\mathcal{D}} p_k \mathrm{log} p_k \, \mathrm{d}\mathbf{r} \tag{16}$$

if $\forall k$ $\mathcal{N}[p_k] = 1$, $A\,[p_k] = A_k$ and $\mathcal{E}[\overline{\omega}] = E_0$, and $S[p, E_0] = -\infty$ otherwise. In the same limit, it is clearly seen from definition (13) and result (16) that there is a concentration of microstates close to the most probable macrostate. The exponential concentration close to this most probable state is a large deviation result, where the entropy appears as the opposite of a large deviation rate function (up to an irrelevant constant).

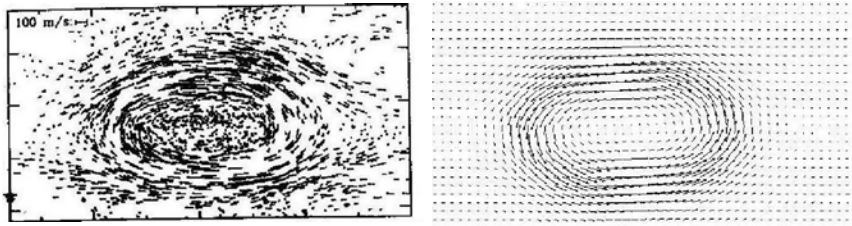
The exponential convergence towards this most probable state also justifies the approximation of the entropy with the entropy of the most probable macrostate. Thus, in the limit $N \to \infty$ we can express the Boltzmann entropy, Eq. (12), as

$$S(E_0) = \sup_{\{p \,\mid\, \mathcal{N}[p]=1\}} \left\{\mathcal{S}[p] \;\mid \mathcal{E}[\bar{\omega}] = E_0, \forall k \; A[p_k] = A_k\right\}, \tag{17}$$

where $p = \{p_1, \ldots, p_K\}$ and $\forall$ $\mathbf{r}$, $\mathcal{N}[p](\mathbf{r}) = \sum_{k=1}^{K} p_k(\mathbf{r}) = 1$ is the local normalization. Furthermore, $A[p^k]$ is the area of the domain corresponding to the vorticity value $\omega = \sigma_k$. The fact that the Boltzmann entropy $S(E_0)$ Eq. (12) can be computed from the variational problem (17) is a powerful non-trivial result of large deviation theory.

3.4 Applications of equilibrium statistical mechanics

In the two previous sections, we have defined the microcanonical measure for the two-dimensional Euler and quasi geostrophic equations, and we have proven that the logarithm of the probability of a macrostate p is given by the macrostate entropy (16). We can conclude that most of the microstates will correspond to the most probable macrostate, the one that actually maximize the variational problem (17). This most probable macrostate is called the equilibrium macrostate. This means that if we take a random microstate, it will nearly surely have the same



Observation (Voyager) Statistical Equilibrium

Fig. 5: Left: the observed velocity field is from Voyager spacecraft data, from Dowling and Ingersoll [25] ; the length of each line is proportional to the velocity at that point. Note the strong jet structure of width of order R, the Rossby deformation radius. Right: the velocity field for the statistical equilibrium model of the Great Red Spot. The actual values of the jet maximum velocity, jet width, vortex width and length fit with the observed ones. The jet is interpreted as the interface between two phases; each of them corresponds to a different mixing level of the potential vorticity. The jet shape obeys a minimal length variational problem (an isoperimetrical problem) balanced by the effect of the deep layer shear.

velocity as the one of the equilibrium macrostate. As a consequence, we conclude that equilibrium macrostates are natural candidates to model self organized large scale turbulent flows, like for instance the Great Red Spot of Jupiter shown on figure (1).

A number of works have considered the comparison of self-organized turbulent flows with equilibrium macrostates. Interested readers will find comparison with experiments and numerical simulations described in the review [64], whereas models of geophysical flows, for instance the Great Red Spot of Jupiter, ocean mesoscale vortices, strong mid basin jets similar to the Gulf Stream or the Kuroshio are discussed in the review [16]. Recent applications to model the vertical structure of oceans can be found in the papers [70; 71].

