ELSEVIER

Physica D 106 (1997) 389–431

PHYSICA D

Chaotic Hamiltonian dynamics of particle's horizontal motion in the atmosphere

Vered Rom-Kedar [a,*] Yona Dvorkin [b], Nathan Paldor [b]

[a] *The Department of Applied Mathematics and Computer Science, The Weizmann Institute of Science, PO Box 26, Rehovot 76100, Israel*
[b] *The Department of Atmospheric Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel*

Received 24 November 1996; revised 20 December 1996; accepted 20 December 1996
Communicated by C.K.R.T. Jones

Abstract

A non-separable, near-integrable, two degrees-of-freedom (d.o.f.) Hamiltonian system arising in the context of particle's dynamics on a geopotential of the atmosphere is studied. In the unperturbed system, homoclinic orbits to periodic motion in the longitude angle exist, giving rise to several types of homoclinic chaos in the perturbed system; a new type of homoclinic chaos is found, exhibiting chaos which is non-uniformly distributed in the longitude angle along the homoclinic loop while it is topologically uniformly distributed in this angle in a neighborhood of the hyperbolic periodic orbit, i.e. the return map to different values of the longitude angle are topologically conjugate only near the hyperbolic periodic orbit. The concept of *colored* energy surfaces and its corresponding energy–momentum map are developed as analytical tools for delineating the regions in the four-dimensional phase space where the various types of homoclinic chaos prevail. Geometrical interpretation of the Melnikov analysis is offered for detecting this non-uniformity in angle. Applying the results of the analysis to the planetary atmosphere with an infinite wavelength perturbation of amplitude ϵ, we find that for eastward going particles with initial velocities $(u_0, v_0) = (\overline{u}, 0)$ the chaotic band occupies two narrow $(O(\epsilon))$ bands on both sides of the equator near latitudes $\phi = \pm \arccos(1 - 2\overline{u})$. By contrast, for westward going particles the chaotic zone is thicker $(O(\sqrt{\epsilon}))$ and is centered on the equator. Moreover, the dependence of the chaotic zone's thickness on the perturbation frequency is much more sensitive to the value of the initial speed for eastward going particles than for those going westward. Qualitative and quantitative differences between the distribution of the homoclinic chaotic motion in the longitude angle for westward and eastward travelling wave perturbations are predicted and numerically confirmed.

Keywords: Non-separable Hamiltonian systems; Homoclinic chaos; Tracer dispersal

1. Introduction

The simplest possible motion on the surface of a rotating sphere, due only to Coriolis force (i.e. in the absence of any body force), is called the inertial motion. This motion corresponds to a fundamental physically realizable system, and its investigation leads to the development of new tools for analyzing two degrees-of-freedom (d.o.f.)

* Corresponding author. Tel.: 972-8-9343170; fax: 972-8-9344122; e-mail: vered@wisdom.weizmann.ac.il.

PII S0167-2789(97)00015-8

systems [1,2]. It may be relevant for describing the motion of satellites, of particles in the atmosphere (such as weather balloons and pollutant clouds) or floats in the deep ocean along a geopotential surface. Indeed, in the context of atmospheric and oceanographic flows, possible solutions to this system have been heuristically described in various textbooks [3]. The introduction of the general equations of motion, their integrals, and the full analysis of their possible solutions has been introduced by Paldor and Killworth [4]. Paldor and Boss [5] showed that this inertial motion may be formulated as an integrable two d.o.f. Hamiltonian flow. Adding a weak travelling wave forcing to the equations of motion they demonstrated numerically that in some cases chaotic trajectories exist and cause exponential dispersion of lines of particles. This addition of forcing to the inertial motion can be viewed as a simple (ad hoc) inclusion of a zonally travelling pressure wave which perturbs the geopotential surface from its mean spherical shape.

Field experiments on the dispersal of clusters of weather balloons flying on a given geopotential at high altitudes [6,7] demonstrate that the rate of particle dispersal is not uniform in time and that the rate of increase of pair separation is highly variable among the various participating pairs. These observations lend credence to the hypothesis that the dispersal of passive particles in the atmosphere is dominated by a process far from simple isotropic turbulence. A chaotic paradigm can provide an alternate explanation for the observed dispersal of such particles. One such paradigm is the Lagrangian advection approach, which has been extensively studied in the last decade, see [8–11] and references therein. The present study deals with a different idealized system of atmospheric dynamics for which the chaotic paradigm for describing pair separation is natural [5].

Mathematically, the equations considered are brought to the form of near integrable non-separable two d.o.f. Hamiltonian system. These are systems with a Hamiltonian of the form:

$$H(x_1, p_1, D, \Lambda) = H_0(x_1, p_1, D) + \epsilon H_1(x_1, p_1, D, \Lambda)$$

with $\partial^2 H_0/\partial D \partial x_1 \not\equiv 0$. D represents the second constant of motion (in the unperturbed system), and Λ is its conjugate variable. Furthermore, we assume that for $\epsilon = 0$, for some open interval U of D, there exists a hyperbolic fixed point in the (x_1, p_1) plane and its stable and unstable manifolds are part of a homoclinic loop. Systems of this form arise, for example, when considering travelling wave perturbations to the motion of a particle in a central field [1], where D represents the angular momentum. These systems exhibit a rich behavior: for some regions in phase space their behavior is similar to one-and-a-half d.o.f. Hamiltonian systems; namely, even though the perturbation depends on Λ, in such regions the perturbed motion is topologically uniform in Λ – return maps to a given Λ_0 value are essentially independent of Λ_0 (more precisely return map to different Λ's are topologically conjugate). For other regions, a non-uniform distribution of the homoclinic chaos in the angle Λ is created, a phenomenon which has not been analyzed previously. We present new tools for delineating the phase space and parameter space according to these different behaviors and describe some of the new dynamics appearing in the non-separable systems. Colored energy surfaces are constructed, from which the dynamics of the perturbed system may be directly understood, and then related to the bifurcation diagram of the energy–momentum map [12,13]. Introducing the information regarding the regions of non-monotonicity into this bifurcation diagram, it supplies complete information of the unperturbed structure from which qualitative understanding of the underlying dynamics for small (yet finite) perturbations may be extracted. This is the main mathematical result of our paper.

Another set of results is obtained from the interpretation of the standard analysis of some reduced systems in terms of the primitive variables (the longitude, latitude and eastward, westward velocities). Since the constant of motion depends non-trivially on the primitive variables, any conclusion regarding the location and extent of the chaotic regions in D must include the correct rescaling in terms of the primitive variables. These simple considerations lead, for example, to non-intuitive results regarding the differences between the dynamics of westward and eastward going particles. The approach and the basic tools we develop for the analysis are general, and may appear in analyzing near-integrable systems of higher dimensions as well.

Homoclinic structures of two d.o.f. Hamiltonian systems have been studied analytically in several contexts. Lerman and Umanskii [13] have classified integrable four-dimensional Hamiltonian systems near singular points by studying the topological structure of the neighborhoods of the singular sets in the momentum map. Lerman and Umanskii [14] and Grotta-Ragazzo [15] have studied the structure of homoclinic orbits to isolated fixed points of the four-dimensional system. Holmes and Marsden [16] have studied first the homoclinic structure to periodic orbits in separable two d.o.f. Hamiltonian system. Then, studying the appearance of horseshoes near homoclinic tangles in non-canonical Hamiltonian systems with symmetries, they presented as one of their simplest examples non-separable two d.o.f. Hamiltonian system with *monotonic* dependence on the angle variable [17]. Their analysis applies here to the monotonic regions in phase space for which the system is reducible to one-and-a-half d.o.f. Hamiltonian system, leading to homoclinic chaos which is uniform in the longitude angle. Kovacic and Wiggins [18], Haller and Wiggins [8], Haller [36] and Kaper and Kovacic [20] have considered perturbations of an integrable two d.o.f. Hamiltonian system having a two-dimensional normally hyperbolic manifold containing a circle of fixed points and having a family of heteroclinic orbits connecting the fixed points on the circle. Their analysis is used here to describe one particular case of homoclinic chaos arising when the wave perturbation travels westward. In this case the homoclinic chaos found is nowhere uniform in the longitude angle.

This paper is organized as follows: in Section 2 we describe the structure of the integrable system and its symmetries, and distinguish between the equations describing infinite and finite wavelength perturbations. In Sections 3.1 and 3.2 we describe the simple (standard) analysis of the infinite wavelength perturbation case. Then in Section 3.3 we present the interpretation and the application of the analysis in terms of dispersion of particles, location of the zones of chaotic motion, and the stretching and transport of particles. The importance of the two d.o.f. origin of the problem to the application of the results is stressed. In Section 4 the finite wavelength perturbation case is analyzed. After introducing the Hamiltonian formulation in Section 4.1, we construct the *colored energy surfaces* and compactify this presentation by plotting the bifurcation diagram of the energy–momentum map in Section 4.2. From these constructions we delineate the phase space according to the expected different behaviors under perturbation as explained in Section 4.3. In Section 4.4 we briefly discuss the structure of the resonances and their implications. In Section 4.5 we study the homoclinic chaos associated with the finite wavelength perturbation. The geometrical interpretation of the Melnikov vector in this fully three-dimensional geometry is discussed. In Section 4.6 we analyze the more exotic resonance, which occurs only for westward propagating pressure waves, between the unstable free motion near the equator and the pressure wave. Conclusions and some preliminary geophysical discussion of the results are presented in Section 5. Appendix A contains the analytical solutions to the unperturbed problem and the reduced equations of motion.

2. The phase space geometry

The non-dimensional Lagrangian momentum equations for the eastward and northward velocity components (u, v) and the rate of change of the longitude (λ – the east/west coordinate – the zonal coordinate) and latitude (ϕ – the north/south coordinate – the meridional coordinate) given a zonally travelling pressure wave (see Fig. 1) are given by:

$$\frac{\mathrm{d}\lambda}{\mathrm{d}t} = \frac{u}{\cos\phi}, \tag{2.1a}$$

$$\frac{\mathrm{d}u}{\mathrm{d}t} = v\sin\phi\left(1+\frac{u}{\cos\phi}\right) - k\epsilon\frac{A(\phi)}{\cos\phi}\cos(k\lambda - \sigma t + \theta). \tag{2.1b}$$

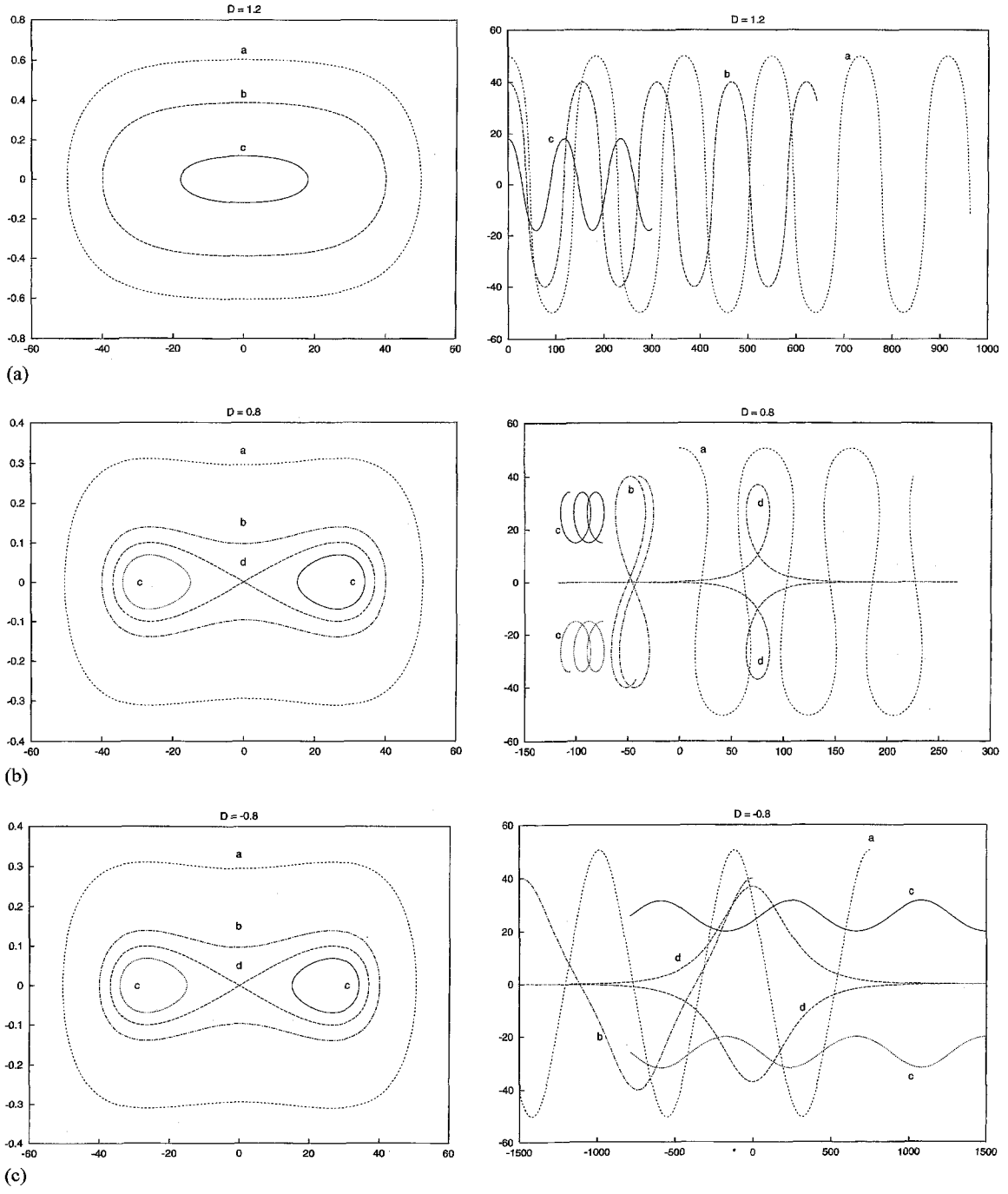

Fig. 1. The inertial motion (left figures correspond to (ϕ, v) plane, right figures to (λ, ϕ) plane, extended to the real line in λ): (a) phase flow $|D| > 1$; (b) phase flow $0 < D < 1$; (c) phase flow $-1 < D < 0$.

$$\frac{d\phi}{dt} = v, \tag{2.1c}$$

$$\frac{dv}{dt} = -u \sin\phi \left(1 + \frac{u}{\cos\phi}\right) - \epsilon A'(\phi) \sin(k\lambda - \sigma t + \theta). \tag{2.1d}$$

The equations for the change in (λ, ϕ) reflect their Lagrangian character (these advect with their velocity). The term $1/\cos\phi$ in the equation for $\dot{\lambda}$ reflects the convergence of the longitude coordinates at the poles. The first term in the r.h.s. of the $\dot{u}$, $\dot{v}$ equations is the Coriolis force, with the geometric correction (the change in local north as one moves eastward) which applies for the global scale on which these equations are considered (the standard "beta-plane-approximation" used for mid-latitude studies corresponds to replacing $\sin\phi$ by its linear expansion near a given latitude ϕ_0, while neglecting the geometrical correction term). The second term corresponds to the perturbing pressure wave (which in the fluid equations of motion is coupled back to the velocities via incompressibility). $A(\phi)$ represents the latitude dependent amplitude of the pressure wave, assumed to be even in ϕ:

$$A(\phi) = A(-\phi), \qquad A(0) = 1, \qquad A'(0) = 0. \tag{2.2}$$

When no pressure waves are present the motion is integrable; two integrals of motion were derived in [4] for the inertial (i.e. the case $\epsilon = 0$) trajectories, corresponding to (twice) the angular momentum D and the kinetic energy E:

$$D = \cos\phi(\cos\phi + 2u), \tag{2.3a}$$

$$E = \tfrac{1}{2}(u^2 + v^2). \tag{2.3b}$$

Using (2.3a) u may be expressed in terms of D and ϕ:

$$u(\phi, D) = \frac{1}{2}\left(\frac{D}{\cos\phi} - \cos\phi\right). \tag{2.4}$$

In subsequent sections it will be useful to reformulate the equations of motion using the angular momentum D as the new variable instead of u in the system (2.1). Notice that small velocities near the equator correspond to $D \approx 1$ whereas near the poles $D \approx 0$. Negative D is realized typically only for large westward initial velocity $(u < -0.5\cos\phi)$.

A few notes regarding (2.1) are now in order:

(1) $A(\phi)$ may represent, for example, the amplitude of a daily tidal ($k = 1, \sigma = \frac{1}{2}$) pressure wave where A is the pressure (divided by the density); this amplitude, $A(\phi)$, may be either taken as a constant ($A(\phi) = 1$) or to vary with latitude (e.g. Paldor and Boss [5] took $A(\phi) = \cos^3(\phi)$). Higher or negative wave numbers may be associated with other waves present in the atmosphere (e.g. Rossby waves).

(2) It follows from (2.3a) that the poles may be reached iff $D = 0$, hence only near the surface $D = 0$ singularities associated with the use of the polar co-ordinates may be encountered. The analysis near $D = 0$ is left for future studies.

(3) Eqs. (2.1) are non-dimensionalized using the radius of the earth for length scale and twice the frequency of the earth's rotation for the frequency scale so that a non-dimensional time unit corresponds to $24h/4\pi \approx 1.901$ h. Order 1 velocities in the non-dimensional variables turn out to be nearly 1 km/s dimensional velocities, so the relevant non-dimensional velocities for atmospheric flows are of order 0.01. For satellites, the relevant scales are of order 1 and for oceanographic flows of order 0.001. We assume (see Section 1) that for the atmospheric flows the relevant scale for the pressure wave is that of the order of the kinetic energy, namely ϵ is $O(10^{-4})$.

When $\epsilon = 0$ the equation for λ decouples, D is conserved, and the phase flow of the two-dimensional system obtained for (ϕ, v) for various values of D is as depicted in Fig. 1. Plotting the same orbits in the (λ, ϕ) plane, we

demonstrate that while the (ϕ, v) phase flow is invariant under reflection of the angular momentum ($D \to -D$), the flow in λ is not. The qualitative description and some notation for this flow is listed below. The more technical formulae are deferred to Appendix A.

Qualitatively, there are two regions of D values (see Fig. 1):

- When $|D| > 1$ the origin ($v = 0 = \phi$) is stable (elliptic) and is surrounded by periodic orbits which visit both hemispheres. The maximal latitude reached by a periodic orbit is denoted by $\phi_{\max} = \phi_{\max}(E, D)$ (it follows from (2.3) that for a given D one can parametrize orbits by $\phi_{\max}$ instead of E).
- When $|D| < 1$ the origin is unstable (hyperbolic) and two elliptic fixed points are created at

$$(\phi_{\text{ell}}, v_{\text{ell}}) = \pm(\arccos\sqrt{|D|}, 0). \tag{2.5}$$

The origin has two homoclinic orbits, [1] one in each hemisphere:

$$q_{\text{h}}^{\pm}(t; \phi_{\text{hmax}}) = \pm(\phi_{\text{h}}(t; \phi_{\text{hmax}}), v_{\text{h}}(t; \phi_{\text{hmax}})), \tag{2.6}$$

where ϕ_{hmax} is the maximal latitude reached by the homoclinic orbit and is given by

$$\phi_{\text{hmax}}(D) = \arccos(|D|), \qquad \phi_{\text{hmax}} \in [0, \tfrac{1}{2}\pi]. \tag{2.7}$$

The homoclinic orbits separate the phase space to three distinct regions – the two identical regions (R_1 and R_2) containing the periodic orbits, denoted by c in Figs. 1(b) and (c), which are restricted to one hemisphere (north and south, respectively). The region R_3 containing orbits a, b in Figs. 1(b) and (c) which are unrestricted periodic orbits with $\phi_{\max} > \phi_{\text{hmax}}$. The area of the restricted regions is:

$$\mu(R_1(\phi_{\text{hmax}})) = \mu(R_2(\phi_{\text{hmax}})) = \sin\phi_{\text{hmax}} - \phi_{\text{hmax}}\cos\phi_{\text{hmax}}. \tag{2.8}$$

The unperturbed system belongs to the exclusive family of dynamical systems for which all solutions are found analytically. In Appendix A we present the analytic formulae for the fixed points and their eigenvalues, the homoclinic orbits to the origin $q_{\text{h}}(t; \phi_{\text{hmax}})$ for $|D| < 1$, the associated homoclinic solution for the longitude position $\lambda_{\text{h}}^{\pm}(t; \phi_{\text{hmax}})$, the period ($P_\phi(\phi_{\max}; \phi_{\text{hmax}}) = P_\phi(E, D)$) and action ($I(\phi_{\max}, D)$) of the periodic orbits in the (ϕ, v) phase space, and the longitude position after completion of one period in ϕ:

$$\lambda_P(\phi_{\max}, \phi_{\text{hmax}}) = \lambda_P(E, D) = \lambda(t; \phi_{\max}, \phi_{\text{hmax}})|_{t=P_\phi(\phi_{\max};\phi_{\text{hmax}})} - \lambda(0; \phi_{\max}, \phi_{\text{hmax}}).$$

Consider the structure of the perturbed problem: formally, (2.1) is a non-autonomous (time-dependent) four-dimensional system. However, it follows from (2.1) that

$$\frac{\mathrm{d}D}{\mathrm{d}t} = -2\epsilon A(\phi)k\cos(k\lambda - \sigma t + \theta), \tag{2.9}$$

therefore, we divide our analysis to two cases:

Case 1: $k = 0$. The angular momentum is still conserved and the (ϕ, v) equations of (2.1) may be written as a time-periodically perturbed one d.o.f. freedom system which depends on the constant of motion D.

Case 2: $k \neq 0$. The phase $\lambda - (\sigma/k)t$ serves as a new dependent variable and (2.1) may be transformed to an autonomous two d.o.f. system.

[1] These are solutions which are asymptotic to the origin as $t \to \pm\infty$.

3. Forcing with infinite wavelength ($k = 0$)

3.1. Equations of motion

This forcing corresponds to a perturbation which does not vary with the longitude. If the amplitude of the pressure wave is independent of latitude ($A(\phi) = \text{const.}$) then the perturbation merely changes periodically the radius of the spherical geopotential and hence has no effect on the motion along it. Therefore, only latitude dependent amplitude ($A'(\phi) \neq 0$) is of interest. Since the angular momentum, D, is conserved in this case, using (2.4) it follows that Eqs. (2.1c) and (2.1d) are independent of λ and may be analyzed using the standard tools of a one-and-a-half d.o.f. Hamiltonian system, depending on the parameters D, σ and on $A'(\phi)$:

$$\frac{d\lambda}{dt} = \frac{1}{2}\left(\frac{D}{\cos^2\phi} - 1\right), \tag{3.1a}$$

$$\frac{dD}{dt} = 0, \tag{3.1b}$$

$$\frac{d\phi}{dt} = v, \tag{3.1c}$$

$$\frac{dv}{dt} = \frac{1}{8}\sin(2\phi)\left(1 - \frac{D^2}{\cos^4\phi}\right) + \epsilon A'(\phi)\sin(\sigma t - \theta), \tag{3.1d}$$

and the time-dependent Hamiltonian (for the (ϕ, v) Hamiltonian system):

$$H(\phi, v, t; D) = \frac{v^2}{2} + \frac{1}{8}\left(\frac{D}{\cos\phi} - \cos\phi\right)^2 - \epsilon A(\phi)\sin(\sigma t - \theta). \tag{3.2}$$

The equations and Hamiltonian written for $k = 0$ clearly demonstrate the relevance of the restriction to latitude dependent amplitude of the perturbation. Next we present the analysis of this case, consisting of standard resonance and Melnikov function calculation. In Section 3.3 we discover new geometrical and physical implications of the Melnikov function calculations which are the consequence of the non-trivial D-dependence of the unperturbed problem.

