Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica D 202 (2005) 155–176
PHYSICA D
www.elsevier.com/locate/physd

Optimal perturbation for enhanced chaotic transport

Sanjeeva Balasuriya*

School of Mathematics & Statistics, University of Sydney, NSW 2006, Australia

Received 9 May 2003; received in revised form 16 September 2004; accepted 17 November 2004

Communicated by I. Mezic

Abstract

The issue of determining the best perturbation which results in optimal chaotic flux across a separatrix is addressed, using the Melnikov function and lobe dynamics. This theoretical analysis is motivated mainly through micro-fluidic devices for which this problem has become important recently. Both two- and three-dimensional flows are analysed. Utilising a Fourier transform representation, the nature of the perturbation which maximises this flux for each frequency value is obtained. The resulting optimally attainable flux is computed. A concise bound on this flux is presented in terms of the supremum norm of the normal component of the perturbing velocity, and the size of the heteroclinic manifold. In this instance where the spatial part of the perturbation is permitted to be chosen based on the frequency, it is shown that greater flux is achievable for smaller frequencies. The theory is illustrated through two examples.

PACS: 05.45.−a; 05.60.−k; 47.52.+j; 05.45.Gg

Keywords: Chaotic flux; Melnikov's method; Optimal mixing; Micro-fluidic devices

1. Introduction

There are many engineering, biological, chemical, and combustion applications in which efficiency of an apparatus improves when solutions are well-mixed. In other applications, a well-mixed solution may be a goal in and of itself. A diverse collection of applications of such mixing characterisation appear in [1–9]. Of specific recent interest is the development of microfluidic devices which are of importance in a variety of applications such as drug delivery, diagnostic devices, chemical synthesis, printing, "lab-on-a-chip", protein analysis, gene expression profiling, cell culture, and chemical testing [10–18]. Enhancing or controlling fluid mixing is often a primary design aspect of the microfluidic device. Since turbulence is suppressed in the low Reynolds number scales associated with

* Tel.: +61 2 9351 4873; fax: +61 2 9351 4534.
E-mail address: sanjeeva@maths.usyd.edu.au.

0167-2789/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.physd.2004.11.018

microfluidic devices, and since diffusion by itself is not an effective mixing mechanism [19,11,16,17], focus has been moving towards exploiting chaotic advection for such mixing [16,20,21,18,22]. Characterising chaotic mixing has of course been extensively studied even within the context of macro-scale fluids (see for example [23–33,8,9,34] for a variety of applications).

Given the recent interest in enhancing mixing in micro-mixers, an important issue could be the determination of how best to perturb a laminar flow in order to achieve optimal mixing. Two obvious questions arise. What is the best frequency at which to perturb the system? What spatial form should the perturbation take?

In analysing this problem, an initial consideration is describing and, in particular, *quantifying*, the mixing present in a given flow. Statistical methods have been used for this purpose for quite some time. Alternative techniques to quantify the chaotic flux are related to Lyapunov exponents, ergodic theory, variational principles, partial separatrices, Markov models for transitions between regions, escape rates, effective diffusivities, minimal flux surfaces, and inter-material contact surfaces (for a recent selection of such ideas, see [35–39,24,40,7,41,34]). Many of these techniques provide diagnostics of chaotic flux. One method which provides a more direct assessment, however, is the lobe dynamics approach utilising Melnikov's method [42–46,25]. This usually has the disadvantage of requiring near-integrability, with the deviation from integrability possessing periodic time-dependence. In a two-dimensional setting the areas of lobes which are transferred are known to be related to an integral, with appropriate limits, of an entity called the Melnikov function [42–44]. This function, originally developed in [47], measures the signed distance between the perturbed manifolds which form the boundary of lobes (see also [48,49,44]). Higher-dimensional Melnikov methods exist [50,51], and indeed have been used to understand the geometry of lobes generated through perturbations [52,33] in three-dimensions. In spite of these advances, *quantifying* the flux using this technique is difficult, and has only been done numerically in a few instances [45,46].

With micro-mixers in mind, a natural idea would be to attempt to *optimise* the chaotic flux. Very few studies [25,1] have made any mathematical progress on this topic. The article [25] is particularly related to the current topic; it utilises a simplex algorithm to optimise the flux in a corner vortex flow. The authors discover that the optimum flux as a function of the frequency is monotonically decreasing, and decreases to a non-zero value at high frequencies [25]. This apparently contradicts the non-monotonic and decaying to zero behaviour expected from *non-optimised* studies [45,53,43]; more discussion on this issue will appear in Section 2.4. The current paper addresses the optimisation problem from a theoretical perspective, and proves that the behaviour in [25] for the optimal flux function is to be expected generically (although the flux definition used here is different). It moreover is able to provide considerable insight into the nature of the perturbation which optimises the flux. The present results differ from [25] in that they are not confined to particular equations, or to numerical algorithms.

Let us state the problem in more mathematical terms. Start with the unperturbed dynamical system

$$\dot{x} = f(x) \tag{1}$$

where $x \in \Omega \subset \mathbb{R}^n$ ($n = 2$ or 3), and $f : \Omega \to \mathbb{R}^n$ is volume-preserving and as smooth as required. Suppose this flows possesses a heteroclinic manifold which forms a separatrix in Ω. Therefore, there is no transport across this separatrix. The intention is to now perturb the flow, and consider the chaotic flux resulting across the separatrix. Perturbations to be considered will take the form

$$\dot{x} = f(x) + \varepsilon\, g(x)\, \cos[\omega(t - \beta)], \tag{2}$$

where $0 < \varepsilon \ll 1$, $\omega \in (0, \infty)$ is the frequency of the perturbation, and $\beta \in [0, 2\pi)$ is a phase constant. The choice of the function $g : \Omega \to \mathbb{R}^n$ in order to optimise the flux is the main focus of the subsequent analysis. Here, g will be permitted to depend on ω—a significant extension of available results [45,46,54]. This is a natural approach from the perspective of obtaining the best flux for a given frequency. Clearly, a "large" function g would generate more chaotic flux, and hence, a C^0-norm bound for g will be imposed in the form

$$||g|| := \sup_{x \in \Omega} |g(x)| \leq G.$$

(The notation $||\cdot||$ will be used consistently to mean the C^0 or supremum norm in this paper.) In terms of G, can one *quantify* the chaotic transport occurring across the separatrix? Secondly, how can g be chosen subject to this given C^0-norm bound, such that the flux is maximised?

The development for the two- and three-dimensional settings will be given separately in Sections 2 and 3, respectively. Examples for each of these cases, related to models of Rayleigh–Bénard convection and Hill's spherical vortex, will be presented in Section 4.

2. Two-dimensions

2.1. Set-up

In two dimensions, volume-preservation implies that the unperturbed flow is Hamiltonian. The velocity field f in (1) then takes the form

$$f(x) = J\nabla H(x),$$

where $x \in \Omega \subset \mathbb{R}^2$,

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \nabla H = \begin{pmatrix} \frac{\partial H}{\partial x_1} \\ \frac{\partial H}{\partial x_2} \end{pmatrix}.$$

It shall be assumed that $H : \Omega \to \mathbb{R}$ is as smooth as needed. Suppose (1) possesses two hyperbolic fixed points a and b, each with one-dimensional stable and unstable manifolds. Suppose moreover that a branch of W_a^u coincides with a branch of W_b^s to form a one-dimensional heteroclinic manifold Γ. In this instance, this would be defined through one heteroclinic trajectory $\bar{x}(t)$ which decays to a and b exponentially in backwards and forwards time, respectively. There is a freedom in the time parametrisation of this heteroclinic trajectory; as shall be shown later, the particular time parametrisation chosen has no relevance in what follows. Notice also that Γ is representable as a level curve of H, connecting two points (a and b) at which $\nabla H = 0$. The manifold Γ can be parametrised with $t \in \mathbb{R}$ by associating with the point $\bar{x}(-t)$ (cf. the heteroclinic coordinates as described in [44]). It moreover is a flow separatrix, forming an impermeable barrier to phase space flux transport in the unperturbed flow (1).

This inclusion of a perturbation as in (2), where now $g : \Omega \to \mathbb{R}^2$, destroys this separatrix in general. It is well-known that heteroclinic tangles generically result from time-harmonic perturbations. The signed distance between the manifolds $W_a^u(\varepsilon)$ and $W_b^s(\varepsilon)$, as measured in the direction of ∇H at each point t on Γ is given by

$$d(t,\varepsilon) = \varepsilon \frac{M(t)}{|\nabla H(\bar{x}(-t))|} + \mathcal{O}(\varepsilon^2)$$

where $M(t)$ is the Melnikov function, to be defined below [48,49,44]. Simple zeroes of M correspond to transverse intersection between $W_a^u(\varepsilon)$ and $W_b^s(\varepsilon)$ near the point on Γ parametrised by t.