As an example, figure 5 shows the comparison of the observed velocity field for the Great Red Spot of Jupiter with the velocity field of an equilibrium macrostate of the quasi-geostrophic model. The theoretical analysis of this equilibrium macrostate [11] is based on an analogy with Van Der Walls–Cahn–Hilliard model of first order transition and the shape of the strong jet obeys a minimal length variational problem (an isoperimetrical problem) balanced by the effect of the deep layer shear (see [16] for more details).

Another example of equilibrium prediction is the phase diagram of statistical equilibria for the two-dimensional Euler equation on a doubly periodic domain (torus). This phase diagram (figure 6) shows that the statistical equilibria are either dipoles (one cyclone and one anticyclone) or parallel flows. This example is further discussed in the work [10] and the review [16]. This equilibrium phase diagram has also been used in order to predict non-equilibrium phase transitions [10] as is discussed in section 4.4.

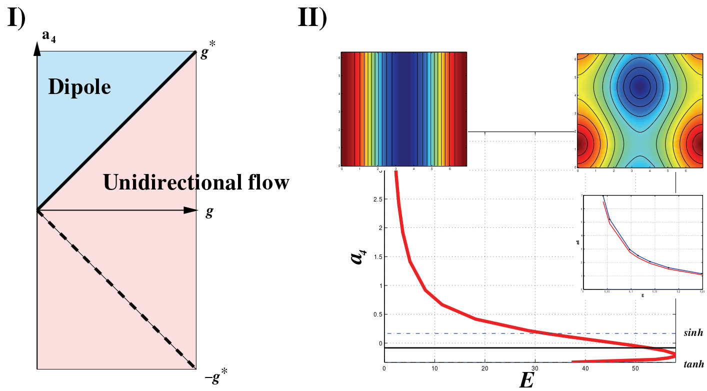


Fig. 6: Bifurcation diagrams for statistical equilibria of the two-dimensional Euler equations in a doubly periodic domain a) in the g-a_4 plane, g is related to the domain aspect ratio and a_4 to the fourth order moment of the vorticity distribution (please see [16]). b) obtained numerically in the $E - a_4$ plane, E is the energy, in the case of doubly periodic geometry with aspect ratio $\delta = 1.1$. The colored insets are streamfunction and the inset curve illustrates good agreement between numerical and theoretical results in the low energy limit.

4 Non equilibrium phase transitions, path integrals, and instanton theory

The aim of this section is to discuss non-equilibrium phase transitions in turbulent flows, more specifically for the dynamics of the two-dimensional Navier–Stokes equations with random forces, quasi-geostrophic dynamics with random forces, or related dynamics. We want to discuss simple examples for which situations with rare transitions between two attractors exist (bistability). We will use path integrals and large deviations in order to compute the most-probable paths for those transitions and the transition rates.

In order to give a pedagogical presentation of path integrals and large deviation theory for stochastic dynamics we first discuss the extremely classical case of the Kramer problem: the over-damped dynamics of a particle in a double-well potential, in section 4.1. We generalize these results to an abstract set of dynamics, called Langevin dynamics, in section 4.2. We apply these results to two-dimensional Euler and Quasi-Geostrophic Langevin dynamics in section 4.3, for which we are able to predict bistability, compute transition rates and the most probable transition paths. Finally we discuss path integral approaches and action minimizer for the stochastic Navier-Stokes equations in a non-equilibrium context in section 4.4.

4.1 Large deviations for the overdamped Langevin dynamics

We wish first to give a pedagogical description of large deviation theory in ***non-equilibrium systems***, more specifically for dynamics consisting of stochastic differential equations. Therefore, we begin by applying large deviation theory to a simple academic example of an over-damped particle in a double-well potential (the Kramer problem) where a large deviation result exists. We will show that we can compute the transition rate for the motion of the particle from one well to the other and that the result is an Arrhenius factor (it is proportional to the exponential of the energy barrier height between the two wells). In fact, this is a large deviation result.

This section develops classical ideas. We use the path integral formalism for stochastic processes [52; 74]. Similar results are discussed by mathematicians in the framework of the Freidlin-Wentzell theory [30; 67]. We are much interested by the time-reversal symmetries of the action and its consequence for the symmetry between relaxation and fluctuation paths, and its consequences for the computation of the most probable transition (instanton). Those symmetries are discuss much less often than the other material, but there are also very classical (some people say it dates from Onsager, we do not know exactly).