3.2. Analysis of the chaotic motion

It follows from (3.1) that the (ϕ, v) system may be treated as a one-and-a-half d.o.f. Hamiltonian system depending on the parameters σ and ϵ and the constant of motion D (D may be taken to be positive for the analysis). Mathematically, one views the perturbed system as three dimensional in (ϕ, v, t) and considers the dynamics of the Poincaré map, a map found by sampling the solutions at $t = 2j\pi/\sigma + \tau_0$ for fixed τ_0. Using the phase θ to vary the Poincaré map, we set $\tau_0 = 0$. By the evenness of $A(\phi)$, the Poincaré map is symmetric with respect to:

(i) $\phi \to -\phi, v \to -v$ for all θ,

(ii) $\phi \to -\phi, t \to -t$ when $\theta = \frac{1}{2}\pi, \frac{3}{2}\pi$.

Hence, it is beneficial to use $\theta = \frac{1}{2}\pi$ for plotting Poincaré maps.

Resonances: For $D > 1$ the solutions to the unperturbed problem are periodic in t. In the Poincaré map, the orbits with periods rationally related to $2\pi/\sigma$ (these are called resonant orbits) appear as periodic orbits of the mapping (each orbit consists of a finite set of distinct points), and those which are irrationally related trace a continuous closed curve. For small enough perturbation "most" of these curves survive (the ones which have periods sufficiently far

from being resonant). When resonance of order n/m occurs, i.e., when the period of the unperturbed periodic motion in the (ϕ, v) plane, $P_\phi(\phi_{\max}, D)$, satisfies

$$P_\phi = \frac{n}{m}\frac{2\pi}{\sigma}, \tag{3.3}$$

a chain of islands in the (ϕ, v) plane appears. Namely, a stable periodic orbit emerges, crossing the equator $2m$ times as t covers the interval $[0, 2n\pi/\sigma]$ and orbits which are sufficiently close to this solution – namely belong to the resonance band – oscillate about it (the resonance band width is proportional to $\sqrt{\epsilon}$ and is exponentially decreasing with n/m). For example, using (A.6), it follows that $P_\phi(\phi_{\max}, D)|_{(0.982,1.2)} = 2\pi$ and $P_\phi(\phi_{\max}, D)|_{(0.508,1.2)} = 4\pi$ for $\sigma = 1$, $D = 1.2$, hence the 1:1 and 1:2 resonance occur near the latitudes $\phi = 0.982 = 56.25°$ and $\phi = 0.508 = 29.10°$, respectively. In Section 3.3 we will confirm these analytic findings by numerically calculating the Poincaré map.

The width of the resonance band is given by

$$\Delta I = 4\sqrt{\frac{\epsilon M^{m/n}(\sigma, D)}{\Omega'(I^{m/n}, D)}},$$

where $M^{m/n}(\sigma, D)$ denotes the magnitude of the sub-harmonic Melnikov function, $I^{m/n}$ the action and Ω' denotes the periodic orbit frequency's derivative, all evaluated on a resonance orbit satisfying (3.3). The action $I = I(\phi_{\max}, D)$ is given in Appendix A (and the resonance condition (3.3) determines $\phi_{\max}$ as a function of $(D, m/n, \sigma)$). The existence of large resonance zones implies the appearance of new, localized structures in the perturbed flow. The localization found here is in the (ϕ, v) plane – namely localization in latitude. Using Eq. (A.7) we may find the associated zonal behavior. For $0 < D < 1$ a similar scenario happens for periodic orbits in each of the regions R_1, R_2, R_3 which are bounded away from the separatrix. Moreover, in this case we expect to find structures which are localized both zonally and meridionally. A particularly strong resonance occurs when $2\pi/\sigma$ is rationally related to the natural frequency of the elliptic points, the inertial frequency, $2\pi/\sqrt{1-|D|}$ (see Appendix A).

The implication of the resonance on dispersion is that in the resonance region there is a strong non-uniformity in ϕ: starting with a line of weather balloons inside the resonance leads to *less dispersion than* the linear one observed in the *unperturbed case* whereas putting the line of weather balloons near the hyperbolic periodic point leads to exponential dispersion, faster than in the unperturbed case. Examples of these effects are presented in Section 3.3.

Homoclinic chaos: For $|D| < 1$ the separatrix breaks down enabling orbits to spend several rounds in the north hemisphere and then hop to the south hemisphere and vice versa. The change in the Hamiltonian along the separatrix is expressed in terms of the Melnikov function $M(t_0, \theta, \sigma, |D|)$ [21–24], a periodic function in t_0 of period $T = 2\pi/\sigma$:

$$M(t_0, \theta, \sigma, |D|) = 2\cos(\sigma t_0 + \theta) \int_0^\infty vA'(\phi)\sin(\sigma t)\big|_{q_h^\pm(t;\arccos(|D|))}\,\mathrm{d}t = M_0(\sigma, |D|)\cos(\sigma t_0 + \theta). \tag{3.4}$$

where t_0 serves as a parametrization along the homoclinic orbit. For fixed D, the distance between the stable and unstable manifolds along the ϕ-axis is given by

$$\mathrm{d}(\theta; \sigma, |D|) = \epsilon\frac{4|D|M(0, \theta, \sigma, |D|)}{|1-D^2|^{3/2}} + \mathrm{O}(\epsilon^2) = d(\sigma, |D|)\cos(\theta) + \mathrm{O}(\epsilon^2), \tag{3.5}$$

where t_0 of (3.4) is set to 0 in (3.5) since here we define the distance to be measured along the ϕ-axis and

$$d(\sigma, |D|) = \epsilon\frac{4|D|M_0(\sigma, |D|)}{|1-D^2|^{3/2}} + \mathrm{O}(\epsilon^2). \tag{3.6}$$

As long as $M_0(\sigma, |D|)$ does not vanish the Melnikov function has simple zeros, implying that the stable and unstable manifolds intersect transversely.

Three remarks are now in order:

(1) Notice that while both M_0 and the normalization factor in (3.6) do not depend on sign (D), the initial westward velocity at the equator does (see (2.4)):

$$u_h(D) = \lim_{t\to\pm\infty} u_h(t) = \tfrac{1}{2}(D-1) = \tfrac{1}{2}(\text{sign}\,(D)\cos(\phi_{hmax}) - 1),$$

and so does the kinetic energy on the unperturbed separatrix:

$$E_h(\phi_{hmax}, \text{sign}\,(D)) = E_h(D) = \tfrac{1}{8}(\text{sign}\,(D)\cos(\phi_{hmax}) - 1)^2 = \tfrac{1}{8}(D-1)^2. \tag{3.7}$$

It follows that a decrease in u_h beyond $-\frac{1}{2}$ (equivalently an increase in E_h beyond 0.125) results in the same chaotic dynamics observed for larger u_h values (smaller energies), or in mathematical terms, M_0 is at least a two-to-one function of u_h and E_h. It follows that the chaotic behavior is exactly the same for initial conditions near the equator with the different westward velocities $u_h(D)$ and $u_h(-D)$. The geophysical interpretation is that these westward velocities lead to the same homoclinic orbits in (ϕ, v) space – one passing east of the north pole and the other west of it (similarly for the south pole). The extent of the chaotic region depends on the maximal latitude reached by the homoclinic orbit and not on its longitude position.

(2) It is customary to interpret the amplitude of the Melnikov function $(M_0(\sigma, |D|))$ shown in Fig. 2 as a measure of the degree of chaos in the system. However, there are several different quantifiers of the chaos in the system (e.g. stretching, maximal mixing, maximal chaotic zone, etc.), which depend on the Melnikov function magnitude *and* on the properties of the unperturbed flow. Since the unperturbed structure depends on $|D|$, larger $M_0(\sigma, |D|)$ does not necessarily imply "more chaos", and one has to be more precise about the interpretation

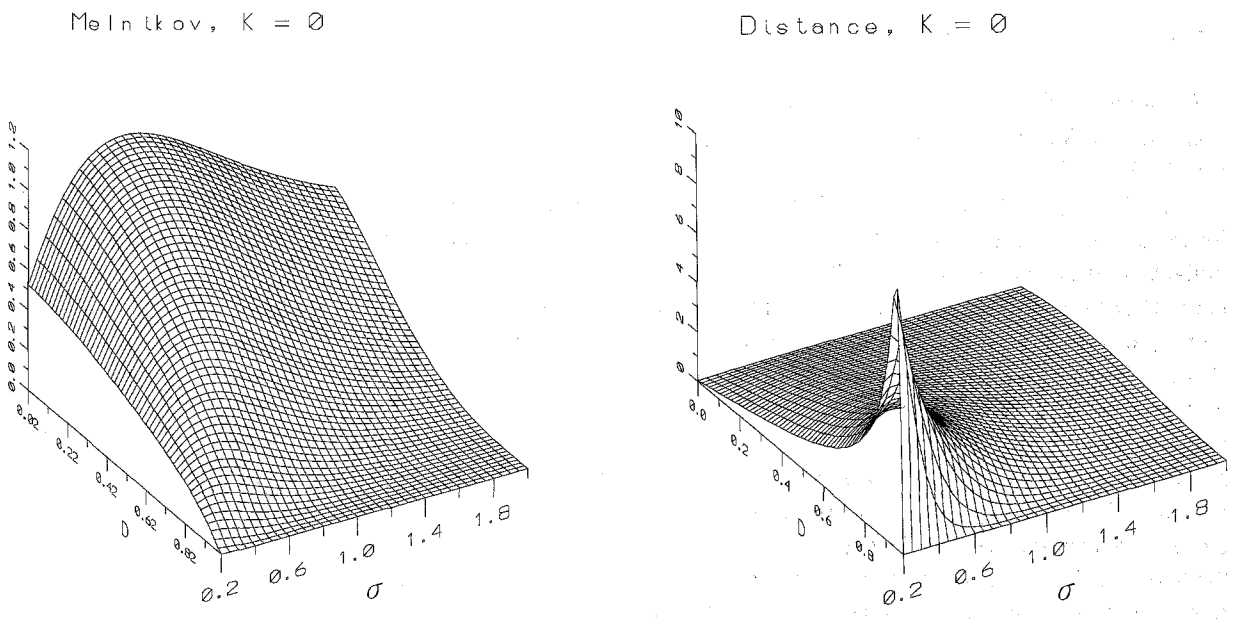

Fig. 2. Separatrix splitting for the infinite wavelength perturbation ($k = 0$): (a) Melnikov integral amplitude; (b) distance function amplitude.

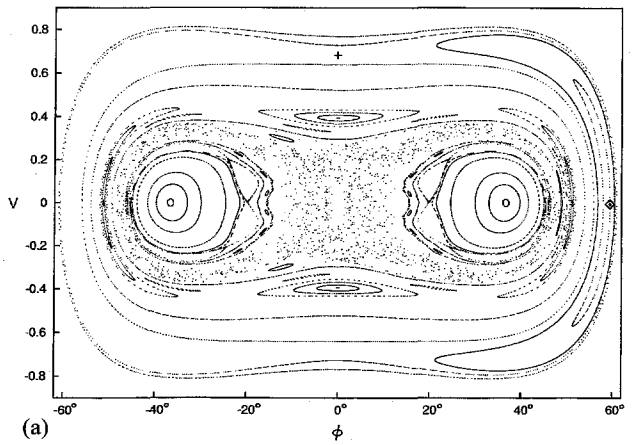

Fig. 3. Resonance influence on stretching: (a) Poincaré map with $k = 0$, $\sigma = 1$, $D = 1.2$, $\epsilon = 0.25$, $A(\phi) = \cos(\phi)$, $+$, $\diamond$ – location of initial conditions for stretching study; (b) location of particles at $t = 40\pi = 125.664$; (c) length of a line connecting the particles; (d) pairwise dispersion.

of the Melnikov function calculation. In Section 3.3 we use the Melnikov function calculation to extract several physical properties of the chaotic flow, including estimates of the extent of the chaotic zone in (ϕ, λ) for a given westward velocity.

(3) Notice that from (3.6), $d(\sigma, 0) = 0$ regardless of the value of $M_0(\sigma, 0)$ and this is apparent in Fig. 2 where $M_0(\sigma, 0) > 0$ while $d(\sigma, 0) = 0$. Indeed, the chaotic zone becomes thinner near ϕ_{hmax} as $|D| \to 0$ while its width near the equator is proportional to $\sqrt{M_0(\sigma, 0)}$ in this limit (see Section 3.3.1 for details).

3.3. Applications of the analysis

Resonances: The existence of resonances established in the preceding section has important consequences. In Fig. 3(a) the Poincaré map at $\theta = \frac{1}{2}\pi$ is shown, where the $1:1$ and $1:2$ resonance bands are clearly seen near $\phi = 0.98 \approx 56°$ and $\phi = 0.5 \approx 28°$, respectively (the symmetry of the perturbed system implies the appearance of two islands of stability for the $1:1$ resonance band). Here the implication of these resonances are presented: for the example of Figs. 3(a) and (b) we demonstrate that a group of initial conditions placed on the equator with initial velocities $(u_0, v_0) \approx (0.1, 0.7)$ disperses approximately the same way when $\epsilon \neq 0$ and $\epsilon = 0$. In fact there is an $O(\sqrt{\epsilon})$ dispersion in the maximal latitude reached by the initially close initial conditions. On the other hand, a group of initial conditions placed near $\phi = 56.25°$ with initial velocities $(u_0, v_0) \approx (0.8, 0.0)$ disperses much slower when $\epsilon \neq 0$ than in the unperturbed case. Moreover, this group of initial condition stays in the northern hemisphere for all times when $\epsilon \neq 0$ whereas for $\epsilon = 0$ the same initial conditions lead to periodic unrestricted motion visiting both hemispheres. These examples demonstrate the two general features of fast separation near the hyperbolic periodic orbits which are created by the resonances and the "over" stability which is created inside the resonance island. In Figs. 3(c) and (d) we present two crude measures of these effects: we pick a set of $N = 100$ close initial conditions ordered along a line of increasing latitude or northward velocities, and measure the length of the line (Fig. 3(c)) and the averaged pairwise distances between the trajectories (Fig. 3(d)) every Poincaré section (the latter measure is common in atmospherical field experiments [6]). We repeat the above numerical experiment for both $\epsilon = 0$ and $\epsilon \neq 0$, and for two sets of initial conditions: we place one set of initial condition inside the $1:1$ resonance zone and one across the hyperbolic points of the $1:1$ resonance (using the symmetry we can locate

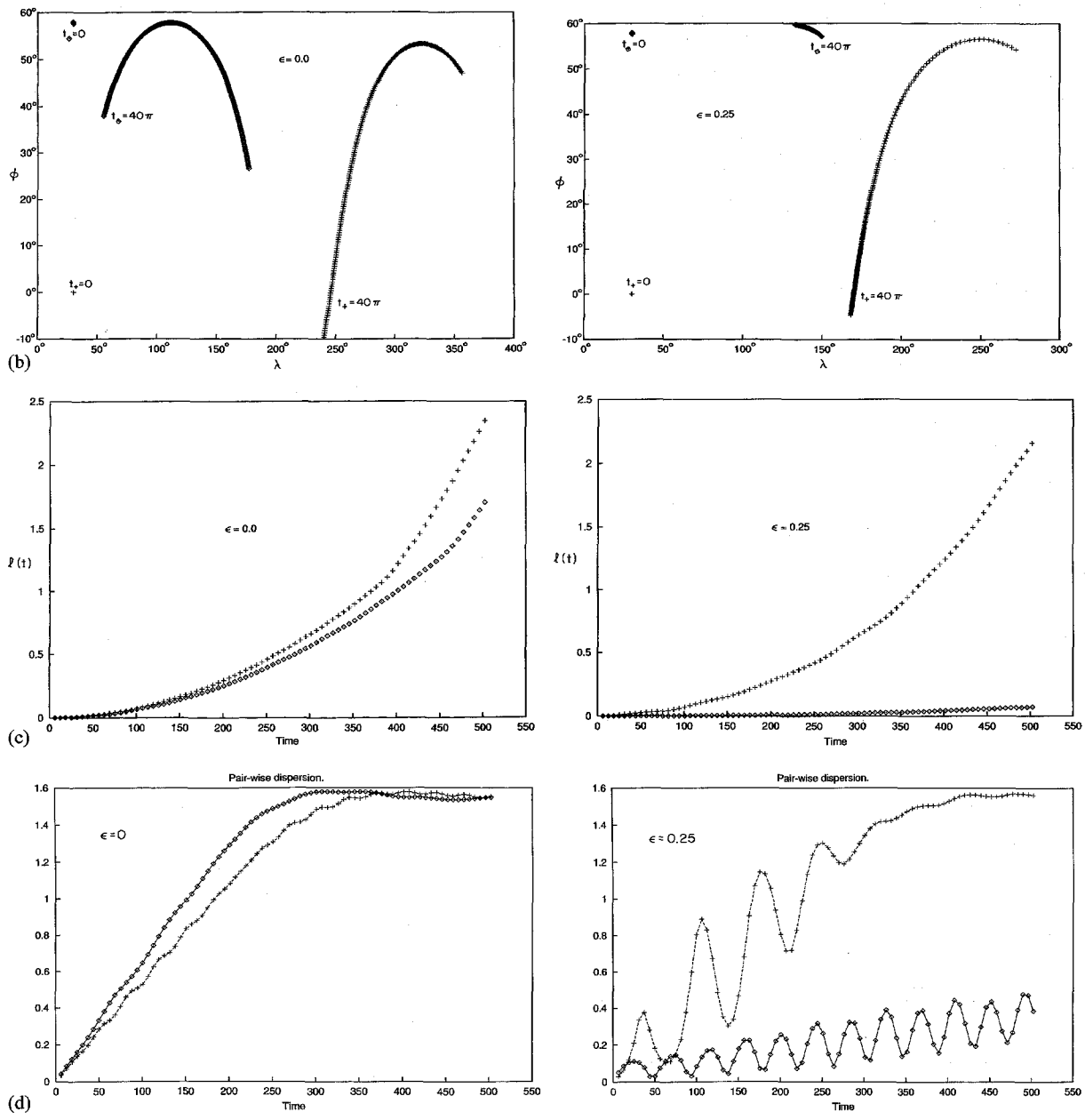

Fig. 3. Continued

these points accurately for $\epsilon \neq 0$), see Fig. 3(a). The over stability property is clearly seen. Theoretically, for $\epsilon \neq 0$ the lines placed across the hyperbolic points should cause exponential stretching and a dispersion over a region of size $O(\sqrt{\epsilon})$. These phenomena are not observable in Fig. 3: the first phenomenon occurs in an exponentially small neighborhood of the hyperbolic point, hence it is not expected to be significant physically, and it is not detected numerically by simple means. The second phenomenon is expected to be significant at longer times – indeed, it

appears after a longer integration time by which the initial conditions complete at least one cycle. By then the presentation of the curves becomes too cluttered, hence the figures showing these phenomena were not included. We also note that for clarity of presentation we have used large velocities and perturbation parameters, yet the same phenomena occur near all resonances appearing for geophysical velocities as well.

Homoclinic chaos: Several physical quantities are estimated using the calculation of the Melnikov function presented in Fig. 2: the extent of the chaotic zone and the zone of motion restricted to one hemisphere, the stretching rates and the mixing rates. These estimates are quite standard for a fixed $|D|$ value, corresponding to the properties of a single trajectory, and appear in various papers dealing with one-and-a-half d.o.f. Hamiltonian system [25–28]. However, fixing D is not physical when ensembles of initial conditions are considered. We extend some of these classical estimates to apply for physical ensembles with *variable D values*.

3.3.1. Chaotic zones in ϕ, λ, v and u

The basic analytical tool used here is the analytical expression for the distance between the stable and unstable manifolds of the equator near the homoclinic orbit at $q_h(-t_0)$, which is given, for fixed D, finite t_0 and ϵ sufficiently small, by

$$d(-t_0, \theta; \epsilon, \sigma, |D|) = \epsilon \frac{M(t_0, \theta, \sigma, |D|)}{\| \nabla H_0(q_h(-t_0)) \|} + O(\epsilon^2) \tag{3.8}$$

(this expression yields (3.5) for $t_0 = 0$). First, we use the estimate of this distance to measure the width of the chaotic zone (also called the stochastic layer) along the ϕ-axis for a fixed D value. The stochastic layer intersects the positive ϕ-axis near $\phi_{hmax} = \arccos(|D|)$ and near the equator. Near ϕ_{hmax}, half the layer width is:

$$\Delta\phi_D(D, \epsilon, \sigma) = K_1(D, \epsilon, \sigma) \max_\theta d(0, \theta; \epsilon, \sigma, D) = \epsilon K_1 \frac{4|DM_0(\sigma, D)|}{(1 - D^2)^{3/2}}(1 + O(\epsilon)), \tag{3.9}$$

where $K_1 = K_1(\sigma, D, \epsilon) \geq 1$ is an unknown, possibly discontinuous function of its arguments. There are some non-rigorous methods for approximating the asymptotic form [25,27,29] of K_1 for small ϵ or large σ, however, finding a rigorous and/or sharp bound of this function is a difficult problem which remains open despite the vast amount of research of one-and-a-half d.o.f. Hamiltonian system. By itself, without knowledge of K_1 formula (3.9) may seem empty. Nonetheless, we will show that various quantities may be related to (3.9) with the same unknown function K_1, hence some quantitative predictions on, for example, the relative size of the chaotic zones, may be found.

To find the width of the stochastic layer near the equator for fixed D we use the observation that the Melnikov function approximates the change in the kinetic energy under one revolution near the separatrix, and relations (2.3) and (2.4) between the energy and (ϕ, v, D) at $(\phi, 0, D)$:

$$\Delta\phi_{D,\text{equator}}(D, \epsilon, \sigma) = \sqrt{\left|\frac{2\epsilon K_1 M_0(\sigma, D)}{\partial^2 H/\partial\phi^2}\right|} = \sqrt{K_1\epsilon}\sqrt{8\frac{|M_0(\sigma, D)|}{(1 - D^2)}}. \tag{3.10}$$

Using the above estimates for the chaotic zones in ϕ for a fixed angular momentum, we find the extent of the chaotic zones in ϕ for a given u_0, v_0, i.e. we pose the following question: given a fixed eastward (westward) velocity $u_0 > 0$ ($u_0 < 0$), and no initial northward velocity $v_0 = 0$, what are the latitude and longitude positions that will lead to chaotic dynamics for small ϵ?

