Eindhoven University of Technology

MASTER

Invariant manifolds and their role in mixing, with applications to point vortex systems

Cottaar, J.

Award date:
2008

Faculteit Technische
Natuurkunde
Groep Transportfysica
Gebouw Cascade
Postbus 513
5600 MB Eindhoven

Title: **Invariant manifolds and their role in mixing, with applications to point vortex systems**

Author: J. Cottaar

Report number: R-1740-A

Date: November 25, 2008

Departments: Turbulence and Vortex Dynamics Group
Department of Applied Physics

Centre for Analysis, Scientific computing and Applications
Department of Mathematics and Computer Science

Eindhoven University of Technology
The Netherlands

Supervisors: dr. ir. L.P.J. Kamp
prof. dr. R.M.M. Mattheij

Abstract

Laminar mixing is an important phenomenon both in geophysics and in industrial mixing applications. However, a method to completely describe the mixing properties of such flows is not yet available. In this work we focus on the relationship between invariant manifolds and mixing in two-dimensional incompressible flows. We develop the theory analytically and demonstrate it numerically using a flow which can often be used as an idealized model of such flows: the point vortex system.

Mixing in 2D incompressible flows tends to be limited to subsets of the phase space. Such a mixing subset is called an ergodic region. A small cloud of particles somewhere in such a region will eventually mix to cover the whole region uniformly. The time this takes depends on the initial location, a property which induces a structure on the phase space: the mixing time distribution. There is an intermediate regime where the cloud is no longer localized but is not distributed uniformly yet. In this regime it exhibits a typical structure independent of the starting location: the partial mixing structure. It turns out that the spatial distribution of the *stable* invariant manifold describes the mixing time distribution, while that of the *unstable* invariant manifold describes the partial mixing structure.

The concept connecting the invariant manifold to the mixing structures is the w-measure. This is a probability measure on the ergodic region based on the spatial variation of finite time Lyapunov exponents. It can be linked directly to the mixing structures, while on the other hand it exactly describes the spatial distribution of the invariant manifolds. This leads to two numerical methods to find this measure, one based on computing Lyapunov exponents and the other on computing invariant manifolds. Both methods are relevant, since the first performs well in regions of *low* w-measure while the second performs well in regions of *high* w-measure.

Contents

1 Introduction

The work described in this report was conducted at the Eindhoven University of Technology between August 2007 and November 2008. It is a continuation of the work of Ronald Otten [60], which focused on the existence and computation of invariant manifolds of distinguished hyperbolic trajectories in 2D incompressible flows. The goal of the current work is to improve our understanding of these manifolds and their physical relevance, particularly their use in describing the mixing process.

In this chapter, a general overview of current research into laminar mixing will first be given (section 1.1). Next, we outline the current work and very briefly introduce the most important concepts (section 1.2). Finally, the structure of the remainder of the report will be described (section 1.3).

1.1 Background

The focus of this work is the study of mixing in laminar flows. Although several concepts had long been known, this field did not really take off until the 1980s with Aref's introduction of the term 'chaotic advection' [6]. The central issue is the appearance of features in advected markers on a scale much smaller than the flow scale, as seen for instance in early experiments by Welander [74]. In other words, a flow can exhibit Lagrangian chaos without Eulerian chaos [23]. Apart from being quite an interesting phenomenon, this has several applications. For instance, in materials processing laminar mixing is necessary when turbulent mixing is simply not possible due to high viscosity, such as with polymer solutions [55]. Another application is in geophysics and environmental sciences, for instance in the description of the large scale transport properties of oceanic flows [65] or the local mixing of pollutants [48].

In this work we consider only two-dimensional incompressible flows. This choice is necessary because it leads to a Hamiltonian description of the particle motion which does not exist for 3D flows. This reduces the applicability to industrial mixing processes since they are often three-dimensional, although it may be possible to generalize our results to apply to *stationary* 3D flows [9, 26, 41]. On the other hand, in atmospheric and oceanic flows the requirement of a 2D incompressible flow is often satisfied. There are three primary reasons for this. Firstly, vertical velocities are generally much smaller than horizontal velocities, as a result of which the velocity field *at a fixed level* is often two-dimensional and incompressible [66]. Secondly, due to background rotation many of these flows are geostrophic and so, by the Taylor-Proudman theorem, are independent of the vertical coordinate [44]. Finally, there is often a density stratification present. This inhibits vertical velocities, since gravity confines fluid elements to layers of constant density [67].

There are several approaches to the study of transport and mixing in these flows. These include the detection of transport barriers [15], characterization as a diffusion process and analysis of the anomalies in such an approach [70], analysis of Poincaré maps [46], computation of Lyapunov exponents and analysis of their spatial distribution [1, 10, 69] and the computation and analysis of the stable and unstable manifolds of hyperbolic trajectories [33, 39, 51]. It is this final approach that forms the starting point for this work.

1.2 Outline

The traditional application of the stable and unstable manifolds mentioned above is in lobe dynamics, where the transport between areas bounded by the stable and unstable manifolds is characterized using turnstile mechanisms [50, 62]. We will use a very different approach. We are not so much interested in the specific shapes of parts of the manifold, but rather in their asymptotic space filling properties. We also do not consider the stable and unstable manifold simultaneously, but show how each one separately describes aspects of the mixing process. The stable manifold does so for forward time, while the unstable manifold describes the process in reverse time. In particular, they will be related to two important structures: the partial mixing structure and the mixing time distribution. We will also demonstrate a quantitative link to Lyapunov exponents [21].

So, the central point is the relationship between the following concepts:

- The invariant manifolds: The stable manifold of a trajectory is the set of all trajectories that converge to it for forward time, while the unstable manifold is the set of all trajectories that converge to it for reverse time. Under certain conditions we may speak of global stable and unstable manifolds of the flow without considering any specific trajectory.

- The w-measure: This is a rigorous probability measure on the phase space or some subset of it, based on the spatial variation of finite time Lyapunov exponents. It can be defined both for forward and for reverse time, leading to different measures.

- The partial mixing structure: This structure is not rigorously defined. It is found by considering a number of particles which have started mixing, but are not yet uniformly distributed. Their distribution turns out to be characteristic of the flow and independent of their starting location. The partial mixing structure is related to the negative time w-measure and to the global unstable manifold.

- The mixing time distribution: The time a cloud of particles takes to mix depends on the starting location. In this way, a structure is induced on the phase space: the mixing time distribution. It is related to the positive time w-measure and to the global stable manifold.

To give an idea of the similarity between these concepts, examples of each of them are shown in Figures 1.1.

There are strong parallels between this work and that of Adrover *et. al.* [2, 4, 16], which also considers the measure theoretical properties of the global unstable manifold. However, we expand on their work in several ways. In the first place, we make a link to the actual mixing process, through the partial mixing structure and the mixing time distribution. Also, we explicitly make the connection to Lyapunov exponents and analytically show their relationship to the spatial distribution of the global invariant manifolds. Finally, our theory and methods are not limited to periodic systems.

Various point vortex systems will be used to demonstrate the theory. These are flows with infinite Reynolds number in which all vorticity is concentrated in a finite number of points according to a δ-distribution. The lack of viscosity, which follows from the infinite Reynolds number, prevents the vorticity from diffusing. Although an idealization, these systems are useful in describing flows consisting of vortices which are small compared to the flow scale.

1.3 Report structure

Since the theory will be supported by the results of numerical simulations in point vortex systems as we go along, we start by considering the properties of these systems in detail in chapter 2. This chapter focuses on the behavior of the point vortices themselves without considering the flow induced by them yet. Considerable attention is devoted to the development of accurate numerical methods for calculating this behavior. Six typical vortex systems are analyzed in detail to serve as examples in later chapters.

Next, a general overview of advection in point vortex systems is presented in chapter 3, including an expansion of the numerical methods to efficiently handle tracer particles. The most important result is the presence in some of the systems of ergodic regions, subsets of the phase space in which mixing takes place. These are considered in more detail in chapter 4. We take a closer look at just what an ergodic region is and what its mixing properties are. We will also encounter the mixing time distribution and the partial mixing structure for the first time here.

Chapter 5 deals with the definition, properties and computation of Lyapunov exponents and the w-measure induced by them. We also demonstrate the link between the w-measure on the one hand and the mixing time distribution and partial mixing structure on the other hand, both analytically and numerically. Invariant manifolds and their relationship to the w-measure are considered in chapter 6. Our conclusions finally are presented in chapter 7, along with some possible avenues for further research.

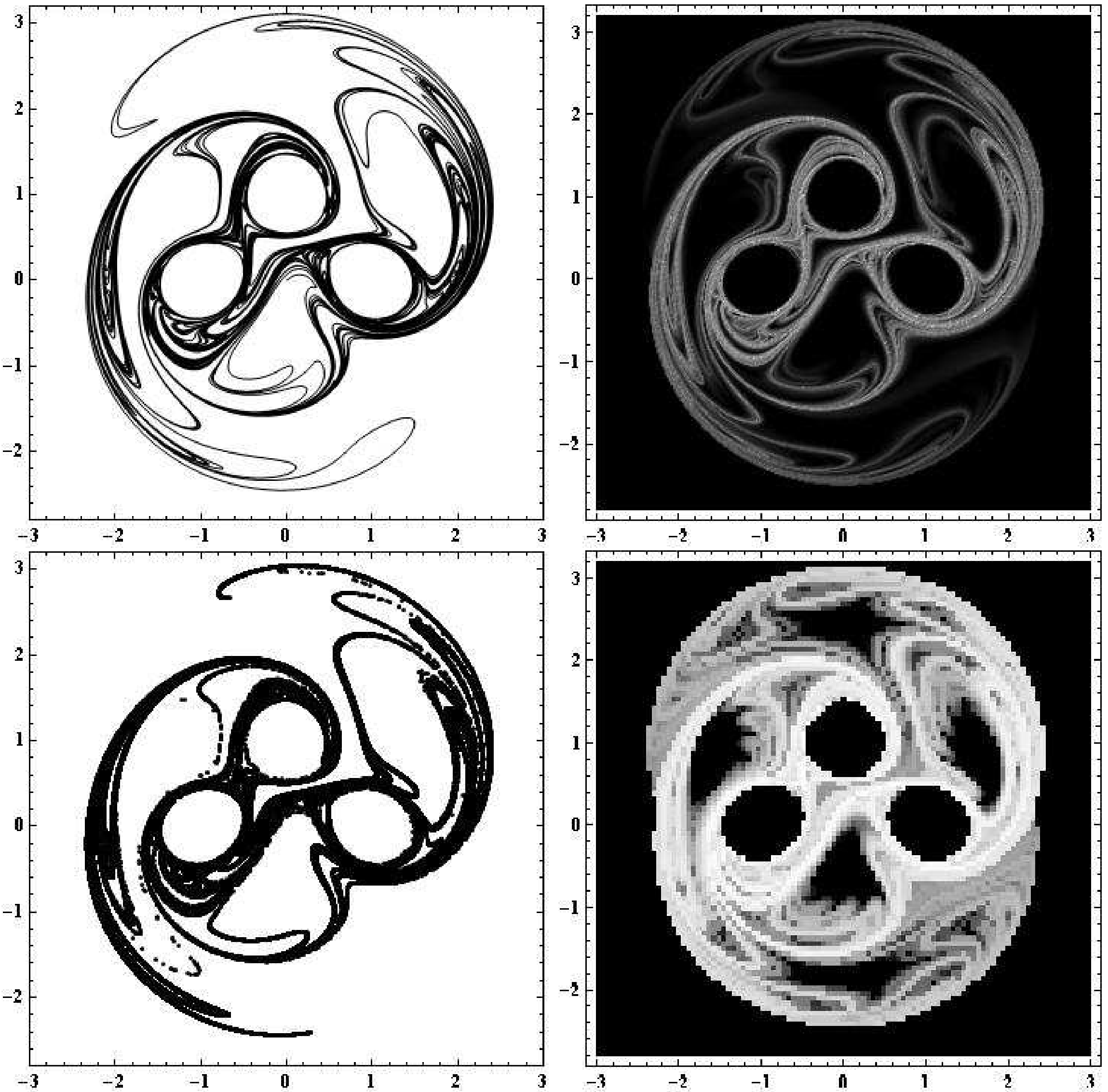

Figure 1.1: Various structures in a typical point vortex system. From the top left going horizontally: stable manifold (Figure 6.4), positive time w-measure (Figure 5.6), partial mixing structure mirrored in the y-axis (Figure 4.1) and mixing time distribution (also Figure 4.1).

2 Point vortices

In this chapter the physical interest and behavior of point vortex systems will be discussed, and numerical methods to obtain the motion of these point vortices will be introduced. This is actually quite a rich topic since as we will see the vortices themselves can exhibit chaotic motion within their phase space. The aim of this report however is the description of the *advection* of passive particles in these systems, and therefore this treatment will be fairly brief.

First the vorticity equation will be derived and discussed (section 2.1.1) and then applied to define and describe the point vortex system (section 2.1.2). Next the numerical methods to evolve such a system in time will be introduced (section 2.2) and finally several three and four point vortex systems which will be used in the rest of this report will be discussed, showcasing various possible behaviors (section 2.3).

2.1 Theory

2.1.1 The vorticity equation

The flow of a fluid in three dimensions is described by the laws of mass and momentum conservation. These are in differential form [44]:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \tag{2.1}$$

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = \frac{1}{\rho}(\nabla p + \mathbf{B}) + \nu \nabla^2 \mathbf{v}, \tag{2.2}$$

where ρ is the density, $\mathbf{v}$ the velocity, p the pressure and ν the kinematic viscosity of the fluid, and $\mathbf{B}$ is the body force density, all of which are in general functions of space and time.

We make several assumptions. In the first place the density ρ is constant, neither depending on time nor space. This means that (2.1) reduces to

$$\nabla \cdot \mathbf{v} = 0. \tag{2.3}$$

The flow is also taken inviscid:

$$\nu = 0, \tag{2.4}$$

eliminating the last term from (2.2). Next the body forces are assumed to be conservative, i.e.

$$\nabla \times \mathbf{B} = 0. \tag{2.5}$$

We limit our attention to two-dimensional flows, implying

$$v_z = 0 \tag{2.6}$$

and

$$\frac{\partial}{\partial z} = 0. \tag{2.7}$$

Define the vorticity $\boldsymbol{\omega}$ as the curl of the velocity field,

$$\boldsymbol{\omega} := \nabla \times \mathbf{v}. \tag{2.8}$$

Note that it follows from (2.6) and (2.7) that the x- and y-components of ω are zero, which means we might as well define it as a scalar:

$$\omega := \omega_z = \frac{\partial}{\partial x} v_y - \frac{\partial}{\partial y} v_x. \tag{2.9}$$

Finally, define the material derivative of a quantity as the time derivative of that quantity along a trajectory of the fluid,

$$\frac{\mathrm{D}}{\mathrm{D}t} := \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla. \tag{2.10}$$

We are now ready to formulate the vorticity equation,

$$\frac{\mathrm{D}\omega}{\mathrm{D}t} = 0. \tag{2.11}$$

This can be verified by plugging it into (2.2) and using the assumptions above. This shows that vorticity is advected by the flow. We can use this as the defining equation for the flow since the velocity field can be determined from the vorticity using

$$\mathbf{v}(\mathbf{x}) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \omega(\mathbf{x}') \frac{\mathbf{e}_z \times (\mathbf{x} - \mathbf{x}')}{||\mathbf{x} - \mathbf{x}'||^2} \mathrm{d}\mathbf{x}', \tag{2.12}$$

assuming that the far field velocity is zero and the vorticity is zero except on a compact set.

2.1.2 Patch and point vortices

Consider a circular patch of positive vorticity a at the origin with radius R, with zero vorticity elsewhere, so working in polar coordinates (r, θ):

$$\omega(r,\theta) = \begin{cases} a & \text{if } r \leq R, \\ 0 & \text{if } r > R. \end{cases} \tag{2.13}$$

Denote the total vorticity by γ, the strength of the vortex:

$$\gamma := \int_{\mathbb{R}^2} \omega(\mathbf{x}')\mathrm{d}\mathbf{x}' = \pi a R^2. \tag{2.14}$$

The velocity field is then given by

$$\mathbf{v}(r,\theta) = \begin{cases} \mathbf{0} & \text{if } r = 0, \\ \frac{ar}{2}\mathbf{e}_\theta & \text{if } 0 < r \leq R, \\ \frac{\gamma}{2\pi r}\mathbf{e}_\theta & \text{if } R < r. \end{cases} \tag{2.15}$$

This can be derived either from (2.12) or more easily using symmetry arguments, but it can be easily verified using (2.8).

Note that the boundary of the patch vortex is a streamline; since vorticity is advected by the flow this means that the vorticity distribution is stationary. Also note that the velocities at the origin and outside the patch do not depend on its radius R (assuming γ is kept constant). This suggests a very simple model: we let R go to zero, concentrating all the vorticity in a point. To test if this simplification is justified, we compared the results of advection in pure point vortex system to those in systems of small patch vortices using the contour dynamics method described in [73]; up to quite large sizes the results were similar (these results are not included in this report). A system of these point vortices can be described easily: the vortices are advected by the flow caused by the other vortices.

Consider a system of N point vortices with identical strength γ, and denote their positions by $\mathbf{x}_i(t) = (x_i(t), y_i(t))$ for $1 \leq i \leq N$ and $t \in \mathbb{R}$. The system itself is written as $\Gamma(t) = (\mathbf{x}_1(t), ..., \mathbf{x}_n(t))$. The time argument will usually be dropped when no confusion can occur. The velocity field is simply the sum of the fields due to the individual vortices, so

$$\mathbf{v}(\mathbf{x},t) = \begin{cases} \frac{\gamma}{2\pi} \sum\limits_{j=1}^{N} \frac{\mathbf{e}_z \times (\mathbf{x}-\mathbf{x}_j)}{||\mathbf{x}-\mathbf{x}_j||^2} & \text{if } \mathbf{x} \neq \mathbf{x}_i \text{ for all } i \text{ with } 1 \leq i \leq N, \\ \frac{\gamma}{2\pi} \sum\limits_{j=1, j\neq i}^{N} \frac{\mathbf{e}_z \times (\mathbf{x}-\mathbf{x}_j)}{||\mathbf{x}-\mathbf{x}_j||^2} & \text{if } \mathbf{x} = \mathbf{x}_i. \end{cases} \tag{2.16}$$

As mentioned the motion of the point vortices is given by the velocity field at their positions since vorticity is advected by the flow, so we have

$$\frac{\mathrm{d}\mathbf{x}_i}{\mathrm{d}t} = \frac{\gamma}{2\pi} \sum_{j=1, j\neq i}^{N} \frac{\mathbf{e}_z \times (\mathbf{x}_i - \mathbf{x}_j)}{||\mathbf{x}_i - \mathbf{x}_j||^2}. \tag{2.17}$$

Or written out by component

$$\begin{aligned} \frac{\mathrm{d}x_i}{\mathrm{d}t} &= -\frac{\gamma}{2\pi} \sum_{j=1, j\neq i}^{N} \frac{y_i - y_j}{(x_i - x_j)^2 + (y_i - y_j)^2}, \\ \frac{\mathrm{d}y_i}{\mathrm{d}t} &= \frac{\gamma}{2\pi} \sum_{j=1, j\neq i}^{N} \frac{x_i - x_j}{(x_i - x_j)^2 + (y_i - y_j)^2}. \end{aligned} \tag{2.18}$$

The point vortex system is Hamiltonian; with canonical variables (x_i, y_i) the Hamiltonian H is given by

$$H = -\frac{\gamma}{4\pi} \sum_{j=2}^{N} \sum_{k=1}^{j-1} \log((x_j - x_k)^2 + (y_j - y_k)^2). \tag{2.19}$$

The equations of motion can than be formulated as

$$\begin{aligned} \frac{\mathrm{d}x_i}{\mathrm{d}t} &= \frac{\partial H}{\partial y_i}, \\ \frac{\mathrm{d}y_i}{\mathrm{d}t} &= -\frac{\partial H}{\partial x_i}. \end{aligned} \tag{2.20}$$

Like in all Hamiltonian systems, H is a constant of motion. Further constants of motion are the linear momenta P_x and P_y and the angular momentum L^2, defined by

$$\begin{aligned} P_x &= \gamma \sum_{j=1}^{N} x_j, \\ P_y &= \gamma \sum_{j=1}^{N} y_j, \\ L^2 &= \frac{1}{2}\gamma \sum_{j=1}^{N} (x_j^2 + y_j^2). \end{aligned} \tag{2.21}$$

Now define the map $\Psi(T)$ as the map moving a system forward by time T, i.e.

$$\Psi(T)(\Gamma(t)) = \Gamma(t + T). \tag{2.22}$$

The reason that t is not needed as an argument of Ψ is that the system is autonomous. Note that

$$\begin{aligned}(\Psi(T)\circ\Psi(S))(\Gamma(t)) =&\Psi(T)(\Psi(S)(\Gamma(t))) = \Psi(T)(\Gamma(t+S)) = \\ &\Gamma(t+S+T) = \Psi(T+S)(\Gamma(t)),\end{aligned}$$

so we find the important property

$$\Psi(T)\circ\Psi(S) = \Psi(T+S). \tag{2.23}$$

Another important property of the map $\Psi(T)$ is that it is *symplectic*, meaning that it conserves the two-form $\mathrm{d}x_i \wedge \mathrm{d}y_i$ for $1 \leq i \leq N$. In other words, areas and volumes in the phase space are conserved under time evolution of the system. This is a general property of Hamiltonian systems [54], and in fact any symplectic mapping corresponds to a Hamiltonian system [53].

2.2 Numerical methods

In this section the numerical method used to approximate $\Psi(T)$ will be developed. We will start with a simple first order Euler method (section 2.2.1) and show why this is not adequate. We will move on to a first order symplectic scheme (section 2.2.2) and then higher order symplectic schemes (section 2.2.3). Finally some further properties of the chosen method, a sixth order scheme, will be discussed (section 2.2.4). All computations in this section were performed on the irregular four vortex system (see section 2.3.2).

2.2.1 Explicit Euler integration

The simplest time integration method we can apply is first order explicit Euler. This simply means that we compute the velocities of each of the point vortices using (2.17) and move every point vortex a distance $\mathbf{v}\Delta t$, where Δt is the timestep. To be more specific, denote the positions at timestep n, so at time $n\Delta t$, by $\mathbf{x}_i^n$. The positions at the next timestep are then computed by

$$\mathbf{x}_i^{n+1} = \mathbf{x}_i^n + \Delta t \frac{\gamma}{2\pi} \sum_{j=1, j\neq i}^{N} \frac{\mathbf{e}_z \times (\mathbf{x}_i^n - \mathbf{x}_j^n)}{||\mathbf{x}_i^n - \mathbf{x}_j^n||^2}. \tag{2.24}$$

This method is consistent and has a discretisation error of order Δt.

We now investigate whether this method is adequate for our needs. To find out, let us have a look at the conserved quantities, the linear momenta P_x and P_y, angular momentum L^2 and Hamiltonian H. The relative change in these quantities for a typical system is shown for two values of Δt in Figures 2.1 and 2.2. Note that the numerical method conserves P_x and P_y but not L^2 and H, although it does better with a smaller timestep.

Of course we can conserve these quantities arbitrarily well by reducing the time step since our method is consistent. Another option would be to use a higher order method, which would reduce the error more efficiently. However, the fact remains that both L^2 and H are diverging from their initial values. We can do better-the method that will now be introduced conserves L^2 and keeps H close to its initial value.

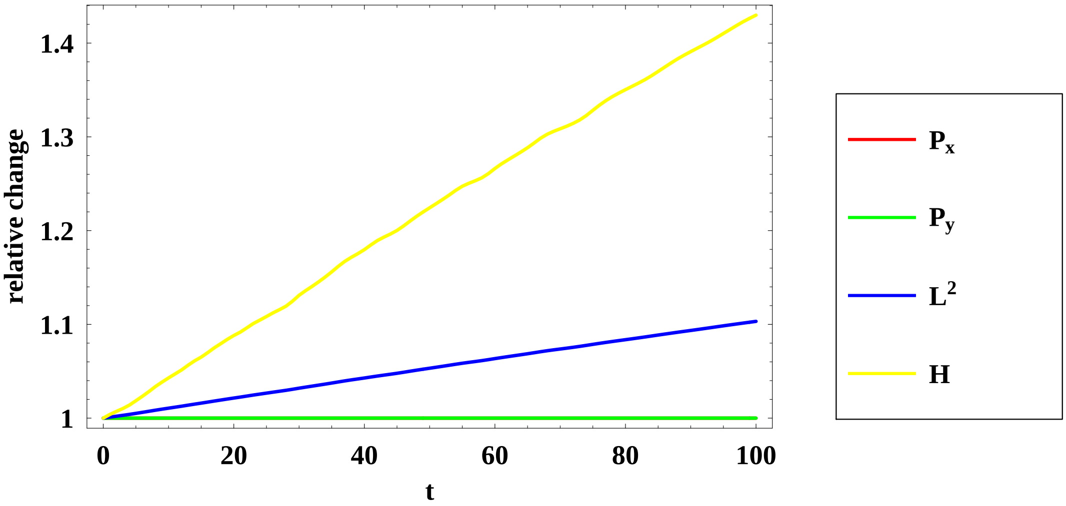

Figure 2.1: Relative change in P_x, P_y, L^2 and H, using explicit Euler integration with $\Delta t = 0.01$. The curves for P_x and P_y overlap.

2.2.2 Symplectic integration

The main issue with explicit Euler integration is that it is not symplectic, i.e. the mapping defined by repeatedly applying (2.24) does not conserve areas and volumes in the phase space. If we could find a consistent symplectic method, it would correspond to a system with a Hamiltonian slightly perturbed from our true Hamiltonian. Since this perturbed Hamiltonian would be conserved to numerical precision, the true Hamiltonian would not diverge [53].

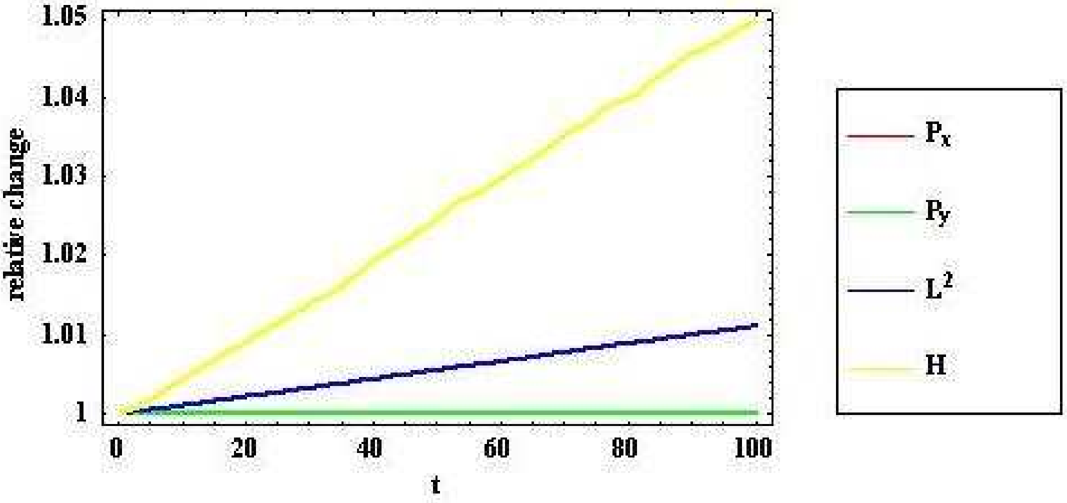

Figure 2.2: Relative change in P_x, P_y, L^2 and H, using explicit Euler integration with $\Delta t = 0.001$. The curves for P_x and P_y overlap.

The method presented here was developed in [45]. We want to approximate $\Psi(\Delta t)$, the operator moving the system Δt forward in time. Using Taylor expansion we can write

$$\begin{aligned}\Psi(\Delta t)(\Gamma(t)) =& \Gamma(t+\Delta t) = \left(\sum_{j=0}^{\infty} \frac{1}{j!}\left(\Delta t \frac{\mathrm{d}}{\mathrm{d}t}\right)^j\right)\Gamma(t) = \\ & \left(\sum_{j=0}^{\infty} \frac{1}{j!}\left(\Delta t \mathbf{H}\right)^j\right)\Gamma(t) = \exp(\Delta t \mathbf{H})\Gamma(t),\end{aligned} \tag{2.25}$$

where $\mathbf{H} := \frac{\mathrm{d}}{\mathrm{d}t}$ is the time differentiation operator corresponding to the Hamiltonian H [68]. Note that H can be written as (see (2.19))

$$H = \sum_{j=2}^{N}\sum_{k=1}^{j-1} \Omega_{jk}, \tag{2.26}$$

where Ω_{jk} is the Hamiltonian corresponding to a dipole consisting of the point vortices j and k,

$$\Omega_{jk} := -\frac{\gamma}{4\pi}\log((x_j - x_k)^2 + (y_j - y_k)^2), \tag{2.27}$$

with corresponding time differentiation operator $\mathbf{\Omega}_{jk}$. The operator

$$\hat{\Omega}_{jk}(\Delta t) := \exp(\Delta t \mathbf{\Omega}_{jk}) \tag{2.28}$$

then moves the dipole ahead by time Δt. This operator can be applied exactly; both vortices simply rotate around their midpoint with angular velocity

$$\omega_{jk} = \frac{\gamma}{\pi ||\mathbf{x}_i - \mathbf{x}_j||^2}. \tag{2.29}$$

Note that

$$\mathbf{H} = \sum_{j=2}^{N} \sum_{k=1}^{j-1} \mathbf{\Omega}_{jk}. \tag{2.30}$$

Now consider the operator

$$\Psi^{(1)}(\Delta t) = \prod_{j=2}^{N} \prod_{k=1}^{j-1} \hat{\Omega}_{jk}(\Delta t), \tag{2.31}$$

where the product of two operators is defined as their composition. What this operator does is take every dipole in turn and evolve it over a time Δt. Note that the updated positions are used at each step. $\Psi^{(1)}(\Delta t)$ approximates $\Psi(\Delta t)$ with first order discretisation error, it is symplectic and it conserves P_x, P_y and L^2 to numerical precision. Let us investigate each of these three claims in turn.

First we will show that the approximation is consistent and of first order. We will need the following theorem, which follows directly from the Baker-Campell-Hausdorff formulas [27].

Theorem 2.1. *Let $\mathbf{X}_i$, $1 \le i \le N$, be a sequence of operators. Then for $h > 0$*

$$\log\left(\prod_{i=1}^{N} \exp\left(h\mathbf{X}_i\right)\right) = h \sum_{i=1}^{N} \mathbf{X}_i + \mathcal{O}(h^2).$$

We now have

$$\begin{aligned} \Psi^{(1)}(\Delta t) &\overset{(2.31)}{=} \prod_{j=2}^{N} \prod_{k=1}^{j-1} \hat{\Omega}_{jk}(\Delta t) \overset{(2.28)}{=} \exp(\log(\prod_{j=2}^{N} \prod_{k=1}^{j-1} \exp(\Delta t \mathbf{\Omega}_{jk}))) \overset{\text{T2.1}}{=} \\ & \exp(\Delta t \sum_{j=2}^{N} \sum_{k=1}^{j-1} \mathbf{\Omega}_{jk} + \mathcal{O}(\Delta t^2)) \overset{(2.30)}{=} \exp(\Delta t \mathbf{H}) + \mathcal{O}(\Delta t^2) \overset{(2.25)}{=} \\ & \Psi(\Delta t) + \mathcal{O}(\Delta t^2), \end{aligned} \tag{2.32}$$

showing that $\Psi^{(1)}(\Delta t)$ is indeed a first order approximation for $\Psi(\Delta t)$.

Next we will show that the mapping is symplectic. This is actually quite simple: each of the elementary maps $\hat{\Omega}_{jk}(\Delta t)$ corresponds to a Hamiltonian (namely Ω_{jk}) and hence is symplectic, conserving areas and volumes in the phase space. Since $\Psi^{(1)}(\Delta t)$ is composed of these elementary maps, it is symplectic as well.

Finally, let us see how $\Psi^{(1)}(\Delta t)$ conserves P_x, P_y and L^2. We will demonstrate this for the angular momentum L^2; the argument for the linear momenta P_x and P_y is similar. We will show that the elementary mappings $\hat{\Omega}_{jk}(\Delta t)$ conserve L^2, which means that $\Psi^{(1)}(\Delta t)$ does as well. We can write L^2 (see (2.21)) as

$$L^2 = \frac{1}{2} \sum_{i=1, i\neq j, i\neq k}^{N} (x_i^2 + y_i^2) + \frac{1}{2}(x_j^2 + y_j^2 + x_k^2 + y_k^2).$$

The first term is not altered by applying $\hat{\Omega}_{jk}(\Delta t)$ since it influences only $\mathbf{x}_j$ and $\mathbf{x}_k$. The second term is simply L^2 for the two vortex system consisting of the vortices j and k; since this is a conserved quantity in this system and $\hat{\Omega}_{jk}(\Delta t)$ is exact, it is not altered either. So L^2 is conserved under the operator $\hat{\Omega}_{jk}(\Delta t)$.

We have now found a consistent symplectic integrator, which additionally conserves all constants of motion with the exception of the Hamiltonian H. The behavior of H using this method is shown in Figure 2.3. Note that H is not constant, but in contrast to the results using explicit Euler integration it does not diverge. This is the main reason to use a symplectic method.

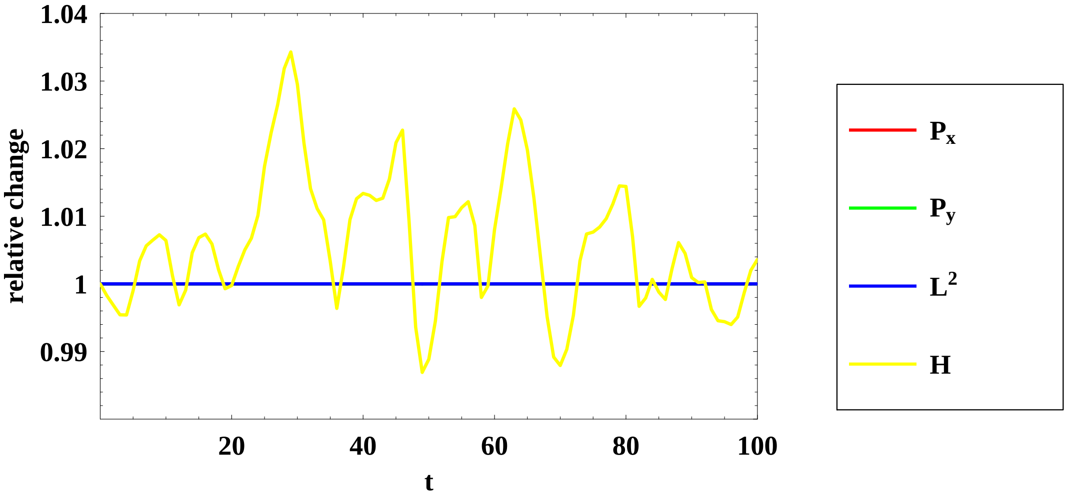

Figure 2.3: Relative change in P_x, P_y, L^2 and H, using first order symplectic integration with $\Delta t = 0.1$. The curves for P_x, P_y and L^2 overlap.

2.2.3 Higher order symplectic integration

Although the first order symplectic integrator found in the previous section has several nice properties, it does not actually perform particularly well compared to our 'naïve' first order Euler method (see Table 2.2). A possible approach to remedy this is to find higher order methods. How to do this is well-known for the explicit Euler method, but we would like to find higher order symplectic methods.

A second order method similar to (2.31) is readily available [45]:

$$\Psi^{(2)}(\Delta t) = \prod_{j=2}^{N}\prod_{k=1}^{j-1}\hat{\Omega}_{jk}(\Delta t/2) \circ \prod_{j=N}^{2}\prod_{k=j-1}^{1}\hat{\Omega}_{jk}(\Delta t/2). \tag{2.33}$$

What this method does is similar to the first order method, which rotates every dipole once. The difference is that we sweep over all dipoles twice, the second time in reverse order. Since it is composed of the same elementary maps as $\Psi^{(1)}(\Delta t)$, $\Psi^{(2)}(\Delta t)$ is symplectic and conserves P_x, P_y and L^2. That it is second order can be shown using a completely similar argument as in (2.32), using a higher order version of Theorem 2.1. We omit the details here as they are rather cumbersome.

This method already performs significantly better than the first order one (see Table 2.2), but we can do even better. A simple way to construct higher order methods is to use Richardson extrapolation. This involves computing the solution using different values of Δt, which allows us to estimate the lowest order error term and then to eliminate it. Although this did work, the resulting method did not turn out to be symplectic. This is not a surprise; after all there is no reason why it should be.

Symplectic higher order methods can be obtained by applying $\Psi^{(2)}$ multiple times, varying the timestep to eliminate error terms. This is based on work by Yoshida [77]. Although he considered only Hamiltonians split as

$$H(x_1, ..., x_N, y_1, ..., y_N) = T(x_1, ..., x_N) + V(y_1, ..., y_N),$$

his methods turn out to work equally well for our more general splitting (2.26). This will be checked in section 2.2.4. A fourth order method is given by

$$\Psi^{(4)}(\Delta t) = \Psi^{(2)}(c_1^{(4)}\Delta t)\Psi^{(2)}(c_2^{(4)}\Delta t)\Psi^{(2)}(c_1^{(4)}\Delta t), \tag{2.34}$$

i.e. three second order steps strung together. The coefficients $c_i^{(4)}$ are given in Table 2.1. A sixth order method could be constructed by stringing three fourth order steps together, involving nine second order steps. However it is also possible using only seven steps:

$$\Psi^{(6)}(\Delta t) = \left(\prod_{j=1}^{3}\Psi^{(2)}(c_j^{(6)}\Delta t)\right)\Psi^{(2)}(c_4^{(6)}\Delta t)\left(\prod_{j=3}^{1}\Psi^{(2)}(c_j^{(6)}\Delta t)\right), \tag{2.35}$$

with coefficients again given in Table 2.1. Finally, an eighth order method is available using fifteen steps:

$$\Psi^{(8)}(\Delta t) = \left(\prod_{j=1}^{7} \Psi^{(2)}(c_j^{(8)} \Delta t) \right) \Psi^{(2)}(c_8^{(8)} \Delta t) \left(\prod_{j=7}^{1} \Psi^{(2)}(c_j^{(8)} \Delta t) \right) . \tag{2.36}$$

The performance of the various methods is compared in Tables 2.2 and 2.3. Note that the performance increases up to sixth order but no further gain is obtained from going to eighth order. Of course if we increase the accuracy requirement the eighth order method will at some point show performance gains, but in most situations encountered in practice this accuracy was not required. For this reason the sixth order method was chosen and used for all computations.

i	$c_i^{(4)}$	$c_i^{(6)}$	$c_i^{(8)}$
1	1.35121	0.78451	0.62903
2	-1.70241	0.23557	1.36935
3		-1.11777	-1.06459
4		-0.80464	1.66336
5			-1.67897
6			-1.55947
7			0.311791
8			1.659000

Table 2.1: Coefficients for Yoshida's higher order symplectic integrators [77]. For the sixth order method, variant A is used (this choice is motivated in section 2.2.4). For the eighth order method, variant C is used.

2.2.4 Further properties

There are some more properties of the chosen method $\Psi^{(6)}(\Delta t)$ worth discussing. First, one last word about the choice of method. There are three variants (A, B and C) given in [77]. Furthermore, there is the sixth order method constructed by stringing together the fourth order method $\Psi^{(4)}$ three times, which we will call variant D. The choice for variant A is motivated in Table 2.4. This is consistent with results found in [53].

Next let us have a look at the computing time required for one timestep. Define an elementary operation as one application of $\hat{\Omega}_{jk}$, i.e. the rotation of one dipole. There

Method	Δt	Computing time (a.u.)
Explicit Euler	$5.7 * 10^{-6}$	100
First order symplectic	$4.3 * 10^{-6}$	133
Second order symplectic	0.020	0.050
Fourth order symplectic	0.20	0.018
Sixth order symplectic	0.50	0.014
Eighth order symplectic	0.50	0.030

Table 2.2: Timestep and computing time needed to calculate $x_1(20)$, the x coordinate of vortex 1 at $t = 20$, to 0.01% accuracy. The computing time is given in arbitrary units.

Method	Δt	Computing time (a.u.)
Second order symplectic	$1.8 * 10^{-5}$	100
Fourth order symplectic	$6.2 * 10^{-3}$	1.03
Sixth order symplectic	0.061	0.21
Eighth order symplectic	0.090	0.30

Table 2.3: Timestep and computing time needed to calculate $x_1(80)$, the x coordinate of vortex 1 at $t = 80$, to 0.01% accuracy. The computing time is given in arbitrary units.

are $N(N-1)/2$ dipoles (recall that N is the number of vortices), and each one is rotated twice for one application of $\Psi^{(2)}$, so applying $\Psi^{(2)}$ takes $N(N-1)$ elementary operations. Since $\Psi^{(6)}(\Delta t)$ consists of seven second order steps, it takes $7N(N-1)$ elementary operations. In particular, the computing time scales as N^2.

Finally, we still need to check that the method is indeed sixth order. This is shown in Figure 2.4, which shows the absolute error in $x_1(20)$ as a function of the timestep Δt. There is clearly a regime ($0.03 \lesssim \Delta t \lesssim 0.8$) where the error is sixth order in Δt. For smaller values of Δt, the error is dominated by computer precision, while for larger values it is dominated by higher order error terms.

One final remark: Figure 2.4 might give the impression that a timestep as large as 0.8 already gives quite accurate results. However, this holds specifically at $t = 20$; due to the chaotic nature of the system (see section 2.3.2) this small error will wreak havoc on the results at a later time. For comparison the error as a function of Δt for $t = 80$ is shown in Figure 2.5; note that for accurate results we need to take Δt at most 0.2. A more detailed analysis of the choice of timestep will be given in section 3.2.1.

Method	Δt	Computing time (a.u.)
Sixth order A	0.061	100
Sixth order B	0.032	191
Sixth order C	0.032	192
Sixth order D	0.032	289

Table 2.4: Timestep and computing time needed to calculate $x_1(80)$, the x coordinate of vortex 1 at $t = 80$, to 0.01% accuracy. The computing time is given in arbitrary units.

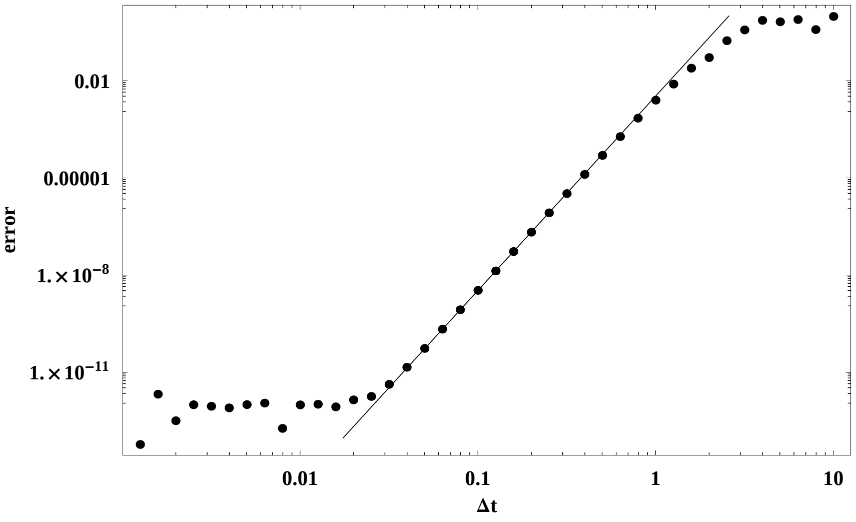

Figure 2.4: Absolute error in $x_1(20)$, the x coordinate of vortex 1 at $t = 20$, as a function of Δt for the sixth order method. The drawn line has a slope of 6.

2.3 Results

In this section several systems will be introduced that showcase various possible behaviors for three and four point vortex systems.

The motion of many of these systems can be simplified by applying a properly chosen background rotation. This simply means that we add a solid body rotation with some angular velocity Ω around some point $\mathbf{r} = (r_x, r_y)$ to the velocity field

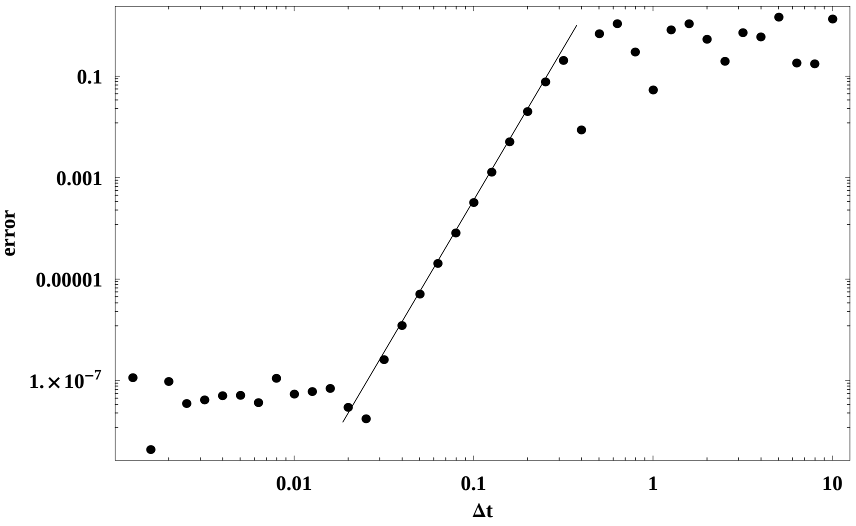

Figure 2.5: Absolute error in $x_1(80)$, the x coordinate of vortex 1 at $t = 80$, as a function of Δt for the sixth order method. The drawn line has a slope of 6.

advecting the point vortices. Note that the equation defining the system, (2.17), is invariant under rotation, so this does not affect the behavior of the system, and we can apply this background rotation at any point without interfering with the numerical method. The full equation of motion becomes:

$$\frac{\mathrm{d}\mathbf{x}_i}{\mathrm{d}t} = \frac{\gamma}{2\pi} \sum_{j=1, j\neq i}^{N} \frac{\mathbf{e}_z \times (\mathbf{x}_i - \mathbf{x}_j)}{||\mathbf{x}_i - \mathbf{x}_j||^2} + \Omega \mathbf{e}_z \times (\mathbf{x}_i - \mathbf{r}). \tag{2.37}$$

We will see three types of motion in this section: stationary, where the point vortices do not move at all; periodic, where they return to their initial configuration after some time T_p; and irregular, where the motion is chaotic in the phase space. The term irregular is used instead of chaotic to avoid confusion later when we consider tracer motion, which can be chaotic even in periodic systems. It is important to mention here that the terms stationary, periodic and irregular apply to the system including background rotation. For instance, periodic systems return to the initial configuration at some point, but rotated at some angle θ which is not in general commensurate with 2π, meaning that the system is actually quasi-periodic. However, when we choose the background rotation such that it provides a counterrotation θ over one period the system is truly periodic.

For all systems presented here, the vortex strength γ is set to 1. There is no loss of generality; other values of γ can be achieved simply by rescaling time (recall that we only consider point vortices of equal strength). A more detailed analysis of the scaling behavior of point vortex systems can be found in [60]. A summary of all systems discussed here, in the order they are introduced, can be found in Table 2.5.

