Accepted to SIAM Journal of Applied Dynamical Systems, 2022

Melnikov theory for two-dimensional manifolds in three-dimensional flows

K.G.D. Sulalitha Priyankara[1,4], Sanjeeva Balasuriya[2], and Erik Bollt[3,4]

[1]Department of Mathematics, Clarkson University, Potsdam, NY 13699, USA.
[2]School of Mathematical Sciences, University of Adelaide, Adelaide, SA 5005, Australia.
[3]Department of Electrical and Computer Engineering, Clarkson University, Potsdam, NY 13699, USA.
[4]Clarkson Center for Complex Systems Science (C^3S^2), Potsdam, NY 13699, USA

July 4, 2022

Abstract

We present a geometric Melnikov method to analyze a two-dimensional stable or unstable manifold associated with a saddle point in three-dimensional non-volume preserving autonomous flows. The time-varying perturbed location of such a manifold is obtained under very general, non-volume preserving and with arbitrary time-dependence, perturbations. We demonstrate the explicit computability of the leading-order spatio-temporal location of the manifold using our formulas. In unperturbed situations with a two-dimensional heteroclinic manifold, we adapt our theory to quantify the splitting into a stable and unstable manifold, and thereby obtain an instantaneous flux quantification in terms of a Melnikov function. The time-varying instantaneous flux theory does not require any intersections between perturbed manifolds, nor rely on descriptions of lobe dynamics. Our theory has specific application to transport in fluid mechanics, where the flow is in three dimensions and flow separators in forward/backward time are two-dimensional stable/unstable manifolds. We demonstrate our theory using both the classical and the swirling versions of Hill's spherical vortex.

1 Introduction

Melnikov methods were originally introduced [45] to analyze how a homoclinic connection (a coincident stable and unstable manifold of a saddle fixed point) splits when an autonomous continuous dynamical system is perturbed time-periodically. This paper specifically develops a Melnikov theory for a flow which is *not* time-periodic, and moreover principally focusses on a situation which has not been studied before, viz., determining locations of a perturbed two-dimensional stable (or unstable) manifold in a three-dimensional flow. This does not require other common aspects usual to Melnikov methods in flows, i.e., time-periodicity of the perturbation, homo- or hetero-clinic situations in which a stable and an unstable manifold coincide, compact manifolds, or Hamiltonian or volume-preserving flows. As a secondary focus, the theory is adapted to quantifying the instantaneous flux in the situation of a broken heteroclinic manifold in a lobe-independent fashion (as is necessary because time-periodicity is not imposed). The emphasis on three dimensions (a less examined situation in Melnikov methods) is crucial because of fluid mechanical implications: two-dimensional time-varying stable/unstable manifolds form important flow barriers [7] for fluid transport, and quantifying transport across a broken barrier as a tool for flux optimization or control.

The time-periodicity of the perturbation in the classical Melnikov approach [45] allows for an interesting interpretation of phase space transport. Rather than viewing as a *flow* (i.e., a continuous dynamical system), one can view the dynamics in terms of a *map* (a discrete dynamical system) which samples the flow at the period T of the perturbation; this is called a Poincaré map [30, 66]. The originally coincident stable and unstable manifold of a fixed point of a flow can then be viewed as the coincident stable and unstable manifold

of a fixed point of a map. Upon perturbing, these stable and unstable manifolds (in the sense of the map) persist for a nearby fixed point, but need no longer coincide. The classical Melnikov approach establishes a function which effectively measures a scaled orthogonal splitting distance between the manifolds, and the simple zeros of this function guarantee transverse intersection points [30, 66]. Indeed, one zero guarantees infinitely many, since the intersection point must itself be mapped to another intersection point because of the invariance of the stable and unstable manifolds. The presence of such zeros serves as one of the few methods for 'proving chaos,' in this instance via the Smale-Birkhoff theorem [1, 30, 66]. (The basic Melnikov function also works for a heteroclinic situation in which the stable and unstable manifolds were from *different* fixed points; however, transverse intersections do not automatically imply chaotic motion in this case except under additional conditions [19, 54].) The intersection regions between the manifolds—often called *lobes*—have fundamental importance: the dynamics of how they transport to one another under iteration of the Poincaré map, i.e., *lobe dynamics*, underlines the chaotic motion. Indeed, lobe dynamics is well-studied purely from the context of maps, without necessarily being generated from a flow [42, 44]. This paper does *not* focus on maps or on time-periodic flows, and thus lobe dynamics is mostly not relevant; however, connections to lobe dynamics will be established under suitable time-periodic restrictions.

Many reinterpretations and extensions of Melnikov theory exist (see the review chapters in [7]). The Melnikov function has been extended for implicitly defined differential equations [16], heteroclinic situations [6, 54, 13], for stochastic perturbations [62, 69], singular perturbations [29, 52], nonhyperbolic situations [65, 72], fixed points at infinity [65], degenerate homoclinics [64] and discontinuous [25, 20, 15, 35] and impulsive [8] vector fields. The standard development is usually extendible to situations in which the heteroclinic manifold is associated with a compact normally hyperbolic invariant set (rather than just a fixed point), by appealing to the concept of exponential dichotomies [23, 53, 17] rather than eigenvalues at the fixed point [22]. Higher-dimensional (greater than two) extensions are also available under some restrictions—usually volume-preserving and/or Hamiltonian flows and/or presence of symmetries [13, 37, 26, 27, 68, 67, 33]—or situations in which the Melnikov function is based on solutions to an equation which cannot be written explicitly [28, 70, 22]. Higher-order Melnikov methods have also been developed [25, 22]. It is also possible to couch the transverse intersection problem in a functional analytic, rather than a geometric, form instead [21, 53, 17, 59, 37, 24, 22]—a method which transforms naturally to higher dimensions. However, in these and other general higher-dimensional cases the kernel of the Melnikov integral contains an *abstract* function which is not in general expressible explicitly for actual computation [28, 70, 22, 21, 53, 17, 59, 37, 24]. There have also been Melnikov methods developed purely in the context of *maps* [38, 39, 40]; in higher-dimensions, again the issue of explicitly determining relevant functions within the 'integral' (an infinite summation in this case [38]) remains [38].

Most applications of Melnikov theory described above relate to defining a Melnikov function whose simple zeros imply the persistence of a homo/hetero-clinic connection [45, 21, 53, 17, 59, 28, 33, 13, 25, 20, 15, 35, 65, 29, 52, 16, 26, 67, 69, e.g.]. (In higher-dimensional situations, this generalizes to a Melnikov *vector* [70, 22] the zeros of whose components represents intersections along each normal vector at a location on an unperturbed heteroclinic trajectory.) In these cases, the Melnikov function does not necessarily express exactly the physical distance between stable and unstable manifolds which separate off the broken homo/hetero-clinic manifold, but rather a nonuniformly scaled signed distance. In the functional analytical developments in particular, this scaling is hidden; it is only the function's zeros which give the pertinent information on where the stable and unstable manifolds intersect. It is less well-known that Melnikov developments can be adapted to characterize the *location* of a perturbed stable or unstable manifold, irrespective of whether it is associated with a homo/hetero-clinic situation [22, 70, 6]. In this paper, we develop the relevant theory to locate two-dimensional time-varying stable (or unstable) manifolds of a hyperbolic fixed point in a three-dimensional flow, due to the inclusion of a perturbation whose spatial derivatives are bounded for all time. Note that we do not require volume preservation in either the unperturbed of perturbed flows, nor time-periodicity in the perturbation. We emphasize that this development does not require a homo- or hetero-clinic situation, and indeed the unperturbed manifold may be noncompact. The theory develops a *leading-order* computable approximation for spatio-temporally varying location for the perturbed manifold, which (as is to be expected) is restricted to finite times and a finite extent. This is the first of the main results of this paper, presented in Theorems 1 and 2, for respectively an unstable or a stable manifold.

Most typically, Melnikov developments seek an integral expression over $\mathbb{R}$ which is often called a Melnikov function. Such an integral is well-known to play a role when one wishes to determine intersections in a broken

homo/heteroclinic situation. A similar definite integral, but with limits not over all of $\mathbb{R}$, also appears when we locate perturbed stable/unstable manifolds. In general, such integrals contain as kernel a particular function the knowledge of which is crucial to represent the Melnikov integral. In Hamiltonian [33, 26, e.g.], as well as in volume-preserving unperturbed situations with a nondegenerate conserved quantity [13, 47], explicit forms for this kernel function can be determined. For more general situations, the kernel function can be expressed in more abstract terms: it is related to the fundamental matrix solution to the adjoint of the variational equation along the relevant homo/hetero-clinic trajectory. Given that this adjoint equation is *nonautonomous*, its solutions *cannot* usually be written down explicitly, unless in special situations such as in two dimensions. Therefore, while a Melnikov function might be expressible for such situations in an abstract sense [28, 21, 37, 70, 22, 24], it is usually not computable. Put another way, most Melnikov developments in dimensions greater than two, or which are not Hamiltonian, provide a theoretical result which is difficult to apply. Knowing an explicit formula for the kernel function is such situations is therefore valuable. In our development, we are able to provide an explicit expression for it in our three-dimensional setting (this appears in the integrals appearing in Theorems 1, 2 and 3), without additional conditions such as volume-preservation or the presence of a conserved quantity. The formula is related to a triple scalar product associated with a parametrization of the two-dimensional manifold. Thus, the Melnikov function that we develop for locating the perturbed version of such a manifold is *computable*, unlike that in most higher-dimensional non-Hamiltonian Melnikov developments.

We have mentioned that Melnikov approaches are more frequently related to determining intersections between the stable and unstable manifolds resulting from a broken homo/hetero-clinic manifold. The new theory that we develop specializes to such a situation as well, and thus we are able to present a computable Meknikov function in a non-Hamiltonian, non-volume-preserving situation, in a dimension greater than two (see Theorem 3). Moreover, we are able to quantify transport across the broken heteroclinic in terms of this Melnikov function, under general time-dependence. This is the second of the main results of this paper, which we present in Theorem 5.

Quantifying transport when a heteroclinic (a flow-separating curve) in two-dimensions is broken is a well-studied problem. In two dimensions, the interweaving of the stable and unstable manifolds which split off the heteroclinic generates lobes, and transport can be characterized via the beautiful theory of lobe dynamics and turnstiles [57, 66], which builds on similar concepts for maps [42, 44]. This theory is confined to two-dimensional flows, and for an area of a lobe to be a well-defined characterizer of the transport engendered across the broken heteroclinic, several other features need to be in place: the flow needs to be area-preserving, and the perturbation 'harmonic' in that it can be written as a spatially-varying two-dimensional function multiplied by a sinusoid in time. The area of a lobe then expresses the amount of fluid transported across the broken heteroclinic during the time-periodicity of the perturbation, and can be expressed in terms of a definite integral of an appropriate Melnikov function [57, 66]. More general time-periodic situations generically do not have well-defined lobe areas because there can be many, differently sized lobes relevant to one iteration of the time-periodic map, or indeed no lobes at all because the perturbed manifolds do not intersect [2]. Obtaining a transport characterization in more general time-*aperiodic* situations therefore requires a slightly different approach, and has been provided in two-dimensional flows via a time-dependent flux idea [4, 7, 10]. As befitting any assessment of transport, this takes into account the *Lagrangian* motion of trajectories, rather than an Eulerian flux. (This terminology stems from fluid mechanics in which 'Lagrangian' refers to following the flow, while 'Eulerian' in this context would mean measuring transport across fixed surfaces in space, without taking into account that these surfaces are themselves moving due to the flow.)

We are able to extend these broken heteroclinic results to our current three-dimensional setting. We obtain in Theorem 5 an expression for the instantaneous flux which is valid for general time-dependence in the perturbation, as well as non-volume-preserving flows. This flux expression—which establishes a direct relationship with an appropriate Melnikov function—is valid whether lobes (generated through intersections of the broken stable and unstable manifolds) exist or not, and is therefore substantially more general than lobe dynamics descriptions. We point out though that in time-harmonic, volume-preserving situations, a nice analog of lobe dynamics is seen to occur; in this case, it is lobe *volumes* rather than areas that is relevant. We specifically obtain an analytic formula for leading-order lobe volume in terms of an appropriate integral of the Melnikov function (Theorem 4), thereby extending a well-known two-dimensional result for lobe areas [57, 66]. This expression is indeed valid in the general situation of time-aperiodicity and non-volume-preservation, but its interpretability to transport in the sense of lobe dynamics requires time-periodicity and

volume-preservation.

We remark that the transport characterization we provide for three-dimensional flows is motivated strongly by fluid mechanics. Realistic flows in fluids are inherently three-dimensional, and internal flow separators must therefore be two-dimensional entities. Two-dimensional stable and unstable manifolds are primary candidates for such flow separators. Locating them and tracking their motion is therefore fundamental in determining boundaries between coherently moving regions of fluids; this is related to the field of 'Lagrangian coherent structures' [14]. In particular, characterizing a flow rate (a flux, i.e., a volume of fluid per unit time) across a broken heteroclinic provides a direct assessment of the transport between two previously separated coherent regions. It is precisely this which we are able to provide with our flux theory. Similar theory has been used extensively for *two*-dimensional flows with one-dimensional flow separators due to the existence of pertinent Melnikov theory [57, 58, 2, 4, e.g.], and can even give insight into how to perturb a flow to optimize mixing [11, 3]. However, genuine fluid flows are *three*-dimensional, and hence our current theory can extend these methods to significantly more realistic flows. While there are other works which study transport in the context of three-dimensional flows, they tend to possess restrictions which we will *not* impose: axisymmetry [33], time-periodicity [33, 51], volume-preservation [33, 51, 47, 46, 13], presence of a conserved quantity [51, 47, 46, 13], or a one-step finite-time flow [50]. Our Melnikov development in this paper is purposefully *geometric* to help provide intuition in this realistic situation where the phase space is equivalent to three-dimensional physical space.

This paper is organized as follows. In Sec. 2, we build the general Melnikov theory for quantifying the spatio-temporal movement of a two-dimensional invariant manifold of a three-dimensional non-volume preserving flow. We develop computable expressions for locating such a manifold under general time-aperiodic perturbation, expressing points on this manifold in parametric form. This is the first of our main results (Theorems 1 and 2), whose application we demonstrate via a simple example in Sec. 2.3. This theory is adapted in Sec. 3 for the particular situation when the unperturbed flow possesses a two-dimensional *heteroclinic* manifold. The Melnikov function we formulate can be used to identify transverse intersections of the perturbed stable and unstable manifolds, as well as to characterize instantaneous flux. We emphasize that there is no requirement for either time-periodicity or volume preservation, neither is it necessary for lobes to form. The flux theory still applies if there are no intersections of perturbed stable and unstable manifolds. This development we use to rationalize the flux explicitly as a time-dependent function, and the accompanying simple formula we obtain in terms of the Melnikov function, is the second of our main results (Theorem 5). We also establish connections to more standard situations in two dimensions (perturbations with separable sinusoidal time-dependence in an overall area-preserving flow) in which lobe dynamics applies [57, 66]; in this case, a nice method of characterizing lobe boundaries (i.e., persistence heteroclinic points) using a Fourier transform procedure is developed. Transport can be described in terms of a version of lobe dynamics, and might be quantified in terms of lobe *volumes*, which we express in terms of the Melnikov function as well. In Sec. 4, we apply the theory to both the classical Hill's spherical vortex [32], and a modification incorporating swirl [61], respectively. These choices are motivated by fluid flows in cylinders with rotating endwalls, which are known to generate such vortices which subsequently breakdown due to perturbations [36, 63, 41, 33, 13]. We conclude in Sec. 5 with some remarks on extensions and applications.

2 Spatiotemporal motion of 2-D manifolds

In this section, we build a Melnikov theory for two-dimensional invariant manifolds that are attached to saddle points in three-dimensional autonomous dynamical systems. We emphasize that the theory does *not* require a homo/hetero-clinic framework, which is the focus of most classical Melnikov approaches. Rather, our theory serves to characterize the *location*, as it varies with time, of a perturbed two-dimensional invariant manifold when the flow is subject to a very general perturbation. We consider the system

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \epsilon\, \boldsymbol{g}(\boldsymbol{x}, t)\,, \tag{1}$$

in which $\boldsymbol{x} \in \Omega \subset \mathbb{R}^3$, $\boldsymbol{f} : \Omega \to \mathbb{R}^3$, $\boldsymbol{g} : \Omega \times \mathbb{R} \to \mathbb{R}^3$ and $0 < \epsilon \ll 1$. The $\epsilon = 0$ system of (1) is considered the *unperturbed* system.

Hypothesis 1. Conditions on the unperturbed flow and the perturbation:

1. The function $\boldsymbol{f} \in \mathbf{C}^2(\Omega)$, and $\boldsymbol{Df}$ is bounded in Ω.

2. The point $\boldsymbol{a} \in \mathbb{R}^3$ is a saddle fixed point of the unperturbed system (i.e., (1) when $\epsilon = 0$). Thus, $\boldsymbol{f}(\boldsymbol{a}) = \boldsymbol{0}$, and the eigenvalues of $\boldsymbol{Df}(\boldsymbol{a})$ fall into one of the following categories:

   - Case 1: one is negative, and the other two have positive real parts, or
   - Case 2: one is positive, and the other two have negative real parts.

3. The eigenvectors associated with the two-dimensional (unstable or stable, corresponding to cases 1 or 2 respectively) subspace of $D\boldsymbol{f}(\boldsymbol{a})$ are linearly independent.

4. The two-dimensional stable or unstable manifold identified above is C^2-smooth.

5. For any $t \in \mathbb{R}$, the perturbing function $\boldsymbol{g}(\boldsymbol{x}, t) \in \mathbf{C}^2(\Omega)$. Additionally, both $\boldsymbol{g}$ and $\boldsymbol{Dg}$ are bounded in $\Omega \times \mathbb{R}$.

In seeking expressions for the perturbed two-dimensional invariant manifold, we will focus on the two possibilities for the eigenvalues separately.

Before proceeding, we remark that the results we present here can be extended to the situation where the unperturbed two-dimensional invariant manifold is of a compact *invariant set* A which is normally hyperbolic with the appropriate attraction/repulsion as intimated by the two cases in the hypothesis. However, we will not develop the theory in this generality because (i) this will require working via exponential dichotomies [23, 53, 71, 70, 22] with implied projection operators stated only in an abstract sense, which will interfere with our quest to provide computable expressions, and (ii) the possible geometries will need to be catagorized and interpreted on a case-by-case basis. This latter issue is because, for example, A may be a periodic orbit or a curve of fixed points.

2.1 Displacement of 2D unstable manifold

First, consider case 1, when the system (1) when $\epsilon = 0$ has one negative eigenvalue and two eigenvalues with positive real parts at the point $\boldsymbol{a}$. So the unperturbed system possesses a one-dimensional stable manifold and a two-dimensional unstable manifold. We are interested in characterizing the impact of the perturbation (i.e., $\epsilon \neq 0$ in (1)) on the two-dimensional manifold, $\Gamma^u(\boldsymbol{a})$. We will identify different trajectories on $\Gamma^u(\boldsymbol{a})$ by the parameter $\alpha \in \mathrm{S}^1$, that is, $\alpha \in [0, 1)$, periodically extended with interval 1. To explain this identification, consider the tangent plane to $\Gamma^u(\boldsymbol{a})$ at $\boldsymbol{a}$, and consider a small circle of radius δ centered at $\boldsymbol{a}$. We can think of α as the angle going around the circle divided by 2π (having chosen an $\alpha = 0$ location), and at each α value, the circle will intersect exactly one trajectory which lies on $\Gamma^u(\boldsymbol{a})$. This is so whether $\boldsymbol{Df}(\boldsymbol{a})$ has two negative, or two complex with negative real part, eigenvalues; in the former case there will be non-spiralling trajectories, and in the latter case there will be spiralling trajectories, going in to $\boldsymbol{a}$. In either situation α as explained can be used to parametrize the choice of trajectory. Next, the time-variation along each trajectory will be parametrized by p. Thus, if $\bar{\boldsymbol{x}}^u(p, \alpha)$ is such a trajectory indexed by α, we have

$$\frac{\partial \bar{\boldsymbol{x}}^u(p, \alpha)}{\partial p} = \boldsymbol{f}\left(\bar{\boldsymbol{x}}^u(p, \alpha)\right) ,$$

because it is a trajectory of (1) with $\epsilon = 0$. Now, the trajectory can extend outwards in various ways (e.g., may be a heteroclinic trajectory and thus approach a different critical point, or may go to infinity). To account for this, we limit p to be in the set $(-\infty, P]$ for any finite P, and choose the p-parametrization of nearby trajectories continuously. Thus, we can use $(p, \alpha) \in [-\infty, P] \times \mathrm{S}^1$ to parametrize a restricted version of $\Gamma^u(\boldsymbol{a})$ which avoids having to specify the limiting behavior of the unstable manifold trajectories $\bar{\boldsymbol{x}}^u(p, \alpha)$ in the limit $p \to \infty$, while realizing that

$$\lim_{p \to -\infty} \bar{\boldsymbol{x}}^u(p, \alpha) = \boldsymbol{a} .$$

Fig. 1 demonstrates the two-dimensional unstable manifold attached to the fixed point $\boldsymbol{a}$, and illustrating the roles of (p, α) in parametrizing the manifold. We assume that the parametrization by (p, α) is C^1-smooth.

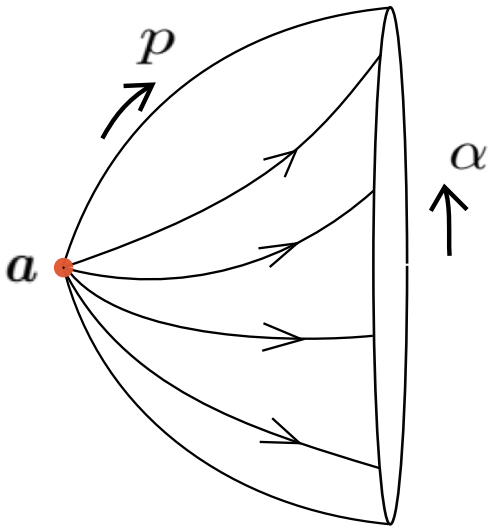


Figure 1: The two-dimensional unperturbed unstable manifold $\Gamma^u(\boldsymbol{a})$ for case 1, with the red dot representing the saddle fixed point. The (p, α)-dependence of trajectories $(\bar{\boldsymbol{x}}^u(p, \alpha))$ can be used to parametrize $\Gamma^u(\boldsymbol{a})$.

This picture corresponds to $\boldsymbol{Df}(\boldsymbol{a})$ having two positive eigenvalues; if they are complex with positive real parts instead, the trajectories will spiral (swirl) away from $\boldsymbol{a}$ instead.

While so far we have described the situation with respect to the system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$, we may consider instead the behavior within the *augmented system* where we append the equation $\dot{t} = 1$. In this situation, the phase space is now $\Omega \times \mathbb{R}$, and the saddle fixed point $\boldsymbol{a}$ becomes a hyperbolic trajectory $(\boldsymbol{a}, t)$. Now, the unstable manifold in the four-dimensional augmented phase space will be parameterized as $(p, \alpha, t) \in (-\infty, P] \times \mathrm{S}^1 \times (-\infty, T]$ for any finite T. Specifically, the point $(\bar{\boldsymbol{x}}^u(p, \alpha), t)$, where α and p are spatial parameters and t is time, now parameterizes the augmented unstable manifold.

Now consider the impact of introducing the perturbation by setting $\epsilon \neq 0$ in (1). Since $\boldsymbol{g}$ and $\boldsymbol{Dg}$ are sufficiently smooth and bounded as per Hypothesis 1, the hyperbolic trajectory $(\boldsymbol{a}, t)$ perturbs to $(\boldsymbol{a}_\epsilon(t), t)$ which is $\mathcal{O}(\epsilon)$-close to $\boldsymbol{a}$ for $t \in [-\infty, T]$. We caution that $\boldsymbol{a}_\epsilon(t)$ *cannot* be obtained by seeking instantaneous fixed points of (1), but instead is defined in terms of exponential dichotomies [23, 53], and is in general difficult to compute (for a perturbative approximation in two-dimensions, see [9, 34]). Next, persistence results of invariant manifolds associated with hyperbolic fixed points [71] indicate the presence of a perturbed unstable manifold, $\Gamma^u_\epsilon(\boldsymbol{a}_\epsilon, t)$ that is $\mathcal{O}(\epsilon)$-close to $\Gamma^u(\boldsymbol{a})$ at finite times t. More specifically, suppose that $t \in (-\infty, T]$ is fixed, and we view the projections of the relevant manifolds on this time-slice t. See Fig. 2, where the unperturbed manifold is in black, and the perturbed manifold is indicated in red. There is a point $\boldsymbol{x}^u(p, \alpha, \epsilon, t)$ on the perturbed manifold which is $\mathcal{O}(\epsilon)$-close to $\bar{\boldsymbol{x}}^u(p, \alpha)$. Our aim is to quantify the distance $d^u(p, \alpha, \epsilon, t)$ obtained by projecting the vector $\boldsymbol{x}^u(p, \alpha, \epsilon, t) - \bar{\boldsymbol{x}}^u(p, \alpha)$ on to the normal vector drawn to $\Gamma^u(\boldsymbol{a})$ at the point $\bar{\boldsymbol{x}}^u(p, \alpha)$. By doing so, we will be able to give the location of the perturbed manifold $\Gamma^u_\epsilon(\boldsymbol{a}_\epsilon, t)$ parametrized by time t and the spatial variables (p, α), to leading-order in ϵ.