For the perturbation as in (2), the Melnikov function is expressible as

$$M(t) = \int_{-\infty}^{\infty} \nabla H(\bar{x}(\tau)) \cdot g(\bar{x}(\tau)) \cos[\omega(t+\tau-\beta)]\mathrm{d}\tau.$$

(See [48,49] for details.) A useful reformulation (see also [46,54]) is now presented. Define the function

$$\lambda(t) := \nabla H(\overline{x}(-t)) \cdot g(\overline{x}(-t)).$$

Adopt the Fourier transform definition

$$\Lambda(\omega) := \mathcal{F}\{\lambda(t)\}(\omega) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \lambda(t) \exp(-\mathrm{i}\omega t)\mathrm{d}t.$$

Simple trigonometric manipulations yield

$$\begin{aligned} M(t) &= \int_{-\infty}^{\infty} \lambda(-\tau)\{\cos[\omega(t-\beta)] \cos\,\omega\tau - \sin[\omega(t-\beta)] \sin\,\omega\tau\}\mathrm{d}\tau \\ &= \sqrt{2\pi}\{\cos[\omega(t-\beta)]\mathrm{Re}(\Lambda(\omega)) - \sin[\omega(t-\beta)]\mathrm{Im}(\Lambda(\omega))\} \\ &= \sqrt{2\pi}|\Lambda(\omega)| \cos[\mathrm{Arg}(\Lambda(\omega)) + \omega(t-\beta)]. \end{aligned} \tag{3}$$

2.2. Flux computation

The Melnikov function (3) has infinitely many equally spaced transverse zeroes, occurring every ω/π. Fig. 1 gives a schematic of this behaviour; the description which follows is available from any standard textbook [48,49,44]. Fig. 1 should be thought of as a picture of the Poincaré map which samples the flow at time intervals of $2\pi/\omega$. The points a and b are perturbed versions of the initial fixed points, but are now fixed points of the Poincaré map. These each possess manifolds with respect to this Poincaré map. To visualise a separatrix in the perturbed setting, identify any one of the primary intersection points (p.i.p.s), p, between the manifolds. Now define a pseudo-separatrix as being formed by the part of the W_a^u between a and p, joined with the part of W_b^s between b and p (cf. [44]). This pseudo-separatrix is indicated in heavy curves in Fig. 1. Now, apart from the two lobes L_1 and L_2 adjacent to p, all other lobes remain on the same side of the pseudo-separatrix upon iteration of the Poincaré map (this is since they retain their relative position with respect to each manifold). However, lobe L_1 maps to L_3, which is on the *opposite* side of the pseudo-separatrix. Similarly, L_4 maps to L_2. All flux transported across the pseudo-separatrix must travel across this *turnstile* consisting of the four lobes L_1, L_2, L_3 and L_4. In particular, note that the fluid that is *exchanged* across the pseudo-separatrix in one iteration of the Poincaré map is simply the area of L_1 or L_2; these two fluid areas simply exchange places.

The flux interchange that occurs across the separatrix per iteration of the map is therefore related to determining lobe areas. The area of a lobe is given by

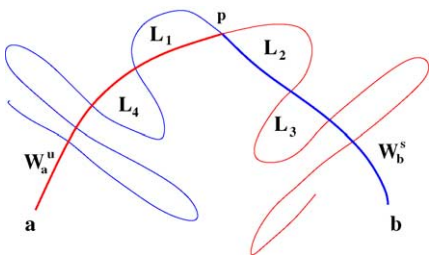

Fig. 1. Lobe dynamics and the pseudo-separatrix (in bold).

$$\text{Area} = \int_{t=t_1}^{t_2} |d(t,\varepsilon)|\,\mathrm{d}l + \mathcal{O}(\varepsilon^2),$$

where $\mathrm{d}l$ is the differential arclength element along the manifold Γ, and t_1 and t_2 are the time-values which characterise the ends of the lobe, say L_1. Since $\mathrm{d}l = |J\nabla H(\bar{x}(-t))|\mathrm{d}t = |\nabla H(\bar{x}(-t))|\mathrm{d}t$, the required lobe area is

$$\begin{aligned}\text{Area} &= \int_{t_1}^{t_2} |d(t,\varepsilon)||\nabla H(\bar{x}(-t))|\,\mathrm{d}t + \mathcal{O}(\varepsilon^2)\\ &= \int_{t_1}^{t_2} \varepsilon\frac{|M(t)|}{|\nabla H(\bar{x}(-t))|}|\nabla H(\bar{x}(-t))|\,\mathrm{d}t + \mathcal{O}(\varepsilon^2)\\ &= \varepsilon\int_{t_1}^{t_2} |M(t)|\,\mathrm{d}t + \mathcal{O}(\varepsilon^2).\end{aligned}$$

The above connection between the Melnikov function and the area was first established in [43]. Define the leading-order (turnstile) lobe area by

$$\mathcal{A} := \int_{t_1}^{t_2} |M(t)|\,\mathrm{d}t.$$

By (3) the Melnikov function is simply a shifted sinusoidal with equally spaced zeroes, and hence the lobe areas L_1 and L_2 must be the same. Since $\mathcal{A}$ is the amount of phase space 'volume' which transfers across the heteroclinic manifold in $2\pi/\omega$ time units, the *leading-order flux* (transported 'volume' of fluid per unit time) would be reasonably defined by [45,53,46]

$$s(\omega) := \frac{\omega}{2\pi}\mathcal{A} = \frac{\omega}{2\pi}\int_{t_1}^{t_2} |M(t)|\,\mathrm{d}t.$$

Substituting from (3),

$$s(\omega) = \frac{\omega}{2\pi}\sqrt{2\pi}|\Lambda(\omega)|\int_{t_1}^{t_2} |\cos[\mathrm{Arg}(\Lambda(\omega)) + \omega(t-\beta)]|\,\mathrm{d}t.$$

Since t_1 and t_2 are adjacent zeroes of (3), they are adjacent zeroes of the integrand above. The quantity $\mathrm{Arg}(\Lambda(\omega)) - \omega\beta$ is a simple phase shift; one may replace the integral above with the integral of $|\cos\omega t|$ between any two adjacent zeroes. Hence,

$$s(\omega) = \frac{\omega}{\sqrt{2\pi}}|\Lambda(\omega)|2\int_0^{\pi/(2\omega)} \cos(\omega t)\mathrm{d}t = \sqrt{\frac{2}{\pi}}|\Lambda(\omega)|.$$

Thus, the flux corresponding to a perturbation of frequency ω, therefore, has a direct relationship to a Fourier transform:

$$s(\omega) = \sqrt{\frac{2}{\pi}}|\mathcal{F}\{\nabla H(\bar{x}(t))\cdot g(\bar{x}(t))\}(\omega)|. \tag{4}$$

The formula (4) is valid for any perturbation of the form (2). The current issue is to determine the form of g to maximise this leading-order flux.

2.3. Choosing g to optimise flux

Let the Fourier transform of $[\nabla H \cdot g](\bar{x}(t))$ have polar representation $R(\omega)\exp[i\theta(\omega)]$. The phase $\theta(\omega)$ has no effect on the flux $s(\omega)$ by (4). Now, for any real α,

$$\mathcal{F}\{(\nabla H \cdot g)(\bar{x}(t-\alpha))\}(\omega) = \mathrm{e}^{\mathrm{i}\omega\alpha}\mathcal{F}\{(\nabla H \cdot g)(\bar{x}(t))\}(\omega) = R(\omega)\mathrm{e}^{\mathrm{i}[\omega\alpha+\theta(\omega)]}.$$

Choosing α corresponds to picking a time parametrisation along Γ. For the choice $\alpha = (-\theta(\omega) + 2m\pi)/\omega$ for any integer m, the Fourier transform is real. Since $\theta(\omega)$ is unknown, the precise value of this time shift is unclear; yet there is a time parametrisation for which this works. Applying (4) for this purely real Fourier transform,

$$s(\omega) = \frac{1}{\pi}\left|\int_{-\infty}^{\infty}(\nabla H \cdot g)(\bar{x}(t))\cos\,\omega t\,\mathrm{d}t\right|.$$

Now $\cos\,\omega t$ is negative in alternating bands of t of width π/ω. To maximise s, it would help if $\nabla H \cdot g$ were positive whenever $\cos\,\omega t$ was, and negative if else. To effect this, define the set

$$\mathbb{T} = \bigcup_{n\in\mathbb{Z}}\left[\frac{(4n-1)\pi}{2\omega}, \frac{(4n+1)\pi}{2\omega}\right),$$

where $\mathbb{Z}$ is the set of integers. Partition the heteroclinic manifold Γ into two alternating portions

$$\Gamma_+ := \{\bar{x}(t) : t \in \mathbb{T}\} \quad \text{and} \quad \Gamma_- := \Gamma \setminus \Gamma_+.$$

Fig. 2 displays these sets. Each comprises countably many segments, diminishing in length as the endpoints a and b are approached. Now define on Γ the function

$$\bar{g}(x) := G\,\frac{\nabla H(x)}{|\nabla H(x)|}[\mathbb{I}_{\Gamma_+}(x) - \mathbb{I}_{\Gamma_-}(x)],$$

where $\mathbb{I}$ is the indicator function. While $\bar{g}$ is not smooth, one can obtain smooth functions g which are arbitrarily close to $\bar{g}$ in the C^0-norm, by suitably connecting the piecewise segments with C^∞-bump functions. Any such function can then be extended smoothly to Ω. The upper bound for the flux $\bar{s}(\omega)$, which can be approached as closely as required, is therefore

$$\bar{s}(\omega) = \frac{G}{\pi}\int_{-\infty}^{\infty}|\nabla H(\bar{x}(t))||\cos\,\omega t|\mathrm{d}t.$$

This expression was derived assuming that a time parametrisation for Γ resulting in a real Fourier transform was made. If a phase shift of $\pi/2$ were instead present, the Fourier transform would be purely imaginary, necessitating the replacement of the cosine above by a sine (the definitions for $\mathbb{T}$, Γ_+ and Γ_- need to be appropriately modified). In practice, it is difficult to know whether the chosen parametrisation relates to these two extremes, or to something intermediate. Hence, in order to handle *all* possible time parametrisations, the general formula for the maximum flux obtainable through a perturbation of frequency ω is

$$\bar{s}(\omega) = \frac{G}{\pi}\sup_{\psi\in[0,\pi)}\int_{-\infty}^{\infty}|\nabla H(\bar{x}(t))||\cos(\omega t + \psi)|\mathrm{d}t. \tag{5}$$

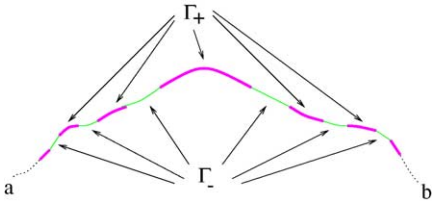

Fig. 2. The subsets Γ_+ and Γ_- of Γ.

