Correcting Numerical Simulations for Known Expectations

Keith W. Myerscough

Correcting Numerical Simulations for Known Expectations

Keith W. Myerscough

cover graphic: depiction of the ordering of vortex pair interactions in a parallel implementation, see Chapter 4

Correcting Numerical Simulations for Known Expectations

ISBN 978-94-6299-131-6
Printing: Ridderprint BV

Correcting Numerical Simulations for Known Expectations

Het corrigeren van numerieke simulaties
voor waarheidsgetrouwe verwachtingswaarden
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht op gezag van
de rector magnificus, prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
maandag 24 augustus 2015 des middags te 4.15 uur

door

Keith Willem Myerscough

geboren op 8 november 1985, te Eindhoven

Promotor: Prof.dr.ir. J.E. Frank

The work described in this thesis was supported by a grant financed by the Netherlands Organisation for Scientific Research (NWO), as part of research programme 613.001.009.

To my daughter
Go for it and enjoy yourself

Preface

It has been an anxious wait, waiting for the reading committee to approve this thesis such that I may publicly defend it. Slightly over four years I have worked towards this thesis, and then it is sent to a panel of established scientists to decide whether the work is worthy for me to prove that I am capable of conducting independent research. The news I received today is a great relief and has also given me a sense of pride about the work that has conveniently motivated me to finally write this very preface.

While this thesis is ostensibly my personal work, the articles that form its main body (Chapters 3–5) were all cooperative efforts. I would like to take this opportunity to thank my co-authors, who were in fact much more than just that.

First and foremost, I wish to thank my supervisor, Jason Frank. I have learned a lot from you about the many aspects of (numerical) mathematics, but also about other aspects of scientific life. You gave me a lot of freedom to do what I wanted, but were also constructively critical of my work.

I am also very grateful for the involvement of Ben Leimkuhler. The periods I spent working with you in Edinburgh led to very interesting new ideas and your enthusiasm is a great motivator.

Then there are many, many people who may not have contributed to this work directly, but who have certainly made my time at CWI very enjoyable.

Wander, wat ben ik blij met jou een kantoor gedeeld te mogen hebben. Je hebt vaak interessante perspectieven geboden op werk, leven en alles daartussenin. Halldóra, ég var verkfræðingur, nú er ég stærðfræðingur. Og þú varst og ert frábær.

I will also greatly miss everyone from our little corridor: Daan, Jānis, Jesse, Nick, Deba, Zsolt and Bart. We had some great times, but more importantly we had a lot of good times. There were too many other kind colleagues at CWI to mention everyone, but I will name a few who have become especially dear to me. Linda, het was fijn dat ik altijd bij je langs kon als ik om een praatje verlegen zat. Klein van stuk maar groot van moed wens ik je al het beste toe. De officieuze huisfilosoof van het CWI heeft mij ook vaak geamuseerd; bedankt, Tom. Christoph deserves a special mention for being instrumental in setting up both the CWI football team and the PhD activity committee. These

both provided some much needed distraction from the occasionally frustrating research.

I would also like to thank my parents for their encouragement and support in all my endeavours. Mijn schoonfamilie dank ik voor hun geduld als ik moeilijk begon te kijken bij vragen over hoe het op mijn werk ging. And finally, having saved the best for last: Sanne, your unconditional love and affection have been of unmeasurable value to me. I could not have done this without you.

Leuven, June 15, 2015

Contents

Summary

Predictions of the future climate require long-time simulation of a chaotic dynamical system. This poses a challenge for numerical simulations, as these do not necessarily capture the correct long-term behaviour of chaotic systems. This problem is exacerbated by the wide range of length scales present in atmospheric and oceanic dynamics. The modeling choices for small scale processes have a large impact on long term statistics of the scales of interest. This thesis studies the dynamics of two different fluid models as a proxy for atmospheric dynamics: a point vortex flow on the sphere and two-dimensional turbulent flow on a torus. We apply gentle dynamical perturbations commonly used in molecular dynamics to these fluid dynamics problems as a means for improving the statistical veracity of low fidelity simulations.

For the point vortex system we study a system consisting of both strong and weak vortices. The strong vortex dynamics are mildly influenced by the weak vortices on a short time scale, but the presence of the weak vortices introduces a variability in the strong vortex energy. We mimic this behaviour in a model with only the strong vortices by gentle perturbations to the equations of motion. The perturbations have a stochastic forcing such that the modified dynamics ergodically sample an invariant measure consistent with observations from the strong vortex system in contact with the weak vortices. We choose the invariant measure as the minimal entropy density consistent with observations. The required Lagrange multipliers can be computed either a priori using a sample set in some prior distribution, or they can be computed on-the-fly using the simulation history as an ensemble. The latter method allows the observations to be updated during runtime, providing the opportunity for data-assimilation.

We construct a Poisson integration method for the aforementioned point vortex dynamics by splitting the Hamiltonian into its constituent vortex pair terms. The method provides exact solutions to a Poisson system with the same bracket as the original dynamical system, but with a modified Hamiltonian function. Different orderings of the pairwise interactions are considered and are also used for the construction of higher order methods. The energy and momentum conservation of the splitting schemes is demonstrated for several test cases. For particular orderings of the pairwise interactions, the

schemes allow scalable parallelization. This results in a linear – as opposed to quadratic – scaling of computation time with system size.

We also explore the direct modification of the pseudo-spectral truncation of two-dimensional, incompressible fluid dynamics on a torus to maintain a prescribed kinetic energy spectrum. The method provides a means of simulating fluid states with defined spectral properties, for the purpose of matching simulation statistics to given information, arising from observations, theoretical predictions or high fidelity simulations. In the scheme outlined here, Nosé-Hoover thermostats – commonly used in molecular dynamics – are introduced as feedback controls applied to energy bands of the Fourier-discretized Navier-Stokes equations. As we demonstrate in numerical experiments, the dynamical properties – quantified using autocorrelation functions – are only modestly perturbed by our device, while ensemble dispersion is significantly enhanced in comparison with simulations of a corresponding truncation using hyperviscosity.

Samenvatting

Voorspellingen van het toekomstige klimaat vereisen lange simulaties van een chaotisch dynamisch systeem. Dit vormt een uitdaging voor numerieke simulaties, omdat deze niet noodzakelijk het juiste langetermijngedrag weergeven van chaotische systemen. Dit probleem wordt verergerd door het brede bereik van lengteschalen in globale lucht- en waterstromingen. De modelkeuzes voor processen op kleine schaal hebben een grote invloed op de langetermijnstatistieken van de lengteschalen van belang. Dit proefschrift bestudeert twee verschillende vloeistofmodellen als surrogaat voor atmospherische dynamica: een puntvortexmodel op een bol en twee-dimensionale turbulentie op een torus. We voegen zachte dynamische verstoringen zoals gebruikt in moleculaire dynamica toe aan deze vloeistofmodellen met als doel het verbeteren van de statistische nauwkeurigheid van laagdimensionale simulaties.

Voor het puntvortexsysteem bestuderen we een systeem dat bestaat uit sterke en zwakke vortices. De beweging van de sterke vortices wordt op korte tijd slechts mild beïnvloed door de zwakkere, maar de aanwezigheid van de zwakke vortices geeft een schommeling in de energie van de sterke vortices. Wij bootsen dit gedrag na in een model met alleen sterke vortices door zachte verstoringen van de bewegingswetten. De verstoringen hebben een stochastische drijfkracht opdat de aangepaste dynamica ergodisch is in een invariante verdeling die consistent is met waarnemingen uit het systeem waar de sterke vortices beïnvloed worden door de zwakke vortices. Wij kiezen hiervoor die verdeling met minimale entropie doch consistent met de gegeven waarnemingen. De benodigde Lagrange-multiplicatoren kunnen dan wel worden berekend met gebruik van een ensemble uit een bepaalde priorverdeling, dan wel kunnen ze à la minute worden berekend met de simulatiegeschiedenis als ensemble. Deze tweede methode staat toe dat de waarnemingen veranderen tijdens de simulatie, wat een mogelijkheid biedt voor data-assimilatietoepassingen.

We construeren een Poisson integratietechniek voor het hierbovengenoemde puntvortexsysteem door de Hamiltoniaan op te delen in zijn paargewijze termen. De methode levert exacte oplossingen voor een Poisson systeem met dezelfde *bracket* als het oorspronkelijke probleem, maar een aangepaste Hamiltoniaan. We beschouwen verschillende volgorden van de paargewijze interac-

ties en construeren ook methodes van hogere orde. Het behoud van energie en impuls wordt getoond voor verschillende testgevallen. Voor bepaalde volgorden van de paargewijze interacties staat het schema schaalbare parallelisatie toe. Dit resulteert in een lineaire – in plaats van kwadratische – toename van rekentijd met systeemgrootte.

We onderzoeken ook de aanpassing van pseudo-spectrale afkapping van twee-dimensionale, onsamendrukbare vloeistofdynamica op een torus om een voorgeschreven kinetisch energiespectrum te behouden. Deze methode biedt een manier om vloeistoffen te simuleren met voorgeschreven spectrale eigenschappen, om simulatiestatistieken gelijk te stellen aan gegeven informatie, uit waarnemingen, theoretische voorspellingen of nauwkeurigere simulaties. In dit schema worden Nosé-Hoover thermostaten – veelvuldig gebruikt in moleculaire dynamica – geïntroduceerd als terugkoppelingsmechanismen toegepast op energiebanden van de Fourier-gediscretiseerde Navier-Stokes vergelijkingen. Zoals we aantonen in numerieke simulaties, zijn de dynamische eigenschappen – gekwantificeerd met autocorrelatiefuncties – slechts weinig aangetast door ons mechanisme, terwijl de ensemblespreiding significant is verbeterd ten opzichte van simulaties met een soortgelijke afkapping en hyperviscositeit.

1

Introduction

Predictions of the future states of dynamical systems allow individuals, groups, or society as a whole to anticipate what is to come and to act accordingly. When the time span of such a prediction is short and the dynamics of the system are well understood, it is possible to make accurate predictions of what will happen when. Over a longer period, however, the behaviour of complex systems will diverge from even the best possible simulations. In such cases we have to relax our goal. Instead of focussing on prediction of exact events on exact dates, we focus on statistical prediction: providing probabilities for events and how these probabilities evolve in time.

An important field of study where such difficulties arise is the behaviour of our climate. While the weather on a fortnightscale shows evidence of predictability, the climate can only be viewed in a probabilistic way. Rather than requiring accurate predictions of events, this needs accurate estimation of the likelihood of events. If events are over– or underestimated probabilistically, the model is said to be biased. The most prominent source of modelling bias is the inability to resolve relevant small-scale dynamics. Because the small scale dynamics play an important role in dissipating kinetic energy from larger scales, they cannot simply be ignored. The choice of model used to represent the effect of the unresolved scales on the resolved scales has severe consequences for the statistics of simulations of the resolved system.

In many other applications there is a similar necessity to correctly model the effect of unresolved components on the dynamics of interest. For instance in molecular dynamics, there is a clear distinction between fast and slow degrees of freedom. The fast degrees of freedom might not be of practical interest, but their variability does play an important role in the dynamics of the other degrees of freedom.

The work in this thesis is aimed at exactly this class of problems where the time scale of interest is long compared to the time scale at which solu-

tions chaotically diverge. At the same time, however, we still wish to simulate dynamical trajectories. This is necessary as some predictions are in fact dynamical events, rather than instantaneous occurrences. Let us illustrate this with an example. The KNMI defines a heat wave as a succession of five "summery days" (temperature over 25°C) of which at least three are "tropical days" (over 30°C) [1]. To predict the likelihood of a heat wave therefore requires not only knowledge of the different temperatures that might occur, but also how the temperature changes from day to day. This means that we need both short term dynamical accuracy and long term statistical veracity. The goal of this work is to reconcile these conflicting long and short term interests in various settings.

The following chapter provides some introductory information on the numerical integration and sampling techniques used throughout Chapters 3–5. In Chapter 3 we discuss the use of observational data in generating simulations of Hamiltonian systems with good long-term statistical properties. Chapter 4 discusses the time integration of a point vortex system, which is used as a model Hamiltonian problem in Chapter 3. The techniques for incorporating observational data into dynamical simulations are applied to two-dimensional turbulence in Chapter 5.

2

Background

The sections in this chapter provide some introductory information on the subjects of the articles presented here as Chapters 3–5. Most of this explanation can also be found in numerous monographs, but it is included here for the convenience of the reader. Some of the topics are also discussed in more detail in the extensive introduction to the PhD thesis of Bajārs [8]. Those already familiar with topics can skip sections without loss of continuity.

2.1 Dynamical systems

In many applications in physics, biology, sociology, economics etc. it is useful to describe the object of study in terms of a dynamical system. A dynamical system consists of two things: the state space D is the collection of all possible states of the system and the map ϕ_τ maps the current state of the system to its state τ time units later. We denote this mapping as

$$\begin{aligned} \phi_\tau &: D \to D \\ & y(t) \mapsto y(t+\tau), \end{aligned} \tag{2.1}$$

where $y(t) \in D$ is the state of the system at time t. The state y can be finite dimensional, in which case the state is given by a state vector $\boldsymbol{y} = (y_1, y_2, \ldots, y_N)^T \in D \subset \mathbb{R}^N$, or infinite dimensional, in which case no such finite representation of the state is possible.

Naturally, to be able to perform computations, we need to be able to represent the system state in terms of a finite set of numbers. This means infinite dimensional problems need to be discretized into a finite representation. The opening sections of Chapter 4 of this thesis are about finding a finite dimensional representation to an infinite dimensional problem. Most of this thesis, however, relates to simulating a given finite dimensional system.

Dynamical systems are often related to an initial value problem for a system of differential equations. Such a problem is defined by a differential equation and an intial condition:

$$\begin{aligned} \dot{y} &= f(y(t), t) \\ y(0) &= y_0. \end{aligned} \tag{2.2}$$

In this thesis all the problems are autonomous, meaning that $f(y(t), t) = f(y(t))$ does not depend explicitly on the time t. The solution to such an initial value problem is given by

$$y(t) = y(0) + \int_0^t f(y(s))\,\mathrm{d}s.$$

As long as such a solution exists, we may write this as

$$y(t) = \phi_t y(0).$$

This provides the dynamical system associated to the differential equation (2.2).

2.1.1 Hamiltonian and Poisson systems

An important class of conservative systems, i.e. systems that do not dissipate energy, can be written in the form

$$\dot{\boldsymbol{y}} = J(\boldsymbol{y})\nabla H(\boldsymbol{y}), \tag{2.3}$$

where $\boldsymbol{y} \in \mathbb{R}^N$ and $H(\boldsymbol{y}) : \mathbb{R}^N \to \mathbb{R}$. The *structure matrix* $J(\boldsymbol{y}) \in \mathbb{R}^{N\times N}$ satisfies the skew-symmetry propety $J(\boldsymbol{y}) = -J^T(\boldsymbol{y})$, for all $y \in D$ and the Jacobi identity:

$$\sum_{\ell=1}^{N} \frac{\partial J_{ij}(\boldsymbol{y})}{\partial \boldsymbol{y}_\ell} J_{\ell k} + \frac{\partial J_{jk}(\boldsymbol{y})}{\partial \boldsymbol{y}_\ell} J_{\ell i} + \frac{\partial J_{ki}(\boldsymbol{y})}{\partial \boldsymbol{y}_\ell} J_{\ell j} = 0,\ \forall i, j, k = 1, \ldots, N.$$

Such a system is called a Poisson system. If the structure matrix does not depend on $\boldsymbol{y}$ it is a Hamiltonian system (and the Jacobi identity is trivially satisfied). In either case $H(\boldsymbol{y})$ is called the Hamiltonian and it is invariant under the motion $\boldsymbol{y}(t)$, i.e. $H(\boldsymbol{y}(t)) = H(\boldsymbol{y}(0)), \forall t$, because

$$\dot{H} = \nabla H \cdot \dot{\boldsymbol{y}} = \nabla H \cdot J(\boldsymbol{y})\nabla H = 0,$$

where the last equality follows from the skew-symmetry of $J(\boldsymbol{y})$.

A Hamiltonian system with the state vector $\boldsymbol{y} = (\boldsymbol{q}, \boldsymbol{p})$ consisting of positions $\boldsymbol{q} \in \mathbb{R}^n$ and momenta $\boldsymbol{p} \in \mathbb{R}^n$ with the structure matrix

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}, \tag{2.4}$$

where I_n is the n-dimensional identity matrix, is a *canonical* Hamiltonian system. These arise when considering the motion of a particle according to Newton's second law under the influence of conservative forces. Two-dimensional point vortex systems on a planar geometry also form a canonical Hamiltonian system.

Associated with the system (2.3) is the *Poisson bracket* defined by

$$\{F, G\} = \nabla F \cdot J(\boldsymbol{y}) \nabla G, \quad F, G : \mathbb{R}^N \to \mathbb{R}.$$

The Poisson bracket is skew-symmetric, i.e. $\{F, G\} = -\{G, F\}$, and satisfies the Jacobi identity

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

The time rate of change for an observable $F(\boldsymbol{y}) : \mathbb{R}^N \to \mathbb{R}$ under the motion $\boldsymbol{y}(t)$ is given by $\{F, H\}$. Invariants of the system[1] are now easily identified as functions $I : \mathbb{R}^N \to \mathbb{R}$ for which $\{I, H\} = 0$. There may or may not be more invariants than the Hamiltonian H. Given invariants $H, I_1, \ldots, I_m$, any $A(\boldsymbol{y}) = A(H(\boldsymbol{y}), I_1(\boldsymbol{y}), \ldots, I_m(\boldsymbol{y}))$ is also an invariant.

A particular type of invariant that may occur are Casimirs C, for which $\{F, C\} = 0, \forall F$, note that the Casimirs follow from the form of Poisson bracket alone, irrespective of the Hamiltonian.

These properties of Poisson systems play an important role in Chapters 3 and 4. The nonlinear interaction of two-dimensional turbulence as discussed in Chapter 5 also possesses a well-known Poisson structure. [105, 131, 7]

2.2 Numerical integration of dynamical systems

The goal of numerical integration is to approximate the state of a system (2.2) at different times $t_1 < t_2 < \ldots$, given the state at some initial time t_0. We restrict this to the case of finding solutions at regular intervals, with the initial condition given at $t_0 = 0$. In that case we have $t_n = n\tau$ for $n = 0, 1, 2, \ldots$.

Most numerical integrators, and certainly those discussed in this thesis, consist of a fixed rule for obtaining the system state at a next time step given the current state. Such integrators provide approximations to the flow map ϕ_τ of (2.1). We denote the approximate map Φ_τ and use it to generate approximations to the exact system state $\boldsymbol{y}(n\tau)$ denoted by $\tilde{\boldsymbol{y}}(n\tau)$

$$\tilde{\boldsymbol{y}}(t + \tau) = \Phi_\tau \tilde{\boldsymbol{y}}(t).$$

Starting from the exact initial condition $\tilde{\boldsymbol{y}}(0) = \boldsymbol{y}(0) = \boldsymbol{y}_0$, the solution after n time steps of the numerical scheme reads

$$\tilde{\boldsymbol{y}}(n\tau) = (\Phi_\tau)^n \boldsymbol{y}_0.$$

[1] or: *first integrals* of the system

The following two sections discuss the accuracy of the numerical integration in different ways.

2.2.1 Forward Error Analysis

The difference between the numerical and exact solutions at some fixed time T provides a measure of accuracy for the numerical method. In particular one is usually interested in how the error depends on the choice of time step τ for the numerical scheme. We write this error as

$$\varepsilon_\tau(T) = || \left(\Phi_\tau\right)^{T/\tau} \boldsymbol{y}_0 - \phi_T \boldsymbol{y}_0 ||, \tag{2.5}$$

assuming that T is an integer multiple of τ. The size of this error depends on the right-hand-side of the original system and on the chosen initial condition $\boldsymbol{y}_0$. In many cases, however, it is possible to make a general statement about how the error is expected to vary with the time step. In particular the *order* of an integration method is given by d in the following

$$\varepsilon_\tau(T) = C_f(T, \boldsymbol{y}_0)\tau^{-d},$$

where $C_f(T, \boldsymbol{y}_0)$ depends on the system, the end time and the initial conditions, but *not* on the chosen step size. Typically, higher order methods require a larger computational effort per time step, but allow for a larger time step for the same accuracy. Chapter 4 contains a comparison between methods of different orders. The optimal choice depends greatly on the required accuracy.

2.2.2 Backward Error Analysis

A large drawback to forward error analysis occurs when one is interested in the behaviour of chaotic systems over long periods of time. Solution trajectories of chaotic systems with the tiniest of perturbations to the initial conditions inevitibly diverge after some time – even exact solutions. Numerical methods therefore require an exponentially small time step for accurate simulations over long time. This is unfeasable, and motivates viewing the accuracy in a different way, that does not result in impractical demands on the numerical method.

An alternative view to accuracy is given by studying the properties of the approximate flow map Φ_τ compared to the exact map ϕ_τ. Crucial to *Backward Error Analysis* is the notion that numerical solutions to a particular problem are in fact exact solutions to some modified problem. Finding the definition of this modified problem – or at least obtaining some of its properties – allows one to make confident statements on the behaviour of the numerical solutions.

Let us illustrate this with the example of a simple harmonic oscillator defined by

$$\begin{aligned}\dot{q} &= p \\ \dot{p} &= -\omega^2 q,\end{aligned}$$

with $q, p \in \mathbb{R}$ and frequency $\omega > 0$. This is a canonical Hamiltonian system with structure matrix as in (2.4) and Hamiltonian $H(q,p) = \frac{1}{2}p^2 + \frac{\omega^2}{2}q^2$. We use four different numerical integrators to simulate trajectories of this system:

1. Explicit Euler

$$\begin{aligned}q^{n+1} &= q^n + p^n \tau \\ p^{n+1} &= p^n - \omega^2 q^n \tau\end{aligned}$$

2. Implicit Euler

$$\begin{aligned}q^{n+1} &= q^n + p^{n+1} \tau \\ p^{n+1} &= p^n - \omega^2 q^{n+1} \tau\end{aligned}$$

3. Semi-Implicit Euler

$$\begin{aligned}q^{n+1} &= q^n + p^n \tau \\ p^{n+1} &= p^n - \omega^2 q^{n+1} \tau\end{aligned}$$

4. Störmer-Verlet

$$\begin{aligned}q^* &= q^n + p^n \frac{1}{2}\tau \\ p^{n+1} &= p^n - \omega^2 q^* \tau \\ q^{n+1} &= q^* + p^{n+1} \frac{1}{2}\tau.\end{aligned}$$

We also present the exact solution for a single step

$$\begin{aligned}q^{n+1} &= q^n \cos(\omega\tau) + \frac{1}{\omega} p^n \sin(\omega\tau) \\ p^{n+1} &= -\omega q^n \sin(\omega\tau) + p^n \cos(\omega\tau).\end{aligned}$$

The exact solution returns to its initial state after a period of $T = \frac{2\pi}{\omega}$. We simulate an ensemble of four initial conditions from $t = 0$ up to $t = T$ using each of the four numerical integrators with time step $\tau = \frac{T}{10}$. Figure 2.1 shows the results every other time step and also shows the exact solution at these times. The four ensemble members are connected by lines to form a quadrilateral. The plots also show dotted contour lines of the Hamiltonian, consistent with the initial values.

The results show that the exact solution returns exactly to the initial conditions and that each of the ensemble members remains exactly on the contour

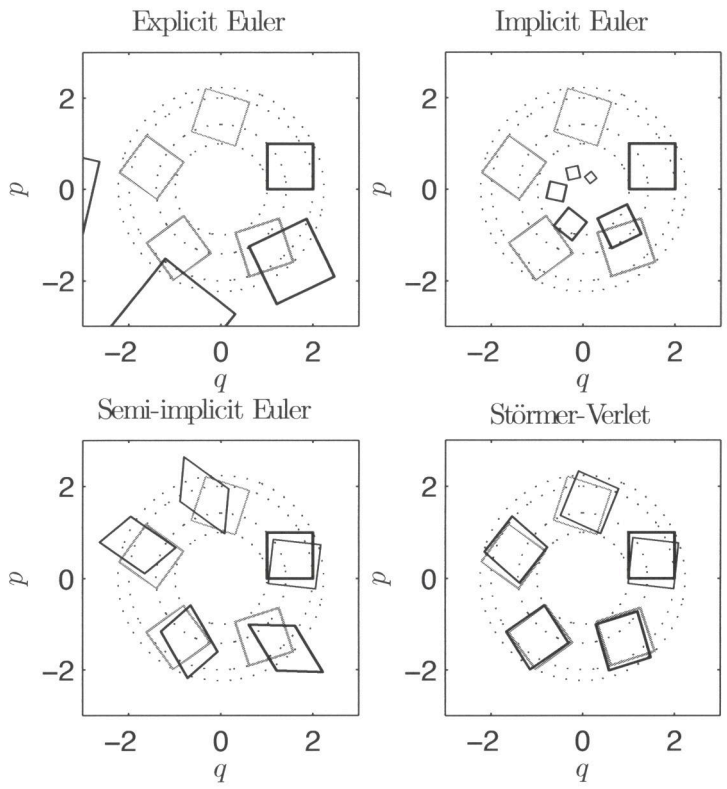


Figure 2.1: Simulations of a Harmonic Oscillator. Four numerical integrators with different properties compared against the exact solution (in grey). Each quadrilateral vertex represents the positions of an ensenble of four simulations. Decreasing line width indicates increasing time.

lines of the Hamiltonian throughout. In fact, the square undergoes rigid body rotation. The Explicit Euler scheme results in a divergence of the trajectories, whereas the trajectories for the Implicit scheme collapse towards the origin. The Semi-Implicit Euler scheme and the Störmer-Verlet scheme remain close to the orbits of the exact solution.

The difference in the stability of these four different integration schemes follows directly from studying the form of the operators Φ_τ. For each of these methods, we may write

$$\tilde{\boldsymbol{y}}^{n+1} = \Phi_\tau \tilde{\boldsymbol{y}}^n = A_\tau \tilde{\boldsymbol{y}}^n,$$

with a 2×2 matrix A that depends on the integration scheme. The exact system ϕ_τ may also be written in such a linear manner.

The long-term behaviour of the system is characterized by the eigenvalues

of these matrices. For the exact solution, this matrix has eigenvalues $\lambda = \cos\omega\tau \pm \sqrt{1-\cos^2\omega\tau}$. The magnitude of both these eigenvalues is exactly one (regardless of step size τ), implying the the origin is a centre, and solutions orbit this centre indefinitely, without growth or decay.

For the Explicit Euler scheme, both eigenvalues have magnitude greater than one, leading to an unconditionally unstable system. The Implicit Euler scheme has eigenvalues with magnitude less than one, leading to the origin being an asymptotically stable point to which all solutions converge. The Semi-Implicit Euler method and the Störmer-Verlet method each have eigenvalues with magnitude exactly unity, under the condition that $\tau^2\omega^2 \leq 4$ for the Semi-Implicit Euler and $\tau^2\omega^2 \leq 2$ for Störmer-Verlet. This implies these trajectories orbit a centre at the origin indefinitely.

We are interested in methods that can be used for long simulations. The unstable growth of the Explicit Euler scheme, and the decay of the Implicit Euler scheme make them both unsuitable for such applications. A smaller time step would slow down the growth or decay, but it cannot prevent it in the long run. Thefore let us focus on the Semi-Implicit Euler and the Störmer-Verlet methods.

The Semi-Implicit Euler scheme advances the solution by first updating the position to q^{n+1} based on the current momentum p^n, and then uses this updated position to find the new momentum p^{n+1}. This effectively splits the dynamical system in two parts and we write

$$\Phi_\tau^{\mathrm{SIE}} = \phi_\tau^{[1]} \circ \phi_\tau^{[2]}, \tag{2.6}$$

where $\phi_\tau^{[1]}$ represents the step updating the position and $\phi_\tau^{[2]}$ that updating the momentum. Note that these operators are evaluated left to right.

The components of this splitting are in fact solutions to two Hamiltonian systems with the same structure matrix as in (2.4) and the two separate terms of the Hamiltonian, i.e. $H^{[1]} = \frac{1}{2}p^2$ and $H^{[2]} = \frac{1}{2}\omega^2 q$. For such a composition, the modified problem is in itself a Hamiltonian system, but with a modified Hamiltonian [62], which we write as a power series expansion

$$\tilde{H} = H + \tau H_2 + \tau^2 H_3 + \tau^3 H_4 + \dots. \tag{2.7}$$

For the splitting (2.6) the correction terms follow from the Poisson bracket between the two separated terms of the Hamiltonian [62] as

$$\begin{aligned}
H_2 &= \frac{1}{2}\left\{H^{[1]}, H^{[2]}\right\} = \frac{1}{2}\omega^2 qp \\
H_3 &= \frac{1}{12}\left\{\left\{H^{[1]}, H^{[2]}\right\}, H^{[2]}\right\} + \frac{1}{12}\left\{\left\{H^{[2]}, H^{[1]}\right\}, H^{[1]}\right\} \\
&= \frac{1}{12}\omega^4 q^2 + \frac{1}{12}\omega^2 p^2 \\
H_4 &= \frac{1}{24}\left\{\left\{\left\{H^{[1]}, H^{[2]}\right\}, H^{[2]}\right\}, H^{[1]}\right\} = \frac{1}{12}\omega^4 qp.
\end{aligned}$$

In this particular example, there is some fortunate cancellation of terms. Observe that $H_3 = \frac{1}{6}\omega^2 H$ and that $H_4 = \frac{1}{6}\omega^2 H_2$. With this we find that $\tau^2 H_3 + \tau^3 H_4 = \frac{1}{6}\tau^2\omega^2 (H + \tau H_2)$. Similar cancellations in the higher order corrections lead to the conclusion that the Semi-Implicit Euler scheme has the modified Hamiltonian

$$\tilde{H}^{\text{SIE}} = \left(H + \frac{1}{2}\tau\omega^2 qp\right) \times \left(1 + \frac{1}{6}\tau^2\omega^2 + \dots\right). \tag{2.8}$$

The Störmer-Verlet method is also a splitting scheme, now with the composition

$$\Phi_\tau^{\text{StV}} = \phi_{\tau/2}^{[1]} \circ \phi_\tau^{[2]} \circ \phi_{\tau/2}^{[1]}. \tag{2.9}$$

The components are solutions to the same Hamiltonian system as before, but now evaluated in a symmetric order. This symmetry implies a cancelation of the terms with odd powers of τ in (2.7). Specifically the terms are given by [62]

$$\begin{aligned} H_2 &= H_4 = H_{2i} = 0 \\ H_3 &= -\frac{1}{24}\left\{\left\{H^{[2]}, H^{[1]}\right\}, H^{[1]}\right\} + \frac{1}{12}\left\{\left\{H^{[1]}, H^{[2]}\right\}, H^{[2]}\right\} \\ &= -\frac{1}{24}\omega^4 p^2 + \frac{1}{12}\omega^4 q^2 \end{aligned}$$

Substituting these terms into (2.7) we find the modified Hamiltonian

$$\tilde{H}^{\text{StV}} = H + \frac{1}{24}\tau^2\left(\omega^4 p^2 + 2\omega^4 q^2\right) + \mathcal{O}(\tau^4). \tag{2.10}$$

The modified Hamiltonians for the Harmonic Oscillator when using the Semi-Implicit Euler or Störmer-Verlet scheme given in (2.8) and (2.10) may not immediately show that solutions remain close to the exact solution in the sense of (2.5), but it does tell us that the dynamics of the modified problem are similar to those of the physical problem of interest. When studying the long time behaviour of chaotic systems the accuracy is lost, this is a much stronger measure of fidelity for a numerical scheme.

The numerical methods detailed in Chapter 4 for studying point vortex systems are splitting schemes just like (2.6) and (2.9). We study their error similarly by computing the correction terms in the modified Hamiltonian.

2.3 Stochastic processes

In a stochastic process the system no longer has a deterministically prescribed path, but all possible paths have a prescribed probability. A stochastic process is thus a sequence of dependent stochastic variables, denoted

$$X = \{X_t : t \in T\},$$

where T is a totally ordered set of times. For any $t \in T$, the process takes a value in the phase space D of the process, i.e. $X_t \in D$. The entire process is an element in the probability space $(\Omega, \mathcal{F}, \mu)$, where Ω is the space containing all possible paths, $\mathcal{F}$ is a sigma-algebra of events and μ is a measure over Ω such that

$$P(X \in A) = \int_A \mathrm{d}\mu \quad \forall A \in \mathcal{F}.$$

In the case that the stochastic process is defined at a finite set of times $T' = \{t_1, t_2, \ldots, t_k\}$, the process $X = \{X_t : t \in T'\}$ is an element of the space $\Omega' = D^k$, i.e. the product of k phase spaces D.

If the history of the process up to some time t is known, the set of possible events is reduced. This set of events is denotes as the *filtration* $\mathcal{F}_t \subset \mathcal{F}$. For $s < t$ it holds that $\mathcal{F}_s \subseteq \mathcal{F}_t$.

A stochastic process is said to be Markovian (or a Markov process) if future probabilities depend only on the present state of the system, and not on its history. More precisely, the process X is Markovian if for any $s, t \in T$ with $s < t$ it holds that

$$P(X_t \in A | \mathcal{F}_s) = P(X_t \in A | X_s) \quad \forall A \in \mathcal{G},$$

where $\mathcal{G}$ is the sigma-algebra associated to D. Many stochastic processes are constructed to be Markov processes. In modeling this choice is usually justified by the fact that the random components of the system are fast components. When there is a dependence on the past states, a Markov process can be recovered by extending the phase space to include previous time steps, similar to the transformation of a system of n ODEs of k-th order to a system of nk first-order ODEs.

2.3.1 The Wiener process

The Wiener process forms the basis for all stochastic processes discussed in this thesis. It is a continuous-time stochastic process denoted by W_t and characterized by three properties [45]

1. $W_0 = 0$,
2. the function $t \to W_t$ is almost surely everywhere continuous,
3. W_t has independent increments with $W_t - W_s$ $(0 \leq s < t)$ normally distributed with mean zero and variance $t - s$.

The independence of subsequent increments implies the Wiener process is Markovian.

2.4 Stochastic differential equations

Stochastic differential equations (SDEs) are stochastic processes that are typically defined by a deterministic term in combination with a noise term. We

write an SDE in its typical form

$$\mathrm{d}X_t = h(X_t, t)\,\mathrm{d}t + \sigma(X_t, t)\,\mathrm{d}W_t, \tag{2.11}$$

which is a shorthand notation of the integral equation

$$X_{t+s} - X_t = \int_t^{t+s} h(X_u, u)\,\mathrm{d}u + \int_t^{t+2} \sigma(X_u, u)\,\mathrm{d}W_u. \tag{2.12}$$

Because the Wiener process is nowhere differentiable, ordinary calculus is insufficient to work with equations (2.11) and (2.12). The study of such systems therefore requires the extension of identities from ordinary calculus to the stochastic realm. Two dominant competing stochastic calculuses exist, named Itō calculus and Stratonovich calculus. In short, the difference is due to the choice of a different type of Riemann sum in the stochastic integral in (2.12). More details on their principles and the practical effect thereof can be found in, for example, the book by Øksendal [115]. Throughout this thesis we will use Itō calculus.

An example where one encounters the need for this extended calculus is when considering the differential of a quantity that is a function of the outcome of a stochastic process such as that in (2.11). Consider a function $f(t, x) : \mathbb{R}^2 \to \mathbb{R}$, twice differentiable in x. To find the differential of f we would ordinarily use the chain rule with respect to its two arguments. Let us take a step back and proceed formally with the Taylor expansion of f

$$\mathrm{d}f = \frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}\,\mathrm{d}t^2 + \frac{\partial^2 f}{\partial t \partial x}\,\mathrm{d}t\,\mathrm{d}x + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\,\mathrm{d}x^2 + \ldots.$$

We substitute the process X_t of (2.11) for $\mathrm{d}x$ to find

$$\mathrm{d}f = \frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial x}\,(h\,\mathrm{d}t + \sigma\,\mathrm{d}W_t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\,(h\,\mathrm{d}t + \sigma\,\mathrm{d}W_t)^2 + \ldots. \tag{2.13}$$

The first two second-order terms (those with $\mathrm{d}t^2$ and $\mathrm{d}t\,\mathrm{d}x$) have been omitted as they yield only negligible products of differentials. The $\mathrm{d}x^2$ term, however, gives rise to a term with $\mathrm{d}W_t^2$. Due to the definition of the Wiener process, this term has expected value $\mathrm{d}t$ and should not be neglected. This is exactly the difference between the "ordinary" chain rule and Itō's lemma:

$$\mathrm{d}f = \frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial x}\,(h\,\mathrm{d}t + \sigma\,\mathrm{d}W_t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma^2\,\mathrm{d}t, \tag{2.14}$$

which follows from the limit of (2.13) as $\mathrm{d}t \to 0$.