We note that the vector $\boldsymbol{f}(\bar{\boldsymbol{x}}^u(p, \alpha))$ lies along a unstable manifold trajectory, since this is the velocity field at the point $\bar{\boldsymbol{x}}^u(p, \alpha)$. Moreover, $\bar{\boldsymbol{x}}^u_\alpha(p, \alpha)$, where the subscript α represents the partial derivative in this instance, is another vector which is tangential to $\Gamma^u(\boldsymbol{a})$. This vector must be transverse to $\boldsymbol{f}(\bar{\boldsymbol{x}}^u(p, \alpha))$; if tangential at any value $(p, \alpha) \in (-\infty, P] \times \mathrm{S}^1$, that would relate to a failure of the trajectories (labelled by α) to foliate $\Gamma^u(\boldsymbol{a})$. Thus, the standard cross product between these two vectors is nonzero, and normal to $\Gamma^u(\boldsymbol{a})$ at $\bar{\boldsymbol{x}}^u(p, \alpha)$. We will use the wedge notation for the cross product. Hence at time $t \in (-\infty, T]$, the distance between $\boldsymbol{x}^u(p, \alpha, \epsilon, t)$ and $\bar{\boldsymbol{x}}^u(p, \alpha)$ is measured perpendicular to original unperturbed manifold can be represented as

$$d^u(p, \alpha, \epsilon, t) = \frac{\boldsymbol{f}(\bar{\boldsymbol{x}}^u(p, \alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(p, \alpha)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}^u(p, \alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(p, \alpha)|} \cdot [\boldsymbol{x}^u(p, \alpha, \epsilon, t) - \bar{\boldsymbol{x}}^u(p, \alpha)] \quad , \quad (p, \alpha, t) \in (-\infty, P] \times \mathrm{S}^1 \times (-\infty, T]\,. \tag{2}$$

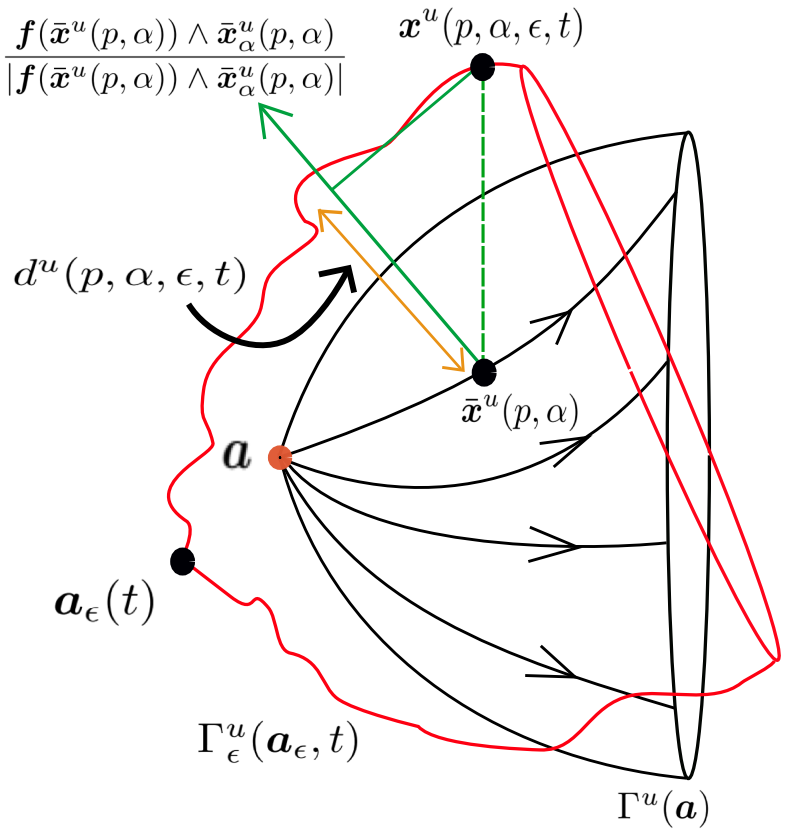


Figure 2: The perturbed unstable manifold $\Gamma^u_\epsilon(\boldsymbol{a}_\epsilon,t)$ (in red) at a general time $t \in (-\infty,T]$, with the unperturbed unstable manifold, $\Gamma^u(\boldsymbol{a})$, shown in black. The distance between perturbed and unperturbed unstable manifold denoted by $d^u(p,\alpha,\epsilon,t)$ is measured perpendicular to the original manifold.

Theorem 1 (Displacement of unstable manifold)**.** *For $(p,\alpha,t) \in (-\infty,P] \times \mathrm{S}^1 \times (-\infty,T]$, the distance $d^u(p,\alpha,\epsilon,t)$ can be expanded in ϵ in the form*

$$d^u(p,\alpha,\epsilon,t) = \epsilon \frac{M^u(p,\alpha,t)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}^u(p,\alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(p,\alpha)|} + \mathcal{O}(\epsilon^2), \tag{3}$$

where the unstable Melnikov function is the convergent improper integral

$$M^u(p,\alpha,t) = \int_{-\infty}^{p} \exp\left[\int_\tau^p \nabla \cdot \boldsymbol{f}(\bar{\boldsymbol{x}}^u(\xi,\alpha))d\xi\right] [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau,\alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau,\alpha)] \cdot \boldsymbol{g}(\bar{\boldsymbol{x}}^u(\tau,\alpha),\tau+t-p)\ \mathrm{d}\tau. \tag{4}$$

Proof. This lengthy proof requires many stages, and is therefore given in Appendix A. Several results which are ingredients in the proof are separated out into additional appendices for clarity. □

For volume-preserving unperturbed flows, the term $\nabla \cdot \boldsymbol{f}(\bar{\boldsymbol{x}}^u(\xi,\alpha))$ is zero, and consequently the integrand of M^u loses the exponential term. Additionally, we note that it is only the normal component of the perturbation $\boldsymbol{g}$, evaluated in appropriate retarded time, that contributes to the leading-order normal displacement which is captured by M^u.

Remark 1 (Approximation of Γ^u_ϵ)**.** Theorem 1 enables a natural approximation for Γ^u_ϵ, with knowledge of the unperturbed flow and the perturbation velocity alone. If $(p,\alpha,t) \in (-\infty,P] \times \mathrm{S}^1 \times (-\infty,T]$ are the parameters for the parametric representation of a general point $\boldsymbol{r}^u$ on Γ^u_ϵ at time t, then

$$\boldsymbol{r}^u(p,\alpha,\epsilon,t) \approx \bar{\boldsymbol{x}}^u(p,\alpha) + \epsilon\, M^u(p,\alpha,t) \frac{\boldsymbol{f}(\bar{\boldsymbol{x}}^u(p,\alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(p,\alpha)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}^u(p,\alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(p,\alpha)|^2} \tag{5}$$

provides a (leading-order in ϵ) parametric representation of Γ^u_ϵ. (While each trajectory on the manifold is known to exhibit a $\mathcal{O}(\epsilon)$ tangential displacement as well, see [6] for a quantification in two-dimensions, in a global view of the manifold as a collection of trajectories, using the normal displacement by itself provides an excellent approximation to the manifold.) Approximation of a perturbed stable or unstable manifold using a Melnikov approach has also been performed by Chow and Yamashita [22] in $\mathbb{R}^n$, but—unlike the expression (5)—has the issue of non-computability outlined in the next remark.

Remark 2 (Kernel of the Melnikov integral)**.** Melnikov functions in general dimensions using functional-analytic or allied approaches [21, 53, 17, 59, 37, 24, 22, 70] for determining persistent heteroclinic intersections usually take the form

$$M \sim \int_{-\infty}^{\infty} \boldsymbol{\Psi}^\top \boldsymbol{g} \, \mathrm{d}\tau \, .$$

Here, the entity $\boldsymbol{\Psi}$ is associated with the fundamental matrix solution of the adjoint of the equation of variations along the heteroclinic trajectory [28, 70, 22, 21, 53, 17, 59, 37, 24], and is usually not computable except in two dimensions, or else if there is a Hamiltonian structure in the unperturbed system [26]. The reason for non-computability in general is that this is a *nonautonomous* linear equation, for which generally solutions cannot be explicitly written down. While the presence of the exponential term is no surprise in instances of non-volume-preservation [28], the structure of the remaining term is usually not known. Hence in such instances the Melnikov approach is a interesting theoretical tool which replaces one issue (finding persistent heteroclinics) with another (finding zeros of a function with a kernel which satisfies a certain property, but which cannot in general be explicitly given by a formula). A similar issue also occurs when using Melnikov approaches for volume-preserving *maps* [38] in which a volume-form related to potentially non-computable adapted vector fields comes into play. Our geometric approach in three dimensions, leading to (4), is the first insight into an explicit form of this kernel function when these conditions are relaxed. (Note however that our limits are not over all of $\mathbb{R}$, because at this stage we are seeking the *location* of the perturbed manifold rather than a persistent heteroclinic connnection.)

2.2 Displacement of 2D stable manifold

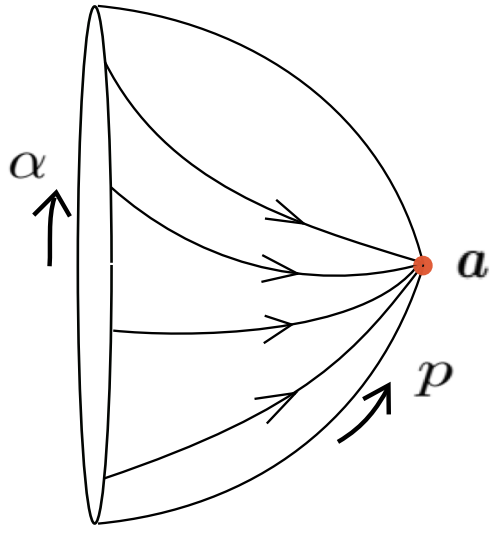


Figure 3: The two-dimensional unperturbed stable manifold $\Gamma^s(\boldsymbol{a})$ for case 2, with the red dot representing the saddle fixed point. The (p, α)-dependence of trajectories $(\bar{\boldsymbol{x}}^s(p, \alpha))$ can be used to parametrize $\Gamma^s(\boldsymbol{a})$.

Secondly, consider case 2, when $\boldsymbol{Df}(\boldsymbol{a})$ has one positive eigenvalue and two eigenvalues with negative real parts at the point $\boldsymbol{a}$. So the unperturbed system posses a one-dimensional unstable manifold and a two-dimensional stable manifold associated with the fixed point $\boldsymbol{a}$. It is once again the displacement of the two-dimensional entity that we capture, in this case, the stable manifold. Rather than repeat the development

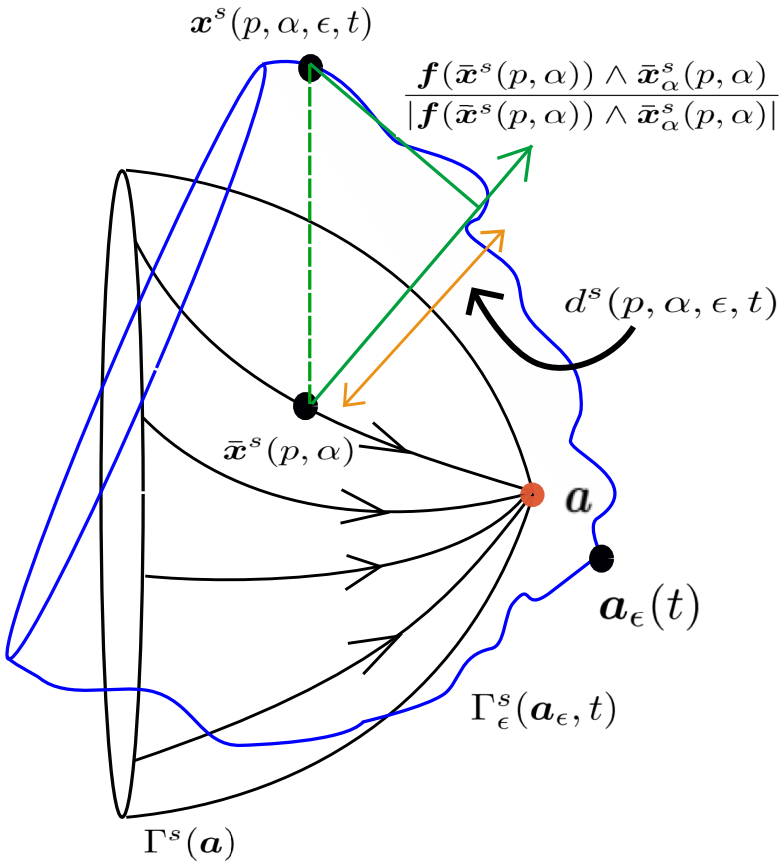


Figure 4: The perturbed stable manifold, which is denoted by $\Gamma^s_\epsilon(\boldsymbol{a}_\epsilon,t)$ (in blue), at a general time $t \in [T,\infty)$, with the unperturbed stable manifold, $\Gamma^s(\boldsymbol{a})$ shown in black. The distance between perturbed and unperturbed stable manifold denoted by $d^s(p,\alpha,\epsilon,t)$ is measured perpendicular to the original manifold.

in detail, we will rely on Figs. 3 and 4 which are exactly analogous to case 1's Figs. 1 and 2. The unperturbed stable manifold $\Gamma^s(\boldsymbol{a})$ is foliated by trajectories $\bar{\boldsymbol{x}}^s(p,\alpha)$ which forward asymptote to $\boldsymbol{a}$ as $p \to \infty$ (Fig. 3). The perturbed stable manifold $\Gamma^s_\epsilon(\boldsymbol{a}_\epsilon,t)$ is attached to the hyperbolic trajectory $(\boldsymbol{a}_\epsilon(t),t)$ (Fig. 4). A time $t \in [T,\infty)$ (where T is finite) is chosen, and then the parameters $p \in [P,\infty)$ (for finite P) and $\alpha \in \mathrm{S}^1$ parametrize $\Gamma^s_\epsilon(\boldsymbol{a}_\epsilon,t)$.

As shown in Fig. 4, the perpendicular distance between perturbed stable manifold ($\Gamma^s_\epsilon(\boldsymbol{a})$) and unperturbed stable manifold ($\Gamma^s(\boldsymbol{a})$) at the location of $\bar{\boldsymbol{x}}^s(p,\alpha)$ in the time instance t is given by

$$d^s(p,\alpha,\epsilon,t) = \frac{\boldsymbol{f}(\bar{\boldsymbol{x}}^s(p,\alpha)) \wedge \bar{\boldsymbol{x}}^s_\alpha(p,\alpha)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}^s(p,\alpha)) \wedge \bar{\boldsymbol{x}}^s_\alpha(p,\alpha)|} \cdot [\boldsymbol{x}^s(p,\alpha,\epsilon,t) - \bar{\boldsymbol{x}}^s(p,\alpha)] \quad , \quad (p,\alpha,t) \in [P,\infty) \times \mathrm{S}^1 \times [T,\infty)\,. \tag{6}$$

Theorem 2 (Displacement of stable manifold)**.** *For $(p,\alpha,t) \in [P,\infty) \times \mathrm{S}^1 \times [T,\infty)$, the distance $d^s(p,\alpha,\epsilon,t)$ can be expanded in ϵ in the form*

$$d^s(p,\alpha,\epsilon,t) = \epsilon \frac{M^s(p,\alpha,t)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}^s(p,\alpha)) \wedge \bar{\boldsymbol{x}}^s_\alpha(p,\alpha)|} + \mathcal{O}(\epsilon^2), \tag{7}$$

where the stable Melnikov function is the convergent improper integral

$$M^s(p,\alpha,t) = -\int_p^\infty \exp\left[\int_\tau^p \nabla \cdot \boldsymbol{f}(\bar{\boldsymbol{x}}^s(\xi,\alpha))d\xi\right] [\boldsymbol{f}(\bar{\boldsymbol{x}}^s(\tau,\alpha)) \wedge \bar{\boldsymbol{x}}^s_\alpha(\tau,\alpha)] \cdot \boldsymbol{g}(\bar{\boldsymbol{x}}^s(\tau,\alpha),\tau+t-p)\ \mathrm{d}\tau\,. \tag{8}$$

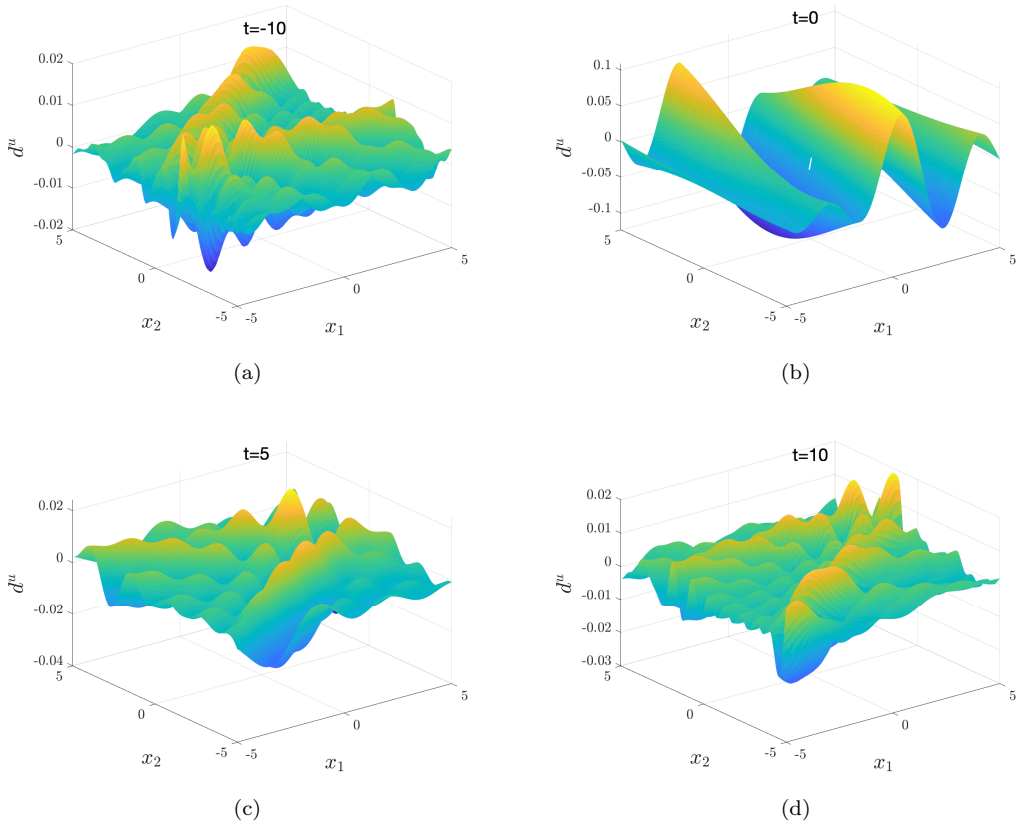


Figure 5: Computed unstable manifold for the system in Sec. 2.3 using the formula (11) for the perturbation (12) with $\epsilon = 0.1$ at several times t.

Proof. The proof is exactly analogous to that for Theorem 1, when we think of simply reversing time. Thus we think of $p \to -p$ (this reverses the time parametrization along an unperturbed trajectory on the 2D manifold), and $t \to -t$. All the ingredients of the proof essentially go through with this understanding. For the sake of brevity, no further details will be given. □

Remark 3 (Approximation of Γ^s_ϵ)**.** If $(p, \alpha, t) \in [P, \infty) \times \mathrm{S}^1 \times [T, \infty)$ are the parameters for the parametric representation of a general point $\boldsymbol{r}^s$ on Γ^s_ϵ at time t, then

$$\boldsymbol{r}^s(p, \alpha, \epsilon, t) \approx \bar{\boldsymbol{x}}^s(p, \alpha) + \epsilon\, M^s(p, \alpha, t) \frac{\boldsymbol{f}(\bar{\boldsymbol{x}}^s(p, \alpha)) \wedge \bar{\boldsymbol{x}}^s_\alpha(p, \alpha)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}^s(p, \alpha)) \wedge \bar{\boldsymbol{x}}^s_\alpha(p, \alpha)|^2} \tag{9}$$

provides a (leading-order in ϵ) parametric representation of Γ^s_ϵ.

2.3 A simple example

We conclude this section with a simple example in which a perturbed unstable manifold is approximated in a non-volume-preserving situation with an aperiodic time-dependence. Consider the unperturbed flow

$$\dot{\boldsymbol{x}} = \begin{bmatrix} 2 & 1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix} \boldsymbol{x}\,, \tag{10}$$

for $\boldsymbol{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$, for which the fixed point at the origin possesses a stable manifold (eigenvalue -3) along the x_3-axis, and an unstable manifold (eigenvalues $2 \pm i$) on the plane $x_3 = 0$. The non-compact

two-dimensional unstable manifold is parametrizable in terms of its spiralling trajectories as

$$\bar{\boldsymbol{x}}^u(p,\alpha) = \begin{bmatrix} e^{2p}\cos(2\pi\alpha - p) \\ e^{2p}\sin(2\pi\alpha - p) \\ 0 \end{bmatrix},$$

where $(p,\alpha) \in \mathbb{R} \times \mathrm{S}^1$. Here, the unperturbed flow violates volume-preservation because $\boldsymbol{\nabla}\cdot\boldsymbol{f} = 1 \neq 0$. If a perturbing vector field $\epsilon\boldsymbol{g}(\boldsymbol{x},t) = \epsilon(g_1,g_2,g_3)$ is added to the right-hand side of (10), then it is possible to use Remark 1 to directly obtain the approximation

$$\boldsymbol{r}^u(p,\alpha,\epsilon,t) \approx \bar{\boldsymbol{x}}^u(p,\alpha) + \epsilon\,\hat{\boldsymbol{e}}_3\,e^{-3p}\int_{-\infty}^{P} e^{3\tau} g_3\left(\bar{\boldsymbol{x}}^u(\tau,\alpha), t+\tau-p\right)\,\mathrm{d}\tau \tag{11}$$

as a parametric representation of the perturbed unstable manifold of the hyperbolic trajectory $\boldsymbol{a}_\epsilon(t)$ which persists $\mathcal{O}(\epsilon)$-close to the origin. Here, $\hat{\boldsymbol{e}}_3$ is the unit vector in the x_3-direction, and the limitations of (11) are that $(p,\alpha,t) \in (-\infty,P]\times\mathrm{S}^1\times(-\infty,T]$ for finite P and T, and $|\epsilon| \ll 1$. As an example, we take

$$g_3(x_1,x_2,x_3,t) = (1+3\,\mathrm{sech}\,t)\sin\left(3x_1 - t^2\right)\cos\left(x_2 t - x_3\right), \tag{12}$$

and note that other components of $\boldsymbol{g}$ are irrelevant in the leading-order approximation (11). We show in Fig. 5 calculations of the perturbed unstable manifold with $\epsilon = 0.1$ using (11) for the vector field (12). While this example distills the essence of the theory for approximating a two-dimensional stable/unstable manifold, a more complicated example will be presented in Sec. 4.

3 Melnikov function for heteroclinic manifolds

Quantifying and understanding transport mechanisms in flows has well-established importance, and in this section we focus on a contribution to the three-dimensional theory. While there are fewer contributions than in two-dimensions, there are still several transport quantifications in three-dimensions [33, 13, 24, 47, 46, 50, 51]. These usually rely on additional conditions such as a specific type of time-periodicity [33, 51], symmetry [33], volume-preservation [33, 13, 47, 46, 50, 51], the presence of a conserved quantity [13, 47, 46, 51], or a single-step transport over a finite time [50].

Instances where an unperturbed system has a codimension-1 *heteroclinic* manifold are interesting in that the manifold forms a flow separator in both forwards and backwards time. Nearby initial conditions on opposite sides of the manifold separate exponentially in time, allowing the manifold to be thought of as a flow barrier [7]. Perturbations have the potential for generating transport across this flow barrier. This setting of a heteroclinic manifold is also the more 'classical' setup of Melnikov theory. However, rather than simply seeking a Melnikov function whose zeros imply persistence of a heteroclinic connection (which we will obtain), we do more in this Section:

- Find ways of characterizing *lobes* bounded by perturbed stable and unstable manifolds at each instance in time (under *general* time-dependent perturbations);
- Obtain a formula in terms of the Melnikov function for the volume of such a lobe;
- In this instance when the perturbation g is *harmonic*, that is when it can be written in the form $\tilde{\boldsymbol{g}}(x)\cos[\omega t + \phi]$, extend the two-dimensional theory of lobe dynamics via a turnstile [57, 66] to three-dimensions;
- In the more general instance in which g has general time-dependence, extend the two-dimensional theory for instantaneous flux across the broken heteroclinic (formerly a flow barrier) to three dimensions;
- In the above instance, express the formula in simple terms in terms of the Melnikov function.

The last two of these issues is particularly important within the context of fluid flows: these are inevitably three-dimensional, and flow barriers are consequently two-dimensional. By rationalizing an instantaneous flux across a flow barrier, we are quantifying an easily computable time-varying transport due to a perturbation

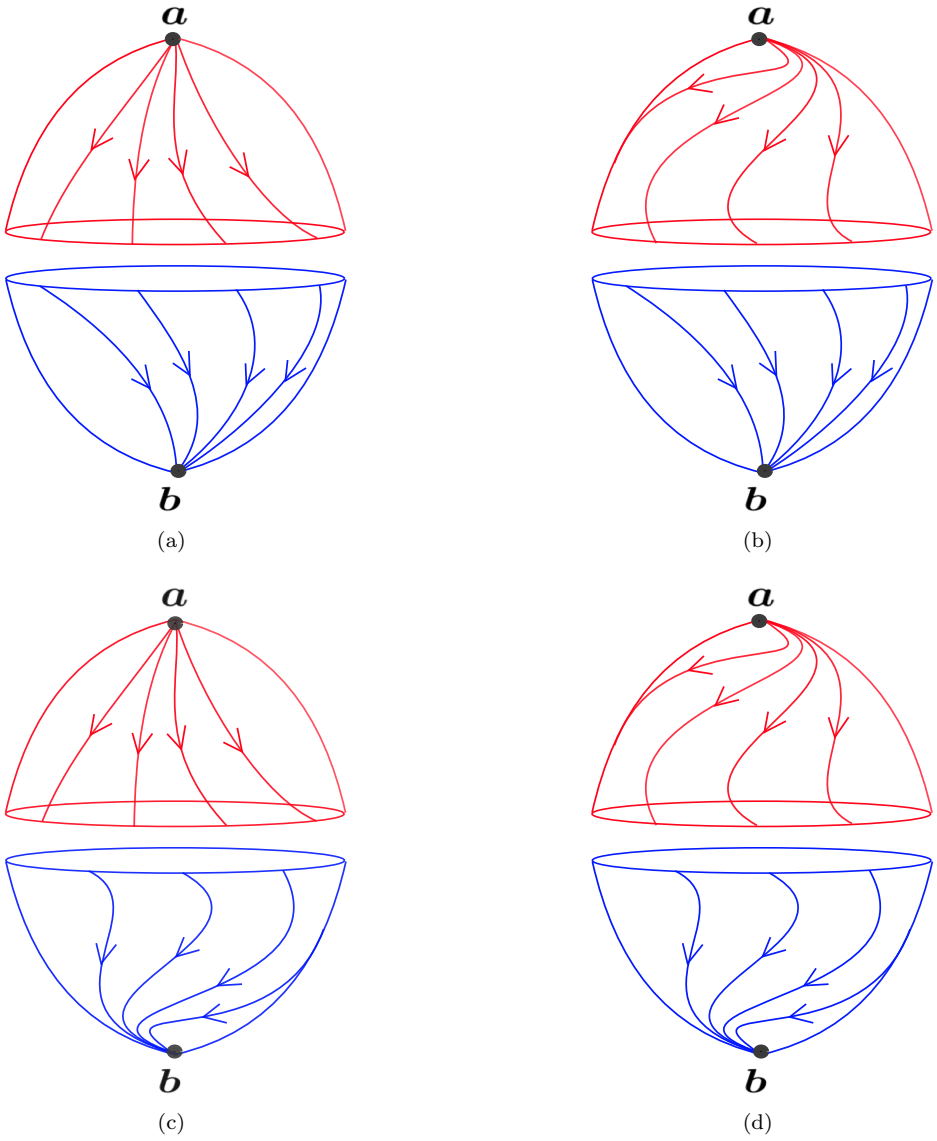


Figure 6: Several trajectories $\bar{\boldsymbol{x}}(p, \alpha)$ on two-dimensional unperturbed stable and unstable manifolds are given for four different situations. The case when no complex eigenvalues are present is shown in (a). The cases in which the complex conjugate eigenvalues are only on the unstable manifold (respectively only on the stable manifold), are given in (b) and (c). Finally, if the eigevalues at both $\boldsymbol{a}$ and $\boldsymbol{b}$ are complex-conjugate, we get the behavior shown in (d).

which has general (aperiodic) time-dependence. This is the principal result of this Section, and is given in Theorem 5.