System name	Vortex positions	$\mathbf{r}$	Ω	Motion type
Stationary three vortex	$(-\frac{1}{2}\sqrt{3}, -\frac{1}{2})$ $(\frac{1}{2}\sqrt{3}, -\frac{1}{2})$ $(0, 1)$	$(0, 0)$	$-\frac{1}{2\pi}$	Stationary
3-periodic three vortex	$(1, 0)$ $(-1, 0)$ $(0, 1)$	$(0, 0.333)$	-0.37014	Periodic $T_p = 11.577$
2-periodic three vortex	$(1, 0)$ $(-1, 0)$ $(0, 5.48689)$	$(0, 1.829)$	-0.016137	Periodic $T_p = 56.708$
Stationary four vortex	$\frac{1}{2}(\sqrt{2}, \sqrt{2})$ $\frac{1}{2}(-\sqrt{2}, -\sqrt{2})$ $\frac{1}{2}(\sqrt{2}, -\sqrt{2})$ $\frac{1}{2}(-\sqrt{2}, \sqrt{2})$	$(0, 0)$	$-\frac{3}{4\pi}$	Stationary
Periodic four vortex	$(1, 0.3)$ $(-1, 0.3)$ $(1, -0.3)$ $(-1, -0.3)$	$(0, 0)$	-0.16365	Periodic $T_p = 5.2934$
Irregular four vortex	$(1, 0)$ $(-1, 0)$ $(0, 1)$ $(0.5, 0.5)$	N/A	0	Irregular

Table 2.5: Summary of the systems introduced in this section. Ω is the angular velocity of the background rotation around point $\mathbf{r}$.

2.3.1 Three vortex systems

The simplest three vortex system has the three vortices on an equilateral triangle; this is the first system we will consider. The three vortices are placed on a circle of radius 1 around the origin, at $(-\frac{1}{2}\sqrt{3}, -\frac{1}{2})$, $(\frac{1}{2}\sqrt{3}, -\frac{1}{2})$ and $(0, 1)$. The motion of the vortices in this system is shown in the left part of Figure 2.6. The configuration is left intact, rotating around the origin. This suggests that a properly chosen background rotation around the origin will make the system stationary; in fact this is the case with $\Omega = -\frac{1}{2\pi}$. Indeed, just filling in (2.37):

$$\begin{aligned}\frac{\mathrm{d}\mathbf{x}_3}{\mathrm{d}t} =& \frac{\gamma}{2\pi}\left(\frac{\mathbf{e}_z\times(\mathbf{x}_3-\mathbf{x}_1)}{||\mathbf{x}_3-\mathbf{x}_1||^2}+\frac{\mathbf{e}_z\times(\mathbf{x}_3-\mathbf{x}_2)}{||\mathbf{x}_3-\mathbf{x}_2||^2}\right)+\Omega\mathbf{e}_z\times(\mathbf{x}_3-\mathbf{r}) = \\ & \frac{1}{2\pi}\left(\frac{\mathbf{e}_z\times((0,1)-(-\frac{1}{2}\sqrt{3},-\frac{1}{2}))}{||(0,1)-(-\frac{1}{2}\sqrt{3},-\frac{1}{2})||^2}+\frac{\mathbf{e}_z\times((0,1)-(\frac{1}{2}\sqrt{3},-\frac{1}{2}))}{||(0,1)-(\frac{1}{2}\sqrt{3},-\frac{1}{2})||^2}\right)- \\ & \frac{1}{2\pi}\mathbf{e}_z\times((0,1)-(0,0)) = \\ & \frac{1}{2\pi}\left((-\frac{1}{2},\frac{1}{6}\sqrt{3})+(-\frac{1}{2},-\frac{1}{6}\sqrt{3})+(1,0)\right) = \mathbf{0}\end{aligned}, \tag{2.38}$$

and similarly $\frac{\mathrm{d}\mathbf{x}_1}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{x}_2}{\mathrm{d}t} = \mathbf{0}$. So with this background rotation the vortices do not move (see the right part of Figure 2.6 as well); for this reason this system will be referred to from here on as the stationary three vortex system (see Table 2.5). The motion is stable, i.e. small perturbations of the equilateral triangle will not grow exponentially [35].

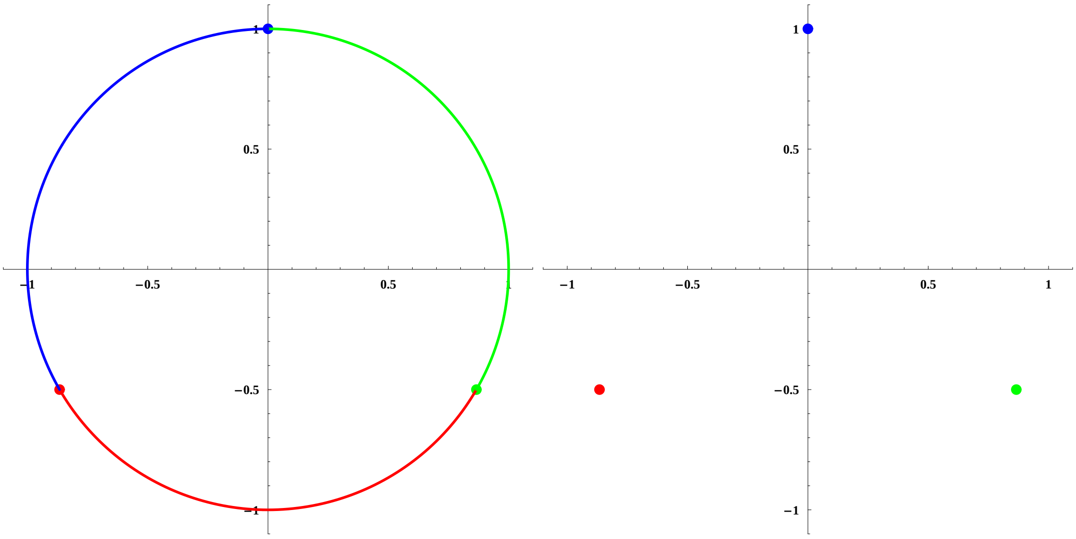

Figure 2.6: Vortex motion from $t = 0$ to $t = 13.159$ for the stationary three vortex system without (left) and with (right) background rotation, with $\Delta t = 0.001$. The dots indicate the initial positions of the vortices.

Now let us consider the general case of three identical point vortices. As derived in [7], the integrals H, L^2 and $P_x^2 + P_y^2$ are in involution. Since there are as many integrals in involution as there are degrees of freedom, namely three, the three point vortex system is integrable [75] and in fact an analytical solution is available [7, 47], although it will not be used here. Note that there are six parameters describing a three point vortex system (x_1, y_1, x_2, y_2, x_3 and y_3). Five of these can be eliminated using five symmetries: translation in either space direction or time, scaling and rotation around the center of mass $\frac{1}{N}(P_x, P_y)$. This means that only one parameter is required to distinguish different systems; a useful one is:

$$\Lambda := \frac{\exp(-2\pi H)}{L^3} = \frac{z_{12}z_{13}z_{23}}{(z_{12}^2 + z_{13}^2 + z_{23}^2)^{3/2}}, \tag{2.39}$$

with $z_{ij} := ||\mathbf{x}_i - \mathbf{x}_j||$ the distance between vortices i and j. Note that

$$0 < \Lambda \leq 1, \tag{2.40}$$

with $\Lambda = 1$ for the equilateral triangle configuration. There are three different regimes: $\Lambda > \Lambda_c := 1/\sqrt{2}$, $\Lambda < \Lambda_c$ and $\Lambda = \Lambda_c$ [47]; the nature of the vortex motion differs between the three as we will now see.

Consider the system with the point vortices at $(-1, 0)$, $(1, 0)$ and $(0, 1)$, on an isosceles triangle. For this system $\Lambda = 0.91856 > \Lambda_c$ and the background rotation needed to achieve periodicity is given by $\Omega = -0.37014$ around $\mathbf{r} = (0, 0.33333)$ (these values were calculated using the method described in [47]). The motion of the vortices is shown in Figure 2.7. Note that at $t = 11.577$ the vortices are again in their initial configuration, but that it is permuted. It would take three of these periods for the vortices to be in their unpermuted initial configuration. Define the period T_p as the first time after $t = 0$ at which the vortices are in their initial configuration, permuted or not. This is a useful definition because we are ultimately interested in the advection in these systems, and since the point vortices are identical it is irrelevant whether the configuration is permuted. So, for our current system we have $T_p = 11.577$. That it takes three periods to return to the unpermuted initial configuration is a general property of three point vortex systems with $\Lambda > \Lambda_c$. It is also why this system will be referred to from here on as the 3-periodic three vortex system (see Table 2.5). Note that all three vortices follow the same trajectory, but at different times.

Next we consider the case $\Lambda < \Lambda_c$, using the system with the point vortices at $(-1, 0)$, $(1, 0)$ and $(0, 5.4869)$. For this system $\Lambda = 0.6 < \Lambda_c$ and the background rotation needed to achieve periodicity is given by $\Omega = -0.016137$ around $\mathbf{r} = (0, 1.8290)$ (again found using the method described in [47]). The motion of the vortices is shown in Figure 2.8. Note the difference with the case $\Lambda > \Lambda_c$: the three vortices no longer follow the same trajectory. Instead, there are two vortices in a like-signed dipole configuration, slightly distorted by a third vortex at a large distance. The two vortices in the dipole do follow the same trajectory, at different times. For this system we have $T_p = 56.708$. Note that it now takes two periods to return to the unpermuted initial configuration; this is a general property of three point vortex systems with $\Lambda < \Lambda_c$.

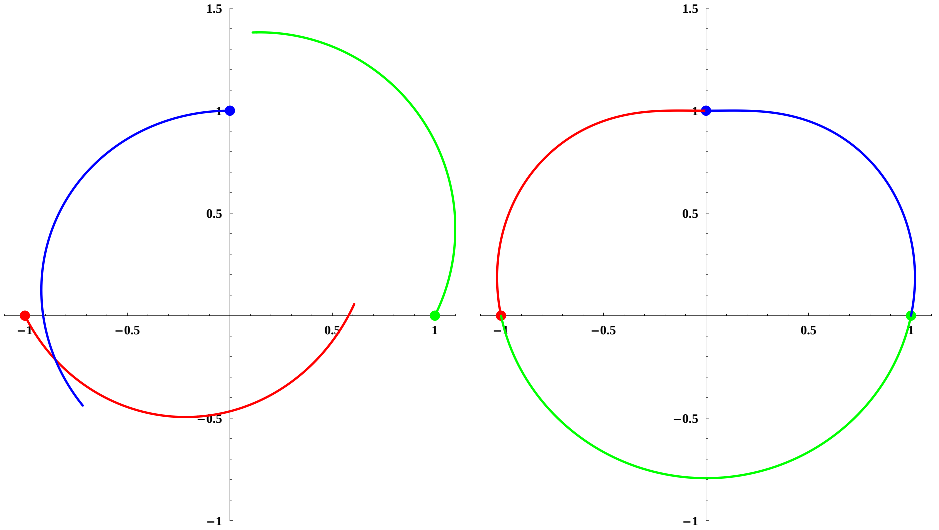

Figure 2.7: Vortex motion from $t = 0$ to $t = 11.577$ for the 3-periodic three vortex system without (left) and with (right) background rotation, with $\Delta t = 0.001$. The dots indicate the initial positions of the vortices.

This system will be referred to from here on as the 2-periodic three vortex system (see Table 2.5).

Finally, the case $\Lambda = \Lambda_c = 1/\sqrt{2}$ corresponds to the stationary but unstable configuration of three point vortices on a line. This system is not considered further here.

2.3.2 Four vortex systems

Unlike three vortex systems, four vortex systems are, with the exception of some specific configurations, irregular [7]. First, let us consider some of these exceptions.

Take the system with vortices at $\frac{1}{2}(\sqrt{2},\sqrt{2})$, $\frac{1}{2}(-\sqrt{2},\sqrt{2})$, $\frac{1}{2}(\sqrt{2},-\sqrt{2})$ and $\frac{1}{2}(-\sqrt{2},-\sqrt{2})$. The vortices form a square and are all on a circle with radius 1 around the origin. The motion of the vortices in this system is shown in Figure 2.9. Just like the equilateral triangle, the motion is stationary with the proper background rotation, in this case with $\Omega = -\frac{3}{4\pi}$ around the origin. This can be checked just as was done in the three vortex case in (2.38). This system will be referred to as the stationary four vortex system (see Table 2.5). This stationary configuration is stable [35].

An example of a periodic configuration is the four vortices on a rectangle [14]. So consider the system with point vortices at $(1, 0.3)$, $(-1, 0.3)$, $(1, -0.3)$ and $(-1, -0.3)$.

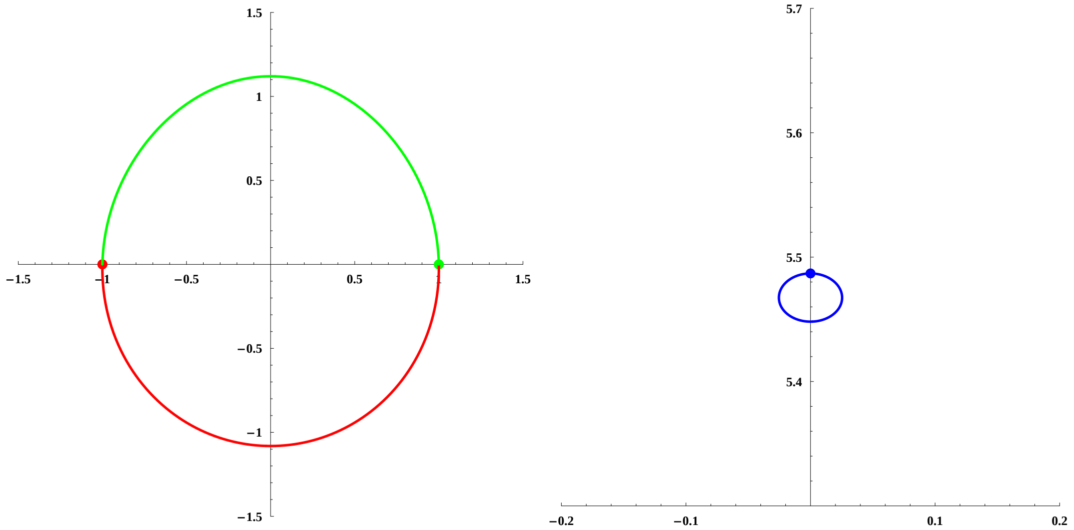

Figure 2.8: Vortex motion from $t = 0$ to $t = 56.708$ for the 2-periodic three vortex system with background rotation, with $\Delta t = 0.001$. The left figure shows the region around the dipole, the right figure the region around the isolated vortex. The dots indicate the initial positions of the vortices.

This system is periodic with period $T_p = 5.2934$ for a background rotation given by $\Omega = -0.16365$ around $\mathbf{r} = (0, 0)$ (the value of T_p was found by trial and error). The motion of the vortices is shown in Figure 2.10. Note that the system can be considered as two like-signed dipoles perturbing each other. It is 2-periodic in the sense described earlier, i.e. it takes two periods to return to the unpermuted original configuration. Since this is the only periodic system we will consider for four vortices, it will be referred to simply as the periodic four vortex system (see Table 2.5).

As mentioned the four vortex system is usually irregular, so if we take a configuration at random we have a good chance of finding an aperiodic system (of course we will need to check it). So let us take the 3-periodic three vortex system and add a fourth vortex halfway between vortices 2 and 3, giving a system with vortices at $(-1, 0)$, $(1, 0)$, $(0, 1)$ and $(0.5, 0.5)$. The motion of the vortices in this system is shown in Figure 2.11. The lack of repeating structures here is already an indication of aperiodicity.

A typical characteristic of aperiodic Hamiltonian systems is that they exhibit exponential growth of small perturbations. In order to quantify these perturbations we will need to define a distance in the phase space between two systems $\Gamma = (\mathbf{x}_1, ..., \mathbf{x}_N)$

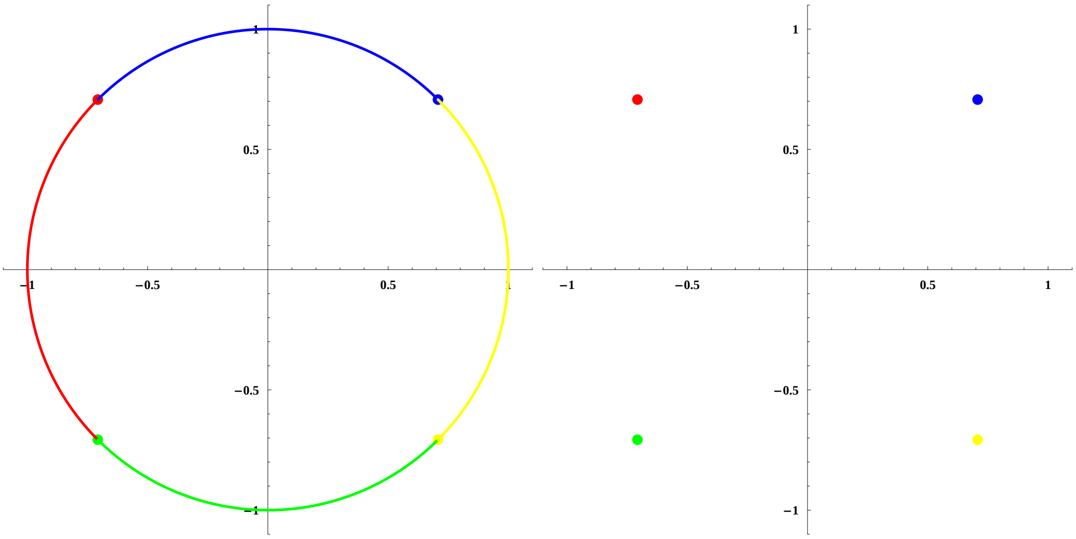

Figure 2.9: Vortex motion from $t = 0$ to $t = 6.5797$ for the stationary four vortex system without (left) and with (right) background rotation, with $\Delta t = 0.001$. The dots indicate the initial positions of the vortices.

and $\Gamma' = (\mathbf{x}'_1, ..., \mathbf{x}'_N)$:

$$||\Gamma - \Gamma'|| = \sqrt{\sum_{j=1}^{N} ||\mathbf{x}_j - \mathbf{x}'_j||^2} = \sqrt{\sum_{j=1}^{N} ((x_j - x'_j)^2 + (y_j - y'_j)^2)} \tag{2.41}$$

Now take $\Gamma(t)$ to be the unperturbed system, and define the perturbed system $\Gamma'(t)$ by taking:

$$\begin{aligned} \mathbf{x}'_1 =& \mathbf{x}_1, \\ \mathbf{x}'_2 =& \mathbf{x}_2 - (10^{-7}, 0), \\ \mathbf{x}'_3 =& \mathbf{x}_3 + (10^{-7}, 0), \\ \mathbf{x}'_4 =& \mathbf{x}_4, \end{aligned} \tag{2.42}$$

i.e. vortex 2 is moved to the left over a distance of 10^{-7} and vortex 3 to the right over the same distance. The growth of the perturbation in time is shown in Figure 2.12. For comparison the same technique is applied to the periodic four vortex system. As is clearly visible the system exhibits exponential growth of the perturbation while the periodic one does not (the linear growth visible is explained by a slight difference in the periods and rotation speeds of the perturbed and unperturbed systems). This shows that our system is indeed irregular and so it will be referred to from here on as the irregular four vortex system (see Table 2.5).

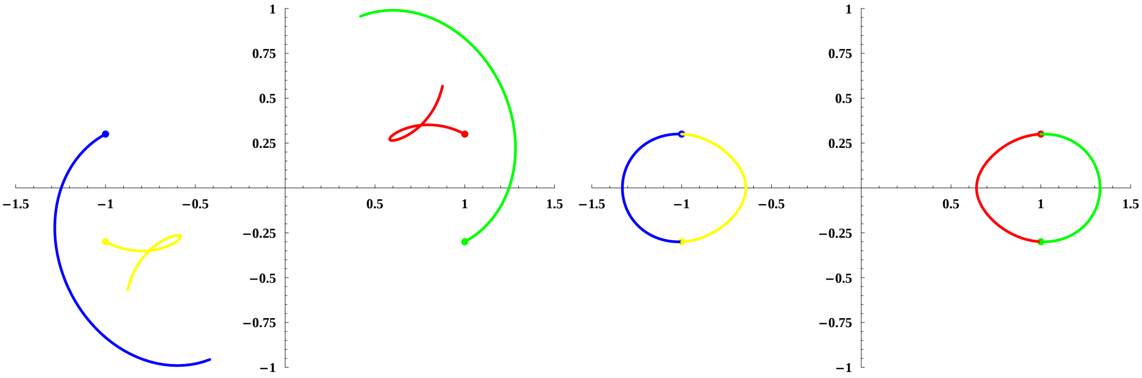

Figure 2.10: Vortex motion from $t = 0$ to $t = 5.2934$ for the periodic four vortex system without (left) and with (right) background rotation, with $\Delta t = 0.001$. The dots indicate the initial positions of the vortices.

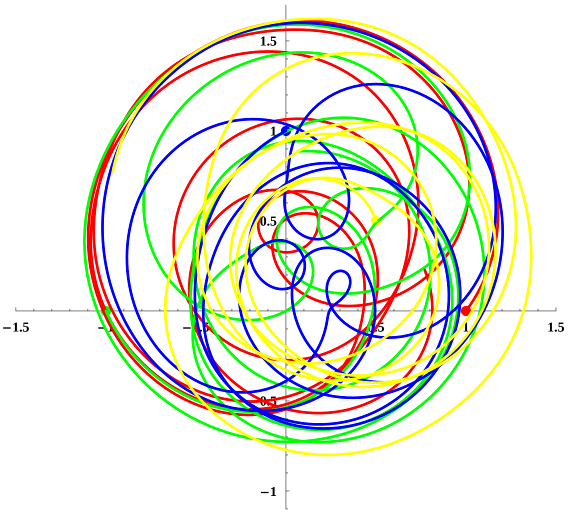

Figure 2.11: Vortex motion from $t = 0$ to $t = 100$ for the irregular four vortex system with no background rotation, with $\Delta t = 0.001$. The dots indicate the initial positions of the vortices.

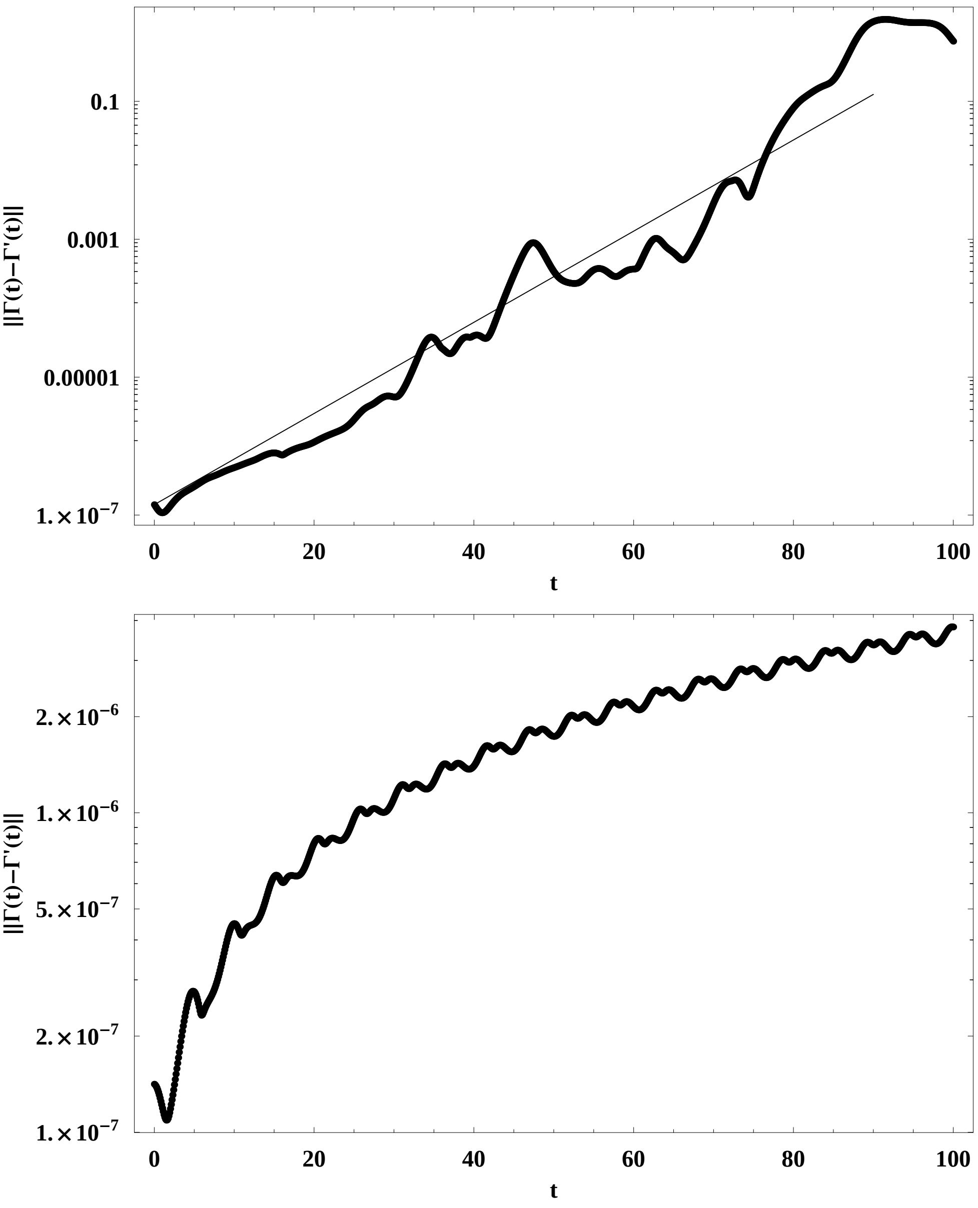

Figure 2.12: Size of the perturbation $||\Gamma(t) - \Gamma'(t)||$ as a function of time for the irregular four vortex system (top) and the periodic four vortex system (bottom). The line drawn in the left figure is a guide to the eye.

3 Tracer advection

In this chapter we will consider the behavior of passive tracers, small particles that are advected by the flow. For the time being we are concerned only with single tracers, so mixing as such is not considered yet.

First we will develop the general theory of tracer advection in point vortex systems (section 3.1). Next the numerical method used to compute tracer motion is introduced (section 3.2) as an extension of the method developed in section 2.2, with particular attention being paid to the choice of timestep (section 3.2.1). Finally, several tracer trajectories will be shown for some of the example systems, with a brief examination of the various regimes present in the phase space (section 3.3).

3.1 Theory

A (passive) tracer is an infinitesimal particle that moves with the flow without influencing it. If the position of a tracer $\mathbf{z}$ is known at a time t_0, so $\mathbf{z}(t_0) = \mathbf{z}_0$, then the trajectory can be found from the ordinary differential equation (ODE)

$$\begin{aligned} \frac{\mathrm{d}\mathbf{z}}{\mathrm{d}t} &= \mathbf{v}(\mathbf{z}(t), t), \\ \mathbf{z}(t_0) &= \mathbf{z}_0, \end{aligned} \tag{3.1}$$

where the velocity field $\mathbf{v}$ is given by (2.16) and a term for the background rotation (see section 2.3):

$$\mathbf{v}(\mathbf{x}, t) = \frac{1}{2\pi} \sum_{j=1}^{N} \frac{\mathbf{e}_z \times (\mathbf{x} - \mathbf{x}_j(t))}{||\mathbf{x} - \mathbf{x}_j(t)||^2} + \Omega \mathbf{e}_z \times (\mathbf{x} - \mathbf{r}). \tag{3.2}$$

Note that the flow is incompressible,

$$\nabla \cdot \mathbf{v} = 0, \tag{3.3}$$

so there is a function $\psi : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$, the stream function, such that

$$\begin{aligned} v_x &= \frac{\partial \psi}{\partial y}, \\ v_y &= -\frac{\partial \psi}{\partial x}. \end{aligned} \tag{3.4}$$

Explicitly, this stream function is given by

$$\psi(x,y,t) = -\frac{1}{4\pi}\sum_{j=1}^{N}\log((x-x_j(t))^2+(y-y_j(t))^2)-\frac{1}{2}\Omega((x-r_x)^2+(y-r_y)^2). \quad (3.5)$$

The equations of motion (3.1) can be combined with (3.4) to obtain, writing $\mathbf{z}(t) = (z_x(t), z_y(t))$:

$$\begin{aligned} \frac{\mathrm{d}z_x}{\mathrm{d}t} =& \frac{\partial\psi}{\partial y}(z_x, z_y, t), \\ \frac{\mathrm{d}z_y}{\mathrm{d}t} =& -\frac{\partial\psi}{\partial x}(z_x, z_y, t), \\ z_x(0) =& z_{0x}, \\ z_y(0) =& z_{0y}. \end{aligned} \quad (3.6)$$

Note the similarity with (2.20), the equation describing the motion of point vortex systems in Hamiltonian form. What we have here are the equations of motion for a tracer as a Hamiltonian system with one degree of freedom, with the Hamiltonian given by the stream function ψ. This has two important consequences.

The first of these is that the system is symplectic. We knew this already though- conservation of areas in the phase space in this case simply means that the flow is incompressible. This does suggest that the symplectic numerical methods developed in section 2.2 will be useful here as well.

To understand the second consequence of the tracer motion being Hamiltonian, note that there are as many degrees of freedom as there are integrals in involution (namely one, the Hamiltonian ψ), which means that the system is integrable [75]. However, this result holds only for autonomous Hamiltonian systems, i.e. where the Hamiltonian does not depend explicitly on time. As can be seen in (3.5), the explicit time dependence results from the time dependence of the vortex positions, so the tracer motion is integrable for stationary systems. However, for non-stationary systems integrability does not have to hold [5]. It is not even necessary for the vortex motion to be irregular; we will see that non-integrable tracer motion occurs even in periodic systems. Although we only consider two-dimensional flows, it is important to note that the this reasoning does not apply to three-dimensional flows as no stream function can be given in that case. Stationary 3D flows *can* exhibit non-integrable tracer motion [6].

For notational convenience we introduce the map Φ which maps a tracer to its position at a later time, similar to the map Ψ (see (2.22)) which does the same for a point vortex system, so:

$$\Phi(t,T)(\mathbf{z}(t)) = \mathbf{z}(t+T), \quad (3.7)$$

for any tracer trajectory $\mathbf{z}$ and $t, T \in \mathbb{R}$. Note that unlike Ψ the map Φ *does* depend on the time t; this is because the tracer motion is in general not autonomous. Since it corresponds to a Hamiltonian system, Φ is symplectic; since the phase space is two-dimensional this simply means that it is area-conserving. Similar to (2.23) we have

$$\Phi(t+T, S) \circ \Phi(t, T) = \Phi(t, T+S). \tag{3.8}$$

Note that in considering $\Phi(t,T)$ T need not be positive. Negative T simply implies solving (3.1) backwards in time. Filling in (3.8) with $S = -T$ we get

$$\Phi(t+T, -T) \circ \Phi(t, T) = \Phi(t, T-T) = \Phi(t, 0) = \mathcal{I},$$

where $\mathcal{I}$ is the identity mapping. This shows that $\Phi(t,T)$ is invertible, with

$$\Phi(t, T)^{-1} = \Phi(t+T, -T). \tag{3.9}$$

This is not surprising; it simply states that if we move a tracer ahead by a time T and then back we return to the initial position. The following result will prove useful.

Theorem 3.1. $\Phi(t,T)$ *is continuous and differentiable almost everywhere for all* $t, T \in \mathbb{R}$.

Proof. Take any $\epsilon > 0$, and consider the set $\mathcal{R}$ defined as $\mathbb{R}^2$ with closed ϵ-balls removed around the location of every point vortex. The velocity field (3.2) is then Lipschitz continuous on $\mathcal{R}$. Since $\Phi(t,T)$ describes the solution of (3.1), it is then also Lipschitz continuous on $\mathcal{R}$ [20]. By Rademacher's theorem [36] $\Phi(t,T)$ is then also differentiable almost everywhere on $\mathcal{R}$. The theorem now follows by letting ϵ go to zero. □

When we are dealing with a periodic system with period T_p, so

$$\mathbf{v}(\mathbf{x}, t+T_p) = \mathbf{v}(\mathbf{x}, t),$$

we have

$$\Phi(t+T_p, T) = \Phi(t, T). \tag{3.10}$$

It is then useful to define the map f which maps a tracer through one period of the system:

$$f := \Phi(0, T_p). \tag{3.11}$$

Note that combining this with (3.10) and (3.8) we find

$$\Phi(mT_p, (m+n)T_p) = f^n, \tag{3.12}$$

for any $m, n \in \mathbb{Z}$.

3.2 Numerical methods

Obviously to compute tracer advection from some time t to $t+T$, we need to compute the vortex motion on the same interval, using the method developed in section 2.2. One possible approach is then to determine the velocity field $\mathbf{v}(\mathbf{x},t)$ from these positions with (3.2), using some form of time interpolation if necessary. The ODE (3.1) can then be solved using a Runge-Kutta method. This is the method used in [60].

A more efficient approach is to combine the computation of the tracer trajectory with that of the point vortex system. Note that a vortex is advected by the flow, just like a tracer. The difference is that a tracer does not affect the other vortices, so it can be considered a zero strength point vortex. Write the system as $\Gamma(t) = (\mathbf{x}_1(t),...,\mathbf{x}_N(t))$ and the system with tracer included as $\Gamma_{\text{trac}}(t) = (\mathbf{x}_1(t),...,\mathbf{x}_N(t),\mathbf{z}(t))$. We will apply the numerical method to this system. This is an abuse both of notation and of the numerical method since they were set up for systems of identical point vortices; the abuse of notation is temporary, but the accuracy of the resulting method will have to be checked.

The method used is the sixth order symplectic method $\Psi^{(6)}$. Recall that one step in this method is simply a sequence of second order steps $\Psi^{(2)}$ with varying timesteps, so let us see how one of these steps affects our system Γ. Using (2.33):

$$\Psi^{(2)}(h) = \prod_{j=2}^{N+1}\prod_{k=1}^{j-1}\hat{\Omega}_{jk}(h/2) \circ \prod_{j=N+1}^{2}\ \prod_{k=j-1}^{1}\hat{\Omega}_{jk}(h/2).$$

Recall that $\hat{\Omega}_{jk}(\Delta t)$ is the exact map that moves the dipole consisting of vortices j and k ahead by Δt. Note that $\hat{\Omega}_{(N+1)k}(\Delta t)$ is a special case: since vortex $N+1$ is our tracer and as such has zero strength, this map does not affect vortex k; it simply rotates the tracer around this vortex.

The equation above can be rewritten as

$$\Psi^{(2)}(h) = \prod_{j=2}^{N}\prod_{k=1}^{j-1}\hat{\Omega}_{jk}(h/2) \circ \prod_{k=1}^{N}\hat{\Omega}_{(N+1)k}(h/2) \circ \prod_{k=N}^{2}\hat{\Omega}_{(N+1)k}(h/2) \circ \prod_{j=N}^{1}\ \prod_{k=j-1}^{1}\hat{\Omega}_{jk}(h/2). \tag{3.13}$$

Let us analyze how this operator affects the point vortex system Γ and the tracer $\mathbf{z}$ separately. Note that the operators $\hat{\Omega}_{(N+1)k}$ do not affect the vortices, so

$$\Psi^{(2)}(h)(\Gamma) = \left(\prod_{j=1}^{N}\prod_{k=1}^{j-1}\hat{\Omega}_{jk}(h/2) \circ \prod_{j=N}^{1}\ \prod_{k=j-1}^{1}\hat{\Omega}_{jk}(h/2)\right)(\Gamma).$$

This is identical to (2.33), so the motion of the vortices themselves is not influenced by the tracer. This is consistent with the physical system; after all the tracer is a passive particle and as such should not have any effect on the vortices.

As described in section 2.2.3 or seen directly in the above equation, $\Psi^{(2)}$ consists of two sweeps over all dipoles, the second in reverse order. Define Γ_{int} as the intermediate system, i.e. after one sweep:

$$\Gamma_{\text{int}} := \left(\prod_{j=N}^{2} \prod_{k=j-1}^{1} \hat{\Omega}_{jk}(h/2) \right) (\Gamma). \tag{3.14}$$

Now define $\Phi^{(2)}(\Gamma_{\text{int}}, h) : \mathbb{R}^2 \to \mathbb{R}^2$ by:

$$\Phi^{(2)}(\Gamma_{\text{int}}, h)(\mathbf{z}) := \left(\prod_{k=1}^{N} \hat{\Omega}_{(N+1)k}(h/2) \circ \prod_{k=N}^{1} \hat{\Omega}_{(N+1)k}(h/2) \right) (\mathbf{z}), \tag{3.15}$$

where the vortex positions required for $\hat{\Omega}_{(N+1)k}$ are taken from Γ_{int}. $\Phi^{(2)}(\Gamma_{\text{int}}, h)$ then describes the effect of $\Psi^{(2)}(h)$ on $\mathbf{z}$, as can be seen directly in (3.13). What this operator does is rotate the tracer around every vortex in turn, and then again in reverse order, using the intermediate positions of the vortices.

So for every application of $\Psi^{(2)}$, we can compute the vortex motion ignoring the tracer and then use the intermediate position, i.e. between the two sweeps, to compute the tracer motion. This shows how we can find an approximation for $\Phi(t, \Delta t)$; simply apply $\Psi^{(6)}(\Delta t)$ to the system at time t; for each of the seven second order steps (see section 2.2.3) store the timestep used, h_j, and the intermediate system, $\Gamma_{\text{int},j}$, for $1 \le j \le 7$. $\Phi(t, \Delta t)$ is then approximated by:

$$\Phi^{(6)}(t, \Delta t) = \prod_{j=7}^{1} \Phi^{(2)}(\Gamma_{\text{int},j}, h_j). \tag{3.16}$$

Note that the intermediate systems $\Gamma_{\text{int},j}$ do not depend on the tracer position, so their computation needs to be done only once, after which any number of tracers can be computed directly using the above equation. The accuracy of this method is shown in Figure 3.1; just like in section 2.2.4 a regime of sixth order convergence is visible.

In periodic systems with period T_p, if we take Δt such that $T_p = n\Delta t$ for some $n \in \mathbb{N}$ we can find an approximation for f (see (3.11)):

$$f^{(6)}_{\Delta t} = \prod_{j=n}^{1} \Phi^{(6)}((j-1)\Delta t, \Delta t). \tag{3.17}$$

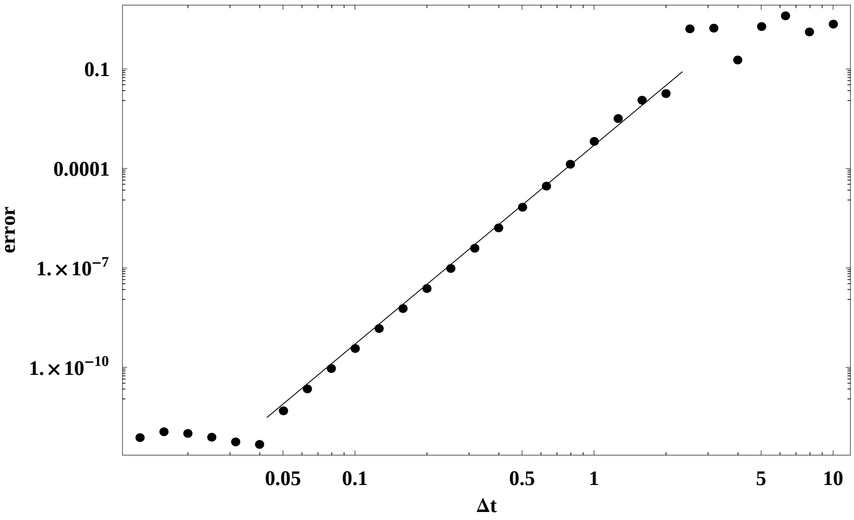

Figure 3.1: Absolute error in $z_x(20)$ as a function of Δt with $\mathbf{z}(0) = (0,0)$, for the irregular four vortex system (see section 2.3.2). The drawn line has a slope of 6.

This can be used to compute tracer trajectories which span more than one period efficiently, since using the above equation combined with (3.12) we don't need to compute the vortex motions for the entire interval.

We will often encounter situations where we need to advect a tracer forward from t to $t+T$ and then back to t. It is important to note that due to possibly chaotic dynamics (3.9) need not apply to the numerical map, so it is possible for the tracer to end up in a completely different position. When the system itself is chaotic, the intermediate vortex positions $\Gamma_{\text{int},j}$ can even diverge for a forward run followed by a backward run. This problem can be solved by locking the system and trajectory, i.e. for the backward run do not recompute $\Gamma_{\text{int},j}$ and the tracer positions, but simply use those from the forward run in reverse order.

3.2.1 Choosing the timestep

For all our systems (see section 2.3) we need to find a timestep Δt which allows us to accurately compute tracer motion. Note that this also means we are accurately computing the vortex motion; if there is a large error here the tracer trajectory will also be affected.

First let us consider the stationary systems. The error in the x-coordinate of a tracer starting at $\mathbf{z}_0 = (0, 0)$ at various times is shown in Figure 3.2 for the stationary three vortex system. Note that for $\Delta t \lesssim 0.1$ the error is caused by numerical accuracy, while for larger Δt the error is mainly caused by the numerical method. So we can take $\Delta t = 0.1$, a reasonable value in the sense that computations with it run in a feasible time. Results for the stationary four vortex system are similar, also leading to $\Delta t = 0.1$.

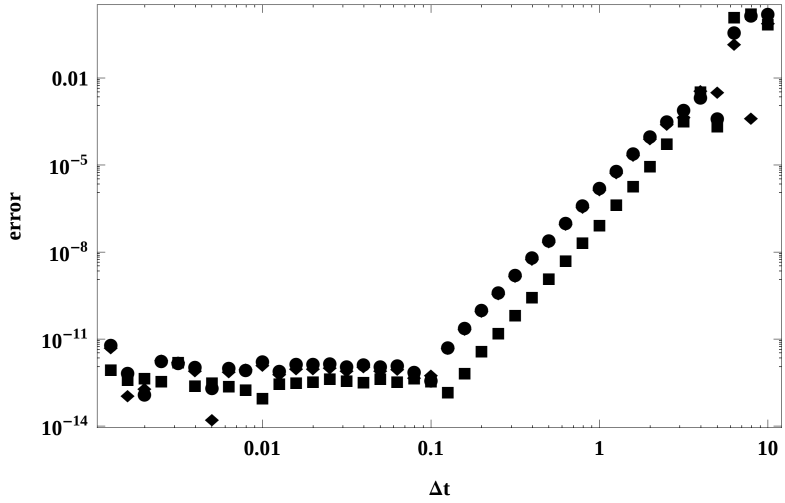

Figure 3.2: Absolute error in $z_x(t)$ as a function of Δt with $\mathbf{z}(0) = (0, 0)$, for the stationary three vortex system. The circles correspond to $t = 100$, the squares to $t = 200$ and the diamonds to $t = 300$.

Next consider the periodic systems. The error in the x-coordinate of a tracer starting at $\mathbf{z}_0 = (0, 0)$ at various times is shown in Figure 3.3 for the 3-periodic three vortex system. Note that unlike the stationary case the error now depends on the time t over which we compute the tracer motion. This is a result of the chaotic dynamics, which causes errors to grow exponentially. However, there is still a clear cutoff in Δt below which numerical accuracy dominates the error; again it is located at $\Delta t = 0.1$. Similar results hold for the other periodic systems, leading to $\Delta t = 0.1$ for the 2-periodic three vortex system and to $\Delta t = 0.02$ for the periodic four vortex system.

Finally, the irregular case. Results for the irregular four vortex system are shown in Figure 3.4. This is similar to the periodic case; the cutoff where numerical accuracy dominates is at $\Delta t = 0.02$. All the timesteps found are summarised in Table 3.1; these will be used in the remainder of this report, unless specified otherwise.

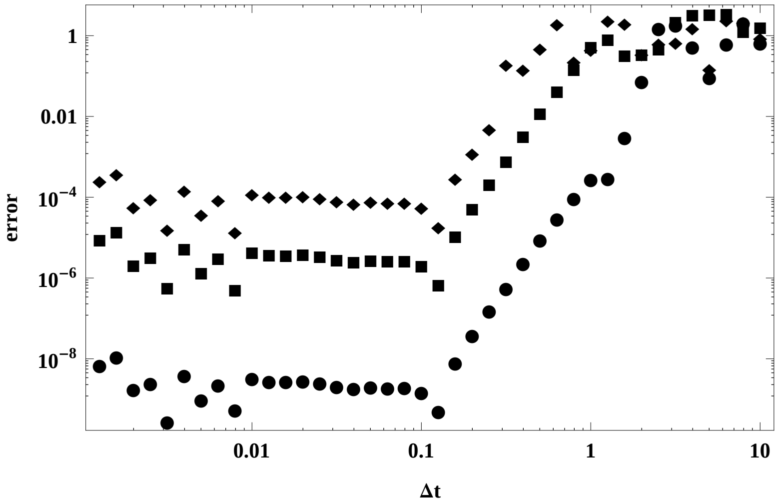

Figure 3.3: Absolute error in $z_x(t)$ as a function of Δt with $\mathbf{z}(0) = (0,0)$, for the 3-periodic three vortex system. The circles correspond to $t = 20T_p$, the squares to $t = 40T_p$ and the diamonds to $t = 60T_p$, with $T_p = 11.577$ the period of the system.

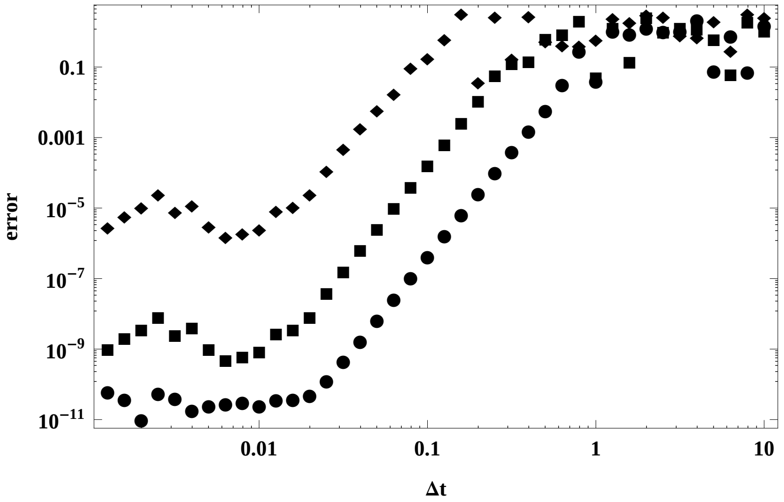

Figure 3.4: Absolute error in $z_x(t)$ as a function of Δt with $\mathbf{z}(0) = (0,0)$, for the irregular four vortex system. The circles correspond to $t = 40$, the squares to $t = 80$ and the diamonds to $t = 120$.

System	Δt
Stationary three vortex	0.1
3-periodic three vortex	0.1
2-periodic three vortex	0.1
Stationary four vortex	0.1
Periodic four vortex	0.02
Irregular four vortex	0.02

Table 3.1: Timestep used for computations, determined by the cutoff value below which numerical accuracy dominates the error.

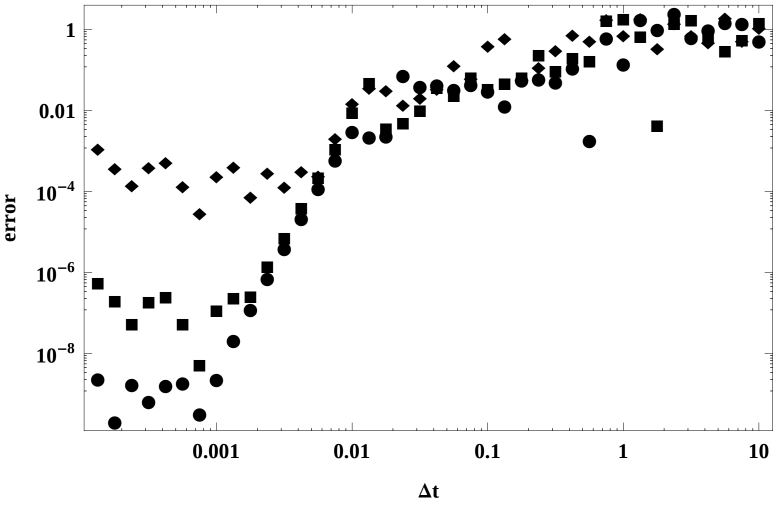

Figure 3.5: Absolute error in $z_x(t)$ as a function of Δt with $\mathbf{z}(0) = (0, 0.96)$, for the irregular four vortex system. The circles correspond to $t = 40$, the squares to $t = 80$ and the diamonds to $t = 120$.

