Nonlinearity **19** (2006) 2775–2795 doi:10.1088/0951-7715/19/12/003

Cross-separatrix flux in time-aperiodic and time-impulsive flows

Sanjeeva Balasuriya

Department of Mathematics, Connecticut College, New London, CT 06320, USA

E-mail: sanjeeva.balasuriya@conncoll.edu

Received 12 July 2006, in final form 2 October 2006
Published 27 October 2006
Online at stacks.iop.org/Non/19/2775

Recommended by K Ohkitani

Abstract

A theory for the fluid flux generated across heteroclinic separatrices under the influence of time-aperiodic perturbations is presented. The flux is explicitly defined as the amount of fluid transferred per unit time, and its detailed time-dependence monitored. The perturbations are allowed to be significantly discontinuous in time, including for example impulsive (Dirac delta type) discontinuities. The flux is characterized in terms of time-varying separatrices, with easily computable formulae (directly related to Melnikov functions) provided.

Mathematics Subject Classification: 34C37, 76R99, 34A37

(Some figures in this article are in colour only in the electronic version)

1. Introduction

This article addresses the quantification, explicitly as the transfer of fluid per unit time, across heteroclinic separatrices in two-dimensional incompressible flows under the influence of perturbations. This is an old problem for time-*periodic* perturbations, where Rom-Kedar and her collaborators have made substantial contributions. The basic approach is to define a Poincaré map and characterize the lobe dynamics via a 'turnstile' mechanism [1, 2]. The relationship of the lobe area to the Melnikov function [2, 3] can then be used to compute lobe areas, which can be considered representative values for flux quantification. Instantaneous flux values can then be imputed by dividing the lobe areas by the periodicity of the perturbation [4,5]. This transport mechanism in time-periodic instances is usually chaotic (in the sense of symbolic dynamics and the Smale–Birkhoff theorem [2]). A downside to this approach is that the time variation of the transport is not captured and that there is no obvious way to extend to time-aperiodic flows.

It has however been argued that the standard Melnikov approach [2,6,7] for establishing the relative positions of perturbed manifolds is not limited to time-periodic perturbations [8–13].

0951-7715/06/122775+21$30.00

For perturbations which are 'nearly' periodic in time (quasiperiodic or almost periodic), suitably modified versions of the Poincaré map have been successfully used [8–10,12,14]. The conclusions in these articles are mainly geared towards establishing chaos in the sense of symbolic dynamics, and not quantifying the flux in fluid flows. There are other studies (for example, [15–17]) which utilize the Melnikov theory to obtain a *qualitative* description of the transport occurring in some instances. However, theoretical methods for *quantitatively* establishing flux measurements are often limited to time-periodic flows [3–5,13,18–20], with genuinely time-aperiodic flux apparently only addressed in [21] (see also [22,23] for numerical calculations). The emphasis in [21] is however on the interaction between the Eulerian and Lagrangian viewpoints, with numerical flux calculations presented based on their results. The current study presents an approach based on which flux across separatrices, explicitly as a transfer of fluid volume per unit time, can be computed for time-aperiodically perturbed flows. Defining an appropriate curve across which the flux computation makes sense is an important first step. A flux *function* is then defined, which characterizes the flux variation with time. The emphasis here is not restricted to *chaotic* fluid flux but on obtaining a description of the time variation of fluid flow across a separatrix which demarcates different fluid regimes. All these ideas can be trivially extended to the quantification of the flux of any passive scalar.

To the author's knowledge, no attempt has been made to quantify flux resulting from time-*discontinuous* perturbations of flows. This may include jump discontinuities in time (unit-step type functions), such as when a certain forcing is turned on at a given time. More serious are perturbations involving *impulses* (Dirac delta type perturbations). Consider for example, a microfluidic device being tapped or vigorously shaken every so often, with non-uniform tapping/shaking protocols in space, which may also be different at each tapping/shaking time. Such an instantaneous change in the boundary position leads to an impulse in the velocity field, which will then affect the flow inside the device. What sort of mixing can occur because of this? This sort of question is important, since the low Reynolds number in such flows impedes good mixing, and at present the suggestions to improve mixing in such devices are mainly *ad hoc*. Conceivably, impulsive perturbations may also be relevant in situations such as determining the fluid responses to underwater earthquakes or explosions, in which the sea floor instantaneously rises. This would cause the flux of passive scalars (for example, heat or nutrient concentration) across a separatrix (say, the boundary of a warm eddy in the colder waters south of the Gulf Stream), for which a quantification would be desirable.

Significant difficulties appear when attempting to build a theory for the flux for such time-discontinuous or time-impulsive perturbations. What sort of solutions exist? *Do* solutions exist? Can manifolds be defined? Across what surfaces does it make sense to compute the flux? Does the Melnikov theory extend to such cases? This study addresses these questions, thereby obtaining a theory for flux quantification for time-aperiodic or time-impulsive perturbations. A formula for the flux function is obtained in all such instances, and its computability illustrated through an example.

2. Melnikov transform

The flows to be considered take the form

$$\dot{x} = J\nabla H(x) + \varepsilon g(x, t), \tag{1}$$

where $x \in \Omega \subset \mathbb{R}^2$, $t \in \mathbb{R}$, J is the symplectic matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $H : \Omega \to \mathbb{R}$ is the Hamiltonian function, and ∇H its gradient in Ω. The quantity ε satisfies $0 < \varepsilon \ll 1$, and $g : \Omega \times \mathbb{R} \to \mathbb{R}^2$ is the perturbing velocity field, which will be taken to have several different smoothness conditions in this article. The $\varepsilon = 0$ flow, however, will be consistently assumed to satisfy the following conditions.

Hypothesis 2.1. The unperturbed ($\varepsilon = 0$) flow has the following properties:

(U1) $H \in \mathrm{C}^3(\Omega)$;
(U2) The flow (1) with $\varepsilon = 0$ possesses two hyperbolic fixed points a and b (which need not be distinct);
(U3) These fixed points each possess one-dimensional stable and unstable manifolds;
(U4) A branch of the unstable manifold of a (denoted W_a^u) coincides with a branch of the stable manifold of b (W_b^s) to form a one-dimensional heteroclinic manifold Γ. This manifold consists of one heteroclinic trajectory $\bar{x}(t)$, $t \in \mathbb{R}$.

Lemma 2.1. *There exist positive constants* α *and* K *such that*

$$|\nabla H(\bar{x}(t))| \leqslant K \mathrm{e}^{-\alpha|t|}$$

uniformly for $t \in \mathbb{R}$.

Proof. This is a consequence of the assumptions on hyperbolicity of the fixed points a and b of the unperturbed flow. This could be more clearly identified in terms of exponential dichotomies [11, 24], from which this lemma is inevitable. □

The perturbing velocity field g is *not* assumed to be time-periodic (the standard hypothesis in a lobe dynamics approach to flux). General time-dependence, with sufficient smoothness and boundedness, is enough. Both smoothness and boundedness will be subsequently relaxed, but initially g will be assumed to be an element of the function space $\mathcal{P}$ defined as follows.

Definition 2.1. The function space $\mathcal{P}$ consists of functions $f(x, t)$ defined from $\Omega \times \mathbb{R}$ to $\mathbb{R}^2$ which satisfy the following two conditions:

(P1) $f(\cdot, t) \in \mathrm{C}^2(\Omega)$ for all $t \in \mathbb{R}$;
(P2) $f(x, \cdot) \in \mathrm{C}^0(\mathbb{R}) \cap \mathrm{L}^\infty(\mathbb{R})$ for all $x \in \Omega$.

As pointed out in a variety of sources, $g \in \mathcal{P}$ is sufficient to ensure the persistence of a perturbed hyperbolic trajectory $(a_\varepsilon(t), t)$ in $\Omega \times \mathbb{R}$, with its attendant manifolds $W_a^u(\varepsilon)$ and $W_a^s(\varepsilon)$ [8–10, 12–14]. A similar result works for $(b_\varepsilon(t), t)$ and its manifolds. Time-periodicity is not a necessary ingredient for this persistence.

Suppose a t-parametrization has been chosen for the unperturbed heteroclinic trajectory $\bar{x}(t)$. In Ω, construct a fixed transversal $\mathcal{T}$ to the flow at $\bar{x}(0)$. This is clearly in the direction $\nabla H(\bar{x}(0))$, since the flow lies along level curves of the Hamiltonian. Now consider the perturbed flow (1) in $\Omega \times \mathbb{R}$. For small enough ε, in any t-slice, the manifolds $W_a^u(\varepsilon)$ $W_b^s(\varepsilon)$ will therefore intersect $\mathcal{T}$; suppose these intersection points are, respectively, $h_\varepsilon^u(t)$ and $h_\varepsilon^s(t)$. Standard Melnikov results [2, 6, 7], along with comments on the applicability in time-aperiodic instances [8–12], show that for $g \in \mathcal{P}$, the distance between the manifolds is given by

$$d(t, \varepsilon) := (h_\varepsilon^u(t) - h_\varepsilon^s(t)) \cdot \frac{\nabla H(\bar{x}(0))}{|\nabla H(\bar{x}(0))|} = \varepsilon \frac{M(t)}{|\nabla H(\bar{x}(0))|} + \mathcal{O}(\varepsilon^2), \tag{2}$$

where the Melnikov function is defined by

$$M(t) = \int_{-\infty}^{\infty} \nabla H(\bar{x}(\tau)) \cdot g(\bar{x}(\tau), t + \tau) \, \mathrm{d}\tau. \tag{3}$$

Motivated by this, as were Meyer and Sell [9] when they analysed almost-periodic flows, the following definition is presented.