We assume that the smoothness hypotheses as stated in Section 2 continue to hold. However, assume now that unperturbed situation of the system (1) when $\epsilon = 0$ has two distinct fixed points $\boldsymbol{a}$ and $\boldsymbol{b}$, such that $\boldsymbol{a}$ possesses a two-dimensional unstable manifold Γ^u, and $\boldsymbol{b}$ a two-dimensional stable manifold Γ^s. Moreover, we assume that these two manifolds coincide, to form a two-dimensional orientable heteroclinic manifold Γ as shown in Fig. 6. Thus, a trajectory $\bar{\boldsymbol{x}}^u(p,\alpha)$ on $\boldsymbol{a}$'s unstable manifold coincides with a trajectory $\bar{\boldsymbol{x}}^s(p,\alpha)$ on $\boldsymbol{b}$'s stable manifold. We denote this heteroclinic trajectory by $\bar{\boldsymbol{x}}(p,\alpha)$, and insist that $(p,\alpha) \in \mathbb{R} \times \mathrm{S}^1$ be chosen such that (p,α) provides a C^2-smooth parameterization of the heteroclinic manifold. Once again, p can be thought of as the time-evolution along the trajectory, with α choosing *which* trajectory. Thus we have the limiting behavior

$$\lim_{p\to-\infty} \bar{\boldsymbol{x}}(p,\alpha) = \boldsymbol{a} \quad \text{and} \quad \lim_{p\to\infty} \bar{\boldsymbol{x}}(p,\alpha) = \boldsymbol{b}$$

for each and every α. We have shown the spiralling situation in Fig. 6(d), corresponding to $\lambda_1^u(\boldsymbol{a})$ and $\lambda_2^u(\boldsymbol{a})$ (eigenvalues with the positive real parts obtained at $\boldsymbol{a}$) being complex conjugates of one another, as are $\lambda_1^s(\boldsymbol{b})$ and $\lambda_2^s(\boldsymbol{b})$ (eigenvalues with the negative real parts obtained at $\boldsymbol{b}$). However, the heteroclinic trajectories need not be spiralling in general; that is, these sets of eigenvalues could be purely positive (for $\boldsymbol{a}$) and purely negative (for $\boldsymbol{b}$). This situation is shown in Fig. 6(a). Indeed, it is allowable to have complex eigenvalues only at one of the endpoints $\boldsymbol{a}$ or $\boldsymbol{b}$ and at the other endpoint, we can have purely positive or negative eigenvalues appropriately. These are given in Figs. 6(b) and (c).

The unperturbed heterclinic manifold Γ is a closed surface, separating phase space into a part which is inside Γ, and another which is outside. To distinguish between the two, we will assume that the α-labelling in the (p,α) parametrization of Γ is chosen such that

$$\frac{\partial \bar{\boldsymbol{x}}(p,\alpha)}{\partial p} \wedge \frac{\partial \bar{\boldsymbol{x}}(p,\alpha)}{\partial \alpha} = \boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)$$

(which we know is normal to Γ) points in the *outwards* direction to Γ at each point $\boldsymbol{x}(p,\alpha)$ on Γ. Thus, the unit outwards-pointing normal vector to Γ at a point parametrized by (p,α) is given by

$$\hat{\boldsymbol{n}}(p,\alpha) := \frac{\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)|}\,. \tag{13}$$

Now, when $\epsilon \neq 0$ but is small in size, we know that $\boldsymbol{a}$ becomes a heteroclinic trajectory, and its unstable manifold persists as a time-parameterized entity $\Gamma_\epsilon^u(\boldsymbol{a}_\epsilon, t)$. This can be parameterized by $(p,\alpha,t) \in (-\infty, P^u] \times \mathrm{S}^1 \times (-\infty, T^u]$ for any finite P^u and T^u. We note that we can take P^u as large as we like (but remaining finite), indicating that we cannot find an approximation to the manifold globally. Intuitively, this means that a perturbed version of the top-hemisphere of Fig. 6 persists; we may approach the south pole, but begin to lose control as we do so. The manifold may continue beyond this region, and if so, our theory is not able to approximate it. Similarly, the stable manifold of $\boldsymbol{b}$ persists as $\Gamma_\epsilon^s(\boldsymbol{b}_\epsilon, t)$, and we can parameterize this by $(p,\alpha,t) \in [P^s, \infty) \times \mathrm{S}^1 \times [T^s, \infty)$ for any finite P^s and T^s, which we can choose to be as negative as we like.

Take a time $t \in [T^s, T^u]$. The unstable manifold of $\boldsymbol{a}_\epsilon$ for time in $(-\infty, t]$ and the stable manifold of $\boldsymbol{b}_\epsilon$ for time in $[t, \infty)$ can both be approximated using our previous results. The manifolds at time t no longer need to coincide, and we show a generic situation in Fig. 7. There is now a perturbed version of $\bar{\boldsymbol{x}}(p,\alpha)$ on $\Gamma_\epsilon^u(\boldsymbol{a}_\epsilon, t)$, which we can call $\boldsymbol{x}^u(p,\alpha,\epsilon,t)$, which exists for $p \in (-\infty, P^u]$. Similarly, there exists a perturbed version of $\bar{\boldsymbol{x}}(p,\alpha)$ on $\Gamma_\epsilon^s(\boldsymbol{b}_\epsilon, t)$, which we call $\boldsymbol{x}^s(p,\alpha,\epsilon,t)$, which exists for $p \in [P^s, \infty)$.

Given a location (p,α) on Γ, we first want to quantify the displacement of between the perturbed unstable and stable manifolds, in the direction normal to Γ:

$$d(p,\alpha,\epsilon,t) = \frac{\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)|} \cdot [\boldsymbol{x}^u(p,\alpha,\epsilon,t) - \boldsymbol{x}^s(p,\alpha,\epsilon,t)] \quad , \quad (p,\alpha,t) \in [P^s, P^u] \times \mathrm{S}^1 \times [T^s, T^u]\,. \tag{14}$$

Given our choice of labelling of the unperturbed heteroclinic trajectories on Γ, we note that a positive $d(p,\alpha,\epsilon,t)$ implies that the unstable manifold is *outside* the stable manifold, while a negative d means that the stable manifold is outside the unstable one at a location $\bar{\boldsymbol{x}}(p,\alpha)$ at a time instance t.

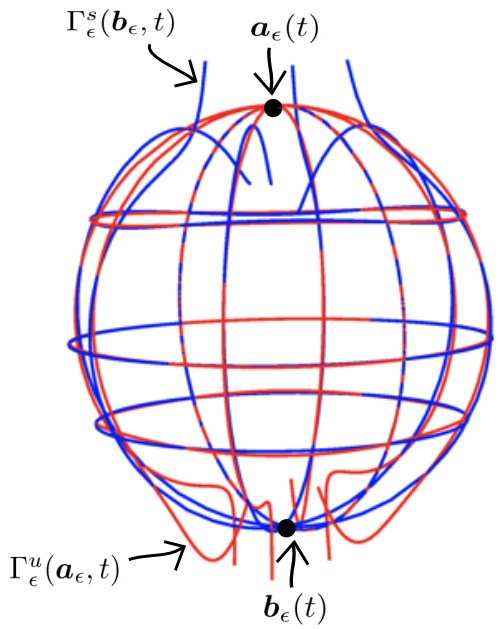


Figure 7: A generic intersection pattern of the perturbed unstable ($\Gamma^u_\epsilon(\boldsymbol{a}_\epsilon, t)$, red) and stable ($\Gamma^s_\epsilon(\boldsymbol{b}_\epsilon, t)$, blue) manifolds of the Hill's spherical vortex with no swirl, which is given in Eq. (36), at time $t = 1$ and $\epsilon = 0.1$. These perturbed unstable and stable manifolds were computed utilizing Eqs. (37) and (38) .

Theorem 3 (Heteroclinic splitting)**.** *For $(p, \alpha, t) \in [P^s, P^u] \times \mathrm{S}^1 \times [T^s, T^u]$, the distance (14) can be expanded in ϵ in the form*

$$d(p, \alpha, \epsilon, t) = \epsilon \frac{M(p, \alpha, t)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}(p, \alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(p, \alpha)|} + \mathcal{O}(\epsilon^2), \tag{15}$$

where the Melnikov function is the convergent improper integral

$$M(p, \alpha, t) = \int_{-\infty}^{\infty} \exp\left[\int_\tau^p \nabla \cdot \boldsymbol{f}(\bar{\boldsymbol{x}}(\xi, \alpha)) d\xi\right] [\boldsymbol{f}(\bar{\boldsymbol{x}}(\tau, \alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(\tau, \alpha)] \cdot \boldsymbol{g}(\bar{\boldsymbol{x}}(\tau, \alpha), \tau + t - p)\ \mathrm{d}\tau\,. \tag{16}$$

Proof. See Appendix E. □

Note that the distance function d in (15) at fixed t can be thought of in the sense of first taking a point on Γ parametrized by (p, α), i.e., the point $\bar{\boldsymbol{x}}(p, \alpha)$, drawing an outward-pointing normal vector to Γ at that point, and determining the signed distance along that normal vector. Thus, we can think of *projecting* this distance information between perturbed stable and unstable manifolds on to Γ.

Remark 4 (Shifts of heteroclinic points)**.** Each heteroclinic intersection point simply 'shift along' the p-location as time t is varied in forward and backward time, because of the limiting behavior expressed by (18). Indeed, (16) shows that

$$M(\tilde{p} + t, \tilde{\alpha}, \tilde{t} + t) = \exp\left[\int_{\tilde{p}}^{\tilde{p}+t} \nabla \cdot \boldsymbol{f}(\bar{\boldsymbol{x}}(\xi, \alpha)) d\xi\right] M(\tilde{p}, \tilde{\alpha}, \tilde{t})$$

for any time-shift t, and hence if there is a zero at a time $\tilde{t}$ at a location encoded by the parameter $\tilde{p}$, then there is correspondingly a zero at the shifted time $\tilde{t} + t$ at a parameter value $\tilde{p} + t$, at the same $\tilde{\alpha}$ value (i.e., traversing along the same unperturbed heteroclinic trajectory). Effectively, tracking the location of this point with time gives the location subtended on Γ of the corresponding perturbed heteroclinic trajectory as a function of time t. This fact is illustrated most strikingly for the volume-preserving case, since M will depend on (p, t) not independently, but in terms of the shift $(p - t)$.

Remark 5 (Transverse intersection points)**.** At a fixed value $\tilde{t}$ of time, a transverse intersection between $\Gamma^u_\epsilon(\boldsymbol{a}_\epsilon)$ and $\Gamma^s_\epsilon(\boldsymbol{b}_\epsilon)$ near a point $\bar{\boldsymbol{x}}(p,\alpha)$ is guaranteed if $M(p,\alpha,t)$ has a simple zero with respect to (p,α); that is, if there exists $(\tilde{p},\tilde{\alpha})$ such that $M(\tilde{p},\tilde{\alpha},\tilde{t})=0$ and

$$\left|\frac{\partial M}{\partial p}(\tilde{p},\tilde{\alpha},\tilde{t})\right| + \left|\frac{\partial M}{\partial \alpha}(\tilde{p},\tilde{\alpha},\tilde{t})\right| \neq 0\,. \tag{17}$$

This is a standard consequence of implicit function theorem type arguments in this setting (see [1, 30]). Each such intersection point $\boldsymbol{x}(\tilde{p},\tilde{\alpha},\tilde{t})$ corresponds to a heteroclinic trajectory, i.e., the trajectory $\boldsymbol{x}(\tilde{p},\tilde{\alpha},t)$ passing through $\boldsymbol{x}(\tilde{p},\tilde{\alpha},\tilde{t})$ at time $\tilde{t}$ satisfies

$$\lim_{t\to-\infty}|\boldsymbol{x}(\tilde{p},\tilde{\alpha},t)-\boldsymbol{a}_\epsilon(t)|=0 \quad\text{and}\quad \lim_{t\to\infty}|\boldsymbol{x}(\tilde{p},\tilde{\alpha},t)-\boldsymbol{b}_\epsilon(t)|=0\,. \tag{18}$$

While this is a generic result in standard Melnikov developments, having an explicit form within the integral in (16) in a dimension larger than 2 without a Hamiltonian structure is new.

Remark 6 (Curves of heteroclinic points)**.** Suppose there exists continuously differentiable functions $\tilde{\alpha}(s)$ and $\tilde{p}(s)$ for $s\in(0,1)$ such that at a fixed t

$$M\left(\tilde{p}(s),\tilde{\alpha}(s),t\right)=0 \quad\text{and}\quad \boldsymbol{\nabla}_{p,\alpha}M\left(\tilde{p}(s),\tilde{\alpha}(s)\right)\neq\boldsymbol{0} \quad \text{for all } s\in(0,1) \quad ,$$

where $\boldsymbol{\nabla}_{p,\alpha}$ is the two-dimensional gradient with respect to (p,α). Assume moreover that there exists a constant H such that

$$\sup_{s\in(0,1)}|\boldsymbol{\nabla}_{p,\alpha}M\left(\tilde{p}(s),\tilde{\alpha}(s)\right)|<H\,.$$

The parametrization $(\tilde{p}(s),\tilde{\alpha}(s))$ represents a non-degenerate *curve* of points, Q, parametrized by $s\in(0,1)$ and at the locations $\bar{\boldsymbol{x}}\left(\tilde{p}(s),\tilde{\alpha}(s)\right)$, along which M is zero. (This is the generic expectation for the intersection between the two two-dimensional perturbed manifolds.) Applying the Banach space version of the implicit function theorem to $d(p,\alpha,\epsilon,t)/\epsilon$ then gives the persistence of a C^1-curve $Q^\star$ for small enough $|\epsilon|$, which is associated with the parametrization $(p^\star(s),\alpha^\star(s))$ which is $\mathcal{O}(\epsilon)$-close to $(\tilde{p}(s),\tilde{\alpha}(s))$. Note that the same implication arises if $s\in\mathrm{S}^1$, i.e., we think of Q as a *closed* curve. In other words, these conditions imply the presence of a curve of heteroclinic points $Q^\star$ which is $\mathcal{O}(\epsilon)$-close to that predicted via the Melnikov function. Thus, each point on $Q^\star$ obeys Remark 5, and is an 'initial condition' for a heteroclinic trajectory of the time-varying problem which satisfies the conditions (18).

A possible intersection between a perturbed stable manifold $\Gamma^s_\epsilon\left(\boldsymbol{b}_\epsilon,t\right)$ and perturbed unstable manifold $\Gamma^u_\epsilon\left(\boldsymbol{a}_\epsilon,t\right)$ at a fixed time t is shown in Fig. 7. Here, we have used red and blue colors to represent the perturbed unstable and stable manifolds respectively, a convention we will follow in the remainder of this paper. We recall that in two-dimensional flows in which a one-dimensional heteroclinic splitting is captured via a Melnikov function, the 'lobe area' between two adjacent intersections of the split manifolds can be obtained by integrating the Melnikov function [57, 66]. An analogous result is available in the present three-dimensional setting. If the Melnikov function M has a closed curve Q on Γ along which M has non-degenerate zeros as explained in Remark 6, then there is a $\mathcal{O}(\epsilon)$-close closed curve $Q^\star$ on Γ which corresponds to the normal projection on to Γ of the intersection ring between perturbed stable and unstable manifolds. If M is sign definite inside Q, the interior of $Q^\star$ is the 'shadow' of the one lobe which is generated by this intersection, i.e., the projection of the lobe on to Γ along normal vectors to Γ. Then, we can give an expression for the lobe volume in terms of the Melnikov function:

Theorem 4 (Lobe volume)**.** *Let $t\in[T^s,T^u]$ be fixed, and suppose that there is an open region R on Γ in which $M(p,\alpha,t)$ is sign definite, and such that $M=0$ and $\boldsymbol{\nabla}_{p,\alpha}M\neq\boldsymbol{0}$ at all points on its boundary Q. Note that Q will generically consist of a finite number of closed curves. Let R' be the region in (p,α)-space corresponding to R. The region between the perturbed stable and manifolds which is subtended by R is a* lobe*, whose volume is given by*

$$\text{Lobe volume}=\epsilon\iint_{R'}|M(p,\alpha,t)|\,\mathrm{d}p\,\mathrm{d}\alpha+\mathcal{O}(\epsilon^2)\,. \tag{19}$$

Proof. See Appendix F. □

Remark 7. The fact that a lobe volume can be represented to leading-order by an integral of the Melnikov function over the p-variable (representing the time-parametrization along a heteroclinic trajectory) is exactly analogous to a well-known similar result for lobe areas in the two-dimensional situation [57, 66]. We point out that it is *not* necessary to impose additional hypotheses such as volume-preservation or time-periodicity for this result to be true; one just needs to know the (p, α)-region associated with the particular lobe of interest.

In the unperturbed situation, the stable and unstable manifold coincide to form a heteroclinic manifold, which separates the 'inside' and the 'outside' flow regimes. However, after perturbation this entity splits into a stable and unstable manifold, and consequently is no longer a flow separator. Hence, transport will now occur across the previously impermeable structure. Understanding the lobe volume would seem to be relevant in quantifying this transport. It turns out, though, that a lobe volume can only be unambiguously assigned as a measure of transport under several additional assumptions, including a restrictive type of time-periodicity of the perturbation $\boldsymbol{g}$, and volume preservation of the unperturbed flow. We show in Section 3.3 how the two-dimensional theory of lobe dynamics [57, 66] can be extended to our situation, to allow for a lobe volume to quantify transport. In more general situations, a lobe volume cannot be used to quantify transport, because it turns out that there may be either many different-sized lobes, or no lobes at all (these are possibilities even under time-periodic perturbations [2]). In this case, an alternative approach, which characterizes an instantaneous flux as a time-varying quantity, is necessary. This more general approach is described in Section 3.1, which builds on a similar approach in two dimensions [4, 10, 7].

3.1 Instantaneous flux for general time-dependence

Here, we develop a genuinely time-dependent approach for quantifying transport. We do *not* require conditions such as time-periodicity [33, 51], symmetry [33], volume-preservation [33, 13, 47, 46, 50, 51], or the presence of a conserved quantity [13, 47, 46, 51] present in other three-dimensional transport quantifications. The theory and quantification is also independent of the concept of lobes [42, 57, 66, 43]. Our approach is analogous to the two-dimensional development of instantaneous flux [4, 10]. We therefore continue to allow $\boldsymbol{g}(\boldsymbol{x}, t)$ to have any *general* temporal behavior (but subject to the smoothness/boundedness assumptions detailed in Section 2), and both $\boldsymbol{f}$ and $\boldsymbol{g}$ to not be divergence-free. We will in Theorem 5 detail our main result, which quantifies the transport engendered under *general* time-variation elegantly in terms of the Melnikov function.

An 'obvious' way of quantifying a time-dependent flux across the previously impermeable Γ might be the Eulerian approach

$$\begin{aligned}\text{Eulerian flux } (t,\epsilon) &= \iint_\Gamma \left[\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha) + \epsilon \boldsymbol{g}\left(\bar{\boldsymbol{x}}(p,\alpha),t\right)\right] \cdot \hat{\boldsymbol{n}}(p,\alpha)\,\mathrm{d}S \\ &= \epsilon \int_0^1 \int_{-\infty}^{\infty} \boldsymbol{g}\left(\bar{\boldsymbol{x}}(p,\alpha),t\right) \cdot \left[\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)\right]\,\mathrm{d}p\,\mathrm{d}\alpha\,. \end{aligned} \tag{20}$$

The term 'Eulerian' is used here in the fluid-mechanics context, in relation to the flux across a fixed surface. The simplification above occurs because $\boldsymbol{f}$ is tangential to Γ, and thus does not contribute.

However, the Eulerian approach does not take into account *transport*. The issue is that the manifolds themselves move, as do trajectories on and adjacent to them. Therefore a velocity flux across Γ does not capture the Lagrangian (following the flow of particle trajectories) transport. As a simple example, imagine that the perturbation is such that the manifold Γ retains its structure as continuing to be heteroclinic, but simply 'puffs out' to be slightly larger than the original Γ. Then, there should be no predicted transport, because the perturbed manifold structure persists. On the other hand, there could be a nonzero Eulerian flux because $\boldsymbol{g}$ need not be zero on Γ, nor perpendicular to Γ. It need only satisfy those conditions on the perturbed version of Γ. Computing the flux across the fixed unperturbed Γ without taking into account the fact that the invariant manifolds have now perturbed, is incorrect. In other words, in a transport computation the fact that the invariant manifold is itself moving must be taken into account, and not simply the time-variation of the vector field.

A *Lagrangian* approach—which takes into account the displacement of the previously impermeable heteroclinic manifold—is thus necessary. On the unperturbed Γ, consider a fixed-p curve, that is $\bigcup_{\alpha\in S^1} \bar{\boldsymbol{x}}(p,\alpha)$.

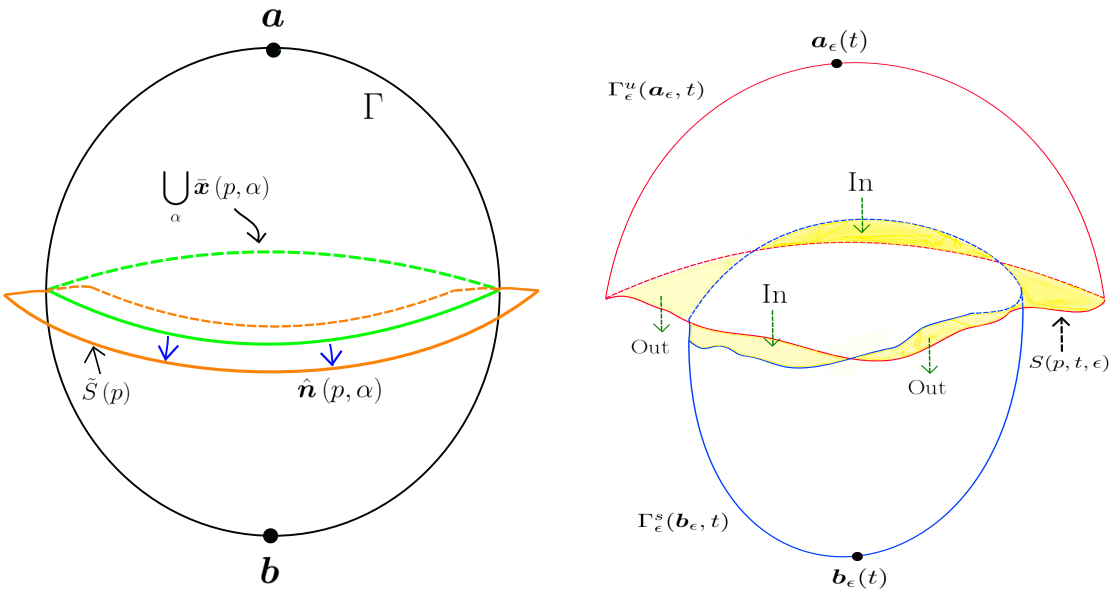


Figure 8: A schematic of the construction of (a) the strip, and (b) the pseudo-separatrix $\tilde{\Gamma}(p,t,\epsilon)$ at time t, associated with making a choice of p, and consisting of the unstable manifold (red) and stable manfold (blue), both extending until intersecting the strip $S(p,\epsilon,t)$, as well as the strip $S(p,\epsilon,t)$.

This goes all the way around Γ, and at each point on it has an outward-pointing normal vector $\hat{\boldsymbol{n}}(p,\alpha)$ as given in (13). We form the unperturbed strip $\tilde{S}(p)$ by taking the union of these normal vectors across $\alpha \in \mathrm{S}^1$, and extending outwards in both the positive and negative directions of $\hat{\boldsymbol{n}}(p,\alpha)$. At this instance, we will not specify how far the extension needs to be done; this will be clear shortly. Thus, $\tilde{S}(p)$ is an oriented strip going all the way around Γ; a portion of this is displayed in Fig. 8(a).

Now consider the perturbed stable and unstable manifolds at a general time t. Using the results from Section 2, we know that the distance to the perturbed unstable manifold along the normal vector direction is given by

$$d^u(p,\alpha,\epsilon,t) = \hat{\boldsymbol{n}}(p,\alpha)\cdot[\boldsymbol{x}^u(p,\alpha,\epsilon,t) - \bar{\boldsymbol{x}}(p,\alpha)]$$

and to the perturbed stable manifold

$$d^s(p,\alpha,\epsilon,t) = \hat{\boldsymbol{n}}(p,\alpha)\cdot[\boldsymbol{x}^s(p,\alpha,\epsilon,t) - \bar{\boldsymbol{x}}(p,\alpha)]$$

in terms of trajectories $\boldsymbol{x}^u$ and $\boldsymbol{x}^s$ respectively on the perturbed manifolds. We now define the strip $S(p,\epsilon,t) \subset \tilde{S}(p)$ by

$$S(p,\epsilon,t) := \bigcup_{\alpha\in\mathrm{S}^1}\bigcup_{s\in[0,1]} \{\bar{\boldsymbol{x}}(p,\alpha) + \hat{\boldsymbol{n}}(p,\alpha)\,[sd^u(p,\alpha,\epsilon,t) + (1-s)d^s(p,\alpha,\epsilon,t)]\}\ . \tag{21}$$

Thus, at each fixed time t and parameter p, the strip $S(p,\epsilon,t)$ is a $\mathcal{O}(\epsilon)$-width 'ribbon' which is attached to the closed curve $\bigcup_{\alpha\in\mathrm{S}^1}\bar{\boldsymbol{x}}(p,\alpha)$ on Γ. This is illustrated by the shaded segment in Fig. 8(b). At each location α, the strip traverses the normal direction to Γ spanning from the stable to the unstable manifold. When the stable and unstable manifolds intersect, then the 'ribbon' has zero width at that point. Therefore, the unperturbed strip $\tilde{S}(p)$ needs to extend out to ensure that $S(p,\epsilon,t) \subset \tilde{S}(p)$ for all times t and perturbative parameters ϵ for which we want to characterize the flux.

We want to define an instantaneous flux across the broken heteroclinic manifold, taking into account the Lagrangian nature of the problem. To do so, we now define the pseudo-separatrix $\tilde{\Gamma}(p,t,\epsilon)$ as the union of the following two-dimensional surfaces:

- The part of the unstable manifold $\Gamma^u_\epsilon(\boldsymbol{a}_\epsilon,t)$ emanating from $\boldsymbol{a}_\epsilon(t)$, until it first hits the strip $S(p,\epsilon,t)$;
- The strip $S(p,\epsilon,t)$; and

- The part of the stable manifold $\Gamma^s_\epsilon(\boldsymbol{b}_\epsilon, t)$ emanating from $\boldsymbol{b}_\epsilon(t)$, until it first hits the strip $S(p, \epsilon, t)$.