The terms h and σ in the equations above are known as the drift and diffusion coefficients, respectively. Equation (2.11) may be generalized to a system of n coupled stochastic differential equations. Wiener increments $\mathrm{d}W_t$ are now taken from $m \leq n$ independent Wiener processes and $\sigma(X_t, t)$ is an $m \times n$ matrix.

2.4.1 Solutions to stochastic differential equations

When studying numerical solutions of deterministic differential equations in Section 2.2, we could safely assume there was a unique analytical solution for given initial conditions that we desire our method to match. With SDEs as in (2.11), the solution can no longer be characterized by a single value as it is a stochastic value at each time. We consider three different ways of providing solutions to an SDE:

1. Monte-Carlo simulation,
2. evolution of expectations of observables, and
3. evolution of the probability density.

Monte-Carlo simulation

In Monte-Carlo simulation the goal is to generate trajectories with a specified probability distribution. The expectation of functions of the stochastic process is then estimated by the average of that function for many such trajectories.

When the stochastic term in the stochastic process depends only on the current system state, Monte-Carlo trajectories can be generated step-by-step similarly to how one would ordinarily simulate a deterministic dynamical system.

As an example let us consider the Ornstein-Uhlenbeck (OU) process, a well-known process that we use in Chapter 3. The one-dimensional OU process is defined by

$$\mathrm{d}X_t = \gamma\left(\mu - X_t\right)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t,$$

where μ is the mean of the invariant distribution, γ defines the strength of the *mean-reverting property* and σ is the diffusion, as before.

A time step of a stochastic process requires an expression for $X_{t+\tau}$ given X_t. Variation of parameters leads to such an expression by considering the function $f(X_s, s) = X_s e^{\gamma s}$ and applying Itō's Lemma (2.14) and finding

$$\mathrm{d}f(X_s, s) = e^{\gamma s}\,\mathrm{d}X_s + \gamma X_s e^{\gamma s}\,\mathrm{d}t = e^{\gamma s}\gamma\mu\,\mathrm{d}t + e^{\gamma s}\sigma\,\mathrm{d}W_s.$$

We integrate this from t to $t+\tau$ and find

$$X_{t+\tau}e^{\gamma(t+\tau)} = X_t e^{\gamma t} + \int_t^{t+\tau} e^{\gamma s}\gamma\mu\,\mathrm{d}t + \int_t^{t+\tau} e^{\gamma s}\sigma\,\mathrm{d}W_s$$

and consequently

$$X_{t+\tau} = X_t e^{-\gamma\tau} + \mu\left(1 - e^{-\gamma\tau}\right) + e^{-t-\tau}\int_t^{t+\tau} e^{\gamma s}\sigma\,\mathrm{d}W_s.$$

The remaining integral is evaluated using the Itō isometry [115] to arrive at

$$X_{t+\tau} = X_t e^{-\gamma\tau} + \mu\left(1 - e^{-\gamma\tau}\right) + \sigma\sqrt{\frac{1 - e^{-2\gamma\tau}}{2\gamma}}\, R, \tag{2.15}$$

where $R \sim \mathcal{N}(0,1)$ is a unit normal, chosen independently for each time step. A Monte-Carlo simulation of the Ornstein-Uhlenbeck process consists of time stepping using equation (2.15), using independent unit normal (pseudo-)random numbers for R at each time step.

Estimates of expected values follow from averaging over an ensemble of Monte-Carlo samples. For a deterministic function $c(X)$ of the stochastic process an unbiased estimator using n sample trajectories is given by

$$\hat{c}_n = \frac{1}{n}\sum_{i=1}^{n} c(X^{[i]}),$$

where the superscript denotes the index of the ensemble member. The observables c can be instantaneous observations depending only on the current system state, i.e. $c(t) = g(X_t)$, or functions of the entire process, e.g. correlation functions. Initial conditions are sampled from the initial distribution of the simulation, which may be a Dirac delta distribution.

In the absence of an exact solution for trajectories of the SDE there are different possible strategies. The simplest approach is the Euler-Maruyama method [81], similar to the Explicit Euler scheme of Section 2.2.2. It defines a time step of the SDE (2.11) as

$$\tilde{X}^{n+1} = \tilde{X}^n + h(\tilde{X}^n, n\tau)\tau + \sigma(\tilde{X}^n, n\tau)\Delta W^n,$$

where the ΔW^n are increments of the Wiener process $\Delta W^n = W_{(n+1)\tau} - W_{n\tau}$. The increments can be taken from a given Wiener process, or generated as independent (pseudo-)random numbers with mean zero and variance τ.

The Euler-Maruyama method converges to the exact solution (with the same Wiener process) with order $\frac{1}{2}$. Milstein [103] provides an improvement using a correction term that increases the order of convergence to 1.

Evolution of observables

Given an SDE (2.11) the *generator* $\mathcal{L}$ is given by [58, 118]

$$\mathcal{L}v = h \cdot \nabla v + \frac{1}{2}\sigma\sigma^T : \nabla\nabla v,$$

where the colon denotes the Frobenius inner product. For some observable $c(t) = \mathbb{E}\left[g(X_t)\right]$ we consider the function

$$u(x,t) = \mathbb{E}\left[g(X_t)|X_0 = x\right].$$

This defines the expectation of an observable as a function of time t and the initial state $X_0 = x$. The evolution of this expectation is given by the *backward Kolmogorov equation*

$$\begin{aligned} \frac{\partial u}{\partial t} &= \mathcal{L}u, \quad \text{for } (x,t) \in D \times (0,\infty) \\ u &= g(x), \quad \text{for } (x,t) \in D \times \{0\}\,. \end{aligned}$$

For an SDE with n degrees of freedom, the backward Kolmogorov equation is an n-dimensional PDE. The complexity of such a system can be prohibitive for finding numerical solutions.

Evolution of the probability density

We define an evolving probability density $\rho(y,t)$ indicating the likelihood of finding the state Y_t near y; more precisely

$$P(Y_t \in A) = \int_A \rho(t,y)\,\mathrm{d}y \quad \forall A \subset D, t.$$

The evolution of the probability density starting from an initial density ρ_0 is given by the *Fokker-Planck equation*[2] [58, 118]

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= \mathcal{L}^*\rho \quad \text{for } (x,t) \in D \times (0,\infty) \\ \rho &= \rho_0, \quad \text{for } (x,t) \in D \times \{0\}\,, \end{aligned}$$

where $\mathcal{L}^*$ is the adjoint of the generator $\mathcal{L}$. For the SDE (2.11) it is defined by

$$\mathcal{L}^*\rho = -\nabla \cdot (\rho h) + \frac{1}{2}\nabla \cdot \nabla \cdot \left(\rho\sigma\sigma^T\right).$$

As with the backward Kolmogorov equation this is an n-dimensional PDE. Numerical solutions representing the evolution of the measure in time are therefore typically not feasible. But the equation is useful for finding invariant measure of the process, as these are solutions to $\mathcal{L}^*\rho = 0$. In Section 2.6 the backward Kolmogorov equation is used to construct processes that have a desired measure as their invariant.

2.5 Information and entropy

The Shannon information entropy, or just entropy, is used as a measure for the amount of information in a system when the likelihood of states is given by the density $\rho(\boldsymbol{y})$.

$$S[\rho] = -\int_D \rho(\boldsymbol{y}) \ln \rho(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}$$

[2] or: *forward Kolmogorov equation*

Orderly states are described using less information, whereas the description of a disordered state requires more information. With the assumption that – without other knowledge about the state – it is more likely to find a system in a disorderly state, the most likely state is found as the maximizer of the Shannon entropy. This maximizer is constrained to the normalization of the probability density function

$$\int_D \rho(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} - 1 = 0. \tag{2.16}$$

This constraint is enforced by introducing a Lagrange multiplier λ_0 in the maximization

$$\rho = \arg\max_{\rho} \left(S[\rho] - \lambda_0 \left(\int_D \rho\,\mathrm{d}\boldsymbol{y} - 1 \right) \right). \tag{2.17}$$

The solution to this variational problem is $\ln\rho + \lambda_0 + 1 = 0$, with λ_0 defined by the normalization (2.16). For systems with compact phase space, this results in

$$\rho(\boldsymbol{y}) = \frac{1}{\mathrm{vol}\,\{D\}},$$

when D is unbounded, additional constraints are required.

If one has further knowledge on the state of the system, this is characterized by expectations c_k of observables $C_k(\boldsymbol{y})$ for $k = 1, 2, \ldots, K$:

$$\mathbb{E}\,[C_k] = \int_D C_k(\boldsymbol{y})\rho(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = c_k.$$

Such knowledge on the system state modifies the entropy maximization in (2.17) to include a Lagrange multiplier for each of the observations

$$\rho = \arg\max_{\rho} \left(S[\rho] - \lambda_0 \left(\int_D \rho\,\mathrm{d}\boldsymbol{y} - 1 \right) - \sum_{k=1}^{K} \lambda_k \left(\int_D C_k \rho\,\mathrm{d}\boldsymbol{y} - c_k \right) \right).$$

The maximizer ρ that follows from this has the form

$$\rho(\boldsymbol{y}) = \exp\left(-\lambda_0 - \sum_{k=1}^{K} \lambda_k C_k(\boldsymbol{y}) \right). \tag{2.18}$$

The values for the Lagrange multipliers are typically no longer available in closed form in this setting. Instead, they must be found via more involved methods. One such approach is detailed in Section 3.2.

The well-known *canonical distribution*[3] specifies the likelihood of states for a system with fixed number of particles, fixed volume and fixed temperature. Its probability density function is given by

$$\rho_{\mathrm{c}}(\boldsymbol{q}, \boldsymbol{p}) = Z^{-1} e^{-\beta H(\boldsymbol{q},\boldsymbol{p})}, \tag{2.19}$$

[3] or: *Gibbs distribution*

where Z is the canonical partition function that acts as a normalization. This distribution was not originally proposed through the use of entropy maximization; however, it is in fact the maximal entropy measure for a system at a prescribed mean energy.

In certain situations there is also prior knowledge on the likelihood of states given in the form of a prior distribution $\pi(\boldsymbol{y})$ over the phases space D. In this case, rather than maximizing the Shannon information entropy, constrained by observations and normalization, it is natural to seek a distribution that is in some sense "close" to the prior distribution. One way of assigning a distance between two measures is provided by the Kullback-Leibler divergence

$$S[\rho, \pi] = \int_D \rho(\boldsymbol{y}) \ln \frac{\rho(\boldsymbol{y})}{\pi(\boldsymbol{y})} \, \mathrm{d}\boldsymbol{y}.$$

The definition sometimes includes a minus sign for similarity with the Shannon information entropy, but we do not include it such that the value is more readily interpreted as a distance between two distribtions. It is not a true distance however, because $S[\rho, \pi] \neq S[\pi, \rho]$, but it does satisfy that $S[\rho, \pi] = 0$ if and only if $\rho = \pi$ and that it is positive otherwise.

The relative entropy is used in the same way as the Shannon information entropy, except now the quantity is to be *minimized* given constraints due to observations. With the same observables C_k used for (2.18), the density with minimal distance to the prior is given by

$$\rho(\boldsymbol{y}) = \exp\left(-\tilde{\lambda}_0 - \sum_{k=1}^{K} \tilde{\lambda}_k C_k(\boldsymbol{y})\right) \pi(\boldsymbol{y}).$$

This result differs from that in (2.17) by the inclusion of the prior distribution *and* the values of the Lagrange multipliers. We stress this by adding a tilde to these Lagrange multipliers.

2.6 Sampling

In situations where the probability density of a system is known from some source of observational or theoretical knowledge, it is useful to be able to generate an ensemble of samples according to this distribution. This task is simple for some distributions, such as uniform, Gaussian or exponential distributions, where samples can be generated via transformation of a set of (pseudo-)random numbers with uniform distribution. For more involved distributions, such as (2.19) with arbitrary potential, we require more elaborate sampling techniques.

We discuss two different classes of methods: Markov-chain sampling and sampling with a modified dynamical system. Both methods construct stochas-

tic processes that have the desired measure as an invariant. The former constructs a discrete-time process. The latter is a modification to a known dynamical model for the system. Before discussing these strategies in detail, we introduce the concept of *ergodicity*, which plays an important role in sampling techniques.

2.6.1 Ergodicity

Loosely speaking a dynamical system is ergodic if any trajectory visits all of the phase space. To state this more precisely, let us first define the concept of an *ergodic set*. Given a process X with the phase space D with sigma-algebra $\mathcal{G}$, an ergodic set $A \in \mathcal{G}$ is a set for which it holds that if $X_0 \in A$ then $X_t \in A$ for all $t > 0$.

The union of two ergodic sets A_1, A_2 is also an ergodic set $A_3 = A_1 \cup A_2$. There may also be ergodic sets $A_5 = A_4 \cup B_4$ that are the union of a smaller ergodic set A_4 with a transient set B_4. The transient set B_4 is such that if $X_0 \in B_4$, then for $s \to \infty$, $X_t \in A_4$ for any $t > s$ with probability 1. In other words, a transient set is a set that the process will almost surely leave and never revisit. Let us call any ergodic set that cannot be decomposed in any such manner a *minimal ergodic set*.

A process X is ergodic with respect to the measure $\nu : \mathcal{G} \to \mathbb{R}$ if all minimal ergodic sets $C \in \mathcal{G}$ have either $\nu(C) = 0$ or $\nu(C) = 1$. This implies initial conditions drawn according to ν will, with probability 1, start in the minimal ergodic set with measure 1. The trajectories subsequently visit almost all of the phase space[4].

A consequence of a system being ergodic is that the time it spends in a certain set $A \in \mathcal{G}$ is proportional to the measure $\nu(\mathcal{G})$, to be precise

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} 1_A(X_{n\tau}) = \nu(A), \text{ for almost any } \tau, X_0,$$

where 1_A is the indicator function for A,

$$1_A(x) \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases}$$

From this it follows that time averages of observables $C : D \to \mathbb{R}$ of the stochastic process equal ensemble averages with respect to the measure ν:

$$\bar{C} := \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} C(X_{n\tau}) = \int_D C(x)\, \mathrm{d}\nu(x) := \langle C \rangle . \tag{2.20}$$

This shows that a stochastic process that is ergodic with respect to a measure ν can be used to produce ensemble averages by computing time averages.

[4] All of the phase space with the exception of the minimal ergodic sets with measure 0.

In the case of a continuous time process, the summation in (2.20) is replaced by an integral and we have

$$\bar{C} := \lim_{T\to\infty} \int_0^T C(X_t)\,\mathrm{d}t = \int_D C(x)\,\mathrm{d}\nu(x) := \langle C \rangle .$$

2.6.2 Markov chain sampling

A Markov chain is a Markov process with discrete times, see Section 2.3. The methods we use for constructing Markov chains are all applications of the Metropolis-Hastings algorithm. This algorithm constructs a Markov chain that has the desired target density as its invariant measure [127]. Each step consists of generating a proposal for the next state and a decision to accept or reject this proposal. For sampling a density ρ using a proposal $\mathbb{P}[y|x] = q(x, y)$, this results in Algorithm 1.

Given the current state X_n
Generate proposal $Y \sim q(X_n, y)$.
Generate $U \sim \mathcal{U}(0, 1)$ and deliver

$$X_{n+1} = \begin{cases} Y & \text{if } U \leq \alpha(X_t, Y) \\ X_n & \text{otherwise} \end{cases},$$

where

$$\alpha(x, y) = \min\left(\frac{\rho(y)q(y, x)}{\rho(x)q(x, y)}, 1\right).$$

Algorithm 1: Metropolis-Hastings algorithm

Iteratively applying Algorithm 1 generates a sequence $X_1, X_2, \ldots$ of dependent random variables. Successful implementation of this approach hinges entirely on the choice of a good proposal probability $q(y, x)$.

A particular type of Metropolis-Hastings algorithm is given by a random walk sampler. Here the proposal is given by

$$Y = X_t + Z,$$

where Z is taken from a suitable symmetric distribution. In this case the probability $q(y, x) = q(x, y)$ and the accept-reject step depends only on the ratio of probabilities of the system state, i.e.

$$\alpha(x, y) = \min\left(\frac{\rho(y)}{\rho(x)}, 1\right).$$

We use a random walk sampler in Chapter 3 for sampling point vortex system states on the sphere according to a given density. Here we generate proposals by selecting a number of vortices and moving them over the sphere.

When generating the initial conditions for the point vortex simulations in Chapter 3, we use a similar strategy, but with a modified accept-reject step to approach a system state with prescribed energy and momentum.

2.6.3 Sampling with dynamics

It is common practice in molecular dynamics to perturb the canonical Hamiltonian system such that simulations of the resulting dynamical system or stochastic process (depending on the perturbation) produce samples according to a desired distribution. In molecular dynamics the goal of these perturbations is usually to maintain the temperature by sampling the canonical distribution of (2.19). Because of this application, these perturbations are dubbed thermostats. We choose to refer to them as thermostats throughout this thesis for want of a better name, even if their function is no longer to maintain a given temperature but to enforce some other desired distribution.

The Langevin thermostat

In molecular dynamics the systems of interest are typically canonical Hamiltonian systems. The Hamiltonian is separable into a kinetic term given by $K(p) = p^T M^{-1} p$, with M the mass matrix, and some potential $V(q)$. A Langevin thermostat adds a stochastic perturbation to the equation for the momenta as follows

$$\begin{aligned} \mathrm{d}\boldsymbol{q} &= M^{-1}\boldsymbol{p}\,\mathrm{d}t \\ \mathrm{d}\boldsymbol{p} &= -V(\boldsymbol{q})\,\mathrm{d}t - \gamma M^{-1}\boldsymbol{p}\,\mathrm{d}t + \sqrt{\frac{2\gamma}{\beta}}\,\mathrm{d}W, \end{aligned}$$

where W is a vector of n independent Wiener processes. The result is a dynamical system that has the canonical distribution (2.19) as its unique invariant measure.

Langevin thermostats are easy to implement and relatively robust in the sense that their ergodicity is not subject to a careful choice of parameters. However, the perturbation to the dynamics is substantial. Even in the limit $\gamma \to 0$, autocorrelation functions of the system are not recovered [54]. Furhtermore, the perturbation may disrupt structural properties of the original system that are ideally maintained by the dynamical sampler.

2.6.4 Gentle thermostats

An alternative approach first introduced by Nosé [113, 114] and later improved by Hoover [70] involves adding a separate thermostat variable ξ. The phase space D is then embedded in the larger $D^* = D \times \mathbb{R}$ and the canonical

measure ρ_c is augmented to

$$\rho_c^*(\boldsymbol{q}, \boldsymbol{p}, \xi) = \rho_c(\boldsymbol{q}, \boldsymbol{p}) e^{-\xi^2}. \tag{2.21}$$

The thermostat variable thus has a unit normal distribution, independent of the system state. Extended systems with this invariant measure will result in the desired marginal distribution for the variables $(\boldsymbol{q}, \boldsymbol{p})$.

The Nosé-Hoover thermostat is constructed to have the augmented distribution (2.21) as an invariant measure. The equations of motion read

$$\begin{aligned}
\mathrm{d}\boldsymbol{q} &= M^{-1}\boldsymbol{p}\,\mathrm{d}t \\
\mathrm{d}\boldsymbol{p} &= -V(\boldsymbol{q})\,\mathrm{d}t - \varepsilon\xi\boldsymbol{p}\,\mathrm{d}t \\
\mathrm{d}\xi &= \varepsilon\left(\beta\boldsymbol{p}^T M^{-1}\boldsymbol{p} - n\right)\,\mathrm{d}t.
\end{aligned}$$

While this system has the extended measure ρ_c^* as an invariant measure, this is not a unique invariant measure. This lack of ergodicity is resolved by including a stochastic perturbation in the equation of motion of the thermostat variable ξ [87]. This results in the Nosé-Hoover-Langevin thermostat

$$\begin{aligned}
\mathrm{d}\boldsymbol{q} &= M^{-1}\boldsymbol{p}\,\mathrm{d}t \\
\mathrm{d}\boldsymbol{p} &= -V(\boldsymbol{q})\,\mathrm{d}t - \varepsilon\xi\boldsymbol{p}\,\mathrm{d}t \\
\mathrm{d}\xi &= \varepsilon\left(\beta\boldsymbol{p}^T M^{-1}\boldsymbol{p} - n\right)\,\mathrm{d}t - \gamma\,\mathrm{d}t + \sqrt{2\gamma}\,\mathrm{d}W,
\end{aligned}$$

where W is a one-dimensional Wiener process. The "gentleness" of these thermostats stems from the fact that the stochastic forcing required for ergodicity is added to auxiliary variables only. As such, the dynamics of the system state $(\boldsymbol{q}, \boldsymbol{p})$ remain close to the original dynamics [54].

This method is generalized to a wider class of systems by the Generalized Bulgac-Kusnezov (GBK) method [44]. For an arbitrary system with divergence-free f we write

$$\begin{aligned}
\mathrm{d}\boldsymbol{y} &= f(\boldsymbol{y})\,\mathrm{d}t + \xi g(\boldsymbol{y})\,\mathrm{d}t \\
\mathrm{d}\xi &= \nabla\cdot g(\boldsymbol{y}) - g(\boldsymbol{y})\nabla A(\boldsymbol{y}) - \gamma\xi + \sqrt{2\gamma}\,\mathrm{d}W.
\end{aligned}$$

This will sample the extended measure $\rho^* \sim e^{-A(y)-\frac{1}{2}\xi^2}$ for any $A(\boldsymbol{y})$ that is a first integral of f, i.e. $\nabla A \cdot f = 0$. Note that such functions can be constructed by taking $A(\boldsymbol{y}) = A\left(H(\boldsymbol{y}), I_1(\boldsymbol{y}), I_2(\boldsymbol{y}), \ldots\right)$, such that it is only a function of $\boldsymbol{y}$ through known first integrals of f, as mentioned in Section 2.1.1. The system may also be extended with multiple thermostats ξ_m, for $m \in \{1, \ldots, M\}$, each with its own perturbation field g_m. In this case, the thermostated system of SDEs reads

$$\begin{aligned}
\mathrm{d}\boldsymbol{y} &= f(\boldsymbol{y})\,\mathrm{d}t + \sum_{m=1}^{M} \xi_m g_m(\boldsymbol{y})\,\mathrm{d}t \\
\mathrm{d}\xi_m &= \nabla\cdot g_m(\boldsymbol{y}) - g_m(\boldsymbol{y})\nabla A(\boldsymbol{y}) - \gamma_m\xi_m + \sqrt{2\gamma_m}\,\mathrm{d}W_m.
\end{aligned}$$

Ergodicity of this sampling scheme depends on the choice for the perturbation fields g_m. We shall construct stochastic processes such that a given measure is an invariant measure to the Fokker-Planck equation associated to the thermostated system. If this measure is strictly positive everywhere, then ergodicity is proven by hypoellipticity of the operator $\mathcal{L}^*$ [63].

Hypoellipticity of the Fokker-Planck equation follows from Hörmanders theorem. For a GBK thermostat applied to a system with phase space $\mathbb{R}^d$, this theorem holds true under the following modified Hörmander condition [87, 8]:

$$\mathbb{R}^d \in \operatorname{span} \mathcal{L}(f, g_1, \ldots, g_M).$$

Here $\mathcal{L}(g_0, g_1, \ldots, g_M)$ denotes the ideal of the vector fields g_m with $m > 0$ within the Lie-algebra generated by all of the g_m, reading

$$\mathcal{L}(g_0, g_1, \ldots, g_M) = \{g_{m_0}, [g_{m_0}, g_{m_1}]\, [g_{m_0}, [g_{m_1}, g_{m_2}]], \ldots\},$$

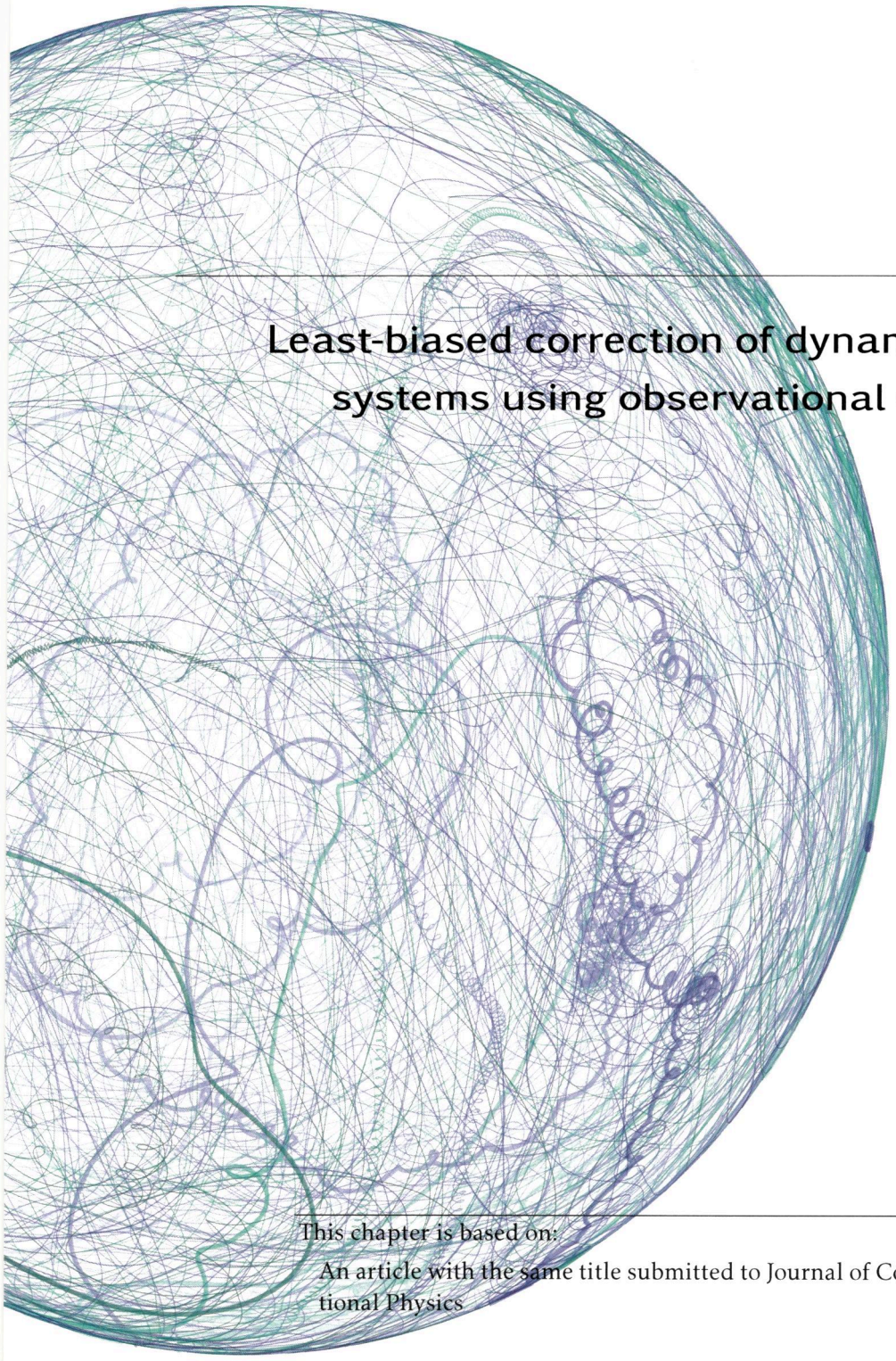
where m_0 takes values in $\{1, \ldots, M\}$ and the other m_k take values in $\{0, \ldots, M\}$. The bracket $[\cdot, \cdot]$ denotes the commutator between two vector fields.

A particularly interesting choice for $g(\boldsymbol{y})$ is to use $g(\boldsymbol{y}) = J^2(\boldsymbol{y})\nabla H(\boldsymbol{y})$. This choice ensures that Casimirs of the original system are not affected by the thermostats perturbation. This "double-bracket thermostat" is used in Chapter 3, where the Casimirs are crucial to the system of interest.

2.7 Fluid dynamics

Much of this thesis deals with the application of thermostats in fluid dynamics settings. As the fluid models are used as is, we do not explain them further here. Instead, we refer the interested reader to a number of monographs regarding specific aspects of fluid dynamics.

A good introduction to aerodynamics in general is provided by Anderson [5] and Chorin et al. [29]. Turbulence in both two and three dimensions is discussed by Frisch [56], who presents the seminal work of Kolmogorov from the forties in a modern way, and also by Holmes et al. [69] and Davidson [34]. Trefethen [138], Canuto et al. [23] and Boyd [17] detail the spectral discretization of fluid flow. Wavelets present an alternative to spectral methods, that retain local information too [101]. Geophysical fluid flows are discussed extensively by Pedlosky [119] and Majda & Wang [95]. The Hamiltonian structure of (two-dimensional) fluid flow is derived by Salmon [131] and Swaters [136].



3

Least-biased correction of dynamical systems using observational data

This chapter is based on:

An article with the same title submitted to Journal of Computational Physics

background graphic: trajectories of strong (thick lines) and weak vortices of positive (purple) and negative (green) circulations

3.1 Introduction

In many applications of modern computational science the physical laws (and equations of motion) are well established yet the detailed behavior is unpredictable on long time scales due to the presence of deterministic chaos. Examples of this arise in molecular dynamics modelling [4, 133] and in the study of turbulent fluids in the atmosphere and ocean [69, 34]. For these problems, long simulations are routinely run, despite the lack of predictability, in the hope that the resulting simulation will yield useful statistical knowledge (e.g. the statistics of rare transitions between basins in molecular dynamics, or slow relaxation processes in fluids). We refer to this approach as *dynamical sampling*, where the name is suggestive of the typical requirement that simulated paths are sufficiently accurate to allow the computation of measures of dynamical mixing such as two-point temporal correlation functions [66].

In Hamiltonian systems such as molecular dynamics, it is common to run canonically prepared ensembles of microcanonical (i.e. constant energy) simulations in order to minimize the perturbation of dynamical properties. For such systems, backward error analysis [62, 86] suggests that the global behavior can best be understood not as the approximation of particular trajectory but rather as an accurate path for a perturbed continuum process described by modified equations. In the case of dynamic sampling of complex systems, the statistics of simulation data are therefore biased in that they sample an invariant measure of the modified equations, i.e. bias arises as an artifact of time discretization. Statistical bias may also arise due to spatial discretization. For example, in the setting of geophysical fluid dynamics, a comparison of discretizations of the quasi-geostrophic equations reveals that the long time mean potential vorticity field and pointwise fluctuation statistics are heavily dependent on discrete conservation laws such as energy, enstrophy, and material conservation of vorticity [2, 42, 43]. It is usually impossible to construct numerical discretizations that automatically preserve all conservation laws of statistical relevance for a given problem, so the discretization necessarily perturbs the statistical distribution. The discretization bias may be reduced by refining the discretization or by incorporating a Metropolis condition [100], but such techniques also typically lead to an increase in computational overhead, which may be unacceptable in large scale applications.

The combination of the need for computations to address both the stationary constraint ("nearness to the steady-state distribution") and to provide accuracy with respect to dynamical processes poses difficult challenges for the simulator. In this paper we consider an approach to perturbing dynamics to correct statistical bias in systems at statistical equilibrium. If the statistical distribution is completely specified via a probability density function (pdf) it can be sampled using a "thermostat." Such thermostats, originating in molecular dynamics, can be extended to handle both smooth [11] and non-

smooth [10] densities and to treat noncanonical Hamiltonian systems. In [44], a thermostat was used as a model reduction technique for a vortex model of a fluid (suppressing the detailed interactions of a few strong vortices with a weak vortex field). In another recent article [39], thermostats have been suggested as a means of sampling incompletely specified systems (with noisy gradients), with applications in learning theory. The standard framework of thermostating used in these and other applications assumes a fixed, known distribution such as the Gibbs-Boltzmann distribution. In this article, we assume that, instead of the pdf, what is available is a partial set of *expectations* of observables with respect to the unknown invariant measure, which may arise from experiment or other types of modelling. In this setting, information theory (in particular entropy maximization [73, 74]) offers tools for constructing least-biased densities, close to some known prior distribution, which are consistent with observations. The iterative method (based on [3, 64, 35]) involves computation of Lagrange multipliers (one for each observable) that modify the probability density. The Lagrange multipliers are computed using an iterative procedure in which each stage represents an ensemble average (with respect to the previous estimate of the density). To make the method practical in situations where the sampling is costly, we consider an adaptive procedure which uses only short-time ensemble bursts to gradually tune the parameters in simulation. At the same time, we are able to show in numerical experiments that autocorrelation functions are only modestly perturbed meaning that we would expect to be able to recover dynamical information such as diffusion and other transport coefficients.

We emphasize that the framework of least-biased estimation is well known but applied here in a novel way. A related technique is used by Majda and Gershgorin [94] to develop a framework for validating computational models and choosing the optimal linear combination of an ensemble of model outputs, so as to minimize the discrepancy between the ensemble distribution and the least-biased estimate, arguing that the latter is the best available measure for comparison, when the true invariant measure is unknown. With the approach we develop here, we enforce exact adherence to the least-biased measure, which is constructed automatically in simulation, by perturbing the dynamics to take full advantage of available information.

The remainder of this article is organized as follows. In the next section, we discuss the maximum entropy framework for correcting the density to reflect thermodynamic constraints. We apply and evaluate the method in the setting of a system of point vortices on the surface of a sphere, which represents a simple geophysical model with multiple statistically relevant first integrals.

3.2 Bias correction method

Our interest is in extended dynamical systems with many degrees of freedom that evolve near statistical equilibrium. Further, we imagine that we are given a simplified dynamical model for the evolution of some projection (i.e. a "coarse graining") of the phase variables (coarse grained variables $y(t) \in \mathbb{R}^d$) . Although the original system is complex and its details unknown, we assume that we can obtain in some way (e.g. through measurement) a collection of "observations" of mean values of functions of the reduced variables. That is there are functions $C_k : \mathbb{R}^d \to \mathbb{R}$, $k = 1, 2 \ldots, K$ and given values c_k, $k = 1, 2, \ldots, K$, such that

$$c_k = \langle C_k(y) \rangle, \quad k = 1, \ldots, K, \tag{3.1}$$

where $\langle C_k(y) \rangle$ represents averaging with respect to the true, empirical invariant measure of the dynamical system. Our goal is to find a perturbed dynamical model for the reduced variables which (a) is compatible with the indicated thermodynamic constraints (3.1), and (b) weakly perturbs the dynamics compared to those of the native model.

Empirical information theory generalizes the principle of insufficient reason, by proposing the least-biased probability density consistent with a set of observations. See the classical work of Jaynes [73, 74], the monographs [64, 38] and an extensive treatment in the geophysical fluid context in the monograph by Majda and Wang [95]. The least-biased density is defined as the probability density $\rho(y)$ that maximizes the information entropy functional

$$\mathcal{S}[\rho] = -\int_{\mathcal{D}} \rho(y) \log \rho(y)\, dy,$$

subject to a set of constraints given by observations. When $\mathcal{D}$ is a compact set and there are no observations, the minimizer is the uniform density $\rho \equiv |\mathcal{D}|^{-1}$. The entropy $\mathcal{S}$ is the unique measure of uncertainty that is positive valued, monotonically increasing as a function of uncertainty, and additive for independent random variables. With observable functions $\{C_k(y) | k = 1, 2, \ldots K\}$ let

$$\mathbb{E}_\rho C_k = \int_{\mathcal{D}} C_k(y) \rho(y)\, dy \tag{3.2}$$

denote expectation in the (as yet undetermined) density ρ. Defining Lagrange multipliers λ_k, $k = 1, \ldots, K$, associated with the observables C_k, the constrained minimization problem is

$$\rho = \arg\max_{\hat{\rho}} \left[\mathcal{S}[\hat{\rho}] - \sum_{k=1}^{K} \lambda_k \left(\mathbb{E}_{\hat{\rho}} C_k(y) - c_k \right) \right].$$

When it exists, the maximum entropy solution satisfies

$$\rho(y) = \lambda_0 \exp\left(-\lambda_1 C_1(y) - \cdots - \lambda_K C_K(y)\right),$$

where λ_0 is chosen to satisfy $\int_{\mathcal{D}} \rho \, dy = 1$, and λ_k is chosen such that $\mathbb{E}_\rho C_k(y) = c_k$.

In some cases, besides the observations, we may be given prior statistical information on the process $y(t)$. The Kullback-Leibler divergence, or relative entropy,

$$\mathcal{S}[\rho(y)] = \int \rho(y) \ln \frac{\rho(y)}{\pi(y)} \, dy$$

which represents a (non-symmetric) distance between measures. It quantifies the information lost in approximating $\rho(y)$ by $\pi(y)$.