One final remark: to determine Δt as above, we have only considered one initial position for the tracer. Using a different initial position would not necessary lead to the same timestep. For instance, Figure 3.5 shows the error for a tracer starting at $(0, 0.96)$ in the irregular four vortex system (instead of $(0, 0)$ as in Figure 3.4), which is very close to the vortex at $(0, 1)$. Note that a much smaller timestep would have to be taken now, $\Delta t \approx 0.001$ as opposed to $\Delta t \approx 0.02$. In general we found that close to the vortices the numerical method is less accurate, possibly as a result of the large velocity gradients there; this will not be a major issue since this regime is not that important (see sections 3.3 and 4.3.1).

3.3 Results

The most direct way to gain insight into the tracer motion is to consider their trajectories. These are shown for the stationary systems in Figures 3.6 and 3.7. Note that the trajectories are confined to closed one-dimensional curves which are level sets of the stream function Ψ, i.e. streamlines. This is consistent with the theoretical results in section 3.1.

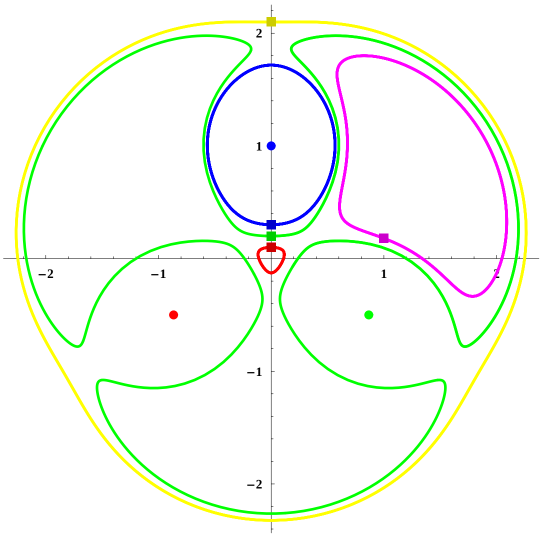

Figure 3.6: Five tracer trajectories for the stationary three vortex system, from $t = 0$ to $t = 100$. The squares indicate the initial positions of the tracers; the dots indicate the positions of the stationary vortices.

Trajectories for the 3-periodic three vortex system are shown in Figure 3.8. We see immediately that the behavior is much more varied than in the stationary case. Four

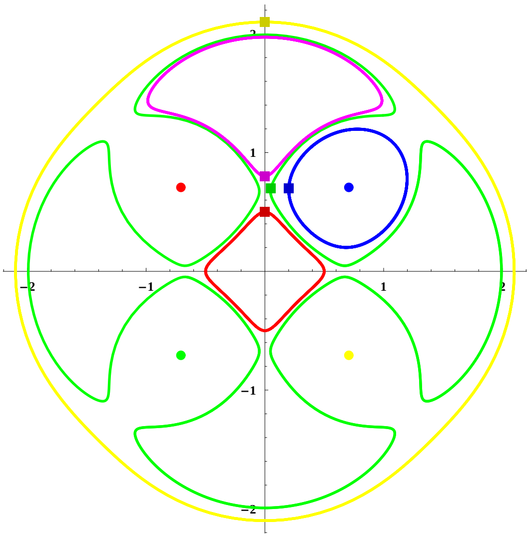

Figure 3.7: Five tracer trajectories for the stationary four vortex system, from $t = 0$ to $t = 400$. The squares indicate the initial positions of the tracers; the dots indicate the positions of the stationary vortices.

types of motion are visible, which can be used to partition the phase space:

- Outer region: The red tracer in the left picture is an example of this. It performs a quasi-circular motion around the vortex system, both as a result of the vortex flow field and of the background rotation.

- Islands: The green and blue tracers in the left picture are examples of this. Although their motion is not strictly periodic, it is confined to small islands in the phase space. For instance, the blue tracer moves between two islands above and below the vortices (for this reason, these islands are referred to as 2-periodic); the green tracer is in a 3-periodic island.

- Vortex cores: The yellow tracer in the left picture is an example of this. It starts near the blue vortex and stays near it (this is not directly visible since the vortex motion is not pictured) while rotating around it. We could call this a 3-periodic island around the vortex; the distinction is made because in any system there are islands at the vortices, while the existence of other islands is not necessary, as we will see later.

- Ergodic region: The purple tracer in the right picture is an example of this. It visits a large region of the phase space, without performing any sort of apparently periodic motion. A more rigorous definition of the ergodic region will be given in the next chapter.

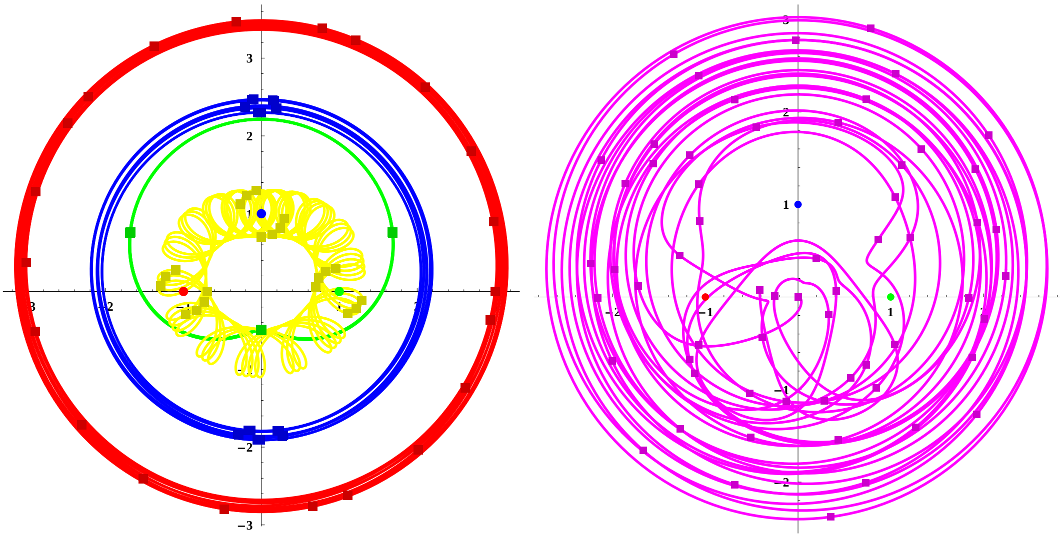

Figure 3.8: Four tracer trajectories for the 3-periodic three vortex system for 20 periods of the system (left) and a fifth one for 60 periods (right). The squares indicate the positions of the tracers at every period of the system; the dots indicate the initial positions of the vortices. To avoid confusion the vortex motion is not pictured; it is shown in Figure 2.7.

Similar results hold for other periodic systems. Consider the periodic four vortex system, for which some trajectories are shown in Figure 3.9. The red tracer is in the outer region, the blue tracer in a 19-periodic island, the yellow tracer in a vortex core and the purple tracer in the ergodic region. In the 2-periodic three vortex case, results are similar (not pictured).

Finally, consider the irregular four vortex system. Three trajectories are shown in Figure 3.10. Three of the regions identified earlier are visible: the green tracer is in the outer region, the red tracer is in the ergodic region and the purple tracer is in the vortex core of the red vortex. This may not be immediately apparent, since it follows an apparently chaotic trajectory similar to the red tracer in the ergodic region. In Figure 3.11 part of the path is shown together with that of the vortex; it is clear that the tracer is 'sticking' to the vortex. Unlike the periodic case there are no islands (see also section 4.3.1).

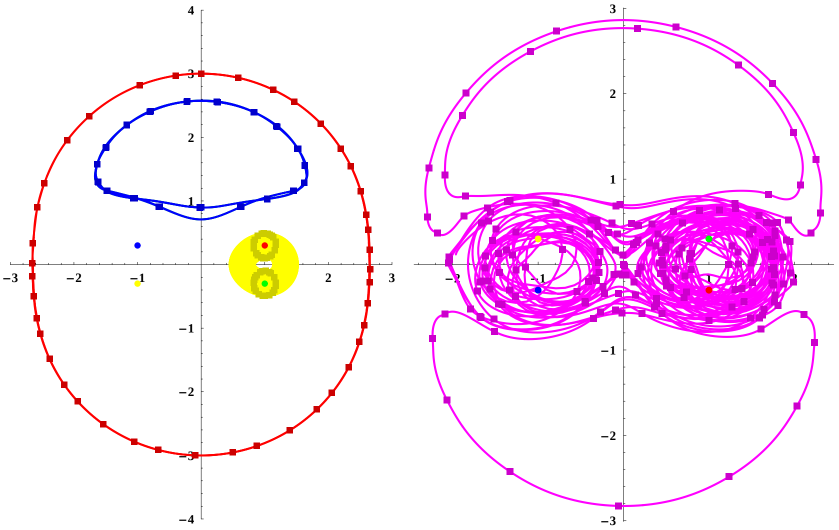

Figure 3.9: Three tracer trajectories for the periodic four vortex system for 40 periods of the system (left) and a fourth one for 200 periods (right). The squares indicate the positions of the tracers at every period of the system; the dots indicate the initial positions of the vortices. To avoid confusion the vortex motion is not pictured; it is shown in Figure 2.10.

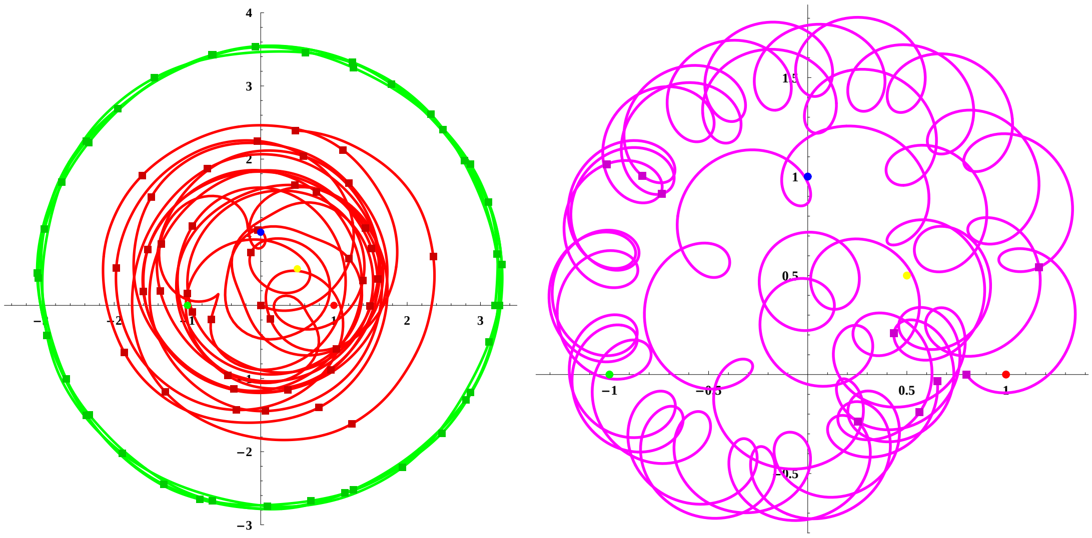

Figure 3.10: Two tracer trajectories for the irregular four vortex system, for $t = 0$ to $t = 400$ (left) and a third one for $t = 0$ to $t = 80$ (right). The squares indicate the positions of the tracers at $t = 0, 10, 20, \ldots$; the dots indicate the initial positions of the vortices. To avoid confusion the vortex motion is not pictured; it is shown in Figure 2.11.

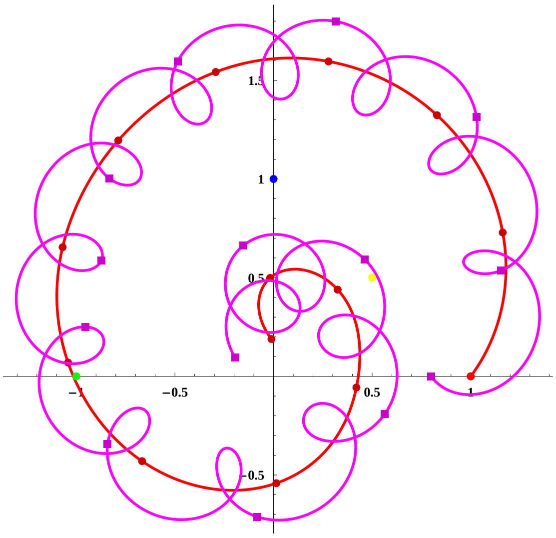

Figure 3.11: Trajectories of a tracer and one of the vortices for the irregular four vortex system, for $t = 0$ to $t = 26$. The squares and red dots indicate the position of the tracer and vortex at t =0, 2, 4, ...; the other dots indicate the initial positions of the remaining vortices.

4 Ergodic regions

In the previous section we analyzed the behavior of tracers in periodic systems by looking directly at their trajectories. Conclusions were drawn based on their positions at every period of the system; for instance, the green tracer in Figure 3.8 is in a 3-periodic island because there are three regions in the phase space it can visit at every period of the system. This suggests that we may learn a lot by considering only these positions, not the full trajectories. This leads to the technique known as Poincaré maps [38], and the ergodic regions they define.

We will start by defining ergodic regions and Poincaré maps and examining the link between the two (section 4.1.1), followed by an analysis of the relationship between ergodic regions and mixing (section 4.1.2). Next, numerical methods are developed to find ergodic regions, using the Poincaré map (section 4.2.1) and a more general technique, tracer cloud spreading, that is also applicable to irregular systems (section 4.2.2). Finally, the ergodic region is computed for all nonstationary example systems (section 4.3.1) and a first look is taken at the actual mixing process (section 4.3.2). We will see two structures important to a description of this process, the partial mixing structure and the mixing time distribution. Both will play an important role in the remainder of this report.

4.1 Theory

4.1.1 Definition and basic properties

Let $(\mathbb{R}^2, \Sigma, \mu)$ be the Lebesgue measure space associated with $\mathbb{R}^2$ [11]. Fix some periodic vortex system with period T_p, defining the periodic mapping $f : \mathbb{R}^2 \to \mathbb{R}^2$ as $f = \Phi(0, T_p)$ (see section 3.1). Recall that f is continuous (Theorem 3.1) and symplectic, i.e. for any $\mathcal{A} \in \Sigma$

$$\mu(f(\mathcal{A})) = \mu(\mathcal{A}). \tag{4.1}$$

Definition 4.1 ([8]). *$\mathcal{R} \in \Sigma$ is an ergodic region if and only if:*

1. *$\mathcal{R}$ is compact;*
2. *$f(\mathcal{R}) = \mathcal{R}$;*

3. $\mu(\mathcal{R}) > 0$*;*

4. For every closed $\mathcal{C} \subset \mathcal{R}$ *such that* $f(\mathcal{C}) = \mathcal{C}$ *and* $\mu(\mathcal{C}) > 0$ *we have* $\mathcal{C} = \mathcal{R}$*.*

So an ergodic region is a compact subset of $\mathbb{R}^2$ invariant under f with positive measure that cannot be partitioned into smaller sets with these properties. Note that $\mathcal{R}$ is allowed to have subsets invariant under f as long as they have zero measure or are not closed; for instance it may contain fixed points or cycles of f. An important property of ergodic regions is that they are invariant under time reversal.

Theorem 4.2. $\mathcal{R} \in \Sigma$ *is an ergodic region under* f *if and only if it is an ergodic region under* f^{-1}*.*

Proof. This is easily checked for each property listed in Definition 4.1 in turn. □

We will now examine the concept of ergodic regions being generated by points within them. This is defined as follows. Define the Poincaré map for $\mathbf{x} \in \mathbb{R}^2$ and $N \in \mathbb{N}$ by [38]:

$$\mathcal{P}(\mathbf{x}, N) := \left\{ f^i(\mathbf{x}) | 0 \leq i \leq N \right\}. \tag{4.2}$$

So the Poincaré map is just the first N iterates of $\mathbf{x}$ under f. N is allowed to run negative, so we consider Poincaré maps for both positive and negative time. We also consider the infinite time variant, for both positive and negative time simultaneously:

$$\mathcal{P}(\mathbf{x}, \infty) := \left\{ f^i(\mathbf{x}) | i \in \mathbb{Z} \right\}. \tag{4.3}$$

Now, define $\mathcal{E}(\mathbf{x})$ as the closure of $\mathcal{P}(\mathbf{x}, \infty)$. If $\mathcal{E}(\mathbf{x})$ is ergodic, then we say that $\mathbf{x}$ generates $\mathcal{E}(\mathbf{x})$. The following two theorems establish the usefulness of this concept.

Theorem 4.3. *Let* $\mathcal{R}$ *be ergodic and* $\mathbf{x} \in \mathcal{R}$ *such that* $\mu(\mathcal{E}(\mathbf{x})) > 0$*. Then* $\mathcal{E}(\mathbf{x}) = \mathcal{R}$*.*

Theorem 4.4. *Let* $\mathcal{R}$ *be ergodic. Then* $\mathcal{E}(\mathbf{x}) = \mathcal{R}$ *for almost all* $\mathbf{x} \in \mathcal{R}$*.*

So an ergodic region $\mathcal{R}$ is generated by almost all points it contains, and in addition for those points $\mathbf{x} \in \mathcal{R}$ that do not generate $\mathcal{R}$ we have $\mu(\mathcal{E}(\mathbf{x})) = 0$, so they do not generate any ergodic region at all by property 3 of Definition 4.1. Several preliminary results, some of which will prove to be quite useful in their own right, will be necessary to prove both theorems.

Lemma 4.5. *If* $\mathbf{e}$ *is a limit point of* $\mathcal{P}(\mathbf{x}, \infty)$ *with* $\mathbf{x} \in \mathbb{R}^2$*, then* $f(\mathbf{e})$ *and* $f^{-1}(\mathbf{e})$ *are limit points of* $\mathcal{P}(\mathbf{x}, \infty)$ *as well.*

Proof. Since $\mathbf{e}$ is a limit point of $\mathcal{P}(\mathbf{x}, \infty)$, there is a sequence (a_n) with $n \in \mathbb{Z}$ and with no repeating elements such that

$$\lim_{n \to \infty} f^{a_n}(\mathbf{x}) = \mathbf{e}.$$

By continuity of f we have:

$$f(\mathbf{e}) = f(\lim_{n\to\infty} f^{a_n}(\mathbf{x})) = \lim_{n\to\infty} f(f^{a_n}(\mathbf{x})) = \lim_{n\to\infty} f^{a_n+1}(\mathbf{x}),$$

so $f(\mathbf{e})$ is the limit of a sequence in $\mathcal{P}(\mathbf{x},\infty)$ with no repeating elements and hence is a limit point. The proof for $f^{-1}(\mathbf{e})$ is similar. □

Corollary 4.6. *For all* $\mathbf{x} \in \mathbb{R}^2$, $\mathcal{P}(\mathbf{x},\infty)$ *contains either no isolated points or only isolated points.*

Proof. This follows directly from Lemma 4.5. □

Proof of theorem 4.3. Let $\mathcal{R}$ be ergodic and $\mathbf{x} \in \mathcal{R}$ such that $\mu(\mathcal{E}(\mathbf{x})) > 0$; we then need to show that $\mathcal{E}(\mathbf{x}) = \mathcal{R}$. Consider $\mathcal{P}(\mathbf{x},\infty)$. If it contains only isolated points, then $\mu(\mathcal{E}(\mathbf{x})) = 0$, contradicting the assumptions. So by Corollary 4.6, $\mathcal{P}(\mathbf{x},\infty)$ contains no isolated points.

Now consider any $\mathbf{e} \in \mathcal{E}(\mathbf{x})$, so $\mathbf{e}$ is the limit of points in $\mathcal{P}(\mathbf{x},\infty)$. Since it contains no isolated points, $\mathbf{e}$ must be a limit point of $\mathcal{P}(\mathbf{x},\infty)$, and so by Lemma 4.5 both $f(\mathbf{e})$ and $f^{-1}(\mathbf{e})$ must be as well, showing $f(\mathbf{e}) \in \mathcal{E}(\mathbf{x})$ and $f^{-1}(\mathbf{e}) \in \mathcal{E}(\mathbf{x})$. This shows that $f(\mathcal{E}(\mathbf{x})) = \mathcal{E}(\mathbf{x})$.

Since $\mathbf{x} \in \mathcal{R}$ and $f(\mathcal{R}) = \mathcal{R}$, we have $\mathcal{P}(\mathbf{x},\infty) \subset \mathcal{R}$. Because $\mathcal{R}$ is closed we then have $\mathcal{E}(\mathbf{x}) \subset \mathcal{R}$. We have already shown that $f(\mathcal{E}(\mathbf{x})) = \mathcal{E}(\mathbf{x})$ and by assumption $\mu(\mathcal{E}(\mathbf{x})) > 0$, so by property 4 of Definition 4.1 $\mathcal{E}(\mathbf{x}) = \mathcal{R}$, completing the proof. □

The following result is necessary for the proof of theorem 4.4.

Lemma 4.7. *Let* $\mathcal{A}$ *be an open subset of an ergodic region* $\mathcal{R}$, $\mathbf{x}$ *an interior point of* $\mathcal{R}$ *and* $N \in \mathbb{N}_0$. *Then there exists* $n \geq N$ *such that* $\mathbf{x} \in f^n(\mathcal{A})$.

Proof. Consider

$$\mathcal{C} = \bigcup_{n=N}^{\infty} f^n(\mathcal{A})$$

and its closure $\mathcal{D}$. We have $f(\mathcal{C}) = \mathcal{C} \setminus f^N(\mathcal{A}) \subset \mathcal{C}$ and by (4.1) $\mu(f(\mathcal{C})) = \mu(\mathcal{C})$; these two properties imply $f(\mathcal{C}) = \mathcal{C}$ because $\mathcal{C}$ is open. By continuity of f we then have $f(\mathcal{D}) = \mathcal{D}$, and since $\mu(\mathcal{D}) \geq \mu(\mathcal{C}) \geq \mu(f^N(\mathcal{A})) \overset{(4.1)}{=} \mu(\mathcal{A}) > 0$ we have $\mathcal{D} = \mathcal{R}$ by property 4 of Definition 4.1. So $\mathbf{x}$ is an interior point of $\mathcal{D}$, which is the closure of $\mathcal{C}$. Since $\mathcal{C}$ is open, we have $\mathbf{x} \in \mathcal{C}$. So there is an $n \geq N$ such that $\mathbf{x} \in f^n(\mathcal{A})$, completing the proof. □

Proof of Theorem 4.4. For $\mathbf{y} \in \mathcal{R}$ and $\delta > 0$, define

$$\mathcal{A}(\mathbf{y}, \delta) = \{\mathbf{x} | \mathbf{x} \in \mathbb{R}^2, \forall n \in \mathbb{Z} : ||f^n(\mathbf{x}) - \mathbf{y}|| > \delta\},$$

i.e. those points for which the Poincaré map stays at least a distance δ from $\mathbf{y}$ for both positive and negative time. Now fix $\mathbf{y}$ and δ; we will write $\mathcal{A}$ instead of $\mathcal{A}(\mathbf{y}, \delta)$ for convenience and show that $\mu(\mathcal{A}) = 0$.

Suppose $\mu(\mathcal{A}) > 0$. Consider

$$\mathcal{B} = \bigcup_{n=-\infty}^{\infty} f^n(\mathcal{A})$$

and its closure $\mathcal{C}$. Clearly we have $\mathcal{B} = f(\mathcal{B})$ and so by continuity of f $\mathcal{C} = f(\mathcal{C})$. Also, we have $\mu(\mathcal{C}) \geq \mu(\mathcal{B}) \geq \mu(\mathcal{A}) > 0$ and so by property 4 of Definition 4.1 we have $\mathcal{C} = \mathcal{R}$. However, by the definition of $\mathcal{A}$ the point $\mathbf{y}$ must be a distance of at least δ from $\mathcal{C}$, contradicting $\mathbf{y} \in \mathcal{R}$.

As a result of the above, we have:

$$\mu(\mathcal{A}(\mathbf{y}, \delta)) = 0 \tag{4.4}$$

for all $\mathbf{y} \in \mathcal{R}$ and $\delta > 0$. Now consider

$$\mathcal{A}(\delta) = \{\mathbf{x} | \mathbf{x} \in \mathbb{R}^2, \exists \mathbf{r} \in \mathcal{R} \forall n \in \mathbb{Z} : ||f^n(\mathbf{x}) - \mathbf{r}|| > \delta\}$$

for $\delta > 0$, i.e. those points for which the Poincaré map stays at least a distance δ from at least one point in $\mathcal{R}$ for both positive and negative time. Now fix δ; we will show that $\mu(\mathcal{A}(\delta)) = 0$.

First take $\{\mathbf{y}_i | 1 \leq i \leq N\}$ for some $N \in \mathbb{N}$ with $\mathbf{y}_i \in \mathcal{R}$ and such that for all $\mathbf{r} \in \mathcal{R}$ there is an i such that $||\mathbf{r} - \mathbf{y}_i|| < \delta/2$; this is possible because $\mathcal{R}$ is compact by property 1 of Definition 4.1. Now let any $\mathbf{x} \in \mathcal{A}(\delta)$ be given. Fix $\mathbf{r}$ such that $||f^n(\mathbf{x}) - \mathbf{r}|| > \delta$ for all $n \in \mathbb{Z}$; this is possible by the definition of $\mathcal{A}(\delta)$. Now, take i such that $||\mathbf{r} - \mathbf{y}_i|| < \delta/2$; this is possible by the definition of $\{\mathbf{y}_i\}$. We have using the triangle inequality and the two inequalities above:

$$||f^n(\mathbf{x}) - \mathbf{y}_i|| \geq ||f^n(\mathbf{x}) - \mathbf{r}|| - ||\mathbf{r} - \mathbf{y}_i|| > \delta - \delta/2 = \delta/2$$

for all $n \in \mathbb{Z}$. So we have $\mathbf{x} \in \mathcal{A}(\mathbf{y}_i, \delta/2)$.

The result above shows that

$$\mathcal{A}(\delta) \subset \bigcup_{i=1}^{N} \mathcal{A}(\mathbf{y}_i, \delta/2),$$

so we have:

$$\mu(\mathcal{A}(\delta)) \leq \mu\left(\bigcup_{i=1}^{N} \mathcal{A}(\mathbf{y}_i, \delta/2)\right) \leq \sum_{i=1}^{N} \mu(\mathcal{A}(\mathbf{y}_i, \delta/2)) = 0,$$

using (4.4). Now, note that $\mathcal{A}(\delta_1) \subset \mathcal{A}(\delta_2)$ for $\delta_1 > \delta_2$, so:

$$\mu\left(\bigcup_{\delta>0} \mathcal{A}(\delta)\right) = \lim_{\delta\to 0} \mu(\mathcal{A}(\delta)) = 0.$$

So for almost all $\mathbf{x} \in \mathcal{R}$, we have that there is no $\delta > 0$ such that $\mathbf{x} \in \mathcal{A}(\delta)$ and so the set

$$\{f^n(\mathbf{x}) | n \in \mathbb{Z}\} = \mathcal{P}(\mathbf{x}, \infty)$$

is dense in $\mathcal{R}$. This shows that $\mathcal{R} \subset \mathcal{E}(\mathbf{x})$ and so by Theorem 4.3 we have $\mathcal{E}(\mathbf{x}) = \mathcal{R}$, completing the proof. □

4.1.2 Relationship to mixing

In the previous section we have seen that one tracer 'fills up' an ergodic region, making these regions quite useful in describing the advection properties. However, does it also tell us anything about the mixing properties of the system? The answer is yes; under certain conditions any open set in an ergodic region will eventually fill the entire region in some sense, as we will now show.

Definition 4.8. *Consider an ergodic region $\mathcal{R}$. If for all open $\mathcal{A} \subset \mathcal{R}$ we have $f^n(\mathcal{A}) \cap \mathcal{A} \neq \emptyset$ asymptotically for all $n \in \mathbb{N}$ then $\mathcal{R}$ is called aperiodic.*

Asymptotically for all $n \in \mathbb{N}$ here means that if $g(N)$ is the number of n with $n < N$ such that $f^n(\mathcal{A}) \cap \mathcal{A} \neq \emptyset$, we have

$$\lim_{N\to\infty} \frac{g(N)}{N} = 1. \tag{4.5}$$

If an ergodic region is not aperiodic, then $f^n(\mathcal{A}) \cap \mathcal{A} = \emptyset$ for some finite proportion of the periods. This implies that tracers spend time in different regions periodically or quasiperiodically. For instance, we would expect this to be the case if the ergodic region contained multiple disconnected components-obviously this would impede mixing. We are now ready to formulate the main result connecting mixing to the ergodic region and Poincaré maps, writing $\mathcal{N}(\mathcal{X}, r)$ for the r-neighbourhood of $\mathcal{X}$:

Theorem 4.9. *If $\mathcal{R}$ is an aperiodic ergodic region, then for any open $\mathcal{A} \subset \mathcal{R}$ and $\epsilon > 0$ we have $\mathcal{R} \subset \mathcal{N}(f^n(\mathcal{A}), \epsilon)$ asymptotically for all $n \in \mathbb{N}$.*

Proof. Let open $\mathcal{A} \subset \mathcal{R}$ and $\epsilon > 0$ be given. Take a finite collection of points $\{\mathbf{y}_i | 1 \leq i \leq N\}$ with $\mathbf{y}_i \in \mathcal{R}$ such that for all $\mathbf{x} \in \mathcal{R}$ there is an i such that $||\mathbf{x} - \mathbf{y}_i|| < \epsilon/2$. This is possible because $\mathcal{R}$ is compact by property 1 of Definition 4.1.

Define $\mathcal{B}_i$ as the open disk centered at $\mathbf{y}_i$ with radius $\epsilon/2$ for $1 \leq i \leq N$. Because $\mathcal{R}$ is ergodic and $\mathcal{A}$ and $\mathcal{B}_i$ are open there must at least one $m \in \mathbb{N}_0$ such that $f^m(\mathcal{A}) \cap \mathcal{B}_i \neq \emptyset$ (Lemma 4.7). For the following, fix i. Note that $f^m(\mathcal{A}) \cap \mathcal{B}_i$ is open so, because $\mathcal{R}$ is aperiodic, we have $f^m(\mathcal{A}) \cap \mathcal{B}_i \cap f^n(f^m(\mathcal{A}) \cap \mathcal{B}_i) \neq \emptyset$ asymptotically for all $n \in \mathbb{N}$. Since $f^m(\mathcal{A}) \cap \mathcal{B}_i \cap f^n(f^m(\mathcal{A}) \cap \mathcal{B}_i) \subset f^{n+m}(\mathcal{A}) \cap \mathcal{B}_i$, we have $f^{n+m}(\mathcal{A}) \cap \mathcal{B}_i \neq \emptyset$ asymptotically for all $n \in \mathbb{N}$ and so $f^n(\mathcal{A}) \cap \mathcal{B}_i \neq \emptyset$ asymptotically for all $n \in \mathbb{N}$.

Since the property derived above holds asymptotically for all $n \in \mathbb{N}$ for individual i, it must hold for all i with $1 \leq i \leq N$ simultaneously, asymptotically for all $n \in \mathbb{N}$. We will now show that for fixed n this implies $\mathcal{R} \subset \mathcal{N}(f^n(\mathcal{A}), \epsilon)$, thereby completing the proof.

Fix n such that $f^n(\mathcal{A}) \cap \mathcal{B}_i \neq \emptyset$ for $1 \leq i \leq N$ and consider any $\mathbf{x} \in \mathcal{R}$. Now fix i such that $||\mathbf{x} - \mathbf{y}_i|| < \epsilon/2$; this is possible by the choice of $\mathbf{y}_i$. Take $\mathbf{z}$ such that $\mathbf{z} \in f^n(\mathcal{A}) \cap \mathcal{B}_i$; then by the definition of $\mathcal{B}_i$ we have $||\mathbf{y}_i - \mathbf{z}|| < \epsilon/2$. Now by the triangle inequality:

$$||\mathbf{x} - \mathbf{z}|| \leq ||\mathbf{x} - \mathbf{y}_i|| + ||\mathbf{y}_i - \mathbf{z}|| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Since $\mathbf{z} \in f^n(\mathcal{A})$ this shows $\mathbf{x} \in \mathcal{N}(f^n(\mathcal{A}), \epsilon)$. Since $\mathbf{x}$ was freely chosen in $\mathcal{R}$, we have $\mathcal{R} \subset \mathcal{N}(f^n(\mathcal{A}), \epsilon)$. □

So any open set will be spread over the full ergodic region $\mathcal{R}$ by repeated applications of f. Of course due to area preservation it can never actually fill $\mathcal{R}$, but it can do so arbitrarily well in the sense described in the above Theorem. Note that due to invariance of $\mathcal{R}$ it is not possible for tracers inside it to escape, so the mixing is strictly limited to $\mathcal{R}$.

We have already seen that no mixing takes place in stationary systems. Using the concepts introduced in this chapter we can show this formally by demonstrating that there are no ergodic regions in these systems. Note that we have to arbitrarily pick a period T_p to define f.

Theorem 4.10. *There are no ergodic regions in stationary vortex systems.*

Proof. If there is an ergodic region, then it has a generator by theorem 4.4, so take $\mathbf{x}$ such that $\mathcal{E}(\mathbf{x})$ is ergodic. Consider the set

$$\mathcal{C} = \{\mathbf{y} | \psi(\mathbf{x}) = \psi(\mathbf{y})\},$$

with ψ the stream function (see (3.5)); the time argument has been omitted since it is constant for a stationary system. Now $\mathcal{P}(\mathbf{x}, \infty) \subset \mathcal{C}$ since Ψ and thus f conserves ψ. ψ is continuous, so $\mathcal{C}$ is closed, showing that $\mathcal{E}(\mathbf{x}) \subset \mathcal{C}$. Now, it can be seen directly that ψ is not constant on any set of positive measure and so $\mathcal{C}$ has zero measure. This shows that $\mu(\mathcal{E}(\mathbf{x})) = 0$, contradicting property 3 of Definition 4.1. So there are no ergodic regions. □

4.2 Numerical methods

4.2.1 Poincaré maps

The results in section 4.1.1 show a simple way to find ergodic regions for periodic systems: simply compute the Poincaré map $\mathcal{P}(\mathbf{x}, N)$ for some $\mathbf{x}$ we suspect to be in the ergodic region and for large N. If this set appears to fill a region with positive measure densely, then we have found an ergodic region. Note that by Theorem 4.3 we are assured that we are seeing the entire ergodic region and not some subset of it. We will not attempt to formally prove that any of the regions we will see is actually ergodic; this is generally quite difficult [78].

To verify the mixing properties, we need to check aperiodicity. The following result connects this to properties of the Poincaré map.

Lemma 4.11. *An ergodic region $\mathcal{R}$ is aperiodic if and only if for all open $\mathcal{A} \subset \mathcal{R}$ and $\mathbf{x} \in \mathcal{R}$ such that $\mathbf{x}$ is a generator of $\mathcal{R}$ we have: asymptotically for all $n \in \mathbb{N}$ there is an $i \in \mathbb{Z}$ such that both $f^i(\mathbf{x})$ and $f^{i+n}(\mathbf{x})$ are in $\mathcal{A}$.*

Proof. We will show that $f^n(\mathcal{A}) \cap \mathcal{A} \neq \emptyset$ if and only if for all generators $\mathbf{x}$ of $\mathcal{R}$ there is an $i \in \mathbb{Z}$ such that both $f^i(\mathbf{x})$ and $f^{i+n}(\mathbf{x})$ are in $\mathcal{A}$.

Suppose $f^n(\mathcal{A}) \cap \mathcal{A} \neq \emptyset$. Since $\mathcal{A}$ is open, so is $f^n(\mathcal{A})$ using continuity of f, and so their intersection is open as well. Since $\mathcal{P}(\mathbf{x}, \infty)$ is dense in $\mathcal{R}$ and $f^n(\mathcal{A}) \cap \mathcal{A} \subset \mathcal{R}$, we can take $j \in \mathbb{Z}$ such that $f^j(\mathbf{x}) \in f^n(\mathcal{A}) \cap \mathcal{A}$. Now take $i = j - n$. Since $f^j(\mathbf{x}) \in f^n(\mathcal{A})$, we have $f^i(\mathbf{x}) = f^{j-n}(\mathbf{x}) \in \mathcal{A}$. Also, $f^{i+n}(\mathbf{x}) = f^j(\mathbf{x}) \in \mathcal{A}$, completing the proof in one direction.

Now suppose there is an $i \in \mathbb{Z}$ such that both $f^i(\mathbf{x})$ and $f^{i+n}(\mathbf{x})$ are in $\mathcal{A}$. Since $f^i(\mathbf{x}) \in \mathcal{A}$, we have $f^{i+n}(\mathbf{x}) \in f^n(\mathcal{A})$ and so $f^{i+n}(\mathbf{x}) \in \mathcal{A} \cap f^n(\mathcal{A})$, implying $f^n(\mathcal{A}) \cap \mathcal{A} \neq \emptyset$. This completes the proof. □

So to check that an ergodic region is aperiodic, we take an open region $\mathcal{A}$ inside it and for a generator $\mathbf{x}$ compute a list of n such that $f^n(\mathbf{x})$ is in $\mathcal{A}$; this can be done directly from the Poincaré map (for $n \leq N$). Then the proportion of $j \leq J$ that occur in this list as the difference between two of its elements should converge to 1 for $J \to \infty$.

The Poincaré map technique is very efficient in the sense that every application of f yields a point in the ergodic region. However, there are some disadvantages to this technique. In the first place, it is not suitable for resolving fine scale structures; we are always forced to evaluate the full ergodic region. To create detailed pictures of parts of the ergodic region, we have to wait for the Poincaré map to visit these parts, meaning we may need many applications of f to find a single point. Another disadvantage is that we learn nothing about time scales or intermediate structures of the mixing. Theorem 4.9 guarantees that mixing will take place eventually, but gives no bounds on the time required to accomplish this and does not describe the process. Finally, the technique is defined in terms of f and so can be applied only to periodic systems. In the following section we will see how this limitation can be overcome.

4.2.2 Tracer cloud spreading

To generalize the method to irregular systems, we need a definition of the ergodic region that does not depend on f. A natural method to accomplish this is to make use of the mixing properties established in section 4.1.2. Write $\mathcal{B}(\mathbf{x}, r)$ for the open disk with radius r centered at $\mathbf{x}$ and $\mathrm{d}(\mathbf{x}, \mathcal{A})$ for the distance of $\mathbf{x}$ to some set $\mathcal{A}$, i.e.

$$\mathrm{d}(\mathbf{x}, \mathcal{A}) = \inf_{\mathbf{a} \in \mathcal{A}} ||\mathbf{x} - \mathbf{a}||. \tag{4.6}$$

Theorem 4.12. *Let $\mathcal{R}$ be an aperiodic ergodic region. Then there exists $\delta > 0$ such that for almost all $\mathbf{x} \in \mathcal{R}$*

$$\mathcal{R} = \{\mathbf{r} | \mathbf{r} \in \mathcal{X}, \liminf_{T \to \pm\infty, T \in \mathcal{T}} \mathrm{d}(\mathbf{r}, \phi(-T, T)\mathcal{B}(\phi(0, -T)\mathbf{x}, \delta)) = 0\},$$

where $\mathcal{T} = \{T | T \in \mathbb{R}_+, \mathcal{B}(\phi(0, -T)\mathbf{x}, \delta) \subset \phi(0, -T)\mathcal{R}\}$.

Proof. First, we need to pick δ. Take $\mathbf{y} \in \mathcal{R}$ and $\delta > 0$ such that $\mathcal{B}(\mathbf{y}, 2\delta) \subset \mathcal{R}$; this is possible because $\mathcal{R}$ is compact and has positive measure. We will show that the theorem holds when $\mathbf{x}$ is a generator of $\mathcal{R}$; by Theorem 4.4 it then holds for almost all $\mathbf{x} \in \mathcal{R}$. So fix some $\mathbf{x}$ such that $\mathbf{x}$ is a generator of $\mathcal{R}$.

First, suppose $\mathbf{r} \in \mathcal{R}$. We need to show that

$$\liminf_{T \to \pm\infty, T \in \mathcal{T}} \mathrm{d}(\mathbf{r}, \phi(-T, T)\mathcal{B}(\phi(0, -T)\mathbf{x}, \delta)) = 0. \tag{4.7}$$

Define $\mathcal{T}' = \{nT_p | n \in \mathbb{N}, f^{-n}(\mathbf{x}) \in \mathcal{B}(\mathbf{y}, \delta/2)\}$ where T_p is the period of the system. Note that we can write $\mathcal{T}' = \{T | T \in \mathbb{R}_+, T/T_p \in \mathbb{N}, \Phi(0, -T)\mathbf{x} \in \mathcal{B}(\mathbf{y}, \delta/2)\}$. Now, $\Phi(0, -T)\mathbf{x} \in \mathcal{B}(\mathbf{y}, \delta/2)$ implies $\mathcal{B}(\Phi(0, -T)\mathbf{x}, \delta) \subset \mathcal{B}(\mathbf{y}, 2\delta) \subset \mathcal{R} = f^{-n}(\mathcal{R}) = \phi(0, -T)\mathcal{R}$, so $\mathcal{T}' \subset \mathcal{T}$. As a result of this

$$\liminf_{T\to\pm\infty, T\in\mathcal{T}} \mathrm{d}(\mathbf{r}, \phi(-T,T)\mathcal{B}(\phi(0,-T)\mathbf{x},\delta)) \leq \liminf_{T\to\pm\infty, T\in\mathcal{T}'} \mathrm{d}(\mathbf{r}, \phi(-T,T)\mathcal{B}(\phi(0,-T)\mathbf{x},\delta)). \tag{4.8}$$

Also as a result of $\Phi(0, -T)\mathbf{x} \in \mathcal{B}(\mathbf{y}, \delta/2)$ we have $\mathcal{B}(\mathbf{y}, \delta/2) \subset \mathcal{B}(\phi(0, -T)\mathbf{x}, \delta))$ for $T \in \mathcal{T}'$. We now have:

$$\liminf_{T\to\pm\infty, T\in\mathcal{T}'} \mathrm{d}(\mathbf{r}, \phi(-T,T)\mathcal{B}(\phi(0,-T)\mathbf{x},\delta)) \leq \liminf_{T\to\pm\infty, T\in\mathcal{T}'} \mathrm{d}(\mathbf{r}, \phi(-T,T)\mathcal{B}(\mathbf{y},\delta/2)) =$$
$$\liminf_{n\to\pm\infty, f^{-n}(\mathbf{x})\in\mathcal{B}(\mathbf{y},\delta/2)} \mathrm{d}(\mathbf{r}, f^n(\mathcal{B}(\mathbf{y},\delta/2))) = 0,$$

where the last inequality follows from Theorem 4.9 and aperiodicity of $\mathcal{R}$. This result together with (4.8) implies (4.7), completing the first part of the proof.

Next, suppose $\mathbf{r}$ is not in $\mathcal{R}$. We need to show that

$$\liminf_{T\to\pm\infty, T\in\mathcal{T}} \mathrm{d}(\mathbf{r}, \phi(-T,T)\mathcal{B}(\phi(0,-T)\mathbf{x},\delta)) \neq 0. \tag{4.9}$$

Write $r = d(\mathbf{r}, \mathcal{R})$; since $\mathcal{R}$ is closed we have $r > 0$. Now suppose for some $T \in \mathcal{T}$ we have

$$\mathrm{d}(\mathbf{r}, \phi(-T,T)\mathcal{B}(\phi(0,-T)\mathbf{x},\delta)) < r. \tag{4.10}$$

Then $\phi(-T,T)\mathcal{B}(\phi(0,-T)\mathbf{x},\delta)$ is not a subset of $\mathcal{R}$, since then we would have $d(\mathbf{r}, \mathcal{R}) < r$. However, this means that $\phi(0,-T)\phi(-T,T)\mathcal{B}(\phi(0,-T)\mathbf{x},\delta) = \mathcal{B}(\phi(0,-T)\mathbf{x},\delta)$ is not a subset of $\phi(0,-T)\mathcal{R}$, contradicting $T \in \mathcal{T}$.

So there is no $T \in \mathcal{T}$ such that (4.10) holds, meaning that

$$\liminf_{T\to\pm\infty, T\in\mathcal{T}} \mathrm{d}(\mathbf{r}, \phi(-T,T)\mathcal{B}(\phi(0,-T)\mathbf{x},\delta)) \geq r > 0.$$

So (4.9) holds, completing the proof. □

This gives us an alternate method to find ergodic regions, which we will call tracer cloud spreading. We take $\mathbf{x}$ which we suspect to be in an ergodic region, and we follow its trajectory backwards over some large time T to $\Phi(0, -T)\mathbf{x}$. Next we consider the disk around this point $\mathcal{B}(\Phi(0, -T)\mathbf{x}, \delta)$ for some δ small compared to a typical length scale of the ergodic region. We can approximate this disk by considering a large amount N of tracers spread out randomly inside it. Finally, we go back to $t = 0$ by applying $\Phi(-T, T)$ to these tracers; we then have a set of points that asymptotically (i.e. for $T, N \to \infty$) fill the ergodic region densely, similar to a Poincaré map. Note that we do need to satisfy the condition

$$\mathcal{B}(\phi(0,-T)\mathbf{x},\delta) \subset \phi(0,-T)\mathcal{R}, \tag{4.11}$$

i.e. $\phi(0,-T)\mathbf{x}$ should be at least a distance δ from any boundary of the image of $\mathcal{R}$ under $\phi(0,-T)$; if we do not satisfy this then part of the tracer cloud is outside the ergodic region. In practice this is not an issue if δ is taken small enough.

Tracer cloud spreading is much less efficient than the Poincaré map-for each point we need to evaluate $\Phi(-T,T)$, typically for T much larger than the period of the system T_p, while for the Poincaré map we only need to evaluate $f = \Phi(0,T_p)$ for every point. Its advantages are that it can be applied to irregular systems and that it is directly related to the mixing process. This means that by considering finite T instead of the limit $T \to \infty$ we can obtain information on the intermediate mixing structures and timescales.

4.3 Results

4.3.1 Computation of ergodic regions

The ergodic region for the 3-periodic three vortex system, computed using the Poincaré map, is shown in Figure 4.1. We see that indeed a region of positive measure is filled densely by the iterates of one point. Other structures such as islands and vortex cores are visible as holes in this ergodic region, while the outer region is simply the entire area outside it.

The ergodic region for the periodic four vortex system is shown in Figure 4.2. Note that there are two large islands above and below the vortices, surrounded by a chain of smaller islands on their boundaries. These islands are 19-periodic (the blue tracer in Figure 3.9 is in such an island) and are in fact surrounded by smaller 152-periodic islands themselves, as can be seen in the detailed picture. This structure might continue on even smaller scales, but obtaining a more detailed figure was not possible due to the computing time involved. Note that in general an island does not necessarily have such an irregular boundary, as can be seen on the right in Figure 4.1.