Definition 2.2. The *Melnikov transform* Φ is defined for $f \in \mathcal{P}$ by

$$\Phi\{f\}(t) := \int_{-\infty}^{\infty} \nabla H(\bar{x}(\tau)) \cdot f(\bar{x}(\tau), t+\tau)\,\mathrm{d}\tau$$
$$:= \int_{-\infty}^{\infty} |\nabla H(\bar{x}(\tau))| f^{\perp}(\bar{x}(\tau), t+\tau)\mathrm{d}\tau,$$

where $f^{\perp} := f \cdot \nabla H / |\nabla H|$ is f's component in the direction of ∇H.

It turns out that this transform has a strong connection to the cross-separatrix flux—more intimate than has been previously stated in the literature—which will be detailed in section 3. An interesting observation is that this transform can be formally computed for *distributions* and not just functions in $\mathcal{P}$. This will be taken advantage of in section 5 for distribution-like perturbations, with a justification of the Melnikov analysis leading to (2) and (3). For the moment, however, some trivial properties of this transform will be stated, which will be useful in detailing the smoothness properties of the flux function defined in section 3. In the following, the notation $\|\cdot\|_p$ will be used for L^p-norms, where $p \in [1, \infty]$, and $\mathrm{W}^{k,p}$ will denote the standard Sobolev spaces.

Lemma 2.2. *For* $f \in \mathcal{P}$, $\Phi\{f\} \in \mathrm{L}^{\infty}(\mathbb{R})$.

Proof. Since $f(x, \cdot) \in \mathrm{L}^{\infty}(\mathbb{R})$ for all $x \in \Omega$, and since the closure of Γ, $\Gamma_c := \Gamma \cup \{a\} \cup \{b\}$ is a compact subset of Ω, f is uniformly bounded on $\Gamma_c \times \mathbb{R}$ (except possibly on a measure zero set of $t \in \mathbb{R}$). By also applying lemma 2.1,

$$\|\Phi\{f\}\|_{\infty} \leqslant K_1 \int_{-\infty}^{\infty} |\nabla H(\bar{x}(\tau))|\,\mathrm{d}\tau \leqslant K_2 \int_{-\infty}^{\infty} \mathrm{e}^{-\alpha|t|}\,\mathrm{d}\tau \leqslant K_3$$

for constants K_1, K_2, K_3 and α. □

Lemma 2.3. *If* $f \in \mathcal{P}$ *and there exists a* $p \in [1, \infty)$ *such that* $f(x, \cdot) \in \mathrm{L}^p(\mathbb{R})$ *for all* $x \in \Omega$, *then* $\Phi\{f\} \in \mathrm{L}^p(\mathbb{R})$.

Proof. Since $\bar{x}(t)$ for $t \in \mathbb{R}$ is contained in Γ_c, the closure of Γ, there exists a function $\tilde{f} : \mathbb{R} \to \mathbb{R}^2$ such that $\tilde{f} \in \mathrm{L}^p(\mathbb{R})$ and

$$|f^{\perp}(\bar{x}(\tau), t)| \leqslant |\tilde{f}(t)|$$

uniformly in τ. Now

$$\|\Phi\{f\}\|_p \leqslant \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\nabla H(\bar{x}(\tau))|^p |f^{\perp}(\bar{x}(\tau), t+\tau)|^p\,\mathrm{d}\tau\,\mathrm{d}t\right)^{1/p}$$
$$\leqslant \left(\int_{-\infty}^{\infty} |\nabla H(\bar{x}(\tau))|^p \left(\int_{-\infty}^{\infty} |\tilde{f}(t+\tau\right)|^p\,\mathrm{d}t)\,\mathrm{d}\tau\right)^{1/p}$$
$$\leqslant \|\tilde{f}\|_p \|\nabla H\|_p.$$

Now, the exponential decay rates of ∇H from lemma 2.1 coupled with H's smoothness ensures that $\nabla H(\bar{x}(\cdot)) \in \mathrm{L}^p(\mathbb{R})$ for any $p \in [1, \infty]$, and the result follows. □

Lemma 2.4. *If there exists a* $p \geqslant 1$ *and* $k \in \{0, 1, 2, \ldots\}$ *such that*

$$\frac{\partial^k f}{\partial t^k}(x, \cdot) \in \mathcal{P} \cap L^p(\mathbb{R})$$

for all $x \in \Omega$, *then* $\Phi\{f\} \in \mathrm{W}^{k,p}(\mathbb{R})$.

Proof. Apply lemma 2.3 to the function $(\partial^k f)/(\partial t^k)$. □

Corollary 2.1. *If the conditions of lemma 2.4 are satisfied and moreover $kp > 1$, then $\Phi\{f\} \in C^0(\mathbb{R})$.*

Proof. This is a direct consequence of the strong $(1 - kp < 0)$ version of the Sobolev embedding theorem [25]. □

Lemma 2.5. *If $f \in \mathcal{P}$ and if there exists a $k \in \mathbb{N}$ such that $f(x, \cdot) \in W^{k,\infty}(\mathbb{R})$ for all $x \in \Omega$, then $\Phi\{f\} \in C^k(\mathbb{R})$.*

Proof. This is an immediate consequence of the dominated convergence theorem, permitting

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k}\Phi\{f\} = \int_{-\infty}^{\infty} \nabla H(\bar{x}(\tau)) \cdot \frac{\partial^k}{\partial t^k} f(\bar{x}(\tau), t + \tau)\, \mathrm{d}\tau,$$

and the right-hand side (RHS) is finite since $(\partial^k f)/(\partial t^k)$ is bounded (except possibly in a zero measure set), and ∇H has exponential decay by lemma 2.1. □

3. Flux

The distance (2) will be now used to rationalize an expression for the cross-separatrix flux in terms of the Melnikov transform. Firstly, a comment on the choice of parametrization of $\bar{x}(t)$ is in order. Suppose an alternative time-parametrization $\bar{x}_\beta(t)$ is chosen, in which

$$\bar{x}_\beta(t) = \bar{x}(t - \beta) \tag{4}$$

for some $\beta \in \mathbb{R}$, where $\bar{x}(t)$ is the original parametrization used in (2). The new transversal, $\mathcal{T}_\beta$, is that drawn at $\bar{x}_\beta(0)$, and the appropriate distance measurement along this transversal is given by (2) with $\bar{x}$ replaced with $\bar{x}_\beta$ and $M(t)$ replaced by

$$\begin{aligned} M_\beta(t) &= \int_{-\infty}^{\infty} \nabla H(\bar{x}_\beta(\tau)) \cdot g(\bar{x}_\beta(\tau), t + \tau)\, \mathrm{d}\tau \\ &= \int_{-\infty}^{\infty} \nabla H(\bar{x}(\tau - \beta)) \cdot g(\bar{x}(\tau - \beta), t + \tau)\, \mathrm{d}\tau \\ &= \int_{-\infty}^{\infty} \nabla H(\bar{x}(\tau)) \cdot g(\bar{x}(\tau), t + \tau + \beta)\, \mathrm{d}\tau \\ &= \Phi\{g\}(t + \beta). \end{aligned}$$

The following is then an obvious consequence of (2).

Lemma 3.1. *Consider a choice of time-parametrization for the unperturbed heteroclinic trajectory encoded through β in (4). Then the signed distance along the transversal $\mathcal{T}_\beta$ measured in a time-slice t is given by*

$$d(t, \beta, \varepsilon) = \varepsilon \frac{\Phi\{g\}(t + \beta)}{|\nabla H(\bar{x}_\beta(0))|} + \mathcal{O}(\varepsilon^2).$$

Remark 3.1. Thus, if $\Phi\{g\}(t + \beta)$ has a simple zero with respect to t at a value $t = t_0$, this implies the existence of a transverse intersection between the perturbed manifolds in a nearby time-slice, somewhere near $\mathcal{T}_\beta$, for sufficiently small ε.

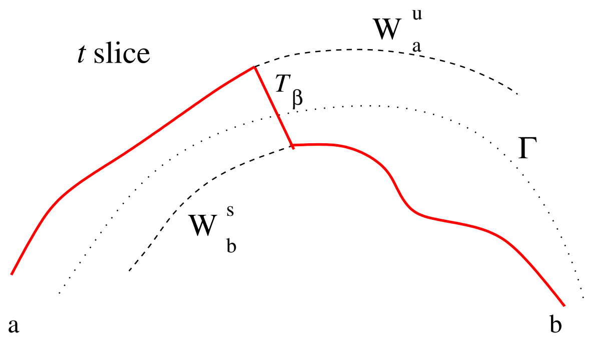

Figure 1. Definition of the separatrix $\Gamma(t, \beta, \varepsilon)$ (heavy curve), in a t-slice. The dotted line shows the position of the unperturbed separatrix Γ.

Remark 3.2. The topological intersection pattern can therefore be found by determining the location of simple zeros of $\Phi\{g\}(t)$, for *any* chosen time-parametrization of $\bar{x}$. There is a one-to-one correspondence between these and the zeros for any choice of β.

Remark 3.3. Note that $d(t, \beta, \varepsilon)$ would be positive if the vector drawn from the stable manifold $W_b^s(\varepsilon)$ to the unstable manifold $W_a^u(\varepsilon)$ along the transversal $\mathcal{T}_\beta$ drawn at the point $\bar{x}_\beta(0)$ in the time-slice t is in the same direction as $\nabla H(\bar{x}_\beta(0))$. It would be negative if in the opposite direction.

The main issue in this section is to use lemma 3.1 to determine the flux resulting from the inclusion of the perturbation in (1). While there is no transport across Γ when $\varepsilon = 0$, transport generically does occur when $\varepsilon \neq 0$. Quantifying this in general is difficult, primarily because of the ambiguity in defining a curve across which transport is to be assessed in the time-dependent case. Lacking the advantage of time-periodicity which generates a natural Poincaré map [2, 3], it is moreover not possible to directly quantify the amount of fluid transferred per unit time. While there is no real difficulty in extending the Melnikov approach to determine the distance between perturbed manifolds under time-aperiodicity [8–12], flux quantification is not obvious. Numerical attempts to do so (for example [23]) tend to compute lobe areas in time-slices, but their specific relationship to the flux is unclear.

