Chaotic Tides*

LEO R. M. MAAS

Netherlands Institute for Sea Research, Texel, Netherlands

ARJEN DOELMAN

Korteweg-de Vries Instituut, Universiteit van Amsterdam, Amsterdam, Netherlands

(Manuscript received 20 March 2001, in final form 11 July 2001)

ABSTRACT

Persistent reports exist that the tides in coastal basins are often accompanied by regular or irregular oscillations of periods ranging from minutes up to several hours. A conceptual model relating the two is proposed here. It employs an almost-enclosed basin, connected to an open (tidal) sea by a narrow strait. Such a basin is a Helmholtz resonator, which is dominated by the ''pumping'' mode. Its response is governed by an ordinary differential equation that is forced by the tide, damped by friction and wave radiation, and whose restoring term is nonlinear due to the sloping bottom. When forced resonantly by a single frequency tide, due to this nonlinearity, the basin may exhibit multiple equilibria. Its response can either be amplified or choked, depending on the precise initial conditions. The presence of a second forcing term may, on a slow timescale, kick the system irregularly from the amplified into the choked regime, yielding a chaotic response. This may happen when either two nearby frequencies, for example, a combination of semidiurnal lunar and solar tides, are near resonance (and the frequency difference provides the beat), or when a small-amplitude, resonant perturbation is modulated by a large-amplitude, low-frequency tide. The aforementioned observations of irregular tides are discussed in the light of analytical and numerical results obtained with this model for these two regimes.

1. Introduction

''Chaotic tides'' sounds like an oxymoron. As the etymology of the word ''tide,'' being time, season, or period according to Webster's (1913) *Revised Unabridged Dictionary* implies, it has, from ancient times on (Pugh 1987), often been held to be synonymous with periodicity. Indeed, this regularity made tidal heights one of the earliest known and best predictable phenomena in physical oceanography, with relative accuracies of prediction often exceeding 90%. The predictive capacities of the corresponding tidal *currents* is, admittedly, less impressive, but is still often in excess of 50%. Less known, however, is the fact that spurious, but persistent, reports on the *irregularity* of the tides exist for over a century (Honda et al. 1908; Krümmel 1911). It is this irregularity in the tide that we here aim to identify as being possibly due to the chaotic response of certain coastal basins to the tide at the open sea.

Tidal predictions have traditionally been made for individual ports on the basis of previously observed tidal elevation records at those locations. Amplitudes and phases of a number of precisely known ''tidal frequencies,'' stemming from the gravitational potential determined by celestial mechanics, are estimated, which can then be used to sum the corresponding harmonic series and thus to predict the future occurrence of tides at that location; see, for example, the review by Godin (1991).

Although the gravitational (or tidal) potential carries only about 5 relevant, independent frequency components, which, through an expansion of the tidal potential in harmonic terms leads to perhaps 11 principal tidal components (Platzman 1971), in the shallow, coastal areas, local nonlinear effects lead to leakage of energy to higher harmonics and sum and difference (combination) frequencies. In this way tidal energy fills in entire spectral bands surrounding the principal components (Pugh 1987, p. 188). As there are, in principle, no difficulties in also resolving the amplitudes and phases of these combination frequencies, running harmonic analysis with 150 frequency components or more has become routine procedure. This empirical *local* analysis, however, does not automatically guarantee that the amplitudes and phases of these independently determined frequencies also show a *spatially* coherent and stationary picture. Quite the contrary.

A spatially coherent description of the tides can be

Corresponding author address: Dr. Leo Maas, Netherlands Institute for Sea Research, P.O. Box 59, 1790 AB Texel, Netherlands.
E-mail: maas@nioz.nl

obtained by solving the governing *dynamical* equations, the Laplace tidal equations (see, e.g., Cartwright 1977). In a realistic ocean domain this is achieved by numerical methods (Schwiderski 1980). Outside continental margins, tidal predictions for these principal components, particularly when constrained by observations from deep-sea tidal stations, are in fair agreement with observations obtained at different seaports and deep sea tidal stations (global root-mean-square difference between modeled and observed elevations of about 10 cm). With the advent of satellite altimetry providing tidal observations, this agreement has in recent years actually been dramatically improved (global root-mean-square difference of about 3 cm); see Andersen et al. (1995).

Except for these directly driven, principal components, together perhaps with their first few harmonics (Davies 1986; Lynch and Werner 1991), spatial coherency is cumbersome and agreement with locally observed tidal components poor, particularly in the shallow, coastal areas. Waves at combination frequencies are generated locally, get amplified, or damped, to a degree depending on the particular resonance properties of that locality, and do not persist as free waves so that they decorrelate quickly. Treating the nonlinearly generated compound waves locally as linearly independent (as harmonic, or Fourier analysis wants it) is therefore somewhat contorted but still defendable when the local nonlinear transfer of energy is *frozen* so that harmonics appear as the result of a stationary process. However, when analyzing different, independent datasets from the same location, variations in tidal amplitudes and phases often still occur, to the extent that some of the minor components appear, in fact, *unresolvable* (Gutiérrez et al. 1981; Godin 1991). Although such variations in locally estimated tidal amplitudes and phases are usually attributed to nonstationary effects due to wind (van Ette and Schoemaker 1966; Gutiérrez et al. 1981), this need not necessarily be its only source. Nonstationarity may also be due to nonlinearity of the hydrodynamic system itself (Pugh 1987), which may not only change the tidal elevation profile, by giving rise to *superharmonics* (overtides, which stay fixed in time), but may also modulate its amplitude, by giving rise to *subharmonics*. Finally, strong nonlinearity may provoke a cascade of such subharmonic bifurcations, giving rise to *chaotic behavior*.

Is there any observational basis for chaotic behavior of the tides, other than the aforementioned unresolvability, or noisiness, of tidal "constants"? In an analysis of the tides in Venice Lagoon, at the head of the Adriatic Sea, where the tides seem to pick up because of near-resonancy of the basin, Vittori (1992) observed that consecutive tidal maxima are highly irregular. She argued this to be indicative of low-dimensional chaos. Whether the low-order dynamics to which this is due is either inherited from the dynamics of the local wind fields or of a genuinely oceanographic nature is not clear. Similar changes in consecutive maxima also appear in long-

term, tidal elevation records in the Wadden Sea basins (J. T. F. Zimmerman 2000, personal communication), which are again close to resonance (see Maas 1997, hereafter referred to as M). Frison et al. (1999) find evidence of nonlinear tides in U.S. coastal estuaries from the nontrivial shape of the attractor obtained in a diagram where observed elevation at some moment is plotted against that observed at some previous instant.

Surprisingly, direct observations of irregular oscillations, accompanying the tides, have been reported for over a century (see Krümmel 1911, 157–185; Defant 1961, p. 187). The suggestion that these "secondary oscillations" were actually related to the "primary," tidal oscillation was firmly put forward in an impressive study by Honda et al. (1908), who had intensively studied the seas around Japan. This study eventually led to a complete classification of bays in terms of "periodic," "quasiperiodic," etc. (Honda et al. 1908; Nakano, 1932). In the absence of a dynamical framework, such as recently offered by the field of nonlinear dynamics, this classification was, however, not further interpreted, and these observations seem, if not completely forgotten, neglected in contemporary tidal literature. One reason for this might be that these secondary oscillations generally seem inconspicuous, as they consist of high-frequency, low-amplitude waves superimposed on a low-frequency, high-amplitude (primary) tide. Typical periods range from some minutes to several hours, and amplitudes do not exceed a few percent of the tidal range. However, an interesting twist is given to this interpretation by recent observations of such irregular tidal elevations in a Norwegian fjord (Golmen et al. 1994). By making simultaneous observations of the currents in the strait connecting the fjord to the sea they found that these irregular small-amplitude elevations were accompanied by irregular $O(1)$ variations of the tidal current (see Fig. 1). As velocities are determined by elevation *differences* over the strait, by inference the *velocity* field can be amplified when, at any moment, the difference in tidal elevation also amounts to just a few percent of the elevation itself, thereby becoming of comparable magnitude to the difference between the (small amplitude) high-frequency elevation that is resonantly excited within the basin but is practically absent at sea. Such $O(1)$ velocity variations are not only important for nautical reasons, but are clearly equally relevant for the flushing of the fjord: the exchange of water, sediments, and dissolved gases or nutrients with the connected sea. It is believed (LeBlond 1991) that this property will transcend the global relevance that both dissipation as well as resonance of tides in the coastal environment may have in setting the boundary conditions for the global tide (Garrett 1975), although Munk and Wunsch (1998) recently made the interesting suggestion that tides might be playing a key role in ocean circulation.

In order to illustrate these issues we will consider the resonant response of a short, deep basin having a sloping
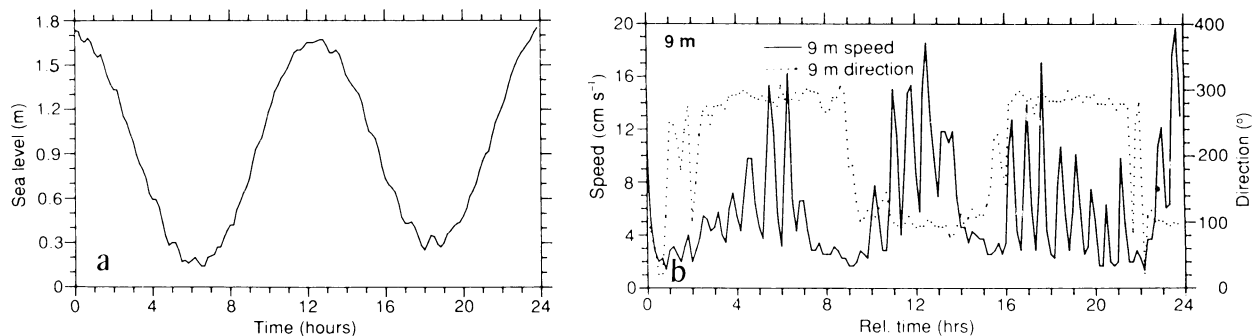
FIG. 1. Observed sea level in Moldefjord and currents through connecting strait; from Golmen et al. (1994).

bottom, to the tide in the adjacent open sea to which it is connected by a narrow strait—a nonlinear Helmholtz resonator (see M). This choice of geometry differs from the mostly wide bays with their quarter-wave resonancies, encountered by Honda et al. (1908), but it ensures that the tide within the basin takes the simplest form possible, as it is governed by the pumping, or Helmholtz mode. This mode is the lowest and generally also the most important frequency in the basin's spectrum, and is characterized by a spatially uniform response. It can thus be represented by a single state variable (the excess volume of water, related to the free surface elevation), whose evolution is therefore described simply by an *ordinary* differential equation. Although nonlinearity can also be present in the frictional damping term (Zimmerman 1992), it is the nonlinearity in the restoring term that gives rise to multiple equilibria and chaos here. The latter nonlinearity is due to the sloping bottom (Green 1992) and can be understood as follows. The current through the strait is driven simply by the elevation difference over the strait. However, the elevation change within the basin, affected by the transport through the strait, clearly depends on the surface area that the incoming water has to cover, which is nonuniform with depth when the walls of the basin are not vertical. The time needed to produce a particular elevation change therefore depends on the preexisting sea level; see Fig. 2. In M its free, forced, and forced-and-damped response was discussed in the case where

the tidal forcing is "pure," that is, contains one frequency component only. In inviscid circumstances, a single frequency forcing can provoke a chaotic response in a small parameter range. However, when damping is added, the chaos fails to persist. Since the inclusion of damping terms is mandatory in any realistic physical context, this "Hamiltonian chaos" appears to have no physical relevance. It is one of the goals of the present paper to determine under what conditions tides, within basins of the kind considered here, can become chaotic, even in the presence of damping. Multiple frequency forcing seems to be a prerequisite for this. Therefore, in light of the observations discussed above, we will consider forcing with either two nearby frequencies, as is the case for a combination of lunar and solar tides, a "mixed" type of tide (see Defant 1961), or with a single low-frequency tide, accompanied by a small-amplitude, high-frequency, resonant perturbation, as is the case in a fjord.

In section 2 the nonlinear tidal Helmholtz resonator will be introduced. It represents the response of a bay to tidal variations at the open sea, to which it is connected by a narrow channel. Its governing ordinary differential equation will be derived. This section also introduces the Poincaré phase plane, which gives a more comprehensive way to follow its evolution than that obtained by direct numerical integration. The tidal response of the bay typically shows a gradual modulation of amplitude and phase, which can be captured by av-
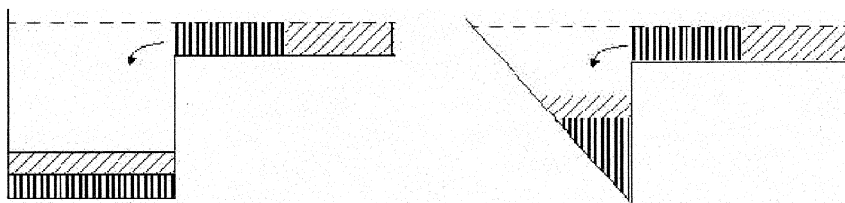

FIG. 2. Sketch showing a side view of two basins that have a channel, connecting it to the sea, at their right-hand sides (Maas 1998). It demonstrates that the elevation change brought about by two equal, consecutively entering packages of water (see arrows) is the same when the basin has a cross-sectional area that is independent of height above the bottom (left), while it is *different* when this cross-sectional area increases with height (right). In the latter case, the response depends on the preexisting water level and is thus nonlinearly related to the incoming flux of water.

eraging the original governing equation (section 3). In section 4 this modulation equation is applied to the case that forcing is at two nearby frequencies. When these are close to resonance, the response of the basin may, under certain conditions, be chaotic. The case that a single, low-frequency tide is accompanied by a small-amplitude, high-frequency resonant perturbation is very similar to this, and is discussed in section 5. In section 6 the results will be summarized.