These entities are shown in Fig. 8(b) for a typical situation. We note that $\tilde{\Gamma}(p, \epsilon, t)$ is $\mathcal{O}(\epsilon)$-close the the unperturbed Γ, which was a genuine flow separator between the its 'inside' and 'outside.' The pseudo-separatrix $\tilde{\Gamma}$ is itself a closed surface for any chosen p, and at any time t. It is one way of attempting to define a semi-separator between the inside and the outside of the perturbed version of Γ, which is in reality the combination of the perturbed stable and unstable manifolds which typically will intersect with each other. Note that the definition takes into account the perturbed versions of both the stable and unstable manifolds, and is hence Lagrangian. We can determine the instantaneous signed flux (volume per unit time) exiting $\tilde{\Gamma}$ at time t, subject to the choice of the parameter p. Key to this quantification is the observation that the stable/unstable manifold parts of $\tilde{\Gamma}(p, \epsilon, t)$ are moving as invariant objects, and hence can have no flux crossing them. Transport across $\tilde{\Gamma}(p, \epsilon, t)$ occurs because of the flux across the strip where the stable and unstable manifolds connect. The instantaneous flux Φ can be quantified elegantly in terms of the Melnikov function:

Theorem 5 (Instantaneous flux)**.** *The instantaneous (signed) flux exiting the pseudo-separatrix $\tilde{\Gamma}(p, t, \epsilon)$ is*

$$\Phi(p, t, \epsilon) = \epsilon \int_0^1 M(p, \alpha, t)\, \mathrm{d}\alpha + \mathcal{O}(\epsilon^2)\,. \tag{22}$$

Proof. See Appendix G. □

We emphasize that this result is for the *general* form of the Melnikov function, without having to make assumptions on time-periodicity or volume-preservation. It is even true if the manifolds do not intersect at all, and hence is independent of the concept of lobes. Theorem 5 is therefore a very general result, for any time-variation of the perturbation (subject to Hypothesis 1), which states that the leading-order instantaneous flux is given by integrating the Melnikov function over all α (representing all the unperturbed heteroclinic trajectories). The instantaneous flux Φ is of course time-dependent; this variation is captured by the t in the Melnikov function. Moreover, Φ is also dependent on the choice of p, the location along which the the perturbed stable and unstable manifolds are joined.

The flux is explicitly the volume of phase space which crosses $\tilde{\Gamma}$ per unit time. This is best rationalized in the fluid mechanical context in which $\boldsymbol{f} + \epsilon\boldsymbol{g}$ is a fluid velocity. Then, the flux represents exactly the volume of fluid per unit time exiting $\tilde{\Gamma}$.

It should also be pointed out that the flux Φ is *signed*. In parts where the unstable manifold is outside the stable manifold on the strip S, there will be flux *exiting* the closed surface $\tilde{\Gamma}$ (see Fig. 8(b)). These will be encoded as positive, reflecting also the fact that a positive Melnikov function implies that $d^u - d^s$, when projected on to the outwards-pointing normal vector direction $\hat{\boldsymbol{n}}(p, \alpha)$, is positive. Similarly, parts where the unstable manifold is inside the stable manifold are associated with flux flowing into $\tilde{\Gamma}$, and thus constitutes negative flux. The expression Φ in Theorem 5 sums all these flux contributions to obtain a signed net flux. If $\Phi > 0$, that means that there is more volume of fluid instantaneously exiting $\tilde{\Gamma}$ than there is entering it.

Remark 8. The instantaneous flux is dependent on the choice of p; that is, the ring of locations on the heteroclinic trajectories at which the strip S is drawn. In view of Remark 4, though, it is clear that if the unperturbed flow is volume-preserving, then the flux Φ's p-dependence is equivalently a shift in t.

While the flux expression in Theorem 5 is very general, it will be useful to view its form under commonly used restrictions—notably a special form of periodicity in $\boldsymbol{g}$—to gain additional insight into the flux. We do this in the next section.

3.2 Flux and geometry under additional conditions

We first obtain simplifications for the Melnikov function and the instantaneous flux under a frequently used condition [57, 66, 33]. We call this *time-harmonicity* or simply *harmonicity*:

Definition 1. *The perturbing function $\boldsymbol{g}$ shall be called* harmonic *if it is representible in the form*

$$\boldsymbol{g}(\boldsymbol{x}, t)] = \tilde{g}\,(\boldsymbol{x}) \cos\left(\omega t + \phi\right)\,, \tag{23}$$

where $\tilde{\boldsymbol{g}}$ and $D\tilde{\boldsymbol{g}}$ are bounded (as per hypothesis), and the frequency $\omega \neq 0$ and the phase shift ϕ are constant.

Remark 9. Harmonicity amounts to the perturbation's time-dependence being separable from the spatial dependence, as well as the time-dependence being sinusoidal. The period of the perturbation is then $T = 2\pi/\omega$. We remark that this is a much stronger restriction than "periodicity in time," which would merely require $\boldsymbol{g}(\boldsymbol{x}, t+T) = \boldsymbol{g}(\boldsymbol{x}, t)$ for all relevant $(\boldsymbol{x}, t)$. (In this periodic case, the results of this section can be extended by following an approach developed for two dimensions [2], which we will avoid because we have a theory for general time-dependence.)

For harmonic $\boldsymbol{g}$, we first show that the Melnikov function takes a remarkably simple form—itself harmonic. To express this, we make the following choice for the definition of the Fourier transform:

$$\mathcal{F}\left\{H(\boldsymbol{.})\right\}(\omega) := \int_{-\infty}^{\infty} e^{-i\omega\tau} H(\tau)\,\mathrm{d}\tau\,,$$

for functions of time in $\mathrm{L}^1(\mathbb{R})$.

Theorem 6 (Melnikov function for harmonic perturbations)**.** *If $\boldsymbol{g}$ is harmonic as given in (23), the Melnikov function is itself harmonic, and expressible as*

$$M(p,\alpha,t) = \left|\mathcal{F}\left\{h(p,\alpha,\boldsymbol{.})\right\}(\omega)\right| \cos\left[\omega\,(t-p) + \phi - \arg\left(\mathcal{F}\left\{h(p,\alpha,\boldsymbol{.})\right\}(\omega)\right)\right] \tag{24}$$

where

$$h(p,\alpha,\tau) := \exp\left[\int_{\tau}^{p} \nabla\cdot\boldsymbol{f}(\bar{\boldsymbol{x}}(\xi,\alpha))d\xi\right]\left[\boldsymbol{f}(\bar{\boldsymbol{x}}(\tau,\alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(\tau,\alpha)\right]\cdot\tilde{\boldsymbol{g}}(\bar{\boldsymbol{x}}(\tau,\alpha))\,. \tag{25}$$

Proof. Under the harmonic assumption, the Melnikov function (16) becomes

$$\begin{aligned} M(p,\alpha,t) &= \int_{-\infty}^{\infty} \exp\left[\int_{\tau}^{p} \nabla\cdot\boldsymbol{f}(\bar{\boldsymbol{x}}(\xi,\alpha))d\xi\right]\left[\boldsymbol{f}(\bar{\boldsymbol{x}}(\tau,\alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(\tau,\alpha)\right]\cdot\tilde{\boldsymbol{g}}(\bar{\boldsymbol{x}}(\tau,\alpha))\cos\left[\omega\,(\tau+t-p)+\phi\right]\,\mathrm{d}\tau \\ &= \int_{-\infty}^{\infty} h(p,\alpha,\tau)\cos\left[\omega\,(\tau+t-p)+\phi\right]\,\mathrm{d}\tau\,. \end{aligned}$$

By treating the other parameters (p,α,t) as constants when taking the Fourier transform of h with respect to τ, the result follows from standard trigonometric manipulations (see [2]). □

The instantaneous flux Φ characterized in Theorem 5 itself simplifies to a harmonic form under the harmonic assumption:

Corollary 1 (Instantaneous flux for time-harmonic perturbations)**.** *If the perturbation $\boldsymbol{g}$ satisfies the time-harmonic assumption (23), then the instantaneous flux of Theorem 5 can be written as*

$$\Phi(p,t,\epsilon) = \epsilon\left|\int_0^1 \mathcal{F}\left\{h(p,\alpha,\boldsymbol{.})\right\}(\omega)\,\mathrm{d}\alpha\right|\cos\left[\omega(t-p)+\phi-\arg\left(\int_0^1 \mathcal{F}\left\{h(p,\alpha,\boldsymbol{.})\right\}(\omega)\,\mathrm{d}\alpha\right)\right] + \mathcal{O}(\epsilon^2)\,, \tag{26}$$

where h is defined in (25).

Proof. See Appendix H. □

The harmonic-in-time nature of the Melnikov function means that, if staying at a location parametrized by (p,α) on Γ, as time evolves the stable and unstable manifold keep flipping their relative ordering periodically. That is, at an instance in time the unstable manifold will be above the stable one when viewing along the normal direction $\hat{\boldsymbol{n}}(p,\alpha)$, but as time evolves the manifolds will get closer, and then the unstable one will be lower than the stable one. The periodic switching will occur at a frequency ω. Given the time-sinusoidal variation, one might quantify an overarching 'size' of the splitting as the amplitude of the Melnikov function, which is simply the modulus of a Fourier transform, i.e., $\left|\mathcal{F}\left\{h(p,\alpha,\boldsymbol{.})\right\}(\omega)\right|$.

The leading-order term of the instantaneous flux is also harmonic with frequency ω. So the phase space volume per unit time crossing the pseudoseparatrix also exhibits a back-and-forth behavior, influenced by the

stable and unstable manifolds continually interchanging their relative locations. A time-averaged measure of the flux could be the amplitude of the harmonic function in (26), i.e.,

$$F_I := \left| \int_0^1 \mathcal{F}\left\{h(p,\alpha,\boldsymbol{.})\right\}(\omega)\,\mathrm{d}\alpha \right| . \tag{27}$$

The harmonic format of the Melnikov function in (24) enables a straightforward (though implicit) characterization of the intersection curves between the stable and unstable manifolds at each fixed time t:

Remark 10 (Intersection curves under harmonicity)**.** By virtue of (24), at each instance in time t, there are intersections between the perturbed stable and unstable manifolds along curves near Γ within $\mathcal{O}(\epsilon)$ of the (p,α) curves defined implicitly by

$$\mathcal{F}\left\{h(p,\alpha,\boldsymbol{.})\right\}(\omega) = 0 \qquad \text{and} \tag{28}$$

$$\omega\,(t-p) + \phi - \arg\left(\mathcal{F}\left\{h(p,\alpha,\boldsymbol{.})\right\}(\omega)\right) = \frac{\pi(2k+1)}{2} \quad ; \quad k \in \mathbb{Z}\,. \tag{29}$$

Remark 11 (Fourier transform under volume-preservation)**.** If the unperturbed flow is volume-preserving, i.e., if $\boldsymbol{\nabla}\cdot\boldsymbol{f} = 0$, then h in (25) loses its p-dependence, and is

$$h(\alpha,\tau) := [\boldsymbol{f}(\bar{\boldsymbol{x}}(\tau,\alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(\tau,\alpha)] \cdot \tilde{\boldsymbol{g}}(\bar{\boldsymbol{x}}(\tau,\alpha))\,. \tag{30}$$

Consequently, the amplitude of the Melnikov function (24) is p-independent, and t and p only occur in the combination $t-p$ within the cosine term. Zero level sets of M arising from the cosine term therefore simply move along in the p-paramatrization direction on Γ as time progresses.

Remark 12 (Intersection curves under harmonicity and volume-preservation in f)**.** If the unperturbed flow is volume-preserving, the intersection curves defined in Remark 10 specialize to

$$\mathcal{F}\left\{h(\alpha,\boldsymbol{.})\right\}(\omega) = 0 \qquad \text{and} \tag{31}$$

$$p = \tilde{p}_k(\alpha,t) := \frac{1}{\omega}\left[\omega t + \phi - \arg\left(\mathcal{F}\left\{h(\alpha,\boldsymbol{.})\right\}(\omega)\right) - \frac{\pi(2k+1)}{2}\right] \quad ; \quad (\alpha,k) \in \mathrm{S}^1 \times \mathbb{Z}\,. \tag{32}$$

Generically, (31) will correspond to a finite number of $\alpha =$ constant lines along Γ (i.e., curves lying along unperturbed trajectories). In general, we may expect these to be at the ordered set of values $\alpha = \alpha_i \in \mathrm{S}^1$, where $i = 1, 2, \ldots, I$ with I being the number of distinct roots to (31). These will generically correspond to transverse intersections since the amplitude can be expected to cross zero as α is varied. Indeed, given that $\alpha \in \mathrm{S}^1$, the expectation then is that I be even. (We note that there not being any solutions is also a possibility.) It is interesting that these lines do *not* change when t is varied; there are therefore a collection of specialized unperturbed heteroclinic trajectories nearby which a curve of heteroclinic points persists for all relevant t. These special trajectories are essentially unaffected by the perturbation; any point on them continues to be a heteroclinic point, and decay in backwards and forwards time to the hyperbolic trajectories $\boldsymbol{a}_\epsilon(t)$ and $\boldsymbol{b}_\epsilon(t)$ respectively, effectively forming a $\mathcal{O}(\epsilon)$-close heteroclinic trajectory to the original unperturbed one.

The second equation in Remark 12 indicates that for any chosen value of $k \in \mathbb{Z}$, $p = \tilde{p}_k(\alpha,t)$ in (32) defines a curve of heteroclinic points which goes across all the α values in S^1. We call this a *heteroclinic ring* Q_k. These points automatically correspond to transverse intersections because of the sinusoidal form of the Melnikov function. Each of the Q_k, $k \in \mathbb{Z}$ is clearly a simple shift of one another in the p-direction, defined according to (32). There being infinitely many heteroclinic rings is the analogous situation to there being infinitely many heteroclinic points in the two-dimensional case. We can visualize the *generic* heteroclinic point distribution most easily in the (α,p) parameter space, at fixed time t¡ as shown in Fig. 9. We have drawn this for the situation where $I = 6$, with the white vertical lines representing the zeros α_i associated with (31). These lines split S^2 into $I = 6$ separate regions. The heteroclinic rings Q_k are given by the white curves $p = \tilde{p}_k(\alpha,t)$ which form an infinite collection spaced apart in the p-direction by π/ω, half the period of the perturbation. The Melnikov function M will switch sign as each of these curves is crossed, and will also switch signs when an $\alpha = \alpha_i$ line is crossed. A positive M indicates that the unstable manifold is

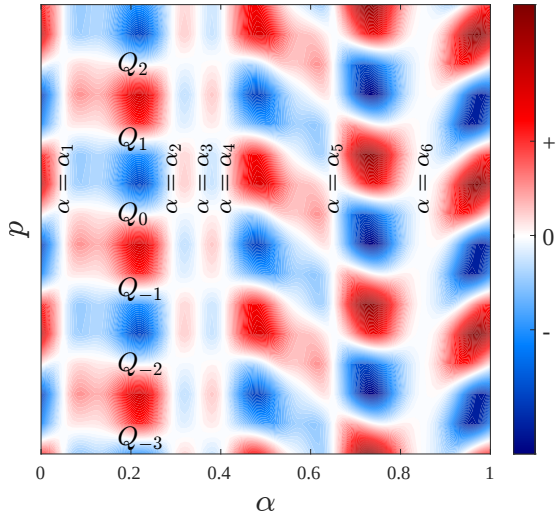


Figure 9: The size of the Melnikov function $M(p, \alpha, t)$ at a fixed time-instance t, with respect to (p, α)-space, under time-harmonic volume-preserving conditions. Darker colours indicate larger amplitudes of M. The (infinitely many) heteroclinic ring curves Q_k correspond to $p = \tilde{p}_k(\alpha, t)$ as defined in (32), at which $M = 0$. The vertical lines $\alpha = \alpha_i$ are defined via (31, and $M = 0$ along these as well. Each region separated by these curves is associated with a lobe.

outside the stable one, and vice versa. Each of the regions bounded by these curves in (p, α)-space subtends a lobe which is bounded by a stable and an unstable manifold, and their intersection at (p, α)-values which defined the bounding curves. Each lobe $L_{ik}(t)$ for $(i, k) \in \{1, \ldots, I\} \times \mathbb{Z}$ at fixed time t therefore has a $\mathcal{O}(\epsilon)$-characterization as

$$L_{ik}(t) = \left\{ \bar{\boldsymbol{x}}(p, \alpha) + \frac{\epsilon \left[s M^u(p, \alpha, t) + (1 - s) M^s(p, \alpha, t) \right] \hat{\boldsymbol{n}}(p, \alpha)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}(p, \alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(p, \alpha)|} \quad : \right.$$
$$\left. \alpha_{i-1} < \alpha < \alpha_i \,,\, \tilde{p}_{k-1}(\alpha, t) \leq p \leq \tilde{p}_k(\alpha, t) \,,\, s \in [0, 1] \right\}, \tag{33}$$

expressed in terms of the unit normal vector $\hat{\boldsymbol{n}}$ at locations $\bar{\boldsymbol{x}}(p, \alpha, t)$ on Γ, as defined in (13), and the unstable and stable Melnikov functions M^u and M^s. Given that $\alpha \in \mathrm{S}^1$, the subscripts to α will be considered modulo I in order to make this definition of a lobe work. We remark that the (p, α)-region demarcating each $L_{ik}(t)$ is not in general equal in area; Fig. 9 shows that the the α_i defined via (31) need not be equally spaced.

Theorem 7 (Lobe volume for time-harmonic volume-preserving flow)**.** *Suppose $\boldsymbol{g}$ is harmonic, the unperturbed flow is volume-preserving (i.e., $\boldsymbol{\nabla} \cdot \boldsymbol{f} = 0$), and the lobe structure is examined at a time t. Then the volume of a lobe $L_{ik}(t)$ for $(i, k) \in \{1, \ldots, I\} \times \mathbb{Z}$ is*

$$\text{Volume}\left(L_{ik}(t)\right) = \frac{2\epsilon}{\omega} \int_{\alpha_{i-1}}^{\alpha_i} \left| \mathcal{F}\left\{ h(\alpha, \boldsymbol{.}) \right\}(\omega) \right| \, \mathrm{d}\alpha + \mathcal{O}(\epsilon^2) \,. \tag{34}$$

Proof. See Appendix I. □

We note that the leading-order lobe volume is independent of k, the indexing associated with the $p = \tilde{p}_k(\alpha, t)$ curves bounding the lobe. It is also independent of the time t. However, it is *not* in general independent of i, which can be see by inspection of Fig. 9: the α_is defined via (31) need not be equally spaced, and the maximum size of the Melnikov function within each (p, α)-region defining a lobe need not be the same. This latter face can moreover be see in (31), given the α-dependence of the amplitude. Consequently, *even under time-harmonicity and volume-preservation*, the generated lobe volumes are *not* equal in general.

3.3 Connections to the lobe dynamics approach

There is a significant existing literature on quantifying transport in *maps*, i.e., in discrete dynamical systems [42, 38, 43], in many cases with additional conditions such as volume-preservation (see [44] for a review). There is limited work on three and higher dimensions in this regard [42, 43], and in particular the recent article by Maelfeyt [43] provides an interesting extension of the concept of symbolic dynamics [31] to three dimensions to allow the Smale–Birkhoff theory of establishing chaos [1, 31] to be adapted. Time-periodic *flows*, i.e., generated from a continuous dynamical system, can be sampled at equally-spaced times (with spacing equal to the period T of the vector field generating the flow) to generate a Poincaré map [1, 31, 66], thereby allowing access to the theory of transport in maps. While Poincaré maps are more restrictive in that they must obey continuous flow topologies, concepts such as lobes and lobe dynamics remain relevant. Consequently, there is a considerable literature on such time-periodic flows, and our goal in this section is to establish some connections to these works, notably to a popular two-dimensional approach which quantifies transport via lobe dynamics, and connects to a Melnikov approach [57, 66]. We emphasize that this section relates to establishing connections to the literature rather than obtaining important results in and of themselves, because we are able to quantify transport more generally in Section 3.1 without the assumptions needed for lobes to be relevant.

It will help to briefly describe the elegant two-dimensional theory [57, 66] first. Suppose the perturbation is periodic in time (with period T). This allows for defining a Poincaré map which samples the flow at time-intervals T, and consequently accesses the relevant theory for maps (discrete dynamical systems) rather than for flows. Then, there is a point $\boldsymbol{a}_\epsilon(t)$ at a time t which is close to $\boldsymbol{a}$ and is a fixed point under the Poincaré map [57, 66]. Alternatively—if thinking in terms of the flow—the $\boldsymbol{a}_\epsilon(t)$ corresponds to the perturbed hyperbolic trajectory's time-dependence. In either interpretation, there will be perturbed stable and unstable manifolds in this time-slice t, and suppose that the manifold intersection pattern is as given in Fig. 10. This picture is drawn under the assumption that the perturbation is *time-harmonic* (see Definition 1; this is stronger than mere time-periodicity), and that there is at least one transverse intersection point $\boldsymbol{q}$ between the perturbed manifolds. (Topologically classifying intersection patterns, as well as cataloguing secondary intersections, can lead to important additional information on transport [55, 56, 49, 48], but this is not our current focus.) There are several methods for quantifying transport due to patterns such as these in general dimensions [60, 43]; these often have specific assumptions such as being generated from maps and/or preserving phase space volume. Our focus is on the perturbative two-dimensional time-periodic theory due to Rom-Kedar and collaborators [57, 66] since this is closest to our setting. It turns out that then the (two-dimensional) Melnikov function which captures intersections is itself sinusoidal with frequency ω, and the implication is that the intersection of the perturbed unstable manifold in relation to the stable one is topologically equivalent to the intersection of the sinusoidal function with the horizontal axis. Thus there are infinitely many isolated transverse intersections.

Next, the idea is to define a 'pseudo-separatrix' [57, 66] at a chosen time t as follows: join the unstable manifold ($\Gamma^u_\epsilon(\boldsymbol{a}_\epsilon, t)$, red) emanating from $\boldsymbol{a}_\epsilon(t)$ up to a transverse intersection point, $\boldsymbol{q}$, to the stable manifold ($\Gamma^s_\epsilon(\boldsymbol{b}_\epsilon, t)$, blue) emanating from $\boldsymbol{b}_\epsilon(t)$ (shown by the heavy curve in Fig. 10). The intersection point $\boldsymbol{q}$ corresponds to a heteroclinic point, because it lies on both manifolds. What this means is that if choosing an 'initial' condition at this point at time t, and defining $\boldsymbol{x}(\tau)$ as being the trajectory going along this point which therefore obeys $\boldsymbol{x}(t) = \boldsymbol{q}$, then $|\boldsymbol{x}(\tau) - \boldsymbol{a}_\epsilon(\tau)|$ decays to zero as $\tau \to -\infty$, and moreover $|\boldsymbol{x}(\tau) - \boldsymbol{b}_\epsilon(\tau)|$ decays to zero as $\tau \to \infty$. Now, one considers a Poincaré map P of time $T = 2\pi/\omega$ (the period of the perturbation) on the phase space at the fixed time t. Given the time-periodicity of the flow, the phase space curves will be exactly the same at times $t + nT$, for $n \in \mathbb{Z}$. Every intersection point in Fig. 10) must map to another under this map, since being on both the stable and the unstable manifolds is an invariant property with respect to the map. Thus, the points $P^{-1}(\boldsymbol{q})$ and $P(\boldsymbol{q})$ are themselves intersection points in Fig. 10, and the region between $P^{-1}(\boldsymbol{q}$ and $\boldsymbol{q}$ (and similarly between $\boldsymbol{q}$ and $P(\boldsymbol{q})$ must possess an intersection pattern topologically equivalent to that of a sinusoidal curve intersecting the horizontal axis over one period. The important region consists of the parts of the stable and unstable manifolds lying between $P^{-1}(\boldsymbol{q})$ and $\boldsymbol{q}$, and also between $\boldsymbol{q}$ and $P(\boldsymbol{q})$. This region is sometimes called the "turnstile" [42, 44]. With respect to the pseudo-separatrix, it is only the lobes in the turnstile region (i.e., A and B in Fig. 10) which are involved in crossing the pseudo-separatrix under the action of P [57, 66] (or P^{-1}). Consequently, determining the areas of these lobes gives a nice assessment of transport.

When is the lobe area a well-defined measure of transport? There are two important assumptions to

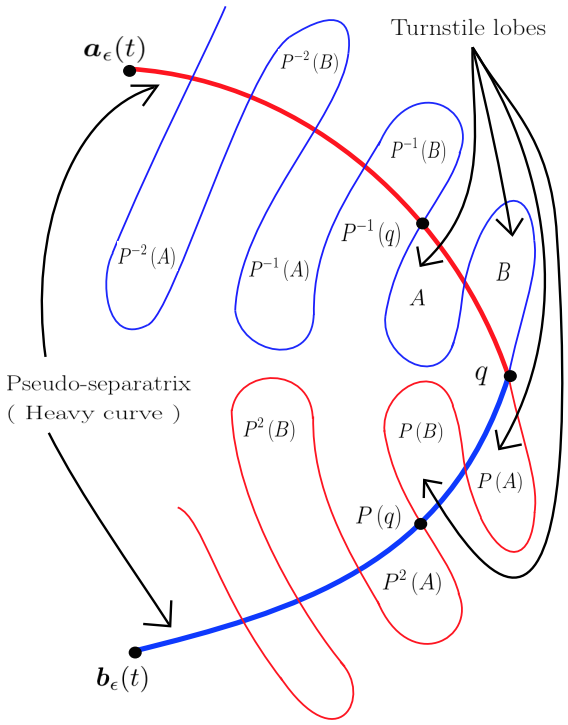


Figure 10: The two-dimensional lobe dynamics scenario [57, 66], occurring due to the intersection of the perturbed unstable manifold of $\boldsymbol{a}_\epsilon(t)$ (red) and stable manifold of $\boldsymbol{b}_\epsilon(t)$ (blue) at at a fixed time t. It is only the *turnstile lobes* between $P^{-1}(\boldsymbol{q})$ and $P(\boldsymbol{q})$ which are involved in crossing the pseudo-separatrix (heavy curve) per iteration of the Poincaré map.

make this work. The first is that the perturbation $\boldsymbol{g}$ is *time-harmonic*, which has already been discussed. This assumption is needed to ensure that the intersection pattern between $P^{-1}(\boldsymbol{q})$ and $\boldsymbol{q}$ is topologically that of a sinusoidal function over one period, and also that the lobes A and B possess the same area to leading-order (this can be shown using the pleasing connection between the Melnikov function and lobe areas [57, 66]). Simply insisting on $\boldsymbol{g}$ being time-periodic is insufficient, because it is possible that then there are many lobes, of different sizes, between an intersection point $\boldsymbol{q}$ and its preimage $P^{-1}(\boldsymbol{q})$ [2]. It is also possible that there are no intersection points at all, and consequently no lobes to speak of [2]. Thus, lobes are not necessarily the optimal method for quantifying transport in flows with *general* periodicity, whereas the instantaneous flux approach continues to be valid. The second assumption is that at least the unperturbed flow is area-preserving. This ensures that when a lobe gets mapped, it maps into a lobe of equal area to leading-order. (If the perturbation is also area-preserving, then the lobes must have equal areas, and not just to leading-order.) Consequently, *the* lobe area (or its leading-order expression) can be used as a well-defined measure of the transport that occurs, because all four of the turnstile lobes must have the same area (at least to leading-order). (We remark that there have been extensions to lobe dynamics approaches to when the perturbation is *quasi-periodic* in time [18]. For general time-*aperiodic* $\boldsymbol{g}$, and general non-area-preserving flows, transport is better quantified in terms of an instantaneous flux [4, 10, 7], eschewing lobes altogether; it is the three-dimensional version of this which we presented in Section 3.1.)

