Distinguished hyperbolic trajectories in time-dependent fluid flows: analytical and computational approach for velocity fields defined as data sets

K. Ide, D. Small, S. Wiggins

Nonlinear Processes in Geophysics (2002) 9: 237–263

Nonlinear Processes in Geophysics

Distinguished hyperbolic trajectories in time-dependent fluid flows: analytical and computational approach for velocity fields defined as data sets

K. Ide[1], **D. Small**[2], **and S. Wiggins**[2]

[1]Department of Atmospheric Sciences, and Institute of Geophysics and Planetary Physics UCLA, Los Angeles, CA 90095-1565, USA
[2]School of Mathematics of Bristol,University Walk, Bristol, BS8 1TW, UK

Received: 6 July 2000 – Accepted: 4 January 2001

Abstract. In this paper we develop analytical and numerical methods for finding special hyperbolic trajectories that govern geometry of Lagrangian structures in time-dependent vector fields. The vector fields (or velocity fields) may have arbitrary time dependence and be realized only as data sets over finite time intervals, where space and time are discretized. While the notion of a hyperbolic trajectory is central to dynamical systems theory, much of the theoretical developments for Lagrangian transport proceed under the assumption that such a special hyperbolic trajectory exists. This brings in new mathematical issues that must be addressed in order for Lagrangian transport theory to be applicable in practice, i.e. how to determine whether or not such a trajectory exists and, if it does exist, how to identify it in a sequence of instantaneous velocity fields. We address these issues by developing the notion of a distinguished hyperbolic trajectory (DHT). We develop an existence criteria for certain classes of DHTs in general time-dependent velocity fields, based on the time evolution of Eulerian structures that are observed in individual instantaneous fields over the entire time interval of the data set. We demonstrate the concept of DHTs in inhomogeneous (or "forced") time-dependent linear systems and develop a theory and analytical formula for computing DHTs. Throughout this work the notion of linearization is very important. This is not surprising since hyperbolicity is a "linearized" notion. To extend the analytical formula to more general nonlinear time-dependent velocity fields, we develop a series of coordinate transforms including a type of linearization that is not typically used in dynamical systems theory. We refer to it as Eulerian linearization, which is related to the frame independence of DHTs, as opposed to the Lagrangian linearization, which is typical in dynamical systems theory, which is used in the computation of Lyapunov exponents. We present the numerical implementation of our method which can be applied to the velocity field given as a data set. The main innovation of our method is that it provides an approximation to the DHT for the entire time-interval of the data set. This offers a great advantage over the conventional methods that require certain regions to converge to the DHT in the appropriate direction of time and hence much of the data at the beginning and end of the time interval is lost.

Correspondence to: S. Wiggins (s.wiggins@bristol.ac.uk)

1 Introduction

Over the past 10 years there has been much work in applying the approach and methods of dynamical systems theory to the study of transport in fluids from the Lagrangian point of view. Suppose one is interested in the motion of a passive tracer in a fluid (e.g. dye, temperature, or any material that can be considered as having a negligible effect on the flow); then, neglecting molecular diffusion, the passive tracer follows fluid particle trajectories which are solutions of

$$\frac{d}{dt}\boldsymbol{x} = \boldsymbol{u}(\boldsymbol{x}, t), \tag{1}$$

where $\boldsymbol{u}(\boldsymbol{x}, t)$ is the velocity field of the fluid flow, $\boldsymbol{x} \in \mathbb{R}^n, n = 2$ or 3. When viewed from the point of view of dynamical systems theory, the phase space of Eq. (1) is actually the physical space in which the fluid flow takes place. Evidently, "structures" in the phase space of Eq. (1) should have some influence on the transport and mixing properties of the fluid. Babiano et al. (1994) and Aref and El Naschie (1994) provide recent reviews of this approach.

To make the connection with the large body of literature on dynamical systems theory more concrete, let us consider a less general fluid mechanical setting. Suppose that the fluid is two-dimensional and incompressible, although our theory can be applied to compressible as well as viscid flows. Then we know that the velocity field can be obtained from the

derivatives of a scalar valued function, $\psi(x_1, x_2, t)$, known as the stream function, as follows

$$\frac{d}{dt}x_1 = \frac{\partial \psi}{\partial x_2}(x_1, x_2, t),$$
$$\frac{d}{dt}x_2 = -\frac{\partial \psi}{\partial x_1}(x_1, x_2, t)\,. \tag{2}$$

In the context of dynamical systems theory, Eq. (2) is a time-dependent Hamiltonian vector field where the stream function plays the role of the Hamiltonian function. If the flow is time-periodic, then the study of Eq. (2) is typically reduced to the study of a two-dimensional area preserving a Poincaré map. Practically speaking, the reduction to a Poincaré map means that rather than viewing a particle trajectory as a curve in continuous time, one views the trajectory only at discrete intervals of time, where the interval of time is the period of the velocity field. The value of making this analogy with Hamiltonian dynamical systems lies in the fact that a variety of techniques in this area have immediate applications to, and implications for, transport and mixing processes in fluid mechanics. For example, the persistence of invariant curves in the Poincaré map (KAM curves) gives rise to barriers to transport and chaos, and Smale horseshoes provide mechanisms for the "randomization" of fluid particle trajectories. An analytical technique, Melnikov's method, allows one to estimate fluxes as well as describe the parameter regimes where chaotic fluid particle motions occur. A relatively new technique, lobe dynamics, enables one to efficiently compute transport between qualitatively different flow regimes.

Dynamical systems techniques were first applied to Lagrangian transport in the context of two-dimensional, time-periodic flows. In recent years these techniques have been extended to include flows having arbitrary time dependence (see Wiggins, 1992; Malhotra and Wiggins, 1998; Haller and Poje, 1998). One aspect of our study is to consider the effect of different types of temporal variability on transport. In recent years the dynamical systems approach has been extended to a number of geophysical fluid dynamics settings (see, for example, Pierrehumbert, 1991a, 1991b; Samelson, 1992; Duan and Wiggins, 1996). These early works mainly involved kinematically defined velocity fields. Some of the first attempts to treat dynamically evolving velocity fields were the works of del Castillo-Negrete and Morrison (1993) and Ngan and Shepherd (1997). They considered special kinematic cases that could be argued to be dynamically consistent, and hence complication provided by dynamical consistency was not really present. The treatment of general dynamically evolving velocity fields became possible with the development of computational techniques to treat velocity fields which only had a numerical representation, i.e. which were the output of the numerical solution of a partial differential equation whose solution was a velocity field. Early work along these lines can be seen in Shariff et al. (1992), Duan and Wiggins (1997), and Miller et al. (1997). Recent work of this type in a geophysical fluid dynamics setting is that of Rogerson et al. (1999), which is concerned with fluid exchange across a barotropic meandering jet. Recent work of Coulliette and Wiggins (2000) allows one to treat transport in basin scale models, such as a wind driven double-gyre system.

Lobe dynamics provides a general theoretical framework, based on invariant manifold ideas from dynamical systems theory, for discussing, describing and quantifying organized structures in a fluid flow and determining their influence on transport. This paper is concerned with issues related to the numerical implementation of the technique of lobe dynamics. To begin with, one must identify certain hyperbolic trajectories (i.e. "moving saddle points") whose stable and unstable manifolds divide the flow into different flow regimes. Earlier work of Malhotra and Wiggins (1998) developed the general mathematical framework. However, implementing these mathematical ideas for practical problems requires one to face a number of new issues.

In order to begin applying dynamical systems theory to the study of transport one needs the right-hand side of Eq. (2), i.e. the velocity field. Until recently, applications have been limited to the cases where the velocity field is expressed as an analytical function of space and time. Then one can compute velocity explicitly once the position and time are given.

This may not be the case when the velocity field is obtained through the solution of some fluid dynamical nonlinear partial differential equations of motion (e.g. the Navier-Stokes equations). In general, such nonlinear partial differential equations cannot be solved analytically, i.e. the right-hand side of Eq. (2) cannot be represented in the form of some elementary or special analytical functions. However, they can often be solved numerically and the velocity field may be given as output of the model simulation at a discrete time sequence which may be also spatially discrete. Another way in which the right-hand side of Eq. (2) can be obtained is through observation. Modern remote sensing techniques (such as high frequency radar arrays) have now been developed to the point where one can measure current fields at a fairly high resolution in space and time.

Whether one obtains the velocity field through numerical simulation of a nonlinear partial differential equation or through observations, the resulting velocity field (i.e. dynamical system) is given as a data file, with gaps in space and time. Moreover, it will only be known for a finite amount of time, which may be at odds with many notions from dynamical systems theory, since dynamical systems theory is often concerned with the asymptotic in time behaviour. Consequently, the fact that the velocity field may only be known for a finite time causes major difficulties with the application of dynamical systems techniques. These difficulties are a central focus of this paper.

The central concern of this paper is to develop the notion of hyperbolic trajectory in a way that it can be applied to finite time data sets, and then develop a numerical search algorithm to find such hyperbolic trajectories for the entire temporal length of the data set. However, as we will see, quite general flows may contain an uncountable infinity of hyperbolic trajectories. In order to clarify this situation, we introduce the notion of a distinguished hyperbolic trajectory

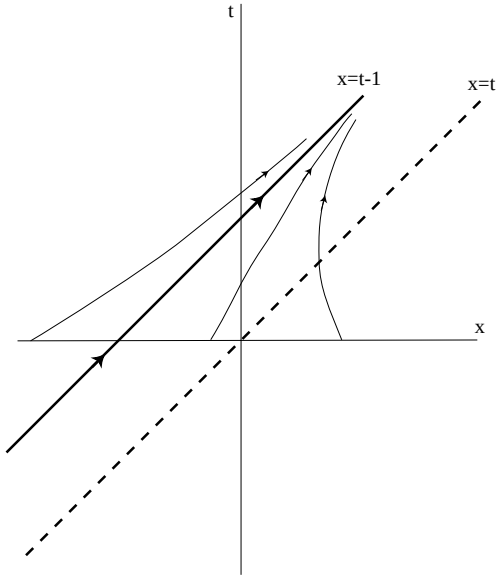

Fig. 1. The trajectories of Eq. (3) plotted in $x-t$ space. The DHT is given by $x(t)=t-1$ and the curve of ISPs is plotted as a dashed line and given by $x=t$.

(henceforth, DHT). Before going further, we want to consider two examples that illustrate in a concrete manner the issues that we will face. These examples are one-(space) dimensional. This may seem far removed from the fluid mechanical applications of interest. However, this is not the case since in many applications the boundary conditions may be free slip and then the issue of saddle type trajectories on one-dimensional boundaries becomes an area of interest.

Example 1. Consider the following example from Szeri et al. (1991):

$$\frac{d}{dt}x=-x+t. \tag{3}$$

The solution through the point x_0 at $t=0$ is given by

$$x(t)=t-1+e^{-t}\,(x_0+1)\,. \tag{4}$$

A typical way to visualize the trajectories of time-dependent vector fields is to consider the instantaneous (or "frozen time") setting, i.e. one fixes time and then considers the resulting instantaneous direction field, instantaneous streamline contours, instantaneous stagnation points (henceforth, ISPs), etc. However, such information can be very misleading if one uses it to try to understand Lagrangian transport issues.

Consider the ISPs for Eq. (3). These are given by

$$x=t. \tag{5}$$

At a fixed t, this is the unique point where the velocity is zero. However, $x=t$ is not a solution of Eq. (3). This is very different from the case of a steady flow where a stagnation point is a solution of the velocity field.

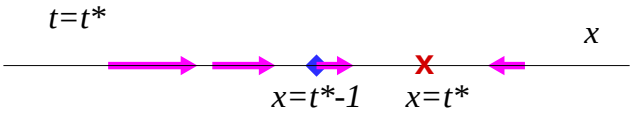

Fig. 2. The instantaneous (or "frozen time") velocity field at $t=t^*$. The DHT and the ISP are indicated by the diamond and the circle, respectively.

Now let us return to the issue of a hyperbolic trajectory. We will define this more formally at the end of this section. Now we will be content with a less mathematically formal description in order to motivate the ideas. A trajectory is said to be hyperbolic if the associated linear equations (linearized about the trajectory in question) have n linearly independent exponentially growing and decaying solutions (as $t\to\infty$), i.e. all solutions of the linearized equations exhibit exponential growth and decay. The linearization of Eq. (3) is given by

$$\dot{\xi}=-\xi, \tag{6}$$

i.e. the linearization of Eq. (3) is the same for any trajectory. Clearly, all trajectories of Eq. (3) are hyperbolic. This brings us to the notion of a DHT. Despite the fact that all trajectories of Eq. (3) are hyperbolic, upon examining the form of the general solution given in Eq. (4) we see that all trajectories decay at an exponential rate to the trajectory

$$x(t)=t-1. \tag{7}$$

This trajectory is our DHT. Note also that it is the only trajectory that does not exhibit exponential growth or decay, which can be clearly seen for the repelling situation as $t\to-\infty$. It remains for us to give it a precise mathematical definition in such a way that it lends itself to numerical computation. However, before doing that let's return to the issue of ISPs and their relationship to DHTs.

In Fig. 1 we plot some of the trajectories of Eq. (3). In particular, we plot the DHT. We show some trajectories converging to it, and we plot the curve of instantaneous stagnation points.

In Fig. 2 we plot the instantaneous velocity field at some time $t=t^*$. In this figure we see something that seems somewhat counterintuitive. Trajectories to the right of the DHT appear to be moving away from the DHT, towards the ISP. However, we know from Eq. (4) that all trajectories decay to the DHT at an exponential rate. What we are "seeing" in Fig. 2 is an artifact of drawing incorrect conclusions from instantaneous velocity fields. Trajectories to the immediate right of the DHT are indeed moving to the right (i.e. away from the DHT). However, the DHT is moving to the right at a faster speed and it eventually overtakes these trajectories. Figure 2 might also lead us to believe that trajectories converge to the ISP. But we know this is not true since we have the exact solutions.

Example 2. The time-dependent inhomogeneous term (or "forcing") on the right-hand side of Eq. (3) is unbounded as

$t \to \infty$. However, we give another example that shows that the phenomena described above is not a consequence of this unboundedness.

Consider the equation

$$\frac{d}{dt}x = -x + \sin t. \tag{8}$$

The general solution through any point x_0 at $t = 0$ is given by

$$x(t) = \frac{1}{2}(\sin t - \cos t) + e^{-t}\left(x_0 + \frac{1}{2}\right). \tag{9}$$

As in the previous example, all solutions are hyperbolic and any solution decays exponentially to the solution

$$x(t) = \frac{1}{2}(\sin t - \cos t),$$

which is the DHT. One can also verify that the ISPs, $x = \sin t$, is not a solution of Eq. (8).

To summarize, these simple examples illustrate the following points:

- A given velocity field can contain an uncountable infinity of hyperbolic trajectories. Indeed, in these examples, all trajectories are hyperbolic with the same decay rates.
- Despite this fact, we see that there may be certain distinguished hyperbolic trajectories. In these examples, this was the one trajectory that all trajectories are attracted to exponentially as $t \to \infty$.
- Due to this abundance of hyperbolic trajectories, we see that a numerical method that is designed just to find hyperbolic trajectories may not be sufficiently refined for applications. For this reason one needs to precisely define the notion of a DHT for an analytically given velocity filed. Then one needs to develop a methodology for numerical identification of the DHT, according to the refined definition, when the velocity field is given as a discrete data set, rather than an analytical function.
- ISPs are not necessarily trajectories of the velocity field and hence are frame dependent. Viewing them in instantaneous velocity fields may lead to misleading information about fluid particle trajectories. However, when we are looking for the DHT associated with a Lagrangian structure with persistent movement, their paths in time may be used as "markers", i.e. regions of the flow which are good (but not certain) candidates for DHTs to exist (e.g. Example 2).
- Numerical methods for locating hyperbolic trajectories that utilize the stretching and contraction properties to allow for certain "test regions" to converge to the hyperbolic trajectory are not adequate for time-dependent velocity fields that are only known for a finite interval of time. In the process of convergence, we "lose" much of the velocity field. Moreover, such methods require a good guess for the "test regions" that somehow bracket the hyperbolic trajectory, as discussed earlier. We have seen that the instantaneous stagnation points do not necessarily provide us with a good guess for the location of the hyperbolic trajectory.

Motivated by the simple examples above, we define two classes of DHTs. The first class considers a velocity field whose linear part is independent and constant so that it closely relates to these examples. The second class considers a general velocity field as an extention of the first class.

Definition 1. Let us consider a velocity field which has the form:

$$\frac{d}{dt}\boldsymbol{y} = \mathbf{D}\boldsymbol{y} + \boldsymbol{g}^{(\mathrm{NL})}(\boldsymbol{y}, t), \tag{10}$$

where $\mathbf{D} \in \mathbb{R}^{n\times n}$ is a constant diagonal matrix for the time-independent linear part and $\boldsymbol{g}^{(\mathrm{NL})}(\boldsymbol{y}, t) \in \mathbb{R}^n$ is the nonlinear time-dependent part. Let $\boldsymbol{y}(t)$ be a trajectory of Eq. (10) that remains in a bounded region for all time. Then $\boldsymbol{y}(t)$ is said to be a distinguished hyperbolic trajectory if:

1. it is hyperbolic,
2. there exists a neighborhood $\mathcal{B}$ in the flow domain having the property that the DHT remains in $\mathcal{B}$ for all time, and all other trajectories starting in $\mathcal{B}$ leave $\mathcal{B}$ in finite time, as time evolves in either a positive or negative sense,
3. it is not a hyperbolic trajectory contained in the chaotic invariant set created by the intersection of the stable and unstable manifolds of another hyperbolic trajectory.

If the data spans only a finite time interval, then the DHT cannot be determined uniquely. Instead, there is a small region in $\mathcal{B}$ where the DHT can exist. We will present a method to obtain an approximation to the DHT, assuming that the time dependence of the velocity field persists outside the time-interval of the data set.

The second part of this definition can also be stated in terms of the stable and unstable manifolds of the DHT. Points on the stable manifold can leave $\mathcal{B}$ in negative time, points on the unstable manifold can leave $\mathcal{B}$ in positive time, and points on neither manifold leave $\mathcal{B}$ in both positive and negative time.

In the case where the DHT does not remain in a bounded region, the definition is more tricky. A definition for linear inhomogeneous systems can be given (provide one does not allow exponential growth or decay in the inhomogeneous term) as in Example 2. Moreover, the linear part of a general velocity field is not necessarily independent or constant. These lead to the second class of DHTs.

Definition 2. Let us consider the general velocity field given by Eq. (1). Let us assume that there exists an invertible coodinate change from $\boldsymbol{x}$ to $\boldsymbol{y}$, i.e. from Eq. (1) to Eq. (10), which is based on the movement of an Eulerian structure in $\boldsymbol{x}$, such as a path of an ISP. Let $\boldsymbol{y}(t)$ be a solution that satisfies the

three conditions given in Definition 1. Then the corresponding $\boldsymbol{x}(t)$ is said to be a distinguished hyperbolic trajectory of Eq. (1).

If $\boldsymbol{y}(t)$ is a DHT in the $\boldsymbol{y}$ coordinates, then the corresponding trajectory $\boldsymbol{x}(t)$ in the original velocity field is also a DHT, because DHTs are frame independent. We will discuss appropriate coordinate changes and frame independence intensively in Sect. 3 and the Appendices. However, the focus in this paper will be mainly on DHTs that remain in bounded regions.

Our task now will be to show that this definition does indeed satisfy the requirement of picking out the important hyperbolic trajectories for the application of Lagrangian transport theory using the method of lobe dynamics. This is the motivation for the third part of Definition 1. If the stable and unstable manifolds of a hyperbolic trajectory intersect transversely, then there is an associated lobe dynamics that describes the motion of trajectories through the homoclinic tangle. A consequence of the transverse intersection of the stable and unstable manifold is the formation of an invariant Cantor set on which the dynamics is chaotic, with all trajectories in the Cantor set being hyperbolic. However, for our purposes, we would not call the hyperbolic trajectories in the Cantor set "distinguished" as the transport of trajectories through this Cantor set is governed by the lobe dynamics associated with the hyperbolic trajectory whose transversely intersecting stable and unstable manifolds give rise to the hyperbolic Cantor set.

Linearization is such a basic analytical method that it would seem that little needs to be said about it. However, there are two types of linearization used throughout this work, each having a definite fluid dynamical interpretation, and we want to alert the reader to these two types of linearization that are interweaved throughout this paper.

Lagrangian versus Eulerian Linearization

Consider the velocity field given by Eq. (1). Let $\bar{\boldsymbol{x}}(t)$ be a trajectory of this velocity field and let $\boldsymbol{x}_0$ be a specified point in the domain. In order to check the stability of the trajectory $\boldsymbol{x}(t)$ we consider the velocity field linearized about the trajectory, i.e.

$$\frac{d}{dt}\boldsymbol{\xi} = \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}}\left(\boldsymbol{x}(t), t\right)\boldsymbol{\xi}.$$

Of course, this is standard in dynamical systems theory. From the point of view of fluid mechanics we are looking at the linearized behaviour around a fluid particle trajectory. This is the Lagrangian point of view in fluid mechanics, which is the reason for the term Lagrangian linearization that we apply in this case.

