Statistical Mechanics and Numerical Modelling of Geophysical Fluid Dynamics

Svetlana Dubinkina

Statistical Mechanics and Numerical Modelling of Geophysical Fluid Dynamics

Svetlana Dubinkina

Statistical Mechanics and Numerical Modelling of Geophysical Fluid Dynamics

Svetlana Borisovna Dubinkina

Printed and bound by Ipskamp Drukkers.

CIP-DATA LIBRARY UNIVERSITEIT VAN AMSTERDAM

Svetlana Borisovna Dubinkina

Statistical Mechanics and Numerical Modelling of Geophysical Fluid Dynamics/
by Svetlana Borisovna Dubinkina.– Amsterdam: Universiteit van Amsterdam,
2010.– Proefschrift.

ISBN 978-90-9025310-7

Statistical Mechanics and Numerical Modelling of Geophysical Fluid Dynamics

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door
het college voor promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op vrijdag 28 mei 2010, te 10:00 uur

door

Svetlana Borisovna Dubinkina

geboren te Bratsk, Rusland

Promotiecommissie

Promotor:	prof. dr. ir. J.E. Frank
Co-promotor:	prof. dr. J.G. Verwer
Overige leden:	prof. dr. R.P. Stevenson
	prof. dr. C.A.J. Klaassen
	prof. dr. P.W. Hemker
	prof. dr. B. Leimkuhler
	prof. dr. S. Reich
	prof. dr. W. Hundsdorfer

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

NWO
Nederlandse Organisatie voor Wetenschappelijk Onderzoek

CWI
Centrum Wiskunde & Informatica

Thomas Stieltjes Institute
for Mathematics

Het onderzoek dat tot dit proefschrift heeft geleid werd mede mogelijk gemaakt door het programma Klimaatvariabiliteit van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek, gebied Aard- en Levenswetenschappen, onder toekenning 854.00.008, en het Centrum Wiskunde & Informatica.

To my parents

Contents

Preface **iii**

1 Introduction **1**
1.1 Hamiltonian systems and geometric integration 2
1.1.1 Hamiltonian systems 2
1.1.2 Geometric integration 6
1.2 Hamiltonian fluid dynamics 8
1.2.1 Hamiltonian PDEs 8
1.2.2 Eulerian and Lagrangian descriptions 10
1.2.3 Numerical methods 12
1.3 Statistical mechanics of fluids 14
1.3.1 Statistical ensembles 16
1.3.2 Information theory 20
1.3.3 Statistical theories for quasigeostrophic flow 22

2 Statistical mechanics of Arakawa's discretizations **27**
2.1 Introduction 27
2.2 The quasigeostrophic model 28
2.3 Spatial semi-discretization 29
2.3.1 Arakawa's discretizations 30
2.3.2 Volume preservation 33
2.4 Energy-enstrophy statistical theory 34
2.4.1 Mean field predictions 35
2.4.2 PV fluctuation predictions 36
2.4.3 Approximation of μ and α 37
2.4.4 Alternative statistical theories 38
2.5 Time integration 39
2.6 Numerical experiments 40
2.7 Conclusions 47

3 Statistical mechanics of the Hamiltonian particle-mesh method **53**
3.1 Introduction 53
3.2 Review of continuum statistical equlibrium theories 55
3.3 Hamiltonian particle-mesh method 57
3.3.1 HPM description 57
3.3.2 Properties of the discretization 59
3.4 A Lagrangian statistical model based on canonical particle distributions 63

3.5 Eulerian statistical model for HPM 65
3.6 Numerical Verification of the HPM Statistical Equilibrium Theories 68
3.6.1 Normally distributed PV 70
3.6.2 Skew PV distributions 70
3.6.3 PV distributions with kurtosis 73
3.7 Conclusions 75

4 A thermostat closure for point vortices 77
4.1 Background 77
4.2 Generalized thermostats 78
4.2.1 Langevin thermostat 80
4.2.2 A generalized Bulgac-Kusnezov method 82
4.3 Statistical mechanics of point vortices 83
4.4 A thermostated integrator for point vortices 86
4.4.1 Infinite and finite reservoir ensembles 86
4.4.2 Choice of s_1 87
4.4.3 Implementation details 87
4.4.4 Computation of temperatures 88
4.5 Numerical experiments 90
4.5.1 Ergodicity tests 90
4.5.2 Momentum conservation 90
4.5.3 Temperature effects 92
4.6 Conclusions 96

Summary 103

Samenvatting 105

Acknowledgements 107

Preface

Statistical mechanics is a powerful approach to understanding complex physical systems. The purpose of statistical mechanics is to construct methods which can handle incompletely known systems; to describe the most likely behaviour of a system; to work with time series data from experiments or numerical models. In this thesis we use statistical mechanics as a tool to show the importance of conservation laws of a numerical method. Depending on what conserved quantities the discrete system has, the statistical theory of this discrete system varies. Therefore, the chosen method has an influence on the statistical results of the simulated model. In the thesis I show that statistical mechanics can be also employed as a tool by a numerical analyst to verify the statistical accuracy of a numerical method. This is very important issue for applications such as climate variability, since in these applications long numerical simulations are run for dynamical systems that are known to be chaotic, and for which it is consequently impossible to simulate a particular solution with any accuracy in the usual sense. Instead, the goal of such simulations is to obtain a data set suitable for computing statistical averages or otherwise to sample the probability distribution associated with the continuous problem. Different numerical discretizations have different discrete dynamics. Therefore it is crucial to establish the influence that a particular choice of method has on the statistical results obtained from simulations.

Another attractive application for statistical mechanics is the modeling of sub-scale motion. Inviscid fluid models are natural in a number of application areas, such as atmosphere and ocean science. These flows are characterized by conservation of energy, sensitive dependence on initial conditions, and the cascade of vorticity to ever finer scales. For the simulation of such flows numerically, the vorticity cascade presents the challenge that any direct discretization of the equation of motion must eventually become underresolved. To address this problem effectively requires modelling the sub-gridscale dynamics and their influence on the coarse scale. In the thesis, I show for a point vortex model that these effects can be parameterized using an adapted 'mathematical thermostat', a technique used in molecular simulations to model a system of particles interacting at a constant temperature. I believe that this methodology can be extended for more complex models with feasible applications in climate variability.

This project was funded in NWO Earth & Life Sciences Council "Climate Variability" program, and it address the problems arising in numerical simulations of geophysical fluid dynamics—statistical accuracy of methods, influence of conserved quantities of a method on statistics of the dynamics; and it proposes

a new technique for modelling the sub-scale motion, which can be employed in atmosphere and ocean science.

This thesis records the numerical mathematics research I conducted between August 2005 and January 2010 in the Modeling, Analysis and Simulation (MAS) department of the Centrum Wiskunde & Informatica (CWI) in Amsterdam.

Chapters 2 through 4 of this thesis have appeared as journal articles:

1. Chapter 2 is based on S. Dubinkina and J. Frank, *Statistical mechanics of Arakawa's discretizations*, Journal of Computational Physics 227, pages 1286–1305, 2007.

2. Chapter 3 is based on S. Dubinkina and J. Frank, *Statistical relevance of vorticity conservation with the Hamiltonian particle-mesh method*, accepted for publication in Journal of Computational Physics, 2010.

3. Chapter 4 is based on S. Dubinkina, J. Frank and B. Leimkuhler, *A thermostat closure for point vortices*, submitted, 2009.

Chapter 1, the introductory chapter collects much of the background material needed to read the rest of the thesis.

Svetlana Borisovna Dubinkina

Amsterdam, January 2010

Chapter 1

Introduction

In numerical simulations of climate variability, the flow of atmosphere and oceans is known to be chaotic, and therefore the quantity of particular interest is not a single solution trajectory, but rather an ensemble of possible motions of the system given certain information about the behavior of this system. Then statistical mechanics becomes useful, since it deals with the probability of a system being in some state under known constraints. To derive a statistical theory for a system, the system should be volume preserving and possess conserved quantities. Ergodicity is usually taken as an assumption to relate dynamics to statistical theory, and the validity of this assumption must be confirmed with numerical experiments. Although a continuous model has conservation laws, it is not clear that after implementing a numerical method, the corresponding discrete model will also have an equivalence of them. This entirely depends on properties of a numerical method. Therefore the statistical results of the discrete model may depend (and they actually do) on the numerical method used. And in an ample range of available numerical methods, it therefore becomes even more crucial and pronounced to understand how the numerical method influences the statistical mechanics of a simulation.

Arakawa's schemes for 2D incompressible flow are a convenient subject to investigate the importance of conservation because they satisfy one or both of a pair of conservation laws, and they are constructed from identical discrete operators, differing only in the order of application of these. If the operators would commute in the discrete case—as they do in the continuous—the schemes would be identical. Nonetheless we observe completely different equilibrium behavior for these methods, see Chapter 2, and for the introduction to the statistical mechanics see Section 1.3.

Symplectic or Poisson discretizations for Hamiltonian partial differential equations are constructed in such a way that the semi-discrete systems are again Hamiltonian and possess an equivalence of conserved quantities of the original continuous Hamiltonian systems. Because Hamiltonian structure plays an important role in statistics, we might expect these methods to perform better than standard methods in terms of statistical performance measures. This motivated us to derive in the thesis statistical theories for a Hamiltonian particle-mesh method in the case of quasigeostrophic flow with topographic forcing, see Chapter 3. For the introduction to the subject of Hamiltonian systems see

Sections 1.1 and 1.2.

Another attractive employment of statistical mechanics is the modelling of sub-gridscale closure of inviscid fluid equations, since such flows, which model essential components of atmosphere and ocean, support a vorticity cascade to ever finer scales. Any finite discretization of these equations must become underresolved, and we can use equilibrium statistical mechanics theory to represent the subgrid motions. In Chapter 4, we show for a point vortex flow on a disc that this can be parameterized by using a mathematical thermostat, a technique from molecular dynamics which allows to model a system in contact with a reservoir of fine scale vorticity. In Section 1.3, we explain statistical mechanics in general, concept of mathematical thermostat, and define statistical temperature.

1.1 Hamiltonian systems and geometric integration

The continuum models studied in this thesis have a common mathematical structure: that of Hamiltonian systems. Hamiltonian structure is also the point of departure for statistical mechanics. We therefore review Hamiltonian structure for ordinary and partial differential equations in this section, and we discuss geometric integrators—numerical methods that preserve Hamiltonian structure. A geometric integrator is a numerical method which preserves some structural properties of a given problem, e.g. symplecticity, first integrals, symmetries. Hamiltonian systems in turn are rich in such properties. Therefore this motivates interest in geometric integration of Hamiltonian systems.

1.1.1 Hamiltonian systems

In the following section we consider Hamilton's principle for finite-dimensional mechanical systems following the derivations described in [3, 42]. In the context of geophysical fluid dynamics see also [57, 73, 74, 82].

The canonical equation

Consider a mechanical system with d degrees of freedom. Its state can be specified by the generalized coordinates $\boldsymbol{q} = (q_1, \ldots, q_d)^T$ as functions of time t. The dynamics of the system are determined by the Lagrangian L, the difference between the kinetic and potential energies. The Lagrangian $L = L(\boldsymbol{q}, \dot{\boldsymbol{q}})$ is a function of the coordinates $\boldsymbol{q}$ and the velocities $\dot{\boldsymbol{q}}$, where the dots on $\boldsymbol{q}$ stand for differentiation with respect to time. The evolution of the system may be determined from Hamilton's principle [42]

$$\delta \int_0^{t_0} L \, dt = 0, \tag{1.1}$$

for arbitrary variations δq_k that vanish at $t = 0$ and $t = t_0$. From this variational problem (1.1) follows Lagrange's equations of the system

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\boldsymbol{q}}}\right) = \frac{\partial L}{\partial \boldsymbol{q}}. \tag{1.2}$$

Define the conjugate momenta as

$$p_k = \frac{\partial L}{\partial \dot{q}_k}(\boldsymbol{q}, \dot{\boldsymbol{q}}), \quad k = 1, \ldots, d \tag{1.3}$$

and the *Hamiltonian*

$$H \equiv \boldsymbol{p}^T \dot{\boldsymbol{q}} - L(\boldsymbol{q}, \dot{\boldsymbol{q}})$$

as a function of $\boldsymbol{q}$ and $\boldsymbol{p}$. This is the Legendre transformation of L with respect to $\dot{\boldsymbol{q}}$. Under the assumption that $\dot{\boldsymbol{q}}$ may be expressed as a function of $\boldsymbol{q}$ and $\boldsymbol{p}$ via (1.3), we can change variables to $(\boldsymbol{q}, \boldsymbol{p})$.

Hamilton's principle in terms of variables $(\boldsymbol{q}, \boldsymbol{p})$ leads to *Hamilton's canonical equations*

$$\dot{q}_k = \frac{\partial H}{\partial p_k}(\boldsymbol{q}, \boldsymbol{p}), \quad \dot{p}_k = -\frac{\partial H}{\partial q_k}(\boldsymbol{q}, \boldsymbol{p}), \quad k = 1, \ldots, d. \tag{1.4}$$

The canonical equation (1.4) continues to hold if $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$, and hence $H(\boldsymbol{q}, \boldsymbol{p}, t)$, contains an explicit time-dependence.

It is useful to combine all the dependent variables in a $2d$-dimensional vector $\boldsymbol{y} = (\boldsymbol{q}, \boldsymbol{p})$. Then (1.4) takes a simple form

$$\dot{\boldsymbol{y}} = J \nabla_{\boldsymbol{y}} H, \tag{1.5}$$

where $J = (J_{ij})$ is the $2d \times 2d$ skew-symmetric matrix

$$J = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix}. \tag{1.6}$$

Here $\mathbf{1}$ and $\mathbf{0}$ represent the unit and zero $d \times d$ matrices, respectively.

The Poisson bracket

An alternative form for a Hamiltonian system is the *Poisson bracket*, a skew-symmetric, bilinear form. For two functions $F(\boldsymbol{y})$ and $G(\boldsymbol{y})$ this is defined as

$$\{F, G\} = \nabla_{\boldsymbol{y}} F(\boldsymbol{y})^T J \nabla_{\boldsymbol{y}} G(\boldsymbol{y}). \tag{1.7}$$

The Poisson bracket of a function $F(\boldsymbol{y})$ with the Hamiltonian $H(\boldsymbol{y})$ is a convenient notation for expressing the time-derivative of F along a solution to the Hamiltonian system (1.5)

$$\frac{d}{dt} F(\boldsymbol{y}(t)) = \{F, H\} = \nabla_{\boldsymbol{y}} F^T J \nabla_{\boldsymbol{y}} H = \nabla_{\boldsymbol{y}} F^T \frac{d\boldsymbol{y}}{dt}.$$

Given a Poisson bracket and a Hamiltonian, the dynamical equations (1.5) can be backed out

$$\dot{y}_k = \{y_k, H\}, \quad k = 1, \ldots, 2d.$$

Generalized Hamiltonian systems are obtained by allowing J to be a skew-symmetric operator, dependent on the coordinates $\boldsymbol{y}$, i.e. $J = J(\boldsymbol{y})$. The Poisson bracket still takes the form (1.7), but is required to satisfy skew symmetry

$$\{F, G\} = -\{G, F\}$$

and the *Jacobi identity*

$$\{E, \{F, G\}\} + \{F, \{G, E\}\} + \{G, \{E, F\}\} = 0, \quad \forall E, F, G.$$

These requirements translate into analogous properties for the matrix representation of $J(\boldsymbol{y})$; respectively, $J(\boldsymbol{y})^T = -J(\boldsymbol{y})$ and

$$J_{im}\frac{\partial J_{jk}}{\partial y_m} + J_{jm}\frac{\partial J_{ki}}{\partial y_m} + J_{km}\frac{\partial J_{ij}}{\partial y_m} = 0, \quad \forall i, j, k,$$

where repeated indices are summed from 1 to d. The Jacobi identity is trivially satisfied by a constant J, so only skew-symmetry is needed in this case. Verification or construction of nonconstant Poisson brackets can be a complicated task [49, 64].

The specific form (1.6) of J corresponding to (1.4) is called the *canonical* form of a Hamiltonian system. Otherwise, the system is *noncanonical*. If J is dependent on $\boldsymbol{y}$, the system is called a *Poisson system*.

First integrals

For an ordinary differential equation $\dot{\boldsymbol{y}} = f(\boldsymbol{y})$, a nonconstant function $I(\boldsymbol{y})$ is a first integral if $I(\boldsymbol{y}(t)) = const.$ along any solution. In other words, each solution is constrained to a level set of I.

$$\frac{d}{dt}I(\boldsymbol{y}(t)) = \nabla_{\boldsymbol{y}} I(\boldsymbol{y})^T f(\boldsymbol{y}) = 0, \quad \forall \boldsymbol{y}.$$

In terms of the Poisson bracket, a first integral is a function whose Poisson bracket with H vanishes

$$\{I, H\} = 0.$$

A system may have more then one first integral. In that case the initial condition defines an intersection of the first integrals and the solution evolves on this intersection. The set of first integrals foliate the phase space constraining the dynamics, restricting the behavoir of a given trajectory. When one derives a statistical theory, the aim is to construct the least biased distribution that still reflects given information about the system. This given information may be formulated in terms of first integrals of the system.

Due to the skew symmetry of the bracket, the Hamiltonian is conserved for any (autonomous) Hamiltonian system

$$\frac{dH}{dt} = \{H, H\} = \nabla_{\boldsymbol{y}} H^T J \nabla_{\boldsymbol{y}} H = 0.$$

In physical applications this often corresponds to conservation of total energy.

Other first integrals in Hamiltonian systems arise due to continuous symmetries in the Hamiltonian, via Noether's theorem [64]. That is, if one can find a one parameter, continuous change of variables that leaves the Hamiltonian invariant, then Noether's theorem identifies a corresponding first integral.

For noncanonical Hamiltonian systems, there are sometimes functions $C(\boldsymbol{y})$, called *Casimirs*, that vanish identically in the Poisson bracket with any other function

$$\{F, C\} = 0, \quad \forall F = F(\boldsymbol{y}) \quad \Leftrightarrow \quad J \nabla_{\boldsymbol{y}} C = 0.$$

Casimirs are obviously first integrals, related to singularity of J. Possession of Casimirs is a property of the Poisson bracket, they do not depend on a particular choice of the Hamiltonian. For canonical systems there are no Casimirs.

Symplectic structure

Consider an ordinary differential equation $\dot{\boldsymbol{y}} = f(\boldsymbol{y})$ with the phase space $\mathbb{R}^{2d}$. Define the flow over time t as a mapping ϕ_t that advances the solution by time t, i.e. $\phi_t(\boldsymbol{y}^0) = \boldsymbol{y}(t, \boldsymbol{y}^0)$, where $\boldsymbol{y}(t, \boldsymbol{y}^0)$ is the solution of the system corresponding to the initial value $\boldsymbol{y}(0) = \boldsymbol{y}^0$. A smooth map ϕ_t on the phase space $\mathbb{R}^{2d}$ is called a *symplectic map* with respect to the (constant and invertible) structure matrix J if its Jacobian $\phi_t'(\boldsymbol{y})$ satisfies

$$\phi_t'(\boldsymbol{y})^T J^{-1} \phi_t'(\boldsymbol{y}) = J^{-1}, \tag{1.8}$$

for all $\boldsymbol{y}$ in the domain of definition of ϕ_t.

Theorem (Poincaré, 1899). The flow map ϕ_t of a Hamiltonian system (1.5) is symplectic.

Symplectic mappings are volume preserving. Taking the determinant of both sides of (1.8) we have

$$|\phi_t'(\boldsymbol{y})|^2 \left| J^{-1} \right| = \left| J^{-1} \right|.$$

Therefore $|\phi_t'(\boldsymbol{y})|$ is either $+1$ or -1. Since at $t = 0$ $|\phi_0'(\boldsymbol{y})| = 1$, a continuity argument shows that $|\phi_t'(\boldsymbol{y})| = 1$ for any t, i.e. the volume is preserved under the symplectic mapping.

The flow map $\phi_t(\boldsymbol{y})$ of a Poisson system is a *Poisson map*, which satisfies

$$\phi_t'(\boldsymbol{y}) J(\boldsymbol{y}) \phi_t'(\boldsymbol{y})^T = J(\phi_t(\boldsymbol{y})). \tag{1.9}$$

This is the generalization of (1.8).

1.1.2 Geometric integration

Geometric integrators are numerical methods that preserve geometric properties of the flow of a differential equation, e.g. symplectic or Poisson integrators for Hamiltonian systems, methods preserving first integrals, etc. Preservation of first integrals plays an important role in statistical mechanics, because one of the aims of statistical mechanics is to construct a proper average of desired quantities of a system under such constraints. Symplecticity gives volume preservation of the Hamiltonian phase flow, and in addition, a numerical solution trajectory on a short time scale stays close to the exact solution of the Hamiltonian system. On ergodic time intervals, the solution drifts over multiple trajectories, while preserving the phase density of solutions.

In the following section we consider processes which evolve continuously in time only, that is having no continuous spatial structure. Therefore they can be described by systems of ordinary differential equations. In this section we follow the derivations described in [35, 45, 79].

Symplectic integrators

Since symplecticity is a characteristic property of Hamiltonian systems according to theorem of Poincaré, it is tempting to search for numerical methods that share this property.

A one-step numerical method $\Phi_h : \boldsymbol{y}^{n+1} = \Phi_h(\boldsymbol{y^n})$, $t_{n+1} = t_n + h$ is called *symplectic* if the map $\boldsymbol{y}^n \to \boldsymbol{y}^{n+1}$ is symplectic whenever the method is applied to a smooth Hamiltonian system.

Here are some examples of symplectic methods: the symplectic Euler rule, the implicit midpoint rule, the Störmer-Verlet scheme, the Gauss collocation methods and other Runge-Kutta methods satisfying the symplecticity condition on the coefficients [77].

Some symplectic integrators preserve first integrals of a system. For example, implicit midpoint preserves any quadratic first integral of the form $I = \frac{1}{2}\boldsymbol{y}^T A\boldsymbol{y} + b^T\boldsymbol{y}$ for a constant symmetric matrix A and an ordinary differential equation $\dot{\boldsymbol{y}} = f(\boldsymbol{y})$. The symplectic Euler method and the generalized leapfrog method preserve any first integral of the form $I = \boldsymbol{q}^T A\boldsymbol{p}$ for the canonical variables $(\boldsymbol{q}, \boldsymbol{p})$ of a Hamiltonian system.

Symplectic methods do not, in general, conserve the Hamiltonian H exactly. In spite of this they do conserve it approximately. This good conservation of H in symplectic integration is related to the existence and exact conservation of a perturbed Hamiltonian as shown by backward error analysis.

Backward error analysis

The origin of backward error analysis dates back to [87] in numerical linear algebra. For the study of numerical differential equations, its importance was recognized much later, see e.g. [4, 34, 69, 78]. In backward error analysis the

numerical one-step map is interpreted as the flow of a modified differential equation, which is constructed as an asymptotic series.

Consider an ordinary differential equation $\dot{\boldsymbol{y}} = f(\boldsymbol{y})$ with exact mapping $\phi_{t,f}$, and a one-step numerical method $\Phi_h(\boldsymbol{y})$: $\boldsymbol{y}^{n+1} = \Phi_h(\boldsymbol{y}^n)$, $t_{n+1} = t_n + h$. A *forward, or common, error analysis* consists of the study of the error $\phi_{h,f}(\boldsymbol{y}^n) - \Phi_h(\boldsymbol{y}^n)$, i.e. the difference between the exact solution and the numerically computed approximation. The idea of *backward error analysis* is to derive a *modified differential equation*

$$\dot{\boldsymbol{y}} = \tilde{f}_i(\boldsymbol{y}, h), \tag{1.10}$$

with

$$\tilde{f}_i(\boldsymbol{y}, h) = f(\boldsymbol{y}) + hf_1(\boldsymbol{y}) + h^2 f_2(\boldsymbol{y}) + \cdots + h^i f_i(\boldsymbol{y}), \tag{1.11}$$

where the expression for $\tilde{f}_i(\boldsymbol{y}, h)$ depends on the numerical method Φ_h of the original system $\dot{\boldsymbol{y}} = f(\boldsymbol{y})$ by using the Taylor expansion of the flow map $\phi_{t,\tilde{f}_i}$ with $t = h$ at $\boldsymbol{y} = \boldsymbol{y}^n$ and comparing its Taylor expansion terms with the terms of the numerical method Φ_h in powers of h. In general for a one-step method, the series (1.11) diverges as $i \to \infty$ but may be optimally truncated as a function of h.

Consider now a Hamiltonian system (1.5) with a smooth Hamiltonian $H : \mathbb{R}^{2d} \to \mathbb{R}$. If a symplectic method $\Phi_h(\boldsymbol{y})$ is applied to it, then the modified vector field $\tilde{f}_i(h)$ in the equation (1.10) is also Hamiltonian with a modified Hamiltonian function $\tilde{H}_i(h)$. More precisely,

$$\tilde{f}_i(\boldsymbol{y}, h) = J\nabla_{\boldsymbol{y}}\tilde{H}_i(\boldsymbol{y}, h),$$

and the modified Hamiltonian $\tilde{H}_i(\boldsymbol{y}, h)$ is close to the originally given Hamiltonian $H(\boldsymbol{y})$, i.e.

$$|\tilde{H}_i(\boldsymbol{y}, h) - H(\boldsymbol{y})| = O(h^p),$$

where $p \geq 1$ is the order of the symplectic method Φ_h [4].

For the Hamiltonian system (1.5), the modified Hamiltonian $\tilde{H}_i(\boldsymbol{y}, h)$ can be written as

$$\tilde{H}_i(\boldsymbol{y}, h) = H(\boldsymbol{y}) + hH_1(\boldsymbol{y}) + h^2 H_2(\boldsymbol{y}) + \cdots + h^i H_i(\boldsymbol{y}),$$

where the set $\{H_i\}$ depends on the symplectic method Φ_h as it was described above, and where each H_i is a Hamiltonian of some Hamiltonian system.

Backward error analysis shows the advantages of using symplectic integrators for Hamiltonian systems. It was shown in [35, 45, 79] that for a symplectic integrator applied to an autonomous Hamiltonian system, modified autonomous Hamiltonian problems exist so that the numerical solution of the original problem is the exact solution of the modified problem. On the contrary, when a nonsymplectic integrator is used the modified system is not in general Hamiltonian anymore. For general Hamiltonians, in [33] it was proven that a symplectic method Φ_h cannot exactly conserve energy. Nevertheless, the symplectic integrators do a good job in preserving Hamiltonians approximately. Consider a

Hamiltonian system with analytic $H : \mathbf{K} \to \mathbb{R}$ (where $\mathbf{K} \subset \mathbb{R}^{2d}$ is an open subset), and apply a symplectic numerical method $\Phi_h(\boldsymbol{y})$ with step size h. If the numerical solution stays in a compact subset of $\mathbf{K}$, then there exists a $\vartheta > 0$ such that

$$\begin{aligned} \tilde{H}_i(\boldsymbol{y}^n, h) &= \tilde{H}_i(\boldsymbol{y}^0, h) + O(e^{-\vartheta/2h}), \\ H(\boldsymbol{y}^n, h) &= H(\boldsymbol{y}^0, h) + O(h^p) \end{aligned}$$

over exponentially long time intervals $nh \leq e^{\vartheta/2h}$. In typical applications, the Hamiltonian oscillates around its initial value with bounded amplitude $O(h^p)$.

In [70], it is proven that whenever the flow map of a given differential equation possess some geometric properties such as existence of first integrals, time reversibility, preservation of volume, symplecticness, and the numerical discretization preserves these properties exactly, then the flow map of the modified differential equation will also satisfy these geometric properties.

Poisson integrators

Poisson integrators generalize of symplectic integrators to Poisson systems.

A numerical one-step method $\boldsymbol{y}^{n+1} = \Phi_h(\boldsymbol{y}^n)$, $t_{n+1} = t_n + h$, is a *Poisson integrator* for the structure matrix $J(\boldsymbol{y})$, if the transformation $\boldsymbol{y}^n \to \boldsymbol{y}^{n+1}$ respects the Casimirs and if it is a Poisson map (1.9) whenever the method is applied to the Poisson system.

According to backward error analysis, if a Poisson integrator Φ_h is applied to the Poisson system, then the modified equation is locally a Poisson system [35].

There is no general technique for constructing Runge-Kutta type Poisson methods. The most generally applicable alternative is splitting methods, e.g. [50].

1.2 Hamiltonian fluid dynamics

The study of the dynamics of fluids is one of the most attractive areas in applied mathematics. The fact that fluid dynamics is an attractive research area is due to many reasons. Perhaps the most important one is the introduction of efficient high resolution numerical simulations into fluid dynamics as a research tool. The significance of this tool is especially pronounced in case of complex behaviour of a system.

1.2.1 Hamiltonian PDEs

The dynamics of fluids has both propagation in time and a spatial structure and, hence, cannot be described by ordinary differential equations (ODEs) anymore but by partial differential equations (PDEs). Many PDEs that arise in physics can be viewed as infinite-dimensional Hamiltonian systems. (Problems described by ordinary differential equations are finite-dimensional.) The numerical integration of infinite-dimensional Hamiltonian systems or Hamiltonian

PDEs is much less explored than that of ODEs, since the solution behaviour of PDEs is much more complex. Nevertheless, there exists a number of papers for Hamiltonian PDEs [12, 14, 17, 24, 26, 27, 31, 32, 51, 83].

The finite-dimensional Hamiltonian system consists of a triple $(\mathbf{K}, \{\cdot,\cdot\}, H)$, where the phase space $\mathbf{K} \subset \mathbb{R}^{2d}$ is an open subset, $H : \mathbf{K} \to \mathbb{R}$ is the Hamiltonian function, and $\{\cdot,\cdot\}$ is a Poisson bracket with structure matrix $J(\boldsymbol{y})$, see (1.7). When the phase space is infinite-dimensional, we write the triple as $(\mathcal{K}, \{\cdot,\cdot\}, \mathcal{H})$, and the Poisson operator as $\mathcal{J}$, according to [64]. Typically $\mathcal{K}$ consists of sets of smooth functions on a finite-dimensional space Y. An element in $\mathcal{K}$ is denoted by $u(\boldsymbol{y})$, $\boldsymbol{y} \in Y$. The Hamiltonian $\mathcal{H} : \mathcal{K} \to \mathbb{R}$ is a functional on this space, and the bracket can be written as

$$\{\mathcal{F}, \mathcal{G}\}[u] = \int_Y \frac{\delta \mathcal{F}}{\delta u} \mathcal{J}(u) \frac{\delta \mathcal{G}}{\delta u}\, d\boldsymbol{y},$$

where $\delta\mathcal{F}/\delta u$ is the variational derivative defined by

$$\lim_{\epsilon \to 0} \frac{\mathcal{F}[u + \epsilon \delta u] - \mathcal{F}[u]}{\epsilon} \equiv \int_Y \frac{\delta \mathcal{F}}{\delta u}\, \delta u\, d\boldsymbol{y}$$

for appropriate δu. $\mathcal{J}(u)$ is, in general, a differential operator, called the *Poisson operator*.

Motivated by the success of symplectic integrators, a reasonable approach to Hamiltonian PDEs is to try to discretize in space while preserving the symplectic or Poisson structure. For canonical structure, it is a simple matter to discretize the Hamiltonian functional with any desired quadrature. The result is a Hamiltonian ODE to which symplectic integrators may be applied. For Poisson systems, it is a significant challenge to derive a discrete bracket that preserves the Jacobi identity. If there are Casimirs, there should be some remnant of these. For fluids, there is an infinite family, and only a finite number of independent integrals can survive, if they constrain the discrete, finite dimensional phase space.

On the other hand, since Lagrangian fluid dynamics *is* canonical, one can approximate the PDE solution with a set of moving particles interacting through an appropriate potential energy function, and a Hamiltonian semi-discretization will be obtained for any quadrature of the Hamiltonian. The set of finite-dimensional Hamiltonian ODEs is then integrated in time using a suitable symplectic or Poisson integrator.

Unfortunately, in the case of a Poisson PDE, unlike canonical Hamiltonian PDEs, it is not possible to establish a common generic approach. For each particular problem one has to develop a proper way of reducing the PDE to a system of ODEs. We will consider the quasigeostrophic potential vorticity equation (a Hamiltonian PDE with Poisson structure), for which we will describe several numerical methods.

1.2.2 Eulerian and Lagrangian descriptions

Geophysical fluid dynamics is the study of fluid motion in the atmosphere and the ocean. The first step in this study is a choice of framework: Eulerian or Lagrangian. The *Eulerian description* is commonly used in literature, e.g. [52, 67], it treats motion of the fluid as a field in which the velocity is to be determined at all positions and times. The *Lagrangian description* regards the fluid as a continuous field of particles, whose positions are to be determined [74].

In the Eulerian description, the independent variables are the space coordinates $\boldsymbol{x} = (x, y, z) \in \mathcal{D}$ and the time t. The dependent variables include the velocity $\boldsymbol{v}(\boldsymbol{x}, t)$ and the mass density $\rho(\boldsymbol{x}, t)$

The noncanonical Poisson bracket for ideal fluid flow in Eulerian variables, according to [58], is

$$\{F, G\} = -\int_{\mathcal{D}} d\boldsymbol{x} \left[\frac{\delta F}{\delta \rho} \nabla_{\boldsymbol{x}} \cdot \frac{\delta G}{\delta \boldsymbol{v}} - \frac{\delta G}{\delta \rho} \nabla_{\boldsymbol{x}} \cdot \frac{\delta F}{\delta \boldsymbol{v}} + \left(\frac{\nabla_{\boldsymbol{x}} \times \boldsymbol{v}}{\rho} \right) \cdot \left(\frac{\delta G}{\delta \boldsymbol{v}} \times \frac{\delta F}{\delta \boldsymbol{v}} \right) \right]$$

with Hamiltonian

$$H[\boldsymbol{v}, \rho] = \int_{\mathcal{D}} d\boldsymbol{x} \left[\frac{\boldsymbol{v} \cdot \boldsymbol{v}}{2\rho} + \rho E(\rho) \right].$$

Here $\boldsymbol{x}$ is an independent variable.

The Poisson bracket has an infinite class of potential vorticity Casimirs of the form

$$\mathcal{C}[\rho] = \int_{\mathcal{D}} d\boldsymbol{x}\, \rho f \left(\frac{\nabla_{\boldsymbol{x}} \times \boldsymbol{v}}{\rho} \right)$$

for an arbitrary function f.

In the Lagrangian description, each fluid particle is assigned a label $\boldsymbol{a} = (a, b, c) \in \mathcal{A}$. For example, the labels may be defined as the positions of particles at the initial time. The independent variables are set of $\boldsymbol{a}$, which are fixed for each particle, and the time t. The dependent variables are the position coordinates $\boldsymbol{x}(\boldsymbol{a}, t)$. The velocity of a particle is given by

$$\boldsymbol{v} = \left(\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}, \frac{\partial z}{\partial t} \right).$$

The mass density ρ is defined via Jacobian matrix $|\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{a}}|$ as

$$\rho = \rho_0 \left| \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{a}} \right|^{-1},$$

where $\rho_0 = \rho_0(\boldsymbol{a})$ does not depend on time t.

Taking derivatives of the expression above leads to the continuity equation

$$\frac{d\rho}{dt} + \rho \nabla_{\boldsymbol{x}} \cdot \boldsymbol{v} = 0.$$

The Poisson bracket for ideal fluid flow in Lagrangian variables is canonical

$$\{F, G\} = \int_{\mathcal{A}} d\boldsymbol{a} \left[\frac{\delta F}{\delta \boldsymbol{x}} \cdot \frac{\delta G}{\delta \boldsymbol{v}} - \frac{\delta G}{\delta \boldsymbol{x}} \cdot \frac{\delta F}{\delta \boldsymbol{v}} \right]$$

with Hamiltonian

$$H[\boldsymbol{x}, \boldsymbol{v}] = \int_{\mathcal{A}} d\boldsymbol{a} \left[\frac{\boldsymbol{v} \cdot \boldsymbol{v}}{2\rho_0} + \rho_0 E \left(\rho_0 \left| \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{a}} \right|^{-1} \right) \right].$$

Here $\boldsymbol{x}$ is a dependent variable.