We now turn to the three-dimensional analog of this process. We have already described the manifold intersection pattern in the previous section under the time-harmonic and volume-preserving conditions, with the intersection curves between the manifolds being given where M is zero. If we go along a constant α-line in Fig. 9, M will periodically vary from positive to negative (and vice versa) as each $p = \tilde{p}_k(\alpha, t)$ curve is crossed. This means that if going along an unperturbed heteroclinic trajectory $\bar{\boldsymbol{x}}(p, \alpha)$ with fixed α on Γ, the perturbed stable and unstable manifold will periodically be intersecting (nearby the locations $\bar{\boldsymbol{x}}\left(\tilde{p}(\alpha, k), \alpha\right)$), with the stable and unstable manifolds interchanging relative locations at each crossing of the curves $p = \tilde{p}_k(\alpha, t)$. So this picture would be topologically equivalent to the two-dimensional picture of Fig. 10, with infinitely

many intersections occurring in approaching the points $\boldsymbol{a}_\epsilon(t)$ and $\boldsymbol{b}_\epsilon(t)$. To understand what happens when α changes, we may imagine rotating Fig. 10 around the axis connecting the points $\boldsymbol{a}_\epsilon(t)$ and $\boldsymbol{b}_\epsilon(t)$. However, in doing so, we must bear in mind that when any $\alpha = \alpha_i$ values are crossed, the stable and unstable manifolds will interchange their relative orientation. This forms a boundary for that lobe, and another lobe (with stable and unstable manifolds in interchanged locations), then appears for α values slightly larger.

We can still apply lobe dynamics, but there are some differences to the two-dimensional situation. Suppose we choose q to be the point associated with the intersection with $p = \tilde{p}_0(\alpha, t)$, and when considered over all $\alpha \in \mathrm{S}^1$ this forms the ring Q_0 of heteroclinic points. The other intersection points visible in Fig. 10 similarly generate heteroclinic rings upon rotation: $P^{-1}(\boldsymbol{q})$ would generate Q_{-2} (since there is another ring, Q_{-1}, in between), while $P(\boldsymbol{q})$ generates Q_2. Reviewing Fig. 9 may also help in understanding these stuctures and their relationship to the Melnikov function. In the full three-dimensional geometry, we can think of the pseudo-separatrix as being the stable manifold emanating from $\boldsymbol{a}_\epsilon(t)$ until it hits Q_0, and the unstable manifold emanating from $\boldsymbol{b}_\epsilon(t)$ until it too hits Q_0. The turnstile region consists of the lobes lying between Q_{-2} and Q_2. Upon iterating the map, it is only lobes in this region which interchange their relative positioning in regards to being inside or outside the pseudo-separatrix. However, unlike the simple statement we made in the two-dimensional case that "B gets mapped from outside the pseudo-separatrix to $P(B)$ inside it," we have the issue that the three-dimensional version of B, as it gets rotated around, actually has the stable and unstable manifold interchanging positioning when crossing $\alpha = \alpha_i$ (this occurs six times for the situation pictured in Fig. 9), for example. Each and every one of these lobes, when mapped by P, will swap its location in relation to the pseudo-separatrix. Given the fact that the lobes do not have to have equal volumes to leading-order even under time-harmonicity and volume-preservation (see Theorem 7), we cannot assign the volume of a particular lobe as unequivocally representing the transport occurring across the pseudo-separatrix. We might instead sum of the volumes of all of the lobes associated with the turnstile region (the analog of B in Fig. 10). Using Theorem 7, this will tell us then that the total transported volume would be

$$\begin{aligned}\text{Birectional transported volume} &= \sum_{i=1}^{I} \frac{2\epsilon}{\omega} \int_{\alpha_{i-1}}^{\alpha_i} \left|\mathcal{F}\left\{h(\alpha, \boldsymbol{.})\right\}(\omega)\right| \, \mathrm{d}\alpha + \mathcal{O}(\epsilon^2) \\ &= \frac{2\epsilon}{\omega} \int_0^1 \left|\mathcal{F}\left\{h(\alpha, \boldsymbol{.})\right\}(\omega)\right| \, \mathrm{d}\alpha + \mathcal{O}(\epsilon^2)\,.\end{aligned}$$

This does 'extend' lobe dynamics in providing a transport measure to the three-dimensional situation, but is a somewhat unsatisfactory quantification, since this adds together lobe volumes which transport in different directions across the pseudo-separatrix.

Remark 13 (Lobes transported per time)**.** As pointed out by Rom-Kedar and Poje [58], given the dependence on the time period of the Poincaré map, the lobe volume by itself may not be a good quantifier of the flux in the sense of fluid volume exchanged per unit time. This is better obtained by dividing the total lobe volume above by the associated time-of-flow $T = 2\pi/\omega$ of the Poincaré map. This yields

$$\text{Flux (harmonic)} = \frac{\epsilon}{\pi} \int_0^1 \left|\mathcal{F}\left\{h(\alpha, \boldsymbol{.})\right\}(\omega)\right| \, \mathrm{d}\alpha + \mathcal{O}(\epsilon^2)\,. \tag{35}$$

Interesting implications related to the flux as a function of the frequency ω can now be made, just as for the two-dimensional case. Generically, the expectation is for the flux to increase at small ω and eventually decay as $\omega \to \infty$, implying the presence of a flux optimizing frequency [58]. For a given unperturbed flow for which the heteroclinic structure is known, the Fourier transform formula in (26) can therefore be employed easily, for example, to find the frequency resulting in the greatest flux (as has been done in two dimensions [5]). Alternatively, the formula can be analyzed with different spatial perturbations $\tilde{\boldsymbol{g}}$ at fixed ω, attempting to find transport-optimizing perturbations (see [3, 11] for two-dimensional implementations).

Remark 14 (Comparison of flux measures for harmonic situations)**.** When $\boldsymbol{g}$ is harmonic, the instantaneous flux description gave us the formula (26), and the amplitude F_I of this harmonic function (27) can therefore be thought of as a time-averaged measure of the leading-order flux. In contrast, the lobe dynamics description gives a measure of the flux as the $\mathcal{O}(\epsilon)$-term (call it F_L) in (35. The two measures F_I (I for instantaneous) and F_L (L for lobes) have some similarities and differences:

- The scaling factor difference of π is inconsequential and related to the time-scaling used in converting the lobe volumes to a flux in (35. If the time-average (rather than the amplitude) of (26) were used instead, a slightly different scaling factor involving ω would result instead.

- The instantaneous flux version F_I is p-dependent (via h), unlike F_L. The reason for this is that Corollory 1 does *not* require a volume-preservation assumption, whereas for the lobe dynamics approach, one needs this to ensure that the "volume of a lobe" is unambiguous to leading-order. Consequently, in the more general instantaneous flux framework, there will be dependence on the location p chosen to define the strip. The lack of volume-preservation means that the flux across two strips located at $p = p_1$ and $p = p_2 \neq p_1$ are not the same. If volume-preservation of $\boldsymbol{f}$ was imposed, however, a straightforward application of the divergence theorem for the region bounded by the strips $S(p_1, t, \epsilon)$ and $S(p_2, t, \epsilon)$, and the perturbed stable and unstable manifolds $\Gamma^u_\epsilon(\boldsymbol{a}_\epsilon, t)$ and $\Gamma^s_\epsilon(\boldsymbol{b}_\epsilon, t)$ implies that the flux across the two strips are the same (to leading-order). Of course, then the Fourier transform $\int_0^1 \mathcal{F}\{h(p, \alpha, \boldsymbol{.})\}(\omega)\, \mathrm{d}\alpha$ reduces to $\int_0^1 \mathcal{F}\{h(\alpha, \boldsymbol{.})\}(\omega)\, \mathrm{d}\alpha$ where the new h is p-independent as given in (30) as opposed to (25).

- The modulus signs are inside the α-integral in F_L, but outside it in F_I. The intuition for this is subtle. In the lobe dynamics approach, a lobe volume is computed by integrating over a region in which the Melnikov function were sign definite, Thus, the amplitude of the harmonic form of the Melnikov function (24) remains the same sign over all α. This is not necessarily so for the general instantaneous flux scenario, in which a ring of constant p is chosen to define the strip which forms the connection between the perturbed stable and unstable manifolds. When going along this, the stable and unstable manifolds may interchange their locations; consequently, the Melnikov function will in general take on both positive and negative values. The instantaneous flux takes all this into account, 'adding everything up' to get a *net* flux, and this is accomplished with the modulus being taking *after* the net impact is computed.

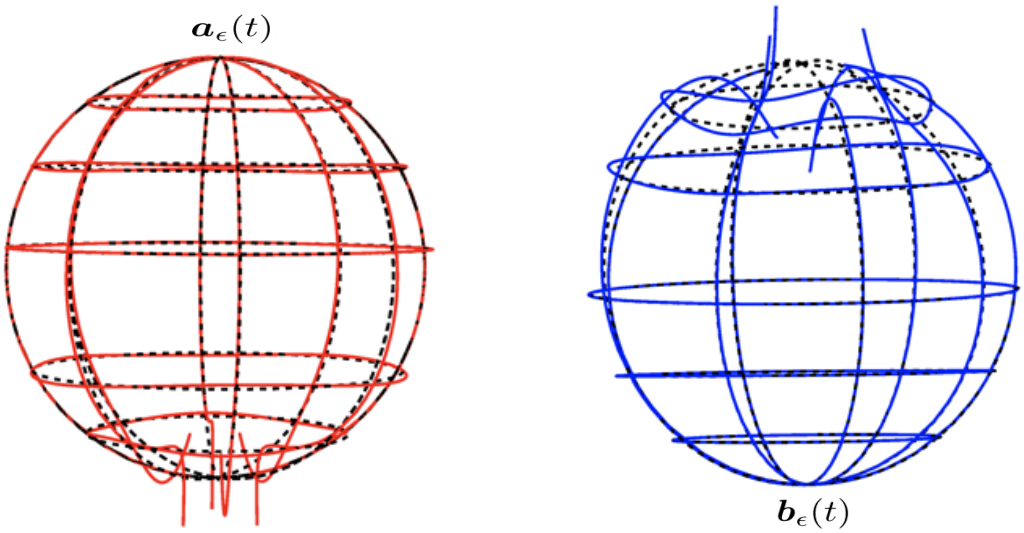
4 Application to Hill's spherical vortex

In this section, our Melnikov theory is applied to Hill's spherical vortex, in particular in quantifying the splitting of the stable and unstable manifolds after perturbation. We consider both the classical (no-swirl) and the rotating (swirl) versions, which correspond respectively to the situations of purely real, and complex-conjugate eigenvalues. Our motivation for using the Hill's vortex is that it is a fundamental solution to the Euler equations of motion in fluid mechanics, and has the kinematic structure which appears when fluid is rotated within a cylinder [36, 63, 41, 33, 13, 46]. The key structures are associated with the saddle fixed points at the 'north pole' and 'south pole,' which have respectively a two-dimensional (one-dimensional) unstable manifold and a one-dimensional (two-dimensional) stable manifold. The two-dimensional time-varying stable manifold may be considered a flow barrier in forwards time, since it is of codimension- 1, and fluid particles which begin near to it on opposite sides will get pushed exponentially away from each other due to the eventual influence of the one-dimensional unstable manifold [7, 12]. Similarly, the two-dimensional time-varying unstable manifold is a flow barrier in backwards time. In the unperturbed (steady) classical Hill's spherical vortex, these coincide to form a two-dimensional spherical surface which can then be considered a flow barrier in both forward and backward time. Under perturbations (arising from instabilities) this separating surface collapses, leading to so-called "vortex breakdown" [36, 63, 41]. In this section, we apply our theory to describe mathematically the impact of general perturbations on these separating surfaces.

4.1 Classical Hill's spherical vortex

The classical Hill's spherical vortex is a solution of Euler's equations of motion for an inviscid fluid. In (r, θ, ϕ) spherical polar coordinates with $r \geq 0$ is the radial distance from the origin, $\theta \in [0, \pi]$ the polar angle and $\phi \in [0, 2\pi)$ the azimuthal angle, the (continuous) velocity field is given by [32, 13]

$$\boldsymbol{f}(r, \theta, \phi) = \begin{cases} \frac{3\cos(\theta)}{2}\left(1 - r^2\right)\hat{\boldsymbol{r}} - \frac{3\sin(\theta)}{2}\left(1 - 2r^2\right)\hat{\boldsymbol{\theta}} & \text{if } r \leq 1 \\ -\frac{\cos(\theta)}{r^3}\left(r^3 - 1\right)\hat{\boldsymbol{r}} + \frac{\sin(\theta)}{2r^3}\left(2r^3 + 1\right)\hat{\boldsymbol{\theta}} & \text{if } r > 1 \end{cases} . \tag{36}$$



(a) (b)

Figure 11: For time $t = 1$ and $\epsilon = 0.1$, (a) the perturbed unstable (red) and (b) stable (in blue) manifolds, in comparison to the unperturbed manifold (in dashed black) for the perturbation (41), illustrated via constant α and p curves. Computations were performed using (37) and (38).

It is easily verifiable that $\boldsymbol{\nabla} \cdot \boldsymbol{f} = 0$ here; the flow is volume-preserving. The globe $r = 1$ is a heteroclinic manifold Γ, associated with saddle points located at the north and south poles. In (r, θ)-coordinates, these points can be written as $\boldsymbol{a} = (1, 0)$ and $\boldsymbol{b} = (1, \pi)$, and the unstable manifold of $\boldsymbol{a}$ coincides with the stable manifold of $\boldsymbol{b}$, to form Γ. This manifold is foliated with heteroclinic trajectories which have a constant ϕ-value (constant longitude), and thus the trajectory-identifying parameter is $\alpha = \phi/(2\pi) \in \mathrm{S}^1$. The p-parametrization along each manifold is thus associated directly with the latitude. The splitting of Γ, using the volume-preserving requirement and a functional-analytic viewpoint, has been previously pursued [13].

Under the condition that the perturbation vector field $\boldsymbol{g}$ has components (g_r, g_θ, g_ϕ) in the (r, θ, ϕ)-coordinate system, we show how we can use our theory to derive that the perturbed unstable manifold pf the hyperbolic trajectory near the north pole is given by

$$\boldsymbol{r}^u(p, \alpha, \epsilon, t) \approx \left[1 + \epsilon \cosh^2 \frac{3p}{2} \int_{-\infty}^{p} \operatorname{sech}^2 \frac{3\tau}{2}\, g_r \left(1, \cos^{-1}\left(-\tanh \frac{3\tau}{2}\right), 2\pi\alpha, \tau + t - p\right) \mathrm{d}\tau\,,\ \cos^{-1}\left(-\tanh \frac{3p}{2}\right), 2\pi\alpha\right], \tag{37}$$

for $(p, \alpha, t) \in (-\infty, P^u] \times \mathrm{S}^1 \times (-\infty, T^u]$ for finite P^u and T^u, in (r, θ, ϕ) component form. Similarly, the perturbed stable manifold of the hyperbolic trajectory near the south pole is given by

$$\boldsymbol{r}^s(p, \alpha, \epsilon, t) \approx \left[1 - \epsilon \cosh^2 \frac{3p}{2} \int_{p}^{\infty} \operatorname{sech}^2 \frac{3\tau}{2}\, g_r \left(1, \cos^{-1}\left(-\tanh \frac{3\tau}{2}\right), 2\pi\alpha, \tau + t - p\right) \mathrm{d}\tau\,,\ \cos^{-1}\left(-\tanh \frac{3p}{2}\right), 2\pi\alpha\right], \tag{38}$$

for $(p, \alpha, t) \in [P^s, \infty) \times \mathrm{S}^1 \times [T^s, \infty)$ for finite P^s and T^s. The Melnikov function takes the form

$$M\,(p, \alpha, t) = 3\pi \int_{-\infty}^{\infty} \operatorname{sech}^2 \frac{3\tau}{2}\, g_r \left(1, \cos^{-1}\left(-\tanh \frac{3\tau}{2}\right), 2\pi\alpha, \tau + t - p\right) \mathrm{d}\tau\,, \tag{39}$$

for $(p, \alpha, t) \in [P^s, P^u] \times \mathrm{S}^1 \times [T^s, T^u]$, and the instantaneous flux is

$$\Phi(p, t, \epsilon) = 3\pi\epsilon \int_0^1 \int_{-\infty}^{\infty} \operatorname{sech}^2 \frac{3\tau}{2}\, g_r \left(1, \cos^{-1}\left(-\tanh \frac{3\tau}{2}\right), 2\pi\alpha, \tau + t - p\right) \mathrm{d}\tau\, \mathrm{d}\alpha + \mathcal{O}(\epsilon^2)\,. \tag{40}$$

See Appendix J for a derivation of these results.

To validate the expressions with our intuition, we will now choose a g_r for which explicit analytical expressions can be obtained for M and Φ, and interpreted. We select

$$g_r\,(r, \theta, \phi, t) = r^2 \sin\theta \sin(3\phi) \cos(4t)\,, \tag{41}$$

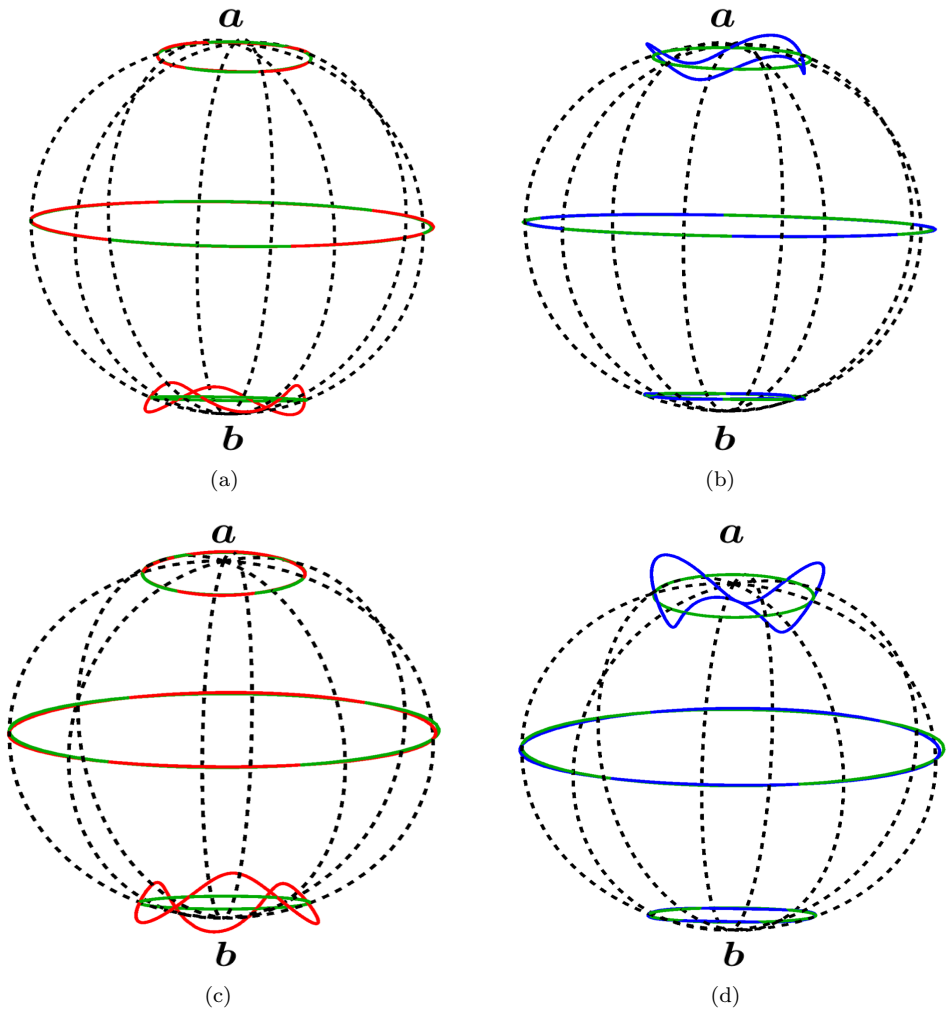


Figure 12: Behavior of the perturbed unstable (red) and stable (blue) manifolds using (37) and (38) with $\epsilon = 0.1$ and the perturbation g_r in (41) for the classical Hill's spherical vortex, at $\theta = 0.4$ (approximate "Arctic circle"), $\theta = \pi/2$ ("equator") and $\theta = \pi - 0.4$ (approximate "Antarctic circle"), with corresponding trajectories on the unperturbed Γ (green), and the rows being at times $t = 0$ (top) and $t = 2$ (bottom).

which we can understand in the context of time-periodic theory. The perturbed manifolds at $t = 1$ and perturbation strength $\epsilon = 0.1$, can be expressed parametrically by the (p, α)-variables as given in $\boldsymbol{r}^{u,s}$. We can visualize the manifolds by evaluating the curves $\alpha =$ constant (with p-varying along each such curve), and also the curves $p =$ constant (with α varying along each curve). We visualize the approximate perturbed manifolds at time $t = 1$ with $\epsilon = 0.1$ in Fig. 11 by plotting $\alpha =$ constant and $p =$ constant curves, in comparison to the unperturbed Γ (black). These pictures illustrate the fact that we lose control of the unstable (resp. stable) manifold as we approach the south (resp. north) poles, with infinite intersections with Γ occurring along constant latitudes which accumulate towards the pole. We illustrate the worsening of this behavior with t in Fig. 12, where the intersections between the manifolds with several constant latitudes are shown at two different t values. The perturbed unstable trajectory approximated at time $t = 2$ near the "Antarctic circle" and the perturbed stable trajectory approximated at time $t = 2$ near the "Arctic circle" demonstrate significant deviation from the unperturbed manifold, an effect which is exacerbated at larger times.

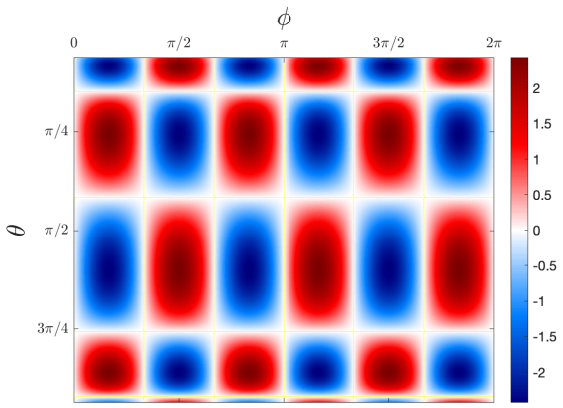


Figure 13: The Melnikov function at $t = 1$ expressed in terms of longitude-latitude coordinates (ϕ, θ) (a Mercator representation) for Hill's spherical vortex, computing using (39) for the perturbation (41). Zero values are indicated by the yellow lines.

For the choice (41), the Melnikov function in (39) is explicitly compoutable to be

$$M(p,\alpha,t) = 6\pi \sin(6\pi\alpha)\cos\left[4(p-t)\right]\int_0^\infty \operatorname{sech}^3\left(\frac{3\tau}{2}\right)\cos\left(4\tau\right)\mathrm{d}\tau = \frac{73\,\pi^2}{9}\operatorname{sech}\frac{4\pi}{3}\,\sin\left(6\pi\alpha\right)\cos\left[4(p-t)\right]\,,$$

which is also consistent with the representation in (24. Consider the perturbed manifolds at a fixed time t. If p satisfies $p - t \neq (2k+1)\pi/8$ for $k \in \mathbb{Z}$, M clearly has simple zeros when $\alpha = 0, 1/6, 1/3, 1/2, 2/3$ and $5/6$, consistent with the condition (31). Thus, the two-dimensional perturbed manifolds intersect along along curves which are $\mathcal{O}(\epsilon)$-close to these six constant longitude lines $\phi = 0, \pi/3, 2\pi/3, \pi, 4\pi/3$ and $5\pi/3$ on Γ. Moreover, at α-values not on these curves, M has simple zeros with respect to p when

$$p = \tilde{p}_k(\boldsymbol{.}, t) := t + (2k+1)\pi/8\,,$$

i.e., on the latitudes defined by

$$\theta = \theta_k := \cos^{-1}\left[-\tanh\left(\frac{3t}{2} + \frac{3(2k+1)\pi}{16}\right)\right] \quad ; \quad k \in \mathbb{Z}\,, \tag{42}$$

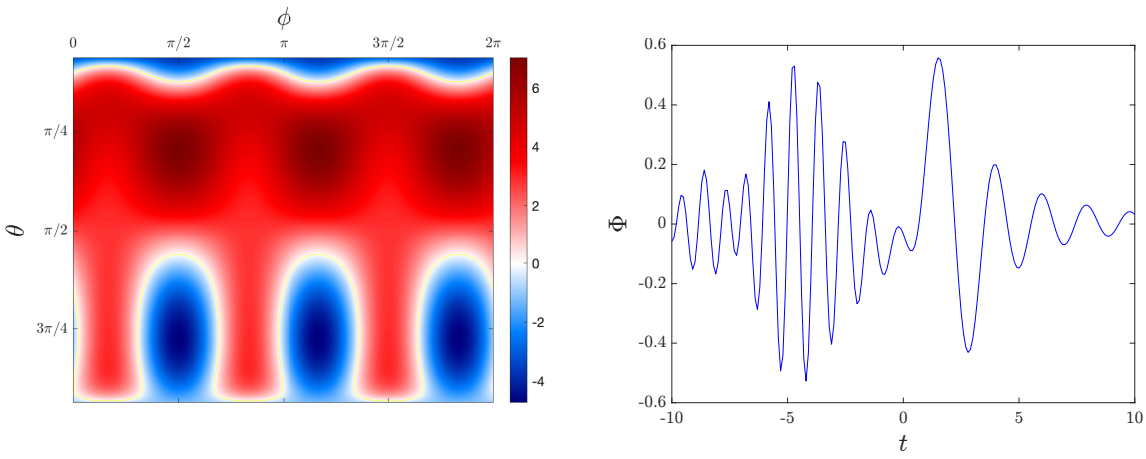
which yields infinitely many unique values for $\theta \in [0, \pi]$, accumulating towards both $\theta = 0$ and π. We illustrate this further in Fig. 13, with the Melnikov function computed directly using (39) for the perturbation (41), rather than appealing to the analytic form above. We refer to this figure as a *Mercator representation* of the Melnikov function in keeping with the fact that it is a representation on the stretched out globe. Fig. 13 is indeed a special case of Fig. 9, since $\alpha \sim \phi$ and $p \sim \theta$. An infinite number of narrower and narrower rows in between the zero (yellow) horizontal lines will accumulate towards the poles.