Since the chaotic dynamics is uniform in λ, we conclude that the chaotic regions will be in the form of strips around the globe, see Fig. 4(a). To find these strips, notice that varying ϕ_0 corresponds to changing D according

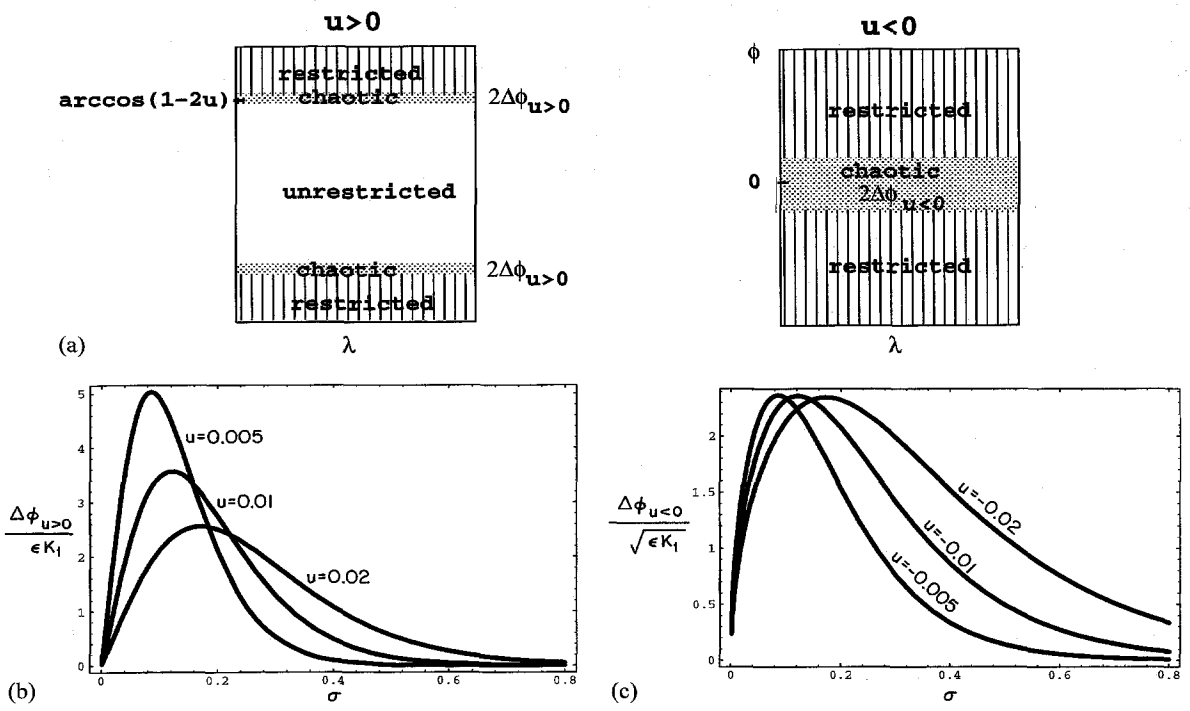

Fig. 4. The extent of the chaotic zone: (a) geometry of the chaotic zone in (λ, ϕ) plane; dashed region–restricted motion, blank–unrestricted motion, dotted – chaotic motion; (b) chaotic extent for $u_0 > 0$; (c) chaotic extent for $u_0 < 0$.

to (2.3a), therefore, for eastward velocities ($0 < u_0 < 1$, with $u_0 \neq \frac{1}{2}$) $\phi_0 > \phi_{\text{ell}}(D(\phi_0, u_0))$ and the chaotic zone occurs for latitudes close to the separatrix latitude (see (2.7))

$$\overline{\phi} = \phi_{\text{hmax}}(1 - 2u_0) = \arccos(|1 - 2u_0|), \tag{3.11}$$

namely the initial condition $(\phi_0, v_0 = 0, \Lambda_0, u_0)$ is chaotic if

$$\phi_0 - \overline{\phi} \in [-\Delta\phi_{u_0>0}, \Delta\phi_{u_0>0}] = [-\Delta\phi_D(D(\overline{\phi} - \Delta\phi_D, u_0), \epsilon, \sigma), \Delta\phi_D(D(\overline{\phi} + \Delta\phi_D, u_0), \epsilon, \sigma)]. \tag{3.12}$$

Expanding (2.3a), (2.7), (3.9) and (3.12) we find

$$\Delta\phi_{u_0>0}(u_0, \epsilon, \sigma) = \frac{\Delta\phi_D(|1 - 2u_0|, \epsilon, \sigma)|}{|1 - 2u_0|} = \frac{\epsilon K_1 |M_0(\sigma, |1 - 2u_0|)|}{2(u_0(1 - u_0))^{3/2}}. \tag{3.13a}$$

Calculating $\Delta\phi$ for $u_0 = 0.005, 0.01, 0.02$ and for $0.001 < \sigma < 0.8$, we find that $\Delta\phi$ is in the form of a hump, which becomes spikier and closer to $\sigma = 0$ as u_0 is decreased, see Fig. 4(b).

For westward velocities ($-1 < u_0 < 0$), $\phi_0 < \phi_{\text{ell}}(D(\phi_0, u_0))$, hence the chaotic zone occurs for latitudes close to the equator namely for $\phi \in [-\Delta\phi_{\text{equator}}, \Delta\phi_{\text{equator}}]$. Using relation similar to (3.12) and expanding (2.3a) and (3.10) we find

$$\Delta\phi_{u_0<0}(u_0, \epsilon, \sigma) \approx \Delta\phi_{D,\text{equator}}(1 + 2u_0, \epsilon, \sigma) = \sqrt{K_1\epsilon}\sqrt{2\frac{|M_0(\sigma, 1 + 2u_0)|}{-u_0(1 + u_0)}}. \tag{3.13b}$$

$\Delta\phi_{u_0<0}(u_0, \epsilon, \sigma)$ is also of the form of a hump which narrows down and approaches $\sigma = 0$ as $|u_0|$ is decreased, yet its maxima does not increase as $u_0 \to 0$, see Fig. 4(c). A preliminary numerical test of the above results has been conducted. We measure the extent of the chaotic zones by computing Lyapunov exponents on a grid of initial conditions with variable (ϕ, λ) and fixed u_0 and $v_0 = 0$. So far we find qualitative agreement with the above estimates (e.g. the insensitivity to $|u_0|$ for $u_0 < 0$ vs. the strong sensitivity to u_0 for $u_0 > 0$), yet quantitative features are still unresolved. In particular the evidently significant role of K_1 on the actual chaotic extent and on the error terms in the above estimates is left for future work.

Now, we estimate the chaotic zones in the velocity field. The width of the chaotic zone in v along the line $\phi = \phi_{\rm ell}(D) = \arccos(\sqrt{|D|})$, corresponding to ensemble with variable v_0 and fixed u_0 ($u_0 = 0$ if $D > 0$ and $u_0 = -\sqrt{|D|}$ if $D < 0$), may be easily found as this ensemble keeps D fixed. The chaotic zone is located near

$$v_{\max}(\phi_{\rm ell}) = \tfrac{1}{2}(\cos(\phi_{\rm ell})^2 - 1) = |\tfrac{1}{2}(\pm D(\phi_{\rm ell}) \mp 1)| = \tfrac{1}{2}(1 - D(\phi_{\rm ell})), \tag{3.14}$$

and its width is found from (3.8) with $q_{\rm h}(-t_0) = (\phi_{\rm ell}(D), v_{\max}(\phi_{\rm ell})$ to be

$$\begin{aligned}\Delta v(D, \epsilon, \sigma) = \Delta v(\phi_{\rm ell}(D), \epsilon, \sigma) &= K_1 \left(\epsilon \frac{M_0(\sigma, \cos^2(\phi_{\rm ell}))}{v_{\max}(\phi_{\rm ell})} + \mathrm{O}(\epsilon^2)\right) \\ &= K_1 \left(\epsilon \frac{M_0(\sigma, \cos^2(\phi_{\rm ell}))}{(1 - \cos(\phi_{\rm ell})^2)/2} + \mathrm{O}(\epsilon^2)\right).\end{aligned} \tag{3.15}$$

Fixing the initial latitude at the equator and $u_0 \in (-1, 0)$, we obtain, as in (3.10), an $\mathrm{O}(\sqrt{\epsilon})$ chaotic zone in v_0, associated with the hyperbolic nature of the equations near the equator:

$$|\Delta v_{\rm equator}| \le \sqrt{\epsilon K_1 M_0(\sigma, 1 - 2|u_0|)}. \tag{3.16}$$

Finally, the chaotic zone in the eastward initial velocity is estimated. Fixing the latitude at ϕ_0 and the initial northward velocity at 0, we conclude that initial eastward velocities leading to chaotic behavior belong to one of the two intervals $(u_{\min,\max}(\phi_0, 0, \sigma) - \Delta u, u_{\min,\max}(\phi_0, 0, \sigma) + \Delta u)$ where

$$\begin{aligned}&u_{\min}(\phi_0, 0, \sigma) = -\tfrac{1}{2}(1 + \cos(\phi_0)), \\ &u_{\max}(\phi_0, 0, \sigma) = \tfrac{1}{2}(1 - \cos(\phi_0)) = u_{\min} + 1.\end{aligned} \tag{3.17}$$

Δu is defined implicitly by requiring the latitude ϕ_0 to belong to the boundary of the chaotic zone with the corresponding angular momentum:

$$\phi_{\rm hmax}(D(\phi_0, u_{\min,\max} \pm \Delta u)) \pm s_{\min,\max} \Delta\phi_D(D(\phi_0, u_{\min,\max} \pm \Delta u), \epsilon, \sigma) = \phi_0, \tag{3.18}$$

where $s_{\min} = -1$, $s_{\max} = 1$ and $D(\phi_0, u)$ is found from (2.3). Expanding (3.18) in ϵ and using (3.17), (2.9) and (2.4) we find:

$$\Delta u = \epsilon K_1 \frac{2|M_0(\sigma, \cos(\phi_0))|}{\sin(\phi_0)^2}. \tag{3.19}$$

Using the same tools one can find an expression (a rather complicated one) for $\Delta u(\phi_0, v_0, \epsilon, \sigma)$ – the chaotic zone in u_0 for a given initial (ϕ_0, v_0).

We conclude with two remarks. First, notice that in each of the above measures of chaos (3.9), (3.13), (3.15) and (3.19) the Melnikov function appears with a different normalization factor which depends on ϕ_0 (or $\phi_{\rm ell}$) in a nontrivial manner, hence there is no one frequency producing the most chaos for all different ensembles. Second, in practice one may be interested in specific range of energies and initial velocities and then the above formulae may be used. A global estimate, over all latitude positions/kinetic energies, are formulated in Section 3.3.4.

3.3.2. Maximal zone of stability in ϕ, v and u

When $\epsilon = 0$, the elliptic fixed point/periodic orbit located at latitude $\phi_{\rm ell} = \arccos(\sqrt{|D|})$ is surrounded by a zone of stability; a particle placed at latitude $\phi_{\rm ell}$ with the same initial westward velocity u_0 as of $\phi_{\rm ell}$ (from (3.1) and (2.4), $u_0 = 0$ and $D = D(\phi_{\rm ell}) = \cos^2(\phi_{\rm ell})$, or $u_0 = -\cos\phi_{\rm ell}$ and $D = -D(\phi_{\rm ell})$) and with a sufficiently small initial northward velocity v_0, will oscillate about $\phi_{\rm ell}$ in a quasi-periodic motion restricted to one hemisphere. For $v_0 \ll 1$ the motion is clearly restricted to one hemisphere and one of its frequencies is close to the inertial frequency. In fact, the motion is restricted in ϕ as long as $|v_0| < v_{\max}(\phi_{\rm ell})$ where $v_{\max}(\phi_{\rm ell})$ denotes the velocity on the separatrix at $\phi = \phi_{\rm ell}$ and $D = \pm D(\phi_{\rm ell})$ (3.14). If $|v_0| > v_{\max}(\phi_{\rm ell})$, the orbit is unrestricted in ϕ – the particle travels to both hemispheres in a quasi-periodic fashion.

Clearly, for small ϵ, the perturbed flow has a zone of stability as well. Mathematically it is simple to observe that by fixing $D = \pm D(\phi_{\rm ell})$ the one-and-a-half d.o.f. Hamiltonian system (3.1c) and (3.1d) has a zone of stability bounded by the largest KAM torus which survives. There are no analytical methods for systematically finding the extent of this KAM torus to a given accuracy – the unknown function K_1 of (3.9) reflects this fact. Instead, we use the previous estimates for the chaotic zone to estimate the restricted motion zones.

For fixed D and $v_0 = 0$, the motion is restricted to one hemisphere for ϕ_0 in the intervals $\pm(\phi_{\min}, \phi_{\max})$:

$$\begin{aligned}
\phi_{\min}(D(\phi_{\rm ell}), \epsilon, \sigma) &= \Delta_{D,\text{equator}}(D, \epsilon, \sigma) = \sqrt{K_1\epsilon}\sqrt{8\frac{|M_0(\sigma, \cos^2(\phi_{\rm ell}))|}{(1-\cos(\phi_{\rm ell})^4)}}, \\
\phi_{\max}(D(\phi_{\rm ell}), \epsilon, \sigma) &= \phi_{\rm hmax}(D) - \Delta\phi_D(D, \epsilon, \sigma) \\
&= \arccos(\cos(\phi_{\rm ell})^2) - \epsilon K_1 \frac{4|M_0(\sigma, \cos^2(\phi_{\rm ell}))|\cos^2(\phi_{\rm ell})}{(1-\cos(\phi_{\rm ell})^4)^{3/2}}.
\end{aligned} \tag{3.20}$$

Fixing $u_0 \in (-1, 1)$ and $v_0 = 0$, the range of ϕ_0 leading to restricted motion is found from (3.12) and (3.13) to be of a different nature (see Fig. 4):

$$\begin{aligned}
&\phi_0 \in \pm(\arccos(|1-2u_0|) + \Delta\phi_{u_0>0}(u_0, \epsilon, \sigma), \tfrac{1}{2}\pi) \quad \text{for } u_0 > 0, \\
&\phi_0 \in \pm(\Delta\phi_{u_0<0}(u_0, \epsilon, \sigma), \tfrac{1}{2}\pi) \quad \text{for } u_0 < 0.
\end{aligned} \tag{3.21}$$

The maximal northward velocity of a particle placed at latitude $\phi_0 = \phi_{\rm ell}$ with the same initial westward velocity as of $\phi_{\rm ell}$ (i.e. $u_0 = 0$ or $u_0 = -\cos(\phi_{\rm ell})$) which is restricted to stay in one hemisphere is found using (3.15):

$$\begin{aligned}
v_{\max}(\phi_{\rm ell}, \epsilon, \sigma) &= \min_\theta v_{\max}(\phi_{\rm ell}, \theta, \epsilon, \sigma) < v_{\max}(\phi_{\rm ell}) - \epsilon K_1 \frac{M_0(\sigma, \cos^2(\phi_{\rm ell}))}{v_{\max}(\phi_{\rm ell})} + \mathrm{O}(\epsilon^2) \\
&= \frac{1}{2}(1-\cos(\phi_{\rm ell})^2)\left(1 - \epsilon K_1 \frac{M_0(\sigma, \cos^2(\phi_{\rm ell}))}{(1-\cos(\phi_{\rm ell})^2)^2/4}\right) + \mathrm{O}(\epsilon^2).
\end{aligned} \tag{3.22}$$

Notice that D is independent of v_0, hence $v_{\max}(\phi_{\rm ell}, \epsilon, \sigma)$ gives the extent in v of the stability zone at $D = \pm\cos(\phi_{\rm ell})^2$.

Finally, fixing ϕ_0 we find the maximal interval of u_0, $(u_{\min}(\phi_0, \epsilon, \sigma), u_{\max}(\phi_0, \epsilon, \sigma))$ such that an initial condition at latitude ϕ_0 with zero northward velocity will be restricted to one hemisphere. Here D varies with u_0 according to (2.4). Using (3.19) we find

$$\begin{aligned}
u_{\min,\max}(\phi_0, \epsilon, \sigma) &= u_{\min,\max}(\phi_0, 0, \sigma) - s_{\min,\max}\epsilon K_1 \frac{d(\sigma, \cos(\phi_0))}{(\partial\phi_{\rm hmax}/\partial D)(\partial D/\partial u)} \\
&= u_{\min,\max}(\phi_0, 0, \sigma) - s_{\min,\max}\epsilon K_1 \frac{2|M_0(\sigma, \cos(\phi_0))|}{\sin(\phi_0)^2}.
\end{aligned} \tag{3.23}$$

3.3.3. Core area

For a given D value, the area of the invariant region is estimated by

$$\mu(A_{\text{core}}(D)) \approx I(\phi_{\max}(D, \epsilon, \sigma), D),$$

where $\phi_{\max}$ is defined by (3.20) and the action $I(\phi_{\max}, D)$ is given in Appendix A (Eq. (A.7)). As discussed above, the condition that D remains fixed renders the problem non-physical. We define instead the core volume V_{core} to be the volume in (ϕ, v, λ, u) space inside of which only restricted motion is possible (these definitions are similar to the quantities defined for two-dimensional quasi-periodic flows, see [30]). Then

$$\begin{aligned} V_{\text{core}}(\epsilon, \sigma) &\approx 2 \int_{-1}^{1} \int_{\phi_{\min}}^{\phi_{\max}} \int_{0}^{2\pi} v \cos(\phi)\, \mathrm{d}\lambda\, \mathrm{d}\phi \frac{\mathrm{d}u}{\mathrm{d}D}\, \mathrm{d}D \\ &= \pi \int_{-1}^{1} \int_{\phi_{\min}}^{\phi_{\max}} v\, \mathrm{d}\phi\, \mathrm{d}D = \pi \int_{-1}^{1} I(\phi_{\max}(D, \epsilon, \sigma), D)\, \mathrm{d}D. \end{aligned} \tag{3.24}$$

We normalize this by the unperturbed core volume – the volume enclosed by the separatrices for all D. Using (2.8), we find

$$V_{\text{core}}(0) = \pi \int_{-1}^{1} \mu(R_1(D))\, \mathrm{d}D = 2\pi \int_{0}^{\pi/2} \sin(\phi)(\sin(\phi) - \phi\cos(\phi))\, \mathrm{d}\phi = \frac{\pi^2}{4}. \tag{3.25}$$

Clearly the prediction gives a smooth dependence of V_{core} on ϵ as it is insensitive to the breakup of KAM tori which causes, in reality, a discontinuous dependence of V_{core} on ϵ and σ (which is lurking in the coefficient K_1).

3.3.4. Area of stochastic zone

Similar to the previous calculation, the area of the chaotic zone for a fixed D may be estimated by

$$\mu(A_{\text{st}}(D)) \geq I(\phi_{\text{hmax}}(D) + K_1 d(\sigma, D)) - 2I(\phi_{\max}, D).$$

Here, since the calculation is valid near the separatrices, the above estimate may be used to approximate the area of the stochastic zone associated with a given westward velocity near the equator. Integrating over all relevant values of the angular momentum we find

$$V_{\text{st}}(\epsilon, \sigma) \approx \pi \int_{-1}^{1} I(\phi_{\text{hmax}}(D) + K_1 d(\sigma, D), D) dD - 2V_{\text{core}}(\epsilon, \sigma).$$

3.3.5. Mixing – Lobe area

Every Poincaré map, the phase-space area transferred between the restricted and unrestricted motion for a fixed D value is given by the area of the turnstile lobes [28,31], which may be approximated by integrating the Melnikov function between two subsequent zeros [28]:

$$\mu(L) = \left| \int_{0}^{\pi/\sigma} \epsilon M_0(\sigma, D) \sin(\sigma t_0)\, \mathrm{d}t \right| = \frac{2\epsilon |M_0(\sigma, D)|}{\sigma}. \tag{3.26}$$

The largest flux (transfer-per-unit-time) between restricted and unrestricted motion is achieved when

$$F(D,\sigma)=\frac{\mu(L)}{2\pi/\sigma} \tag{3.27}$$

is maximal. The probability of transfer between restricted and unrestricted motions may be roughly estimated, for initial conditions in the stochastic layer [31], by

$$P_{\text{trans}}(D,\epsilon,\sigma)=\frac{4\mu(L)}{(2\pi/\sigma)\mu(A_{\text{st}})}. \tag{3.28}$$

Namely, if P_{trans} is large, the orbits in the stochastic layer fluctuate between restricted and unrestricted motions quite frequently whereas small P_{trans} corresponds to the existence of many orbits which travel in one hemisphere for long times, then transfer to unrestricted motion and stay unrestricted for a long time and so on. The probability of transfer from restricted to unrestricted motion in ϕ may be similarly defined and calculated. The relation between the fixed D results and variable D may be found as in Section 3.3.1. More accurate estimates for finite time scales of the transport rates between the north and south hemispheres may be found using the topological approximation method (TAM), a method which uses the templates created by the manifolds structure (the Birkhoff signature) to approximate the transport and the stretching rates [32].

Summary – characteristics of chaos for $k = 0$. The list given above shows that there are several qualitative and quantitative characteristics for the chaotic behavior which may be estimated using the period function, the action and the Melnikov integral. For fixed angular momentum D and variable σ the typical one-and-a-half d.o.f. Hamiltonian system scenario emerges (see Fig. 2) – for large σ values the Melnikov integral approaches zero exponentially – leading to small chaotic regions, small lobes and small stretching rate. For each D there exists an optimal σ for which $M_0(\sigma, D)$ is maximal, and $M_0(\sigma, D) \to 0$ as $\sigma \to 0$. All the measures of the extent of chaos reflect this dependence on σ as these depend monotonically on $K_1 M_0(\sigma, D)$ for fixed D (i.e. fixed $\phi_0, \phi_{\text{ell}}$ in Sections 3.3.1–3.3.5). The resonances, lobe areas and the stretching rates depend explicitly on σ, hence their dependence changes with σ and M accordingly. For example, as σ tends to 0 a delicate balance between the various terms arises – the lobe area increases as σ decreases, the stretching rates increase and other pathological effects, associated with the adiabatic nature of this limit, arise. We are currently working on calculating and interpreting two other classical estimates – the energy deviation estimates and the stretching rates.