This automatically compensates for the original phase angle β as well. Realising that it is only the normal component $g_\perp$ to the manifold Γ which truly contributes to this flux, the implication of (5) for general (unoptimised) perturbations is

$$s(\omega) < \frac{1}{\pi} \sup_{x\in\Gamma} |g_\perp(x)| \sup_{\psi\in[0,\pi)} \int_{-\infty}^{\infty} |\nabla H(\bar{x}(t))| |\cos(\omega t + \psi)| \mathrm{d}t, \tag{6}$$

with equality being approached as $g \to \bar{g}$. Note also from Section 2.2 that $|\nabla H|\mathrm{d}t = \mathrm{d}l$. Let $t = \bar{t}(l)$ be the function expressing the monotonic relationship between the time t and arclength l parametrisation along Γ. Then, (5) has the alternative representation

$$\bar{s}(\omega) = \frac{G}{\pi} \sup_{\psi\in[0,\pi)} \int_{\Gamma} |\cos(\omega \bar{t}(l) + \psi)| \mathrm{d}l. \tag{7}$$

2.4. Optimal flux discussion

The time-periodic switching of g predicted through this analysis is no surprise. The effect of g being in a normal direction to Γ would be to push a manifold in that direction; in the next segment (in which the manifolds have exchanged relative positions), this same manifold would benefit from being pushed in the opposite direction. This would increase the lobe areas, thereby increasing chaotic flux.

The Eq. (5) answers the question as to the maximum flux possible for perturbations of the form given in (2), for a given ω (note that the entities $\mathbb{T}$, Γ_+ and Γ_- are each ω-dependent). Since the flow on the unperturbed heteroclinic is exponentially slow near the critical points a and b, g must switch back and forth in very small regions of phase-space near these points in order to achieve the best chaotic mixing. However, the quantitative effect of this on (5) is quite small, given the exponential damping on $|\nabla H|$ that occurs. The main contributions arise from switchbacks occurring in regions far removed from the critical points. If g were *not* permitted to be chosen depending on ω, the results will be different (indeed, there would be no optimisation problem, as in [45,46]). However, the current question is more natural—it gives a limiting value of the flux achievable for norm-limited perturbations of any given frequency.

While $\bar{s}$ will in reality never be achievable for smooth perturbations, arbitrarily closeness is possible. For example, the normal could be modulated with a cosine term so that g has the form

$$g(x) = G \frac{\nabla H(x)}{|\nabla H(x)|} \cos[\omega \bar{t}(x)]$$

for $x \in \Gamma$. Using many Fourier modes (terms involving $\cos[2\omega\bar{t}(x)]$, $\cos[3\omega\bar{t}(x)]$, etc.) could make this approach $\bar{g}$ as closely as required. This is analogous to the approximation of a discontinuous wave with a Fourier series.

The formulæ (5) and (7) provide the optimal flux for a given (fixed) frequency ω. A natural question would be whether the flux can be optimised over *all* ω. In other words, is there a limit on the flux achievable through any sort

of sinusoidal perturbation? Notice from (7) that

$$\overline{s}(\omega) \leq \overline{s}_m := \frac{G}{\pi}\int_\Gamma \mathrm{d}l = \frac{G\,\mathrm{length}(\Gamma)}{\pi}, \tag{8}$$

which could be more transparently written as

$$\overline{s}(\omega) \leq \frac{\mathrm{length}(\Gamma)}{\pi}\sup_{x\in\Gamma}|g_\perp(x)|. \tag{9}$$

Is the upper bound achievable? Observe from (7) that

$$\overline{s}(0^+) = \frac{G}{\pi}\sup_{\psi\in[0,\pi)}\int_\Gamma |\cos\,\psi|\,\mathrm{d}l = \overline{s}_m.$$

Here, $\omega = 0^+$ is the adiabatic limit corresponding to vanishingly small frequency, or alternatively, infinitely large periodicity. By choosing ω smaller and smaller, the upper limit in (8) can be approached as closely as required. This is qualitatively consistent with the numerically obtained optimal flux for a corner vortex in [25]. However, as pointed out in [45,53], there is an inconsistency in taking the $\omega \to 0^+$ limit in the Melnikov setting, which explicitly presupposes small ε but non-small ω. The difficulty arises in using $s(\omega)$ to approximate the flux, which is actually of the form $\varepsilon s(\omega) + \mathcal{O}(\varepsilon^2)$. As $\omega \to 0^+$, its smallness competes with that of ε, and the $\mathcal{O}(\varepsilon^2)$ terms in the flux may acquire importance. Presumably there is a function $\overline{\omega}(\varepsilon)$ such that $s(\omega)$ and $\overline{s}(\omega)$ are only legitimate representations for the leading-order flux if $\omega > \overline{\omega}(\varepsilon)$. Characterising $\overline{\omega}(\varepsilon)$ is non-trivial, as is the singular limit $\omega \to 0$ in the current context.

For valid (non-infinitesimal) ω, it will now be shown that $\overline{s}(\omega)$ is monotonically decreasing. For each fixed ψ and ω, define a partition of Γ with respect to its arclength parametrisation as follows. Let

$$\mathcal{L}_n(\omega) := [\mathcal{L}_n^-(\omega), \mathcal{L}_n^+(\omega))$$

where $\mathcal{L}_n^\pm(\omega)$ are the length parametrisations l of Γ corresponding to the times

$$t_n^\pm = \frac{1}{\omega}\left[\frac{(2n\pm 1)\pi}{2} - \psi\right].$$

Then, Γ can be represented as the countable union of the sets $\mathcal{L}_n$ for $n \in \mathbb{Z}$. As ω varies this partition changes, yet remains denumerable. The arclength intervals $\mathcal{L}_n$ have a strong connection to the alternating segments of Γ_+ and Γ_- of Section 2.3, with a possible time shifting of ψ additionally included. Now, from (7),

$$\overline{s}(\omega) = \frac{G}{\pi}\sup_{\psi\in[0,\pi)}\bigcup_{n\in\mathbb{Z}}\int_{\mathcal{L}_n(\omega)} |\cos(\omega\overline{t}(l) + \psi)|\,\mathrm{d}l$$

As ω is increased, the length of the segment $\mathcal{L}_n$ decreases, since the time difference between the endpoints $\mathcal{L}_n^-$ and $\mathcal{L}_n^-$ decreases. Now, the integral is of a cosine curve over one-half its period, subject to nonlinear modulations through the monotonic function $\overline{t}(l)$. With the increase in ω, this curve gets squashed horizontally as the length of the interval diminishes. Thus, each integral over $\mathcal{L}_n$ will decrease with the increase of ω, the end result being that the total integral decreases. Therefore, it has been argued that $\overline{s}(\omega)$ is a decreasing function of ω. The underlying principle is that smaller frequencies can generate larger optimal chaotic fluxes.

There is an apparent contradiction with the results of [45,53,43], whose flux functions are non-monotonic in ω. However, the monotonicity has been argued here for the optimal flux function $\overline{s}(\omega)$ (in which g was chosen

depending on ω), and not $s(\omega)$. An optimising problem has been solved here, unlike in [45,53,43] in which g was ω-independent.

The decreasing function $\bar{s}(\omega)$ is bounded from below, and therefore, must have a limit as $\omega \to \infty$. In this limit $|\cos(\omega t + \psi)|$ oscillates so rapidly that its contribution to the integral in (5) is effectively just its average value, $2/\pi$. Thus

$$\lim_{\omega\to\infty} \bar{s}(\omega) = \frac{2}{\pi}\frac{G}{\pi} \sup_{\psi\in[0,\pi)} \int_{-\infty}^{\infty} |\nabla H(\bar{x}(t))|\,\mathrm{d}t = \frac{2\,G\,\text{length}(\Gamma)}{\pi^2} = \frac{2}{\pi}\bar{s}_m.$$

Hence, $\bar{s}(\omega)$ reduces monotonically from a value of $\bar{s}_m$ when $\omega = 0$, to $(2/\pi)\bar{s}_m$ as $\omega \to \infty$. The flux function in [45,53,43,54] was shown to go to zero as $\omega \to \infty$; the current result again does not contradict this since an optimisation approach has been followed. The interesting implication is that, for large frequencies, there is only marginal change in the optimum flux propensity as ω is varied.

The behaviour of the optimal flux function as described above, is completely consistent with the optimising approach of [25]. They numerically solve an optimisation problem for a more heuristic flux definition than used in this paper, and for a specific corner vortex problem. Nevertheless, they argue that the optimal flux function increases as ω decreases (Appendix B), and decreases to a non-zero value as $\omega \to \infty$ (Appendix C). Their numerical results also bear this out. The current theoretical study displays that their results are generically to be expected in such settings, though the technique (unlike in [25]) is confined to time-harmonic perturbations. An extension of the theory to more general time-periodicity (as possible in the numerical algorithm in [25]) seems possible through the usage of the ideas in [46]. However, an important observation that emerges is the distinction between the *optimal* flux function as discussed in [25] and the present article, and the non-optimised flux function of [45,53,43].

The lack of continuity of $\bar{g}$ also leads to the question as to whether the derived theory is applicable in realistic fluid mechanical settings in which incompressibility is also enforced (at this point no such condition has been used for g). Let g_T and g_N represent the tangential and normal components (to Γ) of g, respectively. Incompressibility would mean satisfying

$$\frac{\partial g_T(x)}{\partial T} + \frac{\partial g_N(x)}{\partial N} = 0, \tag{10}$$

for all x. Now, the limiting function $\bar{g}$ has $\bar{g}_T = 0$ on Γ, but has $\bar{g}_N$ undefined at countably many points, and therefore, cannot satisfy this differential form of incompressibility. However, there is no bar to $\bar{g}$ satisfying the *integral* form of incompressibility. One can construct $\bar{g}$ locally (near Γ) by setting up "streams" of constant speed G crossing Γ in opposite directions, and then proceeding out into Ω such that these streams always remain adjacent, non-intersecting, and have constant cross-section. While this contrived situation may be volume-preserving, it violates other fluid mechanical conditions (it has infinite rate-of-strain and vorticity at the stream interfaces).