2. Nonlinear Helmholtz resonator

An almost-enclosed, short basin, connected to a tidal sea by a narrow strait (modeled here as a *pipe*), is generally governed by the pumping (Helmholtz) mode, in which the water level within the basin rises and sinks in unison. Elevation is scaled with (maximum) depth H, volume with $A_0 H$ (where A_0 is the surface area of the basin at rest), while time t is scaled with the Helmholtz frequency of the basin σ_H,

$$\sigma_H = (gHB/A_0 L)^{1/2}, \tag{1}$$

with B and L denoting the strait's width and length, respectively, and g the acceleration of gravity. In M it is shown that the evolution of the (nondimensional) amount of water in excess of that present in the basin in the absence of tides—the *excess* volume, v—is (neglecting quadratic damping terms) governed by

$$\frac{d^2 v}{dt^2} + \zeta(v) = \zeta_e(t) - c\frac{dv}{dt}. \tag{2}$$

Here ζ and ζ_e denote the surface elevation within the basin and in the exterior sea, respectively. The latter is a prescribed function of time. At any time their difference provides the pressure difference that drives the flow u along the connecting strait: $du/dt = \zeta_e - \zeta$. This flow, multiplied by the (unit) cross-sectional area of the strait, equals the rate at which the volume within the basin changes, $u = dv/dt$. From these two equations the inviscid form ($c = 0$) of (2) follows. The basin hypsometry (horizontal area A as a function of vertical coordinate z) determines the excess volume within the basin. With respect to mean sea level $z = 0$, this is defined as

$$v \equiv \int_0^\zeta A(z)\, dz. \tag{3}$$

The inverse of this relation, needed in (2), yields the nonlinear restoring term $\zeta(v)$, which can be made explicit for some simple basin shapes. For a basin with vertical sidewalls, $A = 1$, for instance, it follows from (3) that the restoring term is linear, $\zeta = v$. However, when the basin area increases linearly with depth, $A = 1 + z$ (the deepest point of the basin being at $z = -1$), this restoring term is given by

$$\zeta(v) = (1 + 2v)^{1/2} - 1. \tag{4}$$

This shape of the basin is the only one addressed here

and its hypsometric behavior is thought to be characteristic of a typical basin with shoaling sides (see M). Other basin shapes and the effect of this shape on the dynamics of the basin are addressed elsewhere (Doelman et al. 2001, submitted to *Phys. D,* hereafter DKM). A nonempty basin (and analyticity of the solution) requires $v > -1/2$. In the right hand of (2) we find forcing by the prescribed external tide, $\zeta_e(t)$, and damping, respectively. Although this linear damping formally stands for radiative damping only, we may also consider it to incorporate a linearized version of bottom friction and form drag (see M). For further analysis we treat (2) as a dynamical system by writing for the basin's excess volume $x = v$, and for the strait's flow rate $y = dv/dt$, and obtain, with (4),

$$\begin{cases} \dot{x} = y \\ \dot{y} = 1 - \sqrt{1 + 2x} - cy + \zeta_e(t). \end{cases} \tag{5}$$

As usual, a dot indicates a time derivative. Although the emphasis will here be on the response to quasiperiodic (two frequency) forcing, some results obtained without forcing and with "single frequency" forcing, pertinent to the discussion to follow, will be recapitulated first.

a. Poincaré plane, free response, and periodic forcing

A numerical example of an integration of (5) is given in Fig. 3a for the case without forcing. It presents the evolution over two tidal periods of volume and current, $(x(t), y(t))$, as a function of time, increasing upward. Projections of this helical curve at the back (x–t) and side (y–t) planes give the time evolution of volume and current proper, such as they are commonly presented (except that time now points upward). The third projection, on the bottom x–y plane, suppresses time, which now acts as a parameter *along* such orbits, and is called the phase plane. In this strongly viscous example time can still be discerned in the gradual decay of the oscillations. However, when damping is also absent, solution curves will "circle" around a common center point, which is the only fixed point present in this case. Each solution curve is egg shaped and is described by elliptic functions in terms of which the inviscid, unforced version of (5) can be solved (M). Different curves represent different initial conditions (volume and current). The phase plane is limited by an outer boundary (the limiting orbit) because the argument of the square root in the restoring term of (5) needs to be nonnegative. It can be found in parametric form as

$$x(t) = (3 - t^2)(15 - t^2)/72, \quad y(t) = t(t^2 - 9)/18,$$

for $t \in (-3, 3)$, which is periodic with period 6. Owing to the nonlinearity of the oscillator, the period of such a free, finite-amplitude oscillation is a function of its amplitude. For the present geometry this period spans
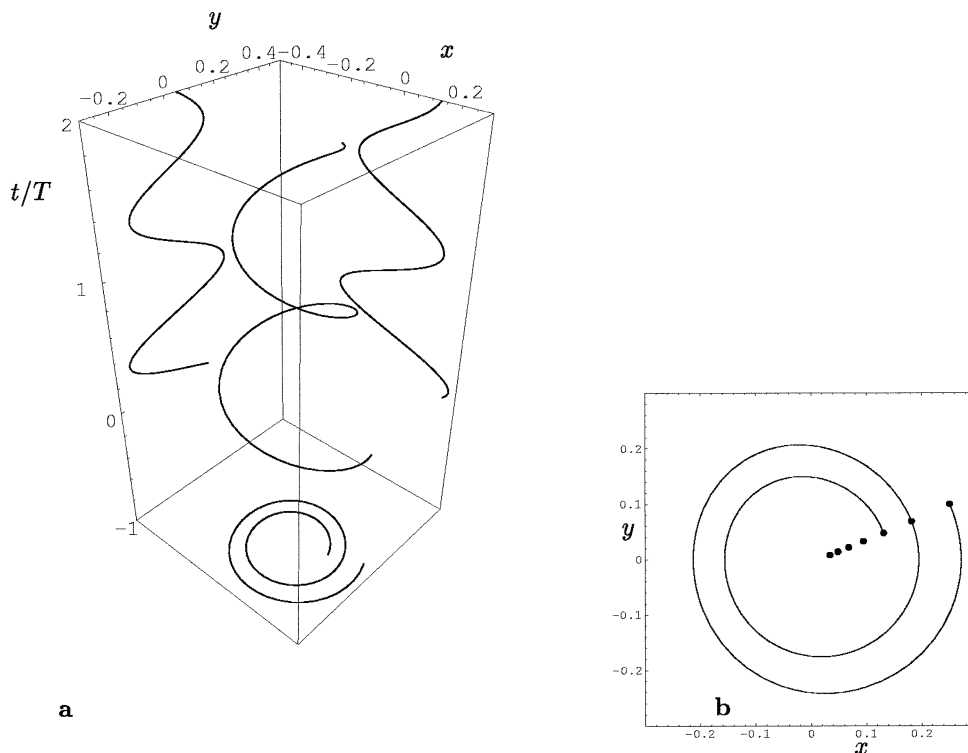
FIG. 3. (a) Example of time (t) evolution of volume (x) and current (y) over two tidal periods T, together with three projections, of the unforced ($\zeta_e = 0$), damped system (5). (b) Projection of helix in (a) on the x–y plane (solid curve), and stroboscopic (tidal) sampling thereof (dots), representing intersections in (a) with horizontal planes at $t/T = 0, 1, 2, \cdots, 6$.

just a narrow range, decreasing from 2π, at the center, down to 6 at the outer boundary. The frequency range of other geometries is considered in DKM.

A comprehensive way of monitoring gradual changes of the tidal response due to forcing, damping, or nonlinearity is obtained by determining subsequent intersections of the helical curve with a sequence of periodically placed, horizontal planes at $t/T = 0, 1, 2, \ldots$, with T indicating the forcing period (so-called Poincaré planes of section). These subsequent intersections define a map, called the Poincaré map, P^T. A strictly periodic signal (of period T) thus yields the same intersection. Compiling each of these subsequent intersections on one and the same horizontal (phase) plane, a stroboscopic impression of the slow evolution of the nearly periodic behavior is obtained. An example of such a Poincaré plane is presented by the dots in Fig. 3b. Here, the spiral is the projection of the continuous helical curve of Fig. 3a on the bottom plane. The phase at which the forcing is sampled can still be varied but is in this study taken to correspond to the moment that the external tide is at low water. From here on we will mostly use this phase-plane representation of the evolution, accepting the suppression of time information. This is illustrated in Fig. 4 with some stroboscopically sampled solutions of (5). When forcing is absent, damping brings any

initial motion to rest, in this Poincaré plane represented by a sequence of points approaching the origin (Fig. 4a). When damping is absent and a single forcing component,

$$\zeta_e = f \cos(\omega t + \theta), \tag{6}$$

is near resonance ($\omega \approx 1$), the amplitude of the basin tide will grow because of the nonlinearity of the restoring term. However, doing so, it changes its period so that the resonator is detuning itself, and the tide at sea will actually damp the basin response. Once it is small again, this process may start afresh, and the whole cycle takes a time span, which one calls the modulation period, that is large compared to the tidal period. This slow growth and decay thus reveals itself in the Poincaré plane as a single closed loop that is traversed in one modulation period. Other confocal loops represent other amplitudes of oscillation. However, when the forcing period is close to resonance, it is clear from the three curves of Fig. 4b that there may be different types of responses, which will be discussed further in the next section. With forcing and relatively strong damping, a regular oscillation of definite amplitude and phase eventually results, so stroboscopic intersections will approach one single spot, away from the orgin (Fig. 4c). When damping is weak and forcing is near resonance,
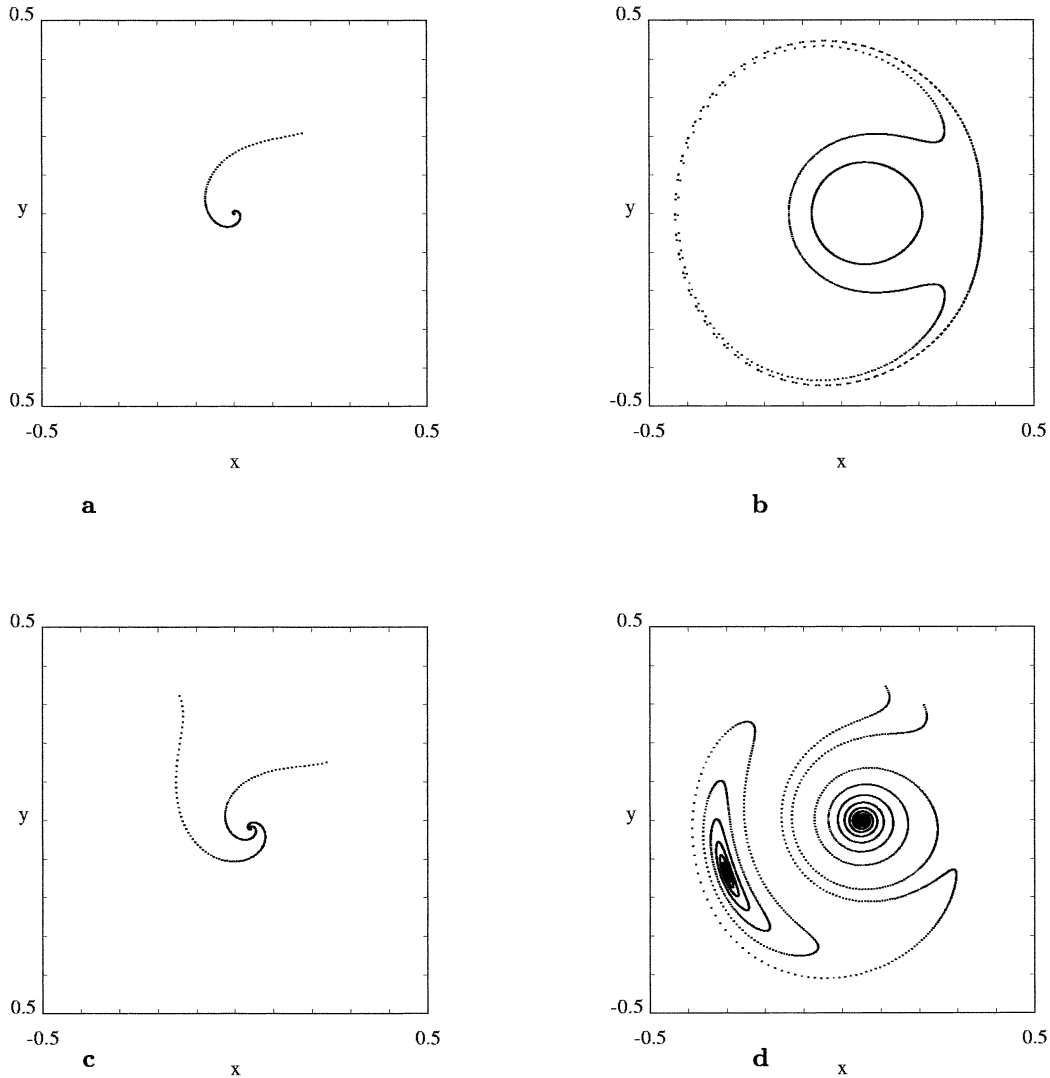
FIG. 4. Stroboscopic plot at $t/T = 0, 1, 2, \cdots, 0$, where $T = 2\pi/\omega$, and forcing $\zeta_e = -f \cos(\omega t)$, $\omega = 1.01$, and forcing amplitude f and damping parameter C given by (a) $f = 0$, $C = 0.01$, (b) $f = 0.001$, $C = 0$, (c) $f = 0.001$, $C = 0.01$, and (d) $f = 0.001$, $C = 0.001$.

depending on the precise initial state, the system may lock into *different* states (Fig. 4d).

To summarize, in this Poincaré plane, a strictly periodic oscillation will be represented by a single (fixed) point, a quasiperiodic oscillation by multiple points lying on a closed curve (limit cycle), and an aperiodic state (encountered in later sections) by an irregularly located set of points (strange attractor).

b. Qualitative effect of quasiperiodic forcing

The multiple equilibria near the Helmholtz resonance correspond to an amplified and a choked tidal response (see the two centers in Fig. 4d, at large and small volumes $|x|$, respectively). These states are stable and attracting when damping is present (Fig. 4d). Without

damping, however, these states would retain their original distances to the fixed points (as the curves in Fig. 4b), and each state can be characterized by its energy level and period (with which it is traversed). The qualitatively different regimes, recognizable in Fig. 4b, are separated from each other by particular special orbits, the separatrices. One outer separatrix separates the outermost orbit from the other two, and one inner separatrix separates the banana-shaped orbit and small circular orbit. The period needed to traverse these separatrices approaches infinity, as the intersection of the two separatrices acts like a saddle. A state can creep up to this relative maximum in potential energy by exchanging some of its "kinetic" energy, allowing it to move forward along its orbit, in favor of "potential" energy.