If instead we were to linearize the velocity field about the specified point $\bar{\boldsymbol{x}}$, we would obtain a linear system having the following form:

$$\frac{d}{dt}\boldsymbol{\xi} = \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}}\left(\bar{\boldsymbol{x}}, t\right)\boldsymbol{\xi} + \boldsymbol{u}\left(\bar{\boldsymbol{x}}, t\right).$$

This has the form of an inhomogeneous or "forced" linear system. When we consider fluid mechanical properties in a fixed region of space (as opposed to following fluid particle trajectories as they evolve in time), this is referred to as the Eulerian point of view of fluid mechanics, which is why we refer to linearization about a specified point as Eulerian linearization.

It will also prove useful to linearize the velocity field about the ISP, $\boldsymbol{x}_{\mathrm{sp}}(t)$. We will refer to this as instantaneous Eulerian linearization, and the linearized equation in this case takes the form:

$$\frac{d}{dt}\boldsymbol{\xi} = \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}}\left(\boldsymbol{x}_{\mathrm{sp}}(t), t\right)\boldsymbol{\xi} - \dot{\boldsymbol{x}}_{\mathrm{sp}}(t).$$

Both types of Eulerian linearization will be important when we search the flow for DHTs, and it is the properties of the inhomogeneous term of the associated linear equation (i.e. $\boldsymbol{u}\left(\bar{\boldsymbol{x}}, t\right)$) that are crucial for the existence of a DHT in a given region. However, hyperbolicity of a given trajectory is a Lagrangian linearization property. Hence, the interplay between Eulerian and Lagrangian linearization is a key element in our development of a constructive theory for DHTs.

Finite Time Velocity Fields

Here we want to re-emphasize that time-dependent velocity fields that are only known for a finite time interval, which we refer to as finite time velocity fields, are our main concern. For finite time velocity fields this method is simply not adequate, since it requires certain regions to converge to the hyperbolic trajectory in the appropriate direction of time. This procedure can "eat up" much of the data set in the process of converging to the hyperbolic trajectory. Moreover, it requires the integration of many trajectories. Even if the velocity field is time-periodic, convergence of the method can require integration of many trajectories through many periods. In the end, this can result in a very complicated geometric object whose complexity may make it difficult to determine the hyperbolic trajectory. The method developed in this paper eliminates these problems by providing an approximation to the DHT for the entire length of the time interval of the data set.

The definition of Hyperbolic Trajectory

Consider a velocity field given by Eq. (1) over a finite time interval $t \in [t_0, t_L]$. Let $\bar{\boldsymbol{x}}(t)$ be a trajectory of this velocity field. Hyperbolicity is a "linear property" in the sense that the hyperbolicity characteristics of a trajectory are determined from the linearization of the vector field about the trajectory. The velocity field linearized about the trajectory is given by:

$$\frac{d}{dt}\boldsymbol{\xi} = \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}}\left(\boldsymbol{x}(t), t\right)\boldsymbol{\xi}, \quad t \in [t_0, t_L]. \tag{11}$$

We let $\mathbf{X}(t, t_0)$ denote the fundamental solution matrix of this linear system, i.e. it is the matrix whose columns consist of linearly independent solutions of the linear system.

In the ordinary differential equations community, a type of finite time hyperbolicity has been known for some time as "exponential dichotomy". Roughly speaking, this means that trajectories can "separate" at an exponential rate. The formulation seems to be due to Massera and Schäffer (1966).

Definition 3. (**Exponential Dichotomy**). Equation (11) is said to possess an exponential dichotomy on $[t_0, t_L]$ if there exists a projection $\mathbf{P}$ (i.e. $\mathbf{P}^2 = \mathbf{P}$), and positive constants K, L, α, and β such that:

$$\begin{aligned} |\mathbf{X}(t,t_0)\mathbf{P}\mathbf{X}^{-1}(s,t_0)| &\le Ke^{-\alpha(t-s)}, \\ \text{for} \quad t \ge s, \quad & t,s \in [t_0,t_L], \\ |\mathbf{X}(t,t_0)(\mathbf{Id}-\mathbf{P})\mathbf{X}^{-1}(s,t_0)| &\le Le^{-\beta(s-t)}, \\ \text{for} \quad s \ge t, \quad & t,s \in [t_0,t_L]. \end{aligned} \tag{12}$$

Further references on exponential dichotomies are Coppel (1978), Henry (1981), and Muldowney (1984).

If the matrix associated with the linearized velocity field (11) is constant (such would be the case for a steady velocity field linearized about a stagnation point), then an exponential dichotomy would be equivalent to the property that none of the eigenvalues of the matrix had zero real parts.

Appendix A gives an equivalent definition of hyperbolicity that is more computationally oriented. In particular, we represent the fundamental solution in the form of a singular value decomposition:

$$\mathbf{X}(t,t_0) = \mathbf{B}(t,t_0)\exp\big(\mathbf{\Sigma}(t,t_0)\big)\mathbf{R}(t,t_0)^T,$$

where $\mathbf{B}(t,t_0)$ and $\mathbf{R}(t,t_0)$ are orthogonal matrices, i.e. $\mathbf{B}(t,t_0)\mathbf{B}(t,t_0)^T = \mathbf{R}(t,t_0)\mathbf{R}(t,t_0)^T = \mathbf{I}$, and $\mathbf{\Sigma}(t,t_0)$ is a diagonal matrix with $\mathbf{\Sigma}(t_0,t_0) = 0$ so that $\exp(\mathbf{\Sigma}(t,t_0))$ is a diagonal matrix with $\exp(\mathbf{\Sigma}(t_0,t_0)) = \mathbf{I}$. We then show that there exists a time-dependent, linear transformation:

$$\boldsymbol{y} = \mathbf{A}(t)\boldsymbol{\xi},$$

where

$$\begin{aligned} \mathbf{A}(t) = &\exp\left((t-t_0)\mathbf{D}\right)\mathbf{R}(t_L,t_0)^T\mathbf{R}(t,t_0) \\ &\cdot\exp\left(-\mathbf{\Sigma}(t,t_0)\right)\mathbf{B}(t,t_0)^T, \end{aligned}$$

which transforms (11) into the following form:

$$\dot{\boldsymbol{y}} = \mathbf{D}\boldsymbol{y},$$

where

$$\mathbf{D} = \frac{1}{t_L - t_0}\mathbf{\Sigma}(t_L,t_0).$$

The equivalence of the two definitions of hyperbolicity is established in Appendix C where it is shown that exponential dichotomy is a frame invariant property.

The outline of this paper is as as follows. In Sect. 2 we develop a quantitative theory for DHTs for inhomogeneous linear systems. The theory yields an analytical formula, and is the basis for the numerical method developed in Sect. 4. It also gives insight into the behaviour of DHTs in nonlinear systems described in Sect. 3. In Sect. 3 we develop a theory for the existence of DHTs for one- and two-dimensional nonlinear velocity fields. The conditions for existence are based on conditions for the instantaneous velocity field (i.e. properties of ISPs). In Sect. 4 we develop a numerical method that can be applied to either flows defined as a data set, or flows that can be expressed in the form of a mathematical formula involving known functions or quadratures. The key aspect of this numerical method is that it allows us to compute the DHT for the entire length of the data set.

2 Theory of distinguished hyperbolic trajectories for forced linear systems

2.1 Motivation for the linear theory

In this section, we develop a theory of distinguished hyperbolic trajectories for forced linear systems when the velocity field is available only over a finite time interval. It may seem rather trivial to study such linear systems. However, hyperbolicity is a property of linearized behaviour and, therefore, we feel it is important to understand it first in the purely linear setting. In particular, the definition of a distinguished hyperbolic trajectory, as well as finiteness of the time interval of the velocity data set, should be first addressed in this context. The linear theory will lead to an analytical formula for the DHT given a finite-time interval of velocity data. This formula will be used as the basis for our numerical method developed in a later section, and it plays an important role in understanding properties of DHTs in nonlinear systems.

We begin by considering a velocity field of the form

$$\frac{d}{dt}\boldsymbol{x} = \mathbf{F}(t)\boldsymbol{x} + \boldsymbol{h}(t), \tag{13}$$

where $\boldsymbol{x} \in \mathbb{R}^n$, $\mathbf{F}(t)$ is a $n \times n$ matrix, and $\boldsymbol{h}(t)$ is a n-vector forcing function. Both $\mathbf{F}(t)$ and $\boldsymbol{h}(t)$ are available only for $t \in [t_0, t_L]$. Applying the coordinate transformation constructed in Appendix A to Eq. (13) gives

$$\frac{d}{dt}\boldsymbol{y} = \mathbf{D}\boldsymbol{y} + \boldsymbol{g}(t), \tag{14}$$

where $\boldsymbol{y} \in \mathbb{R}^n$, $\boldsymbol{g}(t)$ is a n-vector forcing function available only for $t \in [t_0, t_L]$, and $\mathbf{D} \in \mathbb{R}^{n\times n}$ is a diagonal matrix:

$$D_{ij} = \begin{cases} d_i, & \text{for } i = j, \\ 0, & \text{for } i \ne j, \end{cases} \tag{15}$$

with real $d_i \ne 0$. Diagonalization of $\mathbf{F}(t)$ to $\mathbf{D}$ decouples the general n-dimensional problem into n independent, constant-coefficient one-dimensional problems. The general solution of Eq. (14) for y_i with initial condition $y_{i,0}$ at t_0 can be written as:

$$\begin{aligned} y_i(t;\, y_{i,0}, t_0) = &\, y_{i,\mathrm{dht}}(t) + Y_{ii}(t,t_0)y_{i,0} \\ &+ \begin{cases} -\int_{-\infty}^{t_0} Y_{ii}(t,\tau)g_i(\tau)d\tau & \text{for } d_i > 0, \\ \int_{t_0}^{\infty} Y_{ii}(t,\tau)g_i(\tau)d\tau & \text{for } d_i < 0, \end{cases} \end{aligned} \tag{16}$$

where $Y_{ii}(t,t_0)$ are the diagonal elements of the fundamental solution matrix:

$$Y_{ij}(t,t_0) = \begin{cases} \exp\{d_i(t-t_0)\}, & \text{for } i = j, \\ 0, & \text{for } i \ne j. \end{cases} \tag{17}$$

If we let $t_0 \to -\infty$ and $t_L \to \infty$, then it is straightforward to show (Henry,1981) that the unique solution that is bounded for all time is given by:

$$y_{i,\mathrm{dht}}(t) = \begin{cases} \int_{-\infty}^{t} Y_{ii}(t,\tau)g_i(\tau)d\tau & \text{for } d_i > 0, \\ -\int_{t}^{\infty} Y_{ii}(t,\tau)g_i(\tau)d\tau & \text{for } d_i < 0. \end{cases} \tag{18}$$

This solution satisfies Definition 1 in Sect. 1 and is the DHT.

A difficulty arises in computing $y_{i,\mathrm{dht}}(t)$ using Eq. (18) when $g_i(t)$ is available only over a finite time interval. A straightforward approach to estimating the DHT may be to rewrite Eq. (18) into two parts:

$$y_{i,\mathrm{dht}}(t) = \begin{cases} \int_{t_0}^{t} Y_{ii}(t,\tau)g_i(\tau)d\tau \ + \int_{-\infty}^{t_0} Y_{ii}(t,\tau)g_i(\tau)d\tau & \text{for } d_i > 0, \\ -\int_{t}^{t_L} Y_{ii}(t,\tau)g_i(\tau)d\tau \ - \int_{t_L}^{\infty} Y_{ii}(t,\tau)g_i(\tau)d\tau & \text{for } d_i < 0, \end{cases} \tag{19}$$

and use only the first term that can be computed from the available data. However, shortcomings of this approach are most apparent in two ways. One is the systematic error incurred by neglecting the second term. This error corresponds to a loss of uniqueness caused by a lack of data outside the time interval. The other is the unrealistic initial value $y_{i,\mathrm{dht}}(t_0) = 0$ for $d_i > 0$ and the final value $y_{i,\mathrm{dht}}(t_L) = 0$ for $d_i < 0$. It suggests that such an estimate for the DHT critically depends on the length of the data set.

In this paper, we propose an alternate approach for obtaining analytical formulae for the DHT by expressing $\boldsymbol{g}(t)$ as Fourier representation or power series. This has two advantages. One is that it allows us to overcome the finiteness of the data set since expressing $\boldsymbol{g}(t)$ as such a time series is equivalent to extending the time interval to infinity. The other advantage is that it provides an analytical formula for the DHT whose coefficients are determined by the data. However, there is an additional error associated with the difference between $\boldsymbol{g}(t)$ and its series representation. This issue is addressed later when we develop a numerical algorithm based on the linear formula (Sect. 4.3.1) and validate it using a data set (Sect. 4.3.3).

2.2 Time-independent system matrix with bounded forcing

We consider a n-dimensional linear system Eq. (14) whose velocity data set is available over $t \in [t_0, t_L]$. In addition to the hyperbolicity of $\mathbf{D}$, we make the following assumptions on the forcing $\boldsymbol{g}(t)$ Eq. (14).

Assumption 2.1 *Given the data over $t \in [t_0, t_L]$, $\boldsymbol{g}(t)$ is bounded and can be expressed in the following form:*

$$\boldsymbol{g}(t) = \sum_{k=0}^{K} \boldsymbol{g}^{(s,k)}(t) + \boldsymbol{g}^{(c,k)}(t), \tag{20}$$

where

$$\boldsymbol{g}^{(s,k)}(t) = \boldsymbol{b}^{(s,k)} \sin\omega^{(k)}t, \ \boldsymbol{g}^{(c,k)}(t) = \boldsymbol{b}^{(c,k)} \cos\omega^{(k)}t,$$
$$\omega \equiv \frac{2\pi}{t_L - t_0}, \qquad \omega^{(k)} \equiv k\omega, \tag{21}$$

and

$$(\boldsymbol{b}^{(s,k)}, \boldsymbol{b}^{(c,k)}) = \frac{1}{2\pi}\int_{t_0}^{t_L} \boldsymbol{g}(t)(\sin\omega^{(k)}t, \cos\omega^{(k)}t)dt. \tag{22}$$

Note that any bounded forcing $||\boldsymbol{g}(t)|| < g_{max}$ for a time interval $t \subset [t_0, t_L]$ can be expressed as a Fourier series. Moreover, this is equivalent to extending the forcing for an infinite time interval, therefore securing the uniqueness of the DHT.

2.2.1 Instantaneous Stagnation Point (ISP)

As discussed earlier, ISPs do not necessarily follow fluid particle trajectories. However, we now show that there is a quantitative relationship between DHTs and ISPs that can be expressed by an analytical formula. This is significant because computation of ISPs is relatively straightforward, as we shall see in Sect. 4.

By definition, the ISPs are given by

$$\boldsymbol{y}_{\mathrm{sp}}(t) \equiv -\mathbf{D}^{-1}\boldsymbol{g}(t). \tag{23}$$

The temporal mean of the ISP is given by

$$\overline{\boldsymbol{y}}_{\mathrm{sp}} \equiv -\mathbf{D}^{-1}\frac{1}{t_L - t_0}\int_{t_0}^{t_L} \boldsymbol{g}(\tau)d\tau. \tag{24}$$

For the bounded forcing in the form described above, the ISP and its temporal mean take the explicit form:

$$\boldsymbol{y}_{\mathrm{sp}}(t) = -\mathbf{D}^{-1}\sum_{k=0}^{K}(\boldsymbol{g}^{(s,k)}(t) + \boldsymbol{g}^{(c,k)}(t)) \tag{25}$$

$$\overline{\boldsymbol{y}}_{\mathrm{sp}} = -\mathbf{D}^{-1}\boldsymbol{b}^{(c,0)}. \tag{26}$$

Based on these expressions, we can make the following observations:

- The ISP (as a function of time) and the forcing are proportionally related in each variable y_i with factor d_i for $i = 1, \ldots, n$. Accordingly, if $\boldsymbol{y}_{\mathrm{sp}}(t)$ is bounded, so is $\boldsymbol{g}(t)$. This fact implies that the ISP provides a measure of the boundedness of forcing, when the forcing is unknown.
- From Eq. (24), we see that the time mean of the ISP is directly related to the time mean of the forcing as well.

Example 3. Consider the one-dimensional vector field with a sinusoidal forcing

$$\frac{d}{dt}y = dy + b\sin\omega t, \tag{27}$$

where d, b, and ω are real numbers. It is a generalization of Example 2 in Sect. 1. The trajectory through an arbitrary initial condition $x(0)$ at $t = 0$ is given by

$$y(t) = -\frac{b}{\sqrt{d^2+\omega^2}}\sin(\omega t + \alpha) + e^{dt}\left(\frac{\omega b}{d^2+\omega^2} + x(0)\right), \tag{28}$$

and one easily verifies that all trajectories are hyperbolic. The distinguished hyperbolic trajectory is given by

$$y_{\mathrm{dht}}(t) = -\frac{b}{\sqrt{d^2+\omega^2}} \sin(\omega t + \alpha), \tag{29}$$

where the phase shift α is given by

$$\tan\alpha = \frac{\omega}{d}, \tag{30}$$

with $\omega\alpha \geq 0$ (with respect to the general notation, $\omega^{(k)} = \omega^{(1)} = \omega$) and $-\frac{\pi}{2} < \alpha \leq \frac{\pi}{2}$. Similar results for the DHT with a cosine forcing also hold.

2.2.2 The Distinguished Hyperbolic Trajectory and its relation to the instantaneous stagnation points

Since the forcing Eq. (14) is assumed to be represented as a Fourier series (recall Eq. (20)), the superposition principle for linear systems allows us to solve Eq. (18) analytically:

$$\boldsymbol{y}_{\mathrm{dht}}(t) = -\mathbf{D}^{-1} \sum_{k=0}^{K} (\hat{\boldsymbol{g}}^{(s,k)}(t) + \hat{\boldsymbol{g}}^{(c,k)}(t)) \tag{31}$$

where

$$(\hat{g}_i^{(s,k)}(t), \hat{g}_i^{(c,k)}(t)) = r_i^{(k)}(g_i^{(s,k)}(t+\alpha_i^{(k)}), g_i^{(c,k)}(t+\alpha_i^{(k)})), \tag{32}$$

and

$$r_i^{(k)} \equiv \frac{1}{\sqrt{1+(\frac{\omega^{(k)}}{d_i})^2}},$$
$$\tan\alpha_i^{(k)} \equiv \frac{\omega^{(k)}}{d_i}, \tag{33}$$

with $\omega^{(k)}\alpha_i^{(k)} \geq 0$ and $-\frac{\pi}{2} < \alpha_i^{(k)} \leq \frac{\pi}{2}$. for $i = 1, \ldots, n$; d_i and $\omega^{(k)}$ are defined in Eqs. (15) and (21), respectively.

We now make the following observations:

- We see from Eq. (31) that the DHT is given analytically with well-known functions. This will also be very advantageous when we describe our numerical method for finding DHTs in data sets.
- By comparing Eq. (25) and Eq. (31), we see that there is a direct relationship between the ISP and the DHT. This relation between the DHT and ISP is described by the time-scale ratio between the forced and the unforced dynamics, $\omega^{(k)}/d_i$, which controls:
  1. phase-shift $\alpha_i^{(k)}$,
  2. amplitude ratio $r_i^{(k)}$.

  We will see in the next section that when solving for the DHT in a nonlinear system, we will choose one of two localization procedures. The particular choice will depend on the time-scale ratio $\omega^{(k)}/d_i$.