Changing D (by varying the initial westward velocity or the initial latitude) gives rise to a completely different scenario as it corresponds to variations occurring between different unperturbed structures. Here, the magnitudes of the Melnikov function and of K_1 are not sufficient for describing the properties of the chaotic flow. For example, fixing σ, $M_0(\sigma, D)$, attains its maxima at $D = 0$ whereas the maximal distance, $d(\sigma, |D|)$ attains its maxima at some $0 < \overline{D} < 1$. Therefore, given a value of σ, while the kinetic energy variations will be maximal near $D = 0$ ($u_0(\phi = 0, D = 0) = -0.5$), there are two "optimal" westward velocities at the equator $u_0^1 = u(0, -\overline{D}; \sigma) < -0.5 < u_0^2 = u(0, \overline{D}; \sigma)$ for which maximal chaotic zone in ϕ, v plane will be attained (see Fig. 2(b)). The first such velocity, u_0^1, is large in magnitude, hence unphysical for geophysical applications while the other, u_0^2, is directly relevant to flows in the atmosphere or the ocean at values of 0.001–0.01 moving westward at the equator or eastward at $\overline{\phi} = \arccos(1 - 2|u_0^2|)$. We observe that $\overline{D}(\sigma) \to 1$ as $\sigma \to 0$ and $d(\sigma, \overline{D}(\sigma))$ seems to increase with no bound as $\sigma \to 0$ (though the limits: $\sigma \to 0$ with D held fixed, and $D \to 1$ with σ held fixed, lead, as expected, to vanishing d).

4. Forcing with finite wavelength ($k \neq 0$)

4.1. Equations of motion

As the angular momentum, D, varies with time according to (2.9), the equation for (ϕ, v) (latitude and northward velocity) and the one for (λ, D) (longitude and angular momentum) are coupled. Introduce the wave velocity c:

$$c = \sigma/k, \tag{4.1}$$

and define Λ, the conjugate variable of D, by

$$\Lambda = \tfrac{1}{2}(\lambda - ct). \tag{4.2}$$

Thus, Λ is the longitude coordinate in the travelling wave moving frame. The system (2.1) is replaced by

$$\frac{\mathrm{d}\Lambda}{\mathrm{d}t} = \frac{1}{4}\frac{D}{\cos^2\phi} - \frac{1}{2}\left(c + \frac{1}{2}\right) = \frac{\partial H}{\partial D}, \tag{4.3a}$$

$$\frac{\mathrm{d}D}{\mathrm{d}t} = -2k\epsilon A(\phi)\cos(2k\Lambda + \theta) = -\frac{\partial H}{\partial \Lambda}, \tag{4.3b}$$

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = v = \frac{\partial H}{\partial v}, \tag{4.3c}$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{1}{8}\sin(2\phi)\left(1 - \frac{D^2}{\cos^4\phi}\right) - \epsilon A'(\phi)\sin(2k\Lambda + \theta) = -\frac{\partial H}{\partial \phi}, \tag{4.3d}$$

where H, the Hamiltonian, is given by

$$H(\phi, v, \Lambda, D) = \frac{v^2}{2} + \frac{1}{8}\left(\frac{D}{\cos\phi} - \cos\phi\right)^2 - \frac{c}{2}D + \epsilon A(\phi)\sin(2k\Lambda + \theta). \tag{4.4}$$

Using the symmetries of (4.3) and (2.2), and requiring, on physical grounds, that $\sigma \geq 0$, it is easy to show that the relevant parameter ranges for (4.3) are:

$$-\tfrac{1}{2} \leq c < \infty, \quad \operatorname{sign}(k) = \operatorname{sign}(c), \qquad k \in \mathbb{Z}, \ \epsilon \geq 0, \ 0 \leq \theta < 2\pi, \tag{4.5}$$

as all other values may be reduced to the above. Mathematically, the natural period for Λ is π/k, however, physically, it is π. When considering dispersion of particles, this distinction is important, hence we keep $\Lambda \in [0, \pi)$. Note that (4.3) is *not symmetric* with respect to reflections in D: when $D \to -D$ the equation for Λ changes in a non-invariant manner (though for the unperturbed case, the system in (ϕ, v) is symmetric with respect to reflections in D). This asymmetry is due to the unidirectional rotation of the earth.

4.2. The energy surfaces

Eqs. (4.3) constitute a weakly coupled two d.o.f. Hamiltonian system, and their solutions lie on three-dimensional energy surfaces in (v, ϕ, D, Λ) phase space, given by level sets of H of (4.4). The structure of the energy surface serves as a backbone for understanding the solution structure for specific values of the parameters and for the various regions in phase space. Fixing $\Lambda = \theta = 0$, the cross-section of the energy surface is two-dimensional, and is presented in Figs. 5 and 6 for increasing value of the energy level H and for two typical c values. Other cross-sections, corresponding to different values of Λ, may be viewed as changing the value of H by an $O(\epsilon)$ amount and as slight deformation of the surface if $A(\phi)$ is not constant.

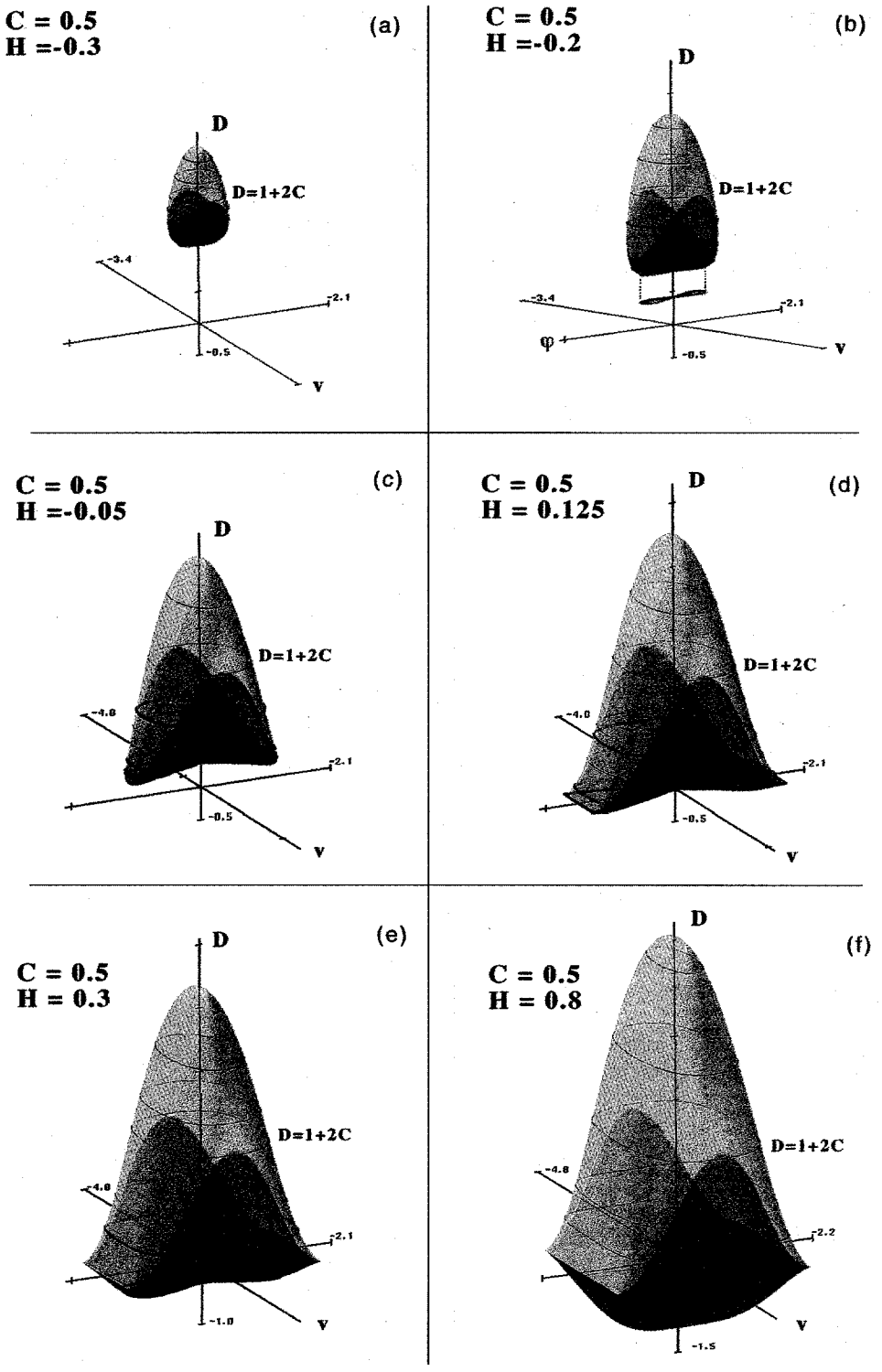

Fig. 5. The energy surfaces for $c > 0$, $\Lambda = 0$: yellow shade $d\Lambda/dt > 0$, blue shade $d\Lambda/dt < 0$, green/red curves $(1, \pm 1)$ resonances of zonally rotational/oscillatory type.

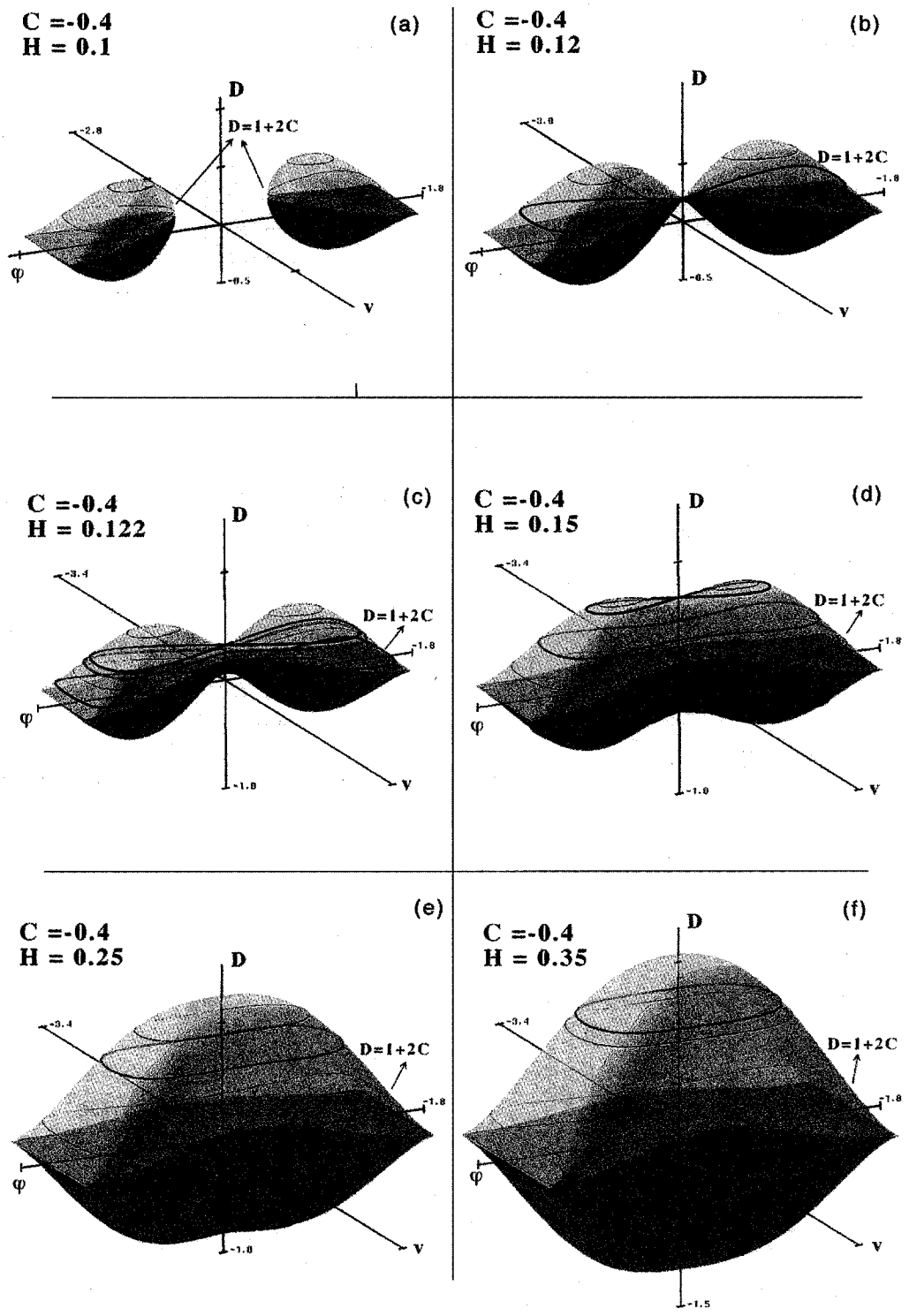

Fig. 6. The energy surfaces for $c < 0$, $\Lambda = 0$; same color-code as in Fig. 5.

The unperturbed motion occurs on constant D lines on the energy surfaces – a few of these level curves are drawn in black on them. To present the unperturbed motion in the Λ direction in a compact fashion we color the regions on the energy surface for which $\mathrm{d}\Lambda/\mathrm{d}t$ is positive along the unperturbed orbits by light (yellow–green) shading and the regions on which it is negative by dark (blue) shading. This coloring scheme enables us to structurally read the four-dimensional flow from the two-dimensional energy surface plot as described in Section 4.2.2. Moreover, we color the D-level set for which strong resonances occur: red line for oscillatory resonance in Λ ($1:0$ resonance) and bright green line for strong rotational ($1:\pm 1$) resonances between ϕ and Λ (see Section 4.4 for details). Fig. 7 presents this information using the energy–momentum bifurcation diagram: the regions in which back-flow (regions in which $\mathrm{d}\Lambda/\mathrm{d}t$ changes its sign along unperturbed orbits) occur are shaded, oscillatory resonances are marked and the stability properties of the singular circles are indicated.

In Section 4.2.1 the construction of the energy surfaces and the corresponding bifurcation diagram of the energy–momentum map (Fig. 7) for $0 < c < \infty$ (Fig. 5) and $-\frac{1}{2} < c < 0$ (Fig. 6) is described. In Section 4.3 we interpret these results and find all typical and degenerate dynamical behavior according to the above bifurcation diagrams.

4.2.1. Construction of the energy surfaces

It follows from (4.4) that at the origin

$$H_{\mathrm{or}}(D) = H(0,0,0,D) = \tfrac{1}{8}(D-1)^2 - \tfrac{1}{2}cD, \tag{4.6}$$

so for a given H, the energy surface touches the surface $\phi = v = 0$ of singular circles at $D_{\pm}(H)$:

$$D_{\pm}(H) = (1+2c) \pm 2\sqrt{2H + c(1+c)}. \tag{4.7}$$

This curve is plotted in Fig. 7 as a dashed line for hyperbolic invariant circle (i.e. for $|D| < 1$) and as solid curve for the non-hyperbolic case ($|D| \geq 1$).

The minimal value of H for which an energy surface includes the origin is given by

$$H_{\min} = \min_{D} H_{\mathrm{or}}(D) = -\tfrac{1}{2}c(1+c), \tag{4.8}$$

with the minimizing D value

$$D_{\mathrm{r}} = 1 + 2c. \tag{4.9}$$

Consider the case $c > 0$. Then $D_{+}(H) > 1$ for all $H > H_{\min}$, and $D_{-}(H) \geq 1$ for $H \in [H_{\min}, H_{-c/2}]$ where (see Fig. 7(a)):

$$H_{-c/2} = \min_{H}\{D_{-\mathrm{sign}(c)} = 1\} = -\tfrac{1}{2}c. \tag{4.10}$$

From (4.4), for $\Lambda = \theta = 0$,

$$\frac{v^2}{2} = H - \frac{1}{8}\left(\frac{D}{\cos\phi} - \cos\phi\right)^2 + \frac{c}{2}D, \tag{4.11}$$

which implies that for $c > 0$ only $H \geq H_{\min}$ need to be considered. For $D \in (D_{-}(H), D_{+}(H))$ the r.h.s. of (4.11) is positive at $\phi = 0$, hence the energy surfaces include only unrestricted orbits for these values of H's and D's, as depicted in Fig. 5(a). In the energy–momentum map (Fig. 7) the allowed regions (the regions for which (4.11) has real solutions) are indicated by stripes.

To color the surfaces, the behavior of $\mathrm{d}\Lambda/\mathrm{d}t$ is investigated. It follows from (4.3a) that

$$\left.\frac{\mathrm{d}\Lambda}{\mathrm{d}t}\right|_{\phi=0} = \tfrac{1}{4}(D-1-2c),$$

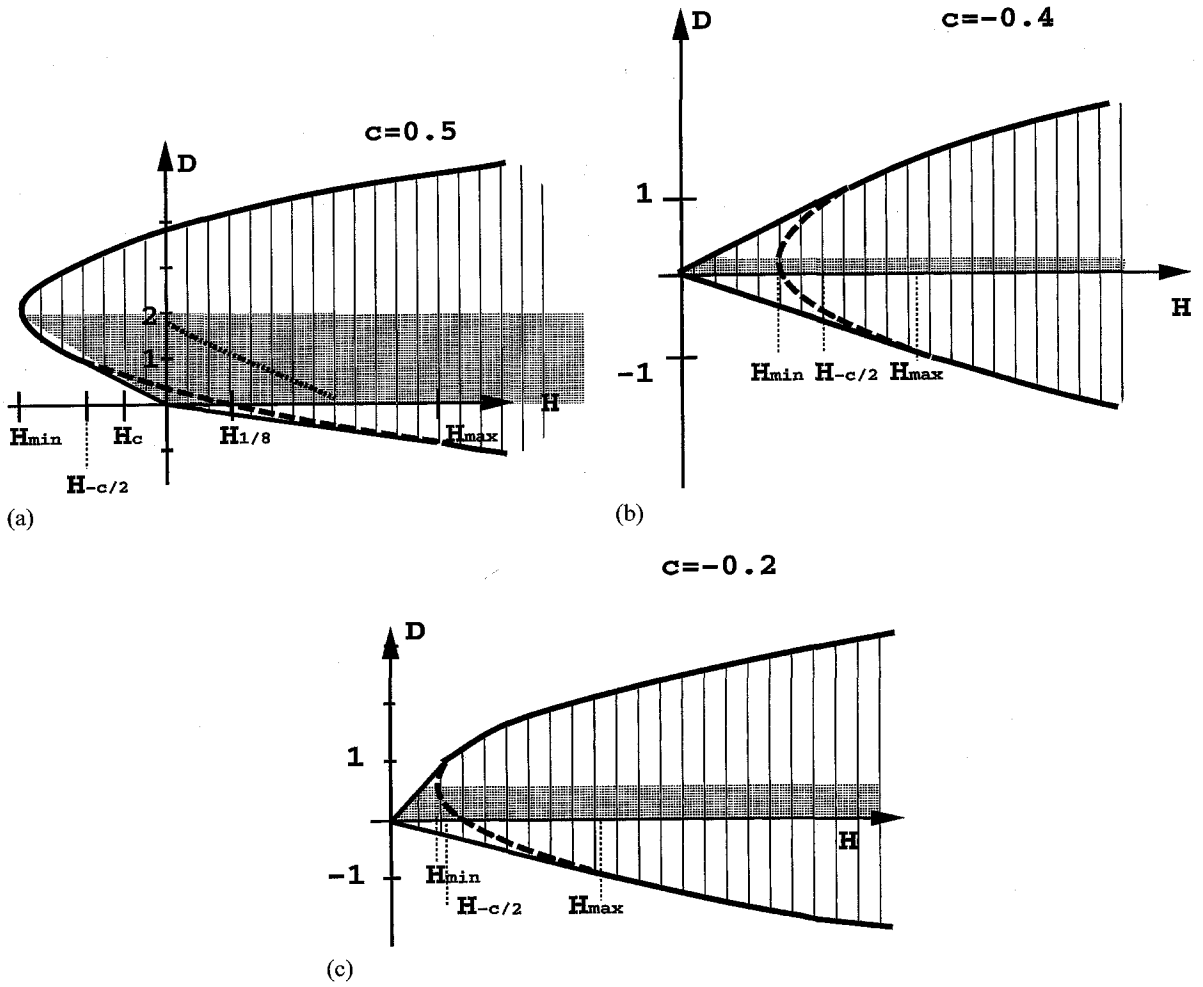

Fig. 7. Energy–momentum bifurcation diagram: striped area – allowed regions for motion: shaded region – backflow in Λ; $(\cdots)$ – oscilattory resonaces; _____ stable periodic orbit; (- - -) – unstable periodic orbit.

hence $\mathrm{d}\Lambda/\mathrm{d}t$ is positive at $\phi = 0$ for $D > 1 + 2c$ and negative for $D < 1 + 2c$. In fact, (4.3) implies that $\mathrm{d}\Lambda/\mathrm{d}t$ changes sign at $\phi = \phi_{\mathrm{r}}$ where

$$\cos^2 \phi_{\mathrm{r}} = \frac{D}{2c+1}, \tag{4.12}$$

hence the line $\phi = \phi_{\mathrm{r}}$ defines the border between yellow and blue on the energy surfaces. It follows that $\mathrm{d}\Lambda/\mathrm{d}t$ may change sign along orbits only if $0 < D < D_{\mathrm{r}}$ where $D_{\mathrm{r}} = 1 + 2c$ as in (4.9). Substituting (4.12) in (4.11) we restrict D so that v^2 is non-negative at $\phi = \phi_{\mathrm{r}}$; since

$$v^2 = 2H + \frac{c(1+c)D}{1+2c},$$

and the r.h.s. vanishes for $D = D_\Delta(H)$,

$$D_\Delta(H) = -\frac{2H}{c}\left(1 + \frac{c}{1+c}\right), \tag{4.13}$$

it follows that $d\Lambda/dt$ vanishes on a given energy surface H (at $\phi = \phi_r$) for $D \in [\max\{0, D_\Delta(H)\}, D_r]$ for $c > 0$ and for $D \in [0, D_r]$ for $c < 0$. These regions on which backflow occurs are shaded on the energy–momentum map (Fig. 7).

For $c = 0.5$, $H_{\min} = -0.375$ and $H_{-c/2} = -0.25$, hence Fig. 5(a) falls into this first range. The changes in the energy surfaces for $c > 0$ as H is further increased from its minimal value $H_{\min}$ are briefly described. At $H = H_{-c/2}$, $D_-(H) = 1$, so that the bifurcation point lies on this energy surface. Increasing H further brings the separatrix and the elliptic points onto the energy surface. Eqs. (2.5) and (4.4) imply, in this case, that the elliptic points on the energy surface H have angular momentum:

$$D_{\text{ell}}(H) = \begin{cases} \dfrac{-2H}{c}, & D > 0, \\ \dfrac{-2H}{c+1}, & D < 0. \end{cases} \tag{4.14}$$

Also, Eq. (4.7) implies that $|D_-(H)| < 1$ for $H_{-c/2} \le H \le H_{\max}$ where

$$H_{\max} = \max_H \{H \mid D_-(H) \ge -1\} = H_o(-1) = \tfrac{1}{2}(c+1), \tag{4.15}$$

hence the separatrices emanating from $(0, 0, D_-(H), \Lambda)$ and the elliptic points are contained in the energy surface with energy H for such H values, as depicted in Figs. 5(b)–(e) and 7.