When friction is introduced, it pulls the state down

into one of the equilibria, except for the odd point sitting right at the saddle and the two orbits that limit exactly on the saddle point in forward time. Note that these orbits merge with the aforementioned separatrices as the friction decreases to 0. Apart from these exceptions, all initial states thus belong to one of two domains of attraction (amplification or choking). When one starts to weakly and slowly "shake" the system (i.e., the saddle and the separatrices) by introducing a large-period perturbation, a new "equilibrium" may arise. This happens because it is clearly quite a sensitive matter near the separatrices to which equilibrium their state will recede. Therefore, when, due to shaking, the separatrices move a little, while the state is slowly receding to one equilibrium, it may find itself a little later in the attraction domain of the other. Making again its way to the other equilibrium, something similar may happen, and, in effect, the neighborhoods of the separatrices may trap the state.

This new "steady" state might turn out, however, to be chaotic. This is because, close to the separatrices, the period of the inviscid orbits (the modulation period of the Helmholtz oscillator) varies greatly, and thus, when the state of the system varies due to the weak forcing, neighboring states fastly diverge from each other.

The way to verify that such a new dynamical equilibrium state turns into existence under the addition of perturbative damping and extra forcing is by performing an energy budget study. For this, one picks one of the original inviscid orbits, characterized by a particular period, and one then evaluates, along that orbit, the net increase or decrease of energy over one period due to both the work done by the perturbative forcing as well as the energy dissipation by the perturbative damping term. Only when these two are in balance (and there is thus no net energy increase or decrease) will the point on average stay near the original, unperturbed orbit. The single orbits considered in the present study are the previously introduced separatrices (homoclinic orbits), along which the period turns to infinity. The function determining the net energy change on the homoclinic orbits is called the Melnikov function. This function is usually given the alternative interpretation of measuring the distance of the perturbed orbits leaving and approaching the saddle. A (traditional) search for its zeros—implying intersections of the two perturbed orbits so that points both belong to the sets of points approaching and leaving the (perturbed) saddle, and are therefore "stuck"—is thus equivalent to the requirement that there is no net change in energy content along its trajectory. Mathematically, all we can show for the moment is that certain points (that belong to this invariant set) get trapped near the unperturbed separatrices, and that they traverse that region in an unpredictable, chaotic way (section 4a). However, we lack the ability to rigorously show that this set is actually at-

tracting (Wiggins 1990): Suggestions that it is follow from numerical analysis (section 4b).

Similar such phenomena appear at other resonances, like near $\omega \approx 1/2 + \varepsilon$, and in particular near $\omega \approx 2 \pm \varepsilon$ (see M).

3. Modulation equations

Much of the behavior obtained by numerically integrating the exact equation can, to a remarkable degree, also be found analytically in the solutions of the modulation equations that will be considered now. The advantage of using modulation equations is that fixed points, homoclinic orbits, Melnikov functions, etc. can be obtained explicitly, which is very helpful in addressing the main issue, namely, whether (and under what conditions) chaos can occur in the nonlinear Helmholtz resonator in the presence of damping. The stroboscopic intersections in Fig. 4 trace out continuous curves, whose locations can be obtained approximately when averaging the governing equations over the forcing period. Doing so, one obtains modulation equations that govern the amplitude dynamics on a slow timescale, which reveal where and under what conditions the dynamical response can become "complex." When the excess-volume fluctuation $v(t)$ is of small amplitude (and when also the forcing and frictional parameters are small), (2) can be approximated by rescaling $v = \varepsilon V$, $c = \varepsilon^2 C$, $f = \varepsilon^3 F$ (with $\varepsilon \ll 1$). With sinusoidal forcing (6), this leads to

$$\frac{d^2V}{dt^2} + V = \frac{\varepsilon}{2}V^2 - \frac{\varepsilon^2}{2}V^3 + \varepsilon^2 F \cos(\omega t + \theta)$$
$$- \varepsilon^2 C \frac{dV}{dt} + O(\varepsilon^3). \quad (7)$$

This equation is solved by a multiple-scale perturbation expansion and has an amplified response at rational values of ω (Nayfeh and Mook 1979, p. 196). By requiring the absence of secular terms in the equations governing the subsequent orders of the perturbation expansion near the primary resonance (which, for consistency, requires $\omega = 1 + \varepsilon^2 \sigma$) the solution is, up to second order, given by

$$V = R \cos(\omega t - \Phi) + \frac{1}{4}\varepsilon R^2 \left[1 - \frac{1}{3}\cos2(\omega t - \Phi)\right]$$
$$+ O(\varepsilon^2). \quad (8)$$

The response is thus frequency locked to the forcing and exhibits a "drift" term (the constant term), providing a net displacement of the mean state. Amplitude R and phase Φ vary on a slow timescale and are governed by

$$\dot{R} = -\frac{C}{2}R + \frac{F}{2}\sin(\Phi + \theta) \quad (9)$$

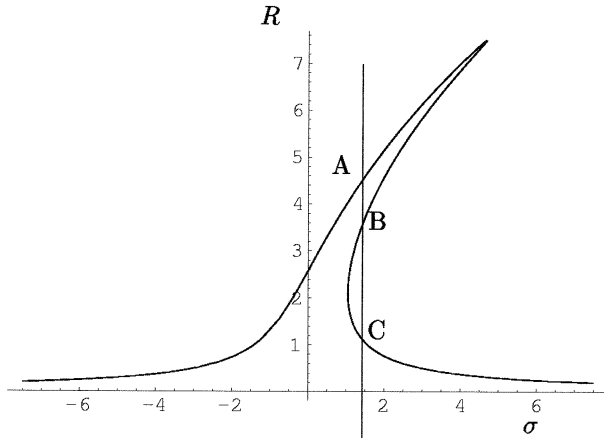$$R\dot{\Phi} = \sigma R - \frac{1}{12}R^3 + \frac{F}{2}\cos(\Phi + \theta), \quad (10)$$

FIG. 5. Amplitude response R against detuned frequency σ. Near resonancy, $1 < \sigma < 5$, it has a bent "resonance horn," indicating multiple equilibria and an increased response. Here forcing $F = 3$ and frictional parameter $C = 0.4$. Near $\sigma \approx 1.5$ (vertical line) three equilibria, A, B, and C, are obtained. Two of these are stable and correspond to a choked (C) and amplified (A) regime.

where a dot indicates differentiating with respect to the slow time variable $T = \varepsilon^2 t$. Equating the right-hand sides to zero determines the steady states and allows one to obtain the frequency response curve for these equations $\sigma(R; F, C)$; see Fig. 5. However, by rescaling $t \rightarrow t/\sigma$, $R \rightarrow \sigma^{1/2} R$, $C \rightarrow \sigma C$, and $F \rightarrow \sigma^{3/2} F$, the detuning frequency σ can be scaled out and can, therefore, without loss of generality, be set equal to one. We may use this property to argue that the actual detuning *defines* the small parameter $\varepsilon \equiv (\omega - 1)^{1/2}$, which would otherwise be left arbitrary. This is consistent with the two rescalings, employed previously. In Cartesian coordinates, $X = R \cos\Phi$ and $Y = R \sin\Phi$, these modulation equations read

$$\dot{X} = -\frac{C}{2} X + \left(\frac{R^2}{12} - 1\right) Y + \frac{F}{2} \sin\theta, \qquad (11)$$

$$\dot{Y} = -\frac{C}{2} Y - \left(\frac{R^2}{12} - 1\right) X + \frac{F}{2} \cos\theta, \qquad (12)$$

where $R^2 = X^2 + Y^2$.

For the single-frequency forcing one can assume, without loss of generality, $\theta = 0$ (which amounts to a rotation of our coordinate system). Its explicit presence is retained here however for later use, when, in the double-frequency forcing case, both F and θ will be varying on the slow timescale. Note that these modulation equations are the same as those found for the Duffing equation [in polar coordinates by the method of multiple scales, see Nayfeh and Mook (1979), and in both polar and Cartesian coordinates by the averaging method, see Guckenheimer and Holmes (1983)]. The extension to two forcing frequencies, discussed later, should thus have relevance to the quasiperiodically forced Duffing equation (Wiggins 1990; Yagasaki 1990,

1993), and extends it to the "double resonant" case of two nearby frequencies, both approximately equal to the natural frequency.

a. Comparison of exact and modulated systems

It is useful to understand the system of modulation equations because this *local* approximation has a strong correspondence with the Poincaré plot of the exact equation. To appreciate this, compare some of its numerically obtained solution curves with Poincaré plots of the exact equation (Fig. 6). Not only is the topology of fixed points and separatrices preserved but, indeed, so is much of their location (when giving proper notice to the rescaling factor ε, which is present in between these two systems). The modulation equations, however, have no restrictions on the region attainable in X–Y phase space, while the exact equations do. It is probably this restricted access of the original phase plane that makes a (formally) local solution perform so well in approximating its "global" dynamics. We will find that all of the behavior found near resonance in the original phase space is mimicked in the phase plane of the modulation equations (including, as we will see, the occurrence of chaos), which therefore forms a useful substitute.

b. Hamiltonian description

In the next section, particular attention will be given to the inviscid ($C = 0$) limit of (11)–(12) since these equations then turn into Hamiltonian form:

$$\dot{X} = \frac{\partial H}{\partial Y} \qquad (13)$$

$$\dot{Y} = -\frac{\partial H}{\partial X} \qquad (14)$$

with (for $\theta = 0$) Hamiltonian

$$H(X, Y) = 3\left[\left(\frac{R^2}{12} - 1\right)^2 - \frac{FX}{6}\right]. \qquad (15)$$

Here a constant has arbitrarily been included to simplify the expression. Multiple equilibria, in this case, exist for $|F| < 8/3$. Because of the symmetry of the modulation equations (11)–(12), $\mathbf{X} \rightarrow -\mathbf{X}$, when $F \rightarrow -F$, it is sufficient to consider F to be single signed only. From here on we take this to be the interval: $-8/3 < F < 0$. Equilibria are, of course, more readily identified from (11)–(12). With $\theta \equiv 0$, they have $Y = 0$, while $X(X^2/12 - 1) = F/2$. With auxiliary variable $\alpha \in (-\pi/2, \pi/2)$, defined through $F = -8 \sin(\alpha)/3$, these steady states are in this case given by

$$X_k = 4 \sin\nu_k, \qquad (16)$$

$$\nu_k = (\alpha + 2\pi k)/3, \qquad (17)$$

and $k \in \{-1, 0, 1\}$, which implies $X_{-1} < X_0 < X_1$.
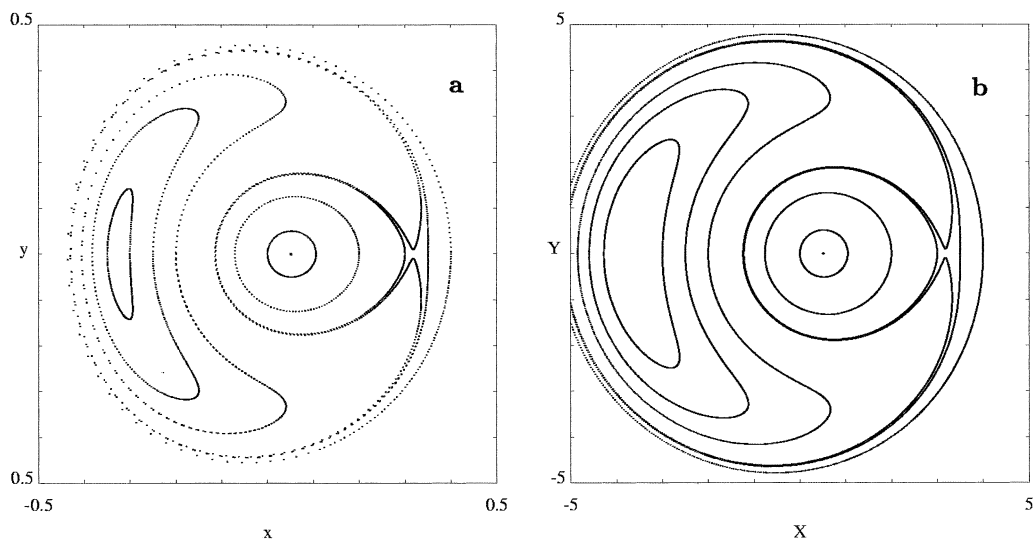
FIG. 6. (a) Poincaré plot of numerical integration of the inviscid ($c = 0$) evolution equation (5) for forcing frequency $\omega = 1.01$, close to resonance. Forcing $f = -10^{-3}$, $\theta = 0$, and initial conditions are $y = 0$, $x \in \{-0.3, -0.25, -0.2, 0.05, 0.2, 0.3, 0.35, 0.4\}$, as well as $y = 0.01$, $x = 0.3185$, which is close to the saddle point. (b) Trajectories of modulation equations (11) and (12) for the corresponding parameters and initial conditions, with $\varepsilon = (\omega - 1)^{1/2} = 0.1$ (and, hence, $\sigma = 1$), after a rescaling $\mathbf{x} = \varepsilon\mathbf{X}$, and employing a rescaled forcing $F = -1$, which is ε^3 times the original f in (a). The two figures are comparable, taking into account the scale factor $\varepsilon = 0.1$.

c. Homoclinic orbits

The motivation for studying the inviscid case ($C = 0$) is related to the fact that the modulation equations can then be solved explicitly. This offers the means to calculate Melnikov functions explicitly, once, at a later stage, weak (slow timescale) forcing and damping are added to the modulation equations. Zeros of the Melnikov functions determine the presence of invariant chaotic sets and suggest the presence of chaotic solutions. The conditions under which zeros of these functions appear will map out regions in parameter space where one may find "chaotic tides."