Suppose y is a random variable with distribution (law) $y \sim \rho$, where ρ is unknown. Suppose further, that we are given a prior distribution π, presumed to be close to ρ, and a set of K observations (3.13). Following Jaynes [73, 74], the least-biased distribution ρ consistent with the observations c_k and prior π solves the constrained minimization problem

$$\rho = \arg\min_{\rho} \left[\mathcal{S} - \lambda_0 \left(1 - \int \rho(y) \, dy\right) - \sum_{k=0}^{K} \lambda_k \left(c_k - \int C_k(y) \rho(y) \, dy\right) \right],$$

where the λ_k are Lagrange multipliers to enforce the condition that the expectations (3.2) agree with the observations (3.13). The solution to the variational problem is

$$\rho(y) = \lambda_0 \exp\left(-\lambda_1 C_1(y) - \cdots - \lambda_K C_K(y)\right) \pi(y), \tag{3.3}$$

where the Lagrange multipliers λ_k are chosen consistently with the observations (3.13) and λ_0 is a normalization constant so that ρ is a probability density function.

Methods for determining the Lagrange multipliers are discussed in [3, 64, 35]. We use the following algorithm based on re-weighting. Assume we are given a sequence of samples y^n, $n = 1, \ldots, N$, distributed according to a known prior distribution $\pi(y)$, i.e. $y^n \sim \pi$. The expectation under $\pi(y)$ of a function $\Phi(y)$ has the consistent and unbiased estimator

$$\widehat{\Phi}^\pi = \frac{1}{N} \sum_{n=1}^{N} \Phi(y^n).$$

Given the posterior distribution $\rho(y)$ of the form (3.3), compute the expectation $\mathbb{E}_\rho \Phi$ by re-weighting of the integral

$$\begin{aligned} \mathbb{E}_\rho \Phi &= \int \Phi(y) \rho(y) \, \mathrm{d}y = \lambda_0 \int \Phi(y) e^{-\sum_{i=1}^{K} \lambda_i C_i(y)} \pi(y) \, \mathrm{d}y \\ &= \lambda_0 \mathbb{E}_\pi \{ \Phi(y) \lambda_0 e^{-\sum_{i=1}^{K} \lambda_i C_i(y)} \}, \end{aligned}$$

yielding an unbiased estimator for $\mathbb{E}_\rho \Phi$ given by

$$\widehat{\Phi}^\rho = \frac{\lambda_0}{N} \sum_{n=1}^{N} \Phi(y^n) e^{-\sum_{i=1}^{K} \lambda_i C_i(y^n)}.$$

We wish to ensure that the observations c_k satisfy

$$c_k = \widehat{C}_k^\rho = \frac{\lambda_0}{N} \sum_{n=1}^{N} C_k(y^n) e^{-\sum_{i=1}^{K} \lambda_i C_i(y^n)}, \quad k = 1, \ldots, K.$$

We can use this fact to define a Newton-Raphson iteration to determine the Lagrange multipliers λ_k. Define the residual r with components

$$r_k(\lambda) = c_k - \frac{\lambda_0}{N} \sum_{n=1}^{N} C_k(y^n) e^{-\sum_{i=1}^{K} \lambda_i C_i(y^n)}, \quad k = 1, \ldots, K,$$

with $\lambda = (\lambda_1, \ldots, \lambda_K)$ and $r = (r_1(\lambda), \ldots, r_K(\lambda))$. Note that λ_0 can be viewed as a function of $\lambda_1, \lambda_2, \ldots, \lambda_K$ chosen from the normalization condition, i.e.,

$$\lambda_0 = \left[\sum_{n=1}^{N} e^{-\sum_{j=1}^{K} \lambda_j C_j(y^n)} \right]^{-1}.$$

The Jacobian matrix $J = (J_{kj})$ of the vector function r is determined as

$$J_{kj}(\lambda) := \frac{\partial r_k}{\partial \lambda_j} = \frac{\lambda_0}{N} \sum_{n=1}^{N} C_k(y^n) C_j(y^n) e^{-\sum_{i=1}^{K} \lambda_i C_i(y^n)} \quad j, k = 1, \ldots, K.$$

The iteration then proceeds as $\lambda^{\alpha+1} \leftarrow \lambda^\alpha - J^{-1}(\lambda^\alpha) r(\lambda^\alpha)$.

3.2.1 Adaptive determination of Lagrange multipliers

In many cases it will be difficult or costly to carry out a complete sampling of the distribution at each iteration step of the Newton procedure. Moreover, the standard framework excludes applications where (i) the statistical knowledge is expected to improve as the simulation progresses, (ii) the average observables are known to vary slowly with time, or (iii) it is unfeasible to constuct a large enough ensemble distributed in the prior. For these cases we consider using the simulation data of a small ensemble (propagated in short bursts of M timesteps) for updating the Lagrange multipliers for mean observation data. This results in an adaptive algorithm for obtaining the Lagrange multipliers "on-the-fly" during simulation.

Consider the following: an ensemble of P simulations (preferably with initial conditions distributed close to $\pi(y)$) is advected $M\Delta t$ in time, where

M is chosen sufficiently large such that the ensemble members sample π well. These ensemble members can be used in an estimator for $\mathbb{E}_{\lambda^1} C_k$ given by

$$\widehat{C}_k^{(1)} = \frac{1}{P}\sum_{p=1}^{P} C_k(y_0^p)\lambda_0^1 \exp\left(-\sum_{i=1}^{K}\lambda_i^1 C_i(y_0^p)\right), \quad k = 1,\dots,K,$$

where the superscript (1) indicates that it is an estimator for a distribution with Lagrange multipliers λ_j^1. A Newton-Raphson iteration to find the first set of Lagrange multipliers such that observations match data has the residual

$$r_k^1 = \widehat{C}_k^{(1)} - c_k = \frac{1}{P}\sum_{p=1}^{P} C_k(y_0^p)\lambda_0^1 \exp\left(-\sum_{i=1}^{K}\lambda_i^1 C_i(y_0^p)\right) - c_k, \quad k = 1,\dots,K.$$

Using this updated value for λ the simulations will sample the distribution $\rho \propto e^{-\sum_{i=1}^{K}\lambda_i^1 c_i(y)}$ after some time $M\Delta t$. (See below for some practical issues associated to this.) Using these samples alongside the initial data, λ_j^2 is found. Iteration of this process leads to the following equation

$$\begin{aligned} r_k^m &= \widehat{C}_k^{(m)} - c_k \\ &= \frac{1}{mP}\sum_{p=1}^{P}\sum_{\ell=0}^{m-1} C_k(y_{\ell M}^p)\lambda_0^\ell \exp\left(\sum_{i=1}^{K}(\lambda_i^l - \lambda_i^m)C_j(y_{\ell M}^p)\right) - c_k, \end{aligned} \tag{3.4}$$

where we remind that, at each stage of iteration, λ_0^l is a function of the multiplier vector λ^l (indices $1\dots K$). In the calculation (3.4) $\lambda_1^0,\dots\lambda_K^0$ would ideally be zero. There are cases where it is impossible to obtain an accurate initial distribution according to the prior, in which case the initial Lagrange multipliers can be chosen different from zero if initial conditions sampling $\pi\lambda_0^0 \exp\left(-\sum_k \lambda_k^0 C_k\right)$ are easier to find than those sampling just the prior π. Solutions of (3.4) are found using Newton-Raphson iteration. The gradient is given by

$$\frac{\partial r_k^m}{\partial \lambda_j^m} = \frac{1}{mP}\sum_{p=1}^{P}\sum_{\ell=0}^{m-1} C_k(y_{\ell M}^p)C_j(y_{\ell M}^p)\lambda_0^\ell \exp\left(\sum_{i=1}^{K}(\lambda_i^l - \lambda_i^m)C_i(y_{\ell M}^p)\right), \tag{3.5}$$

for all $j,k = 1,\dots,K$. In this way the Lagrange multipliers may be found "on-the-fly."

As a convergence result let us consider the case where both P and M may be chose arbitrarily large. For $P \to \infty$ the Lagrange multipliers computing using only the initial data sampling the prior will be correct. Given sufficiently large M the samples after evolving the thermsotated system $M\delta t$ in time will accurately sample the distribution corresponding to these Lagrange multipliers. The ensemble averages will then correspond to the observations, and the Lagrange multipliers no longer need updating.

Adaptive algorithm

There are two important practical modifications to the algorithm described above that are included in the numerical implementation of this method:

- The first modification is limiting the rate of change of the Lagrange multipliers. If the Lagrange multipliers change rapidly, the thermostat may require a long time to equilibriate. This requires a larger value for M, increasing the simulation time required before including new samples. The effect is especially noticable at the beginning of a simulation, due to two factors: (i) the small sample size leads to inaccurate expectations for the observables, and (ii) the Lagrange multipliers may be far from their correct value. By limiting the rate of change of the Lagrange multipliers, these problems are circumvented.
- The second modification regards the number of samples included when updating the Lagrange multipliers. In equations (3.4) and (3.5) *all* previous values λ_k^ℓ are included. In a long simulation, this leads to a growing computational demand. By taking only a fixed number (q) of recent steps the computational demand can be reduced. In the case that the initial samples cannot accurately be drawn from the prior, this has the further advantage that these inaccuracies are eventually forgotten.

The algorithm, including these practical modifications, is summarized in Algorithm 2.

Given initial conditions according to prior $\pi(y)$
Set initial Lagrange multipliers to zero
for $m \leftarrow 1$ **to** n **do**
  for $j \leftarrow 1$ **to** M **do**
    advance simulation one time step using current value for the Lagrange multipliers
  end
  store relevant simulation observables for time step mM.
  while $|\langle C_k(y)\rangle_\lambda - c_k| >$ tolerance **do**
    compute residual using re-weighted samples at times $M \times max(m-q,0), \ldots, mM$
    compute residual gradient using the same data
    update Lagrange multiplier estimation
  end
  limit the change in Lagrange muliplier (if necessary)
end

Algorithm 2: Adaptive determination of Lagrange multipliers "on-the-fly"

3.2.2 Thermostat

We introduce the bias-correction methodology for a Hamiltonian dynamical system

$$\frac{dy}{dt} = f(y) = B(y)\nabla H(y), \qquad y(t) \in \mathcal{D},\ B(y) = -B(y)^T,\ H(y) : \mathcal{D} \to \mathbb{R}, \tag{3.6}$$

possessing a divergence-free vector field[1] $\nabla \cdot f \equiv 0$. Invariance of the Hamiltonian H along solutions of (3.6) follows from $\frac{d}{dt}H(y(t))) = \nabla H \cdot \frac{dy}{dt} = \nabla H \cdot B\nabla H = 0$, due to skew-symmetry of $B(y)$. Additional first integrals are often present: $I_\ell(y) : \nabla I_\ell \cdot f \equiv 0$, $\ell = 1, \ldots, L$. In this paper we consider only the case where all observables of the physical process $\mathcal{Y}(t)$ correspond to functions of the conserved quantities $\{H, I_\ell, \ell = 1, \ldots, L\}$, that is, $C_k(y) = C_k(H(y), I_1(y), \ldots, I_L(y))$, $k = 1, \ldots, K$.

Thermostats are used in molecular dynamics to model the trajectories of molecules in a fluid at constant temperature. From statistical mechanics, it is well known that the trajectories of a system of particles in thermal equilibrium with a reservoir at constant temperature sample the canonical or Gibbs distribution, which has global support. The governing equations are Hamiltonian, however, implying that the trajectories are restricted to a level set of Lebesgue measure zero. Hence, to model a system at constant temperature, it is necessary to perturb the vector field to make trajectories ergodic with respect to the Gibbs distribution. The most common way of achieving this is by adding suitable stochastic and dissipative terms satisfying a fluctuation-dissipation relation (Langevin dynamics). An advantage of Langevin dynamics is provable ergodicity with respect to the Gibbs distribution [97]. However, DelSole [37] warns that direct stochastic forcing of trajectories leads to inaccurate dynamical quantities since autocorrelation functions are strongly perturbed. For smooth deterministic Hamiltonian dynamics, normalized velocity autocorrelation functions are of the form $1 - c\tau^2$, $c > 0$ in the zero-lag limit $\tau \to 0$, whereas the autocorrelation of a variable that is directly forced by white noise must take the form $\exp(-\kappa\tau)$, $\kappa > 0$ in the same limit. This implies that the direct stochastic perturbation leads to auto-correlation functions that have nonzero slope and opposite curvature at zero lag.

An alternative approach, pioneered by Nosé [113, 114] and Hoover [70] proceeds to augment the phase space by one dimension through coupling of (3.6) to an additional thermostat variable $\xi(t)$. The dynamics of ξ are constructed to ensure that the extended dynamics on $\mathbb{R}^{d+1}$ preserves an equilibrium density whose marginal on $\mathbb{R}^d$ is the target (e.g. Gibbs) density. The fully

[1] The latter condition is automatic for systems (3.6) with constant B. Strictly speaking, the approach described here is applicable to any system with divergence-free vector field $\nabla \cdot f \equiv 0$ possessing one or more first integrals.

deterministic thermostats of Nosé and Hoover have no mechanism to guarantee ergodicity with respect to the target density, and hence have been modified by various authors who include stochastic forcing of the thermostat variable ξ, leading to the so-called Nosé-Hoover-Langevin method [132, 87, 88]. A generalization to generic Hamiltonian systems is the Generalized Bulgac-Kusnezov (GBK) [85, 22] thermostat:

$$\begin{aligned} dy &= f(y)dt + \xi\varepsilon g(y)dt \\ d\xi &= \varepsilon h(y)dt - \gamma\xi dt + \sqrt{2\gamma}dw, \end{aligned} \tag{3.7}$$

where $\varepsilon > 0$ and $\gamma > 0$ are parameters, $w(t)$ is a scalar Wiener process, and g and h are discussed below. Given a target density $\rho(y) \propto \exp(-A(y))$, $A : \mathcal{D} \to \mathbb{R}$, denote the augmented product density by $\tilde{\rho}(y, \xi) = \rho(y) \cdot \mu(\xi)$, with μ a univariate normal distribution with mean zero and standard deviation one. It is easily checked that $\tilde{\rho}$ is stationary under the Fokker-Planck operator associated with (3.7) provided

$$h(y) = \nabla \cdot g - g \cdot \nabla A.$$

Furthermore it is argued in [10] that the target measure is ergodic provided the vector fields f and g satisfy a Hörmander condition. In some cases it is also desirable to use the freedom in choosing g to ensure preservation of some first integrals of the vector field f. We will see an example of this later in this paper.

The parameter ε can be used to control the relative strength of the thermostat compared to that of the unperturbed vector field f. This will affect the rate at which the invariant measure is sampled, but has no influence on the measure itself. It has been proved in [88] and observed numerically in [9, 10] that GBK/NHL thermostating leads to a weak perturbation of the original trajectories in the sense that autocorrelation functions preserve the leading terms, i.e. have the form $1 - c\tau^2 + O(\tau^3)$, as $\tau \to 0$. The GBK thermostat is applicable when the vector field f is divergence free $\nabla \cdot f \equiv 0$ and when the target density ρ is a function of first integrals of f.

We therefore propose (1) constructing a least-biased information theoretic target density based on observations of functions of conserved quantities (with or without prior distribution), followed by (2) thermostated perturbation of dynamics to ensure sampling of the target distribution with a GBK thermostat. The thermostating method is incorporated into Algorithm 1 to provide the scheme for sampling the adapted, data-dependent distribution.

3.3 Application to reduced modelling of point vortices

In this section we apply the least-biased correction methodology to the simple model of point vortices on the sphere. We choose this model because it

has a Poisson structure and multiple conserved quantities, including total energy and angular momentum and a set of Casimirs, with various degrees of statistical significance. Although one can construct a point vortex approximation of quasi-geostrophic potential vorticity dynamics, we ignore the effects of topography and finite deformation radius for computational simplicity.

3.3.1 Point vortex system

A simple conceptual model of the atmosphere is given by the quasigeostrophic potential vorticity equation on a rotating sphere:

$$\frac{Dq}{Dt} \equiv \frac{\partial q}{\partial t} + u \cdot \operatorname{grad} q = 0, \quad \operatorname{div} u = 0, \quad q = \operatorname{curl} u + f_0 + h, \tag{3.8}$$

where q denotes the potential vorticity, u is the velocity field in the tangent plane, assumed divergence-free, $f_0 = 2\Omega \sin\theta$ is the local Coriolis force, and h is the surface topography.

The vortex approximation of (3.8) is well known. For algorithms and analysis of the dynamics of point vortex systems, see the books [109, 93]. For advanced modelling and convergence analysis in the continuum limit, see [32]. For numerical computation with point vortices, it is advantageous to embed the sphere in $\mathbb{R}^3$. In the sequel we will denote vectors in $\mathbb{R}^3$ by bold type. For simplicity we neglect topography, taking $h \equiv 0$, under which assumption the quasigeostrophic model is equivalent to the 2D Euler equations. We may then also ignore rotation (i.e. $f_0 \equiv 0$) as it gives rise to a trivial rigid body rotation of the ensuing point vortex system [109]. Since the velocity field is divergence-free in the tangent plane, it can be represented in terms of a stream function ψ as

$$\boldsymbol{u} = \hat{\boldsymbol{k}} \times \nabla\psi$$

where $\hat{\boldsymbol{k}}$ is the unit normal vector on the surface of the sphere. The potential vorticity and stream function are related by $\Delta\psi = q - f_0 - h$ with Δ the Laplace-Beltrami operator (from which it is apparent that topography, if included, would lead to a nonhomogeneous background term in the stream function).

A point vortex system is constructed by taking the vorticity field in (3.8) to be a sum of Dirac distributions

$$q(x,t) = \sum_{i=1}^{M} \Gamma_i \delta(\boldsymbol{x} - \boldsymbol{x}_i),$$

where Γ_i is the vortex strength or circulation of the ith point vortex. The point vortices induce a stream function $\psi(\boldsymbol{x}) = \sum_i \frac{-1}{4\pi}\Gamma_i \ln(2 - 2\boldsymbol{x} \cdot \boldsymbol{x}_i(t))$ as a sum of Green's functions of the Laplacian. The unit normal on the sphere is given

by $\hat{\boldsymbol{k}} = \boldsymbol{x}/|\boldsymbol{x}|$. Because vorticity is materially conserved in the velocity field, the motion of point vortices is given by $\dot{\boldsymbol{x}}_i = \boldsymbol{u}(\boldsymbol{x}_i)$, i.e.,

$$\dot{\boldsymbol{x}}_i = \boldsymbol{x}_i \times \nabla\psi(\boldsymbol{x}_i) \quad i = 1, 2, \ldots, M,$$

where a unit sphere will assumed. The equations of motion may also be written as a Hamiltonian system with Lie-Poisson structure

$$\Gamma_i \dot{\boldsymbol{x}}_i = \boldsymbol{x}_i \times \nabla_{\boldsymbol{x}_i} H \quad i = 1, 2, \ldots, M, \tag{3.9}$$

where the Hamiltonian, defined by $H = \int_{\mathcal{D}} \frac{1}{2}|\boldsymbol{u}|^2 d\boldsymbol{x}$, is given by

$$H = -\sum_{i=1}^{M}\sum_{j=1}^{i-1} \frac{\Gamma_i \Gamma_j}{4\pi} \ln\left(2 - 2\boldsymbol{x}_i \cdot \boldsymbol{x}_j\right).$$

By introducing $y = \left(\boldsymbol{x}_1^T, \boldsymbol{x}_2^T, \ldots, \boldsymbol{x}_M^T\right)^T$, equation (3.9) can be written in the more compact form (3.6) with the block-diagonal structure matrix

$$B(y) = \begin{bmatrix} \Gamma_1^{-1}\widehat{\boldsymbol{x}}_1 & & & \\ & \Gamma_2^{-1}\widehat{\boldsymbol{x}}_2 & & \\ & & \ddots & \\ & & & \Gamma_M^{-1}\widehat{\boldsymbol{x}}_M \end{bmatrix},$$

where $\widehat{\boldsymbol{x}}_i$ denotes the 3×3 skew-matrix satisfying $\widehat{\boldsymbol{x}}_i \boldsymbol{a} := \boldsymbol{x}_i \times \boldsymbol{a}$, for all $\boldsymbol{a} \in \mathbb{R}^3$. The Poisson bracket for the system is given equivalently by

$$\{F, G\} = \sum_{i=1}^{M} \frac{1}{\Gamma_i} \nabla_{\boldsymbol{x}_i} F \cdot (\boldsymbol{x}_i \times \nabla_{\boldsymbol{x}_i} G) \quad \text{or} \quad \{F, G\} = \nabla F(y)^T B(y) \nabla G(y).$$

This Poisson structure is a generalization of the rigid body Poisson structure and also occurs in ferromagnetic spin lattices [49, 53, 55] and elastic rods (e.g. [78]).

The vortex positions are defined in Cartesian coordinates, but initial positions $\boldsymbol{x}_i(0)$ are chosen on the sphere. Because each $|\boldsymbol{x}_i|$ is a Casimir of the Poisson bracket it is ensured that the vortices remain on the sphere. This restricts the effective phase space of the system to the direct product of M spheres S^2. Furthermore, the rotational symmetry of the sphere gives rise to three Noether momenta, which are expressed by the angular momentum vector

$$\boldsymbol{J} = \sum_{i=1}^{M} \Gamma_i \boldsymbol{x}_i. \tag{3.10}$$

When studying the statistics of point vortices on the disk, Bühler [21] did not observe the (planar) angular momentum to be of great importance. On the

sphere, however, the angular momentum *does* play an important role in the statistics.

The GBK thermostat (3.7) is only applicable to nondivergent systems $\nabla \cdot f \equiv 0$. It is straightforward to check that this condition holds for the spherical point vortex model.

3.3.2 Time integration

A numerical integrator can be constructed by splitting the differential equations into integrable subproblems (see related ideas in [144, 117]). We develop a new integrator for the system in Chapter 4 which exactly preserves all Casimir functions of the system. Furthermore, backward error analysis for symplectic integrators can be extended to Poisson systems to explain approximate conservation of the Hamiltonian [62]. Due to the additive form of the angular momentum vector (3.10), it may also be preserved exactly using a pairwise splitting. By expanding the Hamiltonian into its pairwise terms in the dynamics we find

$$\dot{y} = B(y)\nabla H(y) = \sum_{i<j} B(y)\nabla H_{ij}(\boldsymbol{x}_i, \boldsymbol{x}_j), \qquad H_{ij} = \frac{\Gamma_i \Gamma_j}{8\pi} \ln\left(2 - 2\boldsymbol{x}_i \cdot \boldsymbol{x}_j\right)$$

Each pairwise interaction is represented by the dynamical system $\dot{y} = B\nabla H_{ij}$ with the associated time-Δt flow map $\phi^{i,j}_{\Delta t}$. The time-Δt flow map of the dynamics $B\nabla H$ may be approximated by a symmetric composition of pair flows

$$\Phi_{\Delta t} = \prod_{(i,j)\in C} \phi^{i,j}_{\Delta t/2} \circ \prod_{(i,j)\in C^*} \phi^{i,j}_{\Delta t/2},$$

where C is an ordered set of all possible pairs (i, j) with $i < j$ and C^* denotes the reverse ordering. This symmetric splitting yields a consistent numerical method of second order accuracy. The details of the integration procedure involving exact solution of the pairwise interaction is detailed in Chapter 4.

Because the flow map of each vortex pair is the exact solution of the local Poisson system $\dot{y} = B(y)\nabla H_{ij}$ and also respects the Casimirs of the system, the composition $\Phi_{\Delta t}$ is a Poisson integrator [62, p. 247]. Expanding the angular momentum vector as $\boldsymbol{J} = \boldsymbol{J}_{ij} + \sum_{k\neq i,j} \Gamma_k \boldsymbol{x}_k$ we note that the time integration of any pair (i, j) preserves the local angular momentum $\boldsymbol{J}_{ij}$ and leaves the other vortices untouched. Hence the angular momentum is exactly conserved by the splitting method. The Hamiltonian is not exactly conserved under the motion of vortex pair, but the error can be studied by backward error analysis; see e.g. [62]. Figure 3.1a shows the error in the energy for simulations over a range of time step sizes, confirming second order convergence. The angular momentum should be conserved exactly by the Strang splitting. The results displayed in Figure 3.1b confirm this as the errors are always well within machine precision.

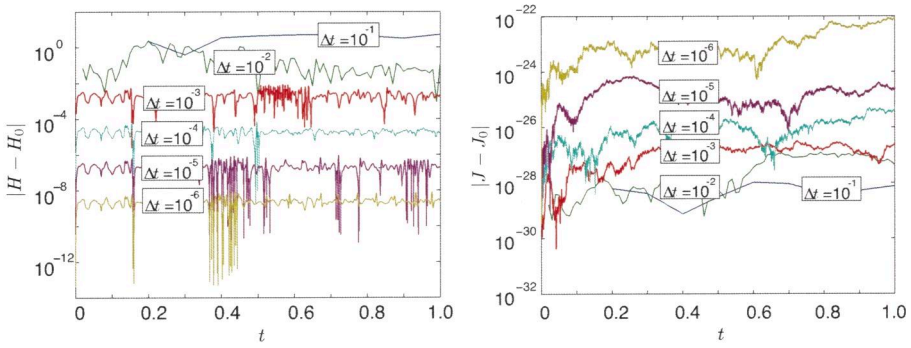


Figure 3.1: Error in the energy (left) and angular momentum (right) for different time step sizes. Both from simulations with 16 vortices, of which 4 with strength ± 5 and the rest with strength ± 1. The momentum error is due to limited floating point accuracy. With decreasing time step the number of time steps increases and the inaccuracies accumulate, but they remain well within any reasonable demand for accuracy.

3.3.3 Thermostat perturbation vector

There is flexibility in the choice of the perturbation vector field $g(y)$ in (3.7). Its selection is determined both by the need for ergodicity with respect to the target measure and the need to preserve some invariants of the vector field $f(y)$. We distinguish between invariants of f whose values are known precisely, due to problem geometry for instance, and those whose values are uncertain and only known in expectation. For point vortices on the sphere, the lengths of the vortex positions $|\boldsymbol{x}_i|$ are Casimir invariants, arising from the embedding of the sphere in $\mathbb{R}^3$, and are not subject to uncertainty. We choose a perturbation vector $g(y)$ that respects these structural invariants.

The double-bracket dissipation developed in [14] preserves Casimirs of the original system and is a candidate for $g(y)$:

$$\tilde{g}_i(\boldsymbol{x}_i) = \sum_{j \neq i} \boldsymbol{x}_i \times \boldsymbol{x}_i \times \frac{\Gamma_j}{4\pi} \frac{\boldsymbol{x}_j}{1 - \boldsymbol{x}_i \cdot \boldsymbol{x}_j}. \tag{3.11}$$

The denominator in (3.11) causes stiffness when like-signed vortices approach one another, restricting the step size of an explicit splitting method. To alleviate these matters we use a modified scheme defined by

$$g_i(\boldsymbol{x}_i) = \sum_{j \neq i} \boldsymbol{x}_i \times \boldsymbol{x}_i \times \frac{\Gamma_j}{4\pi} \boldsymbol{x}_j. \tag{3.12}$$

The desirable properties of the thermostat are unaffected by this modification. Chapter 4 contains a detailed description of the numerical integration of these dynamics.

The thermostat (3.7) is designed to sample a target density $\rho(y) \propto e^{-A(y)}$ on the phase space of y. The thermostat variable ξ is normally distributed, yielding the extended distribution $\rho \propto e^{-A(y)-\frac{1}{2}\xi^2}$. The perturbation vector field g must additionally ensure that the thermostated system is ergodic in the target density. Because the target measure is positive for all open sets on the phase space, hypoellipticity of the Fokker-Planck equation associated with (3.7) is sufficient to prove uniqueness of the invariant measure [10]. Hypoellipticity follows from Hörmander's controllability condition [122]. The condition has been tailored to GBK thermostats in [10], but it is difficult to check in practice. Here we instead check empirically that single trajectories have statistics that agree with the target distribution.

3.3.4 Maximum entropy model

To apply the methodology proposed in Section 3.2 in the setting of a reduced model for point vortices, we use point vortices distributed evenly over the surface of the sphere as the prior π. it remains to specify for which functions of the conserved quantities H and $\boldsymbol{J}$ the expectations will be observed during simulation of the full model.

In [44] a thermostat was used to model a set of point vortices on a disk in the canonical ensemble. To accurately reproduce statistics from a full model with a moderate number of point vortices, it was necessary to modify the canonical density with a term quadratic in the Hamiltonian, that is, a density of the form $\rho(y) \propto \exp(-\beta H(y) - \gamma H(y)^2)$. Motivated by the experience in [44], we choose observations that include linear and quadratic functions of H and $\boldsymbol{J}$.

If the angular momentum of the full system is zero, then there is no directional preference for the angular momentum vector $\boldsymbol{J}$. We consider the following set of observables:

$$C_1 = H, \quad C_2 = |\boldsymbol{J}|^2, \quad C_3 = H^2, \quad C_4 = |\boldsymbol{J}|^4, \quad C_5 = H^2|\boldsymbol{J}|^2, \tag{3.13}$$

and denote the corresponding Lagrange multipliers by $\beta_H, \beta_J, \gamma_H, \gamma_J, \gamma_{HJ}$.

The least-biased density consistent with observations of the $\mathbb{E}C_k$ is

$$\tilde{\rho}(H) = e^{-\beta_H H - \beta_J |\boldsymbol{J}|^2 - \gamma_H H^2 - \gamma_J |\boldsymbol{J}|^4 - \gamma_{HJ} H |\boldsymbol{J}|^2}. \tag{3.14}$$

3.4 Numerical comparison

To verify the methodology proposed in this article for correcting expectations, we apply it to a reduced model of point vortices similar to the configuration used in [20, 44]. We distinguish between three models. The *full model* consists of a system (3.9) of $M_{\text{full}} = 288$ point vortices, of which 8 strong vortices of circulation $\Gamma_j = \pm 1$ and 280 weak vortices of circulation $\Gamma_j = \pm\frac{1}{5}$. Both

strong and weak classes are comprised of equal numbers of positively and negatively oriented point vortices. The *reduced model* consists of (3.9) with just $M = 8$ strong vortices. Finally, the *corrected model* consists of a thermostated system (3.7) with unperturbed vector field f given by (3.9) for $M = 8$ strong vortices, perturbation vector field g given by (3.12), and equilibrium measure defined by the least-biased density (3.14). Additionally, we compare with Metropolis-Hastings samples from the least-biased density (3.14) to help distinguish between errors incurred due to the maximum-entropy model and those due to the thermostat.

We run seven long simulations of the full model with angular momentum vector $\boldsymbol{J}_{\text{full}} = \mathbf{0}$ and total energies chosen from the set $H_{\text{full}} \in \{-2, -1, 0, 1, 2\}$. For each run we determine the time averages of the observables (3.13) for the subset of strong vortices. When computing the Hamiltonian H we include only the internal coupling between strong vortices. The time averages are tabulated in Table 3.1.

Table 3.1: Full model observations and (in parentheses) corrected values of first integrals

	$\langle H \rangle$	$\langle \lvert\boldsymbol{J}\rvert^2 \rangle$	$\langle H^2 \rangle$	$\langle \lvert\boldsymbol{J}\rvert^4 \rangle$	$\langle H \lvert\boldsymbol{J}\rvert^2 \rangle$
$H_{\text{full}} = -2$	-0.33 (-0.38)	4.59 (4.45)	0.22 (0.23)	-0.63 (-0.98)	34.58 (31.45)
$H_{\text{full}} = -1$	-0.11 (-0.18)	4.78 (4.68)	0.10 (0.12)	0.38 (-0.01)	37.55 (35.88)
$H_{\text{full}} = 0$	0.02 (-0.04)	4.63 (4.56)	0.08 (0.08)	0.90 (0.60)	35.26 (34.30)
$H_{\text{full}} = 1$	0.17 (0.15)	4.74 (4.75)	0.13 (0.12)	1.73 (1.61)	37.76 (37.44)
$H_{\text{full}} = 2$	0.31 (0.28)	4.87 (5.00)	0.22 (0.21)	2.49 (2.46)	39.26 (41.74)

Given the time averages, we compute the Lagrange multipliers using the algorithm described in Section 3.2.1 with prior distribution π the uniform distribution on the sphere. The Lagrange multipliers are also recorded in Table 3.2. The magnitude of $\gamma_{\{H,J,HJ\}}$ indicates that all observations are relevant for all but the most negative energy levels.

Subsequently, we run simulations of the corrected model using the computed parameters. Table 3.1 also records expectations from the thermostat-corrected model.

By analogy with canonical statistical mechanics, we may think of the weak vortices that are ignored in the reduced model as forming a reservoir with which our reduced model exchanges energy and angular momentum. Experience with canonical statistical mechanics of point vortices in the plane [20, 44] suggests that for small reservoir sizes the canonical Gibbs distribution must be modified with higher order terms to agree with the full system statistics. Table 3.3 contains a study of the Lagrange multipliers as a function of the full system size M_{full}, confirming that the Lagrange multipliers γ_H, γ_J and γ_{HJ}

Table 3.2: Lagrange multipliers for each energy level.

	β_H	β_J	γ_H	γ_J	γ_{HJ}
$H_{\text{full}} = -2$	5.98	−0.20	0.69	0.41×10^{-3}	−0.04
$H_{\text{full}} = -1$	2.89	−0.03	2.67	9.77×10^{-3}	−0.33
$H_{\text{full}} = 0$	−0.76	0.20	3.38	9.97×10^{-3}	−0.37
$H_{\text{full}} = 1$	−3.54	0.37	4.29	15.31×10^{-3}	−0.54
$H_{\text{full}} = 2$	−6.42	0.53	4.45	14.05×10^{-3}	−0.51

are more significant for smaller M_{full}.

Table 3.3: Lagrange multipliers as a function of M_{full}, all for $H_{\text{full}} = 0$.

	β_H	β_J	γ_H	γ_J	γ_{HJ}
$M_{\text{full}} = 36$	−1.51	1.35	27.75	117.15×10^{-3}	−3.07
$M_{\text{full}} = 72$	−4.27	0.82	8.71	37.79×10^{-3}	−1.12
$M_{\text{full}} = 144$	−0.97	0.32	6.70	25.80×10^{-3}	−0.82
$M_{\text{full}} = 288$	−0.76	0.20	3.38	9.97×10^{-3}	−0.37
$M_{\text{full}} = 576$	−1.09	0.13	0.87	3.08×10^{-3}	−0.10

The energy of the strong vortices may become arbitrarily large because of the singularity in the Hamiltonian as two vortices approach each other. But the same is true for the energy in the reservoir. If there are at least three reservoir vortices and not all those vortices have the same sign, the reservoir can supply or remove any amount of energy.

The condition on the angular momentum is more interesting. The system of strong vortices, all with strength $\pm\Gamma_{\text{strong}}$, has angular momentum satisfying $|\boldsymbol{J}_{\text{red.}}| \leq M\Gamma_{\text{strong}}$. For the reservoir it holds that $|\boldsymbol{J}_{\text{weak}}| \leq (N-M)\Gamma_{\text{weak}}$. It is necessary that the reservoir can supply sufficient angular momentum, that is

$$M\Gamma_{\text{strong}} \leq (M_{\text{full}} - M)\Gamma_{\text{weak}} \Leftrightarrow \frac{M_{\text{full}} - M}{M} \geq \frac{\Gamma_{\text{strong}}}{\Gamma_{\text{weak}}}.$$

In the thermal bath simulations discussed in this section $M = 8$ and $\frac{\Gamma_{\text{strong}}}{\Gamma_{\text{weak}}} = 5$, this means M_{full} should satisfy $M_{\text{full}} \geq 48$. The smallest system considered ($M_{\text{full}} = 36$) does not, explaining its eccentric parameter values in Table 3.3.

3.4.1 Equilibrium results

In this section we compare statistical properties of the corrected model with those of the full and reduced models. In Figures 3.2–3.4 we show histograms of a number of solution features for the 8 vortex model: the distributions of H and $|\boldsymbol{J}|$, as well as typical distances between like- and opposite-signed vortices, a metric used by Bühler [20]. In each histogram, the statistics corresponding to the strong vortices in the full model, the reduced model, thermostat-corrected reduced model, and Metropolis-Hastings samples are displayed. Figures 3.2–3.4 correspond to approximate total energies $H_{\text{full}} \approx -2$, 0 and 2, respectively.

The full and reduced model simulations are performed with a time step of 5×10^{-3} and run up to $T = 5 \times 10^4$, taking 10^5 samples spaced evenly in time. For the Metropolis-Hastings method we use 10^6 samples. The same figures also show results from the thermostated system (dash-dot lines), run with a time step of 10^{-3} up to $T = 10^6$, taking 10^6 samples. The parameters in (3.7) were set to be $\varepsilon = 10$ and $\gamma = 0.1$. These results confirm that the thermostated system samples the least-biased density closely.