The key to this approach is identifying a prospective flow separatrix in the time-dependent flow. The method used here is very close to the ‘gate’ suggested in [21], based on which numerical flux calculations were done in [22]; nevertheless, the ability to impute a powerful flux formula (theorem 3.1) does not seem to have been appreciated. Consider picking time-slice t in the full $\Omega \times \mathbb{R}$ phase space of (1). Form the ‘separatrix’ here by considering the part of $W_a^u(\varepsilon)$ from $a_\varepsilon(t)$ to where it intersects the transversal $\mathcal{T}_\beta$ in the first instance, the part of $W_b^s(\varepsilon)$ from $b_\varepsilon(t)$ to where it similarly intersects $\mathcal{T}_\beta$ and the part of $\mathcal{T}_\beta$ which lies between these two intersection points (figure 1). This pseudo-separatrix shall be defined to be $\Gamma(t, \beta, \varepsilon)$. See figure 1. Now consider the *flux*, explicitly a transfer of fluid area per unit time across this pseudo-separatrix, at an instantaneous time value t. Observe the following.

(i) The area $\mathrm{d}A$ which crosses $\mathcal{T}_\beta$ in an infinitesimal time increment $\mathrm{d}t$ is given by

$$\begin{aligned}\mathrm{d}A &= d(t, \beta, \varepsilon)(|\nabla H(\bar{x}_\beta(0))| + \mathcal{O}(\varepsilon))\,\mathrm{d}t \\ &= (\varepsilon \Phi\{g\}(t + \beta) + \mathcal{O}(\varepsilon^2))\,\mathrm{d}t.\end{aligned}$$

The first equality is since the speed at which fluid is crossing $\mathcal{T}_\beta$ is $|\nabla H(\bar{x}_\beta(0))|$ (correct to $\mathcal{O}(\varepsilon)$). The second is through lemma 3.1.

(ii) There is a directionality in the above; $\mathrm{d}A$ may be negative. By remark 3.3, it can be seen that if $\mathrm{d}A$ is positive, that corresponds to fluid travelling from what was initially a lower H region to a higher H region (see also figure 1). The opposite would happen if $\mathrm{d}A$ were negative. In other words, positivity of $\Phi\{g\}(t+\beta)$ represents an instantaneous fluid transfer across Γ in the 'direction' of ∇H, whereas negativity implies flux in the opposite direction.

(iii) If $\Phi\{g\}(t+\beta)$ has a zero at some t-value, this means that any contributions to $\mathrm{d}A$ are at most $\mathcal{O}(\varepsilon^2)$ in this t-slice.

(iv) There is no transfer of fluid across the part of $\Gamma(t, \beta, \varepsilon)$ composed of portions of the stable and unstable manifolds. This pseudo-separatrix is therefore a natural way of demarcating a time-dependent boundary. Note that this does not mean that fluid on these manifolds does not have a non-tangential component of motion; it may. However, the definition of the pseudo-separatrix is such that it is time-dependent; as one progresses to a different time-slice (say $t + \mathrm{d}t$), $\Gamma(t, \beta, \varepsilon)$ changes appropriately.

Since a flux is defined to be a transfer of fluid area per unit time, it can be *defined* to be the t-dependent function (and also dependent on β and ε):

$$\mathrm{Flux}(t, \beta, \varepsilon) := \frac{\mathrm{d}A}{\mathrm{d}t}. \tag{5}$$

Remark 3.4. This rationalization of the flux is not limited to small ε, just as standard lobe dynamics approaches are not necessarily perturbative [1–3]. For any ε, suppose the flow (1) possesses hyperbolic trajectories $a_\varepsilon(t)$ and $b_\varepsilon(t)$ in $\Omega \times \mathbb{R}$, with appropriate two-dimensional manifolds $W_a^u(\varepsilon)$ and $W_b^s(\varepsilon)$ existing globally. Let $\mathcal{T}$ be a fixed plane parallel to the t-axis which intersects branches of these manifolds in every t-slice. These entities serve to define a pseudo-separatrix, and the flux across this is given exactly as in (5), with the flux arising only through fluid crossing the transversal $\mathcal{T}$.

Remark 3.5. While it is necessary for the perturbation to be sufficiently smooth in t to ensure the presence of *manifolds*, sections 4 and 5 show how this flux definition is extendable to perturbations *non-smooth* in t, for example those which have unit-step functions or Dirac delta impulses in t.

Remark 3.6. Thinking of the rate of change of area as a mixing quantifier has similarly been used in the context of 'adiabatic' (i.e. slowly varying) time-periodic flows [13, 20]. In these situations, one is assessing the rate of change of area enclosed by a 'frozen-time' *homo* clinic separatrix.

Within the current perturbative setting of $0 < \varepsilon \ll 1$ for smooth aperiodic perturbations, an expression for the leading-order flux has now been established.

Theorem 3.1. *The instantaneous flux across the pseudo-separatrix* $\Gamma(t, \beta, \varepsilon)$ *is given by*

$$\mathrm{Flux}(t, \beta, \varepsilon) = \varepsilon \Lambda_\beta(t) + \mathcal{O}(\varepsilon^2)$$

where the leading-order flux function *is defined by*

$$\Lambda_\beta(t) := \Phi\{g\}(t+\beta) = M_\beta(t)\,.$$

Remark 3.7. This means that the Melnikov function in fact provides a much more direct link to the flux than the results in [1–3] suggest. It is not an *integrated* version of the Melnikov function which quantifies the leading-order flux, but it is the Melnikov function itself!

Remark 3.8. Even for periodic flows, this is a much more natural interpretation of the flux than is used in lobe dynamics approaches, which provide a constant flux measure while not detailing its temporal variation [2, 3, 5]. Nevertheless the results in [5, 18] show that such flux measures correspond exactly to the amplitude of the time-periodic flux function that would arise through theorem 3.1.

Remark 3.9. A direct relationship between the rate of change of an area and the (adiabatic) Melnikov function was also noticed by [20] in adiabatic time-periodic flows. The area defined by [20] was formed by considering how frozen-time homoclinic separatrices evolve, in contrast to a direct flux definition across pseudo-separatrices. Nevertheless, the relationship between the Melnikov function and a rate of change of area was argued in [20], and exploited also in [13].

Remark 3.10. To obtain the basic form of the leading-order flux function, it is enough to simply consider *any* time-parametrization on $\bar{x}$. Any other time-parametrization (encoded through β) corresponds exactly to a direct β-shift on $\Lambda_\beta(t)$.

Remark 3.11. Suppose it is required instead to determine the flux of a passive scalar across Γ. If this has concentration $c(x)$ in the unperturbed flow, it is a trivial observation that the $\mathcal{O}(\varepsilon)$ term of the passive scalar flux is therefore

$$c(\bar{x}_\beta(0))\Lambda_\beta(t) = c(\bar{x}_\beta(0))\phi\{g\}(t+\beta) = c(\bar{x}_\beta(0))M_\beta(t).$$

This is an instantaneous passive scalar flux, given as a function of time, and its sign gives the *direction* of flux transfer. This time-varying function therefore possesses detailed information on how the passive scalar crosses the pseudo-separatrix. For example, one might obtain the *net* transfer in the direction ∇H by integrating this over all time.

Remark 3.12. The form of the flux function here has a close relationship to an intuitive physical flux definition. To see this, note that the arclength parametrization ℓ of Γ can be related to a t-parametrization through $\mathrm{d}\ell = |\nabla H(\bar{x}(t))|\,\mathrm{d}t$, and let $t(\ell)$ be their relationship. Now, since $\varepsilon g^\perp$ represents the only velocity component perpendicular to Γ, a direct (and simple-minded) physical definition of the flux across Γ at time τ might be written as

$$\text{Flux-direct}(\tau) := \int_\Gamma \varepsilon g^\perp(\bar{x}(t(\ell)),\tau)\,\mathrm{d}\ell = \varepsilon\int_{-\infty}^{\infty} g^\perp(\bar{x}(t),\tau)\,|\nabla H(\bar{x}(t))|\,\mathrm{d}t. \tag{6}$$

While this directly expresses the flux across Γ, it does not provide a good assessment of the flux transfer across a *separatrix*, since Γ is no longer one. Indeed, the two regions of different characteristics separated by Γ when $\varepsilon = 0$ now have boundaries which are difficult to demarcate without the concept of a pseudo-separatrix. Nevertheless, the leading-order term in the 'flux-direct' function above is very close to the Melnikov function in (2.2); the only difference is the necessity of replacing the τ in the temporal argument of $g^\perp$ with $\tau + t$. In the event that g has slow temporal variation, the simple-minded flux (6) therefore does in fact approach the true flux of theorem 3.1. (Parenthetically, (6) represents precisely the adiabatic Melnikov function related to considering 'frozen-time' [13, 20].) This illustrates the consistency of the flux definition argued in theorem 3.1; such connections to (6) are not obvious in alternative flux representations.

The properties of the previous section can now be applied to characterize the smoothness of the flux function Λ as follows. Its β-dependence will not be explicitly stated unless this trivial (by remark 3.10) shift is important to the discussion.