We can approximate the volume of a lobe whose boundaries are given by adjacent zeros of p and α. For example, consider the lobe whose α limits are between 0 and $1/6$, and p limits are between $k = 1$ and $k = 2$ as defined in p_k. Using (19), we get

$$\text{Lobe volume} = \epsilon\frac{73\,\pi^2}{9}\operatorname{sech}\frac{4\pi}{3}\int_0^{1/6}\sin\left(6\pi\alpha\right)\,\mathrm{d}\alpha\int_{p_1}^{p_2}\cos\left[4(p-t)\right]\,\mathrm{d}p + \mathcal{O}(\epsilon^2) = \epsilon\frac{73\,\pi}{54}\operatorname{sech}\frac{4\pi}{3} + \mathcal{O}(\epsilon^2)\,.$$

In this case, symmetries ensure that the leading-order lobe volumes are identical for all lobes. They are also equal for all times t. The instantaneous flux is

$$\Phi\left(p,t,\epsilon\right) = \frac{73\,\epsilon\,\pi^2}{9}\operatorname{sech}\left(\frac{4\pi}{3}\right)\cos\left[4(p-t)\right]\int_0^1\sin\left(6\pi\alpha\right)\,d\alpha + \mathcal{O}(\epsilon^2) = \mathcal{O}(\epsilon^2).$$



(a) (b)

Figure 14: (a) The Melnikov function at $t = 1$, and (b) the instantaneous flux Φ variation with time with $\epsilon = 0.1$ and pseudo-separatrix on the equator for Hill's spherical vortex subject to the perturbation (43).

The reason for the leading-order instantaneous flux to be zero (for any time t and any choice of gate location p) is because of the symmetry of the splitting of the heteroclinic. Regions along any constant latitude strip in which the unstable manifold is outside the stable manifold are complemented by regions in which the opposite occurs, while the leading-order velocity field along the strip remains constant. The flux across 'in' and 'out' regions across the strip (see Fig. 8 cancel each other out by symmetry and thus there is an identical amount of flux crossing outwards as that crossing inwards. The transport can be explained via lobe dynamics in this case; three lobes in the turnstile region map from inside to outside, while the other three do the opposite (this is an alternative insight into the zero leading-order instantaneous flux). One would get a nonzero flux if the $\sin(3\phi)$ term in g_r (which led to the integral $\int_0^1 \sin(6\pi\alpha)\, d\alpha$ in the above expression for the flux) were replaced by a term which does *not* integrate to zero over all ϕ.

To demonstrate that our theory is *not* confined to time-periodic perturbations, we briefly show computations associated with the alternative choice

$$g_r(r,\theta,\phi,t) = [r\cos(\theta t + 1) + \sin(3\pi)]\tanh(4rt)\cos\left[4\left(t + e^{-0.01t^t}\right)\right]. \tag{43}$$

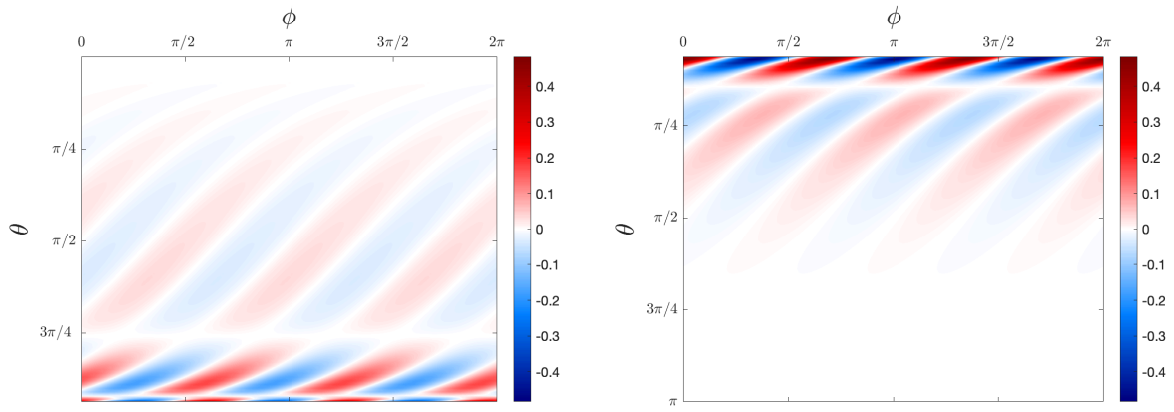
The left panel of Fig. 14 demonstrates the Melnikov function using (39) at time $t = 1$, while the right plots the flux (40) variation with time t with pseudo-separatrix chosen on the equator, with $\epsilon = 0.1$. The general time-aperiodic theory we have developed gives diverse results for different perturbations.

4.2 Hill's spherical vortex with swirl

We now consider Hill's spherical vortex with an additional swirling component in the azimuthal direction, with the far-field flow remaining in the z-direction as for the classical case. This is kinematically consistent with the vortex breakdown phenomenon in a rotating cylinder [36, 63, 41]. The steady solution to the Euler equations in this case corresponds to the continuous velocity field [61]

$$\boldsymbol{f}(r,\theta,\phi) = \begin{cases} \frac{3}{2}(1-r^2)\cos(\theta)\,\hat{\boldsymbol{r}} - \frac{3}{2}(1-2r^2)\sin(\theta)\,\hat{\boldsymbol{\theta}} - \frac{r\sin(\theta)}{2R_0}\,\hat{\boldsymbol{\phi}} & \text{if } r \le 1 \\ -\cos(\theta)\left\{1 - \frac{1}{r^2}\cos\left(\frac{r-1}{R_0}\right) + \frac{R_0}{r^3}\sin\left(\frac{r-1}{R_0}\right)\right\}\,\hat{\boldsymbol{r}} & \\ \quad + \frac{\sin(\theta)}{2r}\left\{2r + \frac{1}{r}\cos\left(\frac{r-1}{R_0}\right) + \left[\frac{1}{R_0} - \frac{R_0}{r^2}\right]\sin\left(\frac{r-1}{R_0}\right)\right\}\,\hat{\boldsymbol{\theta}} - \frac{r\sin(\theta)}{2R_0}\,\hat{\boldsymbol{\phi}} & \text{if } r > 1 \end{cases}, \tag{44}$$

where $R_0 > 0$ is the Rossby number. This velocity field is also divergence-free consonant with the incompressibility assumption, and the classical Hill's spherical vortex is a special case in the limit $R_0 \to \infty$. The north and south poles $(r,\theta) = (1,0)$ and $(1,\pi)$ are still saddle points, and $r = 1$ is the heteroclinic manifold



(a) (b)

Figure 15: Mercator representations of the (a) unstable and (b) stable manifold at $t = 1$ with $\epsilon = 0.1$ for the swirling Hill's spherical vortex subject to the perturbation (41).

Γ connecting them. However, heteroclinic trajectories swirl around the globe, and spiral from the poles because the linearization at the poles results in complex conjugate eigenvalues. As before, we parametrize Γ by $\bar{\boldsymbol{x}}(p,\alpha) = \big(1, \bar{\theta}(p,\alpha), \bar{\phi}(p,\alpha)\big)$ in (r,θ,ϕ)-coordinates, and we observe that since the θ-component of the velocity is identical to the classical case, $\bar{\theta}$ is given by (60), where we have chosen $p = 0$ to be on the equator. Next, since ϕ is changing at the constant rate $-1/(2R_0)$, we have

$$\bar{\phi}(p,\alpha) = \bar{\phi}(0,\alpha) - \frac{p}{2R_0} = 2\pi\alpha - \frac{p}{2R_0}\,,$$

where we choose the parameterization on the equator such that the ϕ-coordinate on the equator divided by 2π gives the trajectory-identifying parameter $\alpha \in \mathrm{S}^1$. Thus, in (r,θ,ϕ)-component form, we have

$$\bar{\boldsymbol{x}}(p,\alpha) = \left(1, \cos^{-1}\left(-\tanh\frac{3p}{2}\right), 2\pi\alpha - \frac{p}{2R_0}\right)\,,$$

which serves to parametrize the heteroclinic manifold $r = 1$. We note that the p-variation in the θ-coordinate is equivalent to the time-variation in this steady situation, and changing α only affects the ϕ-component. Thus, despite $\bar{\boldsymbol{x}}$ having a slight difference, only very slight adjustments to the results for the classical Hill's vortex are necessary. We will not bother to rewrite these equations, apart from stating that the ϕ-coordinate within the g_r function in all the integrals simply needs to change from $2\pi\alpha$ to $2\pi\alpha - \tau/(2R_0)$.

We now show the results of computations using our theory, where we use $R_0 = 0.2$ and $\epsilon = 0.1$. The perturbations of the stable and unstable manifolds using the aforementioned (37) and (38) are shown in Fig. 15 for the perturbation (41). We use the Mercator representation, and the derivation is how the manifold deviations *outwards* from the unperturbed globe. Thus, red regions have protruded, while blue regions intrude in comparison to $r = 1$. The impact of the swirling nature of the unperturbed trajectories is apparent in this Mercator representation.

The Melnikov function can be numerically computed using the modified version of (39) for the perturbation (41), and is shown in Fig. 16. It turns out that an *explicit* analytical form can also be computed, and is given by

$$M(p,\alpha,t) = 6\pi\sqrt{A^2\sin^2(6\pi\alpha) + B^2\cos^2(6\pi\alpha)}\cos\left[4(t-p) - \tan^{-1}\frac{B\cot(6\pi\alpha)}{A}\right]\,,$$

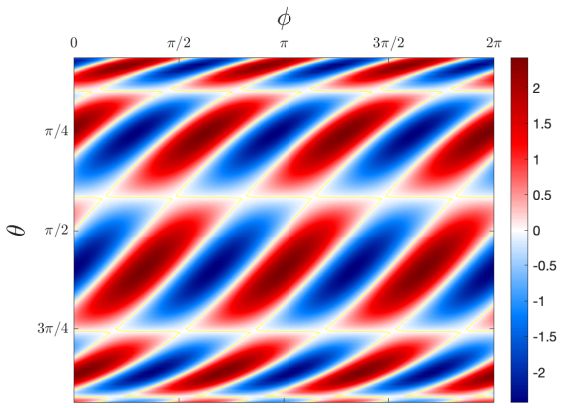


Figure 16: Mercator representation of the Melnikov function $t = 1$ for the swirling Hill's spherical vortex subject to the perturbation (41), with zeros shown by the yellow curves.

where

$$A = \frac{\pi\left[\left(73R_0^2 - 48R_0 + 9\right)\,\mathrm{sech}\,\frac{\pi(8R_0-3)}{6R_0} + \left(73R_0^2 + 48R_0 + 9\right)\,\mathrm{sech}\,\frac{\pi(8R_0+3)}{6R_0}\right]}{108R_0^2}, \quad \text{and}$$

$$B = \frac{\pi\left[\left(73R_0^2 - 48R_0 + 9\right)\,\mathrm{sech}\,\frac{\pi(8R_0-3)}{6R_0} - \left(73R_0^2 + 48R_0 + 9\right)\,\mathrm{sech}\,\frac{\pi(8R_0+3)}{6R_0}\right]}{108R_0^2}.$$

It can then be derived that zeroes of the Melnikov function occur at the (θ, ϕ) coordinates defined via

$$-\cot\left[4\left(t + \frac{2}{3}\tanh^{-1}\cos\theta\right)\right] = \frac{B}{A}\cot\left(3\phi - \frac{1}{R_0}\tanh^{-1}\cos\theta\right). \tag{45}$$

There are consistent with the six zero contour curves (in yellow) in Fig. 16. Finally the instantaneous flux $\Phi(p, t, \epsilon)$ exiting the pseudo-separatrix for the Hill's spherical vortex with swirl is

$$\Phi(p,t,\epsilon) = 3\pi\epsilon\int_{-\infty}^{\infty}\mathrm{sech}^3\left(\frac{3\tau}{2}\right)\cos\left[4(\tau - t + p)\right]\int_0^1\sin\left(6\pi\alpha - \frac{3\tau}{2R_0}\right)\mathrm{d}\alpha\,\mathrm{d}\tau + \mathcal{O}(\epsilon^2) = \mathcal{O}(\epsilon^2);$$

again, the leading-order term is zero to leading-order because of the symmetry of the perturbation.

5 Concluding remarks

In this paper, we have developed a Melnikov theory to examine two-dimensional stable/unstable manifolds of hyperbolic points in three-dimensional flows. Our approach is unashamedly geometric, taking advantage of three-dimensional intuition as is relevant to realistic fluid flows. We require neither volume-preservation (in either the unperturbed or perturbed flow), nor time-periodicity. This latter condition in particular precludes using Melnikov techniques which have been developed for maps [28, 38, 39, 40], since there are no well-defined times at which to sample flow trajectories. Under fairly general conditions on the perturbation, we derive leading-order expressions for the time-varying location of the perturbed two-dimensional stable (or unstable) manifold in Theorems 1 and 2. There is no requirement for the unperturbed manifold to have been homoclinic or heteroclinic in this development.

The second goal of this paper is to characterize transport due a time-varying perturbation breaking apart a two-dimensional heteroclinic manifold in a three-dimensional flow. We can quantify in Theorem 5 the instantaneous flux engendered across the formerly impermeable heteroclinic manifold, building on a similar idea in two dimensions [4]. If thinking in terms of three-dimensional fluid flows, we are thus able to

characterize, in terms of the Melnikov function, the leading-order fluid flux, as a volume of fluid per unit time, crossing the broken heteroclinic manifold. This is in a Lagrangian (as opposed to Eulerian) sense; the flux quantifies the transport of fluid particles following their flow history. Thus, transport between the inside and the outside of the unperturbed heteroclinic manifold is captured by this theory, which allows for both non-volume-preservation and general time-dependence. For the situation of time-harmonicity in the perturbation (a frequently addressed situation [57, 66] for which the theory associated with maps is amenable), we describe the three-dimensional analog of lobe dynamics [57, 66] and develop expressions in terms of the Melnikov function for the leading-order volumes of lobes which lie between the perturbed stable and unstable manifolds at a general time. If the unperturbed flow were volume-preserving, we show that the leading-order volumes of all lobes are identical, thereby allowing for this lobe volume to be a good quantifier of transport across the broken heteroclinic manifold via a three-dimensional version of lobe dynamics.

Higher-dimensional Melnikov methods, which develop a Melnikov function whose zeros are associated with persistent heteroclinic connections, usually have a function inside the integral which is known only as a fundamental solution to the adjoint of the equation of variations along an unperturbed heteroclinic trajectory. This is not an explicit representation (except in the case of Hamiltonian systems, which are of course moreover limited to even dimensions), and therefore the Melnikov function is not *computable*. Through our formulation in three dimensions, we derive this function automatically; this is what is inside the integral for both the general theory of locating two-dimensional manifolds, and in evaluating transport across a broken heteroclinic. We note that this function we derive is valid even if volume is not preserved.

The expressions we derive are based on the original manifold being associated with a saddle fixed point. It would be natural to attempt to extend the theory to two-dimensional manifolds of normally hyperbolic invariant sets such as periodic orbits of a curve of fixed points. In this more general situation, there are several qualitatively different geometries that one would need to consider, and it will not be possible to couch the theory in terms of eigenvalues as we have done in this paper. Nevertheless, the theory of exponential dichotomies [23, 53, 17] allows for the development of the theory with only minor modifications, as has been noted by earlier developments of Melnikov theory [22]. We expect that the formulas will be essentially unchanged, but that the interpretations regarding separating surfaces and flux will need reinterpretation depending on the nature of the normally hyperbolic set. We will study this extension in later work.

The three-dimensional situation, with two-dimensional separating surfaces, is natural to study in the context of fluid flows. As such, our work is expected to be of value in fluid transport: notably in quantifying locations of two-dimensional flow separators, and flux across broken ones. Our formulation in the time-sinusoidal context in particular allows for a tool for optimizing mixing across separating surfaces, analogously to what has been done in optimizing mixing across one-dimensional separators in two-dimensional fluid flows [3, 5, 11]. Thus, applications to either maximizing transport (to empower good mixing of a fluid, or a two-phase fluid, in industrial applications), or minimizing it (to avoid pollutants contaminating a fluid) can be examined using the tools that we have developed in this article.

Acknowledgments: SB acknowledges support from the Australian Research Council under Grant DP200101764. EB is supported by the Army Research Office (N68164-EG) and DARPA, and KGDSP is supported directly by Clarkson University.

A Proof of Theorem 1 (Displacement of unstable manifold)

Suppose $\epsilon \neq 0$. Fix $t \in (-\infty, T]$ and $(p, \alpha) \in (-\infty, P] \times \mathrm{S}^1$. Let $\tau \in (-\infty, t]$ be a general time value. If $\boldsymbol{x}^u(p, \alpha, \epsilon, t)$ (a point on the unstable manifold at time t) is close to the point $\bar{\boldsymbol{x}}^u(\alpha, p)$ on the unperturbed manifold., then we realize that the appropriate parameterization at a general time τ should be chosen such that $\boldsymbol{x}^u(p, \alpha, \epsilon, \tau)$ is the point on perturbed unstable manifold close to the point $\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)$. Since we are thinking of t as fixed, the parameter τ represents varying time, and because (p, α) is also considered fixed, $\boldsymbol{x}^u(p, \alpha, \epsilon, \tau)$ is a trajectory of the perturbed equation (1) which is expected to be close to the trajectory $\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)$ of the unperturbed equation in which $\epsilon = 0$. We define

$$\boldsymbol{z}^u(p, \alpha, \epsilon, \tau) := \frac{1}{\epsilon} \left[\boldsymbol{x}^u(p, \alpha, \epsilon, \tau) - \bar{\boldsymbol{x}}^u(\tau - t + p, \alpha) \right] \quad , \quad \tau \in (-\infty, t] \,. \tag{46}$$

Since T is finite, the difference between the perturbed and unperturbed trajectories is $\mathcal{O}(\epsilon)$, and thus $\boldsymbol{z}^u = \mathcal{O}(1)$, or more precisely there exists $K \in \mathbb{R}$, which is independent of τ and $\epsilon \in [0, \epsilon_0]$ for some ϵ_0 and for

finite T,

$$|\boldsymbol{z}^u(p,\alpha,\epsilon,\tau)| \leq K \quad \text{for } \tau \in (-\infty, T]. \tag{47}$$

Next, we define $\widetilde{M}^u(p,\alpha,\epsilon,\tau)$ as

$$\widetilde{M}^u(p,\alpha,\epsilon,\tau) := [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau-t+p,\alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau-t+p,\alpha)] \cdot \boldsymbol{z}^u(p,\alpha,\epsilon,\tau). \tag{48}$$

We note that $\boldsymbol{z}^u(p,\alpha,\epsilon,t)$ is the $\mathcal{O}(\epsilon)$-distance between the point $\boldsymbol{x}^u(p,\alpha,\epsilon,t)$ on the perturbed unstable manifold $\Gamma^u_\epsilon(\boldsymbol{a}_\epsilon,t)$ time t, and the unperturbed point $\bar{\boldsymbol{x}}^u(p,\alpha)$ which serves to parametrize Γ^u in terms of the parameters (p,α). The quantity $\widetilde{M}^u(p,\alpha,\epsilon,t)$ therefore projects this distance orthogonal to the original unperturbed manifold $\Gamma^u(\boldsymbol{a})$ from a position parametrized by (p,α). If we can determine $\widetilde{M}^u(p,\alpha,\epsilon,t)$ uniquely in the limit as $\epsilon \to 0$, then this implies that the (p,α)-parametrization (at a time t) of the perturbed manifold is legitimate; i.e., that $\Gamma^u_\epsilon(\boldsymbol{a}_\epsilon,t)$ can be realized as a graph of $\Gamma^u(\boldsymbol{a})$ subject to the restrictions $(p,\alpha,t) \in (-\infty,P] \times \mathrm{S}^1 \times (-\infty,T]$. (Once the perturbed manifold ventures beyond being close to $\Gamma^u(\boldsymbol{a})$ which these conditions imply, we will lose control of quantifying it via a the current approach.) Our method of proof will automatically provide this result, under the conditions that we have posed in Hypothesis 1.

Now, given that

$$\begin{aligned}
\widetilde{M}^u(p,\alpha,\epsilon,t) &= [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(p,\alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(p,\alpha)] \cdot \boldsymbol{z}^u(p,\alpha,\epsilon,t) \\
&= [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(p,\alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(p,\alpha)] \cdot \frac{1}{\epsilon}\left[\boldsymbol{x}^u(p,\alpha,\epsilon,t) - \bar{\boldsymbol{x}}^u(p,\alpha)\right],
\end{aligned}$$

we note from (3) that

$$d^u(p,\alpha,\epsilon,t) = \epsilon \frac{\widetilde{M}^u(p,\alpha,\epsilon,t)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}^u(p,\alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(p,\alpha)|}. \tag{49}$$

Consequently, we will determine d^u via an evolution equation for $\widetilde{M}^u$ with respect to the temporal variable τ. Using the subscript as the notation for the partial derivative, taking τ-partial derivative of (48) yields

$$\begin{aligned}
\widetilde{M}^u_\tau(p,\alpha,\epsilon,\tau) &= [D\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau-t+p,\alpha))\ \bar{\boldsymbol{x}}^u_\tau(\tau-t+p,\alpha) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau-t+p,\alpha)] \cdot \boldsymbol{z}^u(p,\alpha,\epsilon,\tau) \\
&\quad + [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau-t+p,\alpha)) \wedge \bar{\boldsymbol{x}}^u_{\alpha\tau}(\tau-t+p,\alpha)] \cdot \boldsymbol{z}^u(p,\alpha,\epsilon,\tau) \\
&\quad + [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau-t+p,\alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau-t+p,\alpha)] \cdot \boldsymbol{z}^u_\tau(p,\alpha,\epsilon,\tau).
\end{aligned} \tag{50}$$

We will build up simplifications for the many terms in (50). Since $\bar{\boldsymbol{x}}^u(\tau-t+p,\alpha)$ is a solution of unperturbed system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$, we have

$$\bar{\boldsymbol{x}}^u_\tau(\tau-t+p,\alpha) = \boldsymbol{f}\left(\bar{\boldsymbol{x}}^u(\tau-t+p,\alpha)\right), \tag{51}$$

Furthermore, since $\boldsymbol{x}^u(p,\alpha,\epsilon,\tau)$ is the solution of perturbed system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \epsilon\boldsymbol{g}(\boldsymbol{x},t)$, we can write

$$\begin{aligned}
\boldsymbol{x}^u_\tau(p,\alpha,\epsilon,\tau) &= \boldsymbol{f}(\boldsymbol{x}^u(p,\alpha,\epsilon,\tau)) + \epsilon\boldsymbol{g}(\boldsymbol{x}^u(p,\alpha,\epsilon,\tau),\tau) \\
&= \boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau-t+p,\alpha)) + \epsilon\ D\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau-t+p,\alpha))\boldsymbol{z}^u(p,\alpha,\epsilon,\tau) \\
&\quad + \frac{\epsilon^2}{2}\left[\boldsymbol{z}^u(p,\alpha,\epsilon,\tau)\right]^\top D^2\boldsymbol{f}(\boldsymbol{y}_1)\boldsymbol{z}^u(p,\alpha,\epsilon,\tau) \\
&\quad + \epsilon\left[\boldsymbol{g}(\bar{\boldsymbol{x}}^u(\tau-t+p,\alpha),\tau) + \epsilon\ D\boldsymbol{g}(\boldsymbol{y}_2,\tau)\ \boldsymbol{z}^u(p,\alpha,\epsilon,\tau)\right],
\end{aligned}$$

where we have applied Taylor's theorem to both $\boldsymbol{f}$ and $\boldsymbol{g}$ about the spatial value $\bar{\boldsymbol{x}}^u(\tau-t+p,\alpha)$ with deviation $\epsilon\boldsymbol{z}^u(p,\alpha,\epsilon,\tau)$. The notation D^2 represents the Hessian matrix, and the unknown values $\boldsymbol{y}_i = \boldsymbol{y}_i(p,\alpha,\epsilon,\tau)$ $(i=1,2)$ are located within ϵK of $\bar{\boldsymbol{x}}^u(\alpha,\tau-t+p)$. Thus

$$\begin{aligned}
\boldsymbol{z}^u_\tau(p,\alpha,\epsilon,\tau) &= \frac{1}{\epsilon}\left[\boldsymbol{x}^u_\tau(p,\alpha,\epsilon,\tau) - \bar{\boldsymbol{x}}^u_\tau(\tau-t+p,\alpha)\right] \\
&= D\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau-t+p,\alpha))\boldsymbol{z}^u(p,\alpha,\epsilon,\tau) + \boldsymbol{g}(\bar{\boldsymbol{x}}^u(\tau-t+p,\alpha),\tau) \\
&\quad + \epsilon\left[\frac{1}{2}\left[\boldsymbol{z}^u(p,\alpha,\epsilon,\tau)\right]^\top D^2\boldsymbol{f}(\boldsymbol{y}_1) + D\boldsymbol{g}(\boldsymbol{y}_2,\tau)\right]\left[\boldsymbol{z}^u(p,\alpha,\epsilon,\tau)\right].
\end{aligned}$$

Next, by the chain rule applied to (51),

$$\bar{\boldsymbol{x}}^u_{\alpha\tau}(\tau - t + p, \alpha) = D\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha))\ \bar{\boldsymbol{x}}^u_\alpha(\tau - t + p, \alpha)\,,$$

By substituting all these into (50), we get by separating out in orders of ϵ,

$$\begin{aligned}
\widetilde{M}^u_\tau(p, \alpha, \epsilon, \tau) &= [D\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha))\ \boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau - t + p, \alpha)] \cdot \boldsymbol{z}^u(p, \alpha, \epsilon, \tau) \\
&+ [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge D\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha))\ \bar{\boldsymbol{x}}^u_\alpha(\tau - t + p, \alpha)] \cdot \boldsymbol{z}^u(p, \alpha, \epsilon, \tau) \\
&+ [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau - t + p, \alpha)] \cdot [D\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha))\ \boldsymbol{z}^u(p, \alpha, \epsilon, \tau)] \\
&+ [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau - t + p, \alpha)] \cdot \boldsymbol{g}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha), \tau) \\
&+ \frac{\epsilon}{2}\ [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau - t + p, \alpha)] \cdot \Big[[\boldsymbol{z}^u(p, \alpha, \epsilon, \tau)]^\top\ D^2\boldsymbol{f}(\boldsymbol{y}_1)\,[\boldsymbol{z}^u(p, \alpha, \epsilon, \tau)]\Big] \\
&+ \epsilon\ [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau - t + p, \alpha)] \cdot [D\boldsymbol{g}(\boldsymbol{y}_2, \tau)\,[\boldsymbol{z}^u(p, \alpha, \epsilon, \tau)]]\ .
\end{aligned}$$

We now apply the identity given in Lemma 1 (Appendix B) to simplify the first three terms (given in the first three lines) above. By choosing $A = D\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\alpha, \tau - t + p))$, $\boldsymbol{b} = \boldsymbol{f}(\bar{\boldsymbol{x}}^u(\alpha, \tau - t + p))$, $\boldsymbol{c} = \bar{\boldsymbol{x}}^u_\alpha(\alpha, \tau - t + p)$ and $\boldsymbol{d} = \boldsymbol{z}^u(p, \alpha, \epsilon, \tau)$, $\widetilde{M}^u_\tau(p, \alpha, \epsilon, \tau)$ can be recast as

$$\begin{aligned}
\frac{\partial}{\partial\tau}\widetilde{M}^u(p, \alpha, \epsilon, \tau) &= \boldsymbol{\nabla} \cdot \boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha))\ \widetilde{M}^u(p, \alpha, \epsilon, \tau) \\
&\quad + [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau - t + p, \alpha)] \cdot \boldsymbol{g}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha), \tau) + \epsilon\, H(\tau)\,.
\end{aligned} \tag{52}$$

where

$$\begin{aligned}
H(\tau) := \frac{1}{2}\ &[\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau - t + p, \alpha)] \cdot \Big[[\boldsymbol{z}^u(p, \alpha, \epsilon, \tau)]^\top\ D^2\boldsymbol{f}(\boldsymbol{y}_1)\,[\boldsymbol{z}^u(p, \alpha, \epsilon, \tau)]\Big] \\
+\ &[\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau - t + p, \alpha)] \cdot [D\boldsymbol{g}(\boldsymbol{y}_2, \tau)\,[\boldsymbol{z}^u(p, \alpha, \epsilon, \tau)]]\ ,
\end{aligned} \tag{53}$$

and we use $\operatorname{Tr} D\boldsymbol{f} = \boldsymbol{\nabla} \cdot \boldsymbol{f}$ (i.e., the divergence of $\boldsymbol{f}$). The differential equation (52) is to be considered with the condition

$$\lim_{\tau\to\infty} \widetilde{M}^u(p, \alpha, \epsilon, \tau) = 0\,, \tag{54}$$

because from (48) we see that $\boldsymbol{f}\,(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \to \boldsymbol{f}(\boldsymbol{a}) = \boldsymbol{0}$ in this limit, with other terms in the definition remaining bounded.