4.1.1 The overdamped Langevin dynamics

We consider a single overdamped particle in a 1D double-well potential $V(x)$ and subjected to random forces due to a small coupling to a thermal bath. For simplicity we considered the overdamped limit, for which the dynamics of the particle position x is governed by the stochastic differential equation

$$\dot{x} = -\frac{\mathrm{d}V}{\mathrm{d}x} + \sqrt{\frac{2}{\beta}}\eta, \tag{18}$$

where η is a random white noise with a Gaussian distribution characterized by $\mathbf{E}\left[\eta(t)\eta(t')\right] = \delta(t-t')$, $V(x)$ is a double well potential (see Fig. 7), and $\beta = 1/k_B T$ where T is the temperature. In the deterministic situation, when $1/\beta = 0$, the particle relaxes to one of the two stable steady states of the potential V, i.e. it converges either to $x = -1$ or to $x = 1$. In the presence of thermal noise, the particle may gain enough energy to jump the potential barrier at $x = 0$ and settle in the other potential well. If the forcing is weak, i.e. $1 \ll \beta \Delta V$, then the jumps between wells will be rare events and will be statistically independent from one another. They will then be described by a Poisson process characterized by a transition rate λ. We will show that one can apply the theory of large deviations in order to compute λ. Moreover the theory of large deviation will lead to the conclusion that most of the transition paths concentrate close to the most probable transition path. As will be discussed more precisely below, this most probable transition path in this situation is called an instanton.

In order to obtain these results, we will use formal computations based on a path integral formulation of the transition probabilities for the stochastic process

(18). Such a path integral formulation is referred as Onsager–Machlup formalism, as Onsager and Machlup first proposed it, few years after the path integral formulation of quantum mechanics by Feynman.

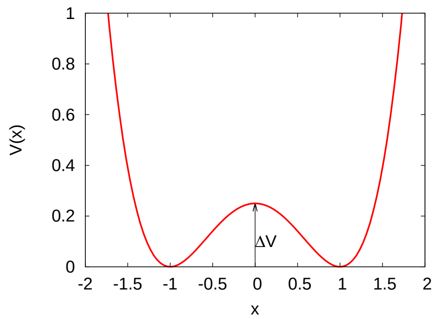


Fig. 7: Graph of the double well potential $V(x) = (x^2 - 1)^2/4$. We observe two stable steady states at $x = \pm 1$ and a saddle at $x = 0$ with height $\Delta V = 1/4$.

4.1.2 The transition probability as a path integral

To give a simple understanding of the Onsager-Machlup formalism, we first consider a vector $\eta = \{\eta_i\}_{1 \le i \le N}$ of independent Gaussian random variables, with zero mean $\mathbf{E}(\eta_i) = 0$ and covariance $\mathbf{E}(\eta_i \eta_j) = \delta_{ij}$. By definition, the probability measure of η is the Gaussian measure

$$\mathrm{d}\mu = \exp\left(-\frac{1}{2}\sum_{i=1}^{N} \eta_i^2\right) \prod_{i=1}^{N} \frac{\mathrm{d}\eta_i}{\sqrt{2\pi}}. \tag{19}$$

The Euler approximation of the Langevin equation (18) is, within the Ito convention,

$$x_i = x_{i-1} - \Delta t \, \frac{\mathrm{d}V}{\mathrm{d}x}(x_{i-1}) + \sqrt{\frac{2\Delta t}{\beta}}\eta_i \tag{20}$$

for $1 \le i \le N$ and with $x_0 = x(0)$ a given initial state. The probability measure of a particular path $x = \{x_i\}_{1 \le i \le N}$ is given by inverting (20) and inserting it in (19),

$$\mathrm{d}\mu = \exp\left(-\frac{\beta}{4}\sum_{i=1}^{N}\left(\frac{x_i - x_{i-1}}{\Delta t} + \frac{\mathrm{d}V}{\mathrm{d}x}(x_{i-1})\right)^2 \Delta t\right) J(\eta|x) \prod_{i=1}^{N} \frac{\mathrm{d}x_i}{\sqrt{2\pi}}. \tag{21}$$

In this expression, $J(\eta|x)$ is the Jacobian of the change of variable $\eta \to x$. In the Ito convention (20), the corresponding matrix is lower-triangular with ones in the diagonal, so that $J(\eta|x) = 1$.