Proposition 3.1. *Consider the flow (1) subject to hypothesis 2.1, and in which $g \in \mathcal{P}$. The corresponding leading-order flux function has the following properties:*

(a) $\Lambda \in L^\infty(\mathbb{R})$*;*
(b) If there is a $p \in [1, \infty)$ *such that* $g(x, \cdot) \in \mathrm{L}^p(\mathbb{R})$ *for all* $x \in \Omega$*, then* $\Lambda \in \mathrm{L}^p(\mathbb{R})$*;*
(c) If there exists $k \in \{0, 1, 2, \ldots\}$ *and* $p \in [1, \infty)$ *such that* $(\partial^k g)/(\partial t^k)$ *for all* $x \in \Omega$*, then* $\Lambda \in \mathrm{W}^{k,p}(\mathbb{R})$*;*
(d) If the conditions of (c) above hold and moreover $kp > 1$*, then* $\Lambda \in C^0(\mathbb{R})$*;*
(e) If there exists a $k \in \mathbb{N}$ *such that* $g(x, \cdot) \in W^{k,\infty}(\mathbb{R})$ *for all* $x \in \Omega$*, then* $\Lambda \in \mathrm{C}^k(\mathbb{R})$*;*

Proof. These are immediate consequences of lemmas 2.2–2.5 and corollary 2.1 from section 2. □

Remark 3.13. A necessary condition for chaos (*vis-à-vis* shift dynamics and the Smale–Birkhoff theorem) to ensue through the $\mathcal{O}(\varepsilon)$ intersection of perturbed manifolds is that Λ has infinitely many zeros [14]. Such usually occurs under perturbations which are time-periodic [1,2,7], and also under mild variations such as quasi-periodicity or almost periodicity [8, 10, 12].

4. Jump discontinuities

Consider now the case where the perturbation g has finitely many jump discontinuities in time. Such would occur naturally in fluid flows in which, say, a perturbing flow is turned on and off at various instances in time, such as in the 'egg-beater flow' which has been proposed as a good mixing strategy (see for example [26]). However, 'manifolds' no longer exist in the traditional sense, since smoothness is compromised. One might not therefore be able to make sense of a flux in any usual way. Nevertheless, the flux theory of section 3 can be extended easily to account for such jump discontinuities. Suppose that g belongs to $\mathcal{P}_J$ defined as follows.

Definition 4.1. The function space $\mathcal{P}_J \subset \mathcal{P}$ (where $\mathcal{P}$ is as in definition 2.1), is the set of functions f satisfying the following additional condition:

(P3) There exists a set $\mathcal{J} = \{t_1, t_2, \cdots, t_n\}$ such that $f(x, \cdot) \in \mathrm{C}^0(\mathbb{R} \setminus \mathcal{J}) \cap \mathrm{L}^\infty(\mathbb{R} \setminus \mathcal{J})$ for all $x \in \Omega$, and moreover

$$\lim_{t \to ti^-} f(x, t) \qquad \text{and} \qquad \lim_{t \to t_i^+} f(x, t)$$

both exist for any $x \in \Omega$ and $i = 1, 2, \ldots, n$.

That is, g is permitted to have jump discontinuities at a finite number of values t_i. Therefore, for sufficiently large $|t|$, g is smooth, and the hyperbolic trajectories $(a_\varepsilon(t), t)$ and $(b_\varepsilon(t), t)$ are well defined, and since their manifolds are defined in terms of exponential decay rates, so are W_a^u and W_b^s in an appropriate t-slice. However, these cease to be defined across the jump discontinuities, since trajectories lose smoothness in time. For discontinuities of this form, trajectories $x(t)$ are in $\mathrm{C}^0(\mathbb{R})$, but lose differentiability at $t \in \mathcal{J}$. Now, although manifolds do not exist in the technical sense, one can define the 'pseudo-manifolds' formed by following the flow in time. That is, define the pseudo-manifold $\tilde{W}_a^u$ for all t by considering W_a^u, which exists for sufficiently negative t, and following its flow for all t. Similarly, define $\tilde{W}_b^s$ by following the flow of W_b^S backwards in time from sufficiently large t.

The $\mathcal{O}(\varepsilon)$ closeness of these pseudo-manifolds to the unperturbed manifolds works on any half-line $(-\infty, t)$ for $\tilde{W}_a^u$, and for any half-line (t, ∞) for $\tilde{W}_b^s$. Therefore, for any finite t,

$\tilde{W}_a^u$ and $\tilde{W}_b^s$ exist as in figure 1. The flux in this instance can therefore be computed exactly as described in section 3, with the term 'pseudo-manifold' replacing 'manifold'. A perusal of the standard Melnikov approaches (say, in [2, 7]) indicates that it still works in this instance, since one simply needs to follow trajectories of the perturbed flow along manifolds, which by definition follow pseudo-manifolds in this case. Extending the Melnikov transform for functions in $\mathcal{P}_J$ is also legitimate. Therefore, no additional work is necessary.

If g is time-periodic and has jump discontinuities in time, it will not be in the function space $\mathcal{P}_J$ since there are infinitely many discontinuities. This occurs, for example, in the 'egg-beater flow' [26]. Nevertheless, the Melnikov transform would still be legitimate in this instance, since one can think of the manifolds (and pseudo-manifolds) in terms of the Poincaré map which exists because of time-periodicity. The theory therefore covers the flux quantification in such situations as well. More serious discontinuities in t could lead to problems in solution and manifold existence—and this is carefully addressed in section 5.

5. Impulsive perturbations

The theory of separatrix splitting does not seem to have been addressed for perturbations *impulsive* in time, for example, those containing terms like the Dirac delta function $\delta(t)$. A clear difficulty is that manifolds no longer exist globally. It is a tempting observation that the Melnikov transform can be formally computed even in such instances. This section addresses the applicability of this approach, thereby rationalizing the usage of the Melnikov function as the flux characterizer for impulsive perturbations.

Firstly, it is necessary to detail the types of permitted perturbations. The intuition is provided through the following expression, to be interpreted symbolically only:

$$\text{`}\dot{x} = \mathrm{J}\nabla H(x) + \varepsilon \sum_{i=1}^{n} \delta(t - t_i) h_i(x, t)\text{'}, \tag{7}$$

where $\mathcal{J} := \{t_1, t_2, \ldots, t_n\}$ is now a *finite* set of jump values at which Dirac delta impulses $\delta(t - t_i)$ are imposed on the flow (not just jump discontinuities as in section 4). The h_i are functions in $\mathcal{P}$. The 'equation' (7) is actually meaningless, since there is ambiguity in the values $h_i(x(t_i), t_i)$, as $x(t)$ exhibits a $\mathcal{O}(\varepsilon)$ jump discontinuity at each t_i. Consider instead the following integral formulation

$$x(t) = x(\alpha) + \int_\alpha^t J\nabla H(x(s))\,\mathrm{d}s + \frac{\varepsilon}{2} \sum_{i=1}^{n} u(\alpha, t; t_i)[h_i x(t_i^-), t_i) + h_i(x(t_i^+), t_i)], \tag{8}$$

where $0 \leqslant \varepsilon \ll 1$, and t and the initial time α live in $\mathbb{R} \setminus \mathcal{J}$. The indicator function $u(\alpha, t; t_i)$ switches on only if the jump-value t_i is between α and t, i.e.

$$u(\alpha, t; t_i) = \begin{cases} 1 & \text{if either } \alpha < t_i < t \text{ or } t < t_i < \alpha, \\ 0 & \text{if else.} \end{cases}$$

This formulation models (7) in the sense that it is the solution obtained if approximating the Dirac delta $\delta(t - t_i)$ as the 'limit' of a square pulse of shrinking width and expanding height centred at t_i, with unit L^1-norm. In representing (8) a particular choice of this limiting procedure had to be made; alternative choices (for example pulses which are non-symmetrical about t_i) lead to *different* integral formulations, each having different solutions. This highlights the illegitimacy of (7). The form chosen for (8) is the most natural choice in the sense that the limit preserves the symmetry that we 'expect' from the Dirac delta distribution.

Definition 5.1. The function space $\mathcal{P}_L \subset \mathcal{P}$ (where $\mathcal{P}$ is defined in definition 2.1) is the set of functions f satisfying the following additional condition:

(P3) There exists a Lipschitz constant η uniform in $t \in \mathbb{R}$ such that

$$|f(x,t) - f(y,t)| \leqslant \eta |x-y| \qquad \text{for all } x, y \in \Omega.$$

Definition 5.2. The space $\mathcal{Q}(R)$ for a finite interval $R \subset \mathbb{R}$ is the set of piecewise continuous (with respect to $\mathcal{J}$) functions from R to $\mathbb{R}$: these functions are in $C^0(R \setminus \mathcal{J})$, and moreover possess left and right hand limits at all points in R.

Lemma 5.1. *Suppose $h_i \in \mathcal{P}_L$ for $i = 1, 2, \ldots, n$. Then, for suitably small ε and T, solutions to (8) exist uniquely in the space $\mathcal{Q}((\alpha - T, \alpha + T))$.*

Proof. For a given $T > 0$, define the linear operator G on $\mathcal{Q}((\alpha - T, \alpha + T))$ by

$$G(x) := x(\alpha) + \int_\alpha^{\alpha+t} J\nabla H(x(s))\,\mathrm{d}s + \frac{\varepsilon}{2}\sum_{i=1}^n u(\alpha, \alpha+t; t_i)[h_i(x(t_i^-), t_i) + h_i(x(t_i^+), t_i)],$$

where $-T < t < T$. Since $x \in C^0$ in each fixed subinterval bounded by the finite number of jump points $\mathcal{J}$, $\mathcal{Q}$ is a Banach space under the essential supremum norm. If $\|\cdot\|_\infty$ denotes the supremum norm on the interval $[\alpha, \alpha+t] \setminus \mathcal{J}$ (or $[\alpha+t, \alpha] \setminus \mathcal{J}$ if $t < 0$),

$$\begin{aligned}
\|G(y) - G(x)\|_\infty \leqslant{}& \left|\int_\alpha^{\alpha+t} |J\nabla H(y(s)) - J\nabla H(x(s))|\,\mathrm{d}s\right| \\
&+ \frac{\varepsilon}{2}\sum_{i=1}^n |h_i(y(t_i^-), t_i) - h_i(x(t_i^-), t_i)| \\
&+ \frac{\varepsilon}{2}\sum_{i=1}^n |h_i(y(t_i^+), t_i) - h_i(x(t_i^+), t_i)| \\
&\leqslant \eta_0 \left|\int_\alpha^{\alpha+t} |y(s) - x(s)|\,\mathrm{d}s\right| + \frac{\varepsilon}{2}\sum_{i=1}^n 2\eta_i \|y - x\|_\infty \\
&\leqslant \left(\eta_0|t| + \varepsilon\sum_{i=1}^n \eta_i\right)\|y - x\|_\infty.
\end{aligned}$$