In the inviscid case ($C = 0$) explicit solutions of (11)–(12) are obtained by introducing

$$S \equiv R^2/12 - 1, \tag{18}$$

which measures the radial "distance" to the "special" radius $R_s = 2\sqrt{3}$. Note that it obeys the equation

$$\frac{dS}{dT} = \frac{FY}{12}, \tag{19}$$

where we have used $\theta \equiv 0$. Hence, by multiplying (11) by $F/6$ one obtains, upon integration,

$$\frac{XF}{6} = S^2 - K, \tag{20}$$

where K is an integration constant related to the energy level. By using $Y = \pm(R^2 - X^2)^{1/2}$ in (19) and then eliminating X from (20), and R^2 from (18), one obtains

$$\begin{aligned}
\frac{dS}{dT} &= \pm\left[\frac{F^2}{12}(S + 1) - \frac{1}{4}(S^2 - K)^2\right]^{1/2} \\
&\equiv \pm\frac{1}{2}(S - S_1)[(S - S_2)(S_3 - S)]^{1/2} \\
&\equiv \pm P_4(S)^{1/2},
\end{aligned} \tag{21}$$

where explicit expressions for $S_{1,2,3}$, in terms of F and K, are found by identification of both expressions (see appendix A). Plots of the graphic $P_4(S)$ appearing in (21) are given in Fig. 7a. These correspond, for any given energy level, to the orbits in the X–Y phase plane, given in Fig. 7b. Of particular interest is the case that the central zeros of the quartic coalesce, and the corresponding orbits represent the homoclinic orbits $\gamma_{\text{in,out}}$. These homoclinic orbits correspond to the separatrices described in section 2b. Equation (21) can be integrated and its solutions can be expressed in terms of elliptic functions. However, we will have no need here for the general expressions and will only describe the homoclinic orbits coming from the saddle point $(X_1, 0)$, which will be used in section 5a to explicitly calculate Melnikov functions.

The solution of (21) for the outer homoclinic orbit, γ_{out}, for which $S_2 < S < S_1$ and $K = -4S_1 - 3S_1^2$, is obtained in appendix A as

$$S = S_1 + \left[\frac{\cosh^2(\nu T)}{S_2'} - \frac{\sinh^2(\nu T)}{S_3'}\right]^{-1}, \tag{22}$$

where $S_{2,3}' = S_{2,3} - S_1$. This shows that $S \to S_1$ as $T \to \pm\infty$ and $S = S_2$ for $T = 0$. Similarly an expression for
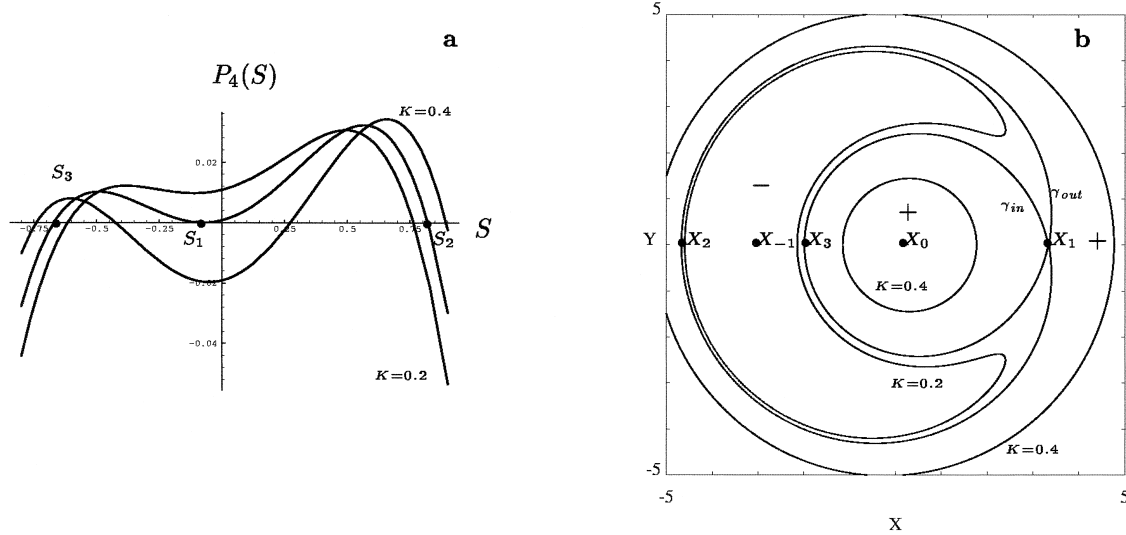
FIG. 7. (a) The quartic $P_4(S)$, Eq. (21), as a function of $S = [(X^2 + Y^2)/12 - 1]$, for $F = -1/2$ and three "energy" levels $K = 0.2$, 0.2832, $\cdots$, 0.4. The central K value corresponds to the presence of a double zero at S_1, which, in the X–Y phase-plane is related to the homoclinic orbits (separatrices) $\gamma_{in,out}$ emanating from the saddle at $(X_1, 0)$. (b) X–Y phase plane of modulation equations (11)–(12), for $F = -1/2$ and energy levels as in (a). The three zeros in (a), S_k for $k \in \{1, 2, 3\}$, in this plane correspond to circles that intersect the $Y = 0$ axis at X_k, and are related by $S_k = X_k^2/12 - 1$. These circles correspond to the maximum and minimum distances that points on the two homoclinic orbits, $\gamma_{in,out}$, have with respect to the "special" radius $R_s = 2\sqrt{3}$.

the inner separatrix can be obtained by simply interchanging S_2 and S_3, which has $S = S_3$ for $T = 0$ and which also approaches the saddle point $S \to S_1$ as $T \to \pm\infty$. With these expressions for $S(T)$, the trajectories $\gamma(T) = (X(T), Y(T))$ of the inner and outer homoclinic orbits follow from (20) and $Y(T) = \pm(R^2 - X^2)^{1/2}$; see Fig. 7b.

4. Forcing at two nearly resonant frequencies

The tidal forcing term $\zeta_e(t)$ is of course not a perfect cosine. Here we consider a somewhat more realistic double-frequency forcing:

$$\zeta_e(t) = f_1 \cos(\omega_1 t + \theta_1) + f_2 \cos(\omega_2 t + \theta_2). \quad (23)$$

The first component of this external tidal forcing might represent the semidiurnal lunar M_2 tide (dimensionally, $\omega_1 = 1.4056 \times 10^{-4}$ s^{-1}); the second component the semidiurnal solar S_2 tide ($\omega_2 = 1.4544 \times 10^{-4}$ s^{-1}). Of course this means that $|\omega_1 - \omega_2|/\omega_1 = 3.47 \times 10^{-2} \ll 1$. As a consequence we cannot apply the ideas for the analysis of weakly, quasiperiodically forced oscillators, developed in, for instance, Beigie et al. (1991). However, we will see that the closeness of the two frequencies ω_1 and ω_2 can provide an additional periodic forcing term in the modulation equations, on the slow timescale. This periodic forcing term will break up the homoclinic orbits of the modulation equations (in the integrable limit) and thus might lead to (transversely) intersecting stable and unstable manifolds for the Poincaré map associated to this new forcing term.

Thus, taking into account an additional solar, and therefore almost resonant, tidal forcing term gives a

mechanism for the construction of chaotic solutions (see also Yagasaki 1990, 1993).

In section 3 we defined the small parameter ε as a detuning. Analogously, we set here (now again employing nondimensional frequencies),

$$\varepsilon^2 = \omega_1 - 1. \quad (24)$$

The above-described mechanism can be constructed if we assume that $\omega_1 - \omega_2 = O(\varepsilon^2)$; that is, if

$$\Delta\omega = \omega_2 - \omega_1 \equiv \varepsilon^2 \Delta\Omega \quad (25)$$

and $\Delta\Omega = O(1)$. The forcing amplitude and the phase now vary on the slow timescale $T = \varepsilon^2 t$:

$$\zeta_e = f(T) \cos[\omega_1 t + \theta(T)], \quad (26)$$

where the combined amplitude and phase are defined by

$$f^2(T) = f_1^2 + f_2^2 + 2f_1 f_2 \cos(\Delta\Omega T + \Delta\theta) \quad \text{and} \quad (27)$$

$$\theta(T) = \theta_1 + \tan^{-1}\left[\frac{f_2 \sin(\Delta\Omega T + \Delta\theta)}{f_1 + f_2 \cos(\Delta\Omega T + \Delta\theta)}\right] \quad (28)$$

with

$$\Delta\theta = \theta_2 - \theta_1. \quad (29)$$

The derivation of the modulation equations is now identical to the derivation given for the single-frequency case in section 3 (now one averages over $2\pi/\omega_1$ instead of $2\pi/\omega$) and the modulation equations are again of the form (9)–(10), or, in Cartesian coordinates, (11)–(12), but now with slowly modulating forcing $F(T) = f(T)/\varepsilon^3$ and phase $\theta(T)$, Eqs. (27) and (28) replacing their constant counterparts (see Nayfeh and Mook 1979, sec-

tion 4.4). The right-hand sides of the modulation equations (11)–(12) now depend on the slow time explicitly, and, in principle, this offers the opportunity for the occurrence of aperiodic solutions.

When the two components are of equal magnitude, $f_2 = f_1$, the phase is simply $\theta = (\Delta\Omega\, T + \theta_1 + \theta_2)/2$, which redefines the effective frequency to be equal to the *average* frequency. The amplitude gradually changes sign over a long timescale $f = 2f_1 \cos[(\Delta\Omega\, T + \Delta\theta)/2]$. The block version of this, $f = 4f_1(1/2 - \Theta\{\cos[(\Delta\Omega\, T + \Delta\theta)/2]\})$, is piecewise integrable. Here $\Theta(x) = 0$ (1) for $x < 0$ (>0) denotes the Heaviside function. It has a saddle point on the X axis, whose position alternates with respect to the center of the X–Y phase plane with subsequent phases of the forcing. It is thus reminiscent of Aref's (1984) blinking vortex with its ensuing chaos. This suggests similar results for a sinusoidal forcing, which is confirmed by numerical analysis (see section 4b), although this cannot be substantiated analytically.

When the second component is but a small perturbation with respect to the first one $f_2/f_1 \equiv \delta \ll 1$, then $f \approx f_1[1 + \delta \cos(\Delta\Omega\, T + \Delta\theta)]$ and $\theta \approx \theta_1 + \delta \sin(\Delta\Omega\, T + \Delta\theta)$. Note that this occurs quite naturally in the context of M_2 and S_2 tides. In contrast to the previous case, this limit is more amendable to further analysis and will be addressed now. Without loss of generality we set $\theta_1 = 0$, fixing the origin of the "fast" time t, and $\Delta\theta = -\pi/2$, fixing the origin of the "slow" time T, whence the amplitude $f = f_1[1 + \delta \sin(\Delta\Omega\, T)]$ and phase $\theta = -\delta \cos(\Delta\Omega\, T)$, so that, with $f = \varepsilon^3 F$, for small values of δ, apart from a factor 2, the forcing terms in (11) and (12) are approximated as $F \sin\theta \approx -\delta F \cos(\Delta\Omega\, T)$ and $F \cos\theta \approx F[1 + \delta \sin(\Delta\Omega\, T)]$, respectively.

a. Perturbative second forcing component and Melnikov analysis

With a weak perturbative second forcing component ($\delta \ll 1$) and weak damping ($C \to \delta C$) the averaged Eqs. (11) and (12) become, after an appropriate shift in the origin of the slow and fast time,

$$
\begin{cases}
\dot{X} = \dfrac{\partial H}{\partial Y} + \delta\left(-\dfrac{C}{2}X - \dfrac{F}{2}\cos\Delta\Omega T\right), \\[2mm]
\dot{Y} = -\dfrac{\partial H}{\partial X} + \delta\left(-\dfrac{C}{2}Y - \dfrac{F}{2}\sin\Delta\Omega T\right),
\end{cases} \tag{30}
$$

with H as in (15). These equations can be written symbolically as

$$
\dot{\mathbf{X}} = \mathbf{g}(\mathbf{X}) + \delta\mathbf{h}(\mathbf{X}, T). \tag{31}
$$

Here a dot denotes differentiation with respect to the slow time T. Note that it is assumed implicitly that $0 < \varepsilon \ll \delta$. One has to include higher-order terms in ε and perform a much more detailed perturbation analysis

when $\delta = O(\varepsilon)$ [see Guckenheimer and Holmes (1983); one has to be careful with applying Melnikovs method in the averaged system in that case; see also DKM for a detailed discussion]. System (31) has a hyperbolic periodic orbit, with period $2\pi/\Delta\Omega$, that is $O(\delta)$ close to the saddle point $P_0 = (X_1, 0)$ of the unperturbed ($\delta = 0$) problem (see Fig. 7b), or, equivalently, the stroboscopic or Poincaré map P^{T_0} associated with (31) has a fixed point $P_\delta^{T_0}$ of saddle type $O(\delta)$ close to P_0 (see also Guckenheimer and Holmes 1983, chapter 4, for more details). Time T_0 defines which section is taken to construct this map. The stable and unstable manifolds of $P_\delta^{T_0}$ are also $O(\delta)$ close to those of P_0 in the unperturbed system. Thus, the perturbation will in general break open both homoclinic loops of the unperturbed system.

Neglecting for the moment the additional forcing term [last term in Eq. (30)], the effect of the damping is to draw trajectories inward across separatrices. Multiplying (30) with ∇H, the rate of change of the Hamiltonian is then obtained, on the separatrix (where $K = -3S_1^2 - 4S_1$), as $dH/dt = -3C(S - S_1)(3(S + S_1) + 4)/2$. Since on γ_{out}, $S \geq S_1$, it has $dH/dt \leq 0$, while on γ_{in}, the inner separatrix, $dH/dt \geq 0$, because $S_3 \leq S \leq S_1$ and $3(S_3 + S_1) + 4 = 4 - 3[-8S_1(1 + S_1)]^{1/2} \geq 0$, for $-2/3 \leq S_1 \leq 0$. Since outside γ_{out} the Hamiltonian has relatively large (K) values (see Fig. 7b) and since its variation is strictly negative ($dH/dt < 0$), this implies that all states outside it are drawn inward. Since the Hamiltonian is again relatively large inside γ_{in}, any trajectories that reach up to that separatrix are further drawn inward, to the choked state. All other states end up in the banana-shaped area in between γ_{in} and γ_{out}, approaching the amplified state (with relatively low values of the Hamiltonian K). With the additional forcing term, multiplication of (30) with ∇H gives the general expression for the rate of change of the energy (i.e., that of the perturbed Hamiltonian, dH/dt). If we evaluate that at the separatrices, we specify $X(T)$ and $Y(T)$ to satisfy the parametric relation corresponding to the particular separatrix (section 3c). The net change of the Hamiltonian, following this trajectory over one period (which is infinite on the separatrices), then amounts to the integral given below. Its vanishing would imply that the net change in energy is zero, and hence forcing and damping are in equilibrium and the state does not drift, representing a new, dynamic equilibrium.