The measure of a Gaussian stochastic process $\eta(t)$ of zero mean $\mathbf{E}[\eta(t)] = 0$ and covariance $\mathbf{E}[\eta(t)\eta(t')] = \delta(t-t')$, on a time interval $[0,T]$ with $T = N\Delta t$, is the formal generalization of the above finite dimensional measure (19),

$$\mathrm{d}\mu = \exp\left(-\frac{1}{2}\int_0^T \eta^2(t)\mathrm{d}t\right)\mathcal{D}[\eta]. \tag{22}$$

The differential element $\mathcal{D}[\eta]$ in the above expression is the formal limit of the finite-dimensional quantity $\prod_{i=1}^N \frac{\mathrm{d}\eta_i}{\sqrt{2\pi}}$ for $N \to \infty$, $\Delta t \to 0$, where $\eta_i = \eta(i\Delta t) = \eta(iT/N)$. People well trained in mathematics know the difficulty to define such an object, but we will keep our discussion at a formal level and state that this formal notation contains all the mathematical subtleties related to the limit $N \to \infty$, $\Delta t \to 0$. Then, the probability measure of a particular trajectory $\{x(t)\}_{0\le t\le T}$ is also the formal limit of (21),

$$\mathrm{d}\mu = \exp\left(-\frac{\beta}{4}\int_0^T \left(\dot{x} + \frac{\mathrm{d}V}{\mathrm{d}x}\right)^2 \mathrm{d}t\right) J[\eta|x]\mathcal{D}[x], \tag{23}$$

where $J[\eta|x]$ is the Jacobian of the change of variable $\eta \to x$, and is also equal to one (we refer to [74] for a more general treatment, noting that [74] actually use the Stratonovich convention).

The transition probability from an initial state x_0 at time 0 to a final state x_T at time T is the sum over all possible paths $\{x(t)\}_{0\le t\le T}$ such that $x(0) = x_0$ and $x(T) = x_T$ of the probability of a single path (23). Such a sum can be formally written as the path integral

$$P(x_T, T; x_0, 0) = \int_{x(0)=x_0}^{x(T)=x_T} \exp\left(-\frac{\beta}{2}\mathcal{A}[x]\right)\mathcal{D}[x], \tag{24}$$

with the action functional

$$\mathcal{A}[x] = \frac{1}{2}\int_0^T \left(\dot{x} + \frac{\mathrm{d}V}{\mathrm{d}x}\right)^2 \mathrm{d}t. \tag{25}$$

From (24), it is clear that the most probable trajectories with prescribed initial and final states are minimizers of the action with prescribed initial and final point. The optimal action is denoted

$$A(x_0, x_T, T) = \min\left\{\mathcal{A}[x] \mid x(0) = x_0,\ x(T) = x_T\right\}.$$

4.1.3 Fluctuation paths

When the initial point $x_0 = x_a$ belongs to an attractor of the deterministic dynamics (for the Kramer problem, if $x_0 = x_a = \pm 1$ is a stable fixed points), it is expected that the action $A(x_a, X, T)$ decreases with time. The action minima starting from one attractor and having an infinite duration will thus play an important role. Moreover, those infinite time action minimizers are essential because the transition probability $P(X,T;x_a,0)$ converges to the stationary distribution of the stochastic process when the time T goes to infinity. Those action minimizers starting from one attractor and with an infinite duration are called fluctuation paths, they solve

$$A(x_a,X,\infty)=\min\left\{\mathcal{A}[x]\,|\lim_{T\to\infty}x(-T)=x_a,\,x(0)=X\right\}.$$

4.1.4 Relaxation paths

We consider a state X that belongs to the basin of attraction of an attractor x_a of the deterministic dynamics. The relaxation path starting at x, denoted $\{x_r(t)\}_{0\leq t\leq T}$ is defined by

$$\dot{x}_r=-\frac{\mathrm{d}V}{\mathrm{d}x}(x_r)$$

with initial conditions $x_r(0)=X$. As the path converges to x_a, we have $x_r(+\infty)=x_a$. Using the expression of the action (25), we see that $\mathcal{A}[x_r]=0$, as the relaxation path is a deterministic solution, and we also notice that $\mathcal{A}[x]\geq 0$ for any path $\{x(t)\}_{0\leq t\leq T}$. As a consequence, relaxation paths are global minimizers of the action $\mathcal{A}[x]$. This is because following the deterministic dynamics x_r in order to reach the attractor x_a starting from X doesn't require any stochastic perturbation, so that the cost is zero and the probability is maximal.