The quasigeostrophic model

The two-dimensional quasigeostrophic potential vorticity (QG) equation [48, 67, 74] describes divergence-free flow over topography by

$$\frac{d}{dt} q = 0, \qquad \Delta \psi(\boldsymbol{x}, t) = q(\boldsymbol{x}, t) - h(\boldsymbol{x}), \tag{1.12}$$

where q is the potential vorticity (PV) field, ψ is the stream function, and h is the topography of the earth. The Laplacian operator is denoted by Δ and the material derivative by $\frac{d}{dt} = \frac{\partial}{\partial t} + \boldsymbol{u} \cdot \nabla$. Here, the divergence-free velocity field $\boldsymbol{u}$ is related to the stream function by $\boldsymbol{u} = \nabla^{\perp} \psi$, where $\nabla^{\perp} = (-\frac{\partial}{\partial y}, \frac{\partial}{\partial x})^T$. We consider the QG equation on a doubly periodic domain

$$\boldsymbol{x} = (x, y) \in \mathcal{D} \equiv [0, 2\pi) \times [0, 2\pi).$$

Define the operator $\mathcal{J}(q, \psi) = q_x \psi_y - q_y \psi_x$. The QG model describes a Hamiltonian PDE with Poisson structure [57],

$$\{\mathcal{F}, \mathcal{G}\} = \int_{\mathcal{D}} q \mathcal{J} \left(\frac{\delta \mathcal{F}}{\delta q}, \frac{\delta \mathcal{G}}{\delta q} \right) d\boldsymbol{x},$$

implying the conservation of the Hamiltonian or total kinetic energy

$$\mathcal{H} = \mathcal{E} = -\frac{1}{2} \int_{\mathcal{D}} \psi \cdot (q - h) \, d\boldsymbol{x}$$

as well as the infinite class of Casimir functionals

$$\mathcal{C}[f] = \int_{\mathcal{D}} f(q) \, d\boldsymbol{x}$$

for any function f for which the integral exists. Of specific interest are the moments of PV

$$\mathcal{C}_r = \int_{\mathcal{D}} q^r \, d\boldsymbol{x}, \quad r = 0, 1, 2, \ldots, \tag{1.13}$$

and in particular the circulation $\mathcal{C}_1$ and enstrophy $\mathcal{C}_2$.

Preservation of the Casimir functionals follows from area-preservation under the divergence-free velocity field [55]. Define a function $G(\sigma, t)$ denoting the measure of that part of the domain $\mathcal{D}$ for which the vorticity is less than σ:

$$G(\sigma, t) = \text{meas}\{\boldsymbol{x} \in \mathcal{D} \,|\, q(\boldsymbol{x}, t) < \sigma\}.$$

We note that due to the divergence-free advection of q, this function is independent of time, $\frac{\partial G}{\partial t} = 0$. Differentiating with respect to σ, the function

$$g(\sigma) = \frac{\partial G}{\partial \sigma} \tag{1.14}$$

is preserved. For the case of a piecewise uniform PV field, $q(\boldsymbol{x}, t) \in \{\sigma_1, \ldots, \sigma_L\}$, this quantity $g_\ell = G(\sigma_{\ell+1}) - G(\sigma_\ell)$ is the measure of the vorticity level set σ_ℓ.

1.2.3 Numerical methods

In the following section we describe several numerical methods to solve the quasigeostrophic model (1.12).

The Zeitlin method

Normal spectral methods for the QG equation preserve the energy and enstrophy at most. However, the Zeitlin method [88] is a spectral method which preserves a Poisson structure, the Hamiltonian and $2M$ Casimirs in a $(2M + 1) \times (2M + 1)$ mode truncation.

The first equation in (1.12) is transformed through two-dimensional Fourier series taking the form of an infinite system of ODEs

$$\frac{d\hat{q}_{\boldsymbol{k}}}{dt} = \sum_{\substack{k_1', k_2' = -\infty \\ \boldsymbol{k}' \neq 0}}^{\infty} \frac{\boldsymbol{k} \times \boldsymbol{k}'}{|\boldsymbol{k}'|^2} \hat{q}_{\boldsymbol{k}+\boldsymbol{k}'} (\hat{q}_{-\boldsymbol{k}'} - \hat{h}_{-\boldsymbol{k}'}). \tag{1.15}$$

Here $\hat{q}_{\boldsymbol{k}}$ denotes the spectral coefficient associated with the two-dimensional wave vector $\boldsymbol{k}$, whose components are integers. The skew-symmetric scalar product $\boldsymbol{k} \times \boldsymbol{k}'$ is $k_1 k_2' - k_2 k_1'$, and the norm $|\boldsymbol{k}|$ is $\sqrt{k_1^2 + k_2^2}$. Since q is real, $\hat{q}_k^* = \hat{q}_{-k}$.

Zeitlin proposed the sine-bracket truncation of the equations. The finite-dimensional set of equations for the Fourier coefficients is then given by

$$\frac{d\hat{q}_{\boldsymbol{k}}}{dt} = \sum_{\substack{k_1', k_2' = -M \\ \boldsymbol{k}' \neq 0}}^{M} \frac{1}{\epsilon} \frac{\sin(\epsilon \boldsymbol{k} \times \boldsymbol{k}')}{|\boldsymbol{k}'|^2} \hat{q}_{\boldsymbol{k}+\boldsymbol{k}'} (\hat{q}_{-\boldsymbol{k}'} - \hat{h}_{-\boldsymbol{k}'}), \quad \epsilon = \frac{2\pi}{2M+1}, \tag{1.16}$$

where all indices are reduced modulo $2M + 1$ to the periodic lattice $-M \leq k_1, k_2 \leq M$. The summation occurs on the $(2M + 1) \times (2M + 1)$ domain of the Fourier coefficients. For $M \to \infty$ and given $\boldsymbol{k}$ and $\boldsymbol{k}'$, $\epsilon^{-1} \sin(\epsilon \boldsymbol{k} \times \boldsymbol{k}') = \boldsymbol{k} \times \boldsymbol{k}' + O(\epsilon^2)$, which gives consistency.

This truncation possesses a Hamiltonian structure with symplectic operator

$$J_{\boldsymbol{k}\boldsymbol{k}'} = -\epsilon^{-1}\sin(\epsilon\boldsymbol{k}\times\boldsymbol{k}')\hat{q}_{\boldsymbol{k}+\boldsymbol{k}'},$$

and Hamiltonian

$$H = E = \frac{1}{2}\sum_{\boldsymbol{k}\neq 0}|\boldsymbol{k}|^2|\hat{\psi}_{\boldsymbol{k}}|^2 = \frac{1}{2}\sum_{\boldsymbol{k}\neq 0}\frac{|\hat{q}_{\boldsymbol{k}}-\hat{h}_{\boldsymbol{k}}|^2}{|\boldsymbol{k}|^2}. \tag{1.17}$$

The symplectic matrix is skew-symmetric and satisfies the Jacobi identity. The sine-bracket truncation (1.16) preserves the Hamiltonian (1.17) and $2M$ independent Casimir invariants corresponding to the first $2M$ moments of potential vorticity. If in the Zeitlin method the Poisson discretization is integrated using the Poisson splitting of McLachlan [50], then these quantities are preserved by the splitting (the energy is only preserved approximately, in the sense of backward error analysis [35]).

Unfortunately, the Zeitlin method is limited to 2D incompressible flows on periodic geometry.

Arakawa's scheme

The Zeitlin method is the only known discretization with Poisson structure for Eulerian fluid models. For more general fluid problems (compressible, non-periodic boundary conditions, etc.) no Poisson discretizations are available. How can we preserve at least some quantities? A well known scheme is Arakawa's scheme [2] which preserves linear and quadratic invariants.

For a start we rewrite the first equation in (1.12) as

$$q_t = \mathcal{J}(q,\psi), \tag{1.18}$$

where the operator $\mathcal{J}$ is defined by

$$\mathcal{J}(q,\psi) = q_x\psi_y - q_y\psi_x.$$

Arakawa's idea consists of several steps. First of all, he uses central differences for x- and y-derivatives. Then he rewrites the continuous $\mathcal{J}$ in three equivalent forms based on the fact that the derivatives with respect to x and y commute in the continuous case, namely $\mathcal{J}(q,\psi) = \partial_x(q\psi_y) - \partial_y(q\psi_x) = \partial_y(q_x\psi) - \partial_x(q_y\psi)$. After discretizing these three equivalent forms of $\mathcal{J}$ and taking their average, one gets four discrete non-equivalent right-hand sides of (1.18), therefore four discretizations. Non-equivalence of the discrete right-hand sides is explained by the fact that the product rule does not hold anymore in the discrete case. It can be shown that one discretization does not conserve anything and, in fact, is unstable; second one conserves only energy; third one—only enstrophy; and fourth one, which is an average of the previous three, conserves both energy and enstrophy. It is worth mentioning that all these discretizations are also volume preserving in the sense of the Liouville property, see Section (1.3). This is a necessary ingredient for a statistical theory.

Arakawa's discretizations are explained in more detail in Chapter 2.

The Hamiltonian particle-mesh method

The Hamiltonian particle-mesh (HPM) method approximates the solution of an ideal fluid flow using a set of moving particles that interact through an appropriate potential energy function. The HPM method was originally designed for rotating shallow water flow with periodic boundary conditions [25] and extended to other physical settings—rotating two-layer shallow water model with rigid-lid constraint, barotropic model, non-hydrostatic vertical slice model, in [16, 17, 28, 83].

HPM is based on the Lagrangian formulation of fluid dynamics, and a set of moving point masses combined with a fixed Eulerian grid. The spatially truncated equations are canonical Hamiltonian and satisfy a Kelvin circulation theorem [25]. Implementation of a splitting method in time to the semi-discrete system of Hamiltonian ODEs gives symplecticity. Hence the HPM method is symplectic, and one can construct a continuum velocity field in which the discrete particle velocities are embedded for all time. Energy is preserved approximately in the sense of backward error analysis, i.e. good long-time energy conservation. Convergence of the method was considered in [56].

The idea of fixing the potential vorticity rather than the mass to a particle was originally proposed in this context for two-dimensional advection under incompressible flow fields in [17]. The result is a regularized point vortex method. There the potential vorticity of a quasigeostrophic model is simply advected in a divergence-free velocity field, obtained by reconstructing the PV field on a uniform grid. The particle motion can be embedded in an area preserving flow on the fluid label space. Hence, the Casimirs (1.13) are trivially conserved if a value of potential vorticity is simply assigned to each particle once and for all. The semi-discrete system is still Hamiltonian, and can be integrated in time using a symplectic integrator. Since the HPM method is symplectic, the phase flow is volume preserving in the sense of the Liouville property, see Section (1.3), which is a necessary ingredient for a statistical theory.

The HPM method for the quasigeostrophic model is explained in more detail in Chapter 3.

1.3 Statistical mechanics of fluids

Statistical mechanics is a powerful tool for understanding complex physical systems, e.g. [23, 30]. There are different purposes of statistical mechanics. One may consider a system for a very long time. Then the quantity of statistical interest is the average behaviour of a system rather than the behaviour of a system at a certain time. An example is a traffic outside of your window. A statistical quantity of interest may be the average speed of each car. Another purpose of statistical mechanics is to design methods to handle systems which are incompletely known. For example, if we do not know the initial conditions of a system, and we want to know its most likely behaviour. This involves an average over all possible states of the system with different initial conditions.

Statistical mechanics is broadly employed in numerical simulations of molecular dynamics and climate variability. Here either a system is very big (10^{23} molecules) or long time simulations are chaotic, or both. Moreover, different numerical discretizations of a dynamical system have different discrete dynamics. Therefore the statistical results can also be distinct.

The main premise of the equilibrium statistical theories of either continuous or discrete dynamics is ergodicity relating the equilibrium distribution to the dynamics. Necessary ingredients are conservation laws and volume preservation of the phase flow. Therefore, to derive a statistical theory of a numerical method the numerical method has to possess conservation laws, and the discrete phase flow should be volume preserving. When one integrates a Hamiltonian system with a symplectic integrator, there is an automatic conservation of discrete analogues of the exact constants of motion and, because of symplecticity, there is volume preservation. Ergodicity is difficult to show for nontrivial systems, so this is usually taken as an assumption or 'approximation' which must be verified with numerics.

The aim of the following section is to explain statistical mechanics, to give a definition of different ensembles and purpose of each, and to discuss the concept of ergodicity, entropy and temperature in statistical sense.

Statistical equilibrium

Consider an ordinary differential equation $\dot{\boldsymbol{y}} = f(\boldsymbol{y})$ with the phase space $\mathbf{K} \subset \mathbb{R}^{2d}$, then a probability distribution function $\rho(\boldsymbol{y}, t)$, $\rho : \mathbf{K} \times \mathbb{R} \to \mathbb{R}$, for example, over a set of uncertain initial conditions, is transported by the flow according to

$$\frac{\partial}{\partial t}\rho + \nabla_{\boldsymbol{y}} \cdot \rho f = 0.$$

Now consider a Hamiltonian system (1.5).

Theorem (Liouville, 1838). The phase flow of a Hamiltonian system (1.5) is volume preserving [3].

For constant J this follows from the skew-symmetry of J, or equivalently, the divergence-free nature of the canonical phase flow $\boldsymbol{y}$. One says that the flow has the *Liouville property*, if this flow satisfies Liouville's theorem, i.e. the flow is volume preserving, i.e. the flow is divergence-free. The Liouville property is a necessary ingredient for a statistical theory, since from it follows that the probability measure is transported under the divergence-free flow. Therefore before deriving a statistical theory, one has to prove that the phase flow is divergence-free.

The transport equation in the case of a Hamiltonian system simplifies to

$$\frac{\partial}{\partial t}\rho + J\nabla_{\boldsymbol{y}} H \cdot \nabla_{\boldsymbol{y}} \rho = 0,$$

which is called the *Liouville equation*.

A steady state of the Liouville equation for the Hamiltonian flow, $\frac{\partial \rho}{\partial t} = 0$, is then

$$J\nabla_{\boldsymbol{y}} H \cdot \nabla_{\boldsymbol{y}} \rho = 0,$$

which is often refered as an *equilibrium* probability distribution function. Note that any function $\rho(\boldsymbol{y}) = \rho(H(\boldsymbol{y}))$ of the Hamiltonian is an equilibrium probability distribution, and to find the proper probability, which corresponds to the dynamics described by the ODE, is the topic of ergodic theory, see e.g. [68, 85].

To be more precise, let us define the time average of a function $F(\boldsymbol{y}(t))$ as

$$\overline{F} \equiv \lim_{T\to\infty} \frac{1}{T} \int_0^T F(\boldsymbol{y}(t))dt,$$

provided that the limit exists.

The ensemble average of F is defined as

$$\langle F \rangle \equiv \int_{\mathbf{K}} F(\boldsymbol{y})\rho(\boldsymbol{y})\, d\boldsymbol{y} \equiv \int_{\mathbf{K}} F\, \nu(d\boldsymbol{y})$$

for a proper measure ν such that $\nu > 0$ and $\int_{\mathbf{K}} \nu(d\boldsymbol{y}) = 1$.

Ergodicity implies that the long time average is equivalent to the ensemble average

$$\overline{F} = \langle F \rangle \tag{1.19}$$

with the probability measure ν or a reasonable approximation to it. Given the probability measure ν it is a challenging task to prove the ergodicity, and we will take it as an assumption.

1.3.1 Statistical ensembles

Let us introduce an important idea of microstate and macrostate of a system. Consider a model in which fluid motion is described by a set of moving particles. Then a *microstate* of such system can be described by positions of these particles, and a *macrostate* is, for example, the observable energy which may correspond to a large number of microstates.

Below we will explain this in more detail following the derivations described in [10].

The microcanonical ensemble

Consider a discrete space $\mathbf{K}$ with a single macrostate given by the energy $H = E$. Let $\boldsymbol{y} \in \mathbf{K}$. The subset

$$\mathbb{D}(E) = \{\boldsymbol{y} \in \mathbf{K} :\ H(\boldsymbol{y}) = E\}$$

consists of discrete states $\boldsymbol{y}$ with the same energy E.

Define $\Omega(E)$ to be the total number of states $\boldsymbol{y} \in \mathbf{K}$ with the energy E. The microcanonical ensemble is the set of all $\boldsymbol{y}$ having $H(\boldsymbol{y}) = E$. Assuming

all such states are equally likely we define the *microcanonical density* for the discrete space

$$\text{Prob}\{\boldsymbol{y} \,|\, H = E\} = \begin{cases} 0, & H(\boldsymbol{y}) \neq E \\ 1/\Omega(E), & H(\boldsymbol{y}) = E. \end{cases}$$

The microcanonical entropy is defined as

$$S(E) = \ln \Omega(E).$$

Consider now a joint Hamiltonian system AB, which consists of two systems A and B, with total energy E such that, the coupling element allows exchange of the energy between systems A and B and adds neither new states to the system nor new terms into the Hamiltonian.

A microstate of system A is defined as $\boldsymbol{y}_A$. If system A is in a microstate $\boldsymbol{y}_A$ with corresponding energy E_A, then since there is total energy conservation, system B should be in a state with the energy $E - E_A$. The probability of the microstate $\boldsymbol{y}_A$ is

$$\text{Prob}\{\boldsymbol{y}_A \,|\, H = E\} = \frac{\Omega_B(E - H_A(\boldsymbol{y}_A))}{N(E)}, \tag{1.20}$$

where $N(E)$ is a normalization constant such that

$$\sum_{\boldsymbol{y}_A \in \mathbf{K}_A} \text{Prob}\{\boldsymbol{y_A} \,|\, H = E\} = 1 \quad \Leftrightarrow \quad N(E) = \sum_{\boldsymbol{y}_A \in \mathbf{K}_A} \Omega_B(E - H_A(\boldsymbol{y}_A)).$$

This is the microcanonical probability of a microstate. The microcanonical probability of the macrostate $H_A = E_A$ is the energy split

$$\text{Prob}\{H_A = E_A \,|\, H = E\} = \frac{\Omega_A(E_A)\Omega_B(E - E_A)}{N(E)}$$

with the normalization constant $N(E)$.

The most probable macrostate can be found by maximizing the number of states with the energy split over all possible states of system A with the energy E_A

$$\max_{E_A} \left[\Omega_A(E_A)\Omega_B(E - E_A)\right],$$

which in terms of entropies gives

$$\max_{E_A} \left[S_A(E_A) + S_B(E - E_A)\right].$$

From the last expression it is clear that the most probable state is the maximizer of the total entropy. When the joint system AB is very large, then it is possible to make transition from the discrete phase space to a continuous one, and then

the maximizer of the total entropy can be found analytically: the maximum occurs at some E_A^* where

$$S'_A(E_A^*) = S'_B(E - E_A^*),$$

and the prime denotes the derivative with respect to the argument. This has motivated the definition of the *microcanonical statistical temperature*

$$S'(E) = \frac{1}{\mathcal{T}}.$$

Note that $\mathcal{T}_A = \mathcal{T}_B$ at the most likely macrostate. In the literature one often uses the *inverse statistical temperature* $\beta = 1/\mathcal{T}$.

Now we show how to derive the microcanonical density for a continuous phase space. The rest of the microcanonical statistical theory for a continuous phase space follows automatically with sums replaced by integrals.

If the energy is the only conserved quantity of the system, then we consider the subspace

$$\mathbb{D}(E, dE) = \{\boldsymbol{y} \in A : \, H(\boldsymbol{y}) \in [E, E + dE]\}, \tag{1.21}$$

with corresponding density

$$\rho(\boldsymbol{y}) = \begin{cases} 0, & H(X) \notin [E, E + dE] \\ 1/\mathrm{vol}\{\mathbb{D}\}, & H(X) \in [E, E + dE]. \end{cases} \tag{1.22}$$

The density (1.22) is a stationary density, since for fixed (E, dE) it depends only on autonomous H. If we take the limit $dE \to 0$, the density ρ is presented only on the surface $H = E$. Then the microcanonical density for the continuous phase space can be written down in terms of Dirac delta functions

$$\rho(\boldsymbol{y}) = \frac{1}{\Omega(E)}\delta(H - E) \quad \text{with} \quad \Omega(E) = \int_{\mathbf{K}} \delta(H - E)\, d\boldsymbol{y},$$

where $\Omega(E)$ is the measure of the surface $H = E$.

We want to underline that the microcanonical statistical mechanics is derived assuming conservation of some quantities, e.g. energy. Therefore all possible states with these constant quantities, i.e. the macrostate, determine the microcanonical ensemble.

The canonical ensemble

Consider again the joint system AB, but with the size of the system B much larger then the size of the system A, the size of the system A may or may not be large compared to unity. In this case the system B is called an energy reservoir for the system A.

Now we would like to derive an equivalence of (1.20) for the described system. First, let us write down (1.20) in terms of entropy:

$$\mathrm{Prob}\{\boldsymbol{y}_A \,|\, H = E\} \sim \exp(S_B(E - E_A)).$$

Since the function S is a slowly varying function in an ample range of possible microstates of the large system B, in contrast to Ω which is not, we can write down the Taylor expansion of $S_B(E - E_A)$, and we truncate the series after the first term

$$S_B(E - E_A) \sim S_B(E) + S'_B(E)(E - E_A - E).$$

Absorbing $S_B(E)$ into the constant of normalization and noticing that $S'_B(E)$ is the inverse temperature β_B, we obtain

$$\text{Prob}\{\boldsymbol{y}_A \,|\, H = E\} \sim \exp(-\beta_B E_A).$$

This motivates the definition of the *canonical probability density* of a system in contact with an energy reservoir in case of a discrete phase space

$$\text{Prob}\{\boldsymbol{y} \,|\, \beta\} = \frac{1}{N(\beta)} \exp(-\beta H(\boldsymbol{y})) \quad \text{and} \quad N(\beta) = \sum_{\boldsymbol{y} \in \mathbf{K}} \exp(-\beta H(\boldsymbol{y})). \quad (1.23)$$

For a continuous phase space $\mathbf{K}$, the sum in (1.23) is replaced by the integral over $\mathbf{K}$.

Sampling

Now that we have defined the statistical ensembles and the probability densities associated with them, a question arises how to ensure sampling of these ensembles. If we consider a system with conserved energy, and having ergodicity, we can sample a microcanonical distribution of constant energy by simulating the dynamics of the system for a long time. Then the ensemble originated by the time series of the dynamics is equivalent to the microcanonical ensemble associated with the constant energy.

There are several approaches to ensure sampling of a canonical distribution. These approaches work in such a way that it either modifies the dynamical system or introduces a stochastic perturbation. The most known classical method for sampling a canonical distribution is the Metropolis algorithm [53]. It is based on a random choice of a state and acceptance of this state depending on the probability which should be sampled. A popular methodology in molecular dynamics to sample a canonical distribution is a mathematical thermostat—a tool to model the system in thermal equilibrium with a reservoir. The thermostat is responsible for the energy exchange between the system and the energy reservoir such that the system stays at a given temperature, which forces sampling of the canonical equilibrium distribution. Here are several thermostat techniques. The classical ones are Langevin dynamics [80], which is a stochastic thermostat, and deterministic thermostats such as the Nosé method [61, 62] and the Nosé-Hoover method [37, 62]. In Langevin dynamics the combination of a damping force and a stochastic term maintains the system at a given temperature. Because of the presence of damping and the introduction of a random forcing, the dynamics are not any longer Hamiltonian. The Nosé and Nosé-Hoover methods preserve the Hamiltonian structure and achieve sampling by working in an extended phase space.

1.3.2 Information theory

We have considered the maximum entropy principle based on the maximization of the number of states of a system and replacing this number by the total entropy. There is an alternative approach to the maximum entropy principle based on Shannon entropy of information theory, which is not derived from physical principles.

Consider an infinite dimensional phase space $\mathbf{K}$ (it could also be finite) with an element of it denoted by $\boldsymbol{y}$. Then the *Shannon entropy*, or information entropy, is

$$S[\rho] = -\int_{\mathbf{K}} \rho \ln \rho \, d\boldsymbol{y}. \tag{1.24}$$

The concept of Shannon entropy plays the central role in information theory, sometimes referred as measure of uncertainty [81]. The probability density function ρ is chosen to maximize S under constraints corresponding to observations on the system. These minimal assumptions lead to the distribution of least bias, i.e. the distribution which is most general and still explains the observations. We show that the microcanonical and canonical distributions are the maximizers of Shannon entropy under suitable constraints.

Suppose that we have only one constraint on the probability density, namely the normalization constraint

$$\int_{\mathbf{K}} \rho \, d\boldsymbol{y} - 1 = 0. \tag{1.25}$$

To ensure this constraint we need to include it in the action principle via Lagrange multiplier θ

$$\delta \left[-\int_{\mathbf{K}} \rho \ln \rho \, d\boldsymbol{y} - \theta \left(\int_{\mathbf{K}} \rho \, d\boldsymbol{y} - 1 \right) \right] = 0$$

for arbitrary variations $\delta\rho$ and $\delta\theta$. After taking variations with respect to ρ we have $\ln \rho = -1 - \theta$ with θ determined by the normalization constraint (1.25). Therefore

$$\rho = \frac{1}{\mathrm{vol}\{\mathbf{K}\}}, \quad \text{and} \quad S^* = \ln \mathrm{vol}\{\mathbf{K}\}$$

is the maximized entropy. Thus, with no further assumptions, the least biased distribution is uniform.

Consider the phase space (1.21) with the microcanonical probability density (1.22). Then it can be shown that this density is the maximizer of Shannon entropy under both the normalization constraint (1.25) and the constraint that the energy can only take values $H \in [E, E+dE]$. The corresponding maximized Shannon entropy is $S^* = \ln \mathrm{vol}\{\mathbb{D}(E, dE)\}$.

The canonical distribution (1.23) is the maximizer of Shannon entropy as well, but under other constraints, namely, under the normalization constraint (1.25) and the constraint of observed mean energy U

$$\int_{\mathbf{K}} H\rho \, d\boldsymbol{y} - U = 0, \tag{1.26}$$

for some fixed U.

The action principle states that

$$\delta\left[-\int_{\mathbf{K}}\rho\ln\rho\,d\boldsymbol{y}-\theta\left(\int_{\mathbf{K}}\rho\,d\boldsymbol{y}-1\right)-\beta\left(\int_{\mathbf{K}}H\rho\,d\boldsymbol{y}-U\right)\right]=0$$

for arbitrary $\delta\rho$, $\delta\theta$ and $\delta\beta$. Here θ and β are Lagrange multipliers corresponding to the normalization and the fixed mean energy constraints respectively.

Taking variations with respect to ρ we obtain

$$\rho=\frac{1}{\exp(1+\theta)}\exp(-\beta H).$$

Since θ here corresponds to the normalization constraint, after renaming it as N, it becomes clear that this is the canonical distribution (1.23).

Therefore the microcanonical and canonical distributions are the maximizers of the Shannon entropy under suitable constraints. This can be extended to a general system with more constraints—more information about the system. For example, if one considers a numerical method with some quantities, say, preservation of energy, or any other information, then this information can be used to construct the least biased density consistent with the observations.

The effectiveness of a density derived this way depends on the detail to which known information about the system is included. For example, the energy-enstrophy statistical theory for the quasigeostrophic model (1.12) based on preservation of energy and enstrophy is a model and is incomplete [1], since it takes into account only linear (circulation) and quadratic (enstrophy and energy) invariants, while we know that the quasigeostrophic model preserves an infinite number of Casimirs (1.13).

Another measure of informational content is *relative entropy*, also known as the Kullback-Leibler's distance or the divergence [41]

$$S[\rho,\Pi]=-\int_{\mathbf{K}}\rho\ln\left(\frac{\rho}{\Pi}\right)\,d\boldsymbol{y},\tag{1.27}$$

where Π is a probability density over $\mathbf{K}$ representing an external bias due to some additional information. For example, in the quasigeostrophic model (1.12) there are an infinite number of Casimirs (1.13). This means that formally one has to consider an infinite number of constraints on the entropy. To find a solution to all these constraints might be a difficult or even impossible task. Instead of this, Π can be chosen such that it reflects the Casimirs, and therefore it gives an external bias on the spatial distribution of PV.

It can be shown that S is non-positive, and S is zero only if $\rho\equiv\Pi$ everywhere. This explains the term *distance* for (1.27). The density Π is often called a *prior distribution* and in a typical application Π is given and S is maximized over possible choices of ρ.

1.3.3 Statistical theories for quasigeostrophic flow

One conclusion of this thesis is that numerical methods can give completely different statistical behaviour depending on their conservation properties. In order to discuss statistical theories of numerical discretizations of the quasigeostrophic equation we want to describe, first, several statistical theories for the continuous quasigeostrophic model. Then the reader can more easily follow the statistical theories derived for Arakawa's scheme in Chapter 2 and for the Hamiltonian particle-mesh method in Chapter 3.

Energy-enstrophy statistical theory

The equilibrium statistical mechanical theory for 2D ideal fluids was developed by Kraichnan [40], Salmon et al. [76], and Carnevale & Frederiksen [11]. It is based on a finite truncation of the spectral decomposition of the equations of motion. Statistical predictions are obtained for the truncated system, and these are extended to the infinite dimensional limit. In this section we go through the derivation of energy-enstrophy statistical theory described in [48] for a finite truncation of the quasigeostrophic equation (1.12).

The Fourier space equation (1.15) of the quasigeostrophic model (1.12) is derived through two-dimensional Fourier series. The standard truncation, the Galerkin truncation of the quasigeostrophic equation (1.12) can be interpreted as the Fourier space equation (1.15) limited to the finite $(2M+1)\times(2M+1)$ domain of the Fourier coefficients

$$\frac{d\hat{q}_{\boldsymbol{k}}}{dt} = \sum_{\substack{k_1',k_2'=-M \\ \boldsymbol{k}'\neq 0}}^{M} \frac{\boldsymbol{k}\times\boldsymbol{k}'}{|\boldsymbol{k}'|^2}\hat{q}_{\boldsymbol{k}+\boldsymbol{k}'}(\hat{q}_{-\boldsymbol{k}'}-\hat{h}_{-\boldsymbol{k}'}) \tag{1.28}$$

with coefficients $\hat{q}_{\boldsymbol{k}}$ having period $(2M+1)$ in $\boldsymbol{k}$. The truncated potential vorticity q_M and stream function ψ_M have to satisfy the Poisson equation, the second equation in (1.12) as well

$$\Delta\psi_M = q_M - h_M. \tag{1.29}$$

Only linear and quadratic conserved quantities survive the truncation. Therefore the truncated energy E_M and the truncated enstrophy Z_M are conserved in the finite-dimensionally truncated dynamics

$$E_M = \frac{1}{2}\sum_{\substack{k_1',k_2'=-M \\ \boldsymbol{k}'\neq 0}}^{M} \frac{|\hat{q}_{\boldsymbol{k}}-\hat{h}_{\boldsymbol{k}}|^2}{|\boldsymbol{k}|^2} = \frac{1}{2}\sum_{\substack{k_1',k_2'=-M \\ \boldsymbol{k}'\neq 0}}^{M} |\boldsymbol{k}|^2|\hat{\psi}_{\boldsymbol{k}}|^2, \tag{1.30}$$

$$Z_M = \frac{1}{2}\sum_{\substack{k_1',k_2'=-M \\ \boldsymbol{k}'\neq 0}}^{M} |\hat{q}_{\boldsymbol{k}}|^2 = \frac{1}{2}\sum_{\substack{k_1',k_2'=-M \\ \boldsymbol{k}'\neq 0}}^{M} |-|\boldsymbol{k}|^2\hat{\psi}_{\boldsymbol{k}}+\hat{h}_{\boldsymbol{k}}|^2. \tag{1.31}$$

An (exact) steady state solution of the truncated equation (1.28) is

$$q_M = \mu\psi_M$$

with scalar μ, and it is nonlinearly stable for $\mu > -1$.

The phase flow of (1.28) satisfies the Liouville property. To simplify the index notation we rewrite $\{\hat{q}_{\boldsymbol{k}} : k_1, k_2 \in [-M, M]\} = \{\hat{q}_1, \dots, \hat{q}_\Lambda\}$ for some Λ. Define $\boldsymbol{y} = \{\operatorname{Re}\hat{q}_1, \operatorname{Im}\hat{q}_1, \dots, \operatorname{Re}\hat{q}_\Lambda, \operatorname{Im}\hat{q}_\Lambda\}$, $\boldsymbol{y} \in \mathbb{R}^{2\Lambda} \equiv \mathbb{R}^d$, then (1.28) can be written in a compact form

$$\frac{d\boldsymbol{y}}{dt} = f(\boldsymbol{y}) \quad \text{with} \quad f = (f_1, \dots, f_d)$$

such that $f_{\boldsymbol{k}}$ does not depend on $y_{\boldsymbol{k}}$ ($f_{\boldsymbol{k}}$ depends on $y_{\boldsymbol{k}} \Leftrightarrow$ in (1.28) $\boldsymbol{k}' = 0$, which is excluded from the sum). This immediately implies the Liouville property, since $\operatorname{div}_{\boldsymbol{y}} f = 0$.

Having the Liouville property, we can derive a statistical theory based on preservation of energy and enstrophy. This is called the *energy-enstrophy statistical theory*.

We maximize the Shannon entropy S of (1.24) under the following constraints

- $\rho(\boldsymbol{y}) \geq 0$;
- normalization constraint $\int_{\mathbb{R}^d} \rho(\boldsymbol{y})\, d\boldsymbol{y} = 1$;
- mean energy constraint $\langle E_M \rangle \equiv \int_{\mathbb{R}^d} E_M(\boldsymbol{y})\rho(\boldsymbol{y})\, d\boldsymbol{y} = \mathcal{E}$ with fixed $\mathcal{E}$;
- mean enstrophy constraint $\langle Z_M \rangle \equiv \int_{\mathbb{R}^d} Z_M(\boldsymbol{y})\rho(\boldsymbol{y})\, d\boldsymbol{y} = \mathcal{Z}$ with fixed $\mathcal{Z}$.

The variational principle for finding the maximizer of S under the above constraints gives

$$\mathcal{G}(\boldsymbol{y}) = N^{-1} \exp\left[-\alpha\left(Z_M + \mu E_M\right)\right], \tag{1.32}$$

where μ and α are Lagrange multipliers corresponding to the mean energy and mean enstrophy constraints, respectively, and N corresponds to normalization. This is a Gibbs-like distribution.

The distribution exists if

$$N = \int_{\mathbb{R}^d} \exp\left[-\alpha\left(Z_M + \mu E_M\right)\right] d\boldsymbol{y} < \infty.$$

To guarantee normalization we need to ensure that the coefficients of the quadratic terms in (1.32) with substituted truncated energy (1.30) and truncated enstrophy (1.31), are negative. According to [48], this implies that

$$\alpha > 0 \quad \text{and} \quad \mu > -1.$$

It can be shown that $\mathcal{G}$ is a product of Gaussians, which allows derivation of the mean field equation for the finite-dimensional dynamics described by (1.28)

$$\langle q_M \rangle = \mu \langle \psi_M \rangle .$$

The mean state has to satisfy the Poisson equation (1.29) as well, thus

$$\Delta \langle \psi_M \rangle + h = \mu \langle \psi_M \rangle .$$

To derive the statistical theory for the infinite dimensional QG equation, one has to take the continuum limit as $M \to \infty$ and check whether all predictions hold.

Miller's equilibrium theory

The energy-enstrophy statistical theory takes into account only linear and quadratic invariants. But as we know the quasigeostrophic model preserves higher Casimirs as well. There are statistical equilibrium theories for ideal fluids, which are based on conservation of all Casimirs, derived by Lynden-Bell [47], Robert & Sommeria [71, 72], and Miller [54]. In this section we consider Miller's equilibrium theory. It was originally developed for the Euler equations, which are mathematically equivalent to the quasigeostrophic equation with trivial topography.

For derivation of Miller's theory it is necessary to consider preservation of the Casimir functionals (1.13) as the area preservation of PV levels (1.14). Then the probability density function associated to the PV value σ at a point $\boldsymbol{x} \in \mathcal{D}$ is denoted by $\rho(\sigma, \boldsymbol{x})$. Incompressibility implies the constraint $\int_{-\infty}^{+\infty} d\sigma\, \rho(\sigma, \boldsymbol{x}) = 1$, and area preservation implies the constraint $\int_{\mathcal{D}} d\boldsymbol{x}\, \rho(\sigma, \boldsymbol{x}) = g(\sigma)$. Miller discretized q on a lattice of fine size a, assuming constant values of PV σ_ℓ on each cell ℓ of a fine mesh. Therefore the permutations of PV values $\{\sigma_\ell\}$ form the microscopic configuration space. The macroscopic vorticity field is the local average of the microscopic field on a coarse mesh. Using combinatorial analysis and letting the size of the fine lattice $a \to 0$, Miller maximizes the entropy for ρ under the above constraints to arrive at the distribution

$$\mathcal{G} = \frac{\exp\left[-\beta\sigma\langle\psi\rangle(\boldsymbol{x}) + \mu(\sigma)\right]}{\int_{-\infty}^{+\infty} d\sigma \exp\left[-\beta\sigma\langle\psi\rangle(\boldsymbol{x}) + \mu(\sigma)\right]},$$

where $\langle\psi\rangle$ is the implicitly defined expectation stream function, β determines the energy and $\mu(\sigma)$ are Lagrange multipliers to ensure the area-preservation constraint.