A geometric interpretation of the Melnikov method is that it measures the splitting distance between the stable and unstable manifolds of $P_\delta^{T_0}$ that used to be part of a homoclinic loop to P_0 in the unperturbed case. In this case there are two Melnikov functions: one that measures the splitting distance in the perturbed, outer loop, $M_{\text{out}}(T_0)$, and one for the inner loop, $M_{\text{in}}(T_0)$. In both situations this distance is for instance measured at the intersections of the stable/unstable manifolds with the $Y = 0$ axis; Fig. 7b. Following, once again Guckenheimer and Holmes (1983, chapter 4) we see that $M_{\text{out,in}}(T_0)$ are, at leading order, given by
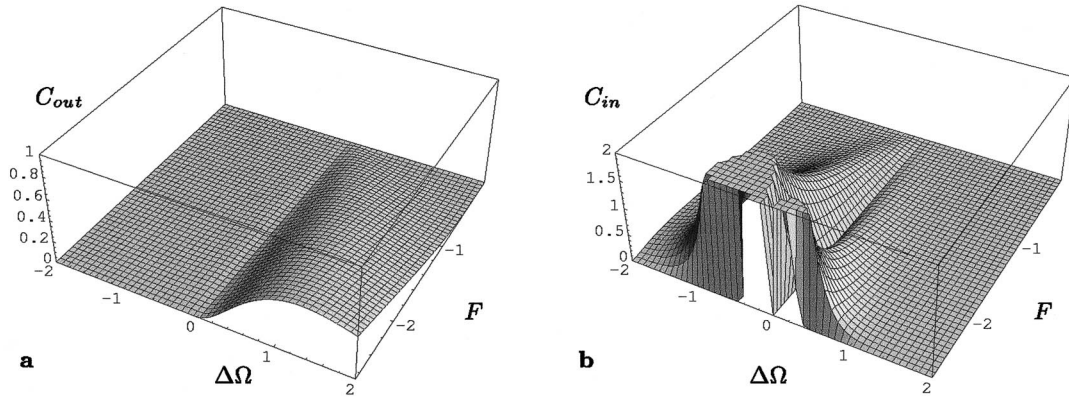
FIG. 8. Surfaces of critical damping as a function of frequency difference $\Delta\Omega$ and forcing amplitude $F \in (-8/3, 0)$ along (a) the outer and (b) the inner homoclinic orbit. For damping values below these surfaces, chaotic states exist in the modulation equations.

$$M_{\text{out,in}}(T_0) = \int_{-\infty}^{\infty} \{g_1[\gamma_0(T - T_0)]h_2[\gamma_0(T - T_0)]$$

$$- g_2[\gamma_0(T - T_0)]h_1[\gamma_0(T - T_0)]\} \, dT, \tag{32}$$

where $\gamma_0(T) = (X_0(T), Y_0(T))$ is either the outer or the inner homoclinic orbit of the unperturbed system (see section 3c); $\mathbf{g} = (g_1, g_2)$ and $\mathbf{h} = (h_1, h_2)$ are defined by (30) and (31).

A (nondegenerate) zero of, for example, $M_{\text{out}}(T_0)$ corresponds to a transversal intersection of the "outer" stable and unstable manifolds of $P_\delta^{T_0}$ for the map P^{T_0}. It is a standard procedure to associate a Smale horseshoe map to this situation (Guckenheimer and Holmes 1983). Thus, it follows that system (31) has chaotic solutions when either $M_{\text{out}}(T_0)$ or $M_{\text{in}}(T_0)$ has zeros. Note, however, that the existence of chaotic solutions does not automatically imply the existence of a chaotic *attractor*; see section 4b.

In this paper we only consider the "classical way" to construct chaotic solutions, namely due to intersections of either the inner stable and unstable homoclinic orbits, or the outer ones. Other possibilities, like the intersection of the *outer* unstable manifold of $P_\delta^{T_0}$ with its *inner* stable manifold, are discussed elsewhere (DKM).

For the outer homoclinic orbit, Eqs. (30), (31), and the definition of the unperturbed homoclinic orbit $\gamma_0(T)$ $= (X_0(T), Y_0(T))$, the Melnikov function (32) is calculated in appendix B. It is obtained as

$$M_{\text{out}}(T_0) = -24\left[C\left(\psi - \frac{3 \tan\psi}{3 + \tan^2\psi}\right)\right.$$

$$\left. + \pi(\Delta\Omega)^2 \cos(\Delta\Omega T_0)\frac{e^{\psi k}}{\sinh\pi k}\right], \tag{33}$$

where auxiliary variables ψ and k are related to forcing F and frequency difference $\Delta\Omega$ by $F = -8\sqrt{3} \cos^2\psi/$

$(1 + 2 \cos^2\psi)^{3/2}$ and $k = -\Delta\Omega(3 + \tan^2\psi)/2 \tan^2\psi$, where $F \in (-8/3, 0)$ for $\psi \in (\pi/2, \pi)$. The derivation of the equivalent of $M_{\text{out}}(T_0)$ for the *inner* homoclinic orbit, $M_{\text{in}}(T_0)$, is completely similar (see appendix B). There is a critical value of the damping $C = C_{\text{out}} = C_{\text{out}}(\Delta\Omega, \psi) = C_{\text{out}}(\Delta\Omega, F)$ for which the Melnikov function (33) vanishes, thus implying that there will be transverse homoclinic intersections for the *outer* orbit and chaotic solutions, when

$$C \leq C_{\text{out}} = \pi(\Delta\Omega)^2 \left|\frac{e^{\psi k}}{\sinh\pi k}\left(\psi - \frac{3 \tan\psi}{3 + \tan^2\psi}\right)^{-1}\right|, \tag{34}$$

depicted in parameter space in Fig. 8a. A similar critical surface C_{in} is obtained in appendix B for the inner orbit; see Fig. 8b.

Note that both C_{out} and C_{in} vanish when the two basic frequencies are equal (i.e., $\Delta\Omega = 0$). In this case both $M_{\text{out}}(T_0)$ and $M_{\text{in}}(T_0)$ are equal to M_1 (for definition of $M_{1,2,3}$ see appendix B), which does not depend on T_0, and thus, in this case there can be no (transversal) intersections of the stable and unstable manifolds for the Poincaré map P^{T_0}. As a consequence there can also be no chaotic solutions. This has been observed earlier in the monofrequency case to which it reduces. The same is true for tidal frequencies that are not both close to the same resonance or, in other words, when $|\Delta\Omega|$ is too large. This follows from (34) and its inner equivalent since the denominator $\sinh\pi k = \sinh[\pi\Delta\Omega/(2\nu)]$ dominates the numerator when $|\Delta\Omega| \gg 1$ (the decay rate decreases to zero when $\psi \uparrow \pi$ ($F \to -8/3$ and $\Delta\Omega > 0$; see Fig. 8a and below). Here an upward (downward) directed arrow implies the limit is approached from below (above). For $\Delta\Omega < 0$ the critical damping curve of the outer orbit is very small, $O(0.01)$, and hence not visibly different from zero on the left-hand side of Fig. 8a.

There are two other physically relevant limits: $F \uparrow 0$ (i.e., $\psi \downarrow \pi/2$ and $\phi \uparrow \pi/2$) and $F \downarrow -8/3$ ($\psi \uparrow \pi$ and $\phi \downarrow 0$). In the first case ("no forcing") one expects once
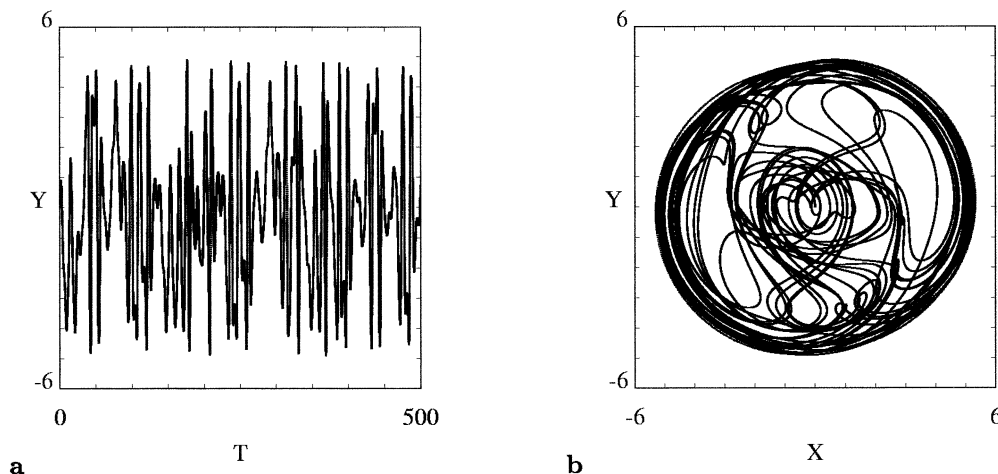
FIG. 9. Numerical integration of modulation equations (11)–(12) when $\delta C = 0.1$ gives an indication of chaos as observed here in current velocity Y vs (a) slow time T and (b) vs excess volume X. The integration is started from the origin ($X(0) = Y(0) = 0$) for forcing $F = 2 \cos\Omega't$, and frequency $\Omega' = \Delta\Omega/2 = 0.5$.

again the disappearance of (possibly) chaotic solutions. Indeed we find that

$$M_{\text{out}}(T_0) \to -12C\pi \quad \text{and} \quad M_{\text{in}}(T_0) \to 12C\pi,$$

when $F \uparrow 0$ [by (B9) and (B11): M_2 and M_3 vanish exponentially fast]. In a sense $M_{\text{out}}(T_0)$ and $M_{\text{in}}(T_0)$ measure the same splitting distance in this case since the outer and inner homoclinic orbits merge in this limit. The difference in sign is explained by the fact that the stable manifold of the inner orbit merges with the unstable manifold of the outer orbit (and vice versa). A similar geometrical point of view gives insight in the limit $F \downarrow -8/3$. In that case the inner homoclinic orbit shrinks to a point, while the outer orbit reaches a well-defined limit. Thus, nothing "singular" will happen to the outer orbit in this limit (see Fig. 8a). However, for the inner orbit, all components $M_{1,2,3}$ will approach 0 as $\phi \downarrow 0$. The behavior of M_1 is quite degenerate: $M_1 = O(\phi^5)$ for small ϕ, but both M_2 and M_3 will approach 0 with an exponential decay rate, for a *fixed* value of $\Delta\Omega$. Thus, we conclude that $C_{\text{in}} \downarrow 0$ for $F \downarrow -8/3$. However, the situation changes if we take the limit $\phi \downarrow 0$ differently, namely along any straight oblique line *through* the "origin" $(\Delta\Omega, \phi) = (0, 0)$: $\Delta\Omega = q\phi$, ($q \neq 0, \infty$), as the exponentially fast decay is then absent and the behavior of C_{in} is dominated by the singularity in it due to the ϕ-dependent factor in M_1. This implies that the critical damping magnitude may increase indefinitely in this limit. This singularity also dominates the magnitude of C_{in} farther away from the origin, in the interior of the $\Delta\Omega$–ϕ plane, and explains the two lobes visible in Fig. 8b on either side of the line $\Delta\Omega = 0$. It may be the reason why chaos can become important, even in quite strongly damped, realistic circumstances.

b. Numerical solutions of modulation and exact equations

Proving that the modulation equations possess a chaotic invariant set in the case when the Melnikov function vanishes is as far as we go analytically. As for the (closely related) Duffing equation, more work is needed to show that this set becomes attractive in some parameter range (Wiggins 1990, 612–613). In the following we will only give some numerical support that the (periodically forced) modulation equations possess such a chaotic (strange) attractor. This, in turn, suggests that the original, quasiperiodically forced Helmholtz oscillator should likewise have a strange attractor in the corresponding parameter regime; a result that we confirm by numerical integration below.

Numerical integration of the modulation equations, (11)–(12), with a fourth-order Runge–Kutta scheme, employing a double-frequency forcing (23), shows that for two nearby frequencies within the frequency range over which multiple equilibria exist, the solution curves appear to be chaotic (Fig. 9). This is suggested both in the (slow) time domain T, by the strait's current speed Y (Fig. 9a), as well as in the phase space of current velocity Y versus excess volume X (Fig. 9b). (Recall that the original **x** and averaged quantities **X** are related by $\mathbf{x} \propto \varepsilon\mathbf{X}$, while the fast t and slow T timescales are related by $t = \varepsilon^{-2}T$.) Here we take forcing amplitudes of the two external tidal components to be equal. One of these components is supposed to be at the (linear) resonance frequency, 1 (the Helmholtz frequency), while the second is at a slightly higher frequency (1.01). The latter frequency is setting the small parameter $\varepsilon = 0.1$. Within the resonance band the theory requires the forcing amplitudes to be small, $O(\varepsilon^3)$, so that we take them $f_1 = f_2 = 10^{-3}$. This leads, from (27), to $F(t) = f/\varepsilon^3 = 2 \cos[(\Delta\Omega\, T + \Delta\theta)/2]$ in the modulation equations (12). In striking contrast to the monofrequency case, where the Hamiltonian chaos disappears with the introduction of even the smallest amount of damping (see M), the present calculations show that chaos survives the addition of damping. Commensurate with the
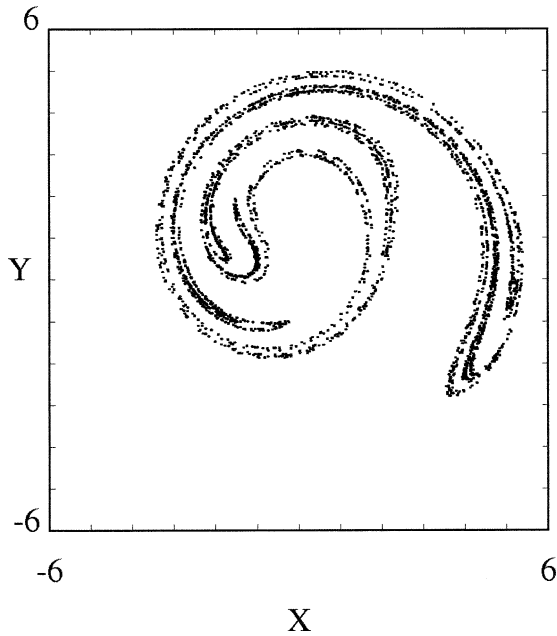
FIG. 10. Stroboscopic (Poincaré) plot of numerically integrated orbit of modulation equations, calculated as in Fig. 9, here extended to an integration over 3000 dimensionless units in time. Sampling is at the long period rate $2\pi/\Omega'$, where $\Omega' = \Delta\Omega/2$.

theoretical requirements we need to take $c \leq O(\varepsilon^2)$ and, in fact, take $c = \varepsilon^3$ here. This amounts to $\delta C = 0.1$. The erratic solution curves suggest that the modulation equations may possess a strange attractor, which is indeed revealed by making a Poincaré plot, by sampling at the modulation period (Fig. 10). Its shape is very similar to the "Japanese attractor" (see Ruelle 1980), which has been encountered for the forced and damped Duffing equation. Notice, in particular, the familiar Cantor-like division of the attractor's branches.