4.1.5 Time-reversal symmetry and the relation between fluctuation and relaxation paths

In order to characterize fluctuations paths and instantons, we will take profit of the time-reversal symmetry of the over-damped Langevin dynamics. We consider a path $\{x(t)\}_{0\leq t\leq T}$ and the reversed path $R[x]=\{x(T-t)\}_{0\leq t\leq T}$. The action of the reversed path reads

$$\mathcal{A}[R[x]]=\frac{1}{2}\int_0^T\left(\frac{\mathrm{d}}{\mathrm{d}t}R[x]+\frac{\mathrm{d}V}{\mathrm{d}x}(R[x])\right)^2\mathrm{d}t=\frac{1}{2}\int_0^T\left(-\dot{x}(t')+\frac{\mathrm{d}V}{\mathrm{d}x}(x(t'))\right)^2\mathrm{d}t',$$

with the change of variable $t'=T-t$. Then, writing

$$\left(\dot{x}-\frac{\mathrm{d}V}{\mathrm{d}x}\right)^2=\left(\dot{x}+\frac{\mathrm{d}V}{\mathrm{d}x}\right)^2-4\dot{x}\frac{\mathrm{d}V}{\mathrm{d}x}=\left(\dot{x}+\frac{\mathrm{d}V}{\mathrm{d}x}\right)^2-4\frac{\mathrm{d}}{\mathrm{d}t}V(x),$$

we get

$$\mathcal{A}[R[x]]=\mathcal{A}[x]-2\left(V(x(T))-V(x(0))\right). \tag{26}$$

Plugging this relation into the path integral expression of the transition probability (24), we obtain

$$P(R[x_T],T;R[x_0],0) = P(x_T,T;x_0,0)\exp\left(\frac{V(x(T))-V(x(0))}{k_BT}\right).$$

We recognize the Gibbs stationary distribution of the over-damped Langevin equation $P_S(x) = \frac{1}{Z}\mathrm{e}^{-V(x)/k_BT}$, so that the above expression gives the detailed balance relation

$$P(x_T,T;x_0,0)P_S(x_0) = P(x_0,T;x_T,0)P_S(x_T).$$

We have thus proven that detailed balance is a consequence of the time-reversal symmetry, as expected on general ground.

We now consider the fluctuation path from one attractor x_a to any point X of its basin of attraction. Using relation (26) and the fact that the action is always positive, we have

$$\mathcal{A}[x] \geq 2\left(V(x(T)) - V(x(0))\right), \tag{27}$$

with equality if and only if x is a minimizer of the reversed action $\mathcal{A}[R[x]]$. If the initial state is an attractor and the final state is another point in the associated basin of attraction, the reversed action $\mathcal{A}[R[x]]$ is naturally minimized by the relaxation path $R[x] = x_r$ that goes from X to the attractor,

$$\frac{\mathrm{d}}{\mathrm{d}t}R[x] = -\frac{\mathrm{d}V}{\mathrm{d}x}(R[x])$$

with $R[x](0) = X$ and $R[x](+\infty) = x_a$. Then the minimizer of $\mathcal{A}[x]$ is the reversed relaxation path. We thus conclude that the fluctuation path from x_a to X, is the time reversed of the relaxation path from X to x_a. This situation is schematically represented in figure 8.

4.1.6 Instanton and large deviation principle

We define the instanton as the most probable path that go from one attractor $x_{-1} = -1$ to the other one $x_1 = 1$ in an infinite time. More precisely we consider x_T the minimizer of variational problem $\min\left\{\mathcal{A}[x] \mid x\left(-\frac{T}{2}\right) = x_{-1},\ \ x\left(\frac{T}{2}\right) = x_1\right\}$, and the instanton is the limit when $T \to \infty$ of x_T. The instanton action is

$$A(x_{-1},x_1) = \lim_{T\to\infty}\min\left\{\mathcal{A}[x] \mid x\left(-\frac{T}{2}\right) = x_{-1} \text{ and } x\left(\frac{T}{2}\right) = x_1\right\},$$

As will soon become clear, instantons are related to the most probable transition paths, and their action to the transition rate λ.