The ergodic region for the 2-periodic three vortex system is shown in Figure 4.3. Note that it is similar to the periodic four vortex system; this may be a result of the dipoles present in both systems. The large islands are now to the left and right of the vortices and again have irregular boundaries, although no island chain along the boundary is clearly visible in this case. Note that there is a second ergodic region

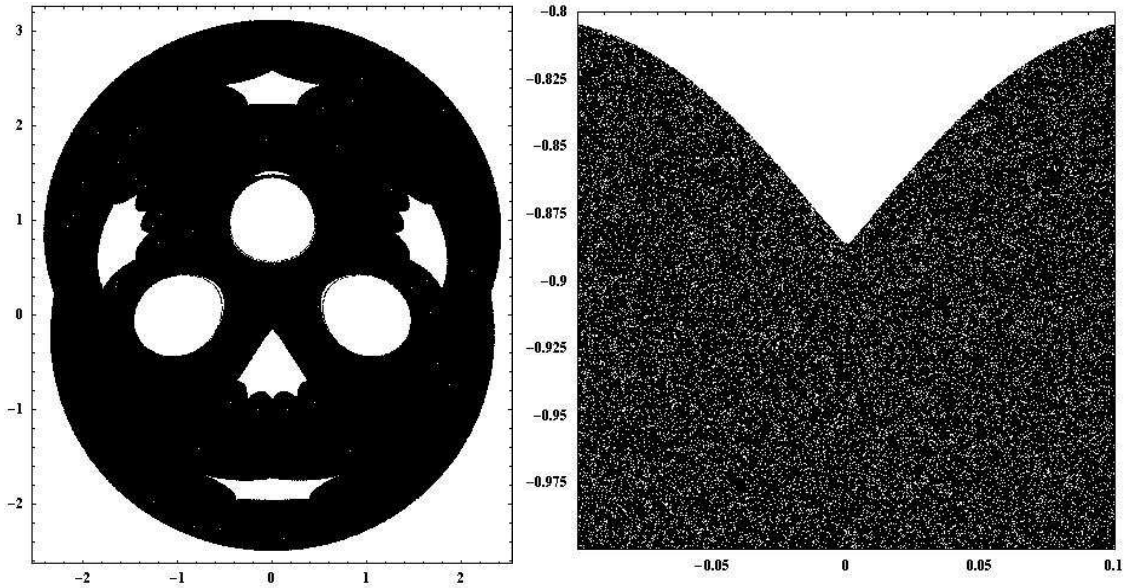

Figure 4.1: Poincaré map of $(0, 0)$ with $N = 10^6$ for the 3-periodic three vortex system. Right: detail of a 3-periodic island boundary.

separating the vortex cores which is not connected to the main one.[1]

Before using tracer cloud spreading to find the ergodic region of the irregular four vortex system, we verify that it is equivalent to the Poincaré map technique by applying it to the 3-periodic three vortex system. This is shown in Figure 4.4; note that the ergodic region is the same as that found earlier using the Poincaré map (see Figure 4.1). The resolution is far lower; 10^5 points are used for tracer cloud spreading as opposed to 10^6 for the Poincaré map. This is because of the much higher computing times involved.

The ergodic region in the irregular four vortex system, found using tracer cloud spreading, is shown in Figure 4.5. Note that the ergodic region is fairly well defined, but there is some fraying along the edges. This has to do with the nature of the advection; near the boundary of the ergodic region and the outer region tracers tend to get trapped for long times in circular paths before at some point returning to the main ergodic region [14], so convergence is very slow here. Using higher T would lead to an improvement but is unfeasible due to computation time. Also note that there are no islands (except possibly near the boundary with the outer region), but vortex

---

[1] Zaslavsky found that the barrier separating the two ergodic regions may be permeable [47]. No evidence of this was found for this particular system.

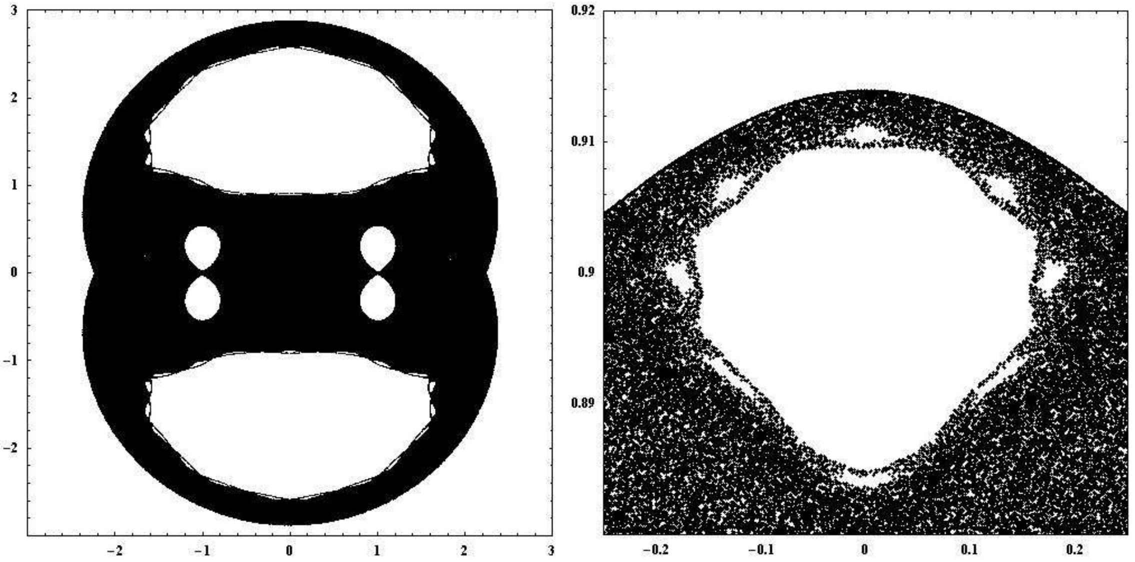

Figure 4.2: Poincaré map of $(0, 0)$ with $N = 10^6$ for the periodic four vortex system. Right: detail of a 19-periodic island.

cores do exist.

Aperiodicity is checked only for the 3-periodic three vortex system; this is done using the method described in section 4.2.1 in Figure 4.6.

An interesting point is that the ergodic regions share the reflection symmetries of the point vortex system. This may appear trivial, but it is not; the velocity field, tracer trajectories and f do not share this property. For instance, consider the 3-periodic three vortex system. The ergodic region shown in Figure 4.1 is symmetric in the y-axis, just like the associated vortex system, but if we look at Figure 3.8 we can clearly see that this does not hold for individual trajectories. This can be understood as follows. Consider what happens when we mirror the system using a reflection symmetry axis of the vortex system. This effectively reverses the rotation direction of the vortices without affecting their positions; tracer trajectories, Poincaré maps and ergodic regions are also mirrored. Now, if we reverse the time direction we restore the original system, since this also reverses the rotation direction of the vortices without affecting their positions. However, this does *not* affect the ergodic region, since it is invariant under time reversal by Theorem 4.2. This shows that the ergodic region is equal to its mirror image even though individual trajectories and Poincaré maps need not be.

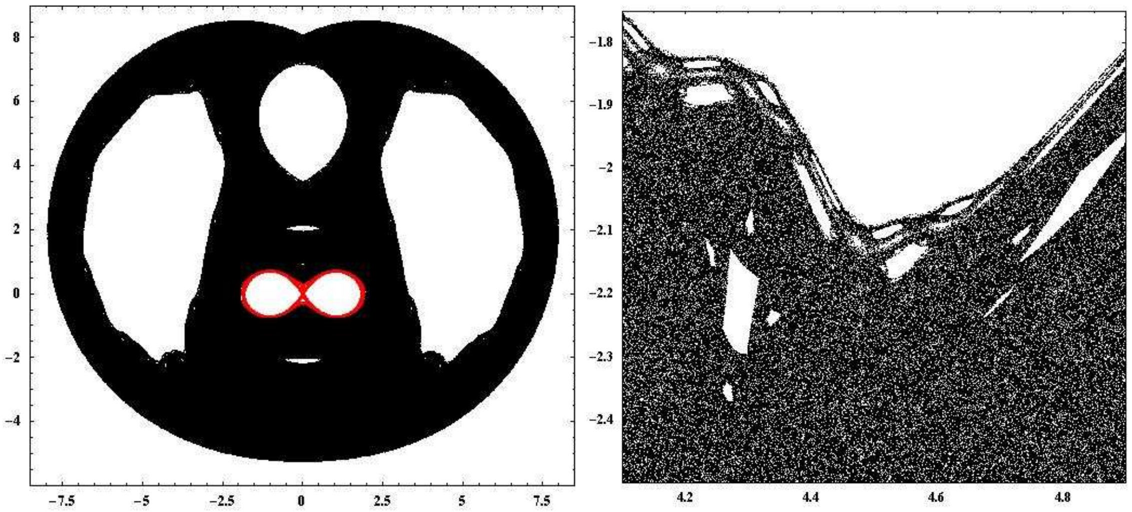

Figure 4.3: Poincaré map of $(0, 3)$ with $N = 10^6$ (black) and of $(0, 0)$ with $N = 10^5$ (red) for the 2-periodic three vortex system. Right: detail of a 1-periodic island boundary.

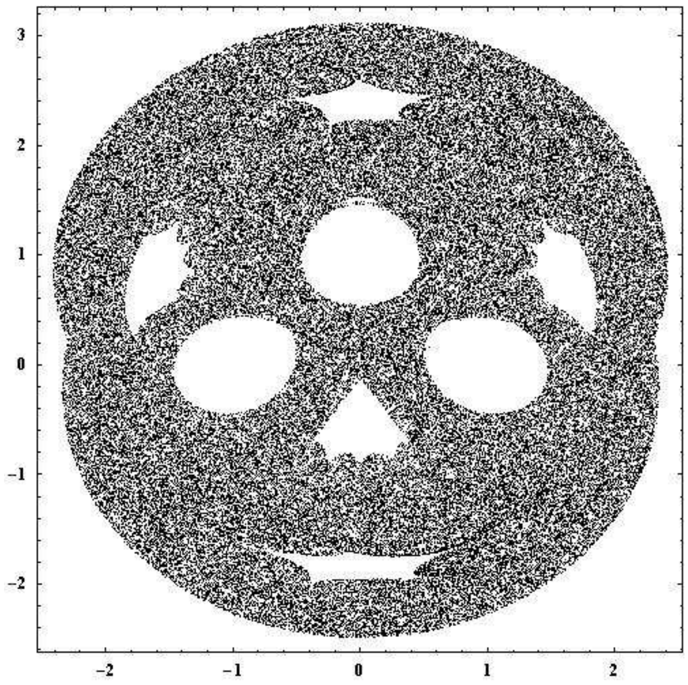

Figure 4.4: Tracer cloud spreading for the 3-periodic three vortex system, using $\mathbf{x} = (0, 0.1)$, $\delta = 0.2$, $T = 1800$ and $N = 10^5$.

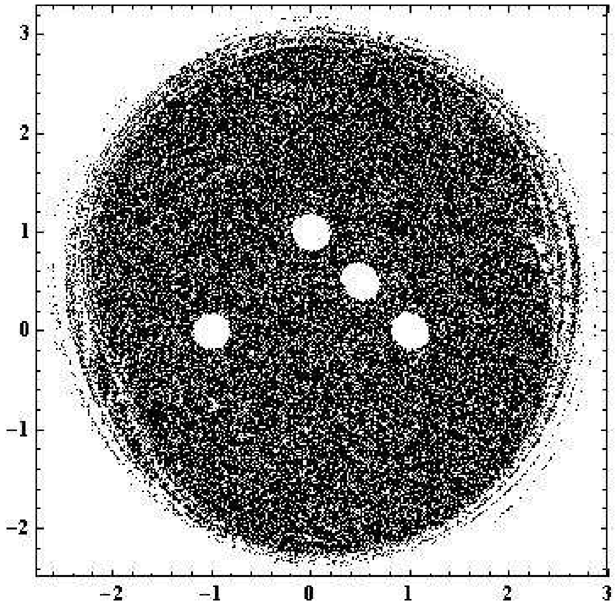

Figure 4.5: Tracer cloud spreading for the irregular four vortex system, using $\mathbf{x} = (0,0)$, $\delta = 0.3$, $T = 3000$ and $N = 10^5$.

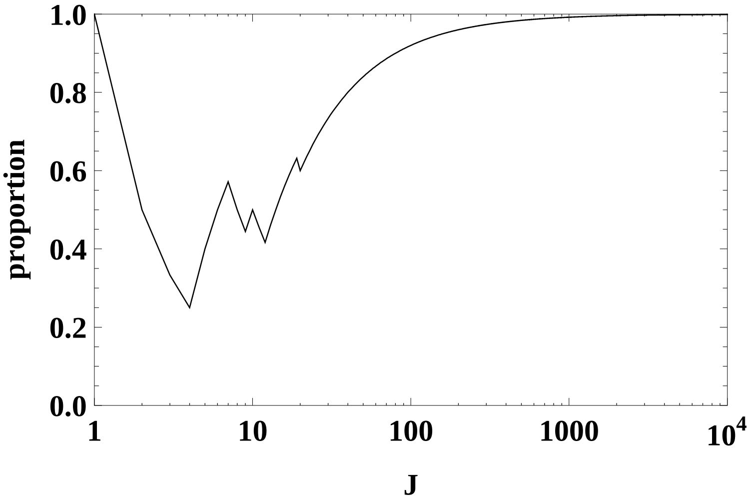

Figure 4.6: Plot of the proportion of j with $0 \leq j \leq J$ such that there is an n with $0 \leq n \leq 10^6 - j$ such that $f^n((0,0))$ and $f^{n+j}((0,0))$ are both in $(-0.12, 0.12) \times (-0.12, 0.12)$. That this proportion converges to 1 for increasing J shows that this property holds for asymptotically all j (see (4.5)). Of course the bound on n (10^6) would have to be increased to consider higher J than shown.

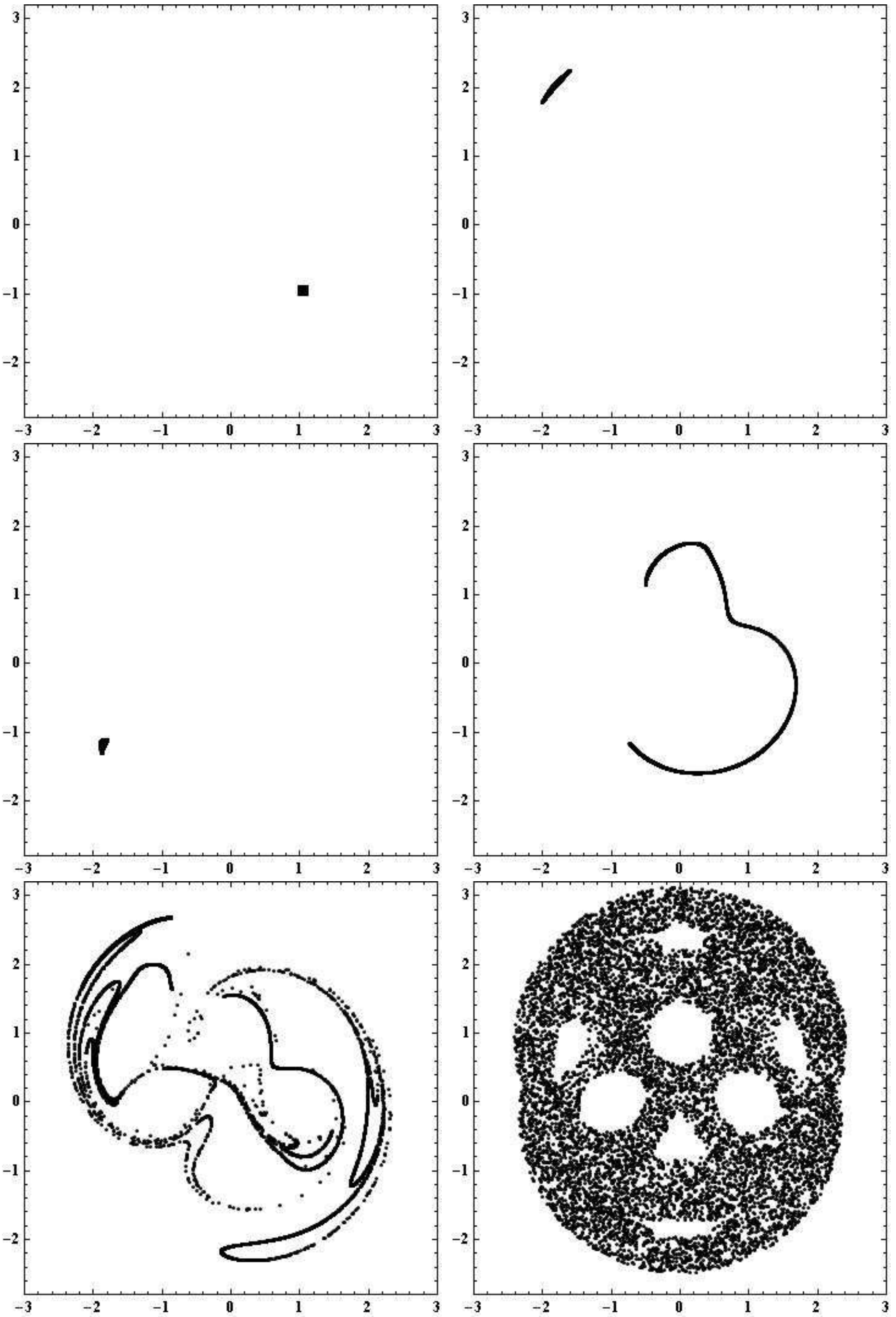

Figure 4.7: Mixing of a tracer cloud in the 3-periodic three vortex system. Initially, 10^4 tracers are uniformly spaced on a grid in $(1, 1.1) \times (-1, -0.9)$. The position of the tracers at various times is shown; from top left going horizontally $t = 0$, $t = 8T_p$, $t = 12T_p$, $t = 16T_p$, $t = 39T_p$ and $t = 99T_p$, with $T_p = 11.572$ the period of the system.

4.3.2 Mixing properties

As mentioned in section 4.2.2, there is another advantage to tracer cloud spreading over Poincaré maps. Since the method is directly connected to the mixing process, we can analyze this process by considering finite T, instead of the limit to infinity used to find the ergodic region. This allows us to find information on the timescales connected to and intermediate structures present during the mixing process. For a general overview, consider Figure 4.7, where the mixing of a tracer cloud in the 3-periodic three vortex system at various times is shown. The behavior seen here is typical: there is an initial period where stretching takes place but no folding (first 3 pictures), after which a partial mixing structure is formed (pictures 4 and 5). Finally, the cloud fills the entire ergodic region (picture 6) as predicted by Theorem 4.9.

The same process occurs for other starting locations of the tracer cloud. An important point is that the partial mixing structure is similar for different starting locations. For instance, the left part of Figure 4.9 shows the partial mixing structure for another starting location; note the similarities with Figure 4.7. What does vary is the time before the mixing process starts, i.e. the time spent in the initial regime where no folding takes place yet. This is shown in Figure 4.8, where the average squared distance between tracers in a cloud as a function of time is shown for various starting locations. This defines another structure in the ergodic region, which we will refer to as the mixing time distribution. It can be computed by covering the ergodic region with a grid and spacing N tracers uniformly in each grid square. We then compute the time required for the average squared distance between the tracers in one square to reach some threshold δ; this then is the mixing time for that grid square. This technique is applied to the 3-periodic three vortex system in the right part of Figure 4.9.

So we have now identified two structures within the ergodic region that are relevant to the mixing properties: the partial mixing structure and the mixing time distribution. These are shown for the other three nonstationary systems in Figures 4.10 through 4.12. The remainder of this report will focus on describing and computing these structures, including a rigorous definition of both as a measure defined on the ergodic region. Note that the two structures are somewhat analogous: the partial mixing structure shows regions that have undergone mixing, while the mixing time distribution shows regions that will undergo mixing. And indeed, we see similarities between both structures if we consider their mirror images using symmetry axes of the vortex systems (recall that time reversal is identical to mirroring the velocity field). This is shown for the 3-periodic three vortex system in Figure 4.13.

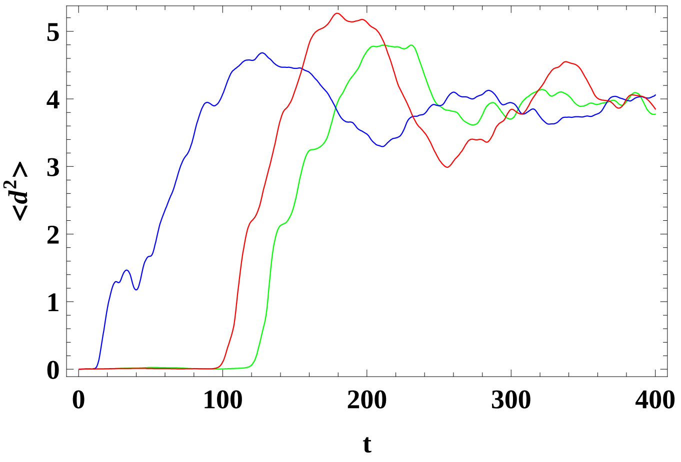

Figure 4.8: Average squared distance between tracers in a cloud of 100 tracers as a function of time, with the tracers initially uniformly spaced on a grid in $(1, 1.1) \times (-1, -0.9)$ (green, see also Figure 4.7), $(-1, -0.9) \times (-1, -0.9)$ (blue, see also Figure 4.9) and $(2, 2.1) \times (-1, -0.9)$ (red).

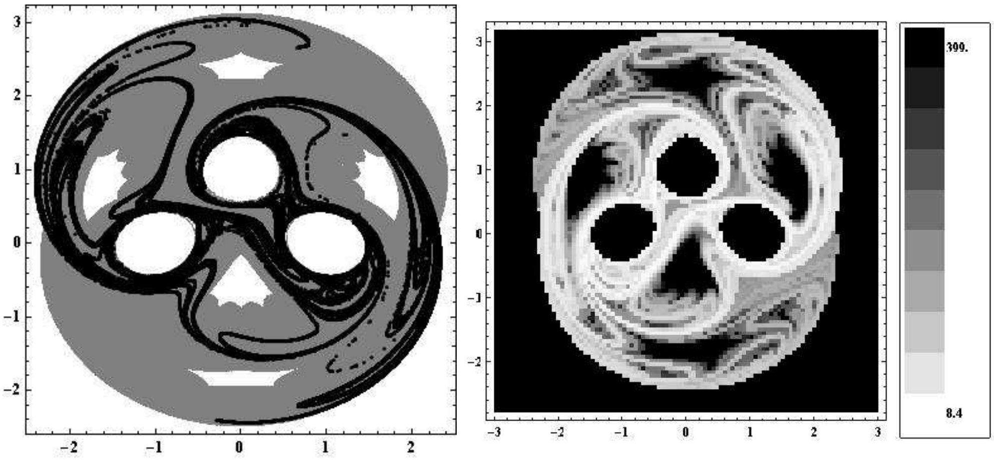

Figure 4.9: Mixing structures in the 3-periodic three vortex system. Left: tracer cloud of 250,000 tracers initially uniformly spaced on a grid in $(-1, -0.9) \times (-1, -0.9)$ at $t = 8T_p$ with $T_p = 11.572$ the period of the system, showing the partial mixing structure. The gray background is the full ergodic region as shown in Figure 4.1. Right: mixing time distribution on a 100x100 grid with $N = 25$ and $\delta = 1$. All regions which have not started mixing at $t = 400$ are colored black.

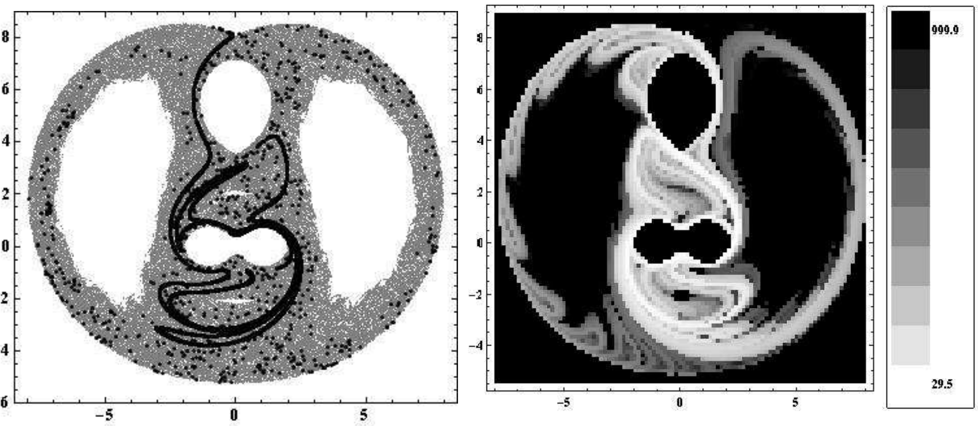

Figure 4.10: Mixing structures in the 2-periodic three vortex system. Left: tracer cloud spreading using $\mathbf{x} = (0, 0.5)$, $\delta = 0.01$, $T = 1600$ and $N = 10^5$, showing the partial mixing structure. The gray background is the full ergodic region as shown in Figure 4.3. Right: mixing time distribution on a 100x100 grid with $N = 25$ and $\delta = 1$. All regions which have not started mixing at $t = 1000$ are colored black.

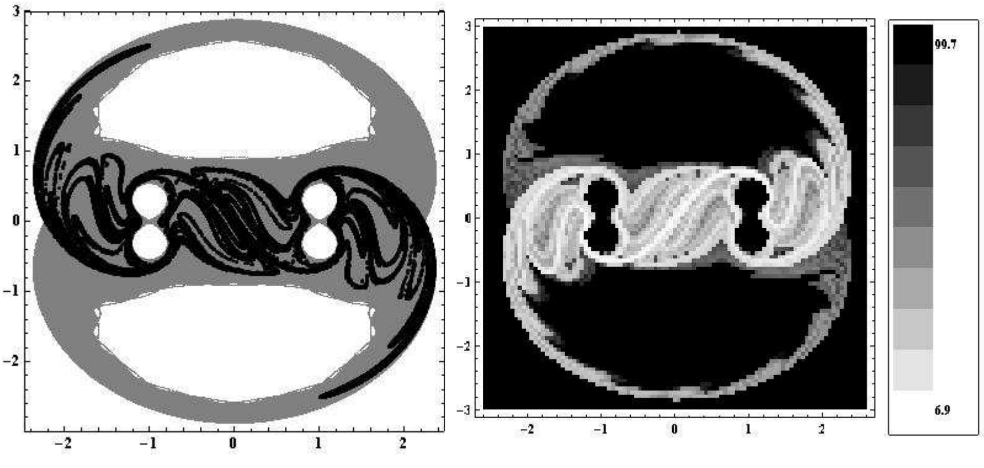

Figure 4.11: Mixing structures in the periodic four vortex system. Left: tracer cloud spreading using $\mathbf{x} = (0, 0)$, $\delta = 0.01$, $T = 30$ and $N = 10^5$, showing the partial mixing structure. The gray background is the full ergodic region as shown in Figure 4.2. Right: mixing time distribution on a 100x100 grid with $N = 25$ and $\delta = 1$. All regions which have not started mixing at $t = 100$ are colored black.

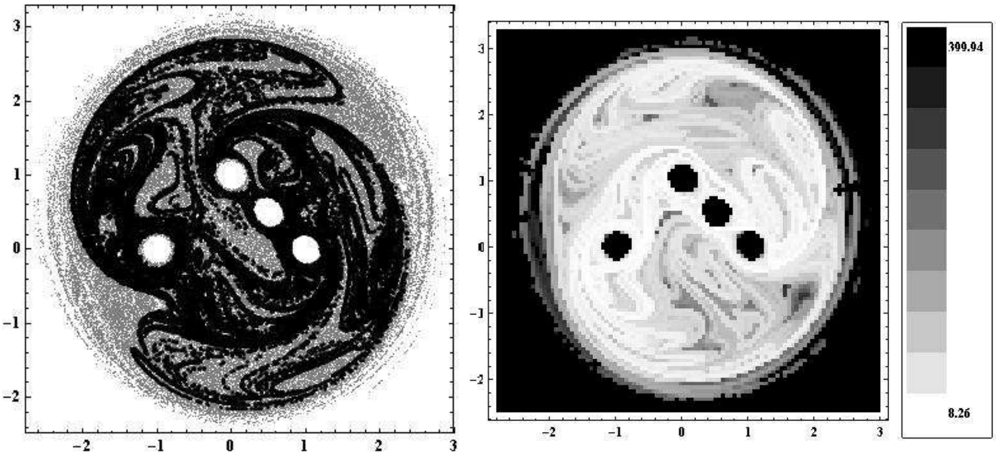

Figure 4.12: Mixing structures in the irregular four vortex system. Left: tracer cloud spreading using $\mathbf{x} = (-0.7, -0.1)$, $\delta = 0.01$, $T = 200$ and $N = 10^5$, showing the partial mixing structure. The gray background is the full ergodic region as shown in Figure 4.5. Right: mixing time distribution on a 100x100 grid with $N = 25$ and $\delta = 1$. All regions which have not started mixing at $t = 400$ are colored black.

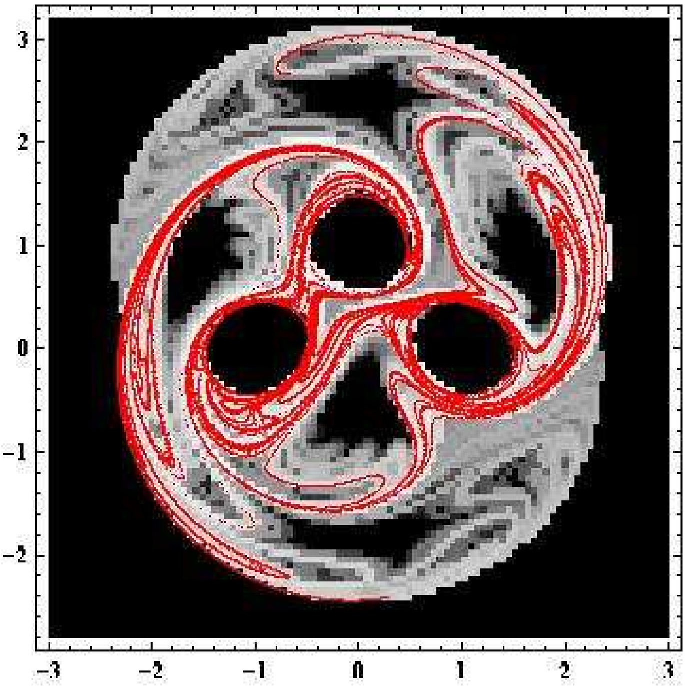

Figure 4.13: Mixing time distribution (background) and mirrored partial mixing structure (red) as shown in Figure 4.9 for the 3-periodic three vortex system.

5 Lyapunov Exponents

In this chapter we will explore the concept of Lyapunov exponents and their application in describing the mixing process in point vortex systems. Lyapunov exponents describe the stretching undergone by infinitesimal fluid elements. This stretching, together with the folding of material lines, drives the mixing process [61]. In the asymptotic case, we can use these exponents to define a timescale for the mixing. For finite times they can be used to describe both the partial mixing structure and the mixing time distribution found in section 4.3.2.

We will first give a detailed description of the general theory of Lyapunov exponents (section 5.1.1), followed by an analysis of their role in describing mixing in point vortex systems (section 5.1.2). Here we will also introduce the the topological entropy, another timescale related to the mixing process, and the w-measure, a rigorously defined probability measure which is directly related to the partial mixing structure and the mixing time distribution. Next the numerical methods used to find the Lyapunov exponents, topological entropy and w-measure will be introduced (section 5.2), followed by their actual computation (section 5.3).

5.1 Theory

5.1.1 Definition and basic properties

Lyapunov theory is based on the stretching of small perturbations of trajectories, so we start with some words on the Jacobian of the flow map $\Phi(t, T)$. Recall that this is the map that takes a tracer from its position at time t to its position at time $t + T$ (see section 3.1). Its Jacobian is the 2-tensor from $\mathbb{R}^2$ to $\mathbb{R}^2$ defined by

$$\nabla\Phi(\mathbf{x}, t, T)\delta\mathbf{x} = \lim_{\epsilon\to 0} \frac{\Phi(t, T)(\mathbf{x} + \epsilon\delta\mathbf{x}) - \Phi(t, T)(\mathbf{x})}{\epsilon}, \tag{5.1}$$

for all $\delta\mathbf{x} \in \mathbb{R}^2$. If we have a trajectory $\mathbf{x}(t)$ and a slightly perturbed one $\mathbf{x}'(t) = \mathbf{x}(t) + \delta\mathbf{x}(t)$, then the Jacobian describes the evolution of this perturbation:

$$\delta\mathbf{x}(t + T) = \nabla\Phi(\mathbf{x}(t), t, T)\delta\mathbf{x}(t) + \mathcal{O}(||\delta\mathbf{x}(t)||^2). \tag{5.2}$$

We can use this equation to derive a useful identity. We have directly from the equation above

$$\delta\mathbf{x}(t+T+S) = \nabla\Phi(\mathbf{x}(t),t,T+S)\delta\mathbf{x}(t) + \mathcal{O}(||\delta\mathbf{x}(t)||^2),$$

but also applying (5.2) twice

$$\begin{aligned}\delta\mathbf{x}(t+T+S) = \nabla\Phi(\mathbf{x}(t+T),t+T,S)\delta\mathbf{x}(t+T) + \mathcal{O}(||\delta\mathbf{x}(t)||^2) = \\ \nabla\Phi(\mathbf{x}(t+T),t+T,S)\nabla\Phi(\mathbf{x}(t),t,T+S)\delta\mathbf{x}(t) + \mathcal{O}(||\delta\mathbf{x}(t)||^2) = \\ \nabla\Phi(\Phi(t,T)\mathbf{x}(t),t+T,S)\nabla\Phi(\mathbf{x}(t),t,T+S)\delta\mathbf{x}(t) + \mathcal{O}(||\delta\mathbf{x}(t)||^2),\end{aligned}$$

where the last equality follows directly from the definition of the flow map (3.7). This gives us:

$$\nabla\Phi(\mathbf{x},t,T+S) = \nabla\Phi(\phi(t,T)(\mathbf{x}),t+T,S)\nabla\Phi(\mathbf{x},t,T). \tag{5.3}$$

This could also have been derived directly from (3.8) using the chain rule. It shows that if we compute the Jacobian of a trajectory from t to $t+T+S$, we can take the product of the Jacobians of the one from t to $t+T$ and the one from $t+T$ to $t+T+S$. Note that because the flow map preserves area we have

$$\det(\nabla\Phi(\mathbf{x},t,T)) = 1. \tag{5.4}$$

This also shows that $\nabla\Phi(\mathbf{x},t,T)$ is invertible. In fact, using (5.3) with $S=-T$:

$$\nabla\Phi(\mathbf{x},t,T)^{-1} = \nabla\Phi(\Phi(t,T)(\mathbf{x}),t+T,-T). \tag{5.5}$$

Where no confusion can occur the arguments $\mathbf{x}$ and t will be omitted. The final argument T however will never be omitted, so $\nabla\Phi(T)$ is short for $\nabla\Phi(\mathbf{x},t,T)$.

So to analyze the stretching of small perturbations we should consider the stretching properties of the Jacobian $\nabla\Phi$. To this end, consider

$$\Lambda(\mathbf{x},t,T) := (\nabla\Phi(\mathbf{x},t,T)^*\nabla\Phi(\mathbf{x},t,T))^{1/2|T|}, \tag{5.6}$$

where * denotes matrix conjugation (the power of $1/2|T|$ is taken for scaling properties which will be needed later). Note that due to (5.4)

$$\det(\Lambda(\mathbf{x},t,T)) = (\det(\nabla\Phi(\mathbf{x},t,T)^*)\det(\nabla\Phi(\mathbf{x},t,T)))^{1/2|T|} = 1. \tag{5.7}$$

$\Lambda(\mathbf{x},t,T)$ admits an orthonormal basis of eigenvectors [64] with positive eigenvalues. Denote these by $\mathbf{e}_1(\mathbf{x},t,T)$ and $\mathbf{e}_2(\mathbf{x},t,T)$ with corresponding eigenvalues $\eta_1(\mathbf{x},t,T)$ and $\eta_2(\mathbf{x},t,T)$, ordered such that $\eta_1(T) \geq \eta_2(T)$. [1] Using (5.7) we have

$$\eta_2(T) = \frac{1}{\eta_1(T)} \leq 1 \leq \eta_1(T). \tag{5.8}$$

---

[1] $\mathbf{e}_1(T)$ and $\mathbf{e}_2(T)$ are not completely defined by this, since their orientation is not fixed. We should actually use the eigenspaces of $\Lambda(T)$, but this would needlessly complicate the notation.

When equality holds we have $\eta_1(T) = \eta_2(T) = 1$, so $\Lambda(T)$ is the identity matrix. In this case $\mathbf{e}_1(T)$ and $\mathbf{e}_2(T)$ can be taken as any orthonormal basis of $\mathbb{R}^2$. These directions are called the (finite time) Lyapunov directions. The relevance to stretching is shown Theorem 5.3 below.

Lemma 5.1. *For all* $\mathbf{x} \in \mathbb{R}^2$ *and* $t, T \in \mathbb{R}$, *we have*

$$\nabla\Phi(\mathbf{x}, t, T)\mathbf{e}_1(\mathbf{x}, t, T) = \eta_1(\mathbf{x}, t, T)^{|T|}\mathbf{e}_2(\Phi(t, T)(\mathbf{x}), t + T, -T)$$
$$\nabla\Phi(\mathbf{x}, t, T)\mathbf{e}_2(\mathbf{x}, t, T) = \eta_1(\mathbf{x}, t, T)^{-|T|}\mathbf{e}_1(\Phi(t, T)(\mathbf{x}), t + T, -T)$$

up to orientation.

Proof. We will show only the first equality; the proof of the second is similar using (5.8). For identity up to orientation we need to show that the scale and direction of both sides are equal. First the scale:

$$\begin{aligned}&||\nabla\Phi(\mathbf{x}, t, T)\mathbf{e}_1(\mathbf{x}, t, T)||^2 = (\nabla\Phi(\mathbf{x}, t, T)\mathbf{e}_1(\mathbf{x}, t, T), \nabla\Phi(\mathbf{x}, t, T)\mathbf{e}_1(\mathbf{x}, t, T)) =\\ &(\mathbf{e}_1(\mathbf{x}, t, T), \nabla\Phi(\mathbf{x}, t, T)^*\nabla\Phi(\mathbf{x}, t, T)\mathbf{e}_1(\mathbf{x}, t, T)) = (\mathbf{e}_1(\mathbf{x}, t, T), \Lambda(\mathbf{x}, t, T)^{2|T|}\mathbf{e}_1(\mathbf{x}, t, T)) =\\ &(\mathbf{e}_1(\mathbf{x}, t, T), \eta_1(\mathbf{x}, t, T)^{2|T|}\mathbf{e}_1(\mathbf{x}, t, T)) = \eta_1(\mathbf{x}, t, T)^{2|T|} =\\ &||\eta_1(\mathbf{x}, t, T)^{|T|}\mathbf{e}_2(\Phi(t, T)(\mathbf{x}), t + T, -T)||^2.\end{aligned}$$

To show equality of direction, note that by definition the right side is an eigenvector of $\Lambda(\Phi(t, T)(\mathbf{x}), t + T, -T)$ with eigenvalue at most 1, so it is an eigenvector of $\Lambda(\Phi(t, T)(\mathbf{x}), t + T, -T)^{-2|T|}$ with eigenvalue at least 1. We will show this same property for the left side. First, note that

$$\begin{aligned}&\Lambda(\Phi(t, T)(\mathbf{x}), t + T, -T)^{-2|T|} =\\ &(\nabla\Phi(\Phi(t, T)(\mathbf{x}), t + T, -T)^*\nabla\Phi(\Phi(t, T)(\mathbf{x}), t + T, -T))^{-1} \overset{(5.5)}{=}\\ &\nabla\Phi(\mathbf{x}, t, T)\nabla\Phi(\mathbf{x}, t, T)^*.\end{aligned}$$

We now have

$$\begin{aligned}&\Lambda(\Phi(t, T)(\mathbf{x}), t + T, -T)^{-2|T|}\nabla\Phi(\mathbf{x}, t, T)\mathbf{e}_1(\mathbf{x}, t, T) =\\ &\nabla\Phi(\mathbf{x}, t, T)\nabla\Phi(\mathbf{x}, t, T)^*\nabla\Phi(\mathbf{x}, t, T)\mathbf{e}_1(\mathbf{x}, t, T) =\\ &\nabla\Phi(\mathbf{x}, t, T)\Lambda(\mathbf{x}, t, T)^{2|T|}\mathbf{e}_1(\mathbf{x}, t, T) = \eta_1(\mathbf{x}, t, T)^{2|T|}\nabla\Phi(\mathbf{x}, t, T)\mathbf{e}_1(\mathbf{x}, t, T).\end{aligned}$$

Since $\eta_1(\mathbf{x}, t, T)^{2|T|} \geq 1$, this completes the proof. □

We will assume from here on that the orientation of the Lyapunov directions is chosen such that the identities in the above Lemma hold directly, i.e. not just up to orientation.

Lemma 5.2. *For all* $\mathbf{x}, \delta\mathbf{x} \in \mathbb{R}^2$ *and* $t, T \in \mathbb{R}$, *we have*

$$\begin{aligned}\nabla\Phi(\mathbf{x},t,T)\delta\mathbf{x} =&(\mathbf{e}_1(\mathbf{x},t,T),\delta\mathbf{x})\eta_1(\mathbf{x},t,T)^{|T|}\mathbf{e}_2(\Phi(t,T)(\mathbf{x}),t+T,-T)+\\ &(\mathbf{e}_2(\mathbf{x},t,T),\delta\mathbf{x})\eta_1(\mathbf{x},t,T)^{-|T|}\mathbf{e}_1(\Phi(t,T)(\mathbf{x}),t+T,-T)\end{aligned}$$

Proof. This follows directly from Lemma 5.1 and orthonormality of $\mathbf{e}_1(\mathbf{x},t,T)$ and $\mathbf{e}_2(\mathbf{x},t,T)$. □

Theorem 5.3. *The identity*

$$||\nabla\Phi(\mathbf{x},t,T)\delta\mathbf{x}|| = \sqrt{(\mathbf{e}_1(\mathbf{x},t,T),\delta\mathbf{x})^2\eta_1(\mathbf{x},t,T)^{2|T|} + (\mathbf{e}_2(\mathbf{x},t,T),\delta\mathbf{x})^2\eta_1(\mathbf{x},t,T)^{-2|T|}}$$

holds for all $\mathbf{x}, \delta\mathbf{x} \in \mathbb{R}^2$ *and* $t, T \in \mathbb{R}$.

Proof. This follows directly from Lemma 5.2 and orthonormality of $\mathbf{e}_1(\Phi(t,T)(\mathbf{x}),t+T,-T)$ and $\mathbf{e}_2(\Phi(t,T)(\mathbf{x}),t+T,-T)$. □

We are now ready to define the finite time Lyapunov exponents $\lambda_1(\mathbf{x},t,T)$ and $\lambda_2(\mathbf{x},t,T)$:

$$\lambda_{1,2}(\mathbf{x},t,T) := \log\eta_{1,2}(\mathbf{x},t,T). \tag{5.9}$$

Although this definition will be used in the remainder of this report, there are non-equivalent alternatives, such as the one found in [15]. Note that by (5.8) we have

$$\lambda_2(T) = -\lambda_1(T) \tag{5.10}$$

and

$$\lambda_1(T) \geq 0. \tag{5.11}$$

So it is enough to consider only $\lambda_1(T)$, which is called the finite time maximum Lyapunov exponents (FTMLE).

We now shift our attention to asymptotic Lyapunov exponents. For this we consider the limit for T to ±∞. This is justified by Oseledec's theorem [22, 59].

Theorem 5.4 (Oseledec). *The limits*

$$\begin{aligned}\Lambda_+(\mathbf{x},t) &:= \lim_{T\to\infty}\Lambda(\mathbf{x},t,T)\\ \Lambda_-(\mathbf{x},t) &:= \lim_{T\to-\infty}\Lambda(\mathbf{x},t,T)\end{aligned} \tag{5.12}$$

exist for all $\mathbf{x} \in \mathbb{R}^2$ *and* $t \in \mathbb{R}$.

From here on we will consider only the positive limit; all results can be easily extended to the negative limit by replacing T by $-T$. We define $\mathbf{e}_{1+}$ and $\mathbf{e}_{2+}$ to be the eigenvectors associated with Λ_+ with respective eigenvalues η_{1+} and η_{2+} ($\eta_{1+} \geq \eta_{2+}$), and define similarly to (5.9)

$$\lambda_{1,2;+}(\mathbf{x}, t) := \log \eta_{1,2;+}(\mathbf{x}, t). \tag{5.13}$$

From their definitions it is clear that $\mathbf{e}_{1,2;+}$, $\eta_{1,2;+}$ and $\lambda_{1,2;+}$ are equal to the limits for $T \to \infty$ of their finite time counterparts introduced above. In particular,

$$\lambda_{1+}(\mathbf{x}, t) = \lim_{T\to\infty} \lambda_{1+}(\mathbf{x}, t, T). \tag{5.14}$$

λ_{1+} is the (asymptotic) maximum Lyapunov exponent (MLE). It is typically defined more simply as the exponential growth rate of small perturbations [15]. The following theorem shows the equivalence of this definition to ours for almost all perturbations.

Theorem 5.5. *The MLE satisfies*

$$\begin{aligned}\lambda_{1+}(\mathbf{x}, t) = \lim_{T\to\infty} \frac{1}{|T|} \log ||\nabla\Phi(\mathbf{x}, t, T)\delta\mathbf{x}|| = \\ \lim_{T\to\infty}\lim_{\epsilon\to 0} \frac{1}{|T|} \log \frac{||\Phi(t, T)(\mathbf{x} + \epsilon\delta\mathbf{x}) - \Phi(t, T)(\mathbf{x})||}{\epsilon||\delta\mathbf{x}||}\end{aligned}$$

for all $\delta\mathbf{x} \in \mathbb{R}^2$ *satisfying* $(\delta\mathbf{x}, \mathbf{e}_{1+}(\mathbf{x}, t)) \neq 0$ *and for all* $\mathbf{x} \in \mathbb{R}^2$ *and* $t \in \mathbb{R}$.

Proof. Fix $\mathbf{x}$, t and $\delta\mathbf{x}$ as above. Define $\alpha_1(T)$ and $\alpha_2(T)$ by

$$\begin{aligned}\alpha_1(T) &:= (\delta\mathbf{x}, \mathbf{e}_1(T))\\ \alpha_2(T) &:= (\delta\mathbf{x}, \mathbf{e}_2(T))\end{aligned}$$

and their corresponding limits for $T \to \infty$ α_1 and α_2. By assumption we have

$$\alpha_1 \neq 0. \tag{5.15}$$

We now have

$$\begin{aligned}&\lim_{T\to\infty}\lim_{\epsilon\to 0} \frac{1}{|T|} \log \frac{||\Phi(t, T)(\mathbf{x} + \epsilon\delta\mathbf{x}) - \Phi(t, T)(\mathbf{x})||}{\epsilon||\delta\mathbf{x}||} = \\ &\lim_{T\to\infty}\lim_{\epsilon\to 0} \left(\frac{1}{|T|} \log \frac{||\Phi(t, T)(\mathbf{x} + \epsilon\delta\mathbf{x}) - \Phi(t, T)(\mathbf{x})||}{\epsilon} - \frac{1}{|T|}\log||\delta\mathbf{x}||\right) \overset{(5.1)}{=} \\ &\lim_{T\to\infty} \frac{1}{|T|} \log||\nabla\Phi(T)\delta\mathbf{x}|| \overset{\text{T5.3}}{=} \lim_{T\to\infty} \frac{1}{2|T|} \log(\alpha_1(T)^2\eta_1(T)^{2|T|} + \alpha_2(T)^2\eta_2(T)^{2|T|}) \overset{(5.8)}{=} \\ &\lim_{T\to\infty} \frac{1}{2|T|} \log(\eta_1(T)^{2|T|}(\alpha_1(T)^2 + \alpha_2(T)^2\eta_1(T)^{-4|T|}) \overset{(5.15)}{=} \lim_{T\to\infty} \log\eta_1(T) \overset{(5.13)}{=} \lambda_{1+},\end{aligned}$$

completing the proof. □

An important property of the MLE is that it is constant along trajectories, which means it makes sense to speak of the Lyapunov exponent of a trajectory without specifying any specific time. This fact is established in the following theorem.