Chaos in the modulation equations suggests the occurrence of chaos in the original equations (5) under double-frequency forcing (23). In the following numerical experiments, the same values for forcing amplitudes and frequencies, as well as damping, were taken as were used in the previous numerical integration of the modulation equations. In order to compare the slow modulation of the numerically computed orbits directly to those obtained from the modulation equations, we sample the former at the tidal period. So, in Fig. 11 Poincaré plots of damped ($c = 0.001$) flow rate y versus time t (panel a) and versus excess volume x (panel b) are shown. Sampling is done once every average period $2\pi/\overline{\omega}$, where $\overline{\omega} = (\omega_1 + \omega_2)/2$. Both figures suggest that chaos is also present in the exact equations. Indeed, this is again made manifest by making a Poincaré plot, but now by sampling at the long-period timescale, $2\pi/\Delta\omega$. Figure 12 shows that it reveals the presence of a strange attractor, not unlike that encountered for the modulation equations; Fig. 10.

5. Forcing at two widely separate frequencies

Estimating the natural frequency of a number of basins occurring in nature leads to a wide range of periods varying from a few minutes to several hours, up to periods of the tides, with a predominancy in the range of several tens of minutes (Honda et al. 1908). For this reason it is useful to consider here the response of the nonlinear Helmholtz resonator to a forcing of the form

$$\zeta_e = a \cos\varepsilon^2 t + b \cos(\omega t + \theta)$$

$$\overset{\text{def}}{=} Z(\varepsilon^2 t) + b \cos(\omega t + \theta), \quad (35)$$

with $0 < b \ll a$: the first term represents the dominant, low-frequency tidal component and the second term a relatively weak perturbation that is almost in resonance
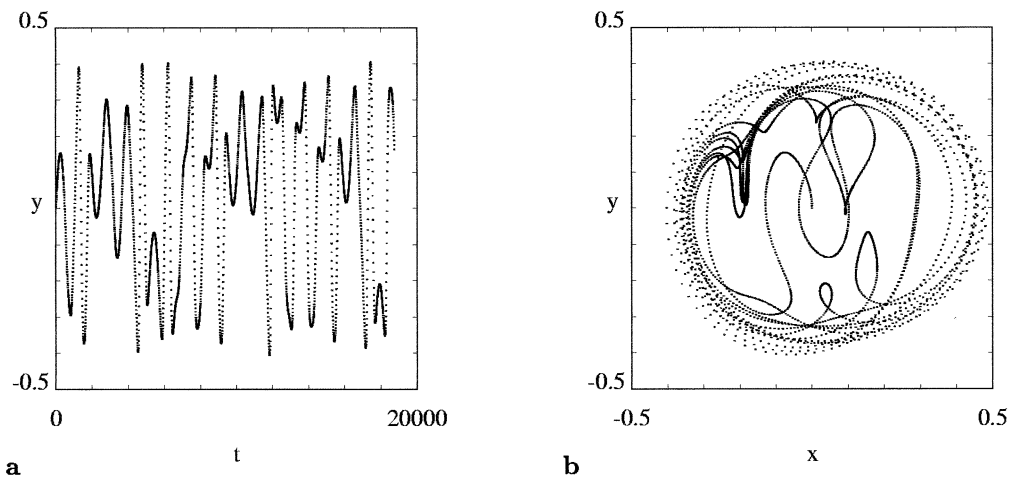

FIG. 11. Poincaré plot of numerically integrated exact equations (5) in case $c = 0.001$ gives an indication of chaos as observed here in current velocity y vs time t (a) and versus excess volume x (b). Integration is started from the origin ($x(0) = y(0) = 0$) for forcing values $f_1 = f_2 = 0.001$, and frequencies $\omega_1 = 1$ and $\omega_2 = 1.01$. Sampling is at the average forcing period $2\pi/\overline{\omega}$.
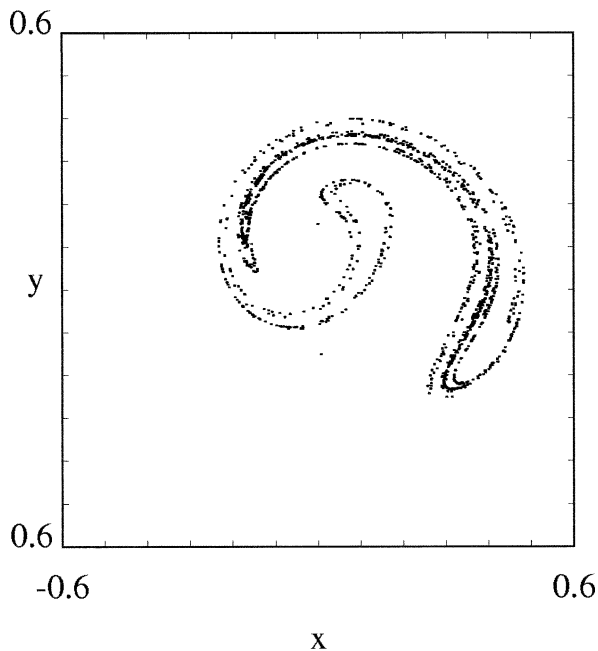
FIG. 12. Stroboscopic (Poincaré) plot of numerically integrated exact orbit calculated as in Fig. 11, here extended to an integration over 3000 dimensionless units. Sampling is at the (long) modulation period $2\pi/\Delta\omega$.

with the Helmholtz frequency of the basin ($\omega \approx 1$). Thus, in this setting, $\varepsilon^2 \ll 1$ measures the ratio of tidal to Helmholtz frequency. Note that we have for simplicity neglected all "intermediate" Fourier components of the external tide: the main purpose of this section is to show that chaos can occur in basins with a high Helmholtz frequency. The mechanism is quite similar to the one described and studied in the previous sections. We will first discuss the effect of such a forcing in qualitative terms and continue this with a more detailed treatment.

Consider first the case in which the perturbation is absent altogether, $b = 0$. Then, when damping is relatively weak because the frequency of the tide $\varepsilon^2 \ll 1$, the elevation within the basin will be able to follow the external tide: $\zeta \approx Z$. Therefore, in the presence of a high-frequency perturbation ($0 \neq b \ll a$), on the fast timescale the dominant component of the tide, Z, acts as a slow, quasi-adiabatic change of the mean depth of the basin. However, the basin depth (among other parameters) determines the Helmholtz frequency; see (1). Therefore, the frequency of the perturbation, assumed to be fixed in absolute measure, is slowly varying when scaled with this Helmholtz frequency. This will provoke the quasi-stationary response to drift along the response curve (see Fig. 5). When this frequency drift, induced by the apparent modulation of water depth, is large enough—covering the frequency range over which multiple equilibria exist, this will lead to a sequence of "catastrophes": rapid changes in the range of the (near) resonant, high-frequency response, which is subject to

periodical collapse and expansion. There will thus be a hysteretic change in the amplitude of the basin elevation. This change may, however, be chaotic because the moments in time at which the amplitude jumps can become randomized. This is perhaps so because the high-frequency response, if not rapidly constrained to a particular phase by strong damping, will still be in a transient adapting phase, recovering from the previous catastrophe, when it undergoes the next one.

It is ironic that this explanation, based on the possibility of undergoing a hysteretic curve requires fairly large, but not too large, damping magnitudes: if damping gets too small, the frequency range over which multiple equilibria exist becomes broader than the range of the modulating, detuned frequency, and the response is "stuck" to a particular branch; if, on the other hand, damping is too strong, the response curve has no multiple equilibria at all, again lacking the ability to shift branches.

In practice, the conditions under which chaos appears may become weaker because not only will equilibria shift location under slow variation of the mean depth, but so, and probably more importantly, will the corresponding domains of attraction. This is perhaps the reason for the apparently chaotic behavior found in Fig. 13 under a mild modulation of the forcing frequency. This figure shows the result of a numerical integration of (2) in phase space with a forcing of type (35). Both results of direct integration (Fig. 13a), as well as a sampling of these on the long, tidal timescale (Fig. 13b) are shown. The direct integration shows that the dominant resonant response on the short Helmholtz timescale vaccilates on the long timescale, between a small- and large-scale response. These small- and large-scale responses correspond to the amplified and choked regimes in the amplitude response diagram of Fig. 5 to which these solutions would tend in the absence of the tide ($a = 0$). Under modulation with a tide of amplitude $a = 3\varepsilon^2$ this provokes a modulation in apparent frequency σ, of approximately 1.5 dimensionless units, which is not big enough to span the entire frequency range over which multiple equilibria exist. Both this figure, as well as the tidally subsampled version in Fig. 13b (with period $2\pi/\varepsilon^2$), however, testify about the irregular nature of the response, the latter figure in particular bearing some resemblance to the earlier Poincaré plots. In order to verify these qualitative ideas one may proceed with a more quantitative analysis. For the sake of brievity, this will however just be sketched here.

Again consider a basin in which the area increases linearly with the depth so that, with (4) and (35), the evolution of the excess volume v is described by

$$\frac{d^2v}{dt^2} + \sqrt{1 + 2v} - [1 + Z(\varepsilon^2 t)]$$
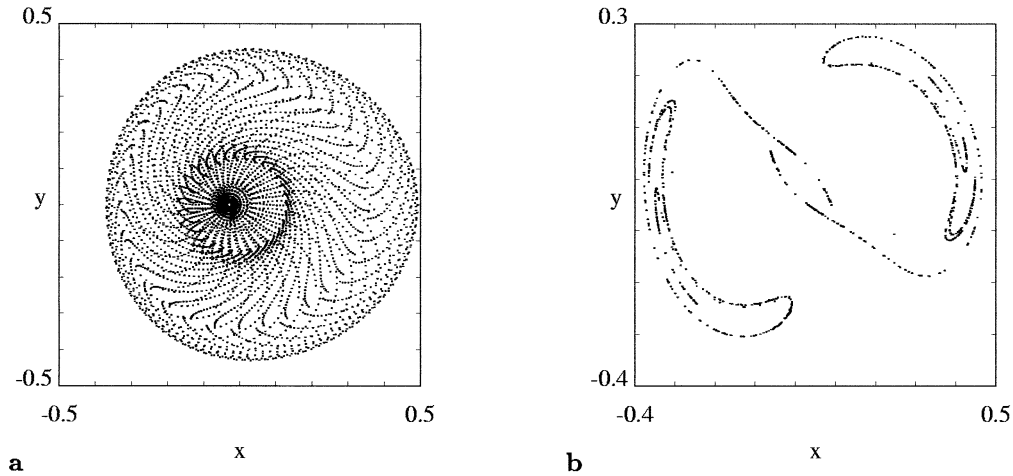
$$= b \cos(\omega t + \theta) - c\frac{dv}{dt} \qquad (36)$$

FIG. 13. Numerical integration of (2) for a forcing (35) that consists of a low-frequency tide and a resonant perturbation. (a) A trajectory in *x*–*y* phase space and (b) the corresponding Poincaré plot, subsampled on the tidal period. Parameters read $\varepsilon = 0.1$, $a = 3\varepsilon^2$, $b = 3\varepsilon^3$, $c = 0.4\varepsilon^2$, and $\omega = 1 + 1.5\varepsilon^2$ (and $\theta = 0$).

[see (2)]. The homogeneous problem ($b = c = 0$) has a "quasi-stationary" equilibrium at $v(\varepsilon^2 t) = Z(\varepsilon^2 t) + \frac{1}{2}Z(\varepsilon^2 t)^2$, so it is natural to make the following shift in v:

$$v = Z + \frac{1}{2}Z^2 + (1 + Z)^2\tilde{v}, \qquad (37)$$

which yields

$$\frac{d^2\tilde{v}}{dt^2} + \frac{1}{1 + Z}(\sqrt{1 + 2\tilde{v}} - 1)$$

$$= \frac{b}{(1 + Z)^2} \cos(\omega t + \theta) - c\frac{d\tilde{v}}{dt}$$

$$- \frac{\varepsilon^2 Z'}{1 + Z}\left(c + 4\frac{d\tilde{v}}{dt} + 2c\tilde{v}\right) + O(\varepsilon^4), \quad (38)$$

where $dZ/dt = \varepsilon^2 dZ/dT = \varepsilon^2 Z'$. As in section 3 we rescale $\tilde{v} = \varepsilon V$, $c = \varepsilon^2 C$, and $b = \varepsilon^3 B$ so that we can start the derivation of a modulation equation for V:

$$\frac{d^2 V}{dt^2} + \frac{1}{1 + Z}V$$

$$= \frac{\varepsilon}{2(1 + Z)}(V^2 - \varepsilon V^3) + \frac{\varepsilon^2 B}{(1 + Z)^2} \cos(\omega t + \theta)$$

$$- \varepsilon^2\left(C + \frac{4Z'}{1 + Z}\right)\frac{dV}{dt} + O(\varepsilon^3) \qquad (39)$$

[see (7)]. Thus, as predicted by the above qualitative arguments, the leading-order Helmholtz frequency $1/\sqrt{1 + Z(\varepsilon^2 t)}$ varies slowly and periodically in time $\varepsilon^2 t = T$. Since we intend to proceed along the lines of the preceeding sections, we need to assume that the perturbation $b \cos(\omega t + \theta)$ is almost in resonance with the Helmholtz frequency of the basin. Hence we write (again) $\omega = 1 + \varepsilon^2\sigma$ and assume that $Z = a \cos\varepsilon^2 t = \varepsilon^2 A \cos\varepsilon^2 t$, so that $1/\sqrt{1 + Z} = 1 - \frac{1}{2}\varepsilon^2 A \cos\varepsilon^2 t +$

$O(\varepsilon^4)$. This last assumption is especially motivated by the mathematical approach, but may be quite realistic in the case of the tides in fjords too. Note that the scalings of $a = \varepsilon^2 A$ and $b = \varepsilon^3 B$ are consistent with the assumption $b \ll a$.