The difference between the first of these three panels, Figs. 5(b), (c) and (d), is the different behavior of $d\Lambda/dt$ in the vicinity of the separatrices; indeed, as long as $D_-(H) < D_\Delta(H)$, $d\Lambda/dt$ may not vanish along the separatrix, as shown in Fig. 5(b). Since $D_-(H) = D_\Delta(H)$ when $\phi_r = \phi_{\text{hmax}}$, we use (2.7) and (4.12) to conclude that at that point $D_-(H) = D_\Delta(H) = D_c$, where

$$D_c = \frac{1}{2c+1}, \tag{4.16}$$

and this determines the critical value of H:

$$H_c = H_o(D_c) = \frac{1}{8}\left(\frac{1}{(1+2c)^2} - 1\right). \tag{4.17}$$

Hence, for $H \in (H_{-c/2}, H_c)$, $d\Lambda/dt$ is negative for all (ϕ_h, v_h) as depicted in Fig. 5(b); for $H = H_c$, $d\Lambda/dt$ vanishes at one point along the separatrix, namely, at $\phi = \phi_{\text{hmax}}$, and otherwise it is negative, and for $H > H_c$, there is a region of ϕ values, $\phi_r < \phi < \phi_{\text{hmax}}$, for which $d\Lambda/dt$ is positive whereas for $\phi < \phi_r$ $d\Lambda/dt$ is negative, as shown in Fig. 5(c) ($H_c(c = 0.5) = -3/32 \approx -0.094$). When $H = 0$, $D_{\text{ell}} = D_\Delta(H) = 0$, corresponding to the degenerate behavior at the poles. When $0 < H < \frac{1}{8}$, $D_-(H) > 0$ and $D_{\text{ell}} < 0$, all orbits with $0 < D < 1$ have regions of backflow where $d\Lambda/dt > 0$, while the periodic orbits with $D < 0$ have none. The intersection of the energy surface with the $D = 0$ surface consists of interior orbits, corresponding to periodic motions restricted to one hemisphere, passing through the poles with $v < \frac{1}{2}$. At $H = \frac{1}{8}$ the separatrix reaches $D = 0$, connecting the equator and the poles (Fig. 5(d)). For $\frac{1}{8} < H$ the intersection of the energy surface with $D = 0$ surface contains exterior orbits, corresponding to periodic motions going through both hemispheres, passing through both the north and south poles with $v > \frac{1}{2}$. These motions correspond to high kinetic energies and are not expected to be encountered in observed oceanic or atmospheric flows. For $\frac{1}{8} < H < H_{\max}$ the separatrix and the elliptic points have negative angular momentum ($0 > D_-(H)$, $D_{\text{ell}}(H) > -1$) where $d\Lambda/dt$ does not change sign (Fig. 5(e)). When $H > H_{\max}$ the energy surface includes only exterior periodic orbits, and the periodic orbits with $0 < D < D_r$ have regions of backflow (Fig. 5(f)). To summarize the above construction, Table 1 includes the list of all critical values of H and D for which the solutions behavior changes.

Table 1
Special values of D and H (for $\epsilon = 0$)

Quantity	Value	Significance
$D_{\pm 1}$	± 1	$(\phi, v) = (0, 0)$ undergoes pitchfork bifurcation
D_r	$1 + 2c$	$d\Lambda/dt$ vanishes at $(\phi, v) = (0, 0)$, and it may change sign along orbits only for $D < D_r$
D_c	$1/(1 + 2c)$	for $c > 0$, and $H = H_c$, $d\Lambda/dt = 0$ at $(\phi, v) = (\phi_{hmax}, 0)$ (separatrix)
$D_\Delta(H)$	$(-2H/c(1 + (c/(1 + c)$	for $c > 0$, $d\Lambda/dt \neq 0$ along orbits on energy surface H for $D < D_\Delta(H)$
$D_\pm(H)$	$1 + 2c \pm 2\sqrt{(1 + c)c + 2H}$	D value at $(\phi, v) = (0, 0)$ on energy surface H
$D_{ell}(H)$	$\begin{cases} -2H/c \\ -2H/(c + 1) \end{cases}$	D at elliptic points for $D > 0$ D at elliptic points for $D < 0$ $(H > 0)$
H_0	0	Minimal value of H for $c < 0$.
H_{min}	$-c(1 + c)/2$	$\min_{(\phi,v)=(0,0)} H$, minimal value of H for $c > 0$
$H_{-c/2}$	$-c/2$	$D_{-\text{sign}(c)} = 1$
H_c	$-\frac{1}{2}[c(1 + c)/(1 + 2c)^2]$	for $c > 0$, $D_-(H_c) = D_\Delta(H_c) = D_c$
$H_{1/8}$	$1/8$	$D_-(H_{1/8}) = 0$
H_{max}	$\frac{1}{2}(c + 1)$	$D_-(H_{max}) = D_{ell}(H_{max}) = -1$

Similar analysis for the $c < 0$ case results in Figs. 6, 7(b) and (c). Clearly new structures of the energy surface appear. Most notably, for $0 < H < H_{min}$, there exists an energy surface consisting of two disconnected components (Fig. 6(a)). At $H = H_{min}$ the two components coalesce at the origin (Fig. 6(b)), and for $H_{min} < H < \min\{-\frac{1}{2}c, \frac{1}{8}\}$ there exist two homoclinic orbits emanating from $D_\pm(H)$ with $0 < D_\pm < 1$ (Fig. 6(c)). For $\frac{1}{8} < H < H_{max}$ the lower homoclinic orbit has $D_-(H) \in (0, -1)$ (Figs. 6(d) and (e)) whereas for $H_{min} < H < -\frac{1}{2}c$ the upper homoclinic orbit has $D_+(H) \in (D_r, 1)$ (Figs. 6(c) and (d)). At $c = -\frac{1}{4}$, since $-\frac{1}{2}c = \frac{1}{8}$,the lower and upper separatrices reach $|D| = 1$ on the same energy surface. The behavior near $D = D_r$ for $H \approx H_{min}$ is of special interest and is analyzed in Section 4.6.

In fact, once the structure of the colored energy surfaces is understood, the energy–momentum map, with the shaded area indicating backflow, and the stability type of the singularities indicated, may be used to encode the structure for given c value. For example in Fig. 7(c) we included the case $c = -0.2$, and the structure of its energy surfaces can be readily understood by comparing it to the case $c = -0.4$.

4.3. *Regimes of analysis*

The series of energy surfaces of Figs. 5 and 6 and the structure of the unperturbed motion on them reveal the nature of the three typical structures of 2 d.o.f. Hamiltonian systems which emerge for small non-vanishing pressure waves. Moreover, the critical values of H and D for which a certain degenerate behavior appears arise naturally as the limiting values for which the boundary separating different subregions are approached. The following classification will be employed in the subsequent analysis.

4.3.1. *Typical structures*

(A) *KAM surfaces:* For most values of D on each energy surface [2] the unperturbed motion is quasi-periodic: the motion is periodic in both the latitude and longitude coordinates with periods which are not rationally related (incommensurate). The black level curves on the energy surfaces (Figs. 5 and 6) present a cross-section of KAM tori at a fixed Λ for $\epsilon = 0$. By the iso-energetic KAM theory [33], even for non-zero perturbation (yet sufficiently small)

[2] With the exception of singular surfaces which may appear for the critical wave speeds, see Section 4.3.2 (D).

most of these tori persist thus guaranteeing stability in the four-dimensional phase space: the motion is restricted to bands on the energy surfaces lying between the surviving KAM tori.

(B) *Resonances:* For a countable number of D values on each energy surface the unperturbed motion is periodic – creating resonances when $\epsilon \neq 0$. Initial conditions which are close, in the full phase space, to a stable resonant periodic orbit, undergo quasi-periodic motion of oscillation around this orbit. Therefore, one effect of the resonance is the enhanced stability of the perturbed motion for some initial conditions. Another effect is the strong non-uniformity in Λ, as the behavior of initial conditions near the unstable periodic orbits will result in a completely different motion. Notice that when $\mathrm{d}\Lambda/\mathrm{d}t$ is non-vanishing the resonances are necessarily of rotational nature in Λ, namely Λ covers the interval $[0, \pi]$. The D values for which rotational resonance of low order occurs (the periods in ϕ equals, in magnitude, to the period in Λ) are denoted by green lines in Figs. 5 and 6. When $\mathrm{d}\Lambda/\mathrm{d}t$ changes sign along an orbit (the orbit passes through two colors in Figs. 5 or 6) there exists the possibility of oscillatory resonance, which, under the perturbation, results in localized structures with respect to the travelling wave. The D values for which oscillatory resonances occur are denoted by a red line on the energy surfaces. The size of the oscillatory resonances is expected to be the largest of all resonances.

(C) *Homoclinic chaos:* For isolated values of D on some of the energy surfaces homoclinic loops appear – these are the solutions emanating from the unstable fixed point at the equator (e.g. Figs. 5(b)–(e)). On the energy–momentum map, these homoclinic loops emanate from the unstable invariant circle which is indicated by the dashed line in Fig. 7. The homoclinic loops which occur for $\epsilon = 0$ are not structurally stable and are expected, in general, to break under perturbation. The structure of the perturbed stable and unstable manifolds and the properties of their intersection set (the homoclinic orbits) depend on the behavior of $\mathrm{d}\Lambda/\mathrm{d}t$.

(i) *Monotonic Λ dependence*: When $\mathrm{d}\Lambda/\mathrm{d}t$ is bounded away from zero over an invariant region which includes the separatrix, the system (4.3) is reduced, as in [16], to a one-and-a-half d.o.f. Hamiltonian system depending on the Hamiltonian value H, where Λ replaces time. The standard theory regarding the dynamics associated with homoclinic chaos in one-and-a-half d.o.f. Hamiltonian systems applies [22]. In particular, the return map to a cross-section defined by $\Lambda = \Lambda_0$ is a Poincaré map and therefore it is independent of Λ_0 (up to topological conjugacy). The reduction and the analysis of the reduced system are presented in Section 4.5. For the monotonic case the separatrix and its neighborhood have only one color on its energy surface (e.g. Figs. 5(b), (e), 6(d) and (e)). From Fig. 7 it is easy to identify the intervals in H for which the unperturbed solutions in Λ near the separatrix are monotonic in t – for these intervals the neighborhood of the dashed line has no shading. It follows (see Figs. 7(a)–(c) or Table 1) that such intervals appear for both westward and eastward perturbation waves.

(ii) *Non-monotonic Λ dependence*: When $\mathrm{d}\Lambda/\mathrm{d}t$ changes sign along orbits in the vicinity of the separatrix (but not at the origin) the system is fully three-dimensional on each energy surface. In particular, the return map to a cross-section defined by $\Lambda = \Lambda_0$ is not orientation preserving and there exist Λ_0, Λ_1 for which the return maps are not topologically conjugate. In Section 4.5 we analyze the behavior of the separatrices in this setting, computing the Melnikov vector, and supply the geometrical interpretation of the results. For the non-monotonic case the separatrix and its neighborhood have two colors on its energy surface (e.g. Figs. 5(c) and 6(c)). In Fig. 7 homoclinic chaos with backflow is identified with the regions in which the dashed line is shaded, appearing for both westward and eastward perturbation waves.

4.3.2. *Special structures*

We identify four types of degenerate behavior in the system. In this paper the analysis of the first one is presented, the others are the subject of future work.

(A) *Hyperbolic resonance:* When $\mathrm{d}\Lambda/\mathrm{d}t$ vanishes on the unstable fixed point $(\phi, v) = (0, 0)$, a strong resonance $(1:0)$ occurs between the inertial trajectories and the forcing, leading to the creation of oscillatory

motion along the equator, with homoclinic and heteroclinic (i.e. orbits with different asymptotes for $t \to \pm\infty$) orbits emanating from it. This structure implies that the motion near the equator consists of oscillatory motion and bursts associated with homoclinic and heteroclinic structures. Using the methods developed by Haller and Wiggins [8], Kovacic and Wiggins [18] and Kaper and Kovacic [20] we find in Section 4.6 the mechanism for these bursts, their magnitude and their zonal behavior as a function of the wave speed and wave number. This behavior may be identified in the energy–momentum mapping with the case at which at the bifurcation value $(H_{\min}, D_r)$ the two invariant circles which are created are hyperbolic – namely the fold of the curve $H_{or}(D)$ occurs on the dashed line segment (Figs. 7(b) and (c)). In terms of the energy surfaces, this behavior corresponds to the case where the two separate components of the energy surface meet at the singular circle, which is therefore included in the border between the two colors (Fig. 6(b)).

(B) *Behavior near parabolic point:* For $\epsilon > 0$ and all values of the other parameters, the behavior near $(\phi, v) = (0, 0)$ at $|D| = 1$ is mathematically non-trivial as it involves perturbations near a parabolic point where the eigenvalues are identically zero (see Appendix A). We expect to find exponentially small splitting of the separatrices near this point [17]. In terms of the energy surfaces, this region corresponds to the behavior near $H = H_{\min}$ and $H = H_{\max}$ where the energy surfaces change their singularity structure. For the energy–momentum map at these points the stability of the invariant circles changes (solid and dashed curves meet) and new elliptic points appear.

(C) *Behavior on the $D = 0$ surface:* The singularities introduced by the spherical coordinates at the poles are important near $D = 0$, where the trajectories may reach the neighborhood of the poles. For example, for $D = 0$, $\epsilon = 0$ all orbits (the restricted and unrestricted periodic orbits and the separatrices) pass through the poles and are discontinuous in v there (reflecting the fact that v changes sign at the poles). Analysis of this case is deferred to future work. Note that excluding the vicinity of the poles, $D = 0$ involves fairly large velocities which are unlikely to occur in geophysical flows.

(D) *Critical wave speeds:*

(1) When $c = 0$, $\mathrm{d}\Lambda/\mathrm{d}t = 0$ at the elliptic fixed points $((\phi, v) = (\pm \arccos\sqrt{D}, 0))$, hence strong resonances are expected to occur. Moreover, since $D_r = D_c = 1$, these resonances end at $D = 1$ where a parabolic resonance appears. All these unperturbed fixed points belong to the same singular energy surface on which no KAM tori exist. This creates strong instabilities for $\epsilon \neq 0$, which vastly influence the behavior at $c \approx 0$ [2,35].
(2) When $c = -\frac{1}{2}$, the hyperbolic resonance occurs for $D_r = 0$, combining the two phenomena (A+C) discussed above.
(3) When $c \to \infty$ the motion is reducible to a one-and-a-half d.o.f. Hamiltonian system (with fast oscillating forcing for which exponentially small splitting is expected to appear) on any finite-H energy surface and for all finite D values which are bounded away from zero, see (4.12) and Table 1. On the other hand, near $D = 0$ the dynamics is non-reducible and a careful analysis is yet to be constructed.

4.4. Elliptic–elliptic resonance

Resonances between the periodic motion in Λ and that in ϕ occur when

$$m P_\phi(\phi_{\max}, D) = n P_\Lambda(\phi_{\max}, D)$$

for some integers n, m where

$$\Lambda(P_\Lambda(\phi_{\max}, D)) = \Lambda(0) + l\frac{\pi}{k}, \quad l = 0, \pm 1.$$

Since $\mathrm{d}\Lambda/\mathrm{d}t$ depends on ϕ in a non-trivial manner, $P_\Lambda(\phi_{\max}, D)$ can be found only implicitly. Instead, we consider the return map of ϕ to the cross-section $\phi = \phi_{\max}$ and find $\Lambda_P = \Lambda(P_\phi(\phi_{\max}, D)) - \Lambda(0)$. It follows from (4.3a)

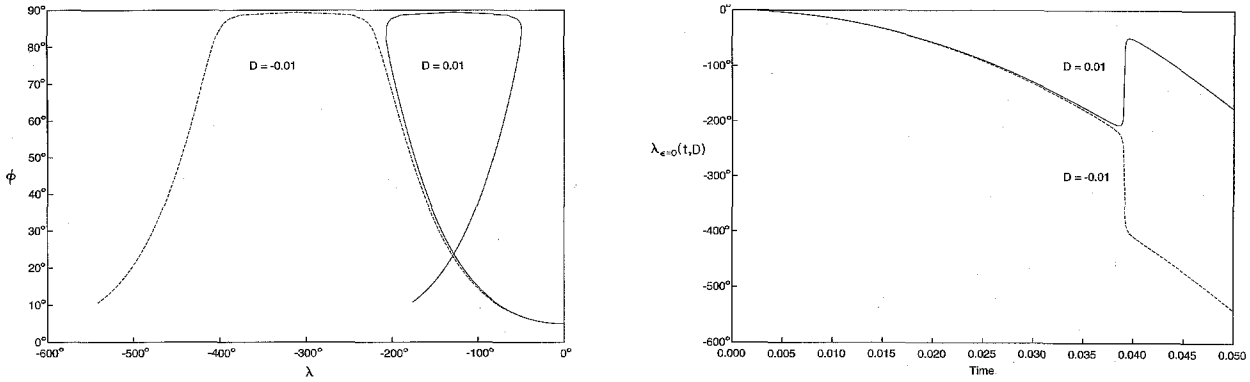

Fig. 8. Change in longitude as $D \to -D$.

that

$$\Lambda(mP_\phi(\phi_{\max}, D)) = m\Lambda_P,$$

hence the resonance condition is

$$\Lambda_P = \frac{nl}{m}\frac{\pi}{k}.$$

The expression for $\Lambda_P(H, D)$ is given by (A.7). The strongest resonance occurs when $\Lambda_P = 0$, corresponding to the creation of a localized structure for which $l = 0$. Such a resonance may occur only in regions where $\mathrm{d}\Lambda/\mathrm{d}t$ changes sign along orbits, hence the colored energy surfaces give a first indication of where such resonances may occur. The red lines on the energy surfaces indicate the level sets of D for which $\Lambda_P = 0$ namely $l = 0$, $m = 1$ and $n = 1$ (resp. $n = 2$) for restricted (resp. unrestricted) orbits in ϕ. The dotted line in Fig. 7 indicates the corresponding resonant values of (H, D) on the energy–momentum bifurcation diagram. For the corresponding resonant orbits, the motion in Λ is oscillatory in nature; Λ does not cover the interval $[0, \pi]$ and in fact is restricted to a π/k interval.

Naively, one would think that for all energy surfaces such an oscillatory resonance must occur since Λ_P is positive for large positive D (the green region) and negative for all $D < 0$. We now explain why this is not the case as is seen in Figs. 6, 7(b),(c) and 5(f), where oscillatory motion does not exist.[3]. In these cases the transfer from positive to negative Λ_P occurs on the surface $D = 0$: for small positive D the orbits encircle the north pole from the east and for small negative D (a larger westward velocity) the orbits encircle the north pole from the west. Therefore, for small D and for motion which is restricted to one hemisphere $\Lambda_P(-D, c) = \Lambda_P(D, c) - \pi$, which implies that the jump in λ near the poles is 360°, as demonstrated in Fig. 8. A similar reasoning shows that for unrestricted periodic motion in ϕ the jump in Λ_P may be π or 2π, depending on the value of c and $\phi_{\max}$.

The green lines on the energy surface indicate the values of D for which $\Lambda_P(D, c) = \pm\pi$ namely $l = \pm 1, m = 1$, $n = k$. Physically such resonances correspond to the appearance of k island chains along the globe, where i.c. in one island jump to the next $n(\mathrm{mod}\, k)$ island after returning to a similar latitude and northward velocity position.

We now discuss the implications of the elliptic–elliptic resonance on the solutions' behavior. The main difference between the resonances of the infinite and finite wavelength cases is that in the latter the D variable is also involved in the resonance creation, hence the phenomena observed in the $k = 0$ case for the ϕ, v variables occur now in the full phase space creating non-uniformity in both Λ and ϕ. It is especially strong near the oscillatory resonances.

[3] We emphasize that even though there are no oscillatory motions present in Fig. 6, there are $c < 0$ values for which oscillatory resonances exist.

Indeed, for $c = 0.5$, $k = 1$, $\epsilon = 0.00025$, near $D = 0.85$ and $H = 0.125$ (Fig. 5(d)) a resonance region of width 0.1 in D appears, concentrated near $\Lambda = \frac{3}{4}\pi$ at the equator.

4.5. Homoclinic chaos

4.5.1. Monotonic Λ dependence

Consider the near separatrix behavior on a given energy surface H for which $\mathrm{d}\Lambda/\mathrm{d}t$ does not vanish near the separatrix for the unperturbed flow, i.e. the dashed line in the energy–momentum bifurcation diagram is not shaded for an open interval of H values near this value of H. Then the system may be reduced near the separatrices to a one-and-a-half d.o.f. Hamiltonian system by substituting Λ for time. The existence of KAM tori guarantees that orbits of the perturbed system, starting close to the unperturbed separatrix, will remain close to it: the orbits are bounded by the smallest unrestricted iso-energetic KAM torus (i.e. the smallest KAM torus on the energy surface H which is not restricted to one hemisphere) and the largest restricted iso-energetic KAM tori. Hence, given a H value in the indicated open intervals, for small enough $\epsilon(H)$, the perturbed orbits have $|\mathrm{d}\Lambda/\mathrm{d}t| > 0$ as well.

The technique described in Section 4.5.2 for studying homoclinic chaos in the non-monotonic-Λ case applies to the monotonic regimes as well. Nonetheless, we find it instructive to include the analysis for the monotonic case separately, so as to emphasize the difference between these cases. We follow the procedure described by Holmes and Marsden [16] to reduce the four-dimensional system (4.3) to a one-and-a-half d.o.f. Hamiltonian system when $\mathrm{d}\Lambda/\mathrm{d}t$ does not vanish. First, we express D as a function of the variables (ϕ, v, Λ) and the Hamiltonian value H:

$$D_{\pm}(\phi, v, \Lambda; H) = (1 + 2c)\cos^2(\phi) \pm 2\cos(\phi)\sqrt{\Delta} = D_{\pm}^0 + \epsilon D_{\pm}^1 + \mathrm{O}(\epsilon^2), \tag{4.18}$$

where the sign in front of the square root is chosen in the different regions so as to guarantee that $|D| < 1$ and that $\mathrm{d}\Lambda/\mathrm{d}t$ does not vanish in the neighborhood of the separatrix, and where

$$\Delta(\phi, v, \Lambda; H) = 2H - v^2 + c(1 + c)\cos^2(\phi) - 2\epsilon A(\phi)\sin(2k\Lambda). \tag{4.19}$$

It follows that

$$\frac{\mathrm{d}\Lambda}{\mathrm{d}t} = \pm\frac{\sqrt{\Delta}}{2\cos\phi} = \frac{\partial H}{\partial D}, \tag{4.20}$$

and, as expected, the zero of $\mathrm{d}\Lambda/\mathrm{d}t$ are identical with those of Δ. Eqs. (4.3) transform to (see Appendix A for the explicit formulae):

$$\frac{\mathrm{d}v}{\mathrm{d}\Lambda} = \frac{\partial D(\phi, v, \Lambda; H)}{\partial\phi}, \qquad \frac{\mathrm{d}\phi}{\mathrm{d}\Lambda} = -\frac{\partial D(\phi, v, \Lambda; H)}{\partial v}. \tag{4.21}$$

These equations describe the motion on one of the energy surfaces described by Figs. 5 or 6 with the Hamiltonian value H. The transformation is valid for orbits on which Δ is bounded away from zero, namely for orbits interior to the one-color regime of the energy surfaces in Figs. 5 and 6.