The reduced model is Hamiltonian and the Poisson integrator ensures that the energy is conserved with a standard deviation of order 10^{-3} and the angular momentum constant to machine precision. Both cases correspond to approximate delta-distributions in the upper histograms in Figures 3.2–3.4. Note that due to the high skewness of the distribution for $|\boldsymbol{J}|$, the observed mean differs significantly from the median and mode, implying some ambiguity in choosing the angular momentum for an appropriate initial condition for the reduced model.

A simple Hamiltonian reduced model is naturally incapable of sampling the energy and angular momentum spectra, since these quantities are first integrals. In turn, the reduced model shows significant bias in statistics such as vortex separation. The thermostat-corrected model faithfully samples the least-biased probability density, as indicated by the good agreement in the histograms of the corrected model and Metropolis-Hastings samples. The least-biased density does a good job of approximating the strong-vortex statistics in the negative to moderate total energy regime. At large positive total energies, the strong vortex energy and angular momentum distributions are still well-represented by the least-biased PDF, but some bias in the vortex separations can be observed. The closeness of the thermostat results to those from the Metropolis-Hastings sampling indicate the error lies in the choice of least-biased density, not in the thermostat sampling.

3.4.2 Dynamic consistency

The results in the previous section confirm that the thermostated simulations lead to equilibrium distributions of observables H and $|\boldsymbol{J}|$ similar to those of the full system. In this section we address the issue of the degree to which our equilibrium correction mechanism disturbs dynamics, as encoded in autocorrelation functions and diffusivity. Diffusivity was considered by [27] for a system of identical point vortices and by [30] for a wide array of problems with scale separation. We emphasize that the values of the thermostat parameters ε and γ have no impact on the equilibrium statistics presented in the previous section, and only affect the rate at which the least-biased PDF is sampled. Faster convergence to the equilibrium distribution correlates with a larger deviation from the unperturbed dynamics and vice-versa.

Autocorrelation functions

Given a sequence of L equally spaced observation times $t_i \in [0, T]$ for $i \in [0, L]$, and the values of the relevant observable (in our case vortex position) $u_i = u(t_i)$ at those times, the discrete autocorrelation function is defined by

$$\nu_i^u = \frac{1}{L-i} \sum_{j=i}^{L} u(t_j) u(t_{j-i}).$$

A normalized autocorrelation function $\hat{\nu}^u$ is given by dividing each ν_i^u by ν_0^u, i.e. $\hat{\nu}_i^u = \nu_i^u / \nu_0^u$.

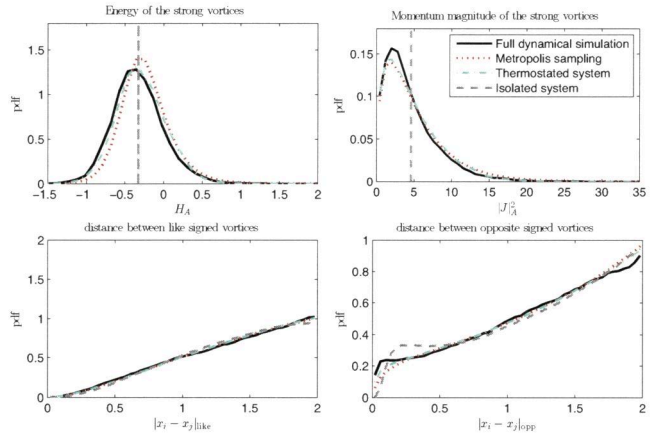


Figure 3.2: Histograms for $H_{\text{full}} \approx -2$. The upper left and right panels compare strong vortex energy and angular momentum magnitude. The lower left (resp. right) panel compares the distance between like (resp. opposite) signed strong vortices. The parameters are specified in the text.

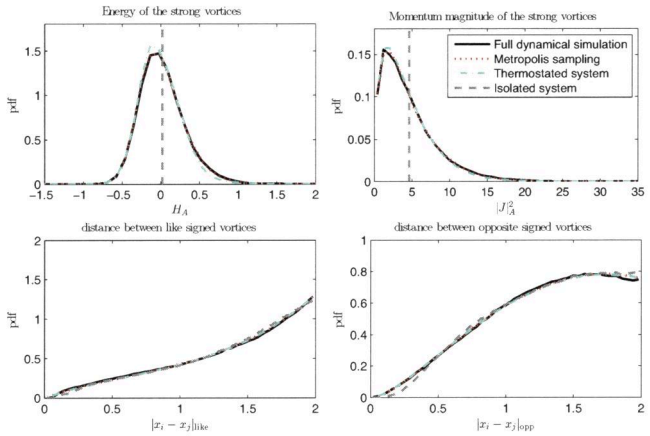


Figure 3.3: Histograms for $H_{\text{full}} \approx 0$. Same panel layout as Figure 3.2.

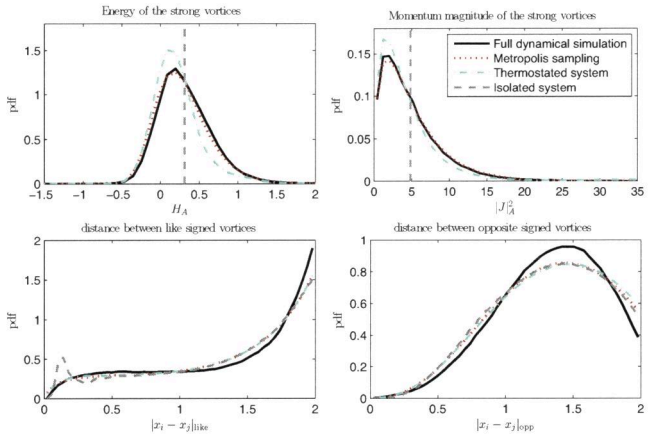


Figure 3.4: Histograms for $H_{\text{full}} \approx 2$. Same panel layout as Figure 3.2.

We average the autocorrelation function over all $3M$ (strong) vortex coordinates. Three symmetries in the problem justify this averaging: the vortex numbering is arbitrary; the choice of reference frame is arbitrary and the sign of the vortices appears in the dynamics as a reversal of time, to which the autocorrelation is insensitive. Additionally, the observables H and $|\boldsymbol{J}|$ are isotropic.

Furthermore we ensure that the phase space is well sampled by averaging the autocorrelation functions over an ensemble of P solutions. The choice of ensemble initial condition is detailed in the Section titled *Dynamical results* below. We then find the average autocorrelation function

$$\nu_i = \frac{1}{3MP(L-i)} \sum_{p=1}^{P} \sum_{m=1}^{M} \sum_{j=i}^{L} x_m^p(t_j) x_m^p(t_{j-i}) + \ldots \\ y_m^p(t_j) y_m^p(t_{j-i}) + z_m^p(t_j) z_m^p(t_{j-i})$$

and the normalized average autocorrelation function

$$\hat{\nu}_i = \frac{1}{MP(L-i)} \sum_{p=1}^{P} \sum_{m=1}^{M} \sum_{j=i}^{L} \boldsymbol{x}_m^p(t_j) \cdot \boldsymbol{x}_m^p(t_{j-i}), \tag{3.15}$$

where a superscript p represents the solution from ensemble member p. The normalized autocorrelation function in (3.15) follows from the Casimirs $C_i = \boldsymbol{x}_i(t) \cdot \boldsymbol{x}_i(t) = 1 \, \forall \, i, t$.

Diffusivity

For general multiscale dynamical systems with a separation of slow and fast dynamics, it is often desirable to model fast forces by a diffusion process, resulting in stochastic differential equation of the form [118, 71]

$$dX = f(X)dt + K(X)dW,$$

where f represents the slow dynamics, W is a Wiener process and K is the diffusivity. The value of the diffusivity K can be estimated by sampling solutions to the original, multiscale, problem and applying Kubo's formula

$$K_{\Delta\tau} = \frac{\langle \Delta X \Delta X \rangle}{\Delta\tau},$$

where ΔX represents displacement during the sampling interval $\Delta\tau$. Choosing the correct sampling interval is a notorious problem; for a comparison see [118].

If we take the average diffusivity for each vortex coordinate we find

$$\begin{aligned} K_{\Delta\tau} &= \frac{1}{6M\Delta\tau}\sum_{m=1}^{M}\langle \Delta x_m \Delta x_m + \Delta y_m \Delta y_m + \Delta z_m \Delta z_m\rangle \\ &= \frac{1}{6M\Delta\tau}\sum_{m=1}^{M}\langle \Delta \boldsymbol{x}_m \cdot \Delta \boldsymbol{x}_m\rangle\,. \end{aligned}$$

We assume the observations are given at the same times t_i as before and that the sampling time is an integer multiple of the observation interval, i.e. $\Delta\tau = i\frac{T}{L}$. With an ensemble of P simulations the diffusivity estimator would then be

$$\begin{aligned} K_{\Delta\tau} &= \frac{1}{6MP\Delta\tau}\sum_{p=1}^{P}\sum_{m=1}^{M}(\boldsymbol{x}_m^p(t_i) - \boldsymbol{x}_m^p(0))\cdot(\boldsymbol{x}_m^p(t_i) - \boldsymbol{x}_m^p(0)) \\ &= \frac{1}{6MP\Delta\tau}\sum_{p=1}^{P}\sum_{m=1}^{M} 2 - 2\boldsymbol{x}_m^p(0)\cdot\boldsymbol{x}_m^p(t_i) \\ &= \frac{1}{3\Delta\tau} - \frac{1}{3MP\Delta\tau}\sum_{p=1}^{P}\sum_{m=1}^{M}\boldsymbol{x}_m^p(0)\cdot\boldsymbol{x}_m^p(t_i), \end{aligned}$$

where again a superscript p denotes the solution from ensemble member p. Averaging over all time series data yields the estimator:

$$\begin{aligned} K_{\Delta\tau}^i &= \frac{1}{3\Delta\tau} - \frac{1}{3MP(L-i)\Delta\tau}\sum_{p=1}^{P}\sum_{j=i}^{L}\sum_{m=1}^{M}\boldsymbol{x}_m^p(t_{j-i})\cdot\boldsymbol{x}_m^p(t_j) \\ &= \frac{1-\hat{\nu}_i}{3\Delta\tau}. \end{aligned}$$

This *shift-averaged estimator* is shown in [30] to improve the quality of the estimator.

Dynamical results

In Figure 3.5 we compare auto-correlation functions for the strong vortices in the full and reduced models as well as for the thermostat-corrected model over a range of parameters ε and γ. The thick solid black line represents the result for an (unthermostated) system in contact with 280 weak ($\Gamma_B = \pm\frac{1}{5}$) vortices, with a total energy $H_{\text{full}} = 0$. The results present the average over an ensemble of 1000 runs. For each simulation the initial placement of each strong vortex was taken uniformly over the sphere and the weak vortices were placed such that the full system satisfied $H_{\text{full}} = 0$ and $\boldsymbol{J}_{\text{full}} = 0$. The thick dashed black line represents the results for an ensemble of simulations of the isolated system, with everything else unchanged. The other lines represent

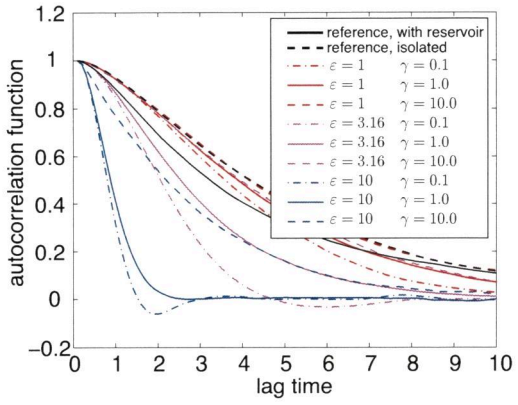


Figure 3.5: Comparison of autocorrelation of the vortex coordinates. The bold lines are two reference cases: the full model (solid) and the reduced model (dashed). The thin lines indicate autocorrelation functions of the thermostat-corrected model for indicated values of parameters ε and γ.

results for thermostated simulations using the parameters as given in Table 3.2 for the case of $H = 0$.

The corresponding diffusivity constants are presented in Figure 3.6. The results are taken from the same simulations as described in the paragraph above. Because this figure is visually more striking, we shall limit our discussion to the diffusion constant.

For ε small, the thermostat perturbation is weak, and both autocorrelation functions and diffusivity approach those of the reduced model with constant H, $\boldsymbol{J}$. Also, the autocorrelations are insensitive to the parameter γ in this regime. For even smaller ε the autocorrelations and diffusivities become indistinguishable from those of the reduced model. Hence even though the dynamics samples the least-biased density on long time scales, its short time dynamics is similar to an unperturbed model. For moderate ε, dependence on γ becomes more pronounced, and a diffusivity closer to that of the full model can be achieved. For even larger values of ε, the diffusivity becomes much more sensitive to the value of γ, as indicated in Figure 3.6(c).

Figure 3.6(a) has been included to illustrate two important properties. Firstly, as the sampling interval goes to zero, the estimator of the diffusivity constant shows linear behavior. This is in agreement with known results for the GBK thermostat[54] and is an improvement on Langevin thermostats, which would erroneously tend to a constant value as the sampling interval is decreased. Secondly, for large sampling interval the estimator shows an inverse linear tendency. This corresponds simply to the decorrelation of the

vortex dynamics.

3.4.3 Adaptive determination of multipliers

Consider the same reduced point vortex model of 8 vortices with $\Gamma = \pm 1$ and assume observations on the energy and momentum are known from a simulation of the full system including the thermal bath. We start such a simulation with an ensemble of $P = 100$ initial conditions drawn from the uniform prior. The time step is chosen as 1×10^{-2} and the method described in Section 3.2.1 for updating the Lagrange multipliers is applied every time unit, i.e. $M = 100$. Between subsequent updates of the multipliers, the maximum difference is limited by $|\Delta\lambda_k| \leq 0.1$. When using equilibrium statistics, this limit only affects the beginning of the simulation, when the small sample size

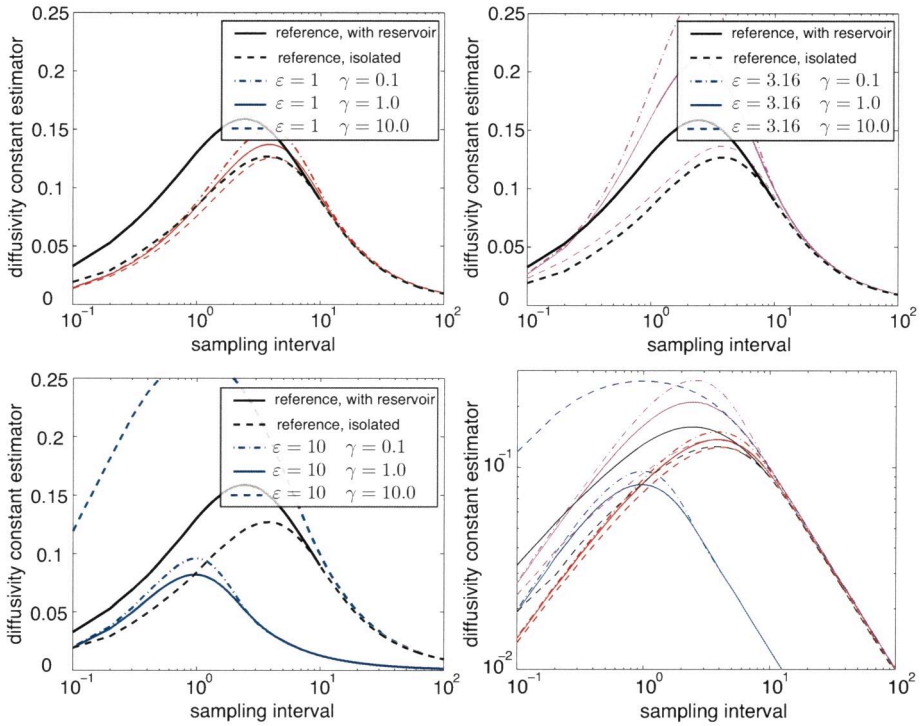


Figure 3.6: Comparison of average diffusivity constant as a function of sampling intervals. In all figures, the bold lines indicate two reference cases: the full model (solid) and the reduced model (dashed). Subfigures (a), (b) and (c) show thermostated simulation results for ε equal to 10^0, $10^{0.5}$ and 10^1 respectively. The value of γ is represented by dash-dot (10^{-1}), solid (10^0) or dashed (10^1) lines. A combined log-log plot of all parameter values is given in subfigure (d).

used leads to a large variance in the estimators.

The target observation values are taken from a simulation of strong vortices interacting with a thermal bath of weak vortices. Three different averages are used.

1. In Figure 3.7 the long time mean is taken and used throughout.
2. In Figure 3.8 the running mean is used. This reflects the situation where we have no a priori knowledge of the observations, and are continuously feeding new real-time data into the simulation.
3. In Figure 3.9 a time-localized average of the observable is used. The averaging has a time-scale of a 100 time units. This also corresponds to feeding the simulation new data, but now the assumption of equilibrium is relaxed.

When using either a long time mean observation or a running mean observation, the simulation results tends towards the correct long-time averages. When using time-local averages the simulation averages appear to tend towards a similar value. In all three cases the instantaneous ensemble mean remains close to the (moving) target for both energy and momentum. This is especially notable for the third case, where the target varies over time, but the simulation ensemble mean follows closely, with only a little lag.

The inaccuracies incurred during the first approximately 100 time units indicate that the prior does not match the observed state well. This results in the (negative) growth of β_H being limited briefly at the beginning of each simulation. Subsequently, both Lagrange multipliers appear to oscillate irregularly about some mean value for the first two cases. In the case of a shifting target, the Lagrange multipliers vary in time more erratically, as is to be expected. This results in the limiter being active for a few brief periods of the simulation.

3.5 Conclusion

In this article we propose a method for perturbing trajectories of numerical simulations to correct for equilibrium observations. We treat the restricted case of a Hamiltonian ODE with observations on the set of first integrals of the system. The approach entails perturbing the solutions using a stochastic thermostat such that they become ergodic in a prescribed invariant measure: the least-biased density corresponding to a maximum entropy treatment of the observations.

We apply the approach to the case of model reduction in a heterogeneous system of weak and strong point vortices on a sphere, in which observations of the energy H and angular momentum magnitude $|\boldsymbol{J}|$ are made on a subsystem consisting of the strong vortices. A reduced model is constructed by neglecting the weak vortices, and the expectations of the reduced model are

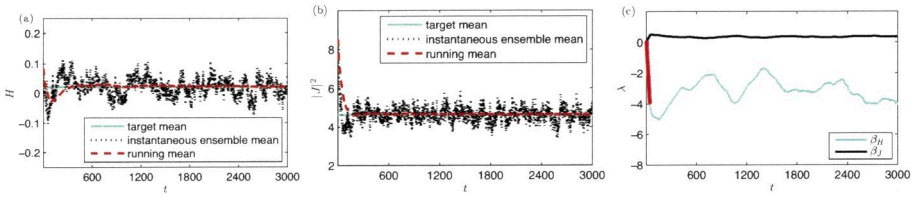


Figure 3.7: Results when using *long-time mean* observations as a target while adaptively determing the Lagrange multipliers. Target observations for Hamiltonian (a) and momentum magnitude (b) are overlaid with the instantaneous ensemble mean (black dotted) and the running ensemble mean (red solid) from simulation. (c): Lagrange multipliers, the red dots indicate time steps at which their rate of change was limited.

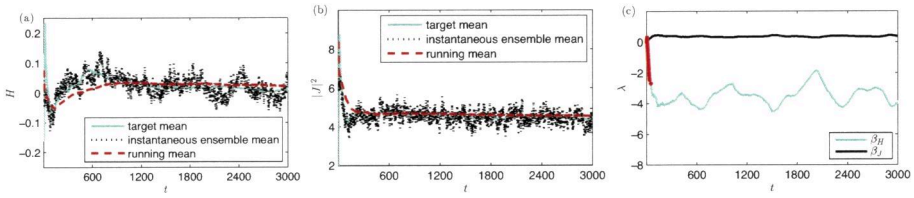


Figure 3.8: Results when using *running mean* observations as a target while adaptively determing the Lagrange multipliers. Target observations for Hamiltonian (a) and momentum magnitude (b) are overlaid with the instantaneous ensemble mean (black dotted) and the running ensemble mean (red solid) from simulation. (c): Lagrange multipliers, the red dots indicate time steps at which their rate of change was limited.

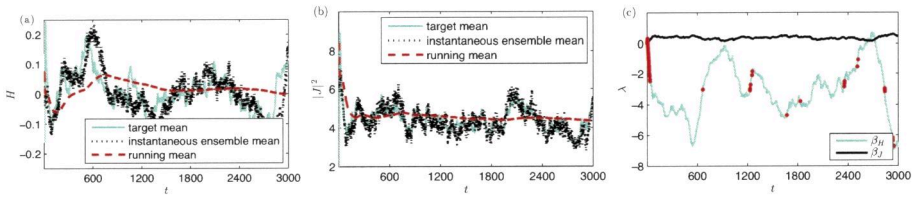


Figure 3.9: Results when using *time-local averaged* observations as a target while adaptively determing the Lagrange multipliers. Target observations for Hamiltonian (a) and momentum magnitude (b) are overlaid with the instantaneous ensemble mean (black dotted) and the running ensemble mean (red solid) from simulation. (c): Lagrange multipliers, the red dots indicate time steps at which their rate of change was limited.

corrected using the proposed methodology.

Numerical experiments confirm that the distributions of the observed quantities H and $|\boldsymbol{J}|$ can be well represented using the thermostat technique. Other equilibrium metrics such as the distribution of distances between like- and opposite-signed vortices are also in agreement across a range of total energy values of the full system, although some discrepancies occur at large positive energies.

We also investigated the degree to which correction of trajectories for expectations may affect dynamical information in the form of autocorrelation functions and diffusivity. By decreasing the perturbation parameter ε of the thermostat, the autocorrelation functions of the unperturbed, reduced system may be precisely recovered. As ε is increased, one may increase the diffusivity to values that agree with the full system. This is consistent with results reported in [54] in the context of molecular dynamics where it was shown that the GBK thermostat used here approaches Langevin dynamics in the limit of large stochastic forcing.

4

Explicit parallel Poisson integration of point vortex systems

This chapter is based on:

An article with the same title submitted to Journal of Computational and Applied Mathematics.

background graphic: depiction of the ordering of vortex pair interactions

4.1 Motivation

A point vortex represents a singular measure solution to the vorticity equation for two-dimensional, incompressible fluid flow. A point vortex model consists of multiple point vortices mutually interacting. The motion of each point vortex is dictated by the flow field induced by the other vortices and by external forcing, e.g. topography. Point vortices were introduced by Helmholtz [67] and have since been the subject of much study; see for example Lamb [84], Saffman [129], Newton [109].

Dynamical studies of point vortex systems provide insight into the (qualitative) behaviour of fluid dynamics. The series of papers by Newton et al. [109, 72, 111, 112] discuss relative equilibria and the conditions for integrability of the dynamics. Vortex dynamics were studied extensively by Aref who compiled an extensive review on their history [6]. Newton [110] discusses the future of point vortex research in the "post-Aref era".

In statistical fluid mechanics, the behaviour of point vortex systems has been studied as a model for two-dimensional turbulence in the limit of an infinite number of vortices. This was first done by Onsager [116], who provided an explanation for the formation of clusters of like-signed vortices in a bounded domain. This research has since been continued by, amongst others, Joyce and Montgommery [77, 76], Pointin & Lundgren [120], Eyink & Spohn [48], and Lions & Majda [91]. Such results are of interest in the fields of geophysical fluid dynamics [61] and stellar dynamics [28]. Some of Onsager's statements were tested numerically by Bühler [20].

Point vortices and their three-dimensional generalization, vortex filament methods, are also used as a discretization of practical fluid flows in engineering applications [31]. By using a large number of point vortices a continuous velocity field is approximated. Such techniques find practical application in the works of Chatelain et al. [25, 26], Rossinelli et al. [126]. Winckelmans et al. [141] and Rossinelli & Koumoutsakos [125] present the fast multipole, vortex-in-cell and hybrid methods that are used for computing these large systems. Regularized approximations to the delta distributions provide more accurate representations of continuous vorticity fields, but their solutions are no longer exact, as the kernel itself ought to deform due to shearing [18, 140].

It is important to develop efficient time integrators for point vortex methods for two reasons. First, the use of very large numbers of point vortices, as required for accurate approximation of continuous fluids, is hampered by the quadratic complexity of the pairwise coupling between vortices, i.e. evaluations of the vector field with N vortices requires N^2 operations. Second, the concept of numerical stability of a system of point vortices on planar geometry is not without ambiguity. Equilibria only exist for certain configurations, and are never asymptotically stable since the dynamics are Hamiltonian. The simplest nontrivial system is a pair of like-signed vortices, whose solution is

periodic. If a contracting method such as backward Euler is employed, the vortices will eventually approach one another, and the derivatives grow unbounded. If an expanding method such as forward Euler is employed, the vortices will drift apart and the trajectories will grow without bound. Hence, even for this simple configuration some degree of energy conservation is necessary to maintain a bounded solution with bounded derivative.

Recently, Vankerschaver & Leok [140] have developed a Poisson integrator for point vortex systems via the construction of a higher dimensional linear Lagrangian. The associated dynamics project down onto solutions of the point vortex equations on the sphere. The resulting integrator exactly conserves the Casimirs and momentum of the point vortex dynamics and also has good conservation of energy. The implicit definition, however, requires the use of an iterative solver.

We give an interpretation of the point vortex method in light of the approach first communicated by McLachlan [98] for discretizing Hamiltonian PDEs; namely as a scheme that discretizes the Poisson structure and Hamiltonian separately. With a vorticity field given as a sum of point vortices, the quadrature of the Hamiltonian functional is evaluated exactly as a sum of pointwise values. We do not consider regularizations of the vortices, but they could be accommodated in the quadrature scheme for the Hamiltonian. The Poisson bracket is discretized exactly for a particular class of functionals.

A numerical integrator for these dynamics follows from splitting the Hamiltonian into its constituent pairwise terms. The scheme developed is Poisson, explicit and allows scalable parallelization. It may also be applied to regularized point vortices, provided the kernel is rotation- and translation-invariant. The method requires an explicit expression for the pairwise flow map for the two-vortex system. Any regularization that maintains a pairwise Hamiltonian form will have three Poisson-commuting first integrals and is thus integrable. Both rotation of the sphere and topography introduce only decoupled, splittable terms in the Hamiltonian.

The remainder of this paper is organized as follows. Section 4.2 describes two-dimensional incompressible fluid flow in Hamiltonian form. Section 4.3 discusses the discretization according the ideas of McLachlan [98]. A Poisson integrator for the resulting point vortex description for fluids is developed in Section 4.4. The parallelization of this method is discussed in Section 4.5. Numerical results and comparisons of computation times are presented in Sections 4.4 and 4.5, respectively. Finally, in Section 4.6 we state conclusions and discuss the extension of the method to practical applications.

4.2 Continuous Hamiltonian description

The barotropic quasi-geostrophic equations on the unit sphere provide a simple model for studying geophysical fluid dynamics [95]. Point vortex representations capture much of the system's dynamics, for instance the formation of coherent vortical structures over long time [116]. This is a consequence of the existence of negative temperature states, that are possible due to the bounded domain. On a disk or on an annulus, the same behaviour can be observed, but these geometries require the inclusion of ghost vortices to maintain the boundary conditions. The boundedness of the domain also implies that solutions remain bounded for almost any initial condition when considering heterogeneous systems, i.e. systems with both positive and negative circulation vortices.

We express the barotropic quasi-geostrophic equations on the sphere [95] in terms of the stream function ψ and potential vorticity q

$$q_t + J(\psi, q) = 0 \tag{4.1}$$

$$q = \Delta_S \psi + 2\Omega z + h, \tag{4.2}$$

where Ω is the angular velocity of the sphere about the z-axis and h represents topography. The Laplace-Beltrami operator on the sphere Δ_S is defined (in spherical coordinates) as

$$\Delta_S \psi = \frac{1}{\cos\theta}\left[\frac{1}{\cos\theta}\psi_{\phi\phi} + \frac{\partial}{\partial\theta}\left(\cos\theta\,\psi_\theta\right)\right],$$

where ϕ is the longitude and θ the latitude. The Jacobian $J(f, g)$ is defined as

$$J(f, g) = \frac{1}{\cos\theta}\left(f_\phi g_\theta - g_\phi f_\theta\right). \tag{4.3}$$

On the sphere the Hamiltonian is given by:

$$\mathcal{H} = -\frac{1}{2}\int_{\mathbb{S}^2} \psi \Delta_S \psi \,\mathrm{d}S = \frac{1}{2}\int_{\mathbb{S}^2} \nabla_S \psi \cdot \nabla_S \psi \,\mathrm{d}S,$$

where the second equality follows from the divergence theorem. Using the rightmost expression we find the first variation of $\mathcal{H}$

$$\begin{aligned}
\delta\mathcal{H} &= \int_{\mathbb{S}^2} \nabla_S \psi \cdot \nabla_S \delta\psi \,\mathrm{d}S \\
&= -\int_{\mathbb{S}^2} \psi \Delta_S \delta\psi \,\mathrm{d}S \\
&= -\int_{\mathbb{S}^2} \psi\delta\left(q - 2\Omega z - h\right) \mathrm{d}S,
\end{aligned}$$

and consequently

$$\frac{\delta\mathcal{H}}{\delta q} = -\psi.$$

The Poisson bracket is given by

$$\{\mathcal{F},\mathcal{G}\}[q] = -\int_{\mathbb{S}^2} \frac{\delta\mathcal{F}}{\delta q} J\left(q, \frac{\delta\mathcal{G}}{\delta q}\right) \mathrm{d}S,$$

where the Jacobian J is given by (4.3).

Point vortex systems represent singular measure solutions to equations (4.1)–(4.2). They assume a potential vorticity field that can be expressed as the sum of Dirac-delta distributions, i.e.

$$q(\boldsymbol{x}) = \sum_{i=1}^{N} \Gamma_i \delta(\boldsymbol{x} - \boldsymbol{x}_i(t)).$$

The vortex centres are represented as vectors $\boldsymbol{x}_i$ embedded in $\mathbb{R}^3$. The geometric structure of the equations of motion preserves $|\boldsymbol{x}_i|$. Numerical integrators must maintain this property, either by construction or by introducing a projection step.

We introduce two new streamfunctions ψ_{C} and ψ_{h} such that $2\Omega z = \Delta_S \psi_{\mathrm{C}}$ and $h = \Delta_S \psi_{\mathrm{h}}$. With these we rewrite (4.2) as

$$\Delta_S \psi = \omega = q - 2\Omega - h = q - \Delta_S \psi_{\mathrm{C}} - \Delta_S \psi_{\mathrm{h}}.$$

We solve this for ψ

$$\psi = \Delta_S^{-1} q - \psi_{\mathrm{C}} - \psi_{\mathrm{h}}, \tag{4.4}$$

where $\Delta_S^{-1} q = \sum_{i=1}^{N} \Gamma_i G(\boldsymbol{x} - \boldsymbol{x}_i(t))$ represents the sum of Green's functions for the Laplace equation on the sphere, given by

$$G(\boldsymbol{x} - \boldsymbol{x}_i(t)) = \frac{1}{4\pi} \ln\left(|\boldsymbol{x} - \boldsymbol{x}_i(t)|^2\right). \tag{4.5}$$

4.3 Discrete Hamiltonian representation

In this section we review the point vortex description on the sphere by interpreting it as a Hamiltonian discretization in the sense of McLachlan [98]. By discretizing the Hamiltonian and Poisson bracket individually, and ensuring that the latter defines a finite dimensional Poisson bracket, it is guaranteed that the finite dimensional approximation is again Poisson, and Poisson integrators may be employed. For point vortices in planar geometry, the bracket is canonical and hence symplectic Runge-Kutta methods are applicable. On the sphere, the Poisson bracket is nontrivial and splitting methods offer the most generic approach.

With the assumption that the vorticity field is a sum of Dirac delta distributions, the integration of $\mathcal{H}$ reduces to a sum over the values of the integrand at the vortex centres

$$\mathcal{H} = -\frac{1}{2}\int_{\mathbb{S}^2} \psi q \,\mathrm{d}S = -\frac{1}{2}\sum_{i=1}^{N} \Gamma_i \psi(\boldsymbol{x}_i).$$

Substituting the inverse Laplacian of (4.4) with Green's function (4.5), we find

$$\mathcal{H} = \frac{1}{2}\sum_{i=1}^{N}\Gamma_i\left(\psi_{\mathrm{C}}(\boldsymbol{x}_i) + \psi_{\mathrm{h}}(\boldsymbol{x}_i) - 2\sum_{j<i}\Gamma_j\frac{1}{4\pi}\ln\left(|\boldsymbol{x}_j - \boldsymbol{x}_i|^2\right)\right) =: H. \quad (4.6)$$

The Hamiltonian can thus be expressed discretely in terms of only the positions of the vortex centres $\boldsymbol{x}_i$. This discrete representation of the dynamics is exact if the point vortices are singular, and hence the discrete H can be defined equal to the functional $\mathcal{H}$. It is assumed that the stream functions associated with the Coriolis and topography terms are known explicitly. For regularized systems, the Green's function may still be known, but the quadrature of $\mathcal{H}$ can no longer be performed exactly and the discretized Hamiltonian will no longer be exact.

The Poisson bracket is discretized separately. First of all, it is useful to rewrite (4.3), because we have defined the point vortex positions as vectors in $\mathbb{R}^3$ rather than in spherical coordinates. For any $\boldsymbol{x} \in \mathbb{R}^3 : |\boldsymbol{x}| = 1$, (4.3) is equivalent to

$$J(f, g, \boldsymbol{x}) = (\boldsymbol{x} \times \nabla f) \cdot \nabla g.$$

The Poisson bracket then follows as

$$\begin{aligned}\{\mathcal{F}, \mathcal{G}\} &= \int \frac{\delta\mathcal{F}}{\delta q}\nabla q \cdot \boldsymbol{x} \times \left(\nabla\frac{\delta\mathcal{G}}{\delta q}\right) \mathrm{d}S \\ &= -\int q\nabla\cdot\left(\frac{\delta\mathcal{F}}{\delta q}\boldsymbol{x} \times \nabla\frac{\delta\mathcal{G}}{\delta q}\right) \mathrm{d}S \\ &= -\int q\nabla\left(\frac{\delta\mathcal{F}}{\delta q}\right)\cdot\left(\boldsymbol{x} \times \nabla\frac{\delta\mathcal{G}}{\delta q}\right) \mathrm{d}S, \end{aligned} \quad (4.7)$$

using first the divergence theorem and then the fact that the divergence of the curl equals zero. The discrete form of the functional $\mathcal{F} = \int f(\boldsymbol{x})\,\mathrm{d}S$ is given by

$$F = \sum_{i=1}^{N} f(\boldsymbol{x}_i) = \int f(\boldsymbol{x})\left(\sum_{i=1}^{N}\delta(\boldsymbol{x} - \boldsymbol{x}_i)\right)\mathrm{d}S.$$

We assume there exists a field λ_q for the vorticity field q with the properties:

$$\begin{aligned}\lambda_q(\boldsymbol{x}_i) &= \Gamma_i^{-1}, \\ \nabla\lambda_q(\boldsymbol{x})|_{\boldsymbol{x}=\boldsymbol{x}_i} &= 0, \\ \lim_{\varepsilon\to 0}\frac{\lambda_{q+\varepsilon v} - \lambda_q}{\varepsilon} &= 0.\end{aligned}$$

With this we write

$$F = \int f(\boldsymbol{x})\lambda(\boldsymbol{x})\left(\sum_{i=1}^{N}\Gamma_i\delta(\boldsymbol{x} - \boldsymbol{x}_i)\right)\mathrm{d}S = \int f(\boldsymbol{x})\lambda(\boldsymbol{x})q(\boldsymbol{x})\,\mathrm{d}S,$$

from which the variational derivative follows

$$\begin{aligned}\left\langle \frac{\delta F}{\delta q}, v \right\rangle &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(\langle f\lambda_{q+\varepsilon v}, (q+\varepsilon v) \rangle - \langle f\lambda_q, q \rangle \right) \\ &= \langle f\lambda_q, v \rangle .\end{aligned}$$

Substitution of this form for the functional $\mathcal{F}$ and $\mathcal{G}$ in (4.7) leads to the discrete form of the Poisson bracket

$$\begin{aligned}\{F, G\} &= -\int q \nabla \frac{\delta F}{\delta q} \cdot \boldsymbol{x} \times \nabla \frac{\delta G}{\delta q} \, \mathrm{d}S \\ &= -\int q \nabla (f\lambda_q) \cdot \boldsymbol{x} \times \nabla (g\lambda_q) \, \mathrm{d}S \\ &= -\int q \lambda_q^2 \nabla f \cdot \boldsymbol{x} \times \nabla g \, \mathrm{d}S \\ &= -\sum_{i=1}^{N} \Gamma_i^{-1} \nabla f(\boldsymbol{x}_i) \cdot \boldsymbol{x}_i \times \nabla g(\boldsymbol{x}_i).\end{aligned}$$

This is a generalization of the well-known Poisson bracket for rigid body rotation [68], also used in models for ferromagnetism [92]. The bracket is in fact equivalent to the bracket for a Heisenberg spin chain [49].