The main obstacle in setting up Miller's statistical theory for a dynamical system is solving the nonlinear relations for $\langle\psi\rangle$.

Prior distribution

In the statistical theories of Lynden-Bell [47], Robert & Sommeria [71, 72], and Miller [54], vorticity invariants are treated microcanonically in the sense that $\mu(\sigma)$ is chosen as a Lagrange multiplier to satisfy constraints on the area distribution function $g(\sigma)$. An alternative approach developed by Ellis, Haven & Turkington [22] treats vorticity invariants canonically, while enforcing the

energy and circulation constraints microcanonically. Canonical treatment of vorticity invariants results in the relative entropy, see (1.27),

$$S[\rho, \Pi] = -\int_{\mathcal{D}} \rho \ln\left(\frac{\rho}{\Pi}\right) \, d\boldsymbol{x},$$

with prior distribution $\Pi(\sigma)$ determined with respect to the vorticity invariants. When a prior distribution is given, the statistical equilibrium state is obtained by maximizing the relative entropy $S[\rho, \Pi]$ at fixed energy, circulation and normalization condition.

The variational principle for finding the maximizer of relative entropy S under the above constraints gives

$$\mathcal{G} = \frac{\Pi(\sigma) \exp\left[(-\beta\langle\psi\rangle(\boldsymbol{x}) + \alpha)\sigma\right]}{\int_{-\infty}^{+\infty} d\sigma \Pi(\sigma) \exp\left[(-\beta\langle\psi\rangle(\boldsymbol{x}) + \alpha)\sigma\right]},$$

where β and α are Lagrange multipliers corresponding to the main energy and circulation constraints, respectively.

Note an equivalence of Miller's and Ellis, Haven & Turkington's distributions, when $e^{\mu(\sigma)} = e^{\alpha\sigma}\Pi$, i.e. $\mu(\sigma) = \ln\left(e^{\alpha\sigma}\Pi\right)$.

Miller's statistical theory is based on the assumption that the flow is described by 2D Euler equations, without forcing and dissipation. However, in geophysical situations, the flow is forced and dissipated at small scales, which destroys the conservation of Casimirs. Ellis, Haven & Turkington's equilibrium theory takes care of this situation by fixing the prior distribution instead of vorticity invariants, since it is more easy to determine the prior distribution from data than to determine higher moments of the atmospheric vorticity accurately. For more comparison between Miller's and Ellis, Haven & Turkington's statistical theories see Chavanis [13].

Overview of thesis

In Chapter 2, we will consider quasigeostrophic flow with topographic forcing. We will construct statistical mechanical theories for the discrete dynamical systems arising from three discretizations due to Arakawa [2] which conserve energy, enstrophy or both. The referees to Sections 1.1 and 1.3 could be helpful for the reader.

In Chapter 3, we will consider the Hamiltonian particle-mesh method for quasigeostrophic flow over topography. We will propose Lagrangian and Eulerian statistical models for the discrete dynamics. The referees to Sections 1.2 and 1.3 could be helpful for the reader.

In Chapter 4, we will consider the point vortex flow on a disk. We will present a closure for incompressible ideal fluid flow in the form of the mathematical thermostat. The referees to Section 1.3 could be helpful for the reader.

Chapter 2

Statistical mechanics of Arakawa's discretizations

2.1 Introduction

In applications such as weather and climate predictions, long numerical simulations are run for dynamical systems that are known to be chaotic, and for which it is consequently impossible to simulate a particular solution with any accuracy in the usual sense of numerical analysis. Instead, the goal of such simulations is to obtain a data set suitable for computing statistical averages or otherwise to sample the probability distribution associated with the continuous problem.

Different numerical discretizations have very different discrete dynamics, however. Recent work on geometric integration [35, 45, 79, 84] relies on backward error analysis, in which the numerical solution generated by a given method is viewed as the exact solution of a perturbed problem. The properties of different discrete dynamics become more pronounced when the numerical map is iterated over a very large number of time steps. Therefore it is important to establish the influence that a particular choice of method has on the statistical results obtained from simulations. Ideally, one would like to determine criteria which a method should satisfy to yield meaningful statistics, and to understand statistical accuracy in terms of discretization parameters.

To that end, in this chapter we consider three related discretizations for an ideal fluid in vorticity-stream function form, originally proposed by Arakawa [2]. The three discretizations conserve discrete approximations of energy, enstrophy, or both. We analyze the three methods through appropriate (trivial) modifications of the statistical mechanics theory of quasigeostrophic flow over topography—based on the original work of Kraichnan [40], Salmon et al. [76] and Carnevale & Frederiksen [11], and recently expounded in Majda & Wang [48]. The resulting theories predict entirely different statistical behavior for the three methods. Numerical experiments with conservative and projected time integrators agree with the statistical predictions, confirming that the conservation properties of a discretization define the backdrop, or climatic mean, against which the dynamics takes place.

It should be mentioned at the outset that the energy-enstrophy statistical

theory is a model and is known to be incomplete. In [1], Abramov & Majda show that nonzero values of the third moment of potential vorticity can cause significant deviation from the statistical predictions. In Section 2.6 we use the numerical setup of [1] to facilitate comparison with their results. We wish to stress, however, that the focus of this thesis is not the statistical mechanics of ideal fluids *per se*, but rather the application of statistical mechanics as a tool for the numerical analysis of discretizations.

The chapter is organized as follows. In Section 2.2 we briefly recall the quasigeostrophic potential vorticity equation and its conservation properties. Section 2.3 we review Arakawa's discretizations, their conservation properties, and prove that all of these define divergence-free vector fields. In Section 2.4, the equilibrium statistical mechanical theories are developed for the three discretizations. Most of this section is simply a summary of material in Chapters 7 and 8 of [48] for the energy-enstrophy theory. Once established, it is a simple matter to extend the results to the cases in which only one of these quantities is conserved, and we do this in Section 2.4.4. Time integration aspects are discussed in Section 2.5. The numerical experiments confirming the statistical predictions are presented in Section 2.6.

2.2 The quasigeostrophic model

This section addresses the statistical mechanics of conservative discretizations of the quasigeostrophic potential vorticity model (QG) on a doubly periodic domain, $\Omega = \{\boldsymbol{x} = (x, y) \,|\, x, y \in [0, 2\pi)\}$. The QG equation [67, 74] is

$$q_t = \mathcal{J}(q, \psi), \tag{2.1a}$$

$$\Delta\psi = q - h, \tag{2.1b}$$

where the potential vorticity (PV) $q(\boldsymbol{x}, t)$, the stream function $\psi(\boldsymbol{x}, t)$, and the orography $h(\boldsymbol{x})$ are scalar fields, periodic in x and y with period 2π. The Laplace operator is denoted by Δ, and the operator $\mathcal{J}$ is defined by

$$\mathcal{J}(q, \psi) = q_x\psi_y - q_y\psi_x. \tag{2.2}$$

The QG equation is a Hamiltonian PDE [57] having Poisson bracket

$$\{\mathcal{F}, \mathcal{G}\} = \int_{\mathcal{D}} q\mathcal{J}\left(\frac{\delta\mathcal{F}}{\delta q}, \frac{\delta\mathcal{G}}{\delta q}\right) d\boldsymbol{x} \tag{2.3}$$

and Hamiltonian functional

$$\mathcal{E}[q] = -\frac{1}{2}\int_{\mathcal{D}} \psi \cdot (q - h)\, d\boldsymbol{x}. \tag{2.4}$$

The Poisson bracket is degenerate with Casimir invariants the generalized enstrophies $\mathcal{C}[f] = \int_{\mathcal{D}} f(q)\, d\boldsymbol{x}$ for arbitrary function f, see Section 1.1. Of particular interest are the PV moments $\mathcal{C}_r = \int_{\mathcal{D}} q^r\, d\boldsymbol{x}$, $r = 1, 2, \ldots$ The most

important of these are the total circulation

$$\mathcal{C} \equiv \mathcal{C}_1[q] = \int_{\mathcal{D}} q \, d\boldsymbol{x}. \tag{2.5}$$

and the second moment of vorticity, i.e. the enstrophy

$$\mathcal{Z} \equiv \frac{1}{2}\mathcal{C}_2[q] = \frac{1}{2}\int_{\mathcal{D}} q^2 \, d\boldsymbol{x}. \tag{2.6}$$

2.3 Spatial semi-discretization

We first consider the discretization of (2.1) in space only. The resulting system of ordinary differential equations will be referred to as the semi-discretization, and we will primarily be concerned with its analysis and statistical mechanics.

When discretizing Hamiltonian PDEs, it is advisable to consider the discretizations of the Poisson bracket and the Hamiltonian separately. As noted in [51], if a discrete Poisson bracket can be constructed to maintain skew-symmetry and satisfy the Jacobi identity, then any quadrature for the Hamiltonian will yield a semi-discretization that is a Hamiltonian ODE, and consequently will conserve energy and (possibly some subclass of) Casimirs. From the point of view of statistical mechanics, it is *also* natural to consider the discretizations of the bracket and the Hamiltonian separately. The bracket ensures the conservation of energy and enstrophy and preservation of volume, which are necessary ingredients for the *existence* of a statistical theory at all. But only the conserved quantities themselves enter into the probability distribution. Thus the *predictions* of the theory depend only on the discretization of these conserved quantities. The discretization of the Hamiltonian (2.4) amounts to a choice for the discrete Laplacian in (2.1b) and will be treated latter. The bracket will be discretized with (generalized) Arakawa schemes in Section 2.3.1.

For Eulerian fluid models, the only known discretization with Poisson structure is the sine-bracket truncation of Zeitlin [88], which is limited to 2D, incompressible flows on periodic geometry, see Section 1.2. This truncation conserves M polynomial enstrophies on an $M \times M$ grid. Its statistics are investigated in [1]. For more general fluid problems, no Poisson discretizations are available. In lieu of a semi-discretization with Poisson structure, one may attempt to construct discretizations which conserve desired first integrals and are volume preserving. The flow of energy is important for statistics, and the spatial discretization determines the local flow. In numerical weather prediction, energy conserving discretizations were advocated by Lorenz in 1960 [46]. Motivated by Lorenz's work, Arakawa [2] constructed discretizations that conserved energy, enstrophy or both. As we will see, these discretizations are also all volume preserving.

We discretize (2.1) on a uniform $M \times M$ grid. Let $\Delta x = \Delta y = 2\pi/M$ and consider a grid function $\boldsymbol{q}(t) \in \mathbb{R}^{M\times M}$, with components $q_{i,j}(t) \approx q(i\Delta x, j\Delta y, t)$, $i, j = 0, \ldots, M-1$, where periodicity is realized by identifying the indices M

and 0. We think of $\boldsymbol{q}$ as a vector in an M^2-dimensional phase space; that is, we identify $\mathbb{R}^{M^2}$ and $\mathbb{R}^{M\times M}$, and use vector notation, e.g., $\boldsymbol{\Psi}^T\boldsymbol{q}$ for the vector inner product of two such vectors.

Spectral solution of the stream function

The linear elliptic PDE (2.1b) is solved using the Fourier spectral method. Let the Fourier transform of $\boldsymbol{q}\in\mathbb{R}^{M\times M}$ be defined by

$$\hat{\boldsymbol{q}}=\mathcal{F}\boldsymbol{q} \quad\Longleftrightarrow\quad \hat{q}_{k,\ell}=\frac{1}{M}\sum_{i,j=0}^{M-1} q_{i,j}e^{-\mathrm{i}(ik+j\ell)},\quad k,\ell=-M/2+1,\dots,M/2. \tag{2.7}$$

The inverse transform is $\mathcal{F}^{-1}=\mathcal{F}^*$, and Parseval's identity reads

$$\sum_{i,j} q_{i,j}^2=\sum_{k,\ell}|\hat{q}_{k,\ell}|^2.$$

Equation (2.1b) is solved exactly in Fourier-space. Denote the discrete Laplace operator by Δ_M:

$$\Delta_M\boldsymbol{\psi}=\boldsymbol{q}-\boldsymbol{h} \iff -(k^2+\ell^2)\hat{\psi}_{k,\ell}=\hat{q}_{k,\ell}-\hat{h}_{k,\ell},\quad k,\ell=-M/2+1,\dots,M/2. \tag{2.8}$$

This relation is solved for stream function field $\boldsymbol{\psi}$ with mean zero. The inverse Laplacian restricted to the hyperplane $\hat{\psi}_{0,0}\equiv 0$ is denoted by Δ_M^{-1}, i.e.

$$\boldsymbol{\psi}=\Delta_M^{-1}(\boldsymbol{q}-\boldsymbol{h}) \quad\Longleftrightarrow\quad \hat{\psi}_{k,\ell}=\begin{cases} 0, & k=\ell=0,\\ -(\hat{q}_{k,\ell}-\hat{h}_{k,\ell})/(k^2+\ell^2), & \text{otherwise.}\end{cases}$$

2.3.1 Arakawa's discretizations

Arakawa [2] constructed finite difference discretizations of (2.2) that preserve discrete versions of energy (2.4), enstrophy (2.6), or both. We consider generalizations to Arakawa's discretizations with the Nambu bracket approach of [75].

Let D_x and D_y denote discretization matrices that (i) are skew symmetric: $D_x^T=-D_x$, $D_y^T=-D_y$, and (ii) approximate the first derivative in x and y, respectively:

$$(D_x\boldsymbol{q})_{i,j}\approx q_x(i\Delta x,j\Delta y),\qquad (D_y\boldsymbol{q})_{i,j}\approx q_y(i\Delta x,j\Delta y).$$

Arakawa's classical discretizations [2] use the central differences

$$(D_x\boldsymbol{q})_{i,j}=\frac{q_{i+1,j}-q_{i-1,j}}{2\Delta x},\qquad (D_y\boldsymbol{q})_{i,j}=\frac{q_{i,j+1}-q_{i,j-1}}{2\Delta y}, \tag{2.9}$$

and these will also be used in our numerical experiments. However, we wish to stress that the statistical predictions in Section 2.4 remain unchanged for a different choice of (skew-symmetric) D_x and D_y.

Denote the element-wise product of two vectors by $(\boldsymbol{u} * \boldsymbol{v})_{i,j} = u_{i,j} v_{i,j}$. The scalar product

$$\boldsymbol{u}^T (\boldsymbol{v} * \boldsymbol{w}) = \sum_{i,j} u_{i,j} v_{i,j} w_{i,j} \tag{2.10}$$

is fully symmetric with respect to the vectors $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$.

Arakawa's discretizations can be viewed as discrete approximations to the equivalent formulations of (2.2)

$$\begin{aligned}
\mathcal{J}(q,\psi) &= q_x \psi_y - q_y \psi_x, \\
\mathcal{J}(q,\psi) &= \partial_x (q \psi_y) - \partial_y (q \psi_x), \\
\mathcal{J}(q,\psi) &= \partial_y (q_x \psi) - \partial_x (q_y \psi),
\end{aligned}$$

and are given by

$$J_0(\boldsymbol{q},\boldsymbol{\psi}) = (D_x \boldsymbol{q}) * (D_y \boldsymbol{\psi}) - (D_y \boldsymbol{q}) * (D_x \boldsymbol{\psi}), \tag{2.11}$$

$$J_E(\boldsymbol{q},\boldsymbol{\psi}) = D_x (\boldsymbol{q} * D_y \boldsymbol{\psi}) - D_y (\boldsymbol{q} * D_x \boldsymbol{\psi}), \tag{2.12}$$

$$J_Z(\boldsymbol{q},\boldsymbol{\psi}) = D_y (\boldsymbol{\psi} * D_x \boldsymbol{q}) - D_x (\boldsymbol{\psi} * D_y \boldsymbol{q}), \tag{2.13}$$

and the average of these

$$J_{EZ}(\boldsymbol{q},\boldsymbol{\psi}) = \frac{1}{3} \left[J_0(\boldsymbol{q},\boldsymbol{\psi}) + J_E(\boldsymbol{q},\boldsymbol{\psi}) + J_Z(\boldsymbol{q},\boldsymbol{\psi}) \right]. \tag{2.14}$$

That is, the semi-discretizations are defined by (2.8) and

$$\frac{d}{dt} \boldsymbol{q} = J(\boldsymbol{q},\boldsymbol{\psi}) \tag{2.15}$$

for J taken to be one of (2.11)–(2.14).

The Arakawa schemes are interesting for us, because they are all based on the standard central difference operators applied in various 'conservation forms' and hence, for short simulations with smooth solutions, there is often little noticeable difference between different discretizations. One might therefore expect that they yield similar statistics. On the contrary, the long-term statistics differ greatly.

The conservation properties of these three discretizations were established for the case of second order differences (2.9) in [2]. The case (2.14) has been generalized using the Nambu bracket formalism [59, 60, 75]. Define the associated bracket (the gradients are with respect to $\boldsymbol{q}$)

$$\{F,G,H\}_0 = -\nabla F^T J_0(\nabla G, \nabla H), \tag{2.16}$$

$$\{F,G,H\}_E = -\nabla F^T J_E(\nabla G, \nabla H), \tag{2.17}$$

$$\{F,G,H\}_Z = -\nabla F^T J_Z(\nabla G, \nabla H), \tag{2.18}$$

$$\{F,G,H\}_{EZ} = \frac{1}{3} \left(\{F,G,H\}_0 + \{F,G,H\}_E + \{F,G,H\}_Z \right) \tag{2.19}$$

for arbitrary differentiable $F(\boldsymbol{q})$, $G(\boldsymbol{q})$, $H(\boldsymbol{q}) : \mathbb{R}^{M^2} \to \mathbb{R}$.

The derivative dF/dt of any function $F(\boldsymbol{q})$ along a solution $\boldsymbol{q}(t)$ to the discrete equations (2.15) is given by the associated bracket of F with Z_M and E_M:

$$\frac{dF}{dt} = \{F, Z_M, E_M\}. \tag{2.20}$$

where E_M and Z_M are discrete approximations to the energy

$$E_M(\boldsymbol{q}) = -\frac{1}{2}\boldsymbol{\psi}^T(\boldsymbol{q} - \boldsymbol{h})\Delta x\Delta y = \frac{1}{2}\sum_{k,\ell}(k^2 + \ell^2)|\hat{\psi}_{k,\ell}|^2\Delta x\Delta y \tag{2.21}$$

and enstrophy

$$Z_M(\boldsymbol{q}) = \frac{1}{2}\boldsymbol{q}^T\boldsymbol{q}\Delta x\Delta y = \frac{1}{2}\sum_{k,\ell}|\hat{q}_{k,\ell}|^2\Delta x\Delta y. \tag{2.22}$$

This fact can be used to prove the conservation properties of the various discretizations.

The proofs rely on the antisymmetry of (2.16) with respect to its last two arguments,

$$\{F, G, H\}_0 = -\{F, H, G\}_0,$$

as well as the identities

$$\{F, G, H\}_E = \{G, H, F\}_0, \qquad \{F, G, H\}_Z = \{H, F, G\}_0,$$

all of which follow from the skew-symmetry of D_x and D_y and the symmetry of (2.10).

Taking $F \equiv E_M$ in (2.20), it follows that for J_E,

$$\frac{dE_M}{dt} = \{E_M, Z_M, E_M\}_E = \{Z_M, E_M, E_M\}_0 = 0.$$

Similarly, taking $F \equiv Z_M$ in (2.20), it follows that for J_Z,

$$\frac{dZ_M}{dt} = \{Z_M, Z_M, E_M\}_Z = \{E_M, Z_M, Z_M\}_0 = 0.$$

The bracket (2.19) is fully antisymmetric in all three arguments (hence it is a proper Nambu bracket), and therefore conserves both E_M and Z_M. Finally, taking $F = C_M = \sum_{i,j} q_{i,j}\,\Delta x\,\Delta y$, one can show that all of the discretizations (2.16)–(2.19) conserve total circulation.

In reference to their conservation properties, we will refer to the discretizations (2.11)-(2.14) as the 0, E, Z and EZ discretizations, respectively.

One can check that a solution of the form $\boldsymbol{q} = \mu\boldsymbol{\psi}$, μ a scalar, is an exact steady state for the 0 and EZ discretizations. Such a solution is not, in general, a steady state solution for the E and Z discretizations. However, the limit cases $\{\boldsymbol{\psi} \equiv 0, \boldsymbol{q} = \boldsymbol{h}\}$ and $\{\boldsymbol{q} \equiv 0, \boldsymbol{\psi} = -\Delta_M^{-1}\boldsymbol{h}\}$ obviously *are* steady states to those discretizations.

2.3.2 Volume preservation

In addition to conservation, a second important ingredient for statistical mechanics is the preservation, by the flow map, of the phase space volume element. In this section we demonstrate that each of the discretizations from Section 2.3.1 is volume preserving. Let us define the matrix $D(\boldsymbol{a}) = \mathrm{diag}(\boldsymbol{a})$ to be the diagonal matrix whose diagonal elements are the components of the vector $\boldsymbol{a}$ (i.e. $D(\boldsymbol{a})_{ij} = a_i \delta_{ij}$).

Recall that for an ODE

$$\boldsymbol{y}' = \boldsymbol{f}(\boldsymbol{y})$$

the divergence of the vector field $\boldsymbol{f}$ satisfies

$$\mathrm{div}\ \boldsymbol{f} = \mathrm{tr}(\boldsymbol{f}'),$$

where $\boldsymbol{f}'$ denotes the Jacobian matrix of $\boldsymbol{f}$. In particular, for a matrix A, $\mathrm{div} A\boldsymbol{f}(\boldsymbol{y}) = \mathrm{tr}(A\boldsymbol{f}')$. Furthermore, for

$$\boldsymbol{y}' = \boldsymbol{f}(\boldsymbol{y}) = \boldsymbol{g}(\boldsymbol{y}) * \boldsymbol{h}(\boldsymbol{y})$$

it holds that

$$\boldsymbol{f}' = D(\boldsymbol{g})\boldsymbol{h}' + D(\boldsymbol{h})\boldsymbol{g}'.$$

In the following calculations we make ready use of the commutative and transpose properties of the trace $\mathrm{tr}(AB) = \mathrm{tr}(BA) = \mathrm{tr}(B^T A^T)$. We also need the following properties of our discretization matrices. The difference operators D_x and D_y are given by symmetric finite difference stencils, are skew-symmetric and commute $D_x D_y - D_y D_x = 0$. The discrete inverse Laplacian matrix Δ_M^{-1} is symmetric and represents a (global) central finite difference stencil. In this case, the matrices $D_x \Delta_M^{-1}$ and $D_y \Delta_M^{-1}$ have zeros on the diagonal.

Let us write the discretizations (2.11)–(2.13) as functions of $\boldsymbol{q}$ only

$$J_0(\boldsymbol{q}) = (D_x \boldsymbol{q}) * (D_y \Delta_M^{-1} \boldsymbol{q}) - (D_y \boldsymbol{q}) * (D_x \Delta_M^{-1} \boldsymbol{q}), \tag{2.23}$$

$$J_E(\boldsymbol{q}) = D_x(\boldsymbol{q} * D_y \Delta_M^{-1} \boldsymbol{q}) - D_y(\boldsymbol{q} * D_x \Delta_M^{-1} \boldsymbol{q}), \tag{2.24}$$

$$J_Z(\boldsymbol{q}) = D_y\left((D_x \boldsymbol{q}) * \Delta_M^{-1} \boldsymbol{q}\right) - D_x\left((D_y \boldsymbol{q}) * \Delta_M^{-1} \boldsymbol{q}\right). \tag{2.25}$$

Proposition 2.1. *The vector fields defined by (2.23)–(2.25) and their average $J_{EZ} = (J_0 + J_E + J_Z)/3$ are divergence free.*

Proof. We calculate, for (2.23),

$$\begin{aligned}\mathrm{div}\ J_0(\boldsymbol{q}) = {} & \mathrm{tr}\left(D(D_y \Delta_M^{-1} \boldsymbol{q}) D_x\right) + \mathrm{tr}\left(D(D_x \boldsymbol{q}) D_y \Delta_M^{-1}\right) \\ & - \mathrm{tr}\left(D(D_x \Delta_M^{-1} \boldsymbol{q}) D_y\right) - \mathrm{tr}\left(D(D_y \boldsymbol{q}) D_x \Delta_M^{-1}\right) = 0, \end{aligned} \tag{2.26}$$

since each term is the trace of the product of a diagonal matrix and a matrix with zero diagonal.

For (2.24),

$$\begin{aligned}\text{div } J_E(\boldsymbol{q}) &= \text{tr}\left(D_x\left[D(\boldsymbol{q})D_y\Delta_M^{-1} + D(D_y\Delta_M^{-1}\boldsymbol{q})\right]\right)\\ &\quad - \text{tr}\left(D_y\left[D(\boldsymbol{q})D_x\Delta_M^{-1} + D(D_x\Delta_M^{-1}\boldsymbol{q})\right]\right)\\ &= \text{tr}(D(\boldsymbol{q})[D_y\Delta_M^{-1}D_x - D_x\Delta_M^{-1}D_y])\\ &\quad + \text{tr}\left(D_xD(D_y\Delta_M^{-1}\boldsymbol{q}) - D_yD(D_x\Delta_M^{-1}\boldsymbol{q})\right) = 0.\end{aligned}$$

The term in bracket in the last expression is identically zero by symmetry considerations.

Similarly, for (2.25) we have

$$\begin{aligned}\text{div } J_Z(\boldsymbol{q}) &= \text{tr}\left(D_y\left[D(\Delta_M^{-1}\boldsymbol{q})D_x + D(D_x\boldsymbol{q})\Delta_M^{-1}\right]\right)\\ &\quad - \text{tr}\left(D_x\left[D(\Delta_M^{-1}\boldsymbol{q})D_y + D(D_x\boldsymbol{q})\Delta_M^{-1}\right]\right)\\ &= \text{tr}\left(D(\Delta_M^{-1}\boldsymbol{q})\left[D_xD_y - D_yD_x\right]\right)\\ &\quad + \text{tr}\left(D(D_x\boldsymbol{q})\Delta_M^{-1}D_y - D(D_y\boldsymbol{q})\Delta_M^{-1}D_y\right) = 0.\end{aligned}$$

Finally, discretization EZ is divergence-free because it is a linear combination of divergence-free vector fields. □

2.4 Energy-enstrophy statistical theory

The equilibrium statistical mechanical theory for 2D ideal fluids was developed by Kraichnan [40], Salmon et al. [76], and Carnevale & Frederiksen [11]. It is based on a finite truncation of the spectral decomposition of the equations of motion. Statistical predictions are obtained for the truncated system, and these are extended to the infinite dimensional limit, see Section 1.3.3.

Here we would like to adapt the analysis to the semi-discretizations outlined in the previous section. For the discretization EZ, which conserves both energy and enstrophy, the analysis is identical to the spectral case developed by Carnevale & Frederiksen [11]. Consequently, most of the material in Sections 2.4.1, 2.4.2 and 2.4.3 is simply summarized from Chapters 7 and 8 of Majda & Wang [48]. In Section 2.4.4 we modify the statistical predictions of the energy-enstrophy theory to the cases of only one quantity conserved.

As previously noted, semi-discretization of (2.1) using the bracket (2.14) yields a system of M^2 ordinary differential equations having the Liouville property and two first integrals that approximate the energy (2.21) and enstrophy (2.22).[1] Due to the Liouville property, one can speak of transport of probability density functions by this semi-discrete flow, and consider equilibrium solutions to Liouville's equation. Any normalized function of the two first integrals is an equilibrium distribution.

[1] All discretizations also conserve the discrete total circulation $C_M = \sum_{i,j} q_{i,j}\Delta x\Delta y$. Since C_M is a linear first integral, any standard integrator will conserve it exactly in time. A nonzero value of C_M will give a constant displacement in (2.32). For a periodic domain one may assume that $C_M = 0$ so that its effects can be ignored. We do so in the numerical experiments.

We note in advance that that a solution of the semidiscrete equations (2.15) is constrained to the intersection of hypersurfaces defined by the relevant first integrals of the the discretization. The probability distributions obtained from the maximum entropy theory have nonzero probability everywhere in phase space, and as such, are a very crude approximation to the statistics of a single trajectory. Nonetheless, we will see that the maximum entropy theory accurately predicts the differences in long term averages observed for the discretizations (2.12)–(2.14). More on this will be said in Section 2.5.

2.4.1 Mean field predictions

The equilibrium distribution of least bias maximizes entropy under the constraints imposed by conservation of energy and enstrophy. Let $\boldsymbol{y}$ parameterize the M^2 dimensional phase space; that is, each $\boldsymbol{y} \in \mathbb{R}^{M\times M}$ corresponds to a particular realization of the grid function (or discrete field) $\boldsymbol{q}$. Consider the class of probability distribution $\rho : \mathbb{R}^{M\times M} \to \mathbb{R}$ on phase space, satisfying

$$\rho(\boldsymbol{y}) \geq 0, \qquad \int_{\mathbb{R}^{M\times M}} \rho(\boldsymbol{y})\, d\boldsymbol{y} = 1. \tag{2.27}$$

The least biased distribution ρ^* maximizes the entropy functional

$$\mathcal{S}\left[\rho\right] = -\int_{\mathbb{R}^{M\times M}} \rho(\boldsymbol{y}) \ln \rho(\boldsymbol{y})\, d\boldsymbol{y} \tag{2.28}$$

under constraints on the ensemble averages of energy:

$$\int_{\mathbb{R}^{M\times M}} E_M(\boldsymbol{y})\rho(\boldsymbol{y})\, d\boldsymbol{y} - E_M^* = 0, \tag{2.29}$$

and enstrophy:

$$\int_{\mathbb{R}^{M\times M}} Z_M(\boldsymbol{y})\rho(\boldsymbol{y})\, d\boldsymbol{y} - Z_M^* = 0, \tag{2.30}$$

where E_M^* and Z_M^* are prescribed values. Additionally, there is the constraint implied by (2.27). Using the method of Lagrange multipliers, the maximizer is the Gibbs-like distribution (i.e. $\rho^* = \mathcal{G}$)

$$\mathcal{G}(\boldsymbol{y}) = N^{-1} \exp\left[-\alpha\left(Z_M(\boldsymbol{y}) + \mu E_M(\boldsymbol{y})\right)\right], \tag{2.31}$$

where N, α and μ are chosen to ensure (2.27), (2.29) and (2.30).

The expected value of a function $F(\boldsymbol{y})$ is the ensemble average of F with the measure $\mathcal{G}$, denoted

$$\langle F \rangle = \int_{\mathbb{R}^{M\times M}} F(\boldsymbol{y})\mathcal{G}(\boldsymbol{y})\, d\boldsymbol{y}.$$

The mean state is obtained from the observation

$$\begin{aligned}\left\langle \frac{\partial Z_M}{\partial \boldsymbol{y}} + \mu \frac{\partial E_M}{\partial \boldsymbol{y}} \right\rangle &= \int_{\mathbb{R}^{M\times M}} \left(\frac{\partial Z_M}{\partial \boldsymbol{y}} + \mu \frac{\partial E_M}{\partial \boldsymbol{y}} \right) N^{-1} e^{-\alpha(Z_M(\boldsymbol{y}) + \mu E_M(\boldsymbol{y}))}\, d\boldsymbol{y} \\ &= -\alpha^{-1} \int_{\mathbb{R}^{M\times M}} \frac{\partial}{\partial \boldsymbol{y}} \mathcal{G}(\boldsymbol{y})\, d\boldsymbol{y} = 0,\end{aligned}$$

assuming $\mathcal{G}$ decays sufficiently fast at infinity. Since $\nabla_{\boldsymbol{q}} E_M = -\boldsymbol{\psi}$ and $\nabla_{\boldsymbol{q}} Z_M = \boldsymbol{q}$, the mean field relation

$$\langle \boldsymbol{q} \rangle = \mu \langle \boldsymbol{\psi} \rangle \tag{2.32}$$

follows. In other words, the ensemble averages of potential vorticity and stream function are linearly related. Combining (2.32) with the second relation of (2.1) yields a modified Helmholtz problem for the mean stream function given μ:

$$(\mu - \Delta_M)\langle \boldsymbol{\psi} \rangle = \boldsymbol{h}. \tag{2.33}$$

2.4.2 PV fluctuation predictions

In this section we adapt the point statistics of Majda & Wang [48] to yield predictions in terms of potential vorticity. The mean state (2.32) is a nonlinearly stable equilibrium [11]. Solutions to (2.1) may be decomposed into mean and fluctuation parts

$$\boldsymbol{q} = \langle \boldsymbol{q} \rangle + \boldsymbol{q}', \qquad \boldsymbol{\psi} = \langle \boldsymbol{\psi} \rangle + \boldsymbol{\psi}', \qquad \langle \boldsymbol{q} \rangle = \mu \langle \boldsymbol{\psi} \rangle.$$

The fluctuation quantities satisfy

$$\boldsymbol{q}'_t = J(\langle \boldsymbol{q} \rangle, \boldsymbol{\psi}') + J(\boldsymbol{q}', \langle \boldsymbol{\psi} \rangle) + J(\boldsymbol{q}', \boldsymbol{\psi}'), \qquad \Delta_M \boldsymbol{\psi}' = \boldsymbol{q}'. \tag{2.34}$$

This differential equation has the first integral

$$I_M(\boldsymbol{q}') = Z'_M + \mu E'_M, \qquad Z'_M = \frac{1}{2}(\boldsymbol{q}')^T \boldsymbol{q}' \Delta x \Delta y, \quad E'_M = -\frac{1}{2}(\boldsymbol{\psi}')^T \boldsymbol{q}' \Delta x \Delta y. \tag{2.35}$$

One can also set up a statistical mechanics for the fluctuation equations and obtain predictions. To do so, let

$$\hat{p}_{k,\ell} = \left(1 + \frac{\mu}{k^2 + \ell^2}\right)^{1/2} \hat{q}'_{k,\ell}. \tag{2.36}$$

Then the Fourier transform of (2.35) gives

$$I_M = \frac{1}{2} \sum_{k,\ell} \left(1 + \frac{\mu}{k^2 + \ell^2}\right) |\hat{q}'_{k,\ell}|^2 \Delta x \Delta y = \frac{1}{2} \sum_{k,\ell} |\hat{p}_{k,\ell}|^2 \Delta x \Delta y = \frac{1}{2} \sum_{i,j} p_{i,j}^2 \Delta x \Delta y. \tag{2.37}$$

The maximum entropy condition for this first integral yields the Gibbs distribution $\mathcal{G}(\boldsymbol{p}) = N^{-1} \exp\left[-\beta I_M(\boldsymbol{p})\right]$, which is the product of identical Gaussian distributions with mean zero and standard deviation

$$\eta_p = \sqrt{\frac{2\langle I_M \rangle}{M^2 \Delta x \Delta y}} = \sqrt{\frac{\langle I_M \rangle}{2\pi^2}}.$$

The energy is equipartitioned.

Let us also assume that the $p_{i,j}$ are independent. Let $P = \boldsymbol{a}^T\boldsymbol{p}$ denote a linear combination of the $p_{i,j}$. Since these are identically distributed, P is Gaussian with variation

$$\eta(P)^2 = \boldsymbol{a}^T\boldsymbol{a}\,\eta_p^2 = |\boldsymbol{a}|^2\eta_p^2.$$

From (2.36) we have

$$\boldsymbol{q}' = \mathcal{F}^{-1}\,\mathrm{diag}\left(\left(1+\frac{\mu}{k^2+\ell^2}\right)^{-1/2}\right)\mathcal{F}\boldsymbol{p} = A\boldsymbol{p},$$

where A is real and symmetric. It follows that the $q'_{i,j}$ at each grid point i,j are identically normally distributed with mean zero and variance

$$\eta_q^2 = |\boldsymbol{a}|^2\eta_p^2 = |\boldsymbol{a}|^2\frac{\langle I_M\rangle}{2\pi^2}, \tag{2.38}$$

where for $\boldsymbol{a}$ we can take any row of A.