The Melnikov integral is calculated by expanding $\sqrt{\Delta}$ in ϵ and integrating the Poisson bracket $\{D_{\pm}^0, D_{\pm}^1\}$ along the unperturbed homoclinic orbit (see Appendix A for details), to obtain

$$M_{\pm}(\Lambda_0, \theta; H) = M(c, k; D_{\pm}(H))\cos(2k\Lambda_0 + \theta). \tag{4.22}$$

We offer the following geometrical interpretation: let C be the invariant circle $\{\phi = v = 0, \Lambda \in [0, \pi), D = D_{\pm}(0, 0, \Lambda; H)\}$. In Fig. 9 the stable (thick solid line) and unstable (thick dashed line) manifolds of this circle are schematically plotted on the energy surface H at a specific Poincaré section $\Lambda = \overline{\Lambda} = 0$ and $\theta = \frac{3}{2}\pi$. The Melnikov function $M_{\pm}(\Lambda_0 - \overline{\Lambda}, \theta, H)$ provides the difference between the value of the angular momentum of the stable and

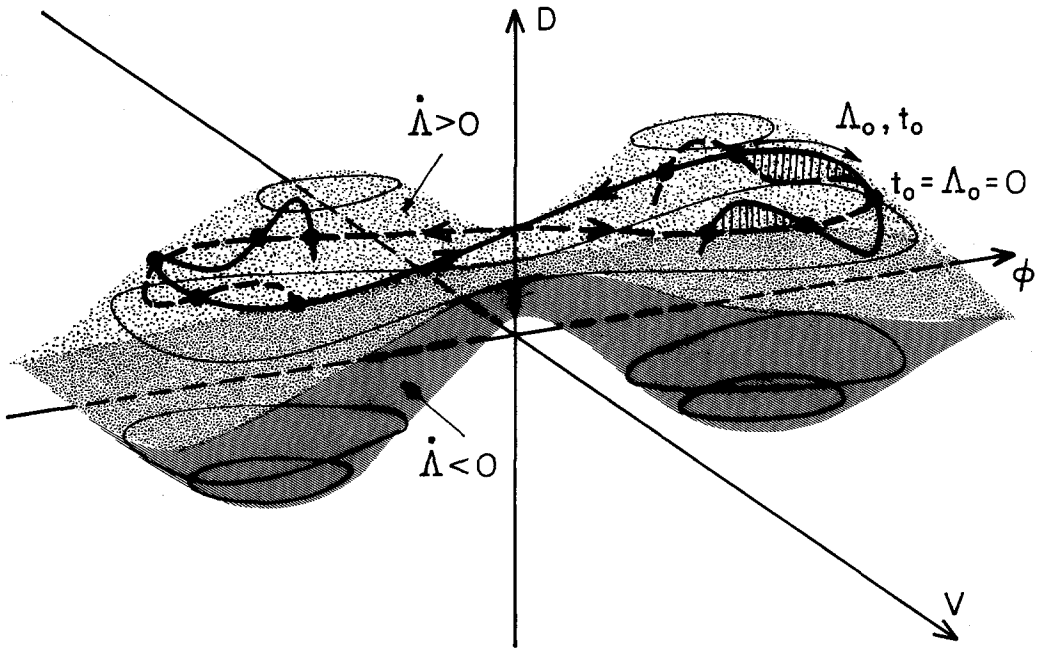

Fig. 9. The homoclinic tangle of the reduced system. Schematic drawing: (_____) stable manifold; (- - -) unstable manifold, energy surface is the grey-scale replica of Fig. 6(c).

unstable manifolds of C near the point $Q_h(-t_0, \Lambda_0) = (\phi, v, D, \Lambda)(-t_0, \Lambda_0)$: when $M > 0$ the unstable manifold is above (has a larger D value than) the stable manifold near $Q_h(-t_0, \Lambda_0)$ at the Poincaré map $\overline{\Lambda} = \Lambda_h(-t_0, \Lambda_0)$. For $t_0 = 0$, this corresponds to the difference between the angular momentum at the cross-section $\Lambda = \Lambda_0$ along the ϕ axis as shown in Fig. 9 (since $M_{\pm}(0, \frac{1}{2}\pi, H) = 0$). Setting $t_0 \neq 0$, corresponds to moving along the homoclinic loop, where the Melnikov function gives the difference between the angular momentum at the Poincaré-section $\overline{\Lambda} = \Lambda_h(-t_0, \Lambda_0)$ near $(\phi_h(-t_0), v_h(-t_0))$. Since $d\Lambda/dt \neq 0$, $\Lambda_h(-t_0, \Lambda_0)$ may be inverted, so that given a $\overline{\Lambda}$ and t_0 a unique Λ_0 in $[0, \pi)$ may be found so that $\Lambda_h(-t_0, \Lambda_0) = \overline{\Lambda}$. It follows that the initial longitude position, Λ_0, and the parametrization along the homoclinic orbit, t_0, can be interchanged, and setting $t_0 = 0$ is convenient (in the next section we deal with the case for which $d\Lambda/dt$ may vanish along the homoclinic orbit, where the above statements cease to hold).

The partial barriers created by the stable and unstable manifolds divide the phase space to restricted and unrestricted regimes as in the $k = 0$ case. The structure of the energy surface supplies a correspondence between the values of D and the type of dynamics – e.g. in Fig. 9 for the upper separatrix larger D values correspond to restricted motion and smaller D values in the vicinity of the separatrix correspond to unrestricted motion, and the converse occurs for the motion near the lower separatrix.

4.5.2. *Non-monotonic Λ dependence*

The vanishing of $d\Lambda/dt$ implies that the cross-section in Λ is not transverse to the flow. Hence, the return map to a cross-section with $\Lambda = \Lambda_0$ is *not* necessarily topologically equivalent to the return map defined for $\Lambda = \Lambda_1 \neq \Lambda_0$, and the motion on the three-dimensional energy surfaces must be analyzed.

It follows from (4.3) and (2.2) that for all ϵ the plane $v = \phi = 0$ is invariant and the dynamics on it is governed by a one d.o.f. Hamiltonian:

$$H(0,0,\Lambda,D)=\tfrac{1}{8}(D-(1+2c))^2+\epsilon\sin(2k\Lambda+\theta)-\tfrac{1}{2}c(c+1), \tag{4.23}$$

i.e. a pendulum with its elliptic fixed point at $D_r=1+2c$, $\Lambda=(\pi-2\theta)/4k$ and a stability zone near its elliptic point which extends up to $D=D_r\pm\sqrt{8\epsilon}$. The behavior for $D=D_r+O(\sqrt{\epsilon})$ is analyzed in Section 4.6. Here we analyze the behavior of the two-dimensional stable and unstable manifolds of the circles $(\phi,v,D,\Lambda)=\{\phi=v=0,\ D=\overline{D}+O(\epsilon),\ \Lambda\in[0,\pi/k)\}$ for $\overline{D}$ values in $(-1,1)$ bounded away from this zonally oscillatory region. Let $(d_{\phi,v}(t_0,\Lambda_0,\overline{D},\theta), d_D(t_0,\Lambda_0,\overline{D},\theta))$ denote the projections of the distance between these manifolds on the (ϕ,v) plane (in the direction perpendicular to the unperturbed separatrix) and on the D axis, respectively, near $Q_h(-t_0,\Lambda_0,\overline{D})$. Then [14,24]

$$d_{\phi,v}(t_0,\Lambda_0,\overline{D},\theta)=\frac{M_1^{\overline{D}}(t_0,\Lambda_0,\theta)}{\|(\mathrm{d}v/\mathrm{d}t,\ \mathrm{d}\phi/\mathrm{d}t)|_{(q_h(-t_0))}\|}, \tag{4.24a}$$

$$M_1^{\overline{D}}(t_0,\Lambda_0,\theta)=\int_{-\infty}^{\infty} vA'(\phi)\cos(2k\Lambda+\theta)\,\mathrm{d}t + \frac{k}{2}\int_{-\infty}^{\infty}\frac{v\overline{D}\sin(2\phi)}{\cos^4\phi}\left\{\int_{-\infty}^{t}A(\phi)\sin(2k\Lambda+\theta)\,\mathrm{d}s\right\}|_{Q_h(t-t_0;\Lambda_0)}\,\mathrm{d}t, \tag{4.24b}$$

and

$$d_D(t_0,\Lambda_0,\theta,\overline{D};c,k)=-2k\cos(2k\Lambda_0+\theta)\lim_{n\to\infty}\int_{-T_n^u}^{T_n^s}A(\phi)\cos 2k\Lambda|_{Q_h(t;\Lambda_0=0)}\,\mathrm{d}t, \tag{4.24c}$$

where the homoclinic orbit $Q_h(t;\Lambda_0)$ is given by (see Appendix A):

$$Q_h(t;\Lambda_0)=(\phi_h(t),v_h(t),\Lambda_h(t,\Lambda_0,\overline{D}),\overline{D}),$$

$$\Lambda_h(t,\Lambda_0,\overline{D})=\Lambda_0+\tfrac{1}{2}(\tfrac{1}{2}\overline{D}-\tfrac{1}{2}-c)t+\tfrac{1}{2}\mathrm{sign}\,(\overline{D})\arctan\left(\frac{\sqrt{1-\overline{D}^2}}{|\overline{D}|}\tanh\frac{t\sqrt{1-\overline{D}^2}}{2}\right). \tag{4.25}$$

The limits of integration in (4.24c), T^{u_n}, T^{s_n} are chosen so as to insure convergence of d_D as this integral is only conditionally converging. Requiring $2k\Lambda=2n\pi$ we obtain

$$T_n^u=T_n^s=\left|\frac{2n\pi}{k(\frac{1}{2}\overline{D}-c-\frac{1}{2})}-\frac{\mathrm{sign}\,(\overline{D})\phi_{\mathrm{hmax}}}{|\frac{1}{2}\overline{D}-c-\frac{1}{2}|}\right|.$$

Since the flow is a two d.o.f. Hamiltonian, it is sufficient to compute the zero of d_D to obtain the homoclinic orbits [14]. Calculating this integral we find

$$d_D(t_0,\Lambda_0,\theta,\overline{D};c,k)=M_D(\overline{D};c,k)\cos(2k\Lambda_0+\theta), \tag{4.26}$$

where $M_D(\overline{D},c,k)$ is presented in Fig. 10 for $k=1,2,-1$. We note that $M_D(\overline{D},c,k)$ is discontinuous at $\overline{D}=0$: $\lim_{\overline{D}\to0+}M_D(\overline{D},c,k)=-\lim_{\overline{D}\to0-}M_D(\overline{D},c,k)$. This discontinuity is the result of the shift $\Lambda(t,D)$ experiences when $D\to-D$ (see Fig. 8). The maxima of M_D for a fixed c value and $k=1$ seems to occur at $D_c=1/(1+2c)$, precisely at the D values for which Λ_h becomes not strictly monotonic. This may be related to the form of the perturbation and will be investigated in more details in the future. Also, notice that for $k=1,2$, as $c\to0$ and $D\to1$ the value of $|M_D(\overline{D},c,k)|$ increases. Preliminary numerical experimentation for $c=0$ and $D\to1$ indicate

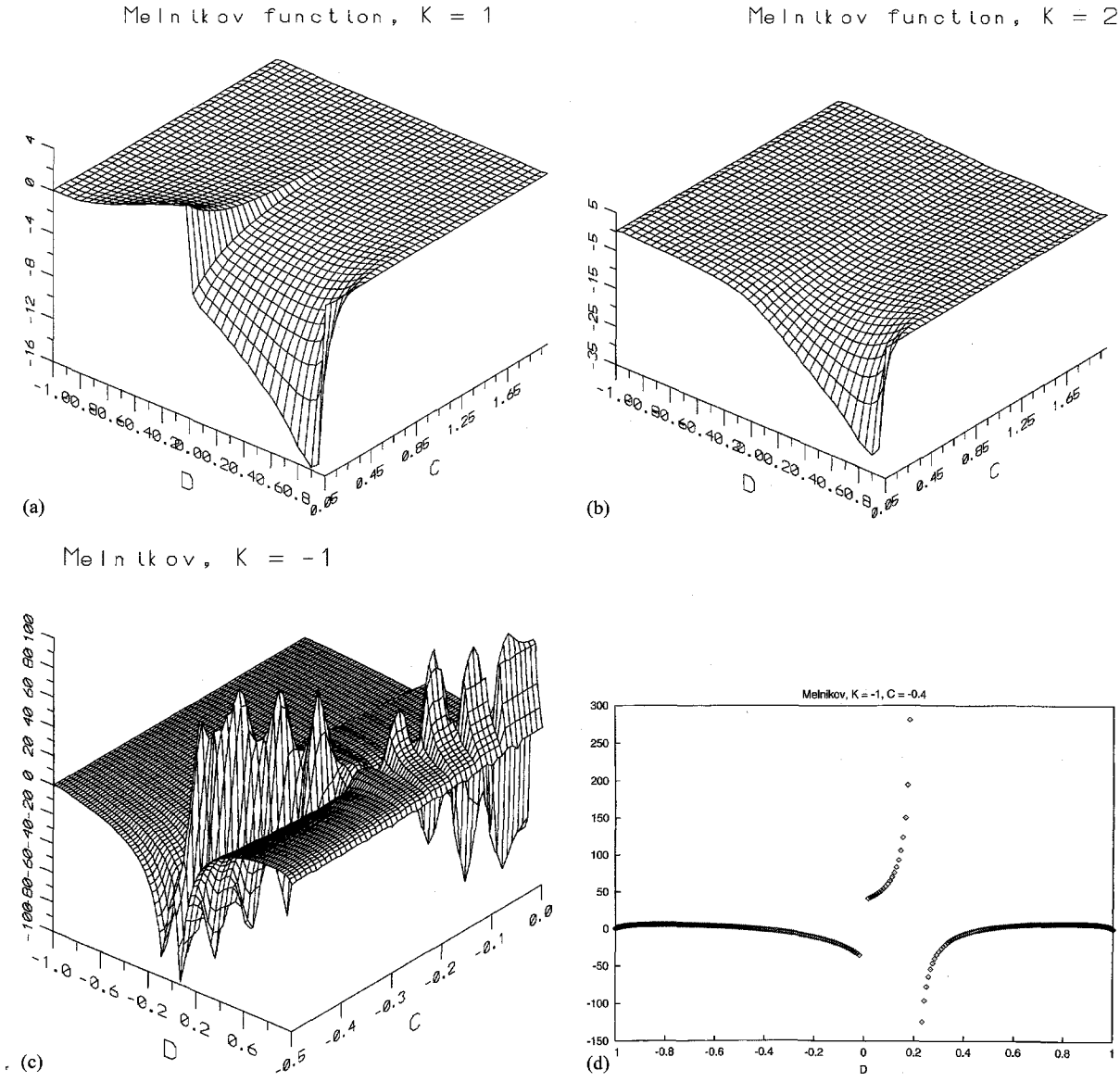

Fig. 10. Melnikov integral amplitude for $k = 1, -1, 2$.

large scale instabilities which are not observed near $D \to -1$, in accordance with the above results. However, the analysis presented here is not valid in the limit $|D| \to 1$ for which the origin in ϕ, v is not hyperbolic, nor in the limit $D \to 1 + 2c$ for which $\mathrm{d}\Lambda/\mathrm{d}t$ vanishes near the origin (see Section 4.6 for this case with $c \neq 0$). This case will be the subject of future work [2]. For $k = -1$ (hence $-\frac{1}{2} < c < 0$) the form of M_D is clearly different. First, it diverges as $D \to D_r = 1 + 2c$, where the circle $\phi = v = 0$ becomes a circle of fixed points (see Section 4.6). Moreover, additional maxima appear for both $D > 0$ and $D < 0$ cases.

From (4.26), it follows that d_D vanishes, as in the monotonic Λ case, at the $2k$ values $\Lambda_i = (\pi(-\frac{1}{2}+i)-\theta)/2k \in [0, \pi)$, $i = 1, \ldots, 2k$. However, the geometrical interpretation is quite different. To understand the geometry of the

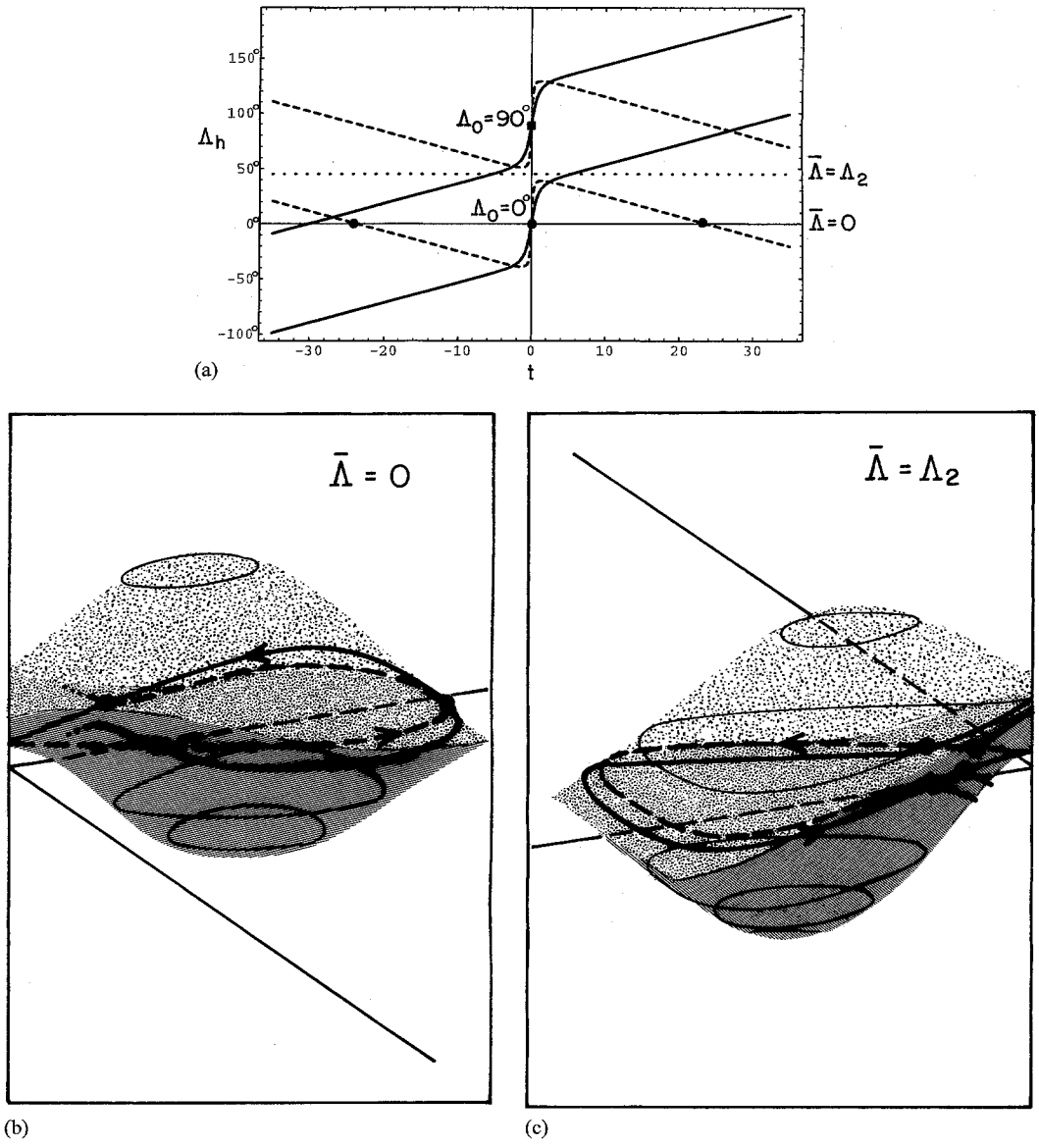

Fig. 11. Geometry of the non-monotonic Λ tangle: (a) (———) $\Lambda_h(t, \Lambda_0, D)$ for the monotonic case, near $D_+(H = 0.122)$ see also Fig. 9, and (- - -) for the non-monotonic case, near $D_-(H = 0.122)$; (b) schematic drawing of the non-uniform homoclinic tangle for D_- at two different cross-sections (since symmetric at these cross-sections, only half the surface is plotted).

intersection, consider the lift of Λ to the real line (i.e. we do not take Λ to be mod π for now, see Fig. 11). It follows from (4.25) that the zero Λ_i produce homoclinic points near $(\phi_h(-t_0), v_h(-t_0))$ at the cross-section $\Lambda_h(-t_0, \Lambda_i)$. In the monotonic case, since $d\Lambda/dt \neq 0$, t_0 and Λ_h are interchangeable and we can take $t_0 = 0$. Here, when $0 < \overline{D} < 1 + 2c$, several different t_0 values can produce the same Λ_h as is shown in Fig. 11. In particular, fixed i and let $-t_m(\overline{D}, c)$ satisfy

$$\frac{d\Lambda_h(t, \Lambda_i, D)}{dt}\Big|_{-t_m(\overline{D},c)} = 0 \tag{4.27}$$

(it follows from (4.12) that $\phi_h(\pm t_m(\overline{D}, c)) = \phi_r$) and let $\bar{t}_m$ satisfy

$$\Lambda_h(\bar{t}_m(\overline{D}, c), \Lambda_i, D) = \Lambda_h(-t_m(\overline{D}, c), \Lambda_i, D) = \Lambda_i - \Lambda_m(\overline{D}, c). \tag{4.28}$$

Then $\Lambda_h(-t_0, \Lambda_i, D) = \Lambda_i + \delta$ has three solutions $t_j \in (-\bar{t}_m, \bar{t}_m)$ for all $|\delta| < \Lambda_m \leq \frac{1}{4}\pi$, and only one solution if $|\delta| > \Lambda_m$. Namely, for cross-sections "close" to Λ_0 the perturbed homoclinic orbit associated with Λ_i crosses the cross-section three times, once with $\phi \in (\phi_r, \phi_{hmax})$ (see (4.12)) whereas the cross-sections with $|\Lambda - \Lambda_0| > \Lambda_m$ are intersected by the homoclinic orbit only once, as in the monotonic Λ case. If the distance between the Λ_i's $(\pi/2k)$, is larger than $2\Lambda_m(\overline{D}, c)$ then for $\Lambda \in (\Lambda_i + \Lambda_m, \Lambda_{i+1} - \Lambda_m)$ there are no homoclinic points for $\phi \in [\phi_r, \phi_{hmax}]$. More generally, the number of additional homoclinic points depends on the ratio between $\Lambda_m(\overline{D}, c)$ and $\pi/4k$. The implications of these phenomena on transport and statistics of the chaotic solutions is yet to be explored. For example, when $\Lambda_m < \pi/4k$, we expect that the probability to find solutions in the chaotic zone near ϕ_{hmax} will be enhanced for $\Lambda \in (\Lambda_i - \Lambda_m, \Lambda_i + \Lambda_m)$.