We introduce the vector $\boldsymbol{y} \in \mathbb{R}^{3N}$ as the concatenation of the $\boldsymbol{x}_i \in \mathbb{R}^3$, $i = 1, 2, \ldots, N$. The dynamics are then

$$\dot{\boldsymbol{y}} = B(\boldsymbol{y}) \nabla H(\boldsymbol{y}), \tag{4.8}$$

with the block-diagonal structure matrix

$$B(\boldsymbol{y}) = \begin{bmatrix} \Gamma_1^{-1} \widehat{\boldsymbol{x}}_1 & & & 0 \\ & \Gamma_2^{-1} \widehat{\boldsymbol{x}}_2 & & \\ & & \ddots & \\ 0 & & & \Gamma_N^{-1} \widehat{\boldsymbol{x}}_N \end{bmatrix},$$

where $\widehat{\boldsymbol{x}}$ is the 3×3 skew-symmetric matrix such that $\widehat{\boldsymbol{x}}\boldsymbol{u} = \boldsymbol{x} \times \boldsymbol{u}$ $\forall \boldsymbol{u} \in \mathbb{R}^3$. The vortex position radii $C_i = |\boldsymbol{x}_i|$ are Casimirs of the Poisson bracket associated with structure matrix $B(\boldsymbol{y})$. That is, for any function $F(\boldsymbol{y})$ and any C_i, one has $\{F, C_i\} \equiv 0$. This property is important as it implies that if the vortex positions initially satisfy $|\boldsymbol{x}_i| = 1$, this is maintained throughout the simulation, ensuring the point vortices remain on the sphere. The numerical integration scheme developed below respects this property inherently, without the need of a projection step.

Due to the rotational symmetries of the sphere, the dynamics exhibit three Noether momenta given in vector form as $\mathcal{J} = \int_{\mathbb{S}^2} \boldsymbol{x} q \, \mathrm{d}S$. In the point vortex discretization, these momenta persist as

$$\boldsymbol{J} = \mathcal{J} = \int_{\mathbb{S}^2} \boldsymbol{x} q \, \mathrm{d}S = \sum_{i=1}^{N} \Gamma_i \boldsymbol{x}_i.$$

We summarize the discrete Hamiltonian representation in Table 4.1.

4.4 Poisson integrator

For Poisson systems such as the point vortex system it is essential to employ a numerical time integrator that maintains the structure of the underlying ordinary differential equations. Standard numerical integrators do not conserve Casimirs. Hence Runge-Kutta or multistep methods will result in point vortices drifting from the sphere. This can be corrected with projections, but as is known from the rigid body equation, doing so can introduce artificial stable equilibria.

Integrators that conserve the geometric structure are of special importance when one is interested in the statistics of long simulations. In geophysical fluid dynamics the long time mean vorticity field and streamfunction, as well as the pointwise statistics, depend heavily on the geometric properties of the numerical integrator [2, 42, 43].

Patrick [117] suggests applying a Poisson splitting method to point vortex dynamics, but does not detail the method. A Poisson integrator preserves Casimirs by definition. We will see that the splitting also preserves the Noether momenta exactly, and the Hamiltonian approximately in the sense of backward error analysis, as detailed in Section 4.4.2.

We expand system (4.8) with Hamiltonian (4.6) as

$$\dot{\boldsymbol{y}} = \sum_{i=1}^{N} B(\boldsymbol{y}) \nabla H_i + \sum_{j<i} B(\boldsymbol{y}) \nabla H_{ij},$$

where

$$H_i = \frac{1}{2}\Gamma_i \left(\psi_{\mathrm{C}}(\boldsymbol{x}_i) + \psi_{\mathrm{h}}(\boldsymbol{x}_i)\right), \text{ and}$$
$$H_{ij} = \frac{1}{4\pi} \ln\left(|\boldsymbol{x}_j - \boldsymbol{x}_i|^2\right) = \frac{1}{4\pi} \ln\left(2 - 2\boldsymbol{x}_i \cdot \boldsymbol{x}_j\right).$$

We treat the dynamics for each of these terms separately. The time-τ flow map associated with each of the H_i terms will be denoted by ϕ_τ^i. The dynamics associated with $\dot{\boldsymbol{y}} = B(\boldsymbol{y})\nabla H_{ij}$ is just that of a two vortex system with time-τ flow map denoted by ϕ_τ^{ij}. This flow is known explicitly as detailed in Section 4.4.1. A splitting method is a composition of the flow maps of all the individual terms in the dynamics.

We initially restrict ourselves to Lie-Trotter splittings and Strang splittings [62], respectively of the form

$$\Phi_\tau^{\mathrm{LT}} = \prod_{(i,j)\in C_N} \phi_\tau^{ij} \circ \prod_{i=1}^{N} \phi_\tau^i \quad \text{and} \quad \Phi_\tau^{S} = \prod_{(i,j)\in C_N} \phi_{\tau/2}^{ij} \circ \prod_{i=1}^{N} \phi_\tau^i \circ \prod_{(i,j)\in C_N^*} \phi_{\tau/2}^{ij}.$$

Table 4.1: Summary of discrete identities compared to continuous identities.

continuous representation	discrete representation
$\mathcal{F}(t) = \int f(\boldsymbol{x}) \sum_{i=1}^{N} \delta\left(\boldsymbol{x} - \boldsymbol{x}_i(t)\right) \mathrm{d}S$ $= \int f(\boldsymbol{x}) \lambda(\boldsymbol{x}) q(\boldsymbol{x}, t) \mathrm{d}S$	$F(t) = \sum_{i=1}^{N} f\left(\boldsymbol{x}_i(t)\right)$
$\mathcal{H}(t) = \int q\psi \mathrm{d}S$	$H(t) = \frac{1}{2} \sum_{i=1}^{N} \Gamma_i \Big(\psi_{\mathrm{C}}(\boldsymbol{x}_i(t)) + \psi_{\mathrm{h}}(\boldsymbol{x}_i(t)) - 2 \sum_{j<i} \Gamma_j \frac{1}{4\pi} \ln\left(\lvert \boldsymbol{x}_j(t) - \boldsymbol{x}_i(t) \rvert^2\right) \Big)$
$\mathcal{J}(t) = \int \boldsymbol{x} q(\boldsymbol{x}, t) \mathrm{d}S$	$J(t) = \sum_{i=1}^{N} \Gamma_i \boldsymbol{x}_i(t)$
$\mathcal{C}_j(t) = \int \lvert \boldsymbol{x} \rvert^2 \delta\left(\boldsymbol{x} - \boldsymbol{x}_j(t)\right) \mathrm{d}S$	$C_j(t) = \lvert \boldsymbol{x}_j(t) \rvert^2$
$\{\mathcal{F}, \mathcal{G}\} = -\int q \nabla \left(\frac{\delta \mathcal{F}}{\delta q} \right) \cdot \left(\boldsymbol{x} \times \nabla \frac{\delta \mathcal{G}}{\delta q} \right) \mathrm{d}S$	$\{F, G\} = -\sum_{i=1}^{N} \Gamma_i^{-1} \nabla f(\boldsymbol{x}_i(t)) \cdot \boldsymbol{x}_i \times \nabla g(\boldsymbol{x}_i(t))$
$\{q, \mathcal{H}\}$	$\left\{x_i^{(1)}, H\right\}, \left\{x_i^{(2)}, H\right\}, \left\{x_i^{(3)}, H\right\}$

(4.9)

For both cases C_N represents an ordering of all the possible pairs (i, j), $i \neq j$. The Strang splitting also uses the reverse ordering, labelled C_N^*, to create a symmetric method. The symmetry results in a cancellation of first order error terms, yielding a method that is second order accurate. The Strang splitting can subsequently be used in the construction of higher order methods [143, 99]. In the remainder we will ignore the effects of the Coriolis force and topography. The corresponding flows ϕ_τ^i are perfectly parallelizable and their evaluation represents an ever smaller fraction of the total workload as the number of vortices increases.

Because each of the pairwise interactions in the splittings in (4.9) is the exact solution to a Poisson dynamical system, each ϕ_τ^{ij} is a *Poisson map* with respect to the bracket $\{F, G\} = \nabla_{\boldsymbol{y}} F^T B(\boldsymbol{y}) \nabla_{\boldsymbol{y}} G$. As the composition of Poisson maps is again Poisson, the splitting schemes are also Poisson maps with respect to the same bracket [62].

4.4.1 Integration of the two vortex system

The dynamics $\dot{\boldsymbol{y}} = B(\boldsymbol{y}) \nabla H_{ij}$ (with flow map ϕ_τ^{ij}) affects only vortices i and j and can thus be expressed as

$$\dot{\boldsymbol{x}}_i = \Gamma_i^{-1} \boldsymbol{x}_i \times \nabla_i H_{ij} = \frac{-\Gamma_j}{4\pi} \frac{\boldsymbol{x}_j \times \boldsymbol{x}_i}{1 - \boldsymbol{x}_i \cdot \boldsymbol{x}_j},$$
$$\dot{\boldsymbol{x}}_j = \Gamma_j^{-1} \boldsymbol{x}_j \times \nabla_j H_{ij} = \frac{-\Gamma_i}{4\pi} \frac{\boldsymbol{x}_i \times \boldsymbol{x}_j}{1 - \boldsymbol{x}_i \cdot \boldsymbol{x}_j}.$$

This two-vortex system has Noether momenta expressed by the vector

$$\boldsymbol{J}_{ij} = \Gamma_i \boldsymbol{x}_i + \Gamma_j \boldsymbol{x}_j.$$

Using the Noether momenta we find

$$\dot{\boldsymbol{x}}_i = \frac{-1}{4\pi} \frac{\boldsymbol{J}_{ij}}{1 - \boldsymbol{x}_i \cdot \boldsymbol{x}_j} \times \boldsymbol{x}_i =: \boldsymbol{a} \times \boldsymbol{x}_i, \tag{4.10}$$
$$\dot{\boldsymbol{x}}_j = \frac{-1}{4\pi} \frac{\boldsymbol{J}_{ij}}{1 - \boldsymbol{x}_i \cdot \boldsymbol{x}_j} \times \boldsymbol{x}_j =: \boldsymbol{a} \times \boldsymbol{x}_j. \tag{4.11}$$

Conservation of the Hamiltonian H_{ij} implies the denominators in (4.10)–(4.11) are constant. This implies the vector $\boldsymbol{a}$ is invariant under the two-vortex dynamics.

Using Rodrigues' formula [86], the solution to equation (4.10) is given by

$$\begin{aligned} \boldsymbol{x}_i(\tau) &= \exp(\widehat{\boldsymbol{a}}\tau) \boldsymbol{x}_i(0) = \boldsymbol{x}_i(0) + \frac{\sin a\tau}{a} \widehat{\boldsymbol{a}} \boldsymbol{x}_i(0) + \frac{1 - \cos a\tau}{a^2} \widehat{\boldsymbol{a}}^2 \boldsymbol{x}_i(0) \\ &= \boldsymbol{x}_i(0) + \sin a\tau \, \tilde{\boldsymbol{a}} \times \boldsymbol{x}_i(0) + (1 - \cos a\tau) \left(\tilde{\boldsymbol{a}} (\tilde{\boldsymbol{a}} \cdot \boldsymbol{x}_i(0)) - \boldsymbol{x}_i(0) \right), \end{aligned}$$

where $\widehat{\boldsymbol{a}}$ is the matrix such that $\widehat{\boldsymbol{a}}\boldsymbol{x} = \boldsymbol{a} \times \boldsymbol{x}$, $a = |\boldsymbol{a}|$ and $\tilde{\boldsymbol{a}} = \boldsymbol{a}/a$. The solution to (4.11) follows by substituting $\boldsymbol{x}_j$ for $\boldsymbol{x}_i$.

This flow map presents an explicit formulation of the exact solution to the two-vortex system of vortices i and j. This pairwise solution is therefore a Poisson system with the same bracket as the N-vortex problem that also preserves the *pairwise* Hamiltonian and momenta exactly. A splitting composed of Poisson flows with identical brackets respects the Casimirs of that bracket. The Noether momenta of the N-vortex system may be written as

$$\boldsymbol{J} = \sum_{i=1}^{N} \Gamma_i \boldsymbol{x}_i = \Gamma_i \boldsymbol{x}_i + \Gamma_j \boldsymbol{x}_j + \sum_{k \neq i,j} \Gamma_k \boldsymbol{x}_k = \boldsymbol{J}_{ij} + \sum_{k \neq i,j} \Gamma_k \boldsymbol{x}_k.$$

This implies that the total momenta $\boldsymbol{J}$ are conserved, because the pairwise flows preserve the pairwise momenta $\boldsymbol{J}_{ij}$ and do not modify the other vortices. The Hamiltonian is not conserved exactly, as the evaluation of pair (i, j) perturbs the values of the Hamiltonian terms H_{ik} and H_{jk} for $k \neq i, j$. This is considered in more detail in the following section.

4.4.2 Modified Hamiltonian

For splitting schemes consisting of exactly integrated Poisson flows with the same bracket, the combined map approximates, to an exponentially high order, a Poisson system with the same bracket, but a modified Hamiltonian. Before considering a point-vortex system, let us recall the simpler problem where the dynamics is given by

$$\dot{y} = B(y)\left(H^{[1]}(y) + H^{[2]}(y)\right). \tag{4.12}$$

For the symmetric Strang splitting

$$\Phi_\tau = \phi^{[1]}_{\tau/2} \circ \phi^{[2]}_{\tau} \circ \phi^{[1]}_{\tau/2}$$

the modified dynamics read

$$\dot{\tilde{y}} = B(\tilde{y})\left(H^{[1]}(\tilde{y}) + H^{[2]}(\tilde{y}) + \tau^2 H_3(\tilde{y}) + \tau^4 H_5(\tilde{y}) + \ldots\right).$$

Throughout the present work we will only consider the first correction term, $H_3(\tilde{y})$, corresponding to an $\mathcal{O}(\tau^2)$ modification to the Hamiltonian. For the Strang splitting of (4.12) this term is given by [62, p. 299]

$$H_3 = -\frac{1}{24}\left\{\left\{H^{[2]}, H^{[1]}\right\}, H^{[1]}\right\} + \frac{1}{12}\left\{\left\{H^{[1]}, H^{[2]}\right\}, H^{[2]}\right\}. \tag{4.13}$$

When the splitting contains more than two different flow maps, the modified Hamiltonian is constructed by applying (4.13) repeatedly "from the inside out", as illustrated by the following three-vortex example.

Consider a system consisting of three vortices that are integrated according to the Strang splitting Φ_τ^S of (4.9). We consider an ordering of pairwise interactions such that the splitting reads

$$\Phi_\tau^{S,3} = \phi_{\tau/2}^{12} \circ \phi_{\tau/2}^{02} \circ \phi_\tau^{01} \circ \phi_{\tau/2}^{02} \circ \phi_{\tau/2}^{12}.$$

Note that the innermost map ϕ_τ^{01} is a composition of two successive maps in the definition of (4.9). We first construct the modified Hamiltonian for the inner map $\phi_\tau^{\text{inner}} = \phi_{\tau/2}^{02} \circ \phi_\tau^{01} \circ \phi_{\tau/2}^{02}$ using (4.13), resulting in

$$H^{\text{inner}} = H_{02} + H_{01} + \tau^2 H_3^{\text{inner}} + \dots,$$

with

$$H_3^{\text{inner}} = \frac{-1}{24} \{\{H_{01}, H_{02}\}, H_{02}\} + \frac{1}{12} \{\{H_{02}, H_{01}\}, H_{01}\}.$$

The modified Hamiltonian of the full step is found by applying (4.13) to $\Phi_\tau^{S,3} = \phi_{\tau/2}^{12} \circ \phi_\tau^{\text{inner}} \circ \phi_{\tau/2}^{12}$, which results in

$$H^{[S]} = H_{12} + H_{02} + H_{01} + \tau^2 H_3^{\text{inner}} + \tau^2 H_3^{\text{outer}} + \dots, \tag{4.14}$$

with

$$H_3^{\text{outer}} = \frac{-1}{24} \{\{H_{01} + H_{02}, H_{12}\}, H_{12}\} + \frac{1}{12} \{\{H_{12}, H_{01} + H_{02}\}, H_{01} + H_{02}\} + \dots.$$

Terms arising from the Poisson bracket of H_3^{inner} with H_{12} are of order τ^4 and are subsequently neglected. Combining both second order corrections we find

$$\begin{aligned} H_3 &= H_3^{\text{inner}} + H_3^{\text{outer}} \\ &= \frac{-1}{24} \left(\{\{H_{01}, H_{02}\}, H_{02}\} + \{\{H_{01}, H_{12}\}, H_{12}\} + \{\{H_{02}, H_{12}\}, H_{12}\} \right) \\ &\quad + \frac{1}{12} \left(\{\{H_{02}, H_{01}\}, H_{01}\} + \{\{H_{12}, H_{01}\}, H_{01}\} + \{\{H_{12}, H_{02}\}, H_{02}\} \right) \\ &\quad + \frac{1}{12} \left(\{\{H_{12}, H_{01}\}, H_{02}\} + \{\{H_{12}, H_{02}\}, H_{01}\} \right). \end{aligned} \tag{4.15}$$

The first two lines of (4.15) consist of Poisson brackets that are all of the form

$$T_{k\ell m}^I := \{\{H_{k\ell}, H_{km}\}, H_{k\ell}\} = -\{\{H_{km}, H_{k\ell}\}, H_{k\ell}\}. \tag{4.16}$$

The last line consists of brackets of the form

$$T_{k\ell m}^{II} := \{\{H_{k\ell}, H_{km}\}, H_{\ell m}\} = -\{\{H_{km}, H_{k\ell}\}, H_{\ell m}\} = -T_{km\ell}^{II}. \tag{4.17}$$

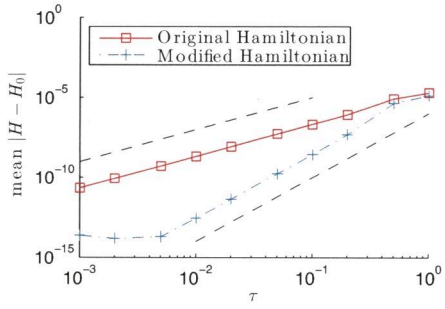


Figure 4.1: Error convergence for the Hamiltonian of a three vortex system in red. In blue is the convergence of the error for the modified Hamiltonian given by (4.14). For reference two dashed lines with slopes τ^2 and τ^4 are given.

With these definitions equation (4.15) is expressed more compactly as

$$H_3 = \frac{-1}{24}\left(-T^I_{021} - T^I_{120} - T^I_{210}\right) + \frac{1}{12}\left(-T^I_{012} - T^I_{102} - T^I_{201} + T^{II}_{120} + T^{II}_{210}\right)$$

We illustrate the accuracy of this correction to the Hamiltonian by simulating a three-vortex system using different time steps. Figure 4.1 compares the L_1 norm of errors in the Hamiltonian and the modified Hamiltonian. We simulate 10 time units starting from arbitrary initial conditions. The second-order convergence of the original Hamiltonian is visible for sufficiently small time steps. The modified Hamiltonian is preserved to fourth order as expected.

In the general N-vortex case, the modified Hamiltonian follows from computing the modified Hamiltonian of the innermost composition and repeatedly working outwards. The ordering C_N consists of $R = \frac{1}{2}N(N-1)$ pairs; we label these $C_N^{[r]}$, with $r \in \{1, \ldots, R\}$. Let $H^{[r]}$ denote the Hamiltonian corresponding to the pair $C_N^{[r]}$. Following the same procedure as for the three-vortex system, only now for a more general number of steps, we find the second-order correction to the Hamiltonian to be

$$H_3 = \sum_{r=1}^{R-1} \sum_{s=r+1}^{R} \left[\frac{-1}{24} \left\{ \left\{ H^{[s]}, H^{[r]} \right\}, H^{[r]} \right\} + \sum_{t=r+1}^{R} \frac{1}{12} \left\{ \left\{ H^{[r]}, H^{[s]} \right\}, H^{[t]} \right\} \right]. \tag{4.18}$$

Many of these terms will be trivial in a large system, as most vortex pairs are disjoint. The other terms in the sum are all distinct, because $t \geq r < s$. This makes it impossible to construct orderings where terms cancel.

When the terms in (4.18) involve only three distinct vortices, they are of a form as in (4.16) or (4.17). Nontrivial terms that involve four distinct vortices

follow either of the following two patterns

$$T^{III}_{k\ell mn} := \{\{H_{k\ell}, H_{km}\}, H_{kn}\} = -\{\{H_{km}, H_{k\ell}\}, H_{kn}\} \tag{4.19}$$
$$T^{IV}_{k\ell mn} := \{\{H_{k\ell}, H_{km}\}, H_{\ell n}\} = -\{\{H_{km}, H_{k\ell}\}, H_{mn}\}. \tag{4.20}$$

Without presenting the derivations, we state that (4.16)–(4.17) and (4.19)–(4.20) are computed as

$$\begin{aligned}
T^{I}_{k\ell m} &= A_{k\ell m} B^2_{k\ell} B_{km} \times \\
&\quad \left[-\Gamma_\ell B_{km} C^2_{\ell km} + (\Gamma_\ell \boldsymbol{x}_m \times \boldsymbol{x}_\ell + \Gamma_k \boldsymbol{x}_k \times \boldsymbol{x}_m) \cdot (\boldsymbol{x}_k \times \boldsymbol{x}_\ell)\right] \\
T^{II}_{k\ell m} &= A_{k\ell m} B_{k\ell} B_{km} B_{\ell m} \times \\
&\quad \left[(\Gamma_\ell B_{km} - \Gamma_m B_{k\ell})\, C^2_{lkm} + (\Gamma_m \boldsymbol{x}_m + \Gamma_\ell \boldsymbol{x}_\ell) \times \boldsymbol{x}_k \cdot (\boldsymbol{x}_m \times \boldsymbol{x}_\ell)\right] \\
T^{III}_{k\ell mn} &= A_{k\ell m} B_{k\ell} B_{km} B_{kn} \Gamma_n \times \\
&\quad [B_{k\ell} C_{\ell km} C_{\ell kn} + B_{km} C_{\ell km} C_{mkn} + (\boldsymbol{x}_m \times \boldsymbol{x}_\ell) \cdot (\boldsymbol{x}_k \times \boldsymbol{x}_n)] \\
T^{IV}_{k\ell mn} &= A_{k\ell m} B_{k\ell} B_{km} B_{\ell n} \Gamma_n \times \\
&\quad [B_{k\ell} C_{\ell km} C_{k\ell n} + (\boldsymbol{x}_k \times \boldsymbol{x}_m) \cdot (\boldsymbol{x}_\ell \times \boldsymbol{x}_n)],
\end{aligned}$$

where

$$\begin{aligned}
A_{k\ell m} &= \frac{-\Gamma_k \Gamma_\ell \Gamma_m}{(4\pi)^3} \\
B_{k\ell} &= \frac{1}{1 - \boldsymbol{x}_k \cdot \boldsymbol{x}_\ell} \\
C_{k\ell m} &= \boldsymbol{x}_k \cdot \boldsymbol{x}_\ell \times \boldsymbol{x}_m.
\end{aligned}$$

From this it follows that the error in the Hamiltonian is dominated by close approaches between vortices, as there the denominator in $B_{k\ell}$ approaches zero. The magnitudes of both $A_{k\ell m}$ and $C_{k\ell m}$ are obviously bounded.

Long time conservation of the Hamiltonian by symplectic methods can be rigorously shown in the case of analytic Hamiltonian H, but is often observed in practice for more general Hamiltonians. For point vortices, the Hamiltonian has singularities when two vortices coincide. The motion of a single pair could inadvertently place one vortex in close proximity to another, effectively a "numerical collision". In practice we have not encountered this. This is only problematic if that vortex pair is evaluated before either of the coincident vortices are moved by a different vortex pair interaction.

4.4.3 Numerical results

Our primary purpose for developing the explicit and symplectic integrator outlined in the present work is to allow performing numerical experiments on moderate to large vortex systems to verify hypothesized statistical behaviour. These simulations must be run over a long time period to yield meaningful

statistics. In this section we demonstrate the approximate energy conservation and exact momentum conservation of the splitting schemes in different settings. We compare the results of the Strang splitting to results for a symplectic, implicit method developed by Vankerschaver & Leok [140] and to a fourth order explicit Runge-Kutta scheme. Both implementations can be found at `https://github.com/jvkersch/hopf_vortices` along with two other implicit schemes.

We investigate the conservation of energy and momentum in a number of different configurations. For each case we compare energy and momentum errors given by $\varepsilon_H(t) = |H(t) - H(0)|$ and $\varepsilon_J(t) = ||\boldsymbol{J}(t) - \boldsymbol{J}(0)||$. We have repeated the experiments of Vankerschaver & Leok [140], but observe that in most of these configurations symmetries play an important role. This symmetry is broken by the splitting methods, resulting in a poorer performance than methods that maintain the symmetry. We also compare results for an arbitrary initial condition at a given energy level in Section 4.4.3. This is a more practical test case for engineering applications and statistical mechanics; it is in fact the setting in which we use this integrator in other work [106].

Collapsing vortices

For certain initial conditions three or more vortices will collapse onto a single point in finite time, while the energy remains bounded. Such initial conditions with three vortices have been studied by Kidambi & Newton [80] and with four vortices by Sakajo [130]. We simulate the same three vortex system as Vankerschaver & Leok [140].

The vortex circulations are $\Gamma_1 = \Gamma_2 = 1$ and $\Gamma_3 = -\frac{1}{2}$. The vortices start at the vertices of a triangle with lengths $l_{12} = \frac{1}{2}\sqrt{2}$, $l_{23} = \frac{1}{2}\sqrt{2}$ and $l_{13} = 1$. These initial conditions result in a collapse of the three vortices onto a single point at $T^- = 4\pi(\sqrt{23} - \sqrt{17}) \approx 8.4537$. At this time the equations of motion become undefined. The numerical methods will not collapse exactly, due to the numerical error – and in the case of the methods in [140], due to regularization of the dynamics. Instead there will be a moment that the vortices approach each other closely. This event is repeated periodically.

We perform this simulation with time steps $\tau = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$ to illustrate how the behaviour changes. Conservation of the Hamiltonian over a short time – enough to show the first collapse event – is illustrated in the left-hand panel of Figure 4.2. As the time step is reduced, the magnitude of the maximum energy error during this part of the simulation does not change, but the length of time over which there is a significant error is greatly reduced by using a smaller time step. In this pathological configuration, the lack of analyticity of the Hamiltonian negates backward error analysis and its prediction of second-order convergence.

The moment of the closest approach is indicated by the largest error in

the Hamiltonian. With smaller time steps this instant approaches the correct time of the true collapse event.

The right-hand panel of Figure 4.2 demonstrates how the energy changes over longer time, including four more near-collapse events. Between the near-collapses, the energy consistently returns close to its initial value. Note also that the results presented by Vankerschaver & Leok [140] include some regularization. This in fact slows down the dynamics around the near-collapses so much that the fifth event does not occur before the end of the simulation at $T = 500$. We should note that the behaviour of this system with a repetition

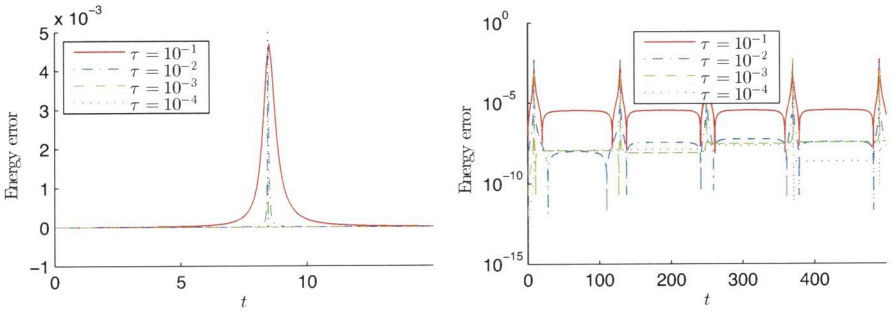


Figure 4.2: Energy error for a system of $N = 3$ collapsing vortices for various time step sizes. Initial conditions that lead to a collapse of the three vortices onto a single point (in finite time) are given by Kidambi & Newton [80]. The exact time of this collapse is indicated by a black dotted line.

of near-collapses does not occur despite numerical error, but because of it. The exact solution becomes undefined at the (first) collapse. When the error is reduced by using a higher order method (results not shown), the dynamics leading up to the first event are more accurate, resulting in a much closer approach between the vortices. This causes a larger energy error after the event than in the Strang splitting.

Stable vortex ring

A ring of N equidistantly placed vortices of equal strength rotates stably around its centre provided $N \leq 7$ and provided that the latitude of the ring (assumed parallel to the equator) is above a certain critical value [121].

We simulate a stable configuration with $N = 6, \Gamma = \frac{1}{N}$ and latitude $\frac{1}{2}\pi - .4$ for 1000 time units. Figure 4.3 compares the energy and momentum errors against those for the Hopf and Runge-Kutta integration methods. All simulations use a time step of 0.05.

The splitting method only approximately conserves energy, but the error is bounded. The momentum is conserved to machine precision throughout the

simulation. Due to the rotational symmetry of this configuration, a method that respects this symmetry will easily exhibit energy conservation. The splitting method does not respect the symmetry due to the influence of the ordering of pairs. Consequently its energy conserving quality in this rather specialized test case is inferior to that of the Hopf integrator. We have also performed the experiment with McLachlan's 6th order composition [99] of the Strang splitting. This shows energy conservation to the same degree as the Hopf method.

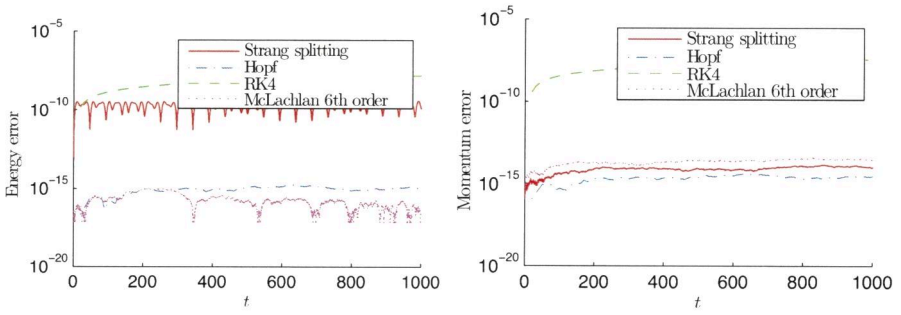


Figure 4.3: Energy (left) and momentum (right) error for the stable vortex ring of Polvani & Dritschel [121] with $N = 6$ vortices. The energy oscillates about a fixed mean, with bounded error. The momentum is conserved to machine precision.

Von Kármán vortex streets

Another relative equilibrium is that of the Von Kármán vortex streets presented by Chamoun et al. [24]. This configuration consists of two staggered rings of N_r vortices placed at latitudes $\theta = \pm\theta_0$ plus one vortex at each pole.

We take $N_r = 5$ vortices per ring, each with $\Gamma = \pm 1$, placed at $\theta = \pm\frac{\pi}{6}$, respectively. The polar vortices satisfy $\Gamma_n = -\Gamma_s = \frac{1}{2}$. This configuration rotates about the z-axis with a period of $T = 10.85$. We simulate this system with a time step of $\tau = .5$ for 10 000 time units.

The splitting scheme and Hopf method both conserve the momenta exactly by construction, and this is reflected in the simulation results. The error in the energy remains bounded throughout the simulation at an accuracy that is somewhat better than that of the Hopf integrator.

This configuration is believed to be inherently unstable [140], making the symmetry of vital importance. The splitting scheme breaks this symmetry, yet the error remains bounded.

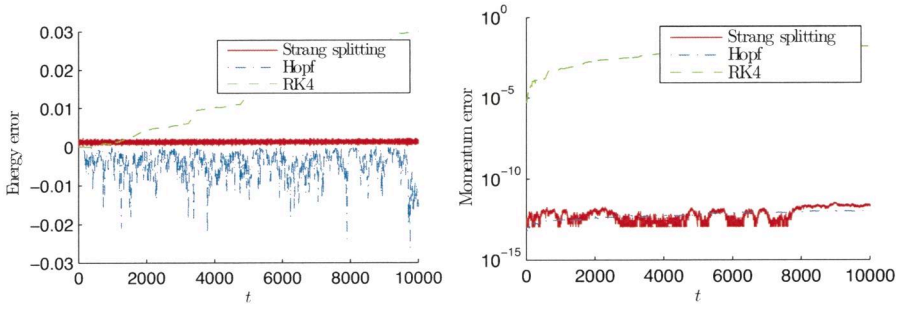


Figure 4.4: Energy (left) and momentum (right) error for the Von Kármán vortex street with two rings of 5 vortices and a vortex at either pole, giving $N = 12$ vortices in total. The energy oscillates rapidly about a slowly varying mean. The momentum is initially conserved to machine precision, but accumulation of arithmetic errors eventually leads to a small drift.

Generic initial conditions

In this final test case we consider a system with 48 vortices, eight with circulation $\Gamma = \pm 1$ and 40 with $\Gamma = \pm\frac{1}{5}$ with equal numbers positive and negative. The initial configuration is drawn randomly from the set of all states with a given energy level and zero angular momentum. We use the same initial condition for the different methods. We consider both large negative energy ($H = -2$), resulting in a strong clustering of like-signed vortices [116, 21], and large positive energy ($H = 2$), leading to a well-mixed configuration with close approaches between opposite signed-vortices. These configurations are extreme in the sense that the specified energy levels lie close to the tails in the distribution of all attainable energy states for this number of vortices with these circulations.

Figure 4.5 compares the energy error over a short time for the Strang splitting method against the Hopf method [140], a fourth order Runge-Kutta scheme and an implicit midpoint method [62], all with time step $\tau = 0.0001$. There is a remarkable difference in the performance of the integrators at positive and negative energies. At negative energies, the Runge-Kutta scheme conserves energy accurately while it exhibits rapid error growth in the positive energy simulation. The Hopf integrator performs poorly at negative energy, and fails to converge at positive energy, even for this modest time step size. In the negative energy simulation, the Strang splitting and implicit midpoint method have roughly the same accuracy. For positive energy, the implicit midpoint is more accurate.

We also compare the results over a longer time period with an increased time step of $\tau = 0.1$. All three implicit methods considered by Vankerschaver

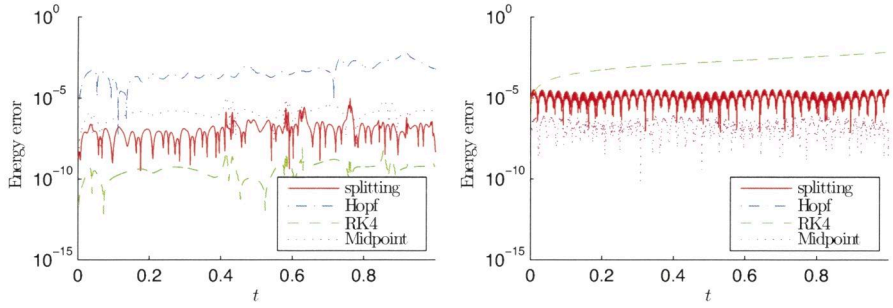


Figure 4.5: Energy error in simulations with negative (left) and positive (right) energies. The Strang splitting is compared against a fourth order Runge-Kutta (RK4) scheme. The Hopf fibration method fails to converge in the positive energy case.

& Leok [140] fail to converge with this time step. Therefore Figure 4.6 shows only the results for the Strang splitting and the fourth order Runge-Kutta scheme. In both cases the Strang splitting shows a smaller energy error. The momentum error (not shown) is within machine precision for the Strang splitting and of order one for the Runge-Kutta approach.

We attribute the higher accuracy of the splitting scheme in the positive energy case to the formation of long lived coherent structures of like-signed vortices. With *exact* conservation of the Hamiltonian, like-signed vortices can only have a close approach if there is a simultaneous close approach between vortices of opposite sign. The coherent structures make the close approach of opposite-signed vortex pairs less likely.