2.4.3 Approximation of μ and α

The ensemble averages of energy and enstrophy can be split into a mean part and a fluctuation part [11, 48]:

$$\langle E_M\rangle = E_M(\langle\boldsymbol{q}\rangle) + E'_M, \qquad \langle Z_M\rangle = Z_M(\langle\boldsymbol{q}\rangle) + Z'_M, \tag{2.39}$$

where, using (2.33),

$$E_M(\langle\boldsymbol{q}\rangle) = -\frac{1}{2}\langle\boldsymbol{\psi}\rangle^T(\langle\boldsymbol{q}\rangle - \boldsymbol{h})\Delta x\Delta y = \frac{1}{2}\sum_{k,\ell=-M/2+1}^{M/2}\frac{(k^2+\ell^2)|\hat{h}_{k,\ell}|^2}{(\mu+k^2+\ell^2)^2}\Delta x\Delta y, \tag{2.40a}$$

$$E'_M = \frac{1}{2\alpha}\sum_{k,\ell=-M/2+1}^{M/2}\frac{1}{\mu+k^2+\ell^2}, \tag{2.40b}$$

and

$$Z_M(\langle\boldsymbol{q}\rangle) = \frac{1}{2}\langle\boldsymbol{q}\rangle^T\langle\boldsymbol{q}\rangle\Delta x\Delta y = \frac{1}{2}\sum_{k,\ell=-M/2+1}^{M/2}\frac{\mu^2|\hat{h}_{k,\ell}|^2}{(\mu+k^2+\ell^2)^2}\Delta x\Delta y, \tag{2.41a}$$

$$Z'_M = \frac{1}{2\alpha}\sum_{k,\ell=-M/2+1}^{M/2}\frac{k^2+\ell^2}{\mu+k^2+\ell^2}. \tag{2.41b}$$

Given guesses for μ and α, it is straightforward to compute $\langle E_M\rangle$ and $\langle Z_M\rangle$ by solving (2.40) and (2.41) and then substituting into (2.39). To estimate μ and α, we proceed iteratively to implicitly solve (2.29) and (2.30) under that assumptions $E_M^* \approx \mathcal{E}_0$ and $Z_M^* \approx \mathcal{Z}_0$.

2.4.4 Alternative statistical theories

In this section we derive alternative statistical models for the cases where either energy or enstrophy, but not both, is conserved numerically.

Energy-based statistical mechanics

For a semi-discretization that only preserves the energy E_M, the least biased distribution (2.31) becomes

$$\mathcal{G}_E(\boldsymbol{y}) = N^{-1} \exp\left[-\lambda E_M(\boldsymbol{y})\right].$$

The mean field prediction (2.32) gives

$$\langle \boldsymbol{\psi} \rangle \equiv 0, \qquad \langle \boldsymbol{q} \rangle = \boldsymbol{h}. \tag{2.42}$$

The fluctuation dynamics (2.34) becomes

$$\boldsymbol{q}'_t = J_E(\boldsymbol{h} + \boldsymbol{q}', \boldsymbol{\psi'}), \qquad \boldsymbol{\psi}' = \Delta_M^{-1} \boldsymbol{q}',$$

which preserves the pseudo-energy

$$I_M = -\frac{1}{2} (\boldsymbol{\psi}')^T \boldsymbol{q}' \Delta x \Delta y = E'_M \approx \mathcal{E}_0$$

We define

$$\hat{p}_{k,\ell} = \frac{\hat{q}'_{k,\ell}}{(k^2 + \ell^2)^{1/2}}.$$

The fluctuation Gibbs distribution is again Gaussian with $\eta_p = (\langle I_M \rangle / 2\pi^2)^{1/2}$. The standard deviation η_q of the fluctuation vorticity is given by (2.38) with $A = (-\Delta_M)^{1/2}$.

Enstrophy-based statistical mechanics

For a semi-discretization that only preserves the enstrophy Z_M, the least biased distribution (2.31) becomes

$$\mathcal{G}_E(\boldsymbol{y}) = N^{-1} \exp\left[-\lambda Z_M(\boldsymbol{y})\right].$$

The mean field prediction (2.32) gives

$$\langle \boldsymbol{q} \rangle \equiv 0, \qquad \langle \boldsymbol{\psi} \rangle = -\Delta_M^{-1} \boldsymbol{h}. \tag{2.43}$$

The fluctuation dynamics (2.34) becomes

$$\boldsymbol{q}' = J_Z(\boldsymbol{q}', \langle \boldsymbol{\psi} \rangle + \boldsymbol{\psi'}), \qquad \boldsymbol{\psi}' = \Delta_M^{-1} \boldsymbol{q}',$$

and the pseudo-energy is just the enstrophy, i.e.

$$I_M = \frac{1}{2} (\boldsymbol{q}')^T \boldsymbol{q}' \Delta x \Delta y = Z'_M \approx \mathcal{Z}_0.$$

The fluctuation Gibbs distribution is Gaussian with $\hat{p}_{k,\ell} = \hat{q}'_{k,\ell}$ and

$$\eta_q = \sqrt{\frac{\langle I_M \rangle}{2\pi^2}}.$$

2.5 Time integration

To test the statistical predictions of the previous section with computations, the semi-discretizations of Section 2.3 must be supplemented with a time stepping scheme. One would prefer to have a scheme that conserves the invariants E_M and Z_M in time whenever these are first integrals of the spatial discretization. Additionally, one would like to have a scheme that preserves volume. There is much literature on the preservation of first integrals under discretization; see [35] for an overview. Much less is known about preserving volume.

Time discretizations

Since both invariants E_M and Z_M of the discretizations are quadratic functions of $\boldsymbol{q}$, they are automatically conserved if the equations are integrated with a Gauss-Legendre Runge-Kutta method [35]. The simplest such method is the implicit midpoint rule

$$\frac{\boldsymbol{q}^{n+1} - \boldsymbol{q}^n}{\Delta t} = J\left(\frac{\boldsymbol{q}^{n+1} + \boldsymbol{q}^n}{2}, \frac{\boldsymbol{\psi}^{n+1} + \boldsymbol{\psi}^n}{2}\right).$$

The discretization is also symmetric, and in the case of zero topography $h(\boldsymbol{x}) \equiv 0$, preserves the time reversal symmetry $t \mapsto -t$, $q \mapsto -q$ of (2.1). Although it is symplectic for Hamiltonian systems with constant structure operators, the midpoint rule is not volume preserving in general. Indeed, it does not preserve volume exactly for our discretizations. However, numerical experiments indicate that volume is approximately conserved on long intervals, even for a relatively large step size.

The implicit midpoint rule requires the solution of a nonlinear system of dimension M^2 at every time step. As a more efficient alternative, we can take any explicit Runge-Kutta method and project the solution onto the integral manifolds as desired. Let the Runge-Kutta method be represented by a map $\boldsymbol{q}^{n+1} = \Phi_{\Delta t}(\boldsymbol{q}^n)$ and compute a predicted step

$$\boldsymbol{q}^* = \Phi_{\Delta t}(\boldsymbol{q}^n).$$

Then project $\boldsymbol{q}^*$ onto the desired constraint manifolds by solving

$$\begin{aligned} \boldsymbol{q}^{n+1} &= \boldsymbol{q}^* + \boldsymbol{g}'(\boldsymbol{q}^*)^T \boldsymbol{\lambda}, \\ \boldsymbol{g}(\boldsymbol{q}^{n+1}) &= \boldsymbol{0} \end{aligned}$$

for $\boldsymbol{\lambda}$, where $\boldsymbol{g}(\boldsymbol{q}) : \mathbb{R}^{M\times M} \to \mathbb{R}^r$, r the number of first integrals, and $\boldsymbol{\lambda} \in \mathbb{R}^r$ is a vector of Lagrange multipliers. For example, we can take ($r = 3$)

$$\boldsymbol{g}(\boldsymbol{q}) = \begin{pmatrix} E_M(\boldsymbol{q}) - \mathcal{E}_0 \\ Z_M(\boldsymbol{q}) - \mathcal{Z}_0 \\ C_M(\boldsymbol{q}) - 0 \end{pmatrix},$$

where the last constraint ensures that there is no drift in total vorticity. At each time step, projection requires solving a small nonlinear problem of dimension r. Projected Runge-Kutta methods will not preserve volume in general.

Time averages

Our interest is in the statistics applied to numerical data obtained from simulations over long times. To apply the theory from the previous sections, we additionally have to assume that the semi-discrete dynamics are ergodic. Denote the time average of a quantity $F(\boldsymbol{q}(t))$ by

$$\overline{F}_T = \frac{1}{T}\int_{t_0}^{t_0+T} F(\boldsymbol{q}(t))\,dt.$$

Then the assumption of ergodicity implies that the long time average converges to the ensemble average

$$\overline{F} = \lim_{T\to\infty} \overline{F}_T = \langle F \rangle.$$

On the other hand, suppose one chooses discrete initial conditions to have a prescribed energy and enstrophy consistent with the continuum problem, i.e.

$$E_M(\boldsymbol{q}(0)) = \mathcal{E}_0, \qquad Z_M(\boldsymbol{q}(0)) = \mathcal{Z}_0.$$

Then it is clear that since $E_M(\boldsymbol{q}(t)) = E_M(\boldsymbol{q}(0))$ and $Z_M(\boldsymbol{q}(t)) = Z_M(\boldsymbol{q}(0))$ are conserved, the dynamics only samples at most a codimension two subspace of $\mathbb{R}^{M\times M}$, so one may ask to what extent the averages will converge. Indeed, one has inequality

$$\overline{E}_M = \langle E_M \rangle \neq \mathcal{E}_0, \qquad \overline{Z}_M = \langle Z_M \rangle \neq \mathcal{Z}_0,$$

in general. By analogy with molecular dynamics, the Gibbs distribution (2.31) determines expectations in the canonical ensemble, whereas a constant energy-enstrophy simulation determines expectations in the microcanonical ensemble (assuming ergodicity). It is only in the 'thermodynamic limit' $M \to \infty$ that these averages coincide, giving equality in the above relations.

2.6 Numerical experiments

For the numerical experiments we use the test problem of [1]. The grid resolution is $M = 22$. The orography is a function of x only, specifically

$$h(x, y) = 0.2\cos x + 0.4\cos 2x.$$

(As a result the predicted mean fields $\overline{\boldsymbol{q}}$ and $\overline{\boldsymbol{\psi}}$ should be functions of x only.) The integrations were carried out using a step size of $\Delta t = 0.1$.

For initial conditions we take a uniformly random field[2] $\boldsymbol{q} = (q_{i,j})$, $i, j = 1, \ldots, M$ and project this onto the constraints

$$E_M(\boldsymbol{q}) = \mathcal{E}_0, \quad Z_M(\boldsymbol{q}) = \mathcal{Z}_0, \quad C_M(\boldsymbol{q}) = 0.$$

The same initial condition is used for all simulations. The discrete energy and enstrophy were taken to be $\mathcal{E}_0 = 7$ and $\mathcal{Z}_0 = 20$.

With these values prescribed, the statistical predictions of Section 2.4 can be computed for the three discretizations (2.12), (2.13), and (2.14). The Lagrange multiplier μ is computed using the procedure described at the end of Section 2.4.3. Fluctuation statistics apply to the time series of PV at an arbitrarily chosen monitor point on the grid $q_{\text{mon}} = q_{3,12}$.

For the energy-enstrophy theory we obtain the mean state (2.32) and estimates

$$EZ: \quad \mu = -0.730, \qquad \langle q_{\text{mon}} \rangle = -0.341, \qquad \eta_q = 0.970. \tag{2.44}$$

For the energy theory of Section 2.4.4 we obtain the mean state (2.42) and estimates

$$E: \quad \langle \psi \rangle \equiv 0, \qquad \langle q_{\text{mon}} \rangle = 0.0740, \qquad \eta_q = 5.36. \tag{2.45}$$

For the enstrophy theory of Section 2.4.4 we obtain the mean state (2.43) and estimates

$$Z: \quad \mu = 0, \qquad \langle q_{\text{mon}} \rangle = 0, \qquad \eta_q = 1.01. \tag{2.46}$$

The discretization (2.11), which conserves neither energy nor enstrophy, was found to be exponentially unstable under time discretization by the implicit midpoint rule, and no experiments with that discretization will be reported here.

Results using implicit midpoint

We first present results obtained using the implicit midpoint discretization in time. The nonlinear relations were solved using fixed point iteration to a tolerance of 10^{-13}, which was the smallest tolerance that gave convergence at each step size for all discretizations. The solutions were averaged over the interval $10^3 \leq t \leq T$, for $T = 10^4$, 10^5 and 10^6. Averages were computed from time $t = 1000$ to allow the initially uniformly random initial condition to de-correlate, and this time is consistent with that used in [1] for a spectral discretization.

Given the average fields $\overline{\boldsymbol{q}}$ and $\overline{\boldsymbol{\psi}}$, the best linear fit to (2.32) yields an estimate of the Lagrange multiplier $\overline{\mu}$, i.e.

$$\overline{\mu} = \frac{\overline{\boldsymbol{\psi}}^T \overline{\boldsymbol{q}}}{\overline{\boldsymbol{\psi}}^T \overline{\boldsymbol{\psi}}}.$$

[2]Experiments with smooth initial conditions typically show no noticable difference, however.

The relative change in energy and enstrophy for each discretization is plotted in Figure 2.1 on the interval $[0, 10^5]$. The relative change is defined as

$$\Delta E_M^n = \frac{E_M^n - \mathcal{E}_0}{\mathcal{E}_0}, \qquad \Delta Z_M^n = \frac{Z_M^n - \mathcal{Z}_0}{\mathcal{Z}_0}.$$

For the EZ discretization, both quantities are conserved up to the tolerance of the fixed point iteration, which leads to a small drift of magnitude 3×10^{-11} (relative) over this interval. For the E discretization, energy is conserved to the tolerance of the fixed point iteration, but enstrophy makes a rapid jump to a mean state roughly 30 times its initial value and subsequently undergoes bounded fluctuations with amplitude about $10 \times Z_0$. In contrast, for the Z discretization, enstrophy is similarly conserved, but energy drifts gradually with a negative trend, to about 25% of its initial value.

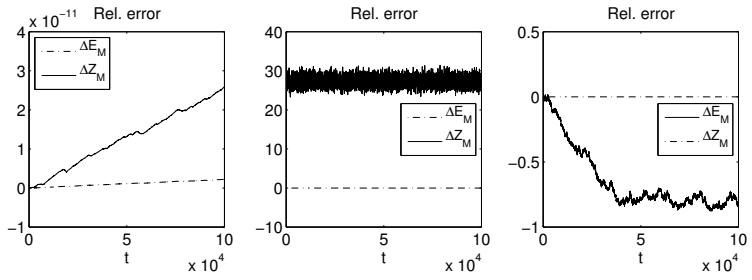

FIGURE 2.1: Relative change in energy and enstrophy with EZ (left), E (middle) and Z (right) discretizations.

Long-time mean fields

The time-averaged stream function $\overline{\boldsymbol{\psi}}$ obtained by averaging over the interval $[10^3, 10^4]$ is shown in Figure 2.2 for the three EZ, E and Z discretizations. Also shown is a scatter plot of the locus $(\overline{\psi}_{i,j}, \overline{q}_{i,j})$ and a linear best fit to this data for the respective discretizations.

For the EZ discretization, the mean stream function is similar to that predicted by the energy-enstrophy statistical theory (2.32), with $\overline{\mu} = -0.734$. For the E discretization, the mean stream function satisfies $\overline{\boldsymbol{\psi}} \approx 0$, consistent with (2.42), and the linear regression is inaccurate. For the Z discretization, we observe a similar mean state with $\overline{\mu} = -0.715$ on this averaging interval, which is inconsistent with prediction (2.46).

In Figures 2.3 and 2.4 we examine more closely the mean fields for the EZ and Z discretizations, for longer averaging times of $T = 10^5$ and $T = 10^6$. For the energy-enstrophy discretization (2.14) in Figure 2.3, the mean field appears to converge to an equilibrium state with $\overline{\mu} \approx -0.732$. The tendency in Figure 2.4 is toward a mean field with zero vorticity, consistent with (2.43).

However the relaxation time is much longer than for the other discretizations. For $T = 10^6$, the mean flow has $\overline{\mu} = -0.0529$. Note that the relation $\overline{\boldsymbol{q}} = \overline{\mu}\overline{\boldsymbol{\psi}}$ approximates the data well for all averaging times, however. Below in this section, we show that the convergence of the Z discretization is in agreement with the EZ predictions on short time intervals, so that we can think of the system staying near statistical equilibrium with slowly drifting energy.

PV fluctuation statistics

In Figure 2.5, the time series for potential vorticity q_{mon} at an arbitrarily chosen grid point $(3, 12)$ is analyzed. As discussed in Section 2.4.2, the statistical theory for fluctuations predicts that the PV should be distributed normally about the mean field according to (2.44)–(2.46). For the longest simulation time of $T = 10^6$, the EZ discretization exhibits Gaussian fluctuations with mean $\overline{q}_{\text{mon}} = -0.395$ and standard deviation $\eta = 0.927$; the E discretization with mean $\overline{q}_{\text{mon}} = -0.0093$ and standard deviation $\eta = 5.35$; and the Z discretization with mean $\overline{q}_{\text{mon}} = -0.0575$ and standard deviation $\eta = 1.05$. These observations are approximately in agreement with (2.44)–(2.46).

We mention that the value $\overline{\mu} = -0.732$, to which the EZ discretization seems to relax, corresponds to a mean energy value of $\langle E_M \rangle = 7.07$. For this value of mean energy, the prediction of Section 2.4.2 gives $\eta = 0.928$, which is much closer to the value observed in Figure 2.5. This indicates that for implicit midpoint, the mean energy is somewhat preturbed from the microcanonical energy $\mathcal{E}_0$.

Time-dependent energy-enstrophy model

In Figure 2.6, the convergence of $\overline{\mu}$ is plotted as a function of averaging interval T for both the Z and EZ discretizations. The EZ dynamics relaxes very rapidly to give $\overline{\mu} \approx -0.73$, whereas the Z dynamics converges rather slowly towards $\overline{\mu} = 0$.

Given the relatively fast relaxation of the energy-enstrophy conserving discretization to statistical equilibrium (2.32) and the slow drift of energy in Figure 2.1 for the enstrophy conserving discretization (2.13), a natural model for the approach to equilibrium would be to consider a state $\overline{q}_T = \overline{\mu}_T \overline{\psi}_T$ with $\overline{\mu}_T$ corresponding to the instantaneous energy $E_M(T)$. To test this idea, we define

$$\overline{\boldsymbol{\psi}}_T = \frac{1}{N_T} \sum_{n=N_0}^{N_T+N_0} \boldsymbol{\psi}^n, \qquad \overline{\boldsymbol{q}}_T = \frac{1}{N_T} \sum_{n=N_0}^{N_T+N_0} \boldsymbol{q}^n,$$

where $T = N_T \cdot \Delta t$, and

$$\overline{\mu}_T = \frac{(\overline{\boldsymbol{\psi}}_T)^T \overline{\boldsymbol{q}}_T}{(\overline{\boldsymbol{\psi}}_T)^T \overline{\boldsymbol{\psi}}_T}.$$

The energy of the associated equilibrium state is denoted $E_M(\overline{\mu}_T)$ and is determined from the relations in Section 2.4.3. This energy is plotted in Figure 2.7

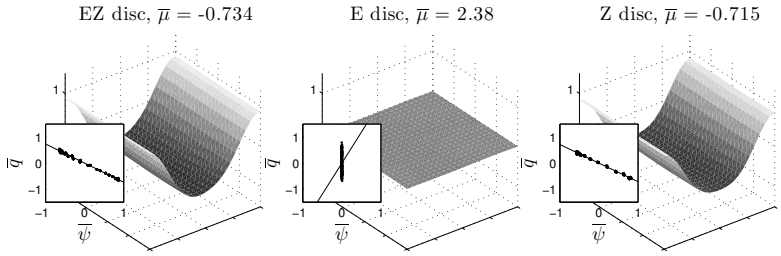

FIGURE 2.2: Mean fields with averaging time 10^4, EZ (left), E (middle), and Z (right). The insets show the best linear fit to the relation $\overline{\psi}_{i,j} = \overline{\mu q}_{i,j}$ at all grid points.

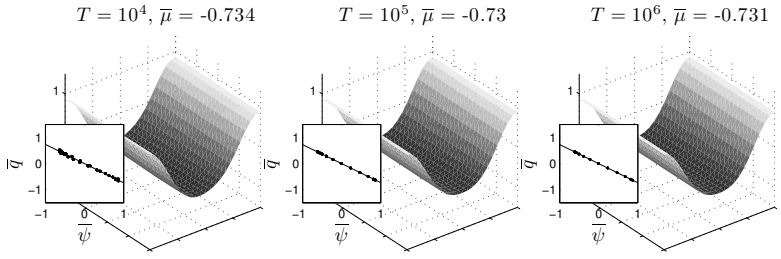

FIGURE 2.3: Mean fields for EZ discretization with averaging times 10^4 (left), 10^5 (middle), and 10^6 (right). The insets show the best linear fit to the relation $\overline{\psi}_{i,j} = \overline{\mu q}_{i,j}$ at all grid points.

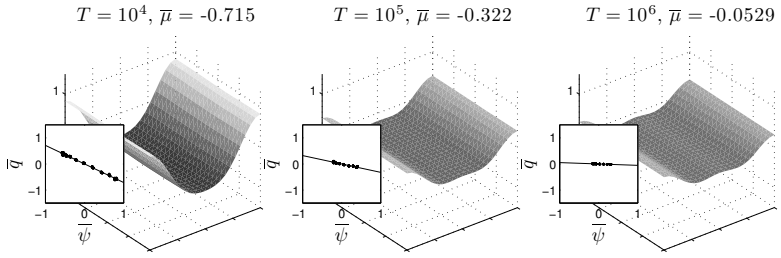

FIGURE 2.4: Mean fields for discretization J_Z with averaging times 10^4 (left), 10^5 (middle), and 10^6 (right). The insets show the best linear fit to the relation $\overline{\psi}_{i,j} = \overline{\mu q}_{i,j}$ at all grid points.

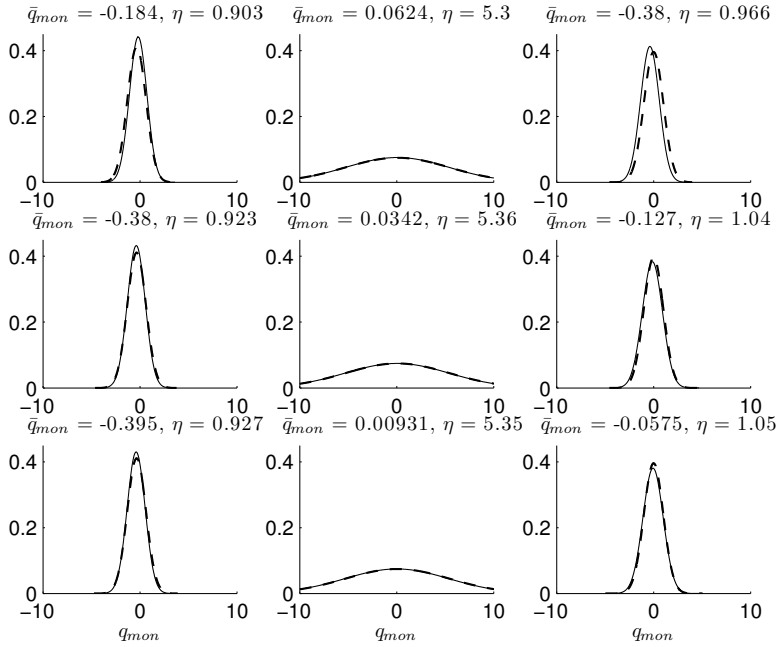

FIGURE 2.5: Fluctuation statistics for the potential vorticity about the predicted mean. Solid line is a Gaussian fit to the numerical observation. Dash line is the predicted distribution. Discretizations EZ, E, and Z in left, middle and right columns. Integration intervals of 10^4, 10^5 and 10^6 in top, middle and bottom rows.

next to the actual discrete energy function, for increasing averaging intervals $T = 10$, $T = 100$ and $T = 1000$. The agreement supports this model. That is, the Z dynamics relaxes on a fast time scale to the statistical equilibrium predicted by energy-enstrophy theory *for the instantaneous energy*, while the energy drifts on a slow time scale towards the equilibrium state predicted by the enstrophy theory.

Results using projected Heun's method

Besides preserving quadratic first integrals exactly, the implicit midpoint rule is symmetric. It is unclear what effect, if any, this may have on statistics. Furthermore, the implicit midpoint rule is fully implicit and therefore not a very practical choice for integrating a nonstiff system such as (2.1). For these reasons we repeat the experiments of the previous section using the second order, explicit Runge-Kutta method due to Heun [36], coupled with projection onto the discrete energy and/or enstrophy manifolds. It should be noted that

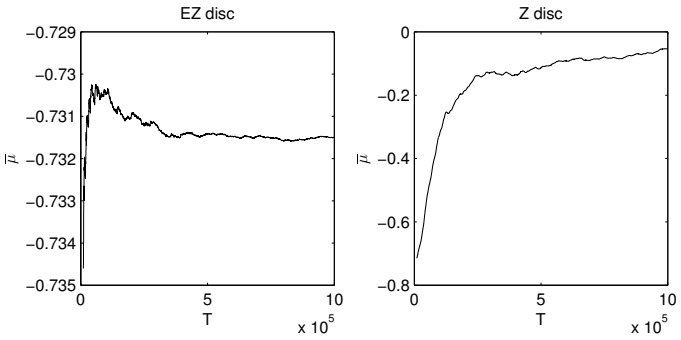

FIGURE 2.6: Convergence of parameter $\overline{\mu}_T$ as a function of the averaging interval T for the EZ and Z discretizations.

Heun's method is linearly unstable with respect to a center equilibrium, and it is only due to projection that we can carry out long integrations with this method.

Figure 2.8 compares the convergence of the parameter $\overline{\mu}_T$ as a function of T for the implicit midpoint and projected Heun integrators for the EZ and Z discretizations. In both cases, it appears that the projected method approaches equilibrium faster than implicit midpoint.

Figures 2.9, 2.10 and 2.11 are analogous to Figures 2.2, 2.3 and 2.4 for implicit midpoint. Again we note that the projected method converges more rapidly and more accurately to the mean states (2.44)–(2.46).

The fluctuation statistics for the projection method are illustrated in Figure 2.12. Here, too, we see that the projection method is very close to the statistically predicted value for mean and standard deviation of PV fluctuations in (2.44)–(2.46). However, it is important to note that since a measure of predictability is the *deviation from the statistical equilibrium*, a numerical method that approaches equilibrium excessively fast is undesirable from a prediction perspective.

Discrete volume preservation

Although the spatial discretizations were shown to be volume preserving, neither the implicit midpoint rule nor the projected Heun integrator preserves volume for the discrete map. To get an impression of the degree of volume contraction, we computed the determinant of the Jacobian of the discrete flow maps, e.g.

$$c^n = \det\left(\frac{d\boldsymbol{q}^{n+1}}{d\boldsymbol{q}^n}\right)$$

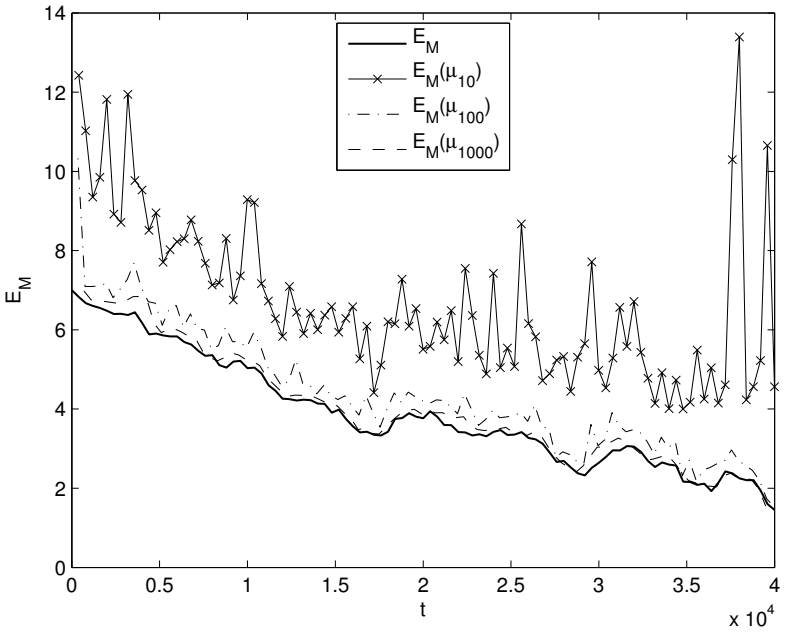

FIGURE 2.7: Energy drift with Z discretization, compared to the energy associated with the best linear fit $\overline{\mu}_T$ with averaging intervals $T = 10$, $T = 100$ and $T = 1000$.

in each time step. The cummulative volume ratio was defined to be

$$V^n = \prod_{m=0}^{n} c^m.$$

This volume ratio is plotted as a function of time in Figure 2.13 for the implicit midpoint and projected Heun methods. In both cases, a grid of size $M = 12$ was and step size $\Delta t = 0.1$ were used. The EZ discretization (2.14) was employed, with in the second case, projection onto the energy and enstrophy manifolds.

Remarkably, the implicit midpoint rule conserves volume to within 3×10^{-3} over the entire interval, exhibiting only a small positive drift.

For the projected method, volume is greatly contracted—to 10^{-4} at time $t = 10$ (100 time steps).

2.7 Conclusions

We have constructed statistical mechanical theories for three conservative discretizations of the quasigeostrophic model due to Arakawa [2], based on conser-

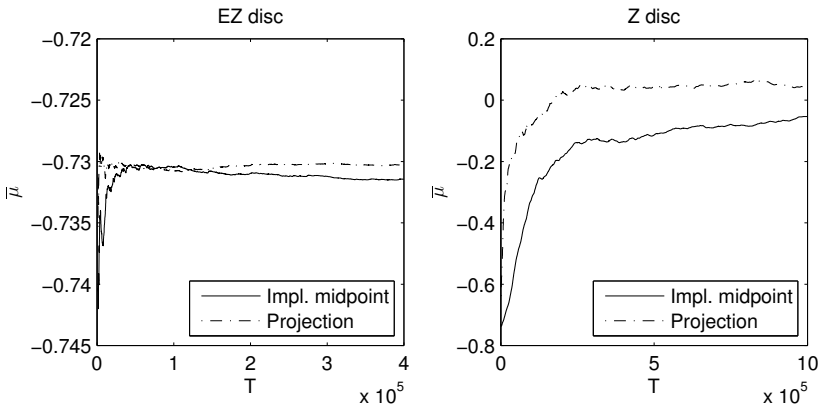

FIGURE 2.8: Convergence of $\overline{\mu}_T$ as a function of averaging interval T for EZ (left) and Z (right) discretizations, comparing the projected Heun's method and implicit midpoint.

vation of energy, enstrophy, or both. Numerical experiments indicate that the statistical theories can give insight into the long time behavior of the discretizations, making this approach a useful tool for numerical analysis.

Time integration of the semi-discretization was done with the symmetric implicit midpoint method—which automatically conserves any quadratic first integrals of the semi-discrete system—and with a projected Runge-Kutta method. Long time averages with the implicit midpoint discretization relax to the predicted equilibrium at a slower rate than for the projected method, suggesting that implicit midpoint has higher potential for prediction. The implicit midpoint rule was also found to approximately conserve volume for long time intervals. This is in stark contrast to the projection method, for which phase space volume is rapidly contracted.

The three statistical theories predict dramatically different behavior, and this is confirmed by the numerical experiments. In other words, *the three discretizations exhibit dramatically different behavior* in simulations over long intervals. The statistical equilibrium states define a backdrop on which the discrete dynamics occurs, and that backdrop depends on the conservation properties of the spatial discretization. Assuming the energy-enstrophy theory to be correct, it is thus essential for any code to preserve both quantities (under semi-discretization) if statistical consistency is desired. The results of this work make a strong argument for the use of conservative discretizations in weather and climate simulations.

On the other hand, it has been shown by Abramov & Majda [1] that the energy-enstrophy theory is incomplete. In [1], the Poisson discretization of [88] is integrated using the Poisson splitting of McLachlan [50]. The semi-discretization preserves, in addition to the Hamiltonian, M Casimirs corre-

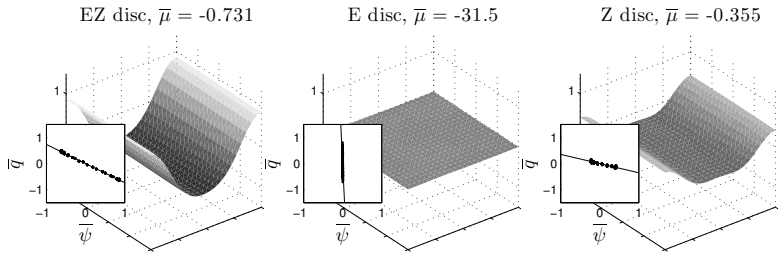

FIGURE 2.9: Same as Figure 2.2, but using projected Heun's method.

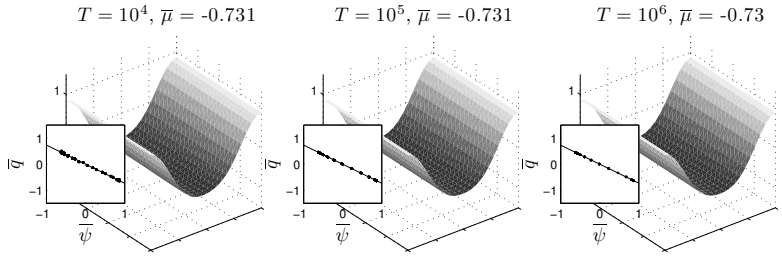

FIGURE 2.10: Same as Figure 2.3, but using projected Heun's method.

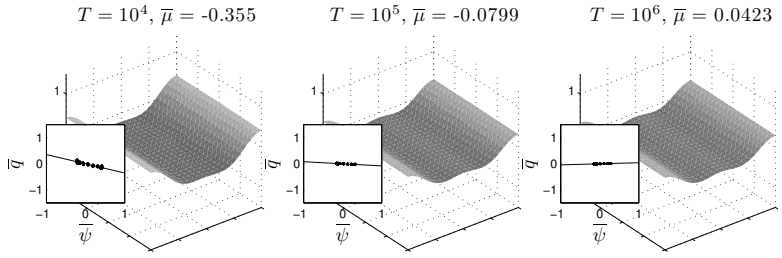

FIGURE 2.11: Same as Figure 2.4, but using projected Heun's method.

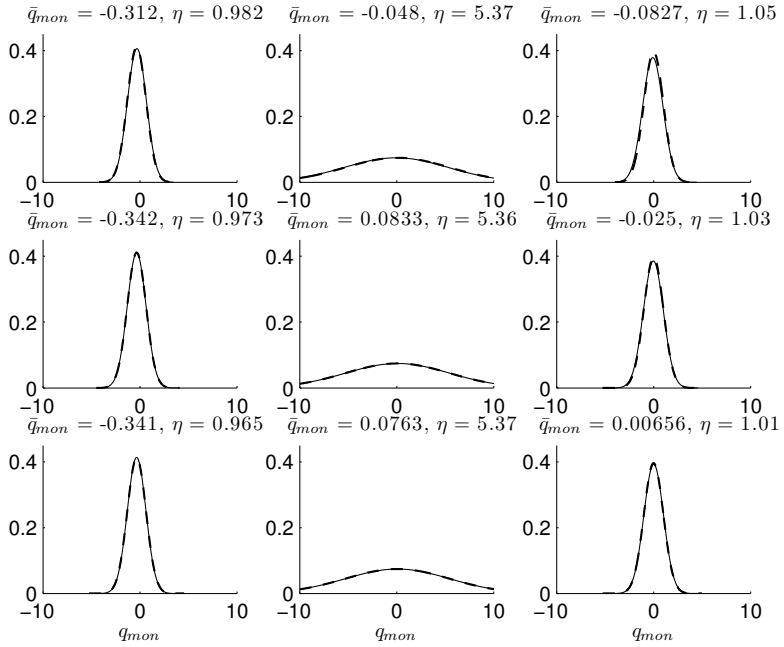

FIGURE 2.12: Same as Figure 2.5, but using projected Heun's method.