Finally, note that $(d_D, d_{\phi,v})$ may be calculated for the monotonic Λ case as well. In fact, though it is not obvious from the formulae, one can check that $M(c, k; D_\pm(H)) = d_D(c, k, \overline{D})$ for $\overline{D} = D_\pm(H)$. The reduced equations were included to assure that the results we obtain are independent of the method and to emphasize the geometrical difference between these cases.

4.5.3. *Implications of the homoclinic chaos*

The homoclinic chaos with finite wavelength perturbation occurs in a three-dimensional region in phase space – a 3D strip of the 3D energy surfaces. This region is composed of a strip of the 2D surfaces of Figs. 5 or 6 crossed with the circle $\Lambda \in [0, \pi)$. Hence the homoclinic tangle is extended in both (ϕ, v) and D directions and, as discussed in Section 4.5, is non-uniform in Λ, see Fig. 11.

The calculation of the Melnikov vector (4.24) supplies an estimate to the distance in (ϕ, v) and in D direction between the stable and unstable manifolds of the hyperbolic circle belonging to the the torus $\phi = v = 0$ (this circle is represented by the fixed point in Fig. 11). The geometrical implications of the above statement lead to the prediction of the type of non-uniformity in Λ of the homoclinic tangle. The influence of these predictions on the extent of the chaotic zone as a function of Λ, i.e. the equivalent graph for $k \neq 0$ of Fig. 4, is under current investigation.

In Fig. 10 we present the calculation of $|M_D|$ for $k = 1, 2, -1$. The dynamical consequences of these results on the extent and location of the stability zone are deferred to future publication. We note that for $k = 1$, and fixed c, M_D seems to be maximal for $D = D_c = 1/(1 + 2c)$, however this property is not true for $k = 2$. When $k = -1$ (and hence $c < 0$), $M_D \to \infty$ for $D \to D_r = 1 + 2c$. This occurs because the regular Melnikov analysis is not valid near D_r as described in Section 4.6.

4.6. *Hyperbolic–elliptic resonance*

When $-\frac{1}{2} < c < 0$, for $D = D_r = 1 + 2c$, there is a particularly degenerate situation in which the circle $(\phi, v, D, \Lambda) = (0, 0, D_r, \Lambda)$ is composed of a continuum of fixed points, hyperbolic in the (ϕ, v) direction and neutral in the (Λ, D) direction. Kovacic and Wiggins [18], Haller and Wiggins [8] and Kaper and Kovacic [20] have analyzed similar situations, finding criteria for the existence of transverse homoclinic and heteroclinic orbits to various solutions on the invariant cylinder. Here, we follow the analysis of Haller and Wiggins [8], including only the essential results and their geometrical interpretations and give some insight into the influence of these results on the dispersion of particles in the vicinity of the cylinder.

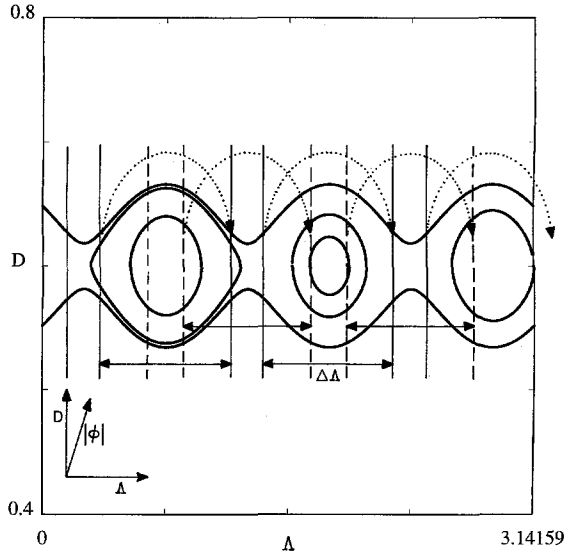

Fig. 12. The resonance band: $c = -0.2$, $k = -3$, $\epsilon = 0.00025$, $D \approx D_r = 0.6$. Numerical trajectories are shown at the $\phi = v = 0$ plane. The lines $z^{\pm}_{\alpha,j}$ and the heteroclinic/homoclinic orbits associated with them are indicated schematically.

For all that follows, assume with no loss of generality that $\theta = 0$, and let $-\frac{1}{2} < c < 0$. It follows from (4.3) and (2.2) that the cylinder $(0, 0, \Lambda, D)$ is preserved for any ϵ and that the motion on it is described by

$$\frac{d\Lambda}{dt} = \tfrac{1}{4}(D - (1 + 2c)), \qquad \frac{dD}{dt} = -2k\epsilon \cos(2k\Lambda). \tag{4.29}$$

Hence, for $\epsilon = 0$ the motion on this cylinder consists of simple rotation in Λ along circles with constant D values, and the frequency of the rotation vanishes for $D = D_r$ (notice that if $c = 0, -\frac{1}{2}$ then $D_r = 1, 0$ and more degenerate situations, which are yet to be analyzed, emerge). In the (ϕ, v) directions these circles are hyperbolic, and for $\epsilon = 0$ the invariant circles possess two families of homoclinic orbits $Q^{\pm}_h(t; \Lambda_0, D)$, rotating in Λ with the invariant circle frequency as $t \to \pm\infty$, see (4.25). For $D = D_r$ these orbits connect fixed points with different Λ values (hence are termed heteroclinic orbits), see Fig. 12. Using (4.25) we find that the heteroclinic orbits connect the fixed points $(0, 0, \Lambda_0, D_r)$, $(0, 0, \Lambda_0 + \Delta\Lambda, D_r)$ where

$$\Delta\Lambda = \Lambda_h(\infty; t_0, \Lambda_0, D_r) - \Lambda_h(-\infty; t_0, \Lambda_0, D_r) = \phi_{hmax}(D_r) = \arccos(1 + 2c), \tag{4.30}$$

for all $\Lambda_0 \in [0, \pi)$. In particular,

$$0 < \Delta\Lambda < \tfrac{1}{2}\pi. \tag{4.31}$$

Now, examine what happens to this structure when $\epsilon \neq 0$. Using the blow-up transformation of Kovacic and Wiggins [18], write

$$D = D_r + \sqrt{\epsilon}\eta, \qquad \tau = \sqrt{\epsilon}t$$

then (4.29) becomes

$$\frac{d\Lambda}{d\tau} = \frac{1}{4}\eta = \frac{\partial h}{\partial \eta}, \quad \frac{d\eta}{d\tau} = -2k\cos(2k\Lambda) = -\frac{\partial h}{\partial \Lambda}, \qquad h = \frac{1}{8}\eta^2 + \sin(2k\Lambda). \tag{4.32}$$

Namely, when $\epsilon \neq 0$ the circle of fixed points resonances with the perturbation oscillations, creating a chain of k pendulum-like islands. There are k elliptic fixed points at $\Lambda_j = (\pi/k)(\frac{3}{4} + j)$, $j = 0, \ldots, k-1$ separated by k hyperbolic points at $\Lambda_j - \pi/2k$ which are connected by two heteroclinic orbits, see Fig. 12. Half the width (maximum of $D - D_r$) of these stability islands is $\sqrt{8\epsilon}$ (i.e. $\sqrt{8}$ in the units of η).

Since $I(k) = \pi/|k|$ is the length of each island, and since $\Delta\Lambda = \phi_{\text{hmax}}$, it follows that the behavior near the islands (in the full phase space) depends on the ratio $\Delta\Lambda/I(k)$: if $\Delta\Lambda < I(k)$, we expect a recurrent motion associated with homoclinic chaos near each island. If $\Delta\Lambda > I(k)$ then we expect to observe heteroclinic motion in which orbits hop between oscillatory motion in different zonal bands.

To establish the above observations and to find the dependence on c of the number of one-hump homoclinic and heteroclinic connections between different resonance zones, we use the energy-phase method with $n = 1$. Let

$$\Delta H = H_1(0, 0, D_r, \Lambda - \Delta\Lambda) - H_1(0, 0, D_r, \Lambda)$$

then, if ΔH has simple zeros in Λ, there are two transverse heteroclinic orbits connecting orbits which are $O(\sqrt{\epsilon})$ close to the base-point (in the (ϕ, v) direction) on the resonance band which have the same η value. The existence of multi-hump homoclinic and heteroclinic orbits follows immediately from the form of the energy-phase criterion, see [20,36]. In fact, Haller [36] shows that a universal bifurcation to a multi-hump tree exists as c is varied. In our case:

$$\Delta H(\Lambda) = 2\cos k(2\Lambda - \Delta\Lambda)\sin k\Delta\Lambda.$$

Hence if

$$\Delta\Lambda \neq \frac{j\pi}{k}, \quad j \in \mathbb{Z}, \tag{4.33}$$

then the equation $\Delta H(\Lambda) = 0$ has two families of simple zero in Λ:

$$\Lambda_{1,j} = \frac{1}{2}\Delta\Lambda + \frac{\pi}{4k} + \frac{j\pi}{k}, \quad j = 0, \ldots, k-1,$$
$$\Lambda_{2,j} = \frac{1}{2}\Delta\Lambda + \frac{3\pi}{4k} + \frac{j\pi}{k}, \quad j = 0, \ldots, k-1.$$

Define the lines:

$$z^+_{1,j} = \{(\eta, \Lambda) | \Lambda = \Lambda_{1,j}\}, \qquad z^+_{2,j} = \{(\eta, \Lambda) | \Lambda = \Lambda_{2,j}\},$$
$$z^-_{1,j} = \{(\eta, \Lambda) | \Lambda = \Lambda_{1,j} - \Delta\Lambda\}, \qquad z^-_{2,j} = \{(\eta, \Lambda) | \Lambda = \Lambda_{2,j} - \Delta\Lambda\}.$$

In Fig. 12 the vertical dashed (respectively solid) lines correspond to $z^\pm_{1,j}$ (resp. $z^\pm_{2,j}$). If a periodic orbit of a reduced system in the resonant region intersects transversely the line $z^-_{\alpha,j}$ and the line $z^+_{\alpha,j}$, then there exists a transverse homoclinic orbit to a nearby ($\sqrt{\epsilon}$-close) periodic orbit on the resonant cylinder, shown as dotted arcs on the figure. If $z^-_{\alpha,j}$ and $z^+_{\alpha,j}$ intersect two different periodic orbits of the reduced system at $\eta = \overline{\eta}$ then there exist two heteroclinic orbits each connecting a different pair of periodic orbits on the resonant cylinder.

Since $I(k) = \pi/|k|$ is the length of each island, it follows from (4.31) and (4.33) that there are three different cases:

(1) $|k| = 1$: Then, there is only one island, $I(k) \geq 2\Delta\Lambda$, and the periodic orbits in the island can be divided to three groups – for the internal ones there are no simple homoclinic loops, for the nearest ring there are four homoclinic loops and for the rest – the eye-shaped region – there are eight homoclinic loops. For the limiting value $c \to -\frac{1}{2}$ the middle ring shrinks to zero, whereas when $c \to 0$ the middle ring occupies most of the island.

(2) $|k| = 2$: Then there are two islands with $I(k) \geq \Delta\Lambda$. Each of these islands is divided by the lines $z_1^{\pm}$, $z_2^{\pm}$ which are centered around the hyperbolic and elliptic fixed points respectively. For c values such that $\Delta\Lambda < \frac{1}{2}I(k)$:

$$c_b^-(0) < c < 0, \qquad c_b^-(0) = \tfrac{1}{2}(\cos\tfrac{1}{4}\pi - 1) = -0.14645$$

the islands are divided by these lines to three regions – internal region in which no simple homoclinic or heteroclinic orbits occur, a ring in which four homoclinic loops occur, and the rest of the island in which, four homoclinic and four heteroclinic orbits occur. When $c = c_b^-(0)$, $\Delta\Lambda = \frac{1}{2}I(k)$ and the lines $z_{1,j}^+$ (resp. $z_{1,j}^-$) and $z_{2,j}^-$ (resp. $z_{2,j-1}^+$) coincide causing the middle region to shrink to zero. For $c < c_b^-(0)$ the lines switch order and the middle ring attains only simple heteroclinic orbits.

(3) $k \geq 3$: The sequence

$$c_k(j) = \frac{1}{2}\left(\cos\frac{j\pi}{k} - 1\right) = -\sin^2\frac{j\pi}{2k}, \quad j = 0, \ldots, \left[\frac{k}{2}\right], \tag{4.34}$$

is a sequence of bifurcation values: for $\delta > 0$ and $\epsilon(\delta)$ sufficiently small, if $c \in (c_k(j) + \delta, c_k(j-1) - \delta)$ then there exist one-hump heteroclinic orbits between island i and the islands $i + j - 1 (\mathrm{mod}\ k)$ and $i + j (\mathrm{mod}\ k)$, for all $0 \leq i < k$.

To justify the above claim, notice that if $c \in (c_k(j), c_k(j-1))$, then $(j-1)I(k) < \Delta\Lambda < jI(k)$ (see (4.30)). Hence, the lines $z_{1,i}^{\pm}$ intersect the islands $i \pm [\frac{1}{2}j]$, respectively, and the lines $z_{2,i}^{\pm}$ intersect the islands $i + [\frac{1}{2}(j-1)] + \frac{1}{2}(1 \pm 1)$, respectively. The conclusion that island i' and $i' + j, i' + j - 1$ follows when considering even and odd j's. In particular, only for $c_k(1) < c < 0$ there exist one-hump homoclinic solutions to oscillatory orbits in Λ.

At $c = c_k(j)$, by definition $\Delta\Lambda = j\pi/k$ and condition (4.33) is violated. Geometrically it implies that to a leading order in ϵ the unstable manifolds of the periodic orbits in resonance i coincide with the stable manifold of the periodic (or heteroclinic) orbits of equal energies (equal η) in resonance $i + j$. Higher-order corrections are expected, generically, to destroy these heteroclinic connections.

In addition, there is another series of bifurcation values $c_b^+(j)$, $c_b^-(j)$ for which the lines $z_{1,i}^+$ (resp. $z_{2,i}^+$) and $z_{2,i+j}^-$ ($z_{1,i+j}^-$) switch order, where

$$c_b^{\pm}(j) = \frac{1}{2}\left(\cos\frac{(2j \pm 1)\pi}{2k} - 1\right) = -\sin^2\left(\frac{j\pi}{2k} \pm \frac{\pi}{4k}\right), \quad j \geq 0.$$

It follows that

$$c_k(j+1) < c_b^-(j+1) = c_b^+(j) < c_k(j), \tag{4.35}$$

and the following scenario occurs as c is decreased from 0: for $c \in (c_b^-(1), 0)$ in each island there is an internal region with no one-hump homoclinic or heteroclinic solutions, encircled by a ring with four homoclinic solutions and the rest of the island (the eye-shaped region) has four homoclinic solutions and four heteroclinic solutions: two incoming solutions from cell $i - 1$ and two outgoing solutions to cell $i + 1$. As c is further decreased so that $c \in (c_k(1), c_b^-(1))$, the ring contains four heteroclinic solutions as described above and the eye-shaped region contains, as before, four homoclinic and four heteroclinic orbits. When $c \in (c_b^+(1), c_k(1))$ there are no homoclinic orbits to oscillatory motion in Λ. The ring in resonance i has four heteroclinic orbits: two heteroclinic orbits which asymptote (resp. are asymptoted by) orbits in resonance $i + 1$ (resp. $i - 1$), and the eye-shaped region has, in addition to these, two heteroclinic orbits to orbits in resonance $i + 2$ and two incoming heteroclinic orbits from resonance $i - 2$. When $c \in (c_k(2), c_b^+(1))$ the ring in resonance i has four heteroclinic orbits to/from orbits in resonance $i \pm 2$. The eye shaped region has, in addition, four heteroclinic

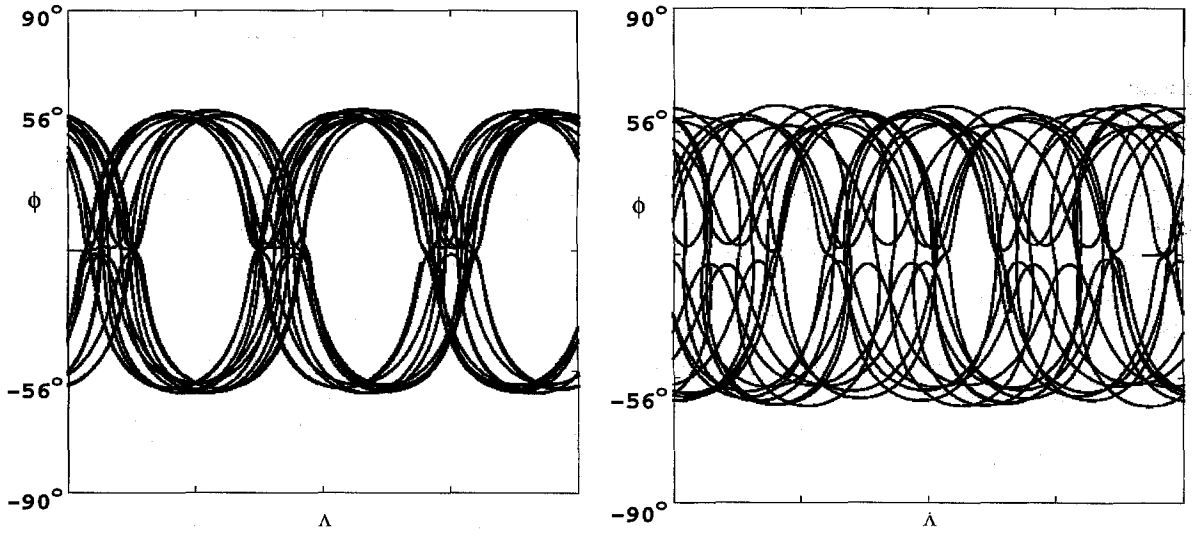

Fig. 13. The hyperbolic-resonance motion on the rotating sphere. $c = -0.2$, $k = -3$, $\epsilon = 0.00025$, $D = 0.62$, $\phi_0 = 0.00001$, $v_0 = 0$, $D_0 = 0.62$ and (a) $\Lambda_0 = 0$, (b) $\Lambda_0 = 2.9$.

orbits to/from resonance $i \pm 1$. The same scenario is repeated when $j \to j + 1$, with heteroclinic orbits to resonance $i + j$ and $i + j - 1$.

To summarize, an initial condition placed exactly on the equator with no northward velocity, and with westward velocity close to that of the pressure wave moves in a longitude dependent (non-chaotic) fashion: depending on its *longitude*, it may either oscillate in one of the k valleys of the pressure wave (i.e. become "captured" by the wave) or miss the wave altogether having a different averaged westward velocity than the wave. However, a tiny northward velocity or, a tiny displacement from the equator will lead to meridional bursts along the homoclinic loops, extending up to $\phi = \arccos(1 + 2c) + O(\sqrt{\epsilon})$. The new feature here is that the zonal distribution of these bursts is not necessarily uniform, and it depends on all initial conditions, and in particular, it depends on the initial longitude position as is demonstrated in Fig. 13. Moreover, the analysis shows that depending on the values of c and k there may be a direct or indirect connection via meridional bursts between neighboring zonal pressure-wave-troughs. The distinction between the physical and mathematical period of Λ arises naturally in this setting – taken mode π/k the above phenomena are not apparent.

5. Conclusions and discussion

The motion of particles moving on a rotating spherical surface under the influence of Coriolis force and a conservative propagating perturbation exhibits, under some conditions, homoclinic chaos. The chaotic motion occurs, for example, when particles are sent with westward velocity ($u_0 = -|u|$) from the vicinity of the equator. Such particles rotate close to the equator for a long time and then burst to latitudes of order $\pm \arccos(1-2|u|)$, return close to the equator, and burst again, with unpredictability with regards to the times and orientation (north–south) of the bursts. The form of the distribution of these bursts with respect to the longitude position and the size of this chaotic zone depend on the precise values of the parameters.

The richness of the problem in hand (with only three standard parameters – amplitude of perturbation, wavelength and wave speed) unfolds using the energy surface classification scheme. For finite wavelength perturbations, several types of homoclinic chaos may appear:

(i) Uniform homoclinic chaos – the chaotic zone is topologically uniform in the longitude angle; the return maps to two different longitude positions are topologically conjugate. While there exist small variations in λ of the exact location of the chaotic zone, the overall behavior is uniform.

(ii) Homoclinic chaos with backflow – the chaotic zone is non-uniform in the longitude angle for O(1) latitudes (more specifically for $\phi = O(\phi_r(u_0, \phi_0))$, see (4.12) and (2.3a)). Then for some longitude regions more visits of chaotic orbits in the high latitudes regimes will be recorded.

(iii) Homoclinic chaos to a resonance band – the homoclinic chaos is non-uniform in the longitude angle for all latitudes, even near the equator. Moreover, the uncertainty in the value of the westward velocity at the equator is of $O(\sqrt{\epsilon})$ whereas for the other cases it is of $O(\epsilon)$.

We believe that the probability distribution of finding particles at a certain latitude at a given time T (resp. averaged on time) will have distinguished features in each of these cases. In the first case it will be approximately (asymptotically) independent of longitude (resp. of Λ – the longitude position w.r.t. the travelling wave), in the second case it will depend on the longitude (resp. on Λ) only for O(1) latitude values and in the third case it will depend on the longitude (resp. on Λ) for all ϕ's in the chaotic zone. Evidence that such behavior occurs for the third case is given in Fig. 13.

To distinguish between these dynamically different scenarios the concept of colored energy surfaces (Figs. 5, 6), and of the corresponding shaded energy–momentum map (Fig. 7) has been developed. The shaded energy–momentum map encodes the features of the colored energy surfaces. The shaded region in the energy–momentum map corresponds to regions in which backflow occurs. When the shaded region intersects the dashed line homoclinic chaos with backflow occurs. When this intersection starts at a dashed fold (Figs. 7(b) and (c)), homoclinic chaos to elliptic resonance occurs.