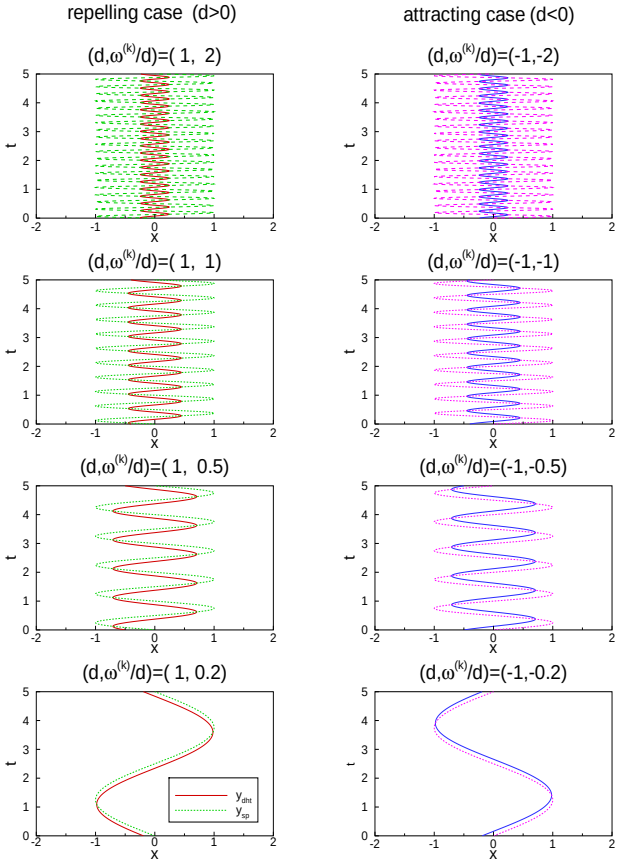

Fig. 3. Graphs of the DHT ($y_{\mathrm{dht}}(t)$) and the ISP ($y_{\mathrm{sp}}(t)$) are shown by solid and thin-dashed lines, respectively, in each panel for various time-scale ratios ($\omega^{(k)}/d$) with a fixed forcing amplitude ($b = 1$). The left and right columns correspond to repelling ($d > 0$) and attracting ($d < 0$) cases, respectively; the absolute value of the time-scale ratios are in decreasing order from the top to the bottom rows, i.e. $|\omega^{(k)}/d|$=2, 1, 0.5 and 0.2.

Here we present two simple examples in the one- and two-dimensional cases that will concretely demonstrate some basic geometrical relations between the DHT and the ISP. We will see later that these relations also hold for nonlinear velocity fields (Sect. 3) and lie at the foundation of our numerical method (Sect. 4).

Example 4. One Dimension: Time-scale dependence for the Repelling ($d > 0$) and Attracting ($d < 0$) Cases. Equation (14) in one dimension is:

$$\frac{d}{dt}y = dy + g(t), \tag{34a}$$

where

$$g(t) = b\cos\omega^{(k)}t. \tag{34b}$$

The geometrical relation between the DHT and ISP, and its dependence on parameters is illustrated in Fig. 3. Specific parameter sets $(d, \omega^{(k)}, b)$ are chosen for fixed $(|d|, b) = (1, 1)$, but with four values of $\omega^{(k)}$ so that the amplitude

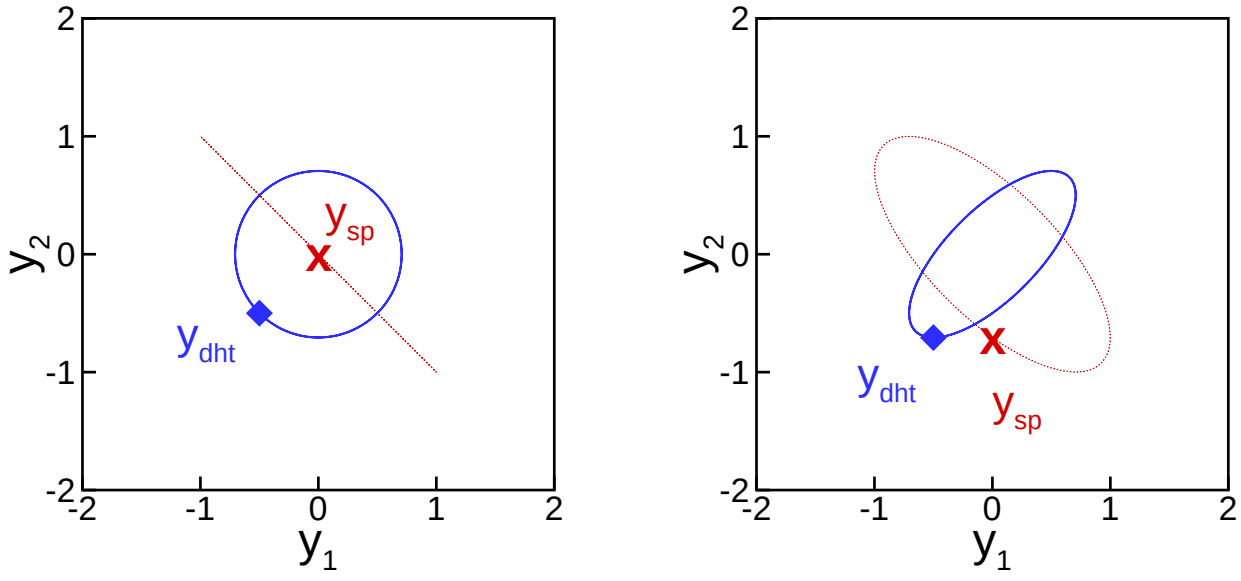

Fig. 4. DHT and path of ISP in time with the corresponding initial position indicated by the diamond (DHT) and X (ISP) for the two different phase values; $(a_1, a_2) = (0, 0)$ and $(a_1, a_2) = (0, -0.25)$ for the left and right panel, respectively. Other parameter values for the saddle are the same for the two panels with $(d_1, \omega_1/d_1, b_1) = (1, 1, 1)$ for the repelling direction and $(d_2, \omega_2/d_2, b_2) = (-1, 1, 1)$ for the attracting direction, and the trajectories move in a clockwise sense.

of the ISP is the same for all cases but the time-scale ratio $\omega^{(k)}/d$, which governs the relation between the DHT and ISP, (i.e. $r^{(k)}$ and $\alpha^{(k)}$) varies.

We make the following observations:

– The DHT has the same harmonic functional form as the ISP, but with an amplitude factor and a phase shift. Both the DHT and the ISP have the same periodicity as the original bounded forcing.

– The amplitude factor $r^{(k)}$ (≤ 1, always) has the following behaviour:

$$r^{(k)} \begin{cases} \to 1, & \text{for } |\omega^{(k)}/d| \to 0^+ \text{ (slow forcing)}, \\ \to 0^+, & \text{for } |\omega^{(k)}/d| \to \infty \text{ (fast forcing)}. \end{cases} \tag{35a}$$

The slower the bounded forcing is, the wider the range of DHT movement. The largest range of the movement must lie within the range of ISP movement. When the forcing is fast, the DHT hardly moves.

– The phase shift exhibits the following behaviour:

$$\alpha^{(k)} \begin{cases} \to \pi/2, & \text{for } \omega^{(k)}/d \to \infty \text{ (fast forcing; divergent dynamics)}, \\ = \pi/4, & \text{for } \omega^{(k)}/d = 1, \\ \to 0^+, & \text{for } \omega^{(k)}/d \to 0^+ \text{ (slow forcing; divergent dynamics)}, \\ \to 0^-, & \text{for } \omega^{(k)}/d \to 0^- \text{ (slow forcing; convergent dynamics)}, \\ = -\pi/4, & \text{for } \omega^{(k)}/d = -1, \\ \to -\pi/2, & \text{for } \omega^{(k)}/d \to -\infty \text{ (fast forcing; convergent dynamics)}. \end{cases} \tag{35b}$$

When the flow dynamics is divergent (convergent), the DHT proceeds (follows) the ISP. This is a general phenomenon that will hold for all one-dimensional velocity fields under general assumptions, as we will show later. Furthermore, the slower the bounded forcing is, the lesser the phase shift; this is because the DHT has more time to adjust to the forcing with which the trace is in phase. As the forcing becomes faster, the DHT and ISP become completely out of phase; the amplitude of the DHT oscillation is nearly zero.

The results in Eqs. (35a) and (35b) can be physically interpreted as an impedance factor between the ISP and DHT. Haller and Poje (1998) observed that when the time-dependence of the velocity field is "slow enough" with respect to the Lagrangian time-scale, the ISP stays near the hyperbolic trajectory. This corresponds to a special case of our theory, i.e. for $\omega^{(k)}/d \to 0^{\pm}$ for a short time interval; see Eqs. (35a) and (35b) above and also the bottom two panels of Fig. 3. However, as the forcing becomes faster, i.e. for $|\omega^{(k)}/d| \to \infty$, the DHT no longer follows the ISP and

becomes almost stationary.

Accordingly, there are two approaches to obtain the DHT in a nonlinear system given an ISP. One is to linearize around the ISP by moving with it when the forcing is slow. Another is to linearize around a stationary point in space given by the time-averaged ISP when the forcing is fast. We will discuss the nonlinear theory and conditions for the choice of approach in the next section.

Example 5. Two-Dimensional Saddle Case: Equation (14) in two dimensions is:

$$\frac{d}{dt}y_1 = d_1 y_1 + g_1(t)$$
$$\frac{d}{dt}y_2 = -d_2 y_2 + g_2(t), \tag{36}$$

where $g_i(t) = b_i \cos(\omega_i t + \pi a_i)$, and $d_i > 0$ for $i = 1, 2$ where the subscript i relate to x_i.

Since this two-dimensional equation has been decoupled into two one-dimensional equations, the same behaviour holds for each component as described in the previous one-dimensional example. However, two-dimensional DHT trajectories can be rather counterintuitive based on the ISP, and the difference may significantly depend on the phase difference in the forcing, as illustrated in Fig. 4.

2.3 Time-independent system matrix with unbounded forcing

Now we will consider the case where the forcing can be unbounded. However, it will only be allowed to grow algebraically in time.

Assumption 2.2

1. *The divergence rate of the forcing is slower than the exponential so as to not conflict with the exponential dichotomy around the DHT. In particular, we assume that given the data set over a finite time $t \in [t_0, t_L]$, the forcing $\boldsymbol{g}(t)$ can be expressed in the following form:*

$$\boldsymbol{g}^{(u)}(t) = \sum_{k=0}^{K} \boldsymbol{g}^{(u,k)}(t) = \sum_{k=0}^{K} \boldsymbol{b}^{(u,k)} t^k, \tag{37}$$

   where $\boldsymbol{b}^{(u,k)}$ is the constant coefficient for polynomial forcing.

2. *Still, realistically speaking in geophysical flows, the fastest unbounded forcing is linear, i.e. $k = 1$. It corresponds to a translating coherent structure system with quasi-uniform velocity.*

The ISPs are given by

$$\boldsymbol{y}_{\mathrm{sp}}(t) = -\mathbf{D}^{-1} \sum_{k=0}^{K} \boldsymbol{b}^{(u,k)} t^k. \tag{38}$$

The distinguished hyperbolic trajectory is obtained by solving Eq. (18) analytically:

$$\boldsymbol{y}_{\mathrm{dht}}(t) = \sum_{k=0}^{K} \hat{\boldsymbol{g}}^{(u,k)}(t), \tag{39}$$

where

$$\hat{g}_i^{(u,k)}(t) = -b_i^{(u,k)} \sum_{m=0}^{k} \frac{k!}{d_i^{(m+1)}(k-m)!} t^k. \tag{40}$$

Example 6: Consider the general case of Example 1 discussed in Sect. 1.

$$\frac{d}{dt}y = dy + bt \tag{41a}$$
$$y_{\mathrm{sp}}^{(u)}(t) = -\frac{b}{d}t \tag{41b}$$
$$y_{\mathrm{dht}}^{(u)}(t) = -\frac{b}{d}t - \frac{b}{d^2}. \tag{41c}$$

Note that the DHT is always a constant distance $-b/d^2$ away from the ISP, as illustrated in Fig. 1.

2.4 Time-dependent system matrix

We now consider the more general case of a system with time-dependent coefficients:

$$\frac{d}{dt}\boldsymbol{x} = \mathbf{F}(t)\boldsymbol{x} + \boldsymbol{h}(t), \tag{42}$$

where $\mathbf{F}(t)$ is the time-dependent Jacobian of the linear system, which is not necessarily diagonal. The ISP is given by:

$$\boldsymbol{x}_{\mathrm{sp}}(t) \equiv -\mathbf{F}(t)^{-1}\boldsymbol{h}(t). \tag{43}$$

where it is assumed that $\mathbf{F}(t)$ is invertible over the time interval of interest, $t \in [t_0, t_L]$.

This linear system has no analytical solution in general, even when $\boldsymbol{h}(t)$ can be expressed as a Fourier series or algebraic function of t. The difficulty in computing the DHT arises from two main sources associated with $\mathbf{F}(t)$, as we shall see in an example. One is dimensionality. Unlike the case of a one-dimensional system, the stability type of a trajectory can be not only hyperbolic or neutral, but also elliptic. Another is time dependence. Unlike the case with a time-independent coefficient matrix, the instantaneous eigenvalues of $\mathbf{F}(t)$, or the instantaneous geometry of the linear velocity field may no longer reflect the stability type over the time interval $t \in [t_0, t_L]$.

To overcome the difficulty and compute the DHT $\boldsymbol{x}_{\mathrm{dht}}(t)$ of Eq. (42), we carry out the following steps:

Step 1. Seek a time-dependent coordinate transform:

$$\boldsymbol{y}(t) = \mathbf{A}(t)\boldsymbol{x}(t), \tag{44}$$

so that the resulting linear system for $\boldsymbol{y}(t)$ is a decoupled, time-independent coefficient system as in Eq. (14).

Step 2. Determine the stability type in the $\boldsymbol{y}$ system based on the time-independent Jacobian matrix $\mathbf{D}$.

Step 3. If the stability type is hyperbolic in $\boldsymbol{y}$, then compute the DHT $\boldsymbol{y}_{\rm dht}(t)$ for the $\boldsymbol{y}$ system using the methodology described in Sects. 2.2 or 2.3. If the stability type is elliptic, then our method fails.

Step 4. Invert the coordinate transformation given by Eq. (44) to obtain the DHT $\boldsymbol{x}_{\rm dht}(t)$ in the $\boldsymbol{x}$ system:

$$\boldsymbol{x}_{\rm dht}(t) = \mathbf{A}(t)^{-1}\boldsymbol{y}_{\rm dht}(t). \tag{45}$$

We show in Appendix A how to accomplish Step 1, i.e. we develop a methodology to compute $\mathbf{A}(t)$ and the resulting system defined by $\mathbf{D}$ and $\boldsymbol{g}(t)$, given a time-dependent coefficient system defined by $\mathbf{F}(t)$ and $\boldsymbol{h}(t)$. In Appendix B, we prove that $\mathbf{A}(t)$ obtained in this way indeed takes a trajectory $\boldsymbol{y}(t)$ of Eq. (14) into a trajectory $\boldsymbol{x}(t)$ of Eq. (42); and hence $\boldsymbol{x}_{\rm dht}(t)$ of Eq. (45) is indeed a DHT in the $\boldsymbol{x}$ system. In other words, DHTs are frame-independent (Haller, 2001). On the contrary, the ISPs are generally frame-dependent; and hence $\boldsymbol{x}_{\rm sp}(t)$ and $\boldsymbol{y}_{\rm sp}(t)$ do not need to correspond to each other under this transformation, i.e. they are frame-dependent:

$$\boldsymbol{x}_{\rm sp}(t) \neq \mathbf{A}(t)^{-1}\boldsymbol{y}_{\rm sp}(t). \tag{46}$$

Example 7: Consider the velocity field given by:

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \delta_1 + \delta_2\cos 2\beta t & \delta_2 \sin 2\beta t - \beta \\ \delta_2 \sin 2\beta t + \beta & \delta_1 - \delta_2 \cos 2\beta t \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} b_1\cos(\omega_1 t + a_1) \\ b_2\cos(\omega_2 t + a_2) \end{pmatrix} \tag{47a}$$

where $\mathbf{F}(t)$ depends on the parameters β, (δ_1, δ_2) and $\boldsymbol{h}(t)$ depends on the parameters (ω_1, ω_2), (a_1, a_2) and (b_1, b_2). Here the subscripts i are related to x_i for $i = 1, 2$. We assume that $\delta_1^2 - \delta_2^2 + \beta^2 \neq 0$ so that an ISP exists:

$$\boldsymbol{x}_{\rm sp}(t) = -\frac{1}{\delta_1^2 - \delta_2^2 + \beta^2}\begin{pmatrix} b_1(\delta_1 - \delta_2\cos 2\beta t)\cos(\omega_1 t + a_1) + b_2(\delta_2 \sin 2\beta t + \beta)\cos(\omega_2 t + a_2) \\ b_1(\delta_2 \sin 2\beta t - \beta)\cos(\omega_1 t + a_1) + b_2(\delta_1 + \delta_2\cos 2\beta t)\cos(\omega_2 t + a_2) \end{pmatrix}. \tag{47b}$$

The two instantaneous eigenvalues $\lambda_{\rm F}^{\pm}$ of $\mathbf{F}(t)$ are time-independent:

$$\lambda_{\rm F}^{\pm} = \delta_1 \pm \sqrt{\delta_2^2 - \beta^2}, \tag{48}$$

and hence the instantaneous flow structure around $\boldsymbol{x}_{\rm sp}(t)$ appears to be hyperbolic for $\delta_1 = 0$, $\delta_2^2 > \beta^2$ and elliptic for $\delta_1 = 0$, $\delta_2^2 < \beta^2$ for any t.

The stability type of the trajectories over a time interval $t \in [t_0, t_L]$ in a linear velocity field can be graphically examined by the evolution of a circle put into the velocity field. If the stability type is hyperbolic, then a circle will deform to an ellipse at exponential rates, where exponential growth of the semi-major axis and exponential decay of the semi-minor axis correspond to twice the finite-time Lyapunov exponents along the principle axis (Appendix A). If the stability type is elliptic, then a circle will remain near a circle and rotate without exponential deformation. The area of the circle remains constant if and only if the trace of the Jacobian $\mathbf{F}(t)$ is zero.

Here we consider the case $\delta_1 = 0$, $\delta_2^2 < \beta^2$ (Fig. 5a) where the instantaneous velocity has an elliptic type structure around $\boldsymbol{x}_{\rm sp}(t)$ at any time t. However, a unit circle put around $\boldsymbol{x} = (0, 0)$ at time $t = 0$ undergoes a counterclockwise rotation and it deforms at an exponential rate in time (Fig. 5b), suggesting that the trajectories are hyperbolic and the DHT may hence exist. We therefore apply the 4-step procedure described above.

Using the methodology in Appendix A, we compute the coordinate change (Step 1):

$$\mathbf{A}(t) = \begin{pmatrix} \cos\beta t & \sin\beta t \\ -\sin\beta t & \cos\beta t \end{pmatrix} \tag{49}$$

and the time-independent coefficient system:

$$\begin{aligned} \frac{d}{dt}y_1 &= d_1 y_1 + b_1\cos(\omega_1 t + a_1)\cos\beta t + b_2\cos(\omega_2 t + a_2)\sin\beta t\ , \\ \frac{d}{dt}y_2 &= d_2 y_2 + b_2\cos(\omega_2 t + a_2)\cos\beta t - b_1\cos(\omega_1 t + a_1)\sin\beta t\ , \end{aligned} \tag{50}$$

where $(d_1, d_2) = (\delta_1 + \delta_2, \delta_1 - \delta_2)$.

For $\delta_1 = 0$, $\delta_1^2 < \delta_2^2$, the two eigenvalues for the $\boldsymbol{y}$ system are real and of opposite signs and the trajectories are hyperbolic (Step 2). In fact, the instantaneous velocity field in the $\boldsymbol{y}$ system has a hyperbolic type structure, as shown in Fig. 6a (compare with Fig. 5a). A unit circle put around $\boldsymbol{y} = (0, 0)$ at $t = 0$ deforms to an ellipse at an exponential rate (Fig. 6b). However, the semi-major and semi-minor axes of the ellipse

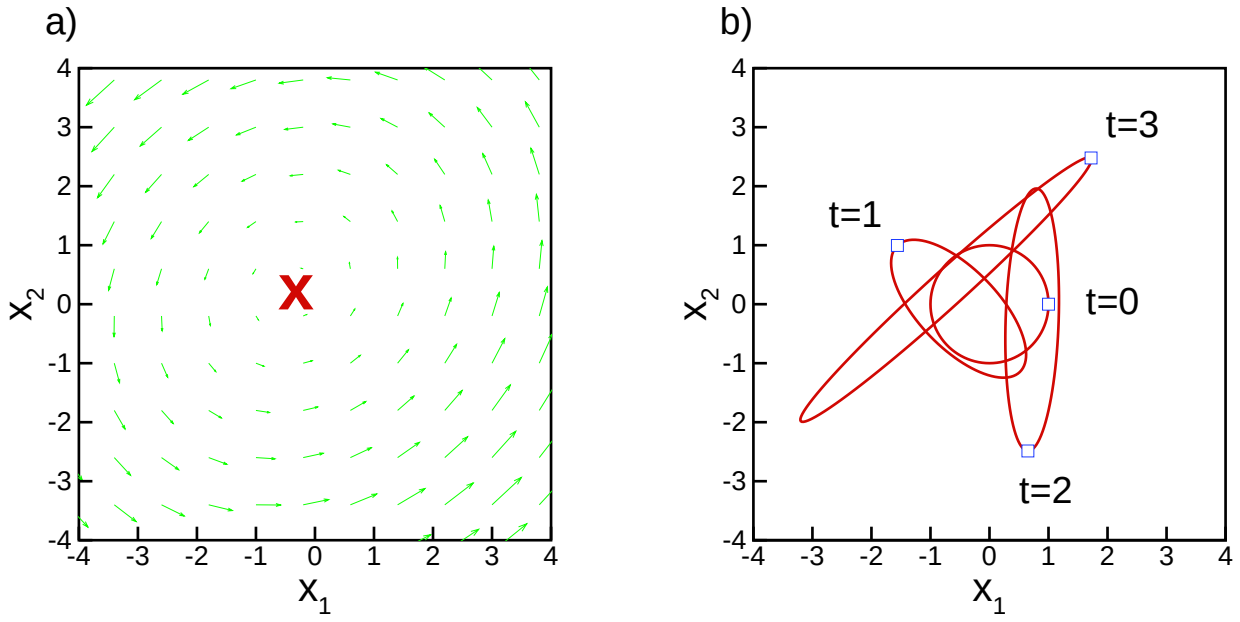

Fig. 5. Time-dependent coefficient system in $\boldsymbol{x}$ for a $\delta_1 = 0$, $\delta_2^2 < \beta^2$ case, with a parameter set $\beta = 2.34$, $(\delta_1, \delta_2) = (0.0, -0.4)$, $(\omega_1, \omega_2) = (3.23, 3)$, $(a_1, a_2) = (0, 0)$, $(b_1, b_2) = (0.8, 0.9)$ a) instantaneous velocity field in $\boldsymbol{x}$ with a red cross for $\boldsymbol{x}_{\rm sp}(0)$; b) evolution of a set of trajectories starting at $t = 0$ on a unit circle around the origin for $t \in [0, 3]$, where the square on each ellipse is at the semi-major axis.

are aligned with the axes of the $\boldsymbol{y}$ system at any t, and the ellipse does not rotate (compare with Fig. 5b). This is one of the two roles played by the coordinate transform $\mathbf{A}(t)$: to suppress the rotation of the principle axis, and to regulate the exponential behaviour to be uniform in time.