As motivated by the above arguments we now introduce a new timescale $\tilde{t} = t/\sqrt{1 + Z} = t(1 - \frac{1}{2}\varepsilon^2 A \cos\varepsilon^2 t + O(\varepsilon^4))$, so that (39) indeed reduces to (7) with t replaced by $\tilde{t}$ and

$$\omega = \omega(\tilde{t}) \stackrel{\text{def}}{=} 1 + \varepsilon^2\sigma\alpha(\varepsilon^2\tilde{t})$$

$$= 1 + \varepsilon^2\sigma\left(\frac{A}{2\sigma} \cos\varepsilon^2\tilde{t}\right) \qquad (40)$$

at leading order on a timescale of $O(1/\varepsilon^2)$. The derivation of the modulation equations is now completely similar to that of (9)–(10) in section 3 (as in section 4) and results in (9)–(10) with σ replaced by the periodically varying function $\sigma\alpha(T)$. As in section 3 we set $\theta = 0$ and scale σ to 1 so that the modulation equations are now given by

$$\begin{cases} \dot{X} = \left(\dfrac{R^2}{12} - 1\right)Y - \left[\dfrac{C}{2}X + \alpha(T)Y\right] \\[4mm] \dot{Y} = -\left(\dfrac{R^2}{12} - 1\right)X + \dfrac{F}{2} - \left[\dfrac{C}{2}Y - \alpha(T)X\right]. \end{cases} \qquad (41)$$

The presence of the periodically oscillating term can indeed be interpreted as a drift along the response curve of Fig. 5. This system can again be analyzed by the Melnikov approach of section 4a if we assume that C, $\alpha = O(\delta)$ with $0 < \varepsilon \ll \delta \ll 1$. Note that in this case the drift along the response curve is also only of order δ. The earlier described emergence of chaos, qualitatively attributed to hysteretic changes, which can only exist for $\delta = O(1)$, is apparently overrestrictive, as a

simple "jitter" of the saddle point seems to be sufficient for the emergence of crossing of stable and unstable homoclinic orbits, and thus for the presence of a chaotic invariant set. For small δ the Melnikov analysis proceeds along the lines of section 4a. The only difference is that the perturbation term $\mathbf{h}(X, T)$ in (31) has changed. It is a straightforward procedure to write down the equivalents of (B1) and redo the calculations of section 4b for the "new" inner and outer Melnikov functions, by which equivalents of (33) can be obtained. Note that (by construction) the structures of $M_{\text{out,in}}(T_0)$ will be the same as in section 4a: $C \times$ (a constant term M_1) + (a T_0 periodic term); M_1 will even be exactly as in (33) and its inner equivalent. We do not present the details of the computations here. Our main observation is that one can thus again show the existence of a Smale horseshoe: it is the same mechanism as that of the quasiperiodically forced case of sections 3 and 4, which creates chaotic solutions to systems with two widely separated frequencies.

Several aspects of the numerical integration in Fig. 13 are still unrealistic. First, the response is dominated by the local resonance and thus the elevation shows a nearly periodic change on the *fast* timescale, instead of on the (long) tidal timescale. Second, the response is much too big, covering (almost) the entire basin depth. More work is therefore needed to bring the response closer to the ordering observed in nature, as for example, in Moldefjord (see Fig. 1), where Helmholtz response, $O(0.1 \text{ m}) \ll$ tidal range, $O(1 \text{ m}) \ll$ total depth, $O(80 \text{ m})$, while currents on the fast Helmholtz timescale are comparable to tidal currents. It should be noticed that the correct ordering can partly be obtained quite simply by choosing model parameters such that all fixed points of the modulation equations are close to the origin. With this choice the homoclinic orbits, and therefore the actual state of the oscillator, are close to the origin of the phase space. Hence, corresponding elevations will be small compared to water depth. Notwithstanding the remaining issues that need clarification, it is believed that the mechanism described in this section may act as a "building block" for the explanation of the appearance of chaotic tides in basins with high Helmholtz frequency, particularly when the tide is extended with additional tidal components and harmonics.

From a practical point of view, any observed persistence of observed secondary oscillations might point at a non-meteorological origin. When their (average) period compares to estimates of the Helmholtz frequency, the model advanced here may be appropriate. Additional support may be obtained by plotting observed elevation (volume) against a strait current. Particularly, a stroboscopic sampling thereof (at the Helmholtz and tidal periods) may elucidate the structure of the underlying attractor. By comparison to predictions from the present model, an estimate of the most elusive factor, the damping, may perhaps be inferred.

6. Summary

Bays and estuaries may co-oscillate with the tides in the adjacent seas (Defant 1961). The extent to which they do generally depends on the geometry of the basin. Two competing mechanisms exist that mainly determine the final state of the coastal tides: the proximity of one of the more prolific tidal frequencies to basin resonances and the amount of friction. Depending on the balance between forcing and damping the coastal tide may be amplified or choked. This twofold nature of the response may actually occur for one and the same set of geometric and frictional parameters, as was shown in the case of the Helmholtz resonator for an almost-enclosed, relatively deep tidal basin with sloping bottom (see M). The slope in the bottom provokes a nonlinear restoring term (Green 1992) that, in turn, is responsible for the occurrence of multiple (stable) equilibria near the (linear) resonance frequency. When forced at the entrance by a singlefrequency tide (located within the resonance band), depending on the initial state, the tide within the basin may either have a small or large range that, after the decay of transients, is stationary.

Here we have shown that the presence of a second tidal frequency component, also within (or near to) the resonance band, may, in contrast, lead to the occurrence of a strange (chaotic) attractor, even with relatively strong damping. The basin tide may thus exhibit a perpetual change, manifested by unstable estimates of the tidal "constants" (amplitude and phase), when based on the analysis of a finite length time series, which may be of relevance to the observation by Gutiérrez et al. (1981) that some tidal components appear "unresolvable." It may also be relevant to chaos in consecutive tidal maxima, observed in Venice Lagoon (Vittori 1992) or in coastal tides (Frison et al. 1999).

Surprisingly, there have been many observations of "irregular" secondary oscillations, accompanying the tide, which date back long before the notion of chaos appeared in the literature. Already in 1908, Honda et al. devoted an extensive study to ascertain the presence of both regular and irregular secondary oscillations in vertical elevation records in some 50, mostly Japanese, basins, usually based on observations covering just a few days. The nature of the basin response is typically, however, a superposition of the tide and a high-frequency oscillation, rendering the previous model, which predicts secondary oscillations on the tidal timescale, less applicable. Periods of these oscillations are observed to be in the range from several minutes up to several hours, depending on the size of the basin. Depending on the basin shape, this high-frequency oscillation can either be identified as the Helmholtz or quarter-wave resonance. The amplitude of the secondary oscillations in vertical elevation is, as the name suggests, observed to be an order of magnitude less than that of the (primary) tide. This may probably explain why, until fairly recently, the interest in secondary oscillations

seems to have waned, and why the topic is held as a curiosity. A recent observation by Golmen et al. (1994), in a Norwegian fjord connected to the sea by a narrow strait, has, however, put the possible relevance of secondary oscillations into a new perspective by noticing that the associated secondary, irregular currents through the connecting strait reach a magnitude *comparable* to that of the tidal current. The reason for this is that, whereas the tide is present both outside and (with some delay) inside of the strait, the secondary elevation is present basically only within the fjord. The elevation *differences,* responsible for the associated currents, may therefore, at any time be of similar magnitude when comparing tidal and secondary oscillations. Hence, also the corresponding currents may be expected to be equally important. The same process may perhaps explain similar highly irregular current observations through lagoon entrances by Kjerfve and Knoppers (1991) and by Smith (1994), and the observed preferential enhancement of current harmonics (relative to that in the elevation field) by Seim and Sneed (1988). The relevance of this observation is that secondary oscillations may greatly alter the corresponding currents and therefore seriously modify the transfer of matter and dissolved substances (the "flushing" of the bay or fjord).

As remarked, these observations require a different kind of model in which tidal and resonance frequency— in our present model the Helmholtz frequency—are widely disparate. By adding a perturbative, but resonant, second forcing term to the primary tidal forcing it was shown that modulation equations are obtained that, to some extent, mimic those derived for the previous case with two tidal frequencies. In both cases, the modulation equations are driven at the long timescale. In the case of two nearby and resonant tidal frequencies this forcing stems from the beat (difference) frequency. In the case when the tidal frequency is much less than that of the resonant perturbative forcing term the tide itself is directly providing a modulation on the long timescale. The presence of a chaotic invariant set that could be ascertained analytically in these case, unfortunately does not guarantee that this set is also attractive. Therefore further support for the presence of a strange attractor was offered by (stroboscopic) Poincaré plots stemming from the numerical integration of the nonlinear Helmholtz resonator, when appropriately forced. The presence of a chaotic response may show up in a numerical simulation (in a basin with sloping shoreline) in much the same way as it reveals itself in nature: by irregularity in the elevation and, particularly, current fields (either within one tidal period, or in between subsequent tidal periods), by broadening of spectral peaks (with poorly resolved phases), and by showing a sensitive dependence on the initial tidal phase (neighboring orbits evolving to different states). The implication is that irregularity in the current and elevation fields is predictable as a phenomenon, but not its exact development.

It requires further effort to quantify the resulting enhanced exchange between sea and basin.

Thus, despite the fact that, with the introduction of satellite-altimetry-derived tidal observations, tidal predictions have reached new unprecedented accuracies, it now seems that, at the same time, the tides may to some extent and in certain locations be intrinsically unpredictable, an unpredictability that reflects the nonlinear nature of the response of such locations.

APPENDIX A

Integrating Eq. (21)

In order for the right-hand of (21) to be real, the argument of the square root, the quartic

$$P_4(S) \equiv \frac{F^2}{12}(S + 1) - \frac{1}{4}(S^2 - K)^2, \qquad \text{(A1)}$$

needs to be positive somewhere, which sets a bound on K, whose particular value is of no relevance here. Then, there can either be one or two intervals over which the quartic is positive, corresponding to the cases that the quartic has two or four real zeros, respectively; see the upper and lower curves in Fig. 7a. For fixed forcing F, these curves have differing values of K. In the former case this value is relatively low; hence the corresponding region in the X–Y parameter plane, in between γ_{in} and γ_{out}, is labeled with a minus sign; see Fig. 7b. In the latter case, it has a relatively high value of K, and there are two corresponding curves in the X–Y plane, one inside γ_{in} and one outside γ_{out}, which regions are therefore labeled with a plus sign. These regions are separated from each other by separatrices $\gamma_{in,out}$ that emanate from the third fixed point: the saddle. These separatrices, or homoclinic orbits, are characterized by the fact that the two central zeros of the quartic coalesce, in which case (21) can be solved algebraically. Equating (A1) to the quartic $-(S - S_1)^2(S - S_2)(S - S_3)/4$, and requiring equivalence of each of the coefficients of the fourth-degree polynomial, leads to four equations from which $S_{1,2,3}$ and K can be determined. Because in the X–Y plane the central, double zero of the quartic corresponds with the saddle point $(X_1; 0)$, whose position is already determined [see (16)–(17)], we find that one of these equations is redundant. The remaining equations yield the two other zeros,

$$S_{2,3} = -S_1 \pm \left(-\frac{F^2}{6S_1} \right)^{1/2}, \qquad \text{(A2)}$$

and the energy level of the separatrix

$$K = S_1^2 - \frac{F^2}{12S_1} = -4S_1 - 3S_1^2, \qquad \text{(A3)}$$

where

$$S_1 = X_1^2/12 - 1 \qquad \text{(A4)}$$

$(Y_1 = 0)$ and $X_1 = 4\sin\nu_1$, where $\nu_1 = (\alpha + 2\pi)/3$ and $\alpha = -\sin^{-1}(3F/8)$. Over the range of F values considered, $-8/3 < F < 0$, one finds $-2/3 < S_1 < 0$.

Replacing the quartic in (21) with its expression in terms of zeros requires the solution of

$$\frac{dS}{dT} = \pm\frac{(S - S_1)}{2}[(S - S_3)(S_2 - S)]^{1/2}, \qquad \text{(A5)}$$

where we have assumed $S_3 < S < S_2$. Now S can still be on either side of S_1, which itself lies in between the lower and upper limits S_3 and S_2, respectively. By taking S either smaller or larger than S_1 we obtain the inner or outer separatrix, respectively.

For definiteness assume $S_1 \leq S \leq S_2$ so that we are constructing the outer separatrix. Introducing

$$S' = S - S_1 > 0, \qquad \text{(A6)}$$

and similarly $S_2' = S_2 - S_1 > 0$ and $S_3' = S_3 - S_1 < 0$, then

$$\frac{dS'}{dT} = \pm\frac{S'}{2}[(S' - S_3')(S_2' - S')]^{1/2}. \qquad \text{(A7)}$$

Introducing $\sigma = 1/S' > 0$ yields

$$\frac{d\sigma}{dT} = \mp\frac{1}{2}[(1 - S_3'\sigma)(S_2'\sigma - 1)]^{1/2}. \qquad \text{(A8)}$$

Multiplying by S_2' and defining $\sigma' = S_2'\sigma > 1$ and $\eta = -S_2'/S_3' > 0$ gives

$$\frac{d\sigma'}{dT} = \mp\frac{(-S_2'S_3')^{1/2}}{2}[(\eta + \sigma')(\sigma' - 1)]^{1/2}. \qquad \text{(A9)}$$

Let $\sigma' = 1 + P^2$ and $\eta = Q^2 - 1$ (with $Q^2 > 1$), then

$$\frac{dP}{dT} = \mp\nu(P^2 + Q^2)^{1/2}, \qquad \text{(A10)}$$

with $\nu \equiv (-S_2'S_3')^{1/2}/4$. Hence, $P = \pm Q\sinh\nu T$, or, following back the substitutions, we find that (21) is solved by (22).