On the other hand, there is no difficulty for a "nearby" smooth function g to satisfy the incompressibility condition above. Note that we only need to specify g_N on Γ, such that it has a value sufficiently close to $\bar{g}$. This leaves the freedom to define g_T on Γ, and also to specify g just off Γ. The only restriction in either of these endeavours is that $|g_N|^2 + |g_T|^2 \leq G^2$. Suppose on Γ we chose g_T (smoothly) which satisfies this. Then $\partial g_T/\partial T$ may be not be zero—but we can extend g_N off Γ to ensure that the incompressibility condition (10) is satisfied on Γ.

A possible practical attempt to make g as close as possible to $\bar{g}$ would be to try to set up "streams" of speed G in opposite directions across Γ as described above. Presumably, the flow will adjust along the stream interfaces to provide a smooth transition, thereby generating a physically viable g which is close to (but not equal to) $\bar{g}$. The locations of these interfaces of reversing direction is given by the transition points of $\Gamma_\pm$, which correspond to t differences of π/ω along the heteroclinic. However, these locations are impossible to determine generally, because of lack of knowledge of the phase shift ψ (occurring in (5), (6) or (7)). It may be that ψ could be varied experimentally in the domain $[0, \pi)$, or in some cases where full information on the unperturbed flow is available, determined analytically (as is possible in the examples in Section 4). Given also that the exponential damping

present in (5) makes the contributions to the flux from the reversals near the critical points small, a finite number of such reversals may give a flux which is near $\overline{s}$. It should be once again emphasised that this g is not $\overline{g}$ because of the inevitable adjustment that must occur along the stream interfaces to prevent violations of fluid laws. This form of g is reminiscent of cross-channel micro-mixers [20,21,1,55]. However, it provides additional information; for example it suggests that the spatial widths of the cross-stream channels should not be the same (but be given by constant time parametrisation changes of π/ω).

3. Three dimensions

3.1. Set up

The ideas in two dimensions can be extended to three dimensions in some cases. Once again the unperturbed flow is as given in (1), where now $x \in \Omega \subset \mathbb{R}^3$, and $f : \Omega \to \mathbb{R}^3$ is as smooth as required. It will be necessary to make the following hypotheses.

(H1) The vector field of (1) is volume-preserving, i.e., div $f = 0$.
(H2) The flow of (1) is axisymmetric, and has no swirl around the axis of symmetry.
(H3) The system (1) possesses two hyperbolic critical points a and b on the axis of symmetry such that W^u_a and W^s_b are two-dimensional.
(H4) A branch of W^u_a coincides with a branch of W^s_b to form a two-dimensional heteroclinic manifold Γ.

The prototypical example that should be kept in mind is that of Hill's spherical vortex [56] from fluid mechanics; more details on this appear in Section 4.2. The volume-preservation and axisymmetry assumptions together give the presence of a *Stokes streamfunction* $H(x)$ [56]. It is convenient to orient the axis of symmetry in the Cartesian z-direction, and employ spherical polar coordinates (r, θ, ϕ). The system (1) is expressible in terms of the Stokes streamfunction as

$$u_r = \frac{1}{r^2 \sin\theta}\frac{\partial H}{\partial \theta}, \quad u_\theta = -\frac{1}{r \sin\theta}\frac{\partial H}{\partial r}, \quad u_\phi = 0,$$

where $f = u_r\hat{r} + u_\theta\hat{\theta} + u_\phi\hat{\phi}$ is the spherical polar representation of the vector field in (1) (cf. [56]).

The basic geometry is shown in Fig. 3. The heteroclinic manifold Γ is topologically a sphere punctured at the poles, and composed of a collection of heteroclinic trajectories which can be indexed by ϕ. Since H is conserved by the flow [56], $H(a) = H(b)$, and since these are the endpoints of any of the heteroclinic trajectories, this also means that H = constant on Γ. Moreover, given the dimensions of the stable and unstable manifolds at a and b, it is clear that ∇H is zero at both a and b. Along any heteroclinic trajectory x_ϕ, however, ∇H is nonzero (though approaching zero as $t \to \pm\infty$) and orientation preserving (consistently points either into or out of Γ). One can parametrise Γ with (t, ϕ) by identification with the point $\overline{x}_\phi(-t)$; ϕ chooses a particular heteroclinic trajectory $\overline{x}_\phi$, and t the reverse time along it. Here, $\phi \in [0, 2\pi)$ and $t \in \mathbb{R}$, and as before, there is ambiguity in this parametrisation related to choosing time-zero.

Consider now (2) as the perturbed flow. Here, conditions such as volume-preservation, axisymmetry, no-swirl, etc, are *not* in general assumed for the perturbed flow (though certain special cases will be considered later). Upon perturbation, W^u_a and W^s_b need no longer coincide, and split apart. A higher-dimensional Melnikov method exists for exactly this situation [50,51]. The splitting between the manifolds is representable through a signed distance

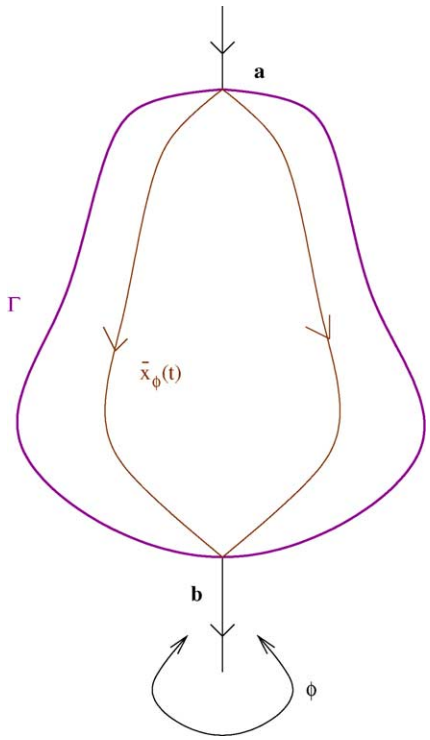

Fig. 3. The heteroclinic manifold Γ, with a couple of heteroclinic trajectories. Notice that the azimuthal coordinate ϕ can be used to index the heteroclinics $\bar{x}_\phi$.

function $d(t, \phi, \varepsilon)$, which measures their separation along a transversal $\nabla H(\bar{x}_\phi(-t))$ at the point (t, ϕ) on Γ. In fact,

$$d(t, \phi, \varepsilon) = \varepsilon \frac{M(t, \phi)}{|\nabla H(\bar{x}_\phi(-t))|} + \mathcal{O}(\varepsilon^2), \tag{11}$$

where $M(t, \phi)$ is the Melnikov function. Not surprisingly, this is defined by

$$M(t, \phi) = \int_{-\infty}^{\infty} \nabla H(\bar{x}_\phi(\tau)) \cdot g(\bar{x}_\phi(\tau)) \cos[w(t + \tau - \beta)] \mathrm{d}\tau. \tag{12}$$

(See [52,33,39,24] for Melnikov transport applications in three dimensions, in addition to the initial theoretical Melnikov developments [50,51].) Through the definition of

$$\lambda_\phi(t) := \nabla H(\bar{x}_\phi(-t)) \cdot g(\bar{x}_\phi(-t))$$

and its Fourier transform $\Lambda_\phi(\omega)$ exactly as in in Section 2.1, the Melnikov function is expressible in the form

$$M(t, \phi) = \sqrt{2\pi} |\Lambda_\phi(\omega)| \cos[\mathrm{Arg}(\Lambda_\phi(\omega)) + \omega(t - \beta)]. \tag{13}$$

In proceeding with the phase space structure determination, the trivial case in which g is itself axisymmetric will be considered first. This will help in extending the flux formula to general separable non-axisymmetry.

3.2. Axisymmetric perturbation

If g is axisymmetric, $\lambda_\phi(t)$ and $\Lambda_\phi(\omega)$ have no dependence on ϕ. The Melnikov function is then independent of the longitude chosen. It is sufficient to picture the topological intersection between W_a^u and W_b^s along one such longitude. As before, M has periodic zeroes, resulting in the heteroclinic tangle illustrated in Fig. 4. This is topologically equivalent to Fig. 1 from the two-dimensional analysis.

However, in reality each turnstile lobe is a three-dimensional entity—it is the shaded region in Fig. 4, rotated about the z-axis to form a *turnstile ring*. It will be necessary to determine the volume of this ring. An apparent complication is that the arclength element along the heteroclinic is no longer given by $|\nabla H(\bar{x}(-\tau))|\,\mathrm{d}\tau$ as before, since $|f| \neq |\nabla H|$. The volume of the turnstile ring is

$$\text{Volume} = \int_{t_1}^{t_2}\int_0^{2\pi} |d(\tau,\varepsilon)|h_\phi\,\mathrm{d}\phi|\,f(\bar{x}(-\tau))|\,\mathrm{d}\tau,$$

where $h_\phi = r\,\sin\theta$ is the scale factor in the $\hat{\phi}$ direction, and t_1 and t_2 are two adjacent zeroes of the Melnikov function. Now,

$$|f| = \sqrt{\left(\frac{1}{r^2\,\sin\theta}\frac{\partial H}{\partial\theta}\right)^2 + \left(\frac{-1}{r\,\sin\theta}\frac{\partial H}{\partial r}\right)^2} = \frac{1}{h_\phi}|\nabla H|,$$

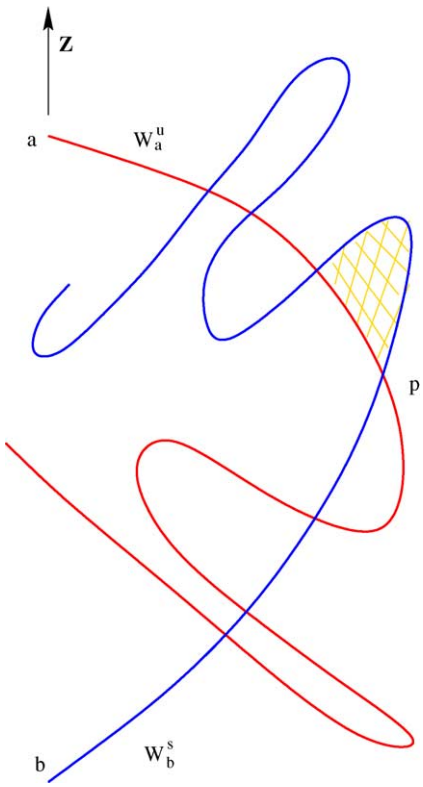

Fig. 4. Typical intersection pattern after axisymmetric perturbations.