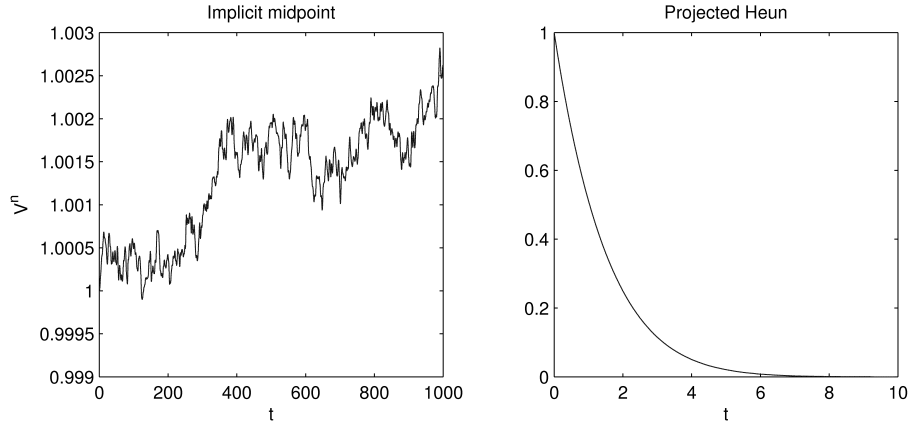

FIGURE 2.13: Volume contraction ratio for implicit midpoint (left) and projected Heun (right) methods, EZ discretization (2.14), $M = 12$, 10^4 time steps.

sponding to the first M moments of potential vorticity (PV), and these are

preserved by the splitting (the energy is only preserved approximately, in the sense of backward error analysis [35]). Abramov & Majda give convincing evidence that nonzero values of the third moment of PV, when conserved by the discretization, can significantly skew the predictions of the standard theory of [11, 40, 48, 76]. In Chapter 3 we will consider statistical theories for the Hamiltonian particle-mesh method, which in addition to the energy preserves any fumction of PV on the particles.

Chapter 3

Statistical mechanics of the Hamiltonian particle-mesh method

3.1 Introduction

Computational applications in atmosphere and ocean science often involve the simulation of geophysical fluids on time intervals much longer than the Lyapunov e-folding time. In these instances the goal of simulation is the evaluation of statistical quantities such as averages and correlations. It is therefore important to investigate the accuracy of numerical discretizations in the context of statistical averages.

The Hamiltonian particle-mesh (HPM) method was originally proposed in the context of rotating shallow water flow in periodic geometry in [25] and extended to other physical settings in [16, 17, 29, 83]. The fluid is discretized on a finite set of Lagrangian particles that transport the mass of the fluid and persist during the flow evolution. The HPM method is symplectic, and one can construct a continuum velocity field in which the discrete particle velocities are embedded for all time. The continuum velocity field satisifies a Kelvin circulation theorem, implying material conservation of potential vorticity and invariance of the infinite family of Casimir functionals, see [28].

For the case of ideal fluid flow in two dimensions, the HPM variant was described in [17]. We will apply the HPM method to the quasigeostrophic potential vorticity equation describing a 2D barotropic flow over topography. Here, it is the potential vorticity (PV), and *not* the mass, that is fixed on each particle and advected in a divergence-free velocity field. In this case we show that the particle motion may be embedded in an area-preserving flow on the fluid label space. Hence, an arbitrary function of PV may be integrated by quadrature over label space and is therefore conserved. On the other hand, this trivial conservation is apparently due to the fact that a value of PV is assigned to each particle once and for all, and does not imply any reduction of the number of degrees of freedom of the flow evolution in the sense of, say, a discrete energy conservation law. In fact, the method makes use of a secondary (coarse-grain)

vorticity field, defined on a uniform grid, and numerical simulations indicate that only energy and linear functionals of PV are conserved at this macroscopic scale.

Consequently, one may question to what degree the PV conservation of the HPM method is meaningful. Long time simulations of nonlinear dynamical systems are typically inaccurate in a pointwise sense and rather are carried out with the goal of sampling an equilibrium probability distribution over the phase space under an assumption of ergodicity. The relevant statistical equilibrium distribution is a function of the conservation laws that restrict the dynamics. Hence, if the PV conservation by HPM is truly trivial, the statistics should adhere to that of an energy-circulation equilibrium theory, whereas a meaningful PV conservation should lead to a richer statistical equilibrium. In this chapter we will show that the PV conservation of HPM significantly influences the statistics of simulation data obtained with the HPM method.

Sophisticated statistical equilibrium theories for ideal fluids are based on conservation of generalized enstrophies—the integrals over the domain of arbitrary functions of PV—or, equivalently, on the area-preservation property of the velocity field. An equilibrium theory with microcanonical treatment of vorticity invariants was developed independently by Lynden-Bell [47], Robert & Sommeria [71, 72], and Miller [54]. An alternative approach that treats vorticity invariants canonically was developed by Ellis, Haven & Turkington [22], see Section 1.3.3. See also Chavanis [13] for a comparison of the two. Our numerical method has features in common with the model used to construct the latter continuum theory (specifically, a natural two-scale structure), and in this chapter we derive analogous *discrete* statistical equilibrium models based on both Lagrangian (Section 3.4) and Eulerian (Section 3.5) fluid considerations and compare these models with simulations (Section 3.6).

We wish to emphasize here that the objective of this chapter is to show that the discrete statistics of the HPM method are in good agreement with predictions of the modern continuum theory. This implies that the contruction of the HPM method respects the dynamical considerations that go into the theory, and that for the numerical experiments included, the discrete flow is sufficiently ergodic to observe convergence of the ensemble averages.

The quasigeostrophic potential vorticity (QG) equation [48, 67, 74] describes barotropic divergence-free flow over topography

$$\frac{d}{dt}q(\boldsymbol{x},t) = 0, \qquad \Delta\psi(\boldsymbol{x},t) = q(\boldsymbol{x},t) - h(\boldsymbol{x}), \tag{3.1}$$

where q is the potential vorticity (PV) field, ψ is the stream function, and h is the topography of the earth. The Laplacian operator is denoted by Δ and the material derivative by $\frac{d}{dt} = \frac{\partial}{\partial t} + \boldsymbol{u}\cdot\nabla$. Here, the divergence-free velocity field $\boldsymbol{u}$ is related to the stream function by $\boldsymbol{u} = \nabla^{\perp}\psi$, where $\nabla^{\perp} = (-\frac{\partial}{\partial y}, \frac{\partial}{\partial x})^T$. In this chapter we consider the QG equation on a doubly periodic domain

$$\boldsymbol{x} = (x, y) \in \mathcal{D} \equiv [0, 2\pi) \times [0, 2\pi).$$

The QG model describes a Hamiltonian PDE with Lie-Poisson structure [74], implying the conservation of the total kinetic energy

$$\mathcal{H} = -\frac{1}{2}\int_{\mathcal{D}} \psi \cdot (q-h)\, d\boldsymbol{x} \tag{3.2}$$

as well as the infinite class of Casimir functionals

$$\mathcal{C}[f] = \int_{\mathcal{D}} f(q)\, d\boldsymbol{x} \tag{3.3}$$

for any function f for which the integral exists. Of particular interest are the PV moments:

$$\mathcal{C}_r = \int_{\mathcal{D}} q^r\, d\boldsymbol{x}, \quad r = 1, 2, \ldots, \tag{3.4}$$

and especially the circulation $\mathcal{C}_1$ and enstrophy $\mathcal{C}_2$.

Preservation of the Casimir functionals follows from area-preservation under the divergence-free velocity field [55]: Define a function $G(\sigma, t)$ denoting the measure of that part of the domain $\mathcal{D}$ for which the vorticity is less than σ:

$$G(\sigma, t) = \text{meas}\{\boldsymbol{x} \in \mathcal{D} \,|\, q(\boldsymbol{x}, t) < \sigma\}. \tag{3.5}$$

We note that due to the divergence-free advection of q, this function is independent of time $\frac{\partial G}{\partial t} = 0$. Differentiating with respect to σ, the function

$$g(\sigma) = \frac{\partial G}{\partial \sigma} \tag{3.6}$$

is preserved. For the case of a piecewise uniform PV field, $q(\boldsymbol{x}, t) \in \{\sigma_1, \ldots, \sigma_\Lambda\}$, the quantity $g_\ell = G(\sigma_{\ell+1}) - G(\sigma_\ell)$ is the measure of the vorticity level set σ_ℓ.

3.2 Review of continuum statistical equlibrium theories

Given a (spatially) discrete approximation $\boldsymbol{q}(t)$ to the solution $q(\boldsymbol{x}, t)$ of (3.1), obtained from a numerical simulation, one would like to analyze the accuracy of computed averages of functions of the solution. For example, the long time average of a function $F(\boldsymbol{q}(t))$ of the PV field is denoted

$$\overline{F} = \lim_{T\to\infty} \frac{1}{T}\int_{t_0}^{t_0+T} F(\boldsymbol{q}(t))dt.$$

If the discrete dynamics is ergodic with respect to a unique invariant measure $p(\boldsymbol{q})$ on the phase space, then the long time average is equivalent to the ensemble average with respect to p,

$$\langle F(\boldsymbol{q})\rangle = \int F(\boldsymbol{q})p(\boldsymbol{q})\, d\boldsymbol{q},$$

where the integral is over the (function) space of PV fields, and it suffices to derive the invariant measure associated with the numerical method, and analyze this with respect to what is known about the continuum model.

In an effort to characterize the long time mean behavior of ideal fluids and explain their tendency to organize into large-scale coherent structures, a number of authors have applied ideas from statistical mechanics. The pioneering work was that of Onsager [65], which addressed the statistical mechanics of a finite point vortex model. He observed that for a bounded domain, the available phase space must eventually decrease as a function of increasing system energy, leading to negative temperature regimes in the microcanonical statistical ensemble. He also predicted that in a heterogeneous system of like-signed and oppositely-signed vortices of varying strengths, the large vortices would tend to cluster so as to achieve maximum disorder with minimum degrees of freedom, like-signed vortices at negative temperatures and vice-versa. These predictions were confirmed numerically in [9], and this problem under the influence of mathematical thermostat will be considered in Chapter 4.

Statistical mechanics theories based on a Fourier-spectral truncation of the Euler equations were proposed by [11, 40, 76], see also Section 1.3.3. The spectral truncation preserves circulation $\mathcal{C}_1$ and the quadratic functions $\mathcal{E}$ and $\mathcal{C}_2$ only. Consequently, a treatment based on constrained maximum entropy (for an introduction, see [48]) yields Gibbs-like distributions $p(q)$ over the vorticity field, with Gaussian distribution of local vorticity fluctuations. The energy-enstrophy theory predicts a linear relation between ensemble average stream function and potential vorticity:

$$\langle q(x) \rangle = \mu \langle \psi(x) \rangle, \tag{3.7}$$

for a scalar μ depending on the observed energy and enstrophy, where the ensemble average is defined with respect to integration over an appropriate function space in this case.

Given that it is not just the enstrophy $\mathcal{C}_2$ but any functional $\mathcal{C}[f]$ that is invariant under the continuum flow of (3.1), it is natural to ask what effect the more general conservation laws have on statistics. Abramov & Majda [1] investigated the statistical significance of the higher PV moments $\mathcal{C}_r$, $r > 2$, numerically using the Poisson discretization of Zeitlin [50, 88], which conserves M Casimirs of an $M \times M$- mode truncation, see Section 1.2. Computing the long time averaged PV and stream function fields, they observed increasing discrepancy relative to the linear mean field theory (3.7), as a function of increasing skewness $\mathcal{C}_3$ of the initial condition, thus proving the statistical relevance of this quantity.

Statistical equilibrium theories incorporating the full family of Casimirs implied by the preservation of area (3.5) were independently proposed by Lynden-Bell [47] in the context of astrophysics, and Miller [54, 55] and Robert & Sommeria [71, 72], see Section 1.3.3. These original theories used a microcanonical treatment of the Casimirs $\mathcal{C}[f]$. That is, the equilibrium distribution is derived by minimizing entropy under constraints of energy and the entire family

of Casimirs. More recently, Ellis, Haven & Turkington [22] proposed an alternative theory featuring canonical treatment of the Casimirs, in which the point statistics of PV is described by a prior distribution, see Section 1.3.3. In all of these papers, a coarse-grain potential vorticity field is described by a probability density over the class of fine-grain PV distributions at each point in the domain.

In Chapter 2 we analyzed energy and enstrophy conserving finite difference methods for the QG model under topographic forcing, and observed that the discrete time-averaged mean fields $\overline{\boldsymbol{q}}$ and $\overline{\boldsymbol{\psi}}$ obtained depend heavily on the conservation properties of the discretizations used. For a discretization that conserves energy only, the predicted mean field is uniformly zero velocity $\langle \boldsymbol{\Psi} \rangle = 0$. Given that the only dynamically conserved quantities of the HPM method are energy and total circulation (see below), any departure from the trivial mean field is an indication of the statistical relevance of the other conserved quantities, namely the area measure (3.5).

3.3 Hamiltonian particle-mesh method

The Hamiltonian particle-mesh (HPM) method is a numerical discretization of inviscid fluid dynamics that retains Hamiltonian structure. The method makes use of a Lagrangian fluid description, to advect fluid particles while conserving energy, and an Eulerian grid for evaluating derivatives using finite differences. The method was adapted for 2D incompressible flow in [17].

3.3.1 HPM description

The PV field is discretized by introducing a set of K discrete particles with fixed potential vorticity Q_k, $k = 1, \ldots, K$. The particles have time-dependent position $\boldsymbol{X}_k(t) \in \mathcal{D}$, and are advected in a divergence-free velocity field according to

$$\frac{d}{dt}\boldsymbol{X}_k = \nabla^\perp \Psi(\boldsymbol{x}, t)\big|_{\boldsymbol{x}=\boldsymbol{X}_k(t)}$$

where the stream function Ψ is described below.

We also make use of a uniform $M \times M$ grid on $\mathcal{D}$, with grid spacing $\Delta x = \Delta y = 2\pi/M$, and denote gridpoints by $\boldsymbol{x}_i$. Given a discrete stream function $\Psi_i(t)$ on the grid, we construct a continuous field via

$$\Psi(\boldsymbol{x}, t) = \sum_i \Psi_i(t)\phi_i(\boldsymbol{x}), \tag{3.8}$$

where $\phi_i(\boldsymbol{x}) = \phi(\frac{\boldsymbol{x}-\boldsymbol{x}_i}{\Delta x})$ is a compactly supported basis function satisfying symmetry, normalization and partition of unity properties, respectively:

$$\phi(\boldsymbol{x}) = \phi(-\boldsymbol{x}), \qquad \int_{\mathcal{D}} \phi(\boldsymbol{x})\, d\boldsymbol{x} = 1, \qquad \sum_i \phi_i(\boldsymbol{x}) = 1, \ \forall \boldsymbol{x} \in \mathcal{D}. \tag{3.9}$$

In our implementation we use the tensor product of normalized cubic B-splines $\phi(\boldsymbol{x}) = \phi_0(x)\,\phi_0(y)$, where

$$\phi_0(r) = \begin{cases} \frac{2}{3} - |r|^2 + \frac{1}{2}|r|^3, & |r| \leq 1, \\ \frac{1}{6}(2 - |r|)^3, & 1 < |r| \leq 2, \\ 0, & \text{otherwise.} \end{cases}$$

The discrete stream function $\Psi_i(t)$ is obtained by solving a Poisson equation on the grid. Given a discrete grid-based PV field $q_i(t)$ we solve

$$\sum_j \Delta_{ij}\Psi_j = q_i - h_i \tag{3.10}$$

where $h_i = h(\boldsymbol{x}_i)$ is the topography function sampled at gridpoints and Δ_{ij} is an appropriate discretization of the Laplacian. In our implementation, we use a spectral approximation and FFTs, but a finite difference formula may be sufficient.

Finally, the PV field on the grid is approximated from the particles using the relation

$$q_i(t) = \sum_k Q_k \phi_i\left(\boldsymbol{X}_k(t)\right). \tag{3.11}$$

In [28] it is shown that the above formula samples the exact solution of a continuity equation of the form $q_t + \nabla \cdot (q\hat{\boldsymbol{u}}) = 0$ with density function $q(\boldsymbol{x}, t) = \sum_k Q_k \phi(\boldsymbol{x} - \boldsymbol{X}_k)$ and auxiliary velocity field $\hat{\boldsymbol{u}}(\boldsymbol{x}, t)$ appropriately defined. In the present case, although the particle velocity field is given by $\nabla^\perp \Psi(\boldsymbol{x}, t)$ and is therefore divergence-free, this will only hold in an approximate sense for the auxiliary velocity field $\hat{\boldsymbol{u}}(\boldsymbol{x}, t)$.

In the present context of vortex dynamics, the HPM method is related to the classical point vortex flow (see [15] and references therein). The singular point vortices have been regularized by convolution with the basis functions ϕ. The Eulerian grid reduces the complexity of vortex-vortex interactions from $\mathcal{O}(K^2)$ to $\mathcal{O}(K \ln K)$ (using FFT). The construction of the method preserves the Hamiltonian structure of the point vortex flow. However, as noted in the introduction, the HPM method was originally in the setting of compressible flow and is in this sense applicable to more general fluids than the point vortex model.

3.3.2 Properties of the discretization

By construction, the numerical method described above defines a Hamiltonian system. The Hamiltonian is

$$H(\boldsymbol{X}) = -\frac{1}{2}\sum_i \Psi_i(q_i - h_i)\,\Delta x^2$$
$$= -\frac{1}{2}\sum_{i,j}\left[\left(\sum_k Q_k\phi_i(\boldsymbol{X}_k)\right) - h_i\right](\Delta^{-1})_{i,j}\left[\left(\sum_\ell Q_\ell\phi_j(\boldsymbol{X}_\ell)\right) - h_j\right]\Delta x^2. \tag{3.12}$$

Introducing phase space coordinates $\boldsymbol{X} = (X_1, \ldots, X_K, Y_1, \ldots, Y_K)^T$ and symplectic two-form structure matrix

$$B = \begin{bmatrix} 0 & -\operatorname{diag}\boldsymbol{Q} \\ \operatorname{diag}\boldsymbol{Q} & 0 \end{bmatrix},$$

the equations of particle motion are described by

$$B\frac{d\boldsymbol{X}}{dt} = \nabla H(\boldsymbol{X}).$$

The Hamiltonian is a first integral of the dynamics and approximates the total kinetic energy. Additionally, the phase flow is symplectic and consequently volume-preserving on $\mathbb{R}^{2K}$.

To integrate the numerical discretization in time we use the implicit midpoint rule:

$$B\frac{\boldsymbol{X}^{n+1} - \boldsymbol{X}^n}{\Delta t} = \nabla H\left(\frac{\boldsymbol{X}^{n+1} + \boldsymbol{X}^n}{2}\right).$$

The numerical map is symplectic for this problem, implying that volume is preserved in the $2K$-dimensional phase space of particle positions. Also the energy is well-preserved, with fluctuations bounded by a term of $\mathcal{O}(\Delta t^2)$ for long times, consistent with theory reported in [35, 45], see also Section 1.1.

Since the particle PV values Q_k, $k = 1, \ldots, K$ are fixed for the duration of the computation, PV is conserved along particle paths for any motion of the particles. However, since the Q_k play the role of parameters in the specification of HPM, their conservation does not imply a reduction in degrees of freedom of the dynamics in the way exact conservation of H does. On the other hand, the motion of the particles is not arbitrary, but area-preserving in the sense described next. The combination of material conservation of PV in an area-preserving flow is the essential feature of the fine scale motion of ideal fluids that enters into the modern statistical mechanics theories.

Given an arbitrary continuous motion of the particles $\boldsymbol{X}_k(t)$, equations (3.11), (3.10) and (3.8) define a continuum approximate stream function $\Psi(\boldsymbol{x}, t)$, with velocity field

$$\boldsymbol{U}(\boldsymbol{x}, t) = \nabla^\perp\Psi(\boldsymbol{x}, t).$$

Let us define label coordinates $\boldsymbol{a} = (a, b)$ on $\mathcal{D}$ and the Lagrangian flow $\boldsymbol{\chi}(\boldsymbol{a}, t) : \mathcal{D} \times \mathbb{R} \to \mathcal{D}$ induced by $\Psi(\boldsymbol{x}, t)$:

$$\frac{\partial}{\partial t}\boldsymbol{\chi}(\boldsymbol{a}, t) = \boldsymbol{U}(\boldsymbol{\chi}(\boldsymbol{a}, t), t), \tag{3.13}$$

Since $\nabla \cdot \boldsymbol{U} \equiv 0$, the Lagrangian flow $\boldsymbol{\chi}$ is area-preserving on $\mathcal{D}$. That is,

$$|\det \frac{\partial \boldsymbol{\chi}}{\partial \boldsymbol{a}}| \equiv 1.$$

This property is retained under temporal semi-discretization with the implicit midpoint rule. That is, the mapping $\boldsymbol{\chi}^n(\boldsymbol{a}) \mapsto \boldsymbol{\chi}^{n+1}(\boldsymbol{a})$ is area-preserving.

On the other hand, for the numerical method, the particle motion $\boldsymbol{X}_k(t)$ is just given by

$$\boldsymbol{X}_k(t) = \boldsymbol{\chi}(\boldsymbol{X}_k(0), t),$$

i.e. the particle motion is embedded in its own Lagrangian flow. Therefore, the discrete particle motion is area-preserving in the sense that it can be embedded in an area-preserving flow.

Typically, we initialize the particles on a uniform grid[1] with spacing $\Delta a = \Delta x/\kappa$ for some positive integer κ. Let $\boldsymbol{a}_k \in \mathcal{D}$ denote the initial position of the kth particle. Let A_k denote the set of labels in the grid cell centered at $\boldsymbol{a}_k$:

$$A_k = \{\boldsymbol{a} \in \mathcal{D} : |\boldsymbol{a} - \boldsymbol{a}_k|_\infty < \frac{\Delta a}{2}\}.$$

Then $\mathcal{D} = \cup_k \overline{A_k}$. Define a piecewise constant initial vorticity field through

$$Q_0(\boldsymbol{a}) = \sum_k Q_k \mathbf{1}_k(\boldsymbol{a}), \tag{3.14}$$

where $\mathbf{1}_k$ is the characteristic function on A_k. This vorticity field is transported by the flow $\boldsymbol{\chi}(\boldsymbol{a}, t)$ via

$$Q(\boldsymbol{\chi}(\boldsymbol{a}, t), t) = Q_0(\boldsymbol{a}). \tag{3.15}$$

Given any function $f(Q)$, we have

$$\begin{aligned}
\int_{\mathcal{D}} f(Q(\boldsymbol{x}, t))\, d\boldsymbol{x} &= \int_{\mathcal{D}} f(Q(\boldsymbol{\chi}(\boldsymbol{a}, t), t))\big|\det \frac{\partial \boldsymbol{\chi}}{\partial \boldsymbol{a}}\big|\, d\boldsymbol{a} \\
&= \int_{\mathcal{D}} f(Q_0(\boldsymbol{a}))\big|\det \frac{\partial \boldsymbol{\chi}}{\partial \boldsymbol{a}}\big|\, d\boldsymbol{a} \\
&= \sum_k f(Q_k) \int_{A_k} \big|\det \frac{\partial \boldsymbol{\chi}}{\partial \boldsymbol{a}}\big|\, d\boldsymbol{a} \\
&= \sum_k f(Q_k)|A_k| \\
&= \Delta a^2 \sum_k f(Q_k),
\end{aligned} \tag{3.16}$$

[1] For an arbitrary initial particle configuration, the subsequent quadrature could be carried out on the Voronoi cells.

which is constant. In particular, the area associated with any particular level set of PV is conserved.

In this sense we see that the fine scale particle flow trivially conserves all Casimirs, and in particular the polynomials functions

$$C_r = \Delta a^2 \sum_k Q_k^r, \qquad r = 1, 2, \ldots.$$

However, this property does not transfer to the gridded PV field $\boldsymbol{q}$. That is, the grid-based analogs

$$\hat{C}_r = \Delta x^2 \sum_i q_i^r, \qquad r = 1, 2, \ldots$$

are not conserved in general. The sole exception is the total circulation $\hat{C}_1$ for which we have, using the third property of (3.9),

$$\hat{C}_1 = \sum_i q_i \Delta x^2 = \sum_i \sum_k Q_k \phi_i(X_k) \kappa^2 \Delta a^2 = \sum_k Q_k \kappa^2 \Delta a^2,$$

which is independent of time. For arbitrary nonlinear $f(q)$, one would not expect the quantity $\sum_i f(q_i)\Delta x^2$ to be invariant in general.

In Figure 3.1 we plot the relative drift

$$\varepsilon_{\text{rel}}[H](t) = \left| \frac{H(t) - H(0)}{H(0)} \right|$$

in the quantities H and $\hat{C}_r$, $r = 2, \ldots 4$ as a functions of time during a typical simulation (the experiment described in Section 3.6.3, for the case $\gamma = 0$, $\delta = 90$). The circulation is preserved to machine precision and is not shown. The energy oscillations are bounded by

$$\varepsilon_{\text{rel}}[H](t) < 2.1 \times 10^{-4},$$

and the bound decreases quadratically with stepsize. The higher order vorticity moments are not preserved, and encounter relative drifts

$$\max_t \varepsilon_{\text{rel}}[\hat{C}_2](t) = 0.74, \quad \max_t \varepsilon_{\text{rel}}[\hat{C}_3](t) = 16.9, \quad \max_t \varepsilon_{\text{rel}}[\hat{C}_4](t) = 16.5.$$

Clearly, these are not conserved.

In some cases, it is useful to consider the bulk motion of the fluid to be prescribed by a time dependent stream function $\Psi(\boldsymbol{x}, t)$, and consider the motion of a typical particle embedded in the flow. The motion of such a particle satisfies a nonautonomous Hamiltonian system. This point of view and its coupling to the dynamics is studied in [5]. The parcel Hamiltonian becomes

$$\tilde{\mathcal{H}} = \int q_0(\boldsymbol{a}) \Psi(\boldsymbol{\chi}(\boldsymbol{a}, t), t)\, d\boldsymbol{a},$$

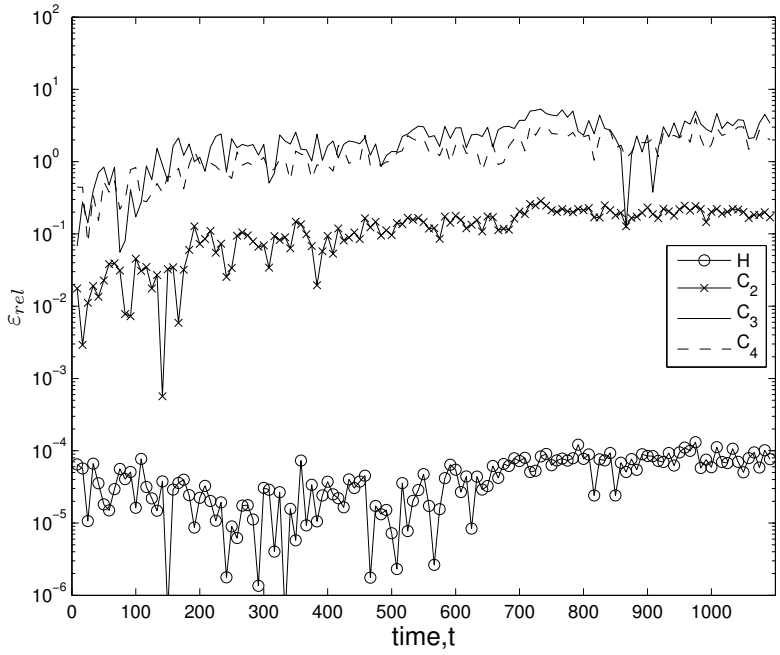

FIGURE 3.1: Relative change in energy H and higher vorticity moments $\hat{C}_2$, $\hat{C}_3$, $\hat{C}_4$ during the simulation described in Section 3.6.3, for the case $\gamma = 0$, $\delta = 90$.

and the dynamics on label space (3.13) satisfy

$$q_0(\boldsymbol{a})\frac{d}{dt}\boldsymbol{\chi}(\boldsymbol{a},t) = J\frac{\delta\tilde{\mathcal{H}}}{\delta\boldsymbol{\chi}} = q_0(\boldsymbol{a})\nabla^{\perp}\Psi(\boldsymbol{\chi}(\boldsymbol{a},t),t), \quad \forall \boldsymbol{a}.$$

where $J = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix}$.

Similarly, for the HPM discretization,

$$Q_k\frac{\partial}{\partial t}\boldsymbol{X}_k = J\nabla_{\boldsymbol{X}_k(t)}\tilde{H} = Q_k\nabla^{\perp}\Psi(\boldsymbol{X}_k(t),t), \tag{3.17}$$

where

$$\tilde{H} = \sum_k Q_k\Psi(\boldsymbol{X}_k(t),t). \tag{3.18}$$

3.4 A Lagrangian statistical model based on canonical particle distributions

Due to its Lagrangian nature, the HPM method has similarities with the point-vortex method whose statistical mechanics was considered by Onsager [65]. The particle motion can be considered a regularized point vortex method. The phase space of the HPM method is bounded: it is simply $\mathcal{D}^K$. In contrast to the point vortex method, the range of energy for the HPM method is also bounded (for finite Q_k, $k = 1, \ldots, K$). If one considers possible configurations for a given energy, as the energy level becomes large enough the available phase space eventually starts to decrease. In other words, the HPM method supports a negative temperature regime.

In this section we construct a statistical equilibrium theory in the natural phase space of HPM particle positions $\boldsymbol{X}_k \in \mathcal{D}$. However, in some cases it may be preferable to directly consider the statistics of the coarse-grain vorticity field $\boldsymbol{q}$ on the grid (3.11), which allows comparision with the existing equilibrium field theories. In the next section we present an Eulerian approximate statistical model from this point of view. To distinguish the two, we refer to the theory in this section as the Lagrangian statistical mechanical model.

Let us consider the statistics of a single distinguished particle in contact with the reservoir formed by all other particles. Recall that the motion of such a particle obeys a nonautonomous Hamiltonian system (3.17) with Hamiltonian (3.18). The energy contribution of particle k is $Q_k\Psi(\boldsymbol{X}_k, t)$. We expand the stream function about the ensemble mean field

$$\Psi(\boldsymbol{x}, t) = \langle\Psi(\boldsymbol{x})\rangle + \delta\Psi(\boldsymbol{x}, t).$$

Neglecting the long time effects of the perturbation part $\delta\Psi$, we obtain the canonical distribution for a distinguished particle (see Section 1.3.1 about canonical sampling)

$$\rho_k(\boldsymbol{x}) = \frac{1}{\zeta_k}\exp\left[-\beta\langle\Psi(\boldsymbol{x})\rangle Q_k\right], \qquad \zeta_k = \int_{\boldsymbol{x}\in\mathcal{D}} \exp\left[-\beta\langle\Psi(\boldsymbol{x})\rangle Q_k\right] d\boldsymbol{x}. \quad (3.19)$$

Figure 3.2 compares typical functions $\rho_k(x)$ with histograms of position data for two arbitrarily chosen particles with $Q_{k^+} = 1.098$ and $Q_{k^-} = -2.165$ obtained from HPM simulations with normally distributed $\{Q_k\}$. We observe good agreement. Due to the choice of topography in Section 3.6 and normally distributed Q_k, the distributions ρ_k are uniform in the y direction.

The one-particle canonical statistics can be used to construct a mean field theory. For particle k the one-particle statistics is (3.19). This quantity gives the probability that $\boldsymbol{X}_k$ is near $\boldsymbol{x} \in \mathcal{D}$. Next consider coordinates $\boldsymbol{\Xi} = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_K) \in \mathcal{D}^K$ on the particle phase space, and the product distribution

$$\rho(\boldsymbol{\Xi}) = \prod_k \rho_k(\boldsymbol{\xi}_k),$$

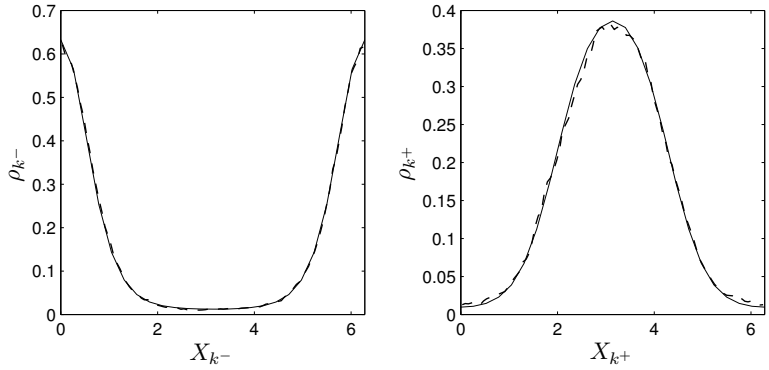

FIGURE 3.2: Histograms of x-component of position for two distinct particles (dash line), compared with the predicted canonical distribution (solid line).

which governs the probability of particle configurations under the modelling assumption that the particle positions are independent.

To each $\boldsymbol{\Xi} \in \mathcal{D}^K$, is an associated grid-based PV field $\boldsymbol{q}(\boldsymbol{\Xi})$ with

$$q_i(\boldsymbol{\Xi}) = \sum_k Q_k \phi_i(\boldsymbol{\xi}_k).$$

The ensemble average PV field is

$$\langle q_i \rangle = \int q_i(\boldsymbol{\Xi}) \rho(\boldsymbol{\Xi}) \, d\boldsymbol{\Xi},$$

which can be simplified as follows

$$\begin{aligned}
\langle q_i \rangle &= \int_{\mathcal{D}^K} \sum_k Q_k \phi_i(\boldsymbol{\xi}_k) \prod_\ell \rho_\ell(\boldsymbol{\xi}_\ell) \, d\boldsymbol{\xi}_1 \cdots d\boldsymbol{\xi}_K \\
&= \sum_k Q_k \int_{\mathcal{D}} \phi_i(\boldsymbol{x}) \rho_k(\boldsymbol{x}) \, d\boldsymbol{x} \\
&= \sum_k Q_k \frac{\int_{\mathcal{D}} \phi_i(\boldsymbol{x}) \exp\left[-\beta \langle \psi(\boldsymbol{x}) \rangle Q_k\right] \, d\boldsymbol{x}}{\int_{\mathcal{D}} \exp\left[-\beta \langle \psi(\boldsymbol{x}) \rangle Q_k\right] \, d\boldsymbol{x}} \\
&= \sum_k Q_k \langle \phi_i \rangle_k
\end{aligned}$$

where $\langle \cdot \rangle_k$ is the ensemble average in the measure (3.19).

If we consider a piecewise constant vorticity distribution with K_ℓ particles with PV σ_ℓ, the relations above can be expressed as

$$\langle q_i \rangle = \sum_\ell \sigma_\ell K_\ell \frac{\int_{\mathcal{D}} \phi_i(\boldsymbol{x}) \exp\left[-\beta \langle \psi(\boldsymbol{x}) \rangle \sigma_\ell\right] \, d\boldsymbol{x}}{\int_{\mathcal{D}} \exp\left[-\beta \langle \psi(\boldsymbol{x}) \rangle \sigma_\ell\right] \, d\boldsymbol{x}} = \sum_\ell \sigma_\ell p_{i,\ell} \tag{3.20}$$

where

$$p_{i,\ell} = K_\ell \langle \phi_i \rangle_\ell. \tag{3.21}$$

This quantity is the proportion of time that a particle with PV σ_ℓ spends within the support of grid point $\boldsymbol{x}_i$, weighted by the kernel function, times the number of such particles.

To compute the mean state, we approximate the integral in (3.20) by quadrature at the grid points

$$\langle q_i \rangle \approx \sum_\ell \sigma_\ell K_\ell \frac{\sum_j \phi_i(\boldsymbol{x}_j) \exp\left[-\beta \langle \psi_j \rangle \sigma_\ell\right]}{\sum_j \exp\left[-\beta \langle \psi_j \rangle \sigma_\ell\right]}.$$

This relation is solved together with

$$\Delta_{ij} \langle \psi_j \rangle = \langle q_i \rangle - h_i,$$

and the constraint relation

$$H(\langle \boldsymbol{q} \rangle) = H_0,$$

which specifies the value of β.

Our Lagrangian statistical theory for the HPM method is analogous to the canonical theory of Ellis, Haven & Turkington [22], inasmuch as the energy is treated microcanonically through the specification of β, and the fine-scale vorticity conservation is treated canonically.