Using the methodology described in Sect. 2.2, we obtain the DHT $\boldsymbol{y}_{\rm dht}(t)$ (Step 3):

$$y_{1,\rm dht}(t) = -\frac{1}{2d_1}\{b_1[r_{11}^+ \cos(\omega_1^+ t + \alpha_{11}^+ + a_1) + r_{11}^- \cos(\omega_1^- t + \alpha_{11}^- + a_1)] + b_2[r_{12}^+ \sin(\omega_2^+ t + \alpha_{12}^+ + a_2) - r_{12}^- \sin(\omega_2^- t + \alpha_{12}^- + a_2)]\}\,, \tag{51a}$$

$$y_{2,\rm dht}(t) = -\frac{1}{2d_2}\{b_2[r_{22}^+ \cos(\omega_2^+ t + \alpha_{22}^+ + a_2) + r_{22}^- \cos(\omega_2^- t + \alpha_{22}^- + a_2)] - b_1[r_{21}^+ \sin(\omega_1^+ t + \alpha_{21}^+ + a_1) - r_{21}^- \sin(\omega_1^- t + \alpha_{21}^- + a_1)]\}\,, \tag{51b}$$

where

$$\omega_j^\pm = \omega_j \pm \beta\,, \quad r_{ij}^\pm = \frac{1}{\sqrt{1 + \left(\frac{\omega_j^\pm}{d_i}\right)^2}}\,, \quad \alpha_{ij}^\pm = \tan^{-1}\frac{\omega_j^\pm}{d_i}\,. \tag{51c}$$

Figure 7 shows the time evolution of $\boldsymbol{y}_{\rm dht}(t)$ and $\boldsymbol{y}_{\rm sp}(t)$ for $t \in [0, 3]$, where $\boldsymbol{y}_{\rm sp}(t)$ is the ISP of Eq. (50). As discussed in Sect. 3.1 for the nonlinear case for each $i = 1, 2$, $y_{i,\rm dht}(t)$ lies within the range defined by the minimum and maximum of $y_{i,\rm sp}(t)$. Also, $y_{i,\rm dht}(t)$ changes direction only when it intersects with $y_{i,\rm sp}(t)$.

Finally, the DHT $\boldsymbol{x}_{\rm dht}(t)$ is obtained through the inverse coordinate transform $\mathbf{A}(t)^{-1}$ (Step 4). A straightforward substitution confirms that $\boldsymbol{x}_{\rm dht}(t)$ is a solution of Eq. (47a) and hence the DHT is frame-independent. Figure 8 shows the time evolution of $\boldsymbol{x}_{\rm dht}(t)$ and $\boldsymbol{x}_{\rm sp}(t)$. In the time-dependent coefficient system, $\boldsymbol{x}_{\rm dht}(t)$ and $\boldsymbol{x}_{\rm sp}(t)$ no longer need to satisfy the relation discussed in Sect. 3.1 and hence it is quite possible that $\boldsymbol{x}_{\rm dht}(t)$ does not necessarily lie within the region where $\boldsymbol{x}_{\rm sp}(t)$ is observed over the time interval.

3 Nonlinear theory of DHTs and their relation to ISPs

In this section we will obtain some results for the existence of DHTs for nonlinear velocity fields with emphasis on their

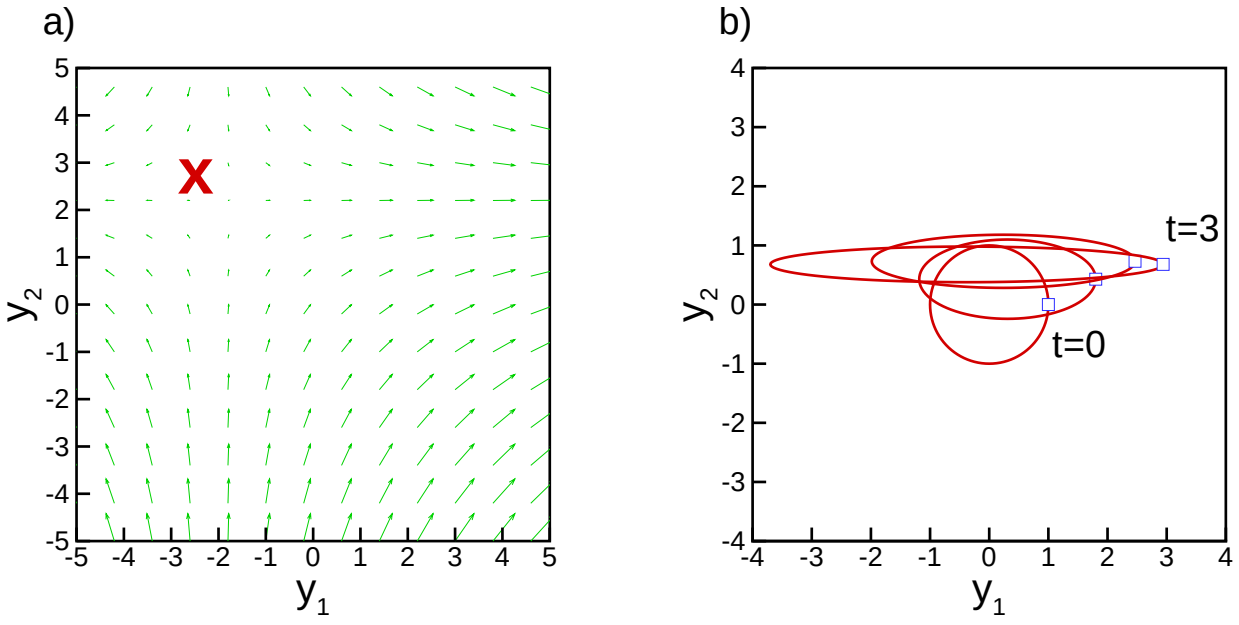

Fig. 6. Time-independent coefficient system in $\boldsymbol{y}$ corresponding to Fig. 5.

relation to ISPs. This is critical in the numerical identification of DHTs, because ISPs are observable in the velocity field but DHTs are not. We will also show how these results are related to the linear results in the previous section, so that these results can be used to develop a numerical method based on there results in the next section. An important theme of our approach is that we are able to deduce information about time varying Lagrangian structures from a sequence of the instantaneous (or "frozen time") velocity field. This will be apparent in the assumptions that we make about the instantaneous velocity field. For one-dimensional velocity fields these assumptions will be manifested in the form of assumptions on the ISPs.

We will first treat one-dimensional velocity fields which is relevant for finding DHTs on the boundaries of two-dimensional flows with free-slip boundary conditions, and then two-dimensional velocity fields.

3.1 One-dimensional velocity fields

We consider a one-dimensional velocity field of the form Eq. (1) defined for $t \in [t_{L^-}, t_{L^+}]$, i.e.

$$\frac{d}{dt}x = u(x,t), \qquad x \in \mathbb{R}. \tag{52}$$

The velocity field should be (at least) continuous in time, and we will require it to be differentiable in space (so that linearization makes sense). In addition, we make the following assumptions on the ISPs.

Existence of an ISP: For $t \subset [t_0, t_L]$ there exists an ISP, denoted by $x_{\rm sp}(t)$, i.e. a function satisfying

$$u(x_{\rm sp}(t), t) \equiv 0. \tag{53}$$

Note that, in general, $x_{\rm sp}(t)$ is not a particle trajectory. In particular, if $x_{\rm sp}(t)$ were a particle trajectory, then it must satisfy the equation

$$\frac{d}{dt}x_{\rm sp}(t) = u(x_{\rm sp}(t), t) = 0,$$

which would imply that $x_{\rm sp}(t)$ is constant in time. In the applications that we will consider, $x_{\rm sp}(t)$ generally changes position as t varies.

Isolated ISP: Let

$$(x_{\rm sp}^{\rm min}, x_{\rm sp}^{\rm max}) = \left(\min x_{\rm sp}(t), \max x_{\rm sp}(t)\right),$$

where the maximum and minimum is taken over $t \in [t_0, t_L]$, and we assume that $x_{\rm sp}^{\rm min}$ and $x_{\rm sp}^{\rm max}$ are bounded. We assume that in the box in the $x-t$ plane defined by

$$x_{\rm sp}^{\rm min} \le x \le x_{\rm sp}^{\rm max}, \qquad t_0 \le t \le t_L, \tag{54}$$

there are no other ISPs.

Hyperbolicity of the box containing the ISP: We will assume that that $\partial u/\partial x$ does not vanish in the whole box as defined above.

Hence, there are two cases:

$$\begin{aligned} &\text{repelling:} \quad \frac{\partial u}{\partial x}(x,t) > 0, \\ &\text{attracting:} \quad \frac{\partial u}{\partial x}(x,t) < 0. \end{aligned} \tag{55}$$

We remark that the assumption that the ISPs are isolated and hyperbolic rule out the subject of bifurcation of ISPs and the consequences of this for DHTs. This is an important topic for applications but is outside the scope of this work.

We are now ready to state and prove the theorem.

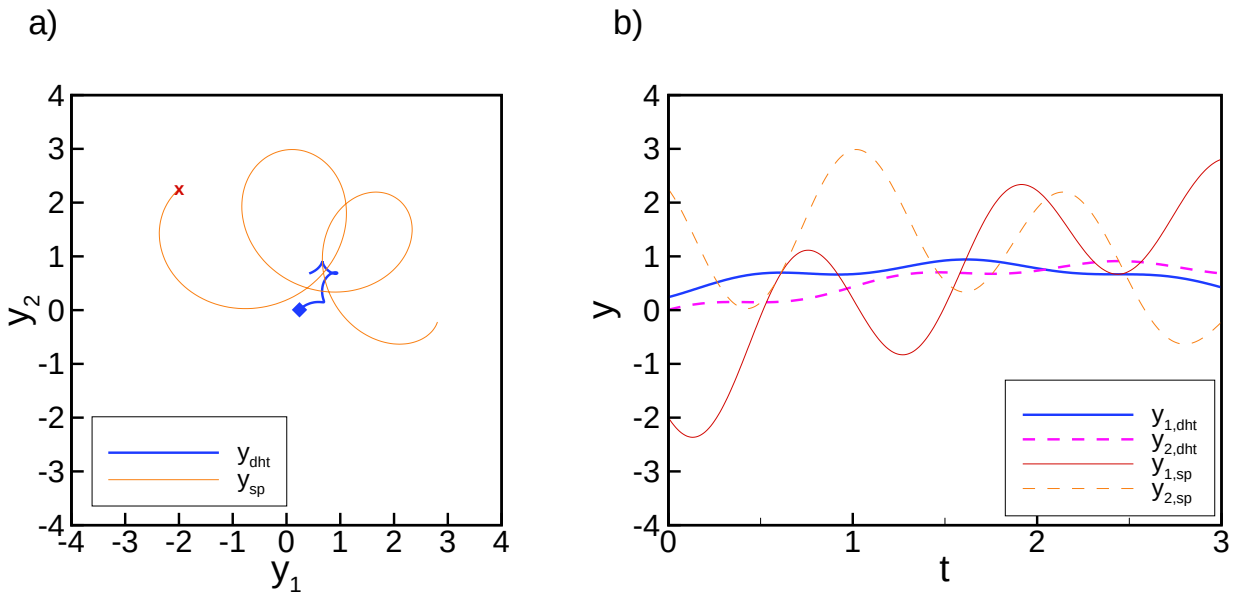

Fig. 7. The DHT and ISP for $t \in [0, 3]$ in $\boldsymbol{x}$: **(a)** phase space plot where the initial locations are indicated by a square for $\boldsymbol{x}_{\rm dht}(0)$ and a cross for $\boldsymbol{x}_{\rm sp}(0)$ (see Fig. 6a); **(b)** time series.

Theorem 3.1 *1. For $t_L = \infty$ there exists a DHT, $x_{\rm dht}(t)$, having the same stability properties as $x_{\rm sp}(t)$. This DHT satisfies the bounds*

$$x_{\rm sp}^{\rm min} \le x_{\rm dht}(t) \le x_{\rm sp}^{\rm max}, \qquad t_0 \le t < \infty,$$

and is unique in the sense that it is the only DHT satisfying these bounds.

2. For t_L finite, essentially the same result as 1 holds. However, in this case, uniqueness of the DHT does not necessarily hold. Rather, there exists an interval of initial conditions, $\mathcal{I}$, satisfying

$$x_{\rm sp}^{\rm min} \le \mathcal{I} \le x_{\rm sp}^{\rm max},$$

such that trajectories in this interval at t_0 satisfy these bounds for $t \in [t_0, t_L]$ and have the same stability properties as $x_{\rm sp}(t)$.

Proof of 1: We prove 1 for the repelling case. The proof for the attracting case is similar.

Since $x_{\rm sp}(t)$ is assumed to be isolated in the box defined by $\{(x, t) \mid x_{\rm sp}^{\rm min} \le x \le x_{\rm sp}^{\rm max},\ t_0 \le t < \infty\}$, we can choose an $\epsilon > 0$ such that $x_{\rm sp}(t)$ is isolated in the box defined by $\{(x, t) \mid x_{\rm sp}^{\rm min} - \epsilon \le x \le x_{\rm sp}^{\rm max} + \epsilon,\ t_0 \le t < \infty\}$, see Fig. 9. Moreover, since $x_{\rm sp}(t)$ is repelling and there are no other ISPs to reverse the direction of the instantaneous velocity than $x_{\rm sp}(t)$ itself, points on the left vertical boundary of the box move off the boundary to the left and points on the right vertical boundary move off the boundary to the right.

Let $x(t, t_0, x_0)$ denote the solution of Eq. (52) through the point $x = x_0$ at $t = t_0$. Let I_{t_0} denote the spatial interval $x_{\rm sp}^{\rm min} - \epsilon \le x \le x_{\rm sp}^{\rm max} + \epsilon$. Then $I_t \equiv x(t, t_0, I_{t_0})$, $t \ge t_0$, denotes the time evolution of the interval I_{t_0} under the trajectories of Eq. (52).

Choose some $T > 0$ and consider the spatial intervals at time $t_0 + jT$

$$I_{t_0+jT} \equiv x(t_0 + jT, t_0, I_{t_0}),\ j = 0, 1, 2, \dots,$$

and

$$I_{-j} \equiv \left\{x_0 \in I_{t_0} \,|\, x(t_0 + jT, t_0, x_0) \in I_{t_0+jT} \cap I_{t_0}\right\}.$$

Since the points of I_{t_0+jT} intersecting the left and right vertical boundaries of the box for a fixed j leave the box and move to the left and right, respectively, as j is increased, we have:

$$I_0 \supset I_{-1} \supset I_{-2} \supset \cdots.$$

All the trajectories whose initial condition is in $\bigcap_{j=0}^{\infty} I_{-j}$ stay in the box for both positive and negative time. We now want to prove that there is only one trajectory staying in this box, or equivalently that

$$\bigcap_{j=0}^{\infty} I_{-j} = \text{ a point}. \tag{56}$$

The proof is by contradiction. Suppose $\lim_{j\to\infty} I_{-j}$ is not a point, but an interval. First, we note that every trajectory with initial condition in this interval must be hyperbolic and repelling. This is seen as follows. Let $\tilde{x}(t)$ denote a trajectory with initial condition in $\lim_{j\to\infty} I_{-j}$. Then stability of this trajectory is determined by the Lyapunov exponent, which is given by:

$$\lim_{t\to\infty} \frac{1}{t} \int_{t_0}^{t} \frac{\partial u}{\partial x}(\tilde{x}(\tau), \tau) d\tau,$$

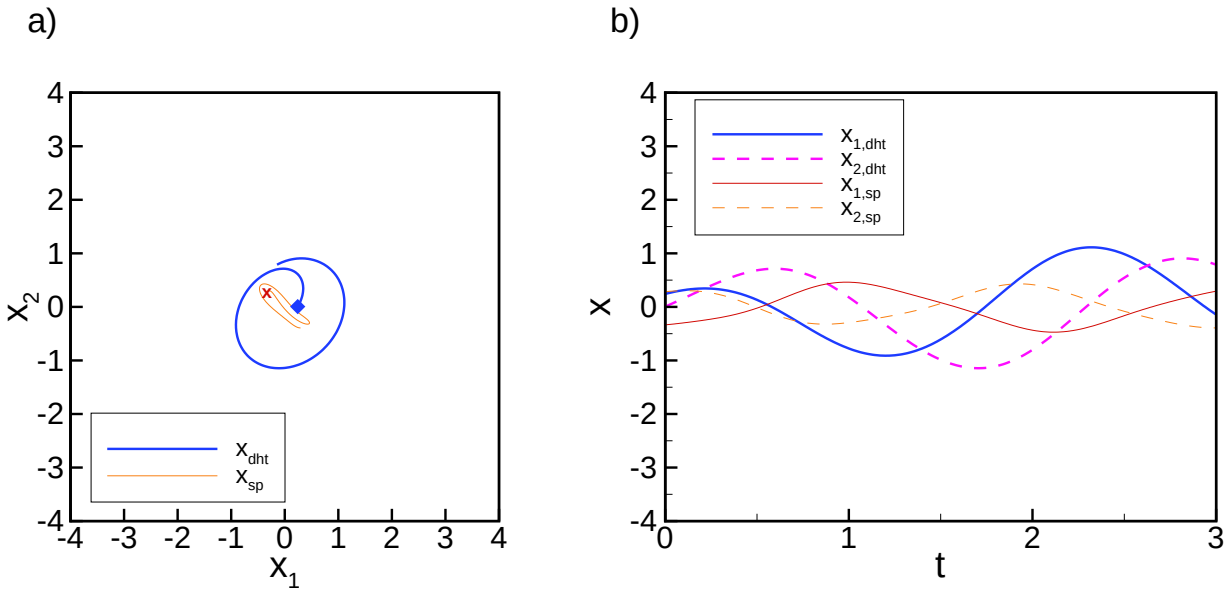

Fig. 8. Similar to Fig. 7, except for the $\boldsymbol{x}$ system.

which must be positive since we have assumed $\frac{\partial u}{\partial x} > 0$.

Returning to our proof by contradiction, if $\lim_{j\to\infty} I_{-j}$ is not a point, but an interval, then the endpoints of the interval are the initial conditions for trajectories that stay in the box for all $t \geq t_0$. These trajectories are repelling by the argument given above. Then there must be an initial condition between these two endpoints corresponding to an attracting trajectory. However, this is impossible since we are assuming that $\frac{\partial u}{\partial x} > 0$ everywhere in the box, and as we argued above, all trajectories with initial conditions in this interval must be repelling. Hence, our assumption that $\lim_{j\to\infty} I_{-j}$ is an interval gives rise to a contradiction. As a result we must have Eq. (56).

This point is the initial condition at $t = t_0$ for a hyperbolic, repelling trajectory that is our DHT. By construction it satisfies the bounds of the theorem.

Proof of 2: Choose j and T such that $t_0 + jT = t_L$. Then we can take $I_{-j} = \mathcal{I}$ since, by the construction above, $I_{-j} = \mathcal{I}$ is the (unique) set of initial conditions that remain in the box until time t_L. □

Therefore, the one-dimensional nonlinear provides us the geometrical relation between the ISP and DHT. Accordingly, the ISPs can be used as a marker for DHT identification.