Theorem 5.6. *The MLE is constant along trajectories:*

$$\lambda_{1+}(\Phi(t,S)(\mathbf{x}),t+S)=\lambda_{1+}(\mathbf{x},t),$$

for all $\mathbf{x}\in\mathbb{R}^2$ *and* $t,S\in\mathbb{R}$.

Proof. Take any $\delta\mathbf{x}\in\mathbb{R}^2$ such that

$$\begin{gathered}(\delta\mathbf{x},\mathbf{e}_{1+}(\mathbf{x},t))\neq 0;\\ (\delta\mathbf{x},\nabla\Phi(\mathbf{x},t,S)^*\mathbf{e}_{1+}(\Phi(t,S)(\mathbf{x}),t+S))\neq 0.\end{gathered}\tag{5.16}$$

We then have

$$(\nabla\Phi(\mathbf{x},t,S)\delta\mathbf{x},\mathbf{e}_{1+}(\Phi(t,S)(\mathbf{x}),t+S))\neq 0,$$

and so by Theorem 5.5:

$$\lambda_{1+}(\Phi(t,S)(\mathbf{x}),t+S)=\lim_{T\to\infty}\frac{1}{|T|}\log||\nabla\Phi(\Phi(t,s)(\mathbf{x}),t+S,T)\nabla\Phi(\mathbf{x},t,S)\delta\mathbf{x}||\overset{(5.3)}{=}$$
$$\lim_{T\to\infty}\frac{1}{|T|}\log||\nabla\Phi(\mathbf{x},t,S+T)\delta\mathbf{x}||=\lim_{T\to\infty}\frac{|T|}{|T-S|}\frac{1}{|T|}\log||\nabla\Phi(\mathbf{x},t,T)\delta\mathbf{x}||\overset{\text{T5.5}}{=}$$
$$\lambda_{1+}(\mathbf{x},t),$$

completing the proof. □

The condition $(\delta\mathbf{x},\mathbf{e}_{1+}(\mathbf{x},t))\neq 0$ in Theorem 5.5 is important. If it does not hold, then $\delta\mathbf{x}$ is perpendicular to $\mathbf{e}_{1+}(\mathbf{x},t)$ and hence oriented along $\mathbf{e}_{2+}(\mathbf{x},t)$, the eigenvector associated with the negative Lyapunov exponent. By the following Theorem it will remain oriented along this direction and hence the perturbation does not grow.

Theorem 5.7. *If the MLE is strictly positive, then* $\mathbf{e}_{2+}$ *is conserved along trajectories, i.e.:*

$$\mathbf{e}_{2+}(\Phi(t,S)\mathbf{x},t+S)=\nabla\Phi(\mathbf{x},t,S)\mathbf{e}_{2+}(\mathbf{x},t)$$

up to scaling, for all $\mathbf{x}\in\mathbb{R}^2$ *and* $t,S\in\mathbb{R}$.

Proof. Since the MLE is strictly positive, we have

$$\eta_{1+}(\mathbf{x},t)>1.\tag{5.17}$$

Consider the finite time stable direction $\mathbf{e}_2(\Phi(t,S)\mathbf{x},t+S,T)$. Since it is the eigenvector with minimal eigenvalue of $\Lambda(\Phi(t,S)\mathbf{x},t+S,T)$, it is the solution of the minimization of [64]

$$F_T(\mathbf{z}_T)=||\nabla\Phi(\Phi(t,S)(\mathbf{x}),t+S,T)\mathbf{z}_T||^2\tag{5.18}$$

for $\mathbf{z}_T \in \mathbb{R}^2$ subject to the constraint

$$||\mathbf{z}_T||^2 = 1. \tag{5.19}$$

Now define

$$\mathbf{a}_T := \frac{\nabla\Phi(\mathbf{x}, t, S)\mathbf{e}_2(\mathbf{x}, t, S+T)}{||\nabla\Phi(\mathbf{x}, t, S)\mathbf{e}_2(\mathbf{x}, t, S+T)||}. \tag{5.20}$$

Since

$$||\nabla\Phi(\mathbf{x}, t, S)\mathbf{e}_2(\mathbf{x}, t, S+T)|| \geq ||\nabla\Phi(\mathbf{x}, t, S)\mathbf{e}_2(\mathbf{x}, t, S)|| \overset{\text{L5.1}}{=} \eta_2(\mathbf{x}, t, S)^S$$

and

$$\begin{aligned}&||\nabla\Phi(\Phi(t, S)(\mathbf{x}), t+S, T)\nabla\Phi(\mathbf{x}, t, S)\mathbf{e}_2(\mathbf{x}, t, S+T)|| \overset{(5.3)}{=}\\ &||\nabla\Phi(\mathbf{x}, t, S+T)\mathbf{e}_2(\mathbf{x}, t, S+T)|| \overset{\text{L5.1}}{=} \eta_2(\mathbf{x}, t, S+T)^{(S+T)},\end{aligned}$$

we have

$$\begin{aligned}F_T(\mathbf{a}_T) =& \frac{||\nabla\Phi(\Phi(t, S)(\mathbf{x}), t+S, T)\nabla\Phi(\mathbf{x}, t, S)\mathbf{e}_2(\mathbf{x}, t, S+T)||^2}{||\nabla\Phi(\mathbf{x}, t, S)\mathbf{e}_2(\mathbf{x}, t, S+T)||^2} \leq\\ &\frac{\eta_2(\mathbf{x}, t, S+T)^{2|S+T|}}{\eta_2(\mathbf{x}, t, S)^{2|S|}}.\end{aligned} \tag{5.21}$$

Since $\mathbf{a}_T$ satisfies (5.19), we have:

$$\begin{aligned}&\lim_{T\to\infty} F_T(\mathbf{e}_2(\Phi(t, S)\mathbf{x}, t+S, T)) \leq \lim_{T\to\infty} F_T(\mathbf{a}_T) \leq \lim_{T\to\infty} \frac{\eta_2(\mathbf{x}, t, S+T)^{2|S+T|}}{\eta_2(\mathbf{x}, t, S)^{2|S|}} \overset{(5.8)}{=}\\ &\lim_{T\to\infty} \frac{\eta_1(\mathbf{x}, t, S)^{2|S|}}{\eta_1(\mathbf{x}, t, S+T)^{2|S+T|}} \overset{(5.17)}{=} 0.\end{aligned} \tag{5.22}$$

Since $\mathbf{e}_1(\mathbf{x}, t, S+T)$ and $\mathbf{e}_2(\mathbf{x}, t, S+T)$ are orthogonal and hence linearly independent and $\nabla\Phi(\mathbf{x}, t, S)$ is invertible, we can write $\mathbf{e}_2(\Phi(t, S)\mathbf{x}, t+S, T)$ as

$$\begin{aligned}&\mathbf{e}_2(\Phi(t, S)\mathbf{x}, t+S, T) =\\ &\alpha_1(T)\nabla\Phi(\mathbf{x}, t, S)\mathbf{e}_1(\mathbf{x}, t, S+T) + \alpha_2(T)\nabla\Phi(\mathbf{x}, t, S)\mathbf{e}_2(\mathbf{x}, t, S+T)\end{aligned} \tag{5.23}$$

for unique $\alpha_1(T), \alpha_2(T) \in \mathbb{R}$. We decompose the asymptotic stable direction in the same way:

$$\mathbf{e}_{2+}(\Phi(t, S)\mathbf{x}, t+S) = \alpha_1\nabla\Phi(\mathbf{x}, t, S)\mathbf{e}_{1+}(\mathbf{x}, t) + \alpha_2\nabla\Phi(\mathbf{x}, t, S)\mathbf{e}_{2+}(\mathbf{x}, t). \tag{5.24}$$

Note that

$$\begin{aligned} \lim_{T\to\infty} \alpha_1(T) &= \alpha_1, \\ \lim_{T\to\infty} \alpha_2(T) &= \alpha_2. \end{aligned} \tag{5.25}$$

If we can show that $\alpha_1 = 0$, then the statement to be proven follows from (5.24). Assume $\alpha_1 \neq 0$; then by (5.25) we have

$$\lim_{T\to\infty} \alpha_1(T)^2 > 0. \tag{5.26}$$

Now:

$$\begin{aligned} &\lim_{T\to\infty} F_T(\mathbf{e}_2(\Phi(t,S)\mathbf{x}, t+S, T)) \overset{(5.18),(5.23)}{=} \\ &\lim_{T\to\infty} ||\nabla\Phi(\Phi(t,S)(\mathbf{x}), t+S, T) \sum_{i=1}^{2} \alpha_i(T) \nabla\Phi(x,t,S)\mathbf{e}_i(\mathbf{x}, t, S+T)||^2 \overset{(5.3)}{=} \\ &\lim_{T\to\infty} || \sum_{i=1}^{2} \alpha_i(T) \nabla\Phi(x,t,S+T)\mathbf{e}_2(\mathbf{x},t,S+T)||^2 \overset{\text{T5.3}}{=} \\ &\lim_{T\to\infty} \sum_{i=1}^{2} \alpha_i^2(T)\eta_i(\mathbf{x},t,S+T)^{2|S+T|} \overset{(5.8)}{=} \\ &\lim_{T\to\infty} \alpha_1(T)^2\eta_1(\mathbf{x},t,S+T)^{2|S+T|} + \alpha_2(T)^2\eta_1(\mathbf{x},t,S+T)^{-2|S+T|} = \infty, \end{aligned}$$

contradicting (5.22). So $\alpha_1 = 0$, completing the proof. □

By time reversal, $\mathbf{e}_{2-}$ is also conserved along a trajectory. It is important to note that the property does not extend to $\mathbf{e}_{1+}$ and $\mathbf{e}_{1-}$; they are in general *not* conserved by the flow. The directions $\mathbf{e}_{2+}$ and $\mathbf{e}_{2-}$ are referred to as the stable Lyapunov direction and the unstable Lyapunov direction respectively, following [29]. Not only are they conserved by the flow, but any perturbation not initially aligned along either of them will converge to one.

Theorem 5.8. *If the MLE is strictly positive and the stable and unstable Lyapunov directions do not coincide, then perturbations not aligned along the stable Lyapunov direction converge to the unstable one, i.e. when* $\lambda_{1+}(\mathbf{x},t) > 0$ *and* $|(\mathbf{e}_{2-}, \mathbf{e}_{2+})| < 1$, *we have*

$$\lim_{T\to\infty}\lim_{\epsilon\to 0}\left(\frac{\Phi(t,T)(\mathbf{x}+\epsilon\delta\mathbf{x}) - \Phi(t,T)(\mathbf{x})}{||\Phi(t,T)(\mathbf{x}+\epsilon\delta\mathbf{x}) - \Phi(t,T)(\mathbf{x})||}, \mathbf{e}_{2-}(\Phi(t,T)\mathbf{x}, t+T)\right) = \pm 1$$

for all $\delta\mathbf{x} \in \mathbb{R}^2$ *satisfying* $(\delta\mathbf{x}, \mathbf{e}_{1+}(\mathbf{x},t)) \neq 0$ *and for all* $\mathbf{x} \in \mathbb{R}^2$ *and* $t \in \mathbb{R}$.

Proof. As in earlier proofs, define $\alpha_i(T)$ and $\beta_i(T)$ by

$$\begin{aligned} \alpha_i(T) &:= (\delta\mathbf{x}, \mathbf{e}_i(\mathbf{x}, t, T)) \\ \beta_i(T) &:= (\mathbf{e}_{2-}(\mathbf{x}, t), \mathbf{e}_i(\mathbf{x}, t, T)) \end{aligned} \tag{5.27}$$

with their corresponding limits for $T \to \infty$ α_i and β_i for $i = 1, 2$. Note that by the assumptions in the statement of the theorem we have

$$\begin{aligned} \alpha_1 &\neq 0; \\ \beta_1 &\neq 0. \end{aligned} \tag{5.28}$$

For convenience, define r to be the limit in the statement of the theorem; we then need to show that $r = \pm 1$. We have

$$\begin{aligned} r = \lim_{T\to\infty} \lim_{\epsilon\to 0} &\left(\frac{\Phi(t,T)(\mathbf{x}+\epsilon\delta\mathbf{x}) - \Phi(t,T)(\mathbf{x})}{||\Phi(t,T)(\mathbf{x}+\epsilon\delta\mathbf{x}) - \Phi(t,T)(\mathbf{x})||}, \mathbf{e}_{2-}(\Phi(t,T)\mathbf{x}, t+T) \right) \overset{(5.1)}{=} \\ \lim_{T\to\infty} \lim_{\epsilon\to 0} &\left(\frac{\nabla\Phi(\mathbf{x},t,T)\delta\mathbf{x}}{||\nabla\Phi(\mathbf{x},t,T)\delta\mathbf{x}||}, \mathbf{e}_{2-}(\Phi(t,T)\mathbf{x}, t+T) \right) \overset{\text{T5.7}}{=} \\ \lim_{T\to\infty} \lim_{\epsilon\to 0} &\left(\frac{\nabla\Phi(\mathbf{x},t,T)\delta\mathbf{x}}{||\nabla\Phi(\mathbf{x},t,T)\delta\mathbf{x}||}, \frac{\nabla\Phi(\mathbf{x},t,T)\mathbf{e}_{2-}(\mathbf{x},t)}{||\nabla\Phi(\mathbf{x},t,T)\mathbf{e}_{2-}(\mathbf{x},t)||} \right). \end{aligned}$$

Now the numerators above can be rewritten using Lemma 5.2, and similarly Lemma 5.3 can be applied to the denominators. Using orthonormality of $\mathbf{e}_1$ and $\mathbf{e}_2$ and simplifying the notation by using (5.27), this leads us to:

$$r = \lim_{T\to\infty} \frac{\eta_1(T)^{2|T|}\alpha_1(T)\beta_1(T) + \eta_1(T)^{2|T|}\alpha_2(T)\beta_2(T)}{\sqrt{\eta_1(T)^{2|T|}\alpha_1(T)^2 + \eta_1(T)^{-2|T|}\alpha_2(T)^2}\sqrt{\eta_1(T)^{2|T|}\beta_1(T)^2 + \eta_1(T)^{-2|T|}\beta_2(T)^2}}.$$

The MLE is strictly positive by assumption, so $\eta_{1+} > 1$. Combining this with (5.28) yields:

$$r = \frac{\alpha_1\beta_1}{|\alpha_1||\beta_1|} = \pm 1,$$

completing the proof. □

A discussion of the process of this convergence along with results for the its speed can be found in [22]. For $T \to -\infty$, perturbations not aligned along the unstable Lyapunov direction converge to the stable one. Note that this does not mean that perturbations converge to some universal direction, since the Lyapunov directions are not constant in time. The condition in the above theorem that the Lyapunov directions should not coincide is generally satisfied in practice (see also section 6.3.1); only some very specifically constructed flow fields violate this property while retaining a positive MLE.

We can now fully describe how a slight perturbation $\delta\mathbf{x}(0)$ at $t = 0$ evolves if we know the Lyapunov exponents and directions. If it is aligned along the stable Lyapunov direction, then it will remain so by Theorem 5.7 and it will be damped,

$$||\delta\mathbf{x}(T)|| \sim \exp(-\lambda T)$$

for $T \to \infty$, with $\lambda > 0$ the maximum Lyapunov exponent. Otherwise, the component aligned along $\mathbf{e}_{1+}$, the direction perpendicular to the stable Lyapunov direction, will dominate. The direction of the perturbation will converge to the unstable Lyapunov direction by Theorem 5.8 and it will grow exponentially,

$$||\delta\mathbf{x}(T)|| \sim \exp(\lambda T)$$

for $T \to \infty$, by Theorem 5.5. Note that this only holds as long as the perturbation remains small.

5.1.2 Lyapunov exponents in periodic point vortex systems

In this section we will consider the relationship between mixing in point vortex systems and Lyapunov exponents, both asymptotic and finite time. Results will not be derived as rigorously as in the preceding chapter; in some cases more formal arguments can be found in the cited literature, but in other cases numerical observations, which will be shown in section 5.3, are the only backup of the informal arguments we use.

Consider a periodic point vortex system, with T_p the period and $\mathcal{R}$ the main ergodic region. Recall that this implies that $\mathcal{R}$ is invariant under f and that mixing takes place inside it, i.e. any open set contained in $\mathcal{R}$ will asymptotically fill it completely in the sense of Theorem 4.9. Note that $\nabla\Phi(t, T)$ is periodic in t for all $T \in \mathbb{R}$, with period T_p. This implies that the Lyapunov exponents and directions introduced in the last section, both asymptotic and finite time, are also periodic in t with period T_p. So it is sufficient to consider only the interval $0 \leq t \leq T_p$; we will go even further and consider only $t = 0$, allowing us to omit the time argument. We will write $\nabla f(\mathbf{x}, n)$ for the Jacobian of the periodic mapping f^n (see section 3.1) with $n \in \mathbb{N}$:

$$\nabla f(\mathbf{x}, n) := \nabla\Phi(\mathbf{x}, nT_p). \tag{5.29}$$

We also make the assumption, already mentioned in the previous section, that

$$(\mathbf{e}_{2-}(\mathbf{x}), \mathbf{e}_{1+}(\mathbf{x})) \neq 0, \tag{5.30}$$

almost everywhere, i.e. that the stable and unstable Lyapunov directions do not coincide.

First, we consider some continuity and boundedness properties. By definition the ergodic region does not contain a vortex; if it did it would be a vortex core (see section 3.3). As a result of this the velocity (3.2) is bounded on $\mathcal{R}$, so $\nabla\Phi(\mathbf{x},t,T)$ is as well for fixed $t,T \in \mathbb{R}$. This means that the finite time Lyapunov exponents $\lambda_1(T)$ and $\lambda_2(T)$ are bounded on $\mathcal{R}$. The same obviously applies to the asymptotic Lyapunov exponents by the existence of the limit (5.14). Since $\Phi(\mathbf{x},t,T)$ is continuously differentiable (Theorem 3.1), $\nabla\Phi(\mathbf{x},t,T)$ is continuous in $\mathbf{x} \in \mathcal{R}$ for all $t,T \in \mathbb{R}$. This also applies to $\Lambda(\mathbf{x},t,T)$ then, and hence the finite time Lyapunov exponents and directions are continuous on $\mathcal{R}$. The asymptotic Lyapunov directions are continuous as well by the stable manifold theorem [63]. In general the asymptotic Lyapunov exponents are not continuous, however as we will now show they are constant almost everywhere on $\mathcal{R}$. Note that this implies continuity almost everywhere, but not necessarily everywhere.

We will now argue that the MLE $\lambda_{1+}(\mathbf{x})$ is constant almost everywhere on $\mathcal{R}$. To do this we will derive an alternative expression for λ_{1+}, which will turn out to be quite useful in its own right. It is based on the fact that a perturbation directed along the unstable Lyapunov direction $\mathbf{e}_{2-}$ remains so. First, note that using Theorem 5.5 we have:

$$\lambda_{1+}(\mathbf{x}) = \lim_{T\to\infty} \frac{1}{|T|} \log ||\nabla\Phi(\mathbf{x},T)\mathbf{e}_{2-}(\mathbf{x})|| = \lim_{n\to\infty} \frac{1}{nT_p} \log ||\nabla f(\mathbf{x},n)\mathbf{e}_{2-}(\mathbf{x})||. \tag{5.31}$$

where the limit for $T \in \mathbb{R}$ is discretised to a limit for $n \in \mathbb{N}$. Note that the condition in the theorem is satisfied due to (5.30). We can decompose the expression in the logarithm as follows:

$$\begin{aligned} ||\nabla f(\mathbf{x},n)\mathbf{e}_{2-}(\mathbf{x})|| &\overset{(5.3)}{=} ||\nabla f(f(\mathbf{x}),n-1)\nabla f(\mathbf{x},1)\mathbf{e}_{2-}(\mathbf{x})|| \overset{\text{T5.7}}{=} \\ &\quad ||\nabla f(f(\mathbf{x}),n-1)\mathbf{e}_{2-}(f(\mathbf{x}))||||\nabla f(\mathbf{x},1)\mathbf{e}_{2-}(\mathbf{x})||. \end{aligned} \tag{5.32}$$

This shows why we need to use a perturbation aligned along the unstable Lyapunov direction: if not we cannot split the norms as done in the second equality above. Now define $h : \mathcal{R} \to \mathbb{R}$ by

$$h(\mathbf{x}) := \frac{1}{T_p} \log ||\nabla f(\mathbf{x},1)\mathbf{e}_{2-}(\mathbf{x})||.$$

Repeatedly applying (5.32) then gives

$$||\nabla f(\mathbf{x},n)\mathbf{e}_{2-}(\mathbf{x})|| = \prod_{i=0}^{n-1} \exp(T_p h(f^n(\mathbf{x}))) = \exp\left(T_p \sum_{i=0}^{n-1} h(f^n(\mathbf{x}))\right). \tag{5.33}$$

Now returning to (5.31):

$$\lambda_{1+}(\mathbf{x}) = \lim_{n\to\infty} \frac{1}{nT_p} \log\exp\left(T_p \sum_{i=0}^{n-1} h(f^n(\mathbf{x}))\right) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} h(f^n(\mathbf{x})). \tag{5.34}$$

So the asymptotic Lyapunov exponent is just the average of h over the trajectory. Now suppose $\mathbf{x}$ is a generator of the ergodic region $\mathcal{R}$ (recall that this holds almost everywhere in $\mathcal{R}$ by Theorem 4.4). Its iterates then 'fill up' the entire set $\mathcal{R}$, asymptotically uniformly because f is symplectic. So the computation of the average in (5.34) can be replaced by a spatial averaging. This result is stated more formally in Birkhoff's ergodic theorem, which in our case states the following [13, 43].

Theorem 5.9 (Birkhoff). *For any L_1-integrable function $g : \mathcal{R} \to \mathbb{R}$, we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} g(f^n(\mathbf{x})) = \frac{1}{\mu(\mathcal{R})} \int_{\mathbf{y}\in\mathcal{R}} g(\mathbf{y})\mathrm{dA}$$

for almost all $\mathbf{x} \in \mathcal{R}$.

Using Theorem 5.3 and boundedness of the finite time Lyapunov exponents, h is bounded on $\mathcal{R}$ and thus L_1-integrable, so

$$\lambda_{1+}(\mathbf{x}) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} h(f^n(\mathbf{x})) = \frac{1}{\mu(\mathcal{R})} \int_{\mathbf{y}\in\mathcal{R}} h(\mathbf{y})\mathrm{dA} \qquad (5.35)$$

for almost all $\mathbf{x} \in \mathcal{R}$. The right side is independent of $\mathbf{x}$, so this shows that λ_{1+} is constant almost everywhere in $\mathcal{R}$, and that it can be written as a time averaging of a finite time stretching exponent. This also allows us to omit the position argument, writing simply λ_{1+} for the MLE of the ergodic region. We will assume throughout that the MLE is strictly positive, i.e.

$$\lambda_{1+} > 0. \qquad (5.36)$$

The fact that the MLE is constant on $\mathcal{R}$ means it is a useful timescale in describing the mixing process, since stretching of material elements is a crucial element of this process. We would also like to be able to use Lyapunov exponents to learn more about the structures described at the end of the previous chapter, the partial mixing structure and the mixing time distribution. We will show that both can be rigorously defined as a measure on the ergodic region.

Obviously the asymptotic Lyapunov exponents will give no information on the mixing structures, as they are constant almost everywhere within the ergodic region. This means we must turn our attention to the finite time Lyapunov exponents. This immediately presents a problem: we are faced with an arbitrary choice of the Lyapunov time T, which is not present in the definition of the mixing structures. So we want to define a measure using the finite time Lyapunov exponents which converges for $T \to \infty$ and for $T \to -\infty$, but not to the trivial constant measure. Similar to the treatment of the previous section consider only the case of positive T; the argument for negative T is similar.

Consider two small subsets of the ergodic region $\mathcal{A}$ and $\mathcal{B}$ of equal measure, and take any $\mathbf{a} \in \mathcal{A}$ and $\mathbf{b} \in \mathcal{B}$. We will try to quantify the difference in mixing times of these two regions using the finite time maximum Lyapunov exponents (FTMLEs). Suppose $\mathcal{A}$ mixes poorly, i.e. has a large mixing time, while $\mathcal{B}$ mixes well. We would then expect $\mathcal{A}$ to undergo less stretching initially, so the FTMLE will be lower. So a first approach would be to consider

$$\frac{\lambda_1(\mathbf{a},T)}{\lambda_1(\mathbf{b},T)}.$$

For properly chosen T this may give some information on the relative mixing properties, but recall that we want to avoid this choice. Taken in the limit for $T \to \infty$, the quantity above converges to 1, since both the numerator and the denominator converge to λ_{1+}, and so doesn't carry the information we are looking for. Instead of using the ratio of the Lyapunov exponents, we could also consider the ratio of the actual stretching:

$$\frac{\exp(\lambda_1(\mathbf{a},T)|T|)}{\exp(\lambda_1(\mathbf{b},T)|T|)}. \tag{5.37}$$

Asymptotically both the numerator and the denominator will behave as $\exp(\lambda_{1+}T)$, but if $\mathcal{A}$ initially undergoes less stretching the numerator will remain smaller than the denominator by some factor. This is assuming of course that this ratio above actually converges for $T \to \infty$. Unfortunately it doesn't, as we will now show in outline (the full derivation is somewhat cumbersome). Assume that we can write $\mathbf{b} = f^m(\mathbf{a})$ with $m \in \mathbb{N}$ and suppose for the moment that (5.37) converges for $T \to \infty$:

$$\begin{aligned}
&\lim_{T\to\infty}\frac{\exp(\lambda_1(\mathbf{a},T)|T|)}{\exp(\lambda_1(\mathbf{b},T)|T|)} = \lim_{n\to\infty}\frac{\exp(\lambda_1(\mathbf{a},nT_p)nT_p)}{\exp(\lambda_1(\mathbf{b},nT_p)nT_p)} \overset{(5.9)}{=} \lim_{n\to\infty}\frac{\eta_1(\mathbf{a},nT_p)^{nT_p}}{\eta_1(\mathbf{b},nT_p)^{nT_p}} \overset{\text{T5.3}}{=}\\
&\lim_{n\to\infty}\frac{||\nabla f(\mathbf{a},n)\mathbf{e}_{2-}(\mathbf{a})||/(\mathbf{e}_1(\mathbf{a},n),\mathbf{e}_{2-}(\mathbf{a}))}{||\nabla f(\mathbf{b},n)\mathbf{e}_{2-}(\mathbf{b})||/(\mathbf{e}_1(\mathbf{b},n),\mathbf{e}_{2-}(\mathbf{b}))} \overset{(5.30)}{=}\\
&\frac{(\mathbf{e}_{1+}(\mathbf{b}),\mathbf{e}_{2-}(\mathbf{b}))}{(\mathbf{e}_{1+}(\mathbf{a}),\mathbf{e}_{2-}(\mathbf{a}))}\lim_{n\to\infty}\frac{||\nabla f(\mathbf{a},n)\mathbf{e}_{2-}(\mathbf{a})||}{||\nabla f(f^m(\mathbf{a}),n)\mathbf{e}_{2-}(f^m(\mathbf{a}))||} \overset{\text{T5.7}}{=}\\
&\frac{(\mathbf{e}_{1+}(\mathbf{b}),\mathbf{e}_{2-}(\mathbf{b}))}{(\mathbf{e}_{1+}(\mathbf{a}),\mathbf{e}_{2-}(\mathbf{a}))}||\nabla f(\mathbf{a},m)\mathbf{e}_{2-}(\mathbf{a})||\lim_{n\to\infty}\frac{1}{||\nabla f(f^n(\mathbf{a}),m)\mathbf{e}_{2-}(f^n(\mathbf{a}))||}.
\end{aligned}$$

The remaining limit is the inverse of the stretching from period n to period $n+m$. This value will clearly fluctuate depending on whether $f^n(\mathbf{a})$ is in a regime of high stretching or low stretching, so the limit does not exist. The crucial issue is that although $\exp(\lambda_1(\mathbf{a},T)|T|))$ does behave as $\exp(\lambda_{1+}T)$ for $T \to \infty$, it does not do this in the strong sense that their ratio converges. Instead of taking just one element out of $\mathcal{A}$ and $\mathcal{B}$, we could of course use a large amount and take the average, but this would not solve the problem. However, if we do not use a finite amount of points we obtain a converging quantity:

$$F(\mathcal{A},\mathcal{B}) := \lim_{T\to\infty} \frac{\int_{\mathbf{y}\in\mathcal{A}} \exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA}}{\int_{\mathbf{y}\in\mathcal{B}} \exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA}}. \tag{5.38}$$

We will argue that this limit exists and show how it is related to the mixing time distribution. First, let us examine the asymptotic properties of these integrals. Consider the behavior of

$$S(T) := \frac{1}{\mu(\mathcal{R})} \int_{\mathbf{y}\in\mathcal{R}_c} \exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA}. \tag{5.39}$$

$\mathcal{R}_c$ is defined as the subset of $\mathcal{R}$ consisting of all $\mathbf{x}$ such that $\lim\limits_{T\to\infty} \lambda_1(\mathbf{x},T) = \lambda_{1+}$. This is chosen for convenience; since $\mathcal{R}_c \subset \mathcal{R}$ and $\mathcal{R} \setminus \mathcal{R}_c$ is of measure zero the value of $S(T)$ would be the same integrated over $\mathcal{R}$. Note that $S(T)$ is simply the average value over $\mathcal{R}$ of $\exp(\lambda_1(T)|T|)$. We are interested in its scaling properties, so define

$$\sigma(T) := \frac{1}{|T|} \log S(T), \tag{5.40}$$

with corresponding limit for $T \to \infty$ σ_+ (its existence is assumed). σ_+ is known as the topological entropy [32, 58, 76]. Actually, the original definition of the topological entropy is rather different from the one given here. The statement that σ_+ is the topological entropy is based mainly on the fact that this turns out to hold numerically. We will return to this point in sections 5.3.1 and 6.1.2. Using Jensen's inequality [40], we find:

$$\begin{aligned}
\sigma_+ = &\lim_{T\to\infty} \frac{1}{|T|} \log \frac{1}{\mu(\mathcal{R})} \int_{\mathbf{y}\in\mathcal{R}_c} \exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA} \geq \\
&\lim_{T\to\infty} \frac{1}{\mu(\mathcal{R})} \int_{\mathbf{y}\in\mathcal{R}_c} \frac{1}{|T|} \log\exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA} = \\
&\lim_{T\to\infty} \frac{1}{\mu(\mathcal{R})} \int_{\mathbf{y}\in\mathcal{R}_c} \lambda_1(\mathbf{y},T)\mathrm{dA} = \frac{1}{\mu(\mathcal{R})} \int_{\mathbf{y}\in\mathcal{R}_c} \lim_{T\to\infty} \lambda_1(\mathbf{y},T)\mathrm{dA} = \\
&\frac{1}{\mu(\mathcal{R})} \int_{\mathbf{y}\in\mathcal{R}_c} \lambda_{1+}\mathrm{dA} = \lambda_{1+}.
\end{aligned} \tag{5.41}$$

So the topological entropy is greater than or equal to the MLE. At first glance it might seem that equality should hold; after all, the integrand behaves as $\exp(\lambda_{1+}t)$ asymptotically almost everywhere in $\mathcal{R}$. So let us see if we can show the opposite inequality:

$$\begin{aligned}
\sigma_+ = &\lim_{T\to\infty}\frac{1}{|T|}\log\frac{1}{\mu(\mathcal{R})}\int_{\mathbf{y}\in\mathcal{R}_c}\exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA} = \\
&\lim_{T\to\infty}\frac{1}{|T|}\log\frac{1}{\mu(\mathcal{R})}\int_{\mathbf{y}\in\mathcal{R}_c}\exp(\lambda_{1+}|T|)\exp((\lambda_1(\mathbf{y},T)-\lambda_{1+})|T|)\mathrm{dA} = \\
&\lambda_{1+}+\lim_{T\to\infty}\frac{1}{|T|}\log\frac{1}{\mu(\mathcal{R})}\int_{\mathbf{y}\in\mathcal{R}_c}\exp((\lambda_1(\mathbf{y},T)-\lambda_{1+})|T|)\mathrm{dA} \leq \\
&\lambda_{1+}+\lim_{T\to\infty}\frac{1}{|T|}\log\frac{1}{\mu(\mathcal{R})}\int_{\mathbf{y}\in\mathcal{R}_c}\exp(\max_{\mathbf{x}\in\mathcal{R}_c}|\lambda_1(\mathbf{x},T)-\lambda_{1+}||T|)\mathrm{dA} = \\
&\lambda_{1+}+\lim_{T\to\infty}\max_{\mathbf{x}\in\mathcal{R}_c}|\lambda_1(\mathbf{x},T)-\lambda_{1+}|.
\end{aligned} \tag{5.42}$$

So if the FTMLE $\lambda_1(\mathbf{x},T)$ converges *uniformly* to the MLE λ_{1+} the opposite inequality holds as well and we have $\sigma_+ = \lambda_{1+}$. This shows that the difference between the topological entropy and the MLE is, in some sense, the degree of non-uniformity in the convergence of the FTMLEs. We will see that in practice σ_+ is strictly greater than λ_{1+}, so the FTMLE converges to the MLE non-uniformly.

Now let us return to $F(\mathcal{A},\mathcal{B})$ as defined in (5.38). Consider the behavior of the integral in the numerator. Initially this will depend on the specific location of $\mathcal{A}$, i.e. whether it is in a region of poor mixing. However, eventually this set will be spread out over the full ergodic region, so we expect it to scale as $\exp(\sigma_+ T)$. There is a crucial difference with the stretching associated with single trajectories as considered in (5.37): after being spread out over the full ergodic region, the trajectories in $\mathcal{A}$ no longer collectively pass through different regimes. This means we can expect the limit

$$\lim_{T\to\infty}\frac{\int_{\mathbf{y}\in\mathcal{A}}\exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA}}{\exp(\sigma_+|T|)} \tag{5.43}$$

to converge, directly showing the convergence of the limit in (5.38). We can now define a non-trivial measure based on the FTMLE:

$$\mu_{w+}(\mathcal{C}) := F(\mathcal{C},\mathcal{R}) = \lim_{T\to\infty}\frac{\int_{\mathbf{y}\in\mathcal{C}}\exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA}}{\int_{\mathbf{y}\in\mathcal{R}}\exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA}}, \tag{5.44}$$

for all $\mathcal{C}\subset\mathcal{R}$. Note that, assuming its existence as argued above, this is a probability measure on $\mathcal{R}$. Following [2], it will be referred to as the (positive time) w-measure. Just like Lyapunov exponents and directions, we can also define it for finite time by omitting the limit:

$$\mu_w(\mathcal{C},T) := \frac{\int_{\mathbf{y}\in\mathcal{C}}\exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA}}{\int_{\mathbf{y}\in\mathcal{R}}\exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA}}. \tag{5.45}$$

Having managed to define this measure, we still need to relate it to the mixing time distribution. This is now fairly straightforward. Consider again the sets $\mathcal{A}$ and $\mathcal{B}$ as defined above; recall that $\mathcal{A}$ mixes poorly while $\mathcal{B}$ mixes well. Consider the behavior of the integrals in (5.38). Using (5.44), we know that *asymptotically* for $T \to \infty$:

$$\int_{\mathbf{y}\in\mathcal{A}} \exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA} = \mu_{w+}(\mathcal{A}) \int_{\mathbf{y}\in\mathcal{R}} \exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA} = C\exp(\sigma_+(T-\tau(\mathcal{A}))), \tag{5.46}$$

defining C as

$$C := \lim_{T\to\infty} \frac{\int_{\mathbf{y}\in\mathcal{R}} \exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA}}{\exp(\sigma_+ T)} \tag{5.47}$$

and $\tau(\mathcal{A})$ as

$$\tau(\mathcal{A}) := -\frac{1}{\sigma_+}\log(\mu_{w+}(\mathcal{A})). \tag{5.48}$$

A similar expression holds for $\mathcal{B}$. Eventually both sets will become indistinguishable as they are mixed by the flow (recall that they have equal measure), so any difference between $\tau(\mathcal{A})$ and $\tau(\mathcal{B})$ will be caused by the initial mixing. From (5.46) we see that it can be considered a lag time before asymptotic behavior sets in. This is consistent with what we have seen in the previous chapter: there is an initial period in which no folding takes place, after which mixing starts according to the partial mixing structure which does not depend on the starting location. From this it becomes clear that $\tau(\mathcal{A}) - \tau(\mathcal{B})$ is the difference in mixing times between $\mathcal{A}$ and $\mathcal{B}$. Note that we cannot make a direct connection between $\tau(\mathcal{A})$ and the mixing time of $\mathcal{A}$ because they are not uniquely defined[2].

In other words, using the measure μ_{w+} we can identify which regions are about to undergo high stretching and thus mixing. Similarly, taking the limit $T \to -\infty$ we find μ_{w-}, with which we can identify regions which *have* undergone high stretching. This establishes the link between μ_{w-} and the partial mixing structure, and formalizes the duality between the mixing time distribution and the partial mixing structure as each others' reverse time equivalents. The relationship between μ_{w-} and the partial mixing structure will be examined more closely in section 6.2.3.

---

[2] In particular, the calculation method for the mixing time described in section 4.3.2 has a somewhat arbitrary parameter δ, and there are other reasonable choices for C in (5.46).

5.2 Numerical methods

5.2.1 Computing Lyapunov exponents

To compute the Lyapunov exponents, topological entropy and w-measure, we first need to find the Jacobian of the flow map $\nabla\Phi(\mathbf{x}, t, T)$ for $\mathbf{x} \in \mathbb{R}^2$ and $t, T \in \mathbb{R}$. We can find its columns by considering its action on the standard basis vectors:

$$\begin{aligned} (\nabla\Phi(\mathbf{x}, t, T))_{\cdot 1} &= \nabla\Phi(\mathbf{x}, t, T)\mathbf{e}_x \overset{(5.1)}{=} \lim_{\epsilon\to 0} \frac{\Phi(t, T)(\mathbf{x} + \epsilon\mathbf{e}_x) - \Phi(t, T)(\mathbf{x})}{\epsilon} \\ (\nabla\Phi(\mathbf{x}, t, T))_{\cdot 2} &= \nabla\Phi(\mathbf{x}, t, T)\mathbf{e}_y \overset{(5.1)}{=} \lim_{\epsilon\to 0} \frac{\Phi(t, T)(\mathbf{x} + \epsilon\mathbf{e}_y) - \Phi(t, T)(\mathbf{x})}{\epsilon}. \end{aligned} \tag{5.49}$$

So we can find the first column of $\nabla\Phi(\mathbf{x}, t, T)$ by following the trajectory of $\mathbf{x}$ and the slightly perturbed trajectory $\mathbf{x}_{(x)} = \mathbf{x} + \delta\mathbf{e}_x$ from t to $t+T$. The result is then simply the perturbation at time $t+T$ divided by δ. The second column is found similarly by taking the perturbation in the y direction. There is a numerical issue however: if the MLE is positive, then these perturbations will grow exponentially. This means that although they may be initially small enough to be in the linear regime, they will not remain so indefinitely, which is a problem when considering large T.[3] The standard solution to this problem is to rescale the perturbation whenever it threatens to escape the linear regime [25].

In practice, the following algorithm is used to determine the first column of $\nabla\Phi(\mathbf{x}, t, T)$. Set

$$\mathbf{x}_{(x)} := \mathbf{x} + \delta\mathbf{e}_x,$$

with δ small (typically, $\delta = 10^{-8}$ is used), and

$$s_x := 0.$$

This variable will be used to keep track of any rescaling done. Apply $\Phi(t, \Delta t)$ to both $\mathbf{x}$ and $\mathbf{x}_{(x)}$, i.e. evolve both trajectories by one timestep. We use s_x to keep track of the growth of the perturbation:

$$s_x := s_x + \log\left(\frac{||\mathbf{x}_{(x)} - \mathbf{x}||}{\delta}\right).$$

Now rescale the perturbation so that its size is set to δ again:

$$\mathbf{x}_{(x)} := \mathbf{x} + \delta\frac{\mathbf{x}_{(x)} - \mathbf{x}}{||\mathbf{x}_{(x)} - \mathbf{x}||},$$

---

[3] Note that we cannot take the perturbation arbitrarily small. Since it is defined as the difference between two trajectories, this would cause numerical issues.

The above step, i.e. applying $\Phi(t, \Delta t)$ and rescaling, is repeated until $t+T$ is reached. At this point, we have found the first column of $\nabla\Phi(\mathbf{x}, t, T)$ as:

$$(\nabla\Phi(\mathbf{x}, t, T))_{\cdot 1} = \exp(s_x)\frac{\mathbf{x}_{(x)} - \mathbf{x}}{\delta} = \begin{pmatrix} \exp(s_x)\cos\theta_x \\ \exp(s_x)\sin\theta_x \end{pmatrix},$$

with θ_x the angle describing the orientation of the perturbation at time $t + T$. The second column is found similarly; note that we do not need to recompute the trajectory of $\mathbf{x}$ this time, but only the perturbed trajectory $\mathbf{x}_{(y)}$. This yields:

$$\nabla\Phi(\mathbf{x}, t, T) = \begin{pmatrix} \exp(s_x)\cos\theta_x & \exp(s_y)\cos\theta_y \\ \exp(s_x)\sin\theta_x & \exp(s_y)\sin\theta_y \end{pmatrix}. \tag{5.50}$$

The matrix is stored as $(s_x, s_y, \theta_x, \theta_y)$; for large values of T the actual values of the entries exceed the numerical limit. A similar method is described in [39].

Having found $\nabla\Phi(\mathbf{x}, t, T)$, we still need to extract the Lyapunov exponents from it. Recall that the FTMLE $\lambda_1(\mathbf{x}, t, T)$ is the logarithm of the maximal eigenvalue of

$$\Lambda(\mathbf{x}, t, T) = (\nabla\Phi(\mathbf{x}, t, T)^*\nabla\Phi(\mathbf{x}, t, T))^{1/2|T|}.$$

Using (5.7), this yields:

$$\lambda_1(\mathbf{x}, t, T) = \frac{\log\left(\frac{1}{2}\left(\mathrm{tr} + \sqrt{\mathrm{tr}^2 - 4}\right)\right)}{2|T|}, \tag{5.51}$$

with

$$\mathrm{tr} = \exp(2s_x) + \exp(2s_y).$$

However, we run into the problem of numerical overflow again here. This can be remedied by using the following expression when $s_x \gg 1$ and/or $s_y \gg 1$:

$$\lambda_1(\mathbf{x}, t, T) \approx \frac{2\max(s_x, s_y) + \log(1 + \exp(-2|s_y - s_x|))}{2|T|}. \tag{5.52}$$

The MLE λ_{1+} is simply the limit for $T \to \infty$ of the FTMLE $\lambda_1(\mathbf{x}, t, T)$ with $\mathbf{x} \in \mathcal{R}$. We have just seen how to compute the FTMLE, so finding the MLE is just a matter of taking this limit. However, this is not the most efficient method. Note that it requires the computation of three trajectories over a large time T-the unperturbed one, and the ones initially perturbed in the x and in the y direction. Using Theorem 5.5 however, we see that we need to consider only one perturbed trajectory-its growth rate, although not equal to the FTMLE for finite T, will converge to the MLE for $T \to \infty$. The algorithm is very similar to the one above, except that no perturbation in the y direction is used. The MLE is simply given by the limit for $T \to \infty$ of:

$$\lambda_{1+,\mathrm{est}}(T) = \frac{s_x(T)}{T}. \tag{5.53}$$

5.2.2 Computing the topological entropy and the w-measure

Recall that the topological entropy is defined as

$$\sigma_+ = \lim_{T\to\infty} \frac{1}{|T|} \log \frac{1}{\mu(\mathcal{R})} \int_{\mathbf{y}\in\mathcal{R}_c} \exp(\lambda_1(\mathbf{y}, T)|T|)\mathrm{dA} \tag{5.54}$$

by (5.39) and (5.40). Computing this integral for finite T is theoretically simple; after all, the integrand is continuous on $\mathcal{R}$ (see section 5.1.2). We simply take a finite but sufficient number of points spread uniformly over $\mathcal{R}$ and compute $\lambda_1(T)$ for each of these points as described in the previous section; we can then directly compute the integral and if T is sufficiently large we have the topological entropy. However, there are a two snags.

Firstly, if we fix the number of points used to compute the integral then (5.54) will simply converge to the MLE λ_{1+}, not the topological entropy, as described in section 5.1.2. So as we increase T we also need to increase the number of points used to compute the integral as the FTMLE field becomes less smooth. This can rapidly increase the computational workload; it would be nice if we can reuse some computations when increasing T. The second issue is that we need to find a set of points spread uniformly over $\mathcal{R}$, requiring very detailed knowledge of the ergodic region if we need many points.

Both of these issues can be resolved for periodic systems by using the Poincaré map. Using Theorem 5.9 (Birkhoff) and discretising the limit for T, we can rewrite (5.54) as:

$$\sigma_+ = \lim_{n\to\infty} \lim_{N\to\infty} \frac{1}{nT_p} \log \frac{1}{N} \sum_{i=0}^{N-1} \exp(\lambda_1(f^i(\mathbf{x}), nT_p)nT_p). \tag{5.55}$$

Recall that T_p is the period of the system. The method now consists of computing $\nabla\Phi(\mathbf{y}, 0, T_p)$ as described in section 5.2.1 for all $\mathbf{y} \in \mathcal{P}(\mathbf{x}, M)$, with $\mathbf{x}$ a generator of $\mathcal{R}$ and M large. Computing (5.55) with $n = 1$ is then straightforward: we can find $\lambda_1(T_p)$ directly from $\nabla\Phi(T_p)$; how to achieve this is also described in section 5.2.1[4]. For higher n we need to find $\nabla\Phi(nT_p)$ for $\mathbf{y} \in \mathcal{P}(\mathbf{x}, M)$; this can be accomplished without recomputing any trajectories. Specifically, using (5.3), we have:

$$\nabla\Phi(\mathbf{y}, 0, nT_p) = \prod_{i=n-1}^{0} \nabla\Phi(f^i(\mathbf{y}), 0, T_p). \tag{5.56}$$

The matrices in the right hand side of this equation are available in the precomputed list. Numerical overflow could theoretically become an issue again here, necessitating rescaling, but in practice n is never large enough for this to occur. So after the precomputation we can compute the topological entropy according to (5.55) for any

---

[4] Note that the unperturbed trajectory is already known from the computation of the Poincaré map.

n and N with $N + n \leq M$ without having to recompute trajectories; this allows for fast convergence checking. Using the notation of (5.53), we can write σ_+ as the limit for $n, N \to \infty$ of:

$$\sigma_{+,\mathrm{est}}(N, n) = \frac{1}{nT_p} \log \frac{1}{N} \sum_{i=0}^{N-1} \exp(\lambda_1(f^i(\mathbf{x}), nT_p)nT_p). \tag{5.57}$$

The limit for N should be taken first, and then the one for n (see (5.55)).