and the volume is given by

$$\text{Volume} = \int_{t_1}^{t_2} \int_0^{2\pi} |d(\tau, \varepsilon)|\, |\nabla H(\bar{x}(-\tau))|\, \mathrm{d}\phi\, \mathrm{d}\tau + \mathcal{O}(\varepsilon^2).$$

Substituting for d from (11), and keeping only the $\mathcal{O}(\varepsilon)$ terms gives the *leading-order volume*

$$\mathcal{V} = \left(\int_0^{2\pi} \mathrm{d}\phi \right) \left(\int_{t_1}^{t_2} |M(\tau)|\, \mathrm{d}\tau \right) = 2\pi \int_{t_1}^{t_2} |M(\tau)|\, \mathrm{d}\tau.$$

Hence the *leading-order flux* is

$$s(\omega) := \frac{\omega}{2\pi} \mathcal{V} = \omega \int_{t_1}^{t_2} |M(\tau)|\, \mathrm{d}\tau.$$

Substituting from (13), and following the identical calculation as in Section 2.2,

$$s(\omega) = 2\sqrt{2\pi} |\mathcal{F}\{\nabla H(\bar{x}(t)) \cdot g(\bar{x}(t))\}(\omega)|,$$

where the ϕ-dependence in $\bar{x}$ has been suppressed since choosing any longitude gives the same result. Not surprisingly, there is no qualitative difference between this formula and the two-dimensional version (4).

3.3. Non-axisymmetric modes

At the next level of complexity, assume that the non-axisymmetry in g consists only of "modes" defined through

$$g(r, \theta, \phi) = \cos(k\phi) h(r, \theta),$$

where k is a positive integer, and h takes values in $\mathbb{R}^3$. This incorporates a mode of wavenumber k in the azimuthal direction. For convenience, let $y \equiv (r, \theta)$, thereby representing each unperturbed heteroclinic trajectory as $\bar{x}_\phi \equiv (\bar{y}, \phi)$. Since the unperturbed flow is axisymmetric, $\nabla H(\bar{x}_\phi(\tau)) = \nabla H(\bar{y}(\tau))$ is independent of ϕ. Thus

$$\nabla H(\bar{x}_\phi(\tau)) \cdot g(\bar{x}_\phi(\tau)) = (\cos\, k\phi) \nabla H(\bar{y}(\tau)) \cdot h(\bar{y}(\tau)),$$

and therefore, $\cos\, k\phi$ can be pulled out of the integral in (12). The function playing the role of $\lambda(t)$ is then seen to be

$$\tilde{\lambda}(t) = \nabla H(\bar{y}(-t)) \cdot h(\bar{y}(-t)),$$

and let its Fourier transform be $\tilde{\Lambda}(\omega)$. Following the process used to obtain (13),

$$M(t, \phi) = \cos\, k\phi \sqrt{2\pi} |\tilde{\Lambda}(\omega)|\, \cos[\mathrm{Arg}(\tilde{\Lambda}(\omega)) + \omega(t - \beta)]. \tag{14}$$

Fix a ϕ such that $\cos\, k\phi \neq 0$. Along this longitude of Γ, the zeroes of M occur in exactly the same way as for the two-dimensional case (3) and the three-dimensional axisymmetric case (13). On the other hand, each fixed t such that the last cosine term in (14) is nonzero corresponds to a fixed latitude of Γ whose zeroes distribution is governed by the term $\cos\, k\phi$. For example, if $k = 3$, there would be six zeroes at $\phi = \pi/6, \pi/2, 5\pi/6, 7\pi/6, 3\pi/2$, and $11\pi/6$. A schematic of the resulting intersection pattern of the perturbed manifolds is shown in Fig. 5. The three-dimensional

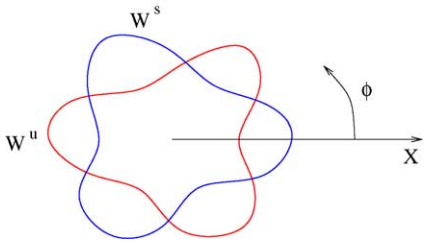

Fig. 5. Generic (exaggerated) intersection pattern in a slice of constant latitude. This picture corresponds to $k = 3$.

picture of manifold intersection is then a combination of Figs. 4 and 5. As one rotates Fig. 4 (corresponding to changing ϕ), the intersecting manifolds change their relative positions at each ϕ value which is a zero of M. The end result is that the perturbed manifolds form a rippling intersection pattern in both the t and the ϕ directions.

To make sense of the volume of transferred fluid at each iteration of the Poincaré map, begin with the simpler axisymmetric picture of Fig. 4. Notice from (14) that this would be the observed intersection in any longitudinal slice ϕ (such that $\cos\, k\phi$ is not identically zero) of the now non-axisymmetric flow. Each intersection point persists under rotation about the initial axis of symmetry, since from (14) the t and ϕ dependencies decouple. Pick one such point, and its rotated circuit, and identify this in $\mathbb{R}^3$ as a *primary intersection circuit* (p.i.c.). Define the pseudo-separatrix by taking the unstable manifold emanating from a up to this p.i.c., and joining this with the stable manifold emanating from b up to this p.i.c. Now, the turnstile structure relates to the two 'rings' adjacent to the p.i.c. In this case the rings are not uniformly solid as one travels around the axis—each 'ring' has k non-connected pieces. This is because the stable and unstable manifolds exchange their relative positions $2k$ times as $\phi : 0 \to 2\pi$; it is only alternating pieces which are inside the pseudo-separatrix. Upon iteration of the Poincaré map, k pieces cross the pseudo-separatrix in one direction, while the other k pieces cross it in the other direction. In the other adjacent ring, the same behaviour occurs in the opposite direction. Thus, the volume of $2k$ pieces is exchanged across the pseudo-separatrix during one periodic cycle.

Each of these pieces has the same volume, and so

$$\text{Volume} = 2k \int_{t_1}^{t_2} \int_{\phi_1}^{\phi_2} |d(\tau,\phi)| h_\phi \,\mathrm{d}\phi |f(\bar{y}(-\tau))| \,\mathrm{d}\tau + \mathcal{O}(\varepsilon^2),$$

where t_1 and t_2 are two adjacent zero points of M with respect to t, and ϕ_1 and ϕ_2 are similarly adjacent zero points with respect to ϕ. But, as in Section 3.2, $h_\phi |f| = |\nabla H|$, leading to

$$\text{Volume} = 2k \int_{t_1}^{t_2} \int_{\phi_1}^{\phi_2} |d(\tau,\phi)| \,\mathrm{d}\phi |\nabla H(\bar{y}(-\tau))| \,\mathrm{d}\tau + \mathcal{O}(\varepsilon)^2 = 2\varepsilon k \int_{t_1}^{t_2} \int_{\phi_1}^{\phi_2} |M(\tau,\phi)| \,\mathrm{d}\phi \,\mathrm{d}\tau + \mathcal{O}(\varepsilon^2).$$

The leading-order volume is the $\mathcal{O}(\varepsilon)$ term of this:

$$\mathcal{V} = 2k\sqrt{2\pi}|\tilde{\Lambda}(\omega)| \left(\int_{\phi_1}^{\phi_2} |\cos\, k\phi| \,\mathrm{d}\phi \right) \left(\int_{t_1}^{t_2} |\cos[\mathrm{Arg}(\tilde{\Lambda}(\omega)) + \omega(\tau - \beta)]| \,\mathrm{d}\tau \right).$$

The two integrals yield $2/k$ and $2/\omega$, respectively. The leading-order flux is therefore

$$s(\omega) := \frac{\omega}{2\pi}\mathcal{V} = 4\sqrt{\frac{2}{\pi}} |\mathcal{F}\{\nabla H(\bar{y}(t)) \cdot h(\bar{y}(t))\}(\omega)|. \tag{15}$$

Notice the somewhat surprising conclusion that the flux is independent of the azimuthal wavenumber k.

The next issue is to analyse the choice of h which would lead to maximising $s(\omega)$. Since $\bar{y}$ is two-dimensional, the construction is the same as in Section 2.3, with h chosen to be a vector of maximum permitted size which alternates in direction between pointing in to, and out of, Γ. However, $||g|| = ||\cos\, k\phi||\, ||h(r,\theta)|| = ||h||$, and hence to optimise choose $||h|| = G$. This leads to the expression for the leading-order optimal flux

$$\bar{s}(\omega) = \frac{4G}{\pi} \sup_{\psi\in[0,\pi)} \int_{-\infty}^{\infty} |\nabla H(\bar{y}(t))|\, |\cos(\omega t + \psi)|\, \mathrm{d}t.$$

3.4. Separable non-axisymmetry

Now consider the more general form

$$g(r,\theta,\phi) = p(\phi)h(r,\theta)$$

for some function p taking values in $\mathbb{R}$, with h being three-dimensional. In this case, the perturbation has a separable scalar azimuthal function, which generalises the modal form of Section 3.3.