3.2 Two-dimensional velocity fields

Now we develop the theory for the existence of DHTs in two-dimensional velocity field of the form Eq. (1) defined for $t \in [t_{L^-}, t_{L^+}]$, i.e.

$$\frac{d}{dt}\boldsymbol{x} = \boldsymbol{u}(\boldsymbol{x}, t), \qquad \boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2. \tag{57}$$

By taking $t_{L^-} \to -\infty$ and/or $t_{L^+} \to \infty$, the results below can be applied to the bi- or semi-infinite time interval.

Our goal here is to provide a theoretical background for the two classes of definitions for the DHT (Definitions 1 and 2 in Sect. 1), and to develop a numerical algorithm for computing the DHT. We accomplish this goal by taking a series of coordinate changes as follows: (1) from $\boldsymbol{x}$ to $\boldsymbol{x}'$, which localizes the velocity field near the ISP; (2) from $\boldsymbol{x}'$ to $\boldsymbol{w}$, which results in a velocity field with an orthogonal, time-independent linear part, so that the linear results developed in Sect. 2 can be applied to the corresponding linearized velocity field; (3) from $\boldsymbol{w}$ to $\boldsymbol{y}$ for additional localization, so that the existence can be proven for the nonlinear velocity field; and then analytically identify the DHT in $\boldsymbol{x}$ by (4) reversing these coordinate changes from $\boldsymbol{w}$ (or $\boldsymbol{y}$) to $\boldsymbol{x}$.

In the one-dimensional case, three assumptions are made concerning the ISPs of the instantaneous velocity field, i.e. existence Eq. (53), uniqueness Eq. (54), and hyperbolicity Eq. (55) in the $x - t$ box. In the two-dimensional case, we first assume only the existence of ISPs in $\boldsymbol{x}$. Additional assumptions will be made later in the $\boldsymbol{w}$ and $\boldsymbol{y}$ coordinates as necessary.

Existence of an ISP: For $t \in [t_{L^-}, t_{L^+}]$, there exists an ISP, denoted by $\boldsymbol{x}_{\rm sp}(t) = \big(x_{{\rm sp},1}(t), x_{{\rm sp},2}(t)\big)$, i.e.

$$\boldsymbol{u}(\boldsymbol{x}_{\rm sp}(t), t) \equiv 0. \tag{58}$$

Note that the same argument as in the one-dimensional case applies to show that $\boldsymbol{x}_{\rm sp}(t)$ is a particle trajectory only if it is constant in time.

Once such a $\boldsymbol{x}_{\rm sp}(t)$ is observed in the velocity field, we use it as a marker to begin a series of coordinate changes as carried out below.

1. **Localization near the ISP (from $\boldsymbol{x}$ to $\boldsymbol{x}'$)**

   In order to prove the existence of a DHT, we must first transform to a coordinate system that is localized about

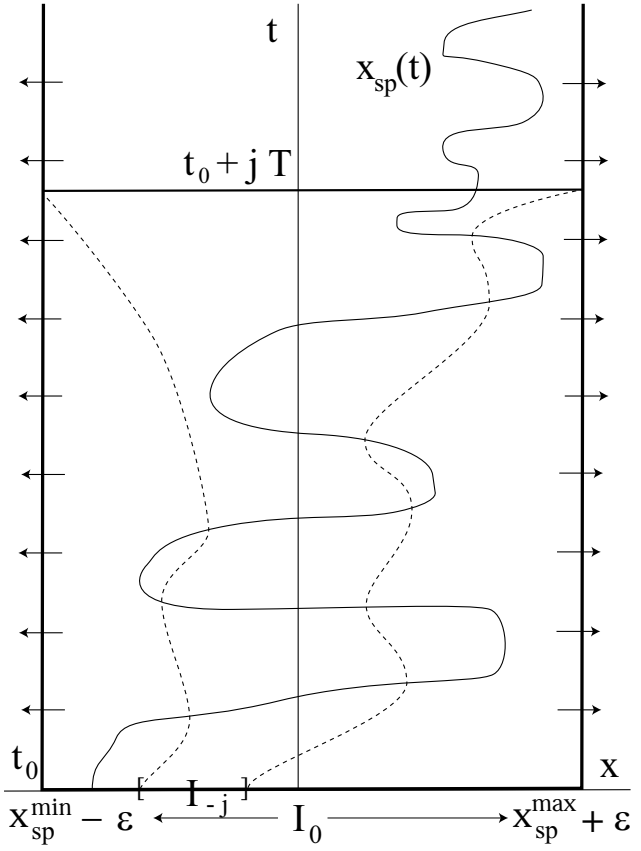

Fig. 9. Graph of the one-dimensional ISP $x_{\rm sp}(t)$ and initial time interval I_{t_0}.

an appropriate structure in the instantaneous velocity field:

$$\boldsymbol{x} = \boldsymbol{x}' + \boldsymbol{x}_{\rm sp}. \tag{59a}$$

The resulting velocity field for $\boldsymbol{x}'$ can be represented as a sum of linear, forcing (i.e. purely time-dependent) and nonlinear terms:

$$\frac{d}{dt}\boldsymbol{x}' = \mathbf{F}(t)\boldsymbol{x}' + \boldsymbol{h}^{(\rm FORCE)}(t) \tag{59b}$$

$$+ \boldsymbol{h}^{(\rm NL)}(\boldsymbol{x}' + \boldsymbol{x}_{\rm sp}, t) \tag{59c}$$

by substituting Eq. (59a) to Eq. (57) and separating the linear part. Note that the superscript "(NL)" denotes nonlinearity in the $\mathbf{x}'$ variable. There are two approaches for $\boldsymbol{x}_{\rm sp}$ that we will pursue below. The choice between the two will be discussed later, and is based on the linear theory such that the next coordinate change from $\boldsymbol{x}'$ to $\mathbf{w}$ results in the most suitable system for DHT identification.

(i) Transformation to a Coordinate System Localized about the Mean Stagnation Point: Eulerian Localization

We localize the velocity field about the mean stagnation point:

$$\boldsymbol{x}_{\rm sp} = \overline{\boldsymbol{x}}_{\rm sp} \equiv \frac{1}{t_{L^+} - t_{L^-}} \int_{t_{L^-}}^{t_{L^+}} \boldsymbol{x}_{\rm sp}(\tau) d\tau. \tag{60a}$$

The three terms in the right-hand side of velocity field Eq. (59b) are:

$$\mathbf{F}(t) = J[\boldsymbol{u}, \boldsymbol{x}]|_{(\overline{\boldsymbol{x}}_{\rm sp}, t)} = \left(\begin{array}{cc} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} \end{array} \right) \Bigg|_{(\overline{\boldsymbol{x}}_{\rm sp}, t)} \tag{60b}$$

$$\boldsymbol{h}^{(\rm FORCE)}(t) = \boldsymbol{u}(\overline{\boldsymbol{x}}_{\rm sp}, t), \tag{60c}$$

$$\boldsymbol{h}^{(\rm NL)}(\boldsymbol{x}' + \overline{\boldsymbol{x}}_{\rm sp}, t) = \boldsymbol{u}(\boldsymbol{x}' + \overline{\boldsymbol{x}}_{\rm sp}, t) - \mathbf{F}(t)\boldsymbol{x}' - \boldsymbol{u}(\overline{\boldsymbol{x}}_{\rm sp}, t). \tag{60d}$$

(ii) Transformation to a Coordinate System Localized about the Curve of Instantaneous Stagnation Points: Instantaneous Eulerian Localization

We localize by moving with the ISP

$$\boldsymbol{x}_{\rm sp} = \boldsymbol{x}_{\rm sp}(t). \tag{61a}$$

The three terms of the velocity field are:

$$\mathbf{F}(t) = J[\boldsymbol{u}, \boldsymbol{x}]|_{(\boldsymbol{x}_{\rm sp}(t), t)} = \left(\begin{array}{cc} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} \end{array} \right) \Bigg|_{(\boldsymbol{x}_{\rm sp}(t), t)}, \tag{61b}$$

$$\boldsymbol{h}^{(\rm FORCE)}(t) = -\dot{\boldsymbol{x}}_{\rm sp}(t), \tag{61c}$$

$$\boldsymbol{h}^{(\rm NL)}(\boldsymbol{x}' + \boldsymbol{x}_{\rm sp}(t), t) = \boldsymbol{u}(\boldsymbol{x}' + \boldsymbol{x}_{\rm sp}(t), t) - \mathbf{F}(t)\boldsymbol{x}'. \tag{61d}$$

1. **Transformation to a system with constant (in time) linear part (from x' to w)**

Once the appropriate type of the localization has been carried out, then we perform a coordinate transformation to make the coefficients of the linear part of the equation constant in time. We now apply the coordinate transformation defined by:

$$\boldsymbol{w}(t) = \mathbf{A}(t)\boldsymbol{x}'(t), \tag{62}$$

where $\mathbf{A}(t)$ is valid for $t \in [t_{L^-}, t_{L^+}]$ as it is derived in Appendix A to Eq. (59b). Note that this transformation is applicable to Eq. (59b), independent of the choice for $\boldsymbol{x}_{\rm sp}$ given by Eq. (60a) and Eq. (61a). After this coordinate transformation, the velocity field then takes the form:

$$\frac{d}{dt}\boldsymbol{w} = \mathbf{D}\boldsymbol{w} + \boldsymbol{g}^{(\rm FORCE)}(t) + \boldsymbol{g}^{(\rm NL)}(\boldsymbol{w}, t), \tag{63a}$$

where $\mathbf{D}$ is a constant matrix and

$$\boldsymbol{g}^{(\rm FORCE)}(t) = \mathbf{A}\boldsymbol{h}^{(\rm FORCE)}(t), \tag{63b}$$

$$\boldsymbol{g}^{(\rm NL)}(\boldsymbol{w}, t) = \mathbf{A}\boldsymbol{h}^{(\rm NL)}\left(\mathbf{A}^{-1}\boldsymbol{w}, t\right). \tag{63c}$$

Neglecting the nonlinear terms of Eq. (63a) gives the associated linear, inhomogeneous system:

$$\frac{d}{dt}\boldsymbol{w} = \mathbf{D}\boldsymbol{w} + \boldsymbol{g}^{(\mathrm{FORCE})}(t). \tag{64}$$

We now impose the second assumption on hyperbolicity so that the corresponding DHT is saddle-type in stability for both nonlinear Eq. (63a) and linearized Eq. (64) systems.

Hyperbolicity: We assume that the matrix $\mathbf{D}$ has the following form:

$$\mathbf{D} = \begin{pmatrix} \bar{\lambda}^+ & 0 \\ 0 & \bar{\lambda}^- \end{pmatrix} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}, \tag{65}$$

where

$$d_2 = \bar{\lambda}^- < -\lambda < 0 < \lambda < d_1 = \bar{\lambda}^+.$$

The system Eq. (64) consists of two independent linear inhomogeneous systems. Therefore, we can immediately write down a formula for the associated linear DHT for each system using the theory from Sect. 2, without the third assumption concerning the uniqueness of the ISP. We will denote such a DHT by $\boldsymbol{w}_{\mathrm{dht}}$.

Moreover, there are two approaches to the coordinate transformation from $\boldsymbol{x}$ to $\boldsymbol{x}'$ effected by localization about: (i) the mean ISP $\bar{\boldsymbol{x}}_{\mathrm{sp}}$ Eq. (60a), or (ii) the ISP $\boldsymbol{x}_{\mathrm{sp}}(t)$ Eq. (61a). The results from the linear theory provide a criterion for the choice between the two as we now describe.

The time-scale ratio and the choice of Eulerian localization: Let us denote the most dominant characteristic time scale of the forcing $\boldsymbol{g}^{(\mathrm{FORCE})}(t)$ by the frequency vector $\omega^{(\mathrm{FORCE})}$ and the eigenvector of the linear dynamics by $\boldsymbol{d} = (d_1, d_2)$. From the linear theory (Sect. 2.2), the time-scale ratio $|\omega^{(\mathrm{FORCE})}|/|\boldsymbol{d}|$ determines whether the DHT is (i) nearly stationary near $\bar{\boldsymbol{x}}_{\mathrm{sp}}$ or (ii) moves closely with the ISP $\boldsymbol{x}_{\mathrm{sp}}(t)$. Therefore, we localize about (i) $\bar{\boldsymbol{x}}_{\mathrm{sp}}$ when $|\omega^{(\mathrm{FORCE})}|/|\boldsymbol{d}| \geq 1$, and (ii) $\boldsymbol{x}_{\mathrm{sp}}(t)$ when $|\omega^{(\mathrm{FORCE})}|/|\boldsymbol{d}| \leq 1$.

2. **Localization of the velocity field about $\boldsymbol{w}_{\mathrm{dht}}(t)$ (from $\boldsymbol{w}$ to $\boldsymbol{y}$)**

   To prove the existence of the corresponding DHT formally in the nonlinear velocity field, we introduce a final coordinate transformation. Let

   $$\boldsymbol{w} = \boldsymbol{w}_{\mathrm{dht}}(t) + \boldsymbol{y}.$$

   Then the nonlinear velocity field Eq. (63a) takes the form:

   $$\frac{d}{dt}\boldsymbol{y} = \mathbf{D}\boldsymbol{y} + \boldsymbol{g}^{(\mathrm{NL})}\left(\boldsymbol{w}_{\mathrm{dht}}(t) + \boldsymbol{y}, t\right). \tag{66}$$

   This is equivalent to the velocity field Eq. (10) for Definition 1. Therefore, we have the following theorem for the DHT of the first class.

   Theorem 3.2 (Existence and Uniqueness of DHT) *Suppose that* $\| \boldsymbol{g}^{(\mathrm{NL})} \|_\infty$, $\| \boldsymbol{g}_{\boldsymbol{y}}^{(\mathrm{NL})} \|_\infty < \infty$ *and*

   $$\| \boldsymbol{g}_{\boldsymbol{y}}^{(\mathrm{NL})} \|_\infty \left(\frac{1}{d_1} - \frac{1}{d_2}\right) < 1.$$

   where $\boldsymbol{g}_{\boldsymbol{y}}^{(\mathrm{NL})} \equiv \frac{\partial \boldsymbol{g}^{(\mathrm{NL})}}{\partial \boldsymbol{y}} \in \mathbb{R}^{N\times n}$ *is the Jacobian matrix Then Eq. (66) has a unique bounded DHT, denoted by* $\boldsymbol{y}_{\mathrm{dht}}(t)$.

   Proof: This proof of this theorem can be obtained through standard iteration or fixed point methods, but it is a bit cumbersome and lengthy to set up and carry out (see Ju et al., 2002).

3. **DHT in the original coordinate $\boldsymbol{x}$**

   Tracing back through the original coordinates, the DHT is given by:

   $$\boldsymbol{x}_{\mathrm{dht}}(t) = \boldsymbol{x}_{\mathrm{sp}} + \mathbf{A}(t)^{-1}\left(\boldsymbol{w}_{\mathrm{dht}}(t) + \boldsymbol{y}_{\mathrm{dht}}(t)\right), \tag{67}$$

   where $\boldsymbol{x}_{\mathrm{sp}}$ is either i) $\overline{\boldsymbol{x}}_{\mathrm{sp}}$ for fast forcing or ii) $\boldsymbol{x}_{\mathrm{sp}}(t)$ for slow forcing. Using Theorem 3.2 along with Appendix C, such $\boldsymbol{x}_{\mathrm{dht}}(t)$ is indeed the DHT of Definition 2.

   In the next section we present a numerical method that gives an approximation of the DHT for the entire length of the data set. In particular, it provides an expression for:

   $$\boldsymbol{x}_{\mathrm{dht}}(t) = \boldsymbol{x}_{\mathrm{sp}} + \mathbf{A}(t)^{-1}\boldsymbol{w}_{\mathrm{dht}}(t). \tag{68}$$

4 A numerical algorithm for computing Distinguished Hyperbolic Trajectories in two-dimensional flow

In this section we describe a numerical algorithm for computing DHTs in flows defined by data sets. The algorithm is based on the linear theory of DHTs developed in the previous section. We consider a flow field defined on a Cartesian coordinate grid, $\{\boldsymbol{x}_{i,j}\} = (x_i, y_j)$ for $i = 1, \ldots, N_x$ and $j = 1, \ldots, N_j$, and at a sequence of times, t_l, $l = 1, \ldots, N_t$. The "discrete" velocity field has the form, $\boldsymbol{u}(\boldsymbol{x}_{i,j}, t_l)$. Note that spatial and temporal interpolation provides a "smooth" velocity field, $\boldsymbol{u}(\boldsymbol{x}, t)$. For convenience in a two-dimensional flow, we use (x, y) instead of (x_1, x_2) in this section.

4.1 Outline of the numerical approach for finding and tracking DHTs

In order to apply the theory developed in Sect. 2 to a velocity field defined by a data set, we must construct a time-dependent linear approximation to the velocity field. For this linear model to calculate the DHT we must linearize about a good estimate for the DHT. As discussed in Sect. 3.2 there are two approaches:

(i) For "fast" forcing, the DHT remains near the time averaged position of the ISP.

(ii) For "slow" forcing, the DHT follows the path of the ISP.

The difference between these approaches given by Eq. (60) and Eq. (61) is only apparent in the forcing term $\boldsymbol{h}^{(\mathrm{FORCE})}(t)$. We therefore use the same notation for both cases in order to simplify the discussion. Let $\boldsymbol{x}_{\mathrm{sp}}$ be the (possibly time-dependent) path of our a priori estimate of the DHT path (see Sect. 3.2). We seek to linearize the velocity field around this path and solve the resulting system for the DHT.

We thus take the following steps:

Step 1: Search for a DHT candidate and its neighborhood: Since the path $\boldsymbol{x}_{\mathrm{sp}}$ is based on the behaviour of an ISP, we first look for persistent and bounded ISPs in the entire data set so as to find neighborhoods where DHTs may exist. Details of each procedure will be discussed in Sect. 4.2.

a. At each time slice t_l for $l = 1, \ldots, N_t$, examine the velocity field and identify all ISPs (an efficient method will discussed in Sect. 4.2.1).

b. Identify each ISP that persists continuously in a bounded area isolated from any other ISPs throughout the entire time series.

c. Define a path, $\boldsymbol{x}_{\mathrm{sp}}$, around which to linearize. As discussed above, this may either be the path of an ISP, $\boldsymbol{x}_{\mathrm{sp}}(t)$, or the time-averaged path of an ISP, $\overline{\boldsymbol{x}}$. The local coordinate system $\boldsymbol{x}'$ has its origin at $\boldsymbol{x}_{\mathrm{sp}}$:

$$\boldsymbol{x} = \boldsymbol{x}_{\mathrm{sp}} + \boldsymbol{x}'. \tag{69}$$

Step 2: Construction of a time-dependent linear model: Each of the possible linearization strategies discussed above defines for each time slice a (i, j) cell which we may use to develop a linear local velocity field:

$$\boldsymbol{u}(\boldsymbol{x}, t_l) = \boldsymbol{u}_x(\boldsymbol{x}_{\mathrm{sp}}, t_l)x' + \boldsymbol{u}_y(\boldsymbol{x}_{\mathrm{sp}}, t_l)y' + \boldsymbol{h}^{(\mathrm{FORCE})}(t), \tag{70}$$

where $\boldsymbol{u}_x(\boldsymbol{x}_{\mathrm{sp}}, t_l)$ and $\boldsymbol{u}_y(\boldsymbol{x}_{\mathrm{sp}}, t_l)$ are to be determined later.

Step 3: Application of the Fourier (linear) DHT theory: This is accomplished in the following steps according to Sect. 2.

a. Construct the coordinate change, $vecw(t) = \mathbf{A}(t)\boldsymbol{x}'(t)$, so that the resulting system in $\mathbf{w}$ has a time-independent system matrix $\mathbf{D}$ (this procedure is discussed in Appendix A).

b. Examine the eigenvalues of $\mathbf{D}$ for hyperbolicity.

c. Solve for the DHT in the $\boldsymbol{w}$ system, $\boldsymbol{w}_{\mathrm{dht}}(t)$, according to the procedure discussed in Sect. 2.2, which immediately leads to the DHT in the $\boldsymbol{x}'$ coordinates, $\boldsymbol{x}'_{\mathrm{dht}}(t) = \mathbf{A}(t)^{-1}\boldsymbol{w}_{\mathrm{dht}}(t)$.

d. Convert the result into the original $\boldsymbol{x}$ coordinates

$$\boldsymbol{x}_{\mathrm{dht}}(t) = \boldsymbol{x}_{\mathrm{sp}} + \boldsymbol{x}'_{\mathrm{dht}}(t).$$

4.2 Construction of the instantaneous linear model from the data

In order to find the ISPs of the flow, we must consider the linearization of the flow velocities within a cell. Let us define the (i, j) rectangle as $\{\boldsymbol{x} | x_i \le x \le x_{i+1}, y_j \le y \le y_{j+1}\}$, defined by the four adjacent grid points $\{\boldsymbol{x}_{i,j}, \boldsymbol{x}_{i+1,j}, \boldsymbol{x}_{i,j+1}, \boldsymbol{x}_{i+1,j+1}\}$. If x lies within the (i, j) cell, then we approximate the velocity within this cell by means of bilinear interpolation

$$\begin{aligned}\boldsymbol{u}(\boldsymbol{x}, t) = \frac{1}{\triangle x_i \triangle y_j}\{&(\triangle x_i - \xi)(\triangle y_j - \eta)\,\boldsymbol{u}_{i,j}(t) \\ &+ (\triangle y_j - \eta)\,\boldsymbol{u}_{i+1,j}(t) + (\triangle x_i - \xi)\,\boldsymbol{u}_{i,j+1}(t) \\ &+ \xi\eta\,\boldsymbol{u}_{i+1,j+1}(t)\},\end{aligned} \tag{71a}$$

where we have used the notation

$$\boldsymbol{u}_{i,j}(t) = \boldsymbol{u}(\boldsymbol{x}_{i,j}, t), \tag{71b}$$

$$(\triangle x_i, \triangle y_j) = (x_{i+1} - x_i, y_{j+1} - y_j), \tag{71c}$$

$$\boldsymbol{\xi} = (\xi, \eta) = (x - x_i, y - y_j). \tag{71d}$$

4.2.1 Efficient search for ISPs

An efficient search for ISPs is conducted by first looking for any (i, j) rectangle which satisfies a necessary condition, and then examining a sufficient condition only for those (i, j) rectangles that satisfy the necessary condition. By definition, bilinear velocity in any (i, j) rectangle velocity satisfies the relation:

$$\begin{aligned}&u_{i,j}^{\min}(t_l) \le u(\boldsymbol{x}, t_l) \le u_{i,j}^{\max}(t_l), \\ &v_{i,j}^{\min}(t_l) \le v(\boldsymbol{x}, t_l) \le v_{i,j}^{\max}(t_l),\end{aligned} \tag{72a}$$

where

$$\begin{aligned}(u_{i,j}^{\min}(t_l), u_{i,j}^{\max}(t_l)) \equiv (min,\ max)\{&u_{i,j}(t_l), \\ &u_{i+1,j}(t_l), u_{i,j+1}(t_l), u_{i+1,j+1}(t_l)\},\end{aligned} \tag{72b}$$

$$\begin{aligned}(v_{i,j}^{\min}(t_l), v_{i,j}^{\max}(t_l)) \equiv (min,\ max)\{&v_{i,j}(t_l), \\ &v_{i+1,j}(t_l), v_{i,j+1}(t_l), v_{i+1,j+1}(t_l)\}.\end{aligned} \tag{72c}$$

Necessary condition: If there exists an ISP in an (i, j) rectangle at time t_l, then the following condition must hold:

$$u_{i,j}^{\min}(t_l) \le 0 \le u_{i,j}^{\max}(t_l), \qquad v_{i,j}^{\min}(t_l) \le 0 \le v_{i,j}^{\max}(t_l). \tag{73}$$

Sufficient condition: The bilinear system given by Eq. (71a) leads to a quadratic equation for $\xi_{\mathrm{sp}}(t_l)$ that can be solved analytically. A sufficient condition is, therefore, that the solution indeed lies in (i, j) rectangle, i.e.

$$0 \le \xi_{\mathrm{sp}}(t_l) \le \triangle x_i, \qquad 0 \le \eta_{\mathrm{sp}}(t_l) \le \triangle y_j. \tag{74}$$

4.2.2 Optimal linear model within a cell

At each time slice t_l, we construct a linear model by optimally eliminating the nonlinear term from the bilinear model Eq. (71a) which can be rewritten as follows:

$$\boldsymbol{u}(\boldsymbol{x}, t_l) = \boldsymbol{u}_a(t_l)\,\frac{\xi}{\triangle x_i}\frac{\eta}{\triangle y_j} + \boldsymbol{u}_b(t_l)\,\frac{\xi}{\triangle x_i} + \boldsymbol{u}_c(t_l)\,\frac{\eta}{\triangle y_j} + \boldsymbol{u}_d(t_l), \tag{75a}$$

where

$$\begin{aligned}
\boldsymbol{u}_a(t_l) &= \quad \boldsymbol{u}_{i,j}(t_l) + \boldsymbol{u}_{i+1,j+1}(t_l) - \boldsymbol{u}_{i+1,j}(t_l) - \boldsymbol{u}_{i,j+1}(t_l),\\
\boldsymbol{u}_b(t_l) &= -\boldsymbol{u}_{i,j}(t_l) + \boldsymbol{u}_{i+1,j}(t_l),\\
\boldsymbol{u}_c(t_l) &= -\boldsymbol{u}_{i,j}(t_l) + \boldsymbol{u}_{i,j+1}(t_l),\\
\boldsymbol{u}_d(t_l) &= \quad \boldsymbol{u}_{i,j}(t_l).
\end{aligned} \tag{75b}$$

Hence for this model to be linear, the following condition must hold:

$$\boldsymbol{u}_a(t_l) \equiv 0. \tag{76}$$

If the local flow is truly linear, then this condition is automatically satisfied. Therefore, condition Eq. (76) can be used as the linearity test of the local flow, so as to apply the Fourier (linear) DHT theory developed in Sect. 2.

We may develop the optimal linear model of Eq. (75) by eliminating the velocity at the grid point furthest away from $\boldsymbol{x}_{\mathrm{sp}}$, using the condition Eq. (76). The resulting linear model leads to Eq. (70), whose time-dependent coefficient vectors $\boldsymbol{u}_x(t_l)$, $\boldsymbol{u}_y(t_l)$ and $\boldsymbol{u}_0(t_l)$ are given in Table 1. Note that the use of this linear model implies that the Jacobian of the velocity field is constant everywhere within the cell, so that we can construct the coefficients in Table 1 without using the local coordinates of $\boldsymbol{x}_{\mathrm{sp}}$ within the (i, j) cell.

4.2.3 The forcing term

In developing a linear model we seek an equation of the form

$$\frac{d\boldsymbol{x}'}{dt} = \mathbf{F}(t)\boldsymbol{x}'(t) + \boldsymbol{h}(t). \tag{77}$$

For the case where our initial estimate $\boldsymbol{x}_{\mathrm{sp}}(t)$ of the location of the DHT is constant, then we may linearize around the lower left-hand coordinate of the cell containing $\boldsymbol{x}_{\mathrm{sp}}$, using the cell coordinates $\boldsymbol{\xi} = (\xi, \eta)$. The forcing term $\boldsymbol{u}_0$ is that given in Table 1:

$$\frac{d}{dt}\boldsymbol{\xi} = \begin{pmatrix} u_x(t) & u_y(t) \\ v_x(t) & v_y(t) \end{pmatrix}\boldsymbol{\xi} + \begin{pmatrix} u_0(t) \\ v_0(t) \end{pmatrix}. \tag{78}$$

When $\boldsymbol{x}_{\mathrm{sp}}(t)$ is the path of the ISP, we must linearize in terms of $\boldsymbol{x}'$, giving:

$$\frac{d}{dt}\boldsymbol{x}' = \begin{pmatrix} u_x(t) & u_y(t) \\ v_x(t) & v_y(t) \end{pmatrix}\boldsymbol{x}' - \frac{d}{dt}\boldsymbol{x}_{\mathrm{sp}}(t), \tag{79}$$

where the coefficients of the Jacobian are deduced from Table 1, which are valid throughout the cell, and the velocity of the ISP path must be calculated numerically.

Finally, a linear, time-dependent system in $\boldsymbol{\xi}$ or $\boldsymbol{x}'$ is obtained by temporal interpolation of each coefficient.

4.3 Numerical reconstruction of DHTs

4.3.1 One-dimensional case

Step 3 of the algorithm described in Sect. 4.1 requires solving for the DHT in the **w** system. In practice this gives rise to a choice of methods; we may either use a frequency-domain or a time-domain solution. This section outlines both methods in the context of a one-dimensional problem and the next discusses the choice between them. We begin by considering the one-dimensional problem

$$\frac{\mathrm{d}y}{\mathrm{d}t} - dy = be^{i\omega t}. \tag{80}$$

This can readily be rewritten as

$$e^{dt}\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-dt}y\right) = be^{i\omega t}. \tag{81}$$

A particular integral (which has no exponential behaviour and is therefore the DHT) for this equation is

$$y = \frac{b}{-d + i\omega}e^{i\omega t}. \tag{82}$$

So, if we represent a forcing function by its Fourier decomposition

$$f(t) = \int_{-\infty}^{\infty} F(\omega)e^{i\omega t}\,dt$$

then the DHT is determined by

$$\hat{g}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} Q(\omega)F(\omega)e^{i\omega t}\,d\omega, \tag{83}$$

where

$$Q(\omega) = \frac{1}{-d + i\omega}.$$

The multiplication in the frequency domain in Eq. (83) is equivalent to a convolution in the time domain, so we can rewrite this as

$$\hat{g}(t) = \int_{-\infty}^{\infty} f(t)q(t-\tau)\,d\tau, \tag{84}$$

where $q(t)$ is the transform of

$$\overline{Q}(\omega) = \frac{1}{-d - i\omega},$$

the complex conjugate of $Q(\omega)$.

Table 1. Linear model on (i, j) grid from data. Nearest grid is one of the four edge points of (i, j) base-grid which is nearest to the ISP. Time variable is omitted for simplicity

nearest grid	$\triangle x_i$ $\boldsymbol{u}_x$		$\triangle y_j$ $\boldsymbol{u}_y$		$\boldsymbol{u}_0$		
(i, j)	$-\boldsymbol{u}_{i,j}$	$+\boldsymbol{u}_{i+1,j}$	$-\boldsymbol{u}_{i,j}$	$+\boldsymbol{u}_{i,j+1}$	$\boldsymbol{u}_{i,j}$		
$(i+1, j)$	$-\boldsymbol{u}_{i,j}$	$+\boldsymbol{u}_{i+1,j}$	$-\boldsymbol{u}_{i+1,j}$	$+\boldsymbol{u}_{i+1,j+1}$	$\boldsymbol{u}_{i,j}$		
$(i, j+1)$	$-\boldsymbol{u}_{i,j+1}$	$+\boldsymbol{u}_{i+1,j+1}$	$-\boldsymbol{u}_{i,j}$	$+\boldsymbol{u}_{i,j+1}$	$\boldsymbol{u}_{i,j}$		
$(i+1, j+1)$	$-\boldsymbol{u}_{i,j+1}$	$+\boldsymbol{u}_{i+1,j+1}$	$-\boldsymbol{u}_{i+1,j}$	$+\boldsymbol{u}_{i+1,j+1}$	$\boldsymbol{u}_{i+1,j}$	$+\boldsymbol{u}_{i,j+1}$	$-\boldsymbol{u}_{i+1,j+1}$

In order to proceed we need a functional representation for $q(t)$. In fact, two such representations are needed; one for each possible sign of d, so that the integrals involved in calculating the Fourier transform of $q(t)$ converge.

For $d > 0$ we have

$$q(t) = \begin{cases} 0 & \text{for } t > 0, \\ -e^{dt} & \text{for } t < 0, \end{cases} \tag{85}$$

while for $d < 0$ we have

$$q(t) = \begin{cases} e^{dt} & \text{for } t > 0, \\ 0 & \text{for } t < 0. \end{cases} \tag{86}$$

This representation of the DHT corresponds to Eq. (17) and Eq. (18) using Fourier decomposition.

4.3.2 A choice of methods

The numerical implementation of the DHT algorithm requires us to choose between the use of Eq. (83), which we call the frequency-domain reconstruction, and that of Eq. (84), which we call the time-domain reconstruction.

In practice the time-domain reconstruction has the considerable advantage of being less affected by the time interval over which data is available. It is clear from Eq. (83) that the DHT is essentially a "low-pass filter" on the spectrum of the forcing function $f(t)$. If $f(t)$ is periodic, then $g(t)$ should also be periodic. However, if the period of $f(t)$ is not known a priori, then the Fourier series representation of $f(t)$ will have high-frequency components (arising from aliasing). Applying Eq. (83) will filter these components out, so that the reconstructed DHT cannot capture the periodicity of $g(t)$ since this would require that the high-frequency components are preserved.

Essentially the frequency-domain reconstruction uses the FFT to represent the function $f(t)$ as a Fourier series over an arbitrary interval, which may require high-frequency components, while simultaneously using it as an approximation to the Fourier transform of $f(t)$, from which the high-frequency components are an undesirable artifact. Using Eq. (84) as the basis for a numerical implementation of the DHT reconstruction allows us to avoid the numerical problems associated with this ambiguity of the FFT, and accordingly gives significantly better results.

4.3.3 Applying the method

In this section we consider the application of the above method.

Example 8: The rotating Duffing equation can be written as:

$$\frac{d}{dt}\boldsymbol{x} = \begin{pmatrix} \sin 2\beta t & \beta + \cos 2\beta t \\ -\beta + \cos 2\beta t & -\sin 2\beta t \end{pmatrix} \boldsymbol{x} + [-(\cos \beta t\, x - \sin \beta t\, y)^3 + \epsilon \sin \omega t] \begin{pmatrix} \sin \beta t \\ \cos \beta t \end{pmatrix}. \tag{87}$$

Using a system given by an explicit formula, but applying methods which use only a discrete data set derived from it, allows us to assess the accuracy (and validity) of our numerical methods.

Equation (87) was discretized on a 40-by-40 grid spatially, with each x and y ranging from -3 to 3. The time discretization used 100 points ranging from 0 to 10.

The implementation closely followed the description in Sect. 4.1; only differences between the code and that section are described here.

Step 1: Instantaneous stagnation points which were not hyperbolic in character in the instantaneous velocity field were discarded. This left no more than one ISP for each time slice. For values of ϵ of 0.3 or greater in Eq. (87), it was common for there to be no hyperbolic ISP for some time slices. In these cases the method was considered to have failed.

Step 3a: In solving the differential equations Eq. (B4) a fourth order Runge-Kutta solver was used, with 20 time-steps between the time slices of the discretized field. Linear interpolation was used to estimate the velocity field between time slices.

Step 3c: The convolution-in-time method described in Sect. 4.3.1 was used to calculate the DHT in η coordinates.

The quadrature in Eq. (84) may be written as

$$I(t) = \int_{t_0}^{t_1} e^{d(t-\tau)} f(\tau)\, d\tau.$$

Using the change of variable, $s = -e^{-d\tau}/d$, this may be written

$$I(t) = e^{-dt} \int_{s_0}^{s_1} \hat{g}(s)\, ds.$$

The integrals of this form were evaluated using the trapezium rule.

In the convolution at each time in the discretized system the range of integration was truncated to the time domain of available data; attempts to "wrap around" from the end to the beginning showed the same problems that arose from using the frequency-domain form of the convolution.

4.3.4 Validation of the method

Clearly it is desirable to have some way of validating the method described here for calculating DHTs. The rotating Duffing system with $\beta = 0$, $\omega = 1$ and small ϵ makes this possible; it is straightforward to check that

$$-\frac{\epsilon}{2}\begin{pmatrix}\sin t\\ \cos t\end{pmatrix}$$

is a trajectory, to first order in ϵ. Since this trajectory is periodic it must be the DHT.

Figure 10 shows the first-order analytical DHT and the results of the numerical DHT reconstruction using both time-domain and frequency-domain methods. It is clear that the time-domain method is more faithful to the known behaviour at the end-points of the trajectory, although the inaccuracy here is still greater than elsewhere. This end-point effect is clearly due to the finiteness of the data set, as discussed in Sect. 2.1.

Behavior of ISP and DHT resembles that observed in Fig. 4a. While a path of persistent ISP can be used to identify the region where a DHT may exist, ISP at any given instance does not offer any direct information concerning DHT.

4.3.5 Blob tracking and non-periodic systems

The method described in this paper was developed specifically to compute DHTs for non-periodic systems. However, such cases are more difficult to validate since neither analytical solutions nor Poincaré map techniques are available. In this case we rely on the circle method; a ring of particles is placed around the position of the candidate DHT and tracked through the time interval. Another ring of particles is placed around the candidate DHT at the end of the time interval and tracked backwards to the beginning of the time interval. If the DHT is within the circles at either end of the time interval, then it will remain within the (increasingly distorted) ring as it evolves. We can thus deduce that the DHT lies within the intersection of the two curves at all times. If the DHT is of saddle-type in stability, then a circle placed around it and evolving forwards in time will contract in the direction of the stable manifold and elongate in the direction of the unstable manifold; the opposite will happen to a circle evolving backward in time. This provides an additional confirmation that the DHT is located inside the evolving blobs.

Animations of this process have been developed for the rotating Duffing system. The particle trajectories were computed using a fourth-order Runge-Kutta scheme with adaptive step doubling, using the explicit formula for the rotating Duffing velocity field. To maintain resolution as the contour deforms additional points were inserted half-way between neighboring points when they moved too far apart.

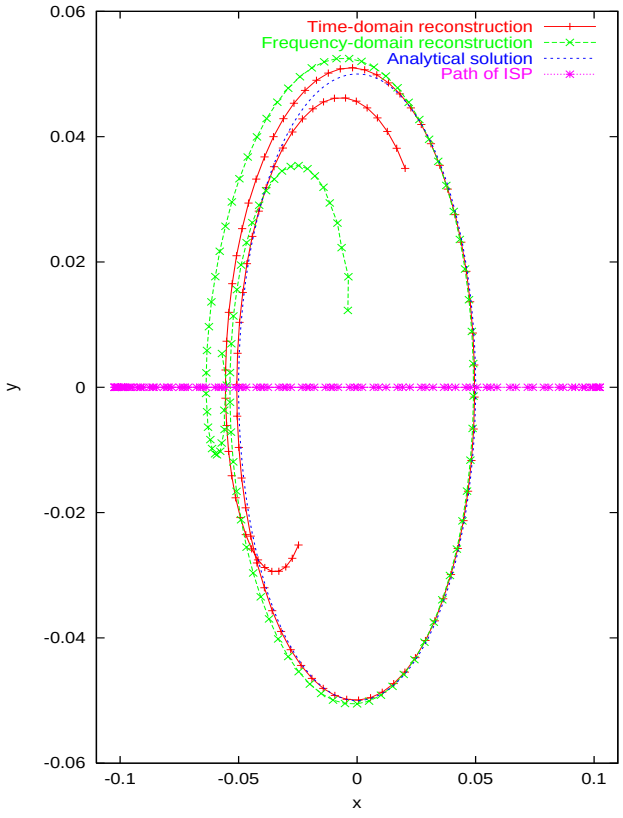

Fig. 10. A comparison of the DHT reconstruction algorithms for a rotating Duffing system with $\epsilon = 0.1$, $\beta = 0$ and $\omega = 1$. It is clear that the time-domain algorithm gives better results than the frequency-domain algorithm. The analytical solution is the first-order solution in ϵ.

Animations for the cases $\epsilon = 0.1, \beta = 0.1, \omega = 1.0$ and $\epsilon = 0.2, \beta = 0.3, \omega = 0.1$ are available on the World Wide Web at http://lacms.maths.bris.ac.uk. In these animations the blobs evolving "backwards in time" from a circle at the end of the time interval are shown in black, the circle evolving forward is in cyan, and the numerically reconstructed DHT is in magenta. It may be seen from these animations that there is good agreement between the numerical reconstruction of the DHT and the real DHT, defined by the intersection of the evolving circles. The circles used are relatively large compared to the region through which the DHT moves, since they evolve from the ends of the time interval where the accuracy of the reconstruction is lowest.

Figures 11 and 12 show the forwards and backwards blobs at the beginning and end of the time interval for the case with $\omega = 1$, $\beta = 0$, $\epsilon = 0.1$. The forward blob begins (at $t = 0$) as a circle of radius 0.1 around the position of the computed DHT at time $t = 0$ and evolves forward in time; the backwards blob begins (at $t = 10$) as a circle of radius 0.1 around the position of the computed DHT at time $t = 10$ and evolves backwards in time. Figures 13 and 14 show the two blobs at the middle of the time interval, together with the analytically derived DHT for this case. Figure 14 shows clearly that the DHT passes through the intersection of the blobs.