3.5 Eulerian statistical model for HPM

The continuum statistical mechanics theories of Miller [54, 55] and Robert [71, 72] can be constructed using a two-level discretization of the continuum vorticity field. The microscopic configuration space consists of permutations of a piecewise uniform vorticity field, assuming constant values on each cell of a fine mesh. The macroscopic vorticity field is the local average of the microscopic field on an embedding coarse mesh. The continuum theory is obtained by first letting the fine mesh size tend to zero for fixed coarse grid mesh size, and subsequently taking the continuum limit of the coarse mesh.

A similar approach—neglecting the continuum limits—can be used to construct a discrete statistical model for the HPM method. Keeping in mind the interpolating continuum flow (3.15), we define p_i^ℓ to be the probability of observing $Q(\boldsymbol{x}_i, t) = \sigma_\ell$ near grid point $\boldsymbol{x}_i$. Then p_i^ℓ has the properties

$$\sum_i p_i^\ell \Delta x^2 = g_\ell, \qquad \sum_\ell p_i^\ell = 1. \tag{3.22}$$

The first of these says the area of PV level sets is conserved and is the discrete analog of (3.6). The second says the flow is incompressible.

It is natural to associate p_i^ℓ with the characteristic functions at the grid points, smoothed by the HPM basis functions ϕ. Denote by $\mathbb{K}_\ell$ the index set of particles with vorticity level σ_ℓ ($\ell = 1, \ldots, \Lambda$, for some $\Lambda \leq K$):

$$\mathbb{K}_\ell = \{1 \leq k \leq K : Q_k = \sigma_\ell\}.$$

Define the function

$$\varphi_i^\ell = \sum_{k \in \mathbb{K}_\ell} \phi_i(\boldsymbol{X}_k) \frac{1}{\kappa^2}.$$

If the particles are initialized on a uniform grid of spacing $\Delta a = \Delta x/\kappa$, for $\kappa \geq 1$ an integer, then φ_i^ℓ has the required properties (3.22). To construct a Miller/Robert theory for the HPM method, we would initialize the particles on such a uniform grid, and consider permutations of the Q_k as an approximation of the configuration space.

The motion of particles in the HPM method conserves energy, as pointed out in the Section 3.3. It is therefore necessary to further restrict the sample space to those permutations of PV that preserve the initial energy to within some tolerance. Define the coarse grain mean potential vorticity by

$$\langle q_i \rangle = \sum_\ell p_i^\ell \sigma_\ell, \tag{3.23}$$

the coarse grain mean stream function by

$$\Delta \langle \boldsymbol{\Psi} \rangle = \langle \boldsymbol{q} \rangle - \boldsymbol{h}, \tag{3.24}$$

and the energy of the mean field by

$$H(\langle \boldsymbol{q} \rangle) = -\frac{1}{2} \langle \boldsymbol{\Psi} \rangle^T (\langle \boldsymbol{q} \rangle - \boldsymbol{h}) \, \Delta x^2.$$

Substituting φ_i^ℓ for p_i^ℓ in (3.23), the above definitions are consistent with the (coarse-grain) grid quantities $\boldsymbol{q}$, $\boldsymbol{\Psi}$ and H given in (3.11), (3.10), and (3.12), respectively.

A microcanonical statistical model analogous to the Miller/Robert approach proceeds at this point by introducing the Shannon information entropy

$$S[p] = -\sum_{i,\ell} p_i^\ell \ln p_i^\ell, \tag{3.25}$$

and maximizing this function with respect to p_i^ℓ subject to contraints of observed values of energy, $H(\langle \boldsymbol{q} \rangle) = H_0$, and the conditions (3.22).

Instead we take here the alternative approach proposed by Ellis, Haven & Turkington [22], which assumes a canonical ensemble with respect to the higher order Casimirs, as determined by a prior distribution over pointwise vorticity, in combination with a microcanonical distribution with respect to $\mathcal{H}$ and $\mathcal{C}_1$. This is consistent with observations of inviscid fluids, where $\mathcal{H}$ and $\mathcal{C}_1$ depend

only on the large scale vorticity whereas the $\mathcal{C}_r$, $r > 1$ depend on the fine scale detailed vorticity and the length scale of averaging. To that end we drop the requirement that p_i^ℓ satisfy the first condition of (3.22).

Given a set of particles initialized on a uniform grid with PV values Q_k, $k = 1, \ldots, K$, we consider the associated piecewise constant continuum vorticity field as described in (3.14)–(3.15). To each vorticity level set σ_ℓ, $\ell = 1, \ldots, \Lambda$ we associate the fractional area

$$\Pi_\ell = \frac{K_\ell \Delta a^2}{|\mathcal{D}|},$$

where K_ℓ is the number of particles with vorticity σ_ℓ and $|\mathcal{D}|$ is the total area of $\mathcal{D}$. Note that $\sum_\ell \Pi_\ell = 1$. We take Π_ℓ to be the prior distribution on PV. Given no other information about the flow, Π_ℓ is the probability of observing PV value σ_ℓ at an arbitrarily chosen point in $\mathcal{D}$. The probability is uniform in space.

To determine the probability distribution p_i^ℓ we maximize the relative entropy

$$S[p, \Pi] = -\sum_{i,\ell} p_i^\ell \ln \frac{p_i^\ell}{\Pi_\ell}. \tag{3.26}$$

Given no other information about the system, we can maximize this entropy as a function of p_i^ℓ to find

$$p_i^\ell = \Pi_\ell,$$

which is the prior distribution at each point on the grid, confirming the earlier statement.

Instead we wish to maximize (3.26) subject to microcanonical constraints on the energy

$$\tilde{E} = H(\langle \boldsymbol{q} \rangle) - H_0 = 0, \tag{3.27}$$

and the circulation

$$\tilde{\Gamma} = \hat{C}_1(\langle \boldsymbol{q} \rangle) - \hat{C}_1(\boldsymbol{q}(0)) = 0, \tag{3.28}$$

as well as the normalization constraints

$$\tilde{N}_i = \sum_\ell p_i^\ell - 1 = 0, \quad \forall i. \tag{3.29}$$

Introducing Lagrange multipliers β, α and λ_i, respectively, for these constraints, we solve

$$\frac{\partial S}{\partial p_i^\ell} + \beta \frac{\partial \tilde{E}}{\partial p_i^\ell} + \alpha \frac{\partial \tilde{\Gamma}}{\partial p_i^\ell} + \sum_j \lambda_j \frac{\partial \tilde{N}_j}{\partial p_i^\ell} = 0.$$

The respective derivatives are

$$\begin{aligned}
\frac{\partial S}{\partial p_i^\ell} &= -(\ln \frac{p_i^\ell}{\Pi_\ell} + 1), \\
\frac{\partial \tilde{E}}{\partial p_i^\ell} &= \sum_j \frac{\partial H}{\partial \langle q_j \rangle} \frac{\partial \langle q_j \rangle}{\partial p_i^\ell} = -\sum_j \langle \Psi_j \rangle \sigma_\ell \delta_{ij} \, \Delta x^2 = -\langle \Psi_i \rangle \sigma_\ell \, \Delta x^2, \\
\frac{\partial \tilde{\Gamma}}{\partial p_i^\ell} &= \sigma_\ell \Delta x^2, \\
\frac{\partial \tilde{N}_j}{\partial p_i^\ell} &= \delta_{ij}.
\end{aligned}$$

Putting this all together, an extreme entropy state must have

$$\ln p_i^\ell = \ln \Pi_\ell - 1 - \beta \langle \Psi_i \rangle \sigma_\ell + \alpha \sigma_\ell + \lambda_i,$$

where a constant Δx^2 has been absorbed into α and β. Solving for p_i^ℓ yields the equilibrium distribution

$$p_i^\ell = N_i^{-1} \exp\left[(-\beta \langle \Psi \rangle_i + \alpha)\, \sigma_\ell\right] \Pi_\ell, \tag{3.30}$$

where β and α can be chosen to satisfy the contraints (3.27) and (3.28), and the partition function N_i is given by

$$N_i = \sum_\ell \exp\left[(-\beta \langle \Psi \rangle_i + \alpha)\, \sigma_\ell\right] \Pi_\ell. \tag{3.31}$$

The relation (3.30) can be combined with (3.23) and (3.24) to solve for prospective mean fields. The mean stream function $\langle \boldsymbol{\Psi} \rangle$ is found by solving

$$\sum_j \Delta_{ij} \langle \Psi_j \rangle = \frac{\sum_\ell \sigma_\ell \exp\left[(-\beta \langle \Psi_i \rangle + \alpha)\, \sigma_\ell\right] \Pi_\ell}{\sum_\ell \exp\left[(-\beta \langle \Psi_i \rangle + \alpha)\, \sigma_\ell\right] \Pi_\ell} - h(\boldsymbol{x}_i). \tag{3.32}$$

together with the constraints (3.27) and (3.28).

The EHT theory is microcanonical with respect to the energy and circulation, in the sense that the parameters β and α are chosen as Lagrange multipliers to ensure that the resulting mean field assumes desired values of these quantities. It is canonical with respect to higher order Casimir's in the sense that the fine scale vorticity is specified as a distribution.

3.6 Numerical Verification of the HPM Statistical Equilibrium Theories

In this section we compare the predicted mean fields $\langle \boldsymbol{q} \rangle$ and $\langle \boldsymbol{\Psi} \rangle$ of the discrete equilibrium statistical models from the previous sections, with long time average mean fields $\overline{\boldsymbol{q}}$ and $\overline{\boldsymbol{\Psi}}$ computed from numerical simulations with the HPM

method, under the assumption that the simulated solution is approximately ergodic. It should be noted that the probability distributions (3.21) and (3.30) predict much more than just the mean states $\langle \boldsymbol{q} \rangle$ and $\langle \boldsymbol{\Psi} \rangle$, so our comparison is necessarily a limited one. Yet from a numerical point of view, correct representation of the mean state is a minimal requirement, as it sets the statistical background for dynamics.

The theoretical mean fields (3.20) and (3.32) based on the Lagrangian and Eulerian statistical models are computed numerically. Due to pointwise conservation of PV on the particles, and the construction (3.11), the space of grid-based PV fields is bounded, as is the partition function (3.31). For a given particle field $\boldsymbol{Q}$ and values for the constraints we solve for the mean fields (3.20) and (3.32) plus associated Lagrange multipliers using a modified Newton iteration. These fields are compared with average fields generated by long time simulations.

For the numerical simulations we use the test setup of Abramov & Majda[1]. We choose grid resolution $M = 24$. The topography is a function of x only, specifically

$$h(x, y) = 0.2 \cos x + 0.4 \cos 2x,$$

which is intended to make departures from Gaussian PV theory readily observable (see below).

The integrations were carried out using a step size of $\Delta t = 2/M$ on the interval $t \in [0, t_0 + T]$. The solutions are averaged over the time interval $t \in [t_0, t_0 + T]$, where t_0 is the time required for decorrelation of the initial condition. In all experiments we use $t_0 = 10^3$ and $T = 10^4$. Longer simulations with $T = 10^6$ were also run with no observable difference in the results. The implicit midpoint rule nonlinear relations were solved to machine precision.

All simulations were carried out with $\kappa = 1$. Together with the low value of M, this implies the simulations were highly under-resolved. This has the double effect of allowing us to stretch the limits of the discrete statistical models, for which various approximations were made, and to allow the system to sample the available phase space (assuming ergodicity) in a reasonably short simulation interval.

We construct initial conditions with a desired prior distribution and energy value. The mean state (3.30) is fully defined by these quantities. If the dynamics is sufficiently ergodic, then the time average mean stream function $\overline{\boldsymbol{\Psi}}$ and mean potential vorticity $\overline{\boldsymbol{q}}$ should agree with the ensemble averages (3.24) and (3.23)

Given a continuous prior distribution on vorticity $\Pi(\sigma)$, we define particle PV values Q_k as follows. The number of particles is $K = \kappa^2 M^2$. We discretize the range of vorticity σ into Λ equal partitions of size $\Delta\sigma$ where

$$\sigma_\ell = \sigma_0 + \ell \Delta\sigma, \quad \ell = 1, \ldots, \Lambda.$$

We choose the number of particles with vorticity $\sigma_{\ell+1/2} = (\sigma_\ell + \sigma_{\ell+1})/2$ to be

$$K_{\ell+1/2} = \lfloor K \int_{\sigma_\ell}^{\sigma_{\ell+1}} \Pi(\sigma)\, d\sigma \rfloor.$$

Any remainder particles are assigned the values of the consecutive most probable level sets.

The particles are initially placed on a uniform grid of spacing $\Delta a = \Delta x/\kappa$ in each direction. Using Monte Carlo simulations, the PV values are randomly permuted until a configuration is found within desired total energy (grid function) $H_0 \pm 0.01$. In all simulations, the target energy was $H_0 = 7$, and the total circulation was $C_1(\boldsymbol{Q}) = 0$, consistent with [1]. The Lagrange multipliers β and α follow from the constraints of total energy and circulation.

3.6.1 Normally distributed PV

From the classical energy-enstrophy theory of Kraichnan and others [11, 40, 76] it is known that if the PV field is normally distributed, the mean field relation should be linear of the form (3.7). To verify this for the HPM method, we draw the particle vorticities from a zero-mean Gaussian prior distribution

$$Q_k \sim \Pi(\sigma) = \exp\left(-\frac{\sigma^2}{2\theta^2}\right).$$

In this case the EHT theory yields (in the semi-discrete case)

$$p_i(\sigma) = N_i^{-1} \exp\left[-\beta\langle\Psi_i\rangle\sigma\right]\Pi(\sigma)$$

which is continuous in the PV σ. This density can be exactly integrated to yield the linear mean field relation

$$\langle q_i\rangle = -\beta\theta^2\langle\Psi_i\rangle.$$

We choose β and θ to specify energy $H_0 = 7$ and enstrophy $C_2 = 40$.

In the left panel of Figure 3.3, the locus of data points $(\overline{\Psi}_i, \overline{q}_i)$ is plotted for the time-averaged fields. The vorticity-stream function relation is nearly linear as predicted. Due to the finite sampling of the Gaussian distribution, the simulation data is not precisely linear. The Eulerian statistical model (3.30) yields a more linear mean field prediction (dash line), but the Lagrangian statistical model (3.21) more precisely fits the simulation data (solid line).

Due to the linearity and isotropy of (3.7) and (3.8), the mean stream function $\langle\boldsymbol{\psi}\rangle$ satisfies a Helmholtz equation and is expected to be independent of y due to the special choice of topography. In the right panel we observe that the mean stream function is indeed independent of y.

3.6.2 Skew PV distributions

In [1], Abramov and Majda show that nonzero values of the third moment of potential vorticity can cause significant deviation from the statistical predictions of the normally distributed PV case. They use the Poisson discretization of Zeitlin [88] to solve the QG model. On an $M \times M$ grid the Zeitlin method conserves energy and approximations of the first M moments of potential vorticity $\hat{C}_r$, $r = 1, \ldots, M$, see Section 1.2

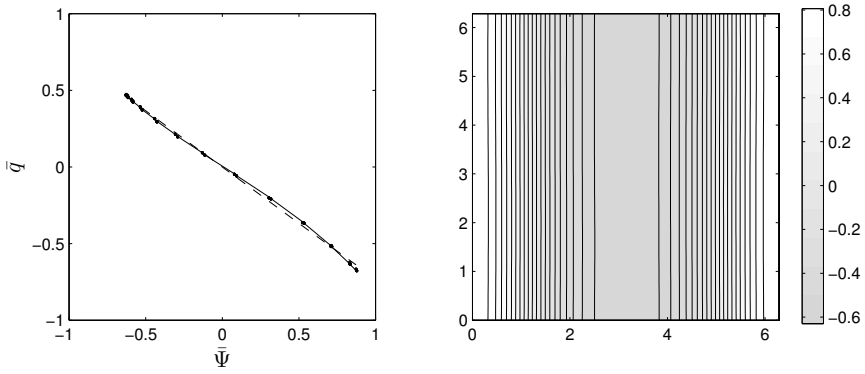

FIGURE 3.3: Normally distributed PV on the particles. The scatter plot of mean fields (left) with linear fit: points are locus $(\overline{q}_i, \overline{\Psi}_i)$, the theoretical prediction based on (3.21) is solid line and on (3.30)—dash line. Mean stream function (right).

We generate initial conditions $\boldsymbol{Q}$ from the shifted gamma-distribution [13]:

$$\Pi(\sigma) = \frac{1}{C_2|\lambda|} R\left(\frac{1}{C_2\lambda}(\sigma + \lambda^{-1}); \frac{1}{C_2\lambda^2}\right),$$

where $R(z;a) = \Gamma(a)^{-1} z^{a-1} \exp(-z)$ for $z \geq 0$ and $R = 0$ otherwise with the Gamma function Γ, and

$$\gamma = \frac{C_3}{C_2^{3/2}} = 2C_2^{1/2}\lambda$$

is the skewness of the distribution. We take $C_2 = 40$ and $\gamma = 0$, 2, 4 and 6 to compare the results of [1] with the HPM method.

Figures 3.4 and 3.5 gives the $(\overline{\Psi}_i, \overline{q}_i)$ loci for the time-averaged fields, for these values of γ. Figure 3.6 illustrates the associated mean stream functions. The solutions are reminiscent of those reported in [1], but there are some differences due to the details of the methods.

For the case $\gamma = 0$, the energy-enstrophy theory predicts a linear relation (3.7) between mean PV and mean stream function, as well as a layered mean stream function. These predictions are confirmed in the upper left panels of Figures 3.4, 3.5 and 3.6. For $\gamma > 0$, there is significant nonlinearity in the mean field relation and vortical structures observable in the mean stream function.

Also shown in Figures 3.4 and 3.5 are the theoretical mean states predicted by the discrete statistical equilibrium theories in Sections 3.4 and 3.5. The Lagrangian statistical model (3.20) is shown in Figure 3.4 and the Eulerian model (3.32) in Figure 3.5. Both models predict the mean states very well.

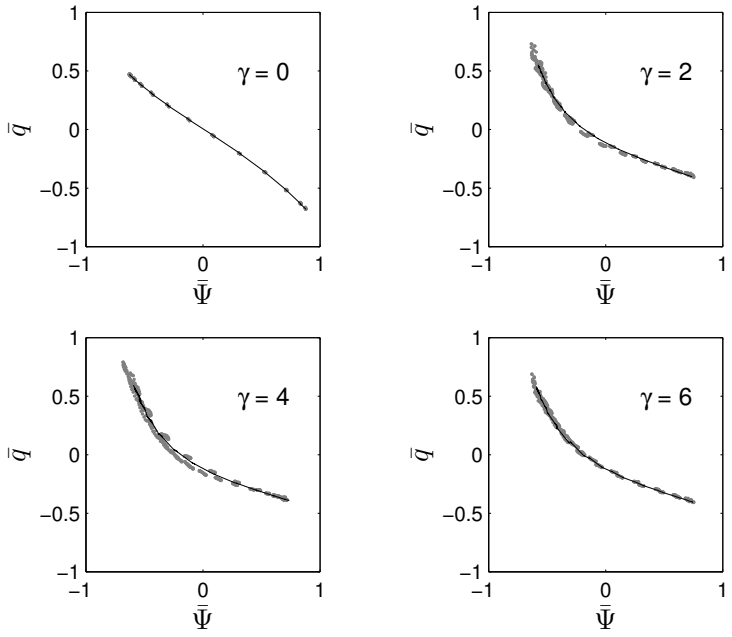

FIGURE 3.4: Locus $(\overline{q}_i, \overline{\Psi}_i)$ for skewed PV distributions, γ=0, 2, 4 and 6 (points). The theoretical prediction based on (3.21) (line).

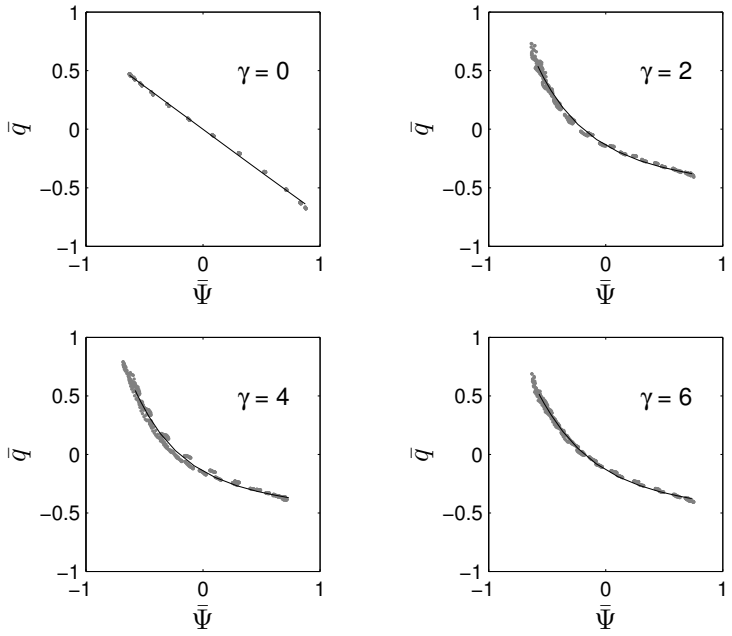

FIGURE 3.5: Locus $(\overline{q}_i, \overline{\Psi}_i)$ for skewed PV distributions, γ=0, 2, 4 and 6 (points). The theoretical prediction based on (3.30) (line).

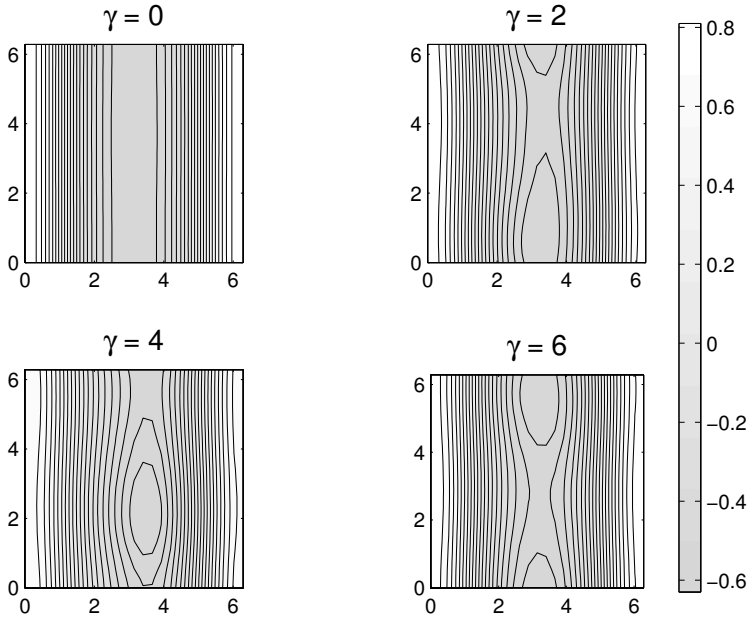

FIGURE 3.6: Mean stream functions (time averages) for skewed PV distributions, γ=0, 2, 4 and 6. For nonzero skewness $\gamma \neq 0$ the stream function is two-dimensional, despite one-dimensional topography.

3.6.3 PV distributions with kurtosis

Abramov & Majda [1] also conjecture that the higher-order moments C_r, $r \geq 4$, are statistically irrelevant for predicting the large-scale mean flow, based on the observation that the experiments agreed well with the energy-enstrophy mean field theory (3.7) in the case $\gamma = 0$, despite the fact that the moments $\hat{C}_r$, $r \geq 4$ were nonzero as arbitrarily determined by their initialization procedure, and conserved by the method.

To investigate this conjecture we choose initial distributions $\boldsymbol{Q}$ having skewness $\gamma = 0$ and nonzero kurtosis (scaled fourth moment of PV),

$$\delta = \frac{C_4}{C_2^2} - 3.$$

In this case we generated the initial particle PV field by first drawing the Q_k from a uniform distribution and then projecting onto the constraint set $\{H_0 = 7,\ C_1 = 0,\ C_2 = 40,\ C_3 = 0,\ C_4 = (\delta + 3)C_2^2\}$.

Figures 3.7 and 3.8 show the mean field relations $(\overline{q}_i, \overline{\Psi}_i)$ for increasing $\delta = 0$, 10, 50 and 90.

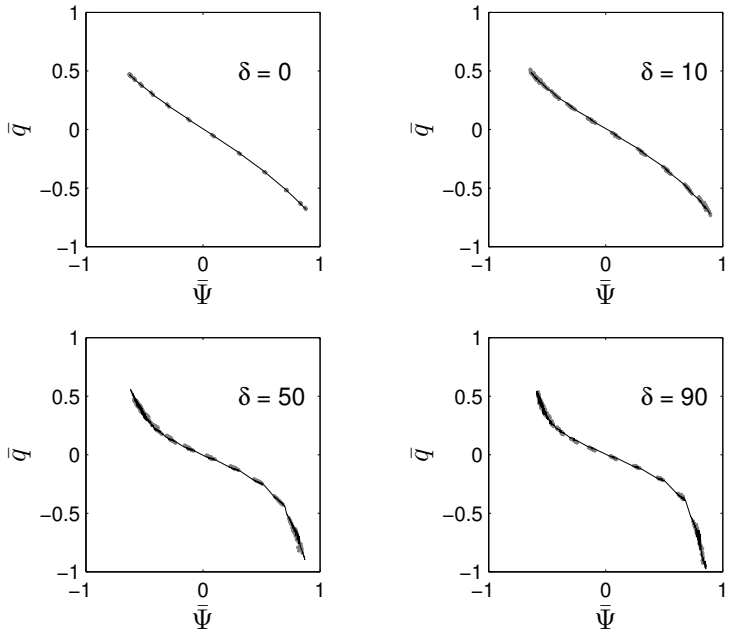

FIGURE 3.7: Locus $(\overline{q}_i, \overline{\Psi}_i)$ for kurtotic PV distributions, δ=0, 10, 50 and 90 (points). The theoretical prediction based on (3.21) (line).

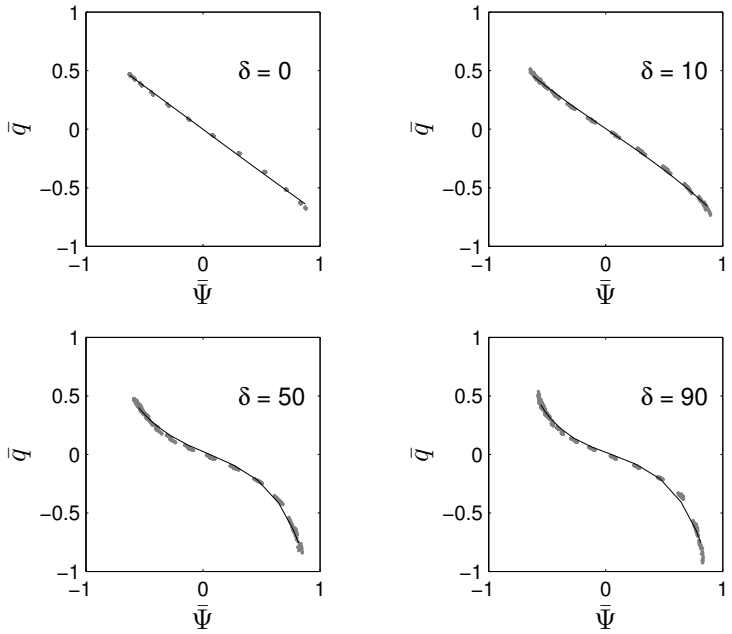

FIGURE 3.8: Locus $(\overline{q}_i, \overline{\Psi}_i)$ for kurtotic PV distributions, δ=0, 10, 50 and 90 (points). The theoretical prediction based on (3.30) (line).

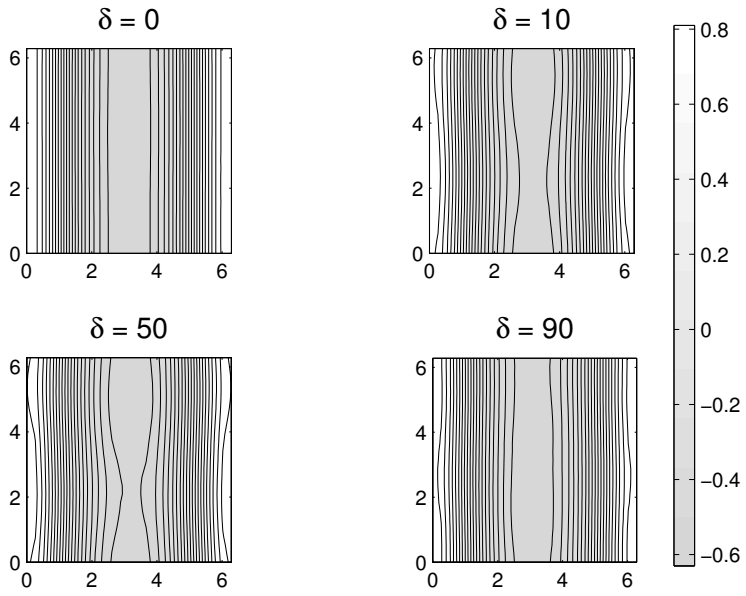

FIGURE 3.9: Mean stream functions (time averages) for kurtotic PV distributions, δ=0, 10, 50 and 90. For nonzero kurtosis $\delta \neq 0$ the stream function is two-dimensional, despite one-dimensional topography.

The corresponding mean stream functions are shown in Figure 3.9. We observe that nontrivial kurtosis may also significantly influence the mean field statistics, which disproves the conjecture of [1].

Again we observe in Figures 3.7 and 3.8 that both (3.20) and (3.32) do an excellent job of predicting the mean states.

3.7 Conclusions

The HPM method, as adapted for 2D incompressible flow, conserves total energy by construction. Each particle is assigned a constant value of potential vorticity at initialization, and this discrete PV field is conserved point-wise, as the particles evolve in the divergence-free flow. In this sense, PV conservation induces no reduction in degrees of freedom on the dynamics. At the coarse scale, the vorticity field on the mesh satisfies conservation of energy and total circulation, but exhibits significant drift for nonlinear PV functionals. This is consistent with what would be observed if a coarse-graining procedure were applied to a real inviscid flow.

A maximum entropy theory based *only* on energy and circulation would predict zero mean flow (i.e. $\langle\psi\rangle \equiv 0$). In contrast to this prediction, we have

demonstrated in this chapter that the HPM method has a much richer statistical mechanics, with nonlinear mean field relations similar to those of [1], and consistent with the canonical EHT theory [22]. In particular, we have demonstrated that both the third and fourth moments of PV ($\mathcal{C}_3$ and $\mathcal{C}_4$) can significantly affect the mean field relation. The latter result disproves a conjecture of Abramov & Majda [1].

We have also presented two statistical mechanics models for the HPM method, a Lagrangian and an Eulerian model. The Eulerian model is analogous to the EHT theory, which uses a canonical treatment of fine-scale vorticity in the form of a prior distribution, and enforces conservation of energy and total circulation through the use of Lagrange multipliers. In the present case, the prior distribution characterizes the particle vorticity field, and the energy and circulation are conserved at the grid scale. The Lagrangian statistical model is constructed on the phase space of particle positions, considering each particle to be immersed in a reservoir defined by the mean flow. The fine scale, particle statistics are given by canonical ensemble distributions, and the temperature parameter is used as a Lagrange multiplier to enforce energy conservation. Mean states computed with both statistical models compare very well with the long time simulation data.

Although PV is simply assigned to particles and its conservation does not imply any dynamic constraint on the evolution, an appeal to the Lagrangian statistical model suggests that a particle's PV value determines its response to the mean flow, and thereby its residence time in any particular region of the flow domain. Via the basis functions ϕ, the local residence time is translated to the grid scale where the coarse-grain dynamics is governed by energy conservation.

The essential ingredients of the Miller/Robert and EHT statistical theories are the fine scale point-wise advection of PV and the coarse scaling associated with the stream function, under the constraint of energy conservation. The HPM method retains these features under discretization, and for this reason its equilibrium statistics are analogous to those theories. From the numerical experiments we can conclude that the HPM method is free of artificial dissipation or other errors that might destroy the equilibrium statistical mechanics. For the experiments conducted, the discrete dynamics is also sufficiently ergodic that the averages are well approximated.

Chapter 4

A thermostat closure for point vortices

4.1 Background

Inviscid fluid models are natural in a number of application areas, such as atmosphere and ocean science, where the Reynolds numbers are so large as to be effectively infinite. These flows are characterized by conservation of total energy, the cascade of vorticity to ever finer scales, and sensitive dependence on initial conditions [74].

For the numerical simulation of such flows, the lack of a viscous diffusion length scale presents the challenge that due to the vorticity cascade, any direct discretization of the equations of motion must eventually become underresolved, as vorticity is transported to scales below the grid resolution. It therefore becomes necessary to close the numerical model by some means. Any finite numerical discretization implies a closure of some kind, whether explicitly modeled or implied by the discretization [21].

The most common approach is the introduction of artificial viscosity, either through modification of the fluid equations to include (hyper-)viscosity, or through the use of stabilized discretizations, for which the viscous terms appear in a modified equation analysis [38]. In either case the viscous length scale must be on the order of the grid resolution to be effective. One disadvantage with a viscous closure model is that it precludes an upscale cascade of vorticity, thereby suppressing hydrodynamic instabilities in geophysical flows, especially in three dimensions. Alternatively, methods can be constructed that preserve the discrete total energy exactly. However, this is achieved via a nonphysical re-injection of the energy from sub-gridscale vorticity at the large scales [63].

A proper closure model should distinguish between resolved and unresolved scales and account for the exchange between these. In this chapter we consider a simple two-scale point vortex flow, consisting of a small number of vortices with large circulation and a very large number of vortices with a much smaller circulation. We seek a simplified computational model for the aggregate behavior of the small-scale point vortices.

The situation is reminiscent of statistical mechanics in the canonical ensem-

ble, in which a system of particles is in thermal equilibrium with a reservoir. This point of view has been exploited by Bühler [9] in a numerical/statistical investigation of the work of Onsager [65], and our goal here is to reproduce the results of [9] without explicitly accounting for the individual motions of the reservoir of small vortices.

A thermostat is a tool used in molecular dynamics to model a system in thermal equilibrium with a reservoir; such thermostats may be either stochastic (e.g. Langevin dynamics) or deterministic. In a Langevin dynamics simulation a stochastic perturbation is introduced in the force field together with a dissipative term; these terms are maintained in balance so as to preserve the canonical ensemble. With a dynamical (or deterministic) thermostat, by contrast, the system is augmented by a few degrees of freedom that model the exchange with the reservoir. The goal of thermostating is to force the system to sample the canonical equilibrium distribution at a given temperature by continually perturbing it. A benefit of the dynamical models is that it is possible to conserve structure (e.g. Hamiltonian structure) in the augmented dynamics. A motivation for using this approach is that if the perturbation is small, the dynamics will still correspond to physical dynamics (in contact with a reservoir) on an intermediate timescale. Some examples of important deterministic thermostating methods are the Nosé method [61, 62], which preserves Hamiltonian structure at the expense of a continuous rescaling of time, the Nosé-Hoover method [62, 37] which recovers the linear time but loses canonical Hamiltonian structure, the Nosé-Poincaré method of Bond, Laird & Leimkuhler [6] which is canonically Hamiltonian, and a generalization of the Nosé-Hoover approach for Hamiltonian systems with Poission structure [8]. Deterministic thermostats have also been coupled with Langevin models in [44] for example.

In this chapter we propose the use of a thermostat to model the unresolved vorticity and its exchange in a simple point vortex model. We will model both an infinite reservoir as in classical thermodynamics, and a finite reservoir as has been used in the experiments of [9]. Statistics of the thermostated dynamics will be compared with the results of Bühler [9]. In §4.2 we make use of a generalized thermostat which can be used to force a Hamiltonian system to sample a general class of equilibrium distributions. The point vortex model and its statistical mechanics is reviewed in §4.3. In §4.4 we present the details of the thermostated numerical methods considered, including the models for finite and infinite reservoirs. Finally, in §4.5 the numerical schemes are verified by comparison with results from the literature.