The w-measure can also be obtained from this list. We rewrite (5.44) as:

$$\mu_{w+}(\mathcal{A}) = \lim_{T\to\infty} \frac{1}{\mu(R)C(T)} \int_{\mathbf{y}\in\mathcal{R}} \chi(\mathbf{y}, \mathcal{A}) \exp(\lambda_1(\mathbf{y}, T)|T|)\mathrm{dA}, \tag{5.58}$$

with χ defined by

$$\chi(\mathbf{y}, \mathcal{A}) = \begin{cases} 1 & \text{if } \mathbf{y} \in \mathcal{A}, \\ 0 & \text{if } \mathbf{y} \notin \mathcal{A}, \end{cases} \tag{5.59}$$

and $C(T)$ used as a normalization constant such that

$$\mu_{w+}(\mathcal{R}) = 1. \tag{5.60}$$

Similar to (5.55), we can rewrite this as:

$$\mu_{w+}(\mathcal{A}) = \lim_{n\to\infty} \lim_{N\to\infty} \frac{1}{NC(nT_p)} \sum_{i=0}^{N-1} \chi(f^i(\mathbf{x}), \mathcal{A}) \exp(\lambda_1(f^i(\mathbf{x}), nT_p)nT_p). \tag{5.61}$$

Again, everything on the right hand side can be obtained from the precomputed list. Note that we are using only those points that are inside $\mathcal{A}$; if we want to compute the w-measure on multiple subsets of $\mathcal{R}$, this filtering can of course be done simultaneously. If these subsets are a partition of $\mathcal{R}$, then C can be determined as well. Note that once the list is calculated, it can be used to find the behavior of small perturbations without having to compute any trajectories; this means it can also be used in calculating the MLE using the method described earlier.

The methods described in this section up to now require periodicity and so cannot be applied to irregular vortex systems. Instead, we need to use (5.44) directly to compute the w-measure. Suppose we have partitioned the relevant part of the phase space (i.e. that part containing the ergodic region) into a grid. We can then compute the Lyapunov exponents within each square, and use (5.44) to find the w-measure from this, treating the denominator as a normalization constant. For low T the integral can be approximated simply by the value in one point; for higher values of T averaging over multiple points within one square is required.

5.3 Results

Several results relating to the theory and methods described earlier in this chapter will now be presented. We first focus on computing the asymptotic time scales, i.e. the maximum Lyapunov exponent and the topological entropy (section 5.3.1). We will then attempt to find the w-measure and relate it to the mixing structures (section 5.3.2).

5.3.1 Asymptotic results

We will first compute the MLE for all non-stationary systems. First, consider the periodic three vortex system. The convergence of the estimate for the MLE using (5.53) up to 10^7 periods is shown in Figure 5.1. Note that the convergence process is very irregular as the trajectory passes through various stretching regimes. This makes it difficult to obtain an accurate value for the MLE, but from the results given we can conclude $\lambda_{1+} = 0.025$. Results for the other periodic systems are not pictured but the values for λ_{1+} can be found in Table 5.1. Convergence results for the irregular four vortex system are shown in Figure 5.2, leading us to conclude $\lambda_{1+} \approx 0.02$ for this system. Note that the Lyapunov exponent does not appear to have converged yet despite the high value of T; this poor convergence is explained by the fact that not only does the trajectory itself pass through different regimes of mixing as in periodic systems, but the entire vortex system passes through different regimes as well.

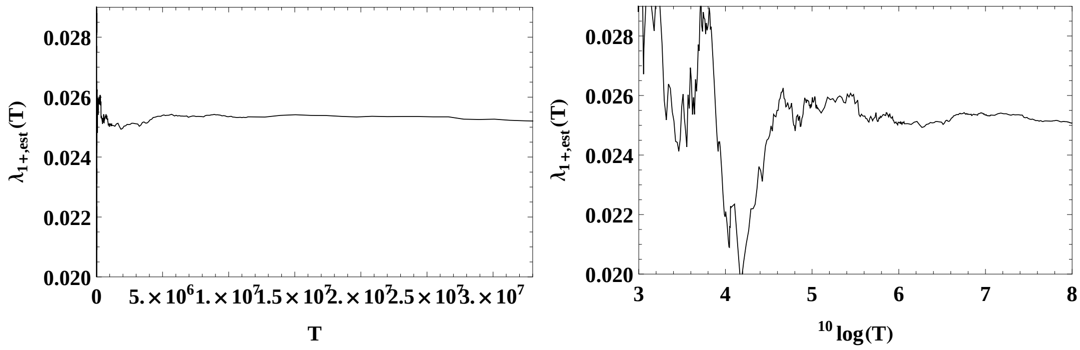

Figure 5.1: Convergence of the estimate for the MLE in the 3-periodic three vortex system as a function of T, using $\mathbf{x}$=(0,0). The right figure is a log-linear plot of the left figure to highlight the behavior of the convergence.

Finding the topological entropy is more difficult since it involves taking two limits. We first focus on the inner limit in (5.57), i.e. fixing n for now. Convergence results for various n are shown in Figure 5.3. Consider $n = 15$ first (bottom right). From

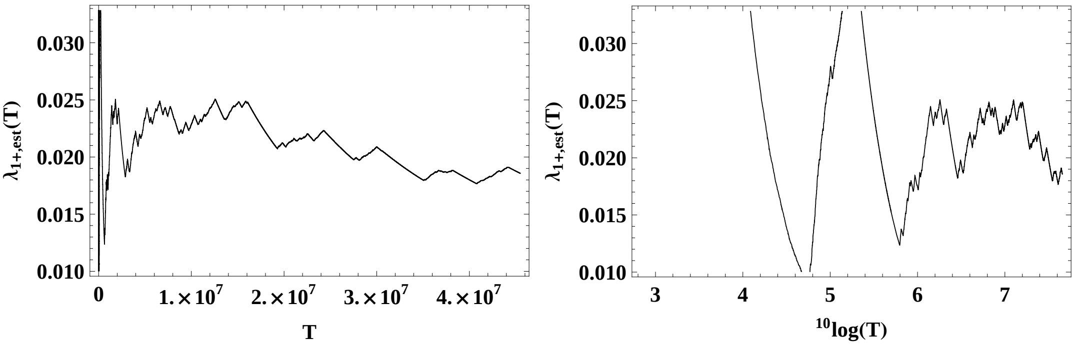

Figure 5.2: Convergence of the estimate for the MLE in the irregular four vortex system as a function of T, using $\mathbf{x}$=(0,0). The right figure is a log-linear plot of the left figure to highlight the behavior of the convergence.

System	N	n	$\lambda_{1+,\text{est}}(NT_p)$	$\sigma_{+,\text{est}}(N,n)$	$\sigma_{+,\text{man}}$
3-periodic three vortex	10^7	9	0.025	0.1	0.10
2-periodic three vortex	10^6	6	0.005	0.02	0.022
Periodic four vortex	10^6	5	0.08	0.3	0.27
Irregular four vortex	$\lambda_{1+,\text{est}}(4.6 * 10^7) = 0.02$			N/A	0.33

Table 5.1: Calculated values of the MLE and the topological entropy for various systems. For the meanings of n and N, see (5.53) and (5.57). The final column shows the value of the topological entropy found using the manifold method, which will be described in section 6.1.2. The method used to compute σ_+ for the periodic systems cannot be applied to the irregular system.

the sharp increases and gradual declines, it is apparent that the integral in (5.54) is dominated by small regions of extreme stretching. We can also see that this effect becomes more and more pronounced as n increases. As a result, convergence is determined by the requirement to 'hit' all these areas in the Poincaré map. For $N = 10^7$, the maximum value we can reasonably compute $\sigma_{+,est}$ for, reasonable convergence is achieved for $n \leq 9$.

Next, we consider the outer limit, i.e. the one for n. As explained above we cannot go above $n = 9$; the estimates for the topological entropy up to this point are shown in Figure 5.4. Convergence is far from apparent, but it still seems reasonable to conclude that $\sigma_+ \approx 0.1$. Estimates for the topological entropy for the other periodic systems can be found in Table 5.1 (the method cannot be applied to the irregular systems as mentioned in section 5.2.2). Note that for these systems $N = 10^6$, an order of magnitude lower than for the 3-periodic three vortex system. This was necessary because of the excessive computing times involved, but it means the other estimates are even less accurate. The main reason to compute them is to compare them with the estimates obtained using the manifold method, which will be discussed in the next chapter. Recall that our current definition of σ_+ is not actually the standard definition of the topological entropy; the manifold method however is based on this standard definition. Therefore the similar results obtained using both methods indicate that σ_+ is indeed the topological entropy. Also, note that the topological entropy is indeed greater than the MLE, indicating non-uniform convergence of the finite time Lyapunov exponents.

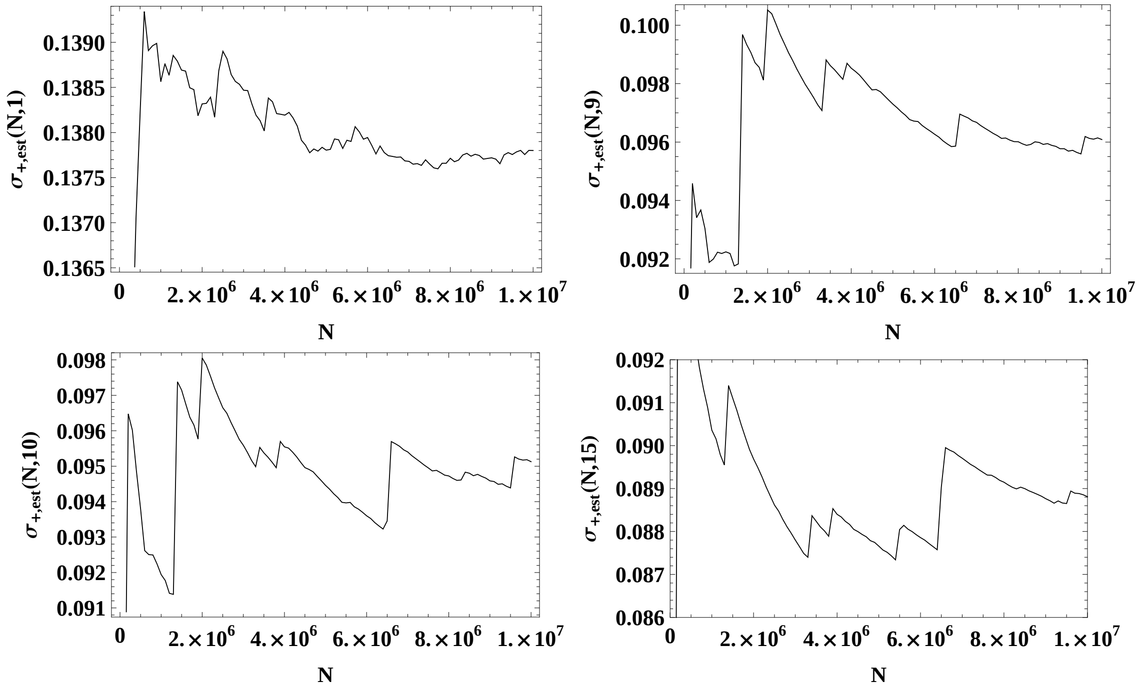

Figure 5.3: Convergence of the estimate for the topological entropy in the 3-periodic three vortex system as a function of N for various n, using $\mathbf{x} = (0, 0)$. From top left to bottom right going horizontally: n=1, n=9, n=10 and n=15.

5.3.2 The w-measure

The convergence behavior of the w-measure μ_{w+} as computed using (5.61) for increasing n is shown in Figure 5.5. Like all plots in this section, the color scale is logarithmic. This is chosen because otherwise features of the distribution in areas of

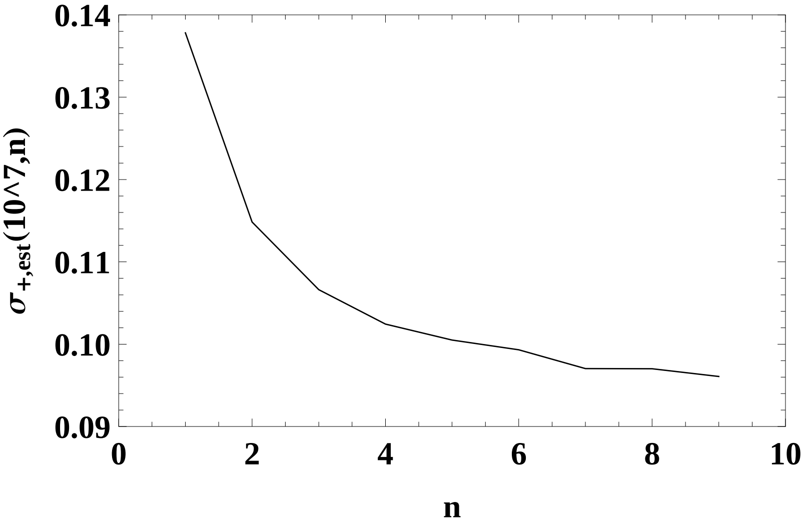

Figure 5.4: Convergence of the estimate for the topological entropy in the 3-periodic three vortex system as a function of n for $N = 10^7$, using $\mathbf{x} = (0,0)$.

low w-measure would not be not visible. It is clearly visible that as n increases the approximation of the w-measure becomes less and less regular. This is not surprising: it is continuous for all finite n (this follows from the definition (5.44) and continuity of finite time Lyapunov exponents), but not in the limit for $n \to \infty$, where a fractal structure appears. A more detailed discussion of the convergence of $\mu_w(T)$ to μ_{w+} and of the properties of this fractal structure can be found in [2, 31, 57].

To compare the w-measure to the mixing time distribution, we compute μ_{w+} from the mixing time distribution using (5.48), and compare this to the results of using (5.61). This comparison is shown in Figure 5.6; note that the agreement of the results is excellent. Similar results for the 2-periodic three vortex system and for the periodic four vortex system are shown in Figures 5.8 and 5.10 respectively. In all three cases, we can see that the features in areas of low w-measure are not as well resolved as those in areas of high w-measure. This is caused by the limited value of n; looking back at Figure 5.5 we see that for increasing n more and more features are visible.

The partial mixing structure is not as clearly defined as the mixing time distribution, so the comparison will be somewhat less detailed. As described in section 5.1.2, we need to use the *negative* time w-measure μ_{w-}. For the periodic systems, this measure can be found easily from μ_{w+}, since for a symmetric point vortex system reversing time just means mirroring in the symmetry axis (see section 4.3.1). Results for all periodic systems are shown in Figures 5.7, 5.9 and 5.11. Again, good agreement between the w-measure and the mixing structure is visible.

The positive time w-measure for the irregular four vortex system is compared with the mixing time distribution in Figure 5.12. The method using the Poincaré map applied to the periodic systems cannot be used here, and the alternative method described in section 5.2.2 is somewhat less efficient. This means we cannot take T as high as necessary to actually consider it as the limit for $T \to \infty$, but nonetheless the agreement with the mixing time distribution is excellent. The negative time w-measure[5] is compared with the partial mixing structure in the left part of Figure 5.13. Although the agreement is good, careful observation reveals that there are some features in the partial mixing structure not present in the w-measure. This is caused by the limited value of T.

---

[5] Note that since the irregular four vortex system is not symmetric we have to compute this separately-we can't just mirror μ_{w+} like before.

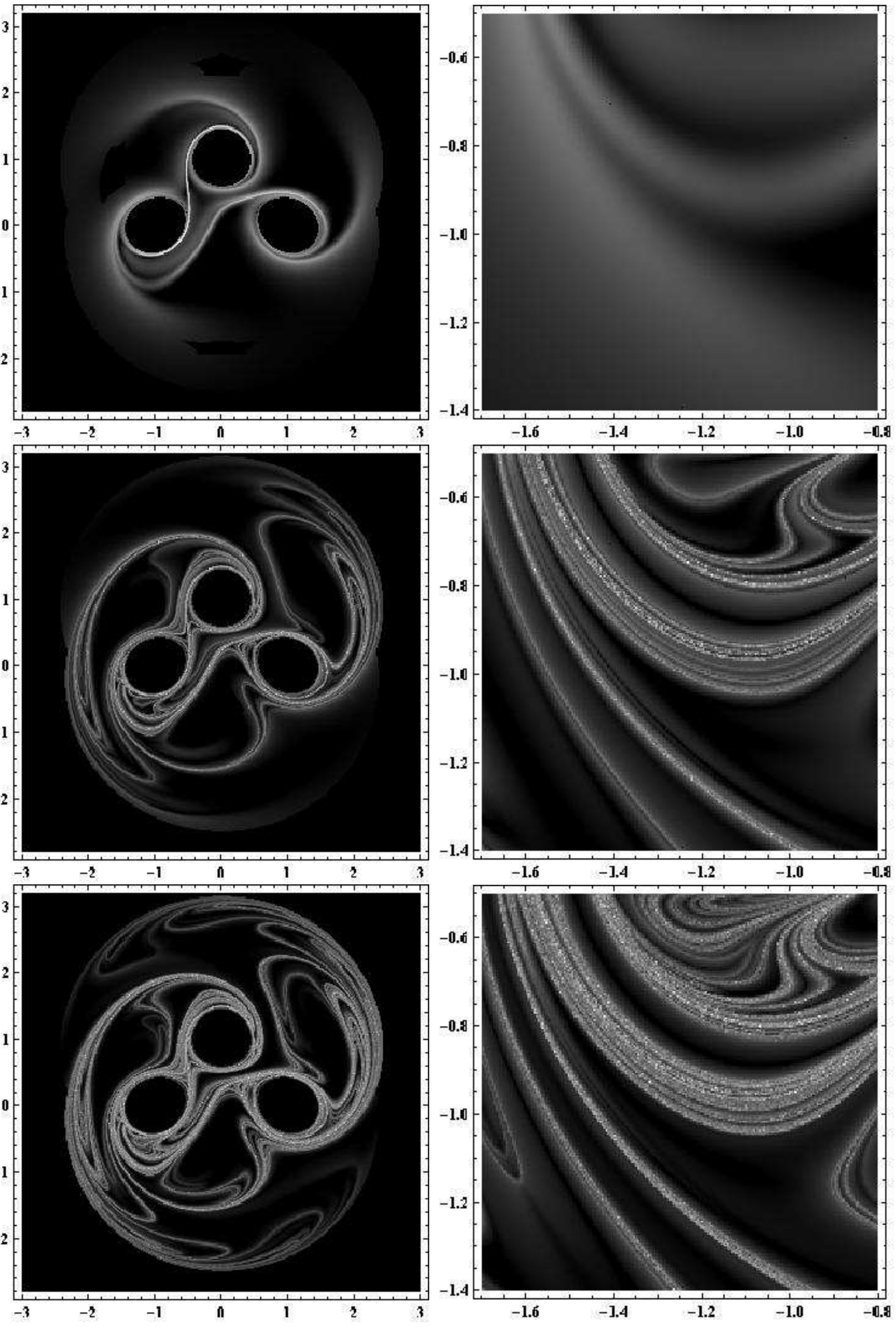

Figure 5.5: Left: w-measure μ_{w+} in the 3-periodic three vortex system computed using (5.61) with $N = 10^7$ and (from top to bottom) $n = 1$, $n = 5$ and $n = 9$ on $(-3, 3) \times (-2.8, 3.2)$ using a 300x300 grid. Right: detail showing $(-1.7, -0.8) \times (-1.4, -0.5)$ using a 200x200 grid.

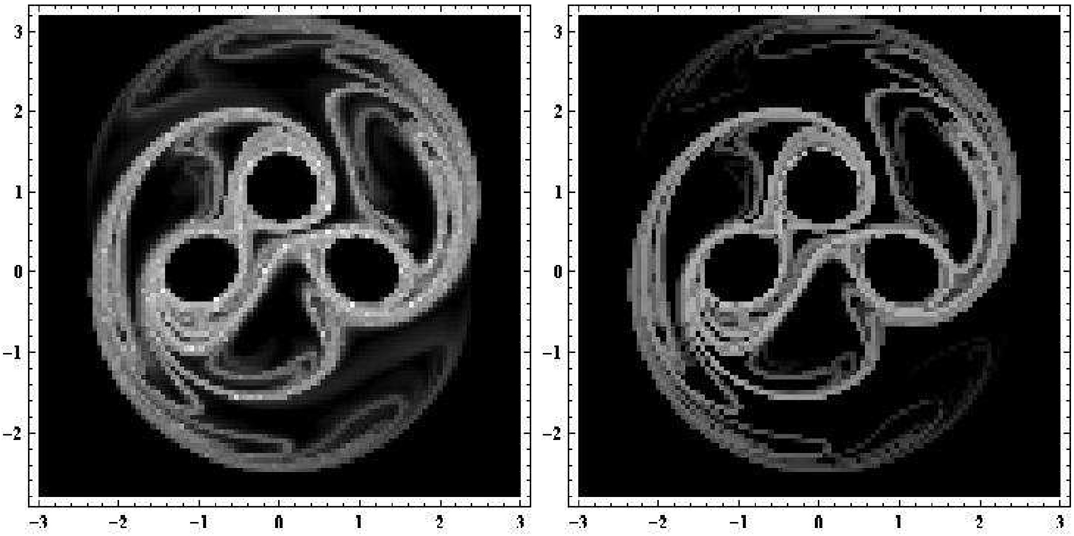

Figure 5.6: Left: w-measure μ_{w+} in the 3-periodic three vortex system computed using (5.61) with $N = 10^7$ and $n = 9$ on $(-3, 3) \times (-2.8, 3.2)$ using a 100x100 grid. Right: μ_{w+} as computed from the mixing time distribution as found in Figure 4.9, using (5.48).

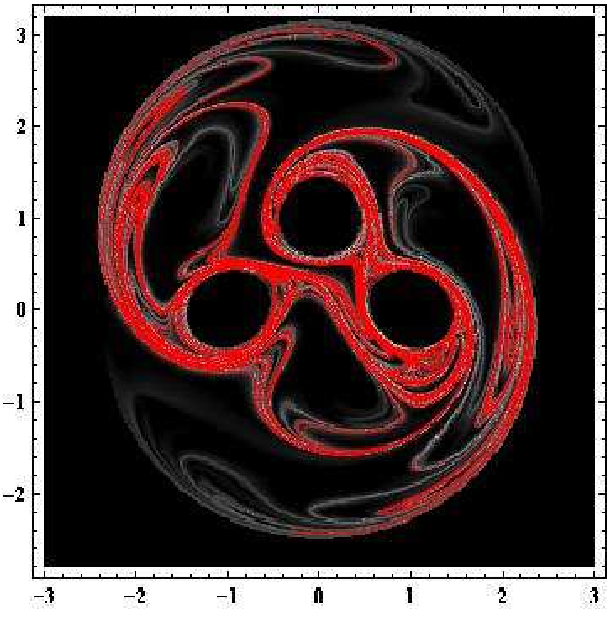

Figure 5.7: Negative time w-measure μ_{w-} in the 3-periodic three vortex system computed using (5.61) with $N = 10^7$ and $n = 9$ on $(-3, 3) \times (-2.8, 3.2)$ using a 300x300 grid. Red overlay: partial mixing structure as found in Figure 4.9.

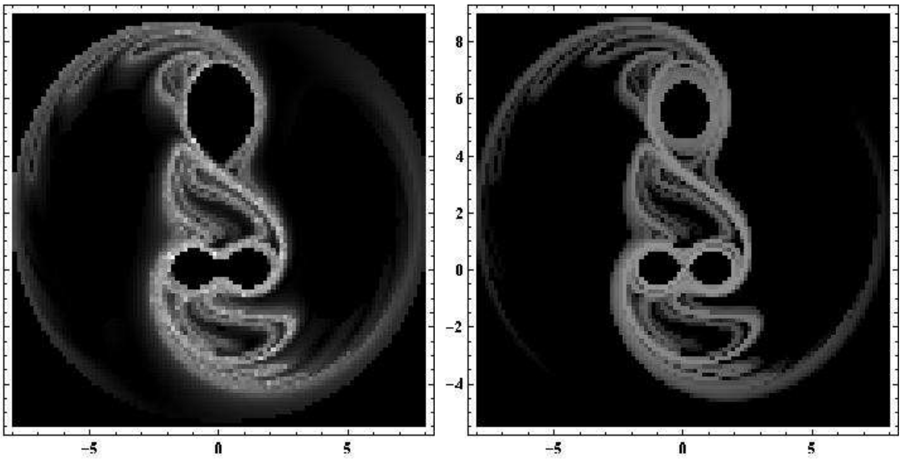

Figure 5.8: Left: w-measure μ_{w+} in the 2-periodic three vortex system computed using (5.61) with $N = 10^6$ and $n = 6$ on $(-8, 8) \times (-5.5, 9)$ using a 100x100 grid. Right: μ_{w+} as computed from the mixing time distribution as found in Figure 4.10, using (5.48).

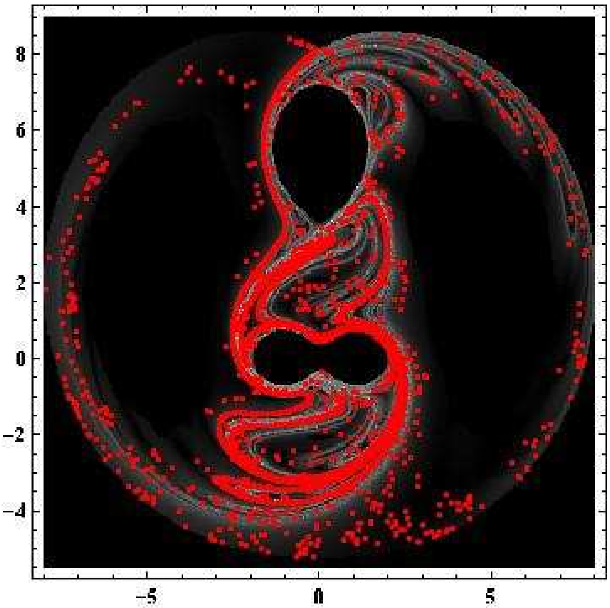

Figure 5.9: Negative time w-measure μ_{w-} in the 2-periodic three vortex system computed using (5.61) with $N = 10^6$ and $n = 6$ on $(-8, 8) \times (-5.5, 9)$ using a 300x300 grid. Red overlay: partial mixing structure as found in Figure 4.10.

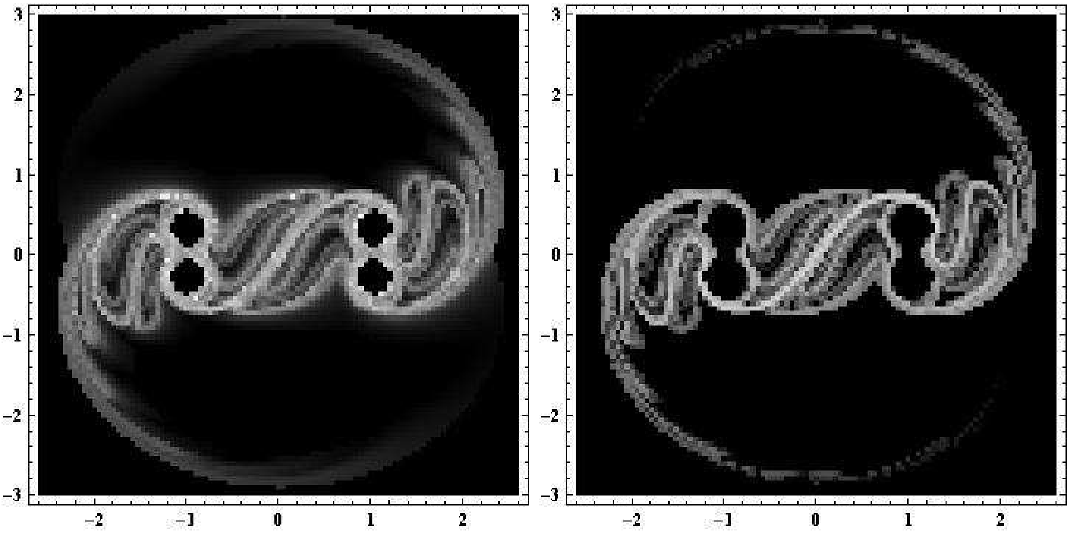

Figure 5.10: Left: w-measure μ_{w+} in the periodic four vortex system computed using (5.61) with $N = 10^6$ and $n = 5$ on $(-2.6, 2.6) \times (3, 3)$ using a 100x100 grid. Right: μ_{w+} as computed from the mixing time distribution as found in Figure 4.11, using (5.48).

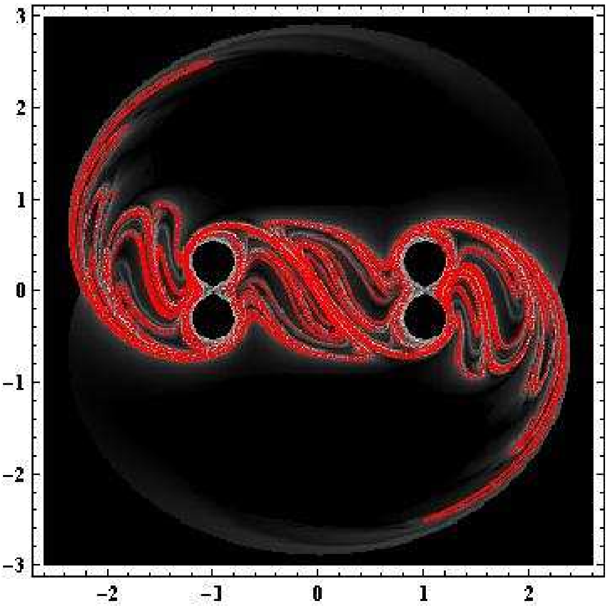

Figure 5.11: Negative time w-measure μ_{w-} in the periodic four vortex system computed using (5.61) with $N = 10^6$ and $n = 5$ on $(-2.6, 2.6) \times (3, 3)$ using a 300x300 grid. Red overlay: partial mixing structure as found in Figure 4.11.

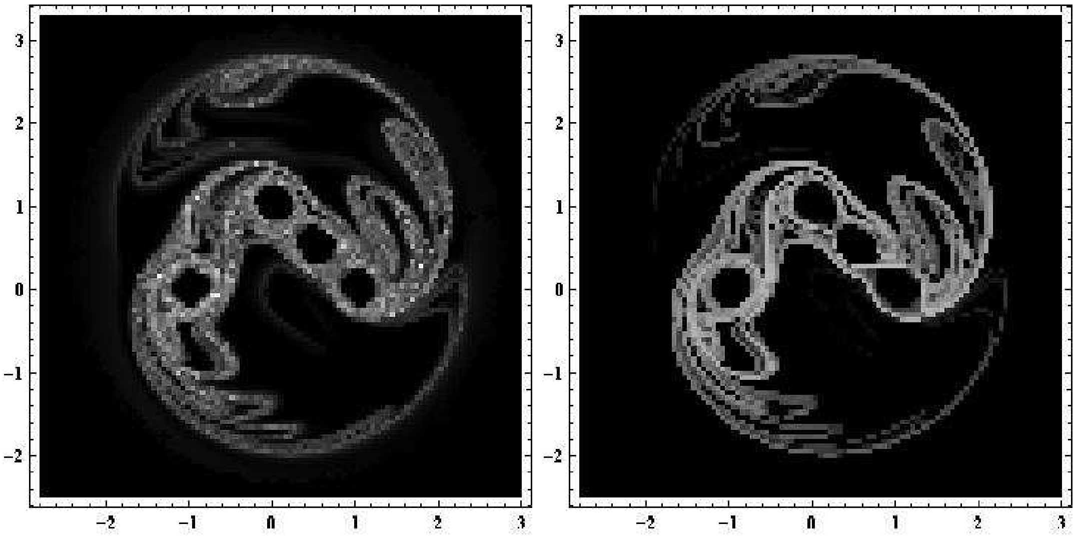

Figure 5.12: Left: w-measure μ_{w+} in the irregular four vortex system computed using (5.44) with $T = 40$ on $(-2.8, 3) \times (-2.5, 3.3)$ using a 100x100 grid, averaging over 16 values for every grid square. Right: μ_{w+} as computed from the mixing time distribution as found in Figure 4.12, using (5.48).

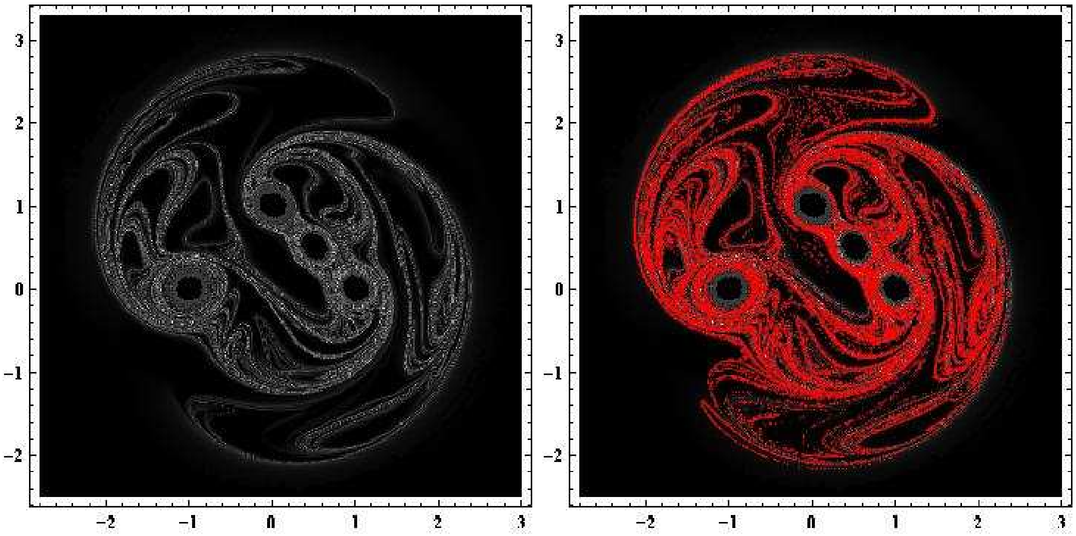

Figure 5.13: Left: Negative time w-measure μ_{w-} in the irregular four vortex system computed using (5.44) with $T = -50$ on $(-2.8, 3) \times (-2.5, 3.3)$ using a 300x300 grid, without averaging inside grid squares. Right: Same, with the red overlay the partial mixing structure as found in Figure 4.12.

6 Invariant manifolds

In the previous chapter, we have seen that the positive and negative time w-measures formalize the concepts of the partial mixing structure and the mixing time distribution respectively. However, actually computing them was a difficult process requiring very long calculations due to slow convergence. In this chapter we will introduce an alternative method based on invariant manifolds. These can be compared to the manifolds of a hyperbolic stagnation point $\mathbf{a}$ in a dynamical system: the stable manifold is then the set of all trajectories that converge to $\mathbf{a}$ in forward time, while the unstable manifold is the set that converges to $\mathbf{a}$ in reverse time. Since our systems do not have points which are stationary for all time, we consider instead the manifolds of *trajectories*. It will turn out that these are non-terminating curves which are dense in the ergodic region. Their spatial distribution is exactly the w-measure.

We will first define invariant manifolds of trajectories and derive several properties (section 6.1.1). We will then limit ourselves to finite segments of these manifolds, and describe how their length changes in time (section 6.1.2). The theory section is concluded with a derivation of the equivalence of the w-measure and the manifold measure, i.e. the spatial distribution of the invariant manifolds (section 6.1.3). Next, the numerical methods used to find invariant manifolds will be described (section 6.2). Finally, several properties of the manifold will be checked using numerically computed invariant manifolds in the 3-periodic three vortex system (section 6.3.1) and the w- and manifold measures for all non-stationary systems are compared (section 6.3.2).

6.1 Theory

6.1.1 Definition and basic properties

We start our discussion from the following definition of invariant manifolds, as found in [60]:

Definition 6.1. *The stable manifold $\mathcal{M}_+(\mathbf{a}, t)$ associated with a point $\mathbf{a} \in \mathbb{R}^2$ at time $t \in \mathbb{R}$ is the set of points of which the trajectory converges to the trajectory of $\mathbf{a}$ asymptotically forward in time, i.e. for $\mathbf{x} \in \mathbb{R}^2$ we have $\mathbf{x} \in \mathcal{M}_+(\mathbf{a}, t)$ if and only if*

$$\lim_{T\to\infty} ||\Phi(t,T)(\mathbf{x}) - \Phi(t,T)(\mathbf{a})|| = 0.$$

The definition of the unstable manifold $\mathcal{M}_-(\mathbf{a},t)$ *is similar, but with the direction of time reversed (i.e. the limit is replaced by* $T \to -\infty$*).*

Note that $\mathbf{a}$ itself is trivially in $\mathcal{M}_+(\mathbf{a},t)$. As usual we will consider only positive time (i.e. stable) manifolds, with the treatment for negative time being similar. Several alternative definitions exist, e.g. those found in [34, 50, 51]; the relationship between some of these and the one we use here is established in [60].

We will now describe various basic properties of invariant manifolds. In the first place, why are they called invariant? This is because they are advected by the flow, as is established by the following Theorem.

Theorem 6.2. *For all* $\mathbf{a} \in \mathbb{R}^2$ *and* $t, S \in \mathbb{R}$*, we have:*

$$\Phi(t,S)\mathcal{M}_+(\mathbf{a},t) = \mathcal{M}_+(\Phi(t,S)(\mathbf{a}), t+S).$$

Proof. Consider any $\mathbf{x} \in \mathcal{M}_+(\mathbf{a},t)$. Then we have:

$$\begin{aligned}&\lim_{T\to\infty} ||\Phi(t+S,T)(\Phi(t,S)\mathbf{x}) - \Phi(t+S,T)(\Phi(t,S)\mathbf{a})|| \overset{(3.8)}{=} \\ &\lim_{T\to\infty} ||\Phi(t,S+T)(\mathbf{x}) - \Phi(t,S+T)(\mathbf{a})|| = \lim_{T\to\infty} ||\Phi(t,T)(\mathbf{x}) - \Phi(t,T)(\mathbf{a})|| \overset{\text{D6.1}}{=} 0,\end{aligned}$$

and by Definition 6.1 this gives us

$$\Phi(t,S)\mathbf{x} \in \mathcal{M}_+(\Phi(t,S)(\mathbf{a}), t+S).$$

From the above, we conclude

$$\Phi(t,S)\mathcal{M}_+(\mathbf{a},t) \subset \mathcal{M}_+(\Phi(t,S)(\mathbf{a}), t+S),$$

and by the same argument

$$\begin{aligned}&\mathcal{M}_+(\Phi(t,S)(\mathbf{a}), t+S) \overset{(3.9)}{=} \Phi(t,S)\Phi(t+S,-S)\mathcal{M}_+(\Phi(t,S)(\mathbf{a}), t+S) \subset \\ &\Phi(t,S)\mathcal{M}_+(\Phi(t+S,-S)(\Phi(t,S)(\mathbf{a})), t+S-S) \overset{(3.9)}{=} \Phi(t,S)\mathcal{M}_+(\mathbf{a},t),\end{aligned}$$

completing the proof. □

Another important property is that it does not matter which point on the stable manifold we choose; any choice will lead to the same manifold.

Theorem 6.3. *Let* $\mathbf{a} \in \mathbb{R}^2$ *and* $t \in \mathbb{R}$*. Then for all* $\mathbf{b} \in \mathcal{M}_+(\mathbf{a},t)$*, we have:*

$$\mathcal{M}_+(\mathbf{b},t) = \mathcal{M}_+(\mathbf{a},t).$$

Proof. Suppose $\mathbf{x} \in \mathcal{M}_+(\mathbf{a}, t)$. We then have:

$$\begin{aligned}
&\lim_{T\to\infty} ||\Phi(t,T)(\mathbf{x}) - \Phi(t,T)(\mathbf{b})|| = \\
&\lim_{T\to\infty} ||\Phi(t,T)(\mathbf{x}) - \Phi(t,T)(\mathbf{a}) + \Phi(t,T)(\mathbf{a}) - \Phi(t,T)(\mathbf{b})|| \leq \\
&\lim_{T\to\infty} ||\Phi(t,T)(\mathbf{x}) - \Phi(t,T)(\mathbf{a})|| + ||\Phi(t,T)(\mathbf{b}) - \Phi(t,T)(\mathbf{a})|| = 0,
\end{aligned}$$

where the last equality follows because both $\mathbf{x}$ and $\mathbf{b}$ are $\mathcal{M}_+(\mathbf{a}, t)$. So $\mathbf{x} \in \mathcal{M}_+(\mathbf{b}, t)$.

From the above, we have

$$\mathcal{M}_+(\mathbf{a}, t) \subset \mathcal{M}_+(\mathbf{b}, t).$$

In particular, $\mathbf{a} \in \mathcal{M}_+(\mathbf{b}, t)$, so we can reverse the above argument to show

$$\mathcal{M}_+(\mathbf{b}, t) \subset \mathcal{M}_+(\mathbf{a}, t),$$

completing the proof. □

So, what kind of sets are these manifolds? We will answer this question by considering *local* properties, based on the Lyapunov exponents. First, suppose $\mathbf{a}$ has a vanishing (positive time) MLE. This will typically be the case for most points in stationary vortex systems or the islands in periodic vortex systems. This means that locally there is no direction in which nearby trajectories are asymptotically attracted, so no nearby points are in $\mathcal{M}_+(\mathbf{a}, t)$. Any other trajectories that asymptotically approach that of $\mathbf{a}$ would eventually have to enter the linear regime. However, the MLE is constant along a trajectory, so these points cannot continue approaching $\mathbf{a}$ for $T \to \infty$. So if the MLE of $\mathbf{a}$ vanishes at time t, then

$$\mathcal{M}_+(\mathbf{a}, t) = \mathbf{a}. \tag{6.1}$$

Note that this does not mean that non-trivial invariant manifolds do not exist in stationary systems. Although most points have zero MLE, there are exceptions, namely stagnation points and any points on their invariant manifolds. A detailed description of these manifolds can be found in [60].

Now suppose $\mathbf{a}$ has positive MLE, as is the case in an ergodic region or a stagnation point in a stationary system. This implies that locally there is a direction along which trajectories are attracted to that of $\mathbf{a}$, namely the stable Lyapunov direction $\mathbf{e}_{2+}$. All other directions are repelled, so the stable manifold is locally a line along $\mathbf{e}_{2+}$. This property holds not just at $\mathbf{a}$, but everywhere along the manifold due to Theorem 6.3. This implies that the global stable manifold is a continuous curve [50]. Using this property we can also alternatively define $\mathcal{M}_+(\mathbf{a}, t)$ as the solution $\boldsymbol{\zeta}(\theta)$ of the ODE [28]:

$$\begin{aligned} \frac{\mathrm{d}\boldsymbol{\zeta}}{\mathrm{d}\theta} &= \mathbf{e}_{2+}(\boldsymbol{\zeta}, t), \\ \boldsymbol{\zeta}(0) &= \mathbf{a}, \end{aligned} \tag{6.2}$$

for $\theta \in \mathbb{R}$. It is important to understand that θ is not a time parameter; the above equation defines the entire manifold at time t. Instead, it is an arc length parameter.

A consequence of the above is that stable manifolds cannot intersect themselves or each other transversally, as this would mean the intersection point would have two attracting directions. There are now two possibilities: either the manifold is a closed curve of finite length, or it has infinite length. Closed manifolds do occur, for instance in stationary systems; we consider only invariant manifolds in ergodic regions from here on. As before, we consider only $t = 0$, allowing us to omit the time argument. Invariant manifolds in ergodic regions are never closed, as is established by the following Theorem in the case of manifolds of generators.

Theorem 6.4. *Let $\mathcal{R}$ be an ergodic region. Then $\mathcal{M}_+(\mathbf{a})$ is not a closed curve for any generator $\mathbf{a}$ of $\mathcal{R}$.*

Proof. Suppose that for some generator $\mathbf{a}$ of $\mathcal{R}$ $\mathcal{M}_+(\mathbf{a})$ is a closed curve and hence has finite length. Note that it cannot cross any boundary of $\mathcal{R}$, since that would imply a trajectory outside $\mathcal{R}$, which will never enter $\mathcal{R}$, approaches the trajectory of $\mathbf{a}$, which will never leave $\mathcal{R}$. So $\mathcal{M}_+(\mathbf{a})$ encloses an open set $\mathcal{A}$ of $\mathcal{R}$ (we do not include $\mathcal{M}_+(\mathbf{a})$ itself in $\mathcal{A}$).

Now, take $n \in \mathbb{N}$ such that $f^n(\mathbf{a}) \in \mathcal{A}$; this is possible by Lemma 4.7 because $\mathbf{a}$ is a generator of $\mathcal{R}$. By Theorem 6.2, we have

$$f^n(\mathcal{M}_+(\mathbf{a})) = \mathcal{M}_+(f^n(\mathbf{a})).$$

By continuity of f, $f^n(\mathcal{M}_+(\mathbf{a}))$ encloses $f^n(\mathcal{A})$, and so $\mathcal{M}_+(f^n(\mathbf{a}))$ encloses $f^n(\mathcal{A})$. Since stable manifolds cannot intersect and $\mathbf{a} \in \mathcal{A}$, $\mathcal{M}_+(f^n(\mathbf{a}))$ is contained in $\mathcal{A}$, and since $\mathbf{a}$ is an internal point of $\mathcal{A}$, we conclude

$$\mu(f^n(\mathcal{A})) < \mu(\mathcal{A}),$$

contradicting symplecticity of f. □

Since invariant manifolds have infinite length it is possible for them to be dense in the ergodic region, and in fact they are. The argument showing this is based on one of the numerical methods for computing manifolds, and will be presented in section

6.2.3, after this method has been introduced. We have already seen that it does not matter which point on the manifold we pick to define it (Theorem 6.3). Now that we know that manifolds are dense in the phase space, we can go even further: it does not matter which point in the ergodic region we pick to define the stable manifold. After all, suppose we have two manifolds $\mathcal{M}_+(\mathbf{a})$ and $\mathcal{M}_+(\mathbf{b})$, with $\mathbf{b} \notin \mathcal{M}_+(\mathbf{a})$. For any $\mathbf{x} \in \mathcal{M}_+(\mathbf{a})$, there is $\mathbf{y} \in \mathcal{M}_+(\mathbf{b})$ arbitrarily close to $\mathbf{x}$ and with the direction of the manifolds arbitrarily parallel (due to continuity of $\mathbf{e}_{2+}$). This means that the manifolds are in a sense equivalent and it makes sense to speak of the global invariant manifold of the ergodic region, without specifying any single point it is based on. This argument is considered in somewhat more detail in [29]. Note that this also means that, at least in this system, there is no need to find distinguished hyperbolic trajectories as defined in [39] to obtain invariant manifolds.

6.1.2 Manifold segments

Since curves of infinite length don't lend themselves well to numerical analysis, we need to restrict ourselves to finite segments of the invariant manifold. We simply define a manifold segment of length L $\mathcal{M}_+(\mathbf{a}, t; L)$ as the solution of (6.2) for $0 \leq \theta \leq L$ (or $L \leq \theta \leq 0$ for $L < 0$). Note that there is some ambiguity at this point since the orientation of $\mathbf{e}_{2+}$ is arbitrary, but as L goes to infinity both sides will tend to the same global invariant manifold anyway.

Three important properties of any invariant manifold $\mathcal{M}_+(\mathbf{a}, t)$ in an ergodic region were discussed in the previous section: it is dense in the ergodic region, it does not matter if we define it based on any other point $\mathbf{b} \in \mathcal{M}_+(\mathbf{a}, t)$ (Theorem 6.3) and it is advected by the flow (Theorem 6.2). We will now examine if and how these properties apply to manifold segments. Obviously, the manifold will not be dense in the ergodic region, since it has finite length. Although it fills the ergodic region arbitrarily well in the limit for $L \rightarrow \infty$, it does not do so for any finite L.

Now assume $\mathbf{b} \in \mathcal{M}_+(\mathbf{a}, t; L)$, assuming for simplicity $L > 0$. Since $\mathbf{a}$ is an endpoint of the manifold segment while $\mathbf{b}$ is not in general, the manifold segment of $\mathbf{b}$ $\mathcal{M}_+(\mathbf{b}, t; L)$ will not be equal to that of $\mathbf{a}$. However, it is not difficult to see that the following property does hold:

$$\mathcal{M}_+(\mathbf{a}, t; L) = \mathcal{M}_+(\mathbf{b}, t; -l) \bigcup \mathcal{M}_+(\mathbf{b}, t; L - l) \tag{6.3}$$

where $0 \leq l \leq L$ is the distance along the manifold from $\mathbf{a}$ to $\mathbf{b}$.