The Melnikov function (14) can be modified by simply replacing $\cos\, k\phi$ with $p(\phi)$; the zeroes of p therefore govern the azimuthal intersection pattern of the perturbed manifolds. Each of two turnstile 'rings' adjacent to the p.i.c. will be broken into pieces as before; these pieces are now defined in terms of whether $p(\phi) > 0$ or $p(\phi) < 0$. In contrast to Section 3.3, these pieces need not have the same volumes. Focus first on one turnstile ring, say the upper one. The pieces corresponding to $p(\phi) > 0$ will move across the pseudo-separatrix in one direction (say inwards) upon iteration of the Poincaré map, whereas the pieces with $p(\phi) < 0$ will go across the other way. In contrast, in the lower turnstile ring, pieces with $p(\phi) > 0$ will move outwards. The total volume transported across the pseudo-separatrix therefore contains both the portions $p(\phi) > 0$ and $p(\phi) < 0$. The fact that there correspond to different turnstile rings does not matter in the volume computation, since the upper and lower turnstile rings have the same volume. Thus, the quantity $2k\int_{\phi_1}^{\phi_2} \cos\, k\phi\, \mathrm{d}\phi$ appearing in the volume expressions in Section 3.3 should be replaced by $\int_0^{2\pi} |p(\phi)|\, \mathrm{d}\phi$. Then, (15) receives the modification

$$s(\omega) = \left(\sqrt{\frac{2}{\pi}} \int_0^{2\pi} |p(\phi)|\, \mathrm{d}\phi\right) |\mathcal{F}\{\nabla H(\bar{y}(t)) \cdot h(\bar{y}(t))\}(\omega)|. \tag{16}$$

Eq. (16) gives an explicit expression for the chaotic flux for perturbations of this specific form. The remaining equations in Section 3.3 for $\bar{s}$ are therefore easily corrected as well, for example the optimum flux

$$\bar{s}(\omega) = \left(\frac{||h||}{\pi} \int_0^{2\pi} |p(\phi)|\, \mathrm{d}\phi\right) \sup_{\psi\in[0,\pi)} \int_{-\infty}^{\infty} |\nabla H(\bar{y}(t))|\, |\cos(\omega t + \psi)|\, \mathrm{d}t. \tag{17}$$

Here, the supremum norm $||h||$ satisfies $||h(r,\theta)||\, ||p(\phi)|| = G$, where $||\cdot||$ indicates C^0-norms associated with the appropriate domains. For unoptimised separable g this implies that

$$s(\omega) < 2 \sup_{x\in\Gamma} |g_\perp(x)| \sup_{\psi\in[0,\pi)} \int_{-\infty}^{\infty} |\nabla H(\bar{y}(t))|\, |\cos(\omega t + \psi)|\, \mathrm{d}t, \tag{18}$$

with equality being approached as $g \to \bar{g}$. Notice that it is possible that p has no zeroes; in this case, the picture is very similar to the axisymmetric perturbation case, since there will not be any topological variation of perturbed manifold structure in the azimuthal direction. If so, the turnstile rings would be solid entities. When p does have transverse zeroes, the topological structure would be a non-uniform analogue of the non-axisymmetric modal case. As argued in Section 2.4 for two-dimensions, the development of the formula for the optimal flux $\bar{s}(\omega)$ is consistent even within an incompressible setting for the perturbing velocity g.

Now, the area element on Γ is given by

$$\mathrm{d}A = (h_\phi \,\mathrm{d}\phi)\left(\frac{|\nabla H(\bar{x}_\phi(-t))|}{h_\phi}\mathrm{d}t\right) = |\nabla H(\bar{y}(-t))|\,\mathrm{d}\phi\,\mathrm{d}t,$$

since $|f| = |\nabla H|/h_\phi$ is the velocity, and the unperturbed flow is axisymmetric. This enables (17) to be reformulated as

$$\bar{s}(\omega) = \frac{G}{\pi||p||}\sup_{\psi\in[0,\pi)}\int\int_\Gamma |p(\phi)|\,|\cos(\omega t+\psi)|\,\mathrm{d}A, \tag{19}$$

analogous to (7) in the two-dimensional setting. Hence

$$\bar{s}(\omega) \le \frac{G}{\pi}\sup_{\psi\in[0,\pi)}\int\int_\Gamma |\cos(\omega t+\psi)|\,\mathrm{d}A,$$

This can be bounded over all ω by

$$\bar{s}(\omega) \le \frac{G\,\mathrm{area}(\Gamma)}{\pi} =: \bar{s}_m. \tag{20}$$

Again, this is more informatively expressed as

$$\bar{s}(\omega) \le \frac{\mathrm{area}(\Gamma)}{\pi}\sup_{x\in\Gamma}|g_\perp(x)|. \tag{21}$$

Note the pleasing similarity between (21) and (9). Now using (19),

$$\bar{s}(0^+) = \frac{G}{\pi||p||}\int\int_\Gamma |p(\phi)|\,\mathrm{d}A$$

which is not necessarily the same as $\bar{s}_m$ in (20). In order for $\bar{s}(0^+)$ to approach the supremum value of $\bar{s}$, one needs to choose $p(\phi)$ to be a constant function. Azimuthal complications apparently do not enhance the flux.

By pulling out the ϕ-dependence, one can use (19) and the argument in Section 2.4 to show that the optimum flux $\bar{s}(\omega)$ is monotonically decreasing in ω from its supremum value at 0^+, and levels out to a limit as $\omega\to\infty$. This limit is

$$\bar{s}(\infty) = \frac{G}{\pi||p||}\frac{2}{\pi}\int\int_\Gamma |p(\phi)|\,\mathrm{d}A = \frac{2}{\pi}s(0^+).$$

4. Examples

4.1. Planar cellular flow

As a two-dimensional example, consider planar cellular flow in which

$$\left.\begin{aligned}\dot{x} &= -\sin(2\pi x)\sin(2\pi y)\\ \dot{y} &= -\cos(2\pi x)\cos(2\pi y)\end{aligned}\right\}. \tag{22}$$

Separatrix map and Melnikov studies of this have appeared recently [57,9,46]; this also forms a simple kinematic model for steady Rayleigh–Bénard convection [58–60]. The flow (22) has a Hamiltonian

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

whose level curves define flow trajectories. The flow consists of periodically repeating square cells of alternating vorticity, separated by linear heteroclinics; see Fig. 6. Given the symmetry in the system, it suffices to address the inter-cellular chaotic transport generated across any one of these linear separatrices. Choose the central lower heteroclinic which goes from $(x, y) = (0, 1/4)$ to (0,3/4); the intention is to now characterise the chaotic transport between the two lower cells in Fig. 6. The heteroclinic trajectory's temporal evolution is given by $\bar{x}(t) = 0$ and

$$\bar{y}(t) = \begin{cases} \frac{1}{2\pi}\cos^{-1}[-\mathrm{sech}(2\pi t)], & \text{if } t \leq 0 \\ 1 - \frac{1}{2\pi}\cos^{-1}[-\mathrm{sech}(2\pi t)], & \text{if } t > 0 \end{cases}$$

for a symmetric choice of time zero. Thus, $|\nabla H(\bar{x}(t), \bar{y}(t))| = \mathrm{sech}(2\pi t)$, which upon insertion into (5) yields the optimum flux

$$\bar{s}(\omega) = \frac{G}{\pi} \sup_{\psi\in[0,\pi)} \int_{-\infty}^{\infty} \mathrm{sech}(2\pi t)|\cos(\omega t + \psi)|\,\mathrm{d}t.$$

Now, given the fact that $\mathrm{sech}(2\pi t)$ is even and unimodal, the choice of ψ to optimise the above must be the symmetric one, which corresponds to $\psi = 0$. Therefore,

$$\bar{s}(\omega) = \frac{2G}{\pi} \int_{0}^{\infty} \mathrm{sech}(2\pi t)|\cos(\omega t)|\,\mathrm{d}t. \tag{23}$$

The arguments of Section 2.4 indicate that (23) approaches its maximum $\bar{s}_m$ as $\omega \to 0^+$, is monotonically decreasing for positive ω, and levels out to a non-zero limit as $\omega \to \infty$. Indeed, from (8), $\bar{s}_m = G/(2\pi)$, and $\bar{s}(\infty) = (2/\pi)\bar{s}_m = G/\pi^2$ are the exact values expected. The results of numerically evaluating (23) with $G = 1$ are shown in Fig. 7, which exhibits all the properties claimed. In particular, the *quantitative* values are borne

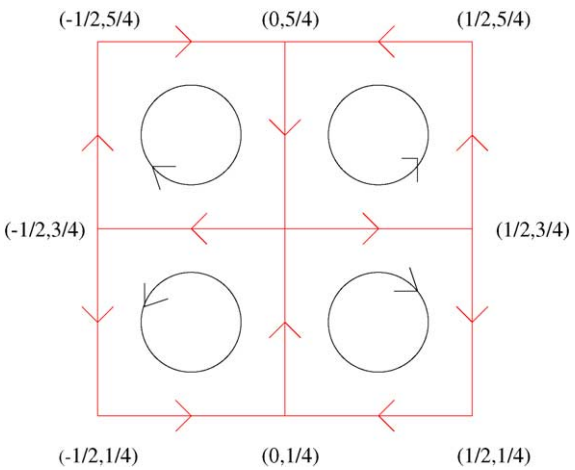

Fig. 6. Phase space for planar cellular flow.

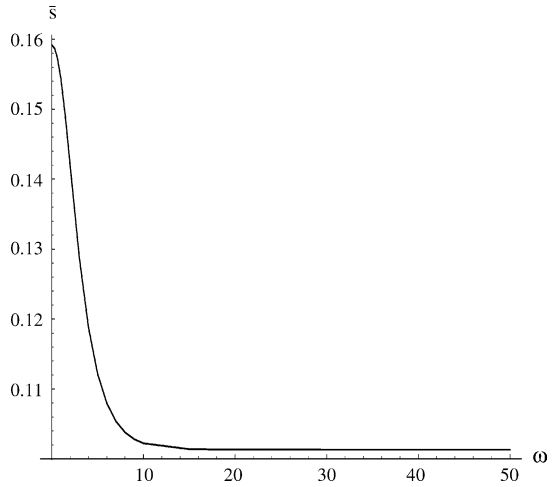

Fig. 7. Optimal flux $\overline{s}$ vs. frequency ω for planar cellular flow with $G = 1$.

out by the following data, correct to six decimals, which were obtained in generating Fig. 7: $\overline{s}(0^+) = 0.159155$, $\overline{s}(30) = 0.101321$, and $\overline{s}(50) = 0.101321$. Hence, the numerics tells us that $\overline{s}_m = 0.159155$, and $\overline{s}(\infty) = 0.101321$. These confirm the theoretical predictions of $1/(2\pi)$ and $1/\pi^2$, respectively.