From the previous discussion, it is easily understood that instantons are decomposed into two parts. First, there is the fluctuation path from x_{-1} to the saddle $x_s = 0$, which is the reverse of the relaxation path from x_s to x_{-1}. The action of this part of the trajectory is given by (26), it is $\mathcal{A}[R[x_r]] = 2\left(V(x_s) - V(x_{-1})\right) = 2\Delta V$, where ΔV is the potential barrier height. The second part of the instanton trajectory is the relaxation path from the saddle x_s to the final attractor x_1. The action of this relaxation path is zero, so that the total instanton action is $\mathcal{A}[x^*] = 2\Delta V$.

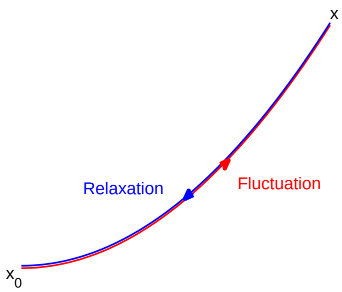


Fig. 8: Schematic representation of the fluctuation and relaxation paths between an attractor of the deterministic dynamics x_0 and another point X in the basin of attraction of x_0, for the over-damped dynamics. The relaxation path is the deterministic trajectory from x to x_0, and the fluctuation path is the time-reversed trajectory. Both trajectories are the most probable paths with the associated initial and final states.

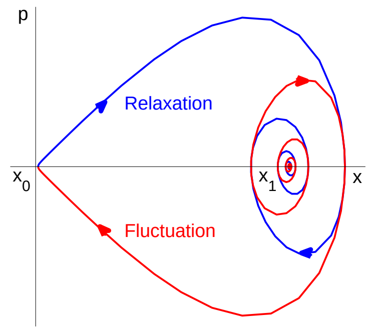


Fig. 9: Fluctuation and relaxation paths between an initial position x_0 and an attractor x_1, for the full Langevin dynamics 30. The fluctuation path (reversed relaxation path) is obtained by reversal of time, so the momentum is changed as $p \to -p$. Both trajectories are the most probable paths with the associated initial and final conditions.

A more precise analysis shows that as both the fluctuation path to the saddle and the relaxation path last for an infinite time (an infinite time is needed to quit the attractor and an infinite time is needed to reach the saddle). This explains the definition of the instanton through the limit of the finite time minimizer x_T. One can also understand that any temporal translation of an instanton is another minimizer from on attractor $x_{-1} = -1$ to the other one $x_1 = 1$ in an infinite time. This degeneracy is related to the notion of a "free-instanton-molecule" gas approximation and has the consequence that for time $T \gg 1$, the transition probability is proportional to time T:

$$P(x_1, T; x_{-1}, 0) \overset{T \gg 1}{\gtrsim} \lambda T.$$

We refer to [19] for a detailed discussion.

In the limit of small forcing $1 \ll \beta \Delta V$, the distribution given by the path integral (24) is concentrated around its most probable state, the instanton we have determined. We can thus apply a saddle-point approximation in order to get the transition probability P_t,

$$\lim_{\beta \to \infty} -\frac{1}{\beta} \log(P_t) = \Delta V. \tag{28}$$

Formula (28) states that the transition probability for observing the rare transition between the two potential wells, in the limit of the weak noise limit, is proportional to the exponential of the barrier height ΔV. Such a result is called a large deviation principle for the probability P_t. We recover the exponential factor of the Arrhenius formula for the transition rate

$$\lambda = \frac{1}{\tau} \exp\left(-\frac{\Delta V}{k_B T}\right) \tag{29}$$

where ΔV is the energy barrier height and $k_B T$ is the temperature.