In the above, η_0 is $J\nabla H$'s Lipschitz constant in the interval, and the η_i are the uniform Lipschitz constants of the h_i. Thus, for suitably small $|t|$ and ε, G is a contraction on $\mathcal{Q}$, and hence has a unique fixed point in $\mathcal{Q}((\alpha - T, \alpha + T))$ for suitably small T and ε. □

Lemma 5.2. *For each jump point t_i,*

$$x(t_i^+) = x(t_i^-) + \frac{\varepsilon}{2}[h_i(x(t_i^-), t_i) + h_i(x(t_i^+, t_i)]. \tag{9}$$

Proof. Apply the integral formulation (8) from an initial value $\alpha = t_i^-$ to a final value $t = t_i^+$. The first integral vanishes, since it is of a bounded function over an interval which shrinks to zero. Only the terms associated with t_i survive in the summation, since this is the only discontinuity, and (9) is easily obtained. □

Based on (9), define the family of mappings I_i on Ω to obey the implicit equation

$$I_i(x) - x - \frac{\varepsilon}{2}[h_i(x, t_i) + h_i(I_i(x), t_i)] = 0. \tag{10}$$

Lemma 5.3. *For ε small enough, and $h_i \in \mathcal{P}_L$, (10) defines I_i on Ω uniquely. Moreover, I_i is as smooth a function on Ω as $h_i(\cdot, t_i)$.*

Proof. Write $y = I_i(x)$, and suppress the t_i dependence for convenience. Then it is required to solve

$$y - x - \frac{\varepsilon}{2}[h_i(y) + h_i(x)] = 0.$$

Note that when $\varepsilon = 0$, a unique solution for $y(x, \varepsilon)$ is $y = x$. Now, the y-derivative of the left-hand side above is

$$d := \left(\mathbb{I} - \frac{\varepsilon}{2} \left(\frac{\partial(h_i^1, h_i^2)}{\partial(y_i^1, y_i^2)} \right) \right),$$

where the superscripts identify components, and $\mathbb{I}$ is the identity matrix. Each term perturbing the above from the identity is bounded by $(\varepsilon/2)\sup_i \eta_i$, and therefore d is a small perturbation from the identity. Hence, for small enough ε, $\det(d) \neq 0$. Thus, for any $x_0 \in \Omega$, there exists an open neighbourhood $B(x_0)$, and also a small interval containing 0 (say, E), such that for $(x, \varepsilon) \in B(x_0) \times E$, y can be solved uniquely as a function of (x, ε), by the implicit function theorem. This moreover establishes that y is as smooth in x as h_i. Since this works for any $x_0 \in \Omega$, a global smooth solution $y(x, \varepsilon)$ exists on $\Omega \times E$. □

What has been shown so far is that each solution to (8) evolves smoothly until a jump value t_i, at which point the solution resets itself to a new value which is $\mathcal{O}(\varepsilon)$ away, and then evolves smoothly until the next jump value. These solutions will now be incorporated to form pseudo-manifolds, essentially as in section 4, which are the next step in attempting to define a flux function for impulsive flows.

When (8) is considered for small $|\varepsilon|$, the finiteness of the jump points means that for $|t|$ sufficiently large, one may as well set $\varepsilon = 0$. Therefore, the perturbed hyperbolic trajectory $a_\varepsilon(t)$ is equal to a for t sufficiently negative, and similarly $b_\varepsilon(t)$ equals b at large t. The hyperbolicity of these entities follows from the hyperbolicity at $\varepsilon = 0$, and therefore $a_\varepsilon(t)$ retains its unstable manifold W_a^u for sufficiently negative t. The definition of this manifolds can be accomplished through exponential decay estimates as $t \to -\infty$ [24]. Similarly, $b_\varepsilon(t)$ retains its stable manifold W_b^s for $t > \sup_i\{t_i \in \mathcal{J}\}$. Now construct the 'pseudo-manifolds' for any finite t not in $\mathcal{J}$ as follows. The pseudo-manifold $\tilde{W}_a^u$ in $\Omega \times (\mathbb{R} \setminus \mathcal{J})$ is obtained by following W_a^u forward in time by the flow (8). Similarly, $\tilde{W}_b^s$ is the entity formed by following W_b^s backwards in time by the flow (8). These pseudo-manifolds are smooth one-dimensional curves in each (non-jump) time-slice and at each jump value are mapped to a curve which retains this smoothness by lemma 5.3. Therefore, the picture in any time-slice which is not a jump value is as given in figure 1. It is then possible to define the flux *exactly* as in section 3, with the pseudo-manifolds playing the role of the manifolds. The flux function $\Lambda(t)$ however is defined on $\mathbb{R} \setminus \mathcal{J}$, and not on $\mathbb{R}$.

Suppose the Melnikov transform of Definition 2.2 is applied *formally* to this instance, by choosing

$$g(x, t) = \sum_{i=1}^{n} \delta(t - t_i) h_i(x, t)$$

(in a distributional sense), as is apparent from (7). If this is valid, the flux function would take the form

$$\begin{aligned}\Lambda(t) = M(t) &= \Phi\{g\}(t)\\ &= \sum_{i=1}^{n} \int_{-\infty}^{\infty} \delta(t+\tau-t_i)\nabla H(\bar{x}(\tau))\cdot h_i(\bar{x}(\tau), t+\tau)\,\mathrm{d}\tau\\ &= \sum_{i=1}^{n} \nabla H(\bar{x}(t_i - t))\cdot h_i(\bar{x}(t_i-t), t_i).\end{aligned}$$

The question is: can this process be justified, using the correct *integral* formulation (8)? The answer fortunately is 'yes' and requires some modifications to the standard process of obtaining the Melnikov function, which usually require the perturbation to be differentiable [2, 7].

Theorem 5.1. *The flux function associated with the perturbation as expressed in (8) is given for $t \in \mathbb{R}\setminus\mathcal{J}$ by any of the forms*

$$\begin{aligned}\Lambda(t) = M(t) = \Phi\left\{\sum_{i=1}^{n}\delta(\tau-t_i)h_i(x,\tau)\right\} &= \sum_{i=1}^{n}\nabla H(\bar{x}(t_i-t))\cdot h_i(\bar{x}(t_i-t),t_i)\\ &= \sum_{i=1}^{n}|\nabla H(\bar{x}(t_i-t))|h_i^{\perp}(\bar{x}(t_i-t),t_i). \qquad (11)\end{aligned}$$

Proof. Begin with expression (2), which gives a definition for the distance between the perturbed pseudo-manifolds, measured in a time-slice t. This measure is obtained along a perpendicular transversal $\mathcal{T}$ to the original heteroclinic at the point $\bar{x}(0)$. It is required to determine the Melnikov function satisfying (2). Applying the standard Melnikov approaches to compute $M(t)$ usually requires working with derivatives of the velocity fields [2, 7], which is not possible in this instance. Nevertheless, obtaining $h_\varepsilon^u(t)$ and $h_\varepsilon^s(t)$ is similarly related to following trajectories of the dynamical system, for which the integral formulation (8) now needs to be used. For notational convenience in this proof only, set $F(x) = J\nabla H(x)$, and let DF be its Jacobian matrix. Now, from (8), if α and τ are not in $\mathcal{J}$,

$$x(\tau) = x(\alpha) + \int_{\alpha}^{\tau} F(x(s))\,\mathrm{d}s + \frac{\varepsilon}{2}\sum_{i=1}^{n} u(\alpha,\tau;t_i)[h_i(x(t_i^+),t_i) + h(x(t_i^-),t_i)]. \qquad (12)$$

Fix $t \in \mathbb{R}\setminus\mathcal{J}$, the time-slice in which the distance is to be measured. Set

$$x^\sigma(\tau) = \bar{x}(\tau - t) + \varepsilon x_i^\sigma(\tau) + \mathcal{O}(\varepsilon^2),$$

where σ is either u or s, and the above expansion is uniformly valid for $\tau \in (-\infty, t]$ for $\sigma = u$ and for $[t,\infty)$ for $\sigma = s$. Thus, x^u represents the trajectory which backwards asymptotes to $a_\varepsilon(\tau)$ and crosses the transversal $\mathcal{T}$ at the point $h_\varepsilon^u(t)$ (the fact that this trajectory jumps by $\mathcal{O}(\varepsilon)$ at intermediate jump values does not impede this expansion). Similarly, x^s is associated with the trajectory passing through $h_\varepsilon^s(t)$ and approaching $b_\varepsilon(\tau)$ as $\tau \to +\infty$. Now, substituting this expansion in (12), and separating out the $\mathcal{O}(\varepsilon^0)$ and $\mathcal{O}(\varepsilon^1)$ terms, one gets, respectively,

$$\bar{x}(\tau - t) = \bar{x}(\alpha - t) + \int_{\alpha}^{\tau} F(\bar{x}(s-t))\,\mathrm{d}s \qquad (13)$$

and

$$\begin{aligned}x_1^\sigma(\tau) = x_1^\sigma(\alpha) &+ \int_{\alpha}^{\tau}\mathrm{DF}(\bar{x}(s-t))x_i^\sigma(s)\,\mathrm{d}s\\ &+ \sum_{i=1}^{n}\frac{u(\alpha,\tau;t_i)}{2}[h_i(\bar{x}(t_i^+ - t),t_i) + h_i(\bar{x}(t_i^- - t),t_i)]. \qquad (14)\end{aligned}$$