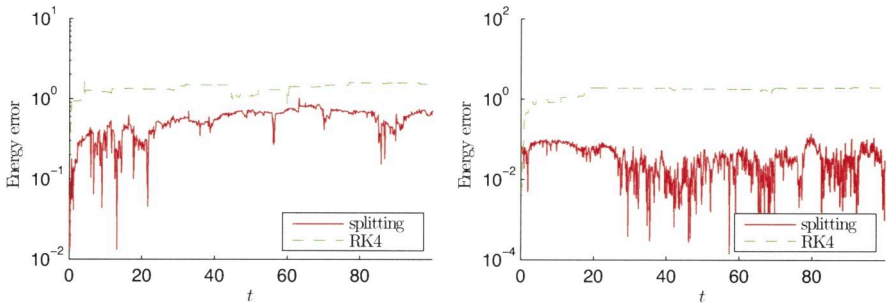


Figure 4.6: Energy error in simulations with negative (left) and positive (right) energies. The Strang splitting is compared against a fourth order Runge-Kutta (RK4) scheme.

Table 4.2: Round-robin scheme for ordering vortex pairs in a four-vortex system.

round	**0**	**1**	**2**
pair	$0-3$	$1-3$	$2-3$
	$1-2$	$2-0$	$0-1$

4.5 Parallelization

In this section we will discuss several different parallelization algorithms for point vortex dynamics. All of them are based on rearranging the pairwise flow maps of (4.9) and grouping together subsets of the composition. By choosing groups that consist of disjoint pairs, the order of evaluation does not affect the result and the pairs may be evaluated in parallel, without loss of accuracy.

As an illustration, let us consider a system with just 4 vortices, labelled[1] $i = 0, 1, 2, 3$. The ordering C must contain all 6 possible vortex pairings once. This is similar to *round robin* tournament scheduling, where all competitors play each other once. Such a scheme is presented in Table 4.2, each round is performed from top to bottom before moving onto the next round. The ordering is constructed by fixing the last vortex, number 3, in position, while the other vortices rotate over the remaining positions. Note that during each "round" each vortex occurs only once, this will be important later on when discussing parallelizations .

The Strang splitting for this ordering is written as

$$\Phi^{\mathrm{S}}_{2\tau} = \phi^{03}_{\tau} \circ \phi^{12}_{\tau} \circ \phi^{13}_{\tau} \circ \phi^{20}_{\tau} \circ \phi^{23}_{\tau} \circ \phi^{01}_{\tau} \circ \phi^{01}_{\tau} \circ \phi^{23}_{\tau} \circ \phi^{20}_{\tau} \circ \phi^{13}_{\tau} \circ \phi^{12}_{\tau} \circ \phi^{03}_{\tau}.$$

For notational convenience we use a time step of 2τ. Because this ordering is symmetrical and each of the ϕ^{ij}_{τ} is self-adjoint, the resulting method $\Phi^{\mathrm{S}}_{2\tau}$ is a second order accurate method.

The more general ordering of N interacting vortices can be arranged in the same way. We write the s-th pair of round r in the ordering C_N as $C^{sr}_N = (A_{sr}, B_{sr})$. Following the same pattern as for the four-vortex system results in the expressions

$$A_{sr} = \mathrm{mod}(s + r, N - 1)$$

$$B_{sr} = \begin{cases} N - 1 & \text{for } s = 0 \\ \mathrm{mod}(N - 1 - s + r, N - 1) & \text{for } s \neq 0. \end{cases}$$

A ten-vortex system exemplifies this ordering in Table 4.3.

[1] For convenience in the ensuing modular arithmetic, we switch to indexing from zero.

Table 4.3: Round-robin scheme for ordering vortex pairs in a ten-vortex system.

round	0	1	2	3	4	5	6	7	8
	$0-9$	$1-9$	$2-9$	$3-9$	$4-9$	$5-9$	$6-9$	$7-9$	$8-9$
	$1-8$	$2-0$	$3-1$	$4-2$	$5-3$	$6-4$	$7-5$	$8-6$	$0-7$
pair	$2-7$	$3-8$	$4-0$	$5-1$	$6-2$	$7-3$	$8-4$	$0-5$	$1-6$
	$3-6$	$4-7$	$5-8$	$6-0$	$7-1$	$8-2$	$0-3$	$1-4$	$2-5$
	$4-5$	$5-6$	$6-7$	$7-8$	$8-0$	$0-1$	$1-2$	$2-3$	$3-4$

4.5.1 Complete parallelization

The arrangement of pairwise vortex interactions into "rounds" in Table 4.2 is not just convenient for notation. Let us denote the composition for each round by $\psi_\tau^{ij,k\ell} = \phi_\tau^{ij} \circ \phi_\tau^{kl}$, for i, j, k, ℓ all different. In this notation the first round of Table 4.2 is the composition $\psi_\tau^{03,12} = \phi_\tau^{03} \circ \phi_\tau^{12}$. But as the evolution of vortex pair $(0,3)$ is independent of that of pair $(1,2)$, the order of the operations is irrelevant, i.e. ϕ_τ^{03} and ϕ_τ^{12} commute. This means that while evolving the system, these two vortex pairs can be evaluated simultaneously, in parallel. Let us stress this fact by using the notation $\chi_\tau^{ij,kl}$ to denote the time-τ flow map of the evolution of vortex pairs (i,j) and (k,l) in *either* order. The Lie-Trotter splitting method where both pairs for each round are evaluated in parallel is then written as

$$\Phi_\tau^{\mathrm{LT}\parallel} = \chi_\tau^{03,12} \circ \chi_\tau^{13,20} \circ \chi_\tau^{23,01},$$

and the Strang splitting

$$\Phi_{2\tau}^{\mathrm{S}\parallel} = \chi_\tau^{03,12} \circ \chi_\tau^{13,20} \circ \chi_\tau^{23,01} \circ \chi_\tau^{23,01} \circ \chi_\tau^{13,20} \circ \chi_\tau^{03,12}.$$

Because each of the $\chi_\tau^{ij,kl}$ compositions is self-adjoint, this method is again second-order accurate. Note that the order of the underlying pairwise interactions is now no longer necessarily symmetric.

When integrating a system with four vortex, each round of the round-robin scheme contains two pairwise interactions that can be performed in parallel. More generally, this scheme allows P processors to evaluate a system with $2P$ vortices. But each processor only evaluates a single vortex pair interaction per round, meaning there is a lot of communication relative to the amount of work done each round.

4.5.2 Reducing communication

It is not necessary for each round to be finished completely before starting evaluating the next. For the s-th vortex pair of round r, (A_{sr}, B_{sr}), to be eval-

Table 4.4: Reduced communication round-robin scheme for ordering vortex pairs in a ten-vortex system.

round	0	1	2	3	4	5	6	7	8
	$0-9$	$4-9$	$8-9$	$3-9$	$7-9$	$2-9$	$6-9$	$1-9$	$5-9$
	$1-8$	$0-8$	$0-7$	$8-7$	$8-6$	$7-6$	$7-5$	$6-5$	$6-4$
pair	$2-7$	$1-7$	$1-6$	$0-6$	$0-5$	$8-5$	$8-4$	$7-4$	$7-3$
	$3-6$	$2-6$	$2-5$	$1-5$	$1-4$	$0-4$	$0-3$	$8-3$	$8-2$
	$4-5$	$3-5$	$3-4$	$2-4$	$2-3$	$1-3$	$1-2$	$0-2$	$0-1$

uated, it is only necessary to wait for these two vortices to have been evaluated in the previous round $r-1$. The evaluation of the remaining vortex pairs commutes with the evaluation of (A_{sr}, B_{sr}).

If thread s in round r has to wait only for vortices A_{sr} and B_{sr} to be done in round $r-1$, it is beneficial to choose the ordering such that one of the two vortices is evaluated on the same thread as in the previous round. This means that each thread has to wait for only one other vortex pair of the previous round. We construct such an ordering based on the construction used in the previous section. Again the vortex $2P-1$ is kept fixed in place, but now the other vortices rotate through $P-1$ positions. In the ten-vortex example, this can be seen as jumping straight to round 4 in Table 4.3 after the first round. In doing so, all vortices on the right under the fixed vortex move to the left, but in reversed order. By subsequently reversing the order of these pairs in all odd rounds, each thread needs to wait only for one other vortex during each round. The resulting ordering $\tilde{C}_N^{sr} = \left(\tilde{A}_{sr}, \tilde{B}_{sr}\right)$ is given by

$$\left(\tilde{A}_{sr}, \tilde{B}_{sr}\right) = \begin{cases} (A_{s\tilde{r}}, B_{s\tilde{r}}) & \text{for } r \text{ even} \\ (A_{s\tilde{r}}, B_{s\tilde{r}}) & \text{for } r \text{ odd}, s = 0 \\ (B_{\tilde{s}\tilde{r}}, A_{\tilde{s}\tilde{r}}) & \text{for } r \text{ odd}, s \neq 0, \end{cases}$$

where $\tilde{r} = \mathrm{mod}\,(r(P-1), N-1)$ and $\tilde{s} = P - s$. An example ordering with 10 vortices is presented in table 4.4.

4.5.3 Hierarchical parallelization

An efficient parallelization finds a balance between reducing the work load per processor and reducing the time required for communication between threads. Without parallelization there is no communication time, but the workload per thread is largest. The parallelization scheme in Section 4.5.1 represents the other extreme – where communication time dominates the total computation time. In this section we develop a scheme that has adjustable

parallelization. Both the "complete parallelization" and the non-parallel computation are special cases of this scheme.

The method we develop here is constructed as a hierarchy of methods that act on a hierarchical system of *leagues* of vortices. Given a number of parallel threads P and a number of vortices N to be evaluated, we place $L = \frac{N}{2P}$ vortices in each lowest-level league in the hierarchy. If the fraction $\frac{N}{2P}$ is not integer, we introduce a number of dummy vortices with zero circulation to increase N such that it is a multiple of $2P$.

We also define a factorization $P = p_1 \times p_2 \times \ldots \times p_n$. We will use the prime factorization, but depending on system architecture a different factorization may be desirable. Given this factorization, we construct a hierarchy of leagues as follows: there are $2p_1$ level-one leagues, each consisting of p_2 level-two leagues, each consisting of p_3 level-three leagues and so on. The level-n leagues are the lowest level and consist of L vortices each.

We denote a level-m league in this tree by $\mathcal{L}_m^{\boldsymbol{k}}$, where the vector $k = (k_1, k_2, \ldots, k_m) \in \mathbb{R}^m$ denotes the ancestry of the league. In other words, k_m denotes the current child of parent league k_{m-1} of grandparent k_{m-2} and so forth.

The hierarchical splitting uses the fact that each vortex pair is either an interaction *between* two leagues, with one vortex from each of the two leagues, or *within* a league, with both vortices from that same league. Looking at the top level first, we see that this means that we have interactions between the $2p_1$ groups, and interactions within each of these groups. One way to do this would be by first evaluating all $p_2(2p_2 - 1)$ possible combinations of level-one leagues according to an ordering as in Section 4.5.2 and then evaluating the interactions within each level-one league separately. This can, however, have the unfortunate effect of wasted computation time if one of the factors of the factorization is odd. Instead we evaluate the interaction between level-one leagues according to the ordering of Section 4.5.2, but we omit the final round. The interactions between these omitted leagues are combined with their interior interactions. For the top level, this forms a system of N/p_1 vortices that has to be evaluated completely.

Let $\Lambda_1[(k, \ell)]$ denote the interaction between all vortices in level-one league $\mathcal{L}_1^k$ and those in $\mathcal{L}_1^\ell$ and let the composition $\Phi_1[(k, \ell)]$ denote the evaluation of *all* interactions within or between level-one leagues $\mathcal{L}_1^k$ *and* $\mathcal{L}_1^\ell$. The Lie-Trotter splitting is then given by (we drop the subscript indicating the time step for notational convenience)

$$\Phi_0^{\mathrm{LT}} = \prod_{r=0}^{2p_1-3} \prod_{s=0}^{p_1-1} \Lambda_1[\tilde{C}_{2p_1}^{sr}] \circ \prod_{s=0}^{p_1-1} \Phi_1[\tilde{C}_{2p_1}^{s,2p_1-2}], \tag{4.21}$$

This composition is represented graphically as the top level of the tree diagram in Figure 4.7. The compositions $\Lambda_1[(k, \ell)]$ and $\Phi_1[(k, \ell)]$ both follow a

Table 4.5: Vortex interaction pairs between vortices $\{0,\dots,3\}$ and $\{4,\dots,7\}$.

round	0	1	2	3
pair	$0-4$	$0-5$	$0-6$	$0-7$
	$1-5$	$1-6$	$1-7$	$1-4$
	$2-6$	$2-7$	$2-4$	$2-5$
	$3-7$	$3-4$	$3-5$	$3-6$

recursive definition detailed below.

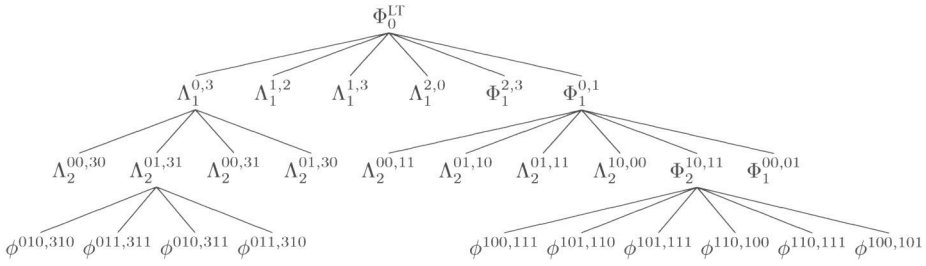


Figure 4.7: Tree representation of the hierarchical parallelization with $L = 2$ and $P = 4$, using prime factorization: $p_1 = p_2 = 2$. Only a few branches are expanded for clarity.

Interactions between leagues

The flow map $\Lambda_q[(\boldsymbol{k},\boldsymbol{\ell})]$ represents all possible interactions with one vortex in league $\mathcal{L}_q^{\boldsymbol{k}}$ and one in league $\mathcal{L}_q^{\boldsymbol{\ell}}$. Using the hierarchical ordering of the vortices, we evaluate all such pairings by evaluating all possible combination between leagues one level down. This is represented graphically by the branches on the left in Figure 4.7.

The ordering D_{2P} contains all possible pairs (i,j) with $i \in \{0,\dots,P-1\}$ and $j \in \{P,\dots,2P-1\}$.

$$D_{2P}^{sr} = (s, P + \mathrm{mod}(s+r, P)), \quad s,r \in \{0,\dots,P-1\}.$$

An example with $P = 4$ is presented in Table 4.5.

With this, we define the recursive definition for the interaction between leagues

$$\Lambda_{q-1}[(\boldsymbol{k},\boldsymbol{\ell})] = \prod_{r=0}^{p_q-1} \prod_{s=0}^{p_q-1} \Lambda_q[I_{\boldsymbol{k},\boldsymbol{\ell}}^{p_q}(D_{2p_q}^{sr})], \tag{4.22}$$

where the mapping

$$I^{p_q}_{\boldsymbol{k},\boldsymbol{\ell}}(k) = \begin{cases} (\boldsymbol{k}, k) & \text{for } k < p_q \\ (\boldsymbol{l}, k - p_q) & \text{for } k \geq p_q \end{cases} \tag{4.23}$$

associates each of the vortices from an ordering with $2p_q$ vortices with children in the groups $\mathcal{L}^{\boldsymbol{k}}_{q-1}$ and $\mathcal{L}^{\boldsymbol{\ell}}_{q-1}$. The lowest level interaction between leagues is given by

$$\Lambda^{\boldsymbol{k},\boldsymbol{\ell}}_n = \prod_{r=0}^{L-1} \prod_{s=0}^{L-1} \phi^{I^L_{\boldsymbol{k}\boldsymbol{\ell}}(D^{sr}_{2L})}. \tag{4.24}$$

Note that we also use the ancestry to enumerate the vortices, rather than linear indexing.

Interactions within leagues

The definition of $\Phi_1[(k,\ell)]$ is given recursively, so it follows the same pattern as $\Phi_{q-1}[(\boldsymbol{k},\boldsymbol{\ell})]$ with $\boldsymbol{k},\boldsymbol{\ell} \in \mathbb{R}^{q-1}$, that is

$$\Phi_{q-1}[(\boldsymbol{k},\boldsymbol{\ell})] = \prod_{r=0}^{2p_q-3} \prod_{s=0}^{p_q-1} \Lambda_q[I^{p_q}_{\boldsymbol{k}\boldsymbol{\ell}}(\tilde{C}^{sr}_{2p_q})] \circ \prod_{s=0}^{p_q-1} \Phi_q[I^L_{\boldsymbol{k}\boldsymbol{\ell}}(\tilde{C}^{s,2p_q-2}_{2p_q})], \tag{4.25}$$

for $q = 1, 2, \ldots, n$. The mapping $I^{p_q}_{\boldsymbol{k}\boldsymbol{\ell}}$ is the same as in (4.23). Note that the definition of Φ^{LT}_0 in (4.21) is in fact equivalent to (4.25) for $q = 1$.

At the lowest level, the league $\mathcal{L}^{\boldsymbol{k}}_n$ no longer consists of leagues, but of L point vortices. So when considering the interaction of all vortices within two lowest-level leagues $\mathcal{L}^{\boldsymbol{k}}_n$ and $\mathcal{L}^{\boldsymbol{\ell}}_n$, this constitutes evaluating a $2L$ point vortex system.

$$\Phi_n[(\boldsymbol{k},\boldsymbol{\ell})] = \prod_{r=0}^{2L-2} \prod_{s=0}^{L-1} \phi^{I^{p_n}_{\boldsymbol{k}\boldsymbol{\ell}}(C^{sr}_L)}. \tag{4.26}$$

By (recursively) substituting equations (4.22)–(4.26) into (4.21) we find a Lie-Trotter splitting for the system with $N = 2LP$ vortices.

Symmetric splitting

To construct a symmetric splitting we need to compose the splitting Φ^{LT}_0 of (4.21) with its adjoint $\Phi^{\mathrm{LT},*}_0$ as

$$\Phi^{\mathrm{S}}_0 = \Phi^{\mathrm{LT}}_0 \circ \Phi^{\mathrm{LT},*}_0.$$

The adjoint of the Lie-Trotter splitting follows from a reversal of the order of the operators – insofar as this is necessary – and taking the adjoint of each of

the interior operators.

$$\Phi_0^{\mathrm{LT},*} = \prod_{s=0}^{p_1-1} \Phi_1^*[\tilde{C}_{2p_1}^{s,2p_1-2}], \circ \prod_{r=2p_1-3}^{0} \prod_{s=0}^{p_1-1} \Lambda_1^*[\tilde{C}_{2p_1}^{sr}]$$

The adjoint of $\Phi_{q-1}[(\boldsymbol{k},\boldsymbol{\ell})]$ for $q \in \{2,\ldots,n\}$ of (4.25) is taken in a similar fashion to be

$$\Phi_{q-1}^*[(\boldsymbol{k},\boldsymbol{\ell})] = \prod_{s=0}^{p_q-1} \Phi_q^*[I_{\boldsymbol{k}\boldsymbol{\ell}}^L(\tilde{C}_{2p_q}^{s,2p_q-2})], \circ \prod_{r=2p_q-3}^{0} \prod_{s=0}^{p_q-1} \Lambda_q^*[I_{\boldsymbol{k}\boldsymbol{\ell}}^{p_n}(\tilde{C}_{2p_q}^{sr})]$$

For the interactions between leagues given by $\Lambda_q[(\boldsymbol{k},\boldsymbol{\ell})]$ in (4.22) we need only reverse the order of the outer product and take the adjoint of the Λ_{q+1} inside. The adjoint of the lowest level operators in (4.24) and (4.26) is achieved by reversing the order of the outside product over r.

The complete parallelization of Section 4.5.1 follows from the choice $L = 1$, $N = 2P = 2p_1$, whereas the non-parallel scheme follows from $P = 1$, $N = 2L$.

4.5.4 Implementation details

All simulations were performed on a desktop Macintosh MacPro running OS X 10.9.5. The system has two Intel Xeon 2.93 GHz processors with six cores each. The system has 32 GB shared memory, 12 MB L3 cache (per processor) and 256 KB L2 cache (per core). The programming code was written in C, compiled into stand-alone applications using Matlab's mex with llvm-gcc-4.2. The motivation for this compiling strategy is to allow for easier transfer of data to Matlab, which was used for all post-processing and data-analysis purposes. All source files are available at `https://github.com/KeithWM/poissonpv`.

4.5.5 Timing experiments

To investigate the practical use for the different splitting schemes and vortex orderings we perform several experiment measuring the required time for different simulations. In all cases the configuration consists of eight strong vortices, four positive ($\Gamma = 1$) and four negative ($\Gamma = -1$), with the remaining vortices weaker with circulation $\pm\frac{1}{5}$ in equal numbers. The initial conditions are chosen such that the total energy and momentum are zero. All timing experiments are performed five times independently, to confirm that the results are not influenced by external factors.

The efficiency of the parallelized splitting method is best represented by studying the scaling of the method to large number of vortices while linearly increasing the number of threads. This implies each thread always operates

on the same number of vortices and consequently the time taken for each round does not change. The number of rounds does change when the number of vortices is increased, but this is only a linear increase.

For these simulations we use a Strang splitting with a time step of $\tau = 0.001$ and simulate up to only $T = 0.01$. This short time makes the timing results for small systems somewhat noisy, but it means that simulating a large system remains feasible, even when using only a few threads. We present results for the splitting as detailed in Section 4.5.3 as well as results for the same principle with a small modification to reduce communication.

In Figure 4.8 we display the time required to simulate a systems with 64 (left) and 1024 (right) vortices per thread, when varying the number of threads from 1 to 12. There is indeed only a linear increase in the workload. In the case with 64 vortices per thread, the benefits of parallelization outweigh the costs roughly from a system size of 192 onwards. With 1024 vortices per thread there is an immediate benefit to parallelization.

The different orderings appear of little to no effect on the speed of computation. We attribute this to the fact that the intended improvement by rearranging the ordering would arise only in situations where there is a significant difference in the time required to evaluate the different pairwise interactions. This is certainly not the case when each thread is evaluating many pairwise interactions each round.

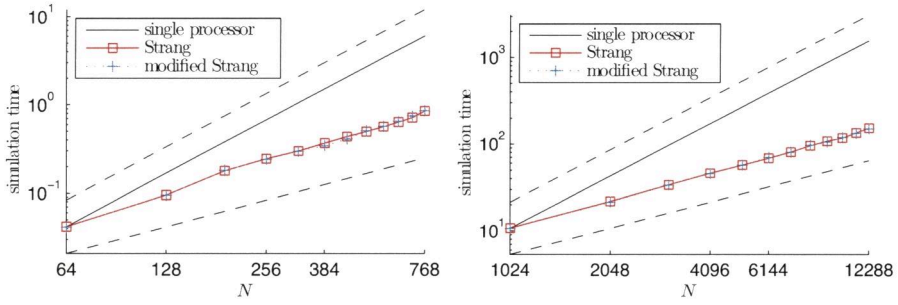


Figure 4.8: Wall clock time required plotted against number of vortices N, keeping the number of vortices per thread fixed (64 on left, 1024 on right). A solid black line represents the time needed to perform the simulations on a single core and dashed black lines represent $\mathcal{O}(N)$ and $\mathcal{O}(N^2)$ scaling.

The same set-up is used to compare the speeds using different combinations of system size and thread counts (N and P respectively). The number of vortices per group is then chosen to be $L = \lceil \frac{N}{2P} \rceil$, leading to the possible introduction of dummy vortices. This will however have minimal effect on computation time in large systems.

Simulation times for system sizes ranging from $N = 24$ to $N = 24000$ and

thread counts from one to twelve are compared in Figure 4.9. The times are normalized by N^2 representing an estimation of the simulation time per vortex interaction. For increasing system size this quantity approaches a fixed constant for a fixed number of threads. This is clearly visible for the smaller thread counts. A general trend that larger systems are evaluated fastest using more threads is apparent, but there are some notable exceptions. Most standing out is the speed of the 12 thread computation of a system with 180 vortices. This is probably a result of the efficiency of evaluating 8 vortices per group due to memory management.

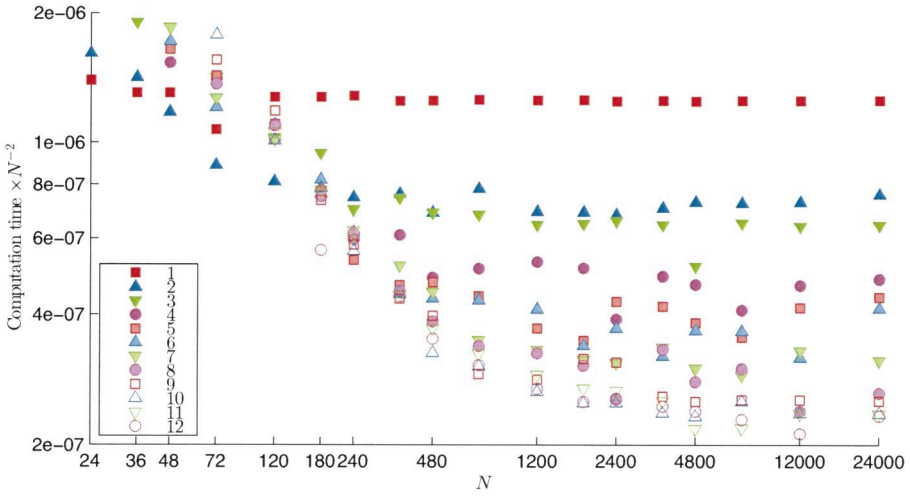


Figure 4.9: Wall clock time plotted against number of vortices N, using different numbers of threads. The times are normalized by N^2, thus estimating the workload per vortex pair. All simulations are performed using a Strang splitting.

We investigate the error convergence for the different orderings in a large system with 360 vortices – 8 strong and 352 weak as before. We run a short simulation up to $T = 0.1$ time units. The time step used is varied from $\tau = 10^{-2}$ down to $\tau = 10^{-5}$. We consider the energy error $\varepsilon_H(t) = |H(t) - H(0)|$ and take the mean $\bar{\varepsilon}_H = \frac{\tau}{T} \sum_{0=1}^{I} \varepsilon_H(i\tau)$ over the simulation interval.

The mean energy error is plotted in Figure 4.10; it is compared against the time step in the left-hand panel, and against the simulation time on the right. As expected the Lie-Trotter splittings show first order convergence and the Strang splittings second order. The heuristic modification made to the separated ordering has little effect on accuracy or speed.

Over the range of time steps considered, Strang splitting outperforms Lie-Trotter splitting not only in terms of accuracy for a given time step, but also

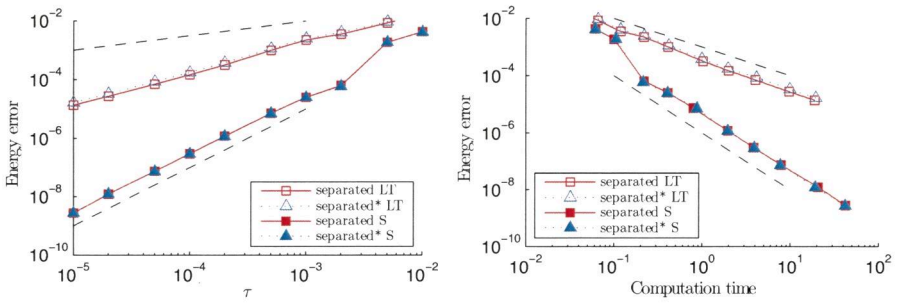


Figure 4.10: Comparison between different orderings in the Lie-Trotter (LT) and Strang (S) splitting schemes for a system with 360 vortices. On the left the error in energy is compared against different time steps, on the right it is compared against simulation time required. All simulations use eight parallel threads. Black dashed lines represent first and second order convergence.

in terms of accuracy against computational cost. This suggests higher order methods could be even more efficient. This is investigated by comparing fourth and sixth order methods against the first order Lie-Trotter and second order Strang splittings for the same problem. The higher order methods follow from a composition of a number of Strang steps of different sizes [62]. For both fourth and sixth order we consider the methods with the minimal number of stages presented by Yoshida [143], as well as the methods of same order but with smaller error coefficients found by McLachlan [99].

The results are shown in Figure 4.11, again with energy error versus time step on the left and energy error versus computational time on the right. All methods exhibit the expected error convergence, albeit for a limited range of time steps. The methods due to McLachlan [99] have considerably smaller error for the same time step. When considering the benefit for the same computational load the difference is much reduced, but still in favour of McLachlan's schemes.

4.6 Conclusion and outlook

Solutions to ideal fluid flow with a singular measure vorticity field result in a Poisson system describing the motion of the vortex centres. By splitting the Hamiltonian of such a point vortex system into the interactions of individual vortex pairs we construct a splitting method. By composing the basic Lie-Trotter splitting with its adjoint (the same method with reversed ordering) a symmetric Strang splitting with second order accuracy is constructed. Solution trajectories from these schemes provide exact solutions to modified Poisson problems with the original bracket, thereby respecting the Casimirs

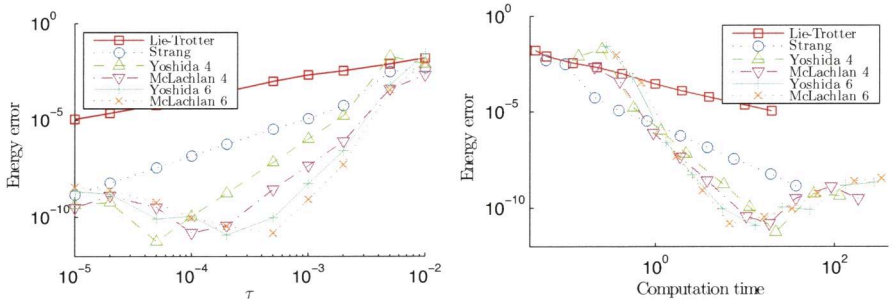


Figure 4.11: Comparison between different splitting schemes for a system with 360 vortices. On the left the error in energy is compared against different time steps, on the right it is compared against simulation time required. All simulations are performed on eight parallel threads.

of the original dynamics. The modified Hamiltonian is studied with the use of backward error analysis, showing that inaccuracies in the Hamiltonian occur with close vortex interactions. The conservation properties are studied in a number of test cases, including those considered by Vankerschaver & Leok [140].

The splitting method can also be rearranged into the interactions between groups of vortices, allowing parallelization of the workload. This reduces the natural quadratic scaling of computation time with system size to linear scaling when the number of processors is increased accordingly. The ordering of the pairwise evaluations can be modified to reduce communication overhead. The method therefore extends well to distributed memory implementations for large systems, allowing the method to be used for engineering applications or for studying statistical mechanics of point vortex dynamics.

The Strang splitting is also used as a basis for constructing higher order methods following Yoshida [143] and McLachlan [99]. When higher accuracy is desired, these methods are more efficient in terms of computational power than Lie-Trotter or Strang splitting.

In the statistical mechanics study of point vortex systems it is usually assumed that the vortex strength decreases as the number of vortices is increased. This follows immediately if it is desired that the enstrophy $Z = \int_{\mathbb{S}^2} \omega^2 \, \mathrm{d}S = \sum_{i-1}^{N} \Gamma_i^2$ matches the enstrophy of some real fluid. The same can be expected of point vortex models used as predictive models. This means that when the number of vortices is increased, the time step need not be decreased, thus preserving the linear increase in computational load.

5

Direct control of the small-scale energy balance in two-dimensional fluids

This chapter is based on:

An article with the same title submitted submitted to Journal of Fluid Mechanics

background graphic: vorticity field in forced-dissipative turbulent fluid flow

5.1 Introduction

In fluid dynamics applications such as ensemble weather prediction and climate simulation, when the time scales of interest are long compared to the Lyapunov time, the goal of simulations is to accurately sample an evolving probability density function of the solution. In general both the intermediate and steady-state distributions are unknown, being the consequence of forcing and dissipation introduced at various scales.

Recent improvements in the understanding of numerical methods, specifically the development of backward error analysis [62, 86], allow interpretation of the numerical solution as the exact solution of a modified system of equations. The modified equations typically admit their own (modified) invariant measure, and numerical truncation errors therefore bias the statistics obtained in simulation according to this altered statistical distribution. Thus numerical methods imply structural bias due to numerical truncation, even when the continuum model is complete.

As a precursor to accurately sampling an *evolving* measure, it would seem essential that the numerical method accurately sample the *stationary* invariant measure in the absence of forcing and dissipation, to allow correct response of the system to perturbations from equilibrium; however, even this requirement is typically not fulfilled, as has been observed in numerical investigations of simple two-dimensional ideal fluids models. In [42] it was shown that the equilibrium statistical mechanics of finite difference discretizations of quasi-geostrophic vorticity flow over topography are sensitive to the preservation of kinetic energy and (quadratic) enstrophy. Even in the idealized setting of unforced, inviscid 2D flow, a correct sampling of non-Gaussian statistics in the Miller-Robert-Sommeria ensemble requires specialized techniques [2, 43], and much less is known about the accuracy of sampling the nonequilibrium steady states treated in this paper.

In the fluid dynamics setting, several equilibrium models are known. For unforced, ideal fluids in two dimensions, the Miller-Robert-Sommeria measure [102, 123, 124], which encodes the area distribution of the vorticity field, is well established [16]. In fluctuating hydrodynamics, the Landau-Lifshitz-Navier-Stokes equations are provably ergodic with respect to the Boltzmann-Gibbs distribution of kinetic energy under a fluctuation-dissipation relation and stochastic forcing [46, 47, 41, 36]. For the geophysically relevant regime of fixed-wavelength stationary forcing and viscous dissipation, fluids in the atmosphere and ocean are believed to sample a nonequilibrium steady state in which the kinetic energy spectrum satisfies a power law over a range of length scales, as posited by Kolmogorov [89, 12, 83, 56, 139]. In this case no equilibrium measure is explicitly known, but in the probabilistic setting a stationary *expectation*—namely the power-law spectrum for kinetic energy—can be observed from measurements in the atmosphere [108].

Power-law kinetic energy spectra can be simulated using forced Navier-Stokes discretizations at operational resolutions, but this typically requires the introduction of a viscosity coefficient that far exceeds that encountered in atmosphere/ocean science applications in nature. In practice, higher order hyperviscosity is used because it has a more localized effect on the spectrum. Excessive numerical viscosity is believed to adversely affect the simulated growth of small-scale physical instabilities as well as inhibiting spread in ensemble simulations [79, 137]. Our simulation experiments bear this out, as we observe a strong influence of numerical viscosity on autocorrelation functions and the information content of ensembles. Turbulent backscatter methods have been introduced in [40, 134, 13] to re-inject kinetic energy at viscous length scales. Alternatively, "superparameterization" methods [142, 60] have been proposed as an intermediate alternative to large eddy simulations. In these, eddy dynamics are modelled by either a simplified dynamics or a stochastic closure model. In this paper we adopt an extreme statistical simplification of the fine-scale model, coupling it via thermostatic controls to directly impose a background power law kinetic energy spectrum at the smallest resolved scales. Our approach allows us to maintain the given target without employing artificially increased viscosity. The energy spectrum we impose can be taken from observational data, theory or higher resolution simulations. In the case of atmospheric turbulence the experiments by Nastrom & Gage [107] provide such data. In two-dimensional forced-dissipated Navier-Stokes it may also be taken from theoretical predictions [89, 12, 83].

The remainder of this paper is organized as follows: In the following section the incompressible 2D Navier-Stokes equations, with forcing and dissipation, are recalled in their vorticity form. Section 5.3 discusses the dynamical perturbations used in molecular dynamics to simulate a molecular gas at constant temperature. The interpretation of such perturbation methods in the context of turbulence is discussed in Section 5.4. A feedback control is then applied to 2D turbulence simulations in Sections 5.5 and 5.6; the former is a simulation with large scale random forcing and forward enstrophy cascade and the latter is a simulation augmented with additional, small scale forcing that is unresolved due to spectral truncation. Both of these sections include statistic and dynamic results of the new approach. A short discussion of similar methods and possible practical applications in Section 5.7 concludes the paper.