The computation of the prefactor $1/\tau$ goes beyond a large deviation result. It was already computed by Kramer, for an overdamped Kramer dynamics. It was the subject of Langer theory for systems with many degrees of freedom. Alternatively, it can be computing in the path integral framework by computing the path integrals at next order, computing the properties of the Gaussian processes close to the instanton, and treating correctly the subtleties related to the instanton degeneracy due to time translation. Such a computation can be found for example in the reference [19]. The result is

$$\tau = 2\pi \left(\frac{d^2V}{dx^2}(x_0)\frac{d^2V}{dx^2}(x_{-1})\right)^{-1/2}.$$

4.1.7 Generalization to the inertial Langevin dynamics

We consider now the dynamics of a particle in the same double-well potential, with random forces, but without the over-damped approximation. The position and momentum of the particle $\{x, p\}$ satisfy

$$\begin{cases} \dot{x} = p \\ \dot{p} = -\frac{\mathrm{d}V}{\mathrm{d}x} - \alpha p + \sqrt{\frac{2\alpha}{\beta}}\eta. \end{cases} \tag{30}$$

In this case, the time-reverse of a given path $\{x(t), p(t)\}_{0 \leq t \leq T}$ is given by $I\left[x, p\right] = \{x(T-t), -p(T-t)\}_{0 \leq t \leq T}$, as represented in figure 9. It is easily proven that the action of the reversed action path satisfies a relation similar to 26. Then, as in the overdamped case, one easily proves that the fluctuation paths is the time reverse of the relaxation paths. As in the over-damped case, instantons from one attractor to the other are composed of a fluctuation path (time reversed relaxation path) from the first attractor $\{x_{-1} = -1, p_{-1} = 0\}$ to the saddle $\{0, 0\}$, and a relaxation path from the saddle to the final attractor $\{1, 0\}$.

4.2 Langevin dynamics with potential $\mathcal{G}$

The aim of this section is to generalize the results discussed for the Kramer model in section 4.1 to a class of dynamics that corresponds to systems coupled with equilibrium (thermal) baths. The consider dynamics with Liouville theorem (for instance Hamiltonian dynamics), with dissipation which are the gradient of a conserved quantity and stochastic forces with Einstein type relations. For those

Langevin dynamics, we prove detailed balance (sometimes in a generalized form), we prove that the fluctuation paths are the time reversed of the relaxation paths, and we describe the instantons.

Whereas such Langevin dynamics are very common in physics, the discussion below is original. As far as we know we are the first to describe this general framework, especially for the case when the potential is not the Hamiltonian but another conserved quantity. The aim is to apply this framework to dynamics that include the two-dimensional Euler and quasi-gestrophic dynamics.

4.2.1 Definition of Langevin dynamics

In this section we consider the deterministic dynamics

$$\frac{\partial q}{\partial t} = \mathcal{F}\left[q\right] \tag{31}$$

where q is either a finite dimensional variable or a field.

If $q \in \mathbf{R}^N$, the dynamics is $\frac{\partial q_i}{\partial t} = \mathcal{F}_i\left[q\right]$. We then assume that this dynamical system conserves the Liouville measure $\prod_{i=1}^N \mathrm{d}q_i$, or equivalently that the divergence of the vector field $\mathcal{F}$ is zero

$$\nabla.\mathcal{F} \equiv \sum_{i=1}^{N} \frac{\partial \mathcal{F}_i}{\partial q_i} = 0.$$

We call this property a Liouville theorem.

If q is a field (for instance a two-dimensional vorticity or potential vorticity field), defined over a domain $\mathcal{D}$, $\mathcal{F}\left[q\right](\mathbf{r})$ is a quantity computed from the field q at any point $\mathbf{r}$. For instance for the Quasi-Geostrophic equation $\mathcal{F}\left[q\right] = -\mathbf{v}\left[q-h\right] \cdot \nabla q\left(\mathbf{r}\right)$. We continue the discussion for a field equation only. For any functional $\mathcal{K}$, $\frac{\delta \mathcal{K}}{\delta q(\mathbf{r})}$ is the functional derivative of $\mathcal{K}$ at point $\mathbf{r}$, a generalization of the usual derivative, such that for any variation δq, at linear order the first variations of $\mathcal{K}$ are given by

$$\delta \mathcal{K} = \int_{\mathcal{D}} \frac{\delta \mathcal{K}}{\delta q(\mathbf{r})} \delta q\left(\mathbf{r}\right)\, \mathrm{d}\mathbf{r}.$$

We assume that a Liouville theorem holds for the dynamics (31), in the sense that the formal generalization of the finite dimensional Liouville theorem

$$\nabla.\mathcal{F} \equiv \int_{\mathcal{D}} \frac{\delta \mathcal{F}}{\delta q(\mathbf{r})}\left(\mathbf{r}\right)\, \mathrm{d}\mathbf{r} = 0,$$

is verified.