While (13) is simply a statement of $\bar{x}$ being a solution to the $\varepsilon = 0$ flow of (12), (14) is representable as

$$x_1^\sigma(\tau) = x_i^\sigma(\alpha) + \int_\alpha^\tau \mathrm{DF}(\bar{x}(s-t))x_i^\sigma(s)\,\mathrm{d}s + \sum_{i=1}^n u(\alpha,\tau;t_i)h_i(\bar{x}(t_i-t),t_i), \tag{15}$$

since h_i and $\bar{x}$ are continuous. Now define the 'wedge' operator for two-dimensional vectors by $(a_1,a_2)\wedge(b_1,b_2) := a_1b_2 - a_2b_1$, and the quantity

$$\Delta^\sigma(\tau) = F(\bar{x}(\tau-t))\wedge x_i^\sigma(\tau)$$

for $\sigma = u, s$. This permits (2) to be written as

$$d(t,\varepsilon) = \varepsilon\frac{\Delta^u(t)-\Delta^s(t)}{|F(\bar{x}(0))|} + \mathcal{O}(\varepsilon^2), \tag{16}$$

since $h_\varepsilon^\sigma(t) = x^\sigma(t) = \bar{x}(0)+\varepsilon x_1^\sigma(t)+\mathcal{O}(\varepsilon^2)$. From (2), this means that the Melnikov function is defined by $M(t) := \Delta^u(t)-\Delta^s(t)$. To proceed further, it is necessary to prove the following expression, valid when $\tau > \alpha$:

$$\Delta^\sigma(\tau) = \Delta^\sigma(\alpha) + \sum_{i=1}^n u(\alpha,\tau;t_i)F(\bar{x}(t_i-t))\wedge h_i(\bar{x}(t_i-t),t_i). \tag{17}$$

The proof of this is relegated to appendix A. Now, take (17) with $\sigma = u$, $\alpha = -\infty$ and $\tau = t$, to get

$$\begin{aligned}\Delta^u(t) &= \lim_{\alpha\to-\infty}\Delta^u(\alpha) + \sum_{i=1}^n u(-\infty,t;t_i)F(\bar{x}(t_i-t))\wedge h_i(\bar{x}(t_i-t),t_i)\\ &= \sum_{i=1}^n u(-\infty,t;t_i)\nabla H(\bar{x}(t_i-t))\cdot h_i(\bar{x}(t_i-t),t_i)\end{aligned}$$

since $\Delta^u(\alpha)\to 0$ as $\alpha\to-\infty$ because $F\to 0$ while x_1^u remains bounded. Similarly considering (17) with $\sigma = s$, $\alpha = t$ and $\tau = \infty$,

$$\lim_{\tau\to\infty}\Delta^s(\tau) = \Delta^s(t) + \sum_{i=1}^n u(t,\infty;t_i)\nabla H(\bar{x}(t_i-t))\cdot h_i(\bar{x}(t_i-t),t_i),$$

and hence

$$\Delta^s(t) = -\sum_{i=1}^n u(t,\infty;t_i)\nabla H(\bar{x}(t_i-t))\cdot h_i(\bar{x}(t_i-t),t_i).$$

Therefore, when computing $M(t) = \Delta^u(t)-\Delta^s(t)$, all jump points are summed over. This gives the result (11). □

Lemma 5.4. *The flux function of theorem 5.1 satisfies*

$$\lim_{t\to t_j^+}\Lambda(t) = \lim_{t\to t_j^-}\Lambda(t)$$

for $t_j\in\mathcal{J}$, enabling the extension of definition of $\Lambda(t)$ to $\mathbb{R}$ with the understanding that for $t\in\mathcal{J}$,

$$\Lambda(t) = \lim_{\tau\to t}\Lambda(\tau).$$

Proof. Since $\bar{x}$, being associated with the smooth $\varepsilon = 0$ flow, is continuous and since all other functions in (11) are by assumption continuous, this is obvious. □

Remark 5.1. (Smoothness properties of the impulsive flux function). The flux function $\Lambda(t)$ is in $C^r(\mathbb{R})$ if each of $\bar{x}$ and h_i is similarly in $C^r(\mathbb{R})$, and $H \in C^{r+1}(\Omega)$. Moreover, since h_i is uniformly bounded (it is in $\mathcal{P}$) and H has exponential decay through Lemma 2.1, $\Lambda(t) \in L^p(\mathbb{R})$ for any $p \in [1, \infty]$.

Remark 5.2. The actual integral formulation used to model the conceptual 'equation' (7) has no effect on the flux computation. That is, representations other than (8) (which was obtained by thinking of the Dirac delta in the limit of symmetric square pulses) are possible. This irrelevance is reflected in (15), which would result irrespective of the limiting process chosen. Thus, the $\mathcal{O}(\varepsilon)$ flux is not affected by this choice, even though individual trajectory details are.

Remark 5.3. As in section 3, an alternative parametrization $\bar{x}_\beta$ chosen for the heteroclinic has the trivial effect on the flux function of shifting it by β.

Remark 5.4. *The Melnikov transform in definition 2.2 therefore provides a direct measure of the cross-separatrix flux not just for perturbations in $\mathcal{P}$ or $\mathcal{P}_J$, but also those representable as in (8) with the h_i in $\mathcal{P}_L$.*

Remark 5.5. The Melnikov summation in theorem 5.1 possesses an illusory similarity with such summations in *discrete* dynamical systems (see, for example [27, 28]), yet the situation here is different. Theorem 5.1 relates to distances and flux between (pseudo)manifolds associated with a *continuous* flow which is reset at a finite number of instances in time.

6. An example

As an illustrative example for quick computations, consider a two-dimensional flow in the variables (x, y), given by

$$\dot{x} = -\sin(2\pi x)\sin(2\pi y), \qquad \dot{y} = -\cos(2\pi x)\cos(2\pi y),$$

which is a frequently used kinematic model for Rayleigh–Bénard-type 'rolls' (see the references in [5]). This has a Hamiltonian function

$$H(x, y) = -\frac{1}{2\pi}\sin(2\pi x)\cos(2\pi y),$$

and a part of the periodic phase-space is shown in figure 2. The intention is to describe the possibility of fluid flux from the left cell to the right, by crossing the heteroclinic separatrix (shown as a heavy line) which goes from $(0, 1/4)$ to $(0, 3/4)$. It is possible in this instance to show that $\bar{y}(t) = (1/4) + (1/\pi)\tan^{-1}\exp(2\pi t)$ and $|\nabla H| = \operatorname{sech}(2\pi t)$ for a symmetrically chosen time-parametrization along the heteroclinic. The flux function for a perturbation $\varepsilon g(x, y, t)$ is therefore

$$\Lambda(t) = \int_{-\infty}^{\infty} \operatorname{sech}(2\pi\tau) g^{\perp}\left(0, \frac{1}{4} + \frac{1}{\pi}\tan^{-1} e^{2\pi\tau}, t + \tau\right) d\tau,$$

by employing the definition of the Melnikov transform. Here, $g^{\perp}$ is the x-component of the perturbation, since ∇H points in the positive x-direction on the heteroclinic. (Thus positive Λ represents motion from the left to the right cell in figure 2.) Computation

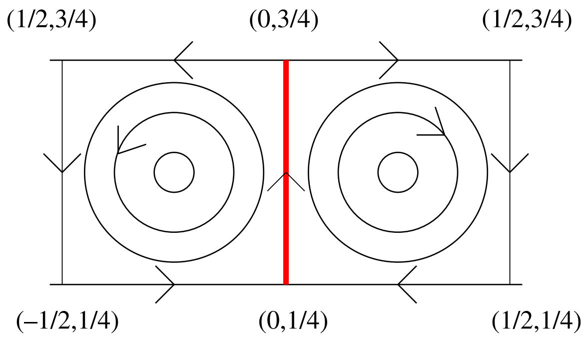

Figure 2. Unperturbed flow for section 6.

of the flux is presented in figures 3 and 4 for a variety of different forms chosen for $g^{\perp}(0, y, t)$:

(a) $\sin(4\pi y)\sin(3t)$,
(b) $\sin(4\pi y)\sin(3t) - 7\cos(12\pi y)\sin(0.37t) + 3\sin(12\pi y)\cos(7t)$,
(c) $\cos(12\pi y)u(t-1/2)$,
(d) $\cos(12\pi y)[u(t-1/3) - u(t-1/2)]$,
(e) $\sin(12\pi y)\delta(t-1)$,
(f) $\sin(12\pi y)\delta(t-1) - \cos^2(12\pi y)\delta(t-3) + 80(y-1/2)^3\delta(t-6)$,
(g) $\sin(12\pi y)\exp(-7|t-1|)$,
(h) $\exp[-7|(t-1)\sin(12\pi y)|]$,
(i) $\mathrm{sech}(t-y) - \mathrm{sech}^2(2y^2 - t^2)$,
(j) $\mathrm{sech}(t-5y)\sin(12\pi y)\delta(t-2)$,
(k) $9\exp[-(t+5y)^2]\cos(4\pi y) - \mathrm{sech}[\sin(12\pi y)t]$,
(l) $\mathrm{sech}[\sin(12\pi y)t]\delta(t-1)$.

In the above, $u(\cdot)$ is the unit-step function, and an abuse of notation (as in (7)) is used whenever a Dirac delta is written; such perturbations in reality should be thought of at the integral equation level (as in (8)). Case (a) in figure 3 is the standard instance of a time-periodic perturbation, and figure 3(*a*) is unsurprising: fluid sloshes back and forth periodically from the left cell to the right (the standard picture of a heteroclinic tangle occurs here, leading to chaotic mixing across the heteroclinic). A quasi-periodic instance is presented in case (b), and the resulting flux function is itself quasi-periodic, with chaotic mixing occurring in the senses described in [8, 9, 10, 12, 14]. In all other examples considered, the flux function has only a finite number of zeros, indicating that the mixing mechanism is *non-chaotic*.