Observe also how steeply the curve decreases at low frequencies before abruptly plateauing. The optimal flux is highly susceptible to variations in ω for small frequencies, but not at all for large ones. For highly wiggling perturbations g, the optimal chaotic flux is essentially G/π^2, a saturation flux. In any case, the main point of this analysis is that for a given frequency ω, the leading-order flux can never be any higher than the supremum norm of g times the corresponding value obtained from the graph of Fig. 7.

4.2. Hill's spherical vortex

A classical ideal fluid flow is that of Hill's spherical vortex [56], which is an axisymmetric entity whose structure is shown in Fig. 8. The point a is the north pole, and b the south, of a manifold Γ which is a spherical separating surface between flow on cylinders (outside), and flow on tori (inside). Since a and b are both fixed points, Γ forms a heteroclinic manifold, being simultaneously part of W^u_a and W^s_b. The heteroclinic trajectories connecting a to b traverse the longitudes of Γ. Perturbations would generically destroy Γ, enabling mixing between the vortex core and its exterior. For related analyses of this particular geometry, and its destruction through a Melnikov approach, see [52,33].

Here, the emphasis is on flux computations. The Stokes streamfunction here is

$$H(r,\theta) = \begin{cases} -\frac{U}{2}\left(r^2 - \frac{c^3}{r}\right)\sin^2\theta, & \text{if } r \geq c, \\ \frac{3U}{4}r^2\left(1 - \frac{r^2}{c^2}\right)\sin^2\theta, & \text{if } r < c, \end{cases}$$

where U and c are positive constants, representing the radius of the spherical surface Γ and a velocity scale, respectively. The corresponding velocity field f, based on the relationship given in Section 3.1, is

$$f(r,\theta,\phi) = \begin{cases} -U\left(1 - \frac{c^3}{r^3}\right)\cos\theta\,\hat{r} + U\left(1 + \frac{c^3}{2r^3}\right)\sin\theta\,\hat{\theta}, & \text{if } r \geq c, \\ \frac{3U}{2}\left(1 - \frac{r^2}{c^2}\right)\cos\theta\,\hat{r} - \frac{3U}{2}\left(1 - \frac{2r^2}{c^2}\right)\sin\theta\,\hat{\theta}, & \text{if } r < c. \end{cases}$$

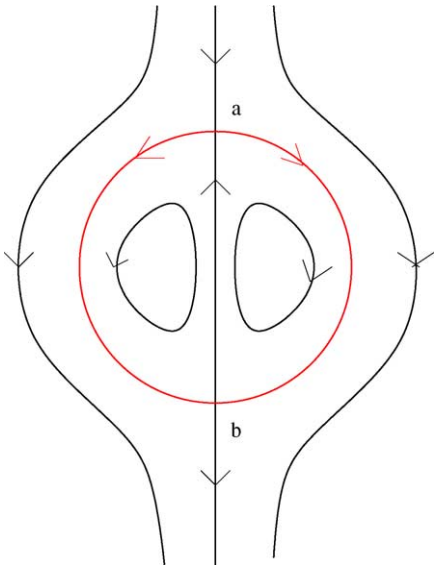

Fig. 8. Hill's spherical vortex in a constant ϕ cross-section.

In spherical polar coordinates, $\bar{x}_\phi(t) = (c, \bar{\theta}(t), \phi)$, and ϕ remains constant for a choice of heteroclinic. The only time-variation is that of the θ-component. Using the velocity field when $r = c$, and noting that the θ-component of velocity is $c\,\mathrm{d}\theta/\mathrm{d}t$,

$$c\frac{\mathrm{d}}{\mathrm{d}t}\bar{\theta} = \frac{3U}{2}\sin\bar{\theta}.$$

Solving this equation subject to the symmetric initial condition (i.e., $\bar{\theta}(0)$ lies on the equator) leads to

$$\sin\bar{\theta}(t) = \operatorname{sech}\left(\frac{3Ut}{2c}\right).$$

Now,

$$\nabla H(x_\phi(t)) = \nabla H(c, \bar{\theta}(t), \phi) = -\frac{3Uc}{2}\sin^2\bar{\theta}(t)\hat{r} = -\frac{3Uc}{2}\operatorname{sech}^2\left(\frac{3Ut}{2c}\right)\hat{r}.$$

See also [33] for the additional details of the above results. The most general separable case of Section 3.4 will be addressed directly; a velocity perturbation in the form $\varepsilon p(\phi)h(r,\theta)\cos[\omega(t-\beta)]$ is hypothesised. The expression for the unoptimised flux from (16) is

$$s(\omega) = \frac{3Uc}{2}\sqrt{\frac{2}{\pi}}\left(\int_0^{2\pi}|p(\phi)|\,\mathrm{d}\phi\right)\left|\mathcal{F}\left\{\operatorname{sech}^2\left(\frac{3Ut}{2c}\right)h_r(c,\bar{\theta}(t))\right\}(\omega)\right|,$$

where h_r is the $\hat{r}$ component of the perturbation h. This can be numerically evaluated for different choices of h_r if necessary.

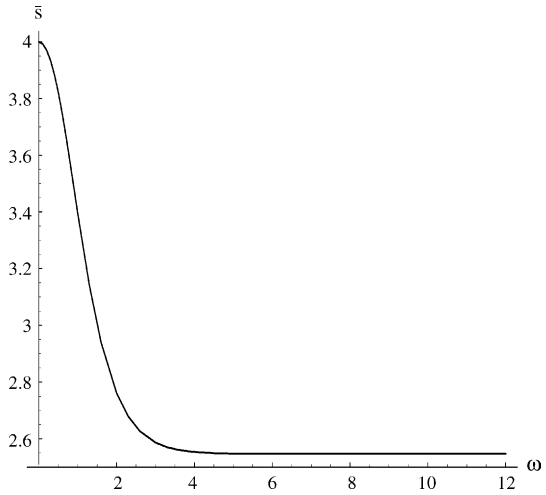

Fig. 9. Optimal flux $\overline{s}$ vs. frequency ω for Hill's spherical vortex with $U = 1$, $c = 1$, $G = 1$ and p constant.

Our interest is in the optimum flux, as given by (17):

$$\overline{s}(\omega) = \frac{3\,U\,c||h_r||}{2\pi}\left(\int_0^{2\pi} |p(\phi)|\,\mathrm{d}\phi\right) \sup_{\psi\in[0,\pi)} \int_{-\infty}^{\infty} \operatorname{sech}^2\left(\frac{3Ut}{2c}\right) |\cos(\omega t + \psi)|\,\mathrm{d}t,$$

for which, as in Section 4.1, the appropriate choice of ψ is 0 by symmetry. This leads to

$$\overline{s}(\omega) = \frac{3\,U\,c||h_r||}{2\pi}\left(\int_0^{2\pi} |p(\phi)|\,\mathrm{d}\phi\right) \int_{-\infty}^{\infty} \operatorname{sech}^2\left(\frac{3Ut}{2c}\right) |\cos(\omega t)|\,\mathrm{d}t, \tag{24}$$

where $||h_r||$ is the supremum norm of the function $h_r(c,\theta)$ over $\theta \in [0,\pi]$. The function p has no qualitative effect on the frequency-dependence of $\overline{s}(\omega)$, which is governed by the improper integral which cannot be evaluated explicitly except in the limits $\omega \to 0^+$ (in which case it is $4c/(3U)$) and $\omega \to \infty$ (in which case it is $(2/\pi)\overline{s}(0^+)$). The same qualitative behaviour as for planar cellular flow is to be expected from $\overline{s}(\omega)$, and this is confirmed by the numerically generated Fig. 9, in which $U = 1$, $c = 1$, p is constant, and $G = 1$ (for this supremum value of g, one can take any consistent values for p and $||h_r||$, for example $p = 1/2$ and $||h_r|| = 2$). A constant p is chosen since as described in Section 3.4, this is the optimal choice. For any perturbation which is azimuthally separable, Fig. 9 can be used to determine the maximum possible chaotic flux which can be approached for each given frequency.

5. Concluding remarks

This study has provided some insights into how best to perturb a system in order to achieve optimum chaotic mixing. Both two- and three-dimensional flows were addressed, and the results hold even within the constraint of incompressibility. Formulæ which bound the chaotic flux at each frequency of perturbation were presented in (6) and (18). The nature of the perturbation which approaches this upper bound was described in detail, and the optimal flux function $\overline{s}(\omega)$ expressed in different ways in (5), (7), (17) and (19).

The behaviour of the optimal flux as a function of frequency was also detailed. It monotonically decreases from a maximum at zero frequency to a non-zero limit at infinite frequency. While being consistent with the results of [25],

this highlights the difference with the *non-optimised* flux function which rises from zero, has at least one maximum, and then decays to zero in the high frequency limit [45,53,43].

The ideas used here are hoped to contribute to the design of devices which exploit chaotic motion to improve mixing. Micro-fluidic mixers (in which turbulence is negligible and diffusion may be too slow) may be particularly relevant applications, and there too the theory seems directly applicable to cross-channel micromixers [20,21,1,55].