We also assume that this dynamical system has a conserved quantity $\mathcal{G}$: $\mathrm{d}\mathcal{G}/\mathrm{dt} = 0$. From (31), we see that this is equivalent to

$$\int_{\mathcal{D}} \mathcal{F}\left[q\right](\mathbf{r}) \frac{\delta \mathcal{G}}{\delta q(\mathbf{r})}\left[q\right] \mathrm{d}\mathbf{r} = 0, \tag{32}$$

for any q. Those hypothesis are verified, for instance if the dynamical system is an Hamiltonian system

$$\mathcal{F}(q) = \left\{q, \mathcal{H}\right\},$$

where $\{.,.\}$ is a Poisson bracket, and $\mathcal{G}$ one of the conserved quantity of the Hamiltonian system, for instance $\mathcal{G} = \mathcal{H}$. We stress however that $\mathcal{G}$ does not need to be $\mathcal{H}$.

If the Liouville hypothesis is verified and $\mathcal{G}$ is a conserved quantity, we call a Langevin dynamics for the potential $\mathcal{G}$ the stochastic dynamics

$$\frac{\partial q}{\partial t} = \mathcal{F}\left[q\right](\mathbf{r}) - \alpha \int_{\mathcal{D}} C(\mathbf{r}, \mathbf{r}') \frac{\delta \mathcal{G}}{\delta q(\mathbf{r}')}\left[q\right] \mathrm{d}\mathbf{r}' + \sqrt{2\alpha\gamma}\eta, \tag{33}$$

where we have introduced a stochastic force η, which we assume to be a Gaussian process, white in time, and correlated as $\mathbf{E}\left[\eta(\mathbf{r},t)\eta(\mathbf{r}',t')\right] = C(\mathbf{r},\mathbf{r}')\delta(t-t')$. As it is a correlation function, C has to be a symmetric positive function: for any function ϕ over $\mathcal{D}$

$$\int_{\mathcal{D}} \int_{\mathcal{D}} \phi\left(\mathbf{r}\right) C(\mathbf{r}, \mathbf{r}') \phi\left(\mathbf{r}'\right) \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \geq 0, \tag{34}$$

and $C(\mathbf{r},\mathbf{r}') = C(\mathbf{r}',\mathbf{r})$. For simplicity, we assume in the following that C is positive definite and has an inverse C^{-1} such that

$$\int_{\mathcal{D}} C(\mathbf{r}, \mathbf{r}_1) C^{-1}(\mathbf{r}_1, \mathbf{r}') \mathrm{d}\mathbf{r}_1 = \delta\left(\mathbf{r} - \mathbf{r}'\right).$$

The major property of a Langevin dynamics is that the stationary probability density functional is known a-priori. It is

$$P_s[q] = \frac{1}{Z} \exp\left(-\frac{\mathcal{G}[q]}{\gamma}\right),$$

where Z is a normalization constant. At a formal level, this can be checked easily by writing the Fokker-Planck equation for the evolution of the probability functionals. Then the fact that P_s is stationary readily follows from the Liouville theorem and the property that $\mathcal{G}$ is a conserved quantity for the deterministic dynamics.

4.2.2 Reversed Langevin dynamics

We consider I a linear involution on the space of fields q (I is a linear functional with $I^2 = \mathrm{Id}$). We define the reversed Langevin dynamics with respect to I as

$$\frac{\partial q}{\partial t} = \mathcal{F}_r\left[q\right](\mathbf{r}) - \alpha \int_{\mathcal{D}} C_r(\mathbf{r}, \mathbf{r}') \frac{\delta \mathcal{G}_r}{\delta q(\mathbf{r}')}\left[q\right] \mathrm{d}\mathbf{r}' + \sqrt{2\alpha\gamma}\eta, \tag{35}$$