In our model, the third scenario of elliptic–hyperbolic resonance occurs only for westward propagating pressure waves ($k, c < 0$), hence we conclude that a change in the direction of the propagation of the pressure wave changes qualitatively and quantitatively the dynamics of near-equator particles with $u_0 \approx -|c|$. The wave number k determines the periodicity interval in Λ: k oscillatory resonance islands appear near the red curves of the energy surfaces for $\epsilon > 0$. Moreover, when $c < 0$, there exists a sequence of k bifurcation values of c associated with homoclinic/heteroclinic tangencies. In particular, for $|k| \geq 3$ there are ranges of c values for which no simple homoclinic orbit exists. Melnikov analysis for the eastward propagating waves shows that low-speed waves increase the size of the chaotic zone for westward going particles with small initial velocities. Moreover, it implies that for a given pressure wave ((k, c)) there exists a velocity for which this chaotic zone is maximal. Standing waves ($c = 0$) lead to strong resonances near the inertial frequencies and to the appearance of parabolic resonances – this case is under current investigation [2,35]. When $c = -\frac{1}{2}$ a hyperbolic resonance with homoclinic orbits reaching the poles exists.

Infinite wavelength perturbations ($k = 0$) result in chaotic motion which is uniform in the longitude angle and which conserves the angular momentum, leading to a two-and-a-half d.o.f. system with a conserved quantity (D). In such systems, physical ensembles of chaotic particles behave differently than a single initial condition since D is not a constant in such ensembles. These considerations lead to non-intuitive results regarding the location and size of the chaotic zone (Fig. 3), e.g. this calculation shows that the size and location of the chaotic zone in (λ, ϕ) of eastward going particles and westward going particles is completely different and that the former are more sensitive to the perturbation frequency than the latter.

Many interesting mathematical and physical questions are left for future works. The mathematical nature of homoclinic chaos with a backflow is not yet understood – we think that for small and finite ϵ values it exhibits topologically different global structure than that obtained from monotonic homoclinic chaos but this is yet to be proven. Recently, we have started to study the impact of parabolic resonances in the general two d.o.f. systems context [2] and in the atmospherical context in particular [35]. The motion near the $D = 0$ surfaces and near critical

wave speed is yet to be studied (see Section 4.3.2), as are the direct implications to atmospherical observations for dispersal of particles in the vicinity of the equator (see discussion below).

One may be tempted to use the above model for studying the motion of the air/water parcels themselves. In fact, the model studied here cannot be directly applied to the fluid (geophysical) flow of air (in the atmosphere) or water (in the ocean) because the continuity equation is not guaranteed to be satisfied by our two-dimensional (in planetary coordinates) model. One simple way to overcome this difficulty is to specify the pressure forcing so as to compensate, through vertical motions and via the hydrostatic balance, for the divergence/convergence in the horizontal flow. Indeed, it has been recently demonstrated numerically (in a different atmospheric model) that, unlike vertical migration, the *horizontal dispersal* is not significantly affected by the inclusion or exclusion of the continuity equation [37].

Another fundamental simplification in our model is that the pressure term can be assumed small. This assumption leads naturally to the paradigm of a perturbed inertial flow in which the inertial flow carries the O(1) dynamics. This inertial force free flow, while certainly being of fundamental importance in any effort to investigate realistic flows on a rotating geopotential, does not, as noted by R. Pierrehumbert, accurately model the average geostrophic flow in the planetary atmosphere and ocean. In order to make the study more relevant to planetary flows which are predominantly in geostrophic balance, the pressure gradient term has to be incorporated in the basic state rather than the perturbed equations only. This can be simply done, without any modification of the analysis itself, by specifying two pressure terms – an O(1) term which is steady and zonally independent and a wave-like perturbation term (solving the unperturbed problem is equivalent to solving an adjustment problem of the velocity field to a given pressure field on a global scale instead of the synoptic one). In terms of the equations presented here, replacing the term $\epsilon A(\phi)\sin(k\lambda - \sigma t + \theta)$ in the Hamiltonian (4.4) by the term $B(\phi) + \epsilon A(\phi)\sin(k\lambda - \sigma t + \theta)$ and specifying $B(\phi)$, $\epsilon A(\phi)$ reverts our quasi-inertial problem to this adjustment process with a perturbing pressure wave. The unperturbed structure is thus slightly altered ($B'(\phi)$ is in fact of O(0.01) for reasonable atmospherical data [35]) and the analysis tools and basic observations stay intact. For $\epsilon = 0$, the zonal geostrophic winds (currents) are given by the steady states of the system. The fact that there exists a Hamiltonian to this system guarantees that those steady states can only be of the types studied here (i.e. no steady or unsteady spirals are possible). In terms of D and u the geostrophic flows are given by

$$D_g = \cos^2(\phi)\sqrt{1 - \frac{8B'(\phi)}{sin(2\phi)}}, \qquad u_g = \frac{1}{2}\cos(\phi)\left(\pm\sqrt{1 - \frac{8B'(\phi)}{sin(2\phi)}} - 1\right).$$

The real adjustment to geostrophy includes the mutual adjustment of both the velocity and the pressure but this is beyond the scope of the problem studied here.

From a dynamical viewpoint the fact that the equilibrium points are either elliptic or hyperbolic implies that these points cannot be reached in finite time by the trajectories. In particular the geostrophic flows are not attracting in this Lagrangian model, unless damping (friction) is introduced. In the frictionless problem, therefore, the geostrophic flow only represents the average flow while the observed flow at any time will inevitably contain the inertial oscillations about this steady state.

Acknowledgements

Discussions with G. Haller, L. Lerman, R. Pierrehumbert, J. Weiss and G.M. Zaslavsky are greatly appreciated. The authors are indebted to Drs. M. Engel and A. Sigalov of the Center for Visualization of Dynamic Systems of the Hebrew University of Jerusalem for their assistance in producing some of the figures contained in this paper. V. Rom-Kedar thanks the hospitality of the Mathematics Department at The University of Chicago where part of

this work has been carried out. NP acknowledges the support provided by the Israel Academy of Sciences via a research grant to the Hebrew University of Jerusalem.

Appendix A. Analytical formulas

A.1. The unperturbed system

It follows from (4.3) that when $\epsilon = 0$ the equation for the λ coordinate decouples, D is conserved, hence may be considered as a parameter, and a two-dimensional system for (ϕ, v) is obtained:

$$\frac{d\phi}{dt} = v, \qquad \frac{dv}{dt} = \frac{1}{8}\sin(2\phi)\left(1 - \frac{D^2}{\cos^4\phi}\right) \tag{A.1}$$

with the Hamiltonian

$$H(\phi, v; D) = \frac{v^2}{2} + \frac{1}{8}\left(\frac{D}{\cos\phi} - \cos\phi\right)^2 = \frac{1}{2}E. \tag{A.2}$$

This system has a fixed point at the origin, with eigenvalues

$$\left(\pm\tfrac{1}{2}\sqrt{1 - D^2}\right).$$

Namely, the origin is unstable (hyperbolic) for $|D| < 1$ and stable (elliptic) for $|D| > 1$. For $|D| < 1$ there exist two additional elliptic fixed points at

$$(\phi_{\text{ell}}, v) = \pm(\arccos\sqrt{|D|}, 0)$$

with eigenvalues

$$(\pm i\sqrt{1 - |D|}).$$

Hence their natural frequency is $2\pi/\sqrt{1 - |D|} = 2\pi/\sin\phi_{\text{ell}}$, which is called the inertial frequency in the atmospheric and oceanic literature.

For $|D| < 1$ there are two homoclinic trajectories to the origin:

$$\begin{aligned} q_{\text{h}}^{\pm}(t; \phi_{\text{hmax}}) &= \pm(\phi_{\text{h}}(t), v_{\text{h}}(t)) \\ &= \pm\left(\arcsin\left(\sin\phi_{\text{hmax}}\,\text{sech}\left(\frac{\sin\phi_{\text{hmax}}\,t}{2}\right)\right), -\frac{1}{2}\frac{\sin^2\phi_{\text{hmax}}\tanh(\sin\phi_{\text{hmax}}\,t/2)}{\sqrt{\cosh^2(\sin\phi_{\text{hmax}}\,t/2) - \sin^2\phi_{\text{hmax}}}}\right), \end{aligned} \tag{A.3}$$

where ϕ_{hmax} is the maximal latitude angle reached by the homoclinic orbit and is related to D by

$$\phi_{\text{hmax}}(D) = |\arccos(|D|)|. \tag{A.4}$$

Integrating (2.1a) along the homoclinic orbit, and using (2.4) and (A.4) in the process we find the longitude homoclinic solution, λ_{h}:

$$\lambda_h^{\pm}(t;D)=\lambda_h(0)+\frac{1}{2}(D-1)t+\operatorname{sign} D\arctan\left(\frac{\sqrt{1-D^2}}{|D|}\tanh\frac{t\sqrt{1-D^2}}{2}\right)$$
$$=\lambda_h(0)+\frac{1}{2}(\operatorname{sign}(D)\cos\phi_{\mathrm{hmax}}-1)\,t+\operatorname{sign}(D)\arctan\left(|\tan(\phi_{\mathrm{hmax}})|\tanh\frac{t|\sin(\phi_{\mathrm{hmax}})|}{2}\right). \tag{A.5}$$

Next we calculate $I(\phi_{\max}, D)$ and $P_\phi(\phi_{\max}, D)$, the action and the period in ϕ of the periodic solutions of (A.1), where $\phi_{\max}$ denotes the maximal latitude reached by the periodic orbit and D denotes the (fixed) angular momentum on the orbit. To find P_ϕ we integrate $\mathrm{d}\phi/v(\phi,\phi_{\max};D)$ where $v(\phi,\phi_{\max};D)=v(\phi,H;D)|_{H=H(\phi_{\max},0;D)}$ is found from (A.2). There are two types of orbits to consider – orbit which encircle the origin, for which the integration in ϕ is performed between 0 and $\phi_{\max}$ and orbits which encircle the elliptic points with $|D|<1$, for which the integration in ϕ is performed between $\phi_{\min}$ and $\phi_{\max}$. Substituting $\cos\phi=x$ and using formula 3.131(4) in Gradshteyn and Ryzhik [38] we find:

$$P_\phi(\phi_{\max},D)=-N\frac{2}{\sqrt{x-z}}K\left(\sqrt{\frac{y-z}{x-z}}\right), \tag{A.6a}$$

where K is the complete elliptic integral of the first kind, and

$$\begin{aligned} &x=1,\quad y=\frac{D^2}{\cos^2(\phi_{\max})},\quad z=\cos^2(\phi_{\max}),\quad N=2 \text{ for restricted orbits;}\\ &x=\frac{D^2}{\cos^2(\phi_{\max})},\quad y=1,\quad z=\cos^2(\phi_{\max}),\quad N=4 \text{ for unrestricted orbits.}\end{aligned} \tag{A.6b}$$

Similarly,

$$I(\phi_{\max},D)=2\int_{\phi_{\min}}^{\phi_{\max}} v\,\mathrm{d}\phi=N\left(\frac{\frac{1}{2}x-4E(\phi_{\max},0,D)-D}{\sqrt{x-z}}K\left(\sqrt{\frac{y-z}{x-z}}\right)\right.$$
$$\left.-\frac{1}{2}\sqrt{x-z}E\left(\sqrt{\frac{y-z}{x-z}}\right)-\frac{2D^2}{z\sqrt{x-z}}\Pi\left(\frac{1}{2}\pi,\frac{z-y}{z},\sqrt{\frac{y-z}{x-z}}\right)\right), \tag{A.7}$$

where $E(\phi_{\max},0,D)$ is defined by (2.3) and K,Π,E are elliptic integrals of the first, second and third kind, respectively.

Let $\Lambda(t;\phi_{\max},D)=\frac{1}{2}(\lambda(t;\phi_{\max},D)-ct)$, where $c=\sigma/k$ (this variable is used when $k\neq 0$, see (4.2) and (4.1)). It is of interest to find resonance behavior between the ϕ and Λ variables. As it is quite complicated to find the period in Λ, we find instead $\Lambda_P(\phi_{\max},D)$, the increment in Λ after one period in ϕ:

$$\Lambda_P(\phi_{\max},D)=\Lambda(P_\phi(\phi_{\max},D);\phi_{\max},D)$$
$$=-\frac{1}{2}\left(c+\frac{1}{2}\right)P_\phi(\phi_{\max},D)+\frac{1}{2}DN\frac{\Pi\left(-(y-z/z),\sqrt{(y-z)/(x-z)}\right)}{z\sqrt{x-z}}, \tag{A.8}$$

where x, y, z and N depend on the orbits nature as in (A.6). Since $\phi_{\max}=0$ when $(\phi,v)=(0,0)$ and $|D|>1$ and $\phi_{\max}=\phi_{\mathrm{ell}}$ on the elliptic equilibrium points, the following limiting values of P_ϕ, Λ_P are found:

$$\begin{aligned} &P_\phi(0,D)=\frac{4\pi}{\sqrt{D^2-1}},\quad \Lambda_P(0,D)=\frac{1}{2}\left(\frac{1}{2}D-\frac{1}{2}-c\right)P_\phi(0,D),\quad |D|>1,\\ &P_\phi(\phi_{\mathrm{ell}},D)=\frac{2\pi}{\sqrt{1-|D|}},\quad \Lambda_P(\phi_{\mathrm{ell}},D)=-\frac{1}{2}\left(c+\frac{1}{2}-\frac{1}{2}\operatorname{sign}(D)\right)P_\phi(\phi_{\mathrm{ell}},D),\quad |D|<1.\end{aligned} \tag{A.9}$$

A.2. *Melnikov integral for the reduced system*

Below, we include the detailed formulae used in the reduced equations of Section 3b. Expanding Δ of (4.19) in ϵ leads to the expansion of the r.h.s. of (4.21) as follows:

$$\begin{aligned}
&\Delta = \Delta_0 + \epsilon\Delta_1 = 2H + c(1+c)\cos^2(\phi) - v^2 + 2\epsilon A(\phi)\sin(2k\Lambda+\theta),\\
&D = D_0 + \epsilon D_1 + \mathrm{O}(\epsilon^2) = (1+2c)\cos^2(\phi) \pm 2\cos(\phi)\sqrt{\Delta_0} \pm \epsilon\cos(\phi)(\Delta_1/\sqrt{\Delta_0}),\\
&\frac{\mathrm{d}v}{\mathrm{d}\Lambda} = \frac{\partial D_0(\phi, v, \Lambda; H)}{\partial\phi} + \epsilon\frac{\partial D_1(\phi, v, \Lambda; H)}{\partial\phi}\\
&\quad = 2\sin\phi(-\cos\phi(1+2c) \mp \sqrt{\Delta_0}) \mp \frac{c(1+c)\cos\phi\sin(2\phi)}{\sqrt{\Delta_0}}\\
&\quad \pm 2\frac{\epsilon\sin(2k\Lambda+\theta)}{\sqrt{\Delta_0}}\left(-A(\phi)\sin\phi + A'(\phi)\cos\phi + \frac{A(\phi)\cos\phi\sin(2\phi)c(1+c)}{\Delta_0}\right)\\
&\frac{\mathrm{d}\phi}{\mathrm{d}\Lambda} = -\frac{\partial D_0(\phi, v, \Lambda; H)}{\partial v} - \epsilon\frac{\partial D_1(\phi, v, \Lambda; H)}{\partial v} = \pm\frac{2v\cos\phi}{\sqrt{\Delta_0}}\left(1 - \epsilon\frac{A(\phi)\sin(2k\Lambda+\theta)}{\Delta_0}\right).
\end{aligned} \tag{A.10}$$

Hence, the Melnikov integral is given by

$$\begin{aligned}
M(\Lambda_0, \theta; H) &= \int_{-\infty}^{\infty} \{D_0, D_1\}|_{q_h(\Lambda, \Lambda+\Lambda_0)}\,\mathrm{d}\Lambda\\
&= \mp\cos(2k\Lambda_0+\theta)\\
&\quad \times \int_{-\infty}^{\infty} \sin(2k\Lambda)v\left(\frac{4A(\phi)\sin(\phi)}{\sqrt{\Delta_0}} \pm (1+2c)\frac{\sin(2\phi)A(\phi)}{\Delta_0} - \frac{2\cos(\phi)A'(\phi)}{\sqrt{\Delta_0}}\right)\Big|_{q_h(t)}\,\mathrm{d}t.
\end{aligned} \tag{A.11}$$

References

[1] V. Rom-Kedar, Some characteristics of two-degrees-of-freedom hamiltonian flows, in: Transport, Chaos and Plasma Physics, II Proc. (Marseille), eds. Benkadda, Doveil and Elskens, (World Scientific, Singapore, 1996) pp. 164–178.

[2] V. Rom-Kedar, Parabolic resonances and instabilities in near-integrable two-degrees-of-freedom Hamiltonian systems, Chaos 7 (1) (1997) 148–158.

[3] W. Von-Arx, An Introduction to Physical Oceanography (Addison-Wesley, Reading, MA, 1974).

[4] N. Paldor and P.D. Killworth, Inertial trajectories on a rotating earth, J. Atm. Sci. 45 (24) (1988) 4013–4019.

[5] N. Paldor and E. Boss, Chaotic trajectories of tidally perturbed inertial oscillations, J. Atm. Sci. 49 (23) (1992) 2306–2318.

[6] P. Jullian, V. Lally, W. Kellogg, V. Suomi and C. Cote, The TWERLE experiment, Bull. Am. Met. Soc. 58 (9) (1977) 936–948 (see especially their Fig. 4).

[7] P. Morel and M. Larcheveque, Relative dispersion of constant-level balloons in the 200 mb general circulation, J. Atmos. Sci. 31 (1974) 2189–2196.

[8] R.T. Pierrehumbert, Chaotic mixing of tracer and vorticity by modulated travelling Rossby waves Geophys, Astr. Fluid Dynamics 58 (1991) 285.

[9] R.T. Pierrehumbert and H. Yang, Global chaotic mixing on isentropic surfaces, J. Atmos. Sci. 50 (1993) 2462–2480.

[10] J.M. Ottino, Mixing, chaotic advection and turbulence, Ann. Rev. Fluid Mech. 22 (1990) 207–253.

[11] D. del-Castillo-Negrete and P.J. Morrison, Phys. Fluids A 5 (4) (1992) 948.

[12] J. Hale and H. Kocak, Dynamics and Bifurcations (Springer New York, 1991).

[13] L.M. Lerman and Ia. L. Umanskii, Classification of four-dimensional integrable Hamiltonian systems and Poisson actions for R^2 in extended neighborhoods of simple singular points. I, Russian Acad. Sci. Sb. Math. 77 (2) (1992) [English trans., 1994].

[14] L.M. Lerman and Ia.L. Umanskii, On the existence of separatrix loops in four dimensional systems similar to integrable Hamiltonian systems, PMM USSR 47 (1984) 335.

[15] C. Grotta-Ragazzo, Dynamics and diffusion near double-resonances in Hamiltonian systems with discrete symmetries, preprint (1995).
[16] P.J. Holmes and J.E. Marsden, Horseshoes in perturbations of Hamiltonian systems with two degrees of freedom, Comm. Math. Phys. 82 (1982) 523–544.
[17] P. Holmes, J. Marsden and J. Scheurle, Exponentially small splittings of separatices with applications to KAM theory and degenerate bifurcations, Cont. Math. 81 (1983) 213–244.
[18] G. Kovacic and S. Wiggins, Orbits homoclinic to resonances, with an application to chaos in a model of the forced and damped sine-Gordon equation, Physica D 57 (1992) 185–225.
[19] G. Haller and S. Wiggins Orbits homoclinic to resonances: the Hamiltonian case, Physica D 66 (1993) 298–346; N-pulse homoclinic orbits in perturbations of resonant Hamiltonian systems. Arch. Rational Mech. 130 (1995) 25–101.
[20] T. Kaper and G. Kovacic, Multi-bump orbits homoclinic to resonance bands, preprint (1994).
[21] V.K. Melnikov, On the stability of the center for time periodic perturbations, Trans. Moscow Math. Soc. 12 (1963) 1.
[22] J. Guckenheimer and P. Holmes, Non-Linear Oscillations, Dynamical Systems and Bifurcations of Vector Fields (Springer, New York, 1983).
[23] S. Wiggins, Introduction to Applied Nonlinear Dynamical Systems and Chaos (Springer, New York, 1990).
[24] S. Wiggins, Global Bifurcations and Chaos – Analytical Methods (Springer, New York, 1989).
[25] B.V. Chirikov, A universal instability of many-dimensional oscillator systems, Phys. Rep. 52 (5) (1979) 263.
[26] D.F. Escande, Hamiltonian chaos and Adiabaticity, in: Plasma Theory, Nonlinear and Turbulent Processes in Physics, (Proc. Int. Workshop (Kiev, 1987), eds. V.G. Bar'yakhtar, V.M. Chernousenko, N.S. Erokhin, A.G. Sitenko and V.E. Zakharov. (World Scientific, Singapore, (1988); Stochasticity in classical Hamiltonian systems: universal aspects, Phys. Rep. 121 (3–4) (1988) 165–261.
[27] G.M. Zaslavsky, R.Z. Sagdeev, D.A. Usikov and A.A. Chernikov, Weak Chaos and Quasiregular Patterns (Cambridge University Press, Cambridge, 1991).
[28] V. Rom-Kedar, A. Leonard and S. Wiggins, An analytical study of transport, mixing and chaos in an unsteady vortical flow, JFM 214 (1990) 347–394.
[29] A.J. Lichtenberg and M.A. Lieberman, Regular and Stochastic Motion (Springer, New York).
[30] D. Beigie, A. Leonard and S. Wiggins, Chaotic transport in the homoclinic and heteroclinic tangle regions of quasiperiodically forced two dimensional dynamical systems, Nonlinearity 4 (1991) 775.
[31] R.S. MacKay, J.D. Meiss and I.C. Percival, Transport in Hamiltonian systems, Physica D 13 (1984) 55.
[32] V. Rom-Kedar, Homoclinic tangles – classification and applications, Nonlinearity 7 (1994) 441–473.
[33] V.I. Arnold, Mathematical Methods of Classical Mechanics (Springer, New York, 1981).
[34] P.J. Holmes and J.E. Marsden, Horseshoes and Arnold diffusion for Hamiltonian systems on Lie groups, Indiana Univ. Math. J. 32 (2) (1983) 273–309.
[35] V. Rom-Kedar and N. Paldor, From the tropic to the poles in forty days, preprint (1996).
[36] G. Haller, Universal homoclinic bifurcations and chaos near double resonances, J. Stat. Phys. (1997), in Press.
[37] B. Trounday, L. Perthuis, S. Strebelle, J.D. Farrara and C.R. Mechoso, Dispersion properties of the flow in the southern stratosphere during winter and spring, J. Geophys. Res. 100 (D7) (1995) 13 901–13 917.
[38] I.S. Gradshteyn and I.M. Ryzhik, Table of Integrals, Series and Products (Academic Press, New York).