5.2 Two-dimensional turbulence

We focus on driven two-dimensional incompressible flow. Ignoring rotation and topographical effects, we work with the Navier-Stokes equations on a doubly periodic domain $\boldsymbol{x} \in \mathbb{T}^2$. The 2D Navier-Stokes equations with forcing

$f(\boldsymbol{x}, t)$ and generalized viscosity model are:

$$\omega_t + J(\psi, \omega) = f + \nu_{-1}\bar{\Delta}^{-1}\omega - \nu_p(-\Delta)^p\omega, \qquad \Delta\psi = \omega, \tag{5.1}$$

where $\psi(\boldsymbol{x}, t)$ is the stream function and $\omega(\boldsymbol{x}, t)$ is the vorticity component normal to the plane, Δ is the Laplace operator and $\bar{\Delta}^{-1}$ denotes scale-selective damping (see below). In this paper we use either physical viscosity $p = 1$ or fourth order hyperviscosity $p = 4$. The nonlinearity $J(\psi, \omega)$ is defined by

$$J(\psi, \omega) = \frac{\partial\psi}{\partial x}\frac{\partial\omega}{\partial y} - \frac{\partial\psi}{\partial y}\frac{\partial\omega}{\partial x}. \tag{5.2}$$

Equation (5.1) is discretized using a pseudo-spectral method [23], expressing the vorticity field in terms of its Fourier components

$$\omega_{\boldsymbol{k}} = \frac{1}{(2\pi)^2}\int_{\mathbb{T}^2} \omega(\boldsymbol{x})e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x}, \quad |\boldsymbol{k}|_\infty \leq K, \tag{5.3}$$

where $\boldsymbol{k} = (k_1, k_2)$, is an index vector, and we denote $|\boldsymbol{k}| = (k_1^2 + k_2^2)^{1/2}$ and $|\boldsymbol{k}|_\infty = \max\{|k_1|, |k_2|\}$. In terms of its Fourier components, equation (5.1) is written

$$\dot{\omega}_{\boldsymbol{k}} + J_{\boldsymbol{k}}(\boldsymbol{\omega}) = f_{\boldsymbol{k}} + \nu_{-1}\bar{\Delta}_{\boldsymbol{k}}^{-1}\omega_{\boldsymbol{k}} - \nu_p\Delta_{\boldsymbol{k}}^p\omega_{\boldsymbol{k}},$$

where $\Delta_{\boldsymbol{k}} = -|\boldsymbol{k}|^2$ and the scale-selective damping is defined by

$$\bar{\Delta}_{\boldsymbol{k}}^{-1} = \begin{cases} -|\boldsymbol{k}|^{-2}, & |\boldsymbol{k}| \leq 3 \\ 0, & \text{otherwise.} \end{cases}$$

The nonlinear term $J_{\boldsymbol{k}}(\boldsymbol{\omega})$ represents the pseudo-spectral evaluation of (5.2) on a uniform $2K \times 2K$ grid, implementing a standard 3/2 filter to avoid aliasing due to quadratic terms [23].

Our computational set-up is similar to that of Gotoh [59]. Scale-selective viscosity is restricted to those modes with $|\boldsymbol{k}| \leq 3$ to curtail the inverse cascade of energy. The forcing is Gaussian white noise in time and applied in a band of energy shells with $3.5 < |\boldsymbol{k}| < 6.5$. For the simulations of Section 5.6, small scale random forcing is additionally applied in the range $202.5 < |\boldsymbol{k}| < 206.5$. The magnitude of the forcing is scaled such that the expected power input matches a given value P as follows. With the energy given by $E(t) = -\frac{1}{2}\sum_{\boldsymbol{k}} \Delta_{\boldsymbol{k}}^{-1}\omega_{\boldsymbol{k}}\omega_{\boldsymbol{k}}^*$, the expected power input due to forcing is equal to the expected change in energy (using Itô's formula)

$$\mathbb{E}\left[\,\mathrm{d}E\right] = \mathbb{E}\left[\frac{\partial E}{\partial t}\,\mathrm{d}t + \sum_{\boldsymbol{k}}\frac{\partial E}{\partial\omega_{\boldsymbol{k}}}\,\mathrm{d}\omega_{\boldsymbol{k}} + \frac{1}{2}\,\mathrm{d}\omega_{\boldsymbol{k}}\frac{\partial^2 E}{\partial\omega_{\boldsymbol{k}}^2}\,\mathrm{d}\omega_{\boldsymbol{k}}\right]. \tag{5.4}$$

For forcing with uniform magnitude across a band $\boldsymbol{k} \in K_f$ of wave numbers, we substitute $\mathrm{d}\omega_{\boldsymbol{k}} = \hat{f}\,\mathrm{d}Z_{\boldsymbol{k}}$, where $\mathrm{d}Z_k = \mathrm{d}A_k + i\,\mathrm{d}B_k$ is a complex Wiener

Physical	Reference	Truncated	Hyperviscosity	Nosé-Hoover
P	0.1	0.1	0.1	0.1
ν_{-1}	2	2	2	2
p	1	1	4	1
ν_p	1.0×10^{-4}	1.0×10^{-4}	4.3×10^{-15}	1.0×10^{-4}
Numerical				
$\Delta t \times 10^3$	1	1	1	1
K	256	85	85	85
ℓ^*	-	-	-	$\{51, 71^\dagger, 81\}$
ε_0	-	-	-	$\{10^{-1}, 10^{-1/2}, 1^\dagger\}$
Results				(case denoted †)
η	4.92	4.92	4.92	4.92
C'	1.15	0.645	1.33	1.14
d	0.789	0.305	0.900	0.779

Table 5.1: Parameters and results for the simulation of two-dimensional turbulence

increment, into (5.4) to find

$$\mathbb{E}[\mathrm{d}E] = -\hat{f}^2 \sum_{\boldsymbol{k} \in K_f} \Delta_{\boldsymbol{k}}^{-1} \,\mathrm{d}t = P \,\mathrm{d}t,$$

whence we compute that the magnitude of the forcing in the forced modes should equal $\hat{f} = \sqrt{\frac{P}{\sum_{\boldsymbol{k} \in K_f} |\boldsymbol{k}|^{-2}}}$. Consequently a time-$\Delta t$ forcing increment is computed via the formula

$$\omega_{\boldsymbol{k}}(t + \Delta t) = \omega_{\boldsymbol{k}}(t) + \sqrt{\frac{P \Delta t}{\sum_{\boldsymbol{k} \in K_f} |\boldsymbol{k}|^{-2}}} (R_{\boldsymbol{k}} + i S_{\boldsymbol{k}}),$$

where $R_{\boldsymbol{k}}, S_{\boldsymbol{k}}$ are unit normal pseudorandom numbers.

The values of the parameters used in the simulations are summarized in table 5.1.

The viscous terms in (5.2) have typical length scales defined by the wave number magnitudes for which the coefficients $\nu_p \Delta_{\boldsymbol{k}}^p$ have magnitude unity. By using $\Delta_{\boldsymbol{k}} = -|\boldsymbol{k}|^2$ we find $k_d = \sqrt{\nu_1}^{-1}$ for $j = 1$ and $k_h = \sqrt{\nu_{-1}}$ for $j = -1$. Assume that these two scales are sufficiently well separated, $k_d \gg k_h$, and that forcing acts primarily at some intermediate length scale. In this setting it is expected that the hypothesis of Kraichnan [82] holds and that there is a steady flux of energy from the forcing wave numbers to larger scales (i.e. the inverse energy cascade), as well as a steady enstrophy flux to smaller scales (i.e. the direct enstrophy cascade). These cascades terminate when the dissipative scales k_h and k_d are reached, but with sufficient separation between

forcing and dissipation scales there persist a range of wave numbers of statistically stationary energy and enstrophy transport. Because the dynamics in these ranges are almost unaffected by damping and forcing, they are dubbed inertial ranges. The steady fluxes of energy (resp. enstrophy) in both regimes yield power law energy spectra.

This means the energy in wavenumbers near k, given by

$$E_k(\boldsymbol{\omega}) = -\frac{1}{2} \sum_{k-\frac{1}{2}<|\boldsymbol{k}|<k+\frac{1}{2}} \Delta_{\boldsymbol{k}}^{-1} \omega_{\boldsymbol{k}} \omega_{\boldsymbol{k}}^*, \tag{5.5}$$

satisfies on average an approximate power law (time average denoted by overbar)

$$\bar{E}_k \approx C\varepsilon^{2/3} k^{-5/3} \text{ for } k \ll k_f,$$

in the inverse energy cascade, and

$$\bar{E}_k \approx C\eta^{2/3} k^{-3} \text{ for } k \gg k_f,$$

in the direct enstrophy cascade. The parameters ε and $\eta \approx k_f^2 \varepsilon$ denote the energy and enstrophy injection rates [15]. Using analysis of structure functions, [90] concludes that a k^{-3} inertial spectrum is plausible under two-dimensional turbulence assumptions, but that a $k^{-5/3}$ range cannot be so explained. In numerical simulations, various other power laws are observed (see references below). The methodology we propose makes no assumptions on the functional form of the kinetic energy spectrum and is therefore applicable to any observed spectrum. We demonstrate this by using the method in a forced-dissipated turbulence cascade in Section 5.5 and in a case with both large and small-scale stochastic forcing in Section 5.6. The versatility of the method also promises straightforward generalization to three-dimensional turbulence.

The effect of spectral truncation on the kinetic energy spectrum is most pronounced in the inertial enstrophy regime. To save computational effort, we design our simulation with small separation between the scales of forcing and large-scale damping. Such a parameter set is given in the column labeled "Reference" of table 5.1. This simulation yields an inertial range power spectrum as depicted in figure 5.1. The figure shows both the instantaneous spectrum and the time-averaged spectrum after 50 time units, corresponding to over 250 eddy turnovers. The computed energy spectrum is steeper than the hypothesized k^{-3} slope. This is common in numerical simulations and is usually attributed to insufficiently large Reynolds number due to limited resolution [15, 59, 19]. Saffman [128] proposes a k^{-4} spectrum due to vortex filamentation and small amplitude front formation. Farazmand et al. [50] suggest that numerical simulations differ from the hypothesized spectra

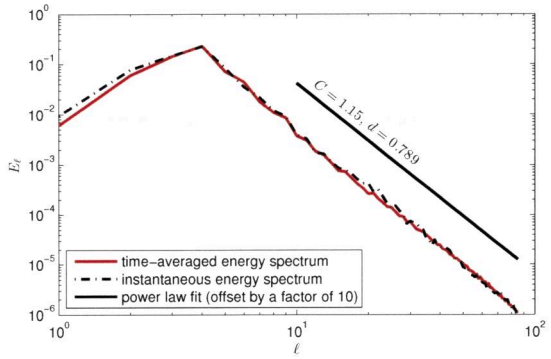


Figure 5.1: Two-dimensional turbulence kinetic energy spectrum and power law fit. Simulation parameters are given in the "reference" column of table 5.1.

not due to insufficient resolution, but due to the choice of forcing. They investigate forcing functions that yield the hypothesized $k^{-5/3}$ and k^{-3} power law regimes. This is in line with the findings of Danilov & Gurarie [33], who demonstrate that power law spectra can only be observed for a special set of external parameters. Farge et al. [51] use wavelet methods to analyse turbulent velocity fields and define a local energy spectrum. They discover that the k^{-3} energy spectrum only holds outside of regions of strong vorticity and shear layers. Inside those regions the energy spectrum scales as k^{-6} and k^{-4} respectively. Sukoriansky et al. [135] find that the forward enstrophy cascade spectral slope depends directly on the chosen large scale drag.

We fit a power law to the observed (steeper) spectrum in the form suggested by Gotoh [59]

$$\bar{E}_k \approx C' \eta^{2/3} k^{-3-d},$$

using a least-squares approach. The newly introduced parameter d indicates a deviation from the theoretical slope. Even though the power law that develops for a given parameter set in (5.1) is different from the theoretical spectrum, it is still independent of the chosen initial conditions. This is evidence that the dynamics are ergodic and motivates consideration of the invariant measure of the dynamical system, analogous to such considerations in molecular dynamics. The following section briefly consider the tools used in molecular dynamics, before their application to 2D turbulence is detailed in Section 5.4.

5.3 Canonical sampling and temperature control

In this section we briefly recall the problem of constant temperature simulations in molecular dynamics, which is the inspiration for the control we propose for kinetic energy in the Navier-Stokes equations. The dynamics of a classical molecular gas are governed by a Hamiltonian system

$$\dot{\boldsymbol{q}} = \boldsymbol{p}, \tag{5.6}$$

$$\dot{\boldsymbol{p}} = -\nabla_{\boldsymbol{q}} V(\boldsymbol{q}), \tag{5.7}$$

where $\boldsymbol{q} \in \mathbb{R}^n$ represents the vector of particle positions, $\boldsymbol{p} \in \mathbb{R}^n$ the vector of particle momenta, and unit mass is assumed. The Hamiltonian $H(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{2}\|\boldsymbol{p}\|^2 + V(q)$ represents the total energy as a sum of kinetic energy $\kappa(\boldsymbol{p}) = \frac{1}{2}\|\boldsymbol{p}\|^2$ and potential energy $V(\boldsymbol{q})$. Given an initial condition with total energy H_0, a solution is restricted to the constant energy surface $H(\boldsymbol{q}(t), \boldsymbol{p}(t)) = H_0$. If the phase flow is ergodic on the constant energy surface, then for almost any initial condition, the time average of an observable function $a(\boldsymbol{q}, \boldsymbol{p})$ of the solution,

$$\bar{a}(\boldsymbol{q}, \boldsymbol{p}) = \lim_{T \to \infty} \frac{1}{T} \int_0^T a(\boldsymbol{q}(t), \boldsymbol{p}(t))\, \mathrm{d}t,$$

is equal to the ensemble average with respect to the *microcanonical* measure π_m,

$$\langle a \rangle = \int a(\boldsymbol{q}, \boldsymbol{p})\, \pi_m(\,\mathrm{d}q,\, \mathrm{d}p) = \int a(\boldsymbol{q}, \boldsymbol{p})\, \delta(H(\boldsymbol{q}, \boldsymbol{p}) - H_0)\, \mathrm{d}\boldsymbol{q}\, \mathrm{d}\boldsymbol{p},$$

where H_0 is the total energy defined by the initial condition.

For a molecular gas in thermal equilibrium with a reservoir of temperature τ, the total energy is no longer conserved. Instead the time averaged kinetic energy satisfies

$$\bar{\kappa} = \frac{n k_B \tau}{2} = \frac{n}{2\beta},$$

where k_B is the Boltzmann constant and β is the inverse temperature. In this case, trajectories of the system sample the *canonical* (Gibbs) measure

$$\pi_c(\,\mathrm{d}\boldsymbol{q},\, \mathrm{d}\boldsymbol{p}) = \rho_c(\boldsymbol{q}, \boldsymbol{p})\, \mathrm{d}\boldsymbol{q}\, \mathrm{d}\boldsymbol{p}, \quad \rho_c(\boldsymbol{q}, \boldsymbol{p}) \propto e^{-\beta H(\boldsymbol{q}, \boldsymbol{p})},$$

and when the flow is ergodic in this measure the temperature is related to the canonical mean

$$\frac{n k_B \tau}{2} = \langle \kappa \rangle = \int \kappa(\boldsymbol{p}) \pi_c(\,\mathrm{d}\boldsymbol{q},\, \mathrm{d}\boldsymbol{p}).$$

To carry out numerical simulations of molecular dynamics at constant energy, there exist numerical methods that (exactly or approximately) preserve

the Hamiltonian. Preservation of a quantity like $\bar{\kappa}$ is more subtle, since it is only conserved 'on average'. Methods for constant temperature molecular dynamics introduce perturbations to the dynamical equations (5.6)–(5.7) called thermostats. The Nosé-Hoover thermostat [113, 114, 70] augments the dynamics with an extra variable ξ that controls the kinetic energy in the system as follows

$$\dot{\boldsymbol{q}} = \boldsymbol{p} \tag{5.8}$$
$$\dot{\boldsymbol{p}} = -\nabla_{\boldsymbol{q}} V(\boldsymbol{q}) - \xi\varepsilon\boldsymbol{p}$$
$$\dot{\xi} = \varepsilon\left(2\beta\kappa(\boldsymbol{p}) - n\right), \tag{5.9}$$

where ε is a coupling parameter. The Nosé-Hoover method can be interpreted in two ways. First, it is straightforward to verify that the Liouville equation associated to the augmented system (5.8)–(5.9) admits the steady state

$$\rho(\boldsymbol{q}, \boldsymbol{p}, \xi) = \rho_c(\boldsymbol{q}, \boldsymbol{p}) e^{-\frac{1}{2}\xi^2},$$

for which the marginal density with respect to $\boldsymbol{q}$ and $\boldsymbol{p}$ is the Gibbs distribution. Hence the Nosé-Hoover method enforces the canonical invariant measure. On the other hand, it is also apparent that the variable ξ acts as a damping coefficient for kinetic energy when $\xi > 0$ and excites kinetic energy for $\xi < 0$. Furthermore, (5.9) shows that ξ will increase (decrease) when $2\kappa(\boldsymbol{p})/n$ exceeds (falls short of) the target temperature $\beta^{-1} = k_B\tau$. Hence, the Nosé-Hoover method can also be interpreted as a *negative feedback control* on temperature. This second interpretation is crucial to our application of the thermostat to fluids, as we explain below.

5.4 Invariant measures and expectations for two-dimensional turbulence

The tendency of forced-dissipated turbulence to develop a power law spectrum independent of initial conditions provides evidence that the dynamics may sample a unique invariant measure. Ergodicity of finite truncations of the forced-dissipated two-dimensional Navier-Stokes equations is proven by E & Mattingly [46] for the case of stochastic forcing of only two long wave modes, $\boldsymbol{k}_1 = (1,1)$ and either $\boldsymbol{k}_2 = (1,0)$ or $\boldsymbol{k}_2 = (0,1)$. Continuing this analysis, it is readily proven that this result extends to the case of forcing arbitrary wave numbers $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ as well as $\boldsymbol{k}_1 + (1,0)$ and $\boldsymbol{k}_2 + (0,1)$. This is the case in our proposed forcing of selected bands in Fourier space.

Our interest lies in the practical case where computational costs prohibit resolving a sufficient number of modes to capture small-scale dissipation. This is the case in many large scale atmosphere and ocean applications. The computational load is determined by restricting the Fourier expansion in (5.3)

to those modes with $|\boldsymbol{k}| \leq K$. Denote the truncated vector of vorticity coefficients by $\boldsymbol{\omega}_K$, and the full Fourier transform by $\boldsymbol{\omega}_\infty$. If the resolution K is insufficiently large to capture the diffusion of enstrophy at the scale k_d, the dynamics of the truncated system will differ greatly from those in a system that is well-resolved.

Stochastic approaches to model the effect of the unresolved degrees of freedom focus on the dynamical interaction between resolved and unresolved modes (see for instance Mori [104], Zwanzig [145], Hasselmann [65], Majda et al. [96] and Fatkullin & Vanden-Eijnden [52]). Here, we instead focus on correcting the statistics of the truncated system, as embodied in its invariant distribution and expectations.

Ideally, for correct sampling we would require that the invariant density ρ_K sampled by the truncated dynamics (i.e., $\boldsymbol{\omega}_K \sim \rho_K$) be equal to the marginal distribution of those same modes in the resolved case. Let us introduce a partition $\boldsymbol{\omega}_\infty = (\hat{\boldsymbol{\omega}}, \tilde{\boldsymbol{\omega}})$ where $\hat{\boldsymbol{\omega}}$ consists of resolved modes with $|\boldsymbol{k}| < K$ and $\tilde{\boldsymbol{\omega}}$ for the unresolved modes. For equivalence between the invariant measures of the two systems, we would require ρ_K to be equal to the marginal density

$$\hat{\rho} = \int \rho_\infty(\hat{\boldsymbol{\omega}}, \tilde{\boldsymbol{\omega}}) \, \mathrm{d}\tilde{\boldsymbol{\omega}}.$$

However, steady solutions to the Fokker-Planck equation for general forced-dissipated turbulence are not explicitly available, due to the inherent complexity of the nonlinear wave interactions. This means we cannot define a perturbation for the truncated system such that the invariant measure is correct.

Instead we focus attention on *expectations* in the unknown measures. If the invariant measures $\bar{\rho}$ and ρ_K *did* match, so would the expectations of arbitrary observables $\langle a(\bar{\omega}) \rangle$. Ergodicity of the systems would then also imply equivalent time averages. We take the energy spectrum as a set of observables. Given a truncation $K \gg k_d$, the modes $\omega_{\boldsymbol{k}}$ with $k_f \ll |\boldsymbol{k}| \ll k_d$ revert to a power law spectrum due to the forward enstrophy cascade. In a system truncated well below the viscous scale $K \ll k_d$, the downscale cascade of enstrophy is terminated abruptly, resulting in an artificial build-up of enstrophy at the smallest resolved scales known as spectral blocking. An inaccurate energy spectrum in the highest wave numbers eventually leads to deviation from the power law spectrum in the energy range, i.e. in the large scales [135].

In this paper, we propose employing the Nosé-Hoover thermostat (5.8)–(5.9) to enforce a power law spectrum on the kinetic energy in the absence of a mean flow. A crucial difference in the application to fluids, compared to constant temperature molecular dynamics, is that in the current context the invariant measure of the extended variable is unknown. In canonical sampling, the distribution for the thermostat variable ξ is known to be normally

distributed with mean zero, independent of the original dynamics. This allows augmentation of the thermostat variable dynamics (5.9) by an Ornstein-Uhlenbeck process without disrupting the invariant measure [132]. The addition of stochastic forcing in the auxiliary variable makes the thermostat ergodic [87]. However, in the current context of forced turbulence, we do not know the distribution of the thermostat variable. Moreover, it is expected that $\xi(t)$ will have nonzero mean, because it must remove excess energy on average, yet add energy in the form of backscatter on occasion. For this reason it is crucial to *exclude* the stochastic process in the thermostated wave numbers. Similar arguments were used in a forced molecular model in Jones & Leimkuhler [75].

5.5 Feedback control of the forward enstrophy cascade

The kinetic energy spectrum consists of the kinetic energy distributed over energy shells in wavenumber space, see (5.5). A Nosé-Hoover thermostat could be applied to each shell to drive its energy to the observed average. However, it is undesirable to artificially perturb the largest scale modes in the system, which are well-resolved and least uncertain. For this reason, only energy shells with wave number $\ell > \ell^*$ above a threshold are equipped with a Nosé-Hoover thermostat. To this end, denote by $\ell(\boldsymbol{k})$ the energy shell containing wave number $\boldsymbol{k}$.

We choose perturbation parameters ε_ℓ to satisfy

$$\varepsilon_\ell = \begin{cases} \bar{E}_\ell^{1/2}\varepsilon_0, & \ell \geq \ell^* \\ 0, & \ell < \ell^*. \end{cases}$$

The discrete equations of motion for two-dimensional Euler flow, extended with the thermostats, reads

$$\begin{aligned} \dot{\omega}_{\boldsymbol{k}} + J_{\boldsymbol{k}}(\boldsymbol{\omega}) &= f_{\boldsymbol{k}} + \nu_{-1}\bar{\Delta}_{\boldsymbol{k}}^{-1}\omega_{\boldsymbol{k}} + \nu_1\Delta_{\boldsymbol{k}}\omega_{\boldsymbol{k}} - \xi_\ell\varepsilon_\ell\partial_{\boldsymbol{k}}E_\ell(\boldsymbol{\omega}), \quad \ell = \ell(\boldsymbol{k}), \forall \boldsymbol{k} \\ \dot{\xi}_\ell &= \varepsilon_\ell\left(E_\ell(\boldsymbol{\omega}) - \bar{E}_\ell\right), \quad \ell = \ell^*, \ldots, \ell_{\max}. \end{aligned} \tag{5.10}$$

The viscosity term $\nu_1\Delta_{\boldsymbol{k}}\omega_{\boldsymbol{k}}$ in (5.10) ensures that the resolved scale modes not directly driven by the thermostats possess a dynamics that is consistent with that of the reference solution.

For each thermostated mode, the energy is driven towards a target value. We emphasize that this value may be taken from physical observations, theoretical predictions or, as here for the purpose of method evaluation, from a high fidelity solution that resolves the physical viscosity. As such the method may be seen as a data assimilation approach that uses statistical data to correct mean statistics of a dynamical simulation.

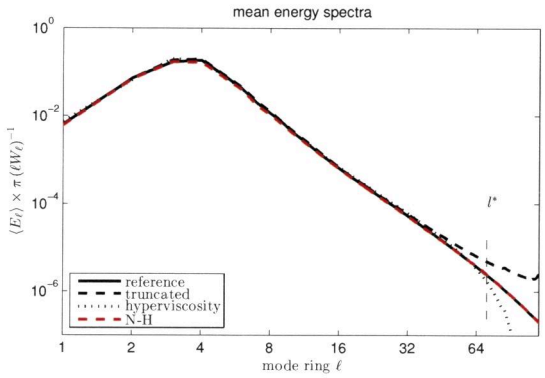


Figure 5.2: The kinetic energy spectrum after 100 time units for the reference simulation (solid), a truncated simulation (dashed), a simulation with hyperviscosity (dotted) and a simulation using the proposed control on the energy spectrum (dashed, red). Parameters are given in table 5.1, where the † indicates the parameters used here for the Nosé-Hoover control.

It should be noted that while the control will certainly drive the system towards correct averages for the energy levels in the thermostated energy shells, the invariant measure sampled by the trajectories remains unknown. An important consequence of this is that the marginal distribution of the thermostat variables is not known a priori. This complicates choosing initial values for the thermostat variables, as initialising them far from their equilibrium will result in a slow relaxation. We perform a pilot simulation in which the thermostated system is allowed to fully equilibrate in order to select initial data for the ξ_ℓ.

5.5.1 Energy spectrum

The energy spectrum for a simulation using a thermostat is compared to a resolved model and an underresolved model with hyperviscosity ($\nu_4 = 4.3 \times 10^{-15}$) in figure 5.2. The kinetic energy per energy shell is multiplied by a correction factor that accounts for (i) the nonuniformity that arises by partitioning of the discrete Fourier space into annular shells and (ii) the incomplete resolution of the highest wave number bands (i.e. those in the corners of the Fourier space). The factor is $\frac{\pi}{\ell W_\ell}$, where $W_\ell = -\frac{1}{2}\sum_{\ell-\frac{1}{2}<|\boldsymbol{k}|<\ell+\frac{1}{2}} \Delta_{\boldsymbol{k}}^{-1}$, and essentially results in smooth spectra in figure 5.2 (vis-à-vis figure 5.1). The artifical (hyper)viscosity model grossly underestimates the kinetic energy in the large wave numbers compared to the well-resolved model. The kinetic energy spectrum in the thermostated model is visually indiscernible from the reference model.

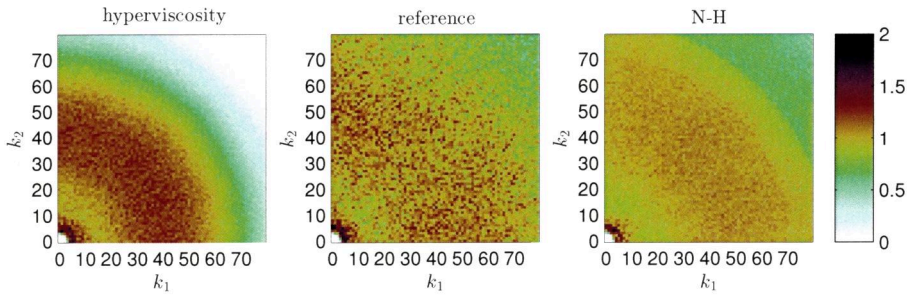


Figure 5.3: A representation of the average energy in each Fourier mode using (left) the model with hyperviscosity, (centre) the reference simulation, and (right) the Nosé-Hoover control, see table 5.1 for details. The value plotted is given by $\hat{C}_{\boldsymbol{k}} = 2\pi\overline{\omega_{\boldsymbol{k}}\omega^*_{\boldsymbol{k}}}/(C'\eta^{2/3}k^{-3-d-1})$, where the parameters C' and d are fit to the reference solution (cf. 5.2). When this value is close to unity, it indicates close local (in Fourier space) agreement to the power law spectrum.

To investigate the energy spectrum more closely, the mean energy per mode is represented in figure 5.3 for the reference, hyperviscosity, and thermostated simulations. The mean energy per mode is multiplied by the cube of the wavenumber magnitude. This will result in areas of equal colour for modes where the energy spectrum scales as k^{-4} and where isotropy can be assumed. Anisotropy would appear as a break of the radial symmetry. This does not occur in the reference simulation and for most of the controlled simulation. However, in figure 5.3 we do observe that, for the thermostated model, kinetic energy appears to be more uniformly distributed across the Fourier modes within a given energy shell than is the case for the reference model. This is most likely due to the heavily reduced dimensionality of the phase space (85^2 versus 256^2) leading to a much faster spread of the stochastic noise from the forcing through the available degrees of freedom.

5.5.2 Vorticity field

In figure 5.4 the computed vorticity fields at $t = 1$ and $t = 10$ are shown for the reference model, the model with hyperviscosity, and model using the Nosé-Hoover control. At $t = 1$, the vorticity structures produced by the thermostat are similar to those of the reference and hyperviscosity models, indicating that the thermostat only weakly perturbs the large scale vorticity. After 10 time units, the solutions have diverged due to the chaotic nature of the dynamics, but the vorticity fields remain qualitatively similar.

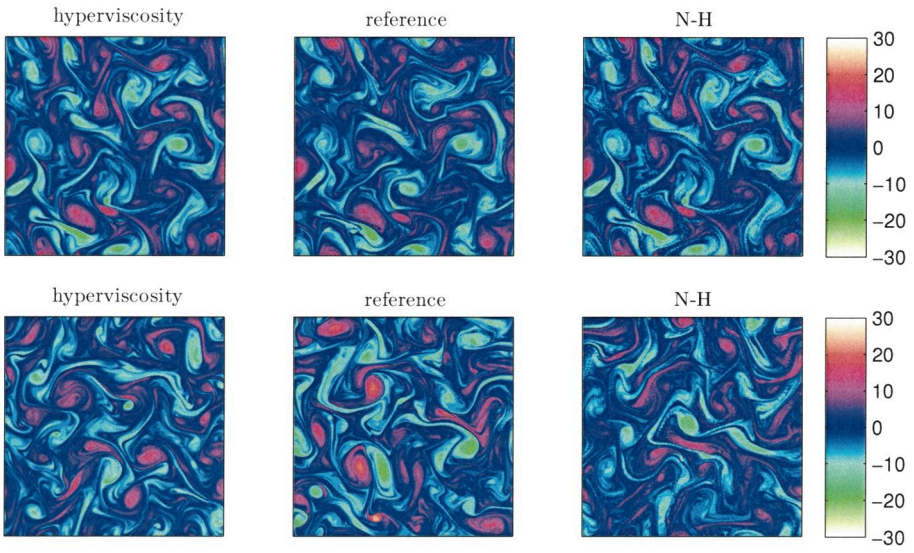


Figure 5.4: Vorticity fields in a classical double cascade, obtained at $t = 1$ (top) and $t = 10$ (bottom) using (left) the model with artificial viscosity added, (centre) the reference simulation, and (right) the Nosé-Hoover method as indicated with † in table 5.1.

5.5.3 Autocorrelation functions

Autocorrelation functions provide dynamical information on the temporal variance at different scales. The autocorrelation function $R_{\omega\omega}(s; x)$ of the vorticity ω at a point x is given by

$$R_{\omega\omega}(s; x) = \frac{1}{T} \int_0^T \omega(t + s, x)\omega(t, x) \, \mathrm{d}t$$

when observed over time T. We study the autocorrelation function of the vorticity at a grid point in the vorticity field. To save computational effort we will focus on the vorticity field corresponding to Fourier modes with $|\boldsymbol{k}|_\infty < 16$. As the vorticity at each grid point is identically distributed assuming homogeneous forcing, we average the autocorrelation functions over space to speed up convergence.

We compare the autocorrelation functions for the thermostated simulations to the fully resolved simulation, a truncated simulation, and a hyperviscosity model. In the left panel of figure 5.5 the autocorrelation function is shown for different perturbation parameters ε_0. The right panel focuses on the short-time behaviour. The agreement is relatively insensitive to perturbation parameter for the range of values shown (an order of magnitude). For

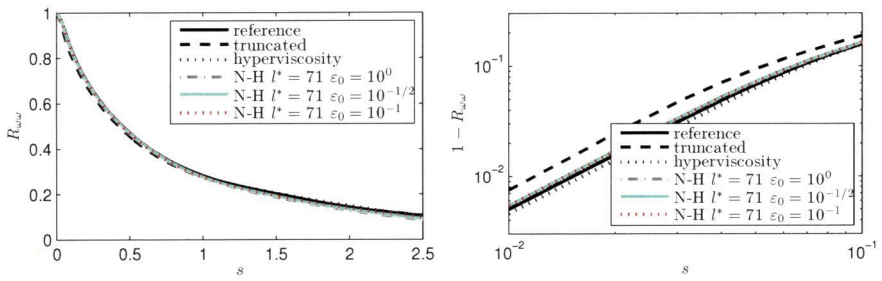


Figure 5.5: Autocorrelation functions $R_{\omega\omega}$ when simulating a classical double cascade. The Nosé-Hoover thermostat using various choices of the perturbation parameter ε_0 is compared to the reference solution (solid black), a truncated simulation (dashed black) and a hyperviscous model (dotted black). The figure on the right shows $1 - R_{\omega\omega}$ on a double logarithmic plot, to focus on short time scale behaviour.

larger values of ε_0, the thermostat acts more strongly, approaching Langevin dynamics in the limit of large ε_0 [54]. For smaller values of ε_0 the thermostat becomes very weak, meaning the relaxation of the spectrum requires averaging on long times. The choice of thermostat threshold ℓ^* has an even smaller effect on the autocorrelation functions and is therefore not shown.

5.5.4 Ensemble dispersion

A 50-member ensemble is created from a single deterministic initial condition by randomizing the phase of all modes with $|\boldsymbol{k}| \geq 50$; in this way each ensemble member has an identical initial kinetic energy spectrum. In figure 5.6 we compare ensembles, simulated up to $t = 10$, for the reference simulation (centre), hyperviscosity (left) and Nosé-Hoover control (right) by studying the phase angle of the (0,1)-mode. Both the hyperviscosity model and the Nosé-Hoover control are less dispersive than the reference solution. Nevertheless, the thermostated ensemble exhibits observably more variance than the hyperviscosity model, and does manage to reflect some of the outlying trajectories of the reference solution.

5.6 Feedback control of a system with subgrid scale forcing

In the previous section the Nosé-Hoover method corrected the energy spectrum in the forward enstrophy region for a truncated system. In this section we deviate from the classical setting of an intermediate forcing that results in two inertial ranges. Here we include a small-scale forcing term. This flattens the energy spectrum in the region between the two forcing scales when

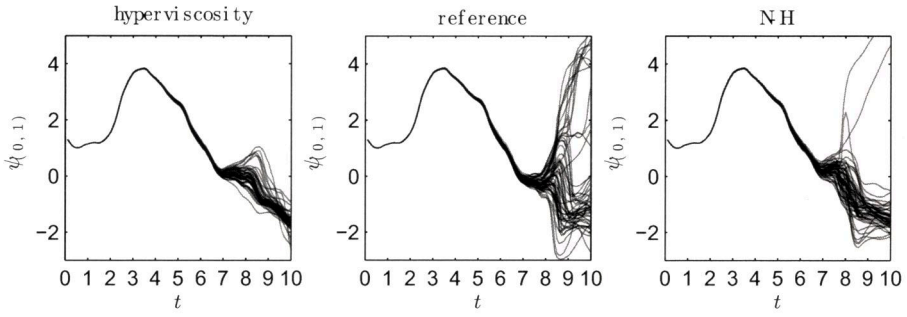


Figure 5.6: Ensemble dispersion in a classical double cascade. Plotted is the phase angle of the (0,1)-mode using (left) the model with artificial viscosity, (centre) the reference simulation, and (right) the Nosé-Hoover method as indicated with † in table 5.1.

compared to the previous case. The form of the forcing is the same as before (Gaussian white noise in the Fourier components), only now the wave numbers with $202.5 < |\boldsymbol{k}| < 206.5$ are additionally forced. The total power input in these modes is equal to that of the low wave number forcing.

A Nosé-Hoover control with remaining parameters as described in table 5.1 is applied to a truncated simulation with $K = 85$, i.e. the small-scale forcing is *not resolved*. Instead, the control target spectrum is observed from the fully resolved simulation after 100 time units. This simulates a scenario in which the fluid is forced at unresolved small scales, and we must attempt to incorporate this forcing given observations at resolved scales. Figure 5.7 shows the mean kinetic energy spectrum after a 100 units for each of the four different simulations of table 5.1. Both the underresolved and hyperviscosity models have no means of sensing the small-scale forcing and will consequently underestimate the energy in the smallest resolved scales. Clearly this is an unfair comparison, but we include results from these models to illustrate the difference. The Nosé-Hoover control acts only on mode bands ℓ^* and beyond ($\ell^* = 71$ in the figure), yet the energy spectrum is accurate over *all* wave numbers.