Finally, we consider the invariance of manifold segments under advection by the flow. It will turn out to be convenient to use the *unstable* manifold for this purpose, although of course similar results hold for the stable manifold. Since $\mathcal{M}_-(\mathbf{a}, t; L) \subset \mathcal{M}_-(\mathbf{a}, t)$, we have

$$\Phi(t, T)\mathcal{M}_-(\mathbf{a}, t; L) \subset \Phi(t, T)\mathcal{M}_-(\mathbf{a}, t) \overset{\text{T6.2}}{=} \mathcal{M}_-(\Phi(t, T)(\mathbf{a}), t + T).$$

Since $\mathcal{M}_-(\mathbf{a}, t; L)$ is connected and $\Phi(t, T)$ is continuous, $\Phi(t, T)\mathcal{M}_-(\mathbf{a}, t; L)$ is also connected and so it is a segment of the unstable manifold $\mathcal{M}_-(\Phi(t, T)(\mathbf{a}), t + T)$ starting at $\Phi(t, T)(\mathbf{a})$. This means we can write:

$$\Phi(t, T)\mathcal{M}_-(\mathbf{a}, t; L) = \mathcal{M}_-(\Phi(t, T)(\mathbf{a}), t + T; \mathcal{L}(T)), \tag{6.4}$$

with $\mathcal{L} : \mathbb{R} \to \mathbb{R}$ a function specific to $\mathbf{a}$ and t describing the evolution of the length of the manifold segment. Note that

$$\mathcal{L}(0) = L. \tag{6.5}$$

We will now discuss the behavior of $\mathcal{L}$, i.e. we examine how the length of a manifold segment evolves in time.

We will see in a moment that the range of $\mathcal{L}$ is $\mathbb{R}_+$, so which value we take for $L = \mathcal{L}(0)$ is not important. We take $0 < L \ll 1$, i.e. the manifold segment is initially in the linear regime. By the discussion in the previous section, it is then a line segment oriented along the unstable Lyapunov direction $\mathbf{e}_{2-}(\mathbf{a}, t)$. Recalling section 5.1.1, for $T \to -\infty$ this perturbation of the trajectory of $\mathbf{a}$ is damped exponentially at rate $\lambda_{1-} > 0$ and so we have:

$$\lim_{T \to \infty} \mathcal{L}(T) = 0. \tag{6.6}$$

We can also see this directly from Definition 6.1: the points on it converge to the trajectory of $\mathbf{a}$ for $T \to -\infty$. It is important to note that none of this implies that $\mathcal{L}$ decreases monotonically with decreasing T.

Next, consider the case $T > 0$. Initially the manifold will remain in the linear regime and grow exponentially at rate λ_{1+}. Again, this does not mean that $\mathcal{L}$ is a monotonically increasing function; it is certainly possible for the length to decrease with increasing T at some point, but the general trend will be exponential growth. This means that at some point it will leave the linear regime. We will not consider the dynamics in this intermediate regime, but will instead skip ahead to the asymptotic behavior for $T \to \infty$. For this purpose we will give a closed form expression for $\mathcal{L}(T)$. Let us fix T, and suppose we lengthen the initial segment by an infinitesimal amount at $\mathbf{a}$, i.e. add $\mathrm{d}L$ to L. What influence will this have on $\mathcal{L}(T)$? The answer to this question is found directly from Theorem 5.3, recalling that the manifold segment is oriented along $\mathbf{e}_{2-}(\mathbf{a}, t)$ at time t:

$$\begin{aligned} &\mathrm{d}\mathcal{L}(T) = \\ &\mathrm{d}L\sqrt{(\mathbf{e}_1(\mathbf{a}, t, T), \mathbf{e}_{2-}(\mathbf{a}, t))^2 \eta_1(\mathbf{a}, t, T)^{2|T|} + (\mathbf{e}_2(\mathbf{a}, t, T), \mathbf{e}_{2-}(\mathbf{a}, t))^2 \eta_1(\mathbf{a}, t, T)^{-2|T|}}. \end{aligned}$$

Asymptotically for $T \to \infty$, the first term will dominate if

$$(\mathbf{e}_{1+}(\mathbf{a},t), \mathbf{e}_{2-}(\mathbf{a},t)) \neq 0.$$

This is assumed as we have done before (see (5.30)). We then have in the asymptotic case:

$$\begin{aligned} \mathrm{d}\mathcal{L}(T) =& \mathrm{d}L(\mathbf{e}_{1+}(\mathbf{a},t), \mathbf{e}_{2-}(\mathbf{a},t))\eta_1(\mathbf{a},t,T)^{|T|} \overset{(5.9)}{=} \\ & \mathrm{d}L(\mathbf{e}_{1+}(\mathbf{a},t), \mathbf{e}_{2-}(\mathbf{a},t)) \exp(\lambda_1(\mathbf{a},t,T)|T|), \end{aligned}$$

so that:

$$\mathcal{L}(T) = \int_{\mathbf{y}\in\mathcal{M}_-(\mathbf{a},t;L)} (\mathbf{e}_{1+}(\mathbf{y},t), \mathbf{e}_{2-}(\mathbf{y},t)) \exp(\lambda_1(\mathbf{y},t,T)|T|)\mathrm{ds}. \tag{6.7}$$

Due to continuity of $\mathbf{e}_{2-}$ and $\mathbf{e}_{1+}$, we have with L small enough:

$$\frac{1}{2} < \frac{(\mathbf{e}_{1+}(\mathbf{y},t), \mathbf{e}_{2-}(\mathbf{y},t))}{(\mathbf{e}_{1+}(\mathbf{a},t), \mathbf{e}_{2-}(\mathbf{a},t))} < 2,$$

for all $\mathbf{y} \in \mathcal{M}_-(\mathbf{a},t;L)$. This allows us to rewrite (6.7) as:

$$\mathcal{L}(T) = C(T) \int_{\mathbf{y}\in\mathcal{M}_+(\mathbf{a},t;L)} \exp(\lambda_1(\mathbf{y},t,T)|T|)\mathrm{ds}, \tag{6.8}$$

where

$$\alpha < C(T) < \beta,$$

with

$$\begin{aligned} \alpha =& \frac{1}{2}(\mathbf{e}_{1+}(\mathbf{a},t), \mathbf{e}_{2-}(\mathbf{a},t)); \\ \beta =& 2(\mathbf{e}_{1+}(\mathbf{a},t), \mathbf{e}_{2-}(\mathbf{a},t)). \end{aligned}$$

We now define the asymptotic growth rate σ_m of the manifold:

$$\sigma_m := \lim_{T\to\infty} \frac{1}{|T|} \log \mathcal{L}(T). \tag{6.9}$$

Using (6.7) and the fact that $C(T)$ is bounded from above and below by strictly positive constants independent of T, we have:

$$\begin{aligned} \sigma_m =& \lim_{T\to\infty} \frac{1}{|T|} \log C(T) \int_{\mathbf{y}\in\mathcal{M}_-(\mathbf{a},t;L)} \exp(\lambda_1(\mathbf{y},t,T)|T|)\mathrm{ds} = \\ & \lim_{T\to\infty} \frac{1}{|T|} \log \int_{\mathbf{y}\in\mathcal{M}_-(\mathbf{a},t;L)} \exp(\lambda_1(\mathbf{y},t,T)|T|)\mathrm{ds}. \end{aligned} \tag{6.10}$$

We have already seen in section 5.1.2 that this limit is equal to the topological entropy. There is one important difference though: we are not integrating over an area now but over a line segment. Still, it seems reasonable to conclude that

$$\sigma_m = \sigma_+. \tag{6.11}$$

This has already been verified by numerical analysis in section 5.3.1; these results can be found in Table 5.1. This also justifies our use of the term 'topological entropy' for σ_+. Actually, we still have not used the original definition, which involves the rate of information loss in dynamic systems [32]; however, the link between that definition and the manifold segment growth rate is well established [57, 76].

In summary, the length of a segment of the unstable manifold goes to zero exponentially for $T \to -\infty$ with rate λ_{1-} and grows exponentially for $T \to \infty$ with rate σ_+. Note that $\mathcal{L}(T)$ takes on the full range of values in $\mathbb{R}_+$ for $T \in \mathbb{R}$, a fact that was used at the start of this discussion. There is an important difference in uniformity between the two limits for T, already briefly mentioned above. For $T \to -\infty$, the segment will be short enough to be entirely in the linear regime, and the evolution of its length is described by the evolution of the size of a small perturbation. This means, recalling the discussion in section 5.1.2, that although we have

$$\lim_{T\to-\infty} \frac{1}{|T|} \log \mathcal{L}(T) = -\lambda_{1-} \tag{6.12}$$

the limit

$$\lim_{T\to-\infty} \frac{\mathcal{L}(T)}{\exp(-\lambda_{1-}|T|)} \tag{6.13}$$

does *not* converge. Now consider the case $T \to \infty$. As the length of the manifold segment increases, it will approach the global unstable manifold (recall that this manifold is independent of the starting location). This means it will not pass through regimes with different mixing properties anymore, so the limit

$$\lim_{T\to\infty} \frac{\mathcal{L}(T)}{\exp(\sigma_+|T|)} \tag{6.14}$$

does converge.

Finally, we consider the behavior of a material curve contained in the ergodic region, not necessarily a manifold segment, for $T \to \pm\infty$. For $T \to \infty$, any parts of the curve not parallel to the stable Lyapunov direction $\mathbf{e}_{2+}$ will align themselves along the unstable Lyapunov direction $\mathbf{e}_{2-}$ by Theorem 5.8 and so the curve will approach the global unstable manifold. For points where the curve *is* aligned along the stable Lyapunov direction, this property will be conserved at that point by Theorem 5.7 and so the curve cannot approach the unstable manifold. Various types of behavior are now possible; this is explored in detail in [17]. We will limit ourselves to stating here that *almost* all material curves converge to the global unstable manifold for $T \to \infty$. Similarly, for $T \to -\infty$ almost all material curves converge to the global stable manifold.

6.1.3 Relationship with the w-measure

We will now investigate the space filling properties of the global stable manifold more closely. We have already seen that it is dense in the ergodic region, but what does its distribution look like? To frame this question we define the manifold measure μ_{m+} using the spatial distribution of the manifold, i.e. for $\mathcal{A} \subset \mathcal{R}$ $\mu_{m+}(\mathcal{A})$ is the fraction of the global stable manifold in $\mathcal{A}$.

Definition 6.5. *The stable manifold measure* μ_{m+} *is defined by*

$$\mu_{m+}(\mathcal{A}, t) := \lim_{L\to\infty} \frac{\int_{\mathbf{y}\in\mathcal{M}_+(\mathbf{a};L)} \chi(\mathbf{y}, \mathcal{A}) \mathrm{ds}}{\int_{\mathbf{y}\in\mathcal{M}_+(\mathbf{a};L)} \mathrm{ds}} = \lim_{L\to\infty} \frac{1}{L} \int_{\mathbf{y}\in\mathcal{M}_+(\mathbf{a};L)} \chi(\mathbf{y}, \mathcal{A}) \mathrm{ds},$$

for any generator $\mathbf{a}$ *of* $\mathcal{R}$ *and open* $\mathcal{A} \subset \mathcal{R}$*, with* $\chi(\mathbf{y}, \mathcal{A})$ *as defined in (5.59), i.e. equal to one when* $\mathbf{y} \in \mathcal{A}$ *and zero otherwise.*

That the measure is independent of the choice of $\mathbf{a}$ follows from the fact that all manifolds of generators are equivalent, as discussed in section 6.1.1. In the remainder of this section we will argue that the stable manifold measure is equal to the positive time w-measure introduced in section 5.1.2.

A defining aspect of the manifold measure is that it is invariant under the dynamical system defined by (6.2). We can show explicitly that μ_{m+} has this property. Similar to $\Phi(t, T)$, the operator taking a trajectory from time t to time $t + T$ in the ODE (3.1), define $w(\theta) : \mathcal{R} \to \mathcal{R}$ as the operator taking an initial condition to θ in (6.2). Note that what this operator does is just move a point a distance θ along its stable manifold. Then the invariance of μ_{m+} simply means that

$$\mu_{m+}(w^{-1}(\theta)\mathcal{A}) = \mu_{m+}(\mathcal{A}), \tag{6.15}$$

for all open $\mathcal{A} \subset \mathcal{R}$ and all $\theta \in \mathbb{R}$. Noting that applying $w(\theta)$ to a manifold just means shifting it over a distance θ, we have:

$$\begin{aligned}\mu_{m+}(w^{-1}(\theta)\mathcal{A}) = \lim_{L\to\infty}\frac{1}{L}\int_{\mathbf{y}\in\mathcal{M}_+(\mathbf{a};L)} \chi(\mathbf{y}, w^{-1}(\theta)\mathcal{A})\mathrm{ds} = \\ \lim_{L\to\infty}\frac{1}{L}\int_{\mathbf{y}\in\mathcal{M}_+(\mathbf{a};L)} \chi(w(\theta)(\mathbf{y}), \mathcal{A})\mathrm{ds} = \\ \lim_{L\to\infty}\frac{1}{L}\int_{\mathbf{y}\in w(\theta)\mathcal{M}_+(\mathbf{a};L)} \chi(\mathbf{y}, \mathcal{A})\mathrm{ds} = \\ \lim_{L\to\infty}\frac{1}{L}\int_{\mathbf{y}\in\mathcal{M}_+(w(\theta)(\mathbf{a});L)} \chi(\mathbf{y}, \mathcal{A})\mathrm{ds} = \mu_{m+}(\mathcal{A}),\end{aligned}$$

showing that (6.15) holds.

So, we need to show that (6.15) holds for the w-measure μ_{w+}. Because (6.2) is autonomous, it suffices to show that

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\mu_{w+}(w^{-1}(\theta)\mathcal{A})\bigg|_{\theta=0} = 0, \tag{6.16}$$

for all open $\mathcal{A}\subset\mathcal{R}$. Notice that this expression is simply the flux of μ_{w+} traveling into $\mathcal{A}$ under the w-system (6.2)[1]. Unfortunately we cannot compute this directly since μ_{w+} is discontinuous, so we will instead show that it holds in the limit using the finite time w-measure $\mu_w(T)$, i.e. that

$$\lim_{T\to\infty} R(T) := \lim_{T\to\infty}\frac{\mathrm{d}}{\mathrm{d}\theta}\mu_w(w^{-1}(\theta)\mathcal{A}, T)\bigg|_{\theta=0} = 0, \tag{6.17}$$

for all open $\mathcal{A}\subset\mathcal{R}$. We use the fact that the finite time w-measure is continuous for all $T\in\mathbb{R}$, i.e. there is a density function $\rho_w(T):\mathcal{R}\to\mathbb{R}$ such that:

$$\mu_w(\mathcal{A}, T) = \int_{\mathbf{y}\in\mathcal{A}} \rho_w(\mathbf{y}; T)\mathrm{dA}, \tag{6.18}$$

for all $\mathcal{A}\subset\mathcal{R}$. Explicitly this density function is given by (see (5.45)):

$$\rho_w(\mathbf{x}; T) = \frac{\exp(\lambda_1(\mathbf{x}, T)|T|)}{\int_{\mathbf{y}\in\mathcal{R}}\exp(\lambda_1(\mathbf{y}, T)|T|)\mathrm{dA}}. \tag{6.19}$$

We will now focus on estimating (6.17) for fixed $T > 0$. As mentioned this is the flux of μ_w flowing into $\mathcal{A}$ under the w-system, so we have writing $\partial\mathcal{A}$ for the boundary of $\mathcal{A}$ and $\mathbf{n}$ for the outward-pointing normal:

$$\begin{aligned}|R(T)| = \left|\frac{\mathrm{d}}{\mathrm{d}\theta}\mu_w(w^{-1}(\theta)\mathcal{A}, T)\bigg|_{\theta=0}\right| = \left|\int_{\mathbf{y}\in\partial\mathcal{A}} \rho_w(\mathbf{y}; T)(\mathbf{e}_{2+}(\mathbf{y})\cdot\mathbf{n})\mathrm{ds}\right| = \\ \left|\int_{\mathbf{y}\in\mathcal{A}} \nabla\cdot(\rho_w(\mathbf{y}; T)\mathbf{e}_{2+}(\mathbf{y}))\mathrm{dA}\right|,\end{aligned}$$

---

[1] The term w-system for (6.2) is taken from [2].

using Gauss' divergence theorem. Continuing:

$$
\begin{aligned}
|R(T)| = &\left|\int_{\mathbf{y}\in\mathcal{A}} \nabla\cdot(\rho_w(\mathbf{y};T)\mathbf{e}_{2+}(\mathbf{y}))\mathrm{dA}\right| \leq \int_{\mathbf{y}\in\mathcal{A}} |\nabla\cdot\rho_w(\mathbf{y};T)\mathbf{e}_{2+}(\mathbf{y})|\mathrm{dA} = \\
&\int_{\mathbf{y}\in\mathcal{A}} |\rho_w(\mathbf{y};T)\nabla\cdot\mathbf{e}_{2+}(\mathbf{y}) + \mathbf{e}_{2+}(\mathbf{y})\cdot\nabla\rho_w(\mathbf{y};T|\mathrm{dA} \overset{(6.19)}{=} \\
&\frac{\int_{\mathbf{y}\in\mathcal{A}} |\exp(\lambda_1(\mathbf{y},T)|T|)\nabla\cdot\mathbf{e}_{2+}(\mathbf{y}) + \mathbf{e}_{2+}(\mathbf{y})\cdot\nabla\exp(\lambda_1(\mathbf{y},T)|T|)|\mathrm{dA}}{\int_{\mathbf{y}\in\mathcal{R}}\exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA}} = \\
&\frac{\int_{\mathbf{y}\in\mathcal{A}} \exp(\lambda_1(\mathbf{y},T)|T|)|\nabla\cdot\mathbf{e}_{2+}(\mathbf{y}) + \mathbf{e}_{2+}(\mathbf{y})\cdot\nabla(\lambda_1(\mathbf{y},T)|T|)|\mathrm{dA}}{\int_{\mathbf{y}\in\mathcal{R}}\exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA}} \leq \\
&\frac{\int_{\mathbf{y}\in\mathcal{A}}\exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA}}{\int_{\mathbf{y}\in\mathcal{R}}\exp(\lambda_1(\mathbf{y},T)|T|)\mathrm{dA}} \int_{\mathbf{y}\in\mathcal{A}} |\nabla\cdot\mathbf{e}_{2+}(\mathbf{y}) + \mathbf{e}_{2+}(\mathbf{y})\cdot\nabla(\lambda_1(\mathbf{y},T)|T|)|\mathrm{dA} = \\
&\mu_w(\mathcal{A},T)\int_{\mathbf{y}\in\mathcal{A}} |\nabla\cdot\mathbf{e}_{2+}(\mathbf{y}) + \mathbf{e}_{2+}(\mathbf{y})\cdot\nabla(\lambda_1(\mathbf{y},T)|T|)|\mathrm{dA}.
\end{aligned}
$$

To proceed, we need to estimate the integrand. This is simple using the following result by Thiffeault [72]:

$$
\lim_{T\to\infty}(\nabla\cdot\mathbf{e}_{2+}(\mathbf{y}) + \mathbf{e}_{2+}(\mathbf{y})\cdot\nabla(\lambda_1(\mathbf{y},T)T)) = 0. \tag{6.20}
$$

This is actually a very powerful constraint on the Lyapunov exponent and it has several interesting consequences for the behavior of material lines advected by the flow [49, 71]. We use it to complete the derivation of (6.17):

$$
\begin{aligned}
\lim_{T\to\infty}|R(T)| \leq &\lim_{T\to\infty}\mu_w(\mathcal{A},T)\int_{\mathbf{y}\in\mathcal{A}} |\nabla\cdot\mathbf{e}_{2+}(\mathbf{y}) + \mathbf{e}_{2+}(\mathbf{y})\cdot\nabla(\lambda_1(\mathbf{y},T)|T|)|\mathrm{dA} = \\
&\mu_{w+}(\mathcal{A})\int_{\mathbf{y}\in\mathcal{A}} \left|\lim_{T\to\infty}(\nabla\cdot\mathbf{e}_{2+}(\mathbf{y}) + \mathbf{e}_{2+}(\mathbf{y})\cdot\nabla(\lambda_1(\mathbf{y},T)T))\right|\mathrm{dA} = 0.
\end{aligned}
$$

This completes the argument that the manifold measure is equal to the w-measure:

$$
\mu_{m+} = \mu_{w+}. \tag{6.21}
$$

So the stable manifold is concentrated in areas where the forward time finite time Lyapunov exponent is high. Similarly, the unstable manifold is concentrated in areas where the reverse time FTLE is high. It is worth noting that the above argument also applies to irregular systems, not just periodic ones. Note that there is some similarity between this method and the approach of Lagrangian coherent structures (LCS) [69], which identifies *ridges* of the FTLE field with approximations of the global invariant manifold.

6.2 Numerical methods

We will focus on computing the unstable manifold in this section; computing the stable manifold is of course similar, with the direction of time reversed. Two definitions of the invariant manifold are given in section 6.1.1. These are Definition 6.1, based on the attracting properties, and the ODE (6.2), based on the alignment along the unstable Lyapunov direction. Both of these definitions give rise to a numerical method for computing the manifold.

Using (6.2) is the simplest of these two. We already know how to approximate $\mathbf{e}_{2-}(\mathbf{y}, t)$, namely as the limit for $T \to -\infty$ of the minimal eigenvector $\mathbf{e}_2(\mathbf{y}, t, T)$ of $\Lambda(\mathbf{y}, t, T)$. How to compute this matrix is described in section 5.2.1. All that is then necessary is to apply a time integration method for θ. We use the fourth order classical Runge-Kutta formula, a description of which can be found in [52]. The 'timestep' $\Delta\theta$ is dynamically chosen to ensure sufficient resolution in areas of high curvature. We will refer to this method as the ODE solving method. It has three important parameters: the maximum timestep $\Delta\theta$, the maximum rotation per timestep α and the time used for determining the unstable Lyapunov direction T (see section 5.2.1). Although this method does find invariant manifolds, it has a very serious drawback: due to small errors in the computation of the unstable Lyapunov direction and/or the time integration, it is possible for the manifold to leave the ergodic region. This behavior is highly unphysical and ruins the important property of the global unstable manifold being dense in the ergodic region.

The underlying problem of the ODE solving method is that it does not make use of the inherent stability of the invariant manifold. We have seen in section 6.1.2 that (almost) any material curve within the ergodic region will converge to the unstable manifold for forward time, which indicates that if we have small errors in the computed manifold these will be damped under forward time evolution. We would like our numerical method to take advantage of this property. This leads to the advection method for finding invariant manifolds.

Suppose we are looking for $\mathcal{M}_-(\mathbf{a}, 0; L)$, i.e. the unstable manifold segment of $\mathbf{a}$ at $t = 0$ with length L. Recall the discussion in section 6.1.2: a small segment of the unstable manifold will grow in length to infinity for increasing T while remaining a manifold segment. So, to obtain the manifold at $t = 0$ we first go backwards in time to $-T$ for some properly chosen[2] T, following the trajectory of $\mathbf{a}$ to obtain $\Phi(0, -T)(\mathbf{a})$. We now define $\mathcal{C}(-T)$ as the line segment connection $\mathbf{a}$ and $\mathbf{a} + (L_0, 0)$ for small L_0, typically using $L_0 = 10^{-4}$. Computing $\Phi(-T, T)\mathcal{C}(-T)$ then gives an

---

[2] This choice does not affect the final manifold, but if it is taken too small the manifold segment found may not have the desired length, while if it is taken too large more time and memory is used by the computation.

approximation for the unstable manifold. What happens is that initially (i.e. at t close to $-T$) $\mathcal{C}(t)$ remains short enough to be in the linear regime and aligns along the unstable Lyapunov direction. Upon leaving the linear regime it is then an actual unstable manifold segment which will continue to grow up to $t = 0$. There are two issues to tackle before we can actually apply this method: how to efficiently compute the image of $\mathcal{C}(-T)$ under $\Phi(-T, T)$, i.e. how to advect a continuous curve, and how to ensure the final manifold segment $\mathcal{C}(0)$ has the desired length L. We will discuss both of these problems in the following subsections.

6.2.1 Advecting a material line

We first describe how to compute the advection of a material line within the flow. This is based on the description found in [51]. We write $T = N\Delta t$ with Δt the timestep and $N \in \mathbb{N}$. This may require slightly adjusting T. The curve $\mathcal{C}(-T + n\Delta t)$ is approximated for $0 \leq n \leq N$ by $M(n)$ points $(\mathbf{x}_{n,j} | 1 \leq j \leq M(n))$. The trick is to make sure we have enough points to have the curve well resolved. The simplest method is of course to ensure we have enough points initially, i.e. represent $\mathcal{C}(-T)$, the initial line segment, by a huge number of points $M(0)$ and simply advect them all individually. However, as we will see the growth of the manifold segment is very irregular, meaning that to resolve the manifold as a whole several areas will have far more points than necessary. It is much more efficient to dynamically add points when the manifold is not well resolved, and to remove them when possible. This raises two questions: how do we determine if the curve is well resolved, and how do we add points if we find this is necessary?

Several possible criteria for determining if a curve is well resolved are discussed in [51]. These were tested, with the best results being found by using the Hobson variant criterion [37] with parameters chosen as described below. This criterion stipulates three properties which every interval between two points should satisfy. In the first place, two consecutive points should never be too far apart, since if they are we might miss the development of features in between these points:

$$||\mathbf{x}_j - \mathbf{x}_{j+1}|| \leq \tau, \tag{6.22}$$

with $\tau = 0.05$ being a typical value. Note that we omit the subscript n for notational convenience. The second property limits the amount of rotation the manifold is allowed to undergo between consecutive points. This ensures areas of high curvature are well resolved. Explicitly, we require:

$$\sin^{-1} \frac{||\overline{\mathbf{x}}_{j-1} - \mathbf{x}_{j-1}||}{2||\mathbf{x}_j - \mathbf{x}_{j-1}||} + \sin^{-1} \frac{||\overline{\mathbf{x}}_j - \mathbf{x}_j||}{2||\mathbf{x}_{j+1} - \mathbf{x}_j||} \leq \alpha, \tag{6.23}$$

where

$$\overline{\mathbf{x}}_j = \mathbf{x}_{j+1} + \frac{\mathbf{x}_{j+1} - \mathbf{x}_{j+2}}{||\mathbf{x}_{j+1} - \mathbf{x}_{j+2}||} ||\mathbf{x}_{j+1} - \mathbf{x}_j||.$$

A typical value is $\alpha = \pi/25$, meaning that a small circle would require 50 points to be considered well resolved. The final requirement is that the distance between two points should never be too small. If it is, then numerical issues might induce artificial curvature. This would then lead to the need to add more points by the second property, which would be even closer together, thus destabilizing the method. So, the third requirement is:

$$||\mathbf{x}_j - \mathbf{x}_{j+1}|| \geq \delta, \tag{6.24}$$

with $\delta = 10^{-7}$ being a typical value. So we consider the curve $\mathcal{C}(-T + n\Delta t)$ well resolved if it satisfies (6.22), (6.23) and (6.24) for all j with $2 \leq j \leq M(n) - 1$.[3]

So suppose we have found a well resolved representation of $\mathcal{C}(-T + n\Delta t)$. We then need to find one of the curve at the next timestep $\mathcal{C}(-T + (n + 1)\Delta t)$. Recall that $\mathcal{C}(-T + n\Delta t)$ is defined by the sequence of points $(\mathbf{x}_{n,j} | 1 \leq j \leq M(n))$. The first step is then to advect these points by one timestep:

$$\mathbf{x}'_{n+1,j} = \Phi(-T + n\Delta t, \Delta t)(\mathbf{x}_{n,j}). \tag{6.25}$$

We then check whether the sequence $(\mathbf{x}'_{n+1,j} | 1 \leq j \leq M(n))$ satisfies the Hobson variant criterion described above. If it is not satisfied by some interval $(\mathbf{x}'_{n+1,j}, \mathbf{x}'_{n+1,j+1})$, then we need to add a point on this interval. We could base this on interpolation in the sequence of image points $(\mathbf{x}'_{n+1,j})$, but by assumption the curve $\mathcal{C}(-T+(n+1)\Delta t)$ is not well resolved by this sequence, meaning that this interpolation cannot be accurate. However, also by assumption the curve $\mathcal{C}(-T + n\Delta t)$ *is* well resolved by the sequence $(\mathbf{x}_{n,j})$. So we find a point $\mathbf{x}_{n,j+\frac{1}{2}}$ by interpolation between $\mathbf{x}_{n,j}$ and $\mathbf{x}_{n,j+1}$ and determine $\mathbf{x}'_{n+1,j+\frac{1}{2}}$ using (6.25). We repeat this until $\mathcal{C}(-T + (n + 1)\Delta t)$ is well resolved, and then remove any points if that removal does not violate (6.22) through (6.24). The reason to delete redundant points is to avoid unnecessary computations. We then obtain the sequence $(\mathbf{x}_{n+1,j})$ simply by renumbering $(\mathbf{x}'_{n+1,j})$ to get rid of any non-integer indices.

One question now remains: just how do we determine where to place $\mathbf{x}_{n,j+\frac{1}{2}}$ by interpolation, i.e. how do we parameterize $\mathcal{C}(-T + n\Delta t)$ between points? This is an important issue; our numerical advection method is accurate enough to ensure that if $\mathbf{x}_{n,j}$ is in the ergodic region then so is the image point $\mathbf{x}'_{n+1,j}$. However, if the interpolation method is not accurate enough it can cause the insertion of points outside the ergodic region, leading to unphysical behavior of the computed manifold segment.

---

[3] The rotation requirement (6.23) cannot be applied at the endpoints. These points require a slightly altered form, which we will not discuss here.

Again, several possible methods are discussed in [51]. In our case Dritschel's curvature method [19] worked best. The formulas do not contain any extra parameters and are easily implemented, but are rather complex and so will not be reproduced here.

6.2.2 Controlling the length of the manifold segment

We now turn to the second issue mentioned earlier: how do we ensure that the manifold segment found at $t = 0$, $\mathcal{C}(0)$, has the desired length L? We have two parameters we can alter to achieve this: the integration time T and the initial length of $\mathcal{C}(-T)$ L_0. Of these L_0 seems to be the easiest to adjust. We simply compute $\mathcal{C}(0)$ for some T and L_0. Say its length L_{final} is ten times too small; we then simply repeat the calculation taking L_0 ten times as large. This does assume that L_{final} is approximately linearly proportional to L_0. Unfortunately, this assumption does not hold, as is shown in Figure 6.1. Adjusting T leads to similar issues (Figure 6.2).

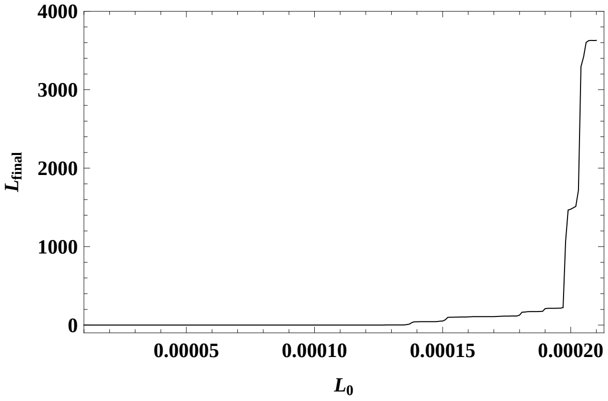

Figure 6.1: Length L_{final} of the unstable manifold segment obtained using the advection method in the 3-periodic three vortex system as a function of L_0, with $\mathbf{a} = (0, 0)$ and $T = 300$.

So an alternative method not based on adjusting L_0 and T is necessary. We will describe one based on the growth properties of the manifold segment as described in section 6.1.2, i.e. irregular growth for small lengths based on the MLE and regular growth for large lengths based on the topological entropy. Define L_{σ_+} as the length above which growth is dominated by the topological entropy; we will assume

$$L > L_{\sigma_+}. \tag{6.26}$$

Note that this means that if the manifold segment has at least length L at any time, it will grow with every timestep. This suggests a simple approach: simply cap the

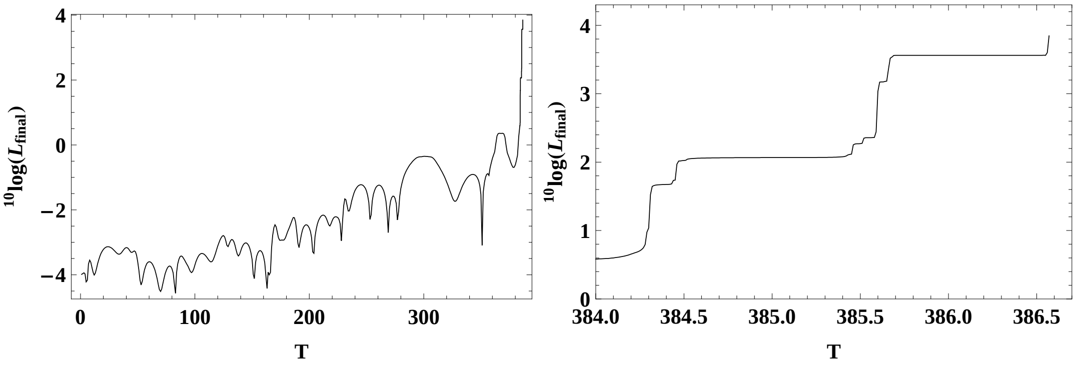

Figure 6.2: Length L_{final} of the unstable manifold segment obtained using the advection method in the 3-periodic three vortex system as a function of T, with $\mathbf{a} = (0, 0)$ and $L_0 = 10^{-4}$. Right: detail of the area around $T = 385$.

length of the manifold segment at L, removing points as necessary. A disadvantage of this method is that if T is very large, the length L will be reached long before $t = 0$ and we have to advect a manifold segment of length L over many timesteps. As a result, the complexity is $\mathcal{O}(LT)$, preventing us from simply taking T very large. However, this is a working reference technique for find a manifold segment of length L.

The method used in practice has two key improvements. One is an adaptive rescaling technique, which we will not discuss here. It reduces the complexity to $\mathcal{O}(L \log T)$. The second improvement is based on the fact that once the length of the manifold segment is at least L_{σ_+}, we can predict its growth. Explicitly, if for

$$T_c = \frac{1}{\sigma_+} \log \frac{L}{L_{\sigma_+}} \tag{6.27}$$

the manifold segment $\mathcal{C}(-T_c)$ has length L_{σ_+}, then the final manifold segment $\mathcal{C}(0)$ will have length approximately L. We can ensure it has length at least L by taking T_c slightly larger than in (6.27). The method now consists of computing T_c, finding $\mathcal{M}_-(\Phi(0, -T_c)(\mathbf{a}), -T_c; L_{\sigma_+})$ using the basic method, advecting this segment to $t = 0$ and removing points beyond length L. The complexity of this approach is $\mathcal{O}(L_{\sigma_+} \log T + L)$.

6.2.3 Further notes

Unless stated otherwise, we always use the parameters as given above (i.e. $L_0 = 10^{-4}$, $\alpha = \pi/25$, $\tau = 0.05$ and $\delta = 10^{-7}$) when applying the advection method. T is simply chosen large enough to find a manifold segment of length L; its actual value does not affect the result as long as this constraint is satisfied. The value of L_{σ_+} also does not affect the result, although it may not reach the desired length L if L_{σ_+} is taken too small. A good value is $L_{\sigma_+} = 500$.

The advection method turns out to be much faster and more reliable than the ODE solving method for finding invariant manifolds. It is not however as closely related to the definition of the manifold, and so it is worthwhile to check that both methods give identical results. This is shown in Figure 6.3. Note that the manifolds found using both methods eventually diverge from each other; however, the points at which this happens depends on the parameters used in the ODE solving method, indicating that this is a result of inaccuracies in this method, and not of any in the advection method. In fact, reducing the parameters for the advection method to yield a more accurate manifold does not produce any visible changes (not pictured).

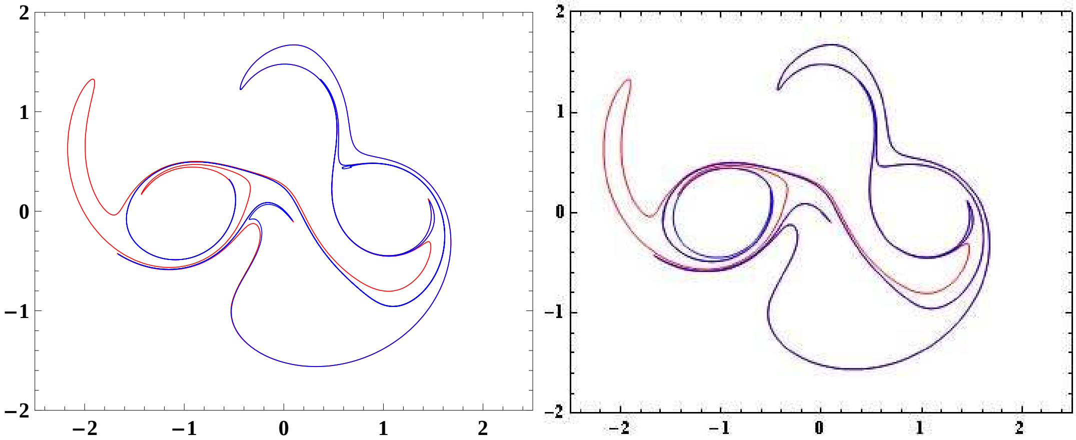

Figure 6.3: Comparison of the advection method and the ODE solving method for finding the invariant manifold. Red, both sides: unstable manifold segment of $(0, 0)$ with L=35 in the 3-periodic three vortex system, found using the advection method. Blue, left: Found using the ODE solving method with $\Delta\theta = 0.01$, $\alpha = \pi/125$ and $T = 200$. Blue, right: Found using the ODE solving method with $\Delta\theta = 0.001$, $\alpha = \pi/1250$ and $T = 400$.

We conclude this section by presenting a further argument, based on considering the manifold as the result of the advection method, for the relationship between the partial mixing structure and the negative time w-measure, which is related to the unstable manifold by section 6.1.3. Imagine a small disk around some point **a** in the ergodic region at $t = -T$ for some sufficiently large T. We want to track the advection of this disk to $t = 0$. There are two approaches: either place a large amount of particles in the disk and advect these to $t = 0$, or follow the evolution of the boundary of the disk. The first method will reveal the partial mixing structure, as discussed in section 4.3.2 (see Figure 4.7 in particular). Eventually the particles will spread out over the full ergodic region, but assuming we have taken enough points the mixing structure will still be visible at $t = 0$. So, the image of the disk under $\Phi(-T, T)$ reveals the partial mixing structure.

The second method mentioned is to follow the evolution of the boundary of the disk, a contour kinematical approach. Numerically this is achieved using the methods described in section 6.2.1. It is basically identical to the advection method if the initial disk is small enough-after all, the circle will be deformed by Lyapunov stretching into two line segments joined by a sharp bend[4]. So the unstable manifold is also the (boundary of the) image of the disk under $\Phi(-T, T)$, which by the discussion above reveals the partial mixing structure. This again establishes the link between the unstable manifold and negative time w-measure on the one hand and the partial mixing structure on the other hand.

The above argument can also be used to show that the global unstable manifold is dense in the ergodic region, a result that has already been used throughout this chapter: the image of the disk under $\Phi(-T, T)$ is asymptotically dense in the ergodic region by Theorem 4.12, so the same applies to its contour which is asymptotically the global unstable manifold.

6.3 Results

6.3.1 Properties of the manifold

In this section we will examine several properties of invariant manifolds discussed earlier. Results will be shown only for the 3-periodic three vortex system, but they hold for all other systems as well, including the irregular system. All manifolds are computed using the advection method.

First things first: let us have a look at the actual manifolds. The stable manifold for the 3-periodic three vortex system is shown in Figure 6.4. The stable and unstable manifold are shown together in Figure 6.5, which includes some detailed pictures of parts of the manifolds. Note that the unstable manifold is the mirror image of the stable manifold in this case, which is again caused by the fact that mirroring the system is equivalent to reversing the direction of time (see section 4.3). In both figures,

---

[4] This approach is actually used in [60] to compute invariant manifolds. It has the advantage that the contour is closed and so has an area which should remain constant by symplecticity, allowing us to dynamically check the accuracy of the numerical method. The disadvantage is the increased computational work since we are basically computing the entire manifold twice, as there are two initial line segments.

several very sharp bends in the manifolds are visible. It is important to note that these are *not* singularities; although the curvature is very high it is still finite. This is a result of the continuity of the asymptotic Lyapunov directions. Another important feature is that all intersections of the stable and unstable manifolds with each other are transversal, justifying the assumption, first encountered in Theorem 5.8, that the asymptotic Lyapunov directions are nowhere identical.

We have stated in section 6.1.1 that the global stable manifold is independent of the point chosen as the starting location. This property is shown in Figure 6.6. Note that although both manifolds are not completely identical, they share the same global structure. However, there is another problem: the manifold should be dense in the ergodic region. Even though we are considering segments that are quite long, up to $L = 20.000$ in Figure 6.4, they leave large areas of the ergodic region uncovered. This is a result of the irregularity of the w-measure, causing the manifold, which is distributed according to this measure, to cluster heavily in certain areas. That more and more areas are filled as L increases is shown in figure 6.8.

Finally, we check the growth properties of the invariant manifold, as described in section 6.1.2. To this end, $\mathcal{L}(T)$ is shown as a function of T in Figure 6.7. Recall that $\mathcal{L}(T)$ is the length of the manifold segment $\Phi(0,T)\mathcal{M}_+(\mathbf{a},0;L_0)$ for specified $\mathbf{a} \in \mathcal{R}$ and $L_0 \in \mathbb{R}$. The behavior described in the theory is clearly visible: irregular, non-monotonic exponential growth with rate λ_{1+}, the maximum Lyapunov exponent, in the linear regime, followed by regular, monotonic exponential growth with rate σ_+, the topological entropy, in the asymptotic regime.

6.3.2 Comparing the w-measure and the manifold measure

In this section we will demonstrate the equivalence of the w-measure and the manifold measure by computing the manifold measure for all systems and comparing this to the w-measure as found in section 5.3.2. Given a manifold segment of sufficient length L, the manifold measure is found directly from Definition 6.5. As in the previous chapter, all results use logarithmic color scales so that the features in area of low w-measure are visible.

Results for the 3-periodic three vortex system are shown in Figure 6.9, using the same color scaling as in section 5.3.2 for both pictures. Note that the areas of high w-measure are identical on both sides, but that the other features present are not visible in the manifold measure. This is easily explained: the manifold is not long enough yet to have passed through these areas. Unfortunately it is not possible to increase L further due to memory constraints. To highlight the similarities, the same results are shown in Figure 6.10, with the color scaling chosen to hide the features

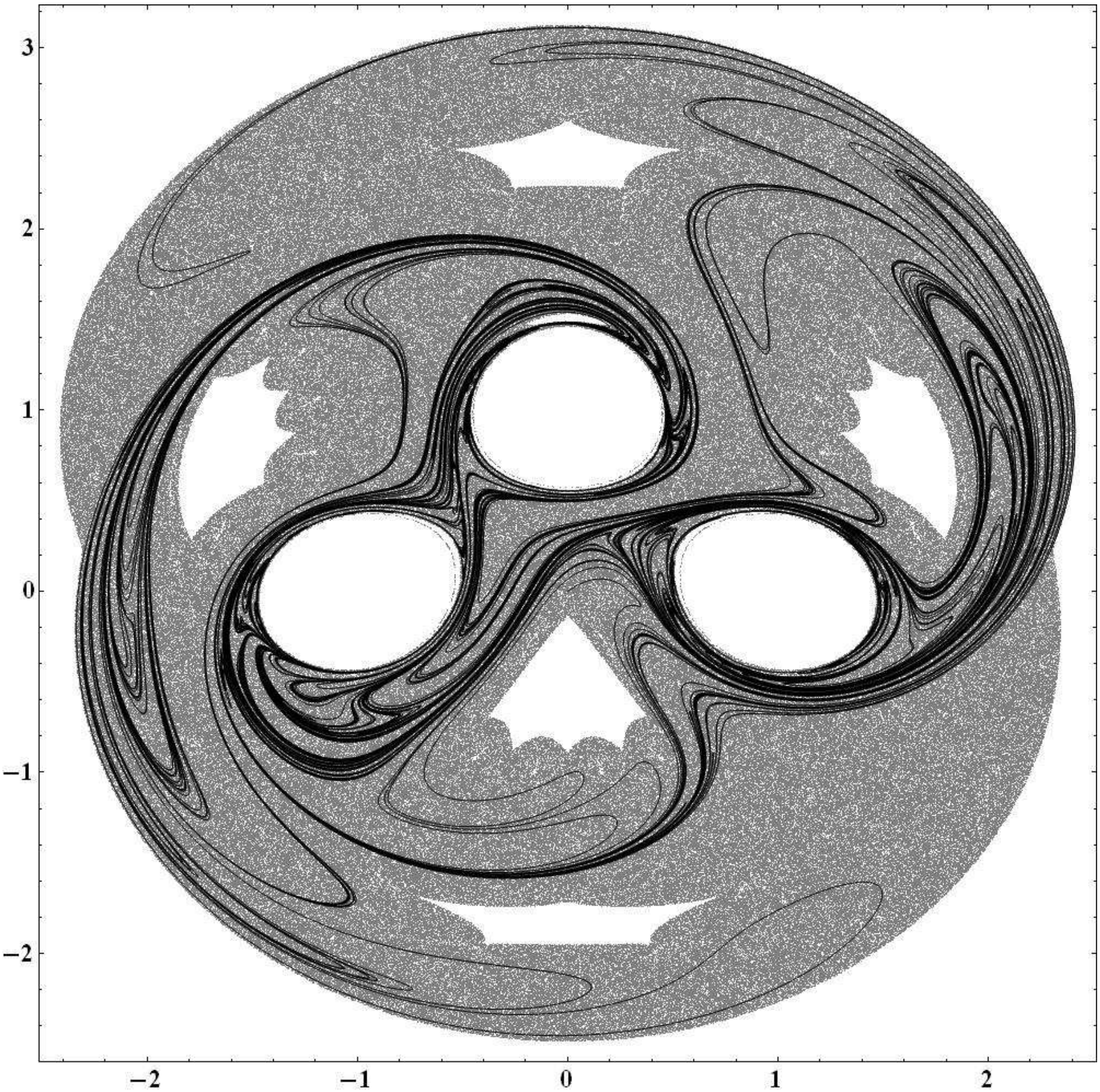

Figure 6.4: Stable manifold segment of $(0, 0)$ with $L = 20.000$ for the 3-periodic three vortex system. The gray background is the full ergodic region as shown in Figure 4.1.