Case (c) is the interesting case when a perturbation is switched on at $t = 1/2$, and the corresponding flux function is essentially zero for $t > 1/2$, but the perturbation has some influence for t just below $1/2$. While this may initially seem contradictory, the evolution of the pseudo-manifolds explains this behaviour. The manifold W_a^u evolves as in the unperturbed flow until $t = 1/2$, whereupon it undergoes a kink which generates $\tilde{W}_a^u$ for larger times. On the other hand, W_b^s is influenced from $+\infty$ backwards in time to $t = 1/2$ by the perturbation. At $t = 1/2$, the perturbation switches off, but now the resulting backwards trajectories are *not* the same as they would have been when $\varepsilon = 0$, since the 'initial' condition at $t = 1/2$ is different. Hence, $\tilde{W}_b^s$ is ε-influenced for $t < 1/2$, whereas $\tilde{W}_a^u$ is not. This results in a difference between them, leading to a non-zero flux. On the other hand, both $\tilde{W}_b^s$ and $\tilde{W}_a^u$ are ε-influenced for $t > 1/2$. Apparently they are influenced similarly to leading order in ε,

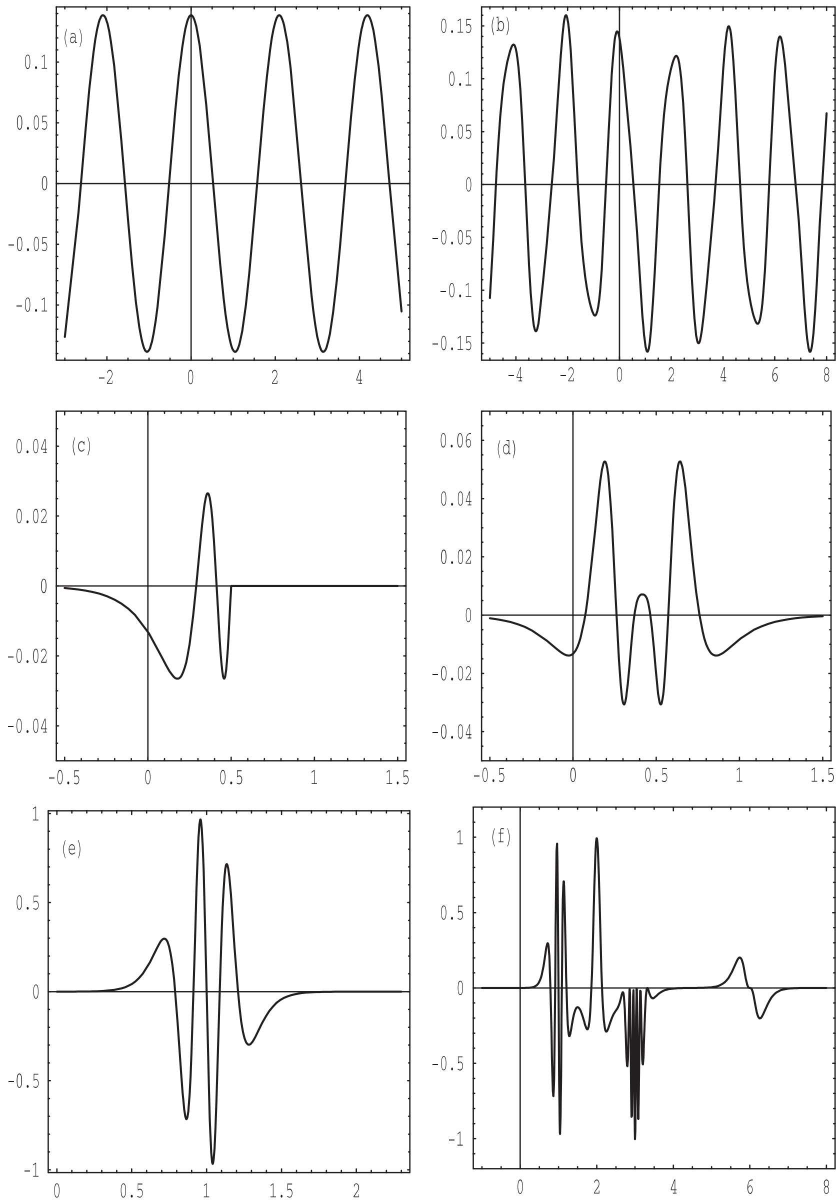

Figure 3. Flux functions ($\Lambda(t)$ versus t), for the various perturbations of section 6.

which results in the $\mathcal{O}(\varepsilon)$ flux being zero. Another issue worthy of note in figure 3(c) is the non-differentiability of $\Lambda(t)$ at $t = 1/2$, caused by $g^{\perp}$'s discontinuity at that value.

In case (d), a pulse is switched on between $t = 1/3$ and $t = 1/2$. Fluid mixing occurs during this interval, and also for nearby times, before its influence dissipates. In case (e), the Dirac delta impulse at $t = 1$ causes a 'soliton' or 'wavelet' type response in the flux function. More oscillations in this function can be created by introducing a more wiggly spatial part to this perturbation, for example $\sin(24\pi y)$. Finally, case (f), in figure 3(*f*), displays the effect of having several Dirac delta impulses. As is obvious from (11), this is obtainable by a linear

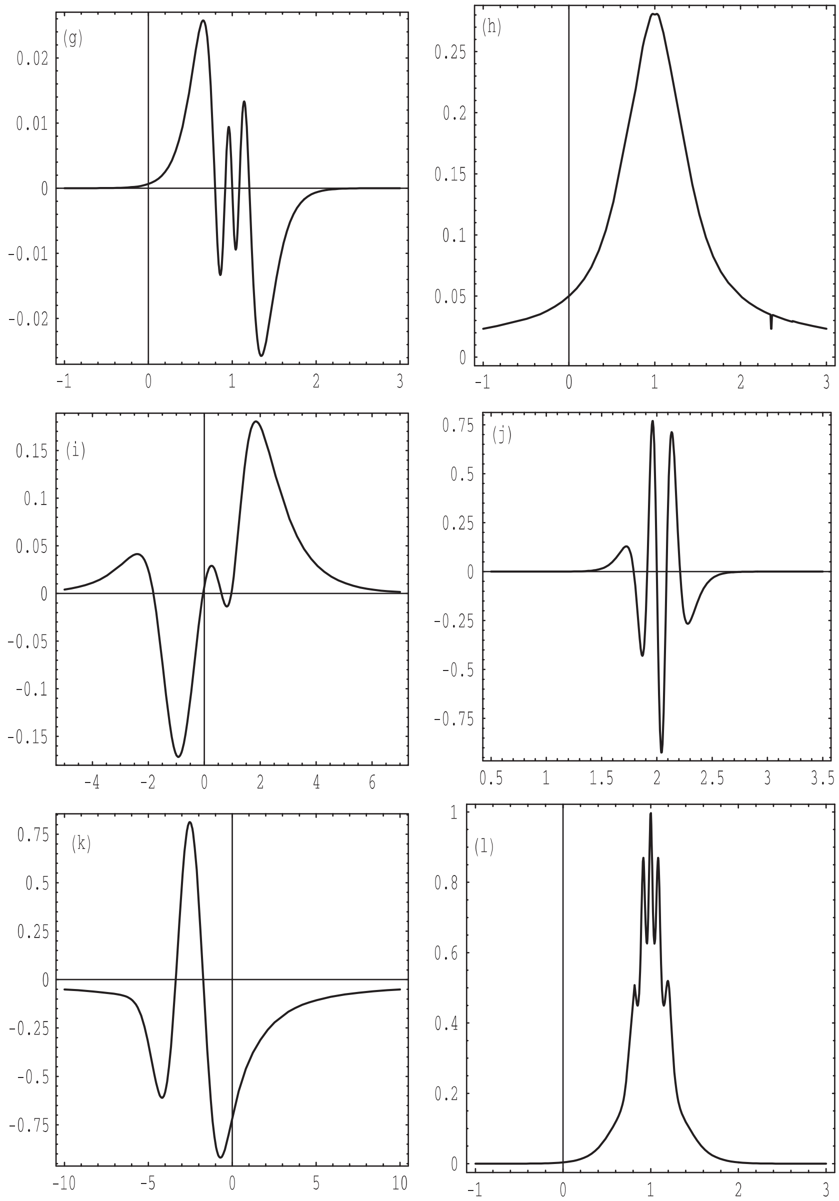

Figure 4. Flux functions ($\Lambda(t)$ versus t), for the various perturbations of section 6.

combination of the 'soliton' structures associated with each impulse. The rapid decay of the hyperbolic secant function means that impulses imposed far away from $t = 0$ have diminished effect; the impulse at $t = 6$ requires a large multiplicative factor to be important.

Case (g) has a perturbation which peaks at $t = 1$, which from figure 4(g) has an important effect near $t = 1$ (where some sloshing back and forth is to be seen), which eventually dies out for large $|t - 1|$. A similar behaviour occurs for case (h), in which however the flux is uni-directional. The cases which follow illustrate how the theory is easily applicable even to *non-separable* perturbations (in which the t and y dependences do not necessarily separate).

case (i) is one such instance, and the flux function has the same qualitative structure as in, say, case (c). In case (j) an apparently significantly more complicated situation is addressed, in which the perturbation has non-separable parts multiplying an impulse. Nevertheless, the flux function is easily computed using theorem 5.1 and has the standard 'soliton'-like structure. The perturbation used for case (k) is also non-separable, and the resulting flux is mainly from the right to the left cell, except for a short time period near $t = 3$ during which there is fluid flow in the opposite direction. Finally, in case (l), an instance in which flux occurs only from the left to the right is shown, with some pulsations in the magnitude near the impulsive time $t = 1$.