4.2 Generalized thermostats

Consider an open subset $\mathcal{D} \subset \mathbb{R}^d$ and a deterministic differential equation

$$\dot{\boldsymbol{X}} = f(\boldsymbol{X}), \qquad \boldsymbol{X}(t) \in \mathcal{D}, \qquad f : \mathcal{D} \to \mathbb{R}^d. \tag{4.1}$$

A probability distribution $\rho(\boldsymbol{X},t) \in \mathcal{D}\times\mathbb{R}\to\mathbb{R}$, $\rho\geq 0$, on $\mathcal{D}$ is transported under the vector field f according to the continuity equation

$$\frac{\partial}{\partial t}\rho(\boldsymbol{X},t)+\nabla\cdot\rho(\boldsymbol{X},t)f(\boldsymbol{X})=0. \tag{4.2}$$

This continuity equation implies that $\int_{\mathcal{D}}\rho\,d\boldsymbol{X}=1$ for all $t>0$ if this holds at $t=0$. An equilibrium distribution is a stationary solution of (4.2). In this chapter we will be concerned primarily with systems of the form

$$\dot{\boldsymbol{X}}=J(\boldsymbol{X})\nabla H(\boldsymbol{X}),\qquad \boldsymbol{X}(t)\in\mathcal{D},\quad J^T=-J,\quad H:\mathcal{D}\to\mathbb{R}. \tag{4.3}$$

The function H is a first integral of (4.3), typically the energy. If J is independent of $\boldsymbol{X}$, then this defines a (generalized) Hamiltonian system. Otherwise, one must also show that $J(\boldsymbol{X})$ satisfies the Jacobi identity, in which case the system is Poisson. We make the weaker assumption that the vector field on the right side of (4.3) is divergence-free, i.e. $\nabla\cdot f(\boldsymbol{X})\equiv 0$, so that the transport equation (4.2) simplifies to the Liouville equation

$$\frac{d}{dt}\rho(\boldsymbol{X},t)=\frac{\partial}{\partial t}\rho(\boldsymbol{X},t)+f(\boldsymbol{X})\cdot\nabla\rho(\boldsymbol{X},t)=0. \tag{4.4}$$

An equilibrium distribution $\frac{\partial\rho}{\partial t}\equiv 0$ must satisfy

$$f(\boldsymbol{X})\cdot\nabla\rho(\boldsymbol{X})\equiv 0.$$

Note that any function $\rho(\boldsymbol{X})=\rho(H(\boldsymbol{X}))$ that depends on $\boldsymbol{X}$ through a first integral is an equilibrium distribution. If (4.1) has additional first integrals $I_2(\boldsymbol{X}),\ldots,I_p(\boldsymbol{X})$, then any distribution $\rho(H,I_2,\ldots,I_p)$ is also an equilibrium distribution. The ensemble average of a function $F(\boldsymbol{X})$ with respect to the equilibrium distribution $\rho(\boldsymbol{X})$ is

$$\langle F\rangle\equiv\int_{\mathcal{D}}F(\boldsymbol{X})\rho(\boldsymbol{X})\,d\boldsymbol{X}.$$

Given their ample supply, the degree to which a given equilibrium distribution is meaningful largely depends on whether the solution to the differential equation is ergodic in that distribution such that the long time average of any function $F(\boldsymbol{X}(t))$ of the solution

$$\overline{F}\equiv\lim_{T\to\infty}\frac{1}{T}\int_0^T F(\boldsymbol{X}(t))\,dt,$$

converges to the ensemble average in the distribution, i.e. satisfies

$$\overline{F}=\langle F\rangle$$

for almost any solution trajectory. If this is the case, the equilibrium distribution characterizes the long time behavior of solutions of the differential equation.

The microcanonical ensemble [39] applies to an isolated system at constant energy, and is the singular measure on the energy level set containing the initial condition

$$\rho_\nu \propto \delta(H(\boldsymbol{X}) - H_0), \tag{4.5}$$

where $H_0 = H(\boldsymbol{X}(0))$. This ensemble is appropriate for a numerical simulation with an energy conserving discretization.

A system in contact with a large reservoir does not conserve energy, but rather exchanges it with the reservoir. If it is in thermal equilibrium with a reservoir of statistical temperature β^{-1}, then the appropriate ensemble is the canonical ensemble [39] with Gibbs measure

$$\rho(\boldsymbol{X}) = N^{-1} \exp\left[-\beta H(\boldsymbol{X})\right], \tag{4.6}$$

where $N = \int_{\mathcal{D}} \exp\left[-\beta H(\boldsymbol{X})\right] d\boldsymbol{X}$. It is clear, however, that a single solution of the system (4.3) will not be ergodic in the Gibbs measure, since with probability one it will sample the constant energy surface containing the initial condition, whereas (4.6) assigns nonzero probability to all energy surfaces. Instead, to model a system in thermal equilibrium with a reservoir, one must devise a method whose dynamics samples phase space with probability given by the canonical distribution (4.6). The development of methods that do just this constitutes an active field of research. A number of techniques have been developed for sampling in a given distribution, including Monte Carlo schemes, which generate random configurations or trajectories according to the chosen distribution; Langevin thermostats, in which the original system of ordinary differential equations is augmented by stochastic forcing and generalized dissipation terms; and deterministic thermostats, in which the reservoir itself is modelled using a small number of additional degrees of freedom. The latter approaches have the advantage that they generate plausible solution behavior, and can be used to compute correlations.

In the next two sections we describe generalized Langevin dynamics and generalized stochastic Bulgac-Kusnezov thermostats for sampling in a wide class of equilibrium distributions for Hamiltonian systems.

4.2.1 Langevin thermostat

If one integrates (4.3) numerically using a symplectic integrator, the Hamiltonian will typically be well-conserved. As a result, the solution will not sample phase space with the measure (4.6) above, but instead will stay near the initial energy level set (at best sampling ρ_ν). For some applications it is desirable to construct a perturbed dynamical system that does sample ρ while retaining something of the dynamical behavior of (4.3). In this way one can construct a plausible (representative) behavior of the system if it were exchanging energy with the reservoir according to ρ.

One approach to sample a given equilibrium distribution augments (4.3) with carefully tuned noise and dissipation terms:

$$\dot{\boldsymbol{X}} = f(\boldsymbol{X}) + g(\boldsymbol{X}) + \Lambda(\boldsymbol{X})\dot{\boldsymbol{w}}(t), \tag{4.7}$$

where $g(\boldsymbol{X}) : \mathcal{D} \to \mathbb{R}^d$, $\Lambda(\boldsymbol{X}) \in \mathbb{R}^{d\times d}$ is a matrix valued function, and $\boldsymbol{w}(t)$ is a vector Wiener process, i.e. the $w_i(t)$, $i = 1, \ldots, d$, are scalar Gaussian random variables with mean zero and increments $w_i(t) - w_i(s) \sim \mathcal{N}(0, t-s)$. Phase space densities are transported by the flow of (4.7) according to the Fokker-Planck equation (see, for example [66])

$$\frac{\partial}{\partial t}\rho(\boldsymbol{X},t) = -\nabla\cdot\rho(\boldsymbol{X},t)(f(\boldsymbol{X})+g(\boldsymbol{X})) + \frac{1}{2}\nabla\cdot\nabla\cdot\rho(\boldsymbol{X},t)\Lambda(\boldsymbol{X})\Lambda^T(\boldsymbol{X}), \quad (4.8)$$

where $g(\boldsymbol{X})$ must be determined such that the desired equilibrium distribution is a stationary solution of (4.8). If ρ depends on $\boldsymbol{X}$ only through its Hamiltonian $\rho(\boldsymbol{X}) = \rho(H(\boldsymbol{X}))$, then the Hamiltonian dynamics drops out of the Fokker-Planck equation, since

$$\nabla\cdot\rho(H(\boldsymbol{X}))J\nabla H(\boldsymbol{X}) = \rho\nabla\cdot J\nabla H - \rho\nabla H\cdot J\nabla H = 0$$

by the divergence free nature of the Hamiltonian flow and conservation of energy. The Fokker-Planck equation reduces to

$$\frac{\partial}{\partial t}\rho(\boldsymbol{X},t) = -\nabla\cdot\rho(\boldsymbol{X},t)g(\boldsymbol{X}) + \frac{1}{2}\nabla\cdot\nabla\cdot\rho(\boldsymbol{X},t)\Lambda(\boldsymbol{X})\Lambda^T(\boldsymbol{X}).$$

The right hand side is zero if

$$-\rho(\boldsymbol{X})g(\boldsymbol{X}) + \frac{1}{2}\nabla\cdot\rho(\boldsymbol{X})\Lambda(\boldsymbol{X})\Lambda^T(\boldsymbol{X}) = \boldsymbol{0}.$$

For the canonical equilibrium distribution (4.6), we can solve this for $g(\boldsymbol{X})$ (suppressing the dependence on $\boldsymbol{X}$ in the notation):

$$\rho g(\boldsymbol{X}) = \frac{1}{2}\rho\nabla\cdot\Lambda\Lambda^T - \frac{1}{2}\beta\rho\Lambda\Lambda^T\nabla H = \rho\left(\frac{1}{2}\nabla\cdot\Lambda\Lambda^T - \frac{1}{2}\beta\Lambda\Lambda^T\nabla H\right).$$

For $\Lambda(\boldsymbol{X})$ constant, this becomes

$$g(\boldsymbol{X}) = -\frac{1}{2}\beta\Lambda\Lambda^T\nabla H(\boldsymbol{X}).$$

Hence, the Langevin dynamics is

$$\dot{\boldsymbol{X}} = J\nabla H(\boldsymbol{X}) - \frac{\beta}{2}\Lambda\Lambda^T\nabla H(\boldsymbol{X}) + \Lambda\dot{\boldsymbol{w}}. \quad (4.9)$$

If the flow map is in addition ergodic with respect to ρ, then the generalized Langevin dynamics (4.7) can be used to sample the canonical distribution at inverse temperature β.

Remark. For $\Lambda = \Lambda(\boldsymbol{X})$ locally defined the noise is multiplicative, one must specify whether the Itô or Stratanovich interpretation is used, and numerical methods must be carefully constructed to maintain accuracy.

4.2.2 A generalized Bulgac-Kusnezov method

The following approach generalizes the Bulgac-Kusnezov method [8] and offers additional flexibility. The method has been proposed for canonical sampling in the molecular dynamics setting in [43]; here we treat an arbitrary smooth ensemble and apply it to the fluid vortex model. We introduce a new variable $\zeta \in \mathbb{R}$ and functions $s(\boldsymbol{X}, \zeta) : \mathcal{D} \times \mathbb{R} \to \mathbb{R}^d$ and $r(\boldsymbol{X}, \zeta) : \mathcal{D} \times \mathbb{R} \to \mathbb{R}$ and form the coupled system

$$\dot{\boldsymbol{X}} = J\nabla H(\boldsymbol{X}) + s(\boldsymbol{X}, \zeta), \tag{4.10}$$

$$\dot{\zeta} = r(\boldsymbol{X}, \zeta). \tag{4.11}$$

We ask that the following extended measure be invariant under the Liouville equation:

$$\tilde{\rho}(\boldsymbol{X}, \zeta) \propto \exp\left[-\beta F(\boldsymbol{X}) - \alpha G(\zeta)\right] \tag{4.12}$$

for F and G appropriately defined functions. In the case of (4.6) we will take $F \equiv H$, but we consider this more general formulation for now. Note that after integration over ζ, this measure is of the form (4.6). The stationarity condition for the transport equation (4.2) is

$$\nabla \cdot \tilde{\rho}\,(f + s) + \partial_\zeta(\tilde{\rho} r) = 0, \quad \text{with } f = J\nabla H(\boldsymbol{X}).$$

Some calculations give

$$\begin{aligned} 0 &= (f+s)\cdot\nabla\tilde{\rho} + \tilde{\rho}\nabla\cdot(f+s) + r\frac{\partial}{\partial\zeta}\tilde{\rho} + \tilde{\rho}\frac{\partial}{\partial\zeta}r \\ &= -\beta\tilde{\rho}\nabla F\cdot(f+s) + \tilde{\rho}\nabla\cdot(f+s) - \alpha\tilde{\rho}r\frac{\partial}{\partial\zeta}G + \tilde{\rho}\frac{\partial}{\partial\zeta}r \\ &= \tilde{\rho}\left(-\beta\nabla F\cdot(f+s) + \nabla\cdot s - \alpha r\frac{\partial}{\partial\zeta}G + \frac{\partial}{\partial\zeta}r\right), \end{aligned}$$

where the divergence-freedom of the Hamiltonian vector field is used in the last inequality.

Next we make some simplifying assumptions. First we assume the thermostat variable ζ to be normally distributed, taking $G(\zeta) = \zeta^2/2$. We also assume that r depends only on $\boldsymbol{X}$, i.e. $r(\boldsymbol{X}, \zeta) = r(\boldsymbol{X})$. The stationarity condition consequently reduces to

$$0 = -\beta\nabla F\cdot(f+s) + \nabla\cdot s - \alpha r\zeta.$$

We wish to use this relation to define r. Note that

$$\zeta r(\boldsymbol{X}) = \frac{1}{\alpha}\left(\nabla\cdot s - \beta\nabla F\cdot(f+s)\right). \tag{4.13}$$

Since ζ may be zero, each term on the right should either vanish or have precisely a factor ζ as on the left. Candidate equilibrium distributions for (4.3) typically

have functional dependence via the Hamiltonian. If we assume that $F(H(\boldsymbol{X}))$, then the skew-symmetry of J implies

$$\nabla F \cdot f = F'(H(\boldsymbol{X}))\nabla H \cdot J\nabla H \equiv 0.$$

If additionally we assume $s(\boldsymbol{X}, \zeta)$ to be linear in ζ, i.e.

$$s(\boldsymbol{X}, \zeta) = s_1(\boldsymbol{X})\zeta, \qquad s_1(\boldsymbol{X}) \in \mathbb{R}^d,$$

then we find that

$$r(\boldsymbol{X}) = \frac{1}{\alpha}\left(\nabla \cdot s_1(\boldsymbol{X}) - \beta \nabla F \cdot s_1(\boldsymbol{X})\right) \tag{4.14}$$

is a solution of (4.13).

Specific choices of the functions $F(\boldsymbol{X})$ and $s_1(\boldsymbol{X})$ will be treated in Section 4.4.

In general, the thermostated dynamics so defined will not be ergodic in the invariant measure (4.12). To improve ergodicity, a Langevin term may be added to (4.11). See also [44].

$$\dot{\boldsymbol{X}} = J\nabla H(\boldsymbol{X}) + s_1(\boldsymbol{X})\zeta, \tag{4.15}$$

$$\dot{\zeta} = r(\boldsymbol{X}) - \frac{\alpha\lambda^2}{2}\zeta + \lambda\dot{w}. \tag{4.16}$$

Since the noise enters through ζ, it influences $\boldsymbol{X}(t)$ only after integration, so its effect on the dynamics is smoothed.

Remark. In the important special case $F(\boldsymbol{X}) = F(H(\boldsymbol{X}))$, if we choose s_1 such that $\nabla \cdot s_1 \equiv 0$, then the system (4.15)-(4.16) can be cast in the form of a generalized Langevin thermostat (4.9) as discussed in the previous section. Define the augmented system

$$\begin{gathered}\tilde{\boldsymbol{X}} = \begin{pmatrix} \boldsymbol{X} \\ \zeta \end{pmatrix}, \qquad \tilde{J}(\tilde{\boldsymbol{X}}) = \begin{bmatrix} \frac{1}{F_H} J & \frac{\beta}{\alpha} s_1(\boldsymbol{X}) \\ -\frac{\beta}{\alpha} s_1(\boldsymbol{X})^T & 0 \end{bmatrix}, \\ \tilde{H}(\boldsymbol{X}, \zeta) = F(H(\boldsymbol{X})) + \frac{\alpha}{2\beta}\zeta^2.\end{gathered} \tag{4.17}$$

Then (4.15)-(4.16) with (4.14) takes the form

$$\frac{d}{dt}\tilde{\boldsymbol{X}} = \tilde{J}\nabla\tilde{H} - \frac{\alpha}{2}\Lambda\Lambda^T\nabla\tilde{H}(\tilde{\boldsymbol{X}}) + \Lambda\dot{\boldsymbol{w}} \tag{4.18}$$

with $\Lambda = \begin{pmatrix} \mathbf{0} \\ \lambda \end{pmatrix}$.

4.3 Statistical mechanics of point vortices

The motion of M point vortices with circulation strengths $\Gamma_i \in \mathbb{R}$, $i = 1, \ldots, M$, and positions $\boldsymbol{x}_i(t) \in \mathbb{R}^2$ is given by the Hamiltonian system

$$\Gamma_i \dot{\boldsymbol{x}}_i = K\frac{\partial H}{\partial \boldsymbol{x}_i}, \qquad i = 1, \ldots, M, \tag{4.19}$$

where $K = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, and the Hamiltonian

$$H = -\frac{1}{4\pi}\sum_{i<j}\Gamma_i\Gamma_j \ln(|\boldsymbol{x}_i - \boldsymbol{x}_j|^2)$$

represents the kinetic energy.

If there are Γ_i with both positive and negative circulations, then the motion of point vortices is unbounded on the plane. A bounded flow can be ensured by imposing periodicity, which alters the Green's function in the Hamiltonian [86]. Alternatively, flow on a disc of radius R can me modeled by defining a set of image vortices

$$\Gamma_i' = -\Gamma_i, \quad \boldsymbol{x}_i' = \boldsymbol{x}_i\frac{R^2}{|\boldsymbol{x}_i|^2}, \quad i = 1, \ldots, M$$

which ensure that the velocity field observed by any point vortex is tangent to the wall. In the disc model, which we adopt in this chapter, the Hamiltonian has three terms due to: the original pair potential, the self-interaction, and the interaction terms of each vortex with the images of the others:

$$\begin{aligned} H = -\frac{1}{4\pi}\sum_{i<j}\Gamma_i\Gamma_j \ln(|\boldsymbol{x}_i - \boldsymbol{x}_j|^2) + \frac{1}{4\pi}\sum_{i}\Gamma_i^2 \ln(R^2 - |\boldsymbol{x}_i|^2) + \\ \frac{1}{4\pi}\sum_{i<j}\Gamma_i\Gamma_j \ln(R^4 - 2R^2\boldsymbol{x}_i\cdot\boldsymbol{x}_j + |\boldsymbol{x}_i|^2|\boldsymbol{x}_j|^2). \end{aligned} \tag{4.20}$$

To cast the system (4.19) in the form (4.3), we define $\boldsymbol{X} = (\boldsymbol{x}_1^T, \cdots, \boldsymbol{x}_M^T)^T$, $H = H(\boldsymbol{X})$, and

$$J = \begin{bmatrix} \Gamma_1^{-1}K & & \\ & \ddots & \\ & & \Gamma_M^{-1}K \end{bmatrix}.$$

Besides the kinetic energy, the point vortex flow on the disc conserves the total angular momentum, defined as

$$\hat{M} = \frac{1}{2\pi}\sum_i \Gamma_i|\boldsymbol{x}_i|^2. \tag{4.21}$$

In general there will be an exchange of momentum between the strong vortices and the reservoir. However on average we would expect the angular momentum of both strong and weak vortex sets to be approximately constant. In fact, it would be straightforward to model the exchange of angular momentum using the thermostat as well. This would require knowledge of the variance of the angular momentum of the reservoir. In this chapter we assume the momentum exchange with the reservoir is negligible, and we can show that $\hat{M}$ is a conserved quantity of the thermostated dynamics. Experiments with Langevin dynamics indicate

significant drift in angular momentum. To correct this, one could construct a projection of the noise term $\Lambda(\boldsymbol{X})$ onto the angular momentum manifold. However, this comes at the cost of multiplicative noise.

The phase space of the point vortex flow consists of the direct product of M copies of the domain. If the domain is bounded, so is the phase space. The energy H is unbounded on the phase space however: as $\boldsymbol{x}_i \to \boldsymbol{x}_j$, the logarithm tends to $-\infty$; if Γ_i and Γ_j are like-signed, $H \to +\infty$, if oppositely-signed, $H \to -\infty$. In particular, if a particle collides with the wall, $H \to -\infty$. As noted by Onsager [65], if we define $\Omega(E)$ to be the measure of the set of configurations in phase space for which $H \in [E, E+dE]$, then we must have $\lim_{E\to\pm\infty} \Omega(E) = 0$. In other words, since the phase space is bounded, the measure of available phase space must eventually decrease as an increasing function of energy. The situation is in contrast to other n-body problems encountered in chemistry and astronomy, where the positive definite kinetic energy terms can accommodate any amount of energy, and the measure of available phase space is a monotone increasing function of energy.

Consequently, the microcanonical entropy $S(E) = \ln \Omega(E)$ must attain a maximum for some E^*. The microcanonical temperature is defined to be $\mathcal{T}_\nu^{-1} = \frac{d}{dE} \ln S(E)$. As temperature varies from $-\infty$ to ∞ along the real line, the associated energy states vary from $E > E^*$ to $E < E^*$ through infinity. In other words temperature passes through zero via a collision. This has important consequences for thermostating, since it implies that that the vortices will collapse to the wall if the temperature changes sign continuously.

Recall that extreme values of the energy H are associated with close approaches between vortices or image-vortices. As noted by Bühler [9], for a homogeneous system with $\Gamma_i = \Gamma$, the energy largely governs the dynamics, since collisions have to occur roughly at constant energy. The situation is more interesting in a heterogeneous system with vortices of greatly differing strength. Onsager predicted that for such systems, extreme values of energy would increase the probability of clustering of like-signed or opposite-signed vortices, with a preference for the strongest ones, such that most of the energy would reside in a few degrees of freedom. As a result, the small vortices would roam aimlessly about, not developing into coherent structures, but exhibiting large entropy.

Bühler discusses Onsager's ideas in the context of the canonical ensemble applied to the strong vortices, which constitute a system in 'thermal' equilibrium with the reservoir of weak vortices. He verifies Onsager's predictions using numerical experiments with a system of 100 point vortices, four having strength $\pm 10\pi$ and the rest having strength $\pm 2\pi$. In each group, half the vortices had positive circulation and half negative. Experiments were carried out for extreme positive, neutral and extreme negative inverse statistical temperatures $\beta_\nu = \mathcal{T}_\nu^{-1}$ in the microcanonical sense. In each case the strong vortices had the same (nearly steady state) initial configuration, so the differences in energy were only due to the random placement of weak vortices. Simulations were run on a long time interval, and statistics were recorded for the distance between like

and opposite signed strong vortices, distance from the wall, and energy in the strong vortices. Bühler distinguishes between a theoretical infinite reservoir and the finite reservoir comprised of the 96 weak vortices. In the infinite reservoir case, the canonical probability measure only exists for a finite interval of inverse temperature β, whose boundary corresponds to collisions. This situation is due to the availability of an infinite amount of energy in the reservoir, and has implications for thermostating in the canonical ensemble. Specifically, if one thermostats in the canonical ensemble and increases β beyond its admissible range, the vortices will collapse onto the boundary. Bühler also points out that the contact with a finite reservoir will suppress this collapse, allowing thermostating at all temperatures. This is because there is a finite amount of energy in the finite reservoir, and this effectively bounds the closeness of approach of any two vortices from below. The probability of a close approach becomes very small. The probability of $H = E$ for a system in contact with a finite reservoir decays like $\exp(-\gamma E^2)$ with E for some $\gamma > 0$.

4.4 A thermostated integrator for point vortices

Our goal is to apply the generalized Bulgac-Kusnezov thermostat from Section 4.2.2 to the point vortex flow of Section 4.3. In this section we fill in the details of the method. First, in Section 4.4.1 we specify two equilibrium distributions corresponding to the cases where the reservoir of small scale vorticity is finite or infinite. In Section 4.4.2 we define a thermostat function s_1 such that the generalized Bulgac-Kusnezov thermostat is a Langevin thermostat. We describe the numerical method used to integrate the model adaptively in time in Section 4.4.3 and the means of computing the temperature in Section 4.4.4.

4.4.1 Infinite and finite reservoir ensembles

As discussed in [9] the behavior of a thermostated point vortex system can vary considerably depending on whether the reservoir is finite or infinite. In the case of an infinite reservoir, as the temperature of the reservoir is pushed toward zero, the subsystem may draw an arbitrarily large amount of energy from the reservoir, leading to collisions between individual vortices or with the wall. For a finite reservoir, there is a limited amount of energy available such that a collision may only occur if a collision with opposite energy occurs at the same time, and this is improbable. Specifically, in the case of a finite reservoir with normally distributed reservoir energy, the equilibrium distribution takes the form

$$\tilde{\rho}(\boldsymbol{X}) = \exp\left[-\beta H(\boldsymbol{X}) - \gamma H(\boldsymbol{X})^2\right].$$

For the generalized thermostat (4.10)–(4.11) we can model both finite and infinite reservoirs. For a finite reservoir we take

$$\begin{aligned} F(\boldsymbol{X}) &= H(\boldsymbol{X}) + \frac{\gamma}{\beta} H(\boldsymbol{X})^2, \\ r(\boldsymbol{X}) &= \frac{1}{\alpha} \left(\nabla \cdot s_1(\boldsymbol{X}) - (\beta + 2\gamma H(\boldsymbol{X})) \nabla H \cdot s_1(\boldsymbol{X}) \right), \end{aligned} \tag{4.22}$$

and for an infinite reservoir $\gamma \equiv 0$ in the expressions above.

4.4.2 Choice of s_1

We make the following choice for the function s_1 in (4.14):

$$s_1(\boldsymbol{X}) = - \begin{pmatrix} K \frac{\boldsymbol{x}_1}{|\boldsymbol{x}_1|} \\ \vdots \\ K \frac{\boldsymbol{x}_M}{|\boldsymbol{x}_M|} \end{pmatrix}. \tag{4.23}$$

The flow of the vector field s_1 preserves the distance of each point vortex from the center of the domain. Consequently the thermostated system (4.15)-(4.16) preserves the angular momentum (4.21).

Furthermore, this choice of s_1 is divergence-free:

$$\nabla \cdot s_1(\boldsymbol{X}) \equiv 0,$$

implying that the thermostated dynamics is a generalized Langevin system (4.18), and that the integral $\tilde{H}$ in (4.17) with $F(H(\boldsymbol{X}))$ from (4.22) is preserved in the limit $\lambda \to 0$ of zero noise and dissipation.

4.4.3 Implementation details

In our numerical implementation, time stepping was done using a splitting approach. We solved alternately the deterministic thermostat system and the stochastic equation for the thermostat variable. The deterministic system is solved with the implicit midpoint rule

$$\frac{\boldsymbol{X}^{n+1} - \boldsymbol{X}^n}{\tau} = J \nabla H(\hat{\boldsymbol{X}}) - s_1(\hat{\boldsymbol{X}}) \hat{\zeta}, \tag{4.24}$$

$$\frac{\zeta^{n+1} - \zeta^n}{\tau} = r(\hat{\boldsymbol{X}}), \tag{4.25}$$

where $\hat{\boldsymbol{X}} = (\boldsymbol{X}^{n+1} + \boldsymbol{X}^n)/2$ and $\hat{\zeta} = (\zeta^{n+1} + \zeta^n)/2$.

The remaining vector field is an Ornstein-Uhlenbeck equation

$$\dot{\zeta} = -\frac{\alpha \lambda^2}{2} \zeta + \lambda \dot{w}$$

with exact solution

$$\zeta^{n+1} = \exp[-\varepsilon\tau]\left(z^n + \lambda\sqrt{\frac{e^{2\varepsilon\tau}-1}{2\varepsilon}}\Delta w\right), \tag{4.26}$$

where $\varepsilon = \alpha\lambda^2/2$ and $\Delta w \sim \mathcal{N}(0,1)$. A full time step of size Δt is constructed by solving (4.26) with $\tau = \Delta t/2$ composed with (4.24)–(4.25) with $\tau = \Delta t$ composed with a second solution of (4.26), $\tau = \Delta t/2$.

During a close approach of two vortices, equivalently when the strong vortex energy is large in magnitude, accuracy and stability considerations motivate the use of an adaptive time-stepping strategy. Given a stepsize Δt_n in the nth time step, the subsequent time step is found by solving

$$\Delta t_n \Delta t_{n+1} = e(\boldsymbol{X}^n)^2 \Delta s^2. \tag{4.27}$$

Here, Δs is a uniform timestep under the time transformation $t = e\cdot s$, and e is a monitor function that measures the stiffness of the local solution. This adaptivity approach is explicit and time-reversible whenever the numerical integrator is symmetric. For our experiments we use

$$e(\boldsymbol{x}) = \min_{i\neq j} |\boldsymbol{x}_i - \boldsymbol{x}_j|,$$

where the minimization is over all vortices and image vortices.

4.4.4 Computation of temperatures

We check the inverse temperature β and reservoir variance γ numerically assuming ergodicity. For some function $a(\boldsymbol{X}) : \mathcal{D} \to \mathbb{R}^d$ and an equilibrium distribution $\rho(\boldsymbol{X}) = \exp\left[-\beta^* \tilde{H}(\boldsymbol{X})\right]$

$$\begin{aligned}\nabla\cdot\rho(\boldsymbol{X})a(\boldsymbol{X}) &= a(\boldsymbol{X})\cdot\nabla\rho(\boldsymbol{X}) + \rho(\boldsymbol{X})\nabla\cdot a(\boldsymbol{X}) \\ &= -\beta^*\rho(\boldsymbol{X})a(\boldsymbol{X})\cdot\nabla\tilde{H} + \rho(\boldsymbol{X})\nabla\cdot a(\boldsymbol{X}).\end{aligned}$$

Formally integrating over phase space

$$\int_{\mathcal{D}} \nabla\cdot\rho(\boldsymbol{X})a(\boldsymbol{X})d\boldsymbol{X} = -\beta^*\int_{\mathcal{D}}\rho(\boldsymbol{X})a(\boldsymbol{X})\cdot\nabla\tilde{H}\,d\boldsymbol{X} + \int_{\mathcal{D}}\rho(\boldsymbol{X})\nabla\cdot a(\boldsymbol{X})d\boldsymbol{X}. \tag{4.28}$$

The expression on the left is zero if either ρ or a is zero on the boundary $\partial\mathcal{D}$ of phase space. The boundary of $\mathcal{D}$ consists of configurations for which at least one point vortex is located on the boundary of the disc. Such a configuration has energy $H \to -\infty$. Likewise, there are points in phase space where two or more point vortices collide and the Hamiltonian tends to $\pm\infty$. The Gibbs distribution (4.6) can be normalized only for β on an open interval [9]:

$$\beta \in \left(\frac{-8\pi}{\Gamma^2 M}, \frac{+4\pi}{\Gamma^2}\right). \tag{4.29}$$

To carry out the integration (4.28), we choose a for the form:

$$a = b/\rho, \qquad \rho(\boldsymbol{X}) = \exp\left[-\beta H(\boldsymbol{X})\right],$$

where β is the desired inverse temperature and $b(\boldsymbol{X})$ is some function with $b = 0$ at the boundary of the phase space. In this case, the expression for β^* simplifies to

$$0 = -\beta^* \langle a \cdot \nabla \tilde{H} \rangle + \langle \nabla \cdot a \rangle.$$

If the flow is ergodic, then the ensemble averages can be replaced with time averages

$$\beta^* = \overline{\nabla \cdot a} / \overline{a \cdot \nabla \tilde{H}},$$

and the disagreement of β^* and β serves as a simple check for nonergodicity. For the infinite reservoir, $\tilde{H} = H$, and for the finite reservoir, $\tilde{H} = H + \gamma^*/\beta^* H^2$, yielding

$$0 = -\beta^* \langle a \cdot \nabla H \rangle - 2\gamma^* \langle a \cdot H \nabla H \rangle + \langle \nabla \cdot a \rangle.$$

Choosing two independent functions $a_1 = b_1/\rho$ and $a_2 = b_2/\rho$, where b_1 and b_2 are identically zero on $\partial\mathcal{D}$, these equations yield a linear system for β^* and γ^*. For our experiments we chose

$$b_1 = \nabla H \prod_i (R^2 - |\boldsymbol{x}_i|^2)|\boldsymbol{x}_i|^2, \quad b_2 = \nabla H \prod_i (R^2 - |\boldsymbol{x}_i|^2)^2 |\boldsymbol{x}_i|^4,$$
$$\rho = \exp(-\beta H - \gamma H^2),$$

where β is one of the three inverse temperatures (4.30) and γ is either 0 for the infinite reservoir or the corresponding reservoir variance (4.31) for the finite reservoir.

Figure 4.1 illustrates convergence of β^* to the values of β (4.30) for both the infinite and finite reservoir, as well as convergence of γ^* to γ (4.31) for the finite reservoir.

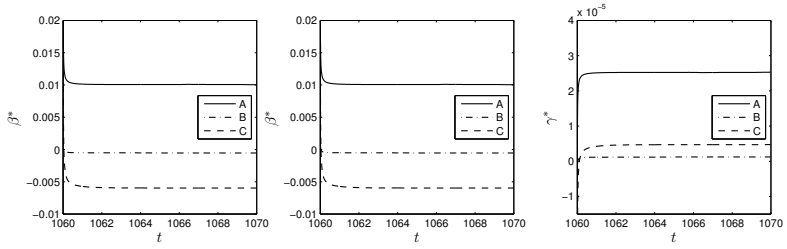

FIGURE 4.1: Convergence of inverse temperature β and reservoir variance γ for high (A), medium (B) and low (C) temperature states. Left: infinite reservoir β^*. Middle: finite reservoir β^*. Right: finite reservoir γ^*.

4.5 Numerical experiments

For all of the numerical experiments using four strong vortices, the initial configuration consists of point vortices with circulations and positions given by [9]:

$$\Gamma_1 = \Gamma_3 = 10\pi, \qquad \Gamma_2 = \Gamma_4 = -10\pi,$$
$$\boldsymbol{x}_1 = (3,0), \quad \boldsymbol{x}_3 = (-3,0), \quad \boldsymbol{x}_2 = (0,3), \quad \boldsymbol{x}_4 = (0,-3).$$

For both the finite and infinite reservoir thermostat we choose negative, neutral and positive inverse temperatures

$$\beta = \{-0.006, -0.00055, 0.01\}. \tag{4.30}$$

Our choice of β is such that it is close to the theoretical upper limit in (4.29) for high temperature, and it is close to the theoretical lower limit in (4.29) for low temperature.

The size of the reservoir is defined by γ. In the case of an infinite reservoir $\gamma \equiv 0$, for a finite reservoir

$$\gamma = \beta/(-2E_0) \quad \text{with} \quad E_0 = \{628, 221, -197\}. \tag{4.31}$$

In all experiments, we take $\alpha = 0.5$ and $\lambda = \sqrt{0.4}$.

We integrated the thermostated dynamics over the interval $t \in [0, T]$ with $T = 12000$ using the time transformation (4.27) and fixed transformed time steps $\Delta s = \Delta t_0 / e(\boldsymbol{X}(0))$ with $\Delta t_0 = 0.001$. The sampling was performed over the time interval $[T_0, T]$ with $T_0 = 1500$, to allow decorrelation of the initial conditions. The resulting time series was sampled uniformly in time in cycles of $\delta t = 0.01$, to produce the histograms shown in Figures 4.6–4.9.

4.5.1 Ergodicity tests

The extended measure (4.12) is Gaussian in the thermostat variable ζ. If the time dynamics is ergodic with respect to (4.12), we expect the time series $\zeta(t)$ to be normally distributed, i.e. $\zeta \in \mathcal{N}(0, \alpha^{-1})$. A histogram of the values of ζ is shown in Figure 4.2 for the neutral case $\beta = -0.00055$. The normal distribution $\rho(\zeta) = \sqrt{\frac{\alpha}{2\pi}} \exp(-\frac{\alpha}{2}\zeta^2)$ is also plotted in the figure. The agreement is good, indicating ergodicity with respect to ζ.

As a second indication of ergodicity, we plot the motion of a single vortex $\boldsymbol{x}_1(t)$ in Figure 4.3. The motion appears well-mixed. The density of points along the trajectory is greater where either the local velocity $\dot{\boldsymbol{x}}_1$ or the local time step Δt_n is small.

4.5.2 Momentum conservation

The function $s_1(\boldsymbol{X})$ in (4.23) is chosen to preserve angular momentum (4.21) of the strong vortex set under the thermostated dynamics. Figure 4.4 shows the angular momentum $\hat{M}$ as a function of time for the three temperatures. We observe that $\hat{M}$ is preserved to the relative precision of the fixed point iteration used to solve (4.24)–(4.25).

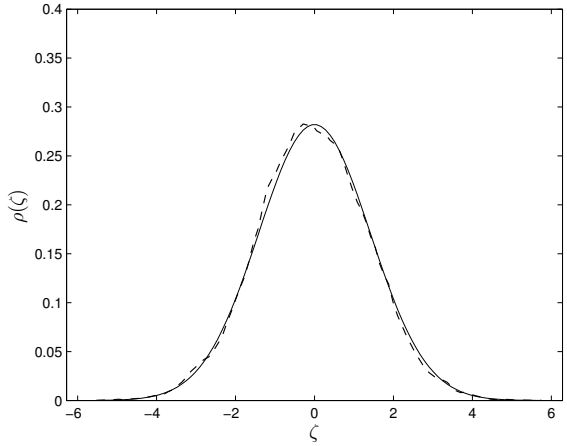

FIGURE 4.2: Distribution of thermostat (dash line), Gaussian fit (solid line).