APPENDIX B

Computing the Melnikov Function

With $\mathbf{g}$, $\mathbf{h}$ from (30) and (31), the Melnikov function (32) is, for the outer homoclinic orbit, defined as

$$M_{\text{out}}(T_0) = -\frac{C}{2}\int_{-\infty}^{\infty} R_0^2\left(\frac{R_0^2}{12} - 1\right) dT + \frac{FC}{4}\int_{-\infty}^{\infty} X_0\, dT + \frac{F}{2}\int_{-\infty}^{\infty} Y_0\left(\frac{R_0^2}{12} - 1\right)\sin\Delta\Omega(T + T_0)\, dT$$

$$+ \int_{-\infty}^{\infty}\left[\frac{F^2}{4} - \frac{FX_0}{2}\left(\frac{R_0^2}{12} - 1\right)\right]\cos\Delta\Omega(T + T_0)\, dT, \qquad \text{(B1)}$$

with a suffix 0 referring to spatial coordinates along an unperturbed homoclinic orbit. Since (18)–(20) apply in general, $S_0 = R_0^2/12 - 1$, $FX_0/6 = S_0^2 - K$, and $FY_0/12 = dS_0/dT$. Thus, along these orbits, the Melnikov function simplifies to

$$M_{\text{out}}(T_0) = -\frac{3C}{2}\int_{-\infty}^{\infty} (3S_0^2 + 4S_0 + K)\, dT + 3\int_{-\infty}^{\infty}\frac{dS_0^2}{dT}\sin[\Delta\Omega(T + T_0)]\, dT$$

$$+ 6\int_{-\infty}^{\infty}\frac{d^2S_0}{dT^2}\cos[\Delta\Omega(T + T_0)]\, dT. \qquad \text{(B2)}$$

We will denote the three integrals in (B2) as $M_1(T_0)$, $M_2(T_0)$, and $M_3(T_0)$, respectively. Note that S_0 approaches the saddle S_1 for $T \to \pm\infty$ along the separatrices, so that the integrands vanish in these limits (recall that $K = -3S_1^2 - 4S_1$). We now define

$$J_n(\Delta\Omega) = \int_{-\infty}^{\infty}\cos(\Delta\Omega T)[S_0(T) - S_1]^n\, dT, \qquad \text{(B3)}$$

so that we can rewrite the $M_i(T_0)$:

$$M_1(T_0) = \frac{3}{2}C[(4 + 6S_1)J_1(0) + 3J_2(0)]$$

$$M_2(T_0) = -3\Delta\Omega\cos(\Delta\Omega T_0)[2S_1 J_1(\Delta\Omega) + J_2(\Delta\Omega)]$$

$$M_3(T_0) = -6(\Delta\Omega)^2\cos(\Delta\Omega T_0)J_1(\Delta\Omega). \qquad \text{(B4)}$$

The expressions for $M_{2,3}(T_0)$ were obtained by expanding $\cos\Delta\Omega(T + T_0)$ and $\sin\Delta\Omega(T + T_0)$ and by partial integration; all integrands involving $\sin(\Delta\Omega T_0)$ "average out" due to the odd/even symmetries of $S_0(T)$ and its derivative. The J_n integrals can be computed explic-

itly by expressing them in an integral $I(a, k)$ that can be evaluated by a contour integral in the complex plane (see appendix C):

$$I(a, k) = \int_0^\infty \frac{\cos kx}{a + \cosh x}\, dx = \frac{\pi}{\sin\phi}\frac{\sinh k\phi}{\sinh k\pi}, \quad (B5)$$

where $a = \cos\phi$. For $J_1(\Delta\Omega)$ this is a rather straightforward procedure, since, by section 3c

$$S_0(T) - S_1 = \frac{1}{\cosh^2 \nu T/S_2' - \sinh^2 \nu T/S_3'}$$

$$\stackrel{\text{def}}{=} \frac{A}{a + \cosh 2\nu T}, \quad (B6)$$

where A, a, and ν can be expressed, in terms of S_2' and S_3', as

$$A = \frac{2S_2'S_3'}{S_3' - S_2'}, \quad a = \frac{S_3' + S_2'}{S_3' - S_2'}, \quad \nu = \frac{(-S_2'S_3')^{1/2}}{4}. \quad (B7)$$

It follows that $-1 < a < 0$, and therefore we introduce ψ by

$$a = \cos\psi \quad \text{with } \psi \in \left(\frac{\pi}{2}, \pi\right) \quad (B8)$$

and hence, from (43),

$$S_1 = -\frac{2}{3 + \tan^2\psi}, \quad \nu = -\frac{\tan\psi}{3 + \tan^2\psi},$$

$$A = 4\nu\sqrt{1 - a^2} \quad (B9)$$

so that

$$J_1(\Delta\Omega) = 4\sqrt{1 - a^2}\,I(a, k) = 4\pi\frac{\sinh\psi k}{\sinh\pi k}, \quad (B10)$$

where

$$k \stackrel{\text{def}}{=} \frac{\Delta\Omega}{2\nu}. \quad (B11)$$

Note that the limit $\Delta\Omega \to 0$, or $k \to 0$, is well defined. Integral $J_2(\Delta\Omega)$ can be obtained by taking the a derivative of $I(a, k)$:

$$J_2(\Delta\Omega) = -16\nu(1 - a^2)\frac{\partial I(a, k)}{\partial a}$$

$$= 16\pi\nu\frac{\sinh\psi k}{\sinh\pi k}\left(\frac{k}{\tanh\psi k} - \frac{1}{\tan\psi}\right). \quad (B12)$$

Once again, the limit $k \to 0$ is well defined. Compiling expressions, the Melnikov function for the outer homoclinic orbit, $M_{\text{out}}(T_0) = M_1 + M_2 + M_3$, is given by (33). The derivation of the equivalent of $M_{\text{out}}(T_0)$ for the *inner* homoclinic orbit, $M_{\text{in}}(T_0)$, is similar to the above analysis. The main difference is that the roles of indices 2 and 3 in the description of the homoclinic

orbit [(B6) and section 3c] are exchanged and thus that $S_0 < S_1$. This amounts to a replacement in all formulas of $\psi \to -\phi$, where $\phi \in (0, \pi/2)$, which is significant, because it now opens the possibility of a singularity in the surface of critical damping; Fig. 8b.

APPENDIX C

Evaluation of an Integral

In order to evaluate Melnikov's distance function, integral $I(a, k)$, Eq. (B5), needs to be obtained for $0 < |a| < 1$. Although this integral can be found in the literature (Prudnikov et al. 1986, p. 470) we here give a short sketch of its derivation by writing it as

$$I(a, k) = \int_{-\infty}^\infty \frac{e^{ikx}}{e^x + 2a + e^{-x}}\, dx$$

$$= \int_0^\infty \frac{p^{ik}}{p^2 + 2ap + 1}\, dp \quad (C1)$$

for $p = \exp(x)$. Hence, by separating the denominator,

$$I(a, k) = \frac{1}{p_+ - p_-}\int_0^\infty \left(\frac{1}{p + p_-} - \frac{1}{p + p_+}\right)p^{ik}\, dp, \quad (C2)$$

where, with $a \equiv \cos\phi$, $p_\pm = \exp(\pm i\phi)$. Recasting the integration in terms of x, the integral reads

$$I(\phi, k) = \frac{1}{2i\sin\phi}\left(\int_{-\infty}^\infty \frac{e^{ikx}}{1 + e^{-i\phi-x}}\, dx\right.$$

$$\left. - \int_{-\infty}^\infty \frac{e^{ikx}}{1 + e^{i\phi-x}}\, dx\right). \quad (C3)$$

For definiteness assume $k > 0$. Then, by extending the integral over the real axis to a complex integration in the upper half-plane over a region enclosed also by a semicircle of infinite radius, this integral can be evaluated by means of Cauchy's theorem. It determines it as $2\pi i$ times the sum of the residues at the poles $x_\pm = \pm i\phi + (2n + 1)\pi i$, for $n = 0, 1, 2, \cdots$. Here, minus and plus signs refer to the poles appearing in the first and second integral, respectively. This yields

$$I(\phi, k) = \frac{\pi}{\sin\phi}(e^{k\phi}S_\infty - e^{-k\phi}S_\infty), \quad (C4)$$

where $S_N = \sum_{n=0}^N q^{2n+1}$ with $q \equiv \exp(-\pi k)$, whence $S_\infty = q/(1 - q^2) = 1/(2 \sinh\pi k)$, as $0 < q < 1$. Therefore the expression in (B5) follows. Note that this expression holds also when $k = 0$.

REFERENCES

Andersen, O. B., P. L. Woodworth, and R. A. Flather, 1995: Intercomparison of recent ocean tide models. *J. Geophys. Res.,* **100,** 25 261–25 282.

Aref, H., 1984: Stirring by chaotic advection. *J. Fluid Mech.,* **143,** 1–21.

Beigie, D., A. Leonard, and S. Wiggins, 1991: Chaotic transport in the homoclinic and heteroclinic tangle regions of quasiperiodically forced two-dimensional systems. *Nonlinearity,* **4,** 775–819.

Cartwright, D. E., 1977: Oceanic tides. *Rep. Prog. Phys.,* **40,** 665.

Davies, A. M., 1986: A three-dimensional model of the northwest European continental shelf, with application to the M_4 tide. *J. Phys. Oceanogr.,* **16,** 797–813.

Defant, A., 1961: *Physical Oceanography.* Vol. II, Pergamon, 598 pp.

Frison, T. W., H. D. I. Abarbanel, M. D. Earle, J. R. Schultz, and W. Scherer, 1999: Chaos and predictability in ocean water levels. *J. Geophys. Res.,* **104,** 7935–7951.

Garrett, C., 1975: Tides in gulfs. *Deep-Sea Res.,* **22,** 23–35.

Godin, G., 1991: The analysis of tides and currents. *Tidal Hydrodynamics,* B. B. Parker, Ed., Wiley and Sons, 675–709.

Golmen, L. G., J. Molvaer, and J. Magnusson, 1994: Sea level oscillations with super-tidal frequencies in a coastal embayment of western Norway. *Cont. Shelf Res.,* **14,** 1439–1454.

Green, T., 1992: Liquid oscillations in a basin with varying surface area. *Phys. Fluids A,* **4,** 630–632.

Guckenheimer, J., and P. Holmes, 1983: *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.* Springer-Verlag, 459 pp.

Gutiérrez, A., F. Mosetti, and N. Purga, 1981: On the indetermination of the tidal harmonic constants. *Nuovo Cimento,* **4,** 563–575.

Honda, K., T. Terada, Y. Yoshida, and D. Isitani, 1908: Secondary undulations of oceanic tides. *J. Coll. Sc. Imp. Univ. Tokyo,* **24,** 1–113 and 95 plates.

Kjerfve, B., and B. A. Knoppers, 1991: Tidal choking in a coastal lagoon. *Tidal Hydrodynamics,* B. B. Parker, Ed., Wiley and Sons, 169–182.

Krümmel, O., 1911: *Handbuch der Ozeanographie.* Verlag von J. Engelhorns, 766 pp.

LeBlond, P. H., 1991: Tides and their interactions with other oceanographic phenomena in shallow water. *Tidal Hydrodynamics,* B. B. Parker, Ed., Wiley and Sons, 675–709.

Lynch, D. R., and F. E. Werner, 1991: Three-dimensional velocities from a finite-element model of English Channel/Southern Bight tides. *Tidal Hydrodynamics,* B. B. Parker, Ed., Wiley and Sons, 183–200.

Maas, L. R. M., 1997: On the nonlinear Helmholtz response of almost-enclosed tidal basins with a sloping bottom. *J. Fluid Mech.,* **349,** 361–380.

——, 1998: On an oscillator equation for tides in almost enclosed basins of non-uniform depth. *Physics of Estuaries and Coastal Seas,* J. Dronkers, and M. Scheffers, Eds., A. A. Balkema, 127–132.

Munk, W., and C. Wunsch, 1998: Abyssal recipes II: Energetics of tidal and wind mixing. *Deep-Sea Res. I,* **45,** 1977–2010.

Nakano, M., 1932: Preliminary note on the accumulation and dissipation of energy of the secondary oscillations in a bay. *Proc. Phys.-Math. Soc. Japan,* **Ser. 3, 14,** 44–56.

Nayfeh, A. H., and D. T. Mook, 1979: *Nonlinear Oscillations.* Wiley-Interscience, 704 pp.

Platzman, G., 1971: Ocean tides and related waves. *Mathematical Problems in the Geophysical Sciences.* Vol. 1, *Geophysical Fluid Dynamics,* W. H. Reid, Ed., American Mathematical Society, 239–291.

Prudnikov, A. P., Yu. A. Brychkov, and O. I. Marichev, 1986: *Integrals and Series.* Vol. I, Gordon and Breach, 798 pp.

Pugh, D. T., 1987: *Tides, Surges and Mean Sea-Level.* John Wiley and Sons, 472 pp.

Ruelle, D., 1980: Strange attractors. *Math. Intelligencer,* **2,** 126–137. [Reprinted in Cvitanović, P., 1984: *Universality in Chaos,* Adam-Hilger, 37–48.]

Schwiderski, E. W., 1980: On charting global ocean tides. *Rev. Geophys. Space Phys.,* **18,** 243–268.

Seim, H. E., and J. E. Sneed, 1988: Enhancement of semidiurnal tidal currents in the tidal inlets to Mississippi Sound. *Hydrodynamics and Sediment Dynamics of Tidal Inlets,* D. G. Aubrey and L. Weishar, Eds., Lecture Notes on Coastal and Estuarine Studies, Vol. 29, Springer-Verlag, 157–168.

Smith, N. P., 1994: Water, salt and heat balance of coastal lagoons. *Coastal Lagoon Processes,* B. Kjerfve, Ed., Elsevier Oceanographic Series, Vol. 60, 69–101.

van Ette, A. C. M., and H. J. Schoemaker, 1966: Harmonic analysis of tides. Essential features and disturbing influences. *Hydrographic Letter,* Vol. 1, Netherlands' Hydrographer Special Publication 2, 1–35.

Vittori, G., 1992: On the chaotic structure of tide elevation in the Lagoon of Venice. *Proc. 23d Int. Conf. on Coastal Engineering,* Venice, Italy, American Society of Civil Engineers, 1826–1839.

Wiggins, S., 1990: *Introduction to Applied Nonlinear Dynamical Systems and Chaos.* Springer-Verlag, 672 pp.

Yagasaki, K., 1990: Second-order averaging and chaos in quasiperiodically forced weakly nonlinear oscillators. *Physica D,* **44,** 445–458.

——, 1993: Chaotic motions near homoclinic manifolds and resonant tori in quasiperiodic perturbations of planar Hamiltonian systems. *Physica D,* **69,** 232–269.

Zimmerman, J. T. F., 1992: On the Lorentz linearization of a nonlinearly damped tidal Helmholtz oscillator. *Proc. K. Ned. Akad. Wet.,* **95,** 127–145.