7. Conclusions

A theory has been established which permits the computation of flux, explicitly as the transfer of fluid per unit time, across a heteroclinic separatrix in two-dimensional flows. The perturbation is permitted to be aperiodic, and significantly discontinuous in the time variable, possibly as nasty as Dirac delta distributions. The derivation of this theory required a rationalization of the flux, which necessitated definitions of pseudo-manifolds and pseudo-separatrices for such instances. The flux was presented in terms of a time-dependent function, which measures the instantaneous transfer of fluid across the separatrix. It was necessary to define an appropriate time-dependent separatrix, which demarcated distinct fluid regimes. The flux formulae, given in theorems 3.1 and 5.1, express this flux directly in terms of the Melnikov transform (definition 2.2) applied to the perturbing function. A pleasing result is the *direct* connection between the Melnikov and the flux function. The extension of validity of the interpretation of the Melnikov transform to discontinuous functions required a re-evaluation of the standard Melnikov approach, which also was presented.

In section 6, computations and interpretations of the flux were shown for an example. The flux calculations were relatively easy, and it is hoped that this approach would increase our understanding and provide significant new analytical and computational tools for time-*aperiodic* perturbations which are moreover permitted to be extremely discontinuous.

Acknowledgments

Comments from Daniel Daners on the proper formulation of the impulsive equations are gratefully acknowledged.

Appendix A. Proof of equation (17)

Start with the identity

$$(Ab) \wedge c + b \wedge (Ac) = \mathrm{Tr}(A)(b \wedge c)$$

valid for 2×2 matrices A and two-dimensional vectors b and c. Now, since $F = J\nabla H$, $A = \mathrm{DF}$ is trace-free, and by also setting $b = F$ and $c = x_1^\sigma$,

$$[(\mathrm{DF})F] \wedge x_1^\sigma + F \wedge [(\mathrm{DF})x_1^\sigma] = 0.$$

Therefore,

$$\int_\alpha^\tau [(\mathrm{DF})(\bar{x}(s-t))F(\bar{x}(s-t))] \wedge x_1^\sigma(s)\,\mathrm{d}s + F(\bar{x}(s-t)) \wedge [(\mathrm{DF}(\bar{x}(s-t)))x_1^\sigma(s)]\,\mathrm{d}s = 0.$$

Now adding the above to the RHS of (17) and employing also the Dirac delta distribution, the RHS of (17) is

$$\mathrm{RHS} = \Delta^\sigma(\alpha) + \sum_{i=1}^{n} \int_\alpha^\tau \delta(s-t_i) F(\bar{x}(s-t)) \wedge h_i(\bar{x}(s-t), t_t)\,\mathrm{d}s$$
$$+ \int_\alpha^\tau [(\mathrm{D}F)F] \wedge x_1^\sigma(s)\,\mathrm{d}s + \int_\alpha^\tau F \wedge [(\mathrm{D}F) x_1^\sigma(s)]\,\mathrm{d}s,$$

where the argument $\bar{x}(s-t)$ for each of F and DF has been omitted for convenience. Therefore,

$$\mathrm{RHS} = \Delta^\sigma(\alpha) + \int_\alpha^\tau [\mathrm{D}F(\bar{x}(s-t))F(\bar{x}(s-t))] \wedge x_1^\sigma(s)\,\mathrm{d}s$$
$$+ \int_\alpha^\tau F(\bar{x}(s-t)) \wedge [\mathrm{D}F(\bar{x}(s-t))x_1^\sigma(s) + \sum_{i=1}^{n} \delta(s-t_i) h_i(\bar{x}(s-t), t_i)]\,\mathrm{d}s. \quad \text{(A.1)}$$

Now apply integration by parts to the second integral in (A.1), by setting $u(s) = F(\bar{x}(s-t))$ and

$$\mathrm{d}v = \left[\mathrm{D}F(\bar{x}(s-t)) + \sum_{i=1}^{n} \delta(s-t_i) h_i(\bar{x}(s-t), t_i) \right] \mathrm{d}s.$$

Then,

$$\mathrm{d}u = \mathrm{D}F(\bar{x}(s-t)) \frac{\mathrm{d}}{\mathrm{d}s}(\bar{x}(s-t)) = \mathrm{D}F(\bar{x}(s-t)) F(\bar{x}(s-t))$$

by (13). Using (15),

$$v(s) = x_1^\sigma(s) - x_1^\sigma(\alpha).$$

Therefore, from (A.1),

$$\mathrm{RHS} = \Delta^\sigma(\alpha) + F(\bar{x}(s-t)) \wedge [x_i^\sigma(s) - x_1^\sigma(\alpha)]]_{s=\alpha}^\tau$$
$$+ \left(\int_\alpha^\tau \mathrm{D}F(\bar{x}(s-t)) F(\bar{x}(s-t))\,\mathrm{d}s \right) \wedge x_1^\sigma(\alpha)$$
$$= \Delta^\sigma(\alpha) + F(\bar{x}(\tau-t)) \wedge x_1^\sigma(\tau) - F(\bar{x}(\tau-t)) \wedge x_1^\sigma(\alpha)$$
$$+ \left(\int_\alpha^\tau \frac{\mathrm{d}}{\mathrm{d}s} F(\bar{x}(s-t))\,\mathrm{d}s \right) \wedge x_1^\sigma(\alpha)$$
$$= F(\bar{x}(\alpha-t)) \wedge x_i^\sigma(\alpha) + F(\bar{x}(\tau-t)) \wedge x_1^\sigma(\tau) - F(\bar{x}(\tau-t)) \wedge x_1^\sigma(\alpha)$$
$$+ F(\bar{x}(\tau-t)) \wedge x_1^\sigma(\alpha) - F(\bar{x}(\alpha-t)) \wedge x_1^\sigma(\alpha)$$
$$= F(\bar{x}(\tau-t)) \wedge x_1^\sigma(\tau) = \Delta^\sigma(\tau)$$

which is the left-hand side of expression (17) as claimed.

References

[1] Rom-Kedar V and Wiggins S 1990 Transport in two-dimensional maps *Arch. Ration. Mech. Anal.* **109** 239–98
[2] Wiggins S 1992 *Chaotic Transport in Dynamical Systems* (New York: Springer)
[3] Rom-Kedar V, Leonard A and Wiggins S 1990 An analytical study of transport, mixing and chaos in an unsteady vortical flow *J. Fluid Mech.* **214** 347–94
[4] Rom-Kedar V and Poje A C 1999 Universal properties of chaotic transport in the presence of diffusion *Phys. Fluids* **11** 2044–57

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

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

[7] Guckenheimer J and Holmes P 1983 *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields* (New York: Springer)

[8] Scheurle J 1986 Chaotic solutions of systems with almost periodic forcing *J. Appl. Math. Phys. (ZAMP)* **37** 12–26

[9] Meyer K R and Sell G R 1989 Melnikov transforms, Bernoulli bundles, and almost periodic perturbations *Trans. Am. Math. Soc.* **314** 63–105

[10] Stoffer D 1988 Transversal homoclinic points and hyperbolic sets for non-autonomous maps II *J. Appl. Math. Phys. (ZAMP)* **39** 783–812

[11] Palmer K J 1984 Exponential dichotomies and transversal homoclinic points *J. Diff. Eqns* **55** 225–56

[12] Wiggins S 1999 Chaos in the dynamics generated by sequences of maps, with applications to chaotic advection in flows with aperiodic time dependence *Z. Angew. Math. Phys.* **50** 585–616

[13] Kaper T J and Wiggins S 1991 Lobe areas in adiabatic Hamiltonian systems *Physica* D **51** 205–12

[14] Lerman L and Shil'nikov L 1992 Homoclinical structures in nonautonomous systems: nonautonomous chaos *Chaos* **2** 447–54

[15] Balasuriya S, Jones C K R T and Sandstede B 1998 Viscous perturbations of vorticity-conserving flows and separatrix splitting *Nonlinearity* **11** 47–77

[16] Sandstede B, Balasuriya S, Jones C K R T and Miller P D 2000 Melnikov theory for finite-time vector fields *Nonlinearity* **13** 1357–77

[17] Grenier E, Jones C K R T, Rousset F and Sandstede B 2005 Viscous perturbations of marginally stable Euler flow and finite-time melnikov theory *Nonlinearity* **18** 465–83

[18] Balasuriya S 2005 Optimal perturbation for enhanced chaotic transport *Physica* D **202** 155–76

[19] Balasuriya S 2005 An approach for maximizing chaotic mixing in microfluidic devices *Phys. Fluids* **17** 118103

[20] Kaper T J and Kovacic G 1994 A geometric criterion for adiabatic chaos *J. Math. Phys.* **35** 1202–18

[21] Haller G and Poje A C 1998 Finite time transport in aperiodic flows *Physica* D **119** 352–80

[22] Poje A C and Haller G 1999 Geometry of cross-stream mixing in a double-gyre ocean model *J. Phys. Oceanography* **29** 1649–65

[23] Miller P D, Jones C K R T, Rogerson A M and Pratt L J 1997 Quantifying transport in numerically generated velocity fields *Physica* D **110** 105–22

[24] Coppel W A 1978 *Dichotomies in Stability Theory* (*Lecture Notes in Mathematics vol 629*) (Berlin: Springer)

[25] Adams R A 1978 *Sobolev Spaces* (San Diego, CA: Academic)

[26] Muzzio F J, Meneveau C, Swanson P D and Ottino J M 1992 Scaling and multifractal properties of mixing in chaotic flows *Phys. Fluids* A **4** 1439–56

[27] Easton R W 1984 Computing the dependence on a parameter of a family of unstable manifolds: generalized Melnikov formulas *Nonlinear Anal. Theory Methods Appl.* **8** 1–4

[28] Delshams A and Ramirez-Ros R 1996 Poincaré–Melnikov–Arnold method for analytic planar maps *Nonlinearity* **9** 1–26