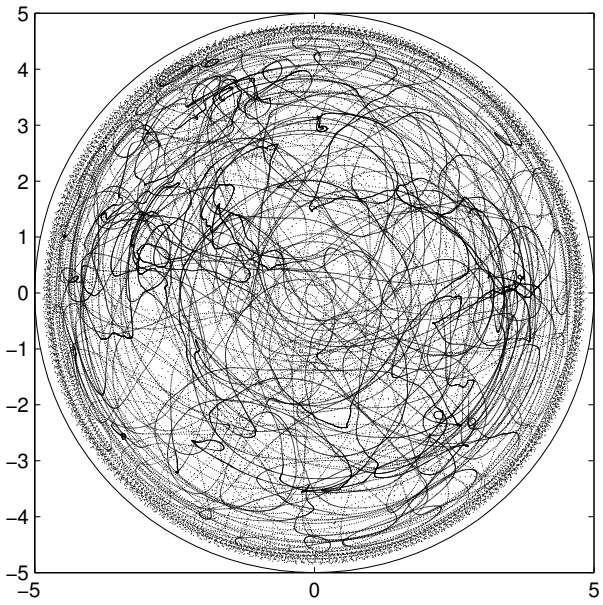

FIGURE 4.3: Motion of a single vortex $x_1(t)$ on the interval $t \in [0, 1000]$ for $\beta = -0.00055$.

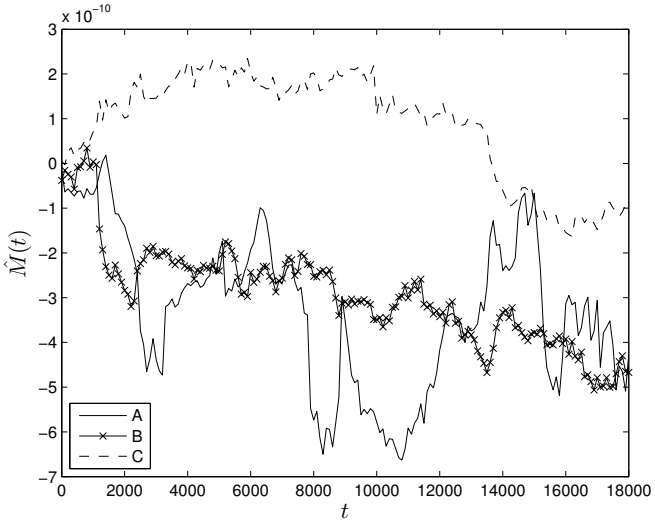

FIGURE 4.4: Momentum for high (A), medium (B) and low (C) temperature states for finite reservoir size. Infinite reservoir gives a similar behavior.

4.5.3 Temperature effects

In this section we attempt to reproduce the experiments of Bühler using thermostated large point vortices. We conduct experiments using both the infinite reservoir canonical distribution ((4.22) with $\gamma \equiv 0$) and the finite reservoir distribution ((4.22) with $\gamma \neq 0$).

The time evolution of the kinetic energy of strong vortices is displayed in Figure 4.5 for both the infinite and finite reservoir models, showing that the thermostat drives the energy evolution towards the desired temperature. Figure 4.6 shows the probability distributions of the kinetic energy of the vortices. For the finite reservoir thermostat, the means and variances are similar to those of [9].

Figure 4.7 displays the histogram of distances $|\boldsymbol{x}_i - \boldsymbol{x}_j|$ between like-signed vortices. Bias in favor of small separations is evident at negative temperatures, consistent with Onsager's predictions. The distributions are very similar to those obtained by Bühler [9]. For the infinite reservoir model, there is a large peak in the distribution at $|\boldsymbol{x}_i - \boldsymbol{x}_j| \approx 1$ which is inconsistent with Bühler's simulations. This occurs because too much energy is drawn from the reservoir. The comparison is recovered in the finite reservoir model.

Figure 4.8 shows the histograms of the distance between opposite-signed vortices. In this case, there is a somewhat milder bias towards close approaches at negative temperatures, in keeping with Onsager's ideas. The bias is less pronounced because the close approaches between a point vortex and its opposite signed image across the domain boundary are not included in this statistic.

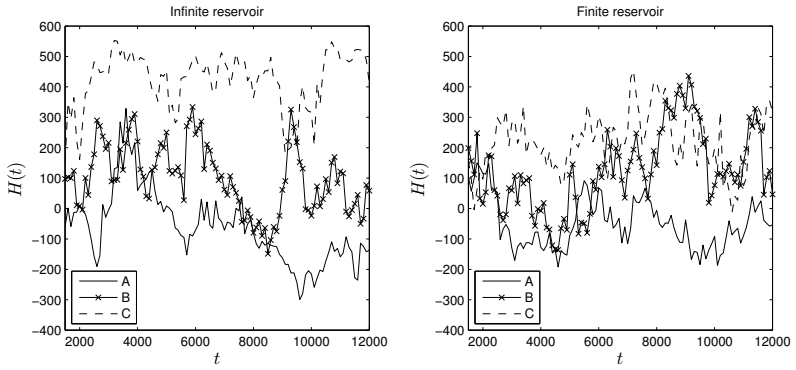

FIGURE 4.5: Time evolution of energy $H(t)$ for infinite (left) and finite (right) reservoirs. Inverse temperatures: $\beta = 0.01$ (solid line), $\beta = -0.00055$ (line with x-mark), $\beta = -0.006$ (dash line).

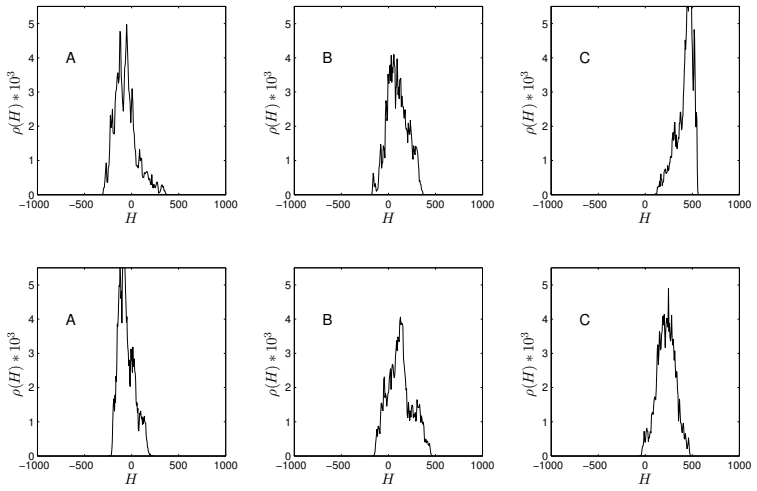

FIGURE 4.6: Distribution of energy for high (A), medium (B) and low (C) temperature states. Top: Infinite reservoir size. Bottom: Finite reservoir size.

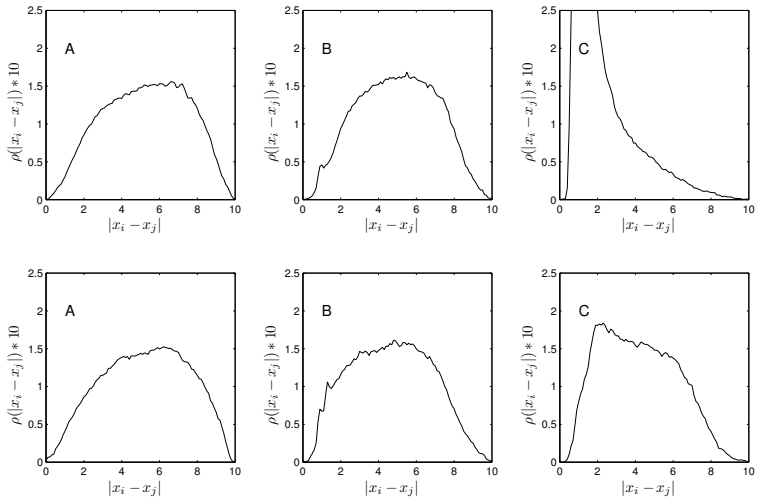

FIGURE 4.7: Interparticle spacing among same-signed vortices for high (A), medium (B) and low (C) temperature states. Top: Infinite reservoir size. Bottom: Finite reservoir size.

Again the histograms are in excellent agreement with the simulation data of [9], for the finite reservoir simulation. For an infinite reservoir, the positive temperature histogram is more peaked.

Figure 4.9 shows histograms of the vortex distance from the origin. For positive temperature, the vortices accumulate near the wall. The finite reservoir figures are in excellent agreement with those of [9]. For the infinite reservoir, the peak at $|\boldsymbol{x}_i| \approx 4.9$ is closer to the wall than for the finite reservoir, indicating that more energy is drawn from the reservoir in this case. At negative temperature, the vortices avoid the wall with high probability.

To observe the effects of temperature on a larger collection of vortices, we also simulated a set of with $M = 12$, under the same conditions as above at the extremal temperatures $\beta = -0.006$ and $\beta = 0.01$. The initial positions in both cases were defined as shown in Figure 4.10 in the left panel. The middle and right panels of Figure 4.10 show characteristic snapshots for each case. The linked animations illustrate the dynamics for positive and negative temperature regimes on a short interval $t \in [1500, 1500.1]$. At positive temperatures, vortices cluster in oppositely signed pairs, or translate parallel to the boundary of the domain. Because oppositely signed pairs translate normal to the dipole axis until they collide with another vortex or the boundary, these pairs are short lived. In contrast, for negative temperatures the vortices separate into two relatively stable regions of positive and negative circulation. Figure 4.11 shows a snapshot of the stream function from the positive and negative temperature

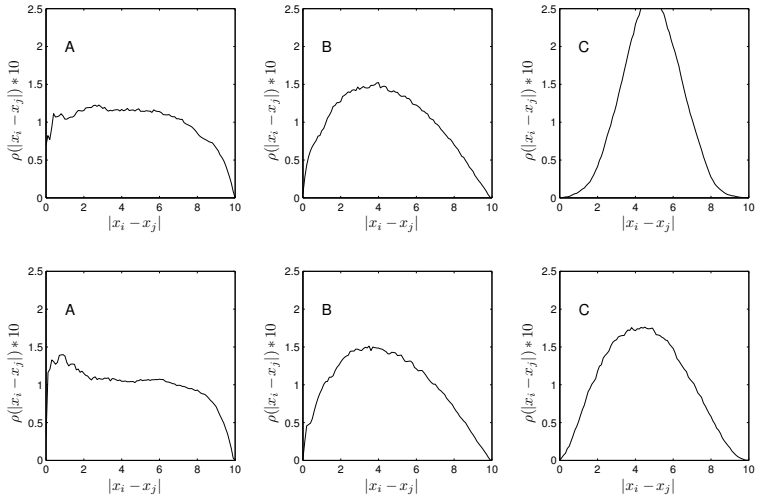

FIGURE 4.8: Interparticle spacing among opposite-signed vortices for high (A), medium (B) and low (C) temperature states. Top: Infinite reservoir size. Bottom: Finite reservoir size.

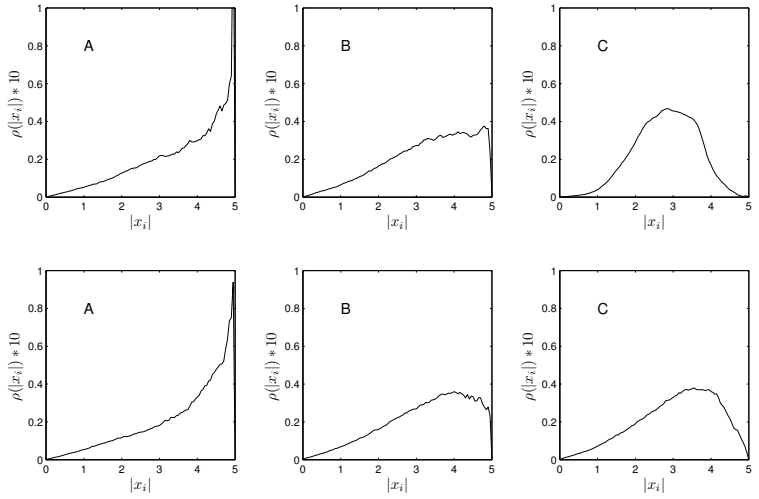

FIGURE 4.9: Distribution of distance from origin for high (A), medium (B) and low (C) temperature states. Top: Infinite reservoir size. Bottom: Finite reservoir size.

simulations. For negative temperatures the vorticity is more concentrated in two counter-rotating patches.

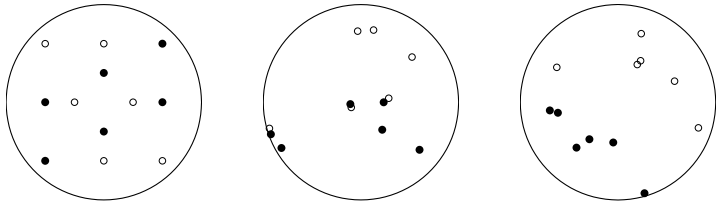

FIGURE 4.10: Snapshots of the case $M = 12$: the initial vortex placement (left), $\beta = 0.01$ (middle) and $\beta = -0.006$ (right), black or white color indicates positive or negative circulation. For positive temperature, clustering occurs pairwise; for negative temperature, large counter-rotating regions occur.

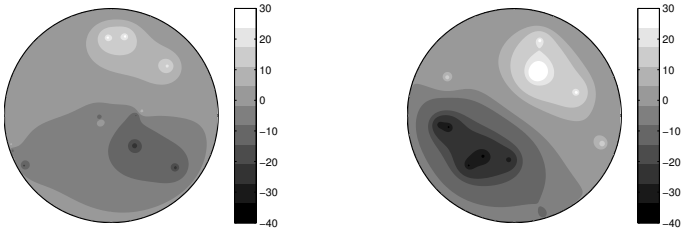

FIGURE 4.11: Snapshots of the stream function for case $M = 12$, $\beta = 0.01$ (left) and $\beta = -0.006$ (right). For negative temperature, clustering of like-signed vortices yields two strong counter-rotating vortices.

4.6 Conclusions

In this chapter we provide proof of concept that the energy exchange between large scale point vortices with a reservoir of small scale point vortices can be well modeled with a simple thermostat device that adds only a single degree of freedom to the phase space of the large scale flow. Specifically, we are able to recover the canonical statistics of the strong vortices, as obtained from direct numerical simulations in [9]. By constructing a thermostat for general energy-dependent equilibrium distributions, we model a canonical ensemble with a finite reservoir.

Bibliography

[1] Abramov, R. and Majda, A.J., *Statistically relevant conserved quantities for truncated quasi-geostrophic flow.* Proc. Natl. Acad. Sci. U.S.A. 100 (7) (2003), 3841–3846.

[2] Arakawa, A., *Computational design for long-term numerical integration of the equations of fluid motion: two-dimensional incompressible flow. Part I.* J. Comput. Phys. 1 (1966), 119–143.

[3] Arnold, V.I., *Mathematical methods of classical mechanics.* Springer-Verlag, New York, second ed., 1989.

[4] Benettin, G. and Giorgilli, A., *On the Hamiltonian interpolation of near to the identity symplectic mappings with application to symplectic integration algorithms.* J. Stat. Phys. 74 (1994), 1117–1143.

[5] Bokhove, O. and Oliver, M., *Parcel Eulerian-Lagrangian fluid dynamics of rotating geophysical flows.* Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 462 (2073) (2006), 2575–2592.

[6] Bond, S.D., Leimkuhler, B.J. and Laird, B.B., *The Nosé-Poincaré method for constant temperature molecular dynamics.* J. Comput. Phys. 151 (1) (1999), 114–134.

[7] Bridges, Th.J. and Reich, S., *Multi-symplectic integrators: numerical schemes for Hamiltonian PDEs that conserve symplecticity.* Phys. Lett. A 284 (2001), 184–193.

[8] Bulgac, A. and Kusnezov, D., *Canonical ensemble averages from pseudomicrocanonical dynamics.* Phys. Rev. A 42 (1990), 5045–5048.

[9] Bühler, O., *Statistical mechanics of strong and weak point vortices in a cylinder.* Physics of Fluids 14 (2002), 2139–2149.

[10] Bühler, O., *A Brief Introduction to Classical, Statistical, and Quantum Mechanics.* Courant Lecture Notes, vol. 13, AMS Bookstore, Rhode Island, 2006.

[11] Carnevale, G.F. and Frederiksen, J.S., *Nonlinear Stability and Statistical Mechanics of Flow Over Topography.* J. Fluid Mech. 175 (1987), 157–181.

[12] Chartier, P. and Faou, E., *Geometric integrators for piecewise smooth Hamiltonian systems.* Math. Model. Numer. Anal. 42 (2) (2008), 223–241.

[13] Chavanis, P.H., *Statistical mechanics of 2D turbulence with a prior vorticity distribution.* Phys. D Nonlinear Phenomena 237 (14-17) (2008), 1998–2002.

[14] Cohen, D., Hairer, E. and Lubich, C. *Conservation of energy, momentum and actions in numerical discretizations of nonlinear wave equations.* Numer. Math. 110 (2008), 113–143.

[15] Cottet, G.H. and Koumoutsakos, P.D., *Vortex Methods: Theory and Practice.* Cambridge University Press, Cambridge, 2000.

[16] Cotter, C., Frank, J. and Reich, S., *Hamiltonian particle-mesh method for two-layer shallow-water equations subject to the rigid-lid approximation.* SIAM J. Appl. Dyn. Syst. 3 (1) (2004), 69–83 (electronic).

[17] Cotter, C. and Reich, S., *Geometric Integration of a Wave-Vortex Model.* Appl. Numer. Math. 48 (2004), 293–305.

[18] Dubinkina, S. and Frank, J., *Statistical mechanics of Arakawa's discretizations.* J. Comput. Phys. 227 (2007), 1286–1305.

[19] Dubinkina, S. and Frank, J., *Statistical relevance of vorticity conservation with the Hamiltonian Particle-Mesh method.* J. Comput. Phys. (2010), published online: doi:10.1016/j.jcp.2009.12.012.

[20] Dubinkina, S., Frank, J. and Leimkuhler, B., *A thermostat closure for point vortices.* submitted (2009).

[21] Durran, D.R., *Numerical methods for wave equations in geophysical fluid dynamics.* Texts in Applied Mathematics, vol. 32, Springer-Verlag, New York, 1999.

[22] Ellis, R.S., Haven, K. and Turkington, B., *Nonequivalent statistical equilibrium ensembles and refined stability theorems for most probable flows.* Nonlinearity 15 (2) (2002), 239–255.

[23] Feynman, R.P., *Statistical mechanics.* Perseus Books, Advanced Book Program, Reading, MA, 1998.

[24] Frank, J., *Geometric space-time integration of ferromagnetic materials.* Appl. Numer. Math. 48 (3-4) (2004), 307–322.

[25] Frank, J., Gottwald, G. and Reich, S., *A Hamiltonian particle-mesh method for the rotating shallow-water equations.* In *Meshfree methods for partial differential equations*, Lect. Notes Comput. Sci. Eng.. 26, Springer, Berlin, 2003, 131–142.

[26] Frank, J., Huang, W. and Leimkuhler, B., *Geometric integrators for classical spin systems.* J. Comput. Phys. 133 (1) (1997), 160–172.

[27] Frank, J., Moore, B.E. and Reich, S. *Linear PDEs and numerical methods that preserve a multisymplectic conservation law.* SIAM J. Sci. Comput. 28 (1) (2006), 260–277 (electronic).

[28] Frank, J. and Reich, S., *Conservation properties of smoothed particle hydrodynamics applied to the shallow water equation.* BIT 43 (1) (2003), 41–55.

[29] Frank, J. and Reich, S., *The Hamiltonian Particle-Mesh Method for the Spherical Shallow Water Equations.* Atmospheric Science Letters 5 (2004), 89–95.

[30] Frenkel, D. and Smit, B., *Understanding Molecular Simulation: from algorithms to applications.* Academic Press, San Diego, 2002.

[31] Frutos, J., Ortega, T. and Sanz-Serna, J.M., *A Hamiltonian, explicit algorithm with spectral accuracy for the 'good' Boussinesq equation.* Comput. Methods Appl. Mech. Egrg. 80 (1990), 417–423.

[32] Frutos, J. and Sanz-Serna, J.M., *An easily implementable fourth-order method for the time integration of wave problems.* J. Comput. Phys. 103 (1992), 160–168.

[33] Ge, Z. and Marsden J.E., *Lie-Poisson Hamilton-Jacobi theory and Lie-Poisson integrators.* Phys. Lett. A 133 (1988), 134–139.

[34] Hairer, E., *Backward analysis of numerical integrators and symplectic methods.* Annals of Numer. Math. 1 (1994), 107–132.

[35] Hairer, E., Lubich, C. and Wanner, G., *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations.* Springer Series in Computational Mathematics, second ed., vol. 31, Springer-Verlag, Berlin, 2006.

[36] Hairer, E., Nørsett, S.P. and Wanner, G., *Solving Ordinary Differential Equations I: Nonstiff Problems.* Springer Series in Computational Mathematics, vol. 8, second ed., Springer-Verlag, Berlin, 1993.

[37] Hoover, W.G., *Canonical dynamics: Equilibrium phase-space distributions.* Phys. Rev. A 31 (1985), 1695–1697.

[38] Hundsdorfer, W. and Verwer, J., *Numerical solution of time-dependent advection-diffusion-reaction equations.* Springer Series in Computational Mathematics, vol. 33, Springer-Verlag, Berlin, 2003.

[39] Khinchin, A.I., *Mathematical Foundations of Statistical Mechanics.* New York: Dover, 1960.

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

[41] Kullback, S. and Leibler, R.A., *On Information and Sufficiency.* Ann. Math. Statist. 22 (1) (1951), 79–86.

[42] Lanczos, C., *The Variational Principles of Mechanics.* Toronto: University of Toronto Press, 1970.

[43] Leimkuhler, B., *Generalized Bulgac-Kusnezov Methods for Sampling of the Gibbs-Boltzmann Measure.* In preparation, 2009.

[44] Leimkuhler, B., Noorizadeh, E. and Theil, F., *A Gentle Stochastic Thermostat for Molecular Dynamics.* J. Stat. Phys. 135 (2) (2009), 261–277.

[45] Leimkuhler, B. and Reich, S., *Simulating Hamiltonian Dynamics*, Cambridge Monographs on Applied and Computational Mathematics, vol. 14, Cambridge University Press, Cambridge, 2004.

[46] Lorenz, E.N., *Energy and numerical weather prediction.* Tellus 12 (1960), 364-373.

[47] Lynden-Bell, D., *Statistical mechanics of violent relaxation in stellar systems.* Mon. Not. R. astr. Soc. 136 (1967), 101–121.

[48] Majda, A.J. and Wang, X., *Non-linear dynamics and statistical theories for basic geophysical flows.* Cambridge University Press, Cambridge, 2006.

[49] Marsden, J.E. and Ratiu, T., *Mechanics and Symmetry.* Springer-Verlag, New York, second ed., 1998.

[50] McLachlan, R.I., *Explicit Lie-Poisson integration and the Euler equations.* Phys. Rev. Lett. 71 (19) (1993), 3043–3046.

[51] McLachlan, R.I., *Symplectic integration of Hamiltonian wave equations.* Numer. Math. 66 (4) (1994), 465–492.

[52] McWilliams, J.C., *Fundamentals of geophysical fluid dynamics.* Cambridge University Press, New York, 2006.

[53] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E., *Equation of State Calculations by Fast Computing Machines.* J. Chem. Phys. 21 (1953), 1087–1092.

[54] Miller, J., *Statistical mechanics of Euler equations in two dimensions.* Phys. Rev. Lett. 65 (17) (1991), 2137–2140.

[55] Miller, J., Weichman, P.B. and Cross, M.C., *Statistical mechanics, Euler's equation, and Jupiter's Red Spot.* Phys. Rev. A 45 (4) (1992), 2328–2359.

[56] Molchanov, V., *Particle-Mesh and Meshless Methods for a Class of Barotropic Fluids.* Thesis, Jacobs University Bremen, Bremen, 2008.

[57] Morrison, P.J., *Hamiltonian description of the ideal fluid.* Rev. Modern Phys. 70 (2) (1998), 467–521.

[58] Morrison, P.J. and Greene, J.M., *Noncanonical Hamiltonian Density Formulation of Hydrodynamics and Ideal Magnetohydrodynamics.* Phys. Rev. Lett. 45 (1980), 790–794.

[59] Nambu, Y., *Generalized Hamiltonian dynamics.* Phys. Rev. D 7 (1973), 2405–2412.

[60] Névir, P. and Blender, R., *A Nambu representation of incompressible hydrodynamics using helicity and enstrophy.* J. Phys. A 26 (22) (1993), L1189–L1193.

[61] Nosé, S., *A molecular dynamics method for simulations in the canonical ensemble.* Mol. Phys. 52 (1984), 255–268.

[62] Nosé, S., *A unified formulation of the constant temperature molecular dynamics methods.* J. Chem. Phys. 81 (1984), 511–519.

[63] Oliver, M. and Bühler, O., *Transparent boundary conditions as dissipative subgrid closures for the spectral representation of scalar advection by shear flows.* J. Math. Phys. 48 (6) (2007), 065502-065502-26.

[64] Olver, P.J., *Applications of Lie Groups to Differential Equations.* Springer-Verlag, New York, 1986.

[65] Onsager, L., *Statistical hydrodynamics.* Nuovo Cimento, Suppl. 6 (2) (1949), 279–287.

[66] Pavliotis, G.A. and Stuart, A.M., *Multiscale Methods: Averaging and Homogenization.* Springer, New York, 2008.

[67] Pedlosky, J., *Geophysical Fluid Dynamics.* Springer, New York, second ed., 1987.

[68] Petersen, K.E., *Ergodic theory.* Cambridge University Press, Cambridge, 1989.

[69] Reich, S., *Numerical integration of the generalized Euler equation.* Technical Report TR 93-20, University of the British Columbia, 1993.

[70] Reich, S., *Backward error analysis for numerical integrators.* SIAM J. Numer. Anal. 36 (1999), 475–491.

[71] Robert, R., *A maximum-entropy principle for two-dimensional perfect fluid dynamics.* J. Statist. Phys. 65 (3-4) (1991), 531–553.

[72] Robert, R. and Sommeria, J., *Statistical equilibrium states for two-dimensional flows.* J. Fluid Mech. 229 (1991), 291–310.

[73] Salmon, R., *Hamiltonian fluid mechanics.* Ann. Rev. Fluid Mech. 20 (1988), 225–256.

[74] Salmon, R., *Lectures on geophysical fluid dynamics.* Oxford University Press, New York, 1998.

[75] Salmon, R., *A general method for conserving quantities related to potential vorticity in numerical models.* Nonlinearity 18 (5) (2005), R1–R16.

[76] Salmon, R., Holloway G. and Hendershott M.C., *The Equilibrium Statistical Mechanics of Simple Quasi-Geostrophic Models.* J. Fluid Mech. 75 (1976), 691–703.

[77] Sanz-Serna, J.M., *Runge-Kutta schemes for Hamiltonian systems.* BIT 28, 877–883.

[78] Sanz-Serna, J.M., *Symplectic integrators for Hamiltonian problems: an overview.* Acta numerica (1992), 243–286.

[79] Sanz-Serna, J.M. and Calvo, M.P., *Numerical Hamiltonian problems.* Applied Mathematics and Mathematical Computation, vol. 7, Chapman & Hall, London, 1994.

[80] Schlick, T., *Molecular Modeling and Simulation: An Interdisciplinary Guide.* Springer, New York, vol. 21, 2002.

[81] Shannon, C.E., *A Mathematical Theory of Communication.* Bell Syst. Tech. J. 27 (1948), 379–423 and 623–656.

[82] Shepherd, T.G., *Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics.* Adv. Geophys. 32 (1990), 287-338.

[83] Shin S. and Reich S., *Hamiltonian particle-mesh simulations for a nonhydrostatic vertical slice model.* Atmos. Sci. Lett. (2009), published online.

[84] Stuart, A.M. and Humphries, A.R., *Dynamical Systems and Numerical Analysis.* Cambridge Monographs on Applied and Computational Mathematics, vol. 2, Cambridge University Press, Cambridge, 1996.

[85] Walters, P., *An Introduction to Ergodic Theory.* Springer-Verlag, New York, 1982.

[86] Weiss, J.B. and McWilliams, J.C., *Nonergodicity of point vortices.* Phys. Fluids A 3 (5 part 1) (1991), 835–844.

[87] Wilkinson, J.H., *Error analysis of floating-point computation.* Numer. Math. 2 (1960), 319–340.

[88] Zeitlin, V., *Finite-mode analogues of 2D ideal hydrodynamics: Coadjoint orbits and local canonical structure* Physica D 49 (3) (1991), 353–362.

Summary

The research in this thesis is devoted to the statistical mechanics and numerical modelling of geophysical fluid dynamics.

The results of statistical analysis of simulation data obtained from long time integrations of geophysical fluid models greatly depend on the conservation properties of the numerical discretization used. In Chapter 2, this is illustrated for quasigeostrophic flow with topographic forcing, for which a well established statistical mechanics exists. We constructed statistical mechanical theories for the discrete dynamical systems arising from three discretizations due to Arakawa (*J. Comput. Phys.*, 1966) which conserve energy, enstrophy or both. Numerical experiments with conservative and projected time integrators have shown that the statistical theories accurately explain the differences observed in statistics derived from the discretizations.

In Chapter 3, we conducted long-time simulations with a Hamiltonian particle-mesh method for ideal fluid flow, to determine the statistical mean vorticity field of the discretization. We proposed Lagrangian and Eulerian statistical models for the discrete dynamics, and compared them against numerical experiments. The observed results are in excellent agreement with theoretical models, as well as with the continuum statistical mechanical theory for ideal flow developed by Ellis, Haven & Turkington (*Nonlinearity*, 2002). In particular the results verified that the apparently trivial conservation of potential vorticity along particle paths using the HPM method significantly influences the mean state. As a side note, the numerical experiments showed that a nonzero fourth moment of potential vorticity can influence the statistical mean.

Inviscid fluid models are characterized by conservation of energy, sensitive dependence on initial conditions, and the cascade of vorticity to ever finer scales. For numerical simulations of such flows, the vorticity cascade presents the challenge that any direct discretization of the equation of motion must eventually become underresolved. To address this problem effectively requires modelling the subgrid scale dynamics and their influence on the coarse scale. In Chapter 4, using the point vortex flow on a disc as a prototype, we presented a closure for incompressible ideal fluid flow in the form of a generalized thermostating device, a technique used in molecular simulations to model a system of particles interacting at a constant temperature. We showed that the thermostat can model either an infinite or finite reservoir with stochastically forced thermostat variables. Numerical experiments are in excellent agreement with the two-scale simulations of Bühler (*Phys. Fluids*, 2002).

Samenvatting

Het verrichte onderzoek, weergegeven in dit proefschrift, gaat over de statistische mechanica en het numeriek modelleren van geofysische vloeistofdynamica.

Uit de statistische analyse van de simulatiegegevens, verkregen door middel van lange tijdintegraties, is gebleken dat de resultaten sterk afhangen van de gebruikte numerieke discretisatie. In Hoofdstuk 2 hebben we dit geillustreerd aan de hand van een quasi-geostrofische stroming met topografische forcing, waarvoor er een goed ontwikkelde theorie in de statistische mechanica bestaat. We hebben statistisch-mechanische theorieën voor de discrete dynamische systemen opgebouwd ten gevolge van de drie discretisaties van Arakawa (*J. Comput. Phys.*, 1966), die de energie, enstrofie, of beide behouden. De numerieke experimenten met de behoudende integrators, evenals de geprojecteerde tijdintegrators, hebben aangetoond dat de statistische theorieen de verschillen in de statistieken, welke van de discretisaties afgeleid zijn, nauwkeurig verklaren.

In Hoofdstuk 3 hebben we lange tijdsimulaties uitgevoerd, waarbij we gebruik gemaakt hebben van een zogenaamde 'Hamiltonian particle-mesh' methode voor de ideale vloeistofstroming, om een statistisch-gemiddelde vorticiteitsveld voor de discretisatie te bepalen. We hebben de statistische modellen van Lagrange en Euler voor de discrete dynamica ontwikkeld en vervolgens getest door middel van numerieke experimenten. De verkregen resultaten zijn in goede overeenstemming met de theoretische voorspellingen, evenals met de continue statistisch-mechanische theorie voor de ideale vloeistofstroming, die werd ontwikkeld door Ellis, Haven & Turkington (*Nonlinearity*, 2002). De resultaten bevestigen met name dat de gemiddelde toestand aanzienlijk wordt beinvloed door het blijkbaar triviale behoud van de potentiele vorticiteit langs de deeltjespaden, bij het gebruik van de HPM-methode. Terzijde willen we tevens opmerken dat, zoals uit de numerieke experimenten is gebleken, een vierde ongelijk-nul moment kan het statistische gemiddelde beinvloeden.

Viscositeitsvrije vloeistofmodellen worden gekenmerkt door het behoud van energie, sterke gevoeligheid voor de beginvoorwaarden, en een cascade aan vorticiteit op een steeds fijnere schaal. Echter, een dergelijke vortisiteitscascade presenteert een uitdaging voor de numerieke simulaties van zulke stromingen, gezien het feit dat de directe discretisatie van de bewegingsvergelijking uiteindelijk geen goed opgeloste oplossingen meer oplevert. Om dit probleem effectief aan te pakken is het vereist om de dynamica op de onder-rooster schaal te modelleren, evenals de invloed daarvan op het grove rooster. In Hoofdstuk 4 hebben we een afsluiting voor de onsamendrukbare ideale vloeistofstroming gepresenteerd, in de vorm van een gegeneraliseerd apparaat voor een gewenste temperatuur, waarbij we gebruik hebben gemaakt van de puntvortex stroming op

een schijf als prototype. Deze techniek wordt gebruikt voor moleculaire simulaties, en met name, om de interacties van een deeltjes-systeem op constante temperatuur te modelleren. We hebben aangetoond dat de termostaat een oneindig of eindig reservoir met stochastisch gedwongen termostaatvariabelen kan modelleren. Numerieke experimenten zijn in uitstekende overeenkomst met de twee-schalige simulaties van Bühler (*Phys. Fluids*, 2002).

Acknowledgements

I would like to take the opportunity to thank my supervisors Jason Frank and Jan Verwer. They taught me a lot and they helped me in different aspects of my professional growth. Without them this thesis would not have existed.

I would like to thank Sebastian Reich and Ben Leimkuhler for fruitful discussions that we had during different conferences and workshops, for their help in writing the proposals.

I wish to express my gratitude to Vladislav Vasil'evich Pukhnachev, my adviser back to Russia, who has introduced me to the exciting world of research.

I am very thankful to my great tutor Vladimir Ivanovich Kocherzhinskiy, who showed me how wonderful mathematics is.

I would like to acknowledge my colleagues at CWI for their help and open-minded discussions. Many thanks to my officemates for creating a nice working atmosphere: Nga Pham Thi, Valeriu Savcenco, Peter van Heijster, Maciej Dobrzynski, Sjors van der Stelt, Mohammad Khalsaraei and Janis Bajars. When it comes to prompt administration of all kind of documents, my regards to Nada Mitrovic, Susanne van Dam, Martine Anholt-Gunzeln, Mike Zonsveld and Irma van Lunenburg. Thanks to Julia Zijlstra for helping me to translate the "Samenvatting".

In these adventure years of my PhD I have been fortuned with many friends. I would like to thank Nga & Thieu, Valera, Maria & Alexey, Valeria, Anna & Vasily, Antonis, Peter, Maciek, Janis, Stephanie & Nils, Giulia & Ovidiu for chattering lunches, lovely evenings, unbeatable dinners, enjoyable board and tennis games, and exciting water events.

I wish to thank my dear husband Denis, who has supported me all these years and who has been traveling to the Netherlands almost every week-end since August 2005.

Lastly, I would like to express my sincere appreciation to my dear mama, papa and my sister. I am very thankful for their support and understanding of my wishes, even though they took me far away from the family hearth.