5.6.1 Vorticity field

The inclusion of small-scale forcing leads to a noisier vorticity field for the reference solution as seen in the centre panels of figure 5.8. Using the Nosé-Hoover control produces similar vorticity fields. At $t = 1$, the large scale structures of all three models are similar, again illustrating that the Nosé-Hoover control only weakly effects the dynamics at small wave numbers. The chaotic nature of the flow leads to decorrelation of solutions over long time,

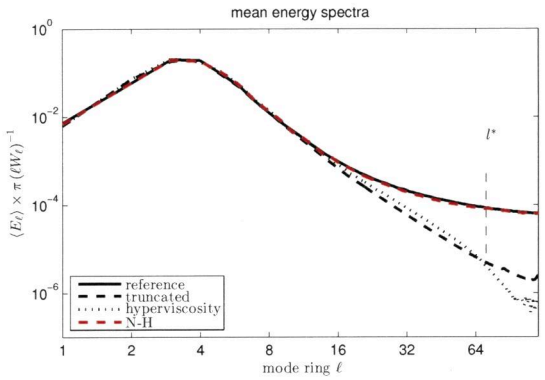


Figure 5.7: The kinetic energy spectrum after 100 time units for the reference simulation (solid), a truncated simulation (dashed), a simulation with hyperviscosity (dotted) and a simulation using the proposed control on the energy spectrum (dashed, red). Parameters are given in table 5.1, where the † indicates the parameters used here for the Nosé-Hoover control. An additional forcing is applied to the wavenumber with $203.5 < |\boldsymbol{k}| < 206.5$, with a power injection equal to that of the large scale forcing.

yet at $t = 10$ the controlled vorticity field is qualitatively still very similar to the reference.

5.6.2 Autocorrelation functions

As in the case with solely large scale forcing, we use autocorrelation functions for comparing dynamical properties. In figure 5.9 we compare the Nosé-Hoover control with different perturbation parameters ($\varepsilon_0 \in \{1, 10^{-1/2}, 10^{-1}\}$) against a reference simulation, a truncated simulation and a hyperviscosity model. The results for the truncated and hyperviscous models show excessive correlation in time. The autocorrelation function for the controlled dynamics depends strongly on the perturbation parameter ε_0 in this case with small scale forcing. For smaller ε_0 the autocorrelation functions approach those of the truncated dynamics. The largest ε_0 considered does decorrelate similarly to the reference solution. The results are insensitive to the wavenumber threshold ℓ^* for the control.

5.6.3 Ensemble dispersion

Again we compare the evolution of the phase angle of the (0,1)-mode in a 50-member ensemble simulation for each of the three models. The ensemble was again prepared by randomizing the phases of modes with $|\boldsymbol{k}| \geq 50$. Subsequently, all ensemble members observed identical Wiener increments

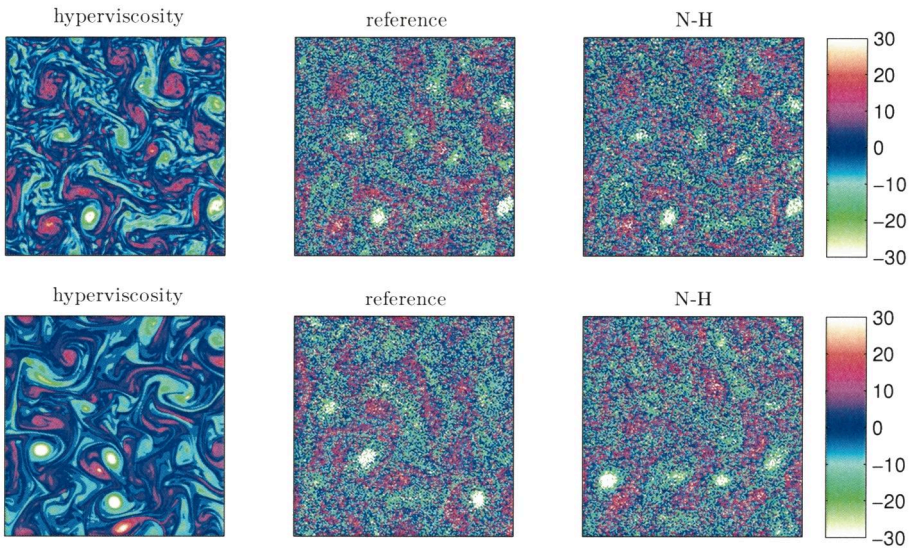


Figure 5.8: Vorticity fields with small scale forcing, obtained at $t = 1$ (top) and $t = 10$ (bottom) using (left) the overly diffusive hyperviscosity model, (centre) the reference simulation, and (right) the Nosé-Hoover controlled method as indicated with † in table 5.1.

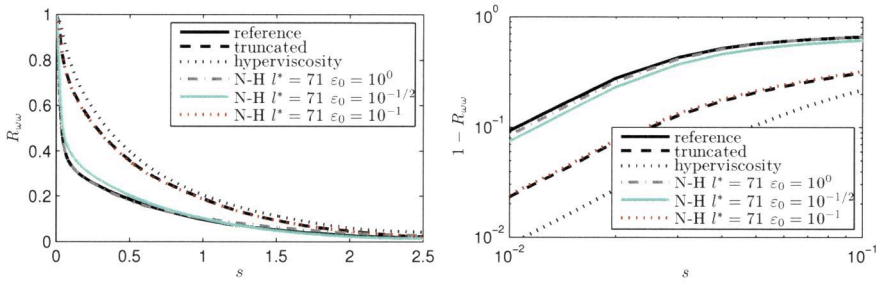


Figure 5.9: Autocorrelation functions $R_{\omega\omega}$ with small scale forcing. The Nosé-Hoover thermostat using various choices of the perturbation parameter ε_0 is compared to the reference solution (solid black), a truncated simulation (dashed black) and a hyperviscous model (dotted black). The figure on the right shows $1 - R_{\omega\omega}$ on a double logarithmic plot, to focus on short time scale behaviour.

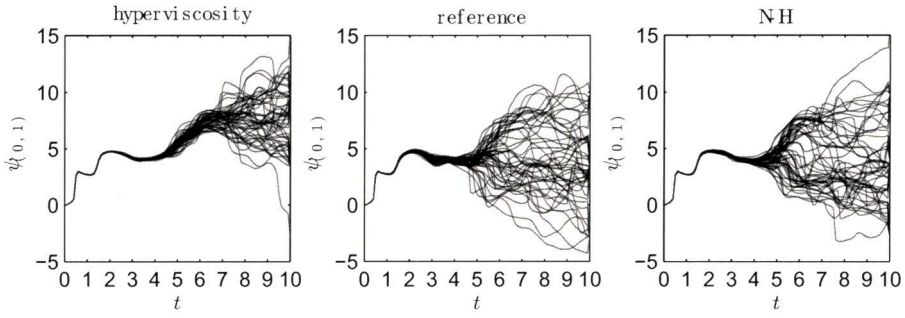


Figure 5.10: Ensemble dispersion in a simulation with small scale forcing. Plotted is the phase angle of the (0,1)-mode using (left) the model with artificial viscosity, (centre) the reference simulation, and (right) the Nosé-Hoover as indicated with † in table 5.1.

for the large scale forcing. Figure 5.10 compares the reference (center) to the hyperviscous (left) and Nosé-Hoover control (right) methods. The reference and Nosé-Hoover controlled ensembles show significant decorrelation at about time $t = 5$, whereas for the model with hyperviscosity, the decorrelation is delayed until time $t = 7$ or $t = 8$. In the reference there is a notable split of the ensembles around $t = 5$ into two main branches. This split can also be observed in the Nosé-Hoover approach, but not in the hyperviscosity model. At the final time $t = 10$, both reduced models are slightly underdispersive.

5.7 Discussion

We have shown that the Nosé-Hoover method can be used to enforce a target background kinetic energy spectrum in 2D turbulence models with stochastic forcing, even when truncated well below the viscous scale. The parameterization comes at the mild cost of one additional dynamic variable for each energy shell controlled.

In the experiments reported in the previous section, the target spectrum was inferred from a high resolution simulation, but it is important to emphasize that the target spectrum could also be taken from experiments or theory. In particular, the method described here could be developed to enforce a k^{-3} spectrum in low resolution models, if so desired. The approach makes no explicit use of two-dimensional structure and hence is potentially extensible to 3D turbulence.

Fröhlich & Schneider [57] simulate two-dimensional turbulent dynamics using a wavelet basis. In this setting applying our control method might be even more effective, as the action can be restricted to the homogenous regions

away from coherent structures.

Sukoriansky et al. [135] observe that the large scale dissipation influences the spectral slope in the forward enstrophy cascade. This means that it might be useful to apply a control on this end of the spectrum to perform simulations that display the hypothesized k^{-3} spectrum in the high wavenumbers.

Curriculum Vitæ

Keith Myerscough was born in November 1985 in Eindhoven, The Netherlands. He attended Stedelijk College Eindhoven from 1998 to 2004, following a bilingual VWO programme. After this, he studied Aerospace Engineering at Delft University of Technology, obtaining a Bachelor of Science in 2007. As a part of the master's programme he did an internship at TNO Automotive Safety Solutions, working on a new gas flow solver for use in simulating airbag inflation. His thesis was on differential equations with stochastic parameters and he received his master's degree, with honours, in October 2010.

Starting the next month, Keith worked towards a PhD in applied mathematics under the supervision of Jason Frank at Centrum Wiskunde & Informatica. Part of this research was conducted at the University of Edinburgh, working together with Ben Leimkuhler and his research group. The topics of his research included, among others, Hamiltonian and Poisson systems, fluid dynamics, geometric numerical integration, stochastic differential equations, and information theory. This work resulted in the three submitted articles that form the basis of this thesis.

After completing the research for his PhD, Keith took a position at KU Leuven in April 2015, researching multiscale problems in the research group of Giovanni Samaey.

Bibliography

[1] Nader verklaard: Hittegolf. http://www.knmi.nl/cms/content/28253/hittegolf. Accessed: 2015-02-25.

[2] R. V. Abramov and A. J. Majda. Statistically relevant conserved quantities for truncated quasigeostrophic flow. *Proceedings of the National Academy of Sciences*, 100(7):3841–3846, 2003.

[3] N. Agmon, Y. Alhassid, and R. D. Levine. An algorithm for finding the distribution of maximal entropy. *Journal of Computational Physics*, 30 (2):250–258, 1979.

[4] M. Allen and D. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, 1989. ISBN:978-0198556459.

[5] J. D. Anderson. *Fundamentals of aerodynamics*. Tata McGraw-Hill Education, 1985.

[6] H. Aref. 150 years of vortex dynamics. *Theoretical and Computational Fluid Dynamics*, 24(1-4):1–7, 2010. ISSN 0935-4964. doi: 10.1007/s00162-009-0178-6.

[7] V. I. Arnold. *Topological methods in hydrodynamics*, volume 125. Springer Science & Business Media, 1998.

[8] J. Bajārs. *Geometric Integration and Thermostat Methods for Hamiltonian Systems*. PhD thesis, Universiteit van Amsterdam, 2012.

[9] J. Bajars, J. Frank, and B. Leimkuhler. Stochastic-dynamical thermostats for constraints and stiff restraints. *Eur. Phys. J. Special Topics*, 200:131–152, 2011.

[10] J. Bajars, J. Frank, and B. Leimkuhler. Weakly coupled heat bath models for Gibbs-like invariant states in nonlinear wave equations. *Nonlinearity*, 26(7):1945, 2013.

[11] E. J. Barth, B. B. Laird, and B. J. Leimkuhler. Generating generalized distributions from dynamical simulation. *The Journal of Chemical Physics*, 118(13):5759–5768, 2003.

[12] G. K. Batchelor. Computation of the energy spectrum in homogeneous two-dimensional turbulence. *Physics of Fluids (1958-1988)*, 12(12):II–233, 1969.

[13] J. Berner, G. J. Shutts, M. Leutbecher, and T. N. Palmer. A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ecmwf ensemble prediction system. *Journal of the Atmospheric Sciences*, 66(3):603–626, 2009.

[14] A. Bloch, P. Krishnaprasad, J. Marsden, and T. Ratiu. The euler-poincarÃĭ equations and double bracket dissipation. *Communications in Mathematical Physics*, 175(1):1–42, 1996.

[15] G. Bofetta and S. Musacchio. Evidence for the double cascade scenario in two-dimensional turbulence. *Physical Review E*, 82, 2010.

[16] F. Bouchet and A. Venaille. Statistical mechanics of two-dimensional and geophysical flows. *Physics Reports*, 515(5):227–295, 2012.

[17] J. P. Boyd. *Chebyshev and Fourier spectral methods*. Courier Dover Publications, 2013.

[18] J. P. Boyd and C. Zhou. Three ways to solve the poisson equation on a sphere with Gaussian forcing. *Journal of Computational Physics*, 228 (13):4702–4713, 2009.

[19] A. Bracco, J. McWilliams, G. Murante, A. Provenzale, and J. Weiss. Revisiting freely decaying two-dimensional turbulence at millennial resolution. *Physics of Fluids (1994-present)*, 12(11):2931–2941, 2000.

[20] O. Bühler. Statistical mechanics of strong and weak point vortices in a cylinder. *Phys. of Fluids*, 14(7):2139–2149, 2002.

[21] O. Bühler. *A Brief Introduction to Classical, Statistical and Quantum Mechanics*. Courant Lect. Notes. Am. Math. Soc./Courant Inst. of Math. Sci., Providence, R.I., first edition, 2006.

[22] A. Bulgac and D. Kusnezov. Canonical ensemble averages from pseudomicrocanonical dynamics. *Phys. Rev. A*, 42:5045–5048, Oct 1990. doi: 10.1103/PhysRevA.42.5045.

[23] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral Methods: Fundamentals in Single Domains*. Springer Verlag, 2006.

[24] G. Chamoun, E. Kanso, and P. K. Newton. Von Kármán vortex streets on the sphere. *Physics of Fluids (1994-present)*, 21(11):116603, 2009.

[25] P. Chatelain, A. Curioni, M. Bergdorf, D. Rossinelli, W. Andreoni, and P. Koumoutsakos. Billion vortex particle direct numerical simulations of aircraft wakes. *Computer Methods in Applied Mechanics and Engineering*, 197(13):1296–1304, 2008.

[26] P. Chatelain, A. Curioni, M. Bergdorf, D. Rossinelli, W. Andreoni, and P. Koumoutsakos. Vortex methods for massively parallel computer architectures. In *High Performance Computing for Computational Science-VECPAR 2008*, pages 479–489. Springer, 2008.

[27] P.-H. Chavanis. Kinetic theory of point vortices: Diffusion coefficient and systematic drift. *Physical Review E*, 64(2):026309, 2001.

[28] P.-H. Chavanis, J. Sommeria, and R. Robert. Statistical mechanics of two-dimensional vortices and collisionless stellar systems. *The Astrophysical Journal*, 471(1):385, 1996.

[29] A. J. Chorin, J. E. Marsden, and J. E. Marsden. *A mathematical introduction to fluid mechanics*, volume 3. Springer, 1990.

[30] C. Cotter and G. Pavliotis. Estimating eddy diffusivities from noisy lagrangian observations. *Communications in Mathematical Sciences*, 7 (4):805–838, 2009.

[31] G.-H. Cottet and P. D. Koumoutsakos. *Vortex methods: theory and practice*. Cambridge university press, 2000.

[32] G.-H. Cottet and P. D. Koumoutsakos. *Vortex Methods*. Camb. Univ. Press, Camb., UK, first edition, 2000.

[33] S. D. Danilov and D. Gurarie. Quasi-two-dimensional turbulence. *Physics-uspekhi*, 43(9):863, 2000.

[34] P. A. Davidson. *Turbulence: An Introduction for Scientists and Engineers*. Oxford University Press, 2004.

[35] S. Davis and G. Gutiérrez. Conjugate variables in continuous maximum-entropy inference. *Physical Review E*, 86(5):051136, 2012.

[36] S. Delong, B. E. Griffith, E. Vanden-Eijnden, and A. Donev. Temporal integrators for fluctuating hydrodynamics. *Physical Review E*, 87(3): 033302, 2013.

[37] T. DelSole. A fundamental limitation of Markov models. *Journal of the Atmospheric Sciences*, 57(13):2158–2168, 2000.

[38] R. C. Dewar, C. H. Lineweaver, R. K. Niven, and K. Regenauer-Lieb, editors. *Beyond the Second Law: Entropy Production and Non-equilibrium Systems*. Understanding Complex Systems. Springer, 2014.

[39] N. Ding, Y. Fang, R. Babbush, C. Chen, R. D. Skeel, and H. Neven. Bayesian sampling using stochastic gradient thermostats. In *Advances in Neural Information Processing Systems (NIPS)*, 2014.

[40] J. A. Domaradzki and E. M. Saiki. Backscatter models for large-eddy simulations. *Theoretical and computational fluid dynamics*, 9(2):75–83, 1997.

[41] A. Donev, E. Vanden-Eijnden, A. Garcia, and J. Bell. On the accuracy of finite-volume schemes for fluctuating hydrodynamics. *Communications in Applied Mathematics and Computational Science*, 5(2):149–197, 2010.

[42] S. Dubinkina and J. Frank. Statistical mechanics of Arakawa's discretizations. *Journal of Computational Physics*, 227(2):1286–1305, 2007.

[43] S. Dubinkina and J. Frank. Statistical relevance of vorticity conservation in the hamiltonian particle-mesh method. *Journal of Computational Physics*, 229(7):2634–2648, 2010.

[44] S. Dubinkina, J. Frank, and B. Leimkuhler. Simplified modelling of a thermal bath system, with application to a fluid vortex system. *Multiscale Model Simul.*, 8(5):1882–1901, 2010.

[45] R. Durrett. *Stochastic calculus: a practical introduction*, volume 6. CRC press, 1996.

[46] W. E and J. C. Mattingly. Ergodicity for the navier-stokes equation with degenerate random forcing: Finite-dimensional approximation. *Communications on Pure and Applied Mathematics*, 54(11):1386–1402, 2001.

[47] W. E, J. Mattingly, and Y. Sinai. Gibbsian dynamics and ergodicity for the stochastically forced navier-stokes equation. *Communications in Mathematical Physics*, 224(1):83–106, 2001.

[48] G. Eyink and H. Spohn. Negative-temperature states and large-scale, long-lived vortices in two-dimensional turbulence. *Journal of statistical physics*, 70(3-4):833–886, 1993.

[49] L. Faddeev and L. Takhtajan. *Hamiltonian Methods in the Theory of Solitons*. Springer Series in Soviet Mathematics. Springer-Verlag, New York, 1987.

[50] M. Farazmand, N.-R. Kevlahan, and B. Protas. Controlling the dual cascade of two-dimensional turbulence. *Journal of Fluid Mechanics*, 668: 202–222, 2011.

[51] M. Farge, N. Kevlahan, V. Perrier, and E. Goirand. Wavelets and turbulence. *Proceedings of the IEEE*, 84(4):639–669, 1996.

[52] I. Fatkullin and E. Vanden-Eijnden. A computational strategy for multiscale systems with applications to lorenz 96 model. *Journal of Computational Physics*, 200(2):605–638, 2004.

[53] J. Frank. Geometric space-time integration of ferromagnetic materials. *Appl. Numer. Math.*, 48(3-4):307–322, 2004. ISSN 0168-9274. doi: 10.1016/j.apnum.2003.11.003.

[54] J. Frank and G. Gottwald. The Langevin limit of the Nosé-Hoover-Langevin thermostat. *Journal of Statistical Physics*, 143(4):715–724, 2011.

[55] J. Frank, W. Huang, and B. Leimkuhler. Geometric integrators for classical spin systems. *J. Comput. Phys.*, 133(1):160–172, 1997. ISSN 0021-9991. doi: 10.1006/jcph.1997.5672.

[56] U. Frisch. *Turbulence: the legacy of AN Kolmogorov*. Cambridge University Press, 1995.

[57] J. Fröhlich and K. Schneider. Computation of decaying turbulence in an adaptive wavelet basis. *Physica D: Nonlinear Phenomena*, 134(3):337–361, 1999.

[58] T. C. Gard. *Introduction to stochastic differential equations*. M. Dekker New York, 1988.

[59] T. Gotoh. Energy spectrum in the inertial and dissipation ranges of two-dimensional steady turbulence. *Physical Review E*, 57(3):2984, 1998.

[60] I. Grooms and A. J. Majda. Efficient stochastic superparameterization for geophysical turbulence. *Proceedings of the National Academy of Sciences*, 110(12):4464–4469, 2013.

[61] M. J. Grote and A. J. Majda. Crude closure dynamics through large scale statistical theories. *Physics of Fluids (1994-present)*, 9(11):3431–3442, 1997.

[62] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*, volume 31. Springer, 2006.

[63] M. Hairer and J. Mattingly. Yet another look at Harris' ergodic theorem for markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, pages 109–117. Springer, 2011.

[64] H. Haken. *Information and Self-Organization: A Macroscopic Approach to Complex Systems*. Springer, 2006.

[65] K. Hasselmann. Stochastic climate models part i. theory. *Tellus*, 28(6): 473–485, 1976.

[66] W. Hayes and K. R. Jackson. A fast shadowing algorithm for high dimensional ODE systems. *SIAM Journal on Scientific Computing*, 29(4): 1738–1758, 2007.

[67] H. Helmholtz. On integrals of the hydrodynamical equations which express vortex motion. *Phil. Mag*, 33:485–512, 1867.

[68] D. D. Holm. *Geometric mechanics*, volume 2. Imperial College Press London, 2008.

[69] P. Holmes, J. L. Lumley, and G. Berkooz. *Turbulence, coherent structures, dynamical systems and symmetry*. Cambridge University Press, 1998.

[70] W. G. Hoover. Canonical dynamics: equilibrium phase-space distributions. *Physical Review A*, 31(3):1695, 1985.

[71] U. Hornung. *Homogenization and porous media*, volume 6. Springer Verlag, 1997.

[72] M. I. Jamaloodeen and P. K. Newton. The N-vortex problem on a rotating sphere. II. heterogeneous platonic solid equilibria. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 462, pages 3277–3299. The Royal Society, 2006.

[73] E. T. Jaynes. Information theory and statistical mechanics. *Physical review*, 106(4):620, 1957.

[74] E. T. Jaynes. Information theory and statistical mechanics. II. *Phys. Rev.*, 108:171–190, Oct 1957. doi: 10.1103/PhysRev.108.171.

[75] A. Jones and B. Leimkuhler. Adaptive stochastic methods for sampling driven molecular systems. *The Journal of Chemical Physics*, 135 (8):084125, 2011. doi: http://dx.doi.org/10.1063/1.3626941.

[76] G. Joyce and D. Montgomery. Simulation of the "negative temperature" instability for line vortices. *Physics Letters A*, 39(5):371–372, 1972.

[77] G. Joyce and D. Montgomery. Negative temperature states for the two-dimensional guiding-centre plasma. *Journal of Plasma Physics*, 10(01): 107–121, 1973.

[78] S. Kehrbaum and J. H. Maddocks. Elasitc rods, rigid bodies, quaternions and the last quadrature. *Phil. Trans. R. Soc. Lond. A*, 355(1732): 2117–2136, 1997.

[79] J. Kent, J. Thuburn, and N. Wood. Assessing implicit large eddy simulation for two-dimensional flow. *Quarterly Journal of the Royal Meteorological Society*, 138(663):365–376, 2012.

[80] R. Kidambi and P. K. Newton. Motion of three point vortices on a sphere. *Physica D: Nonlinear Phenomena*, 116(1):143–175, 1998.

[81] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23. Springer Science & Business Media, 1992.

[82] R. H. Kraichnan. Inertial ranges in two-dimensional turbulence. *Physics of Fluids*, 10:1417–1423, 1967.

[83] R. H. Kraichnan. Inertial-range transfer in two-and three-dimensional turbulence. *Journal of Fluid Mechanics*, 47(03):525–535, 1971.

[84] H. Lamb. *Hydrodynamics*. Dover Publications, New York, sixth edition, 1932.

[85] B. Leimkuhler. Generalized Bulgac-Kusnezov methods for sampling of the Gibbs-Boltzmann measure. *Physical Review E*, 81(2):026703, 2010.

[86] B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*. Camb. Monogr. on Appl. and Comput. Math. Camb. Univ. Press, Camb., U.K., first edition, 2004.

[87] B. Leimkuhler, E. Noorizadeh, and F. Theil. A gentle thermostat for molecular dynamics. *J. Stat. Phys.*, 135(2):261–277, 2009.

[88] B. Leimkuhler, E. Noorizadeh, and O. Penrose. Comparing the efficiencies of stochastic isothermal molecular dynamics models. *J. Stat. Phys.*, 143(5):921–942, 2011.

[89] C. E. Leith. Diffusion approximation for two-dimensional turbulence. *Physics of Fluids (1958-1988)*, 11(3):671–672, 1968.

[90] E. Lindborg. Can the atmospheric kinetic energy spectrum be explained by two-dimensional turbulence? *Journal of Fluid Mechanics*, 388:259–288, 1999.

[91] P.-L. Lions and A. Majda. Equilibrium statistical theory for nearly parallel vortex filaments. *Communications on Pure and Applied Mathematics*, 53(1):76, 2000.

[92] C.-S. Liu, K.-C. Chen, and C.-S. Yeh. A mathematical revision of the Landau-Lifshitz equation. *Journal of Marine Science and Technology*, 17 (3):228–237, 2009.

[93] A. J. Majda and A. L. Bertozzi. *Vorticity and Incompressible Flow*. Camb. Texts in Appl. Math. Camb. Univ. Press, Camb., UK, first edition, 2002.

[94] A. J. Majda and B. Gershgorin. Improving model fidelity and sensitivity for complex systems through empirical information theory. *Proceedings of the National Academy of Sciences*, 108(25):10044–10049, 2011.

[95] A. J. Majda and X. Wang. *Nonlinear Dynamics and Statistical Theories for Basic Geophysical Flows*. Camb. Univ. Press, Camb., UK, first edition, 2006.

[96] A. J. Majda, I. Timofeyev, and E. Vanden Eijnden. A mathematical framework for stochastic climate models. *Communications on Pure and Applied Mathematics*, 54(8):891–974, 2001.

[97] J. C. Mattingly, A. Stuart, and D. D.J. Higham. Ergodicity for sdes and approximations: locally lipschitz vector fields and degenerate noise. *Stochastic processes and their applications*, 101(2):185–232, 2002.

[98] R. I. McLachlan. Symplectic integration of Hamiltonian wave equations. *Numerische Mathematik*, 66(1):465–492, 1993.

[99] R. I. McLachlan. On the numerical integration of ordinary differential equations by symmetric composition methods. *SIAM Journal on Scientific Computing*, 16(1):151–168, 1995.

[100] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.

[101] Y. Meyer. *Wavelets and operators*, volume 1. Cambridge university press, 1995.

[102] J. Miller. Statistical mechanics of euler equations in two dimensions. *Physical Review Letters*, 65(17):2137, 1990.

[103] G. N. Milstein. *Numerical integration of stochastic differential equations*, volume 313. Springer Science & Business Media, 1995.

[104] H. Mori. Transport, collective motion, and brownian motion. *Progress of Theoretical Physics*, 33(3):423–455, 1965.

[105] P. J. Morrison. Poisson brackets for fluids and plasmas. *Mathematical methods in hydrodynamics and integrability in dynamical systems*, 88:13–46, 1982.

[106] K. W. Myerscough, J. Frank, and B. Leimkuhler. Least-biased correction of extended dynamical systems using observational data. *arXiv preprint arXiv:1411.6011*, 2014.

[107] G. Nastrom and K. Gage. A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. *Journal of the atmospheric sciences*, 42(9):950–960, 1985.

[108] G. D. Nastrom, K. S. Gage, and W. H. Jasperson. Kinetic energy spectrum of large-and mesoscale atmospheric processes. *Nature*, 310:36–38, 1984.

[109] P. K. Newton. *The N-Vortex Problem: Analytical Techniques*. Springer-Verlag, New York, first edition, 2001.

[110] P. K. Newton. Point vortex dynamics in the post-Aref era. *Fluid Dynamics Research*, 46(3):031401, 2014.

[111] P. K. Newton and T. Sakajo. The N-vortex problem on a rotating sphere. III. ring configurations coupled to a background field. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 463 (2080):961–977, 2007.

[112] P. K. Newton and T. Sakajo. The N-vortex problem on a rotating sphere: IV. Ring configurations coupled to a background field. *Proc. R. Soc. A*, 463(2080):961–977, 2007.

[113] S. Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molecular physics*, 52(2):255–268, 1984.

[114] S. Nosé. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*, 81(1):511–519, 1984.

[115] B. Øksendal. *Stochastic differential equations*. Springer, 2003.

[116] L. Onsager. Statistical hydrodynamics. *Il Nuovo Cimento (1943-1954)*, 6:279–287, 1949.

[117] G. Patrick. Dynamics of perturbed relative equilibria of point vortices on the sphere or plane. *Journal of Nonlinear Science*, 10(3):401–415, 2000.

[118] G. Pavliotis and A. Stuart. *Multiscale methods: averaging and homogenization*, volume 53. Springer, 2008.

[119] J. Pedlosky. Geophysical fluid dynamics. *New York and Berlin, Springer-Verlag, 1982. 636 p.*, 1, 1982.

[120] Y. B. Pointin and T. S. Lundgren. Statistical mechanics of two-dimensional vortices in a bounded container. *Physics of Fluids*, 19(10): 1459–1470, 1976. doi: http://dx.doi.org/10.1063/1.861347.

[121] L. M. Polvani and D. G. Dritschel. Wave and vortex dynamics on the surface of a sphere. *Journal of Fluid Mechanics*, 255:35–64, 1993.

[122] L. Rey-Bellet. Ergodic properties of markov processes. *Open Quantum Systems II*, pages 1–39, 2006.

[123] R. Robert. A maximum-entropy principle for two-dimensional perfect fluid dynamics. *Journal of Statistical Physics*, 65(3-4):531–553, 1991.

[124] R. Robert and J. Sommeria. Statistical equilibrium states for two-dimensional flows. *Journal of Fluid Mechanics*, 229:291–310, 1991.

[125] D. Rossinelli and P. Koumoutsakos. Vortex methods for incompressible flow simulations on the GPU. *The Visual Computer*, 24(7-9):699–708, 2008.

[126] D. Rossinelli, M. Bergdorf, G.-H. Cottet, and P. Koumoutsakos. GPU accelerated simulations of bluff body flows using vortex particle methods. *Journal of Computational Physics*, 229(9):3316–3333, 2010.

[127] R. Y. Rubinstein and D. P. Kroese. *Simulation and the Monte Carlo method*, volume 707. John Wiley & Sons, 2011.

[128] P. Saffman. On the spectrum and decay of random two-dimensional vorticity distributions at large Reynolds number (energy spectrum and decay of random two dimensional vorticity distributions at large Reynolds number). *Studies in Applied Mathematics*, 50:377–383, 1971.

[129] P. G. Saffman. *Vortex dynamics*. Cambridge university press, 1992.

[130] T. Sakajo. Non-self-similar, partial, and robust collapse of four point vortices on a sphere. *Physical Review E*, 78(1):016312, 2008.

[131] R. Salmon. Hamiltonian fluid mechanics. *Annual review of fluid mechanics*, 20(1):225–256, 1988.

[132] A. A. Samoletov, C. P. Dettmann, and M. A. Chaplain. Thermostats for "slow" configurational modes. *Journal of Statistical Physics*, 128(6): 1321–1336, 2007.

[133] T. Schlick. *Molecular Modeling and Simulation: An Interdisciplinary Guide (Interdisciplinary Applied Mathematics)*. Springer, 2nd edition, 2010. ISBN:978-1441963505.

[134] G. Shutts. A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society*, 131(612):3079–3102, 2005.

[135] S. Sukoriansky, B. Galperin, and A. Chekhlov. Large scale drag representation in simulations of two-dimensional turbulence. *Physics of Fluids (1994-present)*, 11(10):3043–3053, 1999.

[136] G. E. Swaters. *Introduction to Hamiltonian fluid dynamics and stability theory*, volume 102. CRC Press, 1999.

[137] J. Thuburn, J. Kent, and N. Wood. Cascades, backscatter and conservation in numerical models of two-dimensional turbulence. *Quarterly Journal of the Royal Meteorological Society*, 140(679):626–638, 2014.

[138] L. N. Trefethen. *Spectral methods in MATLAB*, volume 10. Siam, 2000.

[139] R. Tulloch and K. Smith. A theory for the atmospheric energy spectrum: Depth-limited temperature anomalies at the tropopause. *Proceedings of the National Academy of Sciences*, 103(40):14690–14694, 2006.

[140] J. Vankerschaver and M. Leok. A novel formulation of point vortex dynamics on the sphere: geometrical and numerical aspects. *Journal of Nonlinear Science*, 24(1):1–37, 2014.

[141] G. Winckelmans, R. Cocle, L. Dufresne, and R. Capart. Vortex methods and their application to trailing wake vortex simulations. *Comptes Rendus Physique*, 6(4):467–486, 2005.

[142] Y. Xing, A. J. Majda, and W. W. Grabowski. New efficient sparse space-time algorithms for superparameterization on mesoscales. *Monthly Weather Review*, 137(12):4307–4324, 2009.

[143] H. Yoshida. Construction of higher order symplectic integrators. *Physics Letters A*, 150(5):262–268, 1990.

[144] M.-Q. Zhang and M.-Z. Qin. Explicit symplectic schemes to solve vortex systems. *Computers Math. Applic.*, 26(5):51–56, 1993.

[145] R. Zwanzig. Memory effects in irreversible thermodynamics. *Physical Review*, 124(4):983, 1961.

Propositions

with the thesis of Keith W. Myerscough

1. Gentle stochastic perturbations to dynamical systems can be used to yield simulations that have truthful statistical qualities on a long time scale, while maintaining dynamical accuracy on a short time scale. (Chapter 3)

2. Lie-Trotter splitting, and derived methods, applied to Poisson systems with pairwise interacting particles give Poisson integrators if the solution to the two-particle system can be expressed exactly. (Chapter 4)

3. Such splittings are easily performed in parallel, as there is a large number of operations involved that commute. (Chapter 4)

4. The small-scale energy balance in truncated two-dimensional turbulence simulations can be controlled by a feedback control that enforces the energy spectrum associated with the forward energy cascade. (Chapter 5)

5. The learning of a programming language often mimics its historical development.

6. A general answer to the question which problems merit a general solution would be invaluable.

7. Birds are intuitively aware that the path of least action in a gravitational field is not a straight line.

8. When learning a foreign language, the moment others reply to you in that language forms a crucial tipping point.

9. If you do not know how it works, social security is primarily a source of insecurity.

10. Being allowed to levy parking fees is a perverse incentive that prevents councils from dissuading car usage and/or implementing alternatives.

These propositions are considered defensible and as such have been approved by the supervisor, prof.dr.ir. J.E. Frank.

Stellingen

bij het proefschrift van Keith W. Myerscough

1. Zachte stochastische aanpassingen aan dynamische systemen kunnen gebruikt worden om simulaties te verkrijgen met waarheidsgetrouwe statistische eigenschappen op een lange tijdschaal, terwijl de dynamica op kleine tijdschaal nauwkeurig blijft. (Hoofdstuk 3)

2. Lie-Trotter splitting, en afgeleide methoden, toegepast op Poisson systemen met paargewijs interagerende deeltjes geven Poisson tijdstappers als men de oplossing van het systeem met twee deeltjes exact kan uitdrukken. (Hoofdstuk 4)

3. Zulke splittings zijn gemakkelijk in parallel uit te voeren, omdat er een groot aantal commuterende bewerkingen bij betrokken is. (Hoofdstuk 4)

4. De energiebalans op kleine schaal in afgekapte twee-dimensionale turbulentiesimulaties kan worden geregeld met een terugkoppeling die het energiespectrum behorende bij de voorwaartse energiecascade vastlegt. (Hoofdstuk 5)

5. Het leren van een programmeertaal weerspiegelt vaak de historische ontwikkeling ervan.

6. Een algemeen antwoord op de vraag welke problemen een algemene aanpak rechtvaardigen, zou van onschatbare waarde zijn.

7. Vogels zijn zich van nature bewust dat het pad van minimale arbeid in een zwaartekrachtveld geen rechte lijn is.

8. Bij het leren spreken van een vreemde taal vormt het moment dat anderen je in die taal beantwoorden een cruciaal kantelpunt.

9. Als je niet weet hoe het werkt, is sociale zekerheid vooral een bron van onzekerheid.

10. Het mogen innen van parkeeropbrengsten is een perverse prikkel die gemeenten weerhoudt autogebruik te ontmoedigen en/of alternatieven te implementeren.

Deze stellingen worden verdedigbaar geacht en zijn als zodanig goedgekeurd door de promotor, prof.dr.ir. J.E. Frank