References

[1] D. D'Alessandro, M. Dahleh, I. Mezić, Control of mixing in fluid flow: a maximum entropy approach, IEEE Trans. Auto. Control 44 (1999) 1852–1863.
[2] T. Shinbrot, L. Bresler, J.M. Ottino, Manipulation of isolated structures in experimental chaotic fluid flows, Exp. Thermal Fluid Sci. 16 (1998) 76–83.
[3] L. Bresler, T. Shinbrot, G. Metcalfe, J. Ottino, Isolated mixing regions: origin, robustness and control, Chem. Eng. Sci. 52 (1997) 1623–1636.
[4] C.M. Coates, A.P. Richardson, S. Wang, Nonpremixed cumbustion in turbulent mixing layers. II: Recirculation, mixing and flame stabilization, Combust. Flame 122 (2000) 271–290.
[5] G. Raman, Using controlled unsteady fluid mass addition to enhance jet mixing, Am. Inst. Aeronaut. Astronaut. J. 35 (1997) 647–656.
[6] G. Ciraolo, R. Lima, M. Vittot, M. Pettini, C. Figarella, P. Ghendrih, Controlling chaotic transport in a Hamiltonian model of interest to magnetized plasmas, J. Phys. A 37 (2004) 3589–3597.
[7] T.H. Solomon, S. Tomas, J.L. Warner, Role of lobes in chaotic mixing of miscible and immiscible impurities, Phys. Rev. Lett. 77 (1996) 2682–2685.
[8] G.O. Fountain, D.V. Khakhar, I. Mezić, J.M. Ottino, Chaotic mixing in bounded three-dimensional flow, J. Fluid Mech. 417 (2000) 265–301.
[9] S.S. Abdullaev, Structure of motion near saddle points and chaotic transport in Hamiltonian systems, Phys. Rev. E 62 (2000) 3508–3528.
[10] D.R. Meldrum, M.R. Holl, Microscale bioanalytic systems, Science 297 (2002) 1197–1198.
[11] H. Stone, S. Kim, Microfluidics: basic issues, applications and challenges, AIChE J. 47 (2001) 1250–1254.
[12] D.D. Cunningham, Fluidics and sample handling in clinical chemical analysis, Anal. Chim. Acta 429 (2001) 1–18.
[13] T.H. Schulte, R.L. Bardell, B.H. Weigl, Microfluidic technologies in clinical diagnostics, Clin. Chim. Acta 321 (2002) 1–10.
[14] B.H. Weigl, R.L. Bardell, C.R. Cabrera, Lab-on-a-chip for drug development, Adv. Drug Deliv. Rev. 55 (2003) 349–377.
[15] D.J. Beebe, G.A. Mensing, G.M. Walker, Physics and applications of microfluidics in biology, Annu. Rev. Biomed. Eng. 4 (2002) 261–286.
[16] M.A. Stremler, F. Haselton, H. Aref, Designing for chaos: applications of chaotic advection at the microscale, Phil. Trans. R. Soc. Lond. A 362 (2004) 1019–1036.
[17] A. Darhuber, J. Chen, J. Davis, S. Troian, A study of mixing in thermocapillary flows on micropatterned surfaces, Phil. Trans. R. Soc. Lond. A 362 (2004) 1037–1058.
[18] M.R. Bringer, C.J. Gerdts, H. Song, J.D. Tice, R.F. Ismagilov, Microfluidic systems for chemical kinetics that rely on chaotic mixing in droplets, Phil. Trans. R. Soc. Lond. A 362 (2004) 1087–1104.
[19] J. Ottino, S. Wiggins, Introduction: mixing in microfluidics, Phil. Trans. R. Soc. Lond. A 362 (1818) (2004) 923–935.
[20] F. Bottausci, I. Mezić, C.D. Meinhart, C. Cardonne, Mixing in the shear superposition micromixer: three dimensional analysis, Phil. Trans. R. Soc. Lond. A 362 (2004) 1001–1018.
[21] P. Taberling, M. Chabert, A. Dodge, C. Julien, F. Okkels, Chaotic mixing in cross-channel micromixers, Phil. Trans. R. Soc. Lond. A 362 (2004) 987–1000.
[22] A.D. Stroock, S.K.W. Dertinger, A. Ajdari, I. Mezić, H.A. Stone, G.M. Whitesides, Chaotic mixer for microchannels, Science 295 (2002) 647–651.
[23] H. Aref, Stirring by chaotic advection, J. Fluid Mech. 143 (1984) 1–21.
[24] I. Mezić, Chaotic advection in bounded Navier–Stokes flows, J. Fluid Mech. 431 (2001) 347–370.
[25] B.R. Noack, I. Mezić, G. Tadmor, A. Banaszuk, Optimal mixing in recirculation zones, Phys. Fluids 16 (2004) 867–888.
[26] A. Vikhansky, Control of stretching rate in time-periodic chaotic flows, Phys. Fluids 14 (2002) 2752–2756.
[27] A.T. Pérez, R. Chicón, A. Castellanos, Lobe dynamics and space charge distribution in nonsteady electroconvection, Phys. Rev. E 59 (1999) 135–142.
[28] P. Miller, L.J. Pratt, K.R. Helfrich, C.K.R.T. Jones, Chaotic transport of mass and potential vorticity for an island recirculation, J. Phys. Oceanogr. 32 (2002) 80–102.
[29] G. Haller, A.C. Poje, Finite-time transport in aperiodic flows, Physica D 119 (1998) 352–380.
[30] S. Balasuriya, C.K.R.T. Jones, B. Sandstede, Viscous perturbations of vorticity-conserving flows and separatrix splitting, Nonlinearity 11 (1) (1998) 47–77.
[31] S. Balasuriya, C.K.R.T. Jones, Diffusive draining and growth of eddies, Nonlin. Proc. Geophys. 8 (2001) 241–251.

[32] B. Sandstede, S. Balasuriya, C.K.R.T. Jones, P.D. Miller, Melnikov theory for finite-time vector fields, Nonlinearity 13 (4) (2000) 1357–1377.
[33] S. Balasuriya, I. Mezić, C.K.R.T. Jones, Weak finite-time Melnikov theory and 3D viscous perturbations of Euler flows, Physica D 176 (1–2) (2003) 82–106.
[34] M. Finn, S. Cox, H. Byrne, Mixing measures for a two-dimensional chaotic Stokes flow, J. Eng. Math. 48 (2004) 129–155.
[35] J.D. Meiss, Symplectic maps, variational principles, and transport, Rev. Modern Phys. 64 (1992) 795–847.
[36] R.S. MacKay, J.D. Meiss, I.C. Percival, Transport in Hamiltonian systems, Physica D 13 (1984) 55–81.
[37] A. Adrover, M. Giona, F.J. Muzzio, S. Cerbelli, M.M. Alvarez, Analytic expression for the short-time rate of growth of the intermaterial contact parameter in two-dimensional chaotic flows and Hamiltonian systems, Phys. Rev. E 58 (1998) 447–457.
[38] M. Giona, A. Adrover, F.J. Muzzio, S. Cerbelli, M.M. Alvarez, The geometry of mixing in time-periodic chaotic flows. i. Asymptotic directionality in physically realizable flows and global invariant properties, Physica D 132 (1999) 298–324.
[39] R.S. MacKay, Transport in 3-D volume-preserving flow, J. Nonlin. Sci. 4 (1993) 329–354.
[40] P. Ashwin, M. Nicol, N. Kirkby, Acceleration of one-dimensional mixing by discontinuous mappings, Physica A 310 (2002) 347–363.
[41] G.A. Voth, G. Haller, J.P. Gollub, Experimental measurements of stretching fields in fluid mixing, Phys. Rev. Lett. 88 (2002) Art. No. 254501.
[42] V. Rom-Kedar, S. Wiggins, Transport in two-dimensional maps, Arch. Rational Mech. Anal. 109 (3) (1990) 239–298.
[43] V. Rom-Kedar, A. Leonard, S. Wiggins, An analytical study of transport, mixing and chaos in an unsteady vortical flow, J. Fluid Mech. 214 (1990) 347–394.
[44] S. Wiggins, Chaotic Transport in Dynamical Systems, Springer-Verlag, New York, 1992.
[45] V. Rom-Kedar, A.C. Poje, Universal properties of chaotic transport in the presence of diffusion, Phys. Fluids 11 (8) (1999) 2044–2057.
[46] S. Balasuriya, Direct chaotic flux quantification in perturbed planar flows: general time-periodicity, SIAM J. Appl. Dyn. Syst., in press.
[47] V.K. Melnikov, On the stability of the centre for time-periodic perturbations, Trans. Mosc. Math. Soc. 12 (1963) 1–56.
[48] J. Guckenheimer, P. Holmes, Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields, Springer-Verlag, New York, 1983.
[49] D.K. Arrowsmith, C.M. Place, An Introduction to Dynamical Systems, Cambridge University Press, Cambridge, 1990.
[50] J. Gruendler, The existence of homoclinic orbits and the method of Melnikov for systems in R^n, SIAM J. Math. Anal. 16 (1985) 907–931.
[51] K.J. Palmer, Exponential dichotomies and transversal homoclinic points, J. Diff. Equations 55 (1984) 225–256.
[52] P. Holmes, Some remarks on chaotic particle paths in time-periodic, three-dimensional swirling flows, Contemporary Math. 28 (1984) 393–404.
[53] V. Rom-Kedar, Frequency spanning homoclinic families, Commun. Nonlin. Sci. Numer. Simul. 8 (3–4) (2003) 149–169 (chaotic transport and complexity in classical and quantum dynamics).
[54] S. Balasuriya, Separatrix splitting through high-frequency non-smooth perturbations, Dyn. Discr. Cont. Impuls. Syst., submitted for publication.
[55] A. Dodge, M. Jullien, Y.-K. Lee, X. Niu, F. Okkels, P. Tabeling, An example of a chaotic micromixer: the cross-channel micromixer, C. R. Physique 5 (2004) 557–563.
[56] D.J. Acheson, Elementary Fluid Dynamics, Oxford University Press, New York, 1990.
[57] T. Ahn, S. Kim, Separatrix-map analysis of chaotic transport in planar periodic vortical flows, Phys. Rev. E 49 (1994) 2900–2912.
[58] S. Chandrasekhar, Hydrodynamics and Hydrodynamic Stability, Dover, New York, 1961.
[59] B.I. Shraiman, Diffusive transport in a Rayleigh–Bénard convection cell, Phys. Rev. A 36 (1987) 261–267.
[60] T.H. Solomon, J.P. Gollub, Chaotic particle transport in time-dependent Rayleigh–Bénard convection, Phys. Rev. A 38 (1988) 6280–6286.