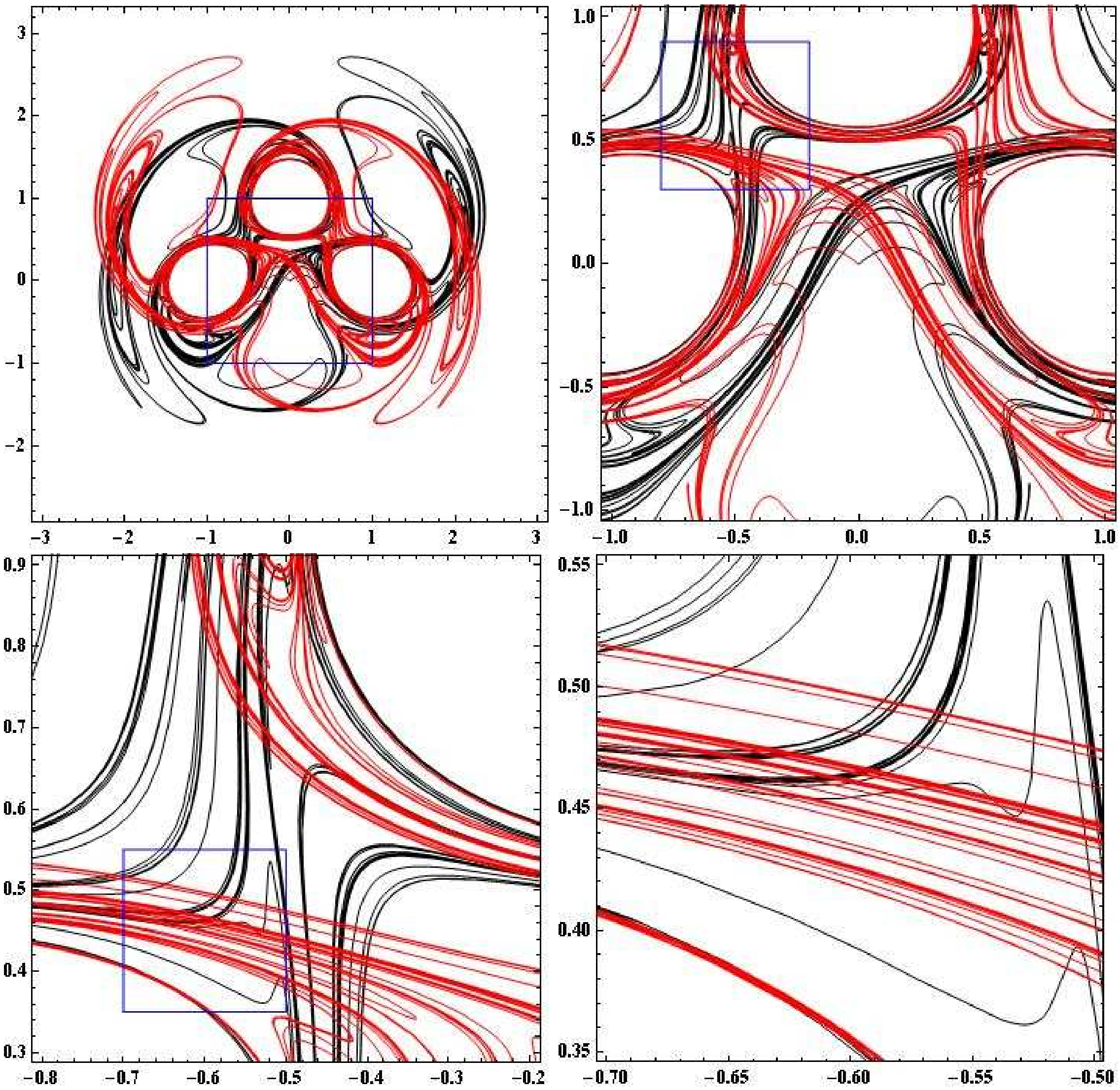

Figure 6.5: Top left: Stable (black) and unstable (red) manifold segments of $(0, 0)$ with $L = 2.000$ for the 3-periodic three vortex system. Going horizontally from there: progressive details.

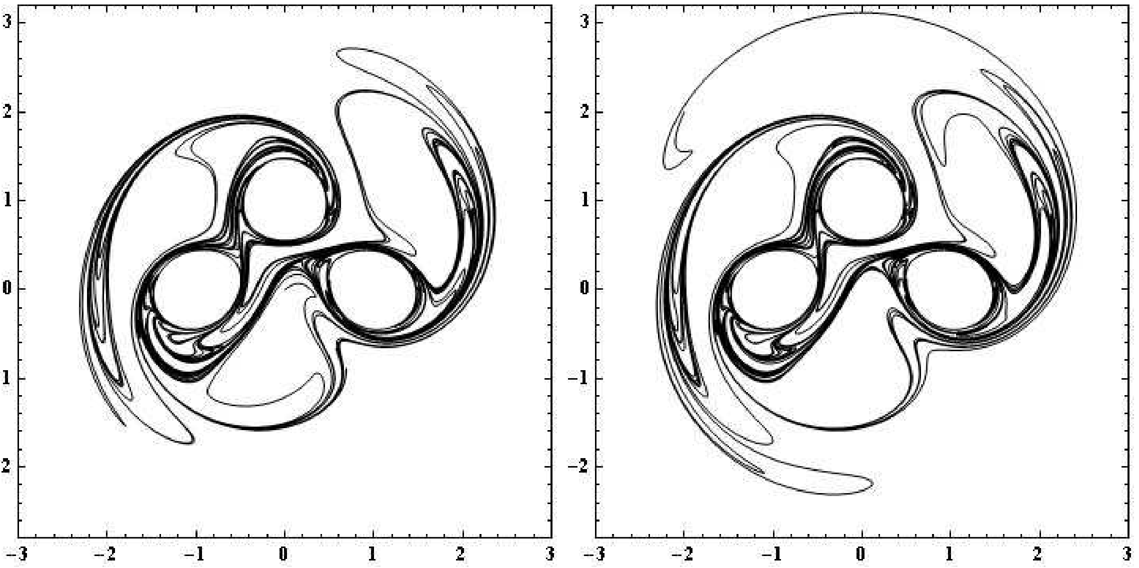

Figure 6.6: Left: stable manifold segment of $(0,0)$ with $L = 2.000$ for the 3-periodic three vortex system. Right: Same, of $(-2,2)$.

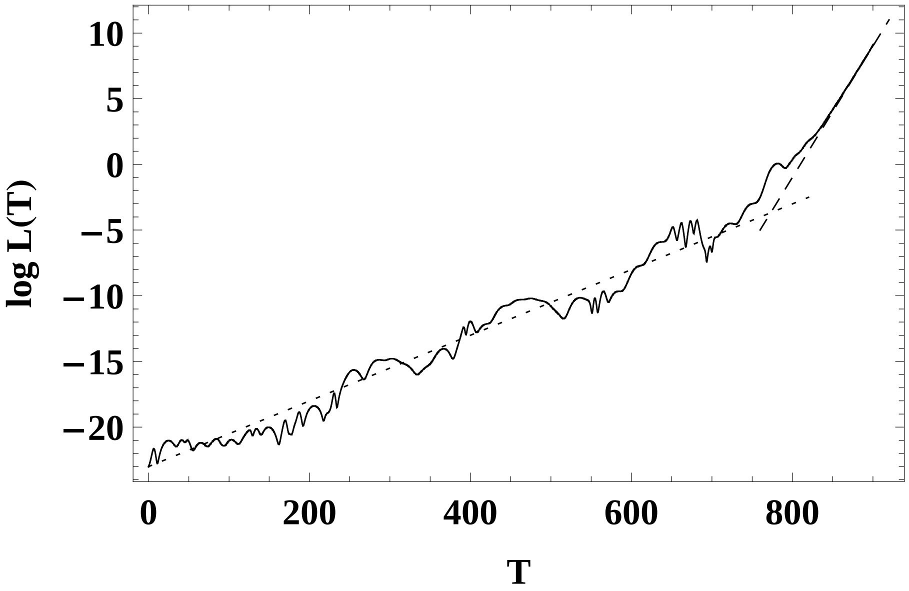

Figure 6.7: $\mathcal{L}(T)$ as a function of T for the 3-periodic three vortex system, with $\mathbf{a} = (0,0)$ and $L_0 = 10^{-10}$. The dotted and dashed lines are guides to the eyes with slopes $\lambda_{1+} = 0.025$ and $\sigma_+ = 0.1$ respectively.

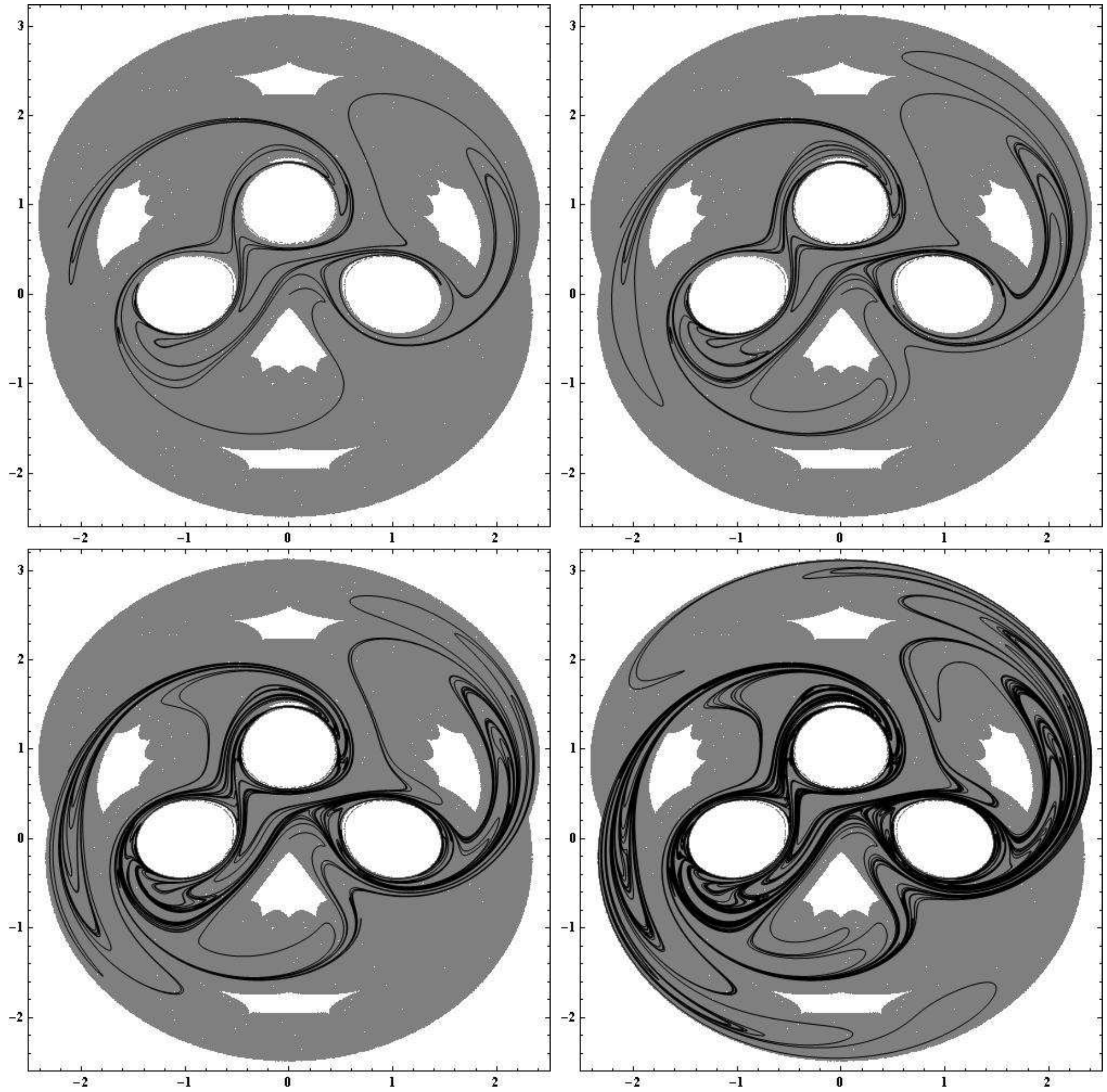

Figure 6.8: Stable manifold segment of $(0,0)$ for the 3-periodic three vortex system for various lengths. From the top left, going horizontally: $L = 200$, $L = 400$, $L = 2.000$ and $L = 20.000$. The gray background is the full ergodic region as shown in Figure 4.1.

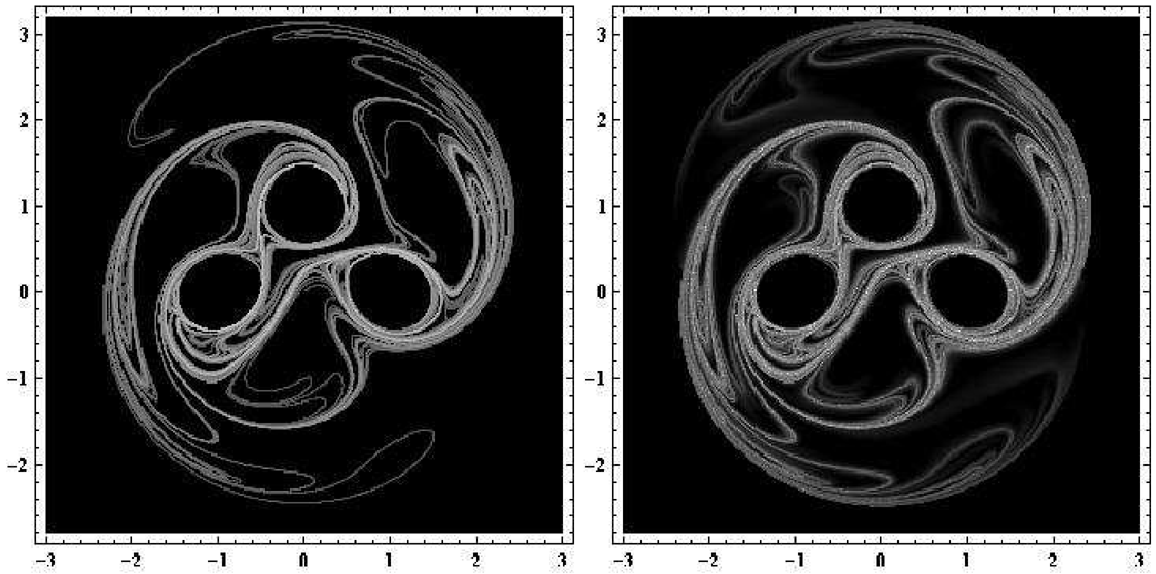

Figure 6.9: Left: Positive time manifold measure for the 3-periodic three vortex system found using the stable manifold segment of $(0, 0)$ with $L = 20.000$ (shown in Figure 6.4). Right: Positive time w-measure as found in Figure 5.6 (using a higher resolution than used there).

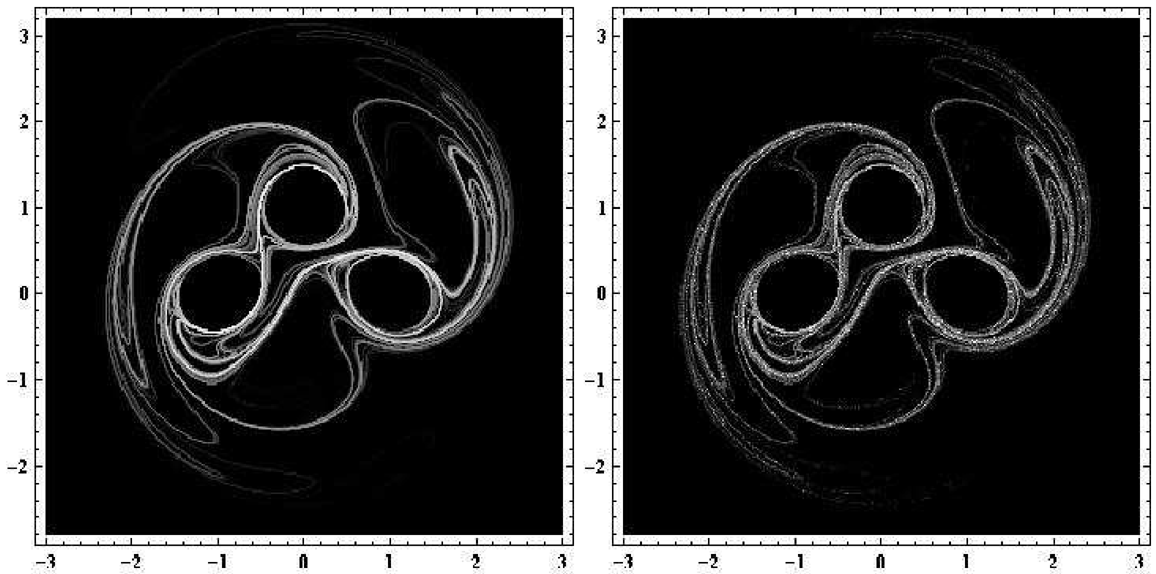

Figure 6.10: Positive time manifold measure (left) and w-measure (right) for the 3-periodic three vortex system as shown in Figure 6.9, with the color scale adjusted to highlight the similarities between both measures.

not present in both measures. It is now clear that in this region both measures are indeed identical. Note also that the w-measure is not as well resolved as the manifold measure. All further pictures will use color scaling to highlight the similarities, so it is important to note that in all cases there are features present in the computed w-measure which are not visible in the computed manifold measure.

Results for the other periodic systems are shown in Figures 6.11 and 6.12. It is not necessary to consider the negative time measures, since these are just the mirror images of the positive times measures for all periodic systems. This is not the case in the irregular four vortex system. Results for this system are shown in Figures 6.13 (positive time) and 6.14 (negative time). In all cases the similarities between both measures are apparent, although the w-measure is not as well resolved as the manifold measure, particularly in the case of the irregular system. The effect is much stronger there because of the lower quality of the computed w-measure.

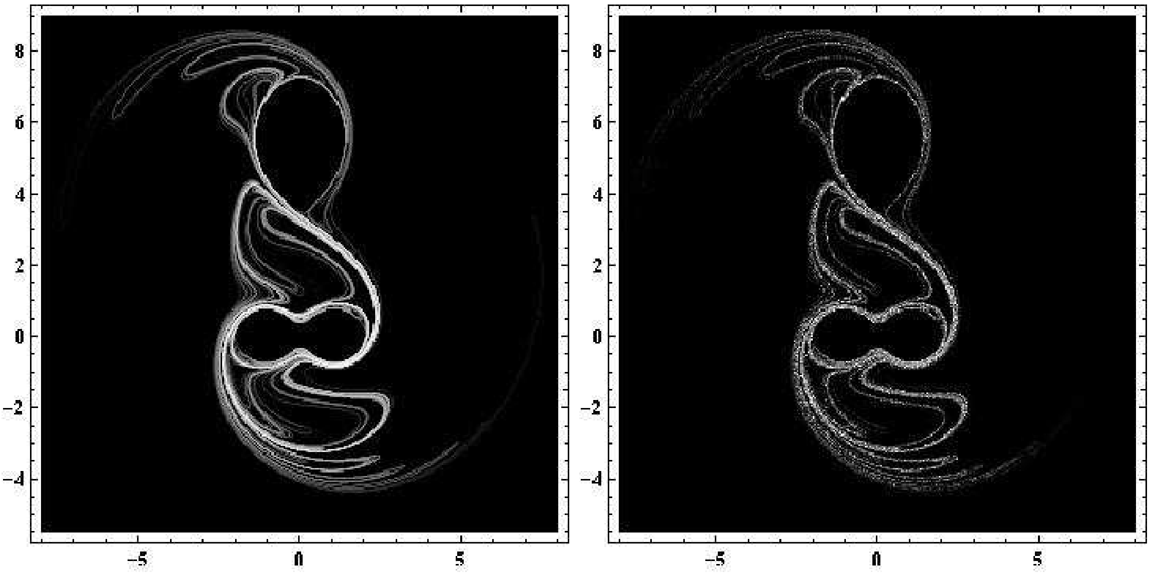

Figure 6.11: Left: Positive time manifold measure for the 2-periodic three vortex system found using the stable manifold segment of $(0, 0)$ with $L = 16.000$. Right: Positive time w-measure as found in Figure 5.8 (using a higher resolution than used there).

Overall however these results support the conclusion of section 6.1.3 that the manifold measure and w-measure are identical. Any differences between the computed measures can be explained by the numerical methods not having converged yet: the computation of the manifold measure is inaccurate in areas of *low* w-measure since

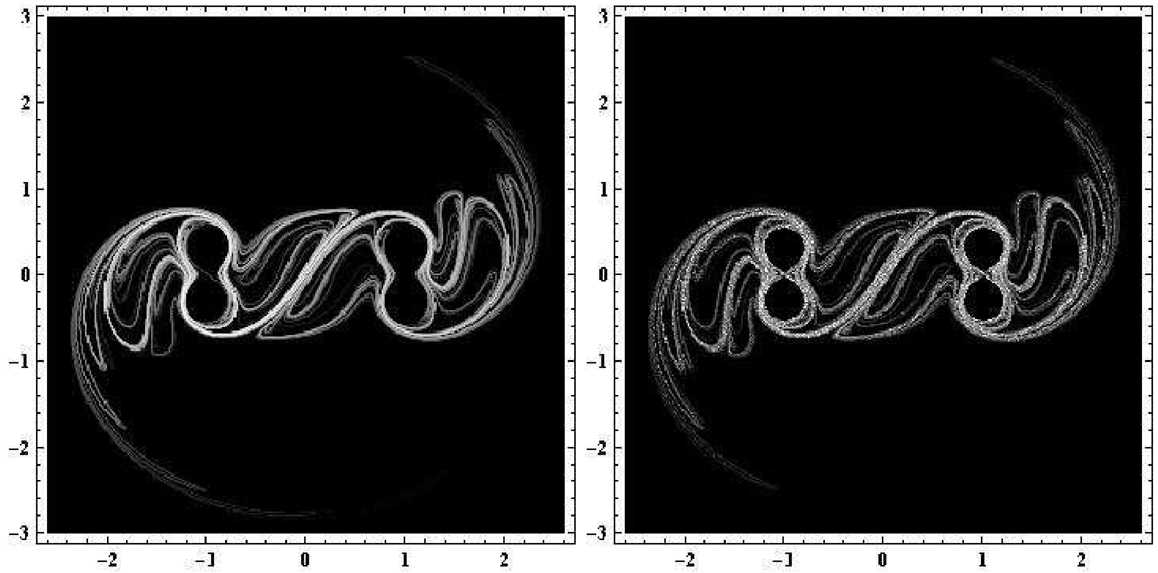

Figure 6.12: Left: Positive time manifold measure for the periodic four vortex system found using the stable manifold segment of $(0, -0.75)$ with $L = 20.000$. Right: Positive time w-measure as found in Figure 5.10 (using a higher resolution than used there).

the manifold is simply unlikely to pass through these areas for finite L, while the computation of the w-measure is inaccurate in areas of *high* w-measure since the measure is highly irregular there, as a result of which any computation using a finite amount of points easily misses some areas of high w-measure. It is also interesting to note that the computation of the manifold measures in this chapter took a few hours, while the computation of the w-measures in the previous chapter took a few days.

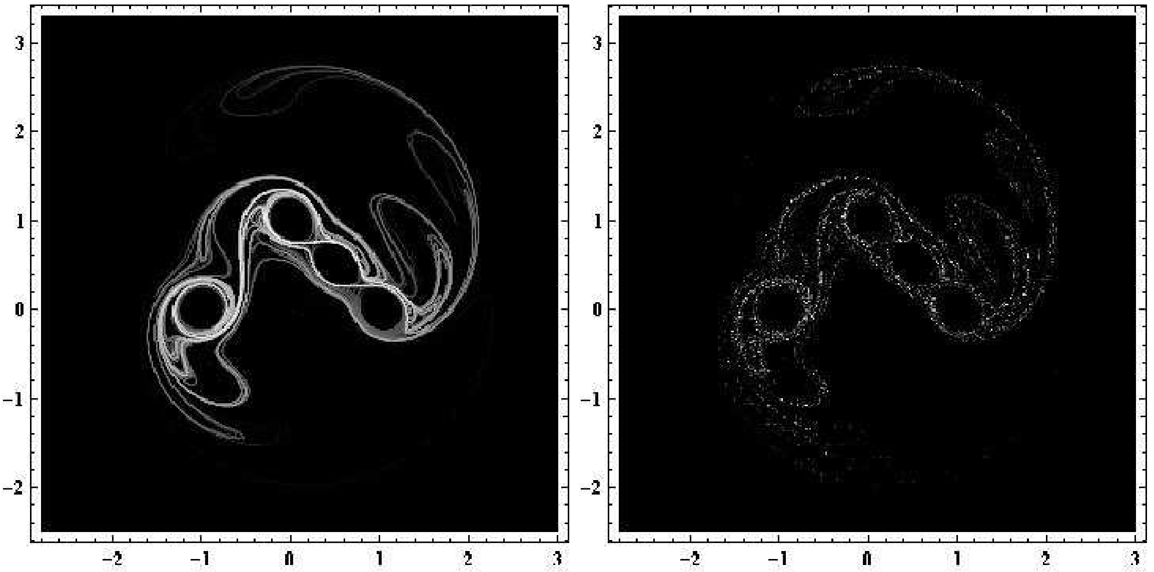

Figure 6.13: Left: Positive time manifold measure for the irregular four vortex system found using the stable manifold segment of $(0, 0)$ with $L = 15.000$. Right: Positive time w-measure as found in Figure 5.12 (using a higher resolution than used there).

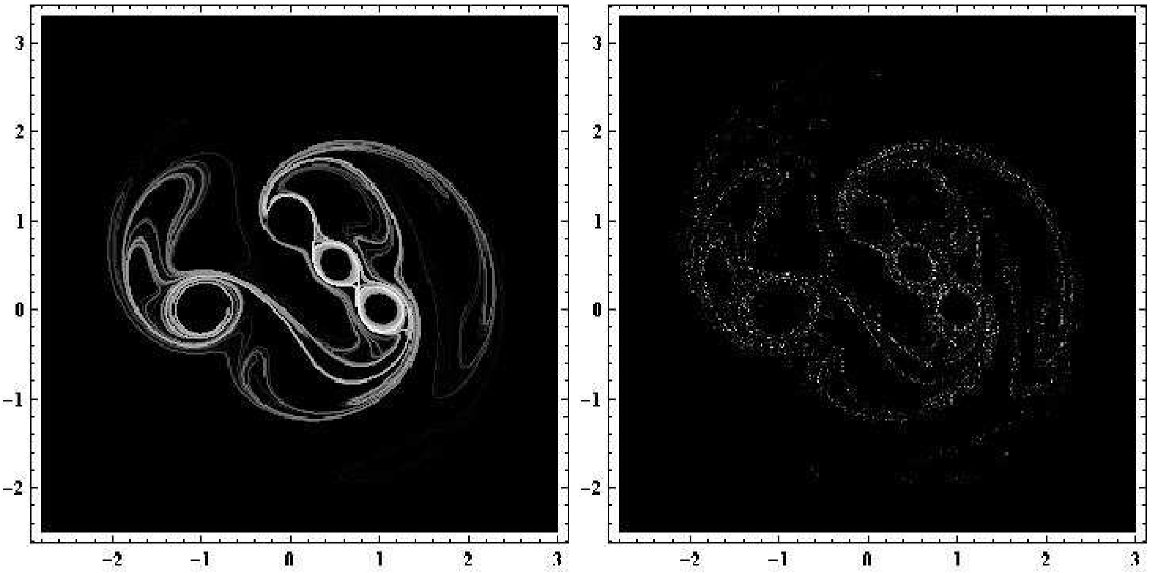

Figure 6.14: Left: Negative time manifold measure for the irregular four vortex system found using the unstable manifold segment of $(0, 0)$ with $L = 15.000$. Right: Negative time w-measure as found in Figure 5.13.

7 Conclusions and recommendations

7.1 Conclusions

The central question of this work, posed at the start of chapter 1, was the relevance of invariant manifolds to the description of the mixing process in two-dimensional incompressible flows. We have shown how, through the w-measure, they describe the mixing time distribution and the partial mixing structure. This allows us to concisely describe how mixing occurs in any given 2D incompressible flow.

In the first place, there are multiple regimes within the phase space with different mixing properties. If the flow is stationary, no mixing occurs anywhere. If it is non-stationary mixing can occur, but that it occurs in some regions of the phase space does not guarantee that the whole phase space mixes. In fact, there are one or multiple ergodic regions, and possibly one or more islands as well. Every ergodic region mixes, in the sense that particles seeded inside them according to some distribution will asymptotically be spread uniformly over this region. Islands are distinguished from ergodic regions by the fact that mixing does *not* occur within them. If there are multiple ergodic regions, they all mix individually but no particles are exchanged between them. Islands are typical of periodic flows, but can occur in irregular flows as well.

Now consider the mixing within an ergodic region. In particular, we will examine what happens to a small cloud of many particles seeded somewhere in this region. We already know that eventually it will spread out over the full ergodic region, but we want to describe the process. Initially, the cloud will stay intact, undergoing some stretching but no folding. The duration of this phase depends on the initial location, a fact which can be used to define a structure in the ergodic region: the mixing time distribution. At some point it will be stretched so much that folding occurs, and a typical structure becomes visible: the partial mixing structure. This structure continues to evolve and become more complex, until at some point the features are too complex for the limited number of particles to capture and all that remains visible is a uniform distribution over the ergodic region. So, in describing the mixing there are two relevant structures in the ergodic region: the mixing time distribution and the partial mixing structure.

These two structures are formalized by the concept of the w-measure. This is based on Lyapunov exponents, the rate of exponential separation of nearby particle trajectories. These can be considered both asymptotically and using finite times. The asymptotic value, known as the maximum Lyapunov exponent (MLE), is constant in the ergodic region and is hence a useful timescale in describing the mixing process. The finite time values can be used to identify regions of high stretching, inducing in this way a well-defined measure on the ergodic region, the w-measure. For positive time, this corresponds to the mixing time distribution; for negative time, to the partial mixing structure.

Another approach is based on the stable and unstable invariant manifolds of trajectories in the ergodic region, i.e. the attracting and repelling sets respectively of these trajectories. These are non-terminating curves in the ergodic region, invariant under the flow. They are also dense in the ergodic region, although not uniformly distributed. In fact, they are distributed exactly according to the w-measure, with the global stable manifold corresponding to the positive time w-measure and hence the mixing time distribution, and the global unstable manifold to the negative time one and hence the partial mixing structure.

Converging numerical methods have been introduced to determine the w-measure based on both of these approaches, i.e. Lyapunov exponents and invariant manifolds. These give identical results asymptotically, but both are necessary because the convergence properties are very different. In particular, the Lyapunov method should be used when we are looking for features in areas of low w-measure, while the manifold method is most accurate in areas of high w-measure.

7.2 Suggestions for further research

There are several avenues of possible further research based on these results, which fall into two categories: ways to expand the methods to analyze more aspects of the flow and the mixing process on the one hand, and applications to physical systems on the other hand.

7.2.1 Further analysis

The relationship between invariant manifolds and Lyapunov exponents is a very important component of this work. There is another approach, already briefly mentioned, that also relies heavily on this relationship, namely the computation of Lagrangian coherent structures (LCS). A detailed description of these structures can be found in [69]. The central point is that invariant manifolds are approximated by ridges in the finite time Lyapunov exponent field. This is very similar to our results, where invariant manifolds are highly concentrated in areas of high Lyapunov exponent. A comparison of the LCS approach to the w-measure approach could lead to a better understanding of both methods and the mixing process in general.

We have seen several examples of *transport barriers*. These are barriers which particles cannot cross. For example, the edges of islands in the ergodic region or of vortex cores are transport barriers. A concept which we have not considered is that of *partial* transport barriers, barriers which are rarely crossed by particles. There are several methods to find them in periodic systems based on the transition matrix, which approximates the fluid motion by a Markov process on a grid [18]. Partial transport barriers are then found by identifying almost invariant sets, using eigenvector analysis [3] or graph algorithms [56]. To give an impression, an eigenvector of the transition matrix is shown for the periodic four vortex system in Figure 7.1. The dark and light regions near the left and right dipoles respectively indicate partial transport barriers (there are other partial barriers visible as well). These play an important role in the mixing process, and so further developing these methods may provide fundamental insights into this process. Another interesting point, also visible in Figure 7.1 when compared with Figure 6.12, is that the structure of the w-measure is also visible in the eigenvector.

Finally, we have only considered the negative and positive time w-measures separately. It is also possible to combine them using the symmetric product measure [30]. This measure describes the spatial distribution of the *intersections* of the stable and unstable manifolds with each other. In peroidic systems it is asymptotically identical to the Bowen measure, which describes the spatial distribution of periodic cycles [42]. The relationship (if any) of this measure to the mixing properties has yet to be established.

7.2.2 Possible applications

The methods developed here can be applied to any incompressible 2D flow as long as we can accurately compute particle trajectories, i.e. we have a good approximation for $\Phi(t, T)$. A specific use could be the prediction of which areas mix the fastest, for instance to determine where a catalyst should be injected to be efficiently mixed in a reactor. Oceanic flows are often approximately incompressible and two-dimensional and so are a frequent target for these analysis methods. Examples of the application of the methods mentioned above can be found in [48] (LCS) and [24] (partial transport barriers).

A new type of industrial mixer, found for instance in 'lab-on-a-chip' implementa-

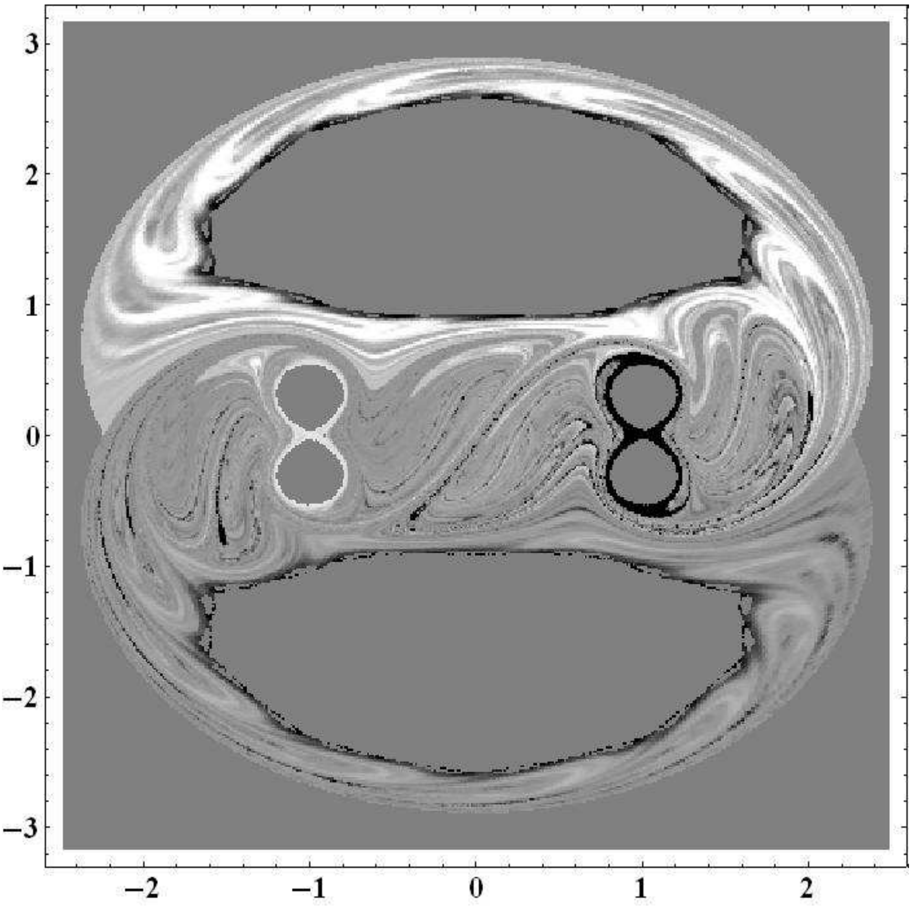

Figure 7.1: Second eigenvector of the transition matrix for the periodic four vortex system.

tions, is the static mixer [12]. The concept is that of a fluid flowing through some suitably shaped pipe which induces mixing. An important point is that the flow is *stationary.* This does not contradict the earlier statement that no mixing takes place in stationary flows, since we are now considering a 3D flow. Still, the methods developed here can be applied; the motion along the pipe takes the role of time [41]. It might also be possible to extend the methods to arbitrary stationary 3D flows, using an extension of symplectic topology to odd-dimensional spaces: contact topology [9, 26].

Bibliography

[1] A. Adrover, M. Giona, *Long-range correlation properties of area-preserving chaotic systems*, Physica A **253**, 143 (1998)

[2] A. Adrover, M. Giona, *Measure-theoretical properties of the unstable foliation of two-dimensional differentiable area-preserving systems*, Physical Review E **60**, 347 (1999)

[3] C. Alpert, A. Kahng, S. Yao, *Spectral partitioning with multiple eigenvectors*, Discrete Applied Mathematics **90**, 3 (1999)

[4] M. Alvarez, F. Muzzio, S. Cerbelli, A. Adrover, M. Giona, *Self-similar spatiotemporal structure of intermaterial boundaries in chaotic flows*, Physical Review Letters **81**, 3395 (1998)

[5] H. Aref, *Stirring by chaotic advection*, Journal of Fluid Mechanics **143**, 1 (1984)

[6] H. Aref, *The development of chaotic advection*, Physics of Fluids **14**, 1315 (2002)

[7] H. Aref, N. Pomphrey, *Integrable and chaotic motions of four vortices I. The case of identical vortices*, Proceedings of the Royal Society of London A **380**, 359 (1982)

[8] V. Arnol'd, A. Avez, *Ergodic problems of classical mechanics*, Benjamin (1968)

[9] V. Arnol'd, *First steps in symplectic topology*, Russian Mathematical Survey **41**, 1 (1986)

[10] V. Artale, G. Boffetta, A. Celani, M. Cencini, A. Vulpiani, *Dispersion of passive tracers in closed basins: beyond the diffusion coefficient*, Physics of Fluids **9**, 3162 (1997)

[11] R. Bartle, *The elements of integration and Lebesgue measure*, Wiley Classics (1995)

[12] D. Beebe, R. Adrian, M. Olsen, M. Stremler, H. Aref, B. Jo, *Passive mixing in microchannels: fabrication and flow experiments*, Mécanique & Industries **2**, 343 (2001)

[13] G. Birkhoff, B. Koopman, *Recent contributions to the ergodic theory*, Proceedings of the National Academy of Sciences **18**, 279 (1932)

[14] S. Boatto, R. Pierrehumbert, *Dynamics of a passive tracer in a velocity field of four identical point vortices*, Journal of Fluid Mechanics **394**, 137 (1999)

[15] G. Boffetta, G. Lacorata, G. Redaelli, A. Vulpiani, *Detecting barriers to transport: a review of different techniques*, Physica D **159**, 58 (2001)

[16] S. Cerbelli, M. Giona, *Characterization of nonuniform chaos in area-preserving nonlinear maps through a continuous archetype*, Chaos, Solitons and Fractals **35**, 13 (2008)

[17] S. Cerbelli, J. Zalc, F. Muzzio, *The evolution of material lines curvature in deterministic chaotic flows*, Chemical Engineering Science **55**, 363 (2000)

[18] M. Dellnitz, O. Junge, W. Koon, F. Lekien, M. Lo, J. Marsden, K. Padberg, R. Preis, S. Ross, B. Thiere, *Transport in dynamical astronomy and multibody problems*, International Journal of Bifurcation and Chaos **15**, 699 (2005)

[19] D. Dritschel, *Contour dynamics and contour surgery: numerical algorithms for extended, high-resolution modelling of vortex dynamics in two-dimensional, inviscid, incompressible flows*, Computer Physics Reports **10**, 77 (1989)

[20] C. van Duijn, M. de Neef, *Analyse van differentiaalvergelijkingen*, Delft University Press (2001)

[21] J. Eckmann, D. Ruelle, *Ergodic theory of chaos and strange attractors*, Reviews of Modern Physics **57**, 617 (1985)

[22] S. Ershov, A. Potapov, *On the concept of stationary Lyapunov basis*, Physica D **118**, 167 (1998)

[23] J. Finn, D. del-Castillo-Negrete, *Lagrangian chaos and Eulerian chaos in shear flow dynamics*, Chaos **11**, 816 (2001)

[24] G. Froyland, K. Padbert, M. England, A. Treguier, *Detection of coherent oceanic structures via transfer operators*, Physical Review Letters **98**, 224503 (2007)

[25] K. Geist, U. Parlitz, W. Lauterborn, *Comparison of different methods for computing Lyapunov exponents: general and mathematical physics*, Progress of Theoretical Physics **83**, 875 (1990)

[26] R. Ghrist, *On the contact topology and geometry of ideal fluids*, Handbook of Mathematical Fluid Dynamics **4**, 1 (2002)

[27] R. Gilmore, *Baker-Campbell-Hausdorff formulas*, Journal of Mathematical Physics **15**, 2090 (1974)

[28] M. Giona, A. Adrover, *Nonuniform stationary measure of the invariant unstable foliation in Hamiltonionan and fluid mixing systems*, Physical Review Letters **81**, 3864 (1998)

[29] M. Giona, A. Adrover, F. Muzzio, S. Cerbelli and M. Alvarez, *The geometry of mixing in time-periodic chaotic flows. I. Asymptotic directionality in physically realizable flows and global invariant properties*, Physica D **132**, 298 (1999)

[30] M. Giona, S. Cerbelli, A. Adrover, *Symmetric product measures: binomial processes and invariant manifold intersections in dynamical systems*, Physica A **356**, 447 (2005)

[31] M. Giona, S. Cerbelli, F. Muzzio, A. Adrover, *Non-uniform stationary measure properties of chaotic area-preserving systems*, Physica A **254**, 451 (1998)

[32] M. Gromov, *Entropy, homology and semialgebraic geometry*, Séminaire N. Bourbaki **663**, 225 (1985)

[33] G. Haller, *Finding finite-time invariant manifolds in two-dimensional velocity fields*, Chaos **10**, 99 (2000)

[34] G. Haller, A. Poje, *Finite time transport in aperiodic flows*, Physica D **119**, 352 (1998)

[35] T. Havelock, *The stability of rectilinear vortices in ring formation*, Philosophical Magazine (7) **11**, 617 (1931)

[36] J. Heinonen, *Lectures on Lipschitz analysis*, Lectures at the 14th Jyväskylä Summer School (2004)

[37] D. Hobson, *An efficient method for computing invariant manifolds in planar maps*, Journal of Computational Physics **104**, 14 (1993)

[38] P. Holmes. *Poincaré, celestial mechanics, dynamical-systems theory and "chaos"*, Physics Reports **3**, 137 (1990)

[39] K. Ide, D. Small, S. Wiggins, *Distinguished hyperbolic trajectories in time-dependent fluid flows: analytical and computational approach for velocity fields defined as data sets*, Nonlinear processes in Geophysics **9**, 237 (2002)

[40] J. Jensen, *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*, Acta Mathematica **30**, 175 (1906)

[41] S. Jones, O. Thomas, H. Aref, *Chaotic advection by laminar flow in a twisted pipe*, Journal of Fluid Mechanics **209**, 335 (1989)

[42] A. Katok, B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, Cambridge University Press (1995)

[43] Y. Katznelson, B. Weiss, *A simple proof of some ergodic theorems*, Israeal Journal of Mathematics **42**, 291 (1982)

[44] P. Kundu, I. Cohen, *Fluid Mechanics*, Third Edition, Elsevier Academic Press (2004)

[45] B. Kuvshinov, T. Schep, *Point-vortex approach in two-dimensional turbulence*, pre-print

[46] L. Kuznetsov, G. Zaslavsky, *Passive particle transport in three-vortex flow*, Physical Review E **61**, 3777 (2000)

[47] L. Kuznetsov, G. Zaslavsky, *Regular and chaotic advection in the flow field of a three-vortex system*, Physical Review E **58**, 7330 (1998)

[48] F. Lekien, C. Coulliette, A. Mariano, E. Ryan, L. Shay, G. Haller, J. Marsen, *Pollution release tied to invariant manifolds: a case study for the coast of Florida*, Physica D **210**, 1 (2005)

[49] A. Leonard, *Curvature and torsion of material lines in chaotic flows*, Fluid Dynamics Research **36**, 261 (2005)

[50] N. Malhotra, S. Wiggins, *Geometric structures, lobe dynamics, and Lagrangian transport in flows with aperiodic time-dependence, with applications to Rossby wave flow*, Journal of Nonlinear Science **8**, 401 (1998)

[51] A. Mancho, D. Small, S. Wiggins, K. Ide, *Computation of stable and usntable manifolds of hyperbolic trajectories in two-dimensional, aperiodically time-dependent vector fields*, Physica D **182**, 188 (2003)

[52] R. Mattheij, S. Rienstra, J. ten Thije Boonkkamp, *Partial differential equations: Modeling, analysis, computation*, Society for Industrial and Applied Mathematics (2005)

[53] R. McLachlan, P. Atela, *The accuracy of symplectic integrators*, Nonlinearity **5**, 541 (1992)

[54] J. Meiss, *Symplectic maps, variational principles, and transport*, Reviews of Modern Physics **64**, 795 (1992)

[55] S. Middleman, *Fundamentals of Polymer Processing*, MacGraw-Hill (1977)

[56] B. Monien, R. Pries, R. Diekmann, *Quality matching and local improvement for multilevel graph-partitioning*, Parallel Computing **26**, 1609 (2000)

[57] F. Muzzio, M. Alvarez, S. Cerbelli, M. Giona, A. Adrover, *The intermaterial area density generated by time- and spatially periodic 2D chaotic flows*, Chemical Engineering Science **55**, 1497 (2000)

[58] S. Newhouse, T. Pignataro, *On the estimation of topological entropy*, Journal of Statistical Physics **72**, 1331 (1993)

[59] V. Oseledec, *Multiplicative ergodic theorem. Lyapunov exponents of dynamical systems*, Transactions of the Moscow Mathematical Society **19**, 197 (1968)

[60] R. Otten, *Hyperbolicity in nonautonomous velocity fields with applications to point vortex systems*, Master thesis, Technische Universiteit Eindhoven (2007)

[61] J. Ottino, *Mixing, chaotic advection, and turbulence*, Annual Review of Fluid Mechanics **22**, 207 (1990)

[62] V. Rom-Kedar, A. Leonard, S. Wiggins, *An analytical study of transport, mixing and chaos in an unsteady vortical flow*, Journal of Fluid Mechanics **214**, 347 (1990)

[63] D. Ruelle, *Ergodic theory of differentiable dynamical systems*, Publications Mathématiques de l'I.H.É.S. **50**, 27 (1979)

[64] Y. Saad, *Iterative methods for sparse linear systems*, Second Edition, Society for Industrial and Applied Mathematics (2003)

[65] R. Samelson, *Chaotic transport by mesocale motions*, in: R. Adler, P. Müller, B. Rozovskii, *Stochastic modelling in Physical Oceanography*, Birkhäuser (1996)

[66] R. Samelson, S. Wiggins, *Lagrangian transport in geophysical jets and waves, the dynamical systems approach*, Springer (2006)

[67] B. Schmid, M. Hengi, U. Stephan, *Salinity-induced density stratification in near-laminar open-channel flows*, Journal of Hydraulic Engineering **130**, 1206 (2004)

[68] B. Schutz, *Geometrical methods of mathematical physics*, Cambridge University Press (1980)

[69] S. Shadden, F. Lekien, J. Marsden, *Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows*, Physica D **212**, 271 (2005)

[70] T. Solomon, E. Weeks, H. Swinney, *Observation of anomalous diffusion and Lévy flights in a two-dimensional rotating flow*, Physical Review Letters **71**, 3975 (1993)

[71] J. Thiffeault, *Derivatives and constraints in chaotic flows: asymptotic behaviour and a numerical method*, Physica D **172**, 139 (2002)

[72] J. Thiffeault, H. Boozer, *Geometrical constraints on finite-time Lyapunov exponents in two and three dimensions*, Chaos **11**, 16 (2001)

[73] P. Vosbeek, *Contour dynamics and applications to 2D vortices*, Ph.D. thesis, Technische Universiteit Eindhoven (1998)

[74] P. Welander, *Studies of the general development of motion in a two-dimensional ideal fluid*, Tellus **7**, 141 (1955)

[75] E. Whittaker, *A treatise on the analytical dynamics of particles and rigid bodies*, Cambridge University Press (1937)

[76] Y. Yomdin, *Volume growth and entropy*, Israel Journal of Mathematics **57**, 285 (1987)

[77] H. Yoshida, *Construction of higher order symplectic integrators*, Physics Letters A **150**, 262 (1990)

[78] L. Young, *Developments in Chaotic Dynamics*, Notices of the AMS **45**, 1318 (1998)