We note that $H(\tau)$ is bounded for $\tau \in (-\infty, t]$ because of the boundedness of $D^2\boldsymbol{f}$, $D\boldsymbol{g}$, $\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)$ (because this converges to $\boldsymbol{a}$ as $\tau \to -\infty$) and $\boldsymbol{z}^u(p, \alpha, \epsilon, \tau)$. Hence it makes sense to also consider (52) with $\epsilon = 0$, i.e.,

$$\begin{aligned}
\frac{\partial}{\partial\tau}M^u(p, \alpha, \tau) &= \boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha))\ M^u(p, \alpha, \tau) \\
&\quad + [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau - t + p, \alpha)] \cdot \boldsymbol{g}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha), \tau)\,,
\end{aligned} \tag{55}$$

whose solution M^u is also subject to the boundary condition (54). It is easy to verify that the linear differential equation (55) with condition (54) has a unique solution

$$M^u(p, \alpha, t) = \int_{-\infty}^{t} e^{\int_\tau^t \nabla\cdot\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\xi - t + p, \alpha))\, d\xi}\, [\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}^u_\alpha(\tau - t + p, \alpha)] \cdot \boldsymbol{g}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha), \tau)\ \mathrm{d}\tau\,. \tag{56}$$

by working via the integrating factor

$$\mu(\tau) = \exp\left[-\int_0^\tau \nabla \cdot \boldsymbol{f}(\bar{\boldsymbol{x}}^u(\xi - t + p, \alpha))\, \mathrm{d}\xi\right]\,. \tag{57}$$

The quantity M^u in (56) is identical to the unstable Melnikov function as defined in (4) after a change of integration variable $\tau - t + p \to \tau$. We show in Appendix C that the improper integral (56) is convergent, and so M^u is a well-defined solution to (52) when $\epsilon = 0$.

Next, we show in Appendix D that the solution $\widetilde{M}^u(p, \alpha, \epsilon, t)$ of (52) is within $\mathcal{O}(\epsilon)$ of $M^u(p, \alpha, t)$. (This is not immediately obvious because integration over the noncompact domain $(-\infty, t]$ is necessary.) This enables the replacement of $\widetilde{M}^u(p, \alpha, \epsilon, t)$ with $M^u(p, \alpha, t) + \mathcal{O}(\epsilon)$ in (49), and consequently proves Theorem 1. In view of our previous remarks, this also establishes the legitimacy of viewing the part of $\Gamma^u_\epsilon(\boldsymbol{a}_\epsilon, t)$ that is close to $\Gamma^u(\boldsymbol{a})$ as being a graph with respect to the parameters (p, α).

B An important identity

We introduce the following elementary identity which is valid for any 3×3 matrix and any 3×1 vectors.

Lemma 1. *The following identity holds for any* 3×3 *matrix* A *and any* $\boldsymbol{b}, \boldsymbol{c}$ *and* $\boldsymbol{d}$ *which are* 3×1 *vectors:*

$$[(A\boldsymbol{b}) \wedge \boldsymbol{c}] \cdot \boldsymbol{d} + [\boldsymbol{b} \wedge (A\boldsymbol{c})] \cdot \boldsymbol{d} + [\boldsymbol{b} \wedge \boldsymbol{c}] \cdot (A\boldsymbol{d}) = \mathrm{Tr}\,(A)]\,[(\boldsymbol{b} \wedge \boldsymbol{c}) \cdot \boldsymbol{d}]\,, \tag{58}$$

where Tr (.) *represents the trace operator.*

Proof. This can be verified by a straightforward though tedious computation, having defined

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \quad \boldsymbol{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \text{ and } \quad \boldsymbol{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}.$$

The term $[(A\boldsymbol{b}) \wedge \boldsymbol{c}] \cdot \boldsymbol{d}$ can be written as

$$\begin{aligned}[(A\boldsymbol{b}) \wedge \boldsymbol{c}] \cdot \boldsymbol{d} = {} & d_1 \left[c_3 \left(a_{21}b_1 + a_{22}b_2 + a_{23}b_3\right) - c_2 \left(a_{31}b_1 + a_{32}b_2 + a_{33}b_3\right)\right] \\ & - d_2 \left[c_3 \left(a_{11}b_1 + a_{12}b_2 + a_{13}b_3\right) - c_1 \left(a_{31}b_1 + a_{32}b_2 + a_{33}b_3\right)\right] \\ & + d_3 \left[c_2 \left(a_{11}b_1 + a_{12}b_2 + a_{13}b_3\right) - c_1 \left(a_{21}b_1 + a_{22}b_2 + a_{23}b_3\right)\right],\end{aligned}$$

and the term $[\boldsymbol{b} \wedge (A\boldsymbol{c})] \cdot \boldsymbol{d}$ is simplified to

$$\begin{aligned}[\boldsymbol{b} \wedge (A\boldsymbol{c})] \cdot \boldsymbol{d} = {} & d_1 \left[b_2 \left(a_{31}c_1 + a_{32}c_2 + a_{33}c_3\right) - b_3 \left(a_{21}c_1 + a_{22}c_2 + a_{23}c_3\right)\right] \\ & - d_2 \left[b_1 \left(a_{31}c_1 + a_{32}c_2 + a_{33}c_3\right) - b_3 \left(a_{11}c_1 + a_{12}c_2 + a_{13}c_3\right)\right] \\ & + d_3 \left[b_1 \left(a_{21}c_1 + a_{22}c_2 + a_{23}c_3\right) - b_2 \left(a_{11}c_1 + a_{12}c_2 + a_{13}c_3\right)\right].\end{aligned}$$

Also, the term $[\boldsymbol{b} \wedge \boldsymbol{c}] \cdot (A\boldsymbol{d})$ can be simplified as

$$\begin{aligned}[\boldsymbol{b} \wedge \boldsymbol{c}] \cdot (A\boldsymbol{d}) = {} & \left(a_{11}d_1 + a_{12}d_2 + a_{13}d_3\right)\left(b_2c_3 - c_2b_3\right) \\ & - \left(a_{21}d_1 + a_{22}d_2 + a_{23}d_3\right)\left(b_1c_3 - c_1b_3\right) \\ & + \left(a_{31}d_1 + a_{32}d_2 + a_{33}d_3\right)\left(b_1c_2 - c_1b_2\right).\end{aligned}$$

So, the left hand side of the identity can be obtained by adding these three equations together, resulting in

$$\begin{aligned}& [(A\boldsymbol{b}) \wedge \boldsymbol{c}] \cdot \boldsymbol{d} \;+\; [\boldsymbol{b} \wedge (A\boldsymbol{c})] \cdot \boldsymbol{d} \;+\; [\boldsymbol{b} \wedge \boldsymbol{c}] \cdot (A\boldsymbol{d}) \\ & = \left(a_{11} + a_{22} + a_{33}\right)\left[d_1 \left(b_2c_3 - c_2b_3\right) - d_2 \left(b_1c_3 - c_1b_3\right) + d_3 \left(b_1c_2 - b_2c_1\right)\right].\end{aligned}$$

The term $(a_{11} + a_{22} + a_{33})$ is $\mathrm{Tr}\,(A)$, and the triple scalar product of $[(\boldsymbol{b} \wedge \boldsymbol{c}) \cdot \boldsymbol{d}]$ is equal to

$$[(\boldsymbol{b} \wedge \boldsymbol{c}) \cdot \boldsymbol{d}] = \left[d_1 \left(b_2c_3 - c_2b_3\right) - d_2 \left(b_1c_3 - c_1b_3\right) + d_3 \left(b_1c_2 - b_2c_1\right)\right],$$

which establishes the required result. □

C Convergence of the unstable Melnikov function M^u

Since $\bar{\boldsymbol{x}}^u$ is a trajectory on the two-dimension unstable manifold of $\boldsymbol{a}$, we know that we are in case 1, where $D\boldsymbol{f}(\boldsymbol{a})$ has two eigenvalues λ_1^u and λ_2^u with positive real part, and one eigenvalue $\lambda^s < 0$. Let is take the (potentially complex valued) eigenvectors $\boldsymbol{v}_1^u$ and $\boldsymbol{v}_2^u$. corresponding to $\lambda_{1,2}^u$ as being normalized. By assumption, $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are linearly independent; this also subsumes the situation of a repeated eigenvalue $\lambda_1^u = \lambda_2^u$ with geometric multiplicity 2. The eigenspace spanned by $\boldsymbol{v}_1^u$ and $\boldsymbol{v}_2^u$ forms the tangent plane to Γ^u at $\boldsymbol{a}$. The deviation of a trajectory from the point $\boldsymbol{a}$ is therefore governed by the linearized flow as $\tau \to -\infty$, i.e.,

$$\left|\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha) - \boldsymbol{a} - A(\alpha)\boldsymbol{v}_1^u e^{\lambda_1^u(\tau - t + p)} - B(\alpha)\boldsymbol{v}_2^u e^{\lambda_2^u(\tau - t + p)}\right| \to 0 \quad \text{as} \quad \tau \to -\infty\,,$$

where $A(\alpha)$ and $B(\alpha)$ are (potentially complex-valued) scalars which are differentiable in α. Thus, as τ approaches to negative infinity,

$$\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha) - \boldsymbol{a} \sim A(\alpha)\boldsymbol{v}_1^u e^{\lambda_1^u(\tau - t + p)} + B(\alpha)\boldsymbol{v}_2^u e^{\lambda_2^u(\tau - t + p)}\,,$$

Since $\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)$ satisfies the equation $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$, we know that

$$\bar{\boldsymbol{x}}_\tau^u(\tau - t + p, \alpha) = \boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)),$$

and so

$$\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \sim A(\alpha)\lambda_1^u\boldsymbol{v}_1^u e^{\lambda_1^u(\tau - t + p)} + B(\alpha)\lambda_2^u\boldsymbol{v}_2^u e^{\lambda_2^u(\tau - t + p)}.$$

Furthermore, the α-partial derivative is then

$$\bar{\boldsymbol{x}}_\alpha^u(\tau - t + p, \alpha) \sim A'(\alpha)\boldsymbol{v}_1^u e^{\lambda_1^u(\tau - t + p)} + B'(\alpha)\boldsymbol{v}_2^u e^{\lambda_2^u(\tau - t + p)}.$$

Lemma 2. *There exists a constant K_2 such that for all $(p, \alpha, \tau) \in (-\infty, P] \times [0, 2\pi) \times (-\infty, t]$, and for all $t \in (-\infty, T]$,*

$$e^{\int_\tau^t \nabla\cdot\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\xi - t + p, \alpha))d\xi}\left|\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}_\alpha^u(\tau - t + p, \alpha)\right| \le K_3 e^{\lambda^s(t-\tau)}\,. \tag{59}$$

Proof. Based on the previous estimates, we have

$$\begin{aligned}
&\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}_\alpha^u(\tau - t + p, \alpha)\\
&\sim \left[A(\alpha)\lambda_1^u\boldsymbol{v}_1^u e^{\lambda_1^u(\tau - t + p)} + B(\alpha)\lambda_2^u\boldsymbol{v}_2^u e^{\lambda_2^u(\tau - t + p)}\right] \wedge \left[A'(\alpha)\boldsymbol{v}_1^u e^{\lambda_1^u(\tau - t + p)} + B'(\alpha)\boldsymbol{v}_2^u e^{\lambda_2^u(\tau - t + p)}\right]\\
&= \left[A(\alpha)B'(\alpha)\lambda_1^u + A'(\alpha)B(\alpha)\lambda_2^u\right] e^{(\lambda_1^u + \lambda_2^u)(\tau - t + p)}\left[\boldsymbol{v}_1^u \wedge \boldsymbol{v}_2^u\right].
\end{aligned}$$

Now, we note that $|\boldsymbol{v}_1^u \wedge \boldsymbol{v}_2^u| \le 1$ (the eigenvectors are normalized), and the functions A and B and its derivatives are bounded on the compact set $\alpha \in 2\pi\,\mathrm{S}^1$. Since we must have $\operatorname{Im}\lambda_1^u = -\operatorname{Im}\lambda_2^u$, we obtain

$$\left|\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}_\alpha^u(\tau - t + p, \alpha)\right| \le K_1 e^{\operatorname{Re}(\lambda_1^u + \lambda_2^u)(\tau - t + p)}\,,$$

for some constant K_1. Moreover, since the trace of $D\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\alpha, \tau - t + p))$ approaches $\operatorname{Tr} D\boldsymbol{f}(\boldsymbol{a}) = \lambda_1^u + \lambda_2^u + \lambda^s = \operatorname{Re}(\lambda_1^u + \lambda_2^u) + \lambda^s$ as $\tau \to -\infty$, we have

$$e^{\int_\tau^t \nabla\cdot\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\xi - t + p, \alpha))d\xi} \sim e^{\int_\tau^t (\operatorname{Re}(\lambda_1^u + \lambda_2^u) + \lambda^s)\mathrm{d}\xi} = e^{(\operatorname{Re}(\lambda_1^u + \lambda_2^u) + \lambda^s)(t-\tau)}\,.$$

Consequently, the exponential term can be bounded by a constant K_2 times the term on the right. We can now estimate the product by

$$\begin{aligned}
e^{\int_\tau^t \nabla\cdot\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\xi - t + p, \alpha))d\xi}\left|\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\tau - t + p, \alpha)) \wedge \bar{\boldsymbol{x}}_\alpha^u(\tau - t + p, \alpha)\right| &\le K_1 e^{\operatorname{Re}(\lambda_1^u + \lambda_2^u)(\tau - t + p)} K_2 e^{(\operatorname{Re}(\lambda_1^u + \lambda_2^u) + \lambda^s)(t-\tau)}\\
&= K_1 K_2 e^{\operatorname{Re}(\lambda_1^u + \lambda_2^u)p} e^{\lambda^s(t-\tau)}\\
&= K_3 e^{\lambda^s(t-\tau)}\,,
\end{aligned}$$

for a constant K_3, as desired. □

Using the result of Lemma 2, since $\boldsymbol{z}^u(p, \alpha, \tau)$ is bounded (say by a constant K_4), from (56), we obtain the bound

$$|M^u(p, \alpha, t)| \le K_3 K_4 e^{\lambda^s t}\int_{-\infty}^t e^{-\lambda^s\tau}\,\mathrm{d}\tau = K_3 K_4 e^{\lambda^s t}\frac{e^{-\lambda^s\tau}}{-\lambda^s}\Big|_{-\infty}^t = \frac{K_3 K_4}{-\lambda^s}\,,$$

where the limit is convergent because $\lambda^s < 0$.

D Proof that $\widetilde{M}^u$ and M^u are $\mathcal{O}(\epsilon)$-close

Let $m(\tau) := \widetilde{M}^u(p,\alpha,\epsilon,\tau) - M^u(p,\alpha,\tau)$ be the difference in the two functions at a general time τ; we need to show that $m(t) = \mathcal{O}(\epsilon)$. Subtracting the equation (55) from (52), and multiplying by the integrating factor $\mu(\tau)$ we get

$$\frac{\partial}{\partial\tau}\left[\mu(\tau)m(\tau)\right] = \epsilon\mu(\tau)H(\tau)$$

subject to the condition $m(-\infty) = 0$. This has a solution

$$m(t) = \epsilon\int_{-\infty}^{t}\frac{\mu(\tau)}{\mu(t)}H(\tau)\,\mathrm{d}\tau = \epsilon\int_{-\infty}^{t}\exp\left[\int_{\tau}^{t}\nabla\cdot\boldsymbol{f}(\bar{\boldsymbol{x}}^u(\xi-t+p,\alpha))\mathrm{d}\xi\right]H(\tau)\,\mathrm{d}\tau\,.$$

Now, $H(\tau)$ in (53) can be factored: one term consists of exactly the left-hand side of (59), whereas the remainder of the terms are bounded because of the boundedness of $\boldsymbol{z}^u$ (as argued in Appendix A), and of $D^2\boldsymbol{f}$ and $D\boldsymbol{g}$ (by hypothesis). Applying Lemma 2, we therefore obtain

$$|m(t)| \le \epsilon K_5\int_{-\infty}^{t}e^{\lambda^s(t-\tau)}\,\mathrm{d}\tau = \frac{\epsilon K_5}{-\lambda^s}\,.$$

for some constant K_5. Hence, $m(t) = \mathcal{O}(\epsilon)$ as desired.

E Proof of Theorem 3 (Heteroclinic manifold splitting)

For fixed (p,α,t) in the relevant domains, we know that d^u in Theorem 1 provides the displacement of $\Gamma^u_\epsilon(\boldsymbol{a}_\epsilon)$ from $\bar{\boldsymbol{x}}(p,\alpha)$ in the direction normal to Γ, and similarly, d^s in Theorem 2 the displacement of $\Gamma^s_\epsilon(\boldsymbol{b}_\epsilon)$ in the same direction. Since $\bar{\boldsymbol{x}} = \bar{\boldsymbol{x}}^u = \bar{\boldsymbol{x}}^s$ in this instance,

$$\begin{aligned}
d(p,\alpha,t,\epsilon &= d^u(p,\alpha,t,\epsilon) - d^s(p,\alpha,t,\epsilon)\\
&= \epsilon\frac{M^u(p,\alpha,t)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha))\wedge\bar{\boldsymbol{x}}_\alpha(p,\alpha)|} - \epsilon\frac{M^s(p,\alpha,t)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha))\wedge\bar{\boldsymbol{x}}_\alpha(p,\alpha)|} + \mathcal{O}(\epsilon^2)\\
&= \epsilon\frac{M^u(p,\alpha,t) - M^s(p,\alpha,t)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha))\wedge\bar{\boldsymbol{x}}_\alpha(p,\alpha)|} + \mathcal{O}(\epsilon^2)\\
&=: \epsilon\frac{M(p,\alpha,t)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha))\wedge\bar{\boldsymbol{x}}_\alpha(p,\alpha)|} + \mathcal{O}(\epsilon^2)\,,
\end{aligned}$$

where from (4) and (8), we get

$$\begin{aligned}
M(p,\alpha,t) &= M^u(p,\alpha,t) - M^s(p,\alpha,t),\\
&= \int_{-\infty}^{p}\exp\left[\int_{\tau}^{p}\nabla\cdot\boldsymbol{f}(\bar{\boldsymbol{x}}(\xi,\alpha))d\xi\right]\left[\boldsymbol{f}(\bar{\boldsymbol{x}}(\tau,\alpha))\wedge\bar{\boldsymbol{x}}_\alpha(\tau,\alpha)\right]\cdot\boldsymbol{g}(\bar{\boldsymbol{x}}(\tau,\alpha),\tau+t-p)\;\mathrm{d}\tau\\
&- \left(-\int_{p}^{\infty}\exp\left[\int_{\tau}^{p}\nabla\cdot\boldsymbol{f}(\bar{\boldsymbol{x}}(\xi,\alpha))d\xi\right]\left[\boldsymbol{f}(\bar{\boldsymbol{x}}(\tau,\alpha))\wedge\bar{\boldsymbol{x}}_\alpha(\tau,\alpha)\right]\cdot\boldsymbol{g}(\bar{\boldsymbol{x}}(\tau,\alpha),\tau+t-p)\;\mathrm{d}\tau\right)\\
&= \int_{-\infty}^{\infty}\exp\left[\int_{\tau}^{p}\nabla\cdot\boldsymbol{f}(\bar{\boldsymbol{x}}(\alpha,\xi))d\xi\right]\left[\boldsymbol{f}(\bar{\boldsymbol{x}}(\alpha,\tau))\wedge\bar{\boldsymbol{x}}_\alpha(\alpha,\tau)\right]\cdot\boldsymbol{g}(\bar{\boldsymbol{x}}(\alpha,\tau),\tau+t-p)\;\mathrm{d}\tau\,,
\end{aligned}$$

as desired.

F Proof of Theorem 4 (Lobe volume)

We note that there is a nearby region, $R^\star$, such that $d(p,\alpha,\epsilon,t)$ sign-definite on $R^\star$, and moreover $R^\star$'s boundary is $Q^\star$, which consists of closed curves which are $\mathcal{O}(\epsilon)$-close to Q. While the lobe volume should

properly be calculated by integrating d over $R^\star$, the error in integrating ϵM over R instead is of higher-order in ϵ. Consequently, the leading-order lobe volume only requires leading-order information.

Since Γ is (p,α)-parametrized by $\bar{\boldsymbol{x}}(p,\alpha)$, we can write the vector surface element on Γ by

$$\boldsymbol{dS} = \bar{\boldsymbol{x}}_p(p,\alpha) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)\,\mathrm{d}p\,\mathrm{d}\alpha = \boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)\,\mathrm{d}p\,\mathrm{d}\alpha$$

However, we know that the signed distance between the perturbed stable and unstable manifolds, measured perpendicular to Γ at $\bar{\boldsymbol{x}}(p,\alpha)$, is given by d in (15). Noting moreover that using R rather than $R^\star$ results in a higher-order error, and d itself is $\mathcal{O}(\epsilon)$, we can write the volume of the lobe lying between the manifolds as

$$\begin{aligned}
\text{Lobe volume} &= \iint_{R^\star} |d(p,\alpha,\epsilon,t)|\;|\boldsymbol{dS}| \\
&= \iint_{R} |d(p,\alpha,\epsilon,t)|\;|\boldsymbol{dS}| + \mathcal{O}(\epsilon^2) \\
&= \iint_{R} \left|\epsilon \frac{M(p,\alpha,t)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)|} + \mathcal{O}(\epsilon^2)\right| \;|\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)|\;\mathrm{d}p\,\mathrm{d}\alpha + \mathcal{O}(\epsilon^2)\,,
\end{aligned}$$

which immediately gives the desired result.

G Proof of Theorem 5 (Instantaneous flux)

The pseudo-separatrix consists of three different segments. There is *no* Lagrangian flux across the stable and unstable manifold parts, because these are invariant objects. They move with time, but remain material surfaces. The only flux that can occur is that crossing the strip. We note from Fig. 8 that in parts of the strip where the unstable manifold is outside the stable one, the flux will be outwards, and hence will be positive. Thus, in (p,α,t) regions in which the Melnikov function is positive, a positive contribution to the flux occurs. Conversely, if the unstable manifold is inside the stable one, the flux is into the closed surface and hence negative, again consonant with the sign of the Melnikov function at such points.

A general point $\boldsymbol{r}$ on the strip S, as given in (21), is

$$\begin{aligned}
\boldsymbol{r}(s,\alpha,\epsilon,t) &= \bar{\boldsymbol{x}}(p,\alpha) + \hat{\boldsymbol{n}}(p,\alpha)\left[s d^u(p,\alpha,\epsilon,t) + (1-s) d^s(p,\alpha,\epsilon,t)\right] \quad ; \quad (s,\alpha) \in [0,1] \times \mathrm{S}^1 \\
&= \bar{\boldsymbol{x}}(p,\alpha) + \epsilon \frac{\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)}{|\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)|} \frac{s M^u(p,\alpha,t) + (1-s) M^s(p,\alpha,t)}{|\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)|} + \mathcal{O}(\epsilon^2)\,.
\end{aligned}$$

The vector surface element on S in terms of the (s,α)-parametrization, chosen so that the outward normal is positive, is therefore

$$\begin{aligned}
\mathbf{d}\boldsymbol{S} &= \boldsymbol{r}_\alpha(s,\alpha,\epsilon,t) \wedge \boldsymbol{r}_s(s,\alpha,\epsilon,t)\,\mathrm{d}s\,\mathrm{d}\alpha \\
&= \left[\bar{\boldsymbol{x}}_\alpha(p,\alpha) + \mathcal{O}(\epsilon)\right] \wedge \epsilon \frac{\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)}{\left|\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)\right|^2} \left[M^u(p,\alpha,t) - M^s(p,\alpha,t)\right]\,\mathrm{d}s\,\mathrm{d}\alpha + \mathcal{O}(\epsilon^2) \\
&= \epsilon \frac{M(p,\alpha,t)\,\bar{\boldsymbol{x}}_\alpha(p,\alpha) \wedge \left[\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)\right]}{\left|\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)\right|^2}\,\mathrm{d}s\,\mathrm{d}\alpha + \mathcal{O}(\epsilon^2).
\end{aligned}$$

Therefore, the Lagrangian flux crossing S is

$$\begin{aligned}
\Phi(p,t,\epsilon) &= \iint_S \left[\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)+\epsilon\boldsymbol{g}\left(\bar{\boldsymbol{x}}(p,\alpha),t\right)\right]\cdot \mathbf{d}\boldsymbol{S} \\
&= \int_0^1\int_0^1 \left[\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)+\epsilon\boldsymbol{g}\left(\bar{\boldsymbol{x}}(p,\alpha),t\right)\right]\cdot \epsilon\frac{M(p,\alpha,t)\,\bar{\boldsymbol{x}}_\alpha(p,\alpha)\wedge\left[\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right)\wedge\bar{\boldsymbol{x}}_\alpha(p,\alpha)\right]}{\left|\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right)\wedge\bar{\boldsymbol{x}}_\alpha(p,\alpha)\right|^2}\,\mathrm{d}s\,\mathrm{d}\alpha+\mathcal{O}(\epsilon^2) \\
&= \epsilon\int_0^1\int_0^1 M(p,\alpha,t)\frac{\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)\cdot\left[\bar{\boldsymbol{x}}_\alpha(p,\alpha)\wedge\left[\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right)\wedge\bar{\boldsymbol{x}}_\alpha(p,\alpha)\right]\right]}{\left|\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right)\wedge\bar{\boldsymbol{x}}_\alpha(p,\alpha)\right|^2}\,\mathrm{d}s\,\mathrm{d}\alpha+\mathcal{O}(\epsilon^2) \\
&= \epsilon\int_0^1\int_0^1 M(p,\alpha,t)\frac{\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)\cdot\left[\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha))\left[\bar{\boldsymbol{x}}_\alpha(p,\alpha)\cdot\bar{\boldsymbol{x}}_\alpha(p,\alpha)\right]-\bar{\boldsymbol{x}}_\alpha(p,\alpha)\left[\boldsymbol{f}(\bar{x}(p,\alpha))\cdot\bar{\boldsymbol{x}}_\alpha(p,\alpha)\right]\right]}{\left|\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right)\wedge\bar{\boldsymbol{x}}_\alpha(p,\alpha)\right|^2}\,\mathrm{d}s\,\mathrm{d}\alpha+\mathcal{O}(\epsilon^2) \\
&= \epsilon\int_0^1\int_0^1 M(p,\alpha,t)\frac{\left|\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)\right|^2\left|\bar{\boldsymbol{x}}_\alpha(p,\alpha)\right|^2-\left|\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)\cdot\bar{\boldsymbol{x}}_\alpha(p,\alpha)\right|^2}{\left|\boldsymbol{f}\left(\bar{\boldsymbol{x}}(p,\alpha)\right)\wedge\bar{\boldsymbol{x}}_\alpha(p,\alpha)\right|^2}\,\mathrm{d}s\,\mathrm{d}\alpha+\mathcal{O}(\epsilon^2) \\
&= \epsilon\int_0^1\int_0^1 M(p,\alpha,t)\,\mathrm{d}s\,\mathrm{d}\alpha+\mathcal{O}(\epsilon^2) \\
&= \epsilon\int_0^1 M(p,\alpha,t)\,\mathrm{d}\alpha+\mathcal{O}(\epsilon^2)\,,
\end{aligned}$$

as required. In this derivation, we have used standard vector identities in three-dimensions: the 'bac-cab' rule and Lagrange's identity.