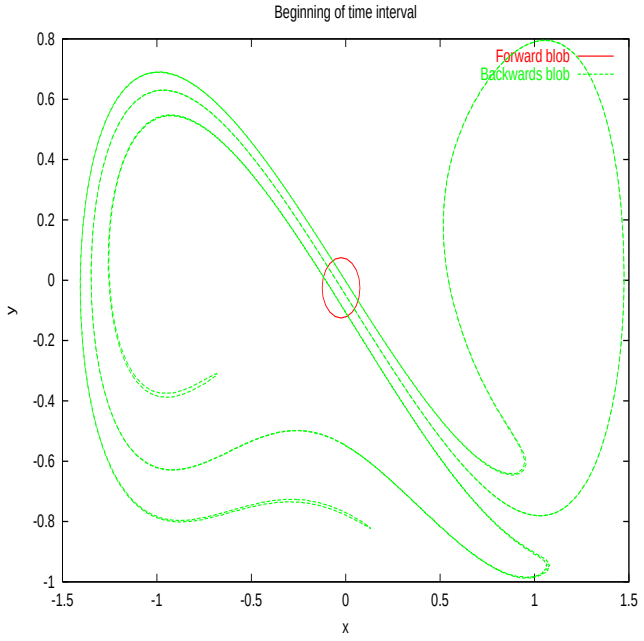

Fig. 11 The forward and backward blobs superimposed at time $t = 0$, for the rotating Duffing problem with $\omega = 1$, $\beta = 0, \epsilon = 0.1$.

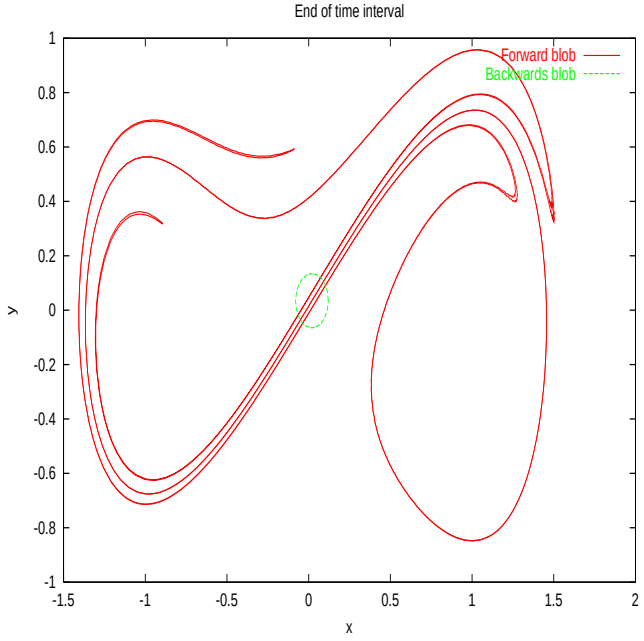

Fig. 12 The forward and backwards blobs superimposed at time $t = 10$, for the same case as Fig. 11.

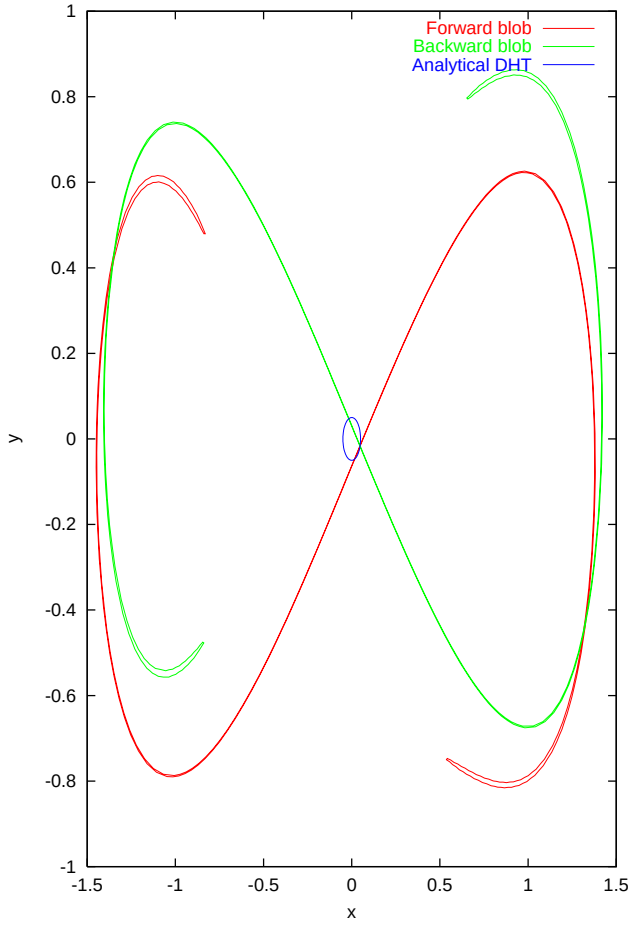

Fig. 13 The forward and backwards blobs superimposed at time $t = 5.05$, for the same case as Fig. 11, together with the analytical DHT for this case.

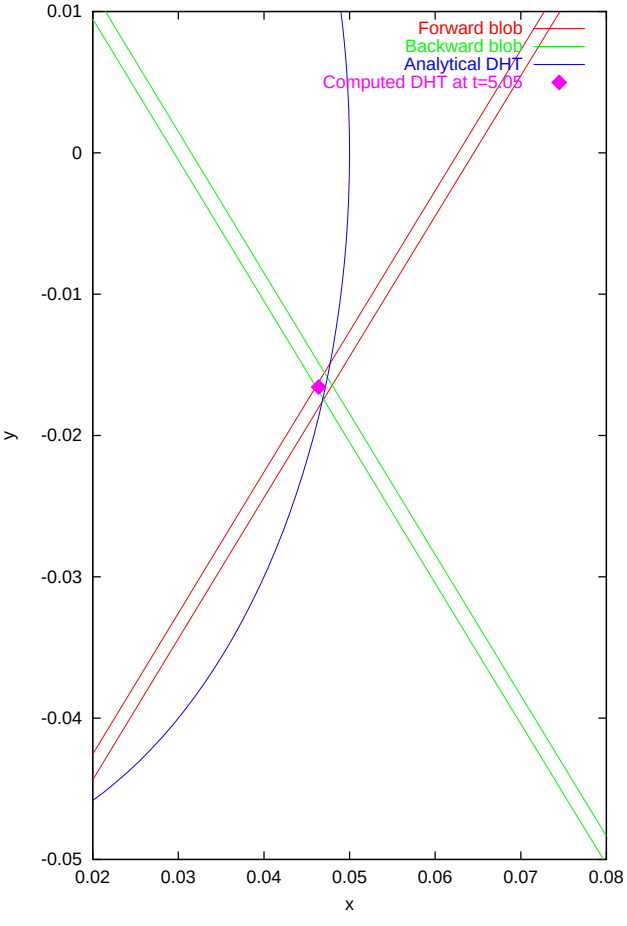

Fig. 14 An expanded version of Fig. 13 showing the DHT and the intersection of the blobs.

Appendix A Derivation of the transformation to a linear system with constant coefficients

The goal of this Appendix is to construct the coordinate transformation $\mathbf{A}(t)$ of Eq. (44) from a time-dependent coefficient system defined by $\mathbf{F}(t)$ and $\boldsymbol{h}(t)$ in Eq. (42) to a time-independent coefficient system defined by $\mathbf{D}$ and $\boldsymbol{g}(t)$ in Eq. (14), so that these two linear systems have the same hyperbolicity over a time interval $t \in [t_0, t_L]$.

Differentiating Eq. (44) with respect to t, substituting Eq. (14) and comparing with Eq. (42) gives:

$$\begin{array}{lll} \frac{d}{dt}\boldsymbol{y} = \left(\frac{d}{dt}\mathbf{A}(t)\right)\boldsymbol{x} + \mathbf{A}(t)\frac{d}{dt}\boldsymbol{x} = \left(\frac{d}{dt}\mathbf{A}(t) + \mathbf{A}(t)\mathbf{F}(t)\right)\boldsymbol{x} + \mathbf{A}(t)\boldsymbol{h}(t) \\ \quad = \mathbf{D}\boldsymbol{y} + \boldsymbol{g}(t) \qquad\qquad = \mathbf{D}\mathbf{A}(t)\boldsymbol{x} \qquad\qquad + \boldsymbol{g}(t), \end{array}$$

from which it follows that:

$$\frac{d}{dt}\mathbf{A}(t) = \mathbf{D}\mathbf{A}(t) - \mathbf{A}(t)\mathbf{F}(t)\ , \tag{A1a}$$

$$\mathbf{A}\boldsymbol{h}(t) = \boldsymbol{g}(t)\ . \tag{A1b}$$

The problem therefore reduces to determining $\mathbf{D}$ and $\mathbf{A}(t)$ for time $t \in [t_0, t_L]$.

At first glance it may appear that this problem is not solveable since Eq. (A1a) is four equations in six unknowns. However, knowledge of $\mathbf{D}$ clearly requires information from the homogeneous dynamics based on $\mathbf{F}(t)$. Therefore, we must also consider the fundamental solution matrices of the two systems:

$$\frac{d}{dt}\mathbf{X}(t, t_0) = \mathbf{F}(t)\ \mathbf{X}(t, t_0)\ , \tag{A2a}$$

$$\frac{d}{dt}\mathbf{Y}(t, t_0) = \mathbf{D}\ \mathbf{Y}(t, t_0)\ , \tag{A2b}$$

with initial condition $\mathbf{X}(t_0, t_0) = \mathbf{Y}(t_0, t_0) = \mathbf{I}$. The fundamental solution matrix $\mathbf{X}(t, t_0)$ can be represented by a singular value decomposition as follows:

$$\mathbf{X}(t, t_0) = \mathbf{B}(t, t_0)\exp\left(\mathbf{\Sigma}(t, t_0)\right)\mathbf{R}(t, t_0)^T, \tag{A3}$$

where $\mathbf{B}(t, t_0)$ and $\mathbf{R}(t, t_0)$ are orthogonal matrices, i.e. $\mathbf{B}(t, t_0)\mathbf{B}(t, t_0)^T = \mathbf{R}(t, t_0)\mathbf{R}(t, t_0)^T = \mathbf{I}$, and $\mathbf{\Sigma}(t, t_0)$ is a diagonal matrix with $\mathbf{\Sigma}(t_0, t_0) = 0$ so that $\exp\left(\mathbf{\Sigma}(t, t_0)\right)$ is a diagonal matrix with $\exp\left(\mathbf{\Sigma}(t_0, t_0)\right) = \mathbf{I}$.

Since $\mathbf{D}$ will be chosen to be a constant matrix the fundamental solution matrix of Eq. (A2b) can be given analytically by Eq. (17).

We leave the choice of $\mathbf{D}$ for the moment, and derive a relationship between $\mathbf{A}(t)$ and the fundamental solution matrices. Given initial conditions related by $\boldsymbol{y}_0 = \mathbf{A}(t_0)\boldsymbol{x}_0$, the solutions of the homogeneous equations satisfy:

$$\boldsymbol{x}(t; \boldsymbol{x}_0, t_0) = \mathbf{X}(t, t_0)\boldsymbol{x}_0\ , \tag{A4a}$$

$$\boldsymbol{y}(t;\ \boldsymbol{y}_0, t_0) = \mathbf{Y}(t, t_0)\boldsymbol{y}_0\ , \tag{A4b}$$

with

$$\boldsymbol{y}(t; \boldsymbol{y}_0, t_0) = \mathbf{A}(t)\boldsymbol{x}(t; \boldsymbol{x}_0, t_0) \tag{A5}$$

from the definition of $\mathbf{A}(t)$. Using Eq. (A4) and the initial condition relation leads to:

$$[\mathbf{Y}(t, t_0)\mathbf{A}(t_0) - \mathbf{A}(t)\mathbf{X}(t, t_0)]\boldsymbol{x}_0 = 0\ , \tag{A6}$$

for any $\boldsymbol{x}_0$. Hence, we have

$$\mathbf{A}(t) = \mathbf{Y}(t, t_0)\mathbf{A}(t_0)\mathbf{X}(t, t_0)^{-1}. \tag{A7}$$

Equation (A7) provides a relation between the transformation $\mathbf{A}(t)$ and the fundamental solution matrices, provided we specify $\mathbf{A}(t_0)$.

Evaluating Eq. (A7) at $t = t_L$, and rearranging terms, gives:

$$\mathbf{X}(t_L, t_0) = \mathbf{A}(t_L)^{-1}\mathbf{Y}(t_L, t_0)\mathbf{A}(t_0). \tag{A8}$$

It is convenient to choose $\mathbf{A}(t_0)$ so that the $\mathbf{X}$ and $\mathbf{Y}$ systems are aligned at $t = t_L$ (we cannot align the systems at $t = t_0$ a priori, since $\mathbf{X}(t_0, t_0) = \mathbf{I}$, and this has no "natural" alignment). This motivates choosing

$$\mathbf{A}(t_0) = \mathbf{R}(t_L, t_0)^T, \tag{A9}$$

so that we have

$$\mathbf{A}(t_L) = \mathbf{B}(t_L, t_0)^T. \tag{A10}$$

Then Eq. (A8) becomes:

$$\begin{aligned} \mathbf{X}(t_L, t_0) &= \mathbf{B}(t_L, t_0)\mathbf{Y}(t_L, t_0)\mathbf{R}(t_L, t_0)^T \\ &= \mathbf{B}(t_L, t_0)\exp\left((t_L - t_0)\mathbf{D}\right)\mathbf{R}(t_L, t_0)^T. \end{aligned} \tag{A11}$$

Now evaluating Eq. (A3) at $t = t_L$ gives:

$$\mathbf{X}(t_L, t_0) = \mathbf{B}(t_L, t_0)\exp\left(\mathbf{\Sigma}(t_L, t_0)\right)\mathbf{R}(t_L, t_0)^T, \tag{A12}$$

It now follows by comparing Eq. (A11) and Eq. (A12) that:

$$\mathbf{D} = \frac{1}{t_L - t_0}\mathbf{\Sigma}(t_L, t_0). \tag{A13}$$

This defines the diagonal matrix $\mathbf{D}$. Furthermore, by substituting Eq. (A9), Eq. (A3), and Eq. (17) into Eq. (A7) gives the following formula for $\mathbf{A}(t)$:

$$\begin{aligned} \mathbf{A}(t) = &\exp\left((t - t_0)\mathbf{D}\right)\mathbf{R}(t_L, t_0)^T\mathbf{R}(t, t_0) \\ &\cdot \exp\left(-\mathbf{\Sigma}(t, t_0)\right)\mathbf{B}(t, t_0)^T. \end{aligned} \tag{A14}$$

Hence, we see that $\mathbf{D}$ and $\mathbf{A}(t)$ are completely determined by the singular value decomposition of the fundamental solution matrix $\mathbf{X}(t, t_0)$.

In theory, if $\mathbf{X}(t, t_0)$ can be computed from Eq. (A2a), then we can take the singular value decomposition of $\mathbf{X}(t_L, t_0)$ once to obtain $\mathbf{A}(t_0)$ from Eq. (A9) and $\mathbf{D}$ from Eq. (A13). The coordinate change $\mathbf{A}(t)$ follows from Eq. (A7) with

Eq. (17). However, a naive numerical integration of Eq. (A2a) in order to compute $\mathbf{X}(t, t_0)$ generally leads to an exponential blowup of the solution. In order to avoid this numerical problem, we compute $\mathbf{B}(t, t_0)$, $\mathbf{R}(t, t_0)$ and $\mathbf{\Sigma}(t, t_0)$ using the algorithm presented in the following Appendix.

Hence, the algorithm for transforming the linear time-dependent system to a linear time-independent system proceeds as follows.

Step 1. Compute $\mathbf{B}(t, t_0)$, $\mathbf{R}(t, t_0)$ and $\mathbf{\Sigma}(t, t_0)$ using the algorithm described in Appendix B.

Step 2. Compute $\mathbf{D}$ using Eq. (A13).

Step 3. Compute $\mathbf{A}(t)$ using Eq. (A14).

Step 4. Compute $\boldsymbol{g}(t)$ using Eq. (A1b).

B An alternative dynamical system for X

For simplicity, we drop the arguments (t, t_0) from the matrices and use $\{\dot{\cdot}\}$ for the time derivative in this Appendix. We wish to solve Eq. (A2a) for $\mathbf{X}$ numerically. However, solutions of this equation may experience exponential growth and a naive implementation risks overflowing machine arithmetic. In order to avoid this we solve Eq. (A2a) in terms of the singular value decomposition given by Eq. (A3). Substituting these relations into Eq. (A2a) gives

$$\dot{\mathbf{B}}\exp(\mathbf{\Sigma})\mathbf{R}^T + \mathbf{B}\dot{\mathbf{\Sigma}}\exp(\mathbf{\Sigma})\mathbf{R}^T + \mathbf{B}\exp(\mathbf{\Sigma})\dot{\mathbf{R}}^T = \mathbf{F}\mathbf{B}\exp(\mathbf{\Sigma}). \tag{B1}$$

Pre-multiplying by $\mathbf{B}^T$ and post-multiplying by $\mathbf{R}\exp(-\mathbf{\Sigma})$ we have

$$\mathbf{B}^T\dot{\mathbf{B}} + \dot{\mathbf{\Sigma}} + \exp(\mathbf{\Sigma})\dot{\mathbf{R}}^T\mathbf{R}\exp(-\mathbf{\Sigma}) = \mathbf{H}, \tag{B2}$$

where $\mathbf{H} = \mathbf{B}^T\mathbf{F}\mathbf{B}$. Since $\mathbf{B}$ and $\mathbf{R}$ are orthogonal we may parameterise them as

$$\mathbf{B} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{B3a}$$

$$\mathbf{R} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}. \tag{B3b}$$

This parameterisation leads to the system

$$\dot{\sigma_1} = H_{11} \tag{B4a}$$

$$\dot{\sigma_2} = H_{22} \tag{B4b}$$

$$\dot{\theta} = \frac{1}{2}(H_{12} - H_{21}) + \frac{1}{2}\coth(\sigma_2 - \sigma_1) \tag{B4c}$$

$$\dot{\phi} = \frac{1}{2}\,\mathrm{cosech}(\sigma_2 - \sigma_1), \tag{B4d}$$

where H_{ij} represents, the components of $\mathbf{H}$, and σ_1 and σ_2 are the diagonal elements of $\mathbf{\Sigma}$.

A singularity for $\dot{\theta}$ and $\dot{\phi}$ occurs at $\sigma_1 = \sigma_2$. This is an artifact of the coordinate transform and may be dealt with by using the original Eq. (A2a) for a short time interval initially. Once they grow apart to become $\sigma_1 \neq \sigma_2$, Eq. (B4) can be used.

C Some relations between particle trajectories, instantaneous stagnation points, and coordinate transformations

Consider the following time-dependent velocity field:

$$\frac{d}{dt}\boldsymbol{x} = \boldsymbol{u}(\boldsymbol{x}, t), \quad \boldsymbol{x} \in \mathbb{R}^n. \tag{C1}$$

We consider a time-dependent transformation of coordinates of the following form:

$$\boldsymbol{x} = \boldsymbol{b}(\boldsymbol{y}, t), \quad \boldsymbol{y} \in \mathbb{R}^n, \tag{C2a}$$

which can be nonlinear in $\boldsymbol{y}$ and/or t. We will assume that for each fixed t, this coordinate transformation is invertible. In particular, for each t, and for all $\boldsymbol{y}$, we assume that

$$\det\left(\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t)\right) \neq 0,$$

where $\boldsymbol{b}_{\boldsymbol{y}} \equiv \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{y}} \in \mathbb{R}^{n\times n}$ denotes the Jacobian matrix of $\boldsymbol{b}$ with respect to the spatial coordinate $\boldsymbol{y}$. We define the inverse of Eq. (C2a) as

$$\boldsymbol{y} = \boldsymbol{a}(\boldsymbol{x}, t) = \boldsymbol{b}^{-1}(\boldsymbol{x}, t)\ , \tag{C2b}$$

then it follows that:

$$\boldsymbol{b}(\boldsymbol{a}(\boldsymbol{x}, t), t) = \boldsymbol{x}\ . \tag{C2c}$$

A special case of Eq. (C2a) is given by the linear time-dependent coordinate transformation in Eq. (44) with $\boldsymbol{b}(\boldsymbol{y}, t) = \mathbf{A}(t)^{-1}\boldsymbol{y}$, with Eq. (C2b) corresponding to $\boldsymbol{a}(\boldsymbol{x}, t) = \mathbf{A}(t)\boldsymbol{x}$.

Now we transform Eq. (C1) into the $\boldsymbol{y}$ coordinates using Eq. (C2a). Towards this end, we have:

$$\frac{d}{dt}\boldsymbol{x} = \boldsymbol{b}_t(\boldsymbol{y}, t) + \boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t)\frac{d}{dt}\boldsymbol{y} = \boldsymbol{u}\left(\boldsymbol{b}(\boldsymbol{y}, t), t\right)\ , \tag{C3}$$

where $\boldsymbol{b}_t \in \mathbb{R}^n$ denotes the partial differentiation of $\boldsymbol{b}$ with respect to time t, and hence the corresponding velocity field in the $\boldsymbol{y}$ coordinates is

$$\frac{d}{dt}\boldsymbol{y} = \left(\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t)\right)^{-1}\left(\boldsymbol{u}\left(\boldsymbol{b}(\boldsymbol{y}, t), t\right) - \boldsymbol{b}_t(\boldsymbol{y}, t)\right)\ . \tag{C4}$$

Now we can ask and answer the first question.

Question C.1 Suppose $\boldsymbol{x}(t)$ is a trajectory of Eq. (C1). Is $\boldsymbol{y}(t) = \boldsymbol{a}(\boldsymbol{x}(t), t)$ a trajectory of Eq. (C4)?

The answer is "yes", as we now show. First, we need a lemma.

Lemma C.1 $\boldsymbol{a}_t(\boldsymbol{x}, t) = -\left(\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t)\right)^{-1}\boldsymbol{b}_t(\boldsymbol{y}, t)$

Proof: Now $\boldsymbol{x} = \boldsymbol{b}(\boldsymbol{y}, t)$ by definition Eq. (C2a). Differentiating this expression with respect to t gives Eq. (C3). Also, $\boldsymbol{y} = \boldsymbol{a}(\boldsymbol{x}, t)$ by definition Eq. (C2b). Differentiating this equation with respect to t gives:

$$\frac{d}{dt}\boldsymbol{y} = \boldsymbol{a}_t(\boldsymbol{x}, t) + \boldsymbol{a}_{\boldsymbol{x}}(\boldsymbol{x}, t)\frac{d}{dt}\boldsymbol{x} \ , \tag{C5}$$

where $\boldsymbol{a}_t \in \mathbb{R}^n$ denotes a differentiation of $\boldsymbol{a}$ with respect to t and $\boldsymbol{a}_{\boldsymbol{x}} \equiv \frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}} \in \mathbb{R}^{n\times n}$ denotes the Jacobian matrix of $\boldsymbol{a}$ with respect to $\boldsymbol{x}$. Substituting Eq. (C5) into Eq. (C3) gives:

$$\frac{d}{dt}\boldsymbol{x} = \boldsymbol{b}_t(\boldsymbol{y}, t) + \boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t)\ \boldsymbol{a}_t(\boldsymbol{x}, t) \tag{C6}$$

$$+\ \boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t)\ \mathbf{a}_{\boldsymbol{x}}(\boldsymbol{x}, t)\ \frac{d}{dt}\boldsymbol{x} \ . \tag{C7}$$

Finally, $\boldsymbol{b}(\boldsymbol{a}(\boldsymbol{x}, t), t) = \boldsymbol{x}$ by Eq. (C2c). Differentiating this expression with respect to $\boldsymbol{x}$ and using the chain rule gives:

$$\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t)\ \boldsymbol{a}_{\boldsymbol{x}}(\boldsymbol{x}, t) = \mathbf{Id} \ , \tag{C8}$$

where **Id** denotes the identity matrix. Substituting Eq. (C8) into Eq. (C7) gives the result after cancelling the $\frac{d}{dt}\boldsymbol{x}$ terms on both sides.

Now we return to our original Question C.1. Suppose $\boldsymbol{x}(t)$ is a trajectory of Eq. (C1). To prove that $\boldsymbol{y}(t) = \boldsymbol{a}(\boldsymbol{x}(t), t)$ is a trajectory of Eq. (C4), it suffices to show that Eq. (C5) represents the same velocity field as Eq. (C4). Rewriting Eq. (C5) using the fact that $\boldsymbol{x}(t)$ is the trajectory of Eq. (C1) leads to:

$$\frac{d}{dt}\boldsymbol{y} = \boldsymbol{a}_t(\boldsymbol{x}, t) + \boldsymbol{a}_{\boldsymbol{x}}(\boldsymbol{x}, t)\boldsymbol{u}(\boldsymbol{b}(\boldsymbol{y}, t), t) \ . \tag{C9}$$

If we apply Lemma C.1, along with $\boldsymbol{a}_{\boldsymbol{x}}(\boldsymbol{x}, t) = \boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t)^{-1}$ from Eq. (C8), Eq. (C9) becomes Eq. (C4). So $\boldsymbol{y}(t) = \boldsymbol{a}(\boldsymbol{x}(t), t)$ is a trajectory of Eq. (C4) when $\boldsymbol{x}(t)$ is a trajectory of Eq. (C1).

Next we consider the question concerning the behaviour of ISPs under time-dependent coordinate transformations.

Question C.2 Suppose $\boldsymbol{x}_{\rm sp}(t)$ is a curve of ISPs of Eq. (C1). Is $\boldsymbol{y}(t) = \boldsymbol{a}(\boldsymbol{x}_{\rm sp}(t), t)$ a curve of ISPs of Eq. (C4)?

Substituting $\boldsymbol{y}(t) = \boldsymbol{a}(\boldsymbol{x}_{\rm sp}(t), t)$ into Eq. (C4) gives:

$$\frac{d}{dt}\boldsymbol{y}\bigg|_{\boldsymbol{y}=\boldsymbol{a}(\boldsymbol{x}_{\rm sp}(t),t)} =$$

$$-\ \boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{a}(\boldsymbol{x}_{\rm sp}(t), t))^{-1}\boldsymbol{b}_t(\boldsymbol{a}(\boldsymbol{x}_{\rm sp}(t), t), t). \tag{C10}$$

In order for the right-hand side to be zero, $\boldsymbol{b}_t(\boldsymbol{a}(\boldsymbol{x}_{\rm sp}(t), t), t)$ must be zero. Therefore, we conclude that ISPs are not preserved under time-dependent coordinate transformation in general. However, note that if the coordinate transformation is independent of time, ISPs are preserved under time-independent coordinate transformations.

Finally, we consider the question of the behaviour of DHTs under time-dependent coordinate transformations.

Question C.3 Suppose $\boldsymbol{x}_{\rm dht}(t)$ is a DHT of Eq. (C1). Is $\boldsymbol{y}(t) = \boldsymbol{a}(\boldsymbol{x}_{\rm dht}(t), t)$ a DHT of Eq. (C4)?

By knowing that $\boldsymbol{y}(t) = \boldsymbol{a}(\boldsymbol{x}_{\rm dht}(t), t)$ is a trajectory from Question C.1, we now show that such $\boldsymbol{y}(t)$ indeed satisfies the definition of a DHT. To do so, we have to consider Eq. (C1) linearized about $\boldsymbol{x}_{\rm dht}(t)$ and Eq. (C4) linearized about $\boldsymbol{y} = \boldsymbol{a}(\boldsymbol{x}_{\rm dht}(t), t)$. These equations are given by:

$$\frac{d}{dt}\boldsymbol{\xi} = \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}}(\boldsymbol{x}_{\rm dht}(t), t)\boldsymbol{\xi}, \tag{C11}$$

and

$$\frac{d}{dt}\boldsymbol{\eta} = \frac{\partial}{\partial \boldsymbol{y}}\Big(\big(\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t)\big)^{-1}\left(\boldsymbol{u}\left(\boldsymbol{b}(\boldsymbol{y}, t), t\right)\right.$$

$$\left.-\boldsymbol{b}_t(\boldsymbol{y}, t)\right)\Big)\bigg|_{\boldsymbol{y}=\boldsymbol{a}(\boldsymbol{x}_{\rm dht}(t),t)}\boldsymbol{\eta}, \tag{C12}$$

respectively. Let $\mathbf{X}(t, t_0)$ denote the fundamental solution matrix of Eq. (C11) and $\mathbf{Y}(t, t_0)$ denotes the fundamental solution matrix of Eq. (C12).

First we must derive a relationship between the fundamental solution matrices of these two linear systems. We evaluate (C2a) on trajectories $\boldsymbol{x}(t; t_0, \boldsymbol{x}_0)$, $\boldsymbol{y}(t; t_0, \boldsymbol{y}_0)$, and then differentiate with respect to $\boldsymbol{y}_0$ to obtain:

$$\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{y}_0} = \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x}_0}\frac{\partial \boldsymbol{x}_0}{\partial \boldsymbol{y}_0} = \frac{\partial}{\partial \boldsymbol{y}_0}\boldsymbol{b}(\boldsymbol{y}, t) = \boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t)\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{y}_0}, \tag{C13}$$

where we leave out the arguments of the trajectories for notational simplicity. Now, using (C2a) and evaluating the arguments of the trajectories at $t = t_0$ gives:

$$\frac{\partial \boldsymbol{x}_0}{\partial \boldsymbol{y}_0} = \boldsymbol{b}_{\boldsymbol{y}_0}(\boldsymbol{y}_0, t_0). \tag{C14}$$

Now for $\boldsymbol{x}(t; t_0, \boldsymbol{x}_0)$ a trajectory of (C1) and $\boldsymbol{y}(t; t_0, \boldsymbol{y}_0)$ a trajectory of (C4), $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x}_0}(t; t_0, \boldsymbol{x}_0) = \mathbf{X}(t, t_0)$ is the fundamental solution matrix of (C11) and $\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{y}_0}(t; t_0, \boldsymbol{y}_0) = \mathbf{Y}(t, t_0)$ is the fundamental solution matrix of (C12). Combining these relations with (C14), (C13) becomes:

$$\mathbf{X}(t, t_0)\boldsymbol{b}_{\boldsymbol{y}_0}(\boldsymbol{y}_0, t_0) = \boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t)\mathbf{Y}(t, t_0), \tag{C15}$$

and from this equation we easily obtain the following relations:

$$\mathbf{Y}(t, t_0) = \big(\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t)\big)^{-1}\mathbf{X}(t, t_0)\boldsymbol{b}_{\boldsymbol{y}_0}(\boldsymbol{y}_0, t_0),$$

$$\mathbf{Y}^{-1}(t, t_0) = \big(\boldsymbol{b}_{\boldsymbol{y}_0}(\boldsymbol{y}_0, t_0)\big)^{-1}\mathbf{X}^{-1}(t, t_0)\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t). \tag{C16}$$

First we describe the general idea why this result should be true; then we provide the rigorous argument. We know that $\mathbf{X}(t, t_0)$ has an exponential dichotomy since $\boldsymbol{x}_{\rm dht}(t)$ is a DHT. Therefore, since Eq. (C15) is an equality, $\mathbf{Y}(t, t_0)$ must have an exponential dichotomy since $\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}, t)$ is bounded in $\boldsymbol{y}$ and exhibits no exponential growth or decay in t. Therefore, $\boldsymbol{y} = \boldsymbol{a}(\boldsymbol{x}_{\rm dht}(t), t)$ is hyperbolic.

We can make this argument rigorous as follows. If $\boldsymbol{x}_{\rm dht}(t)$ is hyperbolic, then Eq. (C11) has an exponential dichotomy,

i.e. there exists a projection operator $\mathbf{P}$ (i.e. $\mathbf{P}^2 = \mathbf{P}$), and positive constants K_X, L_X, α_X and β_X such that:

$$\left|\mathbf{X}(t,t_0)\mathbf{P}\mathbf{X}^{-1}(s,t_0)\right| \leq K_X e^{-\alpha_X(t-s)}, \quad \text{for} \quad t \geq s, \tag{C17a}$$

$$\left|\mathbf{X}(t,t_0)(\mathbf{Id}-\mathbf{P})\mathbf{X}^{-1}(s,t_0)\right| \leq L_X e^{-\beta_X(s-t)}, \quad \text{for} \quad s \geq t. \tag{C17b}$$

Now we argue that Eq. (C12) also has an exponential dichotomy by showing that there exists a projection operator $\hat{\mathbf{P}}$ (i.e. $\hat{\mathbf{P}}^2 = \hat{\mathbf{P}}$) and positive constants K_Y, L_Y, α_Y and β_Y such that:

$$\left|\mathbf{Y}(t,t_0)\hat{\mathbf{P}}\mathbf{Y}^{-1}(s,t_0)\right| \leq K_Y e^{-\alpha_Y(t-s)}, \quad \text{for} \quad t \geq s, \tag{C18a}$$

$$\left|\mathbf{Y}(t,t_0)(\mathbf{Id}-\hat{\mathbf{P}})\mathbf{Y}^{-1}(s,t_0)\right| \leq L_Y e^{-\beta_Y(s-t)}, \quad \text{for} \quad s \geq t. \tag{C18b}$$

We define the projection operator in the $\boldsymbol{y}$ coordinates as:

$$\hat{\mathbf{P}} \equiv \left(\boldsymbol{b}_{\boldsymbol{y}_0}(\boldsymbol{y}_0,t_0)\right)^{-1}\mathbf{P}\boldsymbol{b}_{\boldsymbol{y}_0}(\boldsymbol{y}_0,t_0), \tag{C19}$$

and a simple calculation shows that $\hat{\mathbf{P}}^2 = \hat{\mathbf{P}}$.

Next we make an important assumption of the growth of the spatial derivative of the coordinate transformation with respect to time.

Assumption C.1 *We assume that for all* $\boldsymbol{y}$, $\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y},t)$ *and* $\left(\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y},t)\right)^{-1}$ *satisfy the following bounds:*

$$\left|\left(\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y},t)\right)^{-1}\mathbf{X}(t,t_0)\mathbf{P}\mathbf{X}^{-1}(s,t_0)\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y},s)\right| \leq K_Y e^{-\alpha_Y(t-s)}, \quad \textit{for} \quad t \geq s, \tag{C20}$$

$$\left|\left(\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y},t)\right)^{-1}\mathbf{X}(t,t_0)(\mathbf{Id}-\mathbf{P})\mathbf{X}^{-1}(s,t_0)\,\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y},s)\right| \leq L_Y e^{-\beta_Y(s-t)}, \quad \textit{for} \quad s \geq t, \tag{C21}$$

where K_Y, L_Y, α_Y, *and* β_Y *are positive constants.*

Substituting Eq. (C16) and Eq. (C19) into Eq. (C18a), and using Eq. (C20), gives rise to the following estimate:

$$\begin{aligned}\left|\mathbf{Y}(t,t_0)\hat{\mathbf{P}}\mathbf{Y}^{-1}(s,t_0)\right| &\leq \left|\left(\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y},t)\right)^{-1}\mathbf{X}(t,t_0)\mathbf{P}\mathbf{X}^{-1}(s,t_0)\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y},s)\right| \\ &\leq K_Y e^{-\alpha_Y(t-s)}, \quad \text{for} \quad t \geq s. \end{aligned} \tag{C22}$$

The second condition follows similarly. Substituting Eq. (C16) and Eq. (C19) into Eq. (C18b), and using Eq. (C21)

$$\begin{aligned}\left|\mathbf{Y}(t,t_0)(\mathbf{Id}-\hat{\mathbf{P}})\mathbf{Y}^{-1}(s,t_0)\right| &\leq \left|\left(\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y},t)\right)^{-1}\mathbf{X}(t,t_0)(\mathbf{Id}-\mathbf{P})\mathbf{X}^{-1}(s,t_0)\,\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y},s)\right| \\ &\leq L_Y e^{-\beta_Y(s-t)}, \quad \text{for} \quad s \geq t. \end{aligned} \tag{C23}$$

For the case of the time-dependent linear coordinate change discussed in Appendix A, we have

$$\boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y},t)^{-1} = \mathbf{A}(t), \qquad \boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y},s) = \mathbf{A}(s)^{-1}.$$

By construction, the determinant of $\mathbf{A}(t) > 0$ is bounded over the time interval $t \in [t_0, t_L]$ of the coordinate change. Therefore, if we define

$$\det\mathbf{A}_{\max} = \max\,\{\det\mathbf{A}(t)\} > 0,$$
$$\det\mathbf{A}_{\min} = \min\,\{\det\mathbf{A}(t)\}) > 0,$$

over $t \in [t_0, t_L]$, then a postive constant K' can be chosen such that

$$K' > \frac{\det\mathbf{A}_{\max}}{\det\mathbf{A}_{\min}}\,. \tag{C24}$$

Acknowledgement. We would like to thank the referees for the interest they have taken in this work and the care they have taken in reading the manuscript.

We would also like to acknowledge discussions with Francois Lekien in the early stages of this work.

The work of Kayo Ide was supported by ONR Grant No. N00014-99-1-0020. The work of Des Small and Stephen Wiggins was supported by ONR Grant No. N00014-01-1-0769.

References

Aref, H. and El Naschie, M. S.: (Eds) Chaos Applied to Fluid Mixing, Chaos, Solitons, and Fractals, 4, 6, 1–380 1994.

Babiano, A., Provenzale, A., and Vulpiani, A.: (Eds), Chaotic Advection, Tracer Dynamics, and Turbulent Dispersion. Proceedings of the NATO Advanced Research Workshop and EGS Topical Workshop on Chaotic Advection, Conference Centre Sereno di Gavo, Italy, 24–28 May 1993. Physica D76, 1–329, 1994.

Coppel, W. A.: Dichotomies in Stability Theory, Springer Lecture Notes in Mathematics, 629, Springer-Verlag, New York, Heidelberg, Berlin, 1978.

Coulliette, C. and Wiggins, S.: Intergyre Transport in a Wind-Driven, Quasigeostrophic Double Gyre: An Application of Lobe Dynamics, Non. Proc. Geophys., 7, 59–85, 2000.

del-Castillo-Negrete, D. and Morrison, P. J.: Chaotic transport of Rossby waves in shear flow, Phys. Fluids A, 5(4), 948–965, 1993.

Dijkstra, H. A. and Katsman, C. A.: Temporal variability of the wind-driven quasi-geostrophic double gyre ocean circulation: basic bifurcation diagram, Geophys. Astrophys. Fluid Dynamics, 85, 195–232, 1997.

Duan, J. and Wiggins, S.: Fluid exchange across a meandering jet with quasiperiodic variability, J. Phys. Oceanogr., 26, 7, 1176–1188, 1996.

Duan, J. and Wiggins, S.: Lagrangian transport and chaos in the near wake of the flow around an obstacle: a numerical implementation of lobe dynamics, Non. Proc. Geophys., 4, 125–136, 1997.

Haller, G.: Lagrangian structures and the rate of strain in a partition of two-dimensional turbulence, Physics of Fluids A, accepted, 2001.

Haller, G. and Poje, A.: Finite time transport in aperiodic flows, Physica D, 119, 3/4, 352–380, 1998.

Henry, D.: Geometric theory of semilinear parabolic equations, Lecture Notes in Mathematics, 840, Springer-Verlag, New York, Heidelberg, Berlin, 1981.

Ju, N., Small, D., and Wiggins, S.: Existence and computation of hyperbolic trajectories of aperiodically time dependent vector fields and their approximations, Int. J. Bifurcation and Chaos, accepted, 2002.

Lozier, M. and Riser, S.: Potential vorticity sources and sinks in a quasi-geostrophic ocean: beyond western boundary currents, J. Phys Oceanogr., 20, 1608–1627, 1989.

Malhotra, N. and Wiggins, S.: Geometric Structures, Lobe Dynamics, and Lagrangian Transport in Flows with Aperiodic Time Dependence, with Applications to Rossby Wave Flow, J. Nonlinear Sci., 8, 401–456, 1998.

Massera, J. L. and Schäffer, J. J.: Linear differential equations and function spaces, Academic Press, New York, 1966.

Miller, P. D., Jones, C. K. R. T., Rogerson, A., and Pratt, L. J.: Quantifying transport in numerically generated velocity fields, Physica D, 110, 105–122, 1997.

Muldowney, J. S.: Dichotomies and asymptotic behaviour for linear differential systems, Trans. A. M. S., 283, 2, 465–484, 1984.

Ngan, K. and Shepherd, T. G.: Chaotic mixing and transport in Rossby-wave critical layers, J. Fluid Mech., 334, 315–351, 1997.

Pierrehumbert, R. T.: Chaotic mixing of tracer and vorticity by modulated traveling Rossby waves, Geophys. Astrophys. Fluid Dyn., 58, 285–319, 1991a.

Pierrehumbert, R. T.: Large-scale horizontal mixing in planetary atmospheres, Phys. Fluids A, 3, 5, 1250–1260, 1991b.

Rogerson, A., Miller, P. D., Pratt, L. J., and Jones, C. K. R. T. J.: Lagrangian motion and fluid exchange in a barotropic meandering jet, J. Phys. Oceanogr., 29, 10, 2635–2655, 1999.

Rowley, C.: A Modeling Study of the North Atlantic Current, Ph.D. thesis, University of Rhode Island, Kingston, RI, 1996.

Samelson, R. M.: Fluid exchange across a meandering jet, J. Phys. Oceanogr., 22, 431–440, 1992.

Shariff, K., Pulliam, T., and Ottino, J.: A dynamical systems analysis of kinematics in the time-periodic wake of a circular cylinder, in: Vortex Dynamics and vortex methods, (Eds) Anderson, C. and Greengard, C., Proc. AMS-SIAM Conf., Lectures in Applied Mathematics, American Mathematical Society, Providence, 1992.

Szeri, A., Leal, L. G., and Wiggins, S.: On the Dynamics of Suspended Microstructure in Unsteady, Spatially Inhomogeneous Two-Dimensional Fluid Flows, J. Fluid Mech., 228, 207–241, 1991.

Wiggins, S.: Chaotic Transport in Dynamical Systems, Springer-Verlag, New York, 1992.