H Proof of Corollary 1 (Instantaneous flux for harmonic perturbations)

In this proof, we will use the shorthand notation

$$F(p,\alpha) := \mathcal{F}\left\{h\left(p,\alpha,\boldsymbol{.}\right)\right\}(\omega)\,.$$

Inserting the expression for the time-harmonic Melnikov function (24) into the instantaneous flux formula (22), and performing standard trigonometric manipulations, we get

$$
\begin{aligned}
\Phi(p,t,\epsilon) &= \epsilon \int_0^1 |F(p,\alpha)| \cos\left[\omega(t-p)+\phi - \arg\left(F(p,\alpha)\right)\right] \,\mathrm{d}\alpha + \mathcal{O}(\epsilon^2) \\
&= \epsilon \Bigg\{ \cos\left[\omega(t-p)+\phi\right] \int_0^1 |F(p,\alpha)| \cos\left[\arg\left(F(p,\alpha)\right)\right] \,\mathrm{d}\alpha \\
&\qquad + \sin\left[\omega(t-p)+\phi\right] \int_0^1 |F(p,\alpha)| \sin\left[\arg\left(F(p,\alpha)\right)\right] \,\mathrm{d}\alpha \Bigg\} + \mathcal{O}(\epsilon^2) \\
&= \epsilon \left\{ \cos\left[\omega(t-p)+\phi\right] \int_0^1 \mathrm{Re}\left(F(p,\alpha)\right) \,\mathrm{d}\alpha + \sin\left[\omega(t-p)+\phi\right] \int_0^1 \mathrm{Im}\left(F(p,\alpha)\right) \,\mathrm{d}\alpha \right\} + \mathcal{O}(\epsilon^2) \\
&= \epsilon \sqrt{\left(\int_0^1 \mathrm{Re}\left(F(p,\alpha)\right) \,\mathrm{d}\alpha\right)^2 + \left(\int_0^1 \mathrm{Im}\left(F(p,\alpha)\right) \,\mathrm{d}\alpha\right)^2} \Bigg\{ \frac{\cos\left[\omega(t-p)+\phi\right] \int_0^1 \mathrm{Re}\left(F(p,\alpha)\right) \,\mathrm{d}\alpha}{\sqrt{\left(\int_0^1 \mathrm{Re}\left(F(p,\alpha)\right) \,\mathrm{d}\alpha\right)^2 + \left(\int_0^1 \mathrm{Im}\left(F(p,\alpha)\right) \,\mathrm{d}\alpha\right)^2}} \\
&\qquad + \frac{\sin\left[\omega(t-p)+\phi\right] \int_0^1 \mathrm{Im}\left(F(p,\alpha)\right) \,\mathrm{d}\alpha}{\sqrt{\left(\int_0^1 \mathrm{Re}\left(F(p,\alpha)\right) \,\mathrm{d}\alpha\right)^2 + \left(\int_0^1 \mathrm{Im}\left(F(p,\alpha)\right) \,\mathrm{d}\alpha\right)^2}} \Bigg\} + \mathcal{O}(\epsilon^2) \\
&= \epsilon \left| \int_0^1 F(p,\alpha) \,\mathrm{d}\alpha \right| \Bigg\{ \cos\left[\omega(t-p)+\phi\right] \frac{\mathrm{Re}\left(\int_0^1 F(p,\alpha) \,\mathrm{d}\alpha\right)}{\left|\int_0^1 F(p,\alpha) \,\mathrm{d}\alpha\right|} \\
&\qquad + \sin\left[\omega(t-p)+\phi\right] \frac{\mathrm{Im}\left(\int_0^1 F(p,\alpha) \,\mathrm{d}\alpha\right)}{\left|\int_0^1 F(p,\alpha) \,\mathrm{d}\alpha\right|} \Bigg\} + \mathcal{O}(\epsilon^2) \\
&= \epsilon \left| \int_0^1 F(p,\alpha) \,\mathrm{d}\alpha \right| \Bigg\{ \cos\left[\omega(t-p)+\phi\right] \cos\left[\arg\left(\int_0^1 F(p,\alpha) \,\mathrm{d}\alpha\right)\right] \\
&\qquad + \sin\left[\omega(t-p)+\phi\right] \sin\left[\arg\left(\int_0^1 F(p,\alpha) \mathrm{d}\alpha\right)\right] \Bigg\} + \mathcal{O}(\epsilon^2) \\
&= \epsilon \left| \int_0^1 F(p,\alpha) \,\mathrm{d}\alpha \right| \cos\left[\omega(t-p)+\phi - \arg\left(\int_0^1 F(p,\alpha) \,\mathrm{d}\alpha\right)\right] + \mathcal{O}(\epsilon^2) \,,
\end{aligned}
$$

as required.

I Proof of Theorem 7 (Lobe volume for harmonic perturbations in the volume-preserving situation)

Consider any one of the lobes $L_{ik}(t)$. By Theorem 4, its volume to leading-order in ϵ is given by

$$
\mathrm{Volume}\left(L_{ik}(t)\right) = \epsilon \int_{\alpha_{i-1}}^{\alpha_i} \int_{\tilde{p}(\alpha,k-1)}^{\tilde{p}(\alpha,k)} |M(p,\alpha,t)| \,\mathrm{d}p \,\mathrm{d}\alpha + \mathcal{O}(\epsilon^2) \,.
$$

Employing (24), and under the harmonic assumption in which h is independent of p, we get

$$
\begin{aligned}
\mathrm{Volume}\left(L_{ik}(t)\right) &= \epsilon \int_{\alpha_{i-1}}^{\alpha_i} \int_{\tilde{p}(\alpha,k-1)}^{\tilde{p}(\alpha,k)} \left|\mathcal{F}\left\{h(\alpha,\boldsymbol{.})\right\}(\omega)\right| \left|\cos\left[\omega\left(t-p\right)+\phi - \arg\left(\mathcal{F}\left\{h(\alpha,\boldsymbol{.})\right\}(\omega)\right)\right]\right| \,\mathrm{d}p \,\mathrm{d}\alpha + \mathcal{O}(\epsilon^2) \\
&= \epsilon \int_{\alpha_{i-1}}^{\alpha_i} \left|\mathcal{F}\left\{h(\alpha,\boldsymbol{.})\right\}(\omega)\right| \int_{\tilde{p}(\alpha,k-1)}^{\tilde{p}(\alpha,k)} \left|\cos\left[\omega\left(t-p\right)+\phi - \arg\left(\mathcal{F}\left\{h(\alpha,\boldsymbol{.})\right\}(\omega)\right)\right]\right| \,\mathrm{d}p \,\mathrm{d}\alpha + \mathcal{O}(\epsilon^2)
\end{aligned}
$$

The inner p-integral is between adjacent zeros of the cosine function, specifically as given by the functions $\tilde{p}$. Note that this is of the absolute value of the cosine function of $-\omega p$ plus a phase shift. The phase shift does not affect the integral because it is between adjacent zeros of p. Thus, we can simply shift the integral to be between any two adjacent zeros, and discard the entire phase shift $\omega t + \phi - \arg\left(\mathcal{F}\left\{h(\alpha, \boldsymbol{.})\right\}(\omega)\right)$. We choose ωp to be between $-\pi/2$ and $\pi/2$, i.e., p between $-\pi/(2\omega)$ and $\pi/(2\omega)$. Thus,

$$\begin{aligned}\text{Volume}\left(L_{ik}(t)\right) &= \epsilon \int_{\alpha_{i-1}}^{\alpha_i} \left|\mathcal{F}\left\{h(\alpha, \boldsymbol{.})\right\}(\omega)\right| \int_{-\pi/(2\omega)}^{\pi/(2\omega)} \cos\left[-\omega p\right] \, \mathrm{d}p \, \mathrm{d}\alpha + \mathcal{O}(\epsilon^2) \\ &= \frac{2\epsilon}{\omega} \int_{\alpha_{i-1}}^{\alpha_i} \left|\mathcal{F}\left\{h(\alpha, \boldsymbol{.})\right\}(\omega)\right| \, \mathrm{d}\alpha + \mathcal{O}(\epsilon^2)\,,\end{aligned}$$

whose leading-order term in ϵ is independent of k.

J Derivations related to Hill's spherical vortex

This section outlines the derivation of the results on the manifold displacements, Melnikov function and instantaneous flux for the Hill's spherical vortex outlined in Sec. 4.1. We note that p represents the time-variation along a heteroclinic trajectory $\bar{\boldsymbol{x}}(p, \alpha)$, and is thus functionally related to θ, the latitude coordinate. Each heteroclinic trajectory is given in (r, θ, ϕ)-coordinates as $(1, \bar{\theta}(p), \phi)$, where $\bar{\theta}$ can be found via the velocity along a longitude:

$$\frac{d\bar{\theta}}{dp} = \frac{3}{2}\sin(\bar{\theta}) \quad , \quad \text{and so} \quad \bar{\theta}(p) = \cos^{-1}\left(-\tanh\frac{3p}{2}\right)\,, \tag{60}$$

where we have chosen $\bar{\theta}(0) = \pi/2$, i.e., $p = 0$ at the equator for every heteroclinic trajectory, and this form of inverse trigonometric function gives the principal branch $\theta \in [0, \pi]$ as required. The general heteroclinic-trajectory parameterization of Γ (in (r, θ, ϕ) form) is therefore

$$\bar{\boldsymbol{x}}(p, \alpha) = \left(1, \cos^{-1}\left(-\tanh\frac{3p}{2}\right), 2\pi\alpha\right)\,,$$

in (r, θ, ϕ) components. Given that $\sin(\bar{\theta}(p)) = \operatorname{sech}\left(\frac{3p}{2}\right)$, we note that

$$\boldsymbol{f}(1, \bar{\theta}(p), \phi) = \frac{3}{2}\sin(\bar{\theta}(p))\hat{\boldsymbol{\theta}} = \frac{3}{2}\operatorname{sech}\frac{3p}{2}\,\hat{\boldsymbol{\theta}}\,.$$

Next, the derivative of $\bar{\boldsymbol{x}}$ with respect to α is quantified by

$$\bar{\boldsymbol{x}}_\alpha(p, \alpha) = (1\sin\bar{\theta}(p))\frac{\partial(2\pi\alpha)}{\partial\alpha}\hat{\boldsymbol{\phi}} = 2\pi\operatorname{sech}\frac{3p}{2}\,\hat{\boldsymbol{\phi}}\,,$$

and so

$$\boldsymbol{f}(1, \bar{\theta}(p), \phi) \wedge \bar{\boldsymbol{x}}_\alpha(p, \alpha) = 3\pi\operatorname{sech}^2\frac{3p}{2}\,\hat{\boldsymbol{r}}\,.$$

A general perturbation $\boldsymbol{g}$ in (1) would be expressible in (r, θ, ϕ)-coordinates as

$$\boldsymbol{g}(r, \theta, \phi, t) = g_r(r, \theta, \phi, t)\hat{\boldsymbol{r}} + g_\theta(r, \theta, \phi, t)\hat{\boldsymbol{\theta}} + g_\phi(r, \theta, \phi, t)\hat{\boldsymbol{\phi}}\,.$$

We need neither specify that $\boldsymbol{g}$ be volume-preserving, nor time-periodic; only that the components of $\boldsymbol{g}$ have bounded derivatives. When $\epsilon \neq 0$ but is small, $\boldsymbol{a}$ perturbs to a time-varying hyperbolic trajectory $\boldsymbol{a}_\epsilon(t)$, retaining its unstable manifold $\Gamma^u_\epsilon(\boldsymbol{a}_\epsilon, t)$ which remains close to Γ. Now, its location is associated with the unstable Melnikov function, which using Theorem 1 depends only on one component of $\boldsymbol{g}$:

$$M^u\left(p, \alpha, t\right) = 3\pi\int_{-\infty}^{p} \operatorname{sech}^2\frac{3\tau}{2} g_r\left(1, \cos^{-1}\left(-\tanh\frac{3\tau}{2}\right), 2\pi\alpha, \tau + t - p\right)\mathrm{d}\tau\,.$$

By virtue of Remark 1, this means that the part of Γ_ϵ^u close to Γ can be approximately parameterized by

$$\boldsymbol{r}^u(p,\alpha,\epsilon,t) \approx \bar{\boldsymbol{x}}(p,\alpha) + \epsilon\, M^u(p,\alpha,t)\frac{\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)}{|\boldsymbol{f}(\bar{\boldsymbol{x}}(p,\alpha)) \wedge \bar{\boldsymbol{x}}_\alpha(p,\alpha)|^2}$$

whence (37) results. We highlight that is only the r-component of $\boldsymbol{g}$ which contributes to the leading-order displacement of Γ in the direction normal to it. An analogous approach can be used to obtain that $\boldsymbol{b}_\epsilon(t)$'s stable manifold $\Gamma_\epsilon^s(\boldsymbol{b}_\epsilon, t)$ is approximately parameterizable via (38). The Melnikov function (39) arises as a simple consequence of (16) in Thoerem 3. The instantaneous flux (from the inside to the outside of the sphere) generated by the perturbation is then (40), computed from (22.

References

[1] D. Arrowsmith and C. Place. An introduction to dynamical systems. *University of Cambridge Press, Cambridge, UK*, 1990.

[2] S. Balasuriya. Direct chaotic flux quantification in perturbed planar flows: general time-periodicity. *SIAM J. Appl. Dyn. Sys.*, 4:282–311, 2005.

[3] S. Balasuriya. Optimal perturbation for enhanced chaotic transport. *Phys. D*, 202:155–176, 2005.

[4] S. Balasuriya. Cross-separatrix flux in time-aperiodic and time-impulsive flows. *Nonlinearity*, 19:2775–2795, 2006.

[5] S. Balasuriya. Optimal frequency for microfluidic mixing across a fluid interface. *Phys. Rev. Lett*, 105:064501, 2010.

[6] S. Balasuriya. A tangential displacement theory for locating perturbed saddles and their manifolds. *SIAM J. Appl. Dyn. Sys.*, 10:1100–1126, 2011.

[7] S. Balasuriya. *Barriers and transport in unsteady flows: A Melnikov approach*. Mathematical Modeling and Computation. SIAM Press, Philadelphia, 2016.

[8] S. Balasuriya. Impulsive perturbations to differential equations: stable/unstable pseudo-manifolds, heteroclinic connections, and flux. *Nonlinearity*, 29:3897–3933, 2016.

[9] S. Balasuriya. Local stable and unstable manifolds and their control in nonautonomous finite-time flows. *J. Nonlinear Sci.*, 26:895–927, 2016.

[10] S. Balasuriya. Meridional and zonal wavenumber dependence in tracer flux in Rossby waves. *Fluids*, 1:1–30, 2016.

[11] S. Balasuriya and M. Finn. Energy constrained transport maximization across a fluid interface. *Phys. Rev. Lett*, 108:244503, 2012.

[12] S. Balasuriya, R. Kalampattel, and N. Ouellette. Hyperbolic neighbourhoods as organizers of finite-time exponential stretching. *Journal of Fluid Mechanics*, 807:509–545, 11 2016.

[13] S. Balasuriya, I. Mezić, and C.K.R.T. Jones. Weak finite-time Melnikov theory and 3D viscous perturbations of Euler flows. *Phys. D*, 176:82–106, 2003.

[14] S. Balasuriya, N. Ouellette, and I. Rypina. Generalized Lagrangian coherent structures. *Physica D*, 372:31–51, 2018.

[15] F. Battelli and M. Feckan. Homoclinic trajectories in discontinuous systems. *J Dyn Diff Equations*, 20:337–376, 2008.

[16] F. Battelli and M. Feckan. Melnikov theory for nonlinear implicit ODEs. *Melnikov theory for nonlinear implicit ODEs*, 256:1157–1190, 2014.

[17] F. Battelli and C. Lazzari. Exponential dichotomies, heteroclinic orbits and Melnikov functions. *J. Differential Equations*, 86:342–366, 1986.

[18] D. Beigie, A. Leonard, and S. Wiggins. Chaotic transport in the homoclinic and heteroclinic tangle regions of quasiperiodically forced two-dimensional systems. *Nonlinearity*, 4:775–815, 1991.

[19] A. Bertozzi. Heteroclinic orbits and chaotic dynamics in planar fluid flows. *SIAM J. Math. Anal.*, 19:1271–1294, 1988.

[20] A. Calamai and M. Franca. Mel'nikov methods and homoclinic orbits in discontinuous systems. *J. Dyn. Diff. Equat.*, 25:733–764, 2013.

[21] S-N. Chow, J. K. Hale, and J. Mallet-Paret. An example of bifurcation to homoclinic orbits. *J. Differential Equations*, 37:351–373, 1980.

[22] S.-N. Chow and M. Yamashita. Geometry of the Melnikov vector. In *Mathematics in science and engineering*, volume 185, pages 79–148. Elsevier, 1992.

[23] W. Coppel. *Dichotomies in Stability Theory.* Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1978.

[24] F. Costal and J. Rodriguez. A bifurcation problem to homoclinic orbits for non-autonomous systems. *J. Math. Anal. Appl.*, 105:395–404, 1985.

[25] Z. Du and W. Zhang. Melnikov method for homoclinic bifurcations in nonlinear impact oscillators. *Computers Math. Appl.*, 50:445–458, 2005.

[26] M. Gidea and R. de la Llave. Global Melnikov theory in Hamiltonian systems with general time-dependent perturbations. *J. Nonlin. Sci*, 28:1657–1707, 2018.

[27] M. Gidea, R. de la Llave, and M. Musser. Global effect of non-conservative perturbations on homoclinic orbits. *Qualitative theory of dynamical systems*, 20:1–40, 2021.

[28] J. Gruendler. The existence of homoclinic orbits and the method of Melnikov for systems in $\mathbb{R}^n$. *SIAM J. Math. Anal.*, 16:907–931, 1985.

[29] J. Gruendler. Homoclinic solutions and chaos in ordinary differential equations with singular perturbations. *Trans. Amer. Math. Soc.*, 350:3797–3814, 1998.

[30] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields.* 1983.

[31] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields.* Springer, New York, 1983.

[32] M. J. M. Hill. On a spherical vortex. *Phil. Trans. Roy. Soc. A*, 185:213–245, 1894.

[33] P. Holmes. Some remarks on chaotic particle paths in time-periodic, three-dimensional swirling flows. *Contemp. Math.*, 28:393–404, 1984.

[34] N. Ju, D. Small, and S. Wiggins. Existence and computation of Hyperbolic Trajectories of aperiodically time dependent vector fields and their approximations. *International Journal of Bifurcation and Chaos*, 13(6):1449 – 1457, 6 2003.

[35] P. Kukucka. Mel'nikov method for discontinuous planar systems. *Nonlin. Anal.*, 66:2698–2719, 2007.

[36] S. Leibovich. The structure of vortex breakdown. *Annu. Rev. Fluid Mech.*, 10:221–246, 1978.

[37] X.-B. Lin. Using Melnikov's method to solve Silnikov's problems. *Proc. R. Soc Edin. Sec. A Math.*, 116A:295–325, 1990.

[38] H. Lomelí and J. Meiss. Heteroclinic primary intersections and codimension one Melnikov method for volume-preserving maps. *Chaos*, 10:109–121, 2000.

[39] H. Lomelí and J. Meiss. Heteroclinic intersections between invariant circles of volume-preserving maps. *Nonlinearity*, 16:1573–1595, 2003.

[40] H. Lomelí, J. Meiss, and R. Ramírez-Ros. Canonical Melnikov theory for diffeomorphisms. *Nonlinearity*, 21:485–508, 2008.

[41] J. Lopez and A. Perry. Axisymmetric vortex breakdown. part 3 onset of periodic flow and chaotic advection. *J. Fluid Mech.*, 234:449–471, 1992.

[42] R. Mackay, J. Meiss, and I. Percival. Transport in Hamiltonian systems. *Physica D*, 13:55–81, 1984.

[43] B. Maelfeyt, S. Smith, and K. Mitchell. Using invariant manifolds to construct symbolic dynamics for three-dimensional volume-preserving maps. *SIAM J. Appl. Dyn. Sys.*, 16:729–769, 2017.

[44] J. Meiss. Thirty years of turnstiles and transport. *Chaos*, 25:097602, 2015.

[45] V. Melnikov. On the stability of the centre for time-periodic perturbations. *Trans. Moscow Math. Soc.*, 12:1–56, 1963.

[46] I. Mezić. Chaotic advection in bounded Navier–Stokes flows. *J. Fluid. Mech.*, 431:347–370, 2001.

[47] I. Mezić and S. Wiggins. On the integrability and perturbation of three-dimensional fluid flows with symmetry. *J. Nonlin. Sci.*, 4:157–194, 1994.

[48] K. Mitchell. The topology of nested homoclinic and heteroclinic tangles. *Physica D*, 238:737–763, 2009.

[49] K. Mitchell and J. Delos. A new topological technique for characterizing homoclinic tangles. *Physica D*, 171:170–187, 2006.

[50] B. Mosovsky, M. Speetjens, and J. Meiss. Finite-time transport in volume-preserving flows. *Phys. Rev. Lett.*, 110:214101, 2013.

[51] P. Mullowney, K. Julien, and K. Meiss. Blinking rolls: chaotic advection in three-dimensional flows with an invariant. *SIAM J. Appl. Dyn. Sys.*, 4:159–186, 2005.

[52] H. Oka. Singular perturbations of autonomous ordinary differential equations and heteroclinic bufurcations. In R. Bamon, R. Labarca, J. Lewowicz, and J. Palis, editors, *Dynamical Systems, Santiago de Chile 1990: Pitman Research Notes in Mathematics*, volume 285, pages 159–194, 1993.

[53] K. Palmer. Exponential dichotomies and transversal homoclinic points. *J. Differential Equations*, 55:225–256, 1984.

[54] K. G. D. S. Priyankara, S. Balasuriya, and E. M. Bollt. Quantifying the role of folding in nonautonomous flows: The unsteady double-gyre. *International Journal of Bifurcation and Chaos*, 27, 2017.

[55] V. Rom-Kedar. Homoclinic tangles–classification and applications. *Nonlinearity*, 7:441–473, 1994.

[56] V. Rom-Kedar. Secondary homoclinic bifurcation theorems. *Chaos*, 5:385–401, 1995.

[57] V. Rom-Kedar, A. Leonard, and S. Wiggins. An analytical study of transport, mixing and chaos in an unsteady vortical flow. *J. Fluid Mech.*, 214:347–394, 1990.

[58] V. Rom-Kedar and A. Poje. Universal properties of chaotic transport in the presence of diffusion. *Phys. Fluids*, 11(8):2044–2057, 1999.

[59] B. Sandstede, S. Balasuriya, C.K.R.T. Jones, and P.D. Miller. Melnikov theory for finite-time vector fields. *Nonlinearity*, 13:1357–1377, 2000.

[60] S. Sattari and K. Mitchell. Using periodic orbits to compute chaotic transport rates between resonance zones. *Chaos*, 27:113104, 2017.

[61] M. Scase and H. L. Terry. Hill's spherical vortex in a rotating fluid. *Fluid Dynamics*, 2018.

[62] J. Shen, K. Lu, and W. Zhang. Heteroclinic chaotic behavior driven by a Brownian motion. *J. Differential Equations*, 255:4185–4225, 2013.

[63] F. Sotiropoulos, Y. Ventikos, and T. Lackey. Chaotic advection in three-dimensional stationary vortex-breakdown bubbles: Silnikov's chaos and the devil's staircase. *J. Fluid Mech.*, 444:257–297, 2001.

[64] A. Vanderbauhede. Bifurcation of degenerate homoclinics. *Results Math.*, 21:211–223, 1992.

[65] M. Wechselberger. Extending Melnikov theory to invariant manifolds on non-compact domains. *Dyn. Sys.*, 17:215–233, 2002.

[66] S. Wiggins. *Chaotic Transport in Dynamical Systems*. Springer-Verlag, 1992.

[67] K. Yagasaki. The method of Melnikov for perturbations of multi-degree-of-freedom Hamiltonian systems. *Nolinearity*, 12:799–822, 1999.

[68] K. Yagasaki. Higher-order Melnikov method and chaos for two-degree-of-freedom Hamiltonian systems with saddle-centers. *Discrete Cont. Dyn. Sys.*, 29:387–402, 2011.

[69] K. Yagasaki. Melnikov processes and chaos in randomly perturbed dynamical systems. *Nonlinearity*, 31:3057–3085, 2018.

[70] S. Yamashita. Melnikov vector in higher dimensions. *Nonlinear Analysis*, 18:657–670, 1992.

[71] Y. Yi. A generalized integral manifold theorem. *J. Differential Equations*, 102:153–187, 1993.

[72] H. Zang, M. Han, and D. Xiao. On Melnikov of a homoclinic loop through a nilpotent saddle for planar near-Hamiltonian systems. *J. Differential Equations*, 245:1